Title
Apparent Computational Complexity in Physical Systems

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 15(0)

Authors
Kolen, John F.
Pollack, Jordan B.

Publication Date
1993

Peer reviewed

Apparent Computational Complexity in Physical Systems

John F. Kolen and Jordan B. Pollack
Laboratory for AI Research
Department of Computer and Information Sciences
The Ohio State University
Columbus, OH 43210
kolen-j@cis.ohio-state.edu and pollack@cis.ohio-state.edu

Abstract[1]

Many researchers in AI and Cognitive Science believe that the information processing complexity of a mechanism is reflected in the complexity of a description of its behavior. In this paper, we distinguish two types of complexity and demonstrate that neither one can be an objective property of the underlying physical system. A shift in the method or granularity of observation can cause a system's behavioral description to change in both the number of apparent states and the complexity class. These examples demonstrate how the act of observation itself can suggest frivolous explanations of physical phenomena, up to and including computation.

Introduction

Cognitive Science has generally worked under the assumption that complex behaviors arise from complex computational processes. The generative enterprise in linguistics, for example, maintains that the simplest models of animal behavior - as finite state or stochastic processes - are inadequate for the task of describing language. One needs at least a context-free or context sensitive model to describe or explain language structure.

> There are so many difficulties with the notion of linguistic level based on left-to-right generation, both in terms of complexity of description and lack of explanatory power, that it seems pointless to pursue this approach any further.[2]

Even Newell and Simon's physical symbol system hypothesis (Newell and Simon, 1976) has gone as far as identifying recursive computation as the necessary and sufficient condition for intelligent action. Since the publication of this hypothesis, the general consensus of Cognitive Science has held that the brain is computing something; determining exactly what it *is* computing has emerged as the goal of this new field.

As a corollary to the results presented herein, we believe that questions regarding the contents of the mind's program, or grammar, are fundamentally flawed. Computational complexity, often used to separate cognitive behaviors from other types of animal behavior, will be shown to be dependent upon the observation mechanism as well as the process under examination. While Putnam (1988) has proved that all open physical system can have *post hoc* interpretations as arbitrary abstract finite state machines and Searle (1990) claims that Wordstar must be running on the wall behind him (if only we could pick out the right bits), neither considers the effects of the observer on the complexity *class* of the behavior.

The rest of the paper is organized as follows. We first emphasize the difference between *complexion*, a judgement related to the number of moving parts (or rules, or lines of code) in a system, and the *complexity class*, which may be viewed as the generative capacity of the chosen descriptive framework. Once we recognize that descriptive frameworks apply to measurements of a system's state, we can demonstrate that simple changes of the observation method or measurement granularity can affect either the system's complexion or its class. A mere shift in measurement granularity, in other words, can increase the apparent complexity of a system from a context free language to a context sensitive language. Finally, we discuss the meaning of these results to the Cognitive Science community.

Measurements and Complexity

The foundation of the assumption of the symbolic nature of cognition lies in descriptions of human and animal behavior. For example, a list of moves describes the behavior of the chess player, a transcript records linguistic behavior of conversation, a protocol of introspected states during problem solving describes deliberative means-ends analysis, and a sequence of (x,y) locations over time is a record of eye movement in a study of reading. To construct these descriptions of behavior, one must first collect data from that behavior in the form of *measurements*. We assume that our measurements are discrete since we must be able to write them down.[3] The measurement may be simple, as in the

1. This work was supported by Office of Naval Research grant N00014-92-J-1195.

2. Chomsky (1957, p 24)

case of the cartesian coordinates, or it may be more involved, like the transcript or protocol. To emphasize the creation of discrete symbolic representations of the physical events in the world, we will identify this process as *symbolization*. Transcription of continuous speech, for example, is a symbolization of speech production. It is impossible to avoid symbolization; there is simply too much information inherent in the physical process to pass along without it. Imagine trying to describe the conversation between two people on a street corner. The information generated by such an encounter is infinite due to large number of real dimensions, of movement, sound, time, etc. Researchers avoid these complications by symbolizing the physical action into sequences of measurable events, such as phonemes, words, and sentences.

Information is clearly lost during symbolization. A continuous real value is a bottomless source of binary digits, yet only a small number of distinctions are retained through the transduction of measurement. Of course, no one wants to be shuffling high precision real numbers about when a few bits can do, but it is wrong to believe that the information loss is mere modelling error if, as we show below, it confuses our efforts at understanding the underlying system.

Although judgements of system complexity have no globally accepted methodology, the existing approaches are sharply divided into two groups. The first appeals to the common sense notion which judges the complexity of a system by the number of parts moving around inside it. Thus a system is more complex if it has a larger number of unique states induced by their changeable parts. The term *complexion* has been adopted to refer to systems whose complexity derives from the number of unique moving parts comprising the system (Aida et al, 1984).

The second approach is much more subtle. Imagine a sequence of mechanisms, as specified by a fixed framework, with ever increasing complexion. As the complexion of a device increases, it eventually reaches a limit determined by the framework of changeable parts used in the system. This framework-dependent limit can be modified, however, through the addition or removal of framework constraints. Followers of Chomsky's early work (1957, 1965) in computer science reported on this phenomenon and enshrined the four classes of formal languages, each with a different framework. Regular, context free, context sensitive, and recursive languages

3. This becomes crucial when trying to measure an apparent continuous quantity like temperature, velocity, or mass. Recording continuous signals simply postpones the eventual discretization. Rather than measuring the original event, one measures its analog.

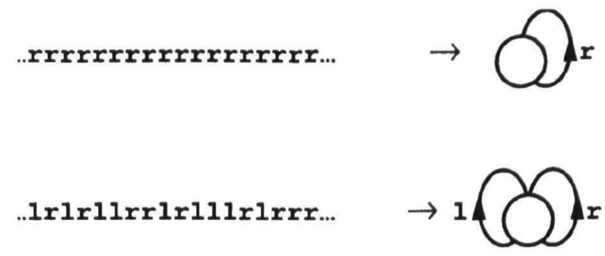

FIGURE 1. Finite state descriptions of equivalent complexity. The first subsequence is from the sequence of all r's. The second subsequence is from a completely random sequence. Both sequences could be generated by a single state generator since each new symbol is independent from all other preceding symbols.

are separated by constraints on the grammars used to specify them, and correlate quite beautifully with automata operating under alternative constraints. Of course, we now know that many other classes are possible by placing different constraints on how the changeable parts interact (see many of the exercises in (Hopcroft and Ullman, 1979)).

These notions of complexity have been traditionally applied only to computational systems. However, recent work by Crutchfield and Young (1991) suggests that one may be able to talk similarly about the complexity class of a physical process. Crutchfield and Young are interested in the problem of finding models for physical systems based solely on measurements of the systems' state. Rather than assuming a stream of noisy numerical measurements, they explore the limitations of taking very crude measurements. The crudest measurement in their eyes is a single decision boundary: either the state of the system is to the left or the right of the boundary. Unlike numerical measurements which can be described mathematically, the binary sequence they collect requires a computational description: i.e. it must have been generated by a particular automaton. Their paper provides two key insights into the problem of recognizing complexity as it arises in nature.

First, the minimality of the induced automaton is important. Crutchfield and Young propose that the minimal finite state generator induced from a sequence of discrete measurements of a system, provides a realistic judgement of the complexity of the physical system under observation. Minimality creates equivalence classes of descriptions based on the amount of *structure* contained in the generated sequence. Consider two systems--the first constantly emits the same symbol, while the second generates a completely random sequence of two different symbols. Both systems can be described by one-state machines capable of generating subse-

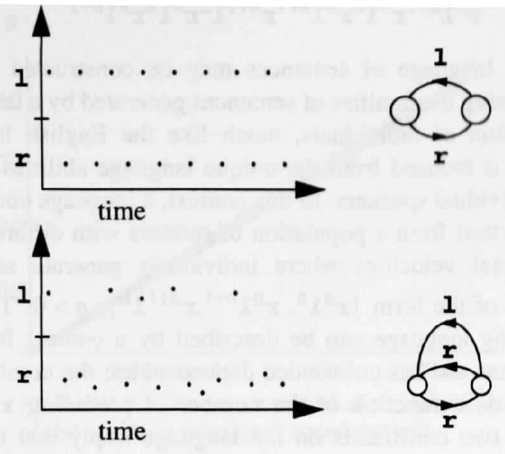

FIGURE 2. The state machines induced from periodic and chaotic systems. Note that the lower system does not produce **11** pairs. This omission is the reason for the increase in number of states.

quences of the observed sequences (Figure 1). In the constant case, the machine has a single transition. The random sequence, on the other hand, has a single state but two stochastic transitions. The ability to describe these sequences with single state generators is equivalent to saying that any subsequence of either sequence will provide no additional information concerning the next symbol in the sequence. Thus total certainty and total ignorance of future events are equivalent in this framework.

Second, they show that physical systems with limit cycles produced streams of bits which appear to be generated by minimal finite state machines with complexion increasing with the period of the cycle. Thus a system with a cycle of period two is held to be as complex as a two-state machine. Similar complexion is exhibited by systems with constrained ergodic behavior, where the number of induced states is determined by the regularities in the visiting of particular bands. These are shown schematically in (Figure 2).

Third, Crutchfield and Young proved that the minimal machines needed to describe the behavior of the simple systems *when tuned to criticality* had an infinite number of states. At criticality, a system displays unbounded dependencies of behavior across space and/or time (Schroeder, 1991). Examples of such self-similar behavior can be found in the spread of forest fires at the percolation threshold of tree density (Bak et al., 1990) and sand pile avalanches (Bak and Chen, 1991). These behaviors are more compactly described as indexed context free languages, a class thought to be consistent with the weak generative capacity of natural languages.

We began to explore the origins of computationally complex behavior in response to the question of generative capacity of certain neural network automata which have finite specifications and yet infinite state spaces (Pollack, 1992). Our original hypothesis began with an assumption that there would be yet another mapping between Chomsky's hierarchy of languages and the state space dynamics of the recurrent neural network. We were encouraged by Crutchfield and Young's paper and a similar conjecture regarding Cellular Automata in the work of Wolfram (1984) and Langton (1990). After many attempts to reconcile our recurrent neural network findings with both the dynamical systems results and the traditional views of intrinsic complexity, we believe that the difficulty of our endeavor lies in the assumption of intrinsic complexity.

Apparent Complexion

While the number of states in the systems studied by Crutchfield and Young can be selected by an external control parameter, the task of merely increasing the number of apparent states of a system is trivial. The key lies in being more sensitive to distinct states. A rock at one level of description never changes it state. By zooming to an atomic level, the rock now enters and exits a myriad of unique states and appears highly complex. Such increases are not interesting since this game can be played with any physical system. Putnam (1988) has proved that an open system has sufficient state generation capacity to support arbitrary finite state interpretations. Searle (1990) has also used this notion to question the relevance of computational models to cognition with his "Wordstar on my wall" example. Thus the apparent complexion, i.e. number of moving parts, fluctuates as the granularity of the observation changes.

Apparent Complexity

Both Putnam and Searle steered clear of the bulwark designed by Chomsky, namely the issue of complexity classes and generative capacity. Is generative capacity also sensitive to manipulation of the observation method? The answer is yes. We will present some simple systems with at least two computational interpretations: a context free language and a context sensitive language.

Consider a point moving in a circular orbit with a fixed rotational velocity, such as the end of a rotating rod spinning around a fixed center, or imagine watching a white dot on a spinning bicycle wheel. We measure the location of the dot in the spirit of Crutchfield and Young, by periodically sampling the location with a single deci-

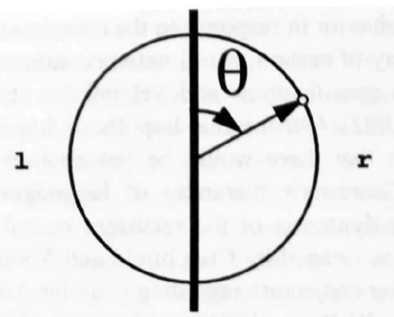

FIGURE 3. Decision regions which induce a context free language. θ is the current angle of rotation. At the time of sampling, if the point is to the left (right) of the dividing line, an **l** (**r**) is generated.

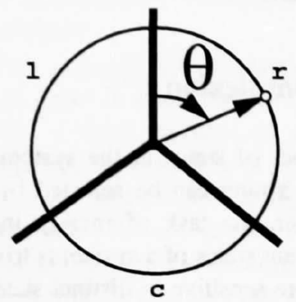

FIGURE 4. Decision regions which induce a context free language.

sion boundary (Figure 2). If the point is to the left of boundary at the time of the sample, we write down an "**l**". Likewise, we write down an "**r**" when the point is on the other side. (The probability of the point landing on the boundary is zero and can arbitrarily be assigned to either category without affecting the results below.) In the limit, we will have recorded an infinite sequence of symbols containing long sequences of **r**'s and **l**'s.

The specific ordering of symbols observed in a long sequence of multiple rotations is dependent upon the initial rotational angle of the system. However, the sequence does possess a number of recurring structural regularities, which we call *sentences*: a run of **r**'s followed by a run of **l**'s. For a fixed rotational velocity (rotations per time unit) and sampling rate, the observed system will generate sentences of the form $\mathbf{r}^n\mathbf{l}^m$ $(n, m > 0)$. (The notation $\mathbf{r}^n$ indicates a sequence of n **r**'s. A more rigorous derivation appears in (Kolen and Pollack, 1993).) For a fixed sampling rate, each rotational velocity specifies up to three sentences whose number of **r**'s and **l**'s differ by at most one. These sentences repeat in an arbitrary manner. Thus a typical subsequence of a rotator which produces sentences $\mathbf{r}^n\mathbf{l}^n$, $\mathbf{r}^n\mathbf{l}^{n+1}$, $\mathbf{r}^{n+1}\mathbf{l}^n$ would look like

$$\mathbf{r}^n\mathbf{l}^{n+1}\mathbf{r}^n\mathbf{l}^n\mathbf{r}^n\mathbf{l}^{n+1}\mathbf{r}^{n+1}\mathbf{l}^n\mathbf{r}^n\mathbf{l}^n\mathbf{r}^n\mathbf{l}^{n+1}.$$

A language of sentences may be constructed by examining the families of sentences generated by a large collection of individuals, much like the English language is induced from the unique language abilities of its individual speakers. In this context, a language could be induced from a population of rotators with different rotational velocities where individuals generate sentences of the form $\{\mathbf{r}^n\mathbf{l}^n, \mathbf{r}^n\mathbf{l}^{n+1}, \mathbf{r}^{n+1}\mathbf{l}^n\}$, $n > 0$. The resulting language can be described by a context free grammar and has unbounded dependencies; the number of **l**'s is a function of the number of preceding **r**'s. These two constraints on the language imply that the induced language is context free.

To show that this complexity class assignment is an artifact of the observational mechanism, consider the mechanism which reports three disjoint regions: **l**, **c**, and **r** (Figure 3). Now the same rotating point will generate sequences of the form

rr...rrcc...ccll...llrr...rrcc...ccll...ll....

For a fixed sampling rate, each rotational velocity specifies up to seven sentences, $\mathbf{r}^n\mathbf{c}^m\mathbf{l}^k$, when n, m, and k can differ no by no more than one. Again, a language of sentences may be constructed containing all sentences in which the number of **r**'s, **c**'s, and **l**'s differs by no more than one. The resulting language is context sensitive since it can be described by a *context sensitive grammar* and cannot be context free as it is the finite union of several context sensitive languages related to $\mathbf{r}^n\mathbf{c}^n\mathbf{l}^n$.

The previous example shows how a population of system behaviors, e.g. the sentential behavior emerging from the family of rotators with different rotational velocities, can be described by computational models from different classes. The two languages observed in the family of rotators can also be observed in the dynamics of a single deterministic system. A slow-moving chaotic dynamical system controlling the rotational velocity parameter in a single system can express the same behavior as a population of rotators with individual rotational velocities. The equations below describes a rotating point with cartesian location (x_0, x_1) and a slowly changing rotational velocity θ controlled by the subsystem defined by x_2, x_3, and x_4.

$$x_0 = \tanh(x_0 - \theta \tanh x_1)$$
$$x_1 = \tanh(x_1 + \theta \tanh x_0)$$
$$x_2 = 4x_2(1 - x_2)$$
$$x_3 = \frac{4}{5}x_1 + \frac{1}{5}x_3$$

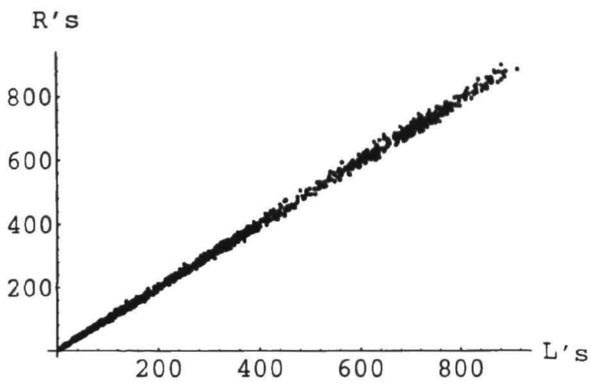

FIGURE 5. The two symbol discrimination of the variable rotational speed system.

$$x_4 = x_4(1 + x_4(x_2 - \frac{1}{2})(\frac{1}{2} - \frac{1}{2}\tanh 5x_3^2))$$
$$\theta = \frac{1}{5}\theta + \frac{4}{5}x_4$$

This system slowly spirals around the (x_0, x_1) origin. The value of x_2 is a chaotic noise generator which is smoothed by the dynamics of x_3 and rotational acceleration of x_4.

As before, we construct two measurement mechanisms and examine the structures in the generated sequence of measurements. The first measurement device outputs an **r** if x_0 is greater than zero, and an **l** otherwise. From this behavior, the graph in Figure 4 plots the number of consecutive **r**'s versus the number of consecutive **l**'s. The diagonal line is indicative of a context free language as a simple corollary to the pumping lemma for regular languages (Hopcroft and Ullman, 1979).

If the underlying language is regular then according to the pumping lemma one would expect to find pumped revisions of $\mathbf{r}^n\mathbf{l}^n$, i.e. there exists some assignment of u, v, and w such that $uvw = \mathbf{r}^n\mathbf{l}^n$ which indicates that the set of strings uv^iw, for $i > 0$, is also in the language. Since the graph plots number of consecutive **r**'s versus the number of consecutive **l**'s, the uv^iw relationship constrains straight lines in the graph to be either vertical, as in the case of v being all **l**'s, or horizontal, as in the case of v being all **r**'s. If v is a string of the form $\mathbf{r}^a\mathbf{l}^b$, then the graph would not contain any straight lines. A formal proof appears in (Kolen and Pollack, 1993).

When the measurement device is changed from two regions to three, we see a parallel change in the Chomsky class of the measurement sequence from context free to context sensitive. Figure 5 shows the relationship between the number of consecutive **r**'s, consecutive **c**'s, and consecutive **l**'s. As in the previous case, one can

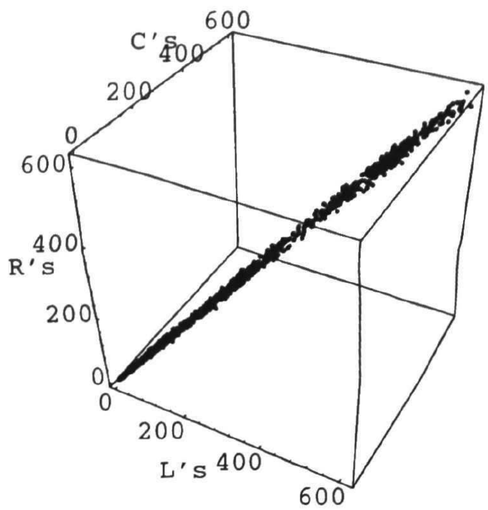

FIGURE 6. The three symbol discrimination of the variable system.

interpret the diagonal line in the graph as the footprint of a context sensitive generator.

Discussion: The Observers' Paradox

The preceding example suggests a paradox: the variable speed rotator, under slightly different measurement granularities, can be seen as both a context free and context sensitive generator. Yet how can this be if computational complexity is an inherent property of a physical system, like mass or temperature? What attribute of the rotator is responsible for the generative capacity of this system? Obviously, there is no explicit stack pushing and popping nonterminal symbols. The rotator does have a particular invariant: equal angles are swept in equal time. By dividing the orbit of the point into two equal halves, we have ensured that the system will spend almost the same amount of time in each decision region, and thus "balancing the parentheses." One may argue that the rotational velocity and the current angle together implement the stack, hence the stack is really there. Such an argument ignores the properties of the stack, namely the ability to arbitrarily push and pop symbols. Likewise, claims regarding an internal Turing machine tape are misguided.

The decision of the observer to break the infinite sequence of symbols into sentences can also affect the complexity class. Similar arguments for sentences of the form $\mathbf{r}^n\mathbf{l}^{n+a}\mathbf{r}^{n+b}$, $(|a|, |b| < C)$ gives rise to a context sensitive language. From this perspective, we can see that Crutchfield and Young accidentally biased the languages they found by assuming closure under substrings, i.e., if string x is in language L then all substrings of x are also in L, which undoubtedly affected the induced minimal automata and criticality languages. Strategic selection of measurement devices can induce an infinite collection of languages from many different complexity classes: the choice of method and granularity of observation "selects" the computational complexity of a physical system.

In other words, **the computational complexity of a physical system cannot be an intrinsic, objective property**; rather, it emerges from the interaction of system state dynamics and measurement as established by an observer.

We believe this result has deep meaning for cognitive science. It suggests that the hierarchy of formal languages and automata is irrelevant to the accounts of complexity in physical systems. Since the brain is a physical system, we cannot know the complexity class of its behavior without establishing an observer. Thus the Physical Symbol System Hypothesis relies on an unmentioned observer to establish that an ant's behavior is not computational while problem-solving by humans is. The necessary and sufficient conditions of universal computation in the Physical Symbol System Hypothesis provide no insight into cognitive behavior; rather, they it implies that humans are capable of writing down behavioral descriptions which require universal computation to simulate. Even the computational intractability of a competence model (e.g. Barton, et al, 1987) is dependent on a particular symbolization of human behavior, not an underlying mechanical capacity, implying that the rejection of mathematical models on the basis of insufficient computational complexity is groundless.

As our ability to establish good measurements has increased, we now know that there are many areas in nature where unbounded dependencies and systematic forms of recursive structuring occur. The coding of the genome, the immunological system, as well in the simple growth of botanical structures, are but a few. These systems are proving as complex as human languages, yet it is only Cognitive science which presumes upon the "specialness" of language and human mental culture to justify a different set of scientific tools and explanations based upon the formal symbol manipulation capacity of the computer. We may have to stop.

Acknowledgments

We wish to thank all those who have made comments and suggestions on the multiple drafts of this paper, including the LAIR connectionists, ConnectFest III participants, and especially Paul Smolensky.

References

Aida, S., et al. (1984) *The Science and Praxis of Complexity*. United Nations University.

Bak, P., and Chen, K. (1991) Self-organized criticality. *Scientific American*. 46-53.

Bak, P., Chen, K., and Creutz, M. (1990) A forest-fire model and some thoughts on turbulence. *Physics Letters*. 147,5-6:297-300.

Barton, G. E., Jr., Berwick, R. C., and Ristad, E. S. (1987) *Computational Complexity and Natural Language*. Cambridge, Mass.: MIT Press.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

Chomsky, N. (1957) *Syntactic Structures*. The Ague: Mounton & Co.

Crutchfield, J., and Young. (1991) Computation at the onset of chaos. In *Entropy, Complexity, and the Physics of Information*. Ed. W. Zurek. Addison-Wesely, Reading.

Hopcroft, J. E., and Ullman, J. D. (1979) *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesely, Reading.

Kolen, J. F., and Pollack, J. B. (1993) The Observers' Paradox. LAIR Technical Report. The Ohio State University.

Langton, (1990) Computation at the edge of chaos: Phase transitions and emergent computation. *Physica*. 42D:12-37.

Newell, A., and Simon, H. A. (1976) Computer Science as Empirical Inquiry: Symbols and Search. *Communications of the ACM*, 19:3.

Pollack, J. B. (1992) The induction of dynamical recognizers. *Machine Learning*. 7:227-252.

Putnam, H. (1988) *Representation and Reality*. Cambridge, Mass.: MIT Press.

Schroeder, M. (1991) *Fractals, Chaos, Power Laws*. New York: W. H. Freeman and Company.

Searle, J. (1990) Is the brain a digital computer? *Proceedings and Addresses of the American Philosophical Association*. 64:21-37.

Wolfram, S. (1984) Universality and complexity in cellular automata. *Physica*. 10D:1-35.