Department of Energy Grant DE-FG03-94ER25231:
Evolving Cellular Automata to Perform Computations

Final Technical Report

Principal Investigators:

James P. Crutchfield and Melanie Mitchell
Santa Fe Institute
1399 Hyde Park Road
Santa Fe, NM 87501

1. Introduction

The overall goals of the project are to determine the usefulness of genetic algorithms (GAs) in designing spatially extended parallel systems to perform computational tasks and to develop theoretical frameworks both for understanding the computation in the systems evolved by the GA and for understanding the evolutionary process by which successful systems are designed.

In our original proposal we scheduled the first year of the project to be devoted to experimental grounding. During the first year we developed the simulation and graphics software necessary for doing experiments and analysis on one dimensional cellular automata (CAs), and we performed extensive experiments and analysis concerning two computational tasks—density classification and synchronization. Details of these experiments and results, and a list of resulting publications, were given in our 1994-1995 report. In addition to those publications, the first year's work made up the bulk of Rajarshi Das's Ph.D. dissertation. Das will officially receive his Ph.D. from Colorado State University this fall.

We scheduled the second year to be devoted to theoretical development. (A third year, to be funded by the National Science Foundation, will be devoted to applications.) Accordingly, most of our effort during the second year was spent on theory, both of GAs and of the CAs that they evolve. A central notion is that of the "computational strategy" of a CA, which we formalize in terms of "domains", "particles", and "particle interactions." This formalization builds on the computational mechanics framework developed by Crutchfield and Hanson for understanding "intrinsic computation" in spatially extended dynamical systems.

We have made significant progress in the following areas:

- *Statistical dynamics of GAs:* We developed a theoretical framework to explain and predict the central phenomenon we observe in our GA simulations—the progression through "evolutionary epochs". This refers to the progression of the population through a series of increasing fitness levels corresponding to increasingly improved computational strategies for solving the task. This progression consists of significant periods of stasis in population fitness which are punctuated by very rapid increases to the next level.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

ph

MASTER

pgs
final

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

**Portions of this document may be illegible
electronic image products. Images are
produced from the best available original
document.**

- *Formalizing particle-based computation in cellular automata:* Our notion of “computational strategy” in the evolved CAs was formalized by a class of models based on domains, particles, and particle interactions. These behavioral components can be cataloged empirically for a given CA and used to model the computation being performed by the CA and to thus estimate the performance of its computational strategy. This work tests our claim that a description in the high-level language of domains, particles, and particle interactions captures a CA’s computational strategy and can predict its performance.
- *Computation in two-dimensional CAs:* We are now extending our analysis techniques to the considerably more difficult case of two dimensions. This research will be essential for the eventual application of our techniques to real-world parallel computation. During the past year we have begun to develop the software necessary for evolving and analyzing two-dimensional CAs. Some preliminary experiments were carried out on both evolving and designing by hand two-dimensional CAs to perform density classification and synchronization, the tasks we have studied in one dimension.

In the next sections we describe in detail the progress we made in each of these areas.

From September 1995 through August 1996 we have published or submitted seven papers related to this project. In addition, members of our group have given 22 invited and contributed presentations on our work. Finally, a doctoral dissertation concerning this project was completed by Rajarshi Das. These accomplishments are listed at the end of the report.

2. Statistical Dynamics of Genetic Algorithms

Our theoretical work on statistical dynamics of GAs was carried out primarily by graduate student Erik van Nimwegen, working with the PIs. This work uses techniques from statistical mechanics, stochastic process theory, and dynamical systems theory to develop a new mathematical approach to understanding the dynamics of genetic algorithms, both for our immediate application of evolving cellular automata to perform computations, and for GAs in general.

We began with a simplified version of the problem: to develop a framework for understanding and predicting the behavior of a simple genetic algorithm using mutation only (no crossover) working on one of the “Royal Road” functions of Forrest, Holland, and Mitchell [2, 8]. These fitness functions bring out very clearly the “punctuated equilibria” phenomena that we observed in our simulations. The first step was to construct a formal operator G that captures the action of the GA on a population fitness-distribution. The eigenvalues of the matrix defined by this operator can be computed, yielding the fixed points of the operator, which translate into evolutionary “epochs”—periods of stasis in the mean fitness of the population punctuated by innovations (jumps in mean fitness). Van Nimwegen has proved that only the last epoch is stable, and has solved the time dynamics of mean fitness the GA for the infinite-population case. This work explains how the metastable evolutionary epochs are caused by the discreteness of the space of population fitness distributions induced by finite populations; they disappear in the infinite-population case. This theory has made predictions that match experimental data very well. This approach and the resulting explanation of GA

behavior is new. To our knowledge, this is the first time that the behavior of a GA working on a non-trivial problem can be predicted so precisely. We recently submitted a short paper on this work to *Physics Letters A* [9]; a longer paper is forthcoming.

Van Nimwegen, working with the PIs, is currently extending this research by developing a new analytic method, using both statistical mechanics and the technique of path integrals, for calculating the lengths of epochs as a function of population size and mutation rate. Once that work is completed, we will be able say that the GA has been completely solved for this simple case, at least in terms of predicting its behavior on fitness distributions for the Royal Road problem. We expect that this work will be completed by the end of the year. These two projects will be the subject of van Nimwegen's doctoral dissertation.

The next step will be to tackle the more difficult and interesting case of a GA that includes both crossover and mutation, as well as more complex fitness functions such as those we use in evolving cellular automata. Over the next year we will extend the initial results in order to predict the evolutionary dynamics that we have observed in our evolving CA experiments. This will require identifying the "constellations" of co-adapted loci in the evolving CA rules that give rise to the observed strategies. Identifying these constellations has been a central part of our analysis of the evolved CAs (e.g., see [7]). Once they are identified, the theoretical framework can apply to predict how the GA moves from one strategy to the next by predicting how the GA moves from one constellation of bits to the next.

We believe these techniques will be a major innovation in applying statistical mechanics and dynamical systems techniques to evolutionary computation methods, and will ultimately be useful for a wide range of problems to which evolutionary computation methods can be applied.

3. Formalizing Particle-Based Computation in Cellular Automata

The mechanisms by which decentralized spatially extended systems such as CAs perform computations is in general not well understood. A major part of our project is to develop a useful framework for describing computations in such systems that will elucidate general mechanisms. To do so, we have built upon the "computational mechanics" framework of Crutchfield and Hanson [1, 3]. This framework identifies "particles" and "particle interactions" in CA space-time behavior as the locus of information transfer and processing. Crutchfield and Hanson's original work did not apply this framework to "useful" computation in spatial systems. In our work we are doing precisely that, by using their framework to explain the computational strategies evolved by the GA.

In Crutchfield and Hanson's framework, the space-time behavior of CAs can be decomposed into two components: (1) regular domains—"computationally homogeneous" regions that can be described in terms of regular languages corresponding to strongly connected finite state automata (e.g., a checkerboard pattern could be described in terms of the regular language $(01)^*$) and (2) particles—localized deviations from or boundaries between regular domains.

Once the regular domains in a space-time configuration are identified, they can be filtered

out, making the particles explicit. Particles can be interpreted as carrying information across space-time, and information processing in a CA is presumed to occur when particles interact. We are investigating whether the “language” of particles and their interactions can serve as a high-level language of computation for spatially extended dynamical systems such as CAs.

In our previous work we have interpreted the evolved CAs in terms of these particles and their interactions. During the previous year, our research with graduate student Wim Hordijk has focused on formalizing this high-level language by constructing models of evolved CAs in terms only of particles and interactions. The behavior of these models can then be compared with that of the original CAs to see how well the particle language can account for the computational performance of the CAs. We have found that the models can account very well, though not perfectly, for the measured computational performance of the CAs. Some discrepancies are due to incompleteness of the model—we are now understanding which aspects of particle behavior are necessary to include. Some discrepancies are due to events during the initial “condensation” period of the CA, during which particles are formed. This period is not part of the model, and we are beginning to understand how and to what extent these initial time periods contribute to the overall computational behavior of the CA.

These results have been written up in preliminary form and have been accepted for presentation at the Fourth Workshop on Physics and Computation [6]. A more complete paper will be submitted by the end of the year to the journal *Physica D*.

Our ultimate goal is to understand both the computational performance of the evolved CAs, as well as that of intermediate evolutionary stages, in terms of the language of particles, and to understand how the GA implicitly manipulates particles and their interactions to improve the performance of the CAs. Such a description will be an important step in understanding how spatially extended parallel systems can compute and how they can be produced by evolutionary processes. This will be the subject of Wim Hordijk’s doctoral dissertation. In addition, Hordijk has been working on developing measures of the “CA landscape” on which the GA searches [5, 4].

4. Evolving Two Dimensional Cellular Automata to Perform Computations

Our previous work on evolving cellular automata has been restricted to one dimension. With graduate student Patrik D’haeseleer, we have been exploring the evolution of computation in two-dimensional CAs. This involves extending the computational mechanics framework to two dimensions. During 1995-1996 we have been looking at the density classification and synchronization tasks in two dimensions. As a first step, we have constructed a series of two-dimensional CAs to perform density classification, using two-dimensional versions of known one-dimensional CAs for performing this task. We have begun to define notions of “particles” and “particle interactions” in the context of two-dimensional CAs, and are attempting to understand their computational performance in these terms. Once we have built up a conceptual framework for understanding computation in these systems we will be ready to use a GA for evolving two-dimensional CAs to perform computations, and to analyze the results.

As part of this preliminary work, we have developed extensions to our genetic algorithm and cellular automaton simulation and graphics software to cover the two-dimensional case.

5. Current Research Activities

Currently, and in the near future, our focus will be on continuing the three projects discussed above. The statistical dynamics of GAs project will be the subject of Erik van Nimwegen's Ph.D. dissertation, and will be finished within the next two years. Formalizing particle-based computation, and using this formalization to describe the evolution of increasingly sophisticated strategies will be the subject of Wim Hordijk's dissertation. Much of the NSF-funded third year of this project will be spent on extending our work to two dimensional CAs, and on applying this work to real-world problems in parallel computation. Rajarshi Das of the IBM T.J. Watson Research Center will spend the next two years at SFI collaborating with us on this project, and we will also collaborate with other members of Das's group at IBM (including Jeff Kephart and James Hanson) who are interested in applying our techniques.

6. Research Group Publications and Presentations

The *Evolving Cellular Automata* group at the Santa Fe Institute during 1995-1996 consisted of the two principal investigators (Melanie Mitchell and Jim Crutchfield), one postdoctoral fellow (Jim Hanson) and four graduate students (Rajarshi Das, Wim Hordijk, Erik van Nimwegen, and Patrik D'haeseleer).

The following is a list of dissertations and papers written and presentations given by our group during 1995-1996

6.1 Dissertations

Das, R. (1996). The Evolution of Emergent Computation in Cellular Automata. Ph.D Thesis. Computer Science Department, Colorado State University, Ft. Collins, CO.

6.2 Papers

Hordijk, W., Crutchfield, J. P., and Mitchell, M. (Submitted). Embedded particle computation in evolved cellular automata. To appear in *Proceedings of the Conference on Physics and Computation—PhysComp96*.

Hordijk, W. (Submitted). The structure of the synchronizing-CA landscape. Submitted to *Proceedings of the Workshop on Control Mechanisms for Complex Systems*, Las Cruces, 1996.

Hordijk, W. (Submitted). Correlation analysis of the synchronizing-CA landscape. Submitted to *Proceedings of the CNLS Conference on Landscape Paradigms in Physics and Biology*, Los Alamos, 1996.

Mitchell, M., Crutchfield, J. P., and Das, R. (1996). Evolving cellular automata to perform computations: A review of recent work. In *Proceedings of the First International Conference on Evolutionary Computation and Its Applications (EvCA'96)*. Moscow, Russia: Russian Academy of Sciences.

Mitchell, M., Crutchfield, J. P., and Das, R. (In press). Evolving cellular automata to perform computations. In Bäck, T., Fogel, D., and Michalewicz, Z. (Eds.), *Handbook of Evolutionary Computation*. Oxford: Oxford University Press.

Mitchell, M. (In press). Computation in cellular automata: A selected review. To appear in Schuster, H. G. and Gramms, T. (editors), *Nonstandard Computation*. Weinheim: VCH Verlagsgesellschaft.

van Nimwegen, E., Crutchfield, J. P., and Mitchell, M. (Submitted). Finite populations induce metastability in evolutionary search. Submitted to *Physics Letters A*.

6.3 Invited Presentations

J. P. Crutchfield:

“Computational Mechanics: Towards a Physics of Complexity”:

Workshop on Theory and Applications of Nonlinear Time Series Analysis, Potsdam. September, 1995.

Symposium on Computational Issues in Learning Dynamical Systems, AAAI Spring Symposium, Stanford University. March 1996.

Santa Fe Institute Summer School on Complex Systems. June, 1996.

Computational Neurobiology Laboratory colloquium, Salk Institute, La Jolla, CA. July, 1996.

“The Evolution of Emergent Computation”: International Conference on Self-Organization of Complex Structures, Berlin. September, 1995.

“How Does Nature Compute?": Joint Colloquium, Keck Center for Integrative Neurobiology and Sloan Center for Theoretical Neurobiology, University of California, San Francisco. December, 1995.

“Forms of Randomness—Embodiments of Computation”: Workshop on Dynamics, Computation, and Cognition, Santa Fe Institute. May, 1996.

“What is a Pattern? Discovering the Hidden Order in Chaos”: Bernard Osher Fellowship public lecture, San Francisco Exploratorium. July, 1996.

W. Hordijk:

“Strategy Performance Estimation for Cellular Automata”: AAAI Spring Symposium on Computational Issues in Learning Models of Dynamical Systems, Stanford University. March, 1996.

M. Mitchell:

“The Evolution of Emergent Computation”: University of Michigan / Santa Fe Institute Seminar. University of Michigan. November, 1995.

"Complex Systems and the Sciences of the 21st Century": Scientia Institute Colloquium on "Approaching the Millennium: Communities, Technologies, Histories", Rice University. December, 1995.

"Emergent Computation and Representation in Dynamical Systems":

Horizon Day: "Representation – The Next Generation". Computer Science Department, Indiana University. January, 1996.

Computer Science Department colloquium, University of New Mexico. January, 1996.

Philosophy, Neuroscience, and Psychology program colloquium, Washington University. February, 1996.

Artificial Intelligence Seminar. University of California, Berkeley. March, 1996.

Institute for Neural Computation Seminar. University of California, San Diego. April, 1996.

Workshop on Dynamics, Computation, and Cognition, Santa Fe Institute. May, 1996.

"Evolving Cellular Automata with Genetic Algorithms": 1996 AAAI Spring Symposium on Computational Issues in Learning Models of Dynamical Systems, Stanford University. March, 1996.

"Adaptive Computation": Conference on Mathematical Geophysics, Santa Fe, NM. June, 1996.

"The Evolution of Emergent Computation: A Review": Keynote lecture, First International Conference on Evolutionary Computation and Its Applications. Russian Academy of Sciences, Moscow. June, 1996.

E. van Nimwegen:

"Statistical Dynamics of Genetic Algorithms". Department of Theoretical Biology, University of Utrecht, the Netherlands, April, 1996.

6.4 Contributed Presentations

W. Hordijk:

"A Measure of Landscapes": Center for Nonlinear Studies 16th Annual Conference, "Landscape Paradigms in Physics and Biology", Los Alamos National Laboratory. May 1996.

6.5 Honors and Awards

Prof. Crutchfield was awarded the Bernard Osher Foundation Fellowship for 1996 to complete work on the San Francisco Exploratorium's NSF-funded exhibition "Turbulent Landscapes: The Natural Forces that Shape Our World", for which he was the principal scientific advisor.

References

- [1] J. P. Crutchfield and J. E. Hanson. Turbulent pattern bases for cellular automata. *Physica D*, 69:279–301, 1993.
- [2] S. Forrest and M. Mitchell. Relative building-block fitness and the Building-Block Hypothesis. In L. D. Whitley, editor, *Foundations of Genetic Algorithms 2*, San Mateo, CA, 1993. Morgan Kaufman.
- [3] J. E. Hanson and J. P. Crutchfield. The attractor-basin portrait of a cellular automaton. *Journal of Statistical Physics*, 66(5/6):1415–1462, 1992.
- [4] W. Hordijk. Correlation analysis of the synchronizing-CA landscape. Submitted to *Proceedings of the CNLS Conference on Landscape Paradigms in Physics and Biology*, Los Alamos, 1996.
- [5] W. Hordijk. The structure of the synchronizing-CA landscape. Submitted to *Proceedings of the Workshop on Control Mechanisms for Complex Systems*, Las Cruces, 1996.
- [6] W. Hordijk, J. P. Crutchfield, and M. Mitchell. Embedded particle computation in evolved cellular automata. In *Proceedings of the Conference on Physics and Computation—PhysComp96*. To appear.
- [7] M. Mitchell, J. P. Crutchfield, and P. T. Hraber. Evolving cellular automata to perform computations: Mechanisms and impediments. *Physica D*, 75:361 – 391, 1994.
- [8] M. Mitchell, J. H. Holland, and S. Forrest. When will a genetic algorithm outperform hill climbing? In J. D. Cowan, G. Tesauro, and J. Alspector (editors), *Advances in Neural Information Processing Systems 6*. San Mateo, CA: Morgan Kaufmann.
- [9] E. van Nimwegen, J. P. Crutchfield, and M. Mitchell. Finite populations induce metastability in evolutionary search. Submitted to *Physics Letters A*.