Why study complex systems?

Some phenomena are inherently complex (that is, the system comprises a large number of components which are interacting in a non linear fashion) and may exhibit chaotic behavior. Exacts mathematical models will often fail at providing a solution, an explanation, a description, of such phenomena. Examples of such systems include the study of crowds (best addressed with a multi-agent system), social media or biological systems.

When studying complex adaptive systems, one must keep in mind that these systems are most likely to be multi-scale (space scales, time scale): the complexity is expressed at different levels and all level of descriptions (from the finest to the largest) must somehow relate. In this case, the tools used to represent the system must scale as well! (using scale laws expressing the functional relationships between quantities, fractals, normalization, …). It is also necessary to understand that behavior propagates across levels (i.e. compartments are not closed or "airtight" and we may be interested in understanding how this propagation occurs. Basically, fine scales behavior influence large scale behavior (and vice and versa) through nonlinear feedback, dissipation (e.g. Bénard cells), amplification (e.g. butterfly effect, enzymatic activity). At larger scales, we may observe the emergence of patterns (e.g. organ development, residential areas, …)

In a complex systems, at a quite fine scale, the interactions of the component may seem random but when the system is observed at a macro level, the patterns observed originate from those interactions. Termites mounds are examples of such emergence. The components of the systems, the termites, are, per see, not very complex but thanks to their collaboration and interactions, amazing large structures can be built; structures that have interconnecting passages, caverns and even ventilation tunnels allowing fresh air to circulate in the mound. Similar kind of amazing behavior can be found in ants societies: their nests are oriented, they exhibit foraging behavior, they sometimes enslave other ants from other species, and last but not least, some ant societies can actually farm fungi. Dozens of complex systems example could be cited (to name but a few: the human bodies, human economy, climate and ecosystems, telecommunication networks, …)

The take home message regarding the notion of emergence is the following: complex systems are continuous systems, always "moving forwards", evolving and exhibiting regularities or sudden changes. The structure of CAS is not "hard-coded" (otherwise the macro behavior would not be said to be emergent). Complex systems can be determinist or non-determinist (e.g. weather forecast is deterministic. If scientists fail at providing accurate weather prediction, it is because the system is chaotic and a slight error measure at time t will cause a great divergence in the trajectory of the model with regards to the trajectory of the actual system). To wrap up, the notion of emergence refers to a bottom-up process where complex behaviors emerge at multiple scale from the interactions of low-level components.

How can we model complex systems?

First, one must ask themselves a few open questions: what is the question to address: predict a phenomenon or understand how a phenomenon occurs; can the CAS be reduced to a model (i.e. a model will, by definition, loose relevance with regards to the actual system and there is no way to know if emergent behaviors can be reproduce: a component may be missing, a component may not be accurately modeled, or everything that allows behaviors to emerge is present in the model but by "lack of chance" (due to the potential stochastic nature of the system), the simulation of the model does not lead to a particular emergent behavior that the modeler is interested in. Lastly, an important question to address as well is how the model can be validated (and how accurate this validation is).

Models can be **predictive** or **explanatory**. In the first case, the idea is to generate a prediction about the future state of the system (what can happen), preferably before the event occurs (it's too late if a hurricane is predicted right before or at the moment it occurs). In this case, the meaning and significance of variables is not important. It is important in the case of an explanatory model: understanding the role of given variables and how significant they are regarding a given outcome is behind the idea of explanatory modeling.

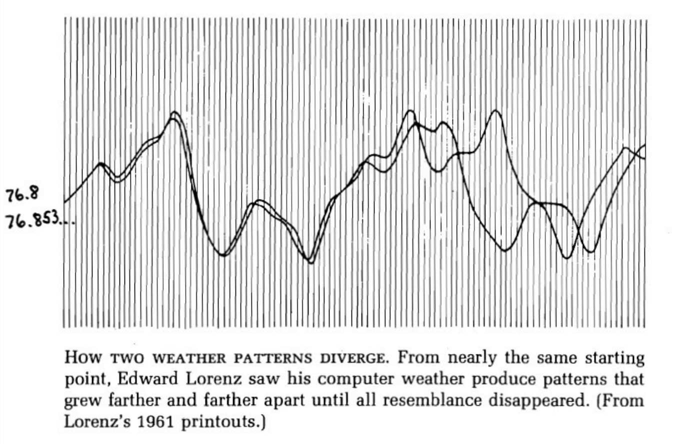

Figure 1: Weather patterns with two slightly different initial conditions

Weather forecasting is an example of a predicting model. However, the chaotic nature of this system and its sensitivity to initial condition makes it very difficult to obtain a model that allows accurate predictions at more than a few days. Illustration 1 below shows how quickly and dramatically trajectories diverge when the initial condition differs very slightly (76.8 v.s. 76.853).

Models are useful to investigate some properties of a system (explanation) or to provide prediction about future states of the systems (prediction) but it is important to keep in mind that a model is (by definition!) and it contains only what the modeler put in it, which means that the model does not necessarily capture all the features of the real system. Modeling intrinsically involves simplification and abstraction: some details must be left out. The modeling process is iterative (fig.2): from empirical data, a model can be built and run. The model is verified against more data and can be adjusted by modifying its parameters, using new data to refine it, or by extending it.

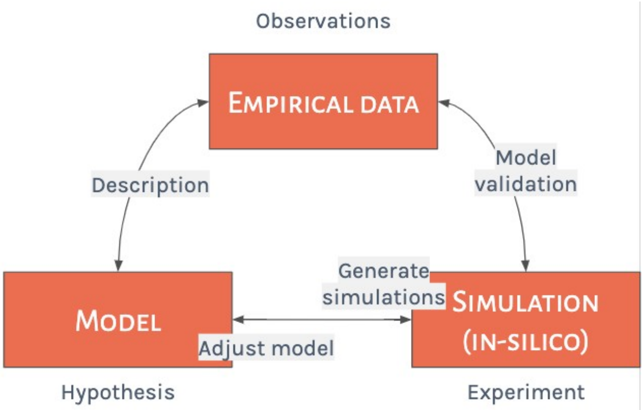

Figure 2: Modeling is an iterative process

Models can be continuous or discrete. In the case of a continuous model, the entities of the model are treated at the population level using differential equations (ODE or PDE). Analytical solutions may be available. If they are not, the model can be numerically integrated. Discrete models involve the modeling of each individual and of the interactions between them. Discrete models include cellular automata or agent based models. In the case of a complex systems, it is of course possible to opt for a hybrid model in order to take multiple scales into account more easily. For instance, a model of cellular tissue could comprise a continuous formalism to model the molecular level and a discrete formalism to model the cellular level. The tissue level need not to be explicit because this is the emergent level of the system.

The **validation** of the model is an important step of the process: an invalid or irrelevant model is a risk for both the modeler or the user of the model. An example of mishap is the use of the Gaussian Copula model that was partially responsible for the 2008 financial crisis[1]. Dedicated to risk prediction, this model was not intended to be used in the context of stock markets because the correlations between market values are too unstable. Yet, it was widely used and contributed to fueling the crisis.

It is also important to limit risks that are related to the sensitivity of a model/system to initial condition. Different scenario can occur: the precision of the model is limited (76.8 and 7.83 are not the same number!), or stochastic processes can be used in the model. In this cases, maybe that some variables are not required and should be removed from the model (decrease its complexity) and seeds can be set so as to be able to reproduce an experiment. It can also be difficult to estimate how sensitive a model / a system is: a measure of uncertainty can be obtained through multiple simulations of the model in order to evaluate to which extent different simulations diverge given some parameter values (other methods involve using a measure of complexity such as the Lyapunov exponent). In general, to ensure a model behaves well enough, it is important to run a large number of simulation for statistical significance purposes.

Risks are also related to the emerging properties of the system: emergence is, by definition, unpredictable. If an “expected” emerging behavior, observed in the system, does not occur in the simulation of the model, it does not necessarily mean that the model is unable to capture it: running multiple simulation may eventually lead to the observation of an “expected” emerging behavior. Upon observing the invert phenomena, that is, the model exhibits an emerging behavior that was not seen in the real system, increasing the number of observation of the system may lead to the observation of the emerging behavior that the model exhibited.

It can be very hard to identify which components of the systems bring about emergent behavior(s) and the intrinsic abstraction in the model may actually set aside essential components that were not identified as essential: in this case, different models can be used to help identifying the essential components.

Finally, risks are also related to the components themselves: the more components and interactions in the model, the more modeling error there is. Re-using previously validated component can help reducing this risk. It should be noted that only one “faulty” component can impair the ability of the model to fulfill its purpose. Validating the components themselves is not enough: even if their individual behavior seem accurate with regard to the systems, it does not guarantee that, when put together, the model will behave as wanted.

Application study: population dynamics

Population dynamics is a field of study in which researcher are interested in understanding how a population evolves over time and how biological or environmental processes influence this evolution (birth rates, death rates, immigration and emigration, among other things). One very

1 https://www.wired.com/2009/02/wp-quant/

simple mathematical model (known as the *logistic map*) for studying population dynamics is the following[2]:

$x_{t+1} = rx_t(1-x_t)$ where x is the population at any given time t and r is the growth rate of the population. This equations states that the population at time t is a function of the growth rate and the population at time t-1. When the growth rate is low, the population dies out and go instinct. When the growth rate is higher, the population might stabilize or fluctuate. For some specific growth rate, the system exhibits c*haotic behavior*.

It means that, for different growth rates values, the evolution of the system will be different. This can be observed on fig. 3. When the growth rate is less than 2, the population dies out (fixed point). When r = 0.5, the population is perfectly stable (fixed point). Above two, different dynamics can be observed and, while $r \leq 3.5$, the population converges towards a periodic pattern (limit cycle). What happens when $r > 3.5$?

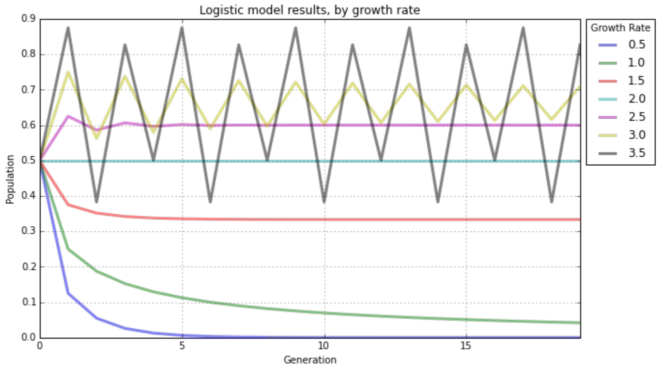

Figure 3: Population evolution dynamics for different growth rates

At the point where r = 3.7, interesting patterns are observable: the population dynamics is no longer stable or periodic: it is **chaotic** (strange attractor). Illustration 4 below shows the chaotic nature of this model: when the growth rate is slightly changed the population dynamics quickly diverge from each other, although the difference in the growth rate is small.

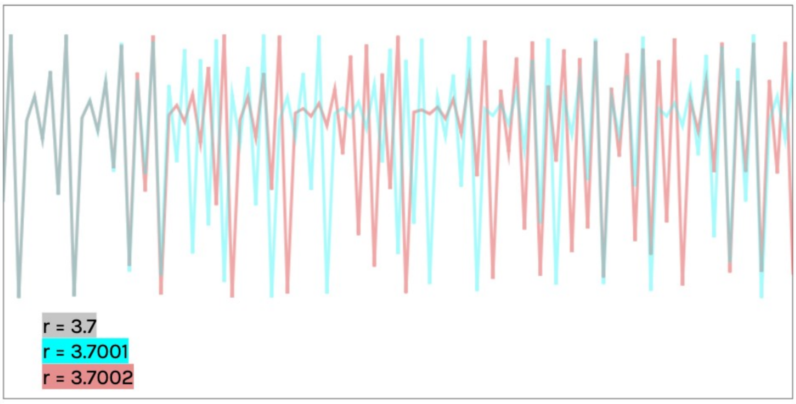

Figure 4: Population dynamics divergence for slight variation of r

An **attractor** of a dynamical system is a value (or several values) towards which the system tends to stabilize over time. There are different types of attractors: fixed points, limit cycles and strange attractors. For instance, the fixed point for r = 0.5 is 0. The fixed point for r = 1 is 0.4. The limit cycle for r = 3.5 comprises 4 values. For r = 3.7, the systems never stabilizes or go back to a value it has already encountered: this is a strange attractor.

2 Visual Analysis of Nonlinear Dynamical Systems: Chaos, Fractals, Self-Similarity and the Limits of Prediction, G. Boeing, Systems 4(4), 37 (2016)

The dynamics of the system can be observed using a *bifurcation diagram*, as depicted in fig. 5.

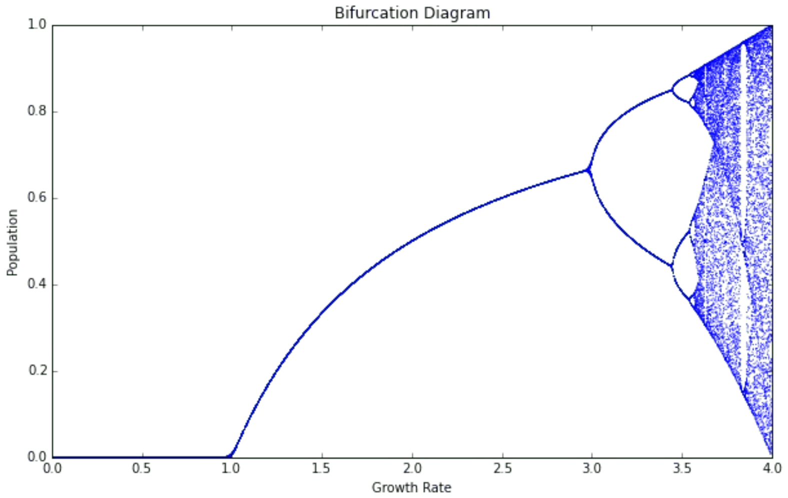

Figure 5: Bifurcation diagram of the logistic map

To plot this diagram, 1000 different growth rates between 0 and 4 were used. This diagram shows 1000 vertical "slices" whose point(s) are the population values that the model stabilized at for a given r (each point of the diagram is the attractor for a given growth rate).

For $r < 1$, the system is never maintained (fixed point = 0). For $1 < r < 3$, the population always stabilizes on one value. For $3.0 < r < 3.5$, the systems oscillates between 2 values. For $r = 3.9$, the population never stabilizes and never reach a same value twice: this is a chaotic regimen.

When r increases, it can be observed that there is a successions of bifurcations between periodic and chaotic behavior. When in a chaotic regiment, the structure of the strange attractor is fractal (i.e. one can zoom infinitely within the diagram an see the very same pattern over an over).

Is chaotic behavior random? Not necessarily! To understand better the chaotic or random nature of a system, it is useful to use a phase diagram, as illustrated in fig. 6.

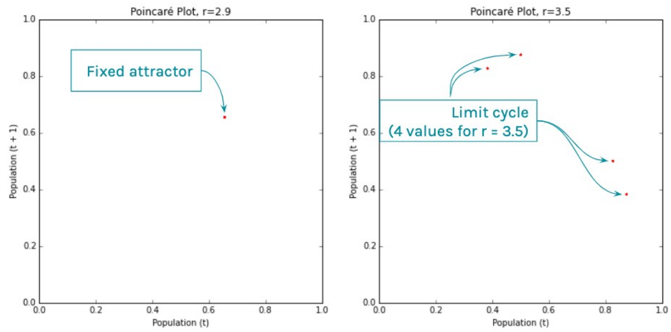

Figure 6: Phase diagram for r=2.9 and r=3.5

On the left diagram, we see that for every t, the system reach population value 0.65: this is a fixed point. On the right diagram, we see that, between t and t+1, the population oscillates between 4 values: this is a limit cycle.

What does the phase diagram look like when the system is in a chaotic regimen?

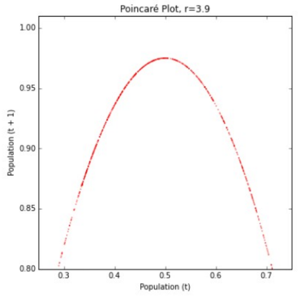

Figure 7: Phase diagram for r=3.9

Fig. 7 shows the values reached by the population for r = 3.9. In this case, the systems reaches different population values, but never the same value twice. This is how chaos and randomness can be distinguished: randomness means that there is a uniform distribution of the values reached by the system. Looking at the phase diagram for r = 3.9, it is clear that the distribution is not uniform: in this case, any values can be reached but not randomly (there is a *structure*). This is the reason why a chaotic system can still be a deterministic system! Phenomena simply never repeat.

Fig. 8 shows the evolution of two time series: one is chaotic, the other is purely random. While it is simply not possible to distinguish the two on the left plot, the phase diagram however shows no ambiguity.

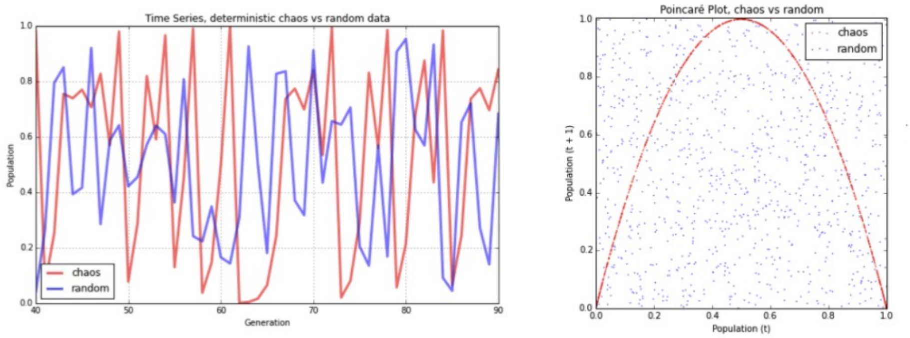

Figure 8: Chaotic and random time series

Basically, when dealing with chaotic systems, knowing how sensitive the system is with regard to its initial condition means that accurate predictions would require infinite precision. In the physical world there is no such things as infinite precisions: measure instrument have a limited precision; computers have limited precision. This is why it is so hard to predict the weather.

Conclusions of chapter 1

The take home message of this chapter is that pretty much every system can be seen as a complex (adaptive) system. Exact mathematics are not necessarily the adequate tool to represent these systems and the theory of complex systems provides an interesting ground for such representations. It is however not a trivial process, and, even formally, it is still a work in progress to characterize the notion of emergence. The rest of this course will show mostly examples of complex systems: dynamical systems and multi-agent systems; (deep) neural networks. But now, you know that we will not have much control over them; that their outcomes will be hard to predict and that they are highly sensitive to the initial parameter values we will use.