Chaos: what is it? is it good or bad? how to detect it? how to process data and make decisions in the presence of chaos?

Vladik Kreinovich, Uyen Hoang Pham, and Olga Kosheleva

Abstract For many real-life phenomena – in particular, for many economic phenomena – it is not possible to exactly predict neither the future sare, not the probabilities of different future states. Such situations are known as chaos. In this chapter, we analyze chaos from the practical viewpoint: what can we predict, how can we make decisions in such situations, and how can we perform computations needed for such predictions and decisions.

1 What is chaos?

Science before chaos. One of the main objectives of science and engineering is to predict the future state of the world – and to make sure that this future state is maximally beneficial for humanity.

At the beginning of science, prediction meant deterministic prediction: based on the current state of the world, we come up with exact predictions. This is how ancient researchers predicted Solar eclipses and future locations of the planets, this is how Newton's physics enables us to predict the trajectory of a spaceship. Usually, the future dynamics is described by differential equations; in this case, predictions are made by solving these equations.

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: vladik@utep.edu

Uyen Hoang Pham
University of Economics and Law – Vietnam National University
Ho Chi Minh City, Vietnam, e-mail: uyenph@uel.edu.vn
Olga Kosheleva
Department of Teacher Education, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: olgak@utep.edu

It later turned out that for some processes, we can make only probabilistic predictions. One of the first such examples is radioactivity. It is not possible to predict when exactly a radioactive atom will go through the corresponding transformation – but we can predict the probability that it will decay by a certain moment of time. Probabilistic predictions are ubiquitous in quantum physics. In quantum physics, we also have differential equations that describe how the state of the world changes with time. The only difference is that in quantum physics, even if we have complete knowledge of the current state of the world, we cannot predict the exact measurement results – we can only predict the probability of different measurement results.

Enter chaos. Until 1970s, researchers believed that we can always predict such probabilities. This all changed with the famous pioneer paper by Edward Lorenz [17] that showed that in weather dynamics, long-terms predictions cannot be made exact – to make exact predictions about future weather, we need to know the initial values with practically impossible accuracy of plus minus 1 atom. A small atom-size change in the initial condition can cause a drastic change in the future weather. Lorenz called this phenomenon *chaos*.

Since 1972, the chaos phenomenon was observed in many areas, including economics. There are many studies of chaos phenomena; see, e.g., [8, 32].

What we do in this paper. In this paper, we focus on practical aspects of chaos: namely, we analyze what can we predict in such situations, how can we make decisions under chaos, and how can we best perform the corresponding computations.

2 Is chaos good or bad?

Why it is important to know this? From the practical viewpoint, why do we need to study chaos?

To answer this question, let us recall what is a practical need for studying different phenomena in the first place:

- Some phenomena we study because of their potential benefits. This is, e.g., why geoscientists look for mineral deposits: to help deliver metals and energy.
- Some phenomena we study because of their potential negative effects. This is, e.g., why geoscientists study earthquakes: to make sure that in each seismic area, all the buildings can withstand the strongest earthquakes possible in this area.

From this viewpoint, it is important to understand whether chaos is good or bad for us. Interestingly, the answer to this question depends on the degree of chaos.

Too much chaos is bad. As we have mentioned, chaos means that instead of a single future value of a physical quantity (or probability), we have a whole range of values. Too much chaos means that this range is very wide – i.e., in effect, that we cannot make any meaningful prediction. If we cannot predict what will happen, then it is difficult to make good decisions.

For example, it has been shown that when the population of a species grows too fast, we cannot really predict consequences of each action, so inevitably, evolution of such a species makes a wrong decision and, as a result, the species becomes extinct; see, e.g., [2, 3, 7, 19, 20, 33]. Similar negative consequences appear when a country's economy starts growing too fast [3].

From this viewpoint, it is desirable to avoid too much chaos – and to study situations when chaos is limited. An example of such situation is transition to democracy: as shown in [14], it leads to a very limited amount of chaos.

A small amount of chaos is often beneficial. In a chaos situation, the whole range of values are possible – if not all of them were possible, we could predict the narrower interval. Thus, the system goes over all possible values from this range in reasonable time. So, if we want to interact with the system, we can select the interaction moment at which the system is in the most beneficial state.

This is a well-known phenomenon. For example, in construction, when a crane lifts some object to an upper floor level, this object does not stand still: it moves around somewhat in a chaotic way. In particular, after a few oscillations, the object gets the closest to the worker who needed this object – so the worker simply grabs the object when it is close. Without such a chaotic motion, delivering objects to upper floors would require much more accurate (and thus, much more expensive and much more time-consuming) crane operations.

Similarly, in economics, prices fluctuate with some degree of chaos. So a person who wants to sell something can naturally wait until the price is the highest within the given range, while the person who need the corresponding object can naturally wait until the price is the lowest within this range. The resulting benefit has the same origin as the benefit of investing money instead of spending them: you get a bonus since you sacrifice the ability to spend the money right away.

An interesting additional benefit of chaos appears in conflict situations. Chaos means that, in particular, that the adversary cannot predict the consequences of your move, which makes counter-measures less effective. For this purpose, additional chaos is added in situations ranging from soccer balls (see, e.g., [9]) to military conflicts (see, e.g., [29]).

The presence of chaos necessitates the use of rule-based systems. Traditionally, dynamics is described by systems of differential equations. Often, such equations enables us to make future predictions. However, as we have mentioned, sometimes, small fluctuations of the initial conditions lead to drastic changes after a moderately small time interval – so we have an example of chaos.

In many practical cases, we do not have a mathematical model describing the system's dynamics, we need to extract this model from the data. In the chaos case, we do not know the exact values of the future quantities, and we do not know the exact values of the future probabilities, so often, we cannot use only the usual technique of differential equations to describe the system's dynamics.

As the above construction example shows, in such cases, we need to supplement (or even replace) differential equations with rules – in the construction example the corresponding rule was "If the object is close, grab it". Such rules have been successfully used in control and other applications as part of the general *fuzzy techniques*; see, e.g., [1, 13, 22, 26, 27, 34].

Comment. The fact that rule-based fuzzy techniques had many practical successes is an indication that the corresponding processes have a significant chaos component. Other examples of such indications are given in the following section.

3 How to detect the presence of chaos?

Sometimes, there is a direct evidence of chaos. By our definition, chaos is the impossibility to predict either the exact values of future quantities or at least the probability of different values. There are many situations when this impossibility is clear by itself; let us mention three of them.

First example: modern AI. The first example – concerning the chaotic character of many real-life processes – comes from the experience Large Language Models (LLMs), the modern Artificial Intelligence (AI); see, e.g., [10]. These models use, in effect, probabilistic approach – in particular, they use it to predict future values and future probabilities. It is well known that in statistics, the more sample one uses for the analysis, the more accurate the predictions.

In general, the prediction accuracy is proportional to $1/\sqrt{N}$, where N is the number of samples; see, e.g., [31]. LLMs use millions and billions of samples, so if the real-life processes were probabilistic, LLMs would predict the probabilities with accuracy of 99.9% or even higher – so the probability of error would be less than 0.1%. In reality, LLMs produce wrong answers in at least 5% of the cases.

This is a clear indication that many real-life processes have a significant chaotic component.

Second example: human experts. A similar example deals with human experts instead of AI; a detailed description of this example is presented in [28]. In 1974, a psychologist Paul Slovic checked how well experts can predict the future – and how do their estimates of their success rate related to their actual success rate. Specifically, Slovic experimented with horse handicappers – experts that earn their living by predicting who will win in horse races. Slovic tested the experts on several situations in which the experts were provided with different amount of information.

As we have mentioned, in general in statistics, the more information we have, the more accurately we can predict. People understand this. In particular, in Slovic's experiment, the experts' degree of confidence in their answers steadily increased when the amount of available information increased. However, interestingly, the actual percentage of correct answers stopped increasing – and stayed at the same 17% level – once the experts were given 5 or more pieces of information.

If the situation was purely probabilistic, the actual success rate would have increased – but it did not. This is an indication that the corresponding process is not purely probabilistic, that it has a significant chaos component.

Third example: stock market. A similar example is specific for economy and finance – it is about the stock market. With numerous experts, the stock prices take into account future possible changes – at least for the major stocks. So, if the stock market was purely probabilistic, its predictions would become more and more accurate as the number of available data increases with time. In reality, there is always a discrepancy between the predicted and actual stock prices, and this discrepancy does not decrease when new tools and new data become available.

This shows that the stock market is not purely probabilistic, that it has a strong chaos component.

Indirect ways to detect chaos. Probabilistic means, in particular, that in a sequence of values corresponding to different moments of time, not only the arithmetic average of the first N elements of the sequence itself tends to the mean value, but also that for any randomly or algorithmically selected subsequence, the mean converges to the same value. This is what distinguishes truly random sequences from sequences like $0101\ldots$, for which the average tends to 0.5, but for subsequences corresponding to odd and to even moments of time, averages are different.

This phenomenon can be used to detect the presence of chaos. Indeed, if a sequence is probabilistic, then the difference between the averages of two non-intersecting N-element subsequences is proportional to $\sqrt{2/N}$ – and for large N, this difference is approximately normally distributed. So, if the average actual difference between such subsequences is larger than $k_0 \cdot \sqrt{2/N}$, then – with the corresponding degree of confidence – we can conclude that there is a chaos component in this process. Here, as usual for normal distributions, $k_0 = 2$ leads to 95% confidence, $k_0 = 3$ leads to 99.9% confidence, and $k_0 = 6$ leads to the confidence level $1 - 10^{-8}$.

4 How to process data and make decisions in the presence of chaos?

What we humans want. We want to use the predicted future values of different quantities $x_1, \ldots, x_n$ to predict the values of related characteristics $y = f(x_1, \ldots, x_n)$. In particular, we want to estimate the future value v of the appropriate objective function $v = J(x_1, \ldots, x_n)$. Computing these values y and v is an important particular case of *data processing*.

Instead of the exact future values, we may know the probability of different combinations $(x_1, \ldots, x_n)$. In this case, we cannot determine the value y, but we would like to know the probability of different values of y. In particular, we want to find the probabilities of different values of v. Computing such probabilities is another case of data processing.

In situations when we can have some way to control the situation, i.e., when the values x_i depend on the parameters $c_1, \ldots, c_k$ of our control action, the resulting value v of the objective function – or the resulting probability distribution – also depends on these control parameters c_i: $v = v(c_1, \ldots, c_k)$.

When we can predict the exact value of the objective function, then we want to select the control values c_i that lead to the largest possible value of the objective function. Selecting these values is an important particular case of *decision making*.

In situations when we only know the probabilities of different values of the utility v, decision theory recommends to select the alternative for which the expected value of the utility is the largest; see, e.g., [5, 6, 15, 18, 24, 25, 30]. This is another case of decision making.

Chaos-related challenge. In the presence of chaos, we cannot predict the future values x_i – and we cannot even compute the probabilities of different combinations $(x_1, \ldots, x_n)$. At best, we can predict the *range* $[\underline{x}_i, \overline{x}_i]$ of each value x_i – or, in the case of probabilities, the range $[\underline{p}, \overline{p}]$ of possible values of each corresponding probability.

In particular, if we describe a probability distribution by a cumulative distribution function (cdf) $F(x) \stackrel{\text{def}}{=} |rmProb(X \le x)$, then in the chaos case, for each x, we can only compute the range $[\underline{F}(x), \overline{F}(x)]$ of possible values of $F(x)$. Such an interval-valued cdf is known as a *probability box*, or, for short, *p-box*; see, e.g., [4].

How can we perform data processing and decision making under such uncertainty?

Data processing in the presence of chaos: interval case. In the presence of chaos, for each of n inputs $x_1, \ldots, x_n$, we only know the interval $[\underline{x}_i, \overline{x}_i]$ of possible values. In this case, the only thing that we can conclude about the quantity $y = f(x_1, \ldots, x_n)$ is that this value y belongs to the range $[\underline{y}, \overline{y}]$ of possible values of $f(x_1, \ldots, x_n)$:

$$[\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) : x_1 \in [\underline{x}_1, \overline{x}_1], \ldots, x_n \in [\underline{x}_n, \overline{x}_n]\}.$$

Computation of such a range is known as *interval computations*, see, e.g., [12, 16, 21, 23]. There are many efficient interval computations algorithms.

In particular, when the input intervals $[\underline{x}_i, \overline{x}_i]$ are narrow – which corresponds to the case when the chaos component is relatively small – we can expand the function $y = f(x_1, \ldots, x_n)$ in Taylor series around the midpoints $\widetilde{x}_i \stackrel{\text{def}}{=} (\underline{x}_i + \overline{x}_i)/2$ and ignore quadratic and higher terms in this expansion. In this case, we get $[\underline{y}, \overline{y}] = [\widetilde{y} - \Delta, \widetilde{y} + \Delta]$, where $\widetilde{y} \stackrel{\text{def}}{=} f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ and

$$\Delta = \sum_{i=1}^{n} \Delta_i \cdot \left| \frac{\partial f}{\partial x_i}_{|x_1 = \tilde{x}_1, \ldots, x_n = \tilde{x}_n} \right|, \text{ where } \Delta_i \stackrel{\text{def}}{=} \frac{\overline{x}_i - \underline{x}_i}{2}.$$

Data processing in the presence of chaos: case of imprecise probabilities. Algorithms are also known for the case of p-box uncertainty; see, e.g., [4] and references therein,

Decision making in the presence of chaos: interval case. Which control $c = (c_1, \ldots, c_k)$ should we select if for each c, we only know the interval $[\underline{v}(c), \overline{v}(c)]$

of possible values of the objective function? In such cases, decision theory recommends to select the control c for which the combination $\alpha \cdot \overline{v}(c) + (1-\alpha) \cdot \underline{v}(c)$ is the largest possible; see, e.g., [11, 15, 18]. The value α should be selected by the decision maker depending on his/her level of optimism/pessimism. Specifically:

- when $\alpha = 1$, the decision maker only takes into account the best-case (= perfectly optimistic) situation;
- when $\alpha = 0$, the decision maker only takes into account the worst-case (= perfectly pessimistic) situation;
- when α is between 0 and 1, the decision maker takes into account both the best-case and the worst-case scenarios.

The use of the above α-combination was first justified by the economist Leo Hurwicz – who later received a Nobel Prize for this research; in his honor, this techniques is known as *Hurwicz optimism-pessimism criterion.*

Decision making in the presence of chaos: case of imprecise probabilities. When the uncertainty is described by imprecise probabilities (e.g., by a p-box), this means that several different probability distributions are consistent with the available information. For each possible probability distribution, we can compute the expected value of the utility. These expected values form an interval $[\underline{v}(c), \overline{v}(c)]$. We can then use Hurwicz's criterion to select the control c.

Acknowledgments

This work was supported by the AT&T Fellowship in Information Technology, by the Institute for Risk and Reliability, Leibniz Universitaet Hannover, Germany, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Focus Program SPP 100+ 2388, Grant Nr. 501624329, by the European Union under the project ROBOPROX (No. CZ.02.01.01/00/22 008/0004590), by the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand, by the Ho Chi Minh City University of Banking, Vietnam, and by Thang Long University, Hanoi, Vietnam.

References

1. R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.
2. A. A. Berryman and J. A. Millstein, "Are biological systems chaotic – and if not, why not?", —it Trends in Ecology and Evolution, 1989, Vol. 4, pp. 26–28.
3. J. Damuth and L. R. Ginzburg, *Nonadaptive selection: An Evolutionary Source of Ecological Laws*, The University of Chicago Press, Chicago, 2025.
4. S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz, *Constructing Probability Boxes and Dempster-Shafer Structures*, Sandia National Laboratories, Report SAND2002-4015, January 2003.

5. P. C. Fishburn, *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York, 1969.
6. P. C. Fishburn, *Nonlinear Preference and Utility Theory*, The John Hopkins Press, Baltimore, Maryland, 1988.
7. L. R. Ginzburg, O. Burger, and J. Damuth, "The May threshold and life-history allometry", *Biological Letters*, 2010, Vol. 6, pp. 850–853.
8. J. Gleick, *Chaos: Making a New Science*, Penguin Books, New York and London, 2008.
9. J. E. Goff, "Balls in the air", *American Scientist*, 2026, Vol. 114, pp. 146–151.
10. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge, Massachusetts, 2016.
11. L. Hurwicz, *Optimality Criteria for Decision Making Under Ignorance*, Cowles Commission Discussion Paper, Statistics, No. 370, 1951.
12. L. Jaulin, M. Kiefer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control, and Robotics*, Springer, London, 2012.
13. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
14. O. Kosheleva and V. Kreinovich, "Does transition to democracy lead to chaos: a theorem", *Aplied Mathematical Sciences*, 2020, Vol. 14, No. 16, pp. 765–769.
15. V. Kreinovich, "Decision making under interval uncertainty (and beyond)", In: P. Guo and W. Pedrycz (eds.), *Human-Centric Decision-Making Models for Social Sciences*, Springer Verlag, 2014, pp. 163–193.
16. B. J. Kubica, *Interval Methods for Solving Nonlinear Constraint Satisfaction, Optimization, and Similar Problems: from Inequalities Systems to Game Solutions*, Springer, Cham, Switzerland, 2019.
17. E. Lorenz, *Predictability: Does the Flap of a Butterfly's Wings in Brazil set off a Tornado in Texas?*, Presented before the American Association for the Advancement of Science, December 29, 1972, see https://ghostarchive.org/archive/OBsPn
18. R. D. Luce and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.
19. R. M. May, "Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos", *Science*, 1974, Vol. 186, pp. 645–647.
20. R. M. May and G. F. Oster, "Bifurcation and dynamic complexity in simple ecological models". *American Naturalist*, 1976, Vol. 110, pp. 573–599.
21. G. Mayer, *Interval Analysis and Automatic Result Verification*, de Gruyter, Berlin, 2017.
22. J. M. Mendel, *Explainable Uncertain Rule-Based Fuzzy Systems*, Springer, Cham, Switzerland, 2024.
23. R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.
24. H. T. Nguyen, O. Kosheleva, and V. Kreinovich, "Decision making beyond Arrow's 'impossibility theorem', with the analysis of effects of collusion and mutual attraction", *International Journal of Intelligent Systems*, 2009, Vol. 24, No. 1, pp. 27–47.
25. H. T. Nguyen, V. Kreinovich, B. Wu, and G. Xiang, *Computing Statistics under Interval and Fuzzy Uncertainty*, Springer Verlag, Berlin, Heidelberg, 2012.
26. H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.
27. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
28. S. Parrish, *Clear Thinking: the Art and Science of Making Better Decisions*, Pneguin Book, New York, 2023.
29. V. Putrenko and N. Pasyynska, "Military situation awareness: Ukrainian experience", *Applied Cybersecurity & Internet Governance*, 2024, Vol. 3, No. 1, pp. 122–146. DOI: https://doi.org/10.60097/ACIG/190341
30. H. Raiffa, *Decision Analysis*, McGraw-Hill, Columbus, Ohio, 1997.
31. D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

32. S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, Westview Press, Boulder, Colorado, 2014.
33. W. R. Thomas, M. J. Pomerantz, and M. E. Gilpin, "Chaos, asymmetric growth, and group selection for dynamical stability", *Ecology*, 1980, Vol. 61, pp. 1312–1320.
34. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.