COMPUTATION BY NATURAL SYSTEMS DEFINED

Ruedi Stoop and Norbert Stoop

Institute of Neuroinformatics, ETH and University of Zurich
Winterthurerstr. 190, Zürich 8057

ABSTRACT

An explicit measure of the computation performed by general systems (electronic circuits, neurons, mechanical devices, etc.) is defined. We propose that the deeper nature of computation, and thus of any measure of computation, is in its reduction of complexity. The latter we understand as the "obstruction against prediction", experienced by an observer. We demonstrate the applicability and usefulness of this concept in different examples, which include some of the most studied families of dynamical systems. The measure can also be computed for higher-dimensional and experimental systems.

1. INTRODUCTION

Real world (="natural") systems have recently received increased scientific attention, because of their obvious efficiency properties. A simple bee, e.g., shows that biological systems can perform perception tasks in a way yet unchallenged by artificial (digital) systems. Often, in this context, it is argued that these natural systems "perform computations". However, this notion of computation generally remains undefined. It may seem that in artificial digital systems, the notion of computation has an intrinsic meaning. Only recently, the fundamentals of computation have been investigated in more details. One line of approach was to define computation in terms of the difficulty of solving classes of fundamental problems, under (mostly: time) constraints. Another line of approach was the extension from intrinsically rational to real weights, as is advocated by biologically motivated neural networks. Depending on the starting point, widely diverging concepts emerged [1, 2, 3]. All that these concepts, however, offer no measure of computation.

The approach that we pursue in this contribution is radically different. It is our understanding that before making extended use of a notion, it first should be defined. In physics, this is generally done by conveying a means for measuring the property (this measure can be as general as verification, which is restricted to the value set $\{0, 1\}$). This is why we are going to define a measure of computation for natural systems. The measure will be applicable to most classes of dynamical systems.

Such a measure of computation cannot be tied to a particular problem to be solved. Rather, it should be defined as an average over all possible problems possible to solve by the device. As the most natural and general enough starting point of computation, we first focus on the initial condition (in the context of artificial digital computation often called operands). Among all possible inputs, low- as well as high-complexity input strings have to be expected, with no particular distribution. The only possible way of implementing this requirement is by taking averages over the input space. Next, a sufficiently general notion of a computing system should be defined, capable of processing the initial condition. In accordance with the intuitive understanding of "computation", we let a general map represent the computational system. This can be motivated by insights from neural network theory, where the simplest tractable representation of a computational unit is a map (general neural networks are universal approximators of functions). More general space-time systems may pose conceptual difficulties. However, many specific aspects will nevertheless be describable in terms of maps. The initial condition to this map (e.g., a single input for one-valued operators, or a string composed from two numbers in the case of a two-valued operator) is then iterated by the map; this is what we take as the representation of the computation. The simplest way is to look upon input as binary digit expansions that are fed into the map. Without restriction of generality, the map hence can be assumed to operate on some subset of the unit interval. The computation halts, when the transient behavior of the map has been passed. (This criterium demands, in principle, a measure of accuracy by the observer. Since in our approach we shall operate on sets of unspecified precision, this assumption will automatically be incorporated.) Because of the general formulation, various results of computation, in particular, periodic, quasiperiodic, as well as chaotic behavior are consistent with the approach. In order to obtain an overall characterization of the computation performed, we will have to average (possibly in a generalized sense) over all initial conditions.

Real systems can be defined as systems that are noisy, with unspecified values on small digits. To define computation in a way applicable to these systems, it makes no sense

to define computation by means of single, isolated, initial conditions. Instead, it implies to operate on the level of sets of neighborhoods, where the linear approximation Df, or f' in the case of dimension one, is the appropriate representation of the map f. This replaces the classical view of computation based on a mapping between points, involving questions of precisions. To arrive at a measure of computation, it appears that the most natural way would be to consider the number of iterations needed, starting from any initial condition, to arrive at any of the possible results. If a non resolution-dependent characterization is to be given, this implies taking the asymptotic limit $n \to \infty$. In this way, as the simplest measure of computation, the quantity $\lim_{n\to\infty} \frac{1}{n} \int \rho(x) \log \mid f^{n'}(x) \mid dx$ emerges, where $\rho(x)$ denotes the natural density. The obtained measure, however, coincides with the Lyapunov exponent of the system, implying that by the notion of computation, no novel quality would be obtained. This is in stark contrast to our intuitive notion, that does not foresee such a coincidence.

In the quest for a more suitable concept, the question emerges what the deeper notion of computation could be. Using a generally accepted perception of computation as the starting point, our answer is, that it should be understood as the reduction of the difficulty of prediction in the statistical sense. This quantity should be averaged over the input space, working on sets of neighboring trajectories because of the finite accuracy required (in this sense, our notion of computation indeed will be vaguely related with the concept of Lyapunov exponents). Computation thus should reduce complexity, when the latter is understood as the obstruction against predictability. In this way, the following rough picture emerges: Consider an arbitrary input, represented as a binary expansion from the unit interval. To this input, already a complexity can be attached, as it may have emerged as the result of some iterated mapping. We shall intrinsically assume that the input complexity can be of any possible complexity (in the Kolmogorov sense [4, 5] or in the sense of a complexity of prediction as defined below). By means of the system that we expect to perform the "computation", this input complexity is transformed into an output complexity (measured, again, by a complexity of prediction). The reduction of the input complexity by means of the map defines our measure of computation.

2. C_S COMPLEXITY

Our outlined concept of a measure of computation thus prerequisites the definition of an appropriate measure of complexity. This measure should express the difficulty of making predictions for real systems based on past observations. As such, it needs to be based on neighborhoods of trajectories in real number space, rather than on individual trajectory in the space of rational numbers. As a consequence, this measure of complexity will be different from the Kolmogorov complexity.

Recently, such a measure of complexity has been worked out [6, 7]. Its main features will be outlined below. Let us consider a dynamical system with discrete time, defined by a map f on some set M in the Euclidean space R^n. Pick an arbitrary point x_0 in the phase space, take some neighborhood $U = U(x_0)$ and consider the orbits $\{f^{(n)}U\}$, $n \in N$. We are interested in observables that relate to measures that are multiplicative along the orbit, i.e., for which the n-step average is evaluated as $(\prod_{k=0}^{n-1} \mu(f^{(k)}(x)))^{1/n}$, where $x \in U$ is the initial state of a particular orbit. Examples of such measures are derivatives, probabilities, etc. Take such a measure $\mu(x)$ and define our observable ν as $\mu(x) =: \exp(\nu(x))$. Our goal is to study the problem of prediction of the next values $\nu(f^{(r)}(x)), r > n$, along the orbits. For the decay of the probability P of retaining a particular measurement value of the observable during a system evolution of n steps, we employ the large deviation ansatz [8] $P(\nu, n)d\nu \sim e^{-ng(\nu)}d\nu$. The thermodynamic formalism implies [8] that

$$g(\nu) = \nu - S(\nu), \tag{1}$$

where $S(\nu)$ is an entropy function. In more detail, the thermodynamic formalism departs from a partition function $Z(n, \beta, \nu)$, where n is the level or depth of the partition and β can be viewed as an inverse temperature. With $Z(n, \beta, \nu)$, a free energy $F(\beta) = \lim_{n\to\infty} \frac{1}{n} \log(Z(n, \beta, \nu))$ is associated, where in $F(\beta)$ we suppressed the dependence on the observable ν. In the absence of phase transitions, an entropy function is obtained by means of the Legendre transform $S(\nu) = \nu\beta - F(\beta)$. Entropy functions $S(\nu)$ have the property of strict convexity, with infinite derivatives at the two end-points of the curve (in the absence of phase transition effects).

A suitable complexity measure is defined as the difficulty of prediction of the observable ν, averaged over all system behaviors. Equation (1) implies that the probability for observing trajectories with a specific value of ν, as a function of n behaves as

$$P(\nu, n)d\nu \sim e^{-n(\nu - S(\nu))}d\nu. \tag{2}$$

As $\nu \geq S(\nu)$, the smaller $\nu - S(\nu)$, the better the prediction based on the past of the orbits will be. Orbits with $\nu = S(\nu)$ will yield perfect long-time prediction. Indeed, this situation characterizes the long-time average of the natural invariant measure (depending on the observable, Lyapunov exponent, information dimension, e.g. [6, 7]). As the complexity is the *difficulty* of making correct predictions, over all length scales, the average

$$\int \frac{S(\nu)}{\nu} d\nu \tag{3}$$

is defined as the measure of complexity.

In order to facilitate the comparison of systems with different topological entropies and to extend the range of applications, the above concept can be refined. ν and $S(\nu)$ can be rescaled as $\tilde{\nu} = \nu/\nu_0$ and $\tilde{S}(\tilde{\nu}) = S(\nu)/\nu_0$, where ν_0 is the topological length scale exponent. Geometrically, this corresponds to a similarity transformation of the entropy function's graph at $(0,0)$, which maps the topological length scale exponent ν_0 to unity. In this case, our complexity measure assumes the form $C_s(\nu) := {\nu_0}^2 \int \frac{\tilde{S}(\tilde{\nu})}{\tilde{\nu}} d\tilde{\nu}$. where on the left-hand side, ν refers to the chosen observable. To obtain a fined-tuned distinction of dynamical systems according to their complexity, we may exponentiate the front factor and the integrand independently. Then the most general form of our measure is obtained as

$$C_s(\gamma,\beta)(\nu) := {\nu_0}^{2\beta} \frac{\nu_1}{\nu_1 - \kappa} \int \left(\frac{\tilde{S}(\tilde{\nu})}{\tilde{\nu}} \right)^{\gamma} d\tilde{\nu}, \quad (4)$$

where γ and β are weightening exponents. To avoid divergence, we require $\gamma > -1$. Most relevant for our purpose, however, will be $C_s(1,0)$, as this characterizes the natural measure of the difficulty of prediction under suppression of lengthscale aspects.

3. COMPUTATION MEASURED

As the desired measure of computation should reflect the ability of the system to reduce the difficulty of prediction, the most natural and straightforward way to define it is to define it as the quantity

$$CO_s = 1/(C_s(1,0) + 1), \quad (5)$$

where $C_s(1,0)$ is the $(1,0)$-complexity measure as defined above and where 1 has been added in the denominator in order to prevent a possible singularity caused by $C_s(1,0) = 0$. This measure of computation is statistical in nature, as it is extracted by means of the thermodynamic formalism. It does not require explicit hierarchical analysis and is non-divergent by construction. If the maximal scaling index ν_{max} is finite, the measure itself will be finite, bounded from below by zero and from above by the value $1 + \log(\nu_{max})$. $CO_s = 1$ indicates a system that performs optimal (decisive) computation, whereas $CO_s \leq \frac{1}{2}$ indicates that the system performs no notable computation. It is worth noting that in this way the measure of computation will be finite, in all realistic cases.

Examples: To get some idea for the implications by this definition, we focus on the the natural partition of the phase space generated by the iteration of a map f and denote the associated observable by $\varepsilon = log\frac{1}{|f'|}$ [6, 7]. According to their computational measure, systems can be divided into classes of increased computation: I ($CO_s < 0.5$), II ($0.5 < CO_s < 1$), III ($CO_s = 1$). The first class is given by systems that almost perform no computation. In this class, we have the intermittent systems, as their complexity when measured by means of $C_s(1,0)$, is highest among the prototypical classes of dynamical systems [6, 7]. This is in agreement with observations that biological complexity is highest at the border between order and chaos [9, 10]. Two of the most prominent 1-d dynamical system classes are the (skewed) tent map family, as the standard representative for hard hyperbolic chaotic systems, and the bungalow tent map as the standard representative for nonhyperbolic chaotic systems. Both systems allow for a simple analytic determination of the computation measure. As a function of the family parameter, they mostly belong to the intermediate, second, class, as is shown in Fig. 1. For the parabola family $y = ax(1-x)$, $a \in [2,4]$, computation can be calculated analytically only for the isolated parameter values. This family belongs to a mixture between the second and the third class, depending on the parameter value. For $a = 4$, as $C_s(1,0)(\varepsilon) = 1/2$, the computation is $2/3$. The third class, with maximal computation $C_s(1,0) = 1$, is characterized by the most simple emergent behavior. As outlined in the introduction, stable periodic orbits are one possible example. This is why in the periodic windows, the parabola falls into this class. However, also classical arithmetic operations are members. In the case of addition where one operand is fixed (e.g., to the value $x = a$), the map $f : y = x + a$ mod 1 can be taken as a representation. Since the complexity of this map is zero, maximal computation $CO_s = 1$ is obtained. The identical result is obtained, if multiplication by a is represented by $f : y = ax$ mod 1.

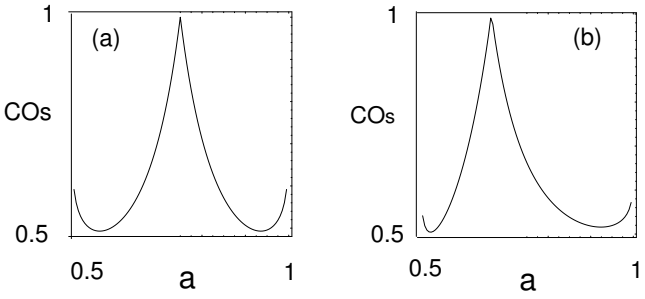

Fig. 1. Computation CO_s. a) Hyperbolic tent map paradigm, parameter a: location of peak. b) Nonhyperbolic bungalow tent map paradigm, parameter a: location of right diagonal point.

4. DISCUSSION

In this paper, we constructed an observable-dependent measure of computation. The range of applications of our method

extends to a very broad class of systems. The evaluation is particularly simple, if a generating or approximate generating partition is available. More generally, our measure can be calculated, whenever an entropy function of scaling exponents can be evaluated. These cases include experimental time series (see, e.g.,[8]).

Our measure provides three main insights. First, the numerical results for 1-d maps point out that nonhyperbolicity *per se* does not have a strong influence on computation. Second, when we compare the fully developed parabola with the symmetric tent map, these systems can be transformed by means of a conjugacy, preserving two points of $S(\varepsilon)$ (the natural measure and the topological measure). For an observer making predictions, they appear as distinct instead. This is captured by our measure, which yields computation $CO_s = \frac{2}{3}$ and $CO_s = 1$, respectively. Third, for intermittent systems, that are at the border between chaos and order, the computation emerges to be smallest. Their apparent job therefore is not to compute, but to provide sufficient complexity, on which computational elements then can do efficient computation.

Our concept of computation also can be interpreted in the context of periodic orbits of dynamical systems. It has recently been shown that unstable orbits can easily be stabilized [11]. Stabilized orbits are among the simplest computational results in our approach. If the stabilization is achieved by means of simple limiter control, it has a simple interpretation in terms of inhibitory neuronal connections. In our framework, this control process is interpreted as a process with maximal computation. We also evaluated our complexity measure for experimental neuron data (unpublished). In an in vivo experiment of cat visual cortex V1, inter-spike intervals ("ISI") between firing events were recorded and analyzed, for two distinct stimulation paradigms [12] (stimulation by noisy patterns moving into the neuron's preferred direction and square-wave stimulation as the optimal stimulus). Using our measure of computation, we obtained the insight that optimal stimulation of the neuron led to larger computation if compared to non-optimal stimulation. Thus, optimal stimulation leads to an improved computation. The vast amount of information arriving at different stages of cortical computation may render the assessment of the computation hardly feasible, if the system is divided into small computational steps. In this case, we propose to replace the whole cascade process by one single map, for which the computation can be evaluated along the lines outlined.

What is the advantage of such a viewpoint, apart from putting computation on the fundament of measureability? Computation in natural systems appears in different stages. Each one taking particular properties into account, ruling out some complexity, but retaining complexity to be combined with results from other computations. With the current approach, the computation performed by individual units in the process as well as the computation performed by the whole can be evaluated and compared. It can be expected that from this comparison, more insight into the relation between computation performed by parts and done by the whole system can be obtained. This knowledge will be required in order to exploit computation by natural, in particular biological, systems for technical purposes.

The authors strongly benefited from discussions with W.-H. Steeb on the nature of computation. R. S. acknowledges the support by a KTI-contract with Phonak AG hearing systems.

5. REFERENCES

[1] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity ans Real Computation*, Springer, 1997.

[2] H. Siegelmann, *Neural Networks and Analog Computation*, Birkhäuser, 1998.

[3] Y. Hardy and W.-H. Steeb, *Classical and Quantum Computing*, Birkhäuser, 2001.

[4] A.N. Kolmogorov, "Three approaches to the quantitive definition of information," *Probl. Inf. Theory*, vol. 1, no. 3, 1965.

[5] M. Li and P. Vitányi, *An Introduction into Kolmogorov Complexity and its Applications*, Springer, 1997.

[6] R. Stoop and N. Stoop, "An integral measure of complexity," in *Control of Oscillations, Proceedings of the Conference Progress in Nonlinear Science, in Honor of 100 Birthday of A.A. Andronov*, V. Shalfeev, Ed., 2002, vol. III, pp. 290–296.

[7] R. Stoop, N. Stoop, and L.A Bunimovich, "Complexity of dynamics as variability of prediction," *J. Stat. Phys.*, p. in press, 2003.

[8] J. Peinke, Parisi J., Roessler O., and Stoop R., *Encounter with Chaos*, Springer, 1992.

[9] C.G. Langton, "Studying artificial life with cellular automata," *Physica D*, vol. 22, pp. 120–140, 1986.

[10] S.A. Kaufmann, *The origin of order*, Oxford University Press, 1993.

[11] R. Stoop and C. Wagner, "Scaling properties of simple limiter control," *Phys. Rev. Letters*, vol. 90, pp. 154101–1–154101–4, 2003.

[12] R. Stoop, D. Blank, A. Kern, J.-J. v.d.Vyver, S. Lecchini, and C. Wagner, "Collective bursting in layer iv," *Cog. Brain Research*, vol. 13, pp. 293, 2002.