Scale

Geoffrey West

Penguin Publishing Group

Annotations by color

100 yellow notes • 64 green notes • 93 blue notes • 39 red notes

Created by Oszkár – Last synced October 5, 2023

Yellow

Scaling simply refers, in its most elemental form, to how a system responds when its size changes. What happens to a city or a company if its size is doubled? Or to a building, an airplane, an economy, or an animal if its size is halved? If the population of a city is doubled, does the resulting city have approximately twice as many roads, twice as much crime, and produce twice as many patents? Do the profits of a company double if its sales double, and does an animal require half as much food if its weight is halved? Addressing such seemingly innocuous questions concerning how systems respond to a change in their size has had remarkably profound consequences across the entire spectrum of science,

26

engineering, and technology and has affected almost every aspect of our lives. Scaling arguments have led to a deep understanding of the dynamics of tipping points and phase transitions (how, for example, liquids freeze into solids or vaporize into gases), chaotic phenomena (the "butterfly effect" in which the mythical flapping of a butterfly's wings in Brazil leads to a hurricane in Florida), the discovery of quarks (the building blocks of matter), the unification of the fundamental forces of nature, and the evolution of the universe after the Big Bang. These are but a few of the more spectacular examples where scaling arguments have been instrumental in illuminating important universal principles or structure. 9

September 17, 2023

A profound consequence of this rule is that on a per gram basis, the larger animal (the woman in this example) is actually more efficient than the smaller one (her dog) because less energy is required to support each gram of her tissue (by about 25 percent). Her horse, by the way, would be even more efficient. This systematic savings with increasing size is known as an economy of scale . Put succinctly, this states that the bigger you are, the less you need per capita (or, in the case of animals, per cell or per gram of tissue) to stay alive. Notice that this is the opposite behavior to the case of increasing returns to scale, or superlinear scaling, manifested in the GDP of cities: in that case, the bigger you are, the more there is per capita, whereas for economies of scale, the bigger you are, the less there is per capita. This kind of scaling is referred to as sublinear scaling. 30

September 17, 2023

To set the stage, I'd like to quote two distinguished thinkers, one a scientist, the other a lawyer. The first is the eminent physicist Stephen Hawking, who in an interview 10 at the turn of the millennium was asked the following question: Some say that while the 20th century was the century of physics, we are now entering the century of biology. What do 31 *you think of this? To which he responded: I think the next century will be the century of complexity. I wholeheartedly agree. As I hope I have already made clear, we urgently need a science of complex adaptive systems to address the host of extraordinarily challenging societal problems we face.*

September 17, 2023

A typical complex system is composed of myriad individual constituents or agents that once aggregated take on collective characteristics that are usually not manifested in, nor could easily be predicted from, the properties of the individual components themselves. For example, you are much more than the totality of your cells and, similarly, your cells are much more than the totality of all of the molecules from which they are composed 33

September 17, 2023

Ant colonies are built without forethought and without the aid of any single mind or any group discussion or consultation. There is no blueprint or master plan. Just thousands of ants working mindlessly in the dark moving millions of grains of earth and sand to create these impressive structures. This feat is accomplished by each individual ant obeying just a few simple rules mediated by chemical cues and other signals, resulting in an extraordinarily coherent collective output. It is almost as if they were programmed to be microscopic operations in a giant computer algorithm. 34

This feat is accomplished by each individual ant obeying just a few simple rules mediated by chemical cues and other signals, resulting in an extraordinarily coherent collective output. It is almost as if they were programmed to be microscopic operations in a giant computer algorithm.

September 17, 2023

In general, then, a universal characteristic of a complex system is that the whole is greater than, and often significantly different from, the simple linear sum of its parts. In many instances, the whole seems to take on a life of its own, almost dissociated from the specific characteristics of its individual building blocks. Furthermore, even if we understood how the individual constituents, whether cells, ants, or people, interact with one another, predicting the systemic behavior of the resulting whole is not usually possible. This collective outcome, in which a system manifests significantly different characteristics from those resulting from simply adding up all of the contributions of its individual constituent parts, is called an emergent behavior . It is a readily recognizable characteristic of economies, financial markets, urban communities, companies, and organisms.

35

September 17, 2023

Closely related to the concepts of emergence and self-organization is another critical characteristic of many complex systems, namely their ability to adapt and evolve in response to changing external conditions. The quintessential example of such a complex adaptive system is, of course, life itself in all of its extraordinary manifestations from cells to cities . The Darwinian theory of natural selection is the scientific narrative that has been developed for understanding and describing how organisms and ecosystems continuously evolve and adapt to changing conditions.

36

September 17, 2023

Scaling up from the small to the large is often accompanied by an evolution from simplicity to complexity while maintaining basic elements or building blocks of the system unchanged or conserved. This is familiar in engineering, economies, companies, cities, organisms, and perhaps most dramatically, evolutionary processes. For example, a skyscraper in a large city is a significantly more complex

37

object than a modest family dwelling in a small town, but the underlying principles of construction and design, including questions of mechanics, energy and information distribution, the size of electrical outlets, water faucets, telephones, laptops, doors, et cetera, all remain approximately the same independent of the size of the building. These basic building blocks do not significantly change when scaling up from my house to the Empire State Building; they are shared by all of us. Similarly, organisms have evolved to have an enormous range of sizes and an extraordinary diversity of morphologies and interactions, which often reflect increasing complexity, yet fundamental building blocks like cells, mitochondria, capillaries, and even leaves do not appreciably change with body size or increasing complexity of the class of systems in which they are embedded.

September 17, 2023

7. YOU ARE YOUR NETWORKS: GROWTH FROM CELLS TO WHALES I began this chapter by pointing out the very surprising and counterintuitive fact that, despite the vagaries and accidents inherent in evolutionary dynamics, almost all of the most fundamental and complex measurable characteristics of organisms scale with size in a remarkably simple and regular fashion. This is explicitly illustrated, for example, in Figure 1, where metabolic rate is plotted against body mass for a sequence of animals.

38

September 17, 2023

Highly complex, self-sustaining structures, whether cells, organisms, ecosystems, cities, or corporations, require the close integration of enormous numbers of their constituent units that need efficient servicing at all scales. This has been accomplished in living systems by evolving fractal-like, hierarchical branching network systems presumed optimized by the continuous "competitive" feedback mechanisms implicit in natural selection. It is the generic physical, geometric, and mathematical properties of these network systems that underlie the origin of these scaling laws,

40

including the prevalence of the one-quarter exponent. As an example, Kleiber's law follows from requiring that the energy needed to pump blood through mammalian circulatory systems, including ours, is minimized so that the energy we devote to reproduction is maximized. Examples of other such networks include the respiratory, renal, neural, and plant and tree vascular systems. These ideas, as well as the concepts of space filling (the need to feed all cells in the body) and fractals (the geometry of the networks), will be elaborated upon in some detail.

September 17, 2023

This process can be analyzed using the network theory to predict a universal quantitative theory of growth curves applicable to any organism, including tumors. A growth curve is simply a graph of the size of the organism plotted as a function of its age. You are probably familiar with growth curves if you have had children, as pediatricians routinely show them to parents so that they can see how their child's development compares with the expectations for the average infant. The growth theory also explains a curious paradoxical phenomenon that you might have pondered, namely, why we eventually stop growing even though we continue to eat. This turns out to be a consequence of the sublinear scaling of metabolic rate and the economies of scale embodied in the network design. In a later chapter, the same paradigm will be applied to the growth of cities, companies, and economies to understand the fundamental question as to the origins of open-ended growth and its possible sustainability. Because networks determine the rates at which energy and resources are delivered to cells, they set the pace of all physiological processes. Because cells are constrained to operate systematically slower in larger organisms relative to smaller ones, the pace of life systematically decreases with increasing size. Thus, large mammals live longer, take longer to mature, have slower heart rates, and cells that don't work as hard as those of small mammals, all to the same predictable degree. Small creatures live life in the fast lane while large ones move ponderously, though more efficiently, through life; think of a scurrying mouse relative to a sauntering elephant. 41

The peace of life.

September 17, 2023

Perhaps even more remarkably they are also scaled socioeconomic versions of one another. Socioeconomic quantities such as wages, wealth, patents, AIDS cases, crime, and educational institutions, which have no analog in biology and did not exist on the planet before humans invented cities ten thousand years ago, also scale with population size but with a superlinear (meaning bigger than one) exponent of approximately 1.15. An example of this is the number of patents produced in a city shown in Figure 3. Thus, on a per capita basis, all of these quantities systematically increase to the same degree as city size increases and, at the same time, there are equivalent savings from economies of scale in all infrastructural quantities. Despite their amazing diversity and complexity across the globe, and despite localized urban planning, cities manifest a surprising coarse-grained simplicity, regularity, and predictability. 15 To put it in simple terms, scaling implies that if a city is twice the size of another city in the same country (whether 40,000 vs. 20,000 or 4 million vs. 2 million), then its wages, wealth, number of patents, AIDS cases, violent crime, and educational institutions all increase by approximately the same degree (by about 15 percent above mere doubling), with similar savings in all of its infrastructure. The bigger the city, the more the average individual systematically owns, produces, and consumes, whether goods, resources, or ideas. The good, the bad, and the ugly are integrated in an approximately predictable package: a person may move to a bigger city drawn by more innovation, a greater sense of "action," and higher wages, but she can also expect to confront an equivalent increase in the prevalence of crime and disease.

43

September 17, 2023

The fact that the same scaling laws are observed for diverse

44

urban metrics in cities and urban systems that evolved independently across the globe strongly suggests that, as in biology, there are underlying generic principles transcending history, geography, and culture and that a fundamental, coarse-grained theory of cities is possible. In chapter 8 I will discuss how the inextricable tension between benefits and costs of social and infrastructural networks has its origins in the underlying universal dynamics of social network structures and group clustering of human interactions. Cities provide a natural mechanism for reaping the benefits of high social connectivity among very different people conceiving and solving problems in a diversity of ways. I will discuss the nature and dynamics of these social network structures and show how scaling laws emerge, including the intriguing link between the 15 percent enhancement of all socioeconomic activities, whether good or bad, and the equivalent 15 percent savings on physical infrastructure.

September 17, 2023

9. COMPANIES AND BUSINESSES It is natural to extend these ideas to ask how they might relate to companies. Could there possibly be a quantitative, predictive science of companies ? Do companies manifest systematic regularities that transcend their size and business character? For example, in terms of sales and assets, are Walmart and Exxon, whose revenues exceed half a trillion dollars, approximately scaled-up versions of smaller companies with sales of less than $10 million? Amazingly, the answer to this is yes, as can be seen from Figure 4: like organisms and cities, companies also scale as simple power laws. Equally surprising is that they scale sublinearly as functions of their size, rather than superlinearly like socioeconomic metrics in cities. In this sense, companies are much more like organisms than cities. The scaling exponent for companies is around 0.9, to be compared with 0.85 for the infrastructure of cities and 0.75 for organisms. However, there is considerably more variation around precise scaling among companies than for organisms or cities. This is especially so in their early stages of development as they jostle for a place in the market. Nevertheless, the surprising regularity manifested in their average behavior suggests that, despite their broad diversity

46

and apparent individuality, companies grow and function under general constraints and principles that transcend their size and business sector.

September 17, 2023

GALILEO'S ARGUMENT ON HOW AREAS AND VOLUMES SCALE To begin, consider one of the simplest possible geometrical objects, namely, a floor tile in the shape of a square, and imagine scaling it up to a larger size; see Figure 5. To be specific let's take the length of its sides to be 1 ft. so that its area, obtained by multiplying the length of two adjacent sides together, is 1 ft. × 1 ft. = 1 sq. ft. Now, suppose we double the length of all of its sides from 1 to 2 ft., then its area increases to 2 ft. × 2 ft. = 4 sq. ft. Similarly, if we were to triple the lengths to 3 ft., then its area would increase to 9 sq. ft., and so on. The generalization is clear: the area increases with the square of the lengths. This relationship remains valid for any two-dimensional geometric shape, and not just for squares, provided that the shape is kept fixed when all of its linear dimensions are increased by the same factor. A simple example is a circle: if its radius is doubled, for instance, then its area increases by a factor of 2 × 2 = 4. A more general example is that of doubling the dimensions of every length in your house while keeping its shape and structural layout the same, in which case the area of all of its surfaces, such as its walls and floors, would increase by a factor of four. 53

September 18, 2023

This kind of scale where instead of increasing linearly as in 1, 2, 3, 4, 5 . . . we increase by factors of 10 as in the Richter scale: 101, 102, 103, 104, 105 . . . is called logarithmic. Notice that it's actually linear in terms of the numbers of orders of magnitude, as indicated by the exponents (the superscripts) on the tens. Among its many attributes, a logarithmic scale allows one to plot quantities that differ by huge factors on the same axis, such as those between the magnitudes of the Valdivia earthquake, the Northridge earthquake, and a stick 62 *of dynamite, which overall cover a range of more than a billion (109). This would be impossible if a linear plot was used because almost all of the events would pile up at the lower end of the graph. To include all earthquakes, which range over five or six orders of magnitude, on a linear plot would require a piece of paper several miles long—hence the invention of the Richter scale.*

September 18, 2023

Let me illustrate this using the weight lifting example. If you look carefully at the graph of Figure 7 you can clearly see that four of the points lie almost exactly on the line, indicating that these weight lifters are lifting almost precisely what they should for their body weights. However, notice that the remaining two, the heavyweight and the middleweight, both lie just a little off the line, one below and one above. Thus, the heavyweight, even though he has lifted more than anyone else, is actually under performing relative to what he should be lifting given his weight, whereas the middleweight is over performing relative to his weight. In other words, from the egalitarian level playing field perspective of a physicist, the strongest man in the world in 1956 was actually the middleweight champion because he was overperforming relative to his weight. Ironically, the weakest of all of the champions from this scientific scaling perspective is the heavyweight, despite the fact that he lifted more than anyone else. 65

September 18, 2023

So how in fact does weight scale with height for human beings? Various statistical analyses of data have led to varying conclusions, ranging from confirmation of the cubic law to more recent analyses suggesting exponents of 2.7 and values that are even smaller and closer to two. 10 To understand why this might be so, we have to remind ourselves of a major assumption that was made in deriving the cubic law, namely that the shape of the system, our bodies in this case, should remain the same when its size 75 *increases. However, shapes change with age, from the extreme case of a baby, with its large head and chunky limbs, to a mature "well-proportioned" adult, and finally to the sagging bodies of people my age. In addition, shapes also depend on gender, culture, and other socioeconomic factors that may or may not be correlated with health and obesity.*

September 18, 2023

8. INNOVATION AND LIMITS TO GROWTH Galileo's deceptively simple argument for why there are limits to the heights of trees, animals, and buildings has profound consequences for design and innovation. Earlier, when explaining his argument I concluded with the remark: Clearly, the structure, whatever it is, will eventually collapse under its own weight if its size is arbitrarily increased. There are limits to size and growth. To which should have been added the critical phrase " unless something changes. " Change and, by implication, innovation, must occur in order to continue growing and avoid collapse. Growth and the continual need to be adapting to the challenges of new or changing environments, often in the form of "improvement" or increasing efficiency, are major drivers of innovation. 76

September 18, 2023

Models of various kinds have been built for centuries, especially in architecture, but these were primarily to illustrate the aesthetic characteristics of the real thing rather than as scale models to test, investigate, or demonstrate the dynamical or physical principles of the system being constructed. And most important, they were almost always "made to scale," meaning that each detailed part was in some fixed proportion to the full size—1:10, for example—just like a map. Each part of the model was a linearly scaled representation of the actual-size ship, cathedral, or city being "modeled." Fine for aesthetics and toys but not much good for learning how the real system works. Nowadays, every conceivable process or physical object, from automobiles, buildings, airplanes, and ships to traffic congestion, 80 *epidemics, economies, and the weather, is simulated on computers as "models" of the real thing. I discussed earlier how specially bred mice are used in biomedical research as scaled-down "models" of human beings. In all of these cases, the big question is how do you realistically and reliably scale up the results and observations of the model system to the real thing? This entire way of thinking has its origins in a sad failure in ship design in the middle of the nineteenth century and the marvelous insights of a modest gentleman engineer into how to avoid it in the future.*

September 18, 2023

Failure and catastrophe can provide a huge impetus and opportunity in stimulating innovation, new ideas, and inventions whether in science, engineering, finance, politics, or one's personal life. 80

September 18, 2023

In other words, a larger ship requires proportionately less fuel to transport each ton of cargo than a smaller ship . Bigger ships are therefore more energy efficient and cost effective than smaller ones—another great example of an economy of scale and one that had enormous consequences for the development of world trade and commerce. 12 85

September 18, 2023

The crucial point recognized by Froude was that because the underlying physics remains the same, objects of different sizes moving at different speeds behave in the same way if their Froude numbers have the same value . Thus, by making the length and speed of the model ship have the same Froude number as that of the full-size version, one can determine the dynamical behavior of the full-size ship before building it. 94

September 19, 2023

11. SIMILARITY AND SIMILITUDE: DIMENSIONLESS AND SCALE-INVARIANT NUMBERS The scaling methodology introduced by Froude has now evolved to become a powerful and sophisticated component of the tool kit of science and engineering and has been applied with great effect to a very broad and diverse range of problems. It was not formalized as a general technique until the beginning of the twentieth century, when the eminent mathematical physicist Lord Rayleigh wrote a provocative and highly influential paper in the journal Nature titled "The Principle of Similitude." 16 This was his term for what we have been calling scaling theory. His major emphasis was on the primary role played in any physical system by special quantities that have the property of being dimensionless.

96

September 19, 2023

This is the simplest example of a dimensionless quantity: it is a "pure" number that does not change when a different system of units is used to measure it. This scale invariance expresses something absolute about the quantities they represent in that the dependence on the arbitrariness of the human choice of units and measurement has been removed. Specific units are convenient inventions of human beings for communicating measures in a standardized language, especially regarding construction, commerce, and the exchange of goods and services. Indeed, the introduction of standardized measures marked a critical stage in the evolution of civilization and the rise of cities, as they were crucial in developing a trustworthy political fabric subject to the rule of law. Perhaps the most famous dimensionless number is pi (π), the ratio of the circumference of a circle to its diameter. This has no units, because it is the ratio of two lengths, and has the same value for all circles everywhere, at all times, no matter how small or how large they are. π therefore embodies the universal quality of "circleness." This concept of "universality" is the reason why the acceleration due to gravity was included in the definition of the Froude number, even though it played no explicit role in how to scale

96

from model ships to the real thing. It turns out that the ratio of the square of the velocity to the length is not dimensionless and so depends on the units used, whereas dividing by gravity renders it dimensionless and therefore scale invariant. But why was gravity chosen and not some other acceleration? Because gravity ubiquitously constrains all motion on Earth. This is pretty evident in our own walking and running where we have to continually fight gravity to raise our legs with each step forward, especially when going uphill. Not quite so obvious is how it enters into the motion of ships, because the buoyancy of water balances gravity (remember Archimedes' principle). However, as a ship moves through water it continually creates wakes and surface waves whose behavior is constrained by the pull of gravity—in fact, the waves you are familiar with on oceans and lakes are technically called gravity waves. So indirectly gravity plays an important role in the motion of ships. Consequently, Froude's number embodies a "universal" quality associated with all motion on Earth transcending the specific details of the object that is moving. Its value is therefore a major determinant not only in the motion of ships, but also for cars, airplanes, and ourselves. Furthermore, it also tells us how these motions on another planet with a different gravitational strength would differ from the same motion on Earth.

Di

September 19, 2023

3 THE SIMPLICITY, UNITY, AND COMPLEXITY OF LIFE As emphasized in the opening chapter, living systems, from the smallest bacteria to the largest cities and ecosystems, are quintessential complex adaptive systems operating over an enormous range of multiple spatial, temporal, energy, and mass scales. In terms of mass alone, the overall scale of life covers more than thirty orders of magnitude (1030) from the molecules that power metabolism and the genetic code up to ecosystems and cities. This range vastly exceeds that of the mass of the Earth relative to that of our entire galaxy, the Milky Way, which covers "only" eighteen orders of magnitude, and is comparable to the mass of an electron

101

relative to that of a mouse.

September 19, 2023

I was a theoretical physicist (and still am) whose primary research interests at that time were in this area. My visceral reaction to the provocative statements concerning the diverging trajectories of physics and biology was that, yes, biology will almost certainly be the predominant science of the twenty-first century, but for it to become truly successful, it will need to embrace some of the quantitative, analytic, predictive culture that has made physics so successful. 105

September 19, 2023

To varying degrees, all theories and models are incomplete. They need to be continually tested and challenged by increasingly accurate experiments and observational data over wider and wider domains and the theory modified or extended accordingly. This is an essential ingredient in the scientific method. Indeed, understanding the boundaries of their applicability, the limits to their predictive power, and the ongoing search for exceptions, violations, and failures has provoked even deeper questions and challenges, stimulating the continued progress of science and the unfolding of new ideas, techniques, and concepts. 134

September 20, 2023

A concept related to the idea of a toy model is that of a "zeroth order" approximation of a theory, in which simplifying assumptions are similarly made in order to give a rough approximation of the exact result. It is usually employed in a quantitative context as, for example, in the statement that "a zeroth order estimate for the population of the Chicago metropolitan area in 2013 is 10 million people." Upon learning a little more about Chicago, one might make what could be called a "first order" estimate of its population 136 *of 9.5 million, which is more precise and closer to the actual number (whose precise value from census data is 9,537,289). One could imagine that after more detailed investigation, an even better estimate would yield 9.54 million, which would be called a "second order" estimate. You get the idea: each succeeding "order" represents a refinement, an improved approximation, or a finer resolution that converges to the exact result based on more detailed investigation and analysis. In what follows, I shall be using the terms "coarse-grained" and "zeroth order" interchangeably.*

September 20, 2023

However, I agree with Sydney Brenner, the distinguished biologist who received the Nobel Prize for his work on the genetic code and who provocatively remarked that "technology gives us the tools to analyze organisms at all scales, but we are drowning in a sea of data and thirsting for some theoretical framework with which to understand it. . . . We need theory and a firm grasp on the nature of the objects we study to predict the rest." His article begins, by the way, with the astonishing pronouncement that "biological research is in crisis." 10

138

September 20, 2023

The rate at which you use energy to pump blood through the vasculature of your circulatory system is called your cardiac power output. This energy expenditure is used to overcome the viscous drag, or friction, on blood as it flows through increasingly narrower and narrower vessels in its journey through the aorta, which is the first artery leaving your heart, down through multiple levels of the network to the tiny capillaries that feed cells. A human aorta is an approximately cylindrical pipe that is about 18 inches long (about 45 cm) and about an inch (about 2.5 cm) in diameter, whereas our capillaries are only about 5 micrometers wide (about a hundredth of an inch), which is somewhat smaller than a hairbreadth. 13 Although a blue whale's aorta is almost a foot in diameter (30 cm), its capillaries are still

147

pretty much the same size as yours and mine. This is an explicit example of the invariance of the terminal units in these networks.

September 20, 2023

For example, the smooth and efficient functioning of social networks, whether in a society, a company, a group activity, and especially in relationships such as marriages and friendships, requires good communication in which information is faithfully transmitted between groups and individuals. When information is dissipated or "reflected," such as when one side is not listening, it cannot be faithfully or efficiently processed, inevitably leading to misinterpretation, a process analogous to the loss of energy when impedances are not matched.

[152](#)

September 21, 2023

But what's really surprising is that blood pressures are also predicted to be the same across all mammals, regardless of their size. Thus, despite the shrew's heart weighing only about 12 milligrams, the equivalent of about 25 grains of salt, and its aorta having a radius of only about 0.1 millimeter and consequently barely visible, whereas a whale's heart weighs about a ton, almost the weight of a Mini Cooper, and its aorta has a radius of about 30 centimeters, their blood pressures are approximately the same. This is pretty amazing—just think of the enormous stresses on the walls of the shrew's tiny aorta and arteries compared with the pressures on yours or mine, let alone those on a whale's. No wonder the poor creature dies after only a year or two.

[155](#)

September 21, 2023

In addition to his work on the cardiovascular system, Young was famous for several other quite diverse and profound discoveries. He is perhaps best known for establishing the

[156](#)

wave theory of light, in which each color is associated with a particular wavelength. But he also contributed to early work in linguistics and Egyptian hieroglyphics, including being the first person to decipher the famous Rosetta Stone now sitting in the British Museum in London. As a fitting tribute to this remarkable man, Andrew Robinson wrote a spirited biography of Young titled The Last Man Who Knew Everything: Thomas Young, the Anonymous Polymath Who Proved Newton Wrong, Explained How We See, Cured the Sick, and Deciphered the Rosetta Stone, Among Other Feats of Genius. I have a certain soft spot for Young because he was born in Milverton, in the county of Somerset in the West of England, just a few short miles from Taunton, where I was born.

September 21, 2023

This repetitive phenomenon is called self-similarity and is a generic characteristic of fractals. Analogous to the repetitive scaling exhibited by broccoli are the infinite reflections in parallel mirrors, or the nesting of Russian dolls (matryoshka) of regularly decreasing sizes inside one another. Long before the concept was invented, self-similarity was poetically expressed by the Irish satirist Jonathan Swift, the author of Gulliver's Travels, in this whimsical quatrain: So, naturalists observe, a flea Hath smaller fleas that on him prey; And these have smaller still to bite 'em; And so proceed ad infinitum. 159

September 21, 2023

So it is with the hierarchical networks we have been discussing. If you cut a piece out of such a network and appropriately scale it up, then the resulting network looks just like the original. Locally, each level of the network essentially replicates a scaled version of the levels adjacent to it. We saw an explicit example of this when discussing the consequences of impedance 159 *matching in the pulsatile regime of the circulatory system where area−preserving branching resulted in the radii of successive vessels decreasing by a constant factor (√2 = 1.41 . . .) with each successive branching. So, for example, if we compare the radii of vessels separated by 10 such branchings, then they are related by a scale factor of (√2)10 = 32. Because our aorta has a radius of about 1.5 centimeters, this means that the radii of vessels at the tenth branching level are only about half a millimeter.*

September 21, 2023

The point of adding the 1 was to connect the idea of fractals to the conventional concept of ordinary dimensions discussed in chapter 2. Recall that a smooth line has dimension 1, a smooth surface dimension 2, and a volume dimension 3. Thus the South African coast is very close to being a smooth line because its fractal dimension is 1.02, which is very close to 1, whereas Norway is far from it because its fractal dimension of 1.52 is so much greater than 1. You could imagine an extreme case of this in which the line is so crinkly and convoluted that it effectively fills an entire area. Consequently, even though it's still a line with "ordinary" dimensions 1, it behaves as if it were an area in terms of its scaling properties, therefore having a fractal dimension of 2. This curious gain of an effective additional dimension is a general feature of space-filling curves, to which I will return in the next chapter. 173

September 21, 2023

In fact, those that are most seriously at risk have fractal dimensions close to one with an uncharacteristically smooth EKG. Thus the fractal dimension of the EKG provides a potentially powerful complementary diagnostic tool for quantifying heart disease and health. 24 The reason that 177 *being healthy and robust equates with greater variance and larger fluctuations, and therefore a larger fractal dimension as in an EKG, is closely related to the resilience of such systems. Being overly rigid and constrained means that there isn't sufficient flexibility for the necessary adjustments needed to withstand the inevitable small shocks and perturbations to which any system is subjected. Think of the stresses and strains your heart is exposed to every day, many of which are unexpected. Being able to accommodate and naturally adapt to these is critical for your long-term survival. These continuous changes and impingements require all of your organs, including your brain as well as its psyche, to be both flexible and elastic and therefore to have a significant fractal dimension.*

September 21, 2023

This is all very satisfying, but a lingering question remains: why does the same ¼ power exponent emerge from each of these different networks rather than, say, a 1/6 power emerging in one, a ⅛ power in another, and so on? In other words, what necessitates that this same set of principles should lead to the same scaling exponents when applied to different network systems with a variety of structures and dynamics? Are there additional design principles that transcend the dynamics that ensure that the ¼ emerges in virtually all organismic groups? This is an important conceptual question, especially for understanding why this universal behavior extends even to systems such as bacteria, where an explicit hierarchical branching network structure is much less obvious. 185

September 22, 2023

A general argument to address this can be made by recognizing that, in addition to minimizing energy loss, natural selection has also led to a maximization of metabolic capacity because metabolism produces the energy and 185 *materials required to sustain and reproduce life. 1 This has been achieved by maximizing surface areas across which resources and energy are transported. These surfaces are in actuality the total surface areas of all the terminal units of the network. For instance, all of our metabolic energy is transmitted across the total surface area of all of our capillaries to fuel our cells, just as the metabolism of a tree is governed by the transmission of energy gathered from sunlight through all of its leaves to fuel photosynthesis and of water from soil through all of the terminal fibers of its root system.*

September 22, 2023

The answer lies in the subtleties of networks and their interplay with physiological limits in the spirit of Galileo's original argument that there are limits to the maximum size of structures. Unlike most biological networks, mammalian circulatory systems are not single self-similar fractals but an admixture of two different ones, reflecting the change in flow from predominantly pulsatile AC to predominantly nonpulsatile DC as blood flows from the aorta to the capillaries. Most of the blood resides in the larger vessels of the upper part of the network where AC dominates, leading to the ¾ power scaling law for metabolic rates. Although the branching changes continuously from one mode to the other, the region over which it changes is relatively narrow and its location (as measured by the number of branchings up from the capillaries) is independent of body size and therefore the same for all mammals. In other words, all mammals have roughly the same number of branching levels, about fifteen, where the flow is predominantly steady nonpulsatile DC. The distinction among mammals as their size increases is the increasing number of levels where the flow is pulsatile AC. For example, we have about seven to eight, the whale has about sixteen to seventeen, and the shrew just one or two. Impedance matching in these vessels ensures that relatively little energy is required to pump blood through them, so the more of them you have the better. Almost all of your cardiac output goes into pumping blood through the much smaller vessels of the nonpulsatile regime, whose number of levels is approximately the same for all mammals. Relatively 190

speaking, then, the proportion of the network where the heart expends most of its energy systematically decreases as the size of a mammal increases, illustrating again that larger mammals are more efficient than smaller ones: the whale needs only one hundredth the amount of energy needed by a shrew to supply blood to one of its cells.

September 22, 2023

This difference in the way these two contributions scale with increasing size plays a central role in growth, so to ensure that you understand what it implies here's a simple example to illustrate the point. Suppose the size of the organism doubles; then the number of cells also doubles, so the amount of energy needed for their maintenance increases by a factor of 2. However, metabolic rate (the supply of energy) increases by only a factor of 2¾ = 1.682 . . . which is less than 2. So the rate at which energy is needed for maintenance increases faster than the rate at which metabolic energy can be supplied, forcing the amount of energy available for growth to systematically decrease and eventually go to zero, resulting in the cessation of growth. In other words, you stop growing because of the mismatch between the way maintenance and supply scale as size increases. 203

September 22, 2023

I want to emphasize just how remarkable this is. The two most important events in an organism's life, its birth and its death, which are usually thought of as being independent, are intimately related to each other: the slopes of these two graphs are determined by exactly the same parameter, the 0.65 eV, representing the average energy needed to produce an ATP molecule. Below I will explore this further when I discuss how a more fundamental theory of aging based on network dynamics explains the mechanistic origins of this temperature dependence. 212

September 23, 2023

Night Thoughts During the Hour of the Wolf According to the ancient Romans, the Hour of the Wolf means the time between night and dawn, just before the light comes, and people believed it to be the time when demons had a heightened power and vitality, the hour when most people died and most children were born, and when nightmares came to one. 12

216

September 23, 2023

This leads to the notion of a maximum possible age that a human being could conceivably live, which turns out to be less than around 125 years. Very few people have even approached this age. The oldest verified person ever was the Frenchwoman Jeanne Calment, who died in 1997 at the remarkable age of 122 years and 164 days. Just to get a sense of how exceptional this is, the next oldest verified person was the American Sarah Knauss, who lived more than three years fewer than Jeanne, dying at the age of 119 years and 97 days. The next super-champs of long life lived almost two years fewer than Sarah, while the oldest person still alive today is the Italian Emma Murano, who is "only" in her 118th year.

227

September 23, 2023

The radical challenge: is it possible to extend life span beyond the apparent maximum limit of approximately 125 years and live, for instance, to 225 years? In a very real sense we're already achieving the first, whereas it is the second that raises serious scientific questions.

228

September 23, 2023

This coarse-grained framework is very general and can incorporate any model of aging based on generalizations of

242

"damage" mechanisms associated with generalized physical or chemical transport phenomena discussed above. The details of the damage mechanism are not important for understanding many of the general features of aging and mortality because the most relevant damage occurs in invariant terminal units of networks (capillaries and mitochondria, for example) whose properties do not appreciably change with the size of the organism. Consequently, the damage per capillary or mitochondrion is approximately the same regardless of the animal. Because these networks are space filling, meaning that they service all cells and mitochondria throughout the body of the organism, damage occurs approximately uniformly and relentlessly throughout the organism, explaining why aging is approximately spatially uniform and progresses approximately linearly with age. This is why at age seventy-five every part of your body has deteriorated to pretty much the same degree, as indicated in Figure 27.

September 24, 2023

260

The growth of a system, whether an economy or a population, is often expressed in terms of a quantity called the doubling time, which is simply the time it takes for the size of the system to double. Exponential growth is characterized by having a constant doubling time, which also sounds fairly harmless until one realizes that it implies, for example, that it would take the same time for a population to double from ten thousand to twenty thousand, thereby adding just ten thousand people, as it would for it to double from 20 million to 40 million, thereby adding a humongous 20 million people. Amazingly, the doubling time of the global population has actually been getting systematically shorter and shorter as was indicated above: it took 300 years from 1500 to 1800 for the population to double from 500 million to a billion, but only 120 years to double to 2 billion, and only another 45 to double again to 4 billion.

September 25, 2023

Questions such as these have fueled a highly spirited ongoing debate that began very soon after the Industrial Revolution set the wheels of exponential growth in motion and has continued right up to the present time. The transition from working the land and artisanal hand production to automated machines and the creation of factories for the mass production of goods, the technological innovation and increased productivity in agriculture, the introduction of new chemical manufacturing and iron production processes, the improved efficiency of water power, and the increasing use of steam fueled by the change from renewable wood energy to fossilized coal energy, all contributed to an inevitable migration of more and more people away from a traditional rural existence to the rapidly expanding urban centers that were perceived as providing greater opportunities for employment. This process continues unabated across the globe to this day. 3

266

September 25, 2023

Thomas Robert Malthus is usually credited with being the first person to recognize the potential threat posed by open-ended exponential growth and connect it to the challenge of resource limitation and availability. Malthus was an English cleric and scholar and an early contributor to the newly emerging fields of economics and demography and their implications for long-term political strategy. He published a highly influential essay in 1798 titled An Essay on the Principle of Population in which he declared that "the power of population is indefinitely greater than the power in the earth to produce subsistence for man." His argument was that the population "multiplies geometrically," meaning that it increases at an exponential rate, whereas the ability to grow and supply food increases only "arithmetically," meaning that it increases at a much slower linear rate, so the size of the population will eventually outstrip the food supply, leading to catastrophic collapse.

271

September 25, 2023

I have met scant few economists who do not automatically dismiss traditional Malthusian-like ideas of eventual or imminent collapse as naive, simplistic, or just plain wrong. On the other hand, I have met scant few physicists or ecologists who think it's nuts to believe otherwise. The late maverick economist Kenneth Boulding perhaps best summed it up when testifying before the U.S. Congress, declaring that "anyone who believes exponential growth can go on forever in a finite world is either a madman or an economist."

274

September 25, 2023

Consequently, the processes that govern the weather and the life history of plants and animals are exponentially sensitive to small changes in the temperature at which they operate. I remind you that a 2°C rise in average temperature leads to a whopping 20 percent increase in these rates.

283

September 26, 2023

Of the trillions of thoughts, ideas, speculations, and proposals for new machines, new products, and new theories, only an infinitesimal minority ever lead to any significance. Almost all fall by the wayside, even though as a totality they all contribute a necessary background noise and weltanschauung for new and innovative phenomena to arise and blossom. All of this requires huge amounts of energy: ex nihilo nihil fit —nothing comes from nothing.

286

September 26, 2023

The lack of progress until relatively recently is quite astonishing given that the basic technology for developing solar energy has been known for more than a hundred years. In 1897, an American engineer, Frank Shuman, built a proof of principle device for utilizing energy from the sun by showing that it could power a small steam engine. His system was eventually patented in 1912, and in 1913 he

288

constructed the world's first solar thermal energy power plant, which was built in Egypt. It generated only about 50 kilowatts (about 65 horsepower) but was able to pump more than 5,000 gallons of water a minute (about 22,000 liters a minute) from the Nile onto adjacent cotton fields. Shuman was an enthusiast and advocate for solar energy and in 1916 was quoted in The New York Times as saying: We have proved the commercial profit of sun power . . . and have more particularly proved that after our stores of oil and coal are exhausted the human race can receive unlimited power from the rays of the sun.

September 26, 2023

Regardless of whether one believes in the innovative capacity of human beings to solve the problems of nuclear energy, whether fusion or fission, or the challenge of affordable and reliable solar technology sufficient to support the energy needs of 10 billion people, or to reverse the amount of carbon we are pumping into our atmosphere, we are still left with the long-term problem of entropy production. Apart from its many other issues, the nuclear option, like that of traditional fossil fuels, keeps us trapped in the paradigm of a closed system, whereas the solar option has the critical capacity for potentially returning us to a truly sustainable paradigm of an open system. 293

September 26, 2023

Furthermore, the slope of the straight line, which is the exponent of the power law, is about 0.85, a little bit higher than the 0.75 (the famous ¾) we saw for the metabolic rate of organisms (Figure 1). Equally intriguing is that this exponent takes on approximately the same value for how gasoline stations scale across all of the countries shown in the figure. This value of around 0.85 is smaller than 1, so in the language developed earlier, the scaling is sublinear , indicating a systematic economy of scale , meaning that the bigger the city the fewer the number of gas stations needed on a per capita basis. Thus, on average, each gas station in a 322 *larger city serves more people and consequently sells more fuel per month than in a smaller one.*

September 29, 2023

Luis was joined in this endeavor by another very bright recruit, José Lobo, an urban economist now in the sustainability program at Arizona State University. When we first met, José was a young faculty member in the Department of City and Regional Planning at Cornell University and had been coming to SFI for several years. Like Luis, José brought a real talent for statistics and sophisticated data analysis to our program and, in addition, brought professional expertise in cities and urbanization, a critical component of our collaboration. 324

September 29, 2023

Luis and José took the lead in assembling and analyzing extensive data sets covering a broad range of metrics about urban systems across the globe, ranging from Spain and the Netherlands in Europe to Japan and China in Asia and Colombia and Brazil in Latin America. This convincingly verified the earlier analyses showing sublinear scaling of infrastructural metrics and strongly supported the universality of systematic economies of scale in cities. Regardless of the specific urban system, whether Japan, the United States, or Portugal, and regardless of the specific metric whether the number of gas stations, the total length of pipes, roads, or electrical wires, only about 85 percent more material infrastructure is needed with every doubling of city size. 2 Thus a city of 10 million people typically needs 15 percent less of the same infrastructure compared with two cities of 5 million each, leading to significant savings in materials and energy use. 3 325

September 29, 2023

So what is the common unifying factor that transcends these differences and underlies this surprising structural and dynamical similarity? I have already strongly hinted at the answer: the great commonality is the universality of social network structures across the globe. Cities are people, and to a large extent people are pretty much the same all over the world in how they interact with one another and how they cluster to form groups and communities. We may look different, dress differently, speak different languages, and have different belief systems, but to a large extent our biological and social organization and dynamics are remarkably similar. After all, we are all human beings sharing pretty much the same genes and the same generic social history. And no matter where we live on the planet, all of us emerged only relatively recently from being mobile hunter-gatherers to becoming predominantly sedentary communal creatures. The underlying commonality that is being expressed by the surprising universality of urban scaling laws is that the structure and dynamics of human social networks are very much the same everywhere. 333

September 29, 2023

Recall that the generic geometric and dynamical properties of biological networks that underlie quarter-power allometric scaling are: (1) they are space filling (so every cell of an organism, for instance, must be serviced by the network); (2) the terminal units, such as capillaries or cells, are invariant within a given design (so, for instance, our cells and capillaries are approximately the same as those of mice and whales); and (3) the networks have evolved to be approximately optimal (so, for instance, the energy our hearts have to use to circulate blood and support our cells is minimized in order to maximize the energy available for reproduction and the rearing of offspring). These properties have direct analogs in the infrastructural networks of cities. For example, our road and transport networks have to be space filling so that every local region of the city is serviced, just as all of the various utility lines have to supply water, gas, and electricity to all of its houses and buildings. It's also natural to extend this concept to social networks: averaged over time, each person interacts with a number of other 336 *people as well as with groups of people in the city in such a way that collectively their network of interactions fills the available "socioeconomic space." Indeed, this urban network of socioeconomic interactions constitutes the cauldron of social activity and interconnectivity that effectively defines what a city is and what its boundaries are. To be part of a city you have to be an ongoing participant in this network. And, of course, the invariant terminal units of these networks, the analogs of capillaries, cells, leaves, and petioles, are people and their houses.*

September 29, 2023

The leading advocate for developing the concept of the fractal city and integrating ideas from complexity theory into traditional urban analysis and planning has been Mike Batty, who runs the Centre for Advanced Spatial Analysis (CASA) at University College London. His work has focused primarily on computer models of the physicality of cities and urban systems. He is enthusiastic about the concept of cities as complex adaptive systems and has consequently become a major proponent of developing a science of cities . His vision is a little different from mine and is summarized in his recent book The New Science of Cities, which emphasizes the more phenomenological traditions of the social sciences, geography, and urban planning as against the more analytic, mathematical traditions of physics based on underlying principles that I've been articulating. 7 Ultimately, both approaches are needed if we are to accomplish the huge challenge of understanding cities. 348

September 29, 2023

A city is not just the aggregate sum of its roads, buildings, pipes, and wires that comprise its physical infrastructure, nor is it just the cumulative sum of the lives and interactions of all of its citizens, but rather it's the amalgamation of all of these into a vibrant, multidimensional living entity. A city is an emergent complex adaptive system resulting from the integration of the flows of energy and resources that sustain 349 *and grow both its physical infrastructure and its inhabitants with the flows and exchange of information in the social networks that interconnect all of its citizenry. The integration and interplay of these two very different networks magically gives rise to increasing economies of scale in its physical infrastructure and simultaneously to a disproportionate increase in social activity, innovation, and economic output.*

September 29, 2023

Over the past twenty years an emerging subfield of network science has blossomed in its own right, leading to a deeper understanding of both the phenomenology of networks in general and also the underlying mechanisms and dynamics that generate them. 8 The subject of network science covers an enormous range of topics including classic community organizations, criminal and terrorist networks, networks of innovation, ecological networks and food webs, health care and disease networks, and linguistic and literary networks. Such studies have provided important insights into a broad range of important societal challenges including devising the most effective strategies for attacking pandemics, terrorist organizations, and environmental issues, for enhancing and facilitating innovative processes, and for optimizing social organizations. A lot of this fascinating work has been carried out or stimulated by many of my colleagues associated with the Santa Fe Institute. 350

September 29, 2023

Steve Strogatz is an eclectic applied mathematician at Cornell University who uses ideas from nonlinear dynamics and complexity theory to analyze and explain a broad range of fascinating problems. For example, he has done some lovely work showing how crickets, cicadas, and fireflies synchronize their behaviors and more recently extended it to show why London's Millennium Bridge was dysfunctional. 11 This latter problem has some interesting lessons for the science of cities, and I want to digress to explain it. 352

September 29, 2023

The provocative work of Milgram and Zimbardo strongly suggests that the conundrum as to why good people can do very bad things originates in peer pressure situations, fear of rejection, and a desire to be part of a group where power and control are conferred on individuals by authority. Zimbardo has become an articulate and vocal advocate for the recognition that this powerful dynamic, which seems to be built into our psyches independent of cultural origins and which has wreaked horrors over the centuries, be explicitly recognized and addressed rather than resorting to our instinctual tendency to put the blame simplistically on individual "bad apples," national characteristics, or cultural norms.

357

September 29, 2023

You will notice that the sequence of numbers that quantify the magnitude of these successive levels of the group hierarchy—5, 15, 50, 150—are sequentially related to each other by an approximately constant scaling factor of about three. This regularity is the familiar fractal-like pattern we saw not only in the network hierarchy of our own circulatory and respiratory systems but also in transport patterns in cities. In addition to the actual flows in these networks, the major geometric difference between them is the value of the branching ratio— the number of units, people in this case, at one level of the hierarchy relative to the next. There is evidence that in social networks this pattern with a branching ratio of three persists beyond the 150 level to groups having sizes of approximately 500, 1,500, and so on. The precise value of these numbers should not be taken too seriously because there is significant variance in the data. The important point for our purposes is that viewed through a coarse-grained lens, social networks exhibit an approximate fractal-like pattern and that this seems to hold true across a broad spectrum of different social organizations. Even though this pattern remains approximately static, the individual members of the network

362

may change with time or may move from one level to another as you become closer or more distant in your relationship with them. For instance, a parent may move out of your inner circle and be replaced by a spouse or a close friend, or you might meet someone casually at a party who subsequently becomes part of your 150. Regardless of such changes, the generic structure of the network with four to six people comprising your core group, and the nested group structure whose discrete size increases by factors of about three up to 150 or so, remains intact.

September 29, 2023

The number of around 150 represents the maximum number of individuals that a typical person can still keep track of and consider casual friends and therefore members of his or her ongoing social network. Consequently, this is the approximate size of a group in which all individuals still know one another sufficiently well for the group to remain coherent and maintain ongoing social relationships. Dunbar found many such examples of functioning social units whose size congregated around this magic number, ranging from bands of hunter-gatherers to army companies in the Roman Empire, in sixteenth-century Spain, and in the twentieth-century Soviet Union. He speculated that this apparent universality has its origins in the evolution of the cognitive structure of the brain: we simply do not have the computational capacity to manage social relationships effectively beyond this size. This suggests that increasing the group size beyond this number will result in significantly less social stability, coherence, and connectivity, ultimately leading to its disintegration. For situations where group identity and cohesiveness are perceived as central for the group to function successfully, recognizing this limitation and the broader implications of social network structure is clearly of great importance. This is especially true in situations where stability, knowledge of other individuals, and social relationships are integral to performance. Companies, the military, government administrations and bureaucracies, universities, and research organizations are among the many institutions where this kind of information and this way of thinking could potentially be beneficial in 364 *improving performance, productivity, and the general well-being of all members of the organization.*

September 29, 2023

connection between brain size and the ability to form social groups is called the social brain hypothesis. Dunbar went much further by suggesting that this relationship is causal in that human intelligence evolved primarily as a response to the challenge of forming large and complex social groups rather than the usual explanation that it is a direct consequence of meeting ecological challenges. 16 Regardless of causation, he used the correlation with brain size to estimate the number 150 as the idealized size of analogous human social groups.

365

September 30, 2023

The lesson is clear: the number of links between people increases much faster than the increase in the number of people in the group and, to a very good approximation, is given by just one half of the square of the number of people in the group .

376

September 30, 2023

This simple nonlinear quadratic relationship between the maximal number of links between people and the size of the group has all sorts of very interesting social consequences. For instance, my wife, Jacqueline, particularly enjoys dinner parties if a single conversation can be sustained by the entire group, so she is reluctant to participate in dinner parties larger than six. With six people there are 6 × 5 ÷ 2 = 15 possible pair-wise independent conversations that have to be "suppressed" for a single collective one to emerge and be maintained. This is just about possible, and it's tempting to speculate that it's because the number of other guests, five, corresponds to Dunbar's number for the group size of an

376

average individual's inner circle. With ten at the table, there are a whopping forty-five such dyadic possibilities, which inevitably leads to a balkanization of the group as it disintegrates into two, three, or more separate conversations. Of course, many people prefer this modality but it is worth remembering that if you want a certain kind of group intimacy, having more than about six people is going to make it quite a challenge.

Dunbar's number 5.

September 30, 2023

When these constraints on the mobility and physical interaction space of people in cities are imposed on the structure of social networks, an important and far-reaching result emerges: the number of interactions with other people in a city that an average person maintains scales inversely to the way that the degree of infrastructure scales with city size. In other words, the degree to which the scaling of infrastructure and energy use is sublinear is predicted to be the same as the degree to which the scaling of the number of an average individual's social interactions is superlinear . Consequently, the exponent controlling social interactions, and therefore all socioeconomic metrics—the universal 15 percent rule for how the good, the bad, and the ugly scale with city size—is bigger than 1 (1.15) to the same degree that the exponent controlling infrastructure and flows of energy and resources is less than 1 (0.85), as observed in the data. Pictorially, the degree to which all of the slopes in Figures 34–38 exceed 1 is the same as the degree to which they are less than 1 in Figure 33. 381

In other words, the degree to which the scaling of infrastructure and energy use is sublinear is predicted to be the same as the degree to which the scaling of the number of an average individual's social interactions is superlinear.

September 30, 2023

395

Zahavi's fascinating observations made a powerful impression on the Italian physicist Cesare Marchetti, who was a senior scientist at the International Institute for Applied Systems Analysis (IIASA) in Vienna. IIASA has been a major player in questions of global climate change, environmental impact, and economic sustainability, and this is where Marchetti's interests and contributions have mostly been. He became intrigued by Zahavi's work and in 1994 published an extensive paper elaborating on the approximate invariance of daily commute times and promoted the idea that the true invariant is actually overall daily travel time, which he called exposure time . 3 So even if an individual's daily commute time is less than an hour, then he or she instinctively makes up for it by other activities such as a daily constitutional walk or jog. In support of this Marchetti wryly remarked, "Even people in prison for a life sentence, having nothing to do and nowhere to go, walk around for one hour a day, in the open." Because walking speed is about 5 kilometers an hour, the typical extent of a "walking city" is about 5 kilometers across (about 3 miles), corresponding to an area of about 20 square kilometers (about 7 square miles). According to Marchetti, "There are no city walls of large, ancient cities (up to 1800), be it Rome or Persepolis, which have a diameter greater than 5km or a 2.5km radius. Even Venice today , still a pedestrian city, has exactly 5km as the maximum dimension of the connected center. " With the introduction of horse tramways and buses, electric and steam trains, and ultimately automobiles, the size of cities could grow but, according to Marchetti, constrained by the one-hour rule. With cars able to travel at 40 kilometers an hour (25 mph), cities, and more generally metropolitan areas, could expand to as much as 40 kilometers or 25 miles across, which is typical of the catchment area for most large cities. This corresponds to an area of about 12 hectares or 450 square miles, more than fifty times the area of a walking city. This surprising observation of the approximately one-hour invariant that communal human beings have spent traveling each day, whether they lived in ancient Rome, a medieval town, a Greek village, or twentieth-century New York, has become known as Marchetti's constant , even though it was originally discovered by Zahavi. As a rough guide it clearly has important implications for the design and structure of cities. As planners begin to design green carless *communities and as more cities ban automobiles from their centers, understanding and implementing the implied constraints of Marchetti's constant becomes an important consideration for maintaining the functionality of the city.*

October 1, 2023

People across the globe behave in pretty much the same way regardless of history, culture, and geography. Thus without having to resort to any fancy mathematical theory, this idea predicts that the number of interactions between people in cities should scale with city size in the same way that all of the diverse socioeconomic quantities scale, namely, as a superlinear power law with an exponent of around 1.15 regardless of the urban system. In other words, the systematic 15 percent increase in socioeconomic activity with every doubling of city size, whether in wages and patent production or crime and disease, should track a predicted 15 percent increase in the interaction between people. 404

October 1, 2023

The image of the city as a large vat in which people are being continually churned, blended, and agitated can be viscerally felt in any of the world's great cities. It is most apparent in the continuous, and sometimes frenetic, movement of people in downtown and commercial areas in what often appears to be an almost random motion much like molecules in a gas or liquid. And in much the same way that bulk properties of gases or liquids, such as their temperature, pressure, color, and smell, result from molecular collisions and chemical reactions, so the properties of cities emerge from social collisions and the chemistry of and between people. Metaphors can be useful but sometimes they can be misleading, and this is one of those instances. For despite appearances, the motion of people in cities is not at all like the random motion of molecules in a gas or particles in a reactor. Instead, it is overwhelmingly systematic and directed. Very few journeys are random. Almost all, regardless of the form of conveyance, involve willful travel 410 *from one specific place to another: mostly from home to work, to a store, to a school or cinema, and so forth . . . and back again. Furthermore, most travelers seek the fastest and shortest route, one that takes the least time and traverses the shortest distance. Ideally, this would mean that everyone would like to travel along straight lines but, given the obvious physical constraints of cities, this is impossible. There is no choice but to follow the meandering roads and rail lines, so, in general, any specific journey involves following a zigzagging route. However, when viewed at a larger scale through a coarse-grained lens by averaging over all journeys for all people over a long enough period of time, the preferred route between any two specific locations approximates a straight line. Loosely speaking this means that on average people effectively travel radially along the spokes of circles whose center is their specific destination, which acts as a hub. With this assumption it is possible to derive an extremely simple but very powerful mathematical result for the movement of people in cities. Here's what it says: Consider any location in a city; this could be a "central place" such as a downtown area or street, a shopping mall or district, but it could just as well be some arbitrary residential area such as where you live. The mathematical theorem predicts how many people visit this location from any distance away and how often they do it. More specifically, it states that the number of visitors should scale inversely as the square of both the distance traveled and the frequency of visitation .*

October 1, 2023

Mathematically, an inverse square law is just a simple version of the kinds of power law scaling we've been discussing throughout the book. In that language the prediction of movement in cities can be restated as saying that the number of people traveling to a specific location scales with both the distance traveled and the visitation frequency as a power law whose exponent is −2. Thus if the number of travelers is 411 *plotted logarithmically against either the distance traveled keeping the visitation frequency fixed, or vice versa, against the visitation frequency keeping the travel distance fixed, a straight line should result in both cases with the same slope of −2 (recall that the minus sign simply means that the line slopes downward). I should emphasize that as with all of the scaling laws, an average over a long enough period of time, such as six months or a year, is presumed, in order to smooth over daily fluctuations, or the differences between weekdays and weekends. As can be readily seen from Figure 47, these predictions are spectacularly confirmed by the data. Indeed, the observed scaling is remarkably tight, with slopes in excellent agreement with the prediction of −2. Particularly satisfying is that the same predicted inverse square law is observed across the globe in diverse cities with very different cultures, geographies, and degrees of development: we see an identical behavior in North America (Boston), Asia (Singapore), Europe (Lisbon), and Africa (Doha). Furthermore, when each of these metropolitan areas is deconstructed into specific locations, the same inverse square law is manifested at each of these within the city, as shown, for example, in Figures 48 and 49 for a sampling of locations in both Boston and Singapore. Let me give a simple example to illustrate how the theorem works. Suppose that on average 1,600 people visit the area around Park Street, Boston, from 4 kilometers away once a month. How many people visit there from twice as far away (8 km) with the same frequency of once a month? The inverse square law tells us that* $\frac{1}{4}$ *(=* $\frac{1}{2}$*2) as many make the visit, so only 400 people (*$\frac{1}{4} \times 1{,}600$*) visit Park*

Street from 8 kilometers away once a month. How about from five times as far away, 20 kilometers? The answer is 1/25 (= 1/5 2) as many, which is just 64 people (1/25 × 1,600) visiting once a month. You get the idea. But there's more: you can likewise ask what happens if you change the frequency of visitation. For instance, suppose we ask how many people visit Park Street from 4 kilometers away but now with a greater frequency of twice a month. This also obeys the inverse square law so the number is $\frac{1}{4}$ (= $\frac{1}{2}$2) as many, namely 400 people. And similarly, if you ask how many people visit there from the same distance of 4 kilometers away but five times a month, the answer is 64 people (1/25 × 1,600). Notice that this is the same number that visit Park Street from five times as far away (20 km) with a frequency of just once a month. Thus the number of people visiting from 4 kilometers away five times a month is the same as the number visiting from five times farther away (20 km) once a month (64 in both cases of our specific example). This result does not depend on the specific numbers I chose for the illustration. It is an example of an amazing general symmetry of mobility: if the distance traveled multiplied by the frequency of visits to any specific location is kept the same, then the number of people visiting also remains the same . In our example we had 4 kilometers × 5 times a month = 20 in the first case and 20 kilometers × once a month = 20 in the second. This invariance is valid for any visiting distance and for any frequency of visitation to any area in any city. These predictions are verified by the data and manifested in the various graphs of Figures 48 and 49 where you can explicitly see that the pattern of visitation remains unchanged when the product of distance times frequency has the *same value.*

In this paragraph exist very important information about how busy people is into determine area from a distance away.

October 1, 2023

My colleagues Luis, José, and Debbie carried out such an analysis for the entire U.S. urban system consisting of 360 Metropolitan Statistical Areas (MSAs) for a suite of metrics. 7 A sample of the results is presented in Figure 50), where the deviations from scaling for personal income and patent production for cities in the United States in 2003 are plotted logarithmically on the vertical axis against the rank order of each city. We called these deviations Scale-Adjusted Metropolitan Indicators (SAMIs). The horizontal axis across the center of these graphs is the line along which the SAMI is zero and there is no deviation from what is predicted from the size of the city. As can be seen, every city deviates to some extent from its expected values. Those to the left denote above -average performance, whereas those to the right denote below -average performance. This provides a meaningful ranking of a city's individuality and uniqueness beyond what is effectively guaranteed just because it's a city of a certain size. Without delving into details of this analysis, I want to make a few salient points about some of the results. First, compared with conventional per capita indicators, which place seven of the largest twenty cities in the top twenty in terms of their GDPs, our science-based metrics rank none of these cities in the top twenty. In other words, once the data are adjusted for the generic superlinear effects of population size, these cities don't fare so well. Mayors of these cities who take credit and boast that their policies have led to economic success as evidenced by their city's being near the top of the per capita GDP rankings are therefore giving a misleading impression. 422

October 2, 2023

Once a city has gained an advantage, or disadvantage, 425 *relative to its scaling expectation, this tends to be preserved over decades. In this sense, either for good or for bad, cities are remarkably robust and resilient—they are hard to change and almost impossible to kill. Think of Detroit and New Orleans, and more drastically of Dresden, Hiroshima, and Nagasaki, all of which have to varying degrees survived what were perceived as major threats to their very existence. All are actually doing fine and will be around for a very long time. A fascinating example of persistent advantage is San Jose, which includes Silicon Valley, the place everyone wants to be. It's hardly a surprise that this is a major overperformer in terms of wealth creation and innovation. But what is a surprise is that San Jose was already overperforming in the 1960s, and almost to the same degree as it is now, as graphically illustrated in Figure 51. This also demonstrates that this overperformance has been sustained and even reinforced for more than forty years despite the short-term boom and bust technological and economic cycle in 1999–2000, at the end of which the city relaxed back to its long-term basal trend. Put slightly differently: apart from a relatively small bump in the late 1990s, the continued success of San Jose was already set well before the birth of Silicon Valley. So rather than seeing Silicon Valley as generating the success of San Jose and lifting it up in the conventional socioeconomic rankings, this suggests that it was the other way around and that it was some intangible in the culture and DNA of San Jose that helped nurture the extraordinary success of Silicon Valley. 8 It takes decades for significant change to be realized. This has serious implications for urban policy and leadership because the timescale of political processes by which decisions about a city's future are made is at best just a few years, and for most politicians two years is infinity. Nowadays, their success depends on rapid returns and instant gratification in order to conform to political pressures and the demands of the electoral process. Very few mayors can afford to think in a time frame of twenty to fifty years and put their major efforts toward promoting strategies that will leave a truly long-term legacy of significant achievement.*

October 2, 2023

Safe water is progressively becoming a source of increased social friction, especially as the climate changes and produces unpredictable periods of severe drought or massive floods, both of which compromise supply and delivery systems. This is already a major issue in many developing countries and hints of it have begun to be seen even in the United States with serious problems arising in supply systems, as in Flint, Michigan, and severe water shortages occurring throughout many of the western states.

426

October 2, 2023

Thus the energy available for growth is just the difference between the rate at which energy can be supplied and the rate that is needed for maintenance. On the supply side, metabolic rate in organisms scales sublinearly with the number of cells (following the generic ¾ power exponent derived from network constraints) while the demand increases approximately linearly . So as the organism increases in size, demand eventually outstrips supply because linear scaling grows faster than sublinear, with the consequence that the amount of energy available for growth continuously decreases, eventually going to zero resulting in the cessation of growth. In other words, growth stops because of the mismatch between the way maintenance and supply scale as size increases. The sublinear scaling of metabolic rate and the associated economies of scale arising from optimizing network performance are therefore responsible for why growth stops and why biological systems exhibit the bounded sigmoidal growth curves that were shown in Figures 15–18 of chapter 4. The same network mechanism that underlies sublinear scaling, economies of scale, and the cessation of growth is also responsible for the systematic slowing down of the pace of biological life as size increases—and for eventual death.

440

October 2, 2023

The mechanisms that have traditionally been suggested for understanding companies can be divided into three broad

450

categories: transaction costs, organizational structure, and competition in the marketplace. Although these are interrelated they have very often been treated separately. In the language of the framework developed in previous chapters these can be expressed as follows: (1) Minimizing transaction costs reflects economies of scale driven by an optimization principle, such as maximizing profits. (2) Organizational structure is the network system within a company that conveys information, resources, and capital to support, sustain, and grow the enterprise. (3) Competition results in the evolutionary pressures and selection processes inherent in the ecology of the marketplace.

October 2, 2023

Axtell, who is at George Mason University in Virginia and is also on the external faculty of the Santa Fe Institute, is a leading expert in agent-based modeling, which is a computational technique used for simulating systems composed of huge numbers of components. 3 Basically, the strategy involves postulating simple rules governing the interactions between individual constituent agents, which could be companies, cities, or people, coupled with an algorithm that specifies how they evolve in time and letting the resulting system run on a computer. More sophisticated versions include rules for learning, adaptation, and even reproduction so as to model more realistic evolutionary processes. 451

October 2, 2023

Nassim Taleb, author of the best-selling, highly influential book The Black Swan, has been particularly harsh on economists despite, or maybe because of, having been trained in business and finance. 5 He has held positions at several distinguished universities including New York University and Oxford and has focused on the importance of coming to terms with outlying events and developing a deeper understanding of risk. He has been brutally outspoken in his condemnation of classical economic thinking 452 *with hyperbolic comments such as: "Years ago, I noticed one thing about economics, and that is that economists didn't get anything right." He has even called for the Nobel Memorial Prize in economics to be withdrawn, saying that the damage from economic theories can be devastating. I may disagree with some of Taleb's ideas and polemics but it's important and healthy to have such outspoken mavericks challenging the orthodoxy, especially when it's had such a poor record and its proclamations have major implications for our lives.*

October 2, 2023

The initial results and conclusions of our investigation into the scaling of companies are very compelling. They provide a powerful basis for developing an understanding of their generic structure and life histories. Figures 60–63 show the sales, incomes, and assets of all 28,853 companies plotted logarithmically against their number of employees. These are the dominant financial characteristics of any company and are standard measures of their fiscal health and dynamics. As these graphs clearly demonstrate, companies do indeed scale following simple power laws and as anticipated they do so with a much greater spread around their average behavior than for either cities or organisms. So in this statistical sense, companies are approximately scaled, self-similar versions of one another: Walmart is an approximately scaled-up version of a much smaller, modest-size company. Even after taking this greater variance into account, this scaling result reveals remarkable regularities in the size and dynamics of companies and is quite surprising given the tremendous variety of different business sectors, locations, and age. 457

October 3, 2023

A crucial aspect of the scaling of companies is that many of their key metrics scale sublinearly like organisms rather than superlinearly like cities. This suggests that companies are more like organisms than cities and are dominated by a version of economies of scale rather than by increasing 461

returns and innovation.

October 3, 2023

This is really quite amazing. After all, when we think of the birth, death, and general life history of companies as they struggle to establish and maintain themselves in the marketplace dealing with the vagaries, uncertainties, and unpredictability of economic life and the myriad specific decisions and accidents that led to the successes and failures that preceded their death, it's hard to believe that collectively they were following such simple general rules. This revelation echoes the surprise that organisms, ecosystems, and cities are likewise subject to generic constraints, despite the apparent uniqueness and individuality of their life histories. 472

October 4, 2023

A more complete analysis therefore needs to include these so-called censored companies, whose life spans are at least as long as and likely longer than the period over which they appear in the data set. This actually involves a sizable number of companies: in the sixty years covered, 6,873 firms were still alive at the end of the window in 2009. Fortunately, there is a well-established sophisticated methodology, called survival analysis, that has been developed precisely for addressing this issue. Survival analysis was developed in medicine in order to estimate survival probabilities for patients who have undergone therapeutic interventions under test conditions. These tests have necessarily to be conducted over a limited time period, leading to the problem we face here, namely, that many subjects die after the test period has ended. The technique commonly used, called the Kaplan-Meier estimator, employs the entire data set and optimizes probabilities assuming that each death event is statistically independent of every other death. 7 473

October 4, 2023

Although there are significant differences, it's hard not to be struck by how similar the growth and death of companies and organisms are when viewed through the lens of scaling—and how dissimilar they both are to cities. Companies are surprisingly biological and from an evolutionary perspective their mortality is an important ingredient for generating innovative vitality resulting from "creative destruction" and "the survival of the fittest." Just as all organisms must die in order that the new and novel may blossom, so it is that all companies disappear or morph to allow new innovative variations to flourish: better to have the excitement and innovation of a Google or Tesla than the stagnation of a geriatric IBM or General Motors. This is the underlying culture of the free market system. 475

October 4, 2023

The great turnover of companies and especially the continual churning of mergers and acquisitions are integral to the market process. And of course this means that the Googles and Teslas, which may seem invincible now, will themselves eventually fade away and disappear. From this point of view we should not lament the passing of any company—it's an essential component of economic life—we should only mourn and be concerned about the fate of the people who often suffer when companies disappear, whether they are the workers, management, or even the owners. If only we could tame the potential brutality and greed of the survival of the fittest and soften some its more egregious consequences by formulating a magic algorithm for how to balance the classic tension between regulation, government intervention, and uncontrolled rampant capitalism. This struggle painfully played itself out as we witnessed the struggle between the death throes of corporations that probably should have died and the desire to save jobs and protect the lives of workers because certain incompetent if not duplicitous corporations were deemed "too big to fail" during the 2008 financial crisis. 475

If only we could tame the potential brutality and greed of the survival of the fittest and soften some its more egregious consequences by formulating a magic algorithm for how to balance the classic tension between regulation, government intervention, and uncontrolled rampant capitalism.

October 4, 2023

There are some wonderful examples of these geriatric survivors. For instance, the oldest shoemaker in Germany is the Eduard Meier company, founded in Munich in 1596, which became purveyors to the Bavarian aristocracy. It still has only a single store that sells, though no longer makes, quality upscale shoes. The oldest hotel in the world according to Guinness World Records is Nishiyama Onsen Keiunkan in Hayakawa, Japan, which was founded in 705. It has been in the same family for fifty-two generations and even in its modern incarnation has only thirty-seven rooms. Its main attraction seems to be its hot springs. The world's oldest company was purported to be Kongo Gumi, founded in Osaka, Japan, in 578. It was also a family business going back many generations, but after almost 1,500 years of continuously being in business it went into liquidation in 2006 and was purchased by the Takamatsu Corporation. And what was the niche market that Kongo Gumi cornered for 1,429 years? Building beautiful Buddhist temples. But sadly, with the changes in Japanese culture following the Second World War the demand for temples dried up and Kongo Gumi was unable to adapt fast enough. 479

October 4, 2023

economies of scale over innovation and idea creation. Companies typically operate as highly constrained top-down organizations that strive to increase efficiency of production and minimize operational costs so as to maximize profits. In contrast, cities embody the triumph of innovation over the hegemony of economies of scale. Cities aren't, of course, driven by a profit motive and have the luxury of being able to balance their books by raising taxes. They operate in a much more distributed fashion, with power spread across multiple 481 *organizational structures from mayors and councils to businesses and citizen action groups. No single group has absolute control. As such, they exude an almost laissez-faire, freewheeling ambience relative to companies, taking advantage of the innovative benefits of social interactions whether good, bad, or ugly. Despite their apparent bumbling inefficiencies, cities are places of action and agents of change relative to companies, which by and large usually project an image of stasis unless they are young.*

October 4, 2023

One of the major challenges of the twenty-first century that will have to be faced is the fundamental question as to whether human-engineered social systems, from economies to cities, which have only existed for the past five thousand years or so, can continue to coexist with the "natural" biological world from which they emerged and which has been around for several billion years. To sustain more than 10 billion people living in harmony with the biosphere at a standard of living and quality of life comparable to what we now have requires that we develop a deep understanding of the principles and underlying system dynamics of this social-environmental coupling. I have argued that a critical component of this is to develop a deeper understanding of cities and urbanization. Continuing to pursue limited and single-system approaches to the many problems we face without developing a unifying framework risks the possibility that we will squander huge financial and social capital and fail miserably in addressing the really big question, resulting in dire consequences. 484

October 4, 2023

We need a broad and more integrated scientific framework that encompasses a quantitative, predictive, mechanistic theory for understanding the relationship between human-engineered systems, both social and physical, and the "natural" environment—a framework I call a grand unified theory of sustainability. It's time to initiate a massive 485

international Manhattan-style project or Apollo-style program dedicated to addressing global sustainability in an integrated, systemic sense. 1

October 4, 2023

So how can such a collapse be avoided, and can it be achieved while still ensuring open-ended growth? The first point to appreciate is that these predictions assume that the parameters of the growth equation do not change. So one clear strategy for forestalling a potential catastrophe is to intervene before reaching the singularity by "resetting" the parameters. Moreover, to maintain open-ended growth with these new settings requires that the driving term in the equation—the "social metabolism"—needs to remain superlinear, meaning that the new dynamic must still be driven by the positive feedback forces of social interaction responsible for innovation, and for wealth and knowledge creation. Such an "intervention" is none other than what is usually referred to as an innovation . A major innovation effectively resets the clock by changing the conditions under which the system has been operating and growth occurring. Thus, to avoid collapse a new innovation must be initiated that resets the clock, allowing growth to continue and the impending singularity to be avoided . 489

October 4, 2023

Unfortunately, however, it's not quite as simple as that. There's yet another major catch, and it's a big one. The theory dictates that to sustain continuous growth the time between successive innovations has to get shorter and shorter. Thus paradigm-shifting discoveries, adaptations, and innovations must occur at an increasingly accelerated pace. Not only does the general pace of life inevitably quicken, but we must innovate at a faster and faster rate! 492

October 4, 2023

Just as the Malthusians ignored the crucial role of innovation, the singularity enthusiasts ignore the crucial role of the entire socioeconomic dynamic of the planet, which in fact is the prime driver of the impending singularity. Neither case is anchored in a broader framework that embraces a quantitative mechanistic theory, so whatever their predictions are it's very hard to evaluate them scientifically. Perhaps the greatest conceptual irony, especially of the singularists, is that their conclusions and speculations are based on exponential growth, which doesn't actually lead to a singularity, at least not in a finite time .

497

October 4, 2023

The great John von Neumann, mathematician, physicist, computer scientist, and polymath, a man whose ideas and accomplishments have had a huge influence on your life, made the following remarkably prescient observation more than seventy years ago: "The ever accelerating progress of technology and changes in the mode of human life . . . gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue." 7 Among von Neumann's many accomplishments before he died at the relatively young age of fifty-three in 1957 are his seminal role in the early development of quantum mechanics, his invention of game theory, which is a major tool in economic modeling, and the conceptual design of modern computers universally referred to as the von Neumann architecture.

499

October 4, 2023

The institute was originally conceived by a small group of distinguished scientists, including several Nobel laureates, most of whom had some association with Los Alamos National Laboratory. They were concerned that the academic landscape had become so dominated by disciplinary stovepiping and specialization that many of the big questions, and especially those that transcend disciplines

510

or were perhaps of a societal nature, were being ignored. The reward system for obtaining an academic position, for gaining promotion or tenure, for securing grants from federal agencies or private foundations, and even for being elected to a national academy, was becoming more and more tied to demonstrating that you were the expert in some tiny corner of some narrow subdiscipline. The freedom to think or speculate about some of the bigger questions and broader issues, to take a risk or be a maverick, was not a luxury many could afford.

October 5, 2023

Data for data's sake, or the mindless gathering of big data, without any conceptual framework for organizing and understanding it, may actually be bad or even dangerous. Just relying on data alone, or even mathematical fits to data, without having some deeper understanding of the underlying mechanism is potentially deceiving and may well lead to erroneous conclusions and unintended consequences. This admonition is closely related to the classic warning that "correlation does not imply causation." Just because two sets of data are closely correlated does not imply that one is the cause of the other. There are many bizarre examples that illustrate this point. 4 For instance, over the eleven-year period from 1999 to 2010 the variation in the total spending on science, space, and technology in the United States almost exactly followed the variation in the number of suicides by hanging, strangulation, and suffocation. It's extremely unlikely that there is any causal connection between these two phenomena: the decrease in spending in science was surely not the cause of the decrease in how many people hanged themselves. However, in many situations such a clear-cut conclusion is not so clear. More generally, correlation is in fact often an important indication of a causal connection but usually it can only be established after further investigation and the development of a mechanistic model. 517

October 5, 2023

A contrary view to this trend was forcibly expressed by the Nobel Prize–winning geneticist Sydney Brenner, whom I quoted in chapter 3 and who was coincidentally director of the famous institute in Cambridge founded by Max Perutz that I mentioned earlier: "Biological research is in crisis. . . . Technology gives us the tools to analyse organisms at all scales, but we are drowning in a sea of data and thirsting for some theoretical framework with which to understand it. Although many believe that 'more is better,' history tells us that 'least is best.' We need theory and a firm grasp on the nature of the objects we study to predict the rest." Not long after the publication of Chris Anderson's article, Microsoft published a fascinating series of essays in a book titled The Fourth Paradigm: Data-Intensive Scientific Discovery. It was inspired by Jim Gray, a computer scientist at Microsoft who was sadly lost at sea in 2007. He envisioned the data revolution as a major paradigm shift in how science would advance in the twenty-first century and called it the fourth paradigm.

520

October 5, 2023

The more one can bring big data into the enterprise the better, provided it is constrained by a bigger-picture conceptual framework that, in particular, can be used to judge the relevance of correlations and their relationship to mechanistic causation. If we are not to "drown in a sea of data" we need a "theoretical framework with which to understand it . . . and a firm grasp on the nature of the objects we study to predict the rest."

525

October 5, 2023

Green

In biology these are controlled and maintained by the process of metabolism. Quantitatively, this is expressed in terms of metabolic rate, which is the amount of energy needed per second to keep an organism alive; for us it's about 2,000 food calories a day, which, surprisingly, corresponds to a rate of only about 90 watts, the equivalent of a standard incandescent lightbulb. 23

September 17, 2023

As social animals now living in cities we still need just a lightbulb equivalent of food to stay alive but, in addition, we now require homes, heating, lighting, automobiles, roads, airplanes, computers, and so on. Consequently, the amount of energy needed to support an average person living in the United States has risen to an astounding 11,000 watts. This social metabolic rate is equivalent to the entire needs of about a dozen elephants. Furthermore, in making this transition from the biological to the social our overall population has increased from just a few million to more than seven billion. No wonder there's a looming energy and resource crisis. 23

September 17, 2023

There is always a price to pay when energy is processed; there is no free lunch. Because energy underlies the transformation and operation of literally everything, no system operates without consequences. Indeed, there is a fundamental law of nature that cannot be transgressed, called the Second Law of Thermodynamics , which says that whenever energy is transformed into a useful form, it also produces "useless" energy as a degraded by-product: "unintended consequences" in the form of inaccessible disorganized heat or unusable products are inevitable. There are no perpetual motion machines. You need to eat to stay alive and maintain and service the highly organized functionality of your mind and body. But after you've eaten, sooner or later you will have to go to the bathroom. This is the physical manifestation of your personal entropy production. 24

September 17, 2023

will be much more precise about what this means and what it implies later but, for the time being, nonlinear behavior can simply be thought of as meaning that measurable characteristics of a system generally do not simply double when its size is doubled. In the example given here, this can be restated as saying that there is a systematic increase in per capita GDP, as well as in average wages, crime rates, and many other urban metrics, as city size increases. This reflects an essential feature of all cities, namely that social activity and economic productivity are systematically enhanced with increasing size of the population. This systematic "value-added" bonus as size increases is called increasing returns to scale by economists and social scientists, whereas physicists prefer the more sexy term superlinear scaling .

29

September 17, 2023

Despite the terrible tragedy that befell Galileo, humanity reaped a wonderful benefit from his incarceration. It may very well have happened anyway, but it was while he was under house arrest that he wrote what is perhaps his finest work, one of the truly great books in the scientific literature, titled Discourses and Mathematical Demonstrations Relating to Two New Sciences. 1 The book is basically his legacy from the preceding forty years during which he grappled with how to systematically address the challenge of understanding the natural world around us in a logical, rational framework. As such, it laid the groundwork for the equally monumental contribution of Isaac Newton and pretty much for all of the science that followed. Indeed, in praising the book, Einstein was not exaggerating when he called Galileo "the Father of Modern Science." 2

52

September 18, 2023

Now, as we scale up a building or an animal, their weights increase in direct proportion to their volumes provided, of course, that the materials they're made of don't change so that their densities remain the same: so doubling the volume doubles the weight. Thus, the weight being supported by a pillar or a limb increases much faster than the corresponding increase in strength, because weight (like volume) scales as the cube of the linear dimensions whereas strength increases only as the square. To emphasize this point, consider increasing the height of a building or tree by a factor of 10 keeping its shape the same; then the weight needed to be supported increases a thousand fold (103) whereas the strength of the pillar or trunk holding it up increases by only a hundred fold (102). Thus, the ability to safely support the additional weight is only a tenth of what it had previously been. Consequently, if the size of the structure, whatever it is, is arbitrarily increased it will eventually collapse under its own weight. There are limits to size and growth. To put it slightly differently: relative strength becomes progressively weaker as size increases . Or, as Galileo so graphically put it: "the smaller the body the greater its relative strength. Thus a small dog could probably carry on his back two or three dogs of his own size; but I believe that a horse could not carry even one of his own size." 56

September 18, 2023

A classic example of some of the challenges and pitfalls is an early study investigating the potentially therapeutic effects of LSD on humans. Although the term "psychedelic" was already coined by 1957, the drug was almost unknown outside of a specialized psychiatric community in 1962, when the psychiatrist Louis West (no relation), together with Chester Pierce at the University of Oklahoma and Warren Thomas, a zoologist at the Oklahoma City zoo, proposed investigating its effects on elephants. Elephants? Yes, elephants and, in particular, Asiatic elephants. Although it may sound somewhat eccentric to use elephants rather than mice as the "model" for studying the effects of LSD, there 68 *were some not entirely implausible reasons for doing so. It so happens that Asiatic elephants periodically undergo an unpredictable transition from their normal placid obedient state to one in which they become highly aggressive and even dangerous for periods of up to two weeks. West and his collaborators speculated that this bizarre and often destructive behavior, known as musth , was triggered by the autoproduction of LSD in elephants' brains. So the idea was to see if LSD would induce this curious condition and, if so, thereby gain insight into LSD's effects on humans from studying how they react. Pretty weird, but maybe not entirely unreasonable. However, this immediately raises an intriguing question: how much LSD should you give an elephant? At that time little was known about safe dosages of LSD. Although it had not yet entered the popular culture it was known that even dosages of less than a quarter milligram would induce a typical "acid trip" for a human being and that a safe dose for cats was about one tenth of a milligram per kilogram of body weight. The investigators chose to use this latter number to estimate how much LSD they should give to Tusko the elephant, their unsuspecting subject that resided at the Lincoln Park Zoo in Oklahoma City. Tusko weighed about 3,000 kilograms, so using the number known to be safe for cats, they estimated that a safe and appropriate dose for Tusko would be about 0.1 milligram per kilogram multiplied by 3,000 kilograms, which comes out to 300 milligrams of LSD. The amount they actually injected was 297 milligrams. Recall that a good hit of LSD for you and me is less than a quarter milligram. The results on Tusko were dramatic and catastrophic. To quote directly from their paper: "Five minutes after the injection he [the elephant] trumpeted, collapsed, fell heavily onto his right side, defecated, and went into status epilepticus." Poor old Tusko died an hour and forty minutes later. Perhaps almost as disturbing as this awful outcome was that the investigators concluded that elephants are "proportionally very sensitive to LSD." The problem, of course, is something we've already stressed several times, namely the seductive* *trap of linear thinking. The calculation of how big a dosage should be used on Tusko was based on the implicit assumption that effective and safe dosages scale linearly with body weight so that the dosage per kilogram of body weight was presumed to be the same for all mammals. The 0.1 milligram per kilogram of body weight obtained from cats was therefore naively multiplied by Tusko's weight, resulting in the outlandish estimate of 297 milligrams, with disastrous consequences. Exactly how doses should be scaled from one animal to another is still an open question that depends, to varying degrees, on the detailed properties of the drug and the medical condition being addressed. However, regardless of details, an understanding of the underlying mechanism by which drugs are transported and absorbed into specific organs and tissues needs to be considered in order to obtain a credible estimate. Among the many factors involved, metabolic rate plays an important role. Drugs, like metabolites and oxygen, are typically transported across surface membranes, sometimes via diffusion and sometimes through network systems. As a result, the dose-determining factor is to a significant degree constrained by the scaling of surface areas rather than the total volume or weight of an organism, and these scale nonlinearly with weight. A simple calculation using the ⅔ scaling rule for areas as a function of weight shows that a more appropriate dose for elephants should be closer to a few milligrams of LSD rather than the several hundred that were actually administered. Had this been done, Tusko would no doubt have lived and a vastly different conclusion about the effects of LSD would have been drawn. The lesson is clear: the scaling of drug dosages is nontrivial, and a naive approach can lead to unfortunate results and mistaken conclusions if not done correctly with due attention being paid to the underlying mechanism of drug transport and absorption. It is clearly an issue of enormous importance, sometimes even a matter of life and death. This is one of the main reasons it takes so long for new drugs to obtain approval for their general use. Lest you think this was some fringe piece of research, the paper on* *elephants and LSD was published in one of the world's most highly regarded and prestigious journals, namely Science . 6*

September 18, 2023

One of his most fascinating innovations was the unique introduction of a broad gauge of 7 feet ¼ inch for the width between tracks. The standard gauge of 4 feet 8½ inches, which was used in all other British railways at that time, was adopted worldwide and is used on almost all railways today. Brunel pointed out that the standard gauge was an arbitrary carryover from the mine railways built before the invention of the world's first passenger trains in 1830. It had simply been determined by the width needed to fit a cart horse between the shafts that pulled carriages in the mines. Brunel rightly thought that serious consideration should be given to determining what the optimum gauge should be and tried to bring some rationality to the issue. He claimed that his calculations, confirmed by a series of trials and experiments, showed that his broader gauge was the optimum size for providing higher speeds, greater stability, and a more comfortable ride to passengers. Consequently, the Great Western Railway was unique in having a gauge that was almost twice as wide as every other railway line. Unfortunately, in 1892, following the evolution of a national railway system, the British Parliament forced the Great Western Railway to conform to the standard gauge, despite its acknowledged inferiority. 82

September 18, 2023

Because the essence of any measurable quantity cannot depend on an arbitrary choice of units made by human beings, neither can the laws of physics. Consequently, all of these and indeed all of the laws of science must be expressible as relationships between scale-invariant dimensionless quantities, even though conventionally they are not typically written that way. This was the underlying message of Rayleigh's seminal paper. His paper elegantly illustrates the technique with many well-chosen examples, including one 98

that provides the scientific explanation for one of the great mysteries of life that all of us have pondered at some time, namely, why is the sky blue? Using an elegant argument based solely on relating purely dimensionless quantities, he shows that the intensity of light waves scattered by small particles must decrease with the fourth power of their wavelength. Thus, when sunlight, which is a combination of all of the colors of the rainbow, scatters from microscopic particles suspended in the atmosphere, the shortest wavelengths, corresponding to blue light, dominate. Actually, Rayleigh had derived this stunning result much earlier in a tour de force based on a masterful mathematical analysis of the problem that provided a detailed mechanistic explanation for the origin of the shift toward the blue end of the spectrum. His point in presenting the simple derivation in his similitude paper was to show that the same result could have been derived "after a few minutes' consideration," as he put it, using a scaling argument in the guise of the "great principle of similitude" without the detailed sophisticated mathematics. His scaling argument showed that the shift to shorter wavelengths is an inevitable outcome of any analysis once you know what the important variables are. What is missing from such a derivation is a deeper understanding of the mechanism by which the result comes about. This is characteristic of many scaling arguments: general results can be derived, but details of their mechanistic origins sometimes remain hidden.

September 19, 2023

All of life functions by transforming energy from physical or chemical sources into organic molecules that are metabolized to build, maintain, and reproduce complex, highly organized systems. This is accomplished by the operation of two distinct but closely interacting systems: the genetic code, which stores and processes the information and "instructions" to build and maintain the organism, and the metabolic system, which acquires, transforms, and allocates energy and materials for maintenance, growth, and reproduction. Considerable progress has been made in elucidating both of these systems at levels from molecules to organisms, and later I will discuss how it can be extended to cities and 101 *companies. However, understanding how information processing ("genomics") integrates with the processing of energy and resources ("metabolics") to sustain life remains a major challenge. Finding the universal principles that underlie the structures, dynamics, and integration of these systems is fundamental to understanding life and to managing biological and socioeconomic systems in such diverse contexts as medicine, agriculture, and the environment.*

September 19, 2023

The search for fundamental principles that govern how the complexity of life emerges from its underlying simplicity is one of the grand challenges of twenty-first-century science. Although this has been, and will continue to be, primarily the purview of biologists and chemists, it is becoming an activity where other disciplines, and in particular physics and computer science, are playing an increasingly important role. 102

The New Kind of Science.

September 19, 2023

Death is an essential feature of life. Indeed, implicitly it is an essential feature of the theory of evolution. A necessary component of the evolutionary process is that individuals eventually die so that their offspring can propagate new combinations of genes that eventually lead to adaptation by natural selection of new traits and new variations leading to the diversity of species. We must all die so that the new can blossom, explore, adapt, and evolve. Steve Jobs put it succinctly 3 : No one wants to die. Even people who want to go to heaven don't want to die to get there. And yet death is the destination we all share. No one has ever escaped it, and that is how it should be, because death is very likely the single best invention of life. It's life's change agent. It clears out the old to make way for the new. 107

September 19, 2023

As I embarked on this new direction, I received unexpected support for my ruminations about biology as a science and its relationship to mathematics from an unlikely source. I discovered that what I had presumed was subversive thinking had been expressed much more articulately and deeply almost one hundred years earlier by the eminent and somewhat eccentric biologist Sir D'Arcy Wentworth Thompson in his classic book On Growth and Form, published in 1917. 4 It's a wonderful book that has remained quietly revered not just in biology but in mathematics, art, and architecture, influencing thinkers and artists from Alan Turing and Julian Huxley to Jackson Pollock. A testament to its continuing popularity is that it still remains in print. The distinguished biologist Sir Peter Medawar, the father of organ transplants, who received the Nobel Prize for his work on graft rejection and acquired immune tolerance, called On Growth and Form "the finest work of literature in all the annals of science that have been recorded in the English tongue."

108

September 19, 2023

Although he was awarded the prestigious Darwin Medal by the Royal Society in 1946, Thompson was critical of conventional Darwinian evolutionary theory because he felt that biologists overemphasized the role of natural selection and the "survival of the fittest" as the fundamental determinants of the form and structure of living organisms, rather than appreciating the importance of the role of physical laws and their mathematical expression in the evolutionary process. The basic question implicit in his challenge remains unanswered: are there "universal laws of life" that can be mathematized so that biology can be formulated as a predictive quantitative Science? He put it this way: It behoves us always to remember that in physics it has taken great men to discover simple things. . . . How far even then mathematics will suffice to describe, and physics to explain, the fabric of the body, no man can foresee. It may be

109

that all the laws of energy, and all the properties of matter, and all the chemistry of all the colloids are as powerless to explain the body as they are impotent to comprehend the soul. For my part, I think it is not so. Of how it is that the soul informs the body, physical science teaches me nothing; and that living matter influences and is influenced by mind is a mystery without a clue. Consciousness is not explained to my comprehension by all the nerve-paths and neurons of the physiologist; nor do I ask of physics how goodness shines in one man's face, and evil betrays itself in another. But of the construction and growth and working of the body, as of all else that is of the earth earthy, physical science is, in my humble opinion, our only teacher and guide.

September 19, 2023

Metabolism is the fire of life . . . and food, the fuel of life . 111

September 19, 2023

Like all of life, we evolved by a process of natural selection, interacting with and adapting to our fellow creatures, whether bacteria and viruses, ants and beetles, snakes and spiders, cats and dogs, or grass and trees, and everything else in a continuously challenging and evolving environment. We have been coevolving together in a never-ending multidimensional interplay of interaction, conflict, and adaptation. Each organism, each of its organs and subsystems, each cell type and genome, has therefore evolved following its own unique history in its own ever-changing environmental niche. The principle of natural selection, introduced independently by Charles Darwin and Alfred Russell Wallace, is key to the theory of evolution and the origin of species. Natural selection, or the "survival of the fittest," is the gradual process by which a successful variation in some inheritable trait or characteristic becomes fixed in a population through the differential reproductive success of organisms that have developed this trait by interacting with their environment. As Wallace expressed it, there is sufficiently broad variation that "there is always material for 112 *natural selection to act upon in some direction that may be advantageous," or as put more succinctly by Darwin: "each slight variation, if useful, is preserved."*

September 19, 2023

If we now consider a cow that is 100 times heavier than the cat, then Kleiber's law predicts that its metabolic rate is likewise 32 times greater than the cat's, and if we extend this to a whale that is 100 times heavier than the cow, its metabolic rate would be 32 times greater than the cow's. This repetitive behavior, the recurrence in this case of the same factor 32 as we move up in mass by the same repetitive factor of 100, is an example of the general self-similar feature of power laws. More generally: if the mass is increased by any arbitrary factor at any scale (100, in the example), then the metabolic rate increases by the same factor (32, in the example) no matter what the value of the initial mass is, that is, whether it's that of a mouse, cat, cow, or whale. This remarkably systematic repetitive behavior is called scale invariance or self-similarity and is a property inherent to power laws. It is closely related to the concept of a fractal, which will be discussed in detail in the following chapter. To varying degrees, fractality, scale invariance, and self-similarity are ubiquitous across nature from galaxies and clouds to your cells, your brain, the Internet, companies, and cities. We just saw that a cat that is 100 times heavier than a mouse requires only about 32 times as much energy to sustain it even though it has approximately 100 times as many cells—a classic example of an economy of scale resulting from the essential nonlinear nature of Kleiber's law . Naive linear reasoning would have predicted the cat's metabolic rate to have been 100 times larger, rather than only 32 times. Similarly, if the size of an animal is doubled it doesn't need 100 percent more energy to sustain it; it needs only about 75 percent more—thereby saving approximately 25 percent with each doubling. Thus, in a systematically predictable and quantitative way, the larger the organism the less energy has to be produced per cell per second to sustain a gram of tissue. Your cells work less hard than your dog's, but your horse's work even less hard. Elephants are roughly 10,000 times heavier than rats but their metabolic 115 *rates are only 1,000 times larger, despite having roughly 10,000 times as many cells to support. Thus, an elephant's cells operate at about a tenth the rate of a rat's, resulting in a corresponding decrease in the rates of cellular damage, and consequently to a greater longevity for the elephant, as will be explained in greater detail in chapter 4. This is an example of how the systematic economy of scale has profound consequences that reverberate across life from birth and growth to death.*

September 19, 2023

Doubling the mass of a mammal increases all of its timescales such as its life span and time to maturity by about 25 percent on average and, concomitantly, decreases all rates, such as its heart rate, by the same amount. Whales live in the ocean, elephants have trunks, giraffes have long necks, we walk on two legs, and dormice scurry around, yet despite these obvious differences, we are all, to a large degree, nonlinearly scaled versions of one another. If you tell me the size of a mammal, I can use the scaling laws to tell you almost everything about the average values of its measurable characteristics: how much food it needs to eat each day, what its heart rate is, how long it will take to mature, the length and radius of its aorta, its life span, how many offspring it will have, and so on. Given the extraordinary complexity and diversity of life, this is pretty amazing. 118

September 19, 2023

Indeed, Huxley coined the term allometric to describe how physiological and morphological characteristics of organisms scale with body size, though his focus was primarily on how that occurred during growth. Allometric was introduced as a generalization of the Galilean concept of isometric scaling, discussed in the previous chapter, where body shape and geometry do not change as size increases, so all lengths associated with an organism increase in the same proportion; iso is Greek for "the same," and metric is derived 121 *from metrikos , meaning "measure." Allometric , on the other hand, is derived from allo, meaning "different," and refers to the typically more general situation where shapes and morphology change as body size increases and different dimensions scale differently. For example, the radii and lengths of tree trunks, or for that matter the limbs of animals, scale differently from one another as size increases: radii scale as the ⅜ power of mass, whereas lengths scale more slowly with ¼ power (that is, as the 2⁄8 power).*

September 20, 2023

Prior to this digression into the interrelationship between the cultures of biology and physics, I argued that the mechanistic origins of scaling laws in biology were rooted in the universal mathematical, dynamical, and organizational properties of the multiple networks that distribute energy, materials, and information to local microscopic sites that permeate organisms, such as cells and mitochondria in animals. I also argued that because the structures of biological networks are so varied and stand in marked contrast to the uniformity of the scaling laws, their generic properties must be independent of their specific evolved design. In other words, there must be a common set of network properties that transcends whether they are constructed of tubes as in mammalian circulatory systems, fibers as in plants and trees, or diffusive pathways as in cells. 138

September 20, 2023

II. The Invariance of Terminal Units This simply means that the terminal units of a given network design, such as the capillaries of the circulatory system that we just discussed, all have approximately the same size and characteristics regardless of the size of the organism. Terminal units are critical elements of the network because they are points of delivery and transmission where energy and resources are exchanged. Other examples are mitochondria within cells, cells within bodies, and petioles (the last branch) of plants 140 *and trees. As individuals grow from newborn to adult, or as new species of varying sizes evolve, terminal units do not get reinvented nor are they significantly reconfigured or rescaled. For example, the capillaries of all mammals, whether children, adults, mice, elephants, or whales, are essentially all the same despite the enormous range and variation of body sizes.*

September 20, 2023

This invariance of terminal units can be understood in the context of the parsimonious nature of natural selection. Capillaries, mitochondria, cells, et cetera, act as "ready-made" basic building blocks of the corresponding networks for new species, which are rescaled accordingly. The invariant properties of the terminal units within a specific design characterize the taxonomic class. For instance, all mammals share the same capillaries. Different species within that class such as elephants, humans, and mice are distinguished from one another by having larger or smaller, but closely related, network configurations. 141

September 20, 2023

However, if you think of a tree as a bundle of fibers all tightly tied together beginning in the trunk and then sequentially spraying out up through its branches, then it's clear that the cross-sectional area must be preserved all the way up through the hierarchy. This is illustrated below, where this fiber bundle structure is compared with the pipe structure of mammals. An interesting consequence of area-preserving branching is that the cross-sectional area of the trunk is the same as the sum of the cross-sectional areas of all the tiny branches at the end of the network (the petioles). Amazingly, this was known to Leonardo da Vinci. I have reproduced the requisite page of his notebook where he demonstrates this fact. Although this simple geometric picture demonstrates why trees obey area-preserving branching, it is in actuality an oversimplification. However, area preserving can be derived from a much more realistic model for trees using the 149 *general network principles of space filling and optimization enunciated above, supplemented with biomechanical constraints that require branches to be resilient against perturbations from wind by bending without breaking. Such an analysis shows that in almost all respects plants and trees scale just like mammals, both within individuals as well as across species, including the ¾ power law for their metabolic rates, even though their physical design is quite different. 14*

September 21, 2023

So if for some perverse reason you wanted to know the radius, length, blood flow rate, pulse rate, velocity, pressure, et cetera, in the fourteenth branch of the circulatory system of the average hippopotamus, the theory will provide you with the answer. In fact, the theory will tell you the answer for any of these quantities for any branch of the network in any animal. As blood flows through smaller and smaller vessels on its way down through the network, viscous drag forces become increasingly important, leading to the dissipation of more and more energy. The effect of this energy loss is to progressively dampen the wave on its way down through the network hierarchy until it eventually loses its pulsatile character and turns into a steady flow. In other words, the nature of the flow makes a transition from being pulsatile in the larger vessels to being steady in the smaller ones. That's why you feel a pulse only in your main arteries—there's almost no vestige of it in your smaller vessels. In the language of electrical transmission, the nature of the blood flow changes from being AC to DC as it progresses down through the network. Thus, by the time blood reaches the capillaries its viscosity ensures that it is no longer pulsatile and that it is moving extremely slowly. It slows down to a speed of only about 1 millimeter per second, which is tiny compared with its speed of 40 centimeters per second when it leaves the heart. This is extremely important because this leisurely speed ensures that oxygen carried by the blood has sufficient time to diffuse efficiently across the walls of the capillaries and be rapidly delivered to feed cells.
Interestingly, the theory predicts that these velocities at the two extremities of the network, the capillaries and the aorta, are the same for all mammals, as observed. You are very

154

likely aware of this huge difference in speeds between capillaries and the aorta. If you prick your skin, blood oozes out very slowly from the capillaries with scant resulting damage, whereas if you cut a major artery such as your aorta, carotid, or femoral, blood gushes out and you can die in just a matter of minutes.

September 21, 2023

Because blood flow changes from pulsatile to nonpulsatile as one progresses down the network, our circulatory system is actually not continuously self-similar nor therefore a precise fractal. In the nonpulsatile domain where the flow is dominated by viscous forces, minimizing the amount of power being dissipated leads to a self-similarity in which the radii of successive vessels decrease by a constant factor of the cube root of two 3√2 (= 1.26 . . .), rather than the square root √2 (= 1.41 . . .) as in the pulsatile region. Thus the fractal nature of the circulatory system subtly changes from the aorta to the capillaries, reflecting the change in the nature of the flow from pulsatile to nonpulsatile. Trees, on the other hand, maintain approximately the same self-similarity from the trunk to their leaves, with radii successively decreasing by the area-preserving ratio of √2. 159

September 21, 2023

His theory did not attempt to explain the fundamental origins of war, that is, why we collectively resort to force and violence to settle our conflicts, but rather to show how the dynamics of arms races escalate, resulting in catastrophic conflict. Although his theory is highly oversimplified, Richardson had some success in comparing his analyses with 165 *data, but more important, he provided an alternative framework for quantitatively understanding the origins of war that could be confronted with data. Furthermore, it had the virtue of showing what parameters were important, especially in providing scenarios under which a potentially peaceful situation could be achieved and sustained. In contrast to conventional, more qualitative theories of conflict, the roles of leadership, cultural and historical animosity, and specific events and personalities play no explicit role in his theory. 18*

September 21, 2023

Well, Benoit Mandelbrot did. He deserves great credit not only for resurrecting Richardson's work but for recognizing its deeper significance. In 1967 he published a paper in the high-profile journal Science with the more transparent title "How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension." 22 This brought Richardson's work to light by expanding on his findings and generalizing the idea. Crinkliness, later to become known as fractality, is quantified by how steep the slopes of the corresponding straight lines are on Richardson's logarithmic plots: the steeper the slope, the more crinkly the curve. These slopes are just the exponents of the power laws relating length to resolution and are the analog of the ¾ exponent relating metabolic rate to mass for organisms. For very smooth traditional curves, like circles, the slope or exponent is zero because its length does not change with increasing resolution but converges to a definite value, as in the living room example. However, for rugged, crinkly coastlines the slope is nonzero. For example, for the west coast of Britain, it's 0.25.

172

September 21, 2023

Perhaps of greater importance is that he realized that these ideas are generalizable far beyond considerations of borders and coastlines to almost anything that can be measured, even including times and frequencies. Examples include our brains, balls of crumpled paper, lightning, river networks,

176

and time series like electrocardiograms (EKGs) and the stock market. For instance, it turns out that the pattern of fluctuations in financial markets during an hour of trading is, on average, the same as that for a day, a month, a year, or a decade. They are simply nonlinearly scaled versions of one another. Thus if you are shown a typical plot of the Dow Jones average over some period of time, you can't tell if it's for the last hour or for the last five years—the distributions of dips, bumps, and spikes is pretty much the same, regardless of the time frame. In other words, the behavior of the stock market is a self-similar fractal pattern that repeats itself across all timescales following a power law that can be quantified by its exponent or, equivalently, its fractal dimension. You might think that with this knowledge you might soon become rich. Although this certainly gives new insight into hidden regularities in stock markets, unfortunately it has predictive power only in an average coarse-grained sense and does not give specific information about the behavior of individual stocks. Nevertheless, it's an important ingredient for understanding the dynamics of the market over different timescales. This has stimulated the development of a new transdisciplinary subfield of finance called econophysics and motivated investment companies to hire physicists, mathematicians, and computer scientists to use these sorts of ideas to develop novel investment strategies. 23 Many have done very well, though it is unclear just how big a role their physics and mathematics actually played in their success.

September 21, 2023

Almost all of the networks that sustain life are approximately self-similar fractals. In the previous chapter, I explained how the nature and origin of these fractal structures are a consequence of generic geometric, mathematical, and physical principles such as optimization and space filling, thereby leading to the derivation for how networks scale both within an average individual as well as across species. 181

September 22, 2023

Thus, instead of scaling with classic ⅓ exponents, as would be the case if they were smooth nonfractal Euclidean objects, they scale with ¼ exponents. Although living things occupy a three-dimensional space, their internal physiology and anatomy operate as if they were four-dimensional. 188

September 22, 2023

This is really weird. Such an animal would have a beating heart but no pulse! This is not only weird, but more important, it would be an extremely inefficient design because it would have entirely lost the advantages of impedance matching and would lead to significant energy being dissipated in all of the vessels throughout its circulatory system. This loss of performance efficiency is reflected in how metabolic rates would scale. Instead of following the classic sublinear ¾ power scaling law, a calculation shows that it would now scale linearly—that is, directly proportional to body mass—and thereby lose the advantages of economies of scale. In this purely DC case the power needed to support a gram of tissue would now be the same regardless of size instead of systematically decreasing with size following quarter-power scaling. Consequently, no evolutionary advantage would be conferred by increasing size. This argument shows that only mammals that are large enough for their circulatory systems to support pulsatile waves through at the very least the first couple of branching levels would have evolved, thereby providing a fundamental reason why there is a minimum size. 3 The theory can be used to derive a formula for when this tipping point occurs. Its actual value depends upon generic quantities such as the density and viscosity of blood and the elasticity of arterial walls. The calculation leads to an approximate value for the size of the smallest mammal of just a couple of grams, comparable to the mass of the Etruscan shrew, which is the smallest known mammal. It is only about 4 centimeters long, easily sitting on the palm of your hand. Its tiny heart beats at more than a thousand times a minute—about twenty times a second—as it pumps blood with the same pressure and speed as you do , and even more astounding, as does a blue whale. 192

And all of this through its minuscule aorta, which is only a couple of millimeters long and an astonishingly couple of tenths of a millimeter wide, not much thicker than a hairbreadth. As I have remarked earlier, no wonder the poor creature doesn't live very long.

September 22, 2023

Let me deconstruct this a little further to give a more mechanistic understanding of what's going on. Recall that the reason metabolic rate scales with a sublinear ¾ power exponent lies in the hegemony of the network. Furthermore, because the entire flow through the network ends up going through all of the capillaries, and because they are invariant across species as well as during ontogeny (capillaries are approximately the same for mice, elephants, their babies and children, as well as for us), their number likewise scales with a ¾ power exponent. So as the organism grows and size increases, each capillary systematically has to service more cells following ¼ power scaling. It is this mismatch at the critical interface between capillaries and cells that controls growth and ultimately leads to its cessation: the increase in the number of supply units (the capillaries) cannot keep up with the demands from the increase in the number of customers (the cells). 203

September 22, 2023

And the theory tells us why: growth is primarily determined by how energy is delivered to cells, and this is constrained by universal properties of networks that transcend design. Among the many other aspects of growth that can be derived from the theory, it predicts how the allocation of metabolic energy between maintenance and growth changes with age. At birth almost all of it is devoted to growth and relatively little to maintenance, whereas beyond maturity all of it is devoted to maintenance, repair, and replacement. 205

September 22, 2023

This leads to the fascinating conclusion that across the spectrum of life all biological rates and times such as those associated with growth, embryonic development, longevity, and evolutionary processes are determined by a joint universal scaling law in terms of just two parameters: the number ¼, arising from the network constraints that control the dependence on mass, and 0.65 eV, originating in the chemical reaction dynamics of ATP production. This result can be restated in a slightly different way: when adjusted for size and temperature, as determined by just these two numbers, all organisms run to a good approximation by the same universal clock with similar metabolic, growth, and evolutionary rates.

211

September 23, 2023

More pertinent, a modest 2°C change in ambient temperature leads to a 20 percent to 30 percent change in growth and mortality rates. 11 This is huge and therein lies our problem. If global warming induces a temperature increase of around 2°C, which it is on track to do, then the pace of almost all biological life across all scales will increase by a whopping 20 percent to 30 percent. This is highly nontrivial and will potentially wreak havoc with the ecosystem. It's analogous to the huge leap that Brunel attempted to make when building his mammoth ship the Great Eastern, which ended in disaster primarily because the science of shipbuilding had not yet been sufficiently well developed. Ships are extremely simple compared with the profound complexity of ecosystems and societies.

213

September 23, 2023

Indeed, the data show that the half-life of publicly traded companies in the United States is only about ten years. So in just fifty years (five half-lives) only (½)5 = 1/32 or about 3 percent are still posting sales. This begs the fascinating question as to whether the same general dynamics underlies

231

the surprising commonality in the mortality of organisms, isotopes, and companies. We will return to speculate about this later.

September 24, 2023

Here's a summary of some of the significant properties of aging and mortality that need to be explained by any theory: Aging and death are "universal": all organisms eventually die. A corollary to this is that there is a maximum life span and a corresponding vanishing survival rate. Semiautonomous subsystems of organisms, such as our various organs, age approximately uniformly. Aging progresses approximately linearly with age. For instance, Figure 27 shows how organ functionality degrades with age. 15 Plotted is the percentage of maximum capacity for various vital functions indicating a linear decline with age beginning almost immediately after maturity at about age twenty. It's slightly depressing to see that on average we are physically optimal (100 percent) for only a very few years and that beginning around age twenty it's literally downhill all the way. Notice also that we build up to our maximum capacity relatively quickly during our growth period. Later, I will suggest that this aging process is under way even during our earliest years prior to maturity but is hidden by the overwhelming predominance of growth. The aging process effectively begins as soon as you are conceived. Bob Dylan got it right when he sang that "he not busy being born is busy dying." Life spans scale with body mass as a power law whose exponent is approximately ¼. As anticipated, there is a large variance in the data, partly because there are no controlled life-history experiments on longevity for mammals, including us. Some of the data are garnered from reports on wild animals, some from zoos, some from domesticated animals, some from research laboratories, each with very different environmental and lifestyle conditions. In addition, some are reports of just one or two animals of a species, and some from large cohorts. Although this lack of control is problematic, there are clear trends and consistencies in the data that statistically point to approximate ¼ power scaling. The number of heartbeats in a lifetime is approximately the same for all mammals, as

236

shown in Figure 2 in the opening chapter. 16 Thus, shrews have heart rates of roughly 1,500 beats a minute and live for about two years, whereas heart rates of elephants are only about 30 beats a minute but they live for about seventy-five years. Despite their vast difference in size, both of their hearts beat approximately one and a half billion times during an average lifetime. This invariance is approximately true for all mammals, even though there are large fluctuations for the reasons I outlined above. The greatest outlier from this intriguing invariance is us: for modern human beings, on average our hearts beat approximately two and a half billion times, which is about twice the number for a typical mammal. However, as I have already emphasized, it is only in the past one hundred years that we have been living this long. Over the entire history of humankind up until relatively recently we lived for approximately half as long as we do now and, like the vast majority of mammals, followed the approximately invariant one and a half billion heartbeats "law." Related to this is another invariant quantity: the total amount of energy used in a lifetime to support a gram of tissue is approximately the same for all mammals and, more generally, for all animals within a specific taxonomic group. 17 For mammals it's about 300 food calories per gram per lifetime. A more fundamental way of expressing this is to note that the number of turnovers during a lifetime of the respiratory machinery responsible for the production of energy in cells is approximately the same for all animals within a specific taxonomic group. For mammals this is about ten thousand trillion times (1016) and translates into the invariance of the number of ATP molecules (our fundamental currency of energy) produced in a lifetime to support one gram of tissue.

September 24, 2023

Damage occurs across multiple scales through many different mechanisms associated with physical or chemical transport phenomena, but loosely speaking it can be separated into two categories: (1) Classic physical wear and tear due to the viscous drag in the flow, analogous to the wear and tear resulting from ordinary friction when two physical objects move over each other like wearing out your 242

shoes or the tires of your car. (2) Chemical damage from free radicals, which are by-products of the production of ATP in respiratory metabolism. A free radical is any atom or molecule that has lost an electron and consequently has a positive electric charge, making it highly volatile. Most of this kind of damage is caused by oxygen radicals that react with vital cellular components. Oxidative damage to DNA may be particularly deleterious, because in nonreplicating cells such as in the brain and musculature, it causes permanent damage to transcriptional, and perhaps most important, regulatory regions of the genome. Although the detailed role and extent of oxidative damage in aging remains unclear, it has stimulated a mini-industry of antioxidant supplements such as vitamin E, fish oil, and red wine as some of those elixirs of life to combat aging.

September 24, 2023

The total number of damages incurred in a lifetime is just the damage rate (that is, the number of damage events per unit time which is proportional to the number of terminal units) multiplied by the life span, and this has to be proportional to the total number of cells, and therefore to body mass. Consequently, life span is proportional to the total number of cells divided by the number of terminal units. But the number of terminal units scales with mass with a ¾ power exponent, while the number of cells scales linearly, resulting in life span scaling as the ¼ power of mass, consistent with data.

245

September 24, 2023

B. HEARTBEATS AND THE PACE OF LIFE: These data also confirm the approximate ¼ power scaling of life spans with mass. Because the theory for the cardiovascular system predicts that heart rates decrease with a ¼ power of mass, the dependence on mass cancels out when we multiply heart rates by life spans: the decrease in one exactly compensates for the increase in the other, resulting in an invariant, a quantity that's the same for all mammals. But multiplying heart rates by life spans just gives the total number of

247

heartbeats in a lifetime, so the theory predicts that this should be the same for all mammals, consistent with the data shown in Figure 2 in chapter 1. This argument can be extended to the fundamental level of the respiratory complexes, the basic units inside mitochondria where ATP is manufactured, to show that the number of times the reaction takes place producing ATP is the same for all mammals.

September 24, 2023

Given this dual nature of cities as, on the one hand, the origin of our major challenges and, on the other, the generator of creativity and ideas and therefore the source of their solutions, it becomes a matter of some urgency to ask whether there can be a "Science of Cities," by which I mean a conceptual framework for understanding their dynamics, growth, and evolution in a quantitatively predictable framework. This is crucial for devising a serious strategy for achieving long-term sustainability, especially given that the overwhelming majority of human beings will be urban dwellers by the second half of this century, many in megacities of unprecedented size. 2

257

September 25, 2023

While it may not be unreasonable to hold the view that collective human ingenuity and innovation facilitated by a free market economy is the secret for maintaining long-term open-ended growth in defiance of potential collapse, I find it somewhat baffling that this is often coupled with a denial or, at the very least, a deep skepticism concerning some of its inevitable consequences. Like many who advocate "innovation" as the panacea for meeting future global socioeconomic challenges, Simon was a vocal skeptic when it came to believing that human activity caused global environmental damage or was the origin of serious health concerns, whether from climate change, pollution, or chemical contamination. The spirit and substance of the Second Law of Thermodynamics and its manifestation in terms of entropy production represent the dark side of open-

278

ended exponential growth. Independent of how superbly innovative we are, ultimately everything is driven and processed by the use of energy, and the processing of energy has inevitable deleterious consequences.

September 25, 2023

To put it slightly differently, the rate at which we need to process energy to sustain our standard of living remained at just a few hundred watts for hundreds of thousands of years, until about ten thousand years ago when we began to form collective urban communities. This marked the beginning of the Anthropocene, in which our effective metabolic rate began its steady rise to its present level of more than 3,000 watts today. But this is just its average value taken across the entire planet. The rate at which energy is used in developed countries is far higher. In the United States it is almost a factor of four larger, at a whopping 11,000 watts, which is more than one hundred times larger than its "natural" biological value. This amount of power is not a lot smaller than the metabolic rate of a blue whale, which is more than one thousand times larger in mass than we are. Thinking of us as an animal using thirty times more energy than we "should" given our physical size, the effective human population of the planet accordingly operates as if it were much larger than the 7.3 billion people who actually inhabit it. In a very real sense, we are operating as if our population were at least thirty times larger, equivalent to a global population in excess of 200 billion people. If the most optimistic of cornucopian thinkers are correct and the world's population reaches 10 billion by the end of the century, all living at a standard comparable to that of the United States, the subsequent effective population would then exceed a trillion people. 280

September 25, 2023

Despite the obviously central role that energy has played in bringing us to this point in the history of the planet and, in particular, in the socioeconomic development of modern 284

human societies, you will be hard-pressed to find even a sentence or two about it in any classic textbook on economics. Remarkably, concepts like energy and entropy, metabolism, and carrying capacity have not found their way into mainstream economics. The continued growth of economies, markets, and populations over the past two hundred years coupled with a parallel increase in standards of living are, not surprisingly, seen as testament to the success of classic economic thinking and as a rejection of neo-Malthusian ideas. There has been no need to think seriously in terms of energy as an underlying driver of economic success or of population growth, let alone to consider entropy as its inevitable consequence. Nor has there been the need to consider the possibility that resources may actually be limited, nor that there might be underlying physical constraints that would question open-ended growth. Until now.

September 26, 2023

No one is more identified with viewing cities through the collective lives of their citizens than the famous urban writer-theorist Jane Jacobs. Her defining book, The Death and Life of Great American Cities, had an enormous influence across the globe on how we think about cities and how we approach "urban planning." 3 It's required reading for anyone interested in cities whether a student, a professional, or just an intellectually curious citizen. I suspect that every mayor of every major city in the world has a copy of Jane's book sitting somewhere on his or her bookshelf and has read at least parts of it. It's a wonderful book, extremely provocative and insightful, highly polemical and personal, very entertaining and well written. 301

September 26, 2023

She was extraordinarily intolerant of urban planners and politicians, and vicious in her attacks on traditional urban planning, especially in regard to its apparent lack of recognition that people, not buildings and highways, were 301 *primary. These classic quotes from her writing are typical of her critical attitude: The pseudoscience of planning seems almost neurotic in its determination to imitate empiric failure and ignore empiric success. In this dependence on maps as some sort of higher reality, project planners and urban designers assume they can create a promenade simply by mapping one in where they want it, then having it built. But a promenade needs promenaders. There is no logic that can be superimposed on the city; people make it, and it is to them, not buildings, that we must fit our plans. . . . We can see what people like. His aim was the creation of self-sufficient small towns, really very nice towns if you were docile and had no plans of your own and did not mind spending your life with others with no plans of their own. As in all Utopias, the right to have plans of any significance belonged only to the planner in charge.*

September 26, 2023

Many writers have picked up this theme, including the urban economists Edward Glaeser and Richard Florida, but none has been as forthright and bold as Benjamin Barber in his book with the provocative title If Mayors Ruled the World: Dysfunctional Nations, Rising Cities. 5 These are indicative of a rising consciousness that cities are where the action is—where challenges have to be addressed in real time and where governance seems to work, at least relative to the increasing dysfunctionality of the nation-state. 311

September 28, 2023

An early recruit to our collaboration was Dirk Helbing, who was the director of the Institute for Transport and Economics at Dresden University of Technology in Germany when I first met him. Dirk had been trained in statistical physics and had applied these techniques to understand highway traffic and pedestrian crowds. He's now at the prestigious Swiss Federal Institute of Technology in Zurich, usually referred to as the ETH, where he runs a large project called the Living Earth Simulator. This is designed to model global-scale systems 321

from economies, governments, and cultural trends to epidemics, agriculture, and technological developments using big data sets and fancy algorithms.

September 29, 2023

Also clearly manifested in these graphs is the equally surprising result that all of the slopes of these various quantities have approximately the same value, clustering around 1.15. Thus these metrics not only scale in an extremely simple fashion following classic power law behavior, but they all do it in approximately the same way with a similar exponent of approximately 1.15 regardless of the urban system. So in marked contrast to infrastructure, which scales sublinearly with population size, socioeconomic quantities—the very essence of a city—scale superlinearly , thereby manifesting systematic increasing returns to scale. The larger the city, the higher the wages, the greater the GDP, the more crime, the more cases of AIDS and flu, the more restaurants, the more patents produced, and so on, all following the "15 percent rule" on a per capita basis in urban systems across the globe. 326

September 29, 2023

It is true that many other creatures, such as herding animals and especially social insects, also discovered economies of scale, but their accomplishments are relatively primitive and static compared with what humans have achieved. The power of language has allowed us to go well beyond classic economies of scale, such as our cells achieve or our hunter-gatherer predecessors attained to evolve and build on those advantages by adapting to new challenges over periods of time that are vastly shorter than typical evolutionary timescales that had hitherto been required for major innovations to be made. Ants brilliantly self-organized to evolve remarkably robust and hugely successful and sophisticated physical and social structures, but it took them millions of years to do so. Furthermore, they accomplished this more than 50 million years ago and have barely evolved 334 *beyond it since. On the other hand, once we had invented verbal language, it took us only tens of thousands of years to evolve from hunting and gathering to becoming sedentary agriculturalists—and even more remarkably, only another ten thousand years to evolve cities, become urbanists, and invent cell phones, airplanes, the Internet, quantum mechanics, and the general theory of relativity.*

September 29, 2023

If cities are thought of solely in terms of their physicality, as just buildings and roads and the multiple network systems of wires and pipes that supply them with energy and resources, then they are indeed quite analogous to organisms, manifesting similar systematic scaling laws encapsulating economies of scale. However, when humans began forming sizable communities they brought a fundamentally new dynamic to the planet beyond that of biology and the discovery of economies of scale. With the invention of language and the consequent exchange of information between people and groups of people via social networks, we discovered how to innovate and create wealth. Cities are therefore much more than giant organisms or anthills: they rely on long-range, complex exchanges of people, goods, and knowledge. They are invariably magnets for creative and innovative individuals, and stimulants for economic growth, wealth production, and new ideas. 335

September 29, 2023

The actual self-similarity of cities more closely reflects the organically evolved hierarchical network structures of transport and utility systems than the rigid hexagonal crystalline structures of Christaller. The city is not a top-down engineered machine dominated by straight lines and classic Euclidean geometry, but rather is much more akin to an organism with its crinkly lines and fractal-like shapes typical of a complex adaptive system—which it is. This is clear from just a casual look at the growth patterns of a typical city with its ever-expanding filigreed infrastructural 344 *network pattern reminiscent of the growth pattern of a bacterial colony, as illustrated here. A careful mathematical analysis of such patterns shows that cities are, in fact, approximate self-similar fractals much like biological organisms or geographical coastlines. For example, if the length of the perceived boundary of a city is measured at different resolutions, analogous to what Lewis Fry Richardson did for coastlines, and these are plotted logarithmically, then approximate straight lines result whose slope is the conventional fractal dimension of the city boundary.*

September 29, 2023

As I explained earlier, fractal dimensions are a measure of an object's degree of crinkliness, which some interpret as a measure of its complexity. Stimulated by the explosive interest in fractals and the incipient development of a science of complexity in the 1980s, the distinguished urban geographer Michael Batty carried out extensive statistical analyses on cities to measure their fractal dimensions. 6 345

See the reference.

September 29, 2023

In conceptualizing this dark psychosocial side of urban life, Milgram borrowed the term "overload" from electrical circuit and systems science theory. In big cities we are continually bombarded with so many sights, so many sounds, so many "happenings," and so many other people at such a high rate that we are simply unable to process the entire barrage of sensory information. If we tried to respond to every stimulus, our cognitive and psychological circuitry would break down and, in a word, we would blow a fuse just like an overloaded electrical circuit. And sadly, some of us do. Milgram suggested that the kinds of "antisocial" behaviors we perceive and experience in large cities are in fact adaptive responses for coping with the sensory onslaught of city life, implying that without such adaptations we'd all blow our 359

fuses.

September 29, 2023

Understanding and deconstructing the hierarchy of social group structures has been a major focus of attention in sociology and anthropology for more than fifty years, but it was only in the last twenty or so that some of their quantitative features have become apparent. Some of this has been driven by the work of the evolutionary psychologist Robin Dunbar and his collaborators, who proposed that an average individual's entire social network can be deconstructed into a hierarchical sequence of discrete nested clusters whose sizes follow a surprisingly regular pattern. 15 The size of the group at each level systematically increases as one progresses up the hierarchy from, say, family to city, while the strength of the bonding between people within the groups systematically decreases. So, for example, most people have a very strong connection with their immediate family members but only a very weak one with the bus driver or members of the city council. 361

September 29, 2023

Now let's go back and take a look at how this plays out in an entire city. If it were possible for everyone to interact meaningfully with everyone else as in one great big happy family, then the above argument would imply that all socioeconomic metrics should scale with the square of the population size. This would mean an exponent of 2, which is certainly superlinear (it's bigger than 1), but significantly larger than 1.15. However, this represents the extreme and totally unrealistic case where the entire population is in a frenzied state of continuous and complete interaction with itself, being churned about much like raisins or nuts in cake dough driven by a super-high-speed electric mixer. This is clearly impossible—and certainly not desirable. Even in a modest-size city of only 200,000 people there are roughly 20 billion possible relationships, and even if each person devoted just one minute a year to each relationship, they would have 377 *to spend their entire waking life relating to other people, leaving no time for anything else. Imagine extending that to a New York or Tokyo. There is also the constraint of the Dunbar number, according to which we even have difficulty sustaining any sort of meaningful relationship with more than about 150 people, let alone a couple of hundred thousand or several million. It is this restriction to a relatively small number of interactions that drives the superlinear exponent to be significantly smaller than its maximum possible value of 2.*

September 30, 2023

In this network scaling sense, it is no accident therefore that the physical and the social mirror each other, so much so that we can think of the physical city —with its networks of buildings, roads, and electrical, gas, and water lines—as an inverse nonlinear representation of the socioeconomic city — with its networks of social interactions. The city is indeed the people. The approximate 15 percent increase in social interactions and therefore in socioeconomic metrics such as income, patents, and crime generated with every doubling of city size can be interpreted as a bonus, or payoff, arising from the 15 percent savings in physical infrastructure and energy use. The systematic increase in social interaction is the essential driver of socioeconomic activity in cities: wealth creation, innovation, violent crime, and a greater sense of buzz and opportunity are all propagated and enhanced through social networks and greater interpersonal interaction. But this can equally well be interpreted in a complementary way by viewing cities as catalytic facilitators or crucibles for social chemistry in which the increase in social interactions enhances creativity, innovation, and opportunity whose dividend is an increase in infrastructural economies of scale. Just as raising the temperature of a gas or liquid increases the rate in the number of collisions between molecules, so increasing the size of a city increases the rate and number of interactions between its citizens. Metaphorically speaking, increasing the size of a city can therefore be thought of as raising its temperature. In this sense, New York, London, Rio, and Shanghai are truly hot cities, especially compared with Santa Fe where I live, and

382

the proverbial image of a "melting pot," originally applied to New York City, is an apt expression of this metaphor.

September 30, 2023

The analysis of these massive data sets of mobile phone calls used to test the theory was brilliantly carried out by a Swiss engineer, Markus Schläpfer, working with the Hungarian physicist Michael Szell, two of the bright young postdocs hired by Carlo Ratti at MIT. Markus later joined us at the Santa Fe Institute in 2013, where we began this particular collaboration. Of the many projects he worked on, a particularly interesting one was with Luis on analyzing how the heights and volumes of buildings relate to city size. Markus has since moved on to the prestigious ETH in Zurich, his hometown, where he is engaged with a large collaborative program called the Future Cities Lab, which is based in Singapore and supported by their government. 416

October 1, 2023

9. THE STRUCTURE OF WEALTH, INNOVATION, CRIME, AND RESILIENCE: THE INDIVIDUALITY AND RANKING OF CITIES How rich, creative, or safe can we expect a city to be? How can we establish which cities are the most innovative, the most violent, or the most effective at generating wealth? How do they rank according to economic activity, the cost of living, the crime rate, the number of AIDS cases, or the happiness of their populations? The conventional answer is to use simple per capita measures as performance indices and rank order of cities accordingly. Almost all official statistics and policy documents on wages, income, gross domestic product (GDP), crime, unemployment rates, innovation rates, cost of living indices, morbidity and mortality rates, and poverty rates are compiled by governmental agencies and international bodies worldwide in terms of both total aggregate and per capita metrics. Furthermore, well-known composite indices of urban performance and the quality of life, such as those assembled by the World Economic Forum and magazines like 420

Fortune , Forbes , and The Economist , primarily rely on naive linear combinations of such measures. 6 Because we have quantitative scaling curves for many of these urban characteristics and a theoretical framework for their underlying dynamics we can do much better in devising a scientific basis for assessing performance and ranking cities. The ubiquitous use of per capita indicators for ranking and comparing cities is particularly egregious because it implicitly assumes that the baseline, or null hypothesis, for any urban characteristic is that it scales linearly with population size. In other words, it presumes that an idealized city is just the linear sum of the activities of all of its citizens, thereby ignoring its most essential feature and the very point of its existence, namely, that it is a collective emergent agglomeration resulting from nonlinear social and organizational interactions. Cities are quintessentially complex adaptive systems and, as such, are significantly more than just the simple linear sum of their individual components and constituents, whether buildings, roads, people, or money. This is expressed by the superlinear scaling laws whose exponents are 1.15 rather than 1.00. This approximately 15 percent increase in all socioeconomic activity with every doubling of the population size happens almost independently of administrators, politicians, planners, history, geographical location, and culture.

October 2, 2023

My colleagues Luis, José, and Debbie carried out an analysis of these data with our postdoc Hyejin Youn taking the lead role. Hyejin was trained in statistical physics in Seoul, South Korea, and had joined SFI to finish her doctorate. She first worked on the origin and structure of languages before joining our collaboration. She has now established herself as an expert on innovation in technology and is currently a fellow of the Institute of New Economic Thinking (INET) at Oxford University—a new program funded by the financier George Soros. As we saw in the analysis of other urban metrics, the data reveal surprisingly simple and unexpected regularities. For instance, the total number of establishments in each city regardless of what business they conduct turns out to be linearly proportional to its population size. Double

431

the size of a city and on average you'll find twice as many businesses. The proportionality constant is 21.6, meaning that there is approximately one establishment for about every 22 people in a city, regardless of the city size . Or to put it slightly differently, on average a new workplace is created each time the population of a city increases by just 22 people, whether in a small town or a large metropolis. This is an unexpectedly small number and usually comes as quite a surprise to most people, even those dealing in business and commerce. Similarly, the data also show that the total number of employees working in these establishments also scales approximately linearly with population size: on average, there are only about 8 employees for every establishment, again regardless of the size of the city. This remarkable constancy of the average number of employees and the average number of establishments across cities of vastly different sizes and characters is not only contrary to previous wisdom, but also rather puzzling when viewed in light of the pervasive superlinear agglomeration effects that underlie all socioeconomic activity including per capita increases in productivity, wages, GDP, and patent production. 10

October 2, 2023

A couple of colloquial examples may help: successful companies and universities attract the smartest people, resulting in their becoming more successful, thereby attracting even smarter people, leading to even greater success and so on, just as wealthy people attract favorable investment opportunities that generate even more wealth which they further invest to get even more wealthy. Hence the catchphrase the rich get richer and its implied, but usually unstated, corollary the poor get poorer to characterize this process. Or, as so articulately put by Jesus according to the Gospel of Matthew in the New Testament: For everyone who has will be given more, and he will have an abundance. Whoever does not have, even what he has will be taken from him. This surprising declaration has been used by some fundamentalist Christians and others as justification for rampant capitalism—a sort of anti–Robin Hood slogan supporting the idea of taking from the poor to give to the

437

rich. But while Jesus's remarks are a good example of preferential attachment , the quote is, not surprisingly, taken out of context. It is often conveniently forgotten that Jesus was actually referring to knowledge of the mysteries of the kingdom of heaven and not to material wealth. He was expressing a spiritual version of the very essence of diligent study, knowledge accumulation, and research and education as expressed by the ancient rabbis: He who does not increase his knowledge decreases it.

October 2, 2023

In its extreme version the underlying philosophy of agent-based modeling is antithetical to the traditional scientific framework, where the primary challenge is to reduce huge numbers of seemingly disparate and disconnected observations down to a few basic general principles or laws: as in biology, where the principle of natural selection applies to all organisms from cells to whales, or in physics, where Newton's laws apply to all motion from automobiles to planets. In contrast, the aim of agent-based modeling is to construct an almost one-to-one mapping of each specific system. General laws and principles that constrain its structure and dynamics play a secondary role. For example, in simulating a specific company, every individual worker, administrator, transaction, sale, cost, et cetera is in effect included and each company consequently treated as a separate, almost unique entity, typically without explicit regard either to its systematic behavior or its relationship to the bigger picture. Clearly, both approaches are needed: the generality and parsimony of "universal" laws and systematic behavior reflecting the big picture and dominant forces shaping general behavior, coupled with and informed by detailed modeling reflecting the individuality and uniqueness of each company. In the case of cities, scaling laws revealed that 80 to 90 percent of their measurable characteristics are determined from just knowing their population size, with the remaining 10 to 20 percent being a measure of their individuality and uniqueness, which can be understood only from detailed studies that incorporate local historical, geographical, and cultural characteristics.

454

October 2, 2023

Let's see how this scenario compares with data. Figure 68 is a wonderful graph showing the growth of sales for all 28,853 companies in the Compustat data set plotted together in real calendar time, adjusted for inflation. To get all of them onto a single manageable graph, the vertical axis representing sales is logarithmic. Despite being a "spaghetti" plot, the graph is surprisingly illuminating. The overall trend is clear: as predicted, many young companies shoot out of the starter's block and grow rapidly before slowing down, while older, more mature ones that have survived continue growing but at a much slower rate. Furthermore, the upward trends of these older, slower-growing companies all follow an approximate straight line with similar shallow slopes. On this semilogarithmic plot, where the vertical axis (sales) is logarithmic but the horizontal one (time) is linear, a straight line means mathematically that sales are growing exponentially with time. Thus, on average, all surviving companies eventually settle down to a steady but slow exponential growth, as predicted. This is very encouraging, but there's a potential pitfall that becomes apparent when the growth of each company is measured relative to the growth of the overall market. In that case, as can be clearly seen in Figure 70 where the overall growth of the market has been factored out, all large mature companies have stopped growing . Their growth curves when corrected for both inflation and the expansion of the market now look just like typical sigmoidal growth curves of organisms in which growth ceases at maturity, as illustrated in Figures 15–18 of chapter 4. This close similarity with the growth of organisms when viewed in this way provides a natural segue into whether this similarity extends to mortality and whether, like us, all companies are destined to die. 463

This close similarity with the growth of organisms when viewed in this way provides a natural segue into whether this similarity extends to mortality and whether, like us, all companies are destined to die.

October 3, 2023

So can we imagine making an innovation as powerful and influential as the invention of the Internet every fifteen, ten, or even five years? This is a classic reductio ad absurdum argument showing that regardless of how ingenious we are, how many marvelous gadgets and devices we invent, we simply won't be able to overcome the threat of the ultimate singularity if we continue business as usual.

499

October 4, 2023

Blue

I suppose that one could view all religion and philosophical reflection as having its origins in how we integrate the inevitable imminence of death into our daily lives.

22

September 17, 2023

This systematic regularity follows a precise mathematical formula which, in technical parlance, is expressed by saying that "metabolic rate scales as a power law whose exponent is very close to the number ¾." I'll explain this in much greater detail later but here I want to give a simple illustration of what it means colloquially. So consider the following: elephants are roughly 10,000 times (four orders of magnitude, 104) heavier than rats; consequently, they have roughly 10,000 times as many cells. The ¾ power scaling law says that, despite having 10,000 times as many cells to support, the metabolic rate of an elephant (that is, the amount of energy needed to keep it alive) is only 1,000 times (three orders of magnitude, 103) larger than a rat's; note the ratio of 3:4 in the powers of ten.

38

September 17, 2023

Parenthetically, it's worth pointing out that the subsequent decrease in the rates of cellular damage from metabolic processes underlies the greater longevity of elephants and provides the framework for understanding aging and mortality. The scaling law can be expressed in the slightly different way that I used earlier: if an animal is twice the size of another (whether 10 lbs. vs. 5 lbs. or 1,000 lbs. vs. 500 lbs.) we might naively expect metabolic rate to be twice as large, reflecting classic linear thinking. The scaling law, however, is nonlinear and says that metabolic rates don't double but, in fact, increase by only about 75 percent, representing a whopping 25 percent savings with every doubling of size. 12 Notice that the ¾ ratio is just the slope of the graph in Figure 1, where the quantities (metabolic rate and mass) are plotted logarithmically— meaning that they increase by factors of ten along both axes. When plotted this way, the slope of the graph is just the exponent of the power law. This scaling law for metabolic rate, known as Kleiber's law after the biologist who first articulated it, is valid across almost all taxonomic groups, including mammals, birds, fish, crustacea, bacteria, plants, and cells. Even more impressive, however, is that similar scaling laws hold for essentially all physiological quantities and life-history events, including growth rate, heart rate, evolutionary rate, genome length, mitochondrial density, gray matter in the brain, life span, the height of trees and even the number of their leaves. Furthermore, when plotted logarithmically this dizzying array of scaling laws all look like Figure 1 and therefore have the same mathematical structure. They are all "power laws" and are typically governed by an exponent (the slope of the graph), which is a simple multiple of ¼, the classic example being the ¾ for metabolic rate. So, for example, if the size of a mammal is doubled, its heart rate decreases by about 25 percent. The number 4 therefore plays a fundamental and almost magically universal role in all of life. 13

38

September 17, 2023

When humans began forming sizable communities they brought a fundamentally new dynamic to the planet. With the invention of language and the consequent exchange of

44

information in social network space we discovered how to innovate and create wealth and ideas, ultimately manifested in superlinear scaling. In biology, network dynamics constrains the pace of life to decrease systematically with increasing size following the ¼ power scaling laws. In contrast, the dynamics of social networks underlying wealth creation and innovation leads to the opposite behavior, namely, the systematically increasing pace of life as city size increases: diseases spread faster, businesses are born and die more often, commerce is transacted more rapidly, and people even walk faster, all following the approximate 15 percent rule. We all sense that life is faster in the big city than in the small town and that it has ubiquitously accelerated during our lifetimes as cities and their economies grew.

September 17, 2023

Just as bounded growth in biology follows from the sublinear scaling of metabolic rate, the superlinear scaling of wealth creation and innovation (as measured by patent production, for example) leads to unbounded, often faster-than-exponential growth consistent with open-ended economies. This is satisfyingly consistent, but there's a big catch, which goes under the forbidding technical name of a finite time singularity . In a nutshell, the problem is that the theory also predicts that unbounded growth cannot be sustained without having either infinite resources or inducing major paradigm shifts that "reset" the clock before potential collapse occurs. We have sustained open-ended growth and avoided collapse by invoking continuous cycles of paradigm-shifting innovations such as those associated on the big scale of human history with discoveries of iron, steam, coal, computation, and, most recently, digital information technology. Indeed, the litany of such discoveries both large and small is testament to the extraordinary ingenuity of the collective human mind. 45

September 17, 2023

Unfortunately, however, there is another serious catch. 46

Theory dictates that such discoveries must occur at an increasingly accelerating pace; the time between successive innovations must systematically and inextricably get shorter and shorter. For instance, the time between the "Computer Age" and the "Information and Digital Age" was perhaps twenty years, in contrast to the thousands of years between the Stone, Bronze, and Iron ages. If we therefore insist on continuous open-ended growth, not only does the pace of life inevitably quicken, but we must innovate at a faster and faster rate. We are all too familiar with its short-term manifestation in the increasingly faster pace at which new gadgets and models appear. It's as if we are on a succession of accelerating treadmills and have to jump from one to another at an ever-increasing rate. This is clearly not sustainable, potentially leading to the collapse of the entire urbanized socioeconomic fabric. Innovation and wealth creation that fuel social systems, if left unchecked, potentially sow the seeds of their inevitable collapse. Can this be avoided or are we locked into a fascinating experiment in natural selection that is doomed to fail?

September 17, 2023

For organisms, the sublinear scaling of metabolic rate underlies their cessation of growth and a size at maturity that remains approximately stable until death. A similar life-history trajectory is at work for companies. They grow rapidly in their early years but taper off as they mature and, if they survive, eventually stop growing relative to the GDP. In their youth, many are dominated by a spectrum of innovative ideas as they seek to optimize their place in the market. However, as they grow and become more established, the spectrum of their product space inevitably narrows and, at the same time, they need to build a significant administration and bureaucracy. Relatively quickly, economies of scale and sublinear scaling, reflecting the challenge of efficiently administering a large and complex organization, dominate innovation and ideas encapsulated in superlinear scaling, ultimately leading to stagnation and to mortality. Half of all the companies in any given cohort of U.S. publicly traded companies disappear within ten years, and a scant few make it to fifty, let alone a 47 *hundred years. 16*

September 17, 2023

A simple example of the first kind of innovation is to use a stronger material such as steel in place of wood for bridges or buildings, while a simple example of the second kind is to use arches, vaults, or domes in their construction rather than just horizontal beams and vertical pillars. The evolution of bridges is, in fact, an excellent example of how innovations in both materials and design were stimulated by the desire, or perceived requirement, to meet new challenges: in this case, to traverse wider and wider rivers, canyons, and valleys in a safe, resilient manner. 77

September 18, 2023

Brunel thought otherwise. His conclusions were based on a simple scaling argument. He realized that the volume of cargo a ship could carry increases as the cube of its dimensions (like its weight), whereas the strength of the drag forces it experiences as it travels through water increases as the cross-sectional area of its hull and therefore only as the square of its dimensions. This is just like Galileo's conclusions for how the strength of beams and limbs scale with body weight. In both cases the strength increases more slowly than the corresponding weight following a ⅔ power scaling law. Thus the strength of the hydrodynamic drag forces on a ship relative to the weight of the cargo it can carry decreases in direct proportion to the length of the ship. 84

September 18, 2023

This is manifested in all kinds of fascinating behaviors and patterns such as those we see in the eddies and whirlpools of rivers and streams, or the wakes left by ships as they move 91 *through water, or the awesome specter of hurricanes and tornadoes and the beauty and infinite variation of ocean waves. All of these are manifestations of turbulence and are encapsulated in the hidden richness of the Navier-Stokes equation. Indeed, studying turbulence gave us the first important mathematical insights into the concept of complexity and its relationship to nonlinearity. Complex systems often manifest chaotic behavior in which a small change or perturbation in one part of the system produces an exponentially enhanced response in some other part. As we discussed earlier, in traditional linear thinking a small perturbation produces a commensurately small response. The highly nonintuitive enhancement in nonlinear systems is popularly expressed as "the butterfly effect," in which the mythical flapping of a butterfly's wings in Brazil produces a hurricane in Florida. Despite 150 years of intense theoretical and experimental study, a general understanding of turbulence remains an unsolved problem in physics even though we have learned an enormous amount about it. Indeed the famous physicist Richard Feynman described turbulence as "the most important unsolved problem of classical physics." 15*

September 19, 2023

Like many new ideas that threaten to change the way we think about an old problem, Froude's endeavors were at first dismissed as irrelevant by the cognoscenti of the time. John Russell, who founded the Institution for Naval Architects in England in 1860 in order to encourage the formal education of ship designers, ridiculed Froude: "You will have on the small scale a series of beautiful, interesting little experiments which, I am sure, will afford Mr. Froude infinite pleasure in the making of them . . . and will afford you infinite pleasure in the hearing of them; but which are quite remote from any practical results upon the large scale." Many of us recognize this kind of rhetoric, often aimed at scholarly or academic research with the implication that it is out of touch with the "real world." Well, no doubt, much of it is. But much of it isn't and, more to the point, it is very often difficult to perceive in the moment the potential impact of some piece of seemingly arcane research. Much of our entire technologically driven

92

society and extraordinary quality of life that many of us are privileged to enjoy is the result of such research. There is a continuing tension in society between supporting what is perceived as pie-in-the-sky basic research with no obvious immediate benefit versus highly directed research focused on "useful, real world" problems.

September 19, 2023

He realized that the primary quantity that determined the character of their relative motion was something that later became known as the Froude number . It is defined as the square of the ship's velocity divided by its length multiplied by the acceleration due to gravity. This is a bit of a mouthful and may sound a little intimidating, but it's actually quite simple because "the acceleration due to gravity" that occurs in this expression is the same for all objects, regardless of their size, shape, or composition. This is just a restatement of Galileo's observations that falling objects of different weights reach the ground in the same time. So in terms of the quantities that actually vary, Froude's number is simply proportional to the velocity squared divided by the length. This ratio plays a central role in all problems involving motion, ranging from speeding bullets and running dinosaurs to flying airplanes and sailing ships. 93

September 19, 2023

The field of physics is concerned with fundamental principles and concepts at all levels of organization that are quantifiable and mathematizable (meaning amenable to computation), and can consequently lead to precise predictions that can be tested by experiment and observation. From this perspective, it is natural to ask if there are "universal laws of life" that are mathematizable so that biology could also be formulated as a predictive, quantitative science much like physics. Is it conceivable that there are yet-to-be-discovered "Newton's Laws of Biology" that would lead, at least in principle, to precise calculations of any biological process so that, for instance, one could 103 *accurately predict how long you and I would live? This seems very unlikely. After all, life is the complex system par excellence, exhibiting many levels of emergent phenomena arising from multiple contingent histories. Nevertheless, it may not be unreasonable to conjecture that the generic coarse-grained behavior of living systems might obey quantifiable universal laws that capture their essential features. This more modest view presumes that at every organizational level average idealized biological systems can be constructed whose general properties are calculable. Thus we ought to be able to calculate the average and maximum life span of human beings even if we'll never be able to calculate our own. This provides a point of departure or baseline for quantitatively understanding actual biosystems, which can be viewed as variations or perturbations around idealized norms due to local environmental conditions or historical evolutionary divergence. I will elaborate on this perspective in much greater depth below, as it forms the conceptual strategy for attacking most of the questions posed in the opening chapter.*

September 19, 2023

Although Thompson's book did not address aging or death, nor was it particularly helpful or sophisticated technically, its philosophy provided support and inspiration for contemplating and applying ideas and techniques from physics to all sorts of problems in biology. In my own thinking, this led me to perceive our bodies as metaphorical machines that need to be fed, maintained, and repaired but which eventually wear out and "die," much like cars and washing machines. However, to understand how something ages and dies, whether an animal, an automobile, a company, or a civilization, one first needs to understand what the processes and mechanisms are that are keeping it alive, and then discern how these become degraded with time. This naturally leads to considerations of the energy and resources that are required for sustenance and possible growth, and their allocation to maintenance and repair for combating the production of entropy arising from destructive forces associated with damage, disintegration, wear and tear, and so on. This line of thinking led me to focus 111

initially on the central role of metabolism in keeping us alive before asking why it can't continue doing so forever.

September 19, 2023

4. UNIVERSALITY AND THE MAGIC NUMBER FOUR THAT CONTROLS LIFE The systematic regularity of Kleiber's law is pretty amazing, but equally surprising is that similar systematic scaling laws hold for almost any physiological trait or life-history event across the entire range of life from cells to whales to ecosystems. In addition to metabolic rates, these include quantities such as growth rates, genome lengths, lengths of aortas, tree heights, the amount of cerebral gray matter in the brain, evolutionary rates, and life spans; a sampling of these is illustrated in Figures 9-12. There are probably well over fifty such scaling laws and—another big surprise—their corresponding exponents (the analog of the ¾ in Kleiber's law) are invariably very close to simple multiples of ¼. For example, the exponent for growth rates is very close to ¾, for lengths of aortas and genomes it's ¼, for heights of trees ¼, for cross-sectional areas of both aortas and tree trunks ¾, for brain sizes ¾, for cerebral white and gray matter 5⁄4, for heart rates minus ¼, for mitochondrial densities in cells minus ¼, for rates of evolution minus ¼, for diffusion rates across membranes minus ¼, for life spans ¼ . . . and many, many more. The "minus" here simply indicates that the corresponding quantity decreases with size rather than increases, so, for instance, heart rates decrease with increasing body size following the ¼ power law, as shown in Figure 10. I can't resist drawing your attention to the intriguing fact that aortas and tree trunks scale in the same way. Particularly fascinating is the emergence of the number four in the guise of the ¼ powers that appear in all of these exponents. It occurs ubiquitously across the entire panoply of life and seems to play a special, fundamental role in determining many of the measurable characteristics of organisms regardless of their evolved design. Viewed through the lens of scaling, a remarkable general universal pattern emerges, strongly suggesting that evolution has been constrained by other general physical principles beyond natural selection. 117

September 19, 2023

Huxley's term allometric was extended from its more restrictive geometric, morphological, and ontogenetic origins to describe the kinds of scaling laws that I discussed above, which include more dynamical phenomena such as how flows of energy and resources scale with body size, with metabolic rate being the prime example. All of these are now commonly referred to as allometric scaling laws . Julian Huxley, himself a very distinguished biologist, was the grandson of the famous Thomas Huxley, the biologist who championed Charles Darwin and the theory of evolution by natural selection, and the brother of the novelist and futurist Aldous Huxley. In addition to the word allometric, Julian Huxley brought several other new words and concepts into biology, including replacing the much-maligned term race with the phrase ethnic group .

122

September 20, 2023

As I have emphasized, no aspect of life can function without energy. Just as every muscle contraction or any activity requires metabolic energy, so does every random thought in your brain, every twitch of your body even while you sleep, and even the replication of your DNA in your cells. At the most fundamental biochemical level metabolic energy is created in semiautonomous molecular units within cells called respiratory complexes . The critical molecule that plays the central role in metabolism goes by the slightly forbidding name of adenosine triphosphate , usually referred to as ATP. The detailed biochemistry of metabolism is extremely complicated but in essence it involves the breaking down of ATP, which is relatively unstable in the cellular environment, from adenosine tri phosphate (with three phosphates) into ADP, adenosine di phosphate (with just two phosphates), thereby releasing the energy stored in the binding of the additional phosphate. The energy derived from breaking this phosphate bond is the source of your metabolic energy and therefore what is keeping you alive. The reverse process converts ADP back into ATP using

124

energy from food via oxidative respiration in mammals such as ourselves (that's why we have to breathe in oxygen), or via photosynthesis in plants. The cycle of releasing energy from the breakup of ATP into ADP and its recycling back from ADP to store energy in ATP forms a continuous loop process much like the charging and recharging of a battery.

September 20, 2023

Given its central role, it's not surprising that the flux of ATP is often referred to as the currency of metabolic energy for almost all of life. At any one time our bodies contain only about half a pound (about 250 g) of ATP, but here's something truly extraordinary that you should know about yourself: every day you typically make about 2 × 1026 ATP molecules—that's two hundred trillion trillion molecules—corresponding to a mass of about 80 kilograms (about 175 lbs.). In other words, each day you produce and recycle the equivalent of your own body weight of ATP! Taken together, all of these ATPs add up to meet our total metabolic needs at the rate of the approximately 90 watts we require to stay alive and power our bodies. 125

September 20, 2023

These little energy generators, the respiratory complexes, are situated on crinkly membranes inside mitochondria , which are potato-shaped objects floating around inside cells. Each mitochondrion contains about five hundred to one thousand of these respiratory complexes . . . and there about five hundred to one thousand of these mitochondria inside each of your cells, depending on the cell type and its energy needs. Because muscles require greater access to energy, their cells are densely packed with mitochondria, whereas fat cells have many fewer. So on average each cell in your body may have up to a million of these little engines distributed among its mitochondria working away night and day, collectively manufacturing the astronomical number of ATPs needed to keep you viable, healthy, and strong. The rate at which the total number of these ATPs is produced is a measure of your 125 *metabolic rate. Your body is composed of about a hundred trillion (1014) cells. Even though they represent a broad range of diverse functionalities from neuronal and muscular to protective (skin) and storage (fat), they all share the same basic features. They all process energy in a similar way via the hierarchy of respiratory complexes and mitochondria. Which raises a huge challenge: the five hundred or so respiratory complexes inside your mitochondria cannot behave as independent entities but have to act collectively in an integrated coherent fashion in order to ensure that mitochondria function efficiently and deliver energy in an appropriately ordered fashion to cells. Similarly, the five hundred or so mitochondria inside each of your cells do not act independently but, like respiratory complexes, have to interact in an integrated coherent fashion to ensure that the 1014 cells that constitute your body are supplied with the energy they need to function efficiently and appropriately. Furthermore, these hundred trillion cells have to be organized into a multitude of subsystems such as your various organs, whose energy needs vary significantly depending on demand and function, thereby ensuring that you can do all of the various activities that constitute living, from thinking and dancing to having sex and repairing your DNA. And this entire interconnected multilevel dynamic structure has to be sufficiently robust and resilient to continue functioning for up to one hundred years!*

September 20, 2023

As I began to ponder what the origins of these surprising scaling laws might be, it became clear that whatever was at play had to be independent of the evolved design of any specific type of organism, because the same laws are manifested by mammals, birds, plants, fish, crustacea, cells, and so on. All of these organisms ranging from the smallest, simplest bacterium to the largest plants and animals depend for their maintenance and reproduction on the close integration of numerous subunits—molecules, organelles, and cells—and these microscopic components need to be serviced in a relatively "democratic" and efficient fashion in order to supply metabolic substrates, remove waste products, and regulate activity. Natural selection has solved 128 *this challenge in perhaps the simplest possible way by evolving hierarchical branching networks that distribute energy and materials between macroscopic reservoirs and microscopic sites. Functionally, biological systems are ultimately constrained by the rates at which energy, metabolites, and information can be supplied through these networks. Examples include animal circulatory, respiratory, renal, and neural systems, plant vascular systems, intracellular networks, and the systems that supply food, water, power, and information to human societies. In fact, when you think about it, you realize that underneath your smooth skin you are effectively an integrated series of such networks, each busily transporting metabolic energy, materials, and information across all scales.*

September 20, 2023

In high energy physics, where we struggle to unravel the basic laws of nature at the most microscopic level, we mostly know what the questions are and most of one's effort goes into trying to be clever enough to carry out the highly technical calculations. In biology I found it to be mostly the other way around: months were spent trying to figure out what the problem actually was that we were trying to solve, the questions we should be asking, and the various relevant quantities that were needed to be calculated, but once that was accomplished, the actual technical mathematics was relatively straightforward. 132

September 20, 2023

The time seems right for revisiting D'Arcy Thompson's challenge: "How far even then mathematics will suffice to describe, and physics to explain, the fabric of the body, no man can foresee. It may be that all the laws of energy, and all the properties of matter, all . . . chemistry . . . are as powerless to explain the body as they are impotent to comprehend the soul. For my part, I think it is not so." Many would agree with the spirit of this remark, though new tools and concepts, including closer collaboration, may well be 138 *needed to accomplish his lofty goal.*

September 20, 2023

The final postulate states that the continuous multiple feedback and fine-tuning mechanisms implicit in the ongoing processes of natural selection and which have been playing out over enormous periods of time have led to the network performance being "optimized." So, for example, the energy used by the heart of any mammal, including us, to pump blood through the circulatory system is on average minimized. That is, it is the smallest it could possibly be given its design and the various network constraints. To put it slightly differently: of the infinite number of possibilities for the architecture and dynamics of circulatory systems that could have evolved, and that are space filling with invariant terminal units, the ones that actually did evolve and are shared by all mammals minimize cardiac output. Networks have evolved so that the energy needed to sustain an average individual's life and perform the mundane tasks of living is minimized in order to maximize the amount of energy available for sex, reproduction, and the raising of offspring. This maximization of offspring is an expression of what is referred to as Darwinian fitness, w hich is the genetic contribution of an average individual to the next generation's gene pool. This naturally raises the question as to whether the dynamics and structure of cities and companies are the result of analogous optimization principles. What, if anything, is optimized in their multiple network systems? Are cities organized to maximize social interactions, or to optimize transport by minimizing mobility times, or are they ultimately driven by the ambition of each citizen and company to maximize their assets, profits, and wealth? I will return to these issues in chapters 8, 9, and 10. 142

September 20, 2023

Their modern formulation is a general mathematical framework in which a quantity called the action , which is loosely related to energy, is minimized. All the laws of 143 *physics can be derived from the principle of least action which, roughly speaking, states that, of all the possible configurations that a system can have or that it can follow as it evolves in time, the one that is physically realized is the one that minimizes its action. Consequently, the dynamics, structure, and time evolution of the universe since the Big Bang, everything from black holes and the satellites transmitting your cell phone messages to the cell phones and messages themselves, all electrons, photons, Higgs particles, and pretty much everything else that is physical, are determined from such an optimization principle. So why not life?*

September 20, 2023

It is important to recognize that the three postulates enunciated above are to be understood in a coarse-grained average sense. Let me explain. It may have occurred to you that there must be variation among the almost trillion capillaries in any individual human body, as there must be across all the species of a given taxonomic group, so strictly speaking capillaries cannot be invariant. However, this variation has to be viewed in a relative scale-dependent way. The point is that any variation among capillaries is extremely small compared with the many orders of magnitude variation in body size. For instance, even if the length of mammalian capillaries varied by a factor of two, this is still tiny compared with the factor of 100 million in the variation of their body masses. Similarly, there is relatively little variation in petioles, the last branch of a tree prior to the leaf, or even in the size of leaves themselves, during the growth of a tree from a tiny sapling to a mature tree that might be a hundred or more feet high. This is also true across species of trees: leaves do vary in size but by a relatively small factor, despite huge factors in the variation of their heights and masses. A tree that is just twenty times taller than another does not have leaves whose diameter is twenty times larger. Consequently, the variation among terminal units within a given design is a relatively small secondary effect. The same goes for possible variations in the other postulates: networks may not be precisely space filling or precisely optimized. Corrections due to such deviations and 144 *variations are considered to be "higher order" effects in the sense we discussed earlier.*

September 20, 2023

Carrying out this strategy proved to be quite a challenge, both conceptually and technically. It took almost a year to iron out all of the details, but ultimately we showed how Kleiber's law for metabolic rates and, indeed, quarter-power scaling in general arises from the dynamics and geometry of optimized space-filling branching networks. Perhaps most satisfying was to show how the magic number four arises and where it comes from. 12 [145]

September 20, 2023

Similarly, the rate at which oxygen is inhaled through our mouths and into the respiratory system is also a measure of metabolic rate. These two systems are tightly coupled together so blood flow rates, respiratory rates, and metabolic rates are all proportional to one another and related by simple linear relationships. Thus, hearts beat approximately four times for each breath that is inhaled, regardless of the size of the mammal. This tight coupling of the oxygen delivery systems is why the properties of the cardiovascular and respiratory networks play such an important role in determining and constraining metabolic rate. [146]

September 20, 2023

To avoid this potential problem and minimize the work our hearts have to do, the geometry of our circulatory systems has evolved so that there are no reflections at any branch point throughout the network. The mathematics and physics of how this is accomplished is a little bit complicated, but the result is simple and [149] *elegant: the theory predicts that there will be no reflections at any branch point if the sum of the cross-sectional areas of the daughter tubes leaving the branch point is the same as the cross-sectional area of the parent tube coming into it . As an example, consider the simple case where the two daughter tubes are identical and therefore have the same cross-sectional areas (which is approximately correct in real circulatory systems). Suppose that the cross-sectional area of the parent tube is 2 square inches; then, in order to ensure that there are no reflections, the cross-sectional area of each daughter has to be 1 square inch. Because the cross-sectional area of any vessel is proportional to the square of its radius, another way of expressing this result is to say that the square of the radius of the parent tube has to be just twice the square of the radius of each of the daughters. So to ensure that there is no energy loss via reflections as one progresses down the network, the radii of successive vessels must scale in a regular self-similar fashion, decreasing by a constant factor of the square root of two (√2) with each successive branching . This so-called area-preserving branching is, indeed, how our circulatory system is constructed, as has been confirmed by detailed measurements across many mammals—and many plants and trees.*

September 21, 2023

It's a lovely thought that the optimum design of our circulatory system obeys the same simple area-preserving branching rules that trees and plants do. It's equally satisfying that the condition of nonreflectivity of waves at 151 *branch points in pulsatile networks is essentially identical to how national power grids are designed for the efficient transmission of electricity over long distances. This condition of nonreflectivity is called impedance matching. It has multiple applications not only in the working of your body but across a very broad spectrum of technologies that play an important part in your daily life. For example, telephone network systems use matched impedances to minimize echoes on long-distance lines; most loudspeaker systems and musical instruments contain impedance matching mechanisms; and the bones in the middle ear provide impedance matching between the eardrum and the inner ear. If you have ever witnessed or been subject to an ultrasound examination you will be familiar with the nurse or technician smearing a gooey gel over your skin before sliding the probe over it. You probably thought that this was for lubrication purposes but in fact it's actually for matching impedances. Without the gel, the impedance mismatch in ultrasound detection would result in almost all of the energy being reflected back from the skin, leaving very little to go into the body to be reflected back from the organ or fetus under investigation. The term impedance matching can be a very useful metaphor for connoting important aspects of social interactions.*

September 21, 2023

This is in marked contrast to what we normally see when, for example, we use a microscope to zoom in on an object using a higher and higher resolution in order to reveal greater detail and new structure that is qualitatively different from those of the whole. Obvious examples are cells in tissue, molecules in materials, or protons in atoms. If, on the other hand, the object is a fractal, no new pattern or detail arises when the resolution is increased: the same pattern repeats itself over and over again. In reality, this is an idealized description for, of course, the images at various levels of resolution differ very slightly from one another and eventually the recursive repetition ceases and new patterns of structural design appear. If you continue breaking down broccoli to increasingly smaller pieces, these eventually lose the geometric characteristics of broccoli and eventually reveal 157 *the structure of its tissue, its cells, and its molecules.*

September 21, 2023

The space-filling requirement that the network must service the entire volume of the organism at all scales also requires it to be self-similar in terms of the lengths of the vessels. To fill the three-dimensional space, lengths of successive vessels have to decrease by a constant factor of 3√2 with each successive branching and, in contrast to radii, this remains valid down through the entire network, including both the pulsatile and nonpulsatile domains. Having determined how networks scale within individuals following these simple rules, the last piece of the derivation is to determine how this connects across species of different weights. This is accomplished from a further consequence of the energy minimization principle: namely, that the total volume of the network—that is, the total volume of blood in the body—must be directly proportional to the volume of the body itself, and therefore proportional to its weight, as observed. In other words, the volume of blood is a constant proportion of the volume of the body, regardless of size. For a tree this is obvious because the network of its vessels constitutes the entire tree—there is no analog of flesh in between all of its branches, so the volume of the network is the volume of the tree. 16 Now, the volume of the network is just the sum of the volumes of all of its vessels or branches, and these can be straightforwardly calculated from knowing how their lengths and radii scale, thereby connecting the self-similarity of the internal network to 160 *body size. It is the mathematical interplay between the cube root scaling law for lengths and the square root scaling law for radii, constrained by the linear scaling of blood volume and the invariance of the terminal units, that leads to quarter-power allometric exponents across organisms. The resulting magic number four emerges as an effective extension of the usual three dimensions of the volume serviced by the network by an additional dimension resulting from the fractal nature of the network. I shall go into this in more detail in the following chapter, where I discuss the general concept of fractal dimension , but suffice it to say here that natural selection has taken advantage of the mathematical marvels of fractal networks to optimize their distribution of energy so that organisms operate as if they were in four dimensions , rather than the canonical three. In this sense the ubiquitous number four is actually 3 + 1. More generally, it is the dimension of the space being serviced plus one. So had we lived in a universe of eleven dimensions, as some of my string theory friends believe, the magic number would have been 11 + 1 = 12, and we would have been talking about the universality of 1/12 power scaling laws rather than $\frac{1}{4}$ power ones.*

September 21, 2023

Stimulated by his passionate pacifism, Richardson embarked on an ambitious program to develop a quantitative theory for understanding the origins of war and international conflict in order to devise a strategy for their ultimate prevention. His aim was nothing less than to develop a science of war. His main thesis was that the dynamics of conflict are primarily governed by the rates at which nations 164 *build up their armaments and that their continued accumulation is the major cause of war. He viewed the accumulation of weapons as a proxy for the collective psychosocial forces that reflect, but transcend, history, politics, economics, and culture and whose dynamics inevitably lead to conflict and instability. Richardson used the mathematics developed for understanding chemical reaction dynamics and the spread of communicable diseases to model the ever-increasing escalation of arms races in which the arsenal of each country increases in response to the increase in armaments of every other country.*

September 21, 2023

In the natural world almost nothing is smooth—most things are crinkly, irregular, and crenulated, very often in a self-similar way. Just think of forests, mountain ranges, vegetables, clouds, and the surfaces of oceans. Consequently, most physical objects have no absolute objective length, and it is crucial to quote the resolution when stating the measurement. So why did it take more than two thousand years for people to recognize something so basic and which now seems almost obvious? Very likely this has its origins in the duality that emerged as we gradually separated from a close connection to the natural world and became more and more distant from the forces of nature that have determined our biology. Once we invented language, learned how to take advantage of economies of scale, formed communities, and began making artifacts, we effectively changed the geometry of our daily world and its immediate surroundings. In designing and manufacturing human-engineered artifacts, whether primitive pots and tools or modern sophisticated automobiles, computers, and skyscrapers, we employed and aspired to the simplicity of straight lines, smooth curves, and smooth surfaces. This was brilliantly formalized and reflected in the development of quantified measurement and the invention of mathematics, manifested, in particular, in the idealized paradigm of Euclidian geometry. This is the mathematics appropriate to the world of artifacts we created around us as we evolved from being a mammal like any other to become social Homo sapiens.

174

September 21, 2023

This can be extended, at least metaphorically, beyond individuals to companies, cities, states, and even life itself. Being diverse and having many interchangeable, adaptable components is another manifestation of this paradigm. Natural selection thrives on and consequently manufactures greater diversity. Resilient ecosystems have greater diversity of species. It is no accident that successful cities are those that offer a greater spectrum of job opportunities and businesses, and that successful companies have a diversity of products and people with the flexibility to change, adapt, and reinvent in response to changing markets. I shall discuss this further in chapters 8 and 9 when I turn to cities and companies.

177

September 21, 2023

Natural selection has taken advantage of the fractal nature of space-filling networks to maximize the total effective surface area of these terminal units and thereby maximize metabolic output. Geometrically, the nested levels of continuous branching and crenulations inherent in fractal-like structures optimize the transport of information, energy, and resources by maximizing the surface areas across which these essential features of life flow. Because of their fractal nature, these effective surface areas are very much larger than their apparent physical size. Let me give you some remarkable examples from your own body to illustrate the point. Even though your lungs are only about the size of a football with a volume of about 5 to 6 liters (about one and a half gallons), the total surface area of the alveoli, which are the terminal units of the respiratory system where oxygen and carbon dioxide are exchanged with the blood, is almost the size of a tennis court and the total length of all the airways is about 2,500 kilometers, almost the distance from Los Angeles to Chicago, or London to Moscow. Even more striking is that if all the arteries, veins, and capillaries of your circulatory system were laid end to end, their total length would be about 100,000 kilometers, or nearly two and a half times around the Earth or over a third of the distance

186

to the moon . . . and all of this neatly fits inside your five-to-six-foot-tall body. It's quite fantastic and yet another amazing feature of your body where natural selection has exploited the wonders of physics, chemistry, and mathematics.

September 22, 2023

As explained in the previous chapter, a crinkly enough line that is space filling can scale as if it's an area. Its fractality effectively endows it with an additional dimension. Its conventional Euclidean dimension, discussed in chapter 2, still has the value 1, indicating that it's a line, but its fractal dimension is 2, indicating that it's maximally fractal and scaling as if it were an area. In a similar fashion an area, if crinkly enough, can behave as if it's a volume, thereby gaining an effective extra dimension: its Euclidean dimension is 2, indicating that it's an area, but its fractal dimension is 3. A familiar example will make this clear. Think of washing sheets. Being sensitive to conserving energy and at the same time wanting to save yourself money and time, you wait several weeks until you have more than a sufficient number of dirty ones to fill the entire tub of your washing machine. So when the time comes you stuff in as much and as many as you possibly can to fill the entire volume of the tub. Now, recall that ordinary volumes scale faster than areas, so if you were to double the size of your washing machine by doubling all of its lengths while keeping its shape the same, its volume would increase by a factor of eight (23) whereas all of its surface areas would increase by a factor of four (22). Naively, you might therefore conclude that because sheets are essentially all area and consequently two-dimensional (their thickness being negligible), you could accommodate four times as many sheets by doubling the size of your washing machine. However, if we stuff all of the sheets into the tub so that they completely fill its entire volume and because this volume has increased by a factor of eight, then it's clear that you can actually accommodate eight times as many sheets, rather than just four times. In other words, the total effective area of two-dimensional sheets filling three-dimensional washing machines scales like a volume rather than an area, so in this sense, we have turned 187 *an area into a volume.*

September 22, 2023

Quarter-power scaling laws are perhaps as universal and as uniquely biological as the biochemical pathways of metabolism, the structure and function of the genetic code, and the process of natural selection. The vast majority of organisms exhibit scaling exponents very close to ¾ for metabolic rate and ¼ for internal times and distances. These are the maximal and minimal values, respectively, for the effective surface area and linear dimensions of a volume-filling fractal-like network. 189

September 22, 2023

Now imagine continuously decreasing the size of the animal. Concomitantly, the number of area-preserving branchings where vessels are large enough to support pulsatile waves decreases until a tipping point is reached where the network can support only nonpulsatile DC flow. At that stage even the major arteries become so small and constricted that they are unable to support pulsatile waves. In such vessels, waves become so overdamped due to the viscosity of blood that they can no longer propagate and the flow shifts to becoming entirely steady DC, just like the flow of water in the pipes of your house: pulsatile waves generated by the beating heart are immediately damped as they enter the aorta. 192

September 22, 2023

The theory has been extended to understand the growth of tumors, plants, insects, and both forest 5 and social insect communities 6 such as ants and bees. These latter applications are forerunners of how we might start thinking about the growth of human organizations such as cities and companies, which I'll be turning to in chapters 8 and 9. Each of these very different systems represents a variation on the 205 *general thematic structure of the growth equation. For instance, tumors are parasitic and use metabolic energy derived from their host to grow, so their vasculature and metabolic rates depend not only on their own size but also on the size of the host. 7 Understanding this provides insight into how to scale up basic properties of tumors as well as potential therapeutic strategies from observations on mice to humans. 8*

September 22, 2023

The overall agreement with theory is extremely satisfying. But much more than that, I find the extraordinary unity and interconnectivity of life that is revealed through this lens to be spiritually elevating in the pantheistic spirit articulated by the philosopher Baruch Spinoza. As Einstein wrote, 10 "We followers of Spinoza see our God in the wonderful order and lawfulness of all that exists and in its soul as it reveals itself in man and animal."

209

September 22, 2023

Because the theory is deliberately simplified, it is inevitable that measurements of real organisms will, to varying degrees, deviate from model predictions. As can be seen in Figure 19 the agreement is surprisingly good with relatively few major outliers that deviate significantly from the idealized growth curve. We as primates are one of those. For instance, we take longer to mature than we "should" given our body weight. This is the result of our rapid evolution from being purely biological to becoming sophisticated socioeconomic creatures. Our effective metabolic rate is now one hundred times greater than what it was when we were truly "biological" animals, and this has had huge consequences for our recent life history. We take longer to mature, we have fewer offspring, and we live longer, all in qualitative agreement with having an effectively larger metabolic rate arising from socioeconomic activity. I shall return to this fascinating development in our history when discussing how these ideas apply to cities.

209

September 22, 2023

This parsimonious formulation of the coarse-grained mass and temperature dependence was introduced as a compact summary of the scaling work in a paper titled "Toward a Metabolic Theory of Ecology" published in the journal Ecology in 2004 and coauthored by Jim Brown and three of our then postdocs—Van Savage, Jamie Gillooly, and Drew Allen—together with me. 211

September 23, 2023

Thus at a deep level, birth, growth, and death are all governed by the same underlying dynamics driven by metabolic rate and encapsulated in the dynamics and structure of networks. 213

September 23, 2023

This is the curse of consciousness. We all know we are going to die. No other organism is burdened with the enormity of the conscious knowledge that it has a finite lifetime and that its individual existence is eventually and inevitably coming to an end. No creature, whether a bacterium, an ant, a rhododendron, or a salmon, "cares" or even "knows" about dying; they live and they die, participating in the continual struggle for existence by propagating their genes into future generations and playing the endless game of the survival of the fittest. So do we. 216

September 23, 2023

Once upon a time, science was referred to as natural philosophy, implying a somewhat broader connotation than the way we think of it today with a greater connection to 219 *philosophical and religious thought. It is no accident that the full title of Newton's famous book the Principia, which introduced his universal laws of nature that revolutionized science, is (in English) The Mathematical Principles of Natural Philosophy . Although Newton held heretical views such as rejecting the classical doctrines of an immortal soul, the existence of devils and demons, and the worship of Christ as God, which he viewed as idolatrous, he saw his work as the revelation of God as a prime mover. Commenting on the Principia, he stated: "When I wrote my treatise about our Systeme I had an eye upon such Principles as might work with considering men for the beliefe of a Deity and nothing can rejoyce me more than to find it useful for that purpose."*

September 23, 2023

How did the universe evolve, what are stars made of, where did all the different animals and plants come from, why is the sky blue, when will the next eclipse occur, and so on and so forth. We understand an enormous amount about the physical universe around us and in many cases with exquisite detail, and we have done it without having to invoke ad hoc or arbitrary arguments that are often the hallmark of religious explanations. Left unanswered, however, are many of the deep questions concerning the very nature of who and what we are as human beings endowed with consciousness and the ability to reflect and reason. We continue to grapple with the nature of mind and consciousness, with psyche and self, with love and hate, and with meaning and purpose. Perhaps all will eventually be understood from the firing of neurons and the complex network dynamics of our brains, but as D'Arcy Thompson proclaimed a hundred years ago, I suspect not. There will always be questions—that is the essence of the human condition—and like Antonius Block, we will never stop asking them even if it is to the great frustration and annoyance of Death. And somehow intertwined with all of this lies the challenge and paradox of understanding aging and mortality, and coming to terms with our collective and individual uneasiness with the finiteness of our own existence. 219

September 23, 2023

The city as the engine for social change and increasing well-being is one of the truly great triumphs of our amazing ability to form social groups and collectively take advantage of economies of scale.

225

September 23, 2023

The concept of a maximum life span is an extremely important one, as it implies that without some major "unnatural" intervention (which is what the elixir believers are seeking), natural processes inextricably limit human life span to about 125 years. Below, I will address what these limiting processes are and present a theoretical framework based on the network theory for determining this number. Before doing so, however, I want to show how classic survivorship curves provide powerful evidence in support of the concept of a maximum human life span. A survivorship curve simply represents the probability that an individual will live to a given age and is determined by plotting the percentage of survivors in a given population as a function of their age. Its converse is called a mortality curve and is the percentage of people who have died at a given age, representing the probability that an individual will die at that age. Biologists, actuaries, and gerontologists have coined the term mortality or death rate to denote the number of deaths in a population that occur in some given period of time (a month, say) relative to the number that are still alive. The general structure of survivorship and mortality curves is pretty obvious: most individuals survive the earliest years, but gradually a larger and larger percentage die until a point is reached where the probability of surviving eventually vanishes while the probability of dying reaches 100 percent. A great deal of statistical analysis has been done on such curves across different societies, cultures, environments, and species .

229

September 23, 2023

One of the surprising results that emerges is that the mortality rate for most organisms remains approximately unchanged with age. In other words, the relative number of individuals that die in any time period is the same at any age. So, for example, if 5 percent of the surviving population dies between ages five and six, then 5 percent of the surviving population will also die between ages forty-five and forty-six and between ninety-five and ninety-six. This sounds nonintuitive, but if we put it a different way it will make more sense. A constant mortality rate means that the number of individuals that die in some time period is directly proportional to how many have survived up until that time. If you go back to the earlier discussion on exponential behavior in chapter 3, you will discover that this is precisely the mathematical definition of the exponential function , which I will discuss in much greater detail in the following chapter. Here it says that survivorship follows a simple exponential curve, meaning that it becomes exponentially less likely that an individual from the original population will survive the older it gets, or equivalently, that it becomes exponentially more likely that an individual will die the older it gets. 230

September 24, 2023

These are in distinct contrast to the ¼ power scaling laws obeyed by organisms, which result from their optimized fractal-like network structures: their metabolic rates (their horsepower) scale with an exponent of ¾ and their heart rates (their RPMs) with an exponent of -¼. The fact that internal combustion engines have no complex network structures and do not follow ¼ power scaling is supporting evidence for the underlying network theory for the origin of ¼ power scaling in biology. Because manufactured engines satisfy classic cubic scaling, one might speculate that their life span increases with the cube root of their weight rather than with a quarter power. Unfortunately, there isn't sufficient data available to test this. However, qualitatively it does predict that bigger automobiles should last longer. In fact, all of the top ten longest-lasting vehicles are either large 240 *trucks or SUVs, while only three regular-size sedans make it into the top twenty. If you're simply looking for longevity, buy big: the Ford F-250 is tops, with the Chevrolet Silverado second and the Suburban third. Cars are now typically expected to last for about 150,000 miles. In fact, like the humans who made them, their longevity has increased dramatically over a relatively short period of time, having almost doubled over the past fifty years. Just to get an idea of what this implies, suppose that when averaged over its lifetime, a typical car moves at 30 miles per hour and that its "heart rate" is 2,500 RPMs, then the total number of "engine beats" during a 150,000-mile lifetime is approximately a billion. Amusingly, this is not so different from the number of times a mammalian heart beats in its lifetime. Is this just a coincidence, or is it telling us something about the commonality of mechanisms responsible for aging?*

El Bocho de Mi Pá.

September 24, 2023

Notice that just as we saw when discussing growth, the mismatch between the scaling of the sources of energy and therefore the sources of damage (the terminal units) and the scaling of the sinks of energy (cells that need to be sustained) has enormous consequences. In the one case, it ensures that we cease growing, and in the other it ensures that larger animals have expanded lifetimes. And all of this follows from the constraint of the networks. 246

September 24, 2023

A. TEMPERATURE AND EXTENDING LIFE SPAN: Because metabolic rate is proportional to the number of terminal units and these are where most damage occurs, we can relate life span directly to metabolic rate. This results in an alternative expression for life span as the ratio of the body mass to its metabolic rate. In other words, life span is inversely proportional to the metabolic rate per unit mass of the organism and therefore inversely proportional to the 246 *average metabolic rate of its cells. We saw just above when discussing the metabolic theory of ecology that this systematically scales with body mass following quarter-power scaling and as an exponential with temperature.*

September 24, 2023

This implies that life span can in principle be extended by lowering the body temperature because this lowers cellular metabolic rate and therefore the rate at which damage is incurred. This is a very large effect: I remind you that a modest 2°C decrease in body temperature can result in a 20 percent to 30 percent increase in life span. 19 So if you were able to artificially lower your body temperature by just 1°C (that's about 1.8°F) you could enhance your life span by about 10 percent to 15 percent. The hitch is that you would have to do this for your entire life in order to reap the "benefits." But more saliently, significantly lowering body temperature may well have many other deleterious, potentially life-threatening outcomes. As I have stressed earlier, changing just one component of a complex adaptive system without fully understanding its multilevel spatiotemporal dynamics usually leads to unintended consequences. 246

September 24, 2023

Given the relative simplicity of the theory, the agreement is surprisingly good. Coupled with the success of its other predictions (including the rate of aging, the allometric scaling of life span, and its temperature dependence) the theory provides a credible coarse-grained baseline for developing a more detailed quantitative theory for understanding aging and mortality. It gives formulae for the rate of aging and maximum life span in terms of generic "universal" biological parameters that show, for instance, how the scale of one hundred years arises from microscopic molecular scales and why mice live for only a very few years. This provides the scientific basis for asking questions about what parameters can be manipulated to extend life and 249

arrest aging, if that is the goal. For example, combining the scaling laws with Figures 23–30 gives quantitative estimates of how long life can be extended by changing body temperature or eating less.

September 24, 2023

A leading critic was the well-known economist Julian Simon, who expressed a fairly extreme version of the view held by many economists that the spectacular growth we have witnessed over the past two hundred years will be sustained "forever" by human ingenuity and our continued ability to innovate. In fact, Simon argued in his 1981 book The Ultimate Resource that a larger population is actually better because it stimulates even more technological innovation, inventiveness, and ingenuity, thereby leading to new ways of exploiting resources and increasing standards of living. 6 As we move into the twenty-first century this vision of a cornucopian "horn of plenty," with its image of limitless barrels of fish being continuously replenished by the magic, not of divine intervention, but of the free expression of human ingenuity and the boundless possibilities of a free market economy, has reemerged as a significant component of corporate and political conceptual thinking. Indeed, Simon's views have effectively been embraced by many in the academic, business, and political communities. A succinct summary of this view was articulately expressed by the economist Paul Romer, one of the founders of endogenous growth theory, which holds that economic growth is driven primarily by investment in human capital, innovation, and knowledge creation. 7 Romer declares that "every generation has perceived the limits to growth that finite resources and undesirable side effects would pose if no new recipes or ideas were discovered. And every generation has underestimated the potential for finding new recipes and ideas. We consistently fail to grasp how many ideas remain to be discovered. Possibilities do not add up. They multiply." Put slightly differently, this proclaims that ideas and innovation increase multiplicatively (that is, exponentially) and not arithmetically (that is, linearly) in tandem with the exponential growth in population and that the process is open-ended and effectively limitless. On the other hand, the 277 *last few decades have also seen a resurfacing of the spiritual successors of The Population Bomb and The Limits to Growth with the rise of the environmental movement and the development of a serious concern for the future of the planet. Closely related to this is a deep concern for the impact of unregulated corporate and political ambition, which has stimulated a perceived need for "corporate social responsibility."*

September 25, 2023

There is no doubt that significant progress has been and will continue to be made, but can it be enough? It is a matter of faith that the free market system geared to open-ended growth, even when tempered by governmental intervention, stimulation, and regulation, can find a meta-stable balance between making significant profits and solving the problem of sustainability. The primary function of business, after all, is not to increase efficiencies, but to make profits. 281

September 26, 2023

In addition, there also seems to be an unexpressed presumption that ideas, the seeds of innovation, cost nothing—after all, they are "just" neural processes in the brains of human beings, and collectively we can produce an almost infinite number of them in our heads. But like everything else, ideas and the innovations they inspire require energy, and lots of it, to support the smart individuals who are thinking and to provide the appropriately stimulating environments and collective experiences we institutionalize in places such as universities, laboratories, parliaments, cafés, concert halls, and conference venues. 285

September 26, 2023

Despite my skepticism, David and Sander convinced me that 296

extending the network-based scaling theory from biology to social organizations was indeed a worthwhile project. They became the prime movers in putting together a broad program that covered our joint interests ranging from innovation and information transfer in both ancient and modern societies to understanding the structure and dynamics of cities and companies, all from a complexity perspective. The program was called the Information Society as a Complex System (ISCOM) and was generously funded by the European Union. Shortly thereafter Denise Pumain, a well-known urban geographer at the Sorbonne in Paris, joined our collaboration and the four of us each ran one component of the project. I assembled a new multidisciplinary collaboration centered on SFI whose first goal was to ask whether cities and companies manifest scaling and, if so, to develop a quantitative principled theory for understanding their structure and dynamics.

September 26, 2023

live in the age of the Urbanocene, and globally the fate of the cities is the fate of the planet. Jane understood this truth more than fifty years ago, and only now are some of the experts beginning to recognize her extraordinary foresight. 311

September 28, 2023

Cities have an organic quality. They evolve and physically grow out of interactions between people. The great metropolises of the world facilitate human interaction, creating that indefinable buzz and soul that is the wellspring of its innovation and excitement and a major contributor to its resilience and success, economically and socially. It is shortsighted and even courting disaster to ignore this critical dimension of urbanization and concentrate only on buildings and infrastructure. 318

September 28, 2023

The first step in carrying out such a program was to ask if cities are approximately scaled versions of one another in a similar way that animals are. In terms of their measurable characteristics, are New York, Los Angeles, Chicago, and Santa Fe scaled versions of one another and, if so, is their relative scaling similar to the way in which Tokyo, Osaka, Nagoya, and Kyoto scale, despite their very different appearances and characters? Does their scaling manifest any analog of the universality we saw in biology where whales, elephants, giraffes, human beings, and mice are approximately scaled versions of one another, all neatly and quantitatively expressed by the preponderance of quarter-power scaling laws? Compared with biology, surprisingly little attention had been paid to such questions regarding cities, urban systems, or companies prior to our work. To some extent this is because urban studies have historically been even less generally quantitative than biology, but also because relatively few computational mechanistic models of cities or companies have been proposed, let alone confronted with data. 320

September 29, 2023

It may not come as such a big surprise to learn that larger cities require fewer gas stations per capita than smaller ones, but what is surprising is that this economy of scale is so systematic: it is approximately the same across all of these countries, obeying the same mathematical scaling law with a similar exponent of around 0.85. What is even more surprising is that other infrastructural quantities associated with transport and supply networks, such as the total length of electrical lines, roads, water and gas lines, all scale in much the same way with approximately the same value of the exponent, namely about 0.85. Furthermore, this systematic behavior appears to be the same across the globe wherever data could be obtained. So as far as their overall infrastructure is concerned, cities behave just like organisms—they scale sublinearly following simple power-law behavior, thereby exhibiting a systematic economy of scale, albeit to a lesser degree as represented by the different values of their exponents (0.75 for organisms vs. 0.85 for 323 *cities).*

September 29, 2023

The extension of this initial probe into whether cities scale to a broader suite of metrics and across a wider range of countries was carried out by a highly talented group of new recruits to the collaboration. These included Luis Bettencourt, whom I first got to know when he was a postdoctoral fellow in astrophysics at Los Alamos working on the evolution of the early universe. He had meanwhile spent a couple of years at MIT before returning to Los Alamos as a staff member in the applied mathematics group. Luis was born, raised, and educated in Portugal, though you'd never know it as he speaks English fluently without any trace of an accent, so much so that when I first met him I thought he was English. He had in fact obtained his doctorate in physics at Imperial College in London, where coincidentally I hold a position in the mathematics department. Luis's fluidity with language is matched by his fluidity with science. He very quickly became engaged with the cities project, gathering and analyzing data from across the globe. He is a passionate adherent of the cause of developing a deep understanding of cities and has now established himself as one of the world's leading experts in this arena.

324

September 29, 2023

This savings leads to a significant decrease in the production of emissions and pollution. Consequently, the greater efficiency that comes with size has the nonintuitive but very important consequence that on average the bigger the city, the greener it is and the smaller its per capita carbon footprint. In this sense, New York is the greenest city in the United States, whereas Santa Fe, where I live, is one of the more profligate ones. On average, each of us in Santa Fe is putting almost twice as much carbon into the atmosphere as New York. This should not be thought of as somehow reflecting the greater wisdom of New York's planners and politicians, nor as the fault of Santa Fe's leadership, but

325

rather as an almost inevitable by-product of the dynamics underlying economies of scale that transcend the individuality of cities as their size increases. These gains are mostly unplanned, though policy makers in cities can certainly play a powerful role in facilitating and enhancing the hidden "natural" processes that are at work. In fact, this is a large part of what their job is. Some cities are very successful at doing this, while others are much less so. I'll discuss the question of relative performance in the next chapter. These results are very encouraging and provide powerful evidence in support of the quest for a possible theory of cities. However, of even greater significance was the surprising discovery that the data also reveal that socioeconomic quantities with no analog in biology such as average wages, the number of professional people, the number of patents produced, the amount of crime, the number of restaurants, and the gross urban domestic product (GDP) also scale in a surprisingly regular and systematic fashion, as illustrated in Figures 34–38.

September 29, 2023

The data convincingly show that despite appearances cities are approximately scaled versions of one another: New York and Tokyo are, to a surprising and predictable degree, nonlinearly scaled-up versions respectively of San Francisco and Nagoya. These extraordinary regularities open a window onto underlying mechanisms, dynamics, and structures common to all cities and strongly suggest that all of these phenomena are in fact highly correlated and interconnected, driven by the same underlying dynamics and constrained by the same set of "universal" principles.

330

September 29, 2023

Consequently, each of these urban characteristics, each metric—whether wages, the length of all the roads, the number of AIDS cases, or the amount of crime—is interrelated and interconnected with every other one and together they form an overarching multiscale

330

quintessentially complex adaptive system that is continuously integrating and processing energy, resources, and information. The result is the extraordinary collective phenomenon we call a city, whose origins emerge from the underlying dynamics and organization of how people interact with one another through social networks. To repeat: cities are an emergent self-organizing phenomenon that has resulted from the interaction and communication between human beings exchanging energy, resources, and information. As urban creatures we all participate in the multiple networks of intense human interaction that is manifested in the metropolitan buzz of productivity, speed, and ingenuity, no matter where we live.

September 29, 2023

To summarize: the bigger the city, the greater the social activity, the more opportunities there are, the higher the wages, the more diversity there is, the greater the access to good restaurants, concerts, museums, and educational facilities, and the greater the sense of buzz, excitement, and engagement. These facets of larger cities have proven to be enormously attractive and seductive to people worldwide who at the same time suppress, ignore, or discount the inevitable negative aspects and the dark side of increased crime, pollution, and disease. Human beings are pretty good at "accentuating the positive and eliminating the negative," especially when it comes to money and material well-being. In addition to the perceived individual benefits coming from increased city size, there are huge collective benefits arising from systematic economies of scale. Coupled together, this remarkable combination of increasing benefits to the individual with systematic increasing benefits for the collective as city size increases is the underlying driving force for the continued explosion of urbanization across the planet. 332

September 29, 2023

Greed is the pejorative image of this insatiable desire for more, but it also has an extremely important, positive flip 339 *side. Metaphorically, it is the social analog of the evolutionary biological drive of animals, including us, to maximize their metabolic power relative to their size. As was discussed in chapter 3, this can be thought of as derivative from the principle of natural selection and underlies the allometric scaling laws that permeate biology. The extension of the concept of the survival of the fittest to the social and political domain has led many thinkers to the controversial concept of Social Darwinism, whose roots go back to Malthus. Regardless of its validity, this idea has been sadly misrepresented, abused, and misused by politicians and social thinkers, sometimes with devastating consequences, to support all sorts of extreme views ranging from eugenics and racism to rampant laissez-faire capitalism. The desire for more can apply to many things beyond wealth and material assets. It is a hugely powerful force in society that poses enormous moral, spiritual, and psychological challenges at both the individual and collective levels. The desire to succeed, whether in sports, business, or academia—to run the fastest, have the most creative company, or generate the most profound and insightful idea—has been a major underlying societal dynamic that has been instrumental in bringing us the extraordinary standard of living and quality of life many of us are privileged to enjoy. At the same time we have tempered our rampant materialistic greed by evolving altruistic and philanthropic behavior that has been integrated into our sociopolitical structures to protect us from excesses.*

September 29, 2023

345

In addition to providing a metric for comparing the complexity of different cities, perhaps one of the more interesting uses of a fractal dimension is as a diagnostic barometer of the health of a city. Typically, the fractal dimension of a healthy robust city steadily increases as it grows and develops, reflecting a greater complexity as more and more infrastructure is built to accommodate an expanding population engaging in more and more diverse and intricate activities. But conversely, its fractal dimension decreases when it goes through difficult economic times or when it temporarily contracts.

September 29, 2023

The question as to how and why good people become evil and do bad things—the human analog to Job's dilemma as to why God allows bad things to happen to good people—has been a fundamental paradox of human behavior since we evolved social consciousness. The question of man's place in relationship to himself—the continuing moral and ethical dilemma of good versus evil—can be viewed as a companion question to that of man's place in relationship to the universe. These are both central issues of human existence that have dominated human contemplation from the time Homo sapiens became conscious, spawning numerous religions, cultures, and philosophies. 357

September 29, 2023

This relationship between an individual and his or her community naturally led him to the broader issue of the psychological dimension of urban life. In 1970 he published a provocative paper in Science titled "The Experience of Living in Cities," which laid the foundations for the newly developing field of urban psychology and which became the central focus of his ensuing interests. 14 Milgram was very much struck by the psychological harshness of life in the big city. The general perception was that outside of the local environment an individual generally went about his or her business avoiding interaction and involvement, rarely acknowledging other people or events that might engender participation or commitment. So much so that most people are very loath to intervene or even call for help when they are witness to crime, violence, or some other crisis event. He devised a sequence of innovative experiments to investigate the apparent lack of trust, the greater sense of fear and anxiety, and the general lack of civility and politeness that seem to characterize life in large metropolises relative to that in small towns. For instance, he had individual investigators ring doorbells saying that they had misplaced the address of a nearby friend and would like to use the phone. He found 358 *that there was a large increase by factors of three to five in the proportion of entries allowed into homes in a small town relative to a large city. Furthermore, 75 percent of the city respondents answered the inquiry by shouting through closed doors or by peering out through peepholes, whereas in small towns 75 percent opened the door.*

September 29, 2023

Partly inspired by work on social primate communities and partly by anthropological studies of human societies ranging from hunter-gatherers to modern corporations, Dunbar discovered that this hierarchy appeared to have a surprisingly regular mathematical structure obeying very simple scaling rules reminiscent of self-similar fractal-like behavior. He and his collaborators found that at the lowest level of the hierarchy the number of people with whom the average individual has his or her strongest relationships is, at any one time, only about five. These are the people we are closest to and care most deeply about; they are usually family members—parents, children, or spouses—but they could be extremely close friends or partners. In surveys designed to measure the size of this core social group one of its defining characteristics was "the set of individuals from whom the respondent would seek personal advice or help in times of severe emotional and financial distress." The next level up contains those you usually refer to as close friends with whom you enjoy spending meaningful time and might still turn to in time of need even if they are not on as intimate terms with you as your inner circle. This typically comprises around fifteen people. In the level above this are people you might still call friends though you would only rarely invite them to dinner but would likely invite them to a party or gathering. This might consist of coworkers, neighbors down the street, or relatives you don't see very often. There are typically about fifty such people in this group. The next level pretty much defines the limit of your social horizon as far as personal interactions are concerned and consists of people you might refer to as "casual friends"—you know their names and remain in social contact with them. This group typically comprises about 150 people. It is this number that is usually referred to as the Dunbar number that has gained a certain 361 *degree of attention in the popular media.*

September 29, 2023

Unlike in biology, surprisingly little attention had been paid to scaling laws for cities, urban systems, or companies prior to our investigations. This may be because few people suspected that such complex, historically contingent man-made systems as these would manifest any sort of systematic quantitative regularity. In addition, there is much less of a tradition of this kind of modeling and of confrontation of theory with data in urban studies than in either biology or physics. There was, however, one major exception to this and that is a famous scaling law known as Zipf's law for the ranking of cities in terms of their population size. This is shown graphically in Figure 39. It's an intriguing observation: in its simplest form, it states that the rank order of a city is inversely proportional to its population size. 366

September 30, 2023

Nevertheless, the fact that these Zipf-like distributions are found across such a diverse set of phenomena suggests that they express some general systemic property that is independent of the character and detailed dynamics of the specific entities under consideration. This is reminiscent of the ubiquitous generality of the "bell-curve" distribution used for describing statistical variations around some average value. Technically this is called a Gaussian or normal distribution and arises mathematically whenever a sequence of events or entities, whatever they are, are randomly distributed, uncorrelated, and independent of one another. So, for example, the average height of men in the United States is about five feet ten inches (1.77 meters) and the frequency distribution of their heights around this mean value—that is, how many men there are of a given height—closely follows a classic Gaussian bell-curve distribution. This tells us what the likelihood is of someone being a specific height. Gaussian statistics are used across all of science, technology, economics, and finance to assign statistical 371

probabilities of various events happening such as in weather forecasting or for drawing conclusions from polling surveys. However, it is sometimes forgotten that these probability estimates are based on the assumption that individual "events," whether comparing today's temperature with the historical record, or one person's height with another's, are independent of one another and can therefore be considered uncorrelated. The canonical Gaussian bell curve is so pervasive and so taken for granted that it is generally assumed without much thought that that's how "everything" is distributed. Consequently, power law distributions like those of Zipf and Pareto barely saw the light of day. It was natural to assume that cities, incomes, and words would be randomly distributed following the classic bell curve. If that were so, it would predict that there are far fewer very large cities, very large companies, and very rich people, and far fewer common words than there actually are, because all of these follow power law distributions that have much longer tails, meaning that there are many more rare events than would be expected had they been random and obeyed Gaussian statistics. This difference is sometimes characterized by saying that power laws have "fat tails." Clearly, words in a book are correlated and not random because they have to form meaningful sentences, as are cities, because they are part of a unified urban system. It's therefore not so surprising that these distributions are not Gaussian. Many of the most interesting phenomena that we have touched upon fall into this category, including the occurrence of disasters such as earthquakes, financial market crashes, and forest fires. All of these have fat-tail distributions with many more rare events, such as enormous earthquakes, large market crashes, and raging forest fires, than would have been predicted by assuming that they were random events following a classic Gaussian distribution. Furthermore, because these are approximately self-similar processes, the same dynamics occur at all scales. Thus the same generic mechanism that leads to a small adjustment in a financial market is at play when that market suffers a major crash. This is in marked contrast to the random nature implicit in Gaussian statistics, where events at different scales are presumed to be independent and uncorrelated. Ironically, economists and financial analysts traditionally use Gaussian statistics in their analysis, ignoring the *preponderance of fat tails and therefore correlations. Caveat emptor!*

September 30, 2023

The two dominant components that constitute a city, its physical infrastructure and its socioeconomic activity, can both be conceptualized as approximately self-similar fractal-like network structures. Fractals are often the result of an evolutionary process that tends toward optimizing specific features, such as ensuring that all cells in an organism or all people in a city are supplied by energy and information, or maximizing efficiency by minimizing transportation times or times for accomplishing tasks with minimal energy. Less obvious is what is being optimized in social networks. There is, for instance, no satisfactory explanation based on underlying principles for understanding the hierarchical structure observed by Dunbar, or for the origin of his sequence of numbers. Even if the social brain hypothesis is correct, it doesn't explain either the origin of the fractal nature of social groups or where the number 150 comes from. There are hints that such generic properties follow from the conjecture made earlier that self-interest—that is, the desire of all individuals and companies to maximize their assets and income—coupled with the concept of maximal filling of social space are the underlying driving forces. There is certainly a great deal yet to be done in constructing a quantitative theory of social networks, and many exciting challenges await future investigation. 373

There is certainly a great deal yet to be done in constructing a quantitative theory of social networks, and many exciting challenges await future investigation.

September 30, 2023

This exercise shows that there is a natural explanation for why social connectivity and therefore socioeconomic quantities scale superlinearly with population size. Socioeconomic quantities are the sum of the interactions or 378 *links between people and therefore depend on how correlated they are. In the extreme case when everyone is interacting with everyone else we saw that this leads to a superlinear power law whose exponent is 2. However, in reality there are significant constraints on the intensity and magnitude of how many people an individual can interact with, and this drastically reduces the value of the exponent to be less than 2.*

September 30, 2023

In this spirit, individuals are considered to be the "invariant terminal units" of social networks, meaning that on average each person operates in roughly the same amount of social and physical space in a city. 380

September 30, 2023

It is perhaps not so surprising that there is a correlation between increased social interaction, socioeconomic activity, and greater economies of scale. What is surprising, however, is that this pivotal interrelationship follows such simple mathematical rules that can be expressed in an elegant universal form: the sublinearity of infrastructure and energy use is the exact inverse of the superlinearity of socioeconomic activity . Consequently, to the same 15 percent degree, the bigger the city the more each person earns, creates, innovates, and interacts—and the more each person experiences crime, disease, entertainment, and opportunity—and all of this at a cost that requires less infrastructure and energy for each of them. This is the genius of the city. No wonder so many people are drawn to them. 383

September 30, 2023

The size of an average individual's modular cluster of acquaintances who interact with one another is an approximate invariant—it doesn't change with city size. For example, the size of the "extended family" of an average 408

individual living in Lisbon, which has more than 500,000 inhabitants, is no larger than that of an average individual living in the small town of Lixa with less than 5,000 inhabitants. So even in large cities we live in groups that are as tightly knit as those in small towns or villages. This is a bit like the invariance of the Dunbar numbers I talked about in the previous chapter and, like those, probably reflects something fundamental about how our neurological structure has evolved to cope with processing social information in large groups. There is, however, an important qualitative difference in the nature of these modular groups in villages relative to those in large cities. In a real village we are limited to a community that is imposed on us by sheer proximity resulting from its small size, whereas in a city we are freer to choose our own "village" by taking advantage of the much greater opportunity and diversity afforded by a greater population and to seek out people whose interests, profession, ethnicity, sexual orientation, and so on are similar to our own. This sense of freedom provided by a greater diversity across many aspects of life is one of the major attractions of urban life and a significant contributor to rapidly increasing global urbanization.

October 1, 2023

This framework was used as a point of departure to give a science-based measure of performance. Four of the six champions lifted loads they should have, given their body weights and the expectations from the scaling law. On the other hand, the middleweight over performed relative to the expectations for his size, whereas the heavyweight under performed. So despite the heavyweight lifting a greater load than anyone else, from a scientific perspective he was actually the weakest of all the champions while the middleweight was the strongest . 418

October 1, 2023

In contrast to the inherent universal properties of cities that are manifested in the scaling laws, these rank-size 436 *distributions of business types reflect the individuality and distinctive characteristics of each specific city as exhibited by the composition of its economic activities. They are the hallmark of each city and obviously do depend on its history, geography, and culture. It is therefore all the more remarkable that, despite the unique admixture of business types for each individual city, the shape and form of their distribution is mathematically the same for all of them. So much so, in fact, that with a simple scale transformation inspired by theory their rank abundances collapse onto a single unique universal curve common to all cities, as shown explicitly in Figure 53. When one considers the vast range of income, density, and population levels, let alone the uniqueness and diverse cultures that vary so widely in cities across the United States, this universality is quite surprising. What is particularly satisfying is that this unexpected universality, as well as the actual form of the universal curve and the logarithmic scaling of diversity, can all be derived from theory. The universality is driven by the constraint that the sum total of all the different businesses in a city scales linearly with population size, regardless of the detailed composition of business types or of the city. The snakelike mathematical form of the actual distribution function in Figure 53 can be understood from a variant of a very general dynamical mechanism that has been successfully used for understanding rank-size distributions in many different areas, ranging from words and genes to species and cities. It goes under many different names including preferential attachment , cumulative advantage , the rich get richer , or the Yule-Simon process . It is based on a positive feedback mechanism in which new elements of the system (business types in this case) are added with a probability proportional to the abundances of how many are already there. The more there are, the more of that type are going to be added, so more frequent types get even more abundant with increasingly higher probability than less frequent types. 11*

October 2, 2023

It is the first of these, the physical infrastructural component, that has close analogies to biology and provides the metaphor of the city as an organism. But as I have 442 *persistently emphasized, the city is much more than its physicality. Consequently, the concept of metabolic rate as the supply-side input that fuels growth and sustains a city has to be expanded to include socioeconomic activity. In addition to the electricity, gas, oil, water, materials, products, artifacts, and so on that are used and generated in a city, we have to add wealth, information, ideas, and social capital. At a more fundamental level all of these, whether physical or socioeconomic, are driven and sustained by the supply of energy. In addition to heating buildings, transporting materials and people, manufacturing goods, and providing gas, water, and electricity, every transaction, every dollar gained or lost, every conversation and meeting, every phone call and text message, every idea and every thought has to be fueled by energy. Furthermore, just as food must be metabolized into a form that is useful for supplying cells and sustaining life, so the incoming energy and resources digested by a city must be transformed into a form that can be used to supply, sustain, and grow socioeconomic activities such as wealth creation, innovation, and the quality of life. No one has articulated this more eloquently than the great urbanist Lewis Mumford 13 : The chief function of the city is to convert power into form, energy into culture, dead matter into the living symbols of art, biological reproduction into social creativity.*

October 2, 2023

The superlinear scaling of metabolism has profound consequences for growth. In contrast to the situation in biology, the supply of metabolic energy generated by cities as they grow increases faster than the needs and demands for its maintenance. Consequently, the amount available for growth, which is just the difference between its social metabolic rate and the requirements for maintenance, continues to increase as the city gets larger. The bigger the city gets, the faster it grows—a classic signal of open-ended exponential growth. A mathematical analysis indeed confirms that growth driven by superlinear scaling is actually faster than exponential: in fact, it's superexponential . Even though the conceptual and mathematical structure of the growth equation is the same for organisms, social insect

444

communities, and cities, the consequences are quite different: sublinear scaling and economies of scale that dominate biology lead to stable bounded growth and the slowing down of the pace of life , whereas superlinear scaling and increasing returns to scale that dominate socioeconomic activity lead to unbounded growth and to an accelerating pace of life.

October 2, 2023

As we saw in chapter 4, sublinear scaling in biology leads to bounded growth and a finite life span, whereas in chapter 8 we saw that the superlinear scaling of cities (and of economies) leads to open-ended growth. Their sublinear scaling therefore suggests that companies also eventually stop growing and ultimately die, hardly the image that many CEOs would cherish. It's actually not quite as simple as that because the prediction for the growth of companies is more subtle than just a simple extrapolation from biology. To explain this I am going to present a simplified version of how the general theory applies to companies focusing on the essential features that determine their growth and mortality. The sustained growth of a company is ultimately fueled by its profits (or net revenue), where these are defined as the difference between sales (or total income) and total expenses; expenses include salaries, costs, interest payments, and so on. To continue growth over a prolonged period, companies must eventually return a profit, part of which is sometimes used to pay dividends to shareholders. Together with other investors, they in turn may buy additional stocks and bonds to help support the future health and growth of the company. However, to understand their generic behavior it is more transparent to ignore dividends and investments, which are primarily important for smaller, younger companies, and concentrate on profits, which are the dominant driver of growth for larger ones. As we've seen, growth in both organisms and cities is fueled by the difference between metabolism and maintenance. Using that language, the total income (or sales) of a company can be thought of as its "metabolism" while expenses can be thought of as its "maintenance" costs. In biology, metabolic rate scales sublinearly with size, so as organisms increase in size the 461 *supply of energy cannot keep up with the maintenance demands of cells, leading to the eventual cessation of growth. On the other hand, the social metabolic rate in cities scales superlinearly, so as cities grow the creation of social capital increasingly outpaces the demands of maintenance, leading to faster and faster open-ended growth. So how does this dynamic play out in companies? Intriguingly, companies manifest yet another variation on this general theme by following a path that sits at the cusp between organisms and cities. Their effective metabolic rate is neither sub- nor superlinear but falls right in the middle by being linear. This is illustrated in Figures 63 and 64, where sales are plotted logarithmically against the number of employees showing a best fit with a slope very close to one. Expenses, on the other hand, scale in a more complicated fashion: they start out sublinearly but, as companies become larger, eventually transition to becoming approximately linear. Consequently, the difference between sales and expenses, which is the driver of growth, also eventually scales approximately linearly. This is good news because, mathematically, linear scaling leads to exponential growth and this is what all companies strive for. Furthermore, this also shows why, on average, the economy continues to expand at an exponential rate because the overall performance of the market is effectively an average over the growth performances of all its individual participating companies. Although this may be good news for the overall economy, it sets a major challenge for each individual company because each one has to keep up with an exponentially expanding market. So even if a company is growing exponentially (the good news), this may not be sufficient for it to survive unless its expansion rate is at least that of the market (the bad news). This primitive version of the "survival of the fittest" for companies is the essence of the free market economy. More good news is that the nonlinear scaling of maintenance expenses in younger companies, buoyed by investments and the ability to borrow large amounts relative to their size, fuels their rapid growth. Consequently, the idealized growth curve of companies has characteristics in common with classic sigmoidal growth in biology in that it starts out relatively rapidly but slows down as companies become larger and maintenance expenses transition to becoming linear. However, unlike biology, whose maintenance costs do not transition to linearity,* *companies do not cease growing but continue to grow exponentially, though at a more modest rate.*

October 3, 2023

In all cases the number of survivors falls rapidly immediately following their initial public offering, with fewer than 5 percent remaining alive after thirty years. Similarly, the mortality curves show that the number that have died reaches almost 100 percent within fifty years, with almost half of them having already disappeared in less than ten. It's tough being a company! The survival curves are well approximated by a simple exponential as shown in Figure 75, where the number of companies that have survived is plotted logarithmically versus their age; plotted this way, exponentials appear as straight lines. You might have thought that these results would depend sensitively on whether death occurred via mergers and acquisitions rather than from bankruptcies and liquidations. However, as you can see, they both follow very similar exponential survival curves with only slightly different values for their mortality. One might also have expected the results to depend on which business sector a company is in. The dynamics and competitive market forces would seem to be quite different, for instance, in the energy sector compared with IT, transportation, or finance. Surprisingly, however, all business sectors show similar characteristic exponential survival curves with similar timescales: no matter which sector or what the stated cause is, only about half of the companies survive for more than ten years. This is consistent with an analysis showing that companies scale in approximately the same way when broken down into separate business categories. Within each sector, power laws are obtained having exponents close to those found for the entire cohort of companies—those shown in Figure 75. In other words, the general dynamics and overall life history of companies are effectively independent of the business sector in which they operate. This strongly suggests that there is indeed a universal dynamic at play that determines their coarse-grained behavior, independent of their commercial activity or whether they are eventually going to go bankrupt or merge with or be bought by another company. In a word, 468 *this strongly supports the idea of a quantitative science of companies.*

October 4, 2023

The survival analysis technique used for dealing with "incomplete observations" such as we have here was invented in 1958 by two statisticians, Edward Kaplan and Paul Meier. It has since been extended to areas outside of medicine and used to estimate, for example, how long people can expect to remain unemployed after a job loss or how long it takes for machine parts to fail. Amusingly, Kaplan and Meier each submitted similar but independent papers to the prestigious Journal of the American Statistical Association for publication, and a wise editor persuaded them to combine them into a single paper. This has since been cited in other scholarly papers more than 34,000 times, which is an extremely large number for an academic paper. 474

October 4, 2023

To achieve greater efficiency in the pursuit of greater market share and increased profits, companies stereotypically add more rules, regulations, protocols, and procedures at increasingly finer levels of organization, resulting in the increased bureaucratic control that is typically needed to administer, manage, and oversee their execution. This is often accomplished at the expense of innovation and R&D (research and development), which should be major components of a company's insurance policy for its long-term future and survivability. It's difficult to obtain meaningful data on "innovation" in companies because it's not straightforward to quantify. Innovation is not necessarily synonymous with R&D, especially as there are significant tax advantages in labeling all sorts of extraneous activities as R&D expenses. Nevertheless, from analyzing the Compustat data set we found that the relative amount allocated to R&D systematically decreases as company size increases, suggesting that support for innovation does not keep up with bureaucratic and administrative expenses as 482 *companies expand.*

October 4, 2023

The increasing accumulation of rules and constraints is often accompanied by stagnating relationships with consumers and suppliers that lead companies to become less agile and more rigid and therefore less able to respond to significant change. In cities we saw that one very important hallmark is that they become ever more diverse as they grow. Their spectrum of business and economic activity is incessantly expanding as new sectors develop and new opportunities present themselves. In this sense cities are prototypically multidimensional, and this is strongly correlated with their superlinear scaling, open-ended growth, and expanding social networks—and a crucial component of their resilience, sustainability, and seeming immortality. While the dimensionality of cities is continually expanding, the dimensionality of companies typically contracts from birth through adolescence, eventually stagnating or even further contracting as they mature and move into old age. When still young and competing for a place in the market, there is a youthful excitement and enthusiasm as new products are developed and ideas bubble up, some may be crazy and unrealistic and some grandiose and visionary. But market forces are at work so that only a few of these are successful as the company gains a foothold and an identity. As it grows, the feedback mechanisms inherent in the market lead to a narrowing of its product space and inevitably to greater specialization. The great challenge for companies is how to balance the positive feedback from market forces, which strongly encourage staying with "tried and true" products versus the long-term strategic need to develop new areas and commodities that may be risky and won't give immediate return. Most companies tend to be shortsighted, conservative, and not very supportive of innovative or risky ideas, happy to stay almost entirely with their major successes while the going is good because these "guarantee" short-term returns. Consequently, they tend toward becoming more and more unidimensional . This reduction in diversity coupled with the predicament described earlier in which companies sit near a critical point is a classic indicator 482 *of reduced resilience and a recipe for eventual disaster. By the time a company realizes its condition it is often too late. Reconfiguring and reinventing become increasingly difficult and expensive. So when a large enough unanticipated fluctuation, perturbation, or shock comes along the company becomes seriously at risk and ripe for a takeover, buyout, or simply going belly-up. In a word, it is, as the Mafiosi put it, il bacio della morte— the kiss of death. 9*

October 4, 2023

This is a wonderfully consistent picture: the same conceptual framework based on underlying network dynamics and geometry with the same mathematical structure leads to quite different outcomes in these two very different cases, and both are strongly supported by a plethora of diverse data and observations. However, there is a big catch with potentially huge consequences. Even though the growth of organisms, cities, and economies follows essentially identical mathematical equations, their resulting solutions have subtle but crucial differences arising from one being driven by sublinear scaling (the economies of scale of organisms) and the other by superlinear scaling (the increasing returns to scale of cities and economies): in the superlinear case, the general solution exhibits an unexpectedly curious property technically known as a finite time singularity, which is a signal of inevitable change, and possibly of potential trouble ahead. A finite time singularity simply means that the mathematical solution to the growth equation governing whatever is being considered—the population, the GDP, the number of patents, et cetera— becomes infinitely large at some finite time , as illustrated in Figure 76. This is obviously impossible, and that's why something has to change. Before addressing some of the consequences of this phenomenon, let me first elaborate on some of its salient features. Simple power laws and exponentials are continuously increasing functions that also eventually become infinitely large, but they take an infinite time to do so. Another way of saying this is that in these cases the "singularity" has been pushed off to an infinite time into the future, thereby rendering it "harmless" relative to the potential impact of a finite time singularity. In the case of growth driven by superlinear 486 *scaling, the approach to the finite time singularity, represented by the solid line in Figure 76, is faster than exponential. This is often referred to as superexponential, a term I've already used earlier when discussing the growth of cities.*

October 4, 2023

This kind of growth behavior is clearly unsustainable because it requires an unlimited, ever-increasing, and eventually infinite supply of energy and resources at some finite time in the future in order to maintain it. Left unchecked, the theory predicts that it triggers a transition to a phase that leads to stagnation and eventual collapse, as illustrated in Figure 77. This scenario sounds just like a rehash of the standard Malthusian argument that has been summarily dismissed by generations of economists: namely, that we won't be able to keep up with demand and that open-ended growth will eventually lead to catastrophe. Which brings us to the crux of the matter. Because of the presence of a finite time singularity resulting from superlinear scaling, this scenario is categorically different from that of Malthus. If growth were purely exponential as assumed by Malthusians, neo-Malthusians, their followers, and critics, then the production of energy, resources, and food could at least in principle keep up with exponential expansion because all of the relevant characteristics of the economy or city remain finite, even if they continue to increase in size and become very large. This cannot be achieved if you are growing superexponentially and approaching a finite time singularity. In this scenario demand gets progressively larger and larger, eventually becoming infinite within a finite period of time. It is simply not possible to supply an infinite amount of energy, resources, and food in a finite time. So if nothing else changes, this inextricably leads to stagnation and collapse, as illustrated in Figure 77. An extensive analysis carried out in 2001 by Didier Sornette and Anders Johansen, then at UCLA, showed that data on population growth and the growth of financial and economic indicators strongly support the theoretical predictions that we have been growing superexponentially and are indeed headed toward such a singularity. 2 I want to emphasize that 487 *this situation is qualitatively quite different from classic Malthusian dynamics, where this is no such singularity. The existence of a singularity signifies that there has to be a transition from one phase of the system to another having very different characteristics, analogous to the way the condensation of steam to water which subsequently freezes to ice epitomizes transitions between different phases of the same system, each having quite different physical properties.*

October 4, 2023

Estimates from the theory suggest that we are due for another paradigm shift within the next twenty to thirty years' time. This is a little shorter than the estimates from fits to data by Johansen and Sornette, who suggest a number more like thirty-five years. The theory cannot, of course, tell us the nature of the change, so we can only make wild speculations as to its nature. It could be something as relatively mundane as driverless cars and associated smart devices or something as dramatic as the kind of science fiction fantasy of Kurzweil and the singularists. Most likely it's none of the above and, if we are able to make the paradigm shift, it'll be something totally unexpected. Perhaps more likely is that we can't make the shift, and that we will need to come to terms with the whole concept of open-ended growth and find some new way of defining "progress" or be content with what we've got and spend our energies raising the entire planet's standard of living to reflect a comparably high quality of life. Now, that would be a truly major paradigm shift! Continuous growth and the consequent ever-increasing acceleration of the pace of life have profound consequences for the entire planet and, in particular, for cities, socioeconomic life, and the process of global urbanization. Until recent times, the time between major innovations far exceeded the productive life span of a human being. Even in my own lifetime it was unconsciously assumed that one would continue working in the same occupation using the same expertise throughout one's life. This is no longer true; a typical human being now lives significantly longer than the time between major innovations, especially in developing and developed countries. Nowadays young people entering the workforce can expect to see several major

499

changes during their lifetime that will very likely disrupt the continuity of their careers. This increasingly rapid rate of change induces serious stress on all facets of urban life. This is surely not sustainable, and, if nothing changes, we are heading for a major crash and a potential collapse of the entire socioeconomic fabric. The challenges are clear: Can we return to an analog of a more "ecological" phase from which we evolved and be satisfied with some version of sublinear scaling and its attendant natural limiting, or no-growth, stable configuration? Is this even possible? Can we have the kind of vibrant, innovative, creative society driven by ideas and wealth creation as manifested by the best of our world's cities and social organizations, or are we destined to a planet of urban slums and the ultimate specter of devastation raised by Cormac McCarthy's novel The Road ? 8 Given the special, unique role of cities as the originators of many of our present problems and their continuing role as the superexponential driver toward potential disaster, understanding their dynamics, growth, and evolution in a scientifically predictable, quantitative framework is crucial to achieving long-term sustainability on the planet. Perhaps of even greater importance for the immediate future is to develop such a theory within the context of a grand unified theory of sustainability by bringing together the multiple studies, simulations, databases, models, theories, and speculations concerning global warming, the environment, financial markets, risk, economies, health care, social conflict, and the myriad other characteristics of man as a social being interacting with his environment.

October 4, 2023

The institute has been internationally recognized as "the formal birthplace of the interdisciplinary study of complex systems" and has played a central role in recognizing that many of the most challenging, exciting, and profound questions facing science and society lie at the boundaries between traditional disciplines. Among these are the origins of life; the generic principles of innovation, growth, evolution, and resilience whether of organisms, ecosystems, pandemics, or societies; network dynamics in nature and society; biologically inspired paradigms in medicine and

509

computation; the interrelationship between information processing, energy, and dynamics in biology and society; the sustainability and fate of social organizations; and the dynamics of financial markets and political conflicts.

October 5, 2023

He identified the first three as (1) empirical observation (pre-Galileo), (2) theory based on models and mathematics (post-Newtonian), and (3) computation and simulation. My impression is that, in contrast to Chris Anderson, Gray viewed this fourth paradigm as an integration of the previous three, namely as a unification of theory, experiment, and simulation, but with an added emphasis on data gathering and analysis. In that sense it's hard to disagree with him because this is pretty much the way science has progressed for the last couple of hundred years—the difference being primarily quantitative: the "data revolution" provides us with a much greater possibility for exploiting and enabling strategies we have been using for a very long time. In this sense this is more like paradigm 3.1 than paradigm 4.0. But there is a new kid on the block that many feel promises more and, like Anderson, potentially subverts the need for the traditional scientific method. This invokes techniques and strategies with names like machine learning, artificial intelligence , and data analytics. There are many versions of these, but all of them are based on the idea that we can design and program computers and algorithms to evolve and adapt based on data input to solve problems, reveal insights, and make predictions. They all rely on iterative procedures for finding and building upon correlations in data without concern for why such relationships exist and implicitly presume that "correlation supersedes causation." This approach has become a huge area of interest and has already had a big impact on our lives. For instance, it is central to how search engines like Google operate, how strategies for investment or operating an organization are devised, and it provides the foundational basis for driverless cars. 521

October 5, 2023

Red

In 2006 the planet crossed a remarkable historical threshold, with more than half of the world's population residing in urban centers, compared with just 15 percent a hundred years ago and still only 30 percent by 1950.

19

September 11, 2023

At the present rate it will be moving the equivalent of the entire U.S. population (more than 300 million people) to cities in the next twenty to twenty-five years. And not far behind are India and Africa. This will be by far the largest migration of human beings to have ever taken place on the planet and will very likely never be equaled in the future. The resulting challenges to the availability of energy and resources and the enormous stress on the social fabric across the globe are mind-boggling . . . and the timescales to address them are very short. Everyone will be affected; there is no hiding place.

20

There is no hiding place.

September 11, 2023

The most brilliantly designed machine, the most creatively organized company, the most beautifully evolved organism cannot escape this grimmest of grim reapers. To maintain order and structure in an evolving system requires the continual supply and use of energy whose by-product is disorder. That's why to stay alive we need to continually eat so as to combat the inevitable, destructive forces of entropy production. Entropy kills.

25

September 17, 2023

A greatly underappreciated case in point is the hidden role that scaling plays in medicine. Much of the research and development on diseases, new drugs, and therapeutic procedures is undertaken using mice as "model" systems. This immediately raises the critical question of how to scale up the findings and experiments on mice to humans. For instance, huge resources are spent each year on investigating cancer in mice, yet a typical mouse develops many more tumors per gram of tissue per year than we do, whereas whales get almost none, begging the question as to the relevance of such research for humans. To put it slightly differently: if we are to gain a deep understanding and solve the challenge of human cancer from such studies we need to know how to reliably scale up from mice to humans, and conversely, down from whales. Dilemmas such as this will be discussed in chapter 4 when addressing scaling issues inherent in biomedicine and health.

27

September 17, 2023

will be much more precise about what this means and what it implies later but, for the time being, nonlinear behavior can simply be thought of as meaning that measurable characteristics of a system generally do not simply double when its size is doubled. In the example given here, this can be restated as saying that there is a systematic increase in per capita GDP, as well as in average wages, crime rates, and many other urban metrics, as city size increases. This reflects an essential feature of all cities, namely that social activity and economic productivity are systematically enhanced with increasing size of the population. This systematic "value-added" bonus as size increases is called increasing returns to scale by economists and social scientists, whereas physicists prefer the more sexy term superlinear scaling .

29

September 17, 2023

It was only in the decade prior to the building of the Great Eastern that the underlying science governing the motion of ships was developed. The field of hydrodynamics was first formalized independently by the French engineer Claude-Louis Navier and the great Irish mathematical physicist George Stokes. The fundamental equation, which is universally known as the Navier-Stokes equation, arises from applying Newton's laws to the motion of fluids, and by extension to the dynamics of physical objects moving through fluids, such as ships through water or airplanes through air.

90

September 19, 2023

Froude may not have fully recognized just how big a challenge he was facing but he did perceive that for applications to shipbuilding a new strategy was needed. It was in this context that he invented the new methodology of modeling and, by extension, the concept of scaling theory for determining how quantitative results from small-scale investigations could be used to help predict how full-size ships would behave. In the spirit of Galileo, Froude realized that almost all scaling is nonlinear, so traditional models based on a faithful 1:1 representation were not useful for determining how the actual system works. His seminal contribution was to suggest a quantitative mathematical strategy for figuring out how to scale from the small-size model to the full-size object.

92

September 19, 2023

So what was irrelevant at one scale can become dominant at another. The challenge at every level of observation is to abstract the important variables that determine the dominant behavior of the system.

135

September 20, 2023

Physicists have coined a concept to help formalize a first step in this approach, which they call a "toy model." The strategy is to simplify a complicated system by abstracting its essential components, represented by a small number of dominant variables, from which its leading behavior can be determined. A classic example is the idea first proposed in the nineteenth century that gases are composed of molecules, viewed as hard little billiard balls, that are rapidly moving and colliding with one another and whose collisions with the surface of a container are the origin of what we identify as pressure.

135

September 20, 2023

I. Space Filling The idea behind the concept of space filling is simple and intuitive. Roughly speaking, it means that the tentacles of the network have to extend everywhere throughout the system that it is serving, as is illustrated in the networks here. More specifically: whatever the geometry and topology of the network is, it must service all local biologically active subunits of the organism or subsystem. A familiar example will make it clear: Our circulatory system is a classic hierarchical branching network in which the heart pumps blood through the many levels of the network beginning with the main arteries, passing through vessels of regularly decreasing size, ending with the capillaries, the smallest ones, before looping back to the heart through the venal network system. Space filling is simply the statement that the capillaries, which are the terminal units or last branch of the network, have to service every cell in our body so as to efficiently supply each of them with sufficient blood and oxygen.

139

September 20, 2023

These postulates underlie the zeroth order, coarse-grained theory for the structure, organization, and dynamics of biological networks, and allow us to calculate many of the essential features of what I referred to as the average idealized organism of a given size. In order to carry out this

145

strategy and calculate quantities such as metabolic rates, growth rates, the heights of trees, or the number of mitochondria in a cell, these postulates have to be translated into mathematics. The goal is to determine the consequences, ramifications, and predictions of the theory and confront them with data and observations. The details of the mathematics depend on the specific kind of network being considered. As discussed earlier, our circulatory system is a network of pipes driven by a beating heart, whereas plants and trees are networks of bundles of thin fibers driven by a steady nonpulsatile hydrostatic pressure. Fundamental to the conceptual framework of the theory is that, despite these completely different physical designs, both kinds of networks are constrained by the same three postulates: they are space filling, have invariant terminal units, and minimize the energy needed to pump fluid through the system .

September 20, 2023

In his desire to create a testable scientific framework, Richardson collected an enormous amount of historical data on wars and conflicts. In order to quantify them he introduced a general concept, which he called the deadly quarrel , defined as any violent conflict between human beings resulting in death. War is then viewed as a particular case of a deadly quarrel, but so is an individual murder. He quantified their magnitudes by the subsequent number of deaths: for an individual murder the size of the deadly quarrel is therefore just one, whereas for the Second World War it is more than 50 million, the exact number depending on how civilian casualties are counted. He then took the bold leap of asking whether there was a continuum of deadly quarrels beginning with the individual and progressing up through gang violence, civil unrest, small conflicts, and ending up with the two major world wars, thereby covering a range of almost eight orders of magnitude. Trying to plot these on a single axis leads to the same challenge we faced earlier when trying to accommodate all earthquakes or all mammalian metabolic rates on a simple linear scale. Practically, it simply isn't possible, and one has to resort to using a logarithmic scale to see the entire spectrum of deadly quarrels. Thus, by analogy with the Richter scale, the 165

Richardson scale begins with zero for a single individual murder and ends with a magnitude of almost eight for the two world wars (eight orders of magnitude would represent a hundred million deaths). In between, a small riot with ten victims would have magnitude one, a skirmish in which one hundred combatants were killed would be two, and so on. Obviously there are very few wars of magnitude seven but an enormous number of conflicts with magnitude zero or one. When he plotted the number of deadly quarrels of a given size versus their magnitude on a logarithmic scale, he found an approximately straight line just like the straight lines we saw when physiological quantities like metabolic rate were plotted in this way versus animal size (see Figure 1). Consequently, the frequency distribution of wars follows simple power law scaling indicating that conflicts are approximately self-similar. 19 This remarkable result leads to the surprising conclusion that, in a coarse-grained sense, a large war is just a scaled-up version of a small conflict, analogous to the way that elephants are approximately scaled-up mice. Thus underlying the extraordinary complexity of wars and conflicts seems to be a common dynamic operating across all scales. Recent work has confirmed such findings for recent wars, terrorist attacks, and even cyberattacks. 20 No general theory has yet been advanced for understanding these regularities, though they very likely reflect the fractal-like network characteristics of national economies, social behavior, and competitive forces. In any case, any ultimate theory of war needs to account for them.

September 21, 2023

coastline using

172

September 21, 2023

However, driven by the forces of natural selection to maximize exchange surfaces, biological networks do achieve maximal space filling and consequently scale like three-dimensional volumes rather than two-dimensional Euclidean

188

surfaces. This additional dimension, which arises from optimizing network performance, leads to organisms' functioning as if they are operating in four dimensions.

September 22, 2023

What is shocking is how few people, even many scientists, appreciate that this sensitivity to temperature is exponential . The reason for this sensitivity is that all chemical reaction rates depend exponentially on temperature. In the previous chapter I showed how metabolism originates in the production of ATP molecules in cells. Consequently, metabolic rate scales exponentially with temperature rather than as a power law as it does with mass. Because metabolic rate—the rate at which energy is supplied to cells—is the fundamental driver of all biological rates and times, all of the central features of life from gestation and growth to mortality are exponentially sensitive to temperature .

210

September 23, 2023

One last note: the underlying physics and chemistry of reaction theory have been known for a very long time, having been developed by the Swedish physicist-turned-chemist Svante August Arrhenius, who won the Nobel Prize in Chemistry in 1903. He holds the distinction of being the first Swede to become a Nobel laureate. Arrhenius was a man of very broad interests whose many novel ideas and contributions to science have been very influential. He was one of the first people to seriously suggest that life on Earth might have originated from the transport of spores having been transported from another planet, a rather speculative theory with a surprisingly large following that now goes by the name of panspermia . Of greater significance is that he was the first scientist to calculate how changes in the levels of carbon dioxide in the atmosphere could alter the surface temperature of the Earth through the greenhouse effect, predicting that the burning of fossil fuels was large enough to cause significant global warming. Most remarkably he did all of this before 1900, which is pretty depressing because it

215

shows that we already were beginning to understand scientifically some of the deleterious consequences of burning fossil fuels well over a hundred years ago and we did almost nothing about it.

September 23, 2023

I had a minor epiphany when I was sixteen years old. Some school friends persuaded me to join them in going to a small arts cinema in the West End of London to see a film much touted by the intelligentsia at that time. This was Ingmar Bergman's extraordinary film The Seventh Seal . It is of Shakespearean grandeur and depth. It tells the story of a medieval knight, Antonius Block, who on his return journey home to Sweden from fighting in the Crusades encounters the personification of Death, who has come to take his life. In his attempt to avoid, or at least delay, the inevitable, Block proposes that they play a game of chess; should he win, his life will be spared. He, of course, eventually loses but only because he is inadvertently tricked into baring his soul to Death, who masquerades as a confessional priest. This allegorical setting provides the stage for delving into the eternal questions concerning the meaning, or pointlessness, of life and its relationship to death. Questions at the very heart of philosophical and religious discourses with which men and women have struggled throughout the centuries are brilliantly depicted by Bergman's genius. Who can forget the final haunting scene in which the black-robed Death leads Antonius and his entourage on an iconic danse macabre silhouetted across a distant hillside to meet their inevitable fate? What an impression this made on an innocent, unconscious, adolescent sixteen-year-old. I think this was my first truly serious inkling that there was more to life than money, sex, and football and began my long-term interest in questions of metaphysics and philosophical thought. I began to voraciously read all of the usual suspects from Socrates, Aristotle, and Job to Spinoza, Kafka, and Sartre, and from Russell and Whitehead to Wittgenstein, A. J. Ayer, and even Colin Wilson, though barely understanding anything of what any of them were saying (especially Wittgenstein, by the way). What I did learn, however, was that although extraordinary men had struggled with the really big 216 *questions for a very long time, there actually were no answers. Just more questions. It speaks to the profundity of Bergman's masterpiece that almost sixty years later the film still makes the same powerful impression, now possibly more nuanced and poignant, on a slightly jaded seventy-five-year-old man as he approaches the final years. At a critical stage in the film Death very reasonably asks Antonius: "Do you never stop questioning?" to which he emphatically responds, "No. I never stop." And neither should we. The fascination with death, coupled with the incessant questioning and search for any meaning to life, permeates human culture but has mostly been manifested and formalized in the multiplicity of religious institutions and experiences that humans have invented. Science has generally placed itself outside of such philosophical meanderings. However, many scientists have seen the quest for understanding and unraveling "the laws of nature," the passion for wanting to know how things work and what they are made of, as an alternative journey in coming to terms with these big questions even if they themselves are neither "religious" nor particularly "philosophical." Somewhere along the line, I realized that I was one of them, finding in science, or at least in physics and mathematics, some version of the spiritual sustenance that seems to be a universal need.*

September 23, 2023

There are two important points I want to emphasize arising from these statistics: (1) The leading causes of death are overwhelmingly associated with damage, whether in organs and tissue (as in heart attacks or stroke) or in molecules (as in cancer)—infectious diseases play a relatively minor role. (2) Even if every cause of death were eliminated, all human beings are destined to die before they reach 125 years old, and the vast majority of us will do so well before we reach that ripe old age. 234

September 24, 2023

Most of the discussion thus far has been about human beings, 235 *but I now want to extend it to other animals in order to connect it to scaling laws and the theoretical framework introduced earlier. The discussion will be in the spirit of a coarse-grained description, so there are undoubtedly outliers and even exceptions to some of these statements. This is especially true for the case of aging and mortality because, unlike most other traits, these are not directly selected for in the evolutionary process. Natural selection only needs to ensure that the majority of individuals in a species survive sufficiently long to produce enough offspring to maximize their evolutionary fitness. Once this has happened and they have performed their evolutionary "duty," how much longer they live is of much less importance, so large variations in individual and species life spans can be expected. Thus human beings have evolved to live for at least forty years so that they can produce ten or so children, at least half of whom will survive to maturity and beyond. It's perhaps no accident, then, that this is the age of a woman's menopause. However, to ensure that enough of us reach this age and reproduce accordingly, we have evolved to be sufficiently "overengineered" that statistically many of us are able to live much longer.*

September 24, 2023

Before exploring this further, it is instructive to compare some of this with the longevity of automobiles. Unfortunately, there have been surprisingly few scaling analyses of automobiles and other machines, especially regarding their longevity. However, Thomas McMahon, an engineer at Harvard, analyzed data on internal combustion engines, ranging from those used in lawn mowers and automobiles to airplanes, and showed that they follow simple isometric Galilean cubic power scaling laws, discussed in chapter 2. For instance, the horsepower rating of these engines (the analog to their metabolic rate) scales linearly with their weight, so to double their power output, you have to double their weight. Thus, unlike organisms, engines do not exhibit an economy of scale as their size increases. McMahon also found that their RPMs (their heart rates) scale with an inverse cubic power of their weight. 18

239

September 24, 2023

Like all organisms, we metabolize energy and resources in a highly efficient way in order to combat the continuous fight against the inevitable production of entropy in the form of waste products and dissipative forces that cause physical damage. As we begin to lose the multiple localized battles against entropy we age, ultimately losing the war and succumbing to death. Entropy kills. Or as the great Russian playwright Anton Chekhov poignantly remarked, "Only entropy comes easy."

241

September 24, 2023

At a more detailed level, this implies that aging within each organ is approximately uniform even though different organs may age at moderately different rates because they have slightly different network characteristics, especially in regard to their potential for repair. Because larger animals metabolize at higher rates following the ¾ power scaling law, they suffer greater production of entropy and therefore greater overall damage, so you might have thought that this would imply that larger animals would have shorter life spans in obvious contradiction to observation. However, we saw in chapter 3 that on a cellular or per unit mass of tissue basis metabolic rate and therefore the rate at which damage is occurring at the cellular and intracellular levels decreases systematically with increasing size of the animal—another expression of economy of scale. Furthermore, as already emphasized, the most significant damage occurs at the terminal units of networks in capillaries, mitochondria, and cells, and their metabolic rates decrease with the size of the organism following power law scaling with an exponent of ¼. Cells in larger animals are systematically processing energy at a slower rate than cells in smaller ones. So at the critical cellular level cells suffer systematically less damage at a slower rate the larger the animal, and this results in a correspondingly longer life span.

243

September 24, 2023

Four companies alone now produce 81 percent of the cattle, 73 percent of the sheep, 57 percent of the pigs, and 50 percent of the chickens consumed in the United States.

275

September 25, 2023

However, it is in business that the concepts and language if not the actual science of ecology and evolutionary biology have seized the imagination, especially in Silicon Valley. The concept of a business ecosystem has become a standard buzzword to connote a sort of Darwinian survival of the fittest in the marketplace. It was introduced in 1993 by James Moore, then at Harvard Law School, in an article he wrote titled "Predators and Prey: A New Ecology of Competition," which won the McKinsey award for article of the year. 1 It's a fairly standard ecological narrative, with individual businesses replacing animals in the evolutionary dynamics of natural selection. In keeping with much of the traditional literature on understanding companies, it's entirely qualitative with no quantitative predictive power. Its great virtue is that it emphasizes the role of community structure, the importance of systemic thinking, and the inevitable processes of innovation, adaptation, and evolution.

295

September 26, 2023

Although its scope and emphasis broadened as progress was made, the vision of the proposal remained pretty much intact over the years. Its motivation was originally expressed as follows: "Because of the obvious analogy with social network systems, such as corporate and urban structures, it is both natural and compelling to investigate the possibility of extending the same sort of analyses used for understanding biological network systems to social organizations," with an added emphasis that "the flow of information in social organizations is as significant as the flow of matter, energy and resources." Many questions were asked, including "What

298

is a social organization? What are the appropriate scaling laws? What constraints must be satisfied by the architecture of the structures that channel social flows of information, matter and energy? In particular, are the relevant constraints all physical, or might there also be social and cognitive constraints that must be taken into account?

September 26, 2023

In terms of the socioeconomic dynamics of cities we can likewise ask what, if anything, is being optimized in urban social networks. This is a tough question to answer definitively, and many scholars have obliquely tried to address it from multiple points of view. 5 If we think of the city as the great facilitator of social interactions or as the great incubator for wealth creation and innovation, it is natural to speculate that its structure and dynamics evolved so as to maximize social capital by optimizing the connectivity between individuals. This suggests that social networks and the entire social fabric of cities and urban systems—that is, who is connected to whom, how much information flows between them, and the nature of their group structure—is ultimately determined by the insatiable drive of individuals, small businesses, and giant companies to always want more. Or, to put it in crass terms, that the socioeconomic machinery that we all participate in is primarily fueled by greed in both its negative and positive connotations as in the sense of the "desire for more." Given the enormous disparities in income distributions that are observed in all cities across the globe, and the apparent drive of most of us to want more despite having plenty, it's not hard to believe that greed in its various forms is an important contributor to the socioeconomic dynamics of cities. To quote Mahatma Gandhi: "The Earth provides enough to satisfy every man's needs, but not every man's greed." 338

September 29, 2023

Zipf's law owes its name to the Harvard linguist George 367

Kingsley Zipf, who popularized it through his fascinating book Human Behavior and the Principle of Least Effort, published in 1949. 17 He first enunciated his law in 1935, not for cities but for the frequency of word use in languages. As originally stated, the law says that the frequency of occurrence of any word in a corpus of written text such as all of Shakespeare's plays, the Bible, or even this book is inversely proportional to its rank in the frequency table. Thus, the most frequent word occurs approximately twice as often as the second most frequent, three times as often as the third, and so on, as shown in Figure 40. For example, analysis of English texts shows that the most frequent word is, not surprisingly, "the," which accounts for roughly 7 percent of all words used, while the second-place word "of" accounts for about half as many, namely 3.5 percent of all words, followed by "and" with about a third as many, namely 2.3 percent, and so on. Even more mystifying is that this same law is valid across an astonishing array of examples, including the rank size distributions of ships, trees, sand particles, meteorites, oil fields, file sizes of Internet traffic, and many more. Figure 41 shows how the distribution of company sizes follows this law. Given its amazing generality and some of its implications, Zipf's law has taken on a curious mystique among many researchers and writers whose imagination has been caught by its surprising simplicity. Zipf and many others following him have pondered its origins, but no generally agreed-upon explanation has yet emerged.

September 30, 2023

Consequently, all of the socioeconomic metrics that reflect such activity and which were discussed earlier when reviewing urban scaling laws are proportional to the number of links, or interactions that take place, between people within the city. If it were possible for everyone to interact with everyone else so that over the period of a year, for example, each person connects meaningfully with every other person in the city, then the total number of interactions between people could be easily calculated from a simple formula: it is the total number of people in the city multiplied by the total number of people that any individual can connect 374 *to in the city. This last number is just the total number of people minus one. For instance, if you are one of ten people in a group, you can connect to only nine others. In addition, you have to divide the answer by two because you don't count the link between you and another person as different from the link between that other person and you. They are symmetric and one and the same. Thus the total possible number of pair-wise links between people in a city is given by the total number of people in the city multiplied by the total number minus one—all divided by two. This may seem a bit of a mouthful but it's actually quite simple, so let me explain with some examples.*

September 30, 2023

Recall that the physical space in which we operate is spanned by space-filling fractal networks, such as roads and utility lines that service infrastructural terminal units such as houses, stores, and office buildings where we reside, work, and interact, and between which we also have to move. The integration of these two kinds of networks, namely, the requirement that socioeconomic interaction represented by space-filling fractal-like social networks must be anchored to the physicality of a city as represented by space-filling fractal-like infrastructural networks , determines the number of interactions an average urban dweller can sustain in a city. And as discussed earlier, it is this number that determines how socioeconomic activity scales with population size. The biological metaphor of the city as a living organism derives primarily from its being perceived in terms of its physicality. This is most apparent in the networks that carry energy and resources in the form of electricity, gas, water, cars, trucks, and people, and it is this component of cities that is the close analog to the networks that proliferate in biology such as our cardiovascular and respiratory systems or the vasculature of plants and trees. Combining the ideas of space filling, invariant terminal units, and optimization (minimizing travel times and energy use, for example) results in these networks also being fractal-like with infrastructural metrics scaling as power laws with sublinear exponents indicative of economies of scale obeying the 15 percent rule. 380

September 30, 2023

One of the great ironies in all of these marvelous inventions (with the possible exception of the gruesome weapons of destruction) is that they all promised to make life easier and more manageable and therefore give us more time. Indeed, when I was a young man, pundits and futurists were speaking of the glorious future anticipated from such time-saving inventions, and a topic that was much discussed was what we would do with all free time that would now be at our disposal. With cheap energy available from nuclear sources and these fantastic machines doing all our manual and mental labor, the workweek would be short and we would have large swaths of time to really enjoy the good life with our families and friends, a little like the boring privileged lives of aristocratic ladies and gentlemen of previous centuries. In 1930 the great economist John Maynard Keynes wrote: For the first time since his creation man will be faced with his real, his permanent problem—how to use his freedom from pressing economic cares, how to occupy the leisure, which science and compound interest will have won for him, to live wisely and agreeably and well. And in 1956, Sir Charles Darwin, grandson of the Charles Darwin, wrote an essay on the forthcoming Age of Leisure in the magazine New Scientist in which he argued: Take it that there are fifty hours a week of possible working time. The technologists, working for fifty hours a week, will be making inventions so the rest of the world need only work twenty-five hours a week. The more leisured members of the community will have to play games for the other twenty-five hours so they may be kept out of mischief. . . . Is the majority of mankind really able to face the choice of leisure enjoyments, or will it not be necessary to provide adults with something like the compulsory games of the schoolboy? 391

October 1, 2023

For instance, an increase of the population by a factor of one hundred from, say, 100,000 to 10 million results in the 433 *addition of one hundred times as many businesses but an increase of only a factor of two in their diversity. To put it slightly differently: doubling the size of a city results in doubling the total number of establishments, but only a meager 5 percent increase in new kind s of businesses. Almost all of this increase in diversity is reflected in a greater degree of specialization and interdependence involving larger numbers of people, both as workers and as clients. This is an important observation because it shows that increasing diversity is closely linked to increasing specialization, and this acts as a major driver of higher productivity following the 15 percent rule.*

October 2, 2023

Given this observation it is natural to ask, as we did for cities and organisms, whether companies scale in terms of their measurable metrics such as their sales, assets, expenses, and profits. Do companies manifest systematic regularities that transcend their size, individuality, and business sector? And if so, could there possibly be a quantitative, predictive science of companies paralleling the science of cities that was developed in previous chapters? Is it possible to understand the general quantitative features of their life histories, how they grow, mature, and eventually die? 448

October 2, 2023

The six points resulting from averaging over each bin are shown as gray dots in the graph. They represent a highly coarse-grained reduction of the data, and as you can see follow a very good straight line supporting the idea that underlying the statistical spread is an idealized power law. Because the size and number of bins used is arbitrary, we could just as well have divided up the entire interval into ten, fifty, or one hundred bins rather than just eight, and test whether the straight line remains robust against increasingly finer resolutions of the data. It does. Although binning is not a rigorous mathematical procedure, the stability of obtaining approximately the same straight-line fit 459 *using different resolutions lends strong support to the hypothesis that on average companies are self-similar and satisfy power law scaling. The graph in Figure 4 at the opening of the book is in fact the result of this binning procedure, as is the graph in Figure 41 taken from Axtell's work on showing that companies follow Zipf's law. These results strongly suggest that companies, like cities and organisms, obey universal dynamics that transcend their individuality and uniqueness and that a coarse-grained science of companies is conceivable.*

October 3, 2023

What is the special property of exponentials that they describe the decay of so many disparate systems? It's simply that they arise whenever the death rate at any given time is directly proportional to the number that are still alive. This is equivalent to saying that the percentage of survivors that die within equal slices of time at any age remains the same. A simple example will make this clear: taking one year as the time slice, this says that the percentage of five-year-old companies that die before they reach six years old is the same as the percentage of fifty-year-old companies that die before they reach fifty-one. In other words: the risk of a company's dying does not depend on its age or size. 472

October 4, 2023

This can be restated as a sort of "theorem": to sustain open-ended growth in light of resource limitation requires continuous cycles of paradigm-shifting innovations , as illustrated in Figure 78. 490

October 4, 2023

To some extent this is in the eyes of the beholder, but it's likely that most of us would agree that certain discoveries and innovations such as printing, coal, the telephone, and 494

computers constitute a major "paradigm shift," whereas railways and cell phones may be more debatable. Unfortunately, there is no established quantitative "science of innovation" and therefore no universally agreed-upon criteria or data relating directly to major innovations and paradigm shifts, let alone to finite time singularities. So in order to confront theory with data we have to rely on informal studies and a certain degree of intuition. This situation may well change as innovation becomes an increasingly active area of investigation, with researchers beginning to grapple with questions such as what is innovation, how do we measure it, how does it happen, and how can it be facilitated? 4

October 4, 2023

498

The time between the "Computer Age" and the "Information and Digital Age" was no more than about thirty years—to be compared with the thousands of years between the Stone, Bronze, and Iron ages. The clock by which we measure time on our watches and digital devices is very misleading; it is determined by the daily rotation of the Earth around its axis and its annual rotation around the sun. This astronomical time is linear and regular. But the actual clock by which we live our socioeconomic lives is an emergent phenomenon determined by the collective forces of social interaction: it is continually and systematically speeding up relative to objective astronomical time. We live our lives on the metaphorical accelerating socioeconomic treadmill. A major innovation that might have taken hundreds of years to evolve a thousand or more years ago may now take only thirty years. Soon it will have to take twenty-five, then twenty, then seventeen, and so on, and like Sisyphus we are destined to go on doing it, if we insist on continually growing and expanding . The resulting sequence of singularities, each of which threatens stagnation and collapse, will continue to pile up, leading to what mathematicians call an essential singularity —a sort of mother of all singularities.

October 4, 2023

Among the classic grand syntheses in modern science are Newton's laws, which taught us that heavenly laws are no different from those on Earth; Maxwell's unification of electricity and magnetism, which brought the ephemeral ether into our lives and gave us electromagnetic waves; Darwin's theory of natural selection, which reminded us that we're just animals and plants after all; and the laws of thermodynamics, which suggest that we can't go on forever. Each of these has had profound consequences not only in changing the way we think about the world, but also in laying the foundations for technological advancements that have led to the standard of living many of us are privileged to enjoy. Nevertheless, they are all to varying degrees incomplete. Indeed, understanding the boundaries of their applicability, the limits to their predictive power, and the ongoing search for exceptions, violations, and failures have provoked even deeper questions and challenges, stimulating the continued progress of science and the unfolding of new ideas, techniques, and concepts.

502

October 4, 2023

Emboldened by its great success, physicists endowed this fantastic vision with the grand title of the Theory of Everything. Demanding mathematical consistency between quantum mechanics and general relativity suggested that the basic building blocks of this universal theory might be microscopic vibrating strings rather than the traditional elementary point particles upon which Newton and all subsequent theoretical developments were based. Consequently this vision took on the more prosaic subtitle "string theory." Like the invention of gods and God, the concept of a Theory of Everything connotes the grandest vision of all, the inspiration of all inspirations, namely that we can encapsulate and understand the entirety of the universe in a small set of precepts, in this case, a concise set of mathematical equations from which literally everything follows. Like the concept of God, however, it is potentially misleading and intellectually dangerous.

504

October 4, 2023

All your annotations

296 notes/highlights • 96 bookmarks

1: THE BIG PICTURE

In 2006 the planet crossed a remarkable historical threshold, with more than half of the world's population residing in urban centers, compared with just 15 percent a hundred years ago and still only 30 percent by 1950.

19

September 11, 2023

September 11, 2023

20

At the present rate it will be moving the equivalent of the entire U.S. population (more than 300 million people) to cities in the next twenty to twenty-five years. And not far behind are India and Africa. This will be by far the largest migration of human beings to have ever taken place on the planet and will very likely never be equaled in the future. The resulting challenges to the availability of energy and resources and the enormous stress on the social fabric across the globe are mind-boggling . . . and the timescales to address them are very short. Everyone will be affected; there is no hiding place.

20

There is no hiding place.

September 11, 2023

I suppose that one could view all religion and philosophical reflection as having its origins in how we integrate the inevitable imminence of death into our daily lives.

September 17, 2023

22

September 17, 2023

22

In biology these are controlled and maintained by the process of metabolism. Quantitatively, this is expressed in terms of metabolic rate, which is the amount of energy needed per second to keep an organism alive; for us it's about 2,000 food calories a day, which, surprisingly, corresponds to a rate of only about 90 watts, the equivalent of a standard incandescent lightbulb.

September 17, 2023

23

As social animals now living in cities we still need just a lightbulb equivalent of food to stay alive but, in addition, we now require homes, heating, lighting, automobiles, roads, airplanes, computers, and so on. Consequently, the amount of energy needed to support an average person living in the United States has risen to an astounding 11,000 watts. This social metabolic rate is equivalent to the entire needs of about a dozen elephants. Furthermore, in making this transition from the biological to the social our overall population has increased from just a few million to more than seven billion. No wonder there's a looming energy and resource crisis.

September 17, 2023

23

There is always a price to pay when energy is processed; there is no free lunch. Because energy underlies the transformation and operation of literally everything, no system operates without consequences. Indeed, there is a

24

fundamental law of nature that cannot be transgressed, called the Second Law of Thermodynamics , which says that whenever energy is transformed into a useful form, it also produces "useless" energy as a degraded by-product: "unintended consequences" in the form of inaccessible disorganized heat or unusable products are inevitable. There are no perpetual motion machines. You need to eat to stay alive and maintain and service the highly organized functionality of your mind and body. But after you've eaten, sooner or later you will have to go to the bathroom. This is the physical manifestation of your personal entropy production.

September 17, 2023

The most brilliantly designed machine, the most creatively organized company, the most beautifully evolved organism cannot escape this grimmest of grim reapers. To maintain order and structure in an evolving system requires the continual supply and use of energy whose by-product is disorder. That's why to stay alive we need to continually eat so as to combat the inevitable, destructive forces of entropy production. Entropy kills. 25

September 17, 2023

Scaling simply refers, in its most elemental form, to how a system responds when its size changes. What happens to a city or a company if its size is doubled? Or to a building, an airplane, an economy, or an animal if its size is halved? If the population of a city is doubled, does the resulting city have approximately twice as many roads, twice as much crime, and produce twice as many patents? Do the profits of a company double if its sales double, and does an animal require half as much food if its weight is halved? Addressing such seemingly innocuous questions concerning how systems respond to a change in their size has had remarkably profound consequences across the entire spectrum of science, engineering, and technology and has affected almost every aspect of our lives. Scaling arguments have led to a deep 26 *understanding of the dynamics of tipping points and phase transitions (how, for example, liquids freeze into solids or vaporize into gases), chaotic phenomena (the "butterfly effect" in which the mythical flapping of a butterfly's wings in Brazil leads to a hurricane in Florida), the discovery of quarks (the building blocks of matter), the unification of the fundamental forces of nature, and the evolution of the universe after the Big Bang. These are but a few of the more spectacular examples where scaling arguments have been instrumental in illuminating important universal principles or structure. 9*

September 17, 2023

A greatly underappreciated case in point is the hidden role that scaling plays in medicine. Much of the research and development on diseases, new drugs, and therapeutic procedures is undertaken using mice as "model" systems. This immediately raises the critical question of how to scale up the findings and experiments on mice to humans. For instance, huge resources are spent each year on investigating cancer in mice, yet a typical mouse develops many more tumors per gram of tissue per year than we do, whereas whales get almost none, begging the question as to the relevance of such research for humans. To put it slightly differently: if we are to gain a deep understanding and solve the challenge of human cancer from such studies we need to know how to reliably scale up from mice to humans, and conversely, down from whales. Dilemmas such as this will be discussed in chapter 4 when addressing scaling issues inherent in biomedicine and health. 27

September 17, 2023

will be much more precise about what this means and what it implies later but, for the time being, nonlinear behavior can simply be thought of as meaning that measurable characteristics of a system generally do not simply double when its size is doubled. In the example given here, this can be restated as saying that there is a systematic increase in 29 *per capita GDP, as well as in average wages, crime rates, and many other urban metrics, as city size increases. This reflects an essential feature of all cities, namely that social activity and economic productivity are systematically enhanced with increasing size of the population. This systematic "value-added" bonus as size increases is called increasing returns to scale by economists and social scientists, whereas physicists prefer the more sexy term superlinear scaling .*

September 17, 2023

will be much more precise about what this means and what it implies later but, for the time being, nonlinear behavior can simply be thought of as meaning that measurable characteristics of a system generally do not simply double when its size is doubled. In the example given here, this can be restated as saying that there is a systematic increase in per capita GDP, as well as in average wages, crime rates, and many other urban metrics, as city size increases. This reflects an essential feature of all cities, namely that social activity and economic productivity are systematically enhanced with increasing size of the population. This systematic "value-added" bonus as size increases is called increasing returns to scale by economists and social scientists, whereas physicists prefer the more sexy term superlinear scaling . 29

September 17, 2023

A profound consequence of this rule is that on a per gram basis, the larger animal (the woman in this example) is actually more efficient than the smaller one (her dog) because less energy is required to support each gram of her tissue (by about 25 percent). Her horse, by the way, would be even more efficient. This systematic savings with increasing size is known as an economy of scale . Put succinctly, this states that the bigger you are, the less you need per capita (or, in the case of animals, per cell or per gram of tissue) to stay alive. Notice that this is the opposite behavior to the case 30 *of increasing returns to scale, or superlinear scaling, manifested in the GDP of cities: in that case, the bigger you are, the more there is per capita, whereas for economies of scale, the bigger you are, the less there is per capita. This kind of scaling is referred to as sublinear scaling.*

September 17, 2023

To set the stage, I'd like to quote two distinguished thinkers, one a scientist, the other a lawyer. The first is the eminent physicist Stephen Hawking, who in an interview 10 at the turn of the millennium was asked the following question: Some say that while the 20th century was the century of physics, we are now entering the century of biology. What do you think of this? To which he responded: I think the next century will be the century of complexity. I wholeheartedly agree. As I hope I have already made clear, we urgently need a science of complex adaptive systems to address the host of extraordinarily challenging societal problems we face. 31

September 17, 2023

A typical complex system is composed of myriad individual constituents or agents that once aggregated take on collective characteristics that are usually not manifested in, nor could easily be predicted from, the properties of the individual components themselves. For example, you are much more than the totality of your cells and, similarly, your cells are much more than the totality of all of the molecules from which they are composed 33

September 17, 2023

Ant colonies are built without forethought and without the aid of any single mind or any group discussion or consultation. There is no blueprint or master plan. Just thousands of ants working mindlessly in the dark moving millions of grains of earth and sand to create these 34 *impressive structures. This feat is accomplished by each individual ant obeying just a few simple rules mediated by chemical cues and other signals, resulting in an extraordinarily coherent collective output. It is almost as if they were programmed to be microscopic operations in a giant computer algorithm.*

This feat is accomplished by each individual ant obeying just a few simple rules mediated by chemical cues and other signals, resulting in an extraordinarily coherent collective output. It is almost as if they were programmed to be microscopic operations in a giant computer algorithm.

September 17, 2023

September 17, 2023 34

In general, then, a universal characteristic of a complex system is that the whole is greater than, and often significantly different from, the simple linear sum of its parts. In many instances, the whole seems to take on a life of its own, almost dissociated from the specific characteristics of its individual building blocks. Furthermore, even if we understood how the individual constituents, whether cells, ants, or people, interact with one another, predicting the systemic behavior of the resulting whole is not usually possible. This collective outcome, in which a system manifests significantly different characteristics from those resulting from simply adding up all of the contributions of its individual constituent parts, is called an emergent behavior . It is a readily recognizable characteristic of economies, financial markets, urban communities, companies, and organisms. 35

September 17, 2023

Closely related to the concepts of emergence and self-organization is another critical characteristic of many 36 *complex systems, namely their ability to adapt and evolve in response to changing external conditions. The quintessential example of such a complex adaptive system is, of course, life itself in all of its extraordinary manifestations from cells to cities . The Darwinian theory of natural selection is the scientific narrative that has been developed for understanding and describing how organisms and ecosystems continuously evolve and adapt to changing conditions.*

September 17, 2023

September 17, 2023 36

Scaling up from the small to the large is often accompanied by an evolution from simplicity to complexity while maintaining basic elements or building blocks of the system unchanged or conserved. This is familiar in engineering, economies, companies, cities, organisms, and, perhaps most dramatically, evolutionary processes. For example, a skyscraper in a large city is a significantly more complex object than a modest family dwelling in a small town, but the underlying principles of construction and design, including questions of mechanics, energy and information distribution, the size of electrical outlets, water faucets, telephones, laptops, doors, et cetera, all remain approximately the same independent of the size of the building. These basic building blocks do not significantly change when scaling up from my house to the Empire State Building; they are shared by all of us. Similarly, organisms have evolved to have an enormous range of sizes and an extraordinary diversity of morphologies and interactions, which often reflect increasing complexity, yet fundamental building blocks like cells, mitochondria, capillaries, and even leaves do not appreciably change with body size or increasing complexity of the class of systems in which they are embedded. 37

September 17, 2023

7. YOU ARE YOUR NETWORKS: GROWTH FROM CELLS TO WHALES I began this chapter by pointing out the very surprising and counterintuitive fact that, despite the vagaries and accidents inherent in evolutionary dynamics, almost all of the most fundamental and complex measurable characteristics of organisms scale with size in a remarkably simple and regular fashion. This is explicitly illustrated, for example, in Figure 1, where metabolic rate is plotted against body mass for a sequence of animals.

38

September 17, 2023

This systematic regularity follows a precise mathematical formula which, in technical parlance, is expressed by saying that "metabolic rate scales as a power law whose exponent is very close to the number ¾." I'll explain this in much greater detail later but here I want to give a simple illustration of what it means colloquially. So consider the following: elephants are roughly 10,000 times (four orders of magnitude, 104) heavier than rats; consequently, they have roughly 10,000 times as many cells. The ¾ power scaling law says that, despite having 10,000 times as many cells to support, the metabolic rate of an elephant (that is, the amount of energy needed to keep it alive) is only 1,000 times (three orders of magnitude, 103) larger than a rat's; note the ratio of 3:4 in the powers of ten.

38

September 17, 2023

Parenthetically, it's worth pointing out that the subsequent decrease in the rates of cellular damage from metabolic processes underlies the greater longevity of elephants and provides the framework for understanding aging and mortality. The scaling law can be expressed in the slightly different way that I used earlier: if an animal is twice the size of another (whether 10 lbs. vs. 5 lbs. or 1,000 lbs. vs. 500 lbs.) we might naively expect metabolic rate to be twice as large, reflecting classic linear thinking. The scaling law, however, is

38

nonlinear and says that metabolic rates don't double but, in fact, increase by only about 75 percent, representing a whopping 25 percent savings with every doubling of size. 12 Notice that the ¾ ratio is just the slope of the graph in Figure 1, where the quantities (metabolic rate and mass) are plotted logarithmically— meaning that they increase by factors of ten along both axes. When plotted this way, the slope of the graph is just the exponent of the power law. This scaling law for metabolic rate, known as Kleiber's law after the biologist who first articulated it, is valid across almost all taxonomic groups, including mammals, birds, fish, crustacea, bacteria, plants, and cells. Even more impressive, however, is that similar scaling laws hold for essentially all physiological quantities and life-history events, including growth rate, heart rate, evolutionary rate, genome length, mitochondrial density, gray matter in the brain, life span, the height of trees and even the number of their leaves. Furthermore, when plotted logarithmically this dizzying array of scaling laws all look like Figure 1 and therefore have the same mathematical structure. They are all "power laws" and are typically governed by an exponent (the slope of the graph), which is a simple multiple of ¼, the classic example being the ¾ for metabolic rate. So, for example, if the size of a mammal is doubled, its heart rate decreases by about 25 percent. The number 4 therefore plays a fundamental and almost magically universal role in all of life. 13

September 17, 2023

40

Highly complex, self-sustaining structures, whether cells, organisms, ecosystems, cities, or corporations, require the close integration of enormous numbers of their constituent units that need efficient servicing at all scales. This has been accomplished in living systems by evolving fractal-like, hierarchical branching network systems presumed optimized by the continuous "competitive" feedback mechanisms implicit in natural selection. It is the generic physical, geometric, and mathematical properties of these network systems that underlie the origin of these scaling laws, including the prevalence of the one-quarter exponent. As an example, Kleiber's law follows from requiring that the energy needed to pump blood through mammalian circulatory *systems, including ours, is minimized so that the energy we devote to reproduction is maximized. Examples of other such networks include the respiratory, renal, neural, and plant and tree vascular systems. These ideas, as well as the concepts of space filling (the need to feed all cells in the body) and fractals (the geometry of the networks), will be elaborated upon in some detail.*

September 17, 2023

This process can be analyzed using the network theory to predict a universal quantitative theory of growth curves applicable to any organism, including tumors. A growth curve is simply a graph of the size of the organism plotted as a function of its age. You are probably familiar with growth curves if you have had children, as pediatricians routinely show them to parents so that they can see how their child's development compares with the expectations for the average infant. The growth theory also explains a curious paradoxical phenomenon that you might have pondered, namely, why we eventually stop growing even though we continue to eat. This turns out to be a consequence of the sublinear scaling of metabolic rate and the economies of scale embodied in the network design. In a later chapter, the same paradigm will be applied to the growth of cities, companies, and economies to understand the fundamental question as to the origins of open-ended growth and its possible sustainability. Because networks determine the rates at which energy and resources are delivered to cells, they set the pace of all physiological processes. Because cells are constrained to operate systematically slower in larger organisms relative to smaller ones, the pace of life systematically decreases with increasing size. Thus, large mammals live longer, take longer to mature, have slower heart rates, and cells that don't work as hard as those of small mammals, all to the same predictable degree. Small creatures live life in the fast lane while large ones move ponderously, though more efficiently, through life; think of a scurrying mouse relative to a sauntering elephant.

41

The peace of life.

September 17, 2023

Perhaps even more remarkably they are also scaled socioeconomic versions of one another. Socioeconomic quantities such as wages, wealth, patents, AIDS cases, crime, and educational institutions, which have no analog in biology and did not exist on the planet before humans invented cities ten thousand years ago, also scale with population size but with a superlinear (meaning bigger than one) exponent of approximately 1.15. An example of this is the number of patents produced in a city shown in Figure 3. Thus, on a per capita basis, all of these quantities systematically increase to the same degree as city size increases and, at the same time, there are equivalent savings from economies of scale in all infrastructural quantities. Despite their amazing diversity and complexity across the globe, and despite localized urban planning, cities manifest a surprising coarse-grained simplicity, regularity, and predictability. 15 To put it in simple terms, scaling implies that if a city is twice the size of another city in the same country (whether 40,000 vs. 20,000 or 4 million vs. 2 million), then its wages, wealth, number of patents, AIDS cases, violent crime, and educational institutions all increase by approximately the same degree (by about 15 percent above mere doubling), with similar savings in all of its infrastructure. The bigger the city, the more the average individual systematically owns, produces, and consumes, whether goods, resources, or ideas. The good, the bad, and the ugly are integrated in an approximately predictable package: a person may move to a bigger city drawn by more innovation, a greater sense of "action," and higher wages, but she can also expect to confront an equivalent increase in the prevalence of crime and disease. 43

September 17, 2023

The fact that the same scaling laws are observed for diverse urban metrics in cities and urban systems that evolved independently across the globe strongly suggests that, as in biology, there are underlying generic principles transcending 44 *history, geography, and culture and that a fundamental, coarse-grained theory of cities is possible. In chapter 8 I will discuss how the inextricable tension between benefits and costs of social and infrastructural networks has its origins in the underlying universal dynamics of social network structures and group clustering of human interactions. Cities provide a natural mechanism for reaping the benefits of high social connectivity among very different people conceiving and solving problems in a diversity of ways. I will discuss the nature and dynamics of these social network structures and show how scaling laws emerge, including the intriguing link between the 15 percent enhancement of all socioeconomic activities, whether good or bad, and the equivalent 15 percent savings on physical infrastructure.*

September 17, 2023

When humans began forming sizable communities they brought a fundamentally new dynamic to the planet. With the invention of language and the consequent exchange of information in social network space we discovered how to innovate and create wealth and ideas, ultimately manifested in superlinear scaling. In biology, network dynamics constrains the pace of life to decrease systematically with increasing size following the ¼ power scaling laws. In contrast, the dynamics of social networks underlying wealth creation and innovation leads to the opposite behavior, namely, the systematically increasing pace of life as city size increases: diseases spread faster, businesses are born and die more often, commerce is transacted more rapidly, and people even walk faster, all following the approximate 15 percent rule. We all sense that life is faster in the big city than in the small town and that it has ubiquitously accelerated during our lifetimes as cities and their economies grew. 44

September 17, 2023

Just as bounded growth in biology follows from the sublinear scaling of metabolic rate, the superlinear scaling of wealth creation and innovation (as measured by patent production, 45 *for example) leads to unbounded, often faster-than-exponential growth consistent with open-ended economies. This is satisfyingly consistent, but there's a big catch, which goes under the forbidding technical name of a finite time singularity . In a nutshell, the problem is that the theory also predicts that unbounded growth cannot be sustained without having either infinite resources or inducing major paradigm shifts that "reset" the clock before potential collapse occurs. We have sustained open-ended growth and avoided collapse by invoking continuous cycles of paradigm-shifting innovations such as those associated on the big scale of human history with discoveries of iron, steam, coal, computation, and, most recently, digital information technology. Indeed, the litany of such discoveries both large and small is testament to the extraordinary ingenuity of the collective human mind.*

September 17, 2023

46

Unfortunately, however, there is another serious catch. Theory dictates that such discoveries must occur at an increasingly accelerating pace; the time between successive innovations must systematically and inextricably get shorter and shorter. For instance, the time between the "Computer Age" and the "Information and Digital Age" was perhaps twenty years, in contrast to the thousands of years between the Stone, Bronze, and Iron ages. If we therefore insist on continuous open-ended growth, not only does the pace of life inevitably quicken, but we must innovate at a faster and faster rate. We are all too familiar with its short-term manifestation in the increasingly faster pace at which new gadgets and models appear. It's as if we are on a succession of accelerating treadmills and have to jump from one to another at an ever-increasing rate. This is clearly not sustainable, potentially leading to the collapse of the entire urbanized socioeconomic fabric. Innovation and wealth creation that fuel social systems, if left unchecked, potentially sow the seeds of their inevitable collapse. Can this be avoided or are we locked into a fascinating experiment in natural selection that is doomed to fail?

September 17, 2023

9. COMPANIES AND BUSINESSES It is natural to extend these ideas to ask how they might relate to companies. Could there possibly be a quantitative, predictive science of companies ? Do companies manifest systematic regularities that transcend their size and business character? For example, in terms of sales and assets, are Walmart and Exxon, whose revenues exceed half a trillion dollars, approximately scaled-up versions of smaller companies with sales of less than $10 million? Amazingly, the answer to this is yes, as can be seen from Figure 4: like organisms and cities, companies also scale as simple power laws. Equally surprising is that they scale sublinearly as functions of their size, rather than superlinearly like socioeconomic metrics in cities. In this sense, companies are much more like organisms than cities. The scaling exponent for companies is around 0.9, to be compared with 0.85 for the infrastructure of cities and 0.75 for organisms. However, there is considerably more variation around precise scaling among companies than for organisms or cities. This is especially so in their early stages of development as they jostle for a place in the market. Nevertheless, the surprising regularity manifested in their average behavior suggests that, despite their broad diversity and apparent individuality, companies grow and function under general constraints and principles that transcend their size and business sector.

46

September 17, 2023

For organisms, the sublinear scaling of metabolic rate underlies their cessation of growth and a size at maturity that remains approximately stable until death. A similar life-history trajectory is at work for companies. They grow rapidly in their early years but taper off as they mature and, if they survive, eventually stop growing relative to the GDP. In their youth, many are dominated by a spectrum of innovative ideas as they seek to optimize their place in the market. However, as they grow and become more established, the spectrum of their product space inevitably narrows and, at the same time, they need to build a

47

significant administration and bureaucracy. Relatively quickly, economies of scale and sublinear scaling, reflecting the challenge of efficiently administering a large and complex organization, dominate innovation and ideas encapsulated in superlinear scaling, ultimately leading to stagnation and to mortality. Half of all the companies in any given cohort of U.S. publicly traded companies disappear within ten years, and a scant few make it to fifty, let alone a hundred years. 16

September 17, 2023

September 17, 2023 49

2: THE MEASURE OF ALL THINGS: An Introduction to Scaling

September 17, 2023 51

52

Despite the terrible tragedy that befell Galileo, humanity reaped a wonderful benefit from his incarceration. It may very well have happened anyway, but it was while he was under house arrest that he wrote what is perhaps his finest work, one of the truly great books in the scientific literature, titled Discourses and Mathematical Demonstrations Relating to Two New Sciences. 1 The book is basically his legacy from the preceding forty years during which he grappled with how to systematically address the challenge of understanding the natural world around us in a logical, rational framework. As such, it laid the groundwork for the equally monumental contribution of Isaac Newton and pretty much for all of the science that followed. Indeed, in praising the book, Einstein was not exaggerating when he called Galileo "the Father of Modern Science." 2

September 18, 2023

September 18, 2023 53

GALILEO'S ARGUMENT ON HOW AREAS AND VOLUMES SCALE To begin, consider one of the simplest possible geometrical objects, namely, a floor tile in the shape of a square, and imagine scaling it up to a larger size; see Figure 5. To be specific let's take the length of its sides to be 1 ft. so that its area, obtained by multiplying the length of two adjacent sides together, is 1 ft. × 1 ft. = 1 sq. ft. Now, suppose we double the length of all of its sides from 1 to 2 ft., then its area increases to 2 ft. × 2 ft. = 4 sq. ft. Similarly, if we were to triple the lengths to 3 ft., then its area would increase to 9 sq. ft., and so on. The generalization is clear: the area increases with the square of the lengths. This relationship remains valid for any two-dimensional geometric shape, and not just for squares, provided that the shape is kept fixed when all of its linear dimensions are increased by the same factor. A simple example is a circle: if its radius is doubled, for instance, then its area increases by a factor of 2 × 2 = 4. A more general example is that of doubling the dimensions of every length in your house while keeping its shape and structural layout the same, in which case the area of all of its surfaces, such as its walls and floors, would increase by a factor of four.

September 18, 2023

September 18, 2023 55

Now, as we scale up a building or an animal, their weights increase in direct proportion to their volumes provided, of course, that the materials they're made of don't change so 56 *that their densities remain the same: so doubling the volume doubles the weight. Thus, the weight being supported by a pillar or a limb increases much faster than the corresponding increase in strength, because weight (like volume) scales as the cube of the linear dimensions whereas strength increases only as the square. To emphasize this point, consider increasing the height of a building or tree by a factor of 10 keeping its shape the same; then the weight needed to be supported increases a thousand fold (103) whereas the strength of the pillar or trunk holding it up increases by only a hundred fold (102). Thus, the ability to safely support the additional weight is only a tenth of what it had previously been. Consequently, if the size of the structure, whatever it is, is arbitrarily increased it will eventually collapse under its own weight. There are limits to size and growth. To put it slightly differently: relative strength becomes progressively weaker as size increases . Or, as Galileo so graphically put it: "the smaller the body the greater its relative strength. Thus a small dog could probably carry on his back two or three dogs of his own size; but I believe that a horse could not carry even one of his own size."*

September 18, 2023

This kind of scale where instead of increasing linearly as in 1, 2, 3, 4, 5 . . . we increase by factors of 10 as in the Richter scale: 101, 102, 103, 104, 105 . . . is called logarithmic. Notice that it's actually linear in terms of the numbers of orders of magnitude, as indicated by the exponents (the superscripts) on the tens. Among its many attributes, a logarithmic scale allows one to plot quantities that differ by huge factors on the same axis, such as those between the magnitudes of the Valdivia earthquake, the Northridge earthquake, and a stick of dynamite, which overall cover a range of more than a billion (109). This would be impossible if a linear plot was used because almost all of the events would pile up at the lower end of the graph. To include all earthquakes, which range over five or six orders of magnitude, on a linear plot would require a piece of paper several miles long—hence the invention of the Richter scale. 62

September 18, 2023

September 18, 2023 62

September 18, 2023 65

Let me illustrate this using the weight lifting example. If you look carefully at the graph of Figure 7 you can clearly see that four of the points lie almost exactly on the line, indicating that these weight lifters are lifting almost precisely what they should for their body weights. However, notice that the remaining two, the heavyweight and the middleweight, both lie just a little off the line, one below and one above. Thus, the heavyweight, even though he has lifted more than anyone else, is actually under performing relative to what he should be lifting given his weight, whereas the middleweight is over performing relative to his weight. In other words, from the egalitarian level playing field perspective of a physicist, the strongest man in the world in 1956 was actually the middleweight champion because he was overperforming relative to his weight. Ironically, the weakest of all of the champions from this scientific scaling perspective is the heavyweight, despite the fact that he lifted more than anyone else. 65

September 18, 2023

September 18, 2023 68

A classic example of some of the challenges and pitfalls is an early study investigating the potentially therapeutic effects of LSD on humans. Although the term "psychedelic" was already coined by 1957, the drug was almost unknown 68 *outside of a specialized psychiatric community in 1962, when the psychiatrist Louis West (no relation), together with Chester Pierce at the University of Oklahoma and Warren Thomas, a zoologist at the Oklahoma City zoo, proposed investigating its effects on elephants. Elephants? Yes, elephants and, in particular, Asiatic elephants. Although it may sound somewhat eccentric to use elephants rather than mice as the "model" for studying the effects of LSD, there were some not entirely implausible reasons for doing so. It so happens that Asiatic elephants periodically undergo an unpredictable transition from their normal placid obedient state to one in which they become highly aggressive and even dangerous for periods of up to two weeks. West and his collaborators speculated that this bizarre and often destructive behavior, known as musth , was triggered by the autoproduction of LSD in elephants' brains. So the idea was to see if LSD would induce this curious condition and, if so, thereby gain insight into LSD's effects on humans from studying how they react. Pretty weird, but maybe not entirely unreasonable. However, this immediately raises an intriguing question: how much LSD should you give an elephant? At that time little was known about safe dosages of LSD. Although it had not yet entered the popular culture it was known that even dosages of less than a quarter milligram would induce a typical "acid trip" for a human being and that a safe dose for cats was about one tenth of a milligram per kilogram of body weight. The investigators chose to use this latter number to estimate how much LSD they should give to Tusko the elephant, their unsuspecting subject that resided at the Lincoln Park Zoo in Oklahoma City. Tusko weighed about 3,000 kilograms, so using the number known to be safe for cats, they estimated that a safe and appropriate dose for Tusko would be about 0.1 milligram per kilogram multiplied by 3,000 kilograms, which comes out to 300 milligrams of LSD. The amount they actually injected was 297 milligrams. Recall that a good hit of LSD for you and me is less than a quarter milligram. The results on Tusko were dramatic and catastrophic. To quote* *directly from their paper: "Five minutes after the injection he [the elephant] trumpeted, collapsed, fell heavily onto his right side, defecated, and went into status epilepticus." Poor old Tusko died an hour and forty minutes later. Perhaps almost as disturbing as this awful outcome was that the investigators concluded that elephants are "proportionally very sensitive to LSD." The problem, of course, is something we've already stressed several times, namely the seductive trap of linear thinking. The calculation of how big a dosage should be used on Tusko was based on the implicit assumption that effective and safe dosages scale linearly with body weight so that the dosage per kilogram of body weight was presumed to be the same for all mammals. The 0.1 milligram per kilogram of body weight obtained from cats was therefore naively multiplied by Tusko's weight, resulting in the outlandish estimate of 297 milligrams, with disastrous consequences. Exactly how doses should be scaled from one animal to another is still an open question that depends, to varying degrees, on the detailed properties of the drug and the medical condition being addressed. However, regardless of details, an understanding of the underlying mechanism by which drugs are transported and absorbed into specific organs and tissues needs to be considered in order to obtain a credible estimate. Among the many factors involved, metabolic rate plays an important role. Drugs, like metabolites and oxygen, are typically transported across surface membranes, sometimes via diffusion and sometimes through network systems. As a result, the dose-determining factor is to a significant degree constrained by the scaling of surface areas rather than the total volume or weight of an organism, and these scale nonlinearly with weight. A simple calculation using the ⅔ scaling rule for areas as a function of weight shows that a more appropriate dose for elephants should be closer to a few milligrams of LSD rather than the several hundred that were actually administered. Had this been done, Tusko would no doubt have lived and a vastly different conclusion about the effects of LSD would have been drawn. The lesson is clear: the scaling of drug dosages* *is nontrivial, and a naive approach can lead to unfortunate results and mistaken conclusions if not done correctly with due attention being paid to the underlying mechanism of drug transport and absorption. It is clearly an issue of enormous importance, sometimes even a matter of life and death. This is one of the main reasons it takes so long for new drugs to obtain approval for their general use. Lest you think this was some fringe piece of research, the paper on elephants and LSD was published in one of the world's most highly regarded and prestigious journals, namely Science . 6*

September 18, 2023

September 18, 2023 71

September 18, 2023 74

So how in fact does weight scale with height for human beings? Various statistical analyses of data have led to varying conclusions, ranging from confirmation of the cubic law to more recent analyses suggesting exponents of 2.7 and values that are even smaller and closer to two. 10 To understand why this might be so, we have to remind ourselves of a major assumption that was made in deriving the cubic law, namely that the shape of the system, our bodies in this case, should remain the same when its size increases. However, shapes change with age, from the extreme case of a baby, with its large head and chunky limbs, to a mature "well-proportioned" adult, and finally to the sagging bodies of people my age. In addition, shapes also depend on gender, culture, and other socioeconomic factors that may or may not be correlated with health and obesity. 75

September 18, 2023

September 18, 2023 [75](#)

8. INNOVATION AND LIMITS TO GROWTH Galileo's deceptively simple argument for why there are limits to the heights of trees, animals, and buildings has profound consequences for design and innovation. Earlier, when explaining his argument I concluded with the remark: Clearly, the structure, whatever it is, will eventually collapse under its own weight if its size is arbitrarily increased. There are limits to size and growth. To which should have been added the critical phrase " unless something changes. " Change and, by implication, innovation, must occur in order to continue growing and avoid collapse. Growth and the continual need to be adapting to the challenges of new or changing environments, often in the form of "improvement" or increasing efficiency, are major drivers of innovation. [76](#)

September 18, 2023

A simple example of the first kind of innovation is to use a stronger material such as steel in place of wood for bridges or buildings, while a simple example of the second kind is to use arches, vaults, or domes in their construction rather than just horizontal beams and vertical pillars. The evolution of bridges is, in fact, an excellent example of how innovations in both materials and design were stimulated by the desire, or perceived requirement, to meet new challenges: in this case, to traverse wider and wider rivers, canyons, and valleys in a safe, resilient manner. [77](#)

September 18, 2023

Models of various kinds have been built for centuries, especially in architecture, but these were primarily to illustrate the aesthetic characteristics of the real thing rather than as scale models to test, investigate, or demonstrate the [80](#)

dynamical or physical principles of the system being constructed. And most important, they were almost always "made to scale," meaning that each detailed part was in some fixed proportion to the full size—1:10, for example—just like a map. Each part of the model was a linearly scaled representation of the actual-size ship, cathedral, or city being "modeled." Fine for aesthetics and toys but not much good for learning how the real system works. Nowadays, every conceivable process or physical object, from automobiles, buildings, airplanes, and ships to traffic congestion, epidemics, economies, and the weather, is simulated on computers as "models" of the real thing. I discussed earlier how specially bred mice are used in biomedical research as scaled-down "models" of human beings. In all of these cases, the big question is how do you realistically and reliably scale up the results and observations of the model system to the real thing? This entire way of thinking has its origins in a sad failure in ship design in the middle of the nineteenth century and the marvelous insights of a modest gentleman engineer into how to avoid it in the future.

September 18, 2023

Failure and catastrophe can provide a huge impetus and opportunity in stimulating innovation, new ideas, and inventions whether in science, engineering, finance, politics, or one's personal life.

80

September 18, 2023

One of his most fascinating innovations was the unique introduction of a broad gauge of 7 feet ¼ inch for the width between tracks. The standard gauge of 4 feet 8½ inches, which was used in all other British railways at that time, was adopted worldwide and is used on almost all railways today. Brunel pointed out that the standard gauge was an arbitrary carryover from the mine railways built before the invention of the world's first passenger trains in 1830. It had simply been determined by the width needed to fit a cart horse between the shafts that pulled carriages in the mines. Brunel

82

rightly thought that serious consideration should be given to determining what the optimum gauge should be and tried to bring some rationality to the issue. He claimed that his calculations, confirmed by a series of trials and experiments, showed that his broader gauge was the optimum size for providing higher speeds, greater stability, and a more comfortable ride to passengers. Consequently, the Great Western Railway was unique in having a gauge that was almost twice as wide as every other railway line. Unfortunately, in 1892, following the evolution of a national railway system, the British Parliament forced the Great Western Railway to conform to the standard gauge, despite its acknowledged inferiority.

September 18, 2023

84

Brunel thought otherwise. His conclusions were based on a simple scaling argument. He realized that the volume of cargo a ship could carry increases as the cube of its dimensions (like its weight), whereas the strength of the drag forces it experiences as it travels through water increases as the cross-sectional area of its hull and therefore only as the square of its dimensions. This is just like Galileo's conclusions for how the strength of beams and limbs scale with body weight. In both cases the strength increases more slowly than the corresponding weight following a ⅔ power scaling law. Thus the strength of the hydrodynamic drag forces on a ship relative to the weight of the cargo it can carry decreases in direct proportion to the length of the ship.

September 18, 2023

85

In other words, a larger ship requires proportionately less fuel to transport each ton of cargo than a smaller ship . Bigger ships are therefore more energy efficient and cost effective than smaller ones—another great example of an economy of scale and one that had enormous consequences for the development of world trade and commerce. 12

September 18, 2023

September 18, 2023 87

September 18, 2023 90

It was only in the decade prior to the building of the Great Eastern that the underlying science governing the motion of ships was developed. The field of hydrodynamics was first formalized independently by the French engineer Claude-Louis Navier and the great Irish mathematical physicist George Stokes. The fundamental equation, which is universally known as the Navier-Stokes equation, arises from applying Newton's laws to the motion of fluids, and by extension to the dynamics of physical objects moving through fluids, such as ships through water or airplanes through air. 90

September 19, 2023

This is manifested in all kinds of fascinating behaviors and patterns such as those we see in the eddies and whirlpools of rivers and streams, or the wakes left by ships as they move through water, or the awesome specter of hurricanes and tornadoes and the beauty and infinite variation of ocean waves. All of these are manifestations of turbulence and are encapsulated in the hidden richness of the Navier-Stokes equation. Indeed, studying turbulence gave us the first important mathematical insights into the concept of complexity and its relationship to nonlinearity. Complex systems often manifest chaotic behavior in which a small change or perturbation in one part of the system produces an exponentially enhanced response in some other part. As we discussed earlier, in traditional linear thinking a small perturbation produces a commensurately small response. 91

The highly nonintuitive enhancement in nonlinear systems is popularly expressed as "the butterfly effect," in which the mythical flapping of a butterfly's wings in Brazil produces a hurricane in Florida. Despite 150 years of intense theoretical and experimental study, a general understanding of turbulence remains an unsolved problem in physics even though we have learned an enormous amount about it. Indeed the famous physicist Richard Feynman described turbulence as "the most important unsolved problem of classical physics." 15

September 19, 2023

Froude may not have fully recognized just how big a challenge he was facing but he did perceive that for applications to shipbuilding a new strategy was needed. It was in this context that he invented the new methodology of modeling and, by extension, the concept of scaling theory for determining how quantitative results from small-scale investigations could be used to help predict how full-size ships would behave. In the spirit of Galileo, Froude realized that almost all scaling is nonlinear, so traditional models based on a faithful 1:1 representation were not useful for determining how the actual system works. His seminal contribution was to suggest a quantitative mathematical strategy for figuring out how to scale from the small-size model to the full-size object.

92

September 19, 2023

Like many new ideas that threaten to change the way we think about an old problem, Froude's endeavors were at first dismissed as irrelevant by the cognoscenti of the time. John Russell, who founded the Institution for Naval Architects in England in 1860 in order to encourage the formal education of ship designers, ridiculed Froude: "You will have on the small scale a series of beautiful, interesting little experiments which, I am sure, will afford Mr. Froude infinite pleasure in the making of them . . . and will afford you infinite pleasure in the hearing of them; but which are quite remote from any

92

practical results upon the large scale." Many of us recognize this kind of rhetoric, often aimed at scholarly or academic research with the implication that it is out of touch with the "real world." Well, no doubt, much of it is. But much of it isn't and, more to the point, it is very often difficult to perceive in the moment the potential impact of some piece of seemingly arcane research. Much of our entire technologically driven society and extraordinary quality of life that many of us are privileged to enjoy is the result of such research. There is a continuing tension in society between supporting what is perceived as pie-in-the-sky basic research with no obvious immediate benefit versus highly directed research focused on "useful, real world" problems.

September 19, 2023

He realized that the primary quantity that determined the character of their relative motion was something that later became known as the Froude number . It is defined as the square of the ship's velocity divided by its length multiplied by the acceleration due to gravity. This is a bit of a mouthful and may sound a little intimidating, but it's actually quite simple because "the acceleration due to gravity" that occurs in this expression is the same for all objects, regardless of their size, shape, or composition. This is just a restatement of Galileo's observations that falling objects of different weights reach the ground in the same time. So in terms of the quantities that actually vary, Froude's number is simply proportional to the velocity squared divided by the length. This ratio plays a central role in all problems involving motion, ranging from speeding bullets and running dinosaurs to flying airplanes and sailing ships. 93

September 19, 2023

The crucial point recognized by Froude was that because the underlying physics remains the same, objects of different sizes moving at different speeds behave in the same way if their Froude numbers have the same value . Thus, by making the length and speed of the model ship have the same Froude 94 *number as that of the full-size version, one can determine the dynamical behavior of the full-size ship before building it.*

September 19, 2023

11. SIMILARITY AND SIMILITUDE: DIMENSIONLESS AND SCALE-INVARIANT NUMBERS The scaling methodology introduced by Froude has now evolved to become a powerful and sophisticated component of the tool kit of science and engineering and has been applied with great effect to a very broad and diverse range of problems. It was not formalized as a general technique until the beginning of the twentieth century, when the eminent mathematical physicist Lord Rayleigh wrote a provocative and highly influential paper in the journal Nature titled "The Principle of Similitude." 16 This was his term for what we have been calling scaling theory. His major emphasis was on the primary role played in any physical system by special quantities that have the property of being dimensionless.

96

September 19, 2023

This is the simplest example of a dimensionless quantity: it is a "pure" number that does not change when a different system of units is used to measure it. This scale invariance expresses something absolute about the quantities they represent in that the dependence on the arbitrariness of the human choice of units and measurement has been removed. Specific units are convenient inventions of human beings for communicating measures in a standardized language, especially regarding construction, commerce, and the exchange of goods and services. Indeed, the introduction of standardized measures marked a critical stage in the evolution of civilization and the rise of cities, as they were crucial in developing a trustworthy political fabric subject to the rule of law. Perhaps the most famous dimensionless number is pi (π), the ratio of the circumference of a circle to its diameter. This has no units, because it is the ratio of two lengths, and has the same value for all circles everywhere, at all times, no matter how small or how large they are. π

96

therefore embodies the universal quality of "circleness." This concept of "universality" is the reason why the acceleration due to gravity was included in the definition of the Froude number, even though it played no explicit role in how to scale from model ships to the real thing. It turns out that the ratio of the square of the velocity to the length is not dimensionless and so depends on the units used, whereas dividing by gravity renders it dimensionless and therefore scale invariant. But why was gravity chosen and not some other acceleration? Because gravity ubiquitously constrains all motion on Earth. This is pretty evident in our own walking and running where we have to continually fight gravity to raise our legs with each step forward, especially when going uphill. Not quite so obvious is how it enters into the motion of ships, because the buoyancy of water balances gravity (remember Archimedes' principle). However, as a ship moves through water it continually creates wakes and surface waves whose behavior is constrained by the pull of gravity—in fact, the waves you are familiar with on oceans and lakes are technically called gravity waves. So indirectly gravity plays an important role in the motion of ships. Consequently, Froude's number embodies a "universal" quality associated with all motion on Earth transcending the specific details of the object that is moving. Its value is therefore a major determinant not only in the motion of ships, but also for cars, airplanes, and ourselves. Furthermore, it also tells us how these motions on another planet with a different gravitational strength would differ from the same motion on Earth.

Di

September 19, 2023

98

Because the essence of any measurable quantity cannot depend on an arbitrary choice of units made by human beings, neither can the laws of physics. Consequently, all of these and indeed all of the laws of science must be expressible as relationships between scale-invariant dimensionless quantities, even though conventionally they are not typically written that way. This was the underlying message of Rayleigh's seminal paper. His paper elegantly illustrates the *technique with many well-chosen examples, including one that provides the scientific explanation for one of the great mysteries of life that all of us have pondered at some time, namely, why is the sky blue? Using an elegant argument based solely on relating purely dimensionless quantities, he shows that the intensity of light waves scattered by small particles must decrease with the fourth power of their wavelength. Thus, when sunlight, which is a combination of all of the colors of the rainbow, scatters from microscopic particles suspended in the atmosphere, the shortest wavelengths, corresponding to blue light, dominate. Actually, Rayleigh had derived this stunning result much earlier in a tour de force based on a masterful mathematical analysis of the problem that provided a detailed mechanistic explanation for the origin of the shift toward the blue end of the spectrum. His point in presenting the simple derivation in his similitude paper was to show that the same result could have been derived "after a few minutes' consideration," as he put it, using a scaling argument in the guise of the "great principle of similitude" without the detailed sophisticated mathematics. His scaling argument showed that the shift to shorter wavelengths is an inevitable outcome of any analysis once you know what the important variables are. What is missing from such a derivation is a deeper understanding of the mechanism by which the result comes about. This is characteristic of many scaling arguments: general results can be derived, but details of their mechanistic origins sometimes remain hidden.*

September 19, 2023

September 19, 2023 101

3: THE SIMPLICITY, UNITY, AND COMPLEXITY OF LIFE

3 THE SIMPLICITY, UNITY, AND COMPLEXITY OF LIFE As 101 *emphasized in the opening chapter, living systems, from the smallest bacteria to the largest cities and ecosystems, are quintessential complex adaptive systems operating over an enormous range of multiple spatial, temporal, energy, and mass scales. In terms of mass alone, the overall scale of life covers more than thirty orders of magnitude (1030) from the molecules that power metabolism and the genetic code up to ecosystems and cities. This range vastly exceeds that of the mass of the Earth relative to that of our entire galaxy, the Milky Way, which covers "only" eighteen orders of magnitude, and is comparable to the mass of an electron relative to that of a mouse.*

September 19, 2023

All of life functions by transforming energy from physical or chemical sources into organic molecules that are metabolized to build, maintain, and reproduce complex, highly organized systems. This is accomplished by the operation of two distinct but closely interacting systems: the genetic code, which stores and processes the information and "instructions" to build and maintain the organism, and the metabolic system, which acquires, transforms, and allocates energy and materials for maintenance, growth, and reproduction. Considerable progress has been made in elucidating both of these systems at levels from molecules to organisms, and later I will discuss how it can be extended to cities and companies. However, understanding how information processing ("genomics") integrates with the processing of energy and resources ("metabolics") to sustain life remains a major challenge. Finding the universal principles that underlie the structures, dynamics, and integration of these systems is fundamental to understanding life and to managing biological and socioeconomic systems in such diverse contexts as medicine, agriculture, and the environment. 101

September 19, 2023

September 19, 2023 102

The search for fundamental principles that govern how the complexity of life emerges from its underlying simplicity is one of the grand challenges of twenty-first-century science. Although this has been, and will continue to be, primarily the purview of biologists and chemists, it is becoming an activity where other disciplines, and in particular physics and computer science, are playing an increasingly important role. 102

The New Kind of Science.

September 19, 2023

The field of physics is concerned with fundamental principles and concepts at all levels of organization that are quantifiable and mathematizable (meaning amenable to computation), and can consequently lead to precise predictions that can be tested by experiment and observation. From this perspective, it is natural to ask if there are "universal laws of life" that are mathematizable so that biology could also be formulated as a predictive, quantitative science much like physics. Is it conceivable that there are yet-to-be-discovered "Newton's Laws of Biology" that would lead, at least in principle, to precise calculations of any biological process so that, for instance, one could accurately predict how long you and I would live? This seems very unlikely. After all, life is the complex system par excellence, exhibiting many levels of emergent phenomena arising from multiple contingent histories. Nevertheless, it may not be unreasonable to conjecture that the generic coarse-grained behavior of living systems might obey quantifiable universal laws that capture their essential features. This more modest view presumes that at every organizational level average idealized biological systems can be constructed whose general properties are calculable. Thus we ought to be able to calculate the average and maximum life span of human beings even if we'll never be able to 103 *calculate our own. This provides a point of departure or baseline for quantitatively understanding actual biosystems, which can be viewed as variations or perturbations around idealized norms due to local environmental conditions or historical evolutionary divergence. I will elaborate on this perspective in much greater depth below, as it forms the conceptual strategy for attacking most of the questions posed in the opening chapter.*

September 19, 2023

September 19, 2023 103

September 19, 2023 104

I was a theoretical physicist (and still am) whose primary research interests at that time were in this area. My visceral reaction to the provocative statements concerning the diverging trajectories of physics and biology was that, yes, biology will almost certainly be the predominant science of the twenty-first century, but for it to become truly successful, it will need to embrace some of the quantitative, analytic, predictive culture that has made physics so successful. 105

September 19, 2023

September 19, 2023 106

Death is an essential feature of life. Indeed, implicitly it is an essential feature of the theory of evolution. A necessary component of the evolutionary process is that individuals eventually die so that their offspring can propagate new 107 *combinations of genes that eventually lead to adaptation by natural selection of new traits and new variations leading to the diversity of species. We must all die so that the new can blossom, explore, adapt, and evolve. Steve Jobs put it succinctly 3 : No one wants to die. Even people who want to go to heaven don't want to die to get there. And yet death is the destination we all share. No one has ever escaped it, and that is how it should be, because death is very likely the single best invention of life. It's life's change agent. It clears out the old to make way for the new.*

September 19, 2023

As I embarked on this new direction, I received unexpected support for my ruminations about biology as a science and its relationship to mathematics from an unlikely source. I discovered that what I had presumed was subversive thinking had been expressed much more articulately and deeply almost one hundred years earlier by the eminent and somewhat eccentric biologist Sir D'Arcy Wentworth Thompson in his classic book On Growth and Form, published in 1917. 4 It's a wonderful book that has remained quietly revered not just in biology but in mathematics, art, and architecture, influencing thinkers and artists from Alan Turing and Julian Huxley to Jackson Pollock. A testament to its continuing popularity is that it still remains in print. The distinguished biologist Sir Peter Medawar, the father of organ transplants, who received the Nobel Prize for his work on graft rejection and acquired immune tolerance, called On Growth and Form "the finest work of literature in all the annals of science that have been recorded in the English tongue." 108

September 19, 2023

Although he was awarded the prestigious Darwin Medal by the Royal Society in 1946, Thompson was critical of conventional Darwinian evolutionary theory because he felt that biologists overemphasized the role of natural selection and the "survival of the fittest" as the fundamental 109 *determinants of the form and structure of living organisms, rather than appreciating the importance of the role of physical laws and their mathematical expression in the evolutionary process. The basic question implicit in his challenge remains unanswered: are there "universal laws of life" that can be mathematized so that biology can be formulated as a predictive quantitative Science? He put it this way: It behoves us always to remember that in physics it has taken great men to discover simple things. . . . How far even then mathematics will suffice to describe, and physics to explain, the fabric of the body, no man can foresee. It may be that all the laws of energy, and all the properties of matter, and all the chemistry of all the colloids are as powerless to explain the body as they are impotent to comprehend the soul. For my part, I think it is not so. Of how it is that the soul informs the body, physical science teaches me nothing; and that living matter influences and is influenced by mind is a mystery without a clue. Consciousness is not explained to my comprehension by all the nerve-paths and neurons of the physiologist; nor do I ask of physics how goodness shines in one man's face, and evil betrays itself in another. But of the construction and growth and working of the body, as of all else that is of the earth earthy, physical science is, in my humble opinion, our only teacher and guide.*

September 19, 2023

Although Thompson's book did not address aging or death, nor was it particularly helpful or sophisticated technically, its philosophy provided support and inspiration for contemplating and applying ideas and techniques from physics to all sorts of problems in biology. In my own thinking, this led me to perceive our bodies as metaphorical machines that need to be fed, maintained, and repaired but which eventually wear out and "die," much like cars and washing machines. However, to understand how something ages and dies, whether an animal, an automobile, a company, or a civilization, one first needs to understand what the processes and mechanisms are that are keeping it alive, and then discern how these become degraded with time. This naturally leads to considerations of the energy and resources that are required for sustenance and possible 111 *growth, and their allocation to maintenance and repair for combating the production of entropy arising from destructive forces associated with damage, disintegration, wear and tear, and so on. This line of thinking led me to focus initially on the central role of metabolism in keeping us alive before asking why it can't continue doing so forever.*

September 19, 2023

Metabolism is the fire of life . . . and food, the fuel of life . 111

September 19, 2023

Like all of life, we evolved by a process of natural selection, interacting with and adapting to our fellow creatures, whether bacteria and viruses, ants and beetles, snakes and spiders, cats and dogs, or grass and trees, and everything else in a continuously challenging and evolving environment. We have been coevolving together in a never-ending multidimensional interplay of interaction, conflict, and adaptation. Each organism, each of its organs and subsystems, each cell type and genome, has therefore evolved following its own unique history in its own ever-changing environmental niche. The principle of natural selection, introduced independently by Charles Darwin and Alfred Russell Wallace, is key to the theory of evolution and the origin of species. Natural selection, or the "survival of the fittest," is the gradual process by which a successful variation in some inheritable trait or characteristic becomes fixed in a population through the differential reproductive success of organisms that have developed this trait by interacting with their environment. As Wallace expressed it, there is sufficiently broad variation that "there is always material for natural selection to act upon in some direction that may be advantageous," or as put more succinctly by Darwin: "each slight variation, if useful, is preserved." 112

September 19, 2023

If we now consider a cow that is 100 times heavier than the cat, then Kleiber's law predicts that its metabolic rate is likewise 32 times greater than the cat's, and if we extend this to a whale that is 100 times heavier than the cow, its metabolic rate would be 32 times greater than the cow's. This repetitive behavior, the recurrence in this case of the same factor 32 as we move up in mass by the same repetitive factor of 100, is an example of the general self-similar feature of power laws. More generally: if the mass is increased by any arbitrary factor at any scale (100, in the example), then the metabolic rate increases by the same factor (32, in the example) no matter what the value of the initial mass is, that is, whether it's that of a mouse, cat, cow, or whale. This remarkably systematic repetitive behavior is called scale invariance or self-similarity and is a property inherent to power laws. It is closely related to the concept of a fractal, which will be discussed in detail in the following chapter. To varying degrees, fractality, scale invariance, and self-similarity are ubiquitous across nature from galaxies and clouds to your cells, your brain, the Internet, companies, and cities. We just saw that a cat that is 100 times heavier than a mouse requires only about 32 times as much energy to sustain it even though it has approximately 100 times as many cells—a classic example of an economy of scale resulting from the essential nonlinear nature of Kleiber's law . Naive linear reasoning would have predicted the cat's metabolic rate to have been 100 times larger, rather than only 32 times. Similarly, if the size of an animal is doubled it doesn't need 100 percent more energy to sustain it; it needs only about 75 percent more—thereby saving approximately 25 percent with each doubling. Thus, in a systematically predictable and quantitative way, the larger the organism the less energy has to be produced per cell per second to sustain a gram of tissue. Your cells work less hard than your dog's, but your horse's work even less hard. Elephants are roughly 10,000 times heavier than rats but their metabolic rates are only 1,000 times larger, despite having roughly 10,000 times as many cells to support. Thus, an elephant's cells operate at about a tenth the rate of a rat's, resulting in a corresponding decrease in the rates of cellular damage, and consequently to a greater longevity for the elephant, as will be explained in greater detail in chapter 4. This is an example *of how the systematic economy of scale has profound consequences that reverberate across life from birth and growth to death.*

September 19, 2023

September 19, 2023 116

4. UNIVERSALITY AND THE MAGIC NUMBER FOUR THAT CONTROLS LIFE The systematic regularity of Kleiber's law is pretty amazing, but equally surprising is that similar systematic scaling laws hold for almost any physiological trait or life-history event across the entire range of life from cells to whales to ecosystems. In addition to metabolic rates, these include quantities such as growth rates, genome lengths, lengths of aortas, tree heights, the amount of cerebral gray matter in the brain, evolutionary rates, and life spans; a sampling of these is illustrated in Figures 9-12. There are probably well over fifty such scaling laws and—another big surprise—their corresponding exponents (the analog of the ¾ in Kleiber's law) are invariably very close to simple multiples of ¼. For example, the exponent for growth rates is very close to ¾, for lengths of aortas and genomes it's ¼, for heights of trees ¼, for cross-sectional areas of both aortas and tree trunks ¾, for brain sizes ¾, for cerebral white and gray matter 5/4, for heart rates minus ¼, for mitochondrial densities in cells minus ¼, for rates of evolution minus ¼, for diffusion rates across membranes minus ¼, for life spans ¼ . . . and many, many more. The "minus" here simply indicates that the corresponding quantity decreases with size rather than increases, so, for instance, heart rates decrease with increasing body size following the ¼ power law, as shown in Figure 10. I can't resist drawing your attention to the intriguing fact that aortas and tree trunks scale in the same way. Particularly fascinating is the emergence of the number four in the guise of the ¼ powers that appear in all of these exponents. It occurs ubiquitously across the entire panoply of life and seems to play a special, fundamental role in determining 117 *many of the measurable characteristics of organisms regardless of their evolved design. Viewed through the lens of scaling, a remarkably general universal pattern emerges, strongly suggesting that evolution has been constrained by other general physical principles beyond natural selection.*

September 19, 2023

Doubling the mass of a mammal increases all of its timescales such as its life span and time to maturity by about 25 percent on average and, concomitantly, decreases all rates, such as its heart rate, by the same amount. Whales live in the ocean, elephants have trunks, giraffes have long necks, we walk on two legs, and dormice scurry around, yet despite these obvious differences, we are all, to a large degree, nonlinearly scaled versions of one another. If you tell me the size of a mammal, I can use the scaling laws to tell you almost everything about the average values of its measurable characteristics: how much food it needs to eat each day, what its heart rate is, how long it will take to mature, the length and radius of its aorta, its life span, how many offspring it will have, and so on. Given the extraordinary complexity and diversity of life, this is pretty amazing. 118

September 19, 2023

September 19, 2023 121

Indeed, Huxley coined the term allometric to describe how physiological and morphological characteristics of organisms scale with body size, though his focus was primarily on how that occurred during growth. Allometric was introduced as a generalization of the Galilean concept of isometric scaling, discussed in the previous chapter, where body shape and geometry do not change as size increases, so all lengths associated with an organism increase in the same 121 *proportion; iso is Greek for "the same," and metric is derived from metrikos , meaning "measure." Allometric , on the other hand, is derived from allo, meaning "different," and refers to the typically more general situation where shapes and morphology change as body size increases and different dimensions scale differently. For example, the radii and lengths of tree trunks, or for that matter the limbs of animals, scale differently from one another as size increases: radii scale as the ⅜ power of mass, whereas lengths scale more slowly with ¼ power (that is, as the 2⁄8 power).*

September 20, 2023

September 19, 2023 122

September 20, 2023 122

Huxley's term allometric was extended from its more restrictive geometric, morphological, and ontogenetic origins to describe the kinds of scaling laws that I discussed above, which include more dynamical phenomena such as how flows of energy and resources scale with body size, with metabolic rate being the prime example. All of these are now commonly referred to as allometric scaling laws . Julian Huxley, himself a very distinguished biologist, was the grandson of the famous Thomas Huxley, the biologist who championed Charles Darwin and the theory of evolution by natural selection, and the brother of the novelist and futurist Aldous Huxley. In addition to the word allometric, Julian Huxley brought several other new words and concepts into biology, including replacing the much-maligned term race with the phrase ethnic group . 122

September 20, 2023

September 20, 2023 122

As I have emphasized, no aspect of life can function without energy. Just as every muscle contraction or any activity requires metabolic energy, so does every random thought in your brain, every twitch of your body even while you sleep, and even the replication of your DNA in your cells. At the most fundamental biochemical level metabolic energy is created in semiautonomous molecular units within cells called respiratory complexes . The critical molecule that plays the central role in metabolism goes by the slightly forbidding name of adenosine triphosphate , usually referred to as ATP. The detailed biochemistry of metabolism is extremely complicated but in essence it involves the breaking down of ATP, which is relatively unstable in the cellular environment, from adenosine tri phosphate (with three phosphates) into ADP, adenosine di phosphate (with just two phosphates), thereby releasing the energy stored in the binding of the additional phosphate. The energy derived from breaking this phosphate bond is the source of your metabolic energy and therefore what is keeping you alive. The reverse process converts ADP back into ATP using energy from food via oxidative respiration in mammals such as ourselves (that's why we have to breathe in oxygen), or via photosynthesis in plants. The cycle of releasing energy from the breakup of ATP into ADP and its recycling back from ADP to store energy in ATP forms a continuous loop process much like the charging and recharging of a battery. 124

September 20, 2023

Given its central role, it's not surprising that the flux of ATP is often referred to as the currency of metabolic energy for almost all of life. At any one time our bodies contain only about half a pound (about 250 g) of ATP, but here's something truly extraordinary that you should know about yourself: every day you typically make about 2 × 1026 ATP molecules—that's two hundred trillion trillion molecules— 125 *corresponding to a mass of about 80 kilograms (about 175 lbs.). In other words, each day you produce and recycle the equivalent of your own body weight of ATP! Taken together, all of these ATPs add up to meet our total metabolic needs at the rate of the approximately 90 watts we require to stay alive and power our bodies.*

September 20, 2023

These little energy generators, the respiratory complexes, are situated on crinkly membranes inside mitochondria , which are potato-shaped objects floating around inside cells. Each mitochondrion contains about five hundred to one thousand of these respiratory complexes . . . and there about five hundred to one thousand of these mitochondria inside each of your cells, depending on the cell type and its energy needs. Because muscles require greater access to energy, their cells are densely packed with mitochondria, whereas fat cells have many fewer. So on average each cell in your body may have up to a million of these little engines distributed among its mitochondria working away night and day, collectively manufacturing the astronomical number of ATPs needed to keep you viable, healthy, and strong. The rate at which the total number of these ATPs is produced is a measure of your metabolic rate. Your body is composed of about a hundred trillion (1014) cells. Even though they represent a broad range of diverse functionalities from neuronal and muscular to protective (skin) and storage (fat), they all share the same basic features. They all process energy in a similar way via the hierarchy of respiratory complexes and mitochondria. Which raises a huge challenge: the five hundred or so respiratory complexes inside your mitochondria cannot behave as independent entities but have to act collectively in an integrated coherent fashion in order to ensure that mitochondria function efficiently and deliver energy in an appropriately ordered fashion to cells. Similarly, the five hundred or so mitochondria inside each of your cells do not act independently but, like respiratory complexes, have to interact in an integrated coherent fashion to ensure that the 1014 cells that constitute your body are supplied with the energy they need to function efficiently and appropriately. Furthermore, these hundred trillion cells have to be 125 *organized into a multitude of subsystems such as your various organs, whose energy needs vary significantly depending on demand and function, thereby ensuring that you can do all of the various activities that constitute living, from thinking and dancing to having sex and repairing your DNA. And this entire interconnected multilevel dynamic structure has to be sufficiently robust and resilient to continue functioning for up to one hundred years!*

September 20, 2023

September 20, 2023 128

128

As I began to ponder what the origins of these surprising scaling laws might be, it became clear that whatever was at play had to be independent of the evolved design of any specific type of organism, because the same laws are manifested by mammals, birds, plants, fish, crustacea, cells, and so on. All of these organisms ranging from the smallest, simplest bacterium to the largest plants and animals depend for their maintenance and reproduction on the close integration of numerous subunits—molecules, organelles, and cells—and these microscopic components need to be serviced in a relatively "democratic" and efficient fashion in order to supply metabolic substrates, remove waste products, and regulate activity. Natural selection has solved this challenge in perhaps the simplest possible way by evolving hierarchical branching networks that distribute energy and materials between macroscopic reservoirs and microscopic sites. Functionally, biological systems are ultimately constrained by the rates at which energy, metabolites, and information can be supplied through these networks. Examples include animal circulatory, respiratory, renal, and neural systems, plant vascular systems, intracellular networks, and the systems that supply food, water, power, and information to human societies. In fact, when you think about it, you realize that underneath your smooth skin you are effectively an integrated series of such networks, each busily transporting metabolic energy, *materials, and information across all scales.*

September 20, 2023

In high energy physics, where we struggle to unravel the basic laws of nature at the most microscopic level, we mostly know what the questions are and most of one's effort goes into trying to be clever enough to carry out the highly technical calculations. In biology I found it to be mostly the other way around: months were spent trying to figure out what the problem actually was that we were trying to solve, the questions we should be asking, and the various relevant quantities that were needed to be calculated, but once that was accomplished, the actual technical mathematics was relatively straightforward. 132

September 20, 2023

To varying degrees, all theories and models are incomplete. They need to be continually tested and challenged by increasingly accurate experiments and observational data over wider and wider domains and the theory modified or extended accordingly. This is an essential ingredient in the scientific method. Indeed, understanding the boundaries of their applicability, the limits to their predictive power, and the ongoing search for exceptions, violations, and failures has provoked even deeper questions and challenges, stimulating the continued progress of science and the unfolding of new ideas, techniques, and concepts. 134

September 20, 2023

So what was irrelevant at one scale can become dominant at another. The challenge at every level of observation is to abstract the important variables that determine the dominant behavior of the system. 135

September 20, 2023

Physicists have coined a concept to help formalize a first step in this approach, which they call a "toy model." The strategy is to simplify a complicated system by abstracting its essential components, represented by a small number of dominant variables, from which its leading behavior can be determined. A classic example is the idea first proposed in the nineteenth century that gases are composed of molecules, viewed as hard little billiard balls, that are rapidly moving and colliding with one another and whose collisions with the surface of a container are the origin of what we identify as pressure. [135](#)

September 20, 2023

A concept related to the idea of a toy model is that of a "zeroth order" approximation of a theory, in which simplifying assumptions are similarly made in order to give a rough approximation of the exact result. It is usually employed in a quantitative context as, for example, in the statement that "a zeroth order estimate for the population of the Chicago metropolitan area in 2013 is 10 million people." Upon learning a little more about Chicago, one might make what could be called a "first order" estimate of its population of 9.5 million, which is more precise and closer to the actual number (whose precise value from census data is 9,537,289). One could imagine that after more detailed investigation, an even better estimate would yield 9.54 million, which would be called a "second order" estimate. You get the idea: each succeeding "order" represents a refinement, an improved approximation, or a finer resolution that converges to the exact result based on more detailed investigation and analysis. In what follows, I shall be using the terms "coarse-grained" and "zeroth order" interchangeably. [136](#)

September 20, 2023

September 20, 2023 138

However, I agree with Sydney Brenner, the distinguished biologist who received the Nobel Prize for his work on the genetic code and who provocatively remarked that "technology gives us the tools to analyze organisms at all scales, but we are drowning in a sea of data and thirsting for some theoretical framework with which to understand it. . . . We need theory and a firm grasp on the nature of the objects we study to predict the rest." His article begins, by the way, with the astonishing pronouncement that "biological research is in crisis." 10 138

September 20, 2023

The time seems right for revisiting D'Arcy Thompson's challenge: "How far even then mathematics will suffice to describe, and physics to explain, the fabric of the body, no man can foresee. It may be that all the laws of energy, and all the properties of matter, all . . . chemistry . . . are as powerless to explain the body as they are impotent to comprehend the soul. For my part, I think it is not so." Many would agree with the spirit of this remark, though new tools and concepts, including closer collaboration, may well be needed to accomplish his lofty goal. 138

September 20, 2023

Prior to this digression into the interrelationship between the cultures of biology and physics, I argued that the mechanistic origins of scaling laws in biology were rooted in the universal mathematical, dynamical, and organizational properties of the multiple networks that distribute energy, materials, and information to local microscopic sites that permeate organisms, such as cells and mitochondria in animals. I also argued that because the structures of biological networks are so varied and stand in marked 138 *contrast to the uniformity of the scaling laws, their generic properties must be independent of their specific evolved design. In other words, there must be a common set of network properties that transcends whether they are constructed of tubes as in mammalian circulatory systems, fibers as in plants and trees, or diffusive pathways as in cells.*

September 20, 2023

September 20, 2023 139

I. Space Filling The idea behind the concept of space filling is simple and intuitive. Roughly speaking, it means that the tentacles of the network have to extend everywhere throughout the system that it is serving, as is illustrated in the networks here. More specifically: whatever the geometry and topology of the network is, it must service all local biologically active subunits of the organism or subsystem. A familiar example will make it clear: Our circulatory system is a classic hierarchical branching network in which the heart pumps blood through the many levels of the network beginning with the main arteries, passing through vessels of regularly decreasing size, ending with the capillaries, the smallest ones, before looping back to the heart through the venal network system. Space filling is simply the statement that the capillaries, which are the terminal units or last branch of the network, have to service every cell in our body so as to efficiently supply each of them with sufficient blood and oxygen.

139

September 20, 2023

September 20, 2023 140

II. The Invariance of Terminal Units This simply means that the terminal units of a given network design, such as the capillaries of the circulatory system that we just discussed, all have approximately the same size and characteristics regardless of the size of the organism. Terminal units are critical elements of the network because they are points of delivery and transmission where energy and resources are exchanged. Other examples are mitochondria within cells, cells within bodies, and petioles (the last branch) of plants and trees. As individuals grow from newborn to adult, or as new species of varying sizes evolve, terminal units do not get reinvented nor are they significantly reconfigured or rescaled. For example, the capillaries of all mammals, whether children, adults, mice, elephants, or whales, are essentially all the same despite the enormous range and variation of body sizes. 140

September 20, 2023

This invariance of terminal units can be understood in the context of the parsimonious nature of natural selection. Capillaries, mitochondria, cells, et cetera, act as "ready-made" basic building blocks of the corresponding networks for new species, which are rescaled accordingly. The invariant properties of the terminal units within a specific design characterize the taxonomic class. For instance, all mammals share the same capillaries. Different species within that class such as elephants, humans, and mice are distinguished from one another by having larger or smaller, but closely related, network configurations. 141

September 20, 2023

The final postulate states that the continuous multiple feedback and fine-tuning mechanisms implicit in the ongoing processes of natural selection and which have been playing out over enormous periods of time have led to the network performance being "optimized." So, for example, the energy used by the heart of any mammal, including us, to pump blood through the circulatory system is on average 142 *minimized. That is, it is the smallest it could possibly be given its design and the various network constraints. To put it slightly differently: of the infinite number of possibilities for the architecture and dynamics of circulatory systems that could have evolved, and that are space filling with invariant terminal units, the ones that actually did evolve and are shared by all mammals minimize cardiac output. Networks have evolved so that the energy needed to sustain an average individual's life and perform the mundane tasks of living is minimized in order to maximize the amount of energy available for sex, reproduction, and the raising of offspring. This maximization of offspring is an expression of what is referred to as Darwinian fitness, w hich is the genetic contribution of an average individual to the next generation's gene pool. This naturally raises the question as to whether the dynamics and structure of cities and companies are the result of analogous optimization principles. What, if anything, is optimized in their multiple network systems? Are cities organized to maximize social interactions, or to optimize transport by minimizing mobility times, or are they ultimately driven by the ambition of each citizen and company to maximize their assets, profits, and wealth? I will return to these issues in chapters 8, 9, and 10.*

September 20, 2023

September 20, 2023 142

Their modern formulation is a general mathematical framework in which a quantity called the action , which is loosely related to energy, is minimized. All the laws of physics can be derived from the principle of least action which, roughly speaking, states that, of all the possible configurations that a system can have or that it can follow as it evolves in time, the one that is physically realized is the one that minimizes its action. Consequently, the dynamics, structure, and time evolution of the universe since the Big Bang, everything from black holes and the satellites transmitting your cell phone messages to the cell phones and 143 *messages themselves, all electrons, photons, Higgs particles, and pretty much everything else that is physical, are determined from such an optimization principle. So why not life?*

September 20, 2023

It is important to recognize that the three postulates enunciated above are to be understood in a coarse-grained average sense. Let me explain. It may have occurred to you that there must be variation among the almost trillion capillaries in any individual human body, as there must be across all the species of a given taxonomic group, so strictly speaking capillaries cannot be invariant. However, this variation has to be viewed in a relative scale-dependent way. The point is that any variation among capillaries is extremely small compared with the many orders of magnitude variation in body size. For instance, even if the length of mammalian capillaries varied by a factor of two, this is still tiny compared with the factor of 100 million in the variation of their body masses. Similarly, there is relatively little variation in petioles, the last branch of a tree prior to the leaf, or even in the size of leaves themselves, during the growth of a tree from a tiny sapling to a mature tree that might be a hundred or more feet high. This is also true across species of trees: leaves do vary in size but by a relatively small factor, despite huge factors in the variation of their heights and masses. A tree that is just twenty times taller than another does not have leaves whose diameter is twenty times larger. Consequently, the variation among terminal units within a given design is a relatively small secondary effect. The same goes for possible variations in the other postulates: networks may not be precisely space filling or precisely optimized. Corrections due to such deviations and variations are considered to be "higher order" effects in the sense we discussed earlier. 144

September 20, 2023

These postulates underlie the zeroth order, coarse-grained 145

theory for the structure, organization, and dynamics of biological networks, and allow us to calculate many of the essential features of what I referred to as the average idealized organism of a given size. In order to carry out this strategy and calculate quantities such as metabolic rates, growth rates, the heights of trees, or the number of mitochondria in a cell, these postulates have to be translated into mathematics. The goal is to determine the consequences, ramifications, and predictions of the theory and confront them with data and observations. The details of the mathematics depend on the specific kind of network being considered. As discussed earlier, our circulatory system is a network of pipes driven by a beating heart, whereas plants and trees are networks of bundles of thin fibers driven by a steady nonpulsatile hydrostatic pressure. Fundamental to the conceptual framework of the theory is that, despite these completely different physical designs, both kinds of networks are constrained by the same three postulates: they are space filling, have invariant terminal units, and minimize the energy needed to pump fluid through the system .

September 20, 2023

Carrying out this strategy proved to be quite a challenge, both conceptually and technically. It took almost a year to iron out all of the details, but ultimately we showed how Kleiber's law for metabolic rates and, indeed, quarter-power scaling in general arises from the dynamics and geometry of optimized space-filling branching networks. Perhaps most satisfying was to show how the magic number four arises and where it comes from. 12 145

September 20, 2023

Similarly, the rate at which oxygen is inhaled through our mouths and into the respiratory system is also a measure of metabolic rate. These two systems are tightly coupled together so blood flow rates, respiratory rates, and metabolic rates are all proportional to one another and related by simple linear relationships. Thus, hearts beat approximately 146 *four times for each breath that is inhaled, regardless of the size of the mammal. This tight coupling of the oxygen delivery systems is why the properties of the cardiovascular and respiratory networks play such an important role in determining and constraining metabolic rate.*

September 20, 2023

September 20, 2023 147

147

The rate at which you use energy to pump blood through the vasculature of your circulatory system is called your cardiac power output. This energy expenditure is used to overcome the viscous drag, or friction, on blood as it flows through increasingly narrower and narrower vessels in its journey through the aorta, which is the first artery leaving your heart, down through multiple levels of the network to the tiny capillaries that feed cells. A human aorta is an approximately cylindrical pipe that is about 18 inches long (about 45 cm) and about an inch (about 2.5 cm) in diameter, whereas our capillaries are only about 5 micrometers wide (about a hundredth of an inch), which is somewhat smaller than a hairbreadth. 13 Although a blue whale's aorta is almost a foot in diameter (30 cm), its capillaries are still pretty much the same size as yours and mine. This is an explicit example of the invariance of the terminal units in these networks.

September 20, 2023

149

To avoid this potential problem and minimize the work our hearts have to do, the geometry of our circulatory systems has evolved so that there are no reflections at any branch point throughout the network. The mathematics and physics of how this is accomplished is a little *bit complicated, but the result is simple and elegant: the theory predicts that there will be no reflections at any branch point if the sum of the cross-sectional areas of the daughter tubes leaving the branch point is the same as the cross-sectional area of the parent tube coming into it . As an example, consider the simple case where the two daughter tubes are identical and therefore have the same cross-sectional areas (which is approximately correct in real circulatory systems). Suppose that the cross-sectional area of the parent tube is 2 square inches: then, in order to ensure that there are no reflections, the cross-sectional area of each daughter has to be 1 square inch. Because the cross-sectional area of any vessel is proportional to the square of its radius, another way of expressing this result is to say that the square of the radius of the parent tube has to be just twice the square of the radius of each of the daughters. So to ensure that there is no energy loss via reflections as one progresses down the network, the radii of successive vessels must scale in a regular self-similar fashion, decreasing by a constant factor of the square root of two (√2) with each successive branching . This so-called area-preserving branching is, indeed, how our circulatory system is constructed, as has been confirmed by detailed measurements across many mammals—and many plants and trees.*

September 21, 2023

However, if you think of a tree as a bundle of fibers all tightly tied together beginning in the trunk and then sequentially spraying out up through its branches, then it's clear that the [149](#)

cross-sectional area must be preserved all the way up through the hierarchy. This is illustrated below, where this fiber bundle structure is compared with the pipe structure of mammals. An interesting consequence of area-preserving branching is that the cross-sectional area of the trunk is the same as the sum of the cross-sectional areas of all the tiny branches at the end of the network (the petioles). Amazingly, this was known to Leonardo da Vinci. I have reproduced the requisite page of his notebook where he demonstrates this fact. Although this simple geometric picture demonstrates why trees obey area-preserving branching, it is in actuality an oversimplification. However, area preserving can be derived from a much more realistic model for trees using the general network principles of space filling and optimization enunciated above, supplemented with biomechanical constraints that require branches to be resilient against perturbations from wind by bending without breaking. Such an analysis shows that in almost all respects plants and trees scale just like mammals, both within individuals as well as across species, including the ¾ power law for their metabolic rates, even though their physical design is quite different. 14

September 21, 2023

September 21, 2023 151

151

It's a lovely thought that the optimum design of our circulatory system obeys the same simple area-preserving branching rules that trees and plants do. It's equally satisfying that the condition of nonreflectivity of waves at branch points in pulsatile networks is essentially identical to how national power grids are designed for the efficient transmission of electricity over long distances. This condition of nonreflectivity is called impedance matching. It has multiple applications not only in the working of your body but across a very broad spectrum of technologies that play an important part in your daily life. For example, telephone network systems use matched impedances to minimize echoes on long-distance lines; most loudspeaker systems and *musical instruments contain impedance matching mechanisms; and the bones in the middle ear provide impedance matching between the eardrum and the inner ear. If you have ever witnessed or been subject to an ultrasound examination you will be familiar with the nurse or technician smearing a gooey gel over your skin before sliding the probe over it. You probably thought that this was for lubrication purposes but in fact it's actually for matching impedances. Without the gel, the impedance mismatch in ultrasound detection would result in almost all of the energy being reflected back from the skin, leaving very little to go into the body to be reflected back from the organ or fetus under investigation. The term impedance matching can be a very useful metaphor for connoting important aspects of social interactions.*

September 21, 2023

152

For example, the smooth and efficient functioning of social networks, whether in a society, a company, a group activity, and especially in relationships such as marriages and friendships, requires good communication in which information is faithfully transmitted between groups and individuals. When information is dissipated or "reflected," such as when one side is not listening, it cannot be faithfully or efficiently processed, inevitably leading to misinterpretation, a process analogous to the loss of energy when impedances are not matched.

September 21, 2023

154

So if for some perverse reason you wanted to know the radius, length, blood flow rate, pulse rate, velocity, pressure, et cetera, in the fourteenth branch of the circulatory system of the average hippopotamus, the theory will provide you with the answer. In fact, the theory will tell you the answer for any of these quantities for any branch of the network in any animal. As blood flows through smaller and smaller vessels on its way down through the network, viscous drag forces become increasingly important, leading to the

dissipation of more and more energy. The effect of this energy loss is to progressively dampen the wave on its way down through the network hierarchy until it eventually loses its pulsatile character and turns into a steady flow. In other words, the nature of the flow makes a transition from being pulsatile in the larger vessels to being steady in the smaller ones. That's why you feel a pulse only in your main arteries—there's almost no vestige of it in your smaller vessels. In the language of electrical transmission, the nature of the blood flow changes from being AC to DC as it progresses down through the network. Thus, by the time blood reaches the capillaries its viscosity ensures that it is no longer pulsatile and that it is moving extremely slowly. It slows down to a speed of only about 1 millimeter per second, which is tiny compared with its speed of 40 centimeters per second when it leaves the heart. This is extremely important because this leisurely speed ensures that oxygen carried by the blood has sufficient time to diffuse efficiently across the walls of the capillaries and be rapidly delivered to feed cells.
Interestingly, the theory predicts that these velocities at the two extremities of the network, the capillaries and the aorta, are the same for all mammals, as observed. You are very likely aware of this huge difference in speeds between capillaries and the aorta. If you prick your skin, blood oozes out very slowly from the capillaries with scant resulting damage, whereas if you cut a major artery such as your aorta, carotid, or femoral, blood gushes out and you can die in just a matter of minutes.

September 21, 2023

But what's really surprising is that blood pressures are also predicted to be the same across all mammals, regardless of their size. Thus, despite the shrew's heart weighing only about 12 milligrams, the equivalent of about 25 grains of salt, and its aorta having a radius of only about 0.1 millimeter and consequently barely visible, whereas a whale's heart weighs about a ton, almost the weight of a Mini Cooper, and its aorta has a radius of about 30 centimeters, their blood pressures are approximately the same. This is pretty amazing—just think of the enormous stresses on the walls of the shrew's tiny aorta and arteries compared with the 155 *pressures on yours or mine, let alone those on a whale's. No wonder the poor creature dies after only a year or two.*

September 21, 2023

In addition to his work on the cardiovascular system, Young was famous for several other quite diverse and profound discoveries. He is perhaps best known for establishing the wave theory of light, in which each color is associated with a particular wavelength. But he also contributed to early work in linguistics and Egyptian hieroglyphics, including being the first person to decipher the famous Rosetta Stone now sitting in the British Museum in London. As a fitting tribute to this remarkable man, Andrew Robinson wrote a spirited biography of Young titled The Last Man Who Knew Everything: Thomas Young, the Anonymous Polymath Who Proved Newton Wrong, Explained How We See, Cured the Sick, and Deciphered the Rosetta Stone, Among Other Feats of Genius. I have a certain soft spot for Young because he was born in Milverton, in the county of Somerset in the West of England, just a few short miles from Taunton, where I was born. 156

September 21, 2023

This is in marked contrast to what we normally see when, for example, we use a microscope to zoom in on an object using a higher and higher resolution in order to reveal greater detail and new structure that is qualitatively different from those of the whole. Obvious examples are cells in tissue, molecules in materials, or protons in atoms. If, on the other hand, the object is a fractal, no new pattern or detail arises when the resolution is increased: the same pattern repeats itself over and over again. In reality, this is an idealized description for, of course, the images at various levels of resolution differ very slightly from one another and eventually the recursive repetition ceases and new patterns of structural design appear. If you continue breaking down broccoli to increasingly smaller pieces, these eventually lose the geometric characteristics of broccoli and eventually reveal 157 *the structure of its tissue, its cells, and its molecules.*

September 21, 2023

September 21, 2023 159

This repetitive phenomenon is called self-similarity and is a generic characteristic of fractals. Analogous to the repetitive scaling exhibited by broccoli are the infinite reflections in parallel mirrors, or the nesting of Russian dolls (matryoshka) of regularly decreasing sizes inside one another. Long before the concept was invented, self-similarity was poetically expressed by the Irish satirist Jonathan Swift, the author of Gulliver's Travels, in this whimsical quatrain: So, naturalists observe, a flea Hath smaller fleas that on him prey; And these have smaller still to bite 'em; And so proceed ad infinitum. 159

September 21, 2023

So it is with the hierarchical networks we have been discussing. If you cut a piece out of such a network and appropriately scale it up, then the resulting network looks just like the original. Locally, each level of the network essentially replicates a scaled version of the levels adjacent to it. We saw an explicit example of this when discussing the consequences of impedance matching in the pulsatile regime of the circulatory system where area−preserving branching resulted in the radii of successive vessels decreasing by a constant factor (√2 = 1.41 . . .) with each successive branching. So, for example, if we compare the radii of vessels separated by 10 such branchings, then they are 159 *related by a scale factor of (√2)10 = 32. Because our aorta has a radius of about 1.5 centimeters, this means that the radii of vessels at the tenth branching level are only about half a millimeter.*

September 21, 2023

Because blood flow changes from pulsatile to nonpulsatile as one progresses down the network, our circulatory system is actually not continuously self-similar nor therefore a precise fractal. In the nonpulsatile domain where the flow is dominated by viscous forces, minimizing the amount of power being dissipated leads to a self-similarity in which the radii of successive vessels decrease by a constant factor of the cube root of two 3√2 (= 1.26 . . .), rather than the square root √2 (= 1.41 . . .) as in the pulsatile region. Thus the fractal nature of the circulatory system subtly changes from the aorta to the capillaries, reflecting the change in the nature of the flow from pulsatile to nonpulsatile. Trees, on the other hand, maintain approximately the same self-similarity from the trunk to their leaves, with radii successively decreasing by the area-preserving ratio of √2. 159

September 21, 2023

The space-filling requirement that the network must service the entire volume of the organism at all scales also requires it to be self-similar in terms of the lengths of the vessels. To fill the three-dimensional space, lengths of successive vessels have to decrease by a constant factor of 3√2 with each successive branching and, in 160 *contrast to radii, this remains valid down through the entire network, including both the pulsatile and nonpulsatile domains. Having determined how networks scale within individuals following these simple rules, the last piece of the derivation is to determine how this connects across species of different weights. This is accomplished from a further consequence of the energy minimization principle: namely, that the total volume of the network—that is, the total volume of blood in the body—must be directly proportional to the volume of the body itself, and therefore proportional to its weight, as observed. In other words, the volume of blood is a constant proportion of the volume of the body, regardless of size. For a tree this is obvious because the network of its vessels constitutes the entire tree—there is no analog of flesh in between all of its branches, so the volume of the network is the volume of the tree. 16 Now, the volume of the network is just the sum of the volumes of all of its vessels or branches, and these can be straightforwardly calculated from knowing how their lengths and radii scale, thereby connecting the self-similarity of the internal network to body size. It is the mathematical interplay between the cube root scaling law for lengths and the square root scaling law for radii, constrained by the linear scaling of blood volume and the invariance of the terminal units, that leads to quarter-power allometric exponents across organisms. The resulting magic number four emerges as an effective extension of the usual three dimensions of the volume serviced by the network by an additional dimension resulting from the fractal nature of the network. I shall go into this in more detail in* *the following chapter, where I discuss the general concept of fractal dimension , but suffice it to say here that natural selection has taken advantage of the mathematical marvels of fractal networks to optimize their distribution of energy so that organisms operate as if they were in four dimensions , rather than the canonical three. In this sense the ubiquitous number four is actually 3 + 1. More generally, it is the dimension of the space being serviced plus one. So had we lived in a universe of eleven dimensions, as some of my string theory friends believe, the magic number would have been 11 + 1 = 12, and we would have been talking about the universality of 1/12 power scaling laws rather than $\frac{1}{4}$ power ones.*

September 21, 2023

September 21, 2023 162

Stimulated by his passionate pacifism, Richardson embarked on an ambitious program to develop a quantitative theory for understanding the origins of war and international conflict in order to devise a strategy for their ultimate prevention. His aim was nothing less than to develop a science of war. His main thesis was that the dynamics of conflict are primarily governed by the rates at which nations build up their armaments and that their continued accumulation is the major cause of war. He viewed the accumulation of weapons as a proxy for the collective psychosocial forces that reflect, but transcend, history, politics, economics, and culture and whose dynamics inevitably lead to conflict and instability. Richardson used the mathematics developed for understanding chemical reaction dynamics and the spread of communicable diseases to model the ever-increasing escalation of arms races in 164 *which the arsenal of each country increases in response to the increase in armaments of every other country.*

September 21, 2023

His theory did not attempt to explain the fundamental origins of war, that is, why we collectively resort to force and violence to settle our conflicts, but rather to show how the dynamics of arms races escalate, resulting in catastrophic conflict. Although his theory is highly oversimplified, Richardson had some success in comparing his analyses with data, but more important, he provided an alternative framework for quantitatively understanding the origins of war that could be confronted with data. Furthermore, it had the virtue of showing what parameters were important, especially in providing scenarios under which a potentially peaceful situation could be achieved and sustained. In contrast to conventional, more qualitative theories of conflict, the roles of leadership, cultural and historical animosity, and specific events and personalities play no explicit role in his theory. 18

165

September 21, 2023

In his desire to create a testable scientific framework, Richardson collected an enormous amount of historical data on wars and conflicts. In order to quantify them he introduced a general concept, which he called the deadly quarrel , defined as any violent conflict between human beings resulting in death. War is then viewed as a particular case of a deadly quarrel, but so is an individual murder. He quantified their magnitudes by the subsequent number of deaths: for an individual murder the size of the deadly quarrel is therefore just one, whereas for the Second World War it is more than 50 million, the exact number depending on how civilian casualties are counted. He then took the bold leap of asking whether there was a continuum of deadly quarrels beginning with the individual and progressing up through gang violence, civil unrest, small conflicts, and ending up with the two major world wars, thereby covering

165

a range of almost eight orders of magnitude. Trying to plot these on a single axis leads to the same challenge we faced earlier when trying to accommodate all earthquakes or all mammalian metabolic rates on a simple linear scale. Practically, it simply isn't possible, and one has to resort to using a logarithmic scale to see the entire spectrum of deadly quarrels. Thus, by analogy with the Richter scale, the Richardson scale begins with zero for a single individual murder and ends with a magnitude of almost eight for the two world wars (eight orders of magnitude would represent a hundred million deaths). In between, a small riot with ten victims would have magnitude one, a skirmish in which one hundred combatants were killed would be two, and so on. Obviously there are very few wars of magnitude seven but an enormous number of conflicts with magnitude zero or one. When he plotted the number of deadly quarrels of a given size versus their magnitude on a logarithmic scale, he found an approximately straight line just like the straight lines we saw when physiological quantities like metabolic rate were plotted in this way versus animal size (see Figure 1). Consequently, the frequency distribution of wars follows simple power law scaling indicating that conflicts are approximately self-similar. 19 This remarkable result leads to the surprising conclusion that, in a coarse-grained sense, a large war is just a scaled-up version of a small conflict, analogous to the way that elephants are approximately scaled-up mice. Thus underlying the extraordinary complexity of wars and conflicts seems to be a common dynamic operating across all scales. Recent work has confirmed such findings for recent wars, terrorist attacks, and even cyberattacks. 20 No general theory has yet been advanced for understanding these regularities, though they very likely reflect the fractal-like network characteristics of national economies, social behavior, and competitive forces. In any case, any ultimate theory of war needs to account for them.

September 21, 2023

coastline using

172

September 21, 2023

September 21, 2023 172

Well, Benoit Mandelbrot did. He deserves great credit not only for resurrecting Richardson's work but for recognizing its deeper significance. In 1967 he published a paper in the high-profile journal Science with the more transparent title "How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension." 22 This brought Richardson's work to light by expanding on his findings and generalizing the idea. Crinkliness, later to become known as fractality, is quantified by how steep the slopes of the corresponding straight lines are on Richardson's logarithmic plots: the steeper the slope, the more crinkly the curve. These slopes are just the exponents of the power laws relating length to resolution and are the analog of the ¾ exponent relating metabolic rate to mass for organisms. For very smooth traditional curves, like circles, the slope or exponent is zero because its length does not change with increasing resolution but converges to a definite value, as in the living room example. However, for rugged, crinkly coastlines the slope is nonzero. For example, for the west coast of Britain, it's 0.25. 172

September 21, 2023

The point of adding the 1 was to connect the idea of fractals to the conventional concept of ordinary dimensions discussed in chapter 2. Recall that a smooth line has dimension 1, a smooth surface dimension 2, and a volume dimension 3. Thus the South African coast is very close to being a smooth line because its fractal dimension is 1.02, which is very close to 1, whereas Norway is far from it because its fractal dimension of 1.52 is so much greater than 1. You could imagine an extreme case of this in which the line is so crinkly and convoluted that it effectively fills an entire area. Consequently, even though it's still a line with "ordinary" dimensions 1, it behaves as if it were an area in terms of its 173 *scaling properties, therefore having a fractal dimension of 2. This curious gain of an effective additional dimension is a general feature of space-filling curves, to which I will return in the next chapter.*

September 21, 2023

In the natural world almost nothing is smooth—most things are crinkly, irregular, and crenulated, very often in a self-similar way. Just think of forests, mountain ranges, vegetables, clouds, and the surfaces of oceans. Consequently, most physical objects have no absolute objective length, and it is crucial to quote the resolution when stating the measurement. So why did it take more than two thousand years for people to recognize something so basic and which now seems almost obvious? Very likely this has its origins in the duality that emerged as we gradually separated from a close connection to the natural world and became more and more distant from the forces of nature that have determined our biology. Once we invented language, learned how to take advantage of economies of scale, formed communities, and began making artifacts, we effectively changed the geometry of our daily world and its immediate surroundings. In designing and manufacturing human-engineered artifacts, whether primitive pots and tools or modern sophisticated automobiles, computers, and skyscrapers, we employed and aspired to the simplicity of straight lines, smooth curves, and smooth surfaces. This was brilliantly formalized and reflected in the development of quantified measurement and the invention of mathematics, manifested, in particular, in the idealized paradigm of Euclidian geometry. This is the mathematics appropriate to the world of artifacts we created around us as we evolved from being a mammal like any other to become social Homo sapiens. 174

September 21, 2023

Perhaps of greater importance is that he realized that these ideas are generalizable far beyond considerations of borders and coastlines to almost anything that can be measured, 176

even including times and frequencies. Examples include our brains, balls of crumpled paper, lightning, river networks, and time series like electrocardiograms (EKGs) and the stock market. For instance, it turns out that the pattern of fluctuations in financial markets during an hour of trading is, on average, the same as that for a day, a month, a year, or a decade. They are simply nonlinearly scaled versions of one another. Thus if you are shown a typical plot of the Dow Jones average over some period of time, you can't tell if it's for the last hour or for the last five years—the distributions of dips, bumps, and spikes is pretty much the same, regardless of the time frame. In other words, the behavior of the stock market is a self-similar fractal pattern that repeats itself across all timescales following a power law that can be quantified by its exponent or, equivalently, its fractal dimension. You might think that with this knowledge you might soon become rich. Although this certainly gives new insight into hidden regularities in stock markets, unfortunately it has predictive power only in an average coarse-grained sense and does not give specific information about the behavior of individual stocks. Nevertheless, it's an important ingredient for understanding the dynamics of the market over different timescales. This has stimulated the development of a new transdisciplinary subfield of finance called econophysics and motivated investment companies to hire physicists, mathematicians, and computer scientists to use these sorts of ideas to develop novel investment strategies. 23 Many have done very well, though it is unclear just how big a role their physics and mathematics actually played in their success.

September 21, 2023

September 21, 2023 177

In fact, those that are most seriously at risk have fractal dimensions close to one with an uncharacteristically smooth EKG. Thus the fractal dimension of the EKG provides a potentially powerful complementary diagnostic tool for 177 *quantifying heart disease and health. 24 The reason that being healthy and robust equates with greater variance and larger fluctuations, and therefore a larger fractal dimension as in an EKG, is closely related to the resilience of such systems. Being overly rigid and constrained means that there isn't sufficient flexibility for the necessary adjustments needed to withstand the inevitable small shocks and perturbations to which any system is subjected. Think of the stresses and strains your heart is exposed to every day, many of which are unexpected. Being able to accommodate and naturally adapt to these is critical for your long-term survival. These continuous changes and impingements require all of your organs, including your brain as well as its psyche, to be both flexible and elastic and therefore to have a significant fractal dimension.*

September 21, 2023

This can be extended, at least metaphorically, beyond individuals to companies, cities, states, and even life itself. Being diverse and having many interchangeable, adaptable components is another manifestation of this paradigm. Natural selection thrives on and consequently manufactures greater diversity. Resilient ecosystems have greater diversity of species. It is no accident that successful cities are those that offer a greater spectrum of job opportunities and businesses, and that successful companies have a diversity of products and people with the flexibility to change, adapt, and reinvent in response to changing markets. I shall discuss this further in chapters 8 and 9 when I turn to cities and companies. 177

September 21, 2023

September 22, 2023 179

September 22, 2023 181

4: THE FOURTH DIMENSION OF LIFE: Growth, Aging, and Death

Almost all of the networks that sustain life are approximately self-similar fractals. In the previous chapter, I explained how the nature and origin of these fractal structures are a consequence of generic geometric, mathematical, and physical principles such as optimization and space filling, thereby leading to the derivation for how networks scale both within an average individual as well as across species.

181

September 22, 2023

This is all very satisfying, but a lingering question remains: why does the same ¼ power exponent emerge from each of these different networks rather than, say, a 1/6 power emerging in one, a ⅛ power in another, and so on? In other words, what necessitates that this same set of principles should lead to the same scaling exponents when applied to different network systems with a variety of structures and dynamics? Are there additional design principles that transcend the dynamics that ensure that the ¼ emerges in virtually all organismic groups? This is an important conceptual question, especially for understanding why this universal behavior extends even to systems such as bacteria, where an explicit hierarchical branching network structure is much less obvious.

185

September 22, 2023

A general argument to address this can be made by recognizing that, in addition to minimizing energy loss, natural selection has also led to a maximization of metabolic

185

capacity because metabolism produces the energy and materials required to sustain and reproduce life. 1 This has been achieved by maximizing surface areas across which resources and energy are transported. These surfaces are in actuality the total surface areas of all the terminal units of the network. For instance, all of our metabolic energy is transmitted across the total surface area of all of our capillaries to fuel our cells, just as the metabolism of a tree is governed by the transmission of energy gathered from sunlight through all of its leaves to fuel photosynthesis and of water from soil through all of the terminal fibers of its root system.

September 22, 2023

186

Natural selection has taken advantage of the fractal nature of space-filling networks to maximize the total effective surface area of these terminal units and thereby maximize metabolic output. Geometrically, the nested levels of continuous branching and crenulations inherent in fractal-like structures optimize the transport of information, energy, and resources by maximizing the surface areas across which these essential features of life flow. Because of their fractal nature, these effective surface areas are very much larger than their apparent physical size. Let me give you some remarkable examples from your own body to illustrate the point. Even though your lungs are only about the size of a football with a volume of about 5 to 6 liters (about one and a half gallons), the total surface area of the alveoli, which are the terminal units of the respiratory system where oxygen and carbon dioxide are exchanged with the blood, is almost the size of a tennis court and the total length of all the airways is about 2,500 kilometers, almost the distance from Los Angeles to Chicago, or London to Moscow. Even more striking is that if all the arteries, veins, and capillaries of your circulatory system were laid end to end, their total length would be about 100,000 kilometers, or nearly two and a half times around the Earth or over a third of the distance to the moon . . . and all of this neatly fits inside your five-to-six-foot-tall body. It's quite fantastic and yet another amazing feature of your body where natural selection has exploited the wonders of physics, chemistry, and *mathematics.*

September 22, 2023

As explained in the previous chapter, a crinkly enough line that is space filling can scale as if it's an area. Its fractality effectively endows it with an additional dimension. Its conventional Euclidean dimension, discussed in chapter 2, still has the value 1, indicating that it's a line, but its fractal dimension is 2, indicating that it's maximally fractal and scaling as if it were an area. In a similar fashion an area, if crinkly enough, can behave as if it's a volume, thereby gaining an effective extra dimension: its Euclidean dimension is 2, indicating that it's an area, but its fractal dimension is 3. A familiar example will make this clear. Think of washing sheets. Being sensitive to conserving energy and at the same time wanting to save yourself money and time, you wait several weeks until you have more than a sufficient number of dirty ones to fill the entire tub of your washing machine. So when the time comes you stuff in as much and as many as you possibly can to fill the entire volume of the tub. Now, recall that ordinary volumes scale faster than areas, so if you were to double the size of your washing machine by doubling all of its lengths while keeping its shape the same, its volume would increase by a factor of eight (23) whereas all of its surface areas would increase by a factor of four (22). Naively, you might therefore conclude that because sheets are essentially all area and consequently two-dimensional (their thickness being negligible), you could accommodate four times as many sheets by doubling the size of your washing machine. However, if we stuff all of the sheets into the tub so that they completely fill its entire volume and because this volume has increased by a factor of eight, then it's clear that you can actually accommodate eight times as many sheets, rather than just four times. In other words, the total effective area of two-dimensional sheets filling three-dimensional washing machines scales like a volume rather than an area, so in this sense, we have turned an area into a volume. 187

September 22, 2023

However, driven by the forces of natural selection to maximize exchange surfaces, biological networks do achieve maximal space filling and consequently scale like three-dimensional volumes rather than two-dimensional Euclidean surfaces. This additional dimension, which arises from optimizing network performance, leads to organisms' functioning as if they are operating in four dimensions.

188

September 22, 2023

Thus, instead of scaling with classic ⅓ exponents, as would be the case if they were smooth nonfractal Euclidean objects, they scale with ¼ exponents. Although living things occupy a three-dimensional space, their internal physiology and anatomy operate as if they were four-dimensional.

188

September 22, 2023

Quarter-power scaling laws are perhaps as universal and as uniquely biological as the biochemical pathways of metabolism, the structure and function of the genetic code, and the process of natural selection. The vast majority of organisms exhibit scaling exponents very close to ¾ for metabolic rate and ¼ for internal times and distances. These are the maximal and minimal values, respectively, for the effective surface area and linear dimensions of a volume-filling fractal-like network.

189

September 22, 2023

The answer lies in the subtleties of networks and their interplay with physiological limits in the spirit of Galileo's original argument that there are limits to the maximum size of structures. Unlike most biological networks, mammalian circulatory systems are not single self-similar fractals but an admixture of two different ones, reflecting the change in flow

190

from predominantly pulsatile AC to predominantly nonpulsatile DC as blood flows from the aorta to the capillaries. Most of the blood resides in the larger vessels of the upper part of the network where AC dominates, leading to the ¾ power scaling law for metabolic rates. Although the branching changes continuously from one mode to the other, the region over which it changes is relatively narrow and its location (as measured by the number of branchings up from the capillaries) is independent of body size and therefore the same for all mammals. In other words, all mammals have roughly the same number of branching levels, about fifteen, where the flow is predominantly steady nonpulsatile DC. The distinction among mammals as their size increases is the increasing number of levels where the flow is pulsatile AC. For example, we have about seven to eight, the whale has about sixteen to seventeen, and the shrew just one or two. Impedance matching in these vessels ensures that relatively little energy is required to pump blood through them, so the more of them you have the better. Almost all of your cardiac output goes into pumping blood through the much smaller vessels of the nonpulsatile regime, whose number of levels is approximately the same for all mammals. Relatively speaking, then, the proportion of the network where the heart expends most of its energy systematically decreases as the size of a mammal increases, illustrating again that larger mammals are more efficient than smaller ones: the whale needs only one hundredth the amount of energy needed by a shrew to supply blood to one of its cells.

September 22, 2023

Now imagine continuously decreasing the size of the animal. Concomitantly, the number of area-preserving branchings where vessels are large enough to support pulsatile waves decreases until a tipping point is reached where the network can support only nonpulsatile DC flow. At that stage even the major arteries become so small and constricted that they are unable to support pulsatile waves. In such vessels, waves become so overdamped due to the viscosity of blood that they can no longer propagate and the flow shifts to becoming entirely steady DC, just like the flow of water in the pipes of your house: pulsatile waves generated by the beating heart 192 *are immediately damped as they enter the aorta.*

September 22, 2023

This is really weird. Such an animal would have a beating heart but no pulse! This is not only weird, but more important, it would be an extremely inefficient design because it would have entirely lost the advantages of impedance matching and would lead to significant energy being dissipated in all of the vessels throughout its circulatory system. This loss of performance efficiency is reflected in how metabolic rates would scale. Instead of following the classic sublinear ¾ power scaling law, a calculation shows that it would now scale linearly—that is, directly proportional to body mass—and thereby lose the advantages of economies of scale. In this purely DC case the power needed to support a gram of tissue would now be the same regardless of size instead of systematically decreasing with size following quarter-power scaling. Consequently, no evolutionary advantage would be conferred by increasing size. This argument shows that only mammals that are large enough for their circulatory systems to support pulsatile waves through at the very least the first couple of branching levels would have evolved, thereby providing a fundamental reason why there is a minimum size. 3 The theory can be used to derive a formula for when this tipping point occurs. Its actual value depends upon generic quantities such as the density and viscosity of blood and the elasticity of arterial walls. The calculation leads to an approximate value for the size of the smallest mammal of just a couple of grams, comparable to the mass of the Etruscan shrew, which is the smallest known mammal. It is only about 4 centimeters long, easily sitting on the palm of your hand. Its tiny heart beats at more than a thousand times a minute—about twenty times a second—as it pumps blood with the same pressure and speed as you do , and even more astounding, as does a blue whale. And all of this through its minuscule aorta, which is only a couple of millimeters long and an astonishingly couple of tenths of a millimeter wide, not much thicker than a hairbreadth. As I have remarked earlier, no wonder the poor creature doesn't live very long.

192

September 22, 2023

September 23, 2023 195

September 22, 2023 195

September 23, 2023 197

September 23, 2023 199

September 23, 2023 200

This difference in the way these two contributions scale with increasing size plays a central role in growth, so to ensure that you understand what it implies here's a simple example to illustrate the point. Suppose the size of the organism doubles; then the number of cells also doubles, so the amount of energy needed for their maintenance increases by a factor of 2. However, metabolic rate (the supply of energy) increases by only a factor of 2¾ = 1.682 . . . which is less than 2. So the rate at which energy is needed for maintenance increases faster than the rate at which metabolic energy can be supplied, forcing the amount of energy available for growth to systematically decrease and eventually go to zero, resulting in the cessation of growth. In other words, you stop growing because of the mismatch between the way maintenance and supply scale as size increases. 203

September 22, 2023

Let me deconstruct this a little further to give a more mechanistic understanding of what's going on. Recall that the reason metabolic rate scales with a sublinear ¾ power exponent lies in the hegemony of the network. Furthermore, because the entire flow through the network ends up going through all of the capillaries, and because they are invariant across species as well as during ontogeny (capillaries are approximately the same for mice, elephants, their babies and children, as well as for us), their number likewise scales with a ¾ power exponent. So as the organism grows and size increases, each capillary systematically has to service more cells following ¼ power scaling. It is this mismatch at the critical interface between capillaries and cells that controls growth and ultimately leads to its cessation: the increase in the number of supply units (the capillaries) cannot keep up with the demands from the increase in the number of customers (the cells). 203

September 22, 2023

And the theory tells us why: growth is primarily determined by how energy is delivered to cells, and this is constrained by universal properties of networks that transcend design. Among the many other aspects of growth that can be derived from the theory, it predicts how the allocation of metabolic energy between maintenance and growth changes with age. At birth almost all of it is devoted to growth and relatively little to maintenance, whereas beyond maturity all of it is devoted to maintenance, repair, and replacement. 205

September 22, 2023

The theory has been extended to understand the growth of tumors, plants, insects, and both forest 5 and social insect communities 6 such as ants and bees. These latter applications are forerunners of how we might start thinking 205 *about the growth of human organizations such as cities and companies, which I'll be turning to in chapters 8 and 9. Each of these very different systems represents a variation on the general thematic structure of the growth equation. For instance, tumors are parasitic and use metabolic energy derived from their host to grow, so their vasculature and metabolic rates depend not only on their own size but also on the size of the host. 7 Understanding this provides insight into how to scale up basic properties of tumors as well as potential therapeutic strategies from observations on mice to humans. 8*

September 22, 2023

September 22, 2023 209

The overall agreement with theory is extremely satisfying. But much more than that, I find the extraordinary unity and interconnectivity of life that is revealed through this lens to be spiritually elevating in the pantheistic spirit articulated by the philosopher Baruch Spinoza. As Einstein wrote, 10 "We followers of Spinoza see our God in the wonderful order and lawfulness of all that exists and in its soul as it reveals itself in man and animal."

209

September 22, 2023

Because the theory is deliberately simplified, it is inevitable that measurements of real organisms will, to varying degrees, deviate from model predictions. As can be seen in Figure 19 the agreement is surprisingly good with relatively few major outliers that deviate significantly from the idealized growth curve. We as primates are one of those. For instance, we take longer to mature than we "should" given our body weight. This is the result of our rapid evolution from being purely biological to becoming sophisticated socioeconomic creatures. Our effective metabolic rate is now

209

one hundred times greater than what it was when we were truly "biological" animals, and this has had huge consequences for our recent life history. We take longer to mature, we have fewer offspring, and we live longer, all in qualitative agreement with having an effectively larger metabolic rate arising from socioeconomic activity. I shall return to this fascinating development in our history when discussing how these ideas apply to cities.

September 22, 2023

September 22, 2023 210

What is shocking is how few people, even many scientists, appreciate that this sensitivity to temperature is exponential . The reason for this sensitivity is that all chemical reaction rates depend exponentially on temperature. In the previous chapter I showed how metabolism originates in the production of ATP molecules in cells. Consequently, metabolic rate scales exponentially with temperature rather than as a power law as it does with mass. Because metabolic rate—the rate at which energy is supplied to cells—is the fundamental driver of all biological rates and times, all of the central features of life from gestation and growth to mortality are exponentially sensitive to temperature . 210

September 23, 2023

This leads to the fascinating conclusion that across the spectrum of life all biological rates and times such as those associated with growth, embryonic development, longevity, and evolutionary processes are determined by a joint universal scaling law in terms of just two parameters: the number ¼, arising from the network constraints that control the dependence on mass, and 0.65 eV, originating in the chemical reaction dynamics of ATP production. This result can be restated in a slightly different way: when adjusted for 211 *size and temperature, as determined by just these two numbers, all organisms run to a good approximation by the same universal clock with similar metabolic, growth, and evolutionary rates.*

September 23, 2023

This parsimonious formulation of the coarse-grained mass and temperature dependence was introduced as a compact summary of the scaling work in a paper titled "Toward a Metabolic Theory of Ecology" published in the journal Ecology in 2004 and coauthored by Jim Brown and three of our then postdocs—Van Savage, Jamie Gillooly, and Drew Allen—together with me. 211

September 23, 2023

I want to emphasize just how remarkable this is. The two most important events in an organism's life, its birth and its death, which are usually thought of as being independent, are intimately related to each other: the slopes of these two graphs are determined by exactly the same parameter, the 0.65 eV, representing the average energy needed to produce an ATP molecule. Below I will explore this further when I discuss how a more fundamental theory of aging based on network dynamics explains the mechanistic origins of this temperature dependence. 212

September 23, 2023

Thus at a deep level, birth, growth, and death are all governed by the same underlying dynamics driven by metabolic rate and encapsulated in the dynamics and structure of networks. 213

September 23, 2023

More pertinent, a modest 2°C change in ambient temperature leads to a 20 percent to 30 percent change in growth and mortality rates. 11 This is huge and therein lies our problem. If global warming induces a temperature increase of around 2°C, which it is on track to do, then the pace of almost all biological life across all scales will increase by a whopping 20 percent to 30 percent. This is highly nontrivial and will potentially wreak havoc with the ecosystem. It's analogous to the huge leap that Brunel attempted to make when building his mammoth ship the Great Eastern, which ended in disaster primarily because the science of shipbuilding had not yet been sufficiently well developed. Ships are extremely simple compared with the profound complexity of ecosystems and societies.

213

September 23, 2023

One last note: the underlying physics and chemistry of reaction theory have been known for a very long time, having been developed by the Swedish physicist-turned-chemist Svante August Arrhenius, who won the Nobel Prize in Chemistry in 1903. He holds the distinction of being the first Swede to become a Nobel laureate. Arrhenius was a man of very broad interests whose many novel ideas and contributions to science have been very influential. He was one of the first people to seriously suggest that life on Earth might have originated from the transport of spores having been transported from another planet, a rather speculative theory with a surprisingly large following that now goes by the name of panspermia . Of greater significance is that he was the first scientist to calculate how changes in the levels of carbon dioxide in the atmosphere could alter the surface temperature of the Earth through the greenhouse effect, predicting that the burning of fossil fuels was large enough to cause significant global warming. Most remarkably he did all of this before 1900, which is pretty depressing because it shows that we already were beginning to understand scientifically some of the deleterious consequences of burning fossil fuels well over a hundred years ago and we did almost nothing about it.

215

September 23, 2023

Night Thoughts During the Hour of the Wolf According to the ancient Romans, the Hour of the Wolf means the time between night and dawn, just before the light comes, and people believed it to be the time when demons had a heightened power and vitality, the hour when most people died and most children were born, and when nightmares came to one. 12

216

September 23, 2023

This is the curse of consciousness. We all know we are going to die. No other organism is burdened with the enormity of the conscious knowledge that it has a finite lifetime and that its individual existence is eventually and inevitably coming to an end. No creature, whether a bacterium, an ant, a rhododendron, or a salmon, "cares" or even "knows" about dying; they live and they die, participating in the continual struggle for existence by propagating their genes into future generations and playing the endless game of the survival of the fittest. So do we.

216

September 23, 2023

I had a minor epiphany when I was sixteen years old. Some school friends persuaded me to join them in going to a small arts cinema in the West End of London to see a film much touted by the intelligentsia at that time. This was Ingmar Bergman's extraordinary film The Seventh Seal . It is of Shakespearean grandeur and depth. It tells the story of a medieval knight, Antonius Block, who on his return journey home to Sweden from fighting in the Crusades encounters the personification of Death, who has come to take his life. In his attempt to avoid, or at least delay, the inevitable, Block proposes that they play a game of chess; should he win, his life will be spared. He, of course, eventually loses but only because he is inadvertently tricked into baring his soul to

216

Death, who masquerades as a confessional priest. This allegorical setting provides the stage for delving into the eternal questions concerning the meaning, or pointlessness, of life and its relationship to death. Questions at the very heart of philosophical and religious discourses with which men and women have struggled throughout the centuries are brilliantly depicted by Bergman's genius. Who can forget the final haunting scene in which the black-robed Death leads Antonius and his entourage on an iconic danse macabre silhouetted across a distant hillside to meet their inevitable fate? What an impression this made on an innocent, unconscious, adolescent sixteen-year-old. I think this was my first truly serious inkling that there was more to life than money, sex, and football and began my long-term interest in questions of metaphysics and philosophical thought. I began to voraciously read all of the usual suspects from Socrates, Aristotle, and Job to Spinoza, Kafka, and Sartre, and from Russell and Whitehead to Wittgenstein, A. J. Ayer, and even Colin Wilson, though barely understanding anything of what any of them were saying (especially Wittgenstein, by the way). What I did learn, however, was that although extraordinary men had struggled with the really big questions for a very long time, there actually were no answers. Just more questions. It speaks to the profundity of Bergman's masterpiece that almost sixty years later the film still makes the same powerful impression, now possibly more nuanced and poignant, on a slightly jaded seventy-five-year-old man as he approaches the final years. At a critical stage in the film Death very reasonably asks Antonius: "Do you never stop questioning?" to which he emphatically responds, "No. I never stop." And neither should we. The fascination with death, coupled with the incessant questioning and search for any meaning to life, permeates human culture but has mostly been manifested and formalized in the multiplicity of religious institutions and experiences that humans have invented. Science has generally placed itself outside of such philosophical meanderings. However, many scientists have seen the quest for understanding and unraveling "the laws of nature," the passion for wanting to know how things work and what they are made of, as an alternative journey in coming to terms with these big questions even if they themselves are neither "religious" nor particularly "philosophical." Somewhere along the line, I *realized that I was one of them, finding in science, or at least in physics and mathematics, some version of the spiritual sustenance that seems to be a universal need.*

September 23, 2023

Once upon a time, science was referred to as natural philosophy, implying a somewhat broader connotation than the way we think of it today with a greater connection to philosophical and religious thought. It is no accident that the full title of Newton's famous book the Principia, which introduced his universal laws of nature that revolutionized science, is (in English) The Mathematical Principles of Natural Philosophy . Although Newton held heretical views such as rejecting the classical doctrines of an immortal soul, the existence of devils and demons, and the worship of Christ as God, which he viewed as idolatrous, he saw his work as the revelation of God as a prime mover. Commenting on the Principia, he stated: "When I wrote my treatise about our Systeme I had an eye upon such Principles as might work with considering men for the beliefe of a Deity and nothing can rejoyce me more than to find it useful for that purpose."

219

September 23, 2023

How did the universe evolve, what are stars made of, where did all the different animals and plants come from, why is the sky blue, when will the next eclipse occur, and so on and so forth. We understand an enormous amount about the physical universe around us and in many cases with exquisite detail, and we have done it without having to invoke ad hoc or arbitrary arguments that are often the hallmark of religious explanations. Left unanswered, however, are many of the deep questions concerning the very nature of who and what we are as human beings endowed with consciousness and the ability to reflect and reason. We continue to grapple with the nature of mind and consciousness, with psyche and self, with love and hate, and with meaning and purpose. Perhaps all will eventually be understood from the firing of neurons and the complex

219

network dynamics of our brains, but as D'Arcy Thompson proclaimed a hundred years ago, I suspect not. There will always be questions—that is the essence of the human condition—and like Antonius Block, we will never stop asking them even if it is to the great frustration and annoyance of Death. And somehow intertwined with all of this lies the challenge and paradox of understanding aging and mortality, and coming to terms with our collective and individual uneasiness with the finiteness of our own existence.

September 23, 2023

September 23, 2023 222

The city as the engine for social change and increasing well-being is one of the truly great triumphs of our amazing ability to form social groups and collectively take advantage of economies of scale. 225

September 23, 2023

September 23, 2023 225

This leads to the notion of a maximum possible age that a human being could conceivably live, which turns out to be less than around 125 years. Very few people have even approached this age. The oldest verified person ever was the Frenchwoman Jeanne Calment, who died in 1997 at the remarkable age of 122 years and 164 days. Just to get a sense of how exceptional this is, the next oldest verified person was the American Sarah Knauss, who lived more than three years fewer than Jeanne, dying at the age of 119 years and 97 days. The next super-champs of long life lived almost two years fewer than Sarah, while the oldest person still alive today is 227

the Italian Emma Murano, who is “only” in her 118th year.

September 23, 2023

The radical challenge: is it possible to extend life span beyond the apparent maximum limit of approximately 125 years and live, for instance, to 225 years? In a very real sense we’re already achieving the first, whereas it is the second that raises serious scientific questions. 228

September 23, 2023

The concept of a maximum life span is an extremely important one, as it implies that without some major “unnatural” intervention (which is what the elixir believers are seeking), natural processes inextricably limit human life span to about 125 years. Below, I will address what these limiting processes are and present a theoretical framework based on the network theory for determining this number. Before doing so, however, I want to show how classic survivorship curves provide powerful evidence in support of the concept of a maximum human life span. A survivorship curve simply represents the probability that an individual will live to a given age and is determined by plotting the percentage of survivors in a given population as a function of their age. Its converse is called a mortality curve and is the percentage of people who have died at a given age, representing the probability that an individual will die at that age. Biologists, actuaries, and gerontologists have coined the term mortality or death rate to denote the number of deaths in a population that occur in some given period of time (a month, say) relative to the number that are still alive. The general structure of survivorship and mortality curves is pretty obvious: most individuals survive the earliest years, but gradually a larger and larger percentage die until a point is reached where the probability of surviving eventually vanishes while the probability of dying reaches 100 percent. A great deal of statistical analysis has been done on such curves across different societies, cultures, environments, and species . 229

September 23, 2023

One of the surprising results that emerges is that the mortality rate for most organisms remains approximately unchanged with age. In other words, the relative number of individuals that die in any time period is the same at any age. So, for example, if 5 percent of the surviving population dies between ages five and six, then 5 percent of the surviving population will also die between ages forty-five and forty-six and between ninety-five and ninety-six. This sounds nonintuitive, but if we put it a different way it will make more sense. A constant mortality rate means that the number of individuals that die in some time period is directly proportional to how many have survived up until that time. If you go back to the earlier discussion on exponential behavior in chapter 3, you will discover that this is precisely the mathematical definition of the exponential function , which I will discuss in much greater detail in the following chapter. Here it says that survivorship follows a simple exponential curve, meaning that it becomes exponentially less likely that an individual from the original population will survive the older it gets, or equivalently, that it becomes exponentially more likely that an individual will die the older it gets. 230

September 24, 2023

Indeed, the data show that the half-life of publicly traded companies in the United States is only about ten years. So in just fifty years (five half-lives) only (½)5 = 1/32 or about 3 percent are still posting sales. This begs the fascinating question as to whether the same general dynamics underlies the surprising commonality in the mortality of organisms, isotopes, and companies. We will return to speculate about this later. 231

September 24, 2023

There are two important points I want to emphasize arising from these statistics: (1) The leading causes of death are overwhelmingly associated with damage, whether in organs and tissue (as in heart attacks or stroke) or in molecules (as in cancer)—infectious diseases play a relatively minor role. (2) Even if every cause of death were eliminated, all human beings are destined to die before they reach 125 years old, and the vast majority of us will do so well before we reach that ripe old age.

234

September 24, 2023

Most of the discussion thus far has been about human beings, but I now want to extend it to other animals in order to connect it to scaling laws and the theoretical framework introduced earlier. The discussion will be in the spirit of a coarse-grained description, so there are undoubtedly outliers and even exceptions to some of these statements. This is especially true for the case of aging and mortality because, unlike most other traits, these are not directly selected for in the evolutionary process. Natural selection only needs to ensure that the majority of individuals in a species survive sufficiently long to produce enough offspring to maximize their evolutionary fitness. Once this has happened and they have performed their evolutionary "duty," how much longer they live is of much less importance, so large variations in individual and species life spans can be expected. Thus human beings have evolved to live for at least forty years so that they can produce ten or so children, at least half of whom will survive to maturity and beyond. It's perhaps no accident, then, that this is the age of a woman's menopause. However, to ensure that enough of us reach this age and reproduce accordingly, we have evolved to be sufficiently "overengineered" that statistically many of us are able to live much longer.

235

September 24, 2023

Here's a summary of some of the significant properties of

236

aging and mortality that need to be explained by any theory: Aging and death are "universal": all organisms eventually die. A corollary to this is that there is a maximum life span and a corresponding vanishing survival rate. Semiautonomous subsystems of organisms, such as our various organs, age approximately uniformly. Aging progresses approximately linearly with age. For instance, Figure 27 shows how organ functionality degrades with age. 15 Plotted is the percentage of maximum capacity for various vital functions indicating a linear decline with age beginning almost immediately after maturity at about age twenty. It's slightly depressing to see that on average we are physically optimal (100 percent) for only a very few years and that beginning around age twenty it's literally downhill all the way. Notice also that we build up to our maximum capacity relatively quickly during our growth period. Later, I will suggest that this aging process is under way even during our earliest years prior to maturity but is hidden by the overwhelming predominance of growth. The aging process effectively begins as soon as you are conceived. Bob Dylan got it right when he sang that "he not busy being born is busy dying." Life spans scale with body mass as a power law whose exponent is approximately ¼. As anticipated, there is a large variance in the data, partly because there are no controlled life-history experiments on longevity for mammals, including us. Some of the data are garnered from reports on wild animals, some from zoos, some from domesticated animals, some from research laboratories, each with very different environmental and lifestyle conditions. In addition, some are reports of just one or two animals of a species, and some from large cohorts. Although this lack of control is problematic, there are clear trends and consistencies in the data that statistically point to approximate ¼ power scaling. The number of heartbeats in a lifetime is approximately the same for all mammals, as shown in Figure 2 in the opening chapter. 16 Thus, shrews have heart rates of roughly 1,500 beats a minute and live for about two years, whereas heart rates of elephants are only about 30 beats a minute but they live for about seventy-five years. Despite their vast difference in size, both of their hearts beat approximately one and a half billion times during an average lifetime. This invariance is approximately true for all mammals, even though there are large *fluctuations for the reasons I outlined above. The greatest outlier from this intriguing invariance is us: for modern human beings, on average our hearts beat approximately two and a half billion times, which is about twice the number for a typical mammal. However, as I have already emphasized, it is only in the past one hundred years that we have been living this long. Over the entire history of humankind up until relatively recently we lived for approximately half as long as we do now and, like the vast majority of mammals, followed the approximately invariant one and a half billion heartbeats "law." Related to this is another invariant quantity: the total amount of energy used in a lifetime to support a gram of tissue is approximately the same for all mammals and, more generally, for all animals within a specific taxonomic group. 17 For mammals it's about 300 food calories per gram per lifetime. A more fundamental way of expressing this is to note that the number of turnovers during a lifetime of the respiratory machinery responsible for the production of energy in cells is approximately the same for all animals within a specific taxonomic group. For mammals this is about ten thousand trillion times (1016) and translates into the invariance of the number of ATP molecules (our fundamental currency of energy) produced in a lifetime to support one gram of tissue.*

September 24, 2023

Before exploring this further, it is instructive to compare some of this with the longevity of automobiles. Unfortunately, there have been surprisingly few scaling analyses of automobiles and other machines, especially regarding their longevity. However, Thomas McMahon, an engineer at Harvard, analyzed data on internal combustion engines, ranging from those used in lawn mowers and automobiles to airplanes, and showed that they follow simple isometric Galilean cubic power scaling laws, discussed in chapter 2. For instance, the horsepower rating of these engines (the analog to their metabolic rate) scales linearly with their weight, so to double their power output, you have to double their weight. Thus, unlike organisms, engines do not exhibit an economy of scale as their size increases. McMahon also found that their RPMs (their heart rates)

239

scale with an inverse cubic power of their weight. 18

September 24, 2023

These are in distinct contrast to the ¼ power scaling laws obeyed by organisms, which result from their optimized fractal-like network structures: their metabolic rates (their horsepower) scale with an exponent of ¾ and their heart rates (their RPMs) with an exponent of -¼. The fact that internal combustion engines have no complex network structures and do not follow ¼ power scaling is supporting evidence for the underlying network theory for the origin of ¼ power scaling in biology. Because manufactured engines satisfy classic cubic scaling, one might speculate that their life span increases with the cube root of their weight rather than with a quarter power. Unfortunately, there isn't sufficient data available to test this. However, qualitatively it does predict that bigger automobiles should last longer. In fact, all of the top ten longest-lasting vehicles are either large trucks or SUVs, while only three regular-size sedans make it into the top twenty. If you're simply looking for longevity, buy big: the Ford F-250 is tops, with the Chevrolet Silverado second and the Suburban third. Cars are now typically expected to last for about 150,000 miles. In fact, like the humans who made them, their longevity has increased dramatically over a relatively short period of time, having almost doubled over the past fifty years. Just to get an idea of what this implies, suppose that when averaged over its lifetime, a typical car moves at 30 miles per hour and that its "heart rate" is 2,500 RPMs, then the total number of "engine beats" during a 150,000-mile lifetime is approximately a billion. Amusingly, this is not so different from the number of times a mammalian heart beats in its lifetime. Is this just a coincidence, or is it telling us something about the commonality of mechanisms responsible for aging? 240

El Bocho de Mi Pá.

September 24, 2023

September 24, 2023 241

Like all organisms, we metabolize energy and resources in a highly efficient way in order to combat the continuous fight against the inevitable production of entropy in the form of waste products and dissipative forces that cause physical damage. As we begin to lose the multiple localized battles against entropy we age, ultimately losing the war and succumbing to death. Entropy kills. Or as the great Russian playwright Anton Chekhov poignantly remarked, "Only entropy comes easy."

241

September 24, 2023

Damage occurs across multiple scales through many different mechanisms associated with physical or chemical transport phenomena, but loosely speaking it can be separated into two categories: (1) Classic physical wear and tear due to the viscous drag in the flow, analogous to the wear and tear resulting from ordinary friction when two physical objects move over each other like wearing out your shoes or the tires of your car. (2) Chemical damage from free radicals, which are by-products of the production of ATP in respiratory metabolism. A free radical is any atom or molecule that has lost an electron and consequently has a positive electric charge, making it highly volatile. Most of this kind of damage is caused by oxygen radicals that react with vital cellular components. Oxidative damage to DNA may be particularly deleterious, because in nonreplicating cells such as in the brain and musculature, it causes permanent damage to transcriptional, and perhaps most important, regulatory regions of the genome. Although the detailed role and extent of oxidative damage in aging remains unclear, it has stimulated a mini-industry of antioxidant supplements such as vitamin E, fish oil, and red wine as some of those elixirs of life to combat aging.

242

September 24, 2023

This coarse-grained framework is very general and can incorporate any model of aging based on generalizations of “damage” mechanisms associated with generalized physical or chemical transport phenomena discussed above. The details of the damage mechanism are not important for understanding many of the general features of aging and mortality because the most relevant damage occurs in invariant terminal units of networks (capillaries and mitochondria, for example) whose properties do not appreciably change with the size of the organism. Consequently, the damage per capillary or mitochondrion is approximately the same regardless of the animal. Because these networks are space filling, meaning that they service all cells and mitochondria throughout the body of the organism, damage occurs approximately uniformly and relentlessly throughout the organism, explaining why aging is approximately spatially uniform and progresses approximately linearly with age. This is why at age seventy-five every part of your body has deteriorated to pretty much the same degree, as indicated in Figure 27.

242

September 24, 2023

At a more detailed level, this implies that aging within each organ is approximately uniform even though different organs may age at moderately different rates because they have slightly different network characteristics, especially in regard to their potential for repair. Because larger animals metabolize at higher rates following the ¾ power scaling law, they suffer greater production of entropy and therefore greater overall damage, so you might have thought that this would imply that larger animals would have shorter life spans in obvious contradiction to observation. However, we saw in chapter 3 that on a cellular or per unit mass of tissue basis metabolic rate and therefore the rate at which damage is occurring at the cellular and intracellular levels decreases systematically with increasing size of the animal—another expression of economy of scale. Furthermore, as already emphasized, the most significant damage occurs at the terminal units of networks in capillaries, mitochondria, and

243

cells, and their metabolic rates decrease with the size of the organism following power law scaling with an exponent of ¼. Cells in larger animals are systematically processing energy at a slower rate than cells in smaller ones. So at the critical cellular level cells suffer systematically less damage at a slower rate the larger the animal, and this results in a correspondingly longer life span.

September 24, 2023

The total number of damages incurred in a lifetime is just the damage rate (that is, the number of damage events per unit time which is proportional to the number of terminal units) multiplied by the life span, and this has to be proportional to the total number of cells, and therefore to body mass. Consequently, life span is proportional to the total number of cells divided by the number of terminal units. But the number of terminal units scales with mass with a ¾ power exponent, while the number of cells scales linearly, resulting in life span scaling as the ¼ power of mass, consistent with data.

245

September 24, 2023

Notice that just as we saw when discussing growth, the mismatch between the scaling of the sources of energy and therefore the sources of damage (the terminal units) and the scaling of the sinks of energy (cells that need to be sustained) has enormous consequences. In the one case, it ensures that we cease growing, and in the other it ensures that larger animals have expanded lifetimes. And all of this follows from the constraint of the networks.

246

September 24, 2023

A. TEMPERATURE AND EXTENDING LIFE SPAN: Because metabolic rate is proportional to the number of terminal units and these are where most damage occurs, we can relate life span directly to metabolic rate. This results in an

246

alternative expression for life span as the ratio of the body mass to its metabolic rate. In other words, life span is inversely proportional to the metabolic rate per unit mass of the organism and therefore inversely proportional to the average metabolic rate of its cells. We saw just above when discussing the metabolic theory of ecology that this systematically scales with body mass following quarter-power scaling and as an exponential with temperature.

September 24, 2023

September 24, 2023 246

This implies that life span can in principle be extended by lowering the body temperature because this lowers cellular metabolic rate and therefore the rate at which damage is incurred. This is a very large effect: I remind you that a modest 2°C decrease in body temperature can result in a 20 percent to 30 percent increase in life span. 19 So if you were able to artificially lower your body temperature by just 1°C (that's about 1.8°F) you could enhance your life span by about 10 percent to 15 percent. The hitch is that you would have to do this for your entire life in order to reap the "benefits." But more saliently, significantly lowering body temperature may well have many other deleterious, potentially life-threatening outcomes. As I have stressed earlier, changing just one component of a complex adaptive system without fully understanding its multilevel spatiotemporal dynamics usually leads to unintended consequences. 246

September 24, 2023

B. HEARTBEATS AND THE PACE OF LIFE: These data also confirm the approximate ¼ power scaling of life spans with mass. Because the theory for the cardiovascular system predicts that heart rates decrease with a ¼ power of mass, 247 *the dependence on mass cancels out when we multiply heart rates by life spans: the decrease in one exactly compensates for the increase in the other, resulting in an invariant, a quantity that's the same for all mammals. But multiplying heart rates by life spans just gives the total number of heartbeats in a lifetime, so the theory predicts that this should be the same for all mammals, consistent with the data shown in Figure 2 in chapter 1. This argument can be extended to the fundamental level of the respiratory complexes, the basic units inside mitochondria where ATP is manufactured, to show that the number of times the reaction takes place producing ATP is the same for all mammals.*

September 24, 2023

Given the relative simplicity of the theory, the agreement is surprisingly good. Coupled with the success of its other predictions (including the rate of aging, the allometric scaling of life span, and its temperature dependence) the theory provides a credible coarse-grained baseline for developing a more detailed quantitative theory for understanding aging and mortality. It gives formulae for the rate of aging and maximum life span in terms of generic "universal" biological parameters that show, for instance, how the scale of one hundred years arises from microscopic molecular scales and why mice live for only a very few years. This provides the scientific basis for asking questions about what parameters can be manipulated to extend life and arrest aging, if that is the goal. For example, combining the scaling laws with Figures 23–30 gives quantitative estimates of how long life can be extended by changing body temperature or eating less.

249

September 24, 2023

September 24, 2023

251

5: FROM THE ANTHROPOCENE TO THE URBANOCENE: A Planet Dominated by Cities

September 25, 2023 — 253

September 25, 2023 — 255

Given this dual nature of cities as, on the one hand, the origin of our major challenges and, on the other, the generator of creativity and ideas and therefore the source of their solutions, it becomes a matter of some urgency to ask whether there can be a "Science of Cities," by which I mean a conceptual framework for understanding their dynamics, growth, and evolution in a quantitatively predictable framework. This is crucial for devising a serious strategy for achieving long-term sustainability, especially given that the overwhelming majority of human beings will be urban dwellers by the second half of this century, many in megacities of unprecedented size. 2

257

September 25, 2023

The growth of a system, whether an economy or a population, is often expressed in terms of a quantity called the doubling time, which is simply the time it takes for the size of the system to double. Exponential growth is characterized by having a constant doubling time, which also sounds fairly harmless until one realizes that it implies, for example, that it would take the same time for a population to double from ten thousand to twenty thousand, thereby adding just ten thousand people, as it would for it to double from 20 million to 40 million, thereby adding a humongous 20 million people. Amazingly, the doubling time of the global population has actually been getting systematically shorter

260

and shorter as was indicated above: it took 300 years from 1500 to 1800 for the population to double from 500 million to a billion, but only 120 years to double to 2 billion, and only another 45 to double again to 4 billion.

September 25, 2023

Questions such as these have fueled a highly spirited ongoing debate that began very soon after the Industrial Revolution set the wheels of exponential growth in motion and has continued right up to the present time. The transition from working the land and artisanal hand production to automated machines and the creation of factories for the mass production of goods, the technological innovation and increased productivity in agriculture, the introduction of new chemical manufacturing and iron production processes, the improved efficiency of water power, and the increasing use of steam fueled by the change from renewable wood energy to fossilized coal energy, all contributed to an inevitable migration of more and more people away from a traditional rural existence to the rapidly expanding urban centers that were perceived as providing greater opportunities for employment. This process continues unabated across the globe to this day. 3

266

September 25, 2023

267

September 25, 2023

Thomas Robert Malthus is usually credited with being the first person to recognize the potential threat posed by open-ended exponential growth and connect it to the challenge of resource limitation and availability. Malthus was an English cleric and scholar and an early contributor to the newly emerging fields of economics and demography and their implications for long-term political strategy. He published a highly influential essay in 1798 titled An Essay on the

271

Principle of Population in which he declared that "the power of population is indefinitely greater than the power in the earth to produce subsistence for man." His argument was that the population "multiplies geometrically," meaning that it increases at an exponential rate, whereas the ability to grow and supply food increases only "arithmetically," meaning that it increases at a much slower linear rate, so the size of the population will eventually outstrip the food supply, leading to catastrophic collapse.

September 25, 2023

I have met scant few economists who do not automatically dismiss traditional Malthusian-like ideas of eventual or imminent collapse as naive, simplistic, or just plain wrong. On the other hand, I have met scant few physicists or ecologists who think it's nuts to believe otherwise. The late maverick economist Kenneth Boulding perhaps best summed it up when testifying before the U.S. Congress, declaring that "anyone who believes exponential growth can go on forever in a finite world is either a madman or an economist." 274

September 25, 2023

Four companies alone now produce 81 percent of the cattle, 73 percent of the sheep, 57 percent of the pigs, and 50 percent of the chickens consumed in the United States. 275

September 25, 2023

September 25, 2023 275

A leading critic was the well-known economist Julian Simon, who expressed a fairly extreme version of the view held by many economists that the spectacular growth we have witnessed over the past two hundred years will be sustained 277 *"forever" by human ingenuity and our continued ability to innovate. In fact, Simon argued in his 1981 book The Ultimate Resource that a larger population is actually better because it stimulates even more technological innovation, inventiveness, and ingenuity, thereby leading to new ways of exploiting resources and increasing standards of living. 6 As we move into the twenty-first century this vision of a cornucopian "horn of plenty," with its image of limitless barrels of fish being continuously replenished by the magic, not of divine intervention, but of the free expression of human ingenuity and the boundless possibilities of a free market economy, has reemerged as a significant component of corporate and political conceptual thinking. Indeed, Simon's views have effectively been embraced by many in the academic, business, and political communities. A succinct summary of this view was articulately expressed by the economist Paul Romer, one of the founders of endogenous growth theory, which holds that economic growth is driven primarily by investment in human capital, innovation, and knowledge creation. 7 Romer declares that "every generation has perceived the limits to growth that finite resources and undesirable side effects would pose if no new recipes or ideas were discovered. And every generation has underestimated the potential for finding new recipes and ideas. We consistently fail to grasp how many ideas remain to be discovered. Possibilities do not add up. They multiply." Put slightly differently, this proclaims that ideas and innovation increase multiplicatively (that is, exponentially) and not arithmetically (that is, linearly) in tandem with the exponential growth in population and that the process is open-ended and effectively limitless. On the other hand, the last few decades have also seen a resurfacing of the spiritual successors of The Population Bomb and The Limits to Growth with the rise of the environmental movement and the development of a serious concern for the future of the planet. Closely related to this is a deep concern for the impact of unregulated corporate and political ambition, which has stimulated a perceived need for "corporate social responsibility."*

September 25, 2023

While it may not be unreasonable to hold the view that collective human ingenuity and innovation facilitated by a free market economy is the secret for maintaining long-term open-ended growth in defiance of potential collapse, I find it somewhat baffling that this is often coupled with a denial or, at the very least, a deep skepticism concerning some of its inevitable consequences. Like many who advocate "innovation" as the panacea for meeting future global socioeconomic challenges, Simon was a vocal skeptic when it came to believing that human activity caused global environmental damage or was the origin of serious health concerns, whether from climate change, pollution, or chemical contamination. The spirit and substance of the Second Law of Thermodynamics and its manifestation in terms of entropy production represent the dark side of open-ended exponential growth. Independent of how superbly innovative we are, ultimately everything is driven and processed by the use of energy, and the processing of energy has inevitable deleterious consequences. [278](#)

September 25, 2023

September 25, 2023 [280](#)

To put it slightly differently, the rate at which we need to process energy to sustain our standard of living remained at just a few hundred watts for hundreds of thousands of years, until about ten thousand years ago when we began to form collective urban communities. This marked the beginning of the Anthropocene, in which our effective metabolic rate began its steady rise to its present level of more than 3,000 watts today. But this is just its average value taken across the entire planet. The rate at which energy is used in developed countries is far higher. In the United States it is almost a factor of four larger, at a whopping 11,000 watts, which is more than one hundred times larger than its "natural" biological value. This amount of power is not a lot smaller than the metabolic rate of a blue whale, which is more than one thousand times larger in mass than we are. [280](#)

Thinking of us as an animal using thirty times more energy than we "should" given our physical size, the effective human population of the planet accordingly operates as if it were much larger than the 7.3 billion people who actually inhabit it. In a very real sense, we are operating as if our population were at least thirty times larger, equivalent to a global population in excess of 200 billion people. If the most optimistic of cornucopian thinkers are correct and the world's population reaches 10 billion by the end of the century, all living at a standard comparable to that of the United States, the subsequent effective population would then exceed a trillion people.

September 25, 2023

There is no doubt that significant progress has been and will continue to be made, but can it be enough? It is a matter of faith that the free market system geared to open-ended growth, even when tempered by governmental intervention, stimulation, and regulation, can find a meta-stable balance between making significant profits and solving the problem of sustainability. The primary function of business, after all, is not to increase efficiencies, but to make profits. 281

September 26, 2023

Consequently, the processes that govern the weather and the life history of plants and animals are exponentially sensitive to small changes in the temperature at which they operate. I remind you that a 2°C rise in average temperature leads to a whopping 20 percent increase in these rates. 283

September 26, 2023

Despite the obviously central role that energy has played in bringing us to this point in the history of the planet and, in particular, in the socioeconomic development of modern human societies, you will be hard-pressed to find even a 284

sentence or two about it in any classic textbook on economics. Remarkably, concepts like energy and entropy, metabolism, and carrying capacity have not found their way into mainstream economics. The continued growth of economies, markets, and populations over the past two hundred years coupled with a parallel increase in standards of living are, not surprisingly, seen as testament to the success of classic economic thinking and as a rejection of neo-Malthusian ideas. There has been no need to think seriously in terms of energy as an underlying driver of economic success or of population growth, let alone to consider entropy as its inevitable consequence. Nor has there been the need to consider the possibility that resources may actually be limited, nor that there might be underlying physical constraints that would question open-ended growth. Until now.

September 26, 2023

In addition, there also seems to be an unexpressed presumption that ideas, the seeds of innovation, cost nothing—after all, they are "just" neural processes in the brains of human beings, and collectively we can produce an almost infinite number of them in our heads. But like everything else, ideas and the innovations they inspire require energy, and lots of it, to support the smart individuals who are thinking and to provide the appropriately stimulating environments and collective experiences we institutionalize in places such as universities, laboratories, parliaments, cafés, concert halls, and conference venues.

285

September 26, 2023

September 26, 2023

285

Of the trillions of thoughts, ideas, speculations, and

286

proposals for new machines, new products, and new theories, only an infinitesimal minority ever lead to any significance. Almost all fall by the wayside, even though as a totality they all contribute a necessary background noise and weltanschauung for new and innovative phenomena to arise and blossom. All of this requires huge amounts of energy: ex nihilo nihil fit —nothing comes from nothing.

September 26, 2023

September 26, 2023 286

288

The lack of progress until relatively recently is quite astonishing given that the basic technology for developing solar energy has been known for more than a hundred years. In 1897, an American engineer, Frank Shuman, built a proof of principle device for utilizing energy from the sun by showing that it could power a small steam engine. His system was eventually patented in 1912, and in 1913 he constructed the world's first solar thermal energy power plant, which was built in Egypt. It generated only about 50 kilowatts (about 65 horsepower) but was able to pump more than 5,000 gallons of water a minute (about 22,000 liters a minute) from the Nile onto adjacent cotton fields. Shuman was an enthusiast and advocate for solar energy and in 1916 was quoted in The New York Times as saying: We have proved the commercial profit of sun power . . . and have more particularly proved that after our stores of oil and coal are exhausted the human race can receive unlimited power from the rays of the sun.

September 26, 2023

293

Regardless of whether one believes in the innovative capacity of human beings to solve the problems of nuclear energy, whether fusion or fission, or the challenge of affordable and reliable solar technology sufficient to support the energy *needs of 10 billion people, or to reverse the amount of carbon we are pumping into our atmosphere, we are still left with the long-term problem of entropy production. Apart from its many other issues, the nuclear option, like that of traditional fossil fuels, keeps us trapped in the paradigm of a closed system, whereas the solar option has the critical capacity for potentially returning us to a truly sustainable paradigm of an open system.*

September 26, 2023

6: PRELUDE TO A SCIENCE OF CITIES

However, it is in business that the concepts and language if not the actual science of ecology and evolutionary biology have seized the imagination, especially in Silicon Valley. The concept of a business ecosystem has become a standard buzzword to connote a sort of Darwinian survival of the fittest in the marketplace. It was introduced in 1993 by James Moore, then at Harvard Law School, in an article he wrote titled "Predators and Prey: A New Ecology of Competition," which won the McKinsey award for article of the year. 1 It's a fairly standard ecological narrative, with individual businesses replacing animals in the evolutionary dynamics of natural selection. In keeping with much of the traditional literature on understanding companies, it's entirely qualitative with no quantitative predictive power. Its great virtue is that it emphasizes the role of community structure, the importance of systemic thinking, and the inevitable processes of innovation, adaptation, and evolution. 295

September 26, 2023

Despite my skepticism, David and Sander convinced me that extending the network-based scaling theory from biology to social organizations was indeed a worthwhile project. They became the prime movers in putting together a broad 296 *program that covered our joint interests ranging from innovation and information transfer in both ancient and modern societies to understanding the structure and dynamics of cities and companies, all from a complexity perspective. The program was called the Information Society as a Complex System (ISCOM) and was generously funded by the European Union. Shortly thereafter Denise Pumain, a well-known urban geographer at the Sorbonne in Paris, joined our collaboration and the four of us each ran one component of the project. I assembled a new multidisciplinary collaboration centered on SFI whose first goal was to ask whether cities and companies manifest scaling and, if so, to develop a quantitative principled theory for understanding their structure and dynamics.*

September 26, 2023

298

Although its scope and emphasis broadened as progress was made, the vision of the proposal remained pretty much intact over the years. Its motivation was originally expressed as follows: "Because of the obvious analogy with social network systems, such as corporate and urban structures, it is both natural and compelling to investigate the possibility of extending the same sort of analyses used for understanding biological network systems to social organizations," with an added emphasis that "the flow of information in social organizations is as significant as the flow of matter, energy and resources." Many questions were asked, including "What is a social organization? What are the appropriate scaling laws? What constraints must be satisfied by the architecture of the structures that channel social flows of information, matter and energy? In particular, are the relevant constraints all physical, or might there also be social and cognitive constraints that must be taken into account?

September 26, 2023

301

No one is more identified with viewing cities through the collective lives of their citizens than the famous urban writer-theorist Jane Jacobs. Her defining book, The Death and Life *of Great American Cities, had an enormous influence across the globe on how we think about cities and how we approach "urban planning." 3 It's required reading for anyone interested in cities whether a student, a professional, or just an intellectually curious citizen. I suspect that every mayor of every major city in the world has a copy of Jane's book sitting somewhere on his or her bookshelf and has read at least parts of it. It's a wonderful book, extremely provocative and insightful, highly polemical and personal, very entertaining and well written.*

September 26, 2023

She was extraordinarily intolerant of urban planners and politicians, and vicious in her attacks on traditional urban planning, especially in regard to its apparent lack of recognition that people, not buildings and highways, were primary. These classic quotes from her writing are typical of her critical attitude: The pseudoscience of planning seems almost neurotic in its determination to imitate empiric failure and ignore empiric success. In this dependence on maps as some sort of higher reality, project planners and urban designers assume they can create a promenade simply by mapping one in where they want it, then having it built. But a promenade needs promenaders. There is no logic that can be superimposed on the city; people make it, and it is to them, not buildings, that we must fit our plans. . . . We can see what people like. His aim was the creation of self-sufficient small towns, really very nice towns if you were docile and had no plans of your own and did not mind spending your life with others with no plans of their own. As in all Utopias, the right to have plans of any significance belonged only to the planner in charge. 301

September 26, 2023

September 26, 2023 304

September 26, 2023 304

September 26, 2023 305

live in the age of the Urbanocene, and globally the fate of the cities is the fate of the planet. Jane understood this truth more than fifty years ago, and only now are some of the experts beginning to recognize her extraordinary foresight. 311

September 28, 2023

Many writers have picked up this theme, including the urban economists Edward Glaeser and Richard Florida, but none has been as forthright and bold as Benjamin Barber in his book with the provocative title If Mayors Ruled the World: Dysfunctional Nations, Rising Cities. 5 These are indicative of a rising consciousness that cities are where the action is—where challenges have to be addressed in real time and where governance seems to work, at least relative to the increasing dysfunctionality of the nation-state. 311

September 28, 2023

September 28, 2023 312

Cities have an organic quality. They evolve and physically grow out of interactions between people. The great metropolises of the world facilitate human interaction, creating that indefinable buzz and soul that is the wellspring of its innovation and excitement and a major contributor to its resilience and success, economically and socially. It is 318 *shortsighted and even courting disaster to ignore this critical dimension of urbanization and concentrate only on buildings and infrastructure.*

September 28, 2023

September 28, 2023 319

7: TOWARD A SCIENCE OF CITIES

September 28, 2023 320

The first step in carrying out such a program was to ask if cities are approximately scaled versions of one another in a similar way that animals are. In terms of their measurable characteristics, are New York, Los Angeles, Chicago, and Santa Fe scaled versions of one another and, if so, is their relative scaling similar to the way in which Tokyo, Osaka, Nagoya, and Kyoto scale, despite their very different appearances and characters? Does their scaling manifest any analog of the universality we saw in biology where whales, elephants, giraffes, human beings, and mice are approximately scaled versions of one another, all neatly and quantitatively expressed by the preponderance of quarter-power scaling laws? Compared with biology, surprisingly little attention had been paid to such questions regarding cities, urban systems, or companies prior to our work. To some extent this is because urban studies have historically been even less generally quantitative than biology, but also because relatively few computational mechanistic models of cities or companies have been proposed, let alone confronted with data. 320

September 29, 2023

An early recruit to our collaboration was Dirk Helbing, who was the director of the Institute for Transport and Economics at Dresden University of Technology in Germany when I first met him. Dirk had been trained in statistical physics and had applied these techniques to understand highway traffic and pedestrian crowds. He's now at the prestigious Swiss Federal Institute of Technology in Zurich, usually referred to as the ETH, where he runs a large project called the Living Earth Simulator. This is designed to model global-scale systems from economies, governments, and cultural trends to epidemics, agriculture, and technological developments using big data sets and fancy algorithms. 321

September 29, 2023

Furthermore, the slope of the straight line, which is the exponent of the power law, is about 0.85, a little bit higher than the 0.75 (the famous ¾) we saw for the metabolic rate of organisms (Figure 1). Equally intriguing is that this exponent takes on approximately the same value for how gasoline stations scale across all of the countries shown in the figure. This value of around 0.85 is smaller than 1, so in the language developed earlier, the scaling is sublinear , indicating a systematic economy of scale , meaning that the bigger the city the fewer the number of gas stations needed on a per capita basis. Thus, on average, each gas station in a larger city serves more people and consequently sells more fuel per month than in a smaller one. 322

September 29, 2023

It may not come as such a big surprise to learn that larger cities require fewer gas stations per capita than smaller ones, but what is surprising is that this economy of scale is so systematic: it is approximately the same across all of these countries, obeying the same mathematical scaling law with a 323 *similar exponent of around 0.85. What is even more surprising is that other infrastructural quantities associated with transport and supply networks, such as the total length of electrical lines, roads, water and gas lines, all scale in much the same way with approximately the same value of the exponent, namely about 0.85. Furthermore, this systematic behavior appears to be the same across the globe wherever data could be obtained. So as far as their overall infrastructure is concerned, cities behave just like organisms—they scale sublinearly following simple power-law behavior, thereby exhibiting a systematic economy of scale, albeit to a lesser degree as represented by the different values of their exponents (0.75 for organisms vs. 0.85 for cities).*

September 29, 2023

September 29, 2023

324

324

The extension of this initial probe into whether cities scale to a broader suite of metrics and across a wider range of countries was carried out by a highly talented group of new recruits to the collaboration. These included Luis Bettencourt, whom I first got to know when he was a postdoctoral fellow in astrophysics at Los Alamos working on the evolution of the early universe. He had meanwhile spent a couple of years at MIT before returning to Los Alamos as a staff member in the applied mathematics group. Luis was born, raised, and educated in Portugal, though you'd never know it as he speaks English fluently without any trace of an accent, so much so that when I first met him I thought he was English. He had in fact obtained his doctorate in physics at Imperial College in London, where coincidentally I hold a position in the mathematics department. Luis's fluidity with language is matched by his fluidity with science. He very quickly became engaged with the cities project, gathering and analyzing data from across the globe. He is a passionate adherent of the cause of developing a deep understanding of cities and has now established himself as one of the world's leading *experts in this arena.*

September 29, 2023

Luis was joined in this endeavor by another very bright recruit, José Lobo, an urban economist now in the sustainability program at Arizona State University. When we first met, José was a young faculty member in the Department of City and Regional Planning at Cornell University and had been coming to SFI for several years. Like Luis, José brought a real talent for statistics and sophisticated data analysis to our program and, in addition, brought professional expertise in cities and urbanization, a critical component of our collaboration. 324

September 29, 2023

Luis and José took the lead in assembling and analyzing extensive data sets covering a broad range of metrics about urban systems across the globe, ranging from Spain and the Netherlands in Europe to Japan and China in Asia and Colombia and Brazil in Latin America. This convincingly verified the earlier analyses showing sublinear scaling of infrastructural metrics and strongly supported the universality of systematic economies of scale in cities. Regardless of the specific urban system, whether Japan, the United States, or Portugal, and regardless of the specific metric whether the number of gas stations, the total length of pipes, roads, or electrical wires, only about 85 percent more material infrastructure is needed with every doubling of city size. 2 Thus a city of 10 million people typically needs 15 percent less of the same infrastructure compared with two cities of 5 million each, leading to significant savings in materials and energy use. 3 325

September 29, 2023

This savings leads to a significant decrease in the production 325 *of emissions and pollution. Consequently, the greater efficiency that comes with size has the nonintuitive but very important consequence that on average the bigger the city, the greener it is and the smaller its per capita carbon footprint. In this sense, New York is the greenest city in the United States, whereas Santa Fe, where I live, is one of the more profligate ones. On average, each of us in Santa Fe is putting almost twice as much carbon into the atmosphere as New York. This should not be thought of as somehow reflecting the greater wisdom of New York's planners and politicians, nor as the fault of Santa Fe's leadership, but rather as an almost inevitable by-product of the dynamics underlying economies of scale that transcend the individuality of cities as their size increases. These gains are mostly unplanned, though policy makers in cities can certainly play a powerful role in facilitating and enhancing the hidden "natural" processes that are at work. In fact, this is a large part of what their job is. Some cities are very successful at doing this, while others are much less so. I'll discuss the question of relative performance in the next chapter. These results are very encouraging and provide powerful evidence in support of the quest for a possible theory of cities. However, of even greater significance was the surprising discovery that the data also reveal that socioeconomic quantities with no analog in biology such as average wages, the number of professional people, the number of patents produced, the amount of crime, the number of restaurants, and the gross urban domestic product (GDP) also scale in a surprisingly regular and systematic fashion, as illustrated in Figures 34–38.*

September 29, 2023

326

Also clearly manifested in these graphs is the equally surprising result that all of the slopes of these various quantities have approximately the same value, clustering around 1.15. Thus these metrics not only scale in an extremely simple fashion following classic power law behavior, but they all do it in approximately the same way with a similar exponent of approximately 1.15 regardless of the urban system. So in marked contrast to infrastructure, which scales sublinearly with population size, socioeconomic *quantities—the very essence of a city—scale superlinearly , thereby manifesting systematic increasing returns to scale. The larger the city, the higher the wages, the greater the GDP, the more crime, the more cases of AIDS and flu, the more restaurants, the more patents produced, and so on, all following the “15 percent rule” on a per capita basis in urban systems across the globe.*

September 29, 2023

The data convincingly show that despite appearances cities are approximately scaled versions of one another: New York and Tokyo are, to a surprising and predictable degree, nonlinearly scaled-up versions respectively of San Francisco and Nagoya. These extraordinary regularities open a window onto underlying mechanisms, dynamics, and structures common to all cities and strongly suggest that all of these phenomena are in fact highly correlated and interconnected, driven by the same underlying dynamics and constrained by the same set of “universal” principles. 330

September 29, 2023

Consequently, each of these urban characteristics, each metric—whether wages, the length of all the roads, the number of AIDS cases, or the amount of crime—is interrelated and interconnected with every other one and together they form an overarching multiscale quintessentially complex adaptive system that is continuously integrating and processing energy, resources, and information. The result is the extraordinary collective phenomenon we call a city, whose origins emerge from the underlying dynamics and organization of how people interact with one another through social networks. To repeat: cities are an emergent self-organizing phenomenon that has resulted from the interaction and communication between human beings exchanging energy, resources, and information. As urban creatures we all participate in the multiple networks of intense human interaction that is manifested in the metropolitan buzz of productivity, speed, 330

and ingenuity, no matter where we live.

September 29, 2023

To summarize: the bigger the city, the greater the social activity, the more opportunities there are, the higher the wages, the more diversity there is, the greater the access to good restaurants, concerts, museums, and educational facilities, and the greater the sense of buzz, excitement, and engagement. These facets of larger cities have proven to be enormously attractive and seductive to people worldwide who at the same time suppress, ignore, or discount the inevitable negative aspects and the dark side of increased crime, pollution, and disease. Human beings are pretty good at "accentuating the positive and eliminating the negative," especially when it comes to money and material well-being. In addition to the perceived individual benefits coming from increased city size, there are huge collective benefits arising from systematic economies of scale. Coupled together, this remarkable combination of increasing benefits to the individual with systematic increasing benefits for the collective as city size increases is the underlying driving force for the continued explosion of urbanization across the planet.

332

September 29, 2023

September 29, 2023

332

So what is the common unifying factor that transcends these differences and underlies this surprising structural and dynamical similarity? I have already strongly hinted at the answer: the great commonality is the universality of social network structures across the globe. Cities are people, and to a large extent people are pretty much the same all over the world in how they interact with one another and how they cluster to form groups and communities. We may look different, dress differently, speak different languages, and

333

have different belief systems, but to a large extent our biological and social organization and dynamics are remarkably similar. After all, we are all human beings sharing pretty much the same genes and the same generic social history. And no matter where we live on the planet, all of us emerged only relatively recently from being mobile hunter-gatherers to becoming predominantly sedentary communal creatures. The underlying commonality that is being expressed by the surprising universality of urban scaling laws is that the structure and dynamics of human social networks are very much the same everywhere.

September 29, 2023

It is true that many other creatures, such as herding animals and especially social insects, also discovered economies of scale, but their accomplishments are relatively primitive and static compared with what humans have achieved. The power of language has allowed us to go well beyond classic economies of scale, such as our cells achieve or our hunter-gatherer predecessors attained to evolve and build on those advantages by adapting to new challenges over periods of time that are vastly shorter than typical evolutionary timescales that had hitherto been required for major innovations to be made. Ants brilliantly self-organized to evolve remarkably robust and hugely successful and sophisticated physical and social structures, but it took them millions of years to do so. Furthermore, they accomplished this more than 50 million years ago and have barely evolved beyond it since. On the other hand, once we had invented verbal language, it took us only tens of thousands of years to evolve from hunting and gathering to becoming sedentary agriculturalists—and even more remarkably, only another ten thousand years to evolve cities, become urbanists, and invent cell phones, airplanes, the Internet, quantum mechanics, and the general theory of relativity. 334

September 29, 2023

If cities are thought of solely in terms of their physicality, as 335

just buildings and roads and the multiple network systems of wires and pipes that supply them with energy and resources, then they are indeed quite analogous to organisms, manifesting similar systematic scaling laws encapsulating economies of scale. However, when humans began forming sizable communities they brought a fundamentally new dynamic to the planet beyond that of biology and the discovery of economies of scale. With the invention of language and the consequent exchange of information between people and groups of people via social networks, we discovered how to innovate and create wealth. Cities are therefore much more than giant organisms or anthills: they rely on long-range, complex exchanges of people, goods, and knowledge. They are invariably magnets for creative and innovative individuals, and stimulants for economic growth, wealth production, and new ideas.

September 29, 2023

Recall that the generic geometric and dynamical properties of biological networks that underlie quarter-power allometric scaling are: (1) they are space filling (so every cell of an organism, for instance, must be serviced by the network); (2) the terminal units, such as capillaries or cells, are invariant within a given design (so, for instance, our cells and capillaries are approximately the same as those of mice and whales); and (3) the networks have evolved to be approximately optimal (so, for instance, the energy our hearts have to use to circulate blood and support our cells is minimized in order to maximize the energy available for reproduction and the rearing of offspring). These properties have direct analogs in the infrastructural networks of cities. For example, our road and transport networks have to be space filling so that every local region of the city is serviced, just as all of the various utility lines have to supply water, gas, and electricity to all of its houses and buildings. It's also natural to extend this concept to social networks: averaged over time, each person interacts with a number of other people as well as with groups of people in the city in such a way that collectively their network of interactions fills the available "socioeconomic space." Indeed, this urban network of socioeconomic interactions constitutes the cauldron of 336 *social activity and interconnectivity that effectively defines what a city is and what its boundaries are. To be part of a city you have to be an ongoing participant in this network. And, of course, the invariant terminal units of these networks, the analogs of capillaries, cells, leaves, and petioles, are people and their houses.*

September 29, 2023

In terms of the socioeconomic dynamics of cities we can likewise ask what, if anything, is being optimized in urban social networks. This is a tough question to answer definitively, and many scholars have obliquely tried to address it from multiple points of view. 5 If we think of the city as the great facilitator of social interactions or as the great incubator for wealth creation and innovation, it is natural to speculate that its structure and dynamics evolved so as to maximize social capital by optimizing the connectivity between individuals. This suggests that social networks and the entire social fabric of cities and urban systems—that is, who is connected to whom, how much information flows between them, and the nature of their group structure—is ultimately determined by the insatiable drive of individuals, small businesses, and giant companies to always want more. Or, to put it in crass terms, that the socioeconomic machinery that we all participate in is primarily fueled by greed in both its negative and positive connotations as in the sense of the "desire for more." Given the enormous disparities in income distributions that are observed in all cities across the globe, and the apparent drive of most of us to want more despite having plenty, it's not hard to believe that greed in its various forms is an important contributor to the socioeconomic dynamics of cities. To quote Mahatma Gandhi: "The Earth provides enough to satisfy every man's needs, but not every man's greed."

338

September 29, 2023

Greed is the pejorative image of this insatiable desire for

339

more, but it also has an extremely important, positive flip side. Metaphorically, it is the social analog of the evolutionary biological drive of animals, including us, to maximize their metabolic power relative to their size. As was discussed in chapter 3, this can be thought of as derivative from the principle of natural selection and underlies the allometric scaling laws that permeate biology. The extension of the concept of the survival of the fittest to the social and political domain has led many thinkers to the controversial concept of Social Darwinism, whose roots go back to Malthus. Regardless of its validity, this idea has been sadly misrepresented, abused, and misused by politicians and social thinkers, sometimes with devastating consequences, to support all sorts of extreme views ranging from eugenics and racism to rampant laissez-faire capitalism. The desire for more can apply to many things beyond wealth and material assets. It is a hugely powerful force in society that poses enormous moral, spiritual, and psychological challenges at both the individual and collective levels. The desire to succeed, whether in sports, business, or academia—to run the fastest, have the most creative company, or generate the most profound and insightful idea—has been a major underlying societal dynamic that has been instrumental in bringing us the extraordinary standard of living and quality of life many of us are privileged to enjoy. At the same time we have tempered our rampant materialistic greed by evolving altruistic and philanthropic behavior that has been integrated into our sociopolitical structures to protect us from excesses.

September 29, 2023

The actual self-similarity of cities more closely reflects the organically evolved hierarchical network structures of transport and utility systems than the rigid hexagonal crystalline structures of Christaller. The city is not a top-down engineered machine dominated by straight lines and classic Euclidean geometry, but rather is much more akin to an organism with its crinkly lines and fractal-like shapes typical of a complex adaptive system—which it is. This is clear from just a casual look at the growth patterns of a typical city with its ever-expanding filigreed infrastructural

344

network pattern reminiscent of the growth pattern of a bacterial colony, as illustrated here. A careful mathematical analysis of such patterns shows that cities are, in fact, approximate self-similar fractals much like biological organisms or geographical coastlines. For example, if the length of the perceived boundary of a city is measured at different resolutions, analogous to what Lewis Fry Richardson did for coastlines, and these are plotted logarithmically, then approximate straight lines result whose slope is the conventional fractal dimension of the city boundary.

September 29, 2023

As I explained earlier, fractal dimensions are a measure of an object's degree of crinkliness, which some interpret as a measure of its complexity. Stimulated by the explosive interest in fractals and the incipient development of a science of complexity in the 1980s, the distinguished urban geographer Michael Batty carried out extensive statistical analyses on cities to measure their fractal dimensions. 6 345

See the reference.

September 29, 2023

In addition to providing a metric for comparing the complexity of different cities, perhaps one of the more interesting uses of a fractal dimension is as a diagnostic barometer of the health of a city. Typically, the fractal dimension of a healthy robust city steadily increases as it grows and develops, reflecting a greater complexity as more and more infrastructure is built to accommodate an expanding population engaging in more and more diverse and intricate activities. But conversely, its fractal dimension decreases when it goes through difficult economic times or when it temporarily contracts. 345

September 29, 2023

September 29, 2023 348

The leading advocate for developing the concept of the fractal city and integrating ideas from complexity theory into traditional urban analysis and planning has been Mike Batty, who runs the Centre for Advanced Spatial Analysis (CASA) at University College London. His work has focused primarily on computer models of the physicality of cities and urban systems. He is enthusiastic about the concept of cities as complex adaptive systems and has consequently become a major proponent of developing a science of cities . His vision is a little different from mine and is summarized in his recent book The New Science of Cities, which emphasizes the more phenomenological traditions of the social sciences, geography, and urban planning as against the more analytic, mathematical traditions of physics based on underlying principles that I've been articulating. 7 Ultimately, both approaches are needed if we are to accomplish the huge challenge of understanding cities. 348

September 29, 2023

A city is not just the aggregate sum of its roads, buildings, pipes, and wires that comprise its physical infrastructure, nor is it just the cumulative sum of the lives and interactions of all of its citizens, but rather it's the amalgamation of all of these into a vibrant, multidimensional living entity. A city is an emergent complex adaptive system resulting from the integration of the flows of energy and resources that sustain and grow both its physical infrastructure and its inhabitants with the flows and exchange of information in the social networks that interconnect all of its citizenry. The integration and interplay of these two very different networks magically gives rise to increasing economies of scale in its physical infrastructure and simultaneously to a disproportionate increase in social activity, innovation, and economic output. 349

September 29, 2023

September 29, 2023 [350](350)

Over the past twenty years an emerging subfield of network science has blossomed in its own right, leading to a deeper understanding of both the phenomenology of networks in general and also the underlying mechanisms and dynamics that generate them. 8 The subject of network science covers an enormous range of topics including classic community organizations, criminal and terrorist networks, networks of innovation, ecological networks and food webs, health care and disease networks, and linguistic and literary networks. Such studies have provided important insights into a broad range of important societal challenges including devising the most effective strategies for attacking pandemics, terrorist organizations, and environmental issues, for enhancing and facilitating innovative processes, and for optimizing social organizations. A lot of this fascinating work has been carried out or stimulated by many of my colleagues associated with the Santa Fe Institute.

350

September 29, 2023

Steve Strogatz is an eclectic applied mathematician at Cornell University who uses ideas from nonlinear dynamics and complexity theory to analyze and explain a broad range of fascinating problems. For example, he has done some lovely work showing how crickets, cicadas, and fireflies synchronize their behaviors and more recently extended it to show why London's Millennium Bridge was dysfunctional. 11 This latter problem has some interesting lessons for the science of cities, and I want to digress to explain it.

352

September 29, 2023

September 29, 2023 353

The question as to how and why good people become evil and do bad things—the human analog to Job's dilemma as to why God allows bad things to happen to good people—has been a fundamental paradox of human behavior since we evolved social consciousness. The question of man's place in relationship to himself—the continuing moral and ethical dilemma of good versus evil—can be viewed as a companion question to that of man's place in relationship to the universe. These are both central issues of human existence that have dominated human contemplation from the time Homo sapiens became conscious, spawning numerous religions, cultures, and philosophies. 357

September 29, 2023

The provocative work of Milgram and Zimbardo strongly suggests that the conundrum as to why good people can do very bad things originates in peer pressure situations, fear of rejection, and a desire to be part of a group where power and control are conferred on individuals by authority. Zimbardo has become an articulate and vocal advocate for the recognition that this powerful dynamic, which seems to be built into our psyches independent of cultural origins and which has wreaked horrors over the centuries, be explicitly recognized and addressed rather than resorting to our instinctual tendency to put the blame simplistically on individual "bad apples," national characteristics, or cultural norms. 357

September 29, 2023

This relationship between an individual and his or her community naturally led him to the broader issue of the psychological dimension of urban life. In 1970 he published a provocative paper in Science titled "The Experience of Living 358 *in Cities," which laid the foundations for the newly developing field of urban psychology and which became the central focus of his ensuing interests. 14 Milgram was very much struck by the psychological harshness of life in the big city. The general perception was that outside of the local environment an individual generally went about his or her business avoiding interaction and involvement, rarely acknowledging other people or events that might engender participation or commitment. So much so that most people are very loath to intervene or even call for help when they are witness to crime, violence, or some other crisis event. He devised a sequence of innovative experiments to investigate the apparent lack of trust, the greater sense of fear and anxiety, and the general lack of civility and politeness that seem to characterize life in large metropolises relative to that in small towns. For instance, he had individual investigators ring doorbells saying that they had misplaced the address of a nearby friend and would like to use the phone. He found that there was a large increase by factors of three to five in the proportion of entries allowed into homes in a small town relative to a large city. Furthermore, 75 percent of the city respondents answered the inquiry by shouting through closed doors or by peering out through peepholes, whereas in small towns 75 percent opened the door.*

September 29, 2023

September 29, 2023 359

359

In conceptualizing this dark psychosocial side of urban life, Milgram borrowed the term "overload" from electrical circuit and systems science theory. In big cities we are continually bombarded with so many sights, so many sounds, so many "happenings," and so many other people at such a high rate that we are simply unable to process the entire barrage of sensory information. If we tried to respond to every stimulus, our cognitive and psychological circuitry would break down and, in a word, we would blow a fuse just like an overloaded electrical circuit. And sadly, some of us do.

Milgram suggested that the kinds of "antisocial" behaviors we perceive and experience in large cities are in fact adaptive responses for coping with the sensory onslaught of city life, implying that without such adaptations we'd all blow our fuses.

September 29, 2023

Understanding and deconstructing the hierarchy of social group structures has been a major focus of attention in sociology and anthropology for more than fifty years, but it was only in the last twenty or so that some of their quantitative features have become apparent. Some of this has been driven by the work of the evolutionary psychologist Robin Dunbar and his collaborators, who proposed that an average individual's entire social network can be deconstructed into a hierarchical sequence of discrete nested clusters whose sizes follow a surprisingly regular pattern. 15 The size of the group at each level systematically increases as one progresses up the hierarchy from, say, family to city, while the strength of the bonding between people within the groups systematically decreases. So, for example, most people have a very strong connection with their immediate family members but only a very weak one with the bus driver or members of the city council.

361

September 29, 2023

Partly inspired by work on social primate communities and partly by anthropological studies of human societies ranging from hunter-gatherers to modern corporations, Dunbar discovered that this hierarchy appeared to have a surprisingly regular mathematical structure obeying very simple scaling rules reminiscent of self-similar fractal-like behavior. He and his collaborators found that at the lowest level of the hierarchy the number of people with whom the average individual has his or her strongest relationships is, at any one time, only about five. These are the people we are closest to and care most deeply about; they are usually family members—parents, children, or spouses—but they

361

could be extremely close friends or partners. In surveys designed to measure the size of this core social group one of its defining characteristics was "the set of individuals from whom the respondent would seek personal advice or help in times of severe emotional and financial distress." The next level up contains those you usually refer to as close friends with whom you enjoy spending meaningful time and might still turn to in time of need even if they are not on as intimate terms with you as your inner circle. This typically comprises around fifteen people. In the level above this are people you might still call friends though you would only rarely invite them to dinner but would likely invite them to a party or gathering. This might consist of coworkers, neighbors down the street, or relatives you don't see very often. There are typically about fifty such people in this group. The next level pretty much defines the limit of your social horizon as far as personal interactions are concerned and consists of people you might refer to as "casual friends"—you know their names and remain in social contact with them. This group typically comprises about 150 people. It is this number that is usually referred to as the Dunbar number that has gained a certain degree of attention in the popular media.

September 29, 2023

362

You will notice that the sequence of numbers that quantify the magnitude of these successive levels of the group hierarchy—5, 15, 50, 150—are sequentially related to each other by an approximately constant scaling factor of about three. This regularity is the familiar fractal-like pattern we saw not only in the network hierarchy of our own circulatory and respiratory systems but also in transport patterns in cities. In addition to the actual flows in these networks, the major geometric difference between them is the value of the branching ratio— the number of units, people in this case, at one level of the hierarchy relative to the next. There is evidence that in social networks this pattern with a branching ratio of three persists beyond the 150 level to groups having sizes of approximately 500, 1,500, and so on. The precise value of these numbers should not be taken too seriously because there is significant variance in the data. The important point for our purposes is that viewed through *a coarse-grained lens, social networks exhibit an approximate fractal-like pattern and that this seems to hold true across a broad spectrum of different social organizations. Even though this pattern remains approximately static, the individual members of the network may change with time or may move from one level to another as you become closer or more distant in your relationship with them. For instance, a parent may move out of your inner circle and be replaced by a spouse or a close friend, or you might meet someone casually at a party who subsequently becomes part of your 150. Regardless of such changes, the generic structure of the network with four to six people comprising your core group, and the nested group structure whose discrete size increases by factors of about three up to 150 or so, remains intact.*

September 29, 2023

September 29, 2023 364

The number of around 150 represents the maximum number of individuals that a typical person can still keep track of and consider casual friends and therefore members of his or her ongoing social network. Consequently, this is the approximate size of a group in which all individuals still know one another sufficiently well for the group to remain coherent and maintain ongoing social relationships. Dunbar found many such examples of functioning social units whose size congregated around this magic number, ranging from bands of hunter-gatherers to army companies in the Roman Empire, in sixteenth-century Spain, and in the twentieth-century Soviet Union. He speculated that this apparent universality has its origins in the evolution of the cognitive structure of the brain: we simply do not have the computational capacity to manage social relationships effectively beyond this size. This suggests that increasing the group size beyond this number will result in significantly less social stability, coherence, and connectivity, ultimately leading to its disintegration. For situations where group

364

identity and cohesiveness are perceived as central for the group to function successfully, recognizing this limitation and the broader implications of social network structure is clearly of great importance. This is especially true in situations where stability, knowledge of other individuals, and social relationships are integral to performance. Companies, the military, government administrations and bureaucracies, universities, and research organizations are among the many institutions where this kind of information and this way of thinking could potentially be beneficial in improving performance, productivity, and the general well-being of all members of the organization.

September 29, 2023

connection between brain size and the ability to form social groups is called the social brain hypothesis. Dunbar went much further by suggesting that this relationship is causal in that human intelligence evolved primarily as a response to the challenge of forming large and complex social groups rather than the usual explanation that it is a direct consequence of meeting ecological challenges. 16 Regardless of causation, he used the correlation with brain size to estimate the number 150 as the idealized size of analogous human social groups.

365

September 30, 2023

September 29, 2023

365

Unlike in biology, surprisingly little attention had been paid to scaling laws for cities, urban systems, or companies prior to our investigations. This may be because few people suspected that such complex, historically contingent man-made systems as these would manifest any sort of systematic quantitative regularity. In addition, there is much less of a tradition of this kind of modeling and of confrontation of

366

theory with data in urban studies than in either biology or physics. There was, however, one major exception to this and that is a famous scaling law known as Zipf's law for the ranking of cities in terms of their population size. This is shown graphically in Figure 39. It's an intriguing observation: in its simplest form, it states that the rank order of a city is inversely proportional to its population size.

September 30, 2023

367

Zipf's law owes its name to the Harvard linguist George Kingsley Zipf, who popularized it through his fascinating book Human Behavior and the Principle of Least Effort, published in 1949. 17 He first enunciated his law in 1935, not for cities but for the frequency of word use in languages. As originally stated, the law says that the frequency of occurrence of any word in a corpus of written text such as all of Shakespeare's plays, the Bible, or even this book is inversely proportional to its rank in the frequency table. Thus, the most frequent word occurs approximately twice as often as the second most frequent, three times as often as the third, and so on, as shown in Figure 40. For example, analysis of English texts shows that the most frequent word is, not surprisingly, "the," which accounts for roughly 7 percent of all words used, while the second-place word "of" accounts for about half as many, namely 3.5 percent of all words, followed by "and" with about a third as many, namely 2.3 percent, and so on. Even more mystifying is that this same law is valid across an astonishing array of examples, including the rank size distributions of ships, trees, sand particles, meteorites, oil fields, file sizes of Internet traffic, and many more. Figure 41 shows how the distribution of company sizes follows this law. Given its amazing generality and some of its implications, Zipf's law has taken on a curious mystique among many researchers and writers whose imagination has been caught by its surprising simplicity. Zipf and many others following him have pondered its origins, but no generally agreed-upon explanation has yet emerged.

September 30, 2023

Nevertheless, the fact that these Zipf-like distributions are found across such a diverse set of phenomena suggests that they express some general systemic property that is independent of the character and detailed dynamics of the specific entities under consideration. This is reminiscent of the ubiquitous generality of the "bell-curve" distribution used for describing statistical variations around some average value. Technically this is called a Gaussian or normal distribution and arises mathematically whenever a sequence of events or entities, whatever they are, are randomly distributed, uncorrelated, and independent of one another. So, for example, the average height of men in the United States is about five feet ten inches (1.77 meters) and the frequency distribution of their heights around this mean value—that is, how many men there are of a given height—closely follows a classic Gaussian bell-curve distribution. This tells us what the likelihood is of someone being a specific height. Gaussian statistics are used across all of science, technology, economics, and finance to assign statistical probabilities of various events happening such as in weather forecasting or for drawing conclusions from polling surveys. However, it is sometimes forgotten that these probability estimates are based on the assumption that individual "events," whether comparing today's temperature with the historical record, or one person's height with another's, are independent of one another and can therefore be considered uncorrelated. The canonical Gaussian bell curve is so pervasive and so taken for granted that it is generally assumed without much thought that that's how "everything" is distributed. Consequently, power law distributions like those of Zipf and Pareto barely saw the light of day. It was natural to assume that cities, incomes, and words would be randomly distributed following the classic bell curve. If that were so, it would predict that there are far fewer very large cities, very large companies, and very rich people, and far fewer common words than there actually are, because all of these follow power law distributions that have much longer tails, meaning that there are many more rare events than would be expected had they been random and obeyed Gaussian statistics. This difference is sometimes characterized by saying that power laws have "fat tails." 371

Clearly, words in a book are correlated and not random because they have to form meaningful sentences, as are cities, because they are part of a unified urban system. It's therefore not so surprising that these distributions are not Gaussian. Many of the most interesting phenomena that we have touched upon fall into this category, including the occurrence of disasters such as earthquakes, financial market crashes, and forest fires. All of these have fat-tail distributions with many more rare events, such as enormous earthquakes, large market crashes, and raging forest fires, than would have been predicted by assuming that they were random events following a classic Gaussian distribution. Furthermore, because these are approximately self-similar processes, the same dynamics occur at all scales. Thus the same generic mechanism that leads to a small adjustment in a financial market is at play when that market suffers a major crash. This is in marked contrast to the random nature implicit in Gaussian statistics, where events at different scales are presumed to be independent and uncorrelated. Ironically, economists and financial analysts traditionally use Gaussian statistics in their analysis, ignoring the preponderance of fat tails and therefore correlations. Caveat emptor!

September 30, 2023

The two dominant components that constitute a city, its physical infrastructure and its socioeconomic activity, can both be conceptualized as approximately self-similar fractal-like network structures. Fractals are often the result of an evolutionary process that tends toward optimizing specific features, such as ensuring that all cells in an organism or all people in a city are supplied by energy and information, or maximizing efficiency by minimizing transportation times or times for accomplishing tasks with minimal energy. Less obvious is what is being optimized in social networks. There is, for instance, no satisfactory explanation based on underlying principles for understanding the hierarchical structure observed by Dunbar, or for the origin of his sequence of numbers. Even if the social brain hypothesis is correct, it doesn't explain either the origin of the fractal nature of social groups or where the number 150 comes from. 373

There are hints that such generic properties follow from the conjecture made earlier that self-interest—that is, the desire of all individuals and companies to maximize their assets and income—coupled with the concept of maximal filling of social space are the underlying driving forces. There is certainly a great deal yet to be done in constructing a quantitative theory of social networks, and many exciting challenges await future investigation.

There is certainly a great deal yet to be done in constructing a quantitative theory of social networks, and many exciting challenges await future investigation.

September 30, 2023

Consequently, all of the socioeconomic metrics that reflect such activity and which were discussed earlier when reviewing urban scaling laws are proportional to the number of links, or interactions that take place, between people within the city. If it were possible for everyone to interact with everyone else so that over the period of a year, for example, each person connects meaningfully with every other person in the city, then the total number of interactions between people could be easily calculated from a simple formula: it is the total number of people in the city multiplied by the total number of people that any individual can connect to in the city. This last number is just the total number of people minus one. For instance, if you are one of ten people in a group, you can connect to only nine others. In addition, you have to divide the answer by two because you don't count the link between you and another person as different from the link between that other person and you. They are symmetric and one and the same. Thus the total possible number of pair-wise links between people in a city is given by the total number of people in the city multiplied by the total number minus one—all divided by two. This may seem a bit of a mouthful but it's actually quite simple, so let me explain with some examples. 374

September 30, 2023

The lesson is clear: the number of links between people increases much faster than the increase in the number of people in the group and, to a very good approximation, is given by just one half of the square of the number of people in the group .

376

September 30, 2023

This simple nonlinear quadratic relationship between the maximal number of links between people and the size of the group has all sorts of very interesting social consequences. For instance, my wife, Jacqueline, particularly enjoys dinner parties if a single conversation can be sustained by the entire group, so she is reluctant to participate in dinner parties larger than six. With six people there are 6 × 5 ÷ 2 = 15 possible pair-wise independent conversations that have to be "suppressed" for a single collective one to emerge and be maintained. This is just about possible, and it's tempting to speculate that it's because the number of other guests, five, corresponds to Dunbar's number for the group size of an average individual's inner circle. With ten at the table, there are a whopping forty-five such dyadic possibilities, which inevitably leads to a balkanization of the group as it disintegrates into two, three, or more separate conversations. Of course, many people prefer this modality but it is worth remembering that if you want a certain kind of group intimacy, having more than about six people is going to make it quite a challenge.

376

Dunbar's number 5.

September 30, 2023

Now let's go back and take a look at how this plays out in an entire city. If it were possible for everyone to interact meaningfully with everyone else as in one great big happy family, then the above argument would imply that all socioeconomic metrics should scale with the square of the population size. This would mean an exponent of 2, which is

377

certainly superlinear (it's bigger than 1), but significantly larger than 1.15. However, this represents the extreme and totally unrealistic case where the entire population is in a frenzied state of continuous and complete interaction with itself, being churned about much like raisins or nuts in cake dough driven by a super-high-speed electric mixer. This is clearly impossible—and certainly not desirable. Even in a modest-size city of only 200,000 people there are roughly 20 billion possible relationships, and even if each person devoted just one minute a year to each relationship, they would have to spend their entire waking life relating to other people, leaving no time for anything else. Imagine extending that to a New York or Tokyo. There is also the constraint of the Dunbar number, according to which we even have difficulty sustaining any sort of meaningful relationship with more than about 150 people, let alone a couple of hundred thousand or several million. It is this restriction to a relatively small number of interactions that drives the superlinear exponent to be significantly smaller than its maximum possible value of 2.

September 30, 2023

This exercise shows that there is a natural explanation for why social connectivity and therefore socioeconomic quantities scale superlinearly with population size. Socioeconomic quantities are the sum of the interactions or links between people and therefore depend on how correlated they are. In the extreme case when everyone is interacting with everyone else we saw that this leads to a superlinear power law whose exponent is 2. However, in reality there are significant constraints on the intensity and magnitude of how many people an individual can interact with, and this drastically reduces the value of the exponent to be less than 2.

378

September 30, 2023

September 30, 2023

378

In this spirit, individuals are considered to be the "invariant terminal units" of social networks, meaning that on average each person operates in roughly the same amount of social and physical space in a city.

380

September 30, 2023

Recall that the physical space in which we operate is spanned by space-filling fractal networks, such as roads and utility lines that service infrastructural terminal units such as houses, stores, and office buildings where we reside, work, and interact, and between which we also have to move. The integration of these two kinds of networks, namely, the requirement that socioeconomic interaction represented by space-filling fractal-like social networks must be anchored to the physicality of a city as represented by space-filling fractal-like infrastructural networks , determines the number of interactions an average urban dweller can sustain in a city. And as discussed earlier, it is this number that determines how socioeconomic activity scales with population size. The biological metaphor of the city as a living organism derives primarily from its being perceived in terms of its physicality. This is most apparent in the networks that carry energy and resources in the form of electricity, gas, water, cars, trucks, and people, and it is this component of cities that is the close analog to the networks that proliferate in biology such as our cardiovascular and respiratory systems or the vasculature of plants and trees. Combining the ideas of space filling, invariant terminal units, and optimization (minimizing travel times and energy use, for example) results in these networks also being fractal-like with infrastructural metrics scaling as power laws with sublinear exponents indicative of economies of scale obeying the 15 percent rule.

380

September 30, 2023

When these constraints on the mobility and physical

381

interaction space of people in cities are imposed on the structure of social networks, an important and far-reaching result emerges: the number of interactions with other people in a city that an average person maintains scales inversely to the way that the degree of infrastructure scales with city size. In other words, the degree to which the scaling of infrastructure and energy use is sublinear is predicted to be the same as the degree to which the scaling of the number of an average individual's social interactions is superlinear . Consequently, the exponent controlling social interactions, and therefore all socioeconomic metrics—the universal 15 percent rule for how the good, the bad, and the ugly scale with city size—is bigger than 1 (1.15) to the same degree that the exponent controlling infrastructure and flows of energy and resources is less than 1 (0.85), as observed in the data. Pictorially, the degree to which all of the slopes in Figures 34–38 exceed 1 is the same as the degree to which they are less than 1 in Figure 33.

In other words, the degree to which the scaling of infrastructure and energy use is sublinear is predicted to be the same as the degree to which the scaling of the number of an average individual's social interactions is superlinear.

September 30, 2023

In this network scaling sense, it is no accident therefore that the physical and the social mirror each other, so much so that we can think of the physical city —with its networks of buildings, roads, and electrical, gas, and water lines—as an inverse nonlinear representation of the socioeconomic city — with its networks of social interactions. The city is indeed the people. The approximate 15 percent increase in social interactions and therefore in socioeconomic metrics such as income, patents, and crime generated with every doubling of city size can be interpreted as a bonus, or payoff, arising from the 15 percent savings in physical infrastructure and energy use. The systematic increase in social interaction is the essential driver of socioeconomic activity in cities: wealth creation, innovation, violent crime, and a greater sense of buzz and opportunity are all propagated and enhanced through social networks and greater interpersonal 382 *interaction. But this can equally well be interpreted in a complementary way by viewing cities as catalytic facilitators or crucibles for social chemistry in which the increase in social interactions enhances creativity, innovation, and opportunity whose dividend is an increase in infrastructural economies of scale. Just as raising the temperature of a gas or liquid increases the rate in the number of collisions between molecules, so increasing the size of a city increases the rate and number of interactions between its citizens. Metaphorically speaking, increasing the size of a city can therefore be thought of as raising its temperature. In this sense, New York, London, Rio, and Shanghai are truly hot cities, especially compared with Santa Fe where I live, and the proverbial image of a "melting pot," originally applied to New York City, is an apt expression of this metaphor.*

September 30, 2023

It is perhaps not so surprising that there is a correlation between increased social interaction, socioeconomic activity, and greater economies of scale. What is surprising, however, is that this pivotal interrelationship follows such simple mathematical rules that can be expressed in an elegant universal form: the sublinearity of infrastructure and energy use is the exact inverse of the superlinearity of socioeconomic activity . Consequently, to the same 15 percent degree, the bigger the city the more each person earns, creates, innovates, and interacts—and the more each person experiences crime, disease, entertainment, and opportunity—and all of this at a cost that requires less infrastructure and energy for each of them. This is the genius of the city. No wonder so many people are drawn to them. 383

September 30, 2023

8: CONSEQUENCES AND PREDICTIONS: From Mobility and the Pace of Life to Social Connectivity, Diversity, Metabolism, and Growth

October 1, 2023 390

One of the great ironies in all of these marvelous inventions (with the possible exception of the gruesome weapons of destruction) is that they all promised to make life easier and more manageable and therefore give us more time. Indeed, when I was a young man, pundits and futurists were speaking of the glorious future anticipated from such time-saving inventions, and a topic that was much discussed was what we would do with all free time that would now be at our disposal. With cheap energy available from nuclear sources and these fantastic machines doing all our manual and mental labor, the workweek would be short and we would have large swaths of time to really enjoy the good life with our families and friends, a little like the boring privileged lives of aristocratic ladies and gentlemen of previous centuries. In 1930 the great economist John Maynard Keynes wrote: For the first time since his creation man will be faced with his real, his permanent problem—how to use his freedom from pressing economic cares, how to occupy the leisure, which science and compound interest will have won for him, to live wisely and agreeably and well. And in 1956, Sir Charles Darwin, grandson of the Charles Darwin, wrote an essay on the forthcoming Age of Leisure in the magazine New Scientist in which he argued: Take it that there are fifty hours a week of possible working time. The technologists, working for fifty hours a week, will be making inventions so the rest of the world need only work twenty-five hours a week. The more leisured members of the community will have to play games for the other twenty-five hours so they may be kept out of mischief. . . . Is the majority of mankind really able to face the choice of leisure enjoyments, or will it not be necessary to provide adults with something like the compulsory games of the schoolboy? 391

October 1, 2023

October 1, 2023 392

Zahavi's fascinating observations made a powerful impression on the Italian physicist Cesare Marchetti, who was a senior scientist at the International Institute for Applied Systems Analysis (IIASA) in Vienna. IIASA has been a major player in questions of global climate change, environmental impact, and economic sustainability, and this is where Marchetti's interests and contributions have mostly been. He became intrigued by Zahavi's work and in 1994 published an extensive paper elaborating on the approximate invariance of daily commute times and promoted the idea that the true invariant is actually overall daily travel time, which he called exposure time . 3 So even if an individual's daily commute time is less than an hour, then he or she instinctively makes up for it by other activities such as a daily constitutional walk or jog. In support of this Marchetti wryly remarked, "Even people in prison for a life sentence, having nothing to do and nowhere to go, walk around for one hour a day, in the open." Because walking speed is about 5 kilometers an hour, the typical extent of a "walking city" is about 5 kilometers across (about 3 miles), corresponding to an area of about 20 square kilometers (about 7 square miles). According to Marchetti, "There are no city walls of large, ancient cities (up to 1800), be it Rome or Persepolis, which have a diameter greater than 5km or a 2.5km radius. Even Venice today , still a pedestrian city, has exactly 5km as the maximum dimension of the connected center. " With the introduction of horse tramways and buses, electric and steam trains, and ultimately automobiles, the size of cities could grow but, according to Marchetti, constrained by the one-hour rule. With cars able to travel at 40 kilometers an hour (25 mph), cities, and more generally metropolitan areas, could expand to as much as 40 kilometers or 25 miles across, which is typical of the catchment area for most large cities. This corresponds to an area of about 12 hectares or 450 square miles, more than fifty times the area of a walking city. This surprising observation of the approximately one-hour invariant that communal human beings have spent traveling each day, whether they lived in ancient Rome, a medieval town, a Greek village, or twentieth-century New 395

York, has become known as Marchetti's constant , even though it was originally discovered by Zahavi. As a rough guide it clearly has important implications for the design and structure of cities. As planners begin to design green carless communities and as more cities ban automobiles from their centers, understanding and implementing the implied constraints of Marchetti's constant becomes an important consideration for maintaining the functionality of the city.

October 1, 2023

October 1, 2023 400

People across the globe behave in pretty much the same way regardless of history, culture, and geography. Thus without having to resort to any fancy mathematical theory, this idea predicts that the number of interactions between people in cities should scale with city size in the same way that all of the diverse socioeconomic quantities scale, namely, as a superlinear power law with an exponent of around 1.15 regardless of the urban system. In other words, the systematic 15 percent increase in socioeconomic activity with every doubling of city size, whether in wages and patent production or crime and disease, should track a predicted 15 percent increase in the interaction between people. 404

October 1, 2023

October 1, 2023 407

The size of an average individual's modular cluster of acquaintances who interact with one another is an approximate invariant—it doesn't change with city size. For example, the size of the "extended family" of an average individual living in Lisbon, which has more than 500,000 408 *inhabitants, is no larger than that of an average individual living in the small town of Lixa with less than 5,000 inhabitants. So even in large cities we live in groups that are as tightly knit as those in small towns or villages. This is a bit like the invariance of the Dunbar numbers I talked about in the previous chapter and, like those, probably reflects something fundamental about how our neurological structure has evolved to cope with processing social information in large groups. There is, however, an important qualitative difference in the nature of these modular groups in villages relative to those in large cities. In a real village we are limited to a community that is imposed on us by sheer proximity resulting from its small size, whereas in a city we are freer to choose our own "village" by taking advantage of the much greater opportunity and diversity afforded by a greater population and to seek out people whose interests, profession, ethnicity, sexual orientation, and so on are similar to our own. This sense of freedom provided by a greater diversity across many aspects of life is one of the major attractions of urban life and a significant contributor to rapidly increasing global urbanization.*

October 1, 2023

The image of the city as a large vat in which people are being continually churned, blended, and agitated can be viscerally felt in any of the world's great cities. It is most apparent in the continuous, and sometimes frenetic, movement of people in downtown and commercial areas in what often appears to be an almost random motion much like molecules in a gas or liquid. And in much the same way that bulk properties of gases or liquids, such as their temperature, pressure, color, and smell, result from molecular collisions and chemical reactions, so the properties of cities emerge from social collisions and the chemistry of and between people. Metaphors can be useful but sometimes they can be misleading, and this is one of those instances. For despite appearances, the motion of people in cities is not at all like the random motion of molecules in a gas or particles in a reactor. Instead, it is overwhelmingly systematic and directed. Very few journeys are random. Almost all, regardless of the form of conveyance, involve willful travel

410

from one specific place to another: mostly from home to work, to a store, to a school or cinema, and so forth . . . and back again. Furthermore, most travelers seek the fastest and shortest route, one that takes the least time and traverses the shortest distance. Ideally, this would mean that everyone would like to travel along straight lines but, given the obvious physical constraints of cities, this is impossible. There is no choice but to follow the meandering roads and rail lines, so, in general, any specific journey involves following a zigzagging route. However, when viewed at a larger scale through a coarse-grained lens by averaging over all journeys for all people over a long enough period of time, the preferred route between any two specific locations approximates a straight line. Loosely speaking this means that on average people effectively travel radially along the spokes of circles whose center is their specific destination, which acts as a hub. With this assumption it is possible to derive an extremely simple but very powerful mathematical result for the movement of people in cities. Here's what it says: Consider any location in a city; this could be a "central place" such as a downtown area or street, a shopping mall or district, but it could just as well be some arbitrary residential area such as where you live. The mathematical theorem predicts how many people visit this location from any distance away and how often they do it. More specifically, it states that the number of visitors should scale inversely as the square of both the distance traveled and the frequency of visitation .

October 1, 2023

Mathematically, an inverse square law is just a simple version of the kinds of power law scaling we've been discussing throughout the book. In that language the prediction of movement in cities can be restated as saying that the number of people traveling to a specific location scales with both the distance traveled and the visitation frequency as a power law whose exponent is −2. Thus if the number of travelers is

411

plotted logarithmically against either the distance traveled keeping the visitation frequency fixed, or vice versa, against the visitation frequency keeping the travel distance fixed, a straight line should result in both cases with the same slope of −2 (recall that the minus sign simply means that the line slopes downward). I should emphasize that as with all of the scaling laws, an average over a long enough period of time, such as six months or a year, is presumed, in order to smooth over daily fluctuations, or the differences between weekdays and weekends. As can be readily seen from Figure 47, these predictions are spectacularly confirmed by the data. Indeed, the observed scaling is remarkably tight, with slopes in excellent agreement with the prediction of −2. Particularly satisfying is that the same predicted inverse square law is observed across the globe in diverse cities with very different cultures, geographies, and degrees of development: we see an identical behavior in North America (Boston), Asia (Singapore), Europe (Lisbon), and Africa (Doha). Furthermore, when each of these metropolitan areas is deconstructed into specific locations, the same inverse square law is manifested at each of these within the city, as shown, for example, in Figures 48 and 49 for a sampling of locations in both Boston and Singapore. Let me give a simple example to illustrate how the theorem works. Suppose that on average 1,600 people visit the area around Park Street, Boston, from 4 kilometers away once a month. How many people visit there from twice as far away (8 km) with the same frequency of once a month? The inverse square law tells us that $\frac{1}{4}$ *(=* $\frac{1}{2}$*2) as many make the visit, so only 400 people (*$\frac{1}{4}$ *× 1,600) visit Park*

Street from 8 kilometers away once a month. How about from five times as far away, 20 kilometers? The answer is 1/25 (= 1/5 2) as many, which is just 64 people (1/25 × 1,600) visiting once a month. You get the idea. But there's more: you can likewise ask what happens if you change the frequency of visitation. For instance, suppose we ask how many people visit Park Street from 4 kilometers away but now with a greater frequency of twice a month. This also obeys the inverse square law so the number is $\frac{1}{4}$ (= $\frac{1}{2}$2) as many, namely 400 people. And similarly, if you ask how many people visit there from the same distance of 4 kilometers away but five times a month, the answer is 64 people (1/25 × 1,600). Notice that this is the same number that visit Park Street from five times as far away (20 km) with a frequency of just once a month. Thus the number of people visiting from 4 kilometers away five times a month is the same as the number visiting from five times farther away (20 km) once a month (64 in both cases of our specific example). This result does not depend on the specific numbers I chose for the illustration. It is an example of an amazing general symmetry of mobility: if the distance traveled multiplied by the frequency of visits to any specific location is kept the same, then the number of people visiting also remains the same . In our example we had 4 kilometers × 5 times a month = 20 in the first case and 20 kilometers × once a month = 20 in the second. This invariance is valid for any visiting distance and for any frequency of visitation to any area in any city. These predictions are verified by the data and manifested in the various graphs of Figures 48 and 49 where you can explicitly see that the pattern of visitation remains unchanged when the product of distance times frequency has the *same value.*

In this paragraph exist very important information about how busy people is into determine area from a distance away.

October 1, 2023

October 1, 2023 414

The analysis of these massive data sets of mobile phone calls used to test the theory was brilliantly carried out by a Swiss engineer, Markus Schläpfer, working with the Hungarian physicist Michael Szell, two of the bright young postdocs hired by Carlo Ratti at MIT. Markus later joined us at the Santa Fe Institute in 2013, where we began this particular collaboration. Of the many projects he worked on, a particularly interesting one was with Luis on analyzing how the heights and volumes of buildings relate to city size. Markus has since moved on to the prestigious ETH in Zurich, his hometown, where he is engaged with a large collaborative program called the Future Cities Lab, which is based in Singapore and supported by their government. 416

October 1, 2023

This framework was used as a point of departure to give a science-based measure of performance. Four of the six champions lifted loads they should have, given their body weights and the expectations from the scaling law. On the other hand, the middleweight over performed relative to the expectations for his size, whereas the heavyweight under performed. So despite the heavyweight lifting a greater load than anyone else, from a scientific perspective he was actually the weakest of all the champions while the middleweight was the strongest . 418

October 1, 2023

October 1, 2023 418

October 1, 2023 419

9. THE STRUCTURE OF WEALTH, INNOVATION, CRIME, AND RESILIENCE: THE INDIVIDUALITY AND RANKING OF CITIES How rich, creative, or safe can we expect a city to be? How can we establish which cities are the most innovative, the most violent, or the most effective at generating wealth? How do they rank according to economic activity, the cost of living, the crime rate, the number of AIDS cases, or the happiness of their populations? The conventional answer is to use simple per capita measures as performance indices and rank order of cities accordingly. Almost all official statistics and policy documents on wages, income, gross domestic product (GDP), crime, unemployment rates, innovation rates, cost of living indices, morbidity and mortality rates, and poverty rates are compiled by governmental agencies and international bodies worldwide in terms of both total aggregate and per capita metrics. Furthermore, well-known composite indices of urban performance and the quality of life, such as those assembled by the World Economic Forum and magazines like Fortune , Forbes , and The Economist , primarily rely on naive linear combinations of such measures. 6 Because we have quantitative scaling curves for many of these urban characteristics and a theoretical framework for their underlying dynamics we can do much better in devising a scientific basis for assessing performance and ranking cities. The ubiquitous use of per capita indicators for ranking and comparing cities is particularly egregious because it implicitly assumes that the baseline, or null hypothesis, for any urban characteristic is that it scales linearly with population size. In other words, it presumes that an idealized city is just the linear sum of the activities of all of its citizens, thereby ignoring its most essential feature and the very point

420

of its existence, namely, that it is a collective emergent agglomeration resulting from nonlinear social and organizational interactions. Cities are quintessentially complex adaptive systems and, as such, are significantly more than just the simple linear sum of their individual components and constituents, whether buildings, roads, people, or money. This is expressed by the superlinear scaling laws whose exponents are 1.15 rather than 1.00. This approximately 15 percent increase in all socioeconomic activity with every doubling of the population size happens almost independently of administrators, politicians, planners, history, geographical location, and culture.

October 2, 2023

422

My colleagues Luis, José, and Debbie carried out such an analysis for the entire U.S. urban system consisting of 360 Metropolitan Statistical Areas (MSAs) for a suite of metrics. 7 A sample of the results is presented in Figure 50), where the deviations from scaling for personal income and patent production for cities in the United States in 2003 are plotted logarithmically on the vertical axis against the rank order of each city. We called these deviations Scale-Adjusted Metropolitan Indicators (SAMIs). The horizontal axis across the center of these graphs is the line along which the SAMI is zero and there is no deviation from what is predicted from the size of the city. As can be seen, every city deviates to some extent from its expected values. Those to the left denote above -average performance, whereas those to the right denote below -average performance. This provides a meaningful ranking of a city's individuality and uniqueness beyond what is effectively guaranteed just because it's a city of a certain size. Without delving into details of this analysis, I want to make a few salient points about some of the results. First, compared with conventional per capita indicators, which place seven of the largest twenty cities in the top twenty in terms of their GDPs, our science-based metrics rank none of these cities in the top twenty. In other words, once the data are adjusted for the generic superlinear effects of population size, these cities don't fare so well. Mayors of these cities who take credit and boast that their policies have led to economic success as evidenced by their city's being near the top of the *per capita GDP rankings are therefore giving a misleading impression.*

October 2, 2023

Once a city has gained an advantage, or disadvantage, relative to its scaling expectation, this tends to be preserved over decades. In this sense, either for good or for bad, cities are remarkably robust and resilient—they are hard to change and almost impossible to kill. Think of Detroit and New Orleans, and more drastically of Dresden, Hiroshima, and Nagasaki, all of which have to varying degrees survived what were perceived as major threats to their very existence. All are actually doing fine and will be around for a very long time. A fascinating example of persistent advantage is San Jose, which includes Silicon Valley, the place everyone wants to be. It's hardly a surprise that this is a major overperformer in terms of wealth creation and innovation. But what is a surprise is that San Jose was already overperforming in the 1960s, and almost to the same degree as it is now, as graphically illustrated in Figure 51. This also demonstrates that this overperformance has been sustained and even reinforced for more than forty years despite the short-term boom and bust technological and economic cycle in 1999–2000, at the end of which the city relaxed back to its long-term basal trend. Put slightly differently: apart from a relatively small bump in the late 1990s, the continued success of San Jose was already set well before the birth of Silicon Valley. So rather than seeing Silicon Valley as generating the success of San Jose and lifting it up in the conventional socioeconomic rankings, this suggests that it was the other way around and that it was some intangible in the culture and DNA of San Jose that helped nurture the extraordinary success of Silicon Valley. 8 It takes decades for significant change to be realized. This has serious implications for urban policy and leadership because the timescale of political processes by which decisions about a city's future are made is at best just a few years, and for most politicians two years is infinity. Nowadays, their success depends on rapid returns and instant gratification in order to conform to political pressures and the demands of the electoral process. Very few mayors can afford to think in a time frame of twenty to fifty 425 *years and put their major efforts toward promoting strategies that will leave a truly long-term legacy of significant achievement.*

October 2, 2023

Safe water is progressively becoming a source of increased social friction, especially as the climate changes and produces unpredictable periods of severe drought or massive floods, both of which compromise supply and delivery systems. This is already a major issue in many developing countries and hints of it have begun to be seen even in the United States with serious problems arising in supply systems, as in Flint, Michigan, and severe water shortages occurring throughout many of the western states. 426

October 2, 2023

October 2, 2023 429

My colleagues Luis, José, and Debbie carried out an analysis of these data with our postdoc Hyejin Youn taking the lead role. Hyejin was trained in statistical physics in Seoul, South Korea, and had joined SFI to finish her doctorate. She first worked on the origin and structure of languages before joining our collaboration. She has now established herself as an expert on innovation in technology and is currently a fellow of the Institute of New Economic Thinking (INET) at Oxford University—a new program funded by the financier George Soros. As we saw in the analysis of other urban metrics, the data reveal surprisingly simple and unexpected regularities. For instance, the total number of establishments in each city regardless of what business they conduct turns out to be linearly proportional to its population size. Double the size of a city and on average you'll find twice as many businesses. The proportionality constant is 21.6, meaning that there is approximately one establishment for about 431 *every 22 people in a city, regardless of the city size . Or to put it slightly differently, on average a new workplace is created each time the population of a city increases by just 22 people, whether in a small town or a large metropolis. This is an unexpectedly small number and usually comes as quite a surprise to most people, even those dealing in business and commerce. Similarly, the data also show that the total number of employees working in these establishments also scales approximately linearly with population size: on average, there are only about 8 employees for every establishment, again regardless of the size of the city. This remarkable constancy of the average number of employees and the average number of establishments across cities of vastly different sizes and characters is not only contrary to previous wisdom, but also rather puzzling when viewed in light of the pervasive superlinear agglomeration effects that underlie all socioeconomic activity including per capita increases in productivity, wages, GDP, and patent production. 10*

October 2, 2023

For instance, an increase of the population by a factor of one hundred from, say, 100,000 to 10 million results in the addition of one hundred times as many businesses but an increase of only a factor of two in their diversity. To put it slightly differently: doubling the size of a city results in doubling the total number of establishments, but only a meager 5 percent increase in new kind s of businesses. Almost all of this increase in diversity is reflected in a greater degree of specialization and interdependence involving larger numbers of people, both as workers and as clients. This is an important observation because it shows that increasing diversity is closely linked to increasing specialization, and this acts as a major driver of higher productivity following the 15 percent rule. 433

October 2, 2023

In contrast to the inherent universal properties of cities that 436

are manifested in the scaling laws, these rank-size distributions of business types reflect the individuality and distinctive characteristics of each specific city as exhibited by the composition of its economic activities. They are the hallmark of each city and obviously do depend on its history, geography, and culture. It is therefore all the more remarkable that, despite the unique admixture of business types for each individual city, the shape and form of their distribution is mathematically the same for all of them. So much so, in fact, that with a simple scale transformation inspired by theory their rank abundances collapse onto a single unique universal curve common to all cities, as shown explicitly in Figure 53. When one considers the vast range of income, density, and population levels, let alone the uniqueness and diverse cultures that vary so widely in cities across the United States, this universality is quite surprising. What is particularly satisfying is that this unexpected universality, as well as the actual form of the universal curve and the logarithmic scaling of diversity, can all be derived from theory. The universality is driven by the constraint that the sum total of all the different businesses in a city scales linearly with population size, regardless of the detailed composition of business types or of the city. The snakelike mathematical form of the actual distribution function in Figure 53 can be understood from a variant of a very general dynamical mechanism that has been successfully used for understanding rank-size distributions in many different areas, ranging from words and genes to species and cities. It goes under many different names including preferential attachment , cumulative advantage , the rich get richer , or the Yule-Simon process . It is based on a positive feedback mechanism in which new elements of the system (business types in this case) are added with a probability proportional to the abundances of how many are already there. The more there are, the more of that type are going to be added, so more frequent types get even more abundant with increasingly higher probability than less frequent types. 11

October 2, 2023

A couple of colloquial examples may help: successful companies and universities attract the smartest people, 437 *resulting in their becoming more successful, thereby attracting even smarter people, leading to even greater success and so on, just as wealthy people attract favorable investment opportunities that generate even more wealth which they further invest to get even more wealthy. Hence the catchphrase the rich get richer and its implied, but usually unstated, corollary the poor get poorer to characterize this process. Or, as so articulately put by Jesus according to the Gospel of Matthew in the New Testament: For everyone who has will be given more, and he will have an abundance. Whoever does not have, even what he has will be taken from him. This surprising declaration has been used by some fundamentalist Christians and others as justification for rampant capitalism—a sort of anti–Robin Hood slogan supporting the idea of taking from the poor to give to the rich. But while Jesus's remarks are a good example of preferential attachment , the quote is, not surprisingly, taken out of context. It is often conveniently forgotten that Jesus was actually referring to knowledge of the mysteries of the kingdom of heaven and not to material wealth. He was expressing a spiritual version of the very essence of diligent study, knowledge accumulation, and research and education as expressed by the ancient rabbis: He who does not increase his knowledge decreases it.*

October 2, 2023

Thus the energy available for growth is just the difference between the rate at which energy can be supplied and the rate that is needed for maintenance. On the supply side, metabolic rate in organisms scales sublinearly with the number of cells (following the generic ¾ power exponent derived from network constraints) while the demand increases approximately linearly . So as the organism increases in size, demand eventually outstrips supply because linear scaling grows faster than sublinear, with the consequence that the amount of energy available for growth continuously decreases, eventually going to zero resulting in the cessation of growth. In other words, growth stops because of the mismatch between the way maintenance and supply scale as size increases. The sublinear scaling of metabolic rate and the associated economies of scale arising 440 *from optimizing network performance are therefore responsible for why growth stops and why biological systems exhibit the bounded sigmoidal growth curves that were shown in Figures 15–18 of chapter 4. The same network mechanism that underlies sublinear scaling, economies of scale, and the cessation of growth is also responsible for the systematic slowing down of the pace of biological life as size increases—and for eventual death.*

October 2, 2023

It is the first of these, the physical infrastructural component, that has close analogies to biology and provides the metaphor of the city as an organism. But as I have persistently emphasized, the city is much more than its physicality. Consequently, the concept of metabolic rate as the supply-side input that fuels growth and sustains a city has to be expanded to include socioeconomic activity. In addition to the electricity, gas, oil, water, materials, products, artifacts, and so on that are used and generated in a city, we have to add wealth, information, ideas, and social capital. At a more fundamental level all of these, whether physical or socioeconomic, are driven and sustained by the supply of energy. In addition to heating buildings, transporting materials and people, manufacturing goods, and providing gas, water, and electricity, every transaction, every dollar gained or lost, every conversation and meeting, every phone call and text message, every idea and every thought has to be fueled by energy. Furthermore, just as food must be metabolized into a form that is useful for supplying cells and sustaining life, so the incoming energy and resources digested by a city must be transformed into a form that can be used to supply, sustain, and grow socioeconomic activities such as wealth creation, innovation, and the quality of life. No one has articulated this more eloquently than the great urbanist Lewis Mumford 13 : The chief function of the city is to convert power into form, energy into culture, dead matter into the living symbols of art, biological reproduction into social creativity. 442

October 2, 2023

The superlinear scaling of metabolism has profound consequences for growth. In contrast to the situation in biology, the supply of metabolic energy generated by cities as they grow increases faster than the needs and demands for its maintenance. Consequently, the amount available for growth, which is just the difference between its social metabolic rate and the requirements for maintenance, continues to increase as the city gets larger. The bigger the city gets, the faster it grows—a classic signal of open-ended exponential growth. A mathematical analysis indeed confirms that growth driven by superlinear scaling is actually faster than exponential: in fact, it's superexponential . Even though the conceptual and mathematical structure of the growth equation is the same for organisms, social insect communities, and cities, the consequences are quite different: sublinear scaling and economies of scale that dominate biology lead to stable bounded growth and the slowing down of the pace of life , whereas superlinear scaling and increasing returns to scale that dominate socioeconomic activity lead to unbounded growth and to an accelerating pace of life.

444

October 2, 2023

October 2, 2023

448

9: TOWARD A SCIENCE OF COMPANIES

Given this observation it is natural to ask, as we did for cities and organisms, whether companies scale in terms of their measurable metrics such as their sales, assets, expenses, and profits. Do companies manifest systematic regularities that transcend their size, individuality, and business sector? And if so, could there possibly be a quantitative, predictive science

448

of companies paralleling the science of cities that was developed in previous chapters? Is it possible to understand the general quantitative features of their life histories, how they grow, mature, and eventually die?

October 2, 2023

The mechanisms that have traditionally been suggested for understanding companies can be divided into three broad categories: transaction costs, organizational structure, and competition in the marketplace. Although these are interrelated they have very often been treated separately. In the language of the framework developed in previous chapters these can be expressed as follows: (1) Minimizing transaction costs reflects economies of scale driven by an optimization principle, such as maximizing profits. (2) Organizational structure is the network system within a company that conveys information, resources, and capital to support, sustain, and grow the enterprise. (3) Competition results in the evolutionary pressures and selection processes inherent in the ecology of the marketplace. 450

October 2, 2023

Axtell, who is at George Mason University in Virginia and is also on the external faculty of the Santa Fe Institute, is a leading expert in agent-based modeling, which is a computational technique used for simulating systems composed of huge numbers of components. 3 Basically, the strategy involves postulating simple rules governing the interactions between individual constituent agents, which could be companies, cities, or people, coupled with an algorithm that specifies how they evolve in time and letting the resulting system run on a computer. More sophisticated versions include rules for learning, adaptation, and even reproduction so as to model more realistic evolutionary processes. 451

October 2, 2023

October 2, 2023 [451]

Nassim Taleb, author of the best-selling, highly influential book The Black Swan, has been particularly harsh on economists despite, or maybe because of, having been trained in business and finance. 5 He has held positions at several distinguished universities including New York University and Oxford and has focused on the importance of coming to terms with outlying events and developing a deeper understanding of risk. He has been brutally outspoken in his condemnation of classical economic thinking with hyperbolic comments such as: "Years ago, I noticed one thing about economics, and that is that economists didn't get anything right." He has even called for the Nobel Memorial Prize in economics to be withdrawn, saying that the damage from economic theories can be devastating. I may disagree with some of Taleb's ideas and polemics but it's important and healthy to have such outspoken mavericks challenging the orthodoxy, especially when it's had such a poor record and its proclamations have major implications for our lives. [452]

October 2, 2023

In its extreme version the underlying philosophy of agent-based modeling is antithetical to the traditional scientific framework, where the primary challenge is to reduce huge numbers of seemingly disparate and disconnected observations down to a few basic general principles or laws: as in biology, where the principle of natural selection applies to all organisms from cells to whales, or in physics, where Newton's laws apply to all motion from automobiles to planets. In contrast, the aim of agent-based modeling is to construct an almost one-to-one mapping of each specific system. General laws and principles that constrain its structure and dynamics play a secondary role. For example, in simulating a specific company, every individual worker, administrator, transaction, sale, cost, et cetera is in effect [454] *included and each company consequently treated as a separate, almost unique entity, typically without explicit regard either to its systematic behavior or its relationship to the bigger picture. Clearly, both approaches are needed: the generality and parsimony of "universal" laws and systematic behavior reflecting the big picture and dominant forces shaping general behavior, coupled with and informed by detailed modeling reflecting the individuality and uniqueness of each company. In the case of cities, scaling laws revealed that 80 to 90 percent of their measurable characteristics are determined from just knowing their population size, with the remaining 10 to 20 percent being a measure of their individuality and uniqueness, which can be understood only from detailed studies that incorporate local historical, geographical, and cultural characteristics.*

October 2, 2023

October 2, 2023 454

The initial results and conclusions of our investigation into the scaling of companies are very compelling. They provide a powerful basis for developing an understanding of their generic structure and life histories. Figures 60–63 show the sales, incomes, and assets of all 28,853 companies plotted logarithmically against their number of employees. These are the dominant financial characteristics of any company and are standard measures of their fiscal health and dynamics. As these graphs clearly demonstrate, companies do indeed scale following simple power laws and as anticipated they do so with a much greater spread around their average behavior than for either cities or organisms. So in this statistical sense, companies are approximately scaled, self-similar versions of one another: Walmart is an approximately scaled-up version of a much smaller, modest-size company. Even after taking this greater variance into account, this scaling result reveals remarkable regularities in the size and dynamics of companies and is quite surprising given the tremendous variety of different business sectors, 457 *locations, and age.*

October 3, 2023

The six points resulting from averaging over each bin are shown as gray dots in the graph. They represent a highly coarse-grained reduction of the data, and as you can see follow a very good straight line supporting the idea that underlying the statistical spread is an idealized power law. Because the size and number of bins used is arbitrary, we could just as well have divided up the entire interval into ten, fifty, or one hundred bins rather than just eight, and test whether the straight line remains robust against increasingly finer resolutions of the data. It does. Although binning is not a rigorous mathematical procedure, the stability of obtaining approximately the same straight-line fit using different resolutions lends strong support to the hypothesis that on average companies are self-similar and satisfy power law scaling. The graph in Figure 4 at the opening of the book is in fact the result of this binning procedure, as is the graph in Figure 41 taken from Axtell's work on showing that companies follow Zipf's law. These results strongly suggest that companies, like cities and organisms, obey universal dynamics that transcend their individuality and uniqueness and that a coarse-grained science of companies is conceivable. 459

October 3, 2023

A crucial aspect of the scaling of companies is that many of their key metrics scale sublinearly like organisms rather than superlinearly like cities. This suggests that companies are more like organisms than cities and are dominated by a version of economies of scale rather than by increasing returns and innovation. 461

October 3, 2023

As we saw in chapter 4, sublinear scaling in biology leads to bounded growth and a finite life span, whereas in chapter 8 we saw that the superlinear scaling of cities (and of economies) leads to open-ended growth. Their sublinear scaling therefore suggests that companies also eventually stop growing and ultimately die, hardly the image that many CEOs would cherish. It's actually not quite as simple as that because the prediction for the growth of companies is more subtle than just a simple extrapolation from biology. To explain this I am going to present a simplified version of how the general theory applies to companies focusing on the essential features that determine their growth and mortality. The sustained growth of a company is ultimately fueled by its profits (or net revenue), where these are defined as the difference between sales (or total income) and total expenses; expenses include salaries, costs, interest payments, and so on. To continue growth over a prolonged period, companies must eventually return a profit, part of which is sometimes used to pay dividends to shareholders. Together with other investors, they in turn may buy additional stocks and bonds to help support the future health and growth of the company. However, to understand their generic behavior it is more transparent to ignore dividends and investments, which are primarily important for smaller, younger companies, and concentrate on profits, which are the dominant driver of growth for larger ones. As we've seen, growth in both organisms and cities is fueled by the difference between metabolism and maintenance. Using that language, the total income (or sales) of a company can be thought of as its "metabolism" while expenses can be thought of as its "maintenance" costs. In biology, metabolic rate scales sublinearly with size, so as organisms increase in size the supply of energy cannot keep up with the maintenance demands of cells, leading to the eventual cessation of growth. On the other hand, the social metabolic rate in cities scales superlinearly, so as cities grow the creation of social capital increasingly outpaces the demands of maintenance, leading to faster and faster open-ended growth. So how does this dynamic play out in companies? Intriguingly, companies manifest yet another variation on this general theme by following a path that sits at the cusp between organisms and cities. Their effective metabolic rate is neither sub- nor superlinear but falls right in the middle by being linear. This 461 *is illustrated in Figures 63 and 64, where sales are plotted logarithmically against the number of employees showing a best fit with a slope very close to one. Expenses, on the other hand, scale in a more complicated fashion: they start out sublinearly but, as companies become larger, eventually transition to becoming approximately linear. Consequently, the difference between sales and expenses, which is the driver of growth, also eventually scales approximately linearly. This is good news because, mathematically, linear scaling leads to exponential growth and this is what all companies strive for. Furthermore, this also shows why, on average, the economy continues to expand at an exponential rate because the overall performance of the market is effectively an average over the growth performances of all its individual participating companies. Although this may be good news for the overall economy, it sets a major challenge for each individual company because each one has to keep up with an exponentially expanding market. So even if a company is growing exponentially (the good news), this may not be sufficient for it to survive unless its expansion rate is at least that of the market (the bad news). This primitive version of the "survival of the fittest" for companies is the essence of the free market economy. More good news is that the nonlinear scaling of maintenance expenses in younger companies, buoyed by investments and the ability to borrow large amounts relative to their size, fuels their rapid growth. Consequently, the idealized growth curve of companies has characteristics in common with classic sigmoidal growth in biology in that it starts out relatively rapidly but slows down as companies become larger and maintenance expenses transition to becoming linear. However, unlike biology, whose maintenance costs do not transition to linearity, companies do not cease growing but continue to grow exponentially, though at a more modest rate.*

October 3, 2023

Let's see how this scenario compares with data. Figure 68 is a wonderful graph showing the growth of sales for all 28,853 companies in the Compustat data set plotted together in real calendar time, adjusted for inflation. To get all of them onto a single manageable graph, the vertical axis representing 463 *sales is logarithmic. Despite being a "spaghetti" plot, the graph is surprisingly illuminating. The overall trend is clear: as predicted, many young companies shoot out of the starter's block and grow rapidly before slowing down, while older, more mature ones that have survived continue growing but at a much slower rate. Furthermore, the upward trends of these older, slower-growing companies all follow an approximate straight line with similar shallow slopes. On this semilogarithmic plot, where the vertical axis (sales) is logarithmic but the horizontal one (time) is linear, a straight line means mathematically that sales are growing exponentially with time. Thus, on average, all surviving companies eventually settle down to a steady but slow exponential growth, as predicted. This is very encouraging, but there's a potential pitfall that becomes apparent when the growth of each company is measured relative to the growth of the overall market. In that case, as can be clearly seen in Figure 70 where the overall growth of the market has been factored out, all large mature companies have stopped growing . Their growth curves when corrected for both inflation and the expansion of the market now look just like typical sigmoidal growth curves of organisms in which growth ceases at maturity, as illustrated in Figures 15–18 of chapter 4. This close similarity with the growth of organisms when viewed in this way provides a natural segue into whether this similarity extends to mortality and whether, like us, all companies are destined to die.*

This close similarity with the growth of organisms when viewed in this way provides a natural segue into whether this similarity extends to mortality and whether, like us, all companies are destined to die.

October 3, 2023

In all cases the number of survivors falls rapidly immediately following their initial public offering, with fewer than 5 percent remaining alive after thirty years. Similarly, the mortality curves show that the number that have died reaches almost 100 percent within fifty years, with almost half of them having already disappeared in less than ten. It's tough being a company! The survival curves are well approximated by a simple exponential as shown in Figure 75,

468

where the number of companies that have survived is plotted logarithmically versus their age; plotted this way, exponentials appear as straight lines. You might have thought that these results would depend sensitively on whether death occurred via mergers and acquisitions rather than from bankruptcies and liquidations. However, as you can see, they both follow very similar exponential survival curves with only slightly different values for their mortality. One might also have expected the results to depend on which business sector a company is in. The dynamics and competitive market forces would seem to be quite different, for instance, in the energy sector compared with IT, transportation, or finance. Surprisingly, however, all business sectors show similar characteristic exponential survival curves with similar timescales: no matter which sector or what the stated cause is, only about half of the companies survive for more than ten years. This is consistent with an analysis showing that companies scale in approximately the same way when broken down into separate business categories. Within each sector, power laws are obtained having exponents close to those found for the entire cohort of companies—those shown in Figure 75. In other words, the general dynamics and overall life history of companies are effectively independent of the business sector in which they operate. This strongly suggests that there is indeed a universal dynamic at play that determines their coarse-grained behavior, independent of their commercial activity or whether they are eventually going to go bankrupt or merge with or be bought by another company. In a word, this strongly supports the idea of a quantitative science of companies.

October 4, 2023

This is really quite amazing. After all, when we think of the birth, death, and general life history of companies as they struggle to establish and maintain themselves in the marketplace dealing with the vagaries, uncertainties, and unpredictability of economic life and the myriad specific decisions and accidents that led to the successes and failures that preceded their death, it's hard to believe that collectively they were following such simple general rules. This 472 *revelation echoes the surprise that organisms, ecosystems, and cities are likewise subject to generic constraints, despite the apparent uniqueness and individuality of their life histories.*

October 4, 2023

What is the special property of exponentials that they describe the decay of so many disparate systems? It's simply that they arise whenever the death rate at any given time is directly proportional to the number that are still alive. This is equivalent to saying that the percentage of survivors that die within equal slices of time at any age remains the same. A simple example will make this clear: taking one year as the time slice, this says that the percentage of five-year-old companies that die before they reach six years old is the same as the percentage of fifty-year-old companies that die before they reach fifty-one. In other words: the risk of a company's dying does not depend on its age or size. 472

October 4, 2023

A more complete analysis therefore needs to include these so-called censored companies, whose life spans are at least as long as and likely longer than the period over which they appear in the data set. This actually involves a sizable number of companies: in the sixty years covered, 6,873 firms were still alive at the end of the window in 2009. Fortunately, there is a well-established sophisticated methodology, called survival analysis, that has been developed precisely for addressing this issue. Survival analysis was developed in medicine in order to estimate survival probabilities for patients who have undergone therapeutic interventions under test conditions. These tests have necessarily to be conducted over a limited time period, leading to the problem we face here, namely, that many subjects die after the test period has ended. The technique commonly used, called the Kaplan-Meier estimator, employs the entire data set and optimizes probabilities assuming that each death event is statistically independent of every other 473 *death. 7*

October 4, 2023

The survival analysis technique used for dealing with "incomplete observations" such as we have here was invented in 1958 by two statisticians, Edward Kaplan and Paul Meier. It has since been extended to areas outside of medicine and used to estimate, for example, how long people can expect to remain unemployed after a job loss or how long it takes for machine parts to fail. Amusingly, Kaplan and Meier each submitted similar but independent papers to the prestigious Journal of the American Statistical Association for publication, and a wise editor persuaded them to combine them into a single paper. This has since been cited in other scholarly papers more than 34,000 times, which is an extremely large number for an academic paper.

474

October 4, 2023

Although there are significant differences, it's hard not to be struck by how similar the growth and death of companies and organisms are when viewed through the lens of scaling—and how dissimilar they both are to cities. Companies are surprisingly biological and from an evolutionary perspective their mortality is an important ingredient for generating innovative vitality resulting from "creative destruction" and "the survival of the fittest." Just as all organisms must die in order that the new and novel may blossom, so it is that all companies disappear or morph to allow new innovative variations to flourish: better to have the excitement and innovation of a Google or Tesla than the stagnation of a geriatric IBM or General Motors. This is the underlying culture of the free market system.

475

October 4, 2023

The great turnover of companies and especially the continual

475

churning of mergers and acquisitions are integral to the market process. And of course this means that the Googles and Teslas, which may seem invincible now, will themselves eventually fade away and disappear. From this point of view we should not lament the passing of any company—it's an essential component of economic life—we should only mourn and be concerned about the fate of the people who often suffer when companies disappear, whether they are the workers, management, or even the owners. If only we could tame the potential brutality and greed of the survival of the fittest and soften some its more egregious consequences by formulating a magic algorithm for how to balance the classic tension between regulation, government intervention, and uncontrolled rampant capitalism. This struggle painfully played itself out as we witnessed the struggle between the death throes of corporations that probably should have died and the desire to save jobs and protect the lives of workers because certain incompetent if not duplicitous corporations were deemed "too big to fail" during the 2008 financial crisis.

If only we could tame the potential brutality and greed of the survival of the fittest and soften some its more egregious consequences by formulating a magic algorithm for how to balance the classic tension between regulation, government intervention, and uncontrolled rampant capitalism.

October 4, 2023

October 4, 2023 476

479

There are some wonderful examples of these geriatric survivors. For instance, the oldest shoemaker in Germany is the Eduard Meier company, founded in Munich in 1596, which became purveyors to the Bavarian aristocracy. It still has only a single store that sells, though no longer makes, quality upscale shoes. The oldest hotel in the world according to Guinness World Records is Nishiyama Onsen Keiunkan in Hayakawa, Japan, which was founded in 705. It has been in the same family for fifty-two generations and even in its modern incarnation has only thirty-seven rooms. Its main *attraction seems to be its hot springs. The world's oldest company was purported to be Kongo Gumi, founded in Osaka, Japan, in 578. It was also a family business going back many generations, but after almost 1,500 years of continuously being in business it went into liquidation in 2006 and was purchased by the Takamatsu Corporation. And what was the niche market that Kongo Gumi cornered for 1,429 years? Building beautiful Buddhist temples. But sadly, with the changes in Japanese culture following the Second World War the demand for temples dried up and Kongo Gumi was unable to adapt fast enough.*

October 4, 2023

economies of scale over innovation and idea creation. Companies typically operate as highly constrained top-down organizations that strive to increase efficiency of production and minimize operational costs so as to maximize profits. In contrast, cities embody the triumph of innovation over the hegemony of economies of scale. Cities aren't, of course, driven by a profit motive and have the luxury of being able to balance their books by raising taxes. They operate in a much more distributed fashion, with power spread across multiple organizational structures from mayors and councils to businesses and citizen action groups. No single group has absolute control. As such, they exude an almost laissez-faire, freewheeling ambience relative to companies, taking advantage of the innovative benefits of social interactions whether good, bad, or ugly. Despite their apparent bumbling inefficiencies, cities are places of action and agents of change relative to companies, which by and large usually project an image of stasis unless they are young. 481

October 4, 2023

To achieve greater efficiency in the pursuit of greater market share and increased profits, companies stereotypically add more rules, regulations, protocols, and procedures at increasingly finer levels of organization, resulting in the increased bureaucratic control that is typically needed to 482 *administer, manage, and oversee their execution. This is often accomplished at the expense of innovation and R&D (research and development), which should be major components of a company's insurance policy for its long-term future and survivability. It's difficult to obtain meaningful data on "innovation" in companies because it's not straightforward to quantify. Innovation is not necessarily synonymous with R&D, especially as there are significant tax advantages in labeling all sorts of extraneous activities as R&D expenses. Nevertheless, from analyzing the Compustat data set we found that the relative amount allocated to R&D systematically decreases as company size increases, suggesting that support for innovation does not keep up with bureaucratic and administrative expenses as companies expand.*

October 4, 2023

The increasing accumulation of rules and constraints is often accompanied by stagnating relationships with consumers and suppliers that lead companies to become less agile and more rigid and therefore less able to respond to significant change. In cities we saw that one very important hallmark is that they become ever more diverse as they grow. Their spectrum of business and economic activity is incessantly expanding as new sectors develop and new opportunities present themselves. In this sense cities are prototypically multidimensional, and this is strongly correlated with their superlinear scaling, open-ended growth, and expanding social networks—and a crucial component of their resilience, sustainability, and seeming immortality. While the dimensionality of cities is continually expanding, the dimensionality of companies typically contracts from birth through adolescence, eventually stagnating or even further contracting as they mature and move into old age. When still young and competing for a place in the market, there is a youthful excitement and enthusiasm as new products are developed and ideas bubble up, some may be crazy and unrealistic and some grandiose and visionary. But market forces are at work so that only a few of these are successful as the company gains a foothold and an identity. As it grows, the feedback mechanisms inherent in the market lead to a 482

narrowing of its product space and inevitably to greater specialization. The great challenge for companies is how to balance the positive feedback from market forces, which strongly encourage staying with "tried and true" products versus the long-term strategic need to develop new areas and commodities that may be risky and won't give immediate return. Most companies tend to be shortsighted, conservative, and not very supportive of innovative or risky ideas, happy to stay almost entirely with their major successes while the going is good because these "guarantee" short-term returns. Consequently, they tend toward becoming more and more unidimensional . This reduction in diversity coupled with the predicament described earlier in which companies sit near a critical point is a classic indicator of reduced resilience and a recipe for eventual disaster. By the time a company realizes its condition it is often too late. Reconfiguring and reinventing become increasingly difficult and expensive. So when a large enough unanticipated fluctuation, perturbation, or shock comes along the company becomes seriously at risk and ripe for a takeover, buyout, or simply going belly-up. In a word, it is, as the Mafiosi put it, il bacio della morte— the kiss of death. 9

October 4, 2023

October 4, 2023 484

10: THE VISION OF A GRAND UNIFIED THEORY OF SUSTAINABILITY

One of the major challenges of the twenty-first century that will have to be faced is the fundamental question as to whether human-engineered social systems, from economies to cities, which have only existed for the past five thousand years or so, can continue to coexist with the "natural" biological world from which they emerged and which has 484 *been around for several billion years. To sustain more than 10 billion people living in harmony with the biosphere at a standard of living and quality of life comparable to what we now have requires that we develop a deep understanding of the principles and underlying system dynamics of this social-environmental coupling. I have argued that a critical component of this is to develop a deeper understanding of cities and urbanization. Continuing to pursue limited and single-system approaches to the many problems we face without developing a unifying framework risks the possibility that we will squander huge financial and social capital and fail miserably in addressing the really big question, resulting in dire consequences.*

October 4, 2023

485

We need a broad and more integrated scientific framework that encompasses a quantitative, predictive, mechanistic theory for understanding the relationship between human-engineered systems, both social and physical, and the "natural" environment—a framework I call a grand unified theory of sustainability. It's time to initiate a massive international Manhattan-style project or Apollo-style program dedicated to addressing global sustainability in an integrated, systemic sense. 1

October 4, 2023

486

This is a wonderfully consistent picture: the same conceptual framework based on underlying network dynamics and geometry with the same mathematical structure leads to quite different outcomes in these two very different cases, and both are strongly supported by a plethora of diverse data and observations. However, there is a big catch with potentially huge consequences. Even though the growth of organisms, cities, and economies follows essentially identical mathematical equations, their resulting solutions have subtle but crucial differences arising from one being driven by sublinear scaling (the economies of scale of organisms) and the other by superlinear scaling (the increasing returns to *scale of cities and economies): in the superlinear case, the general solution exhibits an unexpectedly curious property technically known as a finite time singularity, which is a signal of inevitable change, and possibly of potential trouble ahead. A finite time singularity simply means that the mathematical solution to the growth equation governing whatever is being considered—the population, the GDP, the number of patents, et cetera— becomes infinitely large at some finite time , as illustrated in Figure 76. This is obviously impossible, and that's why something has to change. Before addressing some of the consequences of this phenomenon, let me first elaborate on some of its salient features. Simple power laws and exponentials are continuously increasing functions that also eventually become infinitely large, but they take an infinite time to do so. Another way of saying this is that in these cases the "singularity" has been pushed off to an infinite time into the future, thereby rendering it "harmless" relative to the potential impact of a finite time singularity. In the case of growth driven by superlinear scaling, the approach to the finite time singularity, represented by the solid line in Figure 76, is faster than exponential. This is often referred to as superexponential, a term I've already used earlier when discussing the growth of cities.*

October 4, 2023

October 4, 2023 487

487

This kind of growth behavior is clearly unsustainable because it requires an unlimited, ever-increasing, and eventually infinite supply of energy and resources at some finite time in the future in order to maintain it. Left unchecked, the theory predicts that it triggers a transition to a phase that leads to stagnation and eventual collapse, as illustrated in Figure 77. This scenario sounds just like a rehash of the standard Malthusian argument that has been summarily dismissed by generations of economists: namely, that we won't be able to keep up with demand and that open-

ended growth will eventually lead to catastrophe. Which brings us to the crux of the matter. Because of the presence of a finite time singularity resulting from superlinear scaling, this scenario is categorically different from that of Malthus. If growth were purely exponential as assumed by Malthusians, neo-Malthusians, their followers, and critics, then the production of energy, resources, and food could at least in principle keep up with exponential expansion because all of the relevant characteristics of the economy or city remain finite, even if they continue to increase in size and become very large. This cannot be achieved if you are growing superexponentially and approaching a finite time singularity. In this scenario demand gets progressively larger and larger, eventually becoming infinite within a finite period of time. It is simply not possible to supply an infinite amount of energy, resources, and food in a finite time. So if nothing else changes, this inextricably leads to stagnation and collapse, as illustrated in Figure 77. An extensive analysis carried out in 2001 by Didier Sornette and Anders Johansen, then at UCLA, showed that data on population growth and the growth of financial and economic indicators strongly support the theoretical predictions that we have been growing superexponentially and are indeed headed toward such a singularity. 2 I want to emphasize that this situation is qualitatively quite different from classic Malthusian dynamics, where this is no such singularity. The existence of a singularity signifies that there has to be a transition from one phase of the system to another having very different characteristics, analogous to the way the condensation of steam to water which subsequently freezes to ice epitomizes transitions between different phases of the same system, each having quite different physical properties.

October 4, 2023

So how can such a collapse be avoided, and can it be achieved while still ensuring open-ended growth? The first point to appreciate is that these predictions assume that the parameters of the growth equation do not change. So one clear strategy for forestalling a potential catastrophe is to intervene before reaching the singularity by "resetting" the parameters. Moreover, to maintain open-ended growth with

489

these new settings requires that the driving term in the equation—the "social metabolism"—needs to remain superlinear, meaning that the new dynamic must still be driven by the positive feedback forces of social interaction responsible for innovation, and for wealth and knowledge creation. Such an "intervention" is none other than what is usually referred to as an innovation . A major innovation effectively resets the clock by changing the conditions under which the system has been operating and growth occurring. Thus, to avoid collapse a new innovation must be initiated that resets the clock, allowing growth to continue and the impending singularity to be avoided .

October 4, 2023

This can be restated as a sort of "theorem": to sustain open-ended growth in light of resource limitation requires continuous cycles of paradigm-shifting innovations , as illustrated in Figure 78.

490

October 4, 2023

Unfortunately, however, it's not quite as simple as that. There's yet another major catch, and it's a big one. The theory dictates that to sustain continuous growth the time between successive innovations has to get shorter and shorter. Thus paradigm-shifting discoveries, adaptations, and innovations must occur at an increasingly accelerated pace. Not only does the general pace of life inevitably quicken, but we must innovate at a faster and faster rate!

492

October 4, 2023

To some extent this is in the eyes of the beholder, but it's likely that most of us would agree that certain discoveries and innovations such as printing, coal, the telephone, and computers constitute a major "paradigm shift," whereas railways and cell phones may be more debatable.

494

Unfortunately, there is no established quantitative "science of innovation" and therefore no universally agreed-upon criteria or data relating directly to major innovations and paradigm shifts, let alone to finite time singularities. So in order to confront theory with data we have to rely on informal studies and a certain degree of intuition. This situation may well change as innovation becomes an increasingly active area of investigation, with researchers beginning to grapple with questions such as what is innovation, how do we measure it, how does it happen, and how can it be facilitated? 4

October 4, 2023

October 4, 2023 494

497

Just as the Malthusians ignored the crucial role of innovation, the singularity enthusiasts ignore the crucial role of the entire socioeconomic dynamic of the planet, which in fact is the prime driver of the impending singularity. Neither case is anchored in a broader framework that embraces a quantitative mechanistic theory, so whatever their predictions are it's very hard to evaluate them scientifically. Perhaps the greatest conceptual irony, especially of the singularists, is that their conclusions and speculations are based on exponential growth, which doesn't actually lead to a singularity, at least not in a finite time .

October 4, 2023

498

The time between the "Computer Age" and the "Information and Digital Age" was no more than about thirty years—to be compared with the thousands of years between the Stone, Bronze, and Iron ages. The clock by which we measure time on our watches and digital devices is very misleading; it is determined by the daily rotation of the Earth around its axis and its annual rotation around the sun. This astronomical *time is linear and regular. But the actual clock by which we live our socioeconomic lives is an emergent phenomenon determined by the collective forces of social interaction: it is continually and systematically speeding up relative to objective astronomical time. We live our lives on the metaphorical accelerating socioeconomic treadmill. A major innovation that might have taken hundreds of years to evolve a thousand or more years ago may now take only thirty years. Soon it will have to take twenty-five, then twenty, then seventeen, and so on, and like Sisyphus we are destined to go on doing it, if we insist on continually growing and expanding . The resulting sequence of singularities, each of which threatens stagnation and collapse, will continue to pile up, leading to what mathematicians call an essential singularity —a sort of mother of all singularities.*

October 4, 2023

The great John von Neumann, mathematician, physicist, computer scientist, and polymath, a man whose ideas and accomplishments have had a huge influence on your life, made the following remarkably prescient observation more than seventy years ago: "The ever accelerating progress of technology and changes in the mode of human life . . . gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue." 7 Among von Neumann's many accomplishments before he died at the relatively young age of fifty-three in 1957 are his seminal role in the early development of quantum mechanics, his invention of game theory, which is a major tool in economic modeling, and the conceptual design of modern computers universally referred to as the von Neumann architecture.

499

October 4, 2023

So can we imagine making an innovation as powerful and influential as the invention of the Internet every fifteen, ten, or even five years? This is a classic reductio ad absurdum argument showing that regardless of how ingenious we are,

499

how many marvelous gadgets and devices we invent, we simply won't be able to overcome the threat of the ultimate singularity if we continue business as usual.

October 4, 2023

Estimates from the theory suggest that we are due for another paradigm shift within the next twenty to thirty years' time. This is a little shorter than the estimates from fits to data by Johansen and Sornette, who suggest a number more like thirty-five years. The theory cannot, of course, tell us the nature of the change, so we can only make wild speculations as to its nature. It could be something as relatively mundane as driverless cars and associated smart devices or something as dramatic as the kind of science fiction fantasy of Kurzweil and the singularists. Most likely it's none of the above and, if we are able to make the paradigm shift, it'll be something totally unexpected. Perhaps more likely is that we can't make the shift, and that we will need to come to terms with the whole concept of open-ended growth and find some new way of defining "progress" or be content with what we've got and spend our energies raising the entire planet's standard of living to reflect a comparably high quality of life. Now, that would be a truly major paradigm shift! Continuous growth and the consequent ever-increasing acceleration of the pace of life have profound consequences for the entire planet and, in particular, for cities, socioeconomic life, and the process of global urbanization. Until recent times, the time between major innovations far exceeded the productive life span of a human being. Even in my own lifetime it was unconsciously assumed that one would continue working in the same occupation using the same expertise throughout one's life. This is no longer true; a typical human being now lives significantly longer than the time between major innovations, especially in developing and developed countries. Nowadays young people entering the workforce can expect to see several major changes during their lifetime that will very likely disrupt the continuity of their careers. This increasingly rapid rate of change induces serious stress on all facets of urban life. This is surely not sustainable, and, if nothing changes, we are heading for a major crash and a potential collapse of the

499

entire socioeconomic fabric. The challenges are clear: Can we return to an analog of a more "ecological" phase from which we evolved and be satisfied with some version of sublinear scaling and its attendant natural limiting, or no-growth, stable configuration? Is this even possible? Can we have the kind of vibrant, innovative, creative society driven by ideas and wealth creation as manifested by the best of our world's cities and social organizations, or are we destined to a planet of urban slums and the ultimate specter of devastation raised by Cormac McCarthy's novel The Road ? 8 Given the special, unique role of cities as the originators of many of our present problems and their continuing role as the superexponential driver toward potential disaster, understanding their dynamics, growth, and evolution in a scientifically predictable, quantitative framework is crucial to achieving long-term sustainability on the planet. Perhaps of even greater importance for the immediate future is to develop such a theory within the context of a grand unified theory of sustainability by bringing together the multiple studies, simulations, databases, models, theories, and speculations concerning global warming, the environment, financial markets, risk, economies, health care, social conflict, and the myriad other characteristics of man as a social being interacting with his environment.

October 4, 2023

October 4, 2023 500

Afterword

Among the classic grand syntheses in modern science are Newton's laws, which taught us that heavenly laws are no different from those on Earth; Maxwell's unification of electricity and magnetism, which brought the ephemeral ether into our lives and gave us electromagnetic waves; 502

Darwin's theory of natural selection, which reminded us that we're just animals and plants after all; and the laws of thermodynamics, which suggest that we can't go on forever. Each of these has had profound consequences not only in changing the way we think about the world, but also in laying the foundations for technological advancements that have led to the standard of living many of us are privileged to enjoy. Nevertheless, they are all to varying degrees incomplete. Indeed, understanding the boundaries of their applicability, the limits to their predictive power, and the ongoing search for exceptions, violations, and failures have provoked even deeper questions and challenges, stimulating the continued progress of science and the unfolding of new ideas, techniques, and concepts.

October 4, 2023

Emboldened by its great success, physicists endowed this fantastic vision with the grand title of the Theory of Everything. Demanding mathematical consistency between quantum mechanics and general relativity suggested that the basic building blocks of this universal theory might be microscopic vibrating strings rather than the traditional elementary point particles upon which Newton and all subsequent theoretical developments were based. Consequently this vision took on the more prosaic subtitle "string theory." Like the invention of gods and God, the concept of a Theory of Everything connotes the grandest vision of all, the inspiration of all inspirations, namely that we can encapsulate and understand the entirety of the universe in a small set of precepts, in this case, a concise set of mathematical equations from which literally everything follows. Like the concept of God, however, it is potentially misleading and intellectually dangerous. 504

October 4, 2023

October 4, 2023 507

The institute has been internationally recognized as "the formal birthplace of the interdisciplinary study of complex systems" and has played a central role in recognizing that many of the most challenging, exciting, and profound questions facing science and society lie at the boundaries between traditional disciplines. Among these are the origins of life; the generic principles of innovation, growth, evolution, and resilience whether of organisms, ecosystems, pandemics, or societies; network dynamics in nature and society; biologically inspired paradigms in medicine and computation; the interrelationship between information processing, energy, and dynamics in biology and society; the sustainability and fate of social organizations; and the dynamics of financial markets and political conflicts. 509

October 5, 2023

The institute was originally conceived by a small group of distinguished scientists, including several Nobel laureates, most of whom had some association with Los Alamos National Laboratory. They were concerned that the academic landscape had become so dominated by disciplinary stovepiping and specialization that many of the big questions, and especially those that transcend disciplines or were perhaps of a societal nature, were being ignored. The reward system for obtaining an academic position, for gaining promotion or tenure, for securing grants from federal agencies or private foundations, and even for being elected to a national academy, was becoming more and more tied to demonstrating that you were the expert in some tiny corner of some narrow subdiscipline. The freedom to think or speculate about some of the bigger questions and broader issues, to take a risk or be a maverick, was not a luxury many could afford. 510

October 5, 2023

October 5, 2023 514

Data for data's sake, or the mindless gathering of big data, without any conceptual framework for organizing and understanding it, may actually be bad or even dangerous. Just relying on data alone, or even mathematical fits to data, without having some deeper understanding of the underlying mechanism is potentially deceiving and may well lead to erroneous conclusions and unintended consequences. This admonition is closely related to the classic warning that "correlation does not imply causation." Just because two sets of data are closely correlated does not imply that one is the cause of the other. There are many bizarre examples that illustrate this point. 4 For instance, over the eleven-year period from 1999 to 2010 the variation in the total spending on science, space, and technology in the United States almost exactly followed the variation in the number of suicides by hanging, strangulation, and suffocation. It's extremely unlikely that there is any causal connection between these two phenomena: the decrease in spending in science was surely not the cause of the decrease in how many people hanged themselves. However, in many situations such a clear-cut conclusion is not so clear. More generally, correlation is in fact often an important indication of a causal connection but usually it can only be established after further investigation and the development of a mechanistic model. 517

October 5, 2023

October 5, 2023 519

A contrary view to this trend was forcibly expressed by the Nobel Prize–winning geneticist Sydney Brenner, whom I quoted in chapter 3 and who was coincidentally director of the famous institute in Cambridge founded by Max Perutz that I mentioned earlier: "Biological research is in crisis. . . . 520

Technology gives us the tools to analyse organisms at all scales, but we are drowning in a sea of data and thirsting for some theoretical framework with which to understand it. Although many believe that 'more is better,' history tells us that 'least is best.' We need theory and a firm grasp on the nature of the objects we study to predict the rest." Not long after the publication of Chris Anderson's article, Microsoft published a fascinating series of essays in a book titled The Fourth Paradigm: Data-Intensive Scientific Discovery. It was inspired by Jim Gray, a computer scientist at Microsoft who was sadly lost at sea in 2007. He envisioned the data revolution as a major paradigm shift in how science would advance in the twenty-first century and called it the fourth paradigm.

October 5, 2023

He identified the first three as (1) empirical observation (pre-Galileo), (2) theory based on models and mathematics (post-Newtonian), and (3) computation and simulation. My impression is that, in contrast to Chris Anderson, Gray viewed this fourth paradigm as an integration of the previous three, namely as a unification of theory, experiment, and simulation, but with an added emphasis on data gathering and analysis. In that sense it's hard to disagree with him because this is pretty much the way science has progressed for the last couple of hundred years—the difference being primarily quantitative: the "data revolution" provides us with a much greater possibility for exploiting and enabling strategies we have been using for a very long time. In this sense this is more like paradigm 3.1 than paradigm 4.0. But there is a new kid on the block that many feel promises more and, like Anderson, potentially subverts the need for the traditional scientific method. This invokes techniques and strategies with names like machine learning, artificial intelligence , and data analytics. There are many versions of these, but all of them are based on the idea that we can design and program computers and algorithms to evolve and adapt based on data input to solve problems, reveal insights, and make predictions. They all rely on iterative procedures for finding and building upon correlations in data without concern for why such 521 *relationships exist and implicitly presume that "correlation supersedes causation." This approach has become a huge area of interest and has already had a big impact on our lives. For instance, it is central to how search engines like Google operate, how strategies for investment or operating an organization are devised, and it provides the foundational basis for driverless cars.*

October 5, 2023

October 5, 2023 521

The more one can bring big data into the enterprise the better, provided it is constrained by a bigger-picture conceptual framework that, in particular, can be used to judge the relevance of correlations and their relationship to mechanistic causation. If we are not to "drown in a sea of data" we need a "theoretical framework with which to understand it . . . and a firm grasp on the nature of the objects we study to predict the rest." 525

October 5, 2023

October 5, 2023 528

Postscript and Acknowledgments

October 5, 2023 532

	October 5, 2023	538