The Futurological Model

Foundations of an Artificial Intelligence System Integrating the Scientific Knowledge of Humankind

Maël Donoso

ISBN: 978-2-8399-4106-8

The Futurological Model

Foundations of an Artificial Intelligence System Integrating the Scientific Knowledge of Humankind

Maël Donoso

Contents

1
Foundation

In the *Foundation* series, science fiction writer Isaac Asimov introduces the concept of psychohistory, a science of the future combining mathematics, psychology, and sociology. This fictional science would make it possible to predict historical events based on the knowledge of human behavior, by taking advantage of the fact that in very large groups, individual actions are averaged and statistical trends can be determined. In real life, the study of the future holds different names, such as futurology, futures studies, and futures research. Given the complexity of the human civilization, it is generally assumed that a precise forecasting of historical events would be extremely difficult, and these disciplines are usually more concerned with the exploration of alternative, plausible scenarios based on the current trends. This approach is indeed very useful, but the question of the future seems important enough to motivate further thinking. With the increasing amount of data and computing resources at our disposal, as well as the scientific and technological advances in data science, machine learning, deep learning, and artificial intelligence, it might be interesting to take some time and evaluate again whether a true, real-life science of the future could be possible. At its core, this updated version of futurology would need a predictive model of the world, which we could call a futurological model.

The optimistic view would be to imagine that the most important trends of the future could be forecasted relatively easily, by applying a simple set of laws, equations, or algorithms, waiting to be discovered. In this perspective, a working futurological model would likely consist in some clever mathematical formulation of psychological, sociological, economic, or political principles. Unfortunately, this optimistic view is unlikely to be true. Even if we could discover such principles, it remains difficult to see, for example, how environmental transformations like climate change, deforestation, water scarcity, air pollution, biodiversity loss, or the emergence of new diseases, which could all have a dramatic impact on the human civilization, may be taken into account by a shallow futurological model. It is also unclear whether the fundamentals of human mind and social interactions could be considered stable, in a world where computers, social networks, virtual reality, augmented reality, and soon brain-computer interfaces, will drive the human brain into uncharted territories. Since the lower levels of physics, chemistry, biology, and neuroscience influence the higher levels of psychology, sociology, economics, and politics, it seems likely that a successful futurological model should have a certain depth. In other words, it should cover a wide range of scientific knowledge, and integrate algorithms and data dedicated to the forecasting of natural phenomena. This requirement could be seen as an opportunity. Until now, futurology has been mostly focused on the evolution of human society, but other dimensions are interesting in their own right. If we envision a futurological model that includes physical, chemical, biological, and neurological principles, we might take the first step toward a more general futurology.

The pessimistic view, on the contrary, would be to consider that forecasting the most important trends of the future is essentially impossible, since it would require a complete simulation of the Earth or even the universe. In this perspective, a working futurological model could be achieved in theory, at least if we had a perfect knowledge of the laws of physics, and if these laws turned out to be deterministic. However, it would involve the computation of everything from the level of elementary particles and fundamental interactions, which would be practically impossible in terms of computing power, memory, and because we would need a complete knowledge of the current state of the Earth or universe. Fortunately, this pessimistic view is most likely exaggerated. Each level of organization higher than elementary particles, including atoms, molecules, organelles, cells, tissues, organs, systems, organisms, populations, communities, and ecosystems, has its own emerging order, and can be modeled independently to a certain extent. For example, there is probably no need to simulate a biological cell at the level of elementary particles, or even at the level of atoms, if we can characterize most of the cellular behavior using models from molecular and cellular biology. Therefore, each level of organization offers potential for a simplification of the simulation problem, depending of course on our objective, the precision required, and other elements.

What would be a more realistic view, then? Here are some hypotheses. A successful futurological model should integrate a wide range of scientific knowledge, and leverage the advances in data science, machine learning, deep learning, and artificial intelligence. It should model every level of organization, and the interactions between these levels, with dedicated algorithms and data. At each of these levels, it should pay a particular attention to complex adaptive systems, i.e., systems that can learn from their environment and act on the basis of this learning, since these systems are likely to be particularly interesting for futurology. Overall, most of these ideas are not new. Deep, interdisciplinary, multiscale research projects, using massive computing power and artificial intelligence, are currently a trend in the scientific community. The objective of the Human Brain Project[1], for example, was to “gain an in-depth understanding of the complex structure and function of the human brain with a unique interdisciplinary approach”, using “highly advanced methods from computing, neuroinformatics and artificial intelligence”, and with a focus on “multiscale investigation of brain networks, across different spatial and temporal scales”. Now, the human brain can be studied with a variety of well-established scientific methods, but how could we develop, train, and evaluate a futurological model? The future is a broad concept, and the possible dataset would include, among other sources, all the data ever acquired by scientific research, but even this amount of information would not be guaranteed to be sufficient. Moreover, the fundamental question would still be to determine how the futurological model should work, and it seems difficult to define even a starting point. Evolution followed roughly a trend toward complexity, and an interesting path might be to develop the futurological model in that order too, e.g., to model cells before multicellular organisms, simple nervous systems before complex brains, animal societies before the human civilization. But this is more an intuition than a plan.

The objective of this book is to explore to which extent this more realistic view could be followed, and whether it could be possible to outline the first steps of a plan that might lead to a working futurological model. This preliminary exploration may already seem too ambitious, or too complex, but the benefits of a better knowledge of the future cannot be overstated. Let us visualize time as a tree, where the roots represent the past, the trunk the present, and the branches the future. Some branches lead to positive futures, others to neutral ones, or negative ones, and some branches even lead to the termination of our species. In a period of history characterized by accelerated change, with an unprecedented level of uncertainty regarding the future, the difficulty to anticipate the positive, neutral, or negative outcomes of our decisions, and to avoid the paths that could lead to existential threats, is arguably one of the most complex challenges faced by humankind. Futurology will not define what a desirable future is, but it could give us a better understanding of how the future works, and once we agree on an objective, it could help us to make the right choices in order to attain that particular branch. If we succeeded in integrating the scientific knowledge of humankind into a futurological model, leveraging the advances in data science, machine learning, deep learning, and artificial intelligence, with an appropriate modeling of the different levels of organization and complex adaptive systems, we might be on a good path to build a strong foundation for the future of the human civilization.

[1] https://www.humanbrainproject.eu/en/science-development/scientific-strategy-vision/

2
Levels of Organization

Levels of organization are widely used in science, and in particular in biology, as a framework to structure the continuum of reality. A level of organization, also called an integrative level, could be defined roughly as a set of entities or processes emerging from lower-level components. For example, starting from elementary particles and fundamental interactions, entities such as atoms, molecules, organelles, cells, tissues, organs, systems, organisms, populations, communities, and ecosystems all have their own emerging order, and can be considered as distinct levels of organization. This framework is often intuitive in scientific research, but the idea of levels remains difficult to define precisely. Moreover, it is associated to equally difficult concepts, such as complexity, emergence, and reduction. In Chapter 1, we hypothesized that a successful futurological model should integrate a wide range of scientific knowledge, and model every level of organization, and the interactions between these levels, with dedicated algorithms and data. Therefore, we will now discuss how these levels might be integrated into a futurological model.

One method would be to structure explicitly the futurological model according to the levels of organization, by modeling separately the entities and processes associated with each level. In this design, we would develop exhaustive models of elementary particles, atoms, molecules, organelles, and all the other levels, by taking advantage of the large scientific literature at our disposal. These models should account for the diversity of entities existing at each level, like the different types of cells, or the different shapes of multicellular organisms. They should integrate our knowledge about the interactions occurring inside a particular level, like the chemical reactions between molecules, or the relations between populations. Furthermore, they should integrate our knowledge about the interactions occurring across levels, like the influence of a neurotransmitter on a neuron, or the influence of an invasive species on an ecosystem. We might think of a futurological model following this design as a program written in an object-oriented programming language, with a *Cell* class, an *Organism* class, etc. Each of these classes would come with a number of attributes and methods, which should be large enough to instantiate a series of *Cell* objects, *Organism* objects, etc., reflecting the complex entities and processes of our world. Since we could not expect to model everything down to the level of elementary particles, before running this program, we would need to develop a set of clever protocols, in order to define exactly what should be modeled for the futurological task of interest. Nevertheless, even if we could develop such protocols, given the complexity and diversity inside each level of organization, and the multiple interactions occurring across levels, we might expect that this strategy would lead to an inflation of attributes and methods, similar to the plethora of rules implemented in some early artificial intelligence systems.

Another method would be to avoid levels of organization entirely, and instead, to model directly all the mechanisms reported in the scientific literature, regardless of the levels at which these mechanisms occur. In this design, we would develop a multiplicity of models, spanning our entire scientific knowledge, to account for all known cascades of events. For example, we know that in the human visual pathway, when the retina receives a visual signal, a nervous signal is transmitted through the optic nerve, then the information is relayed by the thalamus, then the visual cortex is activated, etc. Since these events can be understood at several scales, we would also need to include multiscale hierarchies, sometimes called levels of mechanisms, where lower-level entities are components of higher-level processes. For example, the activation of the visual cortex means that neurons produce more action potentials, which means that the cell membranes are depolarized, which means that sodium atoms flow through the ion channels, etc. While levels of organization are defined in absolute terms, levels of mechanisms are only relative, and defined on a case-by-case basis. We might think of a futurological model following this design as a program written in a functional programming language, where a *retina()* function would call an *optic_nerve()* function, which would call a *thalamus()* function, which would call a *visual_cortex()* function, etc., while all of them would also interact with higher-level and lower-level functions. Much like levels of organization, cascades of events and levels of mechanisms are widely used in biology, and in particular in neuroscience, as a framework to understand complex systems, at least when these systems are considered in relative isolation. Now, if our objective is to develop a predictive model of the entire world, these functions should be called multiple times, in multiple contexts, and in a particular order. If we really wanted to avoid the objects of an object-oriented programming paradigm, we might be able to devise a sophisticated system of loops and conditional statements to define when a function should be called, and with what arguments. Nevertheless, given the vast and intricated web of interactions of the world, even if we could develop a set of clever protocols to simplify the problem, we might expect that this strategy would lead to an inflation of function calls and arguments, and could turn out to be even more complicated than the previous method.

Could we think of a better alternative to design a futurological model, then? An interesting perspective is given by William Wimsatt, who defines levels of organization as local maxima of regularity and predictability. This concept is described, for example, in a *Stanford Encyclopedia of Philosophy* article[2], where it is compared with other perspectives on levels of organization. According to the article, "patterns and regularities that can be used as a basis for prediction and explanation are found clustered around certain scales, and such clusters indicate levels of organization". In other words, levels of organization correspond to stable regularities of the world, where predictability is maximized. Of course, maximal predictability would also be a desirable property from the perspective of a futurological model, and this convergence suggests an idea. If we could develop a futurological model where levels of organization are flexible and could be adapted, we might be able to use artificial intelligence to find the most relevant levels, where relevance would be measured by the predictability obtained. For example, let us imagine that we start with a classical series of levels, from elementary particles to atoms, molecules, organelles, cells, tissues, organs, systems, organisms, populations, communities, and ecosystems. From this baseline, an artificial intelligence system could suggest a series of improvements, either by adding, subtracting, or modifying levels, even if some levels end up overlapping or limited in scope. For example, if we think in terms of object-oriented programming, would it be better, perhaps, to model independently *Animal*, *Plant*, and *Fungus* classes, instead of having a single *Organism* class? Would it make sense, perhaps, to add new classes for viruses, prions, macromolecules, multinucleate cells, large-scale brain networks, symbiotic relationships, and insect societies? Furthermore, a futurological model following this design might be regularized by adding a constraint on the total number of classes, attributes, or methods, which could help to control the inflation of rules. The exact way the problem should be framed is not obvious, but since we have an objective function, which is predictability, our best bet might be to use machine learning, including deep learning, to optimize the levels of organization integrated into the futurological model, instead of defining them solely on the basis of our scientific intuition.

Levels of organization are a difficult topic. It may also seem that these considerations are quite abstract, and that we are merely discussing equivalent ways to classify our current scientific knowledge. In fact, we are not. Each scientific model comes with its own hypotheses about the processes that generated the data, and a good choice of hypotheses can have a considerable impact on the accuracy of the predictions. Since the objective of a futurological model would be to forecast the most important trends of the future, and since these trends are likely to be generated by a multiplicity of complex processes, it is reasonable to assume that the structure of this model should be determined carefully. Also, whatever the data at our disposal concerning the past or present state of the world, a futurological model would certainly need to infer a considerable amount of missing information, and this extrapolation might increase the need for a good underlying structure.

Overall, we are most probably facing a very difficult problem, and levels of organization could be a good starting point to progress toward a solution. The *Stanford Encyclopedia of Philosophy* article highlights also the link between levels of organization and evolution: "The issue of evolutionary transitions is closely connected to the debates on levels of selection and levels of organization. Here the focus is on the emergence of new levels of organization through evolutionary processes. The background idea is that the complex hierarchical organization of nature that we observe today must itself be a result of evolution, and therefore requires an evolutionary explanation." We will discuss these ideas further in the next chapters, and eventually, we will explore how they might be extended from biological evolution to human history.

[2] https://plato.stanford.edu/entries/levels-org-biology/

3
Complex Adaptive Systems

Complex adaptive systems are used as a framework to characterize a range of dynamic entities and processes. A complex adaptive system could be defined as a system that can learn from its environment, and act on the basis of this learning. Often, such a system consists of many elements or agents, interacting in flexible, non-linear ways, with an emergent order that appears difficult to predict based on the individual interactions. For example, populations, cells, immune systems, nervous systems, and animal societies can be considered as complex adaptive systems. One of the best definitions, perhaps, is given by Murray Gell-Mann in his book *The Quark and the Jaguar*: "The common feature of all these processes is that in each one a complex adaptive system acquires information about its environment and its own interaction with that environment, identifying regularities in that information, condensing those regularities into a kind of 'schema' or model, and acting in the real world on the basis of that schema. In each case, there are various competing schemata, and the results of the action in the real world feed back to influence the competition among those schemata. Each of us humans functions in many different ways as a complex adaptive system." In Chapter 1, we hypothesized that a successful futurological model should pay a particular attention to complex adaptive systems. Therefore, we will start with a quick review of some complex adaptive systems that have emerged during evolution, before discussing the opportunities and challenges that these systems might represent for a futurological model.

Evolution, including prebiotic evolution, is the basis of biological complexity on Earth, and its fundamental adaptation process is natural selection. In a population, the individuals better adapted to a particular environment are more likely to survive and reproduce, transmitting their genes to the next generation. Therefore, a population can be considered as a complex adaptive system, where the genomes represent highly compressed models allowing the individuals to interact with their environment, and where natural selection influences the competition between these models. Moreover, driven by natural selection, a series of other complex adaptive systems have emerged during evolution. Cells can adapt to their chemical environment by modulating the expression of their genes. Immune systems can recognize specific pathogens and enhance the immune response when appropriate. Nervous systems can integrate a wide range of sensory information, perform actions, and evaluate the results of these actions. In the perspective of a futurological model, nervous systems are probably the most important complex adaptive systems to study, since they served as a foundation for other important complex adaptive systems, such as animal societies. In particular, the human brain served as a foundation for the human civilization based on language, and for all the complex adaptive systems found in human society, such as economic markets, political entities, the scientific enterprise, and artificial intelligence. As Murray Gell-Mann points out in his book, complex adaptive systems have a general tendency to generate other complex adaptive systems: "All these systems keep exploring possibilities, opening up new paths, discovering gateways, and occasionally spawning new types of complex adaptive system. Just as new ecological niches keep turning up in biological evolution, so new ways to make a living continue to be discovered in economies, new kinds of theories are invented in the scientific enterprise, and so on." The human brain has been particularly prolific in this respect.

A working futurological model should be able to predict, to a certain extent, the behavior of complex adaptive systems. Let us focus on the human brain, an adaptive system of extraordinary complexity, composed of approximately eighty-six billion neurons, and a similar number of glia cells. Each neuron has, on average, approximately seven thousand synaptic connections to other neurons. Moreover, while artificial neurons in deep learning models perform only simple calculations, biological neurons are much more complex computational units, capable of processing information in complicated ways. Nevertheless, we do not need to take all this complexity into account to make predictions about the human brain activity, or the resulting human behavior. In many neuroscience experiments, very simple models of cognitive functions, based for example on reinforcement learning or Bayesian inference, can provide relatively good predictions, despite all the underlying intricacies of neurons, glia cells, axons, dendrites, and neurotransmitters that support these functions. Is it not surprising that the human brain, with its extraordinary complexity, could be modeled so easily by a set of minimalist algorithms? In fact, the unreasonable effectiveness of neuroscience models seems understandable. Natural selection exerted strong evolutionary pressures on nervous systems, optimizing their capacity to learn and make decisions in a series of environments. Nervous systems, including the human brain, fulfilled this function by reusing and adapting the complex biological structures provided by evolution, but the same tasks could be expressed in a much simpler form, for example as machine learning algorithms running on a computer. In other words, evolutionary pressures toward optimization have shaped complex adaptive systems, constraining them to function in a relatively regular and predictable manner. Of course, we are not suggesting that the neuroscience models currently at our disposal are even remotely sufficient to explain human brain activity, or the resulting human behavior, in the most complicated real-life situations. Nevertheless, even if future neuroscience models become increasingly complex, it seems plausible that they will remain much simpler than the underlying biological reality of the human brain.

Whenever a complex adaptive system emerges, we might expect that evolutionary pressures will take place, and that this system will evolve as an optimized computational unit for learning and decision-making, adapted to its particular environment. Therefore, regardless of the actual processes behind this system, we might be able to predict its behavior, to a certain extent, using simplified algorithms. In Chapter 2, we suggested that a futurological model could be structured according to the levels of organization widely used in science, and in particular in biology. For example, if we think in terms of object-oriented programming, we could have a *Molecule* class, a *Cell* class, an *Organ* class, etc. We also suggested that levels of organization could be considered as stable regularities of the world, where predictability is maximized, and that we might be able to use artificial intelligence to find the most relevant levels, where relevance would be measured by the predictability obtained. Following a similar reasoning, if we consider that evolutionary pressures toward optimization have shaped complex adaptive systems into relatively regular and predictable entities, we might be able to use artificial intelligence to find the most relevant systems, where relevance would also be measured by the predictability obtained. The ideas of levels of organization and complex adaptive systems are closely linked, since when a new complex adaptive system emerges, it often creates a new level of organization. This seems logical, because the better guarantee that a new entity will survive and reproduce may be its capacity to learn from its environment, and to act on the basis of this learning. Taking this conceptual proximity into account, how could we integrate complex adaptive systems into a futurological model? Let us imagine a *Neurotransmitter* class inheriting from *Molecule*, a *Neuron* class inheriting from *Cell*, and a *Brain* class inheriting from *Organ*. When we consider a *Brain* object as a complex adaptive system, we are actually making the hypothesis that this *Brain* object can be modeled directly, as a computational unit for learning and decision-making, without the need to understand how its properties emerge from *Neuron* objects, *Neurotransmitter* objects, or any other lower-level instance. We simply expect that evolutionary pressures have arranged *Neuron* objects, *Neurotransmitter* objects, etc., in order for the *Brain* object to function as expected. Therefore, in the design of a futurological model, complex adaptive systems might be seen as possible modeling shortcuts, where we assume that several levels of organization work together inside an integrated entity.

Now, a deeper question would be whether we could anticipate the future steps in evolution, and predict the emergence of new levels of organization and complex adaptive systems. Let us assume, for a moment, that future neuroscience models will become accurate enough to predict the behavior of the human brain in most real-life situations. However, what if complex brain-computer interfaces were invented, allowing several human brains to exchange their ideas directly, and creating an entirely new ecosystem of shared minds, distributed emotions, and collective consciousness? Most neuroscience models considering the human brain as an independent complex adaptive system might become obsolete. In this new world, to forecast the most important trends of the future, perhaps we would need to focus more attention on the lower-level entities of the human brain, e.g., brain networks, cortical columns, neurons, glia cells, axons, dendrites, and neurotransmitters, in order to understand the new neurological processes generated by the interaction with the brain-computer interfaces. Eventually, we might be able to characterize the new, higher-order levels of organization and complex adaptive systems that would emerge through the integration of several human brains. This particular scenario is still in the realm of science fiction, but it suggests a concept for our futurological model, which we could call the computational continuum. Entities or processes emerging from lower-level components can be described directly, at their own level of organization, with a computationally light model. Or they can be described more completely, through several levels of organization, with a computationally heavy model. The trade-off between simple and complex models is very common in scientific research, but if our objective is to develop a predictive model of the entire world, being able to switch automatically between shallow and deep representations might be one of our core requirements. Complex adaptive systems seem particularly interesting as possible modeling shortcuts, since their behavior might be relatively well predicted by computationally light models, but this is only true when they live between certain frontiers, in the territory for which they were shaped by evolutionary pressures. When this assumption does not hold anymore, computationally heavy models might be needed. Ideally, we should avoid predictability gaps between levels of organization, in the sense that higher-level entities and processes should be explainable by the properties of lower-level entities and processes. Complex adaptive systems might be particularly challenging in this respect, and further scientific research is most certainly needed to improve our understanding of the most complex systems, such as the human brain.

Nevertheless, given the scientific knowledge already acquired by humankind, and the advances in data science, machine learning, deep learning, and artificial intelligence, we could start to envision a futurological model capable of selecting, for each element of the world, the right representation in the computational continuum, in order to make the best possible predictions, while using the least possible amount of computational resources.

4
Evolution

Charles Darwin ends his book *On the Origin of Species* with this famous sentence: "There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved." Highly ordered processes, such as the orbit of Earth mentioned by Charles Darwin, can be easy to predict. Highly disordered processes, such as the trajectories of individual molecules in a turbulent fluid, can be almost impossible to predict, but their average behavior can be described statistically. Of course, predictability depends on the spatial and temporal scale, and on many other variables. But as a general rule, interesting things, including life, evolution, and complex adaptive systems, tend to happen at the frontier between order and disorder, where prediction is possible yet challenging. In Chapter 1, we discussed the fact that since the lower levels of physics, chemistry, biology, and neuroscience influence the higher levels of psychology, sociology, economics, and politics, it seems likely that a successful futurological model should integrate algorithms and data dedicated to the forecasting of natural phenomena. In Chapters 2 and 3, we highlighted the importance of an evolutionary perspective, and wondered whether we could anticipate the future steps in evolution, and predict the emergence of new levels of organization and complex adaptive systems. Therefore, we will now explore how a futurological model could be applied to biological evolution, before being extended to human history.

In Chapter 2, we suggested that a futurological model could be structured according to the levels of organization widely used in science, and in particular in biology. For example, if we think in terms of object-oriented programming, we could have a *Molecule* class, a *Cell* class, an *Organ* class, etc. We also suggested that it might be possible to develop exhaustive models of all the levels of organization, by taking advantage of the large scientific literature at our disposal. However, we did not give any detail about these models. For example, what attributes and methods should we include in a *Cell* class? Well, it seems almost inevitable that a realistic cellular model should include the idea of genetic information. Genes occupy a central position in biology, and as such, they should certainly be integrated into a futurological model. Nevertheless, their role is not limited to the cellular level. Genes are among the most important attributes of cells, but their differential activation also determines the physiology of tissues, organs, systems, and organisms. Furthermore, the primary effect of natural selection is to influence the distribution of genes in a population. In Chapter 3, we suggested that evolutionary pressures toward optimization have shaped complex adaptive systems, constraining them to function in a relatively regular and predictable manner, and cellular processes of gene expression are a perfect example of this regularity and predictability. We do not need to simulate a cell down to the atomic or molecular level, with a computationally heavy model, to predict that a particular gene will code for a particular protein. We can simply assume that gene expression will follow an optimized path using a robust, and almost universal, genetic code. Given the central position of genes in biology, how could we integrate them into a futurological model? Should we perhaps consider genes as a distinct level of organization, and define a *Gene* class, on which the *Cell* class could fall back when necessary? However, other heritable units of information exist at the cellular level, such as epigenetic markers. Therefore, should we also define a *Marker* class? As we discussed in Chapter 2, our best bet might be to use machine learning, including deep learning, to optimize the levels of organization integrated into the futurological model, using the predictability obtained as our objective function. While the exact way the problem should be framed is still not obvious, we can already expect that it would require running realistic, multiscale simulations of biological processes. Further scientific research is most certainly needed to improve our understanding of gene expression, and other complex biological processes, but artificial intelligence has the potential to accelerate scientific discoveries. For example, AlphaFold, an artificial intelligence system developed by DeepMind, can already predict the tridimensional structure of a protein, based on its amino acid sequence.

Extrapolating from the current trends in scientific research and artificial intelligence, it does not seem unreasonable to envision a futurological model capable of simulating evolution, to a certain extent and in a given environment, assuming the core biological processes of information transmission remain constant. Now, what if this assumption does not hold anymore? Throughout evolution, a small number of major transitions have transformed the way in which information is stored, translated, and transmitted from one generation to the next. These critical events include, in particular, the transition from independent replicators to chromosomes, from prokaryotes to eukaryotes, from asexual reproduction to sexual reproduction, from unicellular organisms to multicellular organisms, from solitary animals to animal societies, and from primate societies to the human civilization. The emergence of multicellular organisms occurred at least three times, for animals, plants, and fungi. The emergence of animal societies occurred multiple times, for example for ants, termites, bees, and wasps. However, all the other major transitions might have been unique events, occurring in only one lineage, and this uniqueness is particularly challenging. In fact, the existence of such unique events, whether in biological evolution or in human history, might be one of the most profound challenges for futurology. We have only one world, and therefore, only one time series. We can design predictive models for processes that occurred on multiple occasions, such as the formation of a new species, the adaptation of a species to a new environment, and the extinction of a species, but how could we model the probability of an event that occurred only once? It might be difficult to find a more subtle solution than running multiple simulations through several levels of organization. If we had to fall back on this brute force strategy, our method could be the following. Starting from a set of initial conditions corresponding to the prebiotic environment, we might be able to run a futurological model multiple times, in order to evaluate the probability that life would appear. Then, assuming the existence of life, we might be able to evaluate the probability that eukaryotes would appear. Then, assuming the existence of eukaryotes, we might be able to evaluate the probability that sexual reproduction would appear, and so on for multicellular organisms, animal societies, and the human civilization. After each major transition, we would obtain a tree of future worlds, where only one branch, and not necessarily the most probable one, would lead to the next major transition. In other words, the timeline predicted by the futurological model would consist of successive ramifications. It is interesting to note that all the major transitions might have been accidental, with a low probability of occurrence. If a single one of these transitions had not happened, or if other, exotic transitions had occurred, our planet would probably be extremely different today.

Shortly before the end of *On the Origin of Species*, Charles Darwin writes an equally famous paragraph: "In the distant future I see open fields for far more important researches. Psychology will be based on a new foundation, that of the necessary acquirement of each mental power and capacity by gradation. Light will be thrown on the origin of man and his history." The human civilization was not created from emptiness. This elaborated form of animal society is based on the possibilities of the human brain, which is itself the product of evolution. In fact, the study of animal behavior can offer insights into almost all human abilities, including decision-making, learning, communication, cooperation, deception, conflict, territoriality, sexual behavior, and cultural transmission. Nevertheless, the high level of development of some cognitive functions, such as language, reasoning, and planning, seem to be typically human. It might be interesting to think beyond the major transitions in evolution, and imagine a more specific series of critical events, which we could call the major neurological and cognitive transitions, in the lineage leading to humankind. Throughout evolution, these events would have transformed the way in which the vertebrate brain, the mammalian brain, the primate brain, and eventually the human brain, interact with their environment. Once again, if a single one of these transitions had not happened, or if other, exotic transitions had occurred, the human brain as we know it would probably not exist. We are still quite far from fully understanding the evolution of the brain, let alone from running realistic, multiscale simulations of neurological processes, precisely enough to simulate the emergence of cognitive functions. Nevertheless, this is a path that we might need to follow, if our objective is to extend the idea of evolutionary transitions from biological evolution to human history. While genes and epigenetic markers occupy a central position in biology, neuronal connections and brain networks occupy a central position in neuroscience, and should certainly be integrated as well into a futurological model. However, there seem to be no perfect equivalent of the genetic code, no robust and universal dictionary that we could use to interpret neuronal connections. Should we consider the possibility of more abstract levels of organization? For example, if we model the transmission of genetic information with a *Gene* class, should we model the transmission of cognitive information, or ideas, with a more abstract *Meme* class? Climbing the abstraction ladder might be necessary if we intend, eventually, to extend the futurological model to human organizations, such as economic markets and political entities. Overall, the human brain might be the perfect laboratory to refine our ideas about levels of organization and complex adaptive systems, before extending them from biological evolution to human history.

In Chapter 1, we suggested that an interesting path to develop a futurological model might be to follow the evolutionary trend toward complexity, e.g., to model cells before multicellular organisms, simple nervous systems before complex brains, animal societies before the human civilization. We would naturally follow this path if our strategy consisted in trying to anticipate the future steps in evolution, major transition after major transition, in order to integrate more and more levels of organization and complex adaptive systems into the futurological model. Now, forecasting the future steps in evolution would be a difficult task, but discovering the past steps is not easy either. As we mentioned earlier, we have only one world, and therefore, only one time series. Moreover, this time series is severely incomplete. Our picture of evolution is based on several sources, such as fossils, geological data, and molecular data, but our knowledge is most probably very limited, if we consider the high diversity of species that may have lived on Earth. As a consequence, a futurological model trained on evolutionary data would have multiple ways of filling the gaps, since multiple timelines would be compatible with the current state of the world. The future is uncertain, but so is the past. In fact, even the current state of the world is not known perfectly. For example, we do not know all the species on Earth, and we have certainly not sequenced all the genes, or studied all the proteins, of the species we know. Therefore, we might guess that a futurological model would need to run on multiple scenarios, representing different sets of initial conditions, and use machine learning approaches to extrapolate from the data at our disposal. In the tree of future worlds, we have many possible branches ahead of us, but also many possible roots behind us.

5
Major Transitions in Evolution

In their book *The Major Transitions in Evolution*, John Maynard Smith and Eörs Szathmáry discuss the increase of biological complexity throughout evolution: "Our thesis is that the increase has depended on a small number of major transitions in the way in which genetic information is transmitted between generations. Some of these transitions were unique: for example, the origin of the eukaryotes from the prokaryotes, of meiotic sex, and of the genetic code itself. Other transitions, such as the origin of multicellularity, and of animal societies, have occurred several times independently." In their book *Evolution in Four Dimensions*, Eva Jablonka and Marion J. Lamb adopt a relatively similar approach, centered on the transmission of information, but extend the concept of heredity beyond the genetic system, in order to include the epigenetic, behavioral, and symbolic systems. In Chapter 4, we suggested that the existence of unique critical events, whether in biological evolution or in human history, might be one of the most profound challenges for futurology. We can design predictive models for processes that occurred on multiple occasions, but how could we model the probability of an event that occurred only once? We hypothesized that it might be difficult to find a more subtle solution than running multiple simulations through several levels of organization, a brute force strategy that might be applied after each major transition, in order to predict the probability of the next major transition. However, we did not develop this idea further, or give any specific example of the difficulties involved. Therefore, we will now explore the challenges of applying a futurological model to the major transitions in evolution, by discussing the emergence of eukaryotes, sexual reproduction, multicellular organisms, nervous systems, and the human civilization.

Let us imagine a biologist equipped with a working, but limited, predictive model, and observing the Earth in real time, starting from the beginning of the Precambrian. The first living creatures in the primitive ocean are the bacteria, a diversity of prokaryotic, asexual, unicellular organisms, maybe accompanied by semi-living creatures such as viruses. Extrapolating from this situation, our biologist would certainly predict that life on Earth will continue this trend, and simply generate an increasing diversity of bacterial species. Right? Wrong. Instead, eukaryotes emerge, with a new cellular architecture including a nucleus, and different types of organelles and specialized cellular structures, such as mitochondria, chloroplasts, and the Golgi apparatus. Eukaryotic cells are typically much more complex than prokaryotic cells, and they can harbor a much larger genome, opening new paths for life. Now, with only a limited number of organelles and specialized cellular structures, the possibilities of this new system seem underexploited. Our biologist would certainly predict that life on Earth will continue this trend, and generate even more complex eukaryotic cells, with an increasing diversity of specialized cellular components. Right? Wrong. A limited number of organelles and specialized cellular structures turned out to be a good trade-off. Instead, sexual reproduction emerges, with a new type of cellular division, the meiosis. In sexual populations, genetic recombination can increase diversity, allowing for a faster adaptation to a changing environment. When the gametes have different sizes or forms, two sexes are defined, male and female. Now, again, with only two sexes, the possibilities of this new system seem underexploited. Our biologist would certainly predict that life on Earth will continue this trend, and generate even more complex types of sexual reproduction, with three sexes, four sexes, or more. Right? Wrong. Two sexes turned out to be a good trade-off. Instead, multicellular organisms emerge, with a new organism architecture consisting of several cells. The emergence of multicellular organisms occurred at least three times, for animals, plants, and fungi. During development, multicellular organisms acquire their particular structure, with a cellular specialization allowing for complex physiological functions. Now, again, with only a limited number of forms, the possibilities of this new system seem underexploited. Our biologist would certainly predict that life on Earth will continue this trend, and generate an increasing diversity of physiological structures and functions. Right? Wrong. We are starting to see the pattern here.

Some forms of multicellular organisms, such as the tetrapod organization for amphibians, reptiles, birds, and mammals, turned out to be a good trade-off. Instead, nervous systems of increasing complexity emerge, and serve as a foundation for a diversity of cognitive functions, including decision-making, learning, reasoning, planning, and strategic thinking, as well as the ability to use language. In fact, John Maynard Smith and Eörs Szathmáry consider the emergence of human language as one of the major transitions in evolution. Now, again, with only a limited brain complexity, the possibilities of nervous systems seem underexploited. Our biologist would certainly predict that life on Earth will continue this trend, and generate even more complex brains, with even more diverse cognitive functions. Right? Wrong. Despite its limitations, the human brain turned out to be a good trade-off, or at least, sufficient to serve as a foundation for the emergence of a civilization. Hopefully, this short overview of the major transitions in evolution might provide an intuition of the challenges that we should face, if our objective is to build a working futurological model. Linear extrapolation from a trend is one of the easiest ways to predict the future, but this strategy would be far from sufficient here. As we discussed previously, it might be difficult to find a more subtle solution than running multiple simulations through several levels of organization. Nevertheless, once a major transition happens, it might open some opportunities for simplification. As John Maynard Smith and Eörs Szathmáry point out in their book, many of the major transitions share a common feature: "Entities that were capable of independent replication before the transition could afterwards replicate only as part of a larger whole." Organelles such as mitochondria and chloroplasts originated most likely from symbiotic prokaryotes, but now they can replicate only as part of the whole eukaryotic cell. Eukaryotic cells, in a multicellular organism, can only transmit their genes to the next generation if the whole organism does. Multicellular organisms, in an animal society, can only replicate as part of the whole social group. This common feature might be interesting for a futurological model, since it suggests that when a new level of organization or complex adaptive system emerges, additional evolutionary pressures might be exerted into its lower-level components, constraining them to function in a more regular and predictable manner. This might provide some explanation to our repeated observation that the trends opened by major transitions never seem to be completely continued. Instead, biological systems seem to evolve until a good trade-off is found, serving as a basis for the next major transition. Once they reach this stage, their evolution becomes limited by the constraints exerted by the new, higher-order biological systems.

Building a futurological model would be a difficult task, since it would involve integrating the scientific knowledge of humankind into a common framework. We might guess that each major transition in evolution would add a new level of complexity to this project, and extend the range of scientific knowledge that should be integrated, from biochemistry to molecular biology, cellular biology, developmental biology, neuroscience, and so on. Therefore, we should be looking for every simplification opportunity, and explore several possible ways to build the futurological model, in order to take advantage of all the regularities of the world that could be used as a basis for prediction. In Chapter 3, we introduced the concept of the computational continuum, as a way to describe the trade-off between simple and complex models. Entities or processes emerging from lower-level components can be described directly, at their own level of organization, with a computationally light model. Or they can be described more completely, through several levels of organization, with a computationally heavy model. Now, instead of "computationally light" and "computationally heavy", we could use the terms "heuristic" and "algorithmic". Entities or processes can be described with a heuristic model, in the sense that their behavior can be computed as a global approximation. Or they can be described with an algorithmic model, in the sense that their behavior can be computed as the exact result of the interactions of their lower-level components, considered as axioms. In a series of algorithmic-heuristic relations, the heuristic model of one level can serve as an axiom for the algorithmic model of a higher level. Of course, an algorithmic model does not guarantee a perfect accuracy, or even a better accuracy than a heuristic model, since it is likely that our knowledge of the lower-level components and their interactions would be limited. Nevertheless, the algorithmic approach might be necessary when a new entity or process is emerging, and until it is stable enough. We have mostly discussed the major transitions in evolution as punctual events, but in a *PNAS* article[3], Eörs Szathmáry further decomposes these transitions into phases of origin, maintenance, and transformation of the higher-level units. During the early phases, we might expect that a newly formed higher-level unit would still be unstable, and that we might need to fall back more often on an algorithmic model, at the expense of a greater computational cost. Then, once the higher-level unit stabilizes, we might be able to rely more systematically on a heuristic model. We use the terms "heuristic" and "algorithmic" in a rather special sense, but hopefully this terminology would convey the idea that entities and processes of our world can be modeled using shallow or deep representations, and that a futurological model should be able to switch automatically between both, in order to find a good compromise between accuracy and computational cost.

Understanding all the implications of the major transitions in evolution could require a considerable amount of knowledge in physics, chemistry, biology, and neuroscience. Nevertheless, an interesting way to strengthen our intuition about these transitions might be to imagine the alternative, exotic events that could have occurred, and try to predict how life on Earth would have evolved in these alternative worlds. Could we imagine living forms based on silicon, instead of carbon? A genetic code with eight bases? Giant eukaryotic cells with a thousand types of specialized organelles? Sexual reproduction with ten sexes? An ecosystem dominated by walking trees and swimming mushrooms? A society of insect societies, federating ants, termites, and bees? A civilization ruled by a symbiotic alliance of talking owls and parrots? These scenarios might seem overly speculative, but they reveal what could be one of the most interesting applications of a futurological model: the search for extraterrestrial life. If living forms exist elsewhere in the universe, their biology, and even their chemistry, could be extremely different from ours. As we discussed in Chapter 4, all the major transitions on Earth might have been accidental, with a low probability of occurrence. On other planets, satellites, or asteroids, living forms could have evolved following entirely distinct paths. In a *Scientific American* article[4], some researchers argue that even the words “life” or “biology” could have little sense outside Earth, and that we should instead search for more general “astro life” or “lyfe”. Even if the fundamentals of extraterrestrial life were very similar to ours, a single evolutionary accident could already create a major divergence. For example, in their book, Eva Jablonka and Marion J. Lamb describe the fictional planet Jaynus, where evolution is mostly driven by the epigenetic inheritance system, allowing the organisms to be very diverse despite sharing exactly the same genome. Now, if we really want to advance the search for extraterrestrial life, we might need to explore more systematically the range of possibilities. We might need to integrate our knowledge of physics, chemistry, biology, and neuroscience into a common framework, in order to simulate multiple timelines and evolutionary paths. Eventually, we might be able to compute some preliminary probabilities, and obtain a better evaluation of the types of living forms that might be found in the universe. In other words, the search for extraterrestrial life might require a futurological model.

[3] https://www.pnas.org/doi/10.1073/pnas.1421398112

[4] https://www.scientificamerican.com/article/the-search-for-extraterrestrial-life-as-we-dont-know-it/

6
Strategist Brain

Nothing in biology or neuroscience makes sense, except in the light of evolution. However, nothing in psychology, sociology, economics, or politics makes sense either, except in the light of the human brain. The human brain can be considered from two different perspectives, as our way to understand the world and choose our objectives for the future, or as an element of the world that should be integrated into a futurological model. This duality, the brain as a subject or the brain as an object, might be particularly challenging for futurology, and should most certainly motivate further thinking. Nevertheless, in this chapter, we will focus on the second perspective, and explore how our knowledge of the human brain might be integrated into a futurological model. As we discussed in Chapter 3, while the human brain is an adaptive system of extraordinary complexity, we do not need to take all this complexity into account to make predictions about its activity, or the resulting human behavior. In many neuroscience experiments, very simple models of cognitive functions, based for example on reinforcement learning or Bayesian inference, can provide relatively good predictions, despite all the underlying intricacies of neurons, glia cells, axons, dendrites, and neurotransmitters that support these functions. We explained this unreasonable effectiveness of neuroscience models by arguing that natural selection exerted strong evolutionary pressures on nervous systems, optimizing their capacity to learn and make decisions in a series of environments, and constraining them to function in a relatively regular and predictable manner. However, we did not develop this idea further, or give any specific example. Therefore, we will now explore how several cognitive functions could be modeled with a relatively simple framework, based on our knowledge of the human brain. We will focus on the executive function, i.e., the set of cognitive processes allowing the brain to select and execute actions in order to attain chosen objectives, and we will consider only one level of organization, the level of brain regions.

Let us imagine a neuroscientist trying to integrate our knowledge of the human brain into a futurological model. He or she would certainly like to model the executive function with a biologically plausible approach, based on the functions of actual brain regions. For example, a reinforcement learning model based on our knowledge of basal ganglia could provide relatively good predictions, at least for very simple and limited tasks and environments. Basal ganglia include several subcortical structures, with the dorsal striatum more involved in decision-making and task-switching, and the ventral striatum more involved in reward processing and reinforcement learning. Now, in order to extend the model to more abstract and complex tasks and environments, our neuroscientist would likely try to integrate the role of several cortical regions, in particular regions from the prefrontal cortex. At this point, a promising approach could be to gradually extend the model by adding successive modules, with each module corresponding to a particular cognitive function, implemented in a particular brain region. If our neuroscientist decided to follow this path, what could be a good starting point? Value. The human brain can go beyond simple rewards, and estimate the value of complex behavioral strategies. When several strategies are possible, the ventromedial prefrontal cortex (VMPFC) encodes the value of the chosen option, and the reliability of this choice. Now, what could be next? Exploration. The human brain can go beyond the current choice, and estimate the value of a limited number of alternative options. The frontopolar cortex (FPC) encodes the value of these alternative options. Now, what could be next? Learning. The human brain can go beyond simple reinforcement, and estimate the value of strategies using Bayesian inference. The anterior cingulate cortex (ACC) guides behavior according to the history of actions taken and results obtained, and evaluates whether the current choice should be reconsidered. Now, what could be next? Decision-making. The human brain can go beyond simple stimuli, and consider the context before choosing an action or strategy. The dorsolateral prefrontal cortex (DLPFC) controls the implementation of abstract rules, and guides behavior according to a hierarchy of sensory, contextual, and episodic information. Now, what could be next? Reasoning. At this point, our neuroscientist would likely benefit from a common framework integrating value, exploration, learning, and decision-making.

In a *Science* article[5] that I authored together with Anne G. E. Collins and Étienne Koechlin, we proposed a brain system that could serve as a foundation for human reasoning. Importantly, this system should explain the human ability to select behavioral strategies in uncertain, changing, and open-ended environments, while relying on a biologically plausible computational model. The major characteristics of this computational model are the following. At each point in time, the current strategy, called the actor, is adapted using reinforcement learning, while the absolute reliability of every monitored strategy is evaluated using Bayesian inference. When the actor is considered reliable, the system is in an exploitation period. However, if the actor becomes unreliable, the system switches to an exploration period, and a new strategy, called the probe, is formed as a weighted mixture of strategies stored in long-term memory. This probe becomes the new actor, and while it typically starts as an unreliable strategy, it learns and can improve its reliability over time. The model can return to an exploitation period in two ways. Either one of the alternative strategies becomes reliable, in which case the probe is rejected while this alternative strategy becomes the new actor. Or the probe becomes reliable, in which case it is confirmed and consolidated into long-term memory. Therefore, the trade-off between exploration and exploitation relies on three algorithmic transitions: switch, rejection, and confirmation. Using model-based functional magnetic resonance imaging (fMRI), we tested the predictions of this computational model in a reasoning task, with an experimental design introducing the concept of algorithmic transitions in fMRI research. The results were the following. The VMPFC, known to encode the value of the chosen option, correlates indeed with the reliability of the actor. The FPC, known to encode the value of the alternative options, correlates indeed with the reliability of the alternative strategies. The ACC, known to guide behavior according to the history of actions taken and results obtained, correlates indeed with the switch transition. The DLPFC, known to control the implementation of abstract rules, correlates indeed with the rejection transition, which implies the recovery of an alternative strategy. One of the most significant findings might be that the ventral striatum, known for its involvement in reward processing and reinforcement learning, correlates with the confirmation transition, which implies the validation of the probe. In other words, the activation of the ventral striatum indicates that a new strategy has successfully emerged. This particular function of the reward system had never been observed before, and while more research is certainly needed, it does not seem unreasonable to think that this process might serve as a foundation for the intellectual pleasure of discovery, and might have something to do with the origin of scientific curiosity. Another interesting finding was that the FPC correlates with the reliability of several alternative strategies, and not just one, as was previously suggested. The number of monitored strategies predicted by the computational model was consistent with the capacity limit of human working memory: three or four strategies. Why only three or four? Maybe, like we observed repeatedly in Chapter 5, the executive function of the human brain evolved until a good trade-off was found, serving as a basis for further adaptations.

The human brain is a strategist, but strategy is not limited to control. In fact, the executive system can withdraw some control during the exploration periods, in order to obtain the flexibility required to test several alternative solutions. It is during these exploration periods that a trade-off takes place between the recovery of a previous actor and the validation of the probe, or in other words, between the old and the new, between the past and the future. In Chapter 4, we noticed that interesting things tend to happen at the frontier between order and disorder, and the three algorithmic transitions, i.e., switch, rejection, and confirmation, stand exactly on this frontier. Now, the brain system we just described could serve as a foundation for human reasoning, but integrating the executive function of the human brain into a futurological model would most certainly require additional modules. What could be next? Planning. The human brain can go beyond simple actions, and plan sequences of delayed actions. If our neuroscientist could integrate reasoning and planning into a single framework, he or she would certainly come a step closer to understanding strategic thinking. Nevertheless, at a certain point, such a framework would certainly become increasingly complex, and the approach consisting in adding successive modules might yield diminishing returns, in particular because the possible interactions between modules would increase exponentially. If our objective is to build a comprehensive model of the human brain, we might need to think deeper, and try to develop a better framework, for example by reorganizing the modules in a more integrated manner. The natural way to do this would be to consider the question from an evolutionary perspective, and investigate whether some modules can serve as a basis for others. In Chapter 4, we suggested that it might be interesting to think beyond the major transitions in evolution, and imagine a more specific series of critical events, the major neurological and cognitive transitions, in the lineage leading to humankind. Throughout evolution, these events would have transformed the way in which the vertebrate brain, the mammalian brain, the primate brain, and eventually the human brain, interact with their environment. We have presented a modular view of the executive function, and discussed successively value, exploration, learning, decision-making, reasoning, and planning, but we have not specified if this particular order makes sense in an evolutionary perspective. In fact, it probably does not. For example, the ability to monitor the reliability of alternative strategies, as implemented in the FPC, seems to be a very recent function, and might even be uniquely human. Developing a better framework would probably require an evolutionary model of how the cognitive functions of the human brain have emerged, taking into consideration that all the intermediate forms should have been functional. Furthermore, this evolutionary model might need to be deeper than the framework we discussed, since in order to understand the emergence of a new cognitive function, we might need to model levels of organization lower than brain regions. As we discussed in Chapter 5, a newly formed higher-level unit could be unstable, forcing us to fall back more often on an algorithmic model, at the expense of a greater computational cost. Overall, as we might have guessed, modeling a complex adaptive system as deeply integrated as the human brain is far from trivial, and poses a series of profound challenges.

The time seems right to introduce what could be seen as a Copernican Revolution in our conception of a futurological model. A working futurological model would require the integration of a wide range of scientific knowledge, but it might also become a powerful scientific platform for testing new theories. For example, some hypotheses about the human brain or its evolution should be more probable than others, more compatible with our knowledge in molecular biology, cellular biology, ecology, linguistics, psychology, or economics. An artificial intelligence system integrating the scientific knowledge of humankind could automatically evaluate this compatibility, and constrain the set of possible theories in one domain by the knowledge acquired in all domains. In other words, scientific research on the human brain, and other deeply integrated complex adaptive systems, might strongly benefit from a futurological model. Thinking strategically, we might envision an iterative process, where even a limited progress in one domain could automatically improve the predictive power in other domains, unlocking a virtuous circle. For example, even a simple model of the human brain, such as the one we discussed, could improve our knowledge of human psychology, which could improve our knowledge of behavioral economics, which may give us better insights into the origin of currency, the Neolithic economy, the Neolithic ecology, the Paleolithic ecology, the evolution of primate societies during the Paleolithic, and eventually, the evolution of the human brain. In Chapter 1, we mentioned the fact that deep, interdisciplinary, multiscale research projects, using massive computing power and artificial intelligence, are currently a trend in the scientific community. Everywhere we look, scientific research faces increasingly difficult problems, which are impossible to solve with a simple set of laws, equations, or algorithms. Eventually, and almost inevitably, the solution requires the integration of different predictive models, at different levels of organization, and often with an evolutionary or historical perspective. Whatever the starting point, we end up working toward futurology, with increasingly deep, interdisciplinary, and multiscale scientific models, gradually overlapping in scope. Until now, these efforts have been disparate, but if these scientific projects could join forces in order to develop a common scientific platform, we could be able to improve our knowledge of the world at an unprecedented rate. Therefore, while building a futurological model would be a difficult task, this project might also be the natural horizon of scientific research. Of course, further advances in data science, machine learning, deep learning, and artificial intelligence are most certainly needed. Interestingly, these advances are often inspired, or validated, by discoveries in neuroscience. In a *Neuron* article[6], Demis Hassabis, Dharshan Kumaran, Christopher Summerfield, and Matthew Botvinick review some of the contributions of neuroscience to artificial intelligence, and argue that such contributions could become even more important in the future. When human intelligence and artificial intelligence work together, nothing seems impossible, not even building a futurological model.

[5] https://www.science.org/doi/abs/10.1126/science.1252254

[6] https://www.cell.com/fulltext/S0896-6273(17)30509-3

7
Civilization

The human civilization is best understood by those who decided to simulate it. Here is what we can read in the *Civilization II* video game manual: "The major dynamic of change throughout the history of civilization has been the continuing advance and accumulation of knowledge. As humankind progressed by fits and starts through the ages, civilizations rose and fell, their success or failure due to what knowledge they acquired and how they employed it. Those who first acquire new knowledge are often able to employ it to build a more powerful position, but there are many cases of civilizations that obtained some new invention first, then failed to use it to their advantage. The pace at which a society develops and implements new knowledge depends on many factors, including its social organization, economic organization, geographic location, leadership, and competition." In Chapters 3 and 4, we mentioned the fact that all the complex adaptive systems found in human society, such as economic markets, political entities, the scientific enterprise, and artificial intelligence, are based on the possibilities of the human brain, which is itself the product of evolution. In Chapter 6, we described a brain system that could serve as a foundation for human reasoning, and explain our ability to select behavioral strategies in uncertain, changing, and open-ended environments. Importantly, we mentioned that a particular brain structure might be associated with the intellectual pleasure of discovery, a process critical to the acquisition and transmission of knowledge in the human civilization. We will now explore how the complex adaptive systems found in human society could be integrated into a futurological model, taking into consideration that all these systems rely on a particular organization of knowledge.

When we think about futurology, most of the time, our focus is on human history, not biological evolution. Nevertheless, there is a good reason why, in this book, we took a long detour through natural phenomena. In Chapter 1, we discussed the fact that since the lower levels of physics, chemistry, biology, and neuroscience influence the higher levels of psychology, sociology, economics, and politics, it seems likely that a successful futurological model should have a certain depth. Now, the comparison between biological evolution and human history reveals an interesting pattern. In the natural world, levels of organization and complex adaptive systems are organized more “vertically”, with a relatively clear hierarchy of entities and processes building on top of each other, for example elementary particles, atoms, molecules, organelles, cells, tissues, organs, systems, organisms, populations, communities, and ecosystems. By contrast, in the human world, levels of organization and complex adaptive systems are organized more “horizontally”, with almost all entities and processes built on top of one building block: the human brain. Of course, the history of human societies has always depended on multiple variables, such as geography, climate, mineral resources, plant species, and animal species. Nevertheless, the acquisition and transmission of knowledge by human brains is arguably the major driving force of the human civilization, and this observation might have strong implications for the design of a futurological model. Economic markets, political entities, and other complex adaptive systems found in human society are certainly predictable, to a certain extent, using our knowledge in sociology, economics, and politics. However, companies, markets, cities, and states only exist because a group of human brains share compatible representations of what these companies, markets, cities, and states should be, and because they act according to these representations. Google, the dollar, Paris, the French Republic, the European Union: none of these entities would make any sense without these coordinated representations and actions, and the same is true for all the hierarchical levels inside these entities. As Yuval Noah Harari states in his book *Sapiens*: “There are no gods, no nations, no money and no human rights, except in our collective imagination.” Furthermore, while economic markets, political entities, and other complex adaptive systems found in human society can adapt to their respective environments, the primary component of these environments is always the same: a group of human brains, whether acting as workers, consumers, citizens, voters, or in any other role. In Chapter 5, we suggested that entities or processes can be described with a heuristic model, in the sense that their behavior can be computed as a global approximation, or with an algorithmic model, in the sense that their behavior can be computed as the exact result of the interactions of their lower-level components, considered as axioms. In the human civilization, most algorithmic models would fall back on the same axiom: the human brain.

Therefore, if our objective is to apply a futurological model to the human civilization, integrating a model of the human brain might be our first requirement. Our second requirement might be to extend this model in order to include the acquisition and transmission of knowledge, which is the basis of all the complex social and cultural behaviors associated with civilization. In fact, sometimes, it almost feels like a science is missing. Somewhere between neuroscience, psychology, sociology, and linguistics, we could have expected the emergence of a comprehensive science of ideas, whose object of study would have been the formation, adaptation, and transmission of cognitive information between animal brains, and in particular between human brains. In Chapter 4, we suggested that the transmission of genetic information could be modeled with a *Gene* class, and wondered whether the transmission of cognitive information, or ideas, might be modeled with a more abstract *Meme* class. Given the increasing amount of neurological, psychological, sociological, and linguistic data at our disposal, the emergence of a comprehensive science of ideas might just be a matter of time. If a futurological model integrating our knowledge in physics, chemistry, biology, and neuroscience could be extended to include the acquisition and transmission of knowledge, we might be able to explore systematically human history, following the same strategy as we did for biological evolution. Starting from a set of initial conditions corresponding to the Neolithic, including geography, climate, mineral resources, plant species, and animal species, we might be able to run the futurological model multiple times, in order to evaluate the probability of emergence of different technologies, such as agriculture, masonry, or pottery. Then, assuming the existence of these advances, we might be able to evaluate the probability of further advances, such as metallurgy, the wheel, or writing, and so on for all the major transitions in civilization. As human society progressed through the ages, harnessing the power of entities and processes from different levels of organization, the amount of acquired knowledge has increased, as well as the artificialization of the world. In the perspective of a futurological model, the emergence of computers, in particular, might appear as a turning point. In a world where human brains interact increasingly often with computers, and where these machines serve as a basis for an increasing number of important operations, computing systems should definitely be integrated into a futurological model. Now, this integration might pose interesting challenges. Computers have their own levels of organization, or more appropriately, levels of abstraction, including hardware platforms, machine languages, assembly languages, virtual machines, high-level programming languages, operating systems, and computer networks. Theoretically, if we had access to every hardware blueprint, and every source code from every level of abstraction, we should be able to emulate all the computers in the world, and simulate a large part of the human civilization. Practically, it should be possible to simulate some characteristics of computing systems without taking all their technological complexity into account, just as we can simulate some characteristics of the human brain despite all the underlying intricacies of neurons, glia cells, axons, dendrites, and neurotransmitters. In his book *An Introduction to Mathematics*, Alfred North Whitehead writes: "Civilization advances by extending the number of important operations which we can perform without thinking about them." Computing systems have considerably extended this number of operations, and an important question for futurology could be to determine if computers have made our civilization less predictable, more predictable, or predictable in a different way.

Two complex adaptive systems found in human society may require further thinking. The first one is the scientific enterprise. In the *Civilization II* video game, the player can research a series of civilization advances, organized in a technology tree. However, in the real world, we do not know the technology tree in advance.

Even a futurological model may not be able to predict the next scientific discoveries, or the next technological inventions, since such predictions may be undistinguishable from actually making these discoveries or inventions. Nevertheless, a futurological model could constrain the set of possible discoveries or inventions in one domain by the knowledge acquired in all domains. For example, based on our current knowledge in physics, chemistry, and biology, we might be able to engineer a bacterium capable of degrading plastic in the ocean, but it is more difficult to imagine that we will discover a virus capable of performing nuclear fusion. This example is intuitive, but an artificial intelligence system might be able to go beyond human intuition, and evaluate the probabilities of a wide range of discoveries or inventions. Furthermore, a futurological model might be able to evaluate which research directions are likely to have the most positive impact on the human civilization, and this evaluation could help us to choose the future orientations of the scientific enterprise. The second complex adaptive system that may require further thinking is artificial intelligence. Artificial intelligence could serve as a foundation for futurology, but it is also a variable that should be integrated into a futurological model. Furthermore, the possibility that an artificial general intelligence could surpass the human brain, and eventually replace the human civilization, has been discussed both in science fiction and in academic publications. The hypothetical point in time after which the technological acceleration would become irreversible, for example in a scenario where an artificial general intelligence would acquire the power to upgrade itself, is known as the singularity. Of course, the termination of our species is usually considered to be an undesirable event. Humankind is still in its technological adolescence, and most adolescents do not feel ready to have children of their own, especially if these children turn out to be killer robots. Indeed, from an evolutionary perspective, there is a good reason why we should advance carefully toward artificial general intelligence. There are many differences between the human brain and artificial intelligence, but if we had to choose a single distinctive trait between them, what could it be? Motivation. The human brain emerged after billions of years of evolution, and was shaped to ensure survival and reproduction. It may not be the perfect thinking machine for ruling a civilization, but it turned out to be a good trade-off, and was able to stand the test of time. By contrast, computers were engineered to solve any problem, to pursue whatever objective their programmers define. Since artificial intelligence systems do not currently have motivations of their own, it is difficult to imagine what kind of consciousness they would develop, or what kind of world they would build. In fact, even when artificial intelligence systems are supposed to pursue an objective defined by humans, this objective should be carefully specified in order to prevent undesirable outcomes, a challenge known as the alignment problem. For the foreseeable future, and until this motivation issue is resolved, it seems reasonable to keep consciousness in the realm of humankind and other biological forms, and instead, to use artificial intelligence to discover the diversity of future worlds standing in front of us, and the multiplicity of possible trajectories that the human civilization could follow. In other words, it may be wiser, for the moment, to choose the plurality over the singularity.

How would the human civilization evolve if a working futurological model was built? In Chapter 1, we visualized time as a tree, where the roots represent the past, the trunk the present, and the branches the future. Some branches lead to positive futures, others to neutral ones, or negative ones, and some branches even lead to the termination of our species. Now, while scientific objectivity is a valuable ideal, almost all dimensions of this tree of future worlds would require subjective decisions. The predictions of a futurological model would most likely be expressed as probabilities, and the incredibly complex set of possible future worlds should be more similar to a continuous cloud than to a discrete tree. Therefore, we should decide which elements of the world are particularly relevant for the futurological task of interest, and discretize the predictions according to these elements, in order to obtain well-defined scenarios for the future. We could call this process: futurological perception. Furthermore, the predictions of a futurological model would be of little practical use, unless we have some leverage to attain a particular future world. Therefore, we should decide which actions could be performed in order to influence the future, whether by public investment, economic incentives, direct regulation, or any other means. We could call this process: futurological action. Intuitively, the more resources we could engage in a futurological action, the more leverage we would have for reaching the future world of our choice, taking into consideration that some future worlds could have a higher futurological cost, in the sense that more resources could be needed to attain them from our current position in time. Furthermore, choosing a branch in the tree of future worlds would require an agreement about what a desirable future is. Therefore, we should decide by which political and institutional means this agreement should be reached. We could call this process: futurological decision. Eventually, the possibilities of a futurological model might need to be voluntarily limited, since some developments could be problematic in terms of privacy, free will, and human rights. Would it be right to use a futurological model to predict individual behavior? Would it be wise to use a futurological model to predict the future desires of humankind, or the future requests that we could address to a futurological model? To which extent could we simulate the human brain without actually creating an artificial consciousness? These deep questions, among others, should most certainly motivate further thinking. We could call the limits voluntarily fixed to a futurological model: futurological frontiers. Overall, the evolution of the human civilization if a working futurological model was built would depend on a series of subjective decisions along all these dimensions. Nevertheless, a futurological model might become a powerful self-corrective system, and evolve gradually toward more scientific objectivity. Futurology will not define what a desirable future is, but futurological perception, action, decision, and frontiers could be adapted once we gain a better understanding of how the future works, and a better knowledge of the most important variables to consider. Moreover, building a futurological model may not be an infallible method to ensure the future of the human civilization, but it might end up being a surprisingly strong solution. If the introduction of a futurological model leads to undesirable side effects on the human society, the same model could help us to predict and prevent these effects. If a futurological model reveals that the concentration of economic or political power among a few individuals renders the human society chaotic and unpredictable, or that the decisions of these individuals have a disproportionate impact on the future of humankind, the same model could help us to build a better economic or political system, where the power would be more reasonably distributed. In his book *Elements of Philosophy of Right*, Georg Wilhelm Friedrich Hegel states: “The owl of Minerva spreads its wings only with the falling of dusk.” The common interpretation of this statement is that we can only understand a series of events, for example a historical period, once we have reached its end, but futurology might challenge this assumption. With a working futurological model, we might be able to gain valuable insights into a historical period from its very beginning, or even before it begins. Using this knowledge, we might be able to anticipate the positive, neutral, or negative outcomes of our decisions, to avoid the paths that could lead to existential threats, and to build gradually a better and more peaceful society. While this *Pax Futurologica* may seem distant, adding the advance “Futurology” to our technology tree might turn out to be the victory condition in our real-world *Civilization* game. We should imagine an owl spreading its wings at dawn.

8
Futurology

In this book, we explore whether it could be possible to build a futurological model, a predictive model of the world that may be used to forecast the most important trends of the future. In Chapter 1, we hypothesized that a successful futurological model should integrate a wide range of scientific knowledge, and leverage the advances in data science, machine learning, deep learning, and artificial intelligence. In Chapter 2, we suggested that if we could develop a futurological model where levels of organization are flexible and could be adapted, we might be able to use artificial intelligence to find the most relevant levels, where relevance would be measured by the predictability obtained. In Chapter 3, we observed that complex adaptive systems might be seen as possible modeling shortcuts, where we assume that several levels of organization work together inside an integrated entity, and introduced the concept of the computational continuum, as a way to describe the trade-off between simple and complex models. In Chapter 4, we hypothesized that the probability of unique evolutionary events might be evaluated using a brute force strategy, in which case the timeline predicted by a futurological model would consist of successive ramifications. In Chapter 5, we suggested that entities or processes can be described with a heuristic model, in the sense that their behavior can be computed as a global approximation, or with an algorithmic model, in the sense that their behavior can be computed as the exact result of the interactions of their lower-level components, considered as axioms. In Chapter 6, we discussed how our knowledge of the human brain might be integrated into a futurological model, and suggested that building a futurological model might be the natural horizon of scientific research. Finally, in Chapter 7, we explored how a futurological model might be extended from biological evolution to human history, discussed the special cases of the scientific enterprise and artificial intelligence, and wondered how the human civilization would evolve if a working futurological model was built. Outside the scientific realm, the most famous historical institution that comes to mind when we think about futurology might be the Oracle of Delphi, and this Oracle was allegedly established after Apollo mastered Python, the mythological snake. Interestingly, if our objective is to build a futurological model, the scientists involved in this project might need to master another Python: the Python programming language, particularly popular for artificial intelligence. Throughout this book, we explored some ideas that might serve as a basis for an artificial intelligence system integrating the scientific knowledge of humankind, and in this eighth and last chapter, we will propose the first steps of a plan that might lead to a working futurological model. While this plan is only intended as a starting point, it might give us some insights into the challenges that we would be facing if we decided to develop a true, real-life science of the future.

The first step of the plan would be to develop a unified representation of the scientific knowledge of humankind, beyond the object-oriented programming classes that we used as placeholders throughout this book. Every entity of the world has a certain number of important properties, and this number may depend on the type of entity. If we consider that each of these important properties can be encoded as a value inside a vector, a natural solution to represent our scientific knowledge could be to model all entities of the same type with vectors of the same dimension. However, we will take a different path, and consider that every possible entity can be represented as a particular instance of a vector of arbitrarily high dimension, which we could call the futurological vector. In other words, a hydrogen atom, a human neuron, a population of ants, and every other conceivable entity should be represented as a particular instance of the futurological vector, or expressed more simply, as a particular futurological vector. Therefore, the positions of the futurological vector should encode every known or inferred property of every past, present, or future entity of the world that we would like to model. For example, if our objective is to model elementary particles, some positions might be used to encode location, speed, mass, electrical charge, and spin. Then, if our next objective is to model atoms, some of the latter positions could still be useful, but other positions of the futurological vector might be used to encode the number of protons, neutrons, electrons, and covalent bonds. Overall, the structure of the futurological vector should reveal a cartography of the scientific knowledge of humankind, with a first set of positions used to encode physical properties, a second set to encode chemical properties, a third set to encode biological properties, and so on for all the levels of organization. Of course, for each particular futurological vector, only a few positions would contain relevant information, while the other positions could be set to zero, or to a random number. At every time step, all the particular futurological vectors, assumed to be in column form, should be assembled into a two-dimensional array, which we could call the futurological matrix, representing the current state of the world that we are trying to model. For example, if we are trying to model the communication inside a beehive, the futurological matrix might consist of all the particular futurological vectors representing the individual bees. Now, in order to transition from one time step to the next, a matrix function of arbitrary complexity, which we could call the futurological function, should be executed on the futurological matrix. The futurological function should include every law, equation, or algorithm from every scientific domain, and its result should be the futurological matrix of the next time step. In other words, the chemical reactions of photosynthesis, the mechanisms of cell differentiation during the development of the heart, the mating strategies of owls, and all other conceivable processes should be included in the futurological function. In its simplest form, we could imagine a series of conditional statements, which would be applied when relevant, depending on the content of the futurological matrix. For example, if our objective is to model chemical reactions, some conditional statements inside the futurological function might detect if the futurological matrix contains representations of atoms or molecules. If this is the case, the appropriate rules should be applied, in order to compute the probability of a chemical reaction, and its products. Between time steps, the futurological function should be able to change the values of the futurological matrix, but also the number of columns. In other words, it should be able to create and remove particular futurological vectors. Overall, the structure of the futurological function should reflect the structure of the futurological vector, with a first set of rules applying to physical phenomena, a second set to chemical phenomena, a third set to biological phenomena, and so on for all the levels of organization. Therefore, while the futurological vector would be a comprehensive representation of all possible entities, the futurological function would be a comprehensive representation of all possible processes.

The second step of the plan would be to optimize this particular representation of the scientific knowledge of humankind for the futurological task of interest. We already have at our disposal powerful artificial intelligence tools, in particular deep learning tools such as large language models, which could be used to process the literature from every scientific domain. Nevertheless, a considerable amount of work, or very clever algorithms, might be needed in order to improve both the futurological vector and the futurological function. In particular, it might be necessary to merge redundant entities, processes, properties, and rules. Whenever possible, entities and processes should be associated with specific levels of organization, and it should also be indicated whether these entities and processes form part of a complex adaptive system. If several competing scientific models exist for a single entity or process, relying on different properties or rules, the alternative versions should be appropriately marked. The existence of competing scientific models might render the futurological vector and the futurological function more complex, but this additional layer of complexity should be considered as an opportunity for testing alternative scientific explanations. Furthermore, since some scientific models should be simpler than others, a limit could be defined, for example, on the total complexity of the scientific models actually used, which would be a form of regularization. The third step of the plan, closely linked to the second one, would be to train an artificial intelligence system to select, at each time step, the right representations for the futurological task of interest, using the predictability obtained as the objective function. Indeed, one of the core requirements of a futurological model might be to determine the frontier between order and disorder, to choose which complex adaptive systems are stable enough, which levels of organization can serve as a basis for prediction, and to make these choices automatically, in a diversity of scenarios. Let us consider the example of the human brain. While in some scenarios, a high-level model of the whole human brain might provide relatively good predictions, in other scenarios, we might need to fall back on lower-level models, and consider brain regions, brain networks, cortical columns, or even neurons, glia cells, axons, dendrites, and neurotransmitters. If we assume that the whole human brain can be represented as a particular futurological vector, falling back on lower-level models means that the futurological function should be able to remove this representation, and replace it with a series of particular futurological vectors corresponding to lower-level components. When necessary, the futurological function should also be able to perform the inverse operation, and replace a series of particular futurological vectors with a single one, corresponding to a higher-level entity. More generally, the futurological function should be capable of selecting, for each element of the world, the right representation in the computational continuum, in order to make the best possible predictions, while using the least possible amount of computational resources. Importantly, this capacity might be trained using machine learning algorithms. We might be able to evaluate, over our entire scientific knowledge, in which scenarios the futurological function should change the number of columns of the futurological matrix, using the predictability obtained as our objective function. If some representations are considered unreliable, in the sense that they lead to poor predictions or require too many computational resources, the artificial intelligence system should try to replace them with lower-level or higher-level representations, and evaluate whether there is an improvement. Of course, replacing a heuristic model with an algorithmic model, or the other way around, would not be trivial. Theoretically, if a representation of the whole human brain needed to be replaced by eighty-six billion representations of neurons, we would need to estimate the values of eighty-six billion particular futurological vectors. Practically, if an entity or process is repeated in multiple places, or occurs multiple times, a working futurological model would probably perform some sort of spatial or temporal sampling in order to spare computational resources, and a similar sampling might be used to alleviate the cost of modeling multiple lower-level components. Nevertheless, the problem would still be challenging, and could require multiple iterations between the second and the third steps, as well as extensive datasets spanning our entire scientific knowledge.

The fourth step of the plan would be to train an artificial intelligence system to discover the relevant levels of organization and complex adaptive systems by itself. Fundamentally, for this step, we would need to develop an artificial intelligence system capable of performing independent scientific research and discovery. This ambitious goal would require the automatization of several scientific processes, such as the determination of knowledge gaps in the literature, the formulation and testing of hypotheses, and the creation of scientific models. While we are still at a very early stage, the development of an artificial intelligence system capable of scientific research and discovery is starting to emerge as a possible long-term objective for the scientific community. Of course, we might not have enough data to test every possible hypothesis, and the automatization of experimentation and data acquisition could come with its own challenges. Furthermore, the issues of openness, explainability, and interpretability would obviously be particularly important. Nevertheless, if such an artificial intelligence system becomes available in the future, scientific models could be created and evaluated at an unprecedented rate, and a futurological model might greatly benefit from this automatization. As we discussed in Chapters 2 and 4, would it be better, perhaps, to model independently animals, plants, and fungi, instead of having a single organism model? Would it make sense, perhaps, to add new models for viruses, prions, macromolecules, multinucleate cells, large-scale brain networks, symbiotic relationships, and insect societies? Should we consider genes, epigenetic markers, and ideas as distinct levels of organization? An artificial intelligence system capable of scientific research and discovery might be able to answer these questions, and many others, by formulating and testing multiple hypotheses. This would translate into further adaptations of the futurological vector and the futurological function, and hopefully into better predictions. In return, a futurological model might be the perfect scientific platform for running this automatized process of scientific research and discovery over a wide range of scientific knowledge, and could open entirely new horizons for the scientific community. Now, if we were able to complete this step, would we have at least a proof of concept for a working futurological model? Well, we might. We would have built an artificial intelligence system integrating the scientific knowledge of humankind, a predictive model of the world, which could be extensively trained on all the data ever acquired by scientific research, including our records of biological evolution and human history. The next thing to do would be to run this model into the future with different sets of parameters, for as many steps as we want, as many times as necessary, and to analyze the maëlström of predicted timelines. Eventually, this artificial intelligence system, trained to discover levels of organization and complex adaptive systems in the past, might be able to predict the emergence of new levels of organization and complex adaptive systems in the future. Of course, many things would remain very difficult, almost impossible to predict. The objective of futurology should not be to anticipate everything, but to make the best possible use of our scientific knowledge, in order to gain a better understanding of how the future works. Furthermore, while levels of organization, and to a lesser extent complex adaptive systems, are widely used in science, and while they provide an intuitive way to apprehend regularity and adaptation, these frameworks might just be the beginning. The process of building a futurological model may challenge our assumptions about how to structure the continuum of reality, and we might need to invent more abstract building blocks, and a more general ontology. The issues of explainability and interpretability might become even more important, at least if we want to avoid a futurological black box. Once again, the plan discussed in this chapter is only intended as a starting point, and this book in general should only be considered as a preliminary exploration. If our objective is to develop a true, real-life science of the future, we should definitely think further, and explore several alternative ways to build a futurological model.

Since this book is about predictability, it may come as a surprise that we said nothing about randomness, and barely mentioned chaos. We did not base our discussion on these concepts, because chaos and apparent randomness are just the reason why a futurological model is needed in the first place. Also, for the futurological tasks of interest, we implicitly assumed that the laws of physics were deterministic, and that the time steps could be defined uniformly. If we needed to extend the futurological model in order to include quantum mechanics and relativity phenomena, some design changes would probably be required. Now, we will conclude this preliminary exploration with three additional remarks. First remark: a futurological model might need to anticipate the reactions to its predictions. Moreover, it might need to take these anticipated reactions into account before making its predictions public. Indeed, humans and other agents can change their behavior in response to a prediction, and in some cases, these changes might modify the world to the point that this prediction becomes immediately obsolete. Making a prediction that carries the seeds of its own destruction would be useless. In these particular cases, a futurological model could refuse to make a prediction, or reveal this prediction only to certain agents, but another solution might be to recursively integrate the anticipated reactions into a series of hidden predictions, before making the last prediction public if an equilibrium is found. More precisely, a futurological model could start by making a first hidden prediction, and anticipate what would be the reactions to this prediction if it was made public. Then, the second hidden prediction could include the anticipated reactions to the first hidden prediction, the third hidden prediction could include the anticipated reactions to the second hidden prediction, and so on, until an equilibrium is found and a prediction can be made public, knowing that the reactions to this final prediction should not substantially modify what is predicted. We could call this process: futurological recursion. Since there might be a thin line between anticipating the reactions to a prediction, predicting the future desires of humankind, and predicting the future requests that we could address to a futurological model, this question should most certainly motivate further thinking. Second remark: if a futurological model is built by humans, it will be run by humans, and we better make sure that these humans work for the common good. We already have a few millennia of human history behind us, and it is increasingly annoying to witness the number of good, well-intentioned scientific and technological ideas that end up misunderstood, misused, and exploited for dominating others. Of course, domination can be achieved by easier and more effective means, and a futurological model would probably be much more useful for a democratic society, where the future is relatively open, than for a despotic society, where the future is imposed by force. Nevertheless, since we are discussing the future of the human civilization, the importance of following a good scientific method cannot be overstated. Whenever possible, the algorithms, data, and results of a futurological model should be made public, in order both to advance scientific research and to benefit the whole humankind. Third remark: choosing a branch in the tree of future worlds would require an agreement about what a desirable future is, but the shape of the tree might be important as well. If a branch leads to a relatively deterministic path, i.e., a path with too few ramifications for a long period of time, we might be willing to reconsider the trajectory, and choose instead a branch with more degrees of freedom. If a branch leads to a relatively chaotic path, i.e., a path with too many ramifications in a short period of time, we might be willing to reconsider the trajectory as well, and choose instead a branch with less degrees of freedom. Regarding the shape of the tree, an optimal trajectory might stand somewhere between order and disorder, a pattern that we observed repeatedly in this book. Furthermore, the rate of divergence between several branches might also be an important variable to consider. If some branches diverge slowly, the frontier between the corresponding future worlds should remain thin: the future would be unstable, and the world might easily toggle between several different scenarios. If some branches diverge rapidly, the frontier between the corresponding future worlds should quickly become large: the future would stabilize, and the world might enter an almost irreversible path. Whatever happens, and even if the human species faces extinction, the tree of future worlds will continue growing, and new ramifications will appear. Until the end of time, one of the few things certain in the future is the future itself.

For a long time, I thought that the best way to ensure the future of the human civilization would be the foundation of a World Republic. Today, I still believe that the emergence of a global political entity could be the next step in human history, and that the integration of the whole humankind into a common democracy, with a World Parliament, a World Government, and a World Court, would be an excellent idea. When we look at Earth from space, we can only be impressed by the finitude of our planet, and overwhelmed by the idea that the entire human history takes place on the surface of a fragile blue globe. Viewed from above, the very existence of political frontiers may seem absurd, in a time where the challenges faced by humankind require a global governance. If a World Republic was founded, nuclear weapons, and other weapons of mass destruction, could be kept more easily under control. We could have a better chance of preventing wars, ending extreme poverty, overcoming environmental crises, and solving the other urgent problems of humankind. Furthermore, a World Republic could implement policies granting a universal access to education, health, and culture, and federate enough intelligence and resources to develop ambitious projects, such as the restoration of ecosystems, the development of nuclear fusion, the construction of a space elevator, and the colonization of space. Overall, I believe that the future of the human civilization would be brighter if a World Republic was founded, and if peace, freedom, and progress were protected at the planetary scale. I also believe that the European Union is currently the most interesting laboratory for these federalist ideas, and that the strengthening of the United Nations could serve this objective as well. Nevertheless, I could be wrong, whether about the objective, or about the method. In fact, we could all be wrong. We all have some intuition about what the future could be, or should be, but this intuition is severely limited, because the tree of future worlds is simply too complex to be intuitively apprehended by the human brain. For example, it might turn out that despite our best efforts, a World Republic would be corrupted, or undermined by strong economic inequalities, or destabilized by discriminations based on race, gender, age, or political orientation. When we envision a desirable future, we need to keep in mind that this future is only one possible branch, and not necessarily the best one, or the most probable one. Other scenarios, waiting to be discovered, could turn out to be better, or more probable, or could lead to more interesting outcomes later in the timeline. As the world becomes increasingly complex, focusing on one branch, or on a few branches, is simply not enough: we need to envision the entire tree. In a sense, a futurological model is a very simple idea. We live in a period of history characterized by accelerated change, with an unprecedented level of uncertainty regarding the future. The difficulty to anticipate the positive, neutral, or negative outcomes of our decisions, and to avoid the paths that could lead to existential threats, is arguably one of the most complex challenges faced by humankind. In a world with so much uncertainty, how could we choose the best path for our future? Well, we could develop a science whose object is the future itself, leverage the advances in data science, machine learning, deep learning, and artificial intelligence, and try to build a futurological model, starting probably with the lower levels of physics, chemistry, biology, and neuroscience, before working on the higher levels of psychology, sociology, economics, and politics. This endeavor could be seen as a long-term investment. Even if futuristic cities are eventually built in space, and if humankind becomes an interstellar species, we will still need to anticipate the outcomes of our decisions, and we could still benefit from the insights of a futurological model. After all, while terrestrial life and intelligence might eventually expand beyond Earth, the objective of the human civilization has never been to stand the test of space. From the beginning, the objective of the human civilization has always been to stand the test of time.

For the moment, we are still living on Earth, and one of my favorite places on this planet is certainly the Jardin des Plantes, in Paris. The Jardin des Plantes is the main botanical garden in France, and it is part of one of the largest natural history museums in the world, the Muséum National d'Histoire Naturelle. I like walking on the long alleys covered in white gravel, surrounded by trees, colorful flowers, benches, water sprinklers, minerals, and beautiful buildings. When I walk alongside the Grandes Serres, the large greenhouses filled with exotic plants, and the Ménagerie, one of the oldest zoological gardens in the world, I often think about biological evolution and human history. At a certain moment in the past, inorganic components such as white gravel, water, and minerals have served as the foundation for life, and later for intelligence: life and intelligence have evolved into a diversity of forms, including the trees, the colorful flowers, the plants of the Grandes Serres, the animals of the Ménagerie, and the humans walking on the alleys. At a certain moment in the past, humans have served as the foundation for civilization, and later for the scientific enterprise: civilization and the scientific enterprise have also taken a diversity of forms, including the Muséum National d'Histoire Naturelle, with its gallery of evolution, its gallery of geology and mineralogy, and its many other scientific collections. At a certain moment in the future, we might develop the foundations of an artificial intelligence system integrating the scientific knowledge of humankind, and this predictive model of the world, this futurological model, could change everything. Of course, walking and wondering about the world is a scientific tradition: Isaac Newton imagined himself as a child playing on the beach, and Charles Darwin contemplated an entangled bank populated by many plants and animals. Nevertheless, scientific crossroads such as the Quartier Latin in Paris might be among the best places to walk if we want to reflect on the scientific enterprise itself. From the alleys of the Jardin des Plantes, we can see the Zamansky Tower, better known as the Jussieu Tower, which marks the location of the scientific university formerly named Université Pierre et Marie Curie, today integrated into the larger establishment Sorbonne Université. Before any of these establishments existed, the historical institution Université de Paris, better known as the Sorbonne, chose as its motto *Hic et ubique terrarum*, which could be translated as: here and everywhere on Earth. Eventually, as the scientific enterprise continues to progress, and as our predictive models continue to improve, we might consider replacing *hic et ubique*, here and everywhere, by something like *nunc et semper*, now and always, in the motto of our future universities. Furthermore, the new equilibrium between time and space could be reflected in a new equilibrium between knowledge and power. Until now, the path for understanding the world and the path for governing the world have been very different, but a futurological model could change the rules of the game. The Jardin des Plantes and the Jussieu Tower stand at some distance from the places of power of the French Republic, such as the Assemblée Nationale, the Sénat, the Palais de l'Élysée, and the Hôtel de Matignon. Nevertheless, if an artificial intelligence system integrating the scientific knowledge of humankind was able to forecast the most important trends of the future, including the economic and political trends, this artificial intelligence system might be able, given a certain set of objectives, to generate a better political platform than any politician, think tank, or party. Such a scientific and technological revolution would raise deep questions for our democracies, and force us to reevaluate the very foundations of our political institutions. It might not be the end of history, but it could be the end of political power as we know it, and perhaps an opportunity to build a better economic or political system, where the power would be more reasonably distributed. In this futurological new world, this *Terra Nova*, both scientific institutions and political institutions might need to evolve, to engage in a closer cooperation, and to find new ways of including citizens in the reflections about the future. Now, since we have started this book with a literary reference, we will end with another. In the *His Dark Materials* trilogy, whose story happens to begin in a university, and to end in a botanical garden, fantasy writer Philip Pullman describes a universe consisting of many parallel worlds. The protagonists Lyra and Will travel between these worlds, helped by a truth-telling instrument called the alethiometer, and witness a war between the Kingdom of Heaven, a tyranny spanning multiple worlds and ruled by the Authority, and the rebel army trying to build the Republic of Heaven, an ideal civilization based on free will. We live in a different reality than Lyra and Will, but the challenges we face share some intriguing similarities. We also have many worlds standing in front of us, not parallel worlds at different points in space, but sequential worlds at different points in time. We can choose to walk in the dark, guided by fear and resignation, and react to the uncertainty of the world by following any authority that pretends to know the only legitimate future. Or we can shed some light on our path, be guided by reason and curiosity, and react to the uncertainty of the world by advancing our knowledge, exploring the tree of future worlds, and using our free will to progress toward a desirable outcome. While a futurological model might be the natural horizon of scientific research, it might also be the instrument we need for governing the world, in the etymological sense of steering: steering the human civilization across the branches of the tree of future worlds, like a vessel navigating an ocean of possibilities. In this sense, running a futurological model would be much like holding a helm, or using the alethiometer. Whatever the difficulties, we should try to build a strong foundation for the future of the human civilization, and to make the best possible use of our scientific knowledge, so that our objectives could be both freely determined and realistically achieved. We should try to build the Republic of Heaven.

Maël Donoso

Ph.D. in Cognitive and Computational Neuroscience

linkedin.com/in/maeldonoso
github.com/maeldonoso

In this book, we explore whether it could be possible to build a futurological model, a predictive model of the world that may be used to forecast the most important trends of the future, by leveraging the advances in data science, machine learning, deep learning, and artificial intelligence. This futurological model should integrate the scientific knowledge of humankind, starting probably with the lower levels of physics, chemistry, biology, and neuroscience, before working on the higher levels of psychology, sociology, economics, and politics. We focus on a few core ideas, such as levels of organization, complex adaptive systems, and evolution, highlight the importance of improving our understanding of the human brain, and suggest that building a futurological model might be the natural horizon of scientific research. Eventually, we propose the first steps of a plan that might lead to a working futurological model, and discuss the possible implications for the human civilization.

9 782839 941068 >