The EvoGrid: A Framework for Distributed Artificial Chemistry Cameo Simulations Supporting Computational Origins of Life Endeavors

Bruce Damer[1], Peter Newman[1], Richard Gordon[2], Tom Barbalet[3], David W. Deamer[4] and Ryan Norkus[1]

[1]DigitalSpace Corporation
[2]University of Manitoba
[3]Noble Ape and Biota.org, The Artificial Life Project
[4]University of California at Santa Cruz
bdamer@digitalspace.com

Abstract

The Evolution Grid, or EvoGrid is a computer simulation framework for distributed artificial chemistry (AC) supporting computational origins of life (COoL) research. The EvoGrid consists of a number of small experiments running on short time scales pruned by aggressive tree-branching searches supported by random parametric re-seeding and temporal back-tracking. The EvoGrid is designed to converge upon the observation of "cameo" simulations of key pre-biotic or simple biological structures or behaviors. These cameo simulations can then inform and feed larger AC simulations operating over biologically relevant time scales. In addition, the framework is designed to plug into a heterogeneous set of engines ranging from high fidelity molecular dynamics (MD) to more abstract AC techniques on the same set of data. The EvoGrid also provides shared web-based simulation management services and uniform, open standards for execution, storage and data analysis. We conclude by describing the first prototype implementation of the EvoGrid, early results, next steps and open questions in this and other COoL endeavors.

Introduction

In their seminal paper Open Problems in Artificial Life (Bedau et al., 2000) the authors set a challenge in the second open problem to "achieve the transition to life in an artificial chemistry *in silico*" (p. 364) while also identifying that "[b]etter algorithms and understanding may well accelerate progress… [and] combinations of… simulations… would be more powerful than any single simulation approach" (p. 367-68). The authors also point out that while the digital medium is very different from molecular biology, it "has considerable scope to vary the type of 'physics' underlying the evolutionary process" and that this would permit us to "unlock the full potential of evolution in digital media" (p. 369).

All of this potential awaits further progress in the computational challenges of high fidelity (i.e. accurate and predictive) artificial chemistries. Current state-of-the-art artificial chemistries (AC) (Dittrich, et al., 2001) including molecular dynamics (MD) projects utilize large centralized general-purpose computer clusters or, more recently, purpose built hardware, such as Anton, an MD supercomputer (Shaw, et al., 2009). Simulating tens of thousands of atoms for days to weeks on a commodity cluster will produce a number of nanoseconds of real-time equivalent chemistry. Optimized software running on Anton promises milliseconds of real-time equivalent ACs in weeks of computation (Shaw, et al., 2008).

To meet these challenges, proposals to unify efforts into larger computational origins of life (COoL) endeavors have been brought forth. Shenhav and Lancet (2004) propose utilizing the Graded Autocatalysis Replication Domain (GARD) statistical chemistry framework (Segre and Lancet, 1999, 2000). These authors have developed a hybrid scheme merging MD with stochastic chemistry. In GARD many short MD computations would be conducted to compute rate parameters or constraints for subsequent stochastic simulations. Thus, a federation of simulations and services was conceived which would also involve interplay with *in vitro* experiments. It is this vision for unifying efforts in COoL that has inspired our own work to build a framework for distributing and searching a large number of small chemistry simulation experiments.

As stated by Shenhav and Lancet, "the prebiotic milieu could best be characterized by a dense network of weak interactions among relatively small molecules" (p. 182). Simulating such a soup represents yet another scale of complexity beyond the targets set by even the builders of Anton. While the simulating of the full pathway to life *in silico* seems like a journey of a thousand miles, the first few steps can be taken and may become less daunting when helped along by some innovative algorithmic and architectural short cuts.

A fundamental property of large scale (in time duration and population of objects) simulations is that for the most part they use a homogeneous approach to optimize computation. On the opposite end of the spectrum we propose to run a large number of small simulations. Such an approach would in theory support a heterogeneous network of simulation techniques which vary physics, levels of abstraction and could even employ selection methods and replication of results inspired by the process of evolution. This is the approach

taken by the authors in developing the Evolution Grid (or EvoGrid), to be discussed next.

EvoGrid Search Function

The basic concept behind the EvoGrid is what we are terming *cameo simulations*. Cameo simulations are comprised of no more than a few hundred or thousand particles representing atoms and small molecules running over short time scales and in multiple instances. The existence of those instances is governed by a search tree function which permits variations of initial conditions and the branching of multiple, parallel simulations. Variation of parameters and branching are under control of an analysis step which looks for interesting structures or behaviors within each cameo simulation *frame*. Frames deemed less interesting may be terminated so as to permit other branches to be explored to a greater extent. This approach is inspired by the class of genetic algorithms (GA) combined with hill climbing algorithms widely used in Artificial Intelligence (Russell and Norvig, 2003). It is a form of importance sampling (Kalos and Whitlock, 2008), and its relationship to Maxwell's Demon requires careful scrutiny (Maruyama et al., 2009).

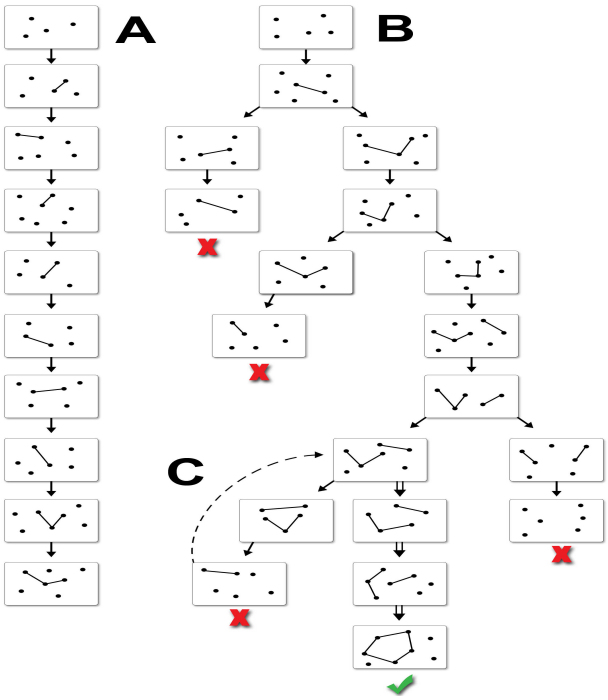

Figure 1: Illustration of the hill climbing search tree method employed by the EvoGrid

Figure 1 illustrates this method for a Control (A) which depicts a typical linear time sequence simulation and Test (B) which depicts the arising of simulation branches in this case due to selection for the phenomenon of more densely interconnected points. This illustration depicts another optimization called temporal back-tracking. If the simulation states of each frame can be stored through time, then a failed branch may be rolled back to the point at which "interesting" frames were still occurring. With a random seed applied, a new branch is started. This branch may yield a complex phenomenon forgone in the failed branch. In the example illustrated abstractly by C, that phenomenon might be a ring structure, as shown in the frame with the check mark. In this way, improbable occurrences may be guided across valleys of highly probable failure.

Genes of Emergence

Efforts to bridge nonliving and living matter and develop protocells from scratch (Rasmussen et al., 2003) will rely on bottom-up self assembly with commensurate self organization of classes of molecules. The development of repeatable self assembly experiments *in silico* (Rajagopalan, 2001) could serve as an important aid to *in vitro* protocell research. Self assembly in simulation may be purposefully designed into the experiment or may be an emergent phenomenon discovered by a directed search through multiple trial simulations. The initial conditions for a simulation could be equated to the coding sequences of a genetic algorithm (GA), and the simulation outputs seen as its expressed phenotype. The EvoGrid's search for self-assembly and other phenomena in cameo simulations is therefore a search for what we might term "genes of emergence" (GoE).

GoEs may be derived from within many different types of simulation, not just in the computationally intensive MD world. More abstract simulation modalities may yield shorter pathways to the production of important emergent phenomena than through computationally complex ACs (Barbalet et al., 2009). One could then see that the EvoGrid represents a "discovery system" operating on a continuum of techniques which might include: the execution of simulation modules that code for abstract universes yielding interesting results, to be then swapped out for a simple AC within which we would hope to reproduce the results, and finally, carrying the GoEs one step further into high fidelity MD, then which could inform validation through full scale *in vitro* experimentation.

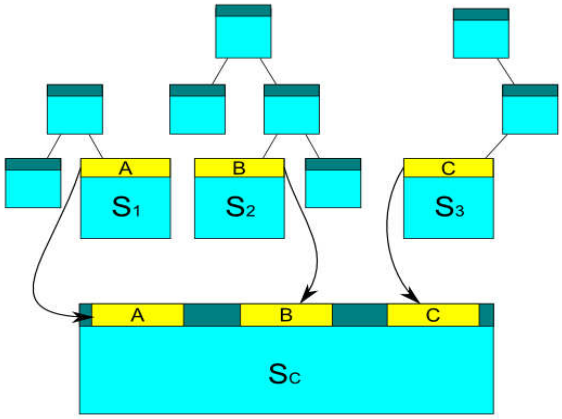

Figure 2: Illustration of the concept of cameo simulations feeding a larger composite simulation.

Figure 2 graphically illustrates the first two stages of this continuum. In the first stage, hill-climbing search functions (represented here as trees) process through a number of small cameo AC simulations. The end-point simulations, shown here as S1, S2 and S3, each meet some criteria for generating a structure or behavior of relevance to a larger composite simulation Sc. In the second stage, Sc is constructed from a mixture of content from each of the "feeder" cameo simulations and is driven by an amalgamation of the individual simulation experimental parameters A, B and C. The hope is that this amalgamation in simulation Sc, running with a much larger content store and over biologically significant time scales, would generate a rich mixture of phenomena, such as the formation of membranes, emergence of replicators, or the observation of autocatalytic reaction pathways. It is this enriched simulation environment which could be the basis for more ambitious computational origin of life endeavors. In another twist, an interesting phenomenon observed in Sc could be captured, its parameters and local contents extracted and cameo simulations run to characterize and fine tune the phenomenon more closely, enabling another ratchet in the emergent power of the larger simulation.

EvoGrid Design and Operation

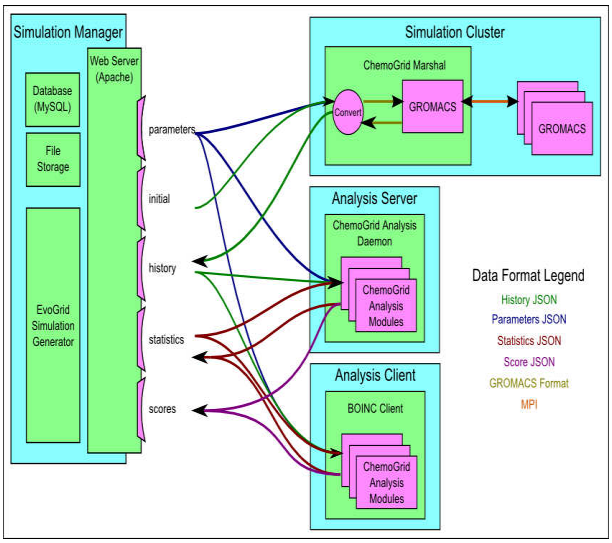

Figure 3: High level design and data flow of the EvoGrid

As depicted in Figure 3, the modular design of the EvoGrid encapsulates an MD simulation engine, in this case GROMACS (Van der Spoel, 2005), which we found to have good performance and was suitable to run as a plug-in component. GROMACS could be swapped out for other suitable simulation systems or the EvoGrid would support these systems running in parallel on the same data set. This architecture is designed to meet the challenge posed by Bedau et al. (2000) in which combinations of different simulation approaches might be a pathway to significant progress.

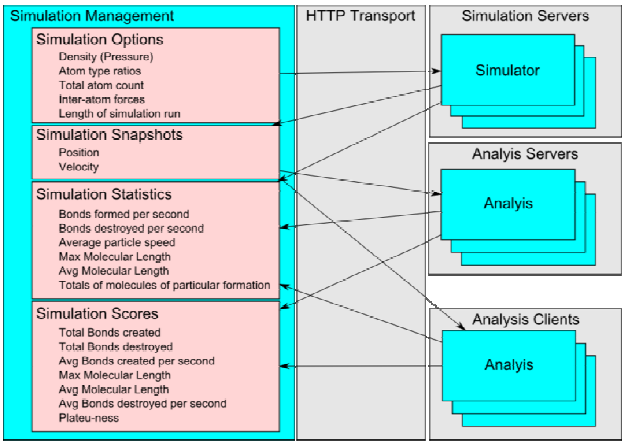

Figure 4: Lower level sequencing of data types through the EvoGrid

Other abstracted components depicted include an Analysis Server and an Analysis Client. Both of these components process inputs and outputs to the Simulation Cluster using the compact JSON format. The Simulation Manager running via HTTP/Web services sequences the simulation of and the analysis of individual frames (Figure 4). MD simulations typically have heavy compute loads in executing the time-steps for each force interaction of artificial atoms. In the EvoGrid, tens of thousands of frames are being executed and replicated through new branches. This generates terabytes of stored states for analysis. This could eventually call for a fully distributed simulation network, such as provided by the BOINC network (Anderson, 2004). BOINC supports many computationally intensive scientific applications, such as Folding@home (Pande et al., 2003). However, at this time we are relying on the centralized analysis server.

EvoGrid Prototype Runs and Results

A prototype of the EvoGrid architecture was built in 2009. Frames of 1,000 simulated atoms were run for 1,000 time steps within the GROMACS module with a uniform heat bath applied.

Initial conditions for GROMACS were:

- Density in particles per Angstrom: 0.01 - 0.1
- Temperature in Kelvin: 200 – 300, used for initial velocity and temperature bath
- Bond outer threshold in Angstrom: 0.1 - 1.0, distance, used for bond creation

The atoms ranged between three and ten randomly generated types. All their parameters (mass, charge, force interaction with other types, radius and volume) were selected from a uniformly distributed random range.

Forces between atom types included:
Pre-computed components of the Lennard-Jones force function:

- c6 0.0 - 0.1
- c12 0.0 - 0.00001

Covalently bonded (pre-computed components of the harmonic bond force function):

- rA 0.0 - 2.0
- krA 0.0 - 2.0
- rB 0.0 - 2.0
- krB 0.0 - 2.0

As an initial test case on a single instance of GROMACS when a bond was created, the Lennard-Jones forces would cease applying, and no new forces were applied. This was done to minimize real world constraints prior to having access to a computer cluster supporting covalent bond computations. The main focus of this prototype was to be able to test the architecture, not faithfully simulate the chemistry.

The position and velocity data was dumped every 1000 cycles and a naïve bonding applied to all atoms or atom-molecule or molecule-molecule objects. After a thousand of these dumps, this collected history was processed by the analysis server. Table 1 represents the scoring for frame number 144,204, the final frame in our trial run. The analysis was set up to look for the formation of "larger" virtual molecules, which in our simplistic interpretation meant a simple count of the greatest number of bonds between any two atoms. Employing Monte Carlo methodologies, the maximum search score reached in the trial was a simple sum of the entries in Table 1.

Measured values	**Final simulation scores**
Average molecular size	2.2303
Maximum average molecular size	4.47307
Average maximum molecular size	9.355
Maximum individual molecular size	17
Final maximum search score	33.0584

Table 1: Scoring produced by prototype analysis server for final simulation frame

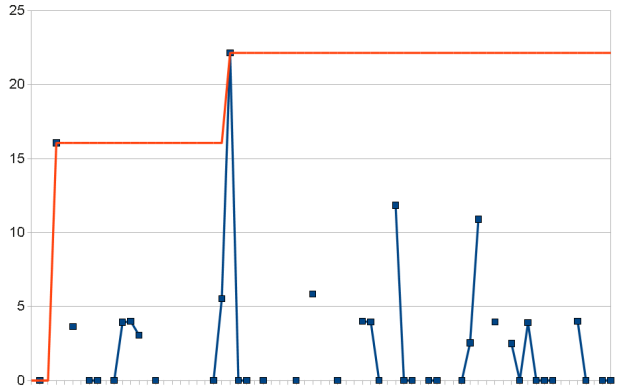

Figure 5: Scoring of experiments in "control" mode (random regeneration with no search tree function)

Figure 5 shows the "control" case (A) from figure 1 in which a random initial frame is simply run with a randomly seeded restarting of GROMACS for a duration of one thousand internal simulation steps (atom-atom interactions) with a thousand state dumps without the search function applied. As we can see, while there were some highly scored frames (red line), there is no maintained trend. Please note that the missing lines indicate cases where our software generated impossible simulation configurations and the execution was halted. This illustrated an area for improvement of how we were operating the GROMACS engine.

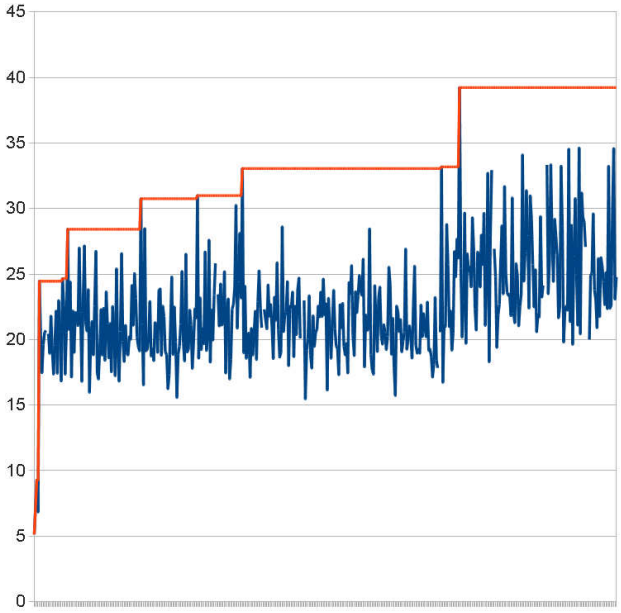

Figure 6: "Test" run showing trend toward higher "fitness" utilizing the search tree function

In Figure 6, the "test" case (B) from Figure 1 applies the search function, which clearly takes the initially high value produced by the same starting frame generated for the control case and improves on it over time. The strength of the search function is that subsequently generated frames eventually climb to a higher score-generating capacity ("fitness") over randomly generated control case frames. The search function will restart with lower performing simulations if all the potentially better options are exhausted. As seen in Figure 6, this causes a period where the evaluated simulation fitness (blue line) remains less than the best observed fitness (orange line). In this manner, the search function is operating as a Stochastic Hill Climbing algorithm in that the system has the ability to find its way out of traps set by local maxima.

EvoGrid Next Steps: Questions for the Computational Origins of Life

This very preliminary work poses far more questions than provides answers. However, as an early exemplar of computational origins of life (COoL) endeavors, the EvoGrid prototype and its proposed development path could serve as a roadmap to more fully functional platforms of the future. This roadmap also summons some broader issues, which might be considered a good start to a list of *open problems in computational origins of life.*

The greatest limitation in the EvoGrid prototype is our use of a naïve model of chemistry including the abstractness of our atom types, bond formation and the resulting "molecular

structures". Bonds are formed by simple proximity calculations using the positions, velocities and other data for objects exported from GROMACS. This situation may be improved by using the MOPAC7 library (Stewart, 2008) employed by GROMACS for covalent bond formation and the representation of other molecular affinities such as those produced by electrostatic and van der Waals forces:

1. Related to this first limitation is the need to go beyond the initial proof of concept prototype which is restricted to abstract atoms assembling into molecules. Our next steps must involve molecules assembling into larger structures that have the potential to exhibit properties of evolution. When this capability is prepared, a "real" set of experiments for testing the capabilities of the EvoGrid architecture should be attempted. Some proposed experiments include support for MD or coarse-grained simulation of lipid bilayer assembly reproducing the work of Fellerman (2009) using LAMMPS (Plimpton, 1995). Another good early test case would be to reproduce a simplified version of the groundbreaking experimental work by Bartel and Szostak (1993) in the isolation of new ribozymes from a large pool of random sequences.

2. The storage of frame states will be implemented in the near future. Temporal back-tracking is now being improved which will enhance the selective power of the search tree function. In addition, the computing resources of CALIT2 at the University of California at San Diego have been offered to the project, giving us critical storage and multiprocessor clusters for the next testing of the framework. A full work-up of computing and storage resources required by this architecture operating at different levels of simulation would be of value. Axes on a plot of EvoGrid computational complexity might include: number of particles and types of interactions handled for volume and time frame simulated, and desired level of fidelity to chemistry.

3. Another significant test of this concept would be the integration of simulation platforms other than GROMACS within the EvoGrid architecture to support heterogeneous simulations. For example, numerous engines, along the continuum of artificial chemistries from the highly abstract to the highly faithful to chemistry, are candidates to be integrated. In no particular order, candidate platforms are: The Organic Builder (Hutton, 2009), Avida (Adami and Brown, 1994), GARD (Segre and Lancet, 1999), NAMD (Philips et al., 2005), Desmond from Shaw et al (2008), and possible tie-ins to GPU-based hardware platforms (Anderson, 2008).

4. Bedau et al (2000) call for creating frameworks for synthesizing dynamical hierarchies at all scales. The heterogeneous nature of EvoGrid simulations would allow for coarse-graining procedures to focus simulation from lower levels to higher ones, saving computing resources by shutting off the less critical, more detailed simulations below. An example of this would be to switch to coarse grained simulation of an entire lipid vesicle, ceasing simulation of individual vesicle wall molecules. Conversely, fine grained simulations could be turned on for locally important details, such as diffusion of molecules through vesicle membranes. As exciting as this all sounds, a decade in the world of 3D simulation platforms has taught the authors of this paper that interfacing different software engines and representations of simulation space is extremely difficult. Running the same simulation space at multiple scales employing multiscale physics (e.g. from MD to dissipative particle dynamics, and beyond to smooth particle hydrodynamics) is also a very challenging problem that awaits future research.

5. A general theory of so-called cameo simulations needs to be developed to understand the minimum number of interacting objects and physical simulation properties required in these simulations for the emergence of "interesting" phenomena pertinent to life's building blocks. Our hypothesis that the GoEs in cameo simulations would apply to larger simulations also needs to be tested in the context of more ambitious COoL efforts capable of supporting artificial evolution thereby giving credence to the "Evo" in EvoGrid.

6. The EvoGrid cannot escape the meta-problem of all designed simulation environments: if we set up and simulate a system acting in the ways we accept as probable, then that system is much less likely to act in improbable and potentially informative ways, as results are always constrained by the abstractions and assumptions used. Another way of stating this very central conundrum is that as long as we do not know how chemical molecules might be able to exhibit emergence of important characteristics such as replication we will not be able to design the fitness functions to actually select for these molecules or their precursors. The fitness-function generation problem is as yet unsolved. However, the EvoGrid framework is being built to: 1) allow each potential experimenter to code in their own definition of fitness, accumulating knowledge applicable to the problem in an iterative fashion; and 2) support a more exotic solution in which the search functions themselves 'evolve' or 'emerge' alongside the simulation being searched. Actually building the second option would first require a much more extensive treatment from the field of information theory.

7. There are the deeper considerations that reach back to Langton who coined the term "artificial life" (Langton, 1986) and envisaged an investigation of *life as it could be*. COoL systems need not be constrained to models of the emergence of life on Earth. More abstract simulations may shine a light on *life as it might be* out in the universe (Gordon and Hoover, 2007), as a tool for use in the search for extraterrestrial intelligence (SETI) (Damer, 2010), or as a *technogenesis* within computing or robotic worlds.

8. A critic of theories of chemical evolution, cosmologist Sir Fred Hoyle used the statement about a ready-to-fly 747 aircraft being assembled by a tornado passing through a junk yard of parts (Hoyle 1984) to ridicule the idea of spontaneous generation of life at its origin. This idea today fuels creationist claims for irreducible complexity as one of their strongest arguments for the existence of a Creator. Like it or not, this flavor of debate will find its way to practitioners of COoL efforts. Gordon (2008), Damer (2008) and Barbalet and Daigle (2008) take this theme head on within a compendium of dialogues between creationists and scientists.

9. A corollary to Gordon's prediction (Gordon, 2008, p. 359) that Alife enthusiasts have an opportunity to solve the "Origin of Artificial Life" problem well before the chemists will solve the "Origin of Life" problem, is the very question of "what defines something as being life?". In the case of an *in silico genesis* we would ask "when will we know something is artificially alive?" Given latitude to speculate about these grand questions from such lofty heights of ignorance, it will be no surprise if emerging COoL endeavors attract a wide and vocal variety of converts and critics alike.

10. In the end the key question must be asked is: of what relevance is digital simulation to real chemistry or biology? Any given computational system might be able to show fascinating emergent phenomena but such discoveries might well stay trapped *in silico* and never transition over to inform experimentation *in vitro*. This would indeed be a shame and as such should motivate builders of systems like the EvoGrid to keep their eye on the ultimate prize: the transfer of concepts developed digitally into chemical experimentation. The inevitable marrying of these two media will produce one of the most powerful new tools for science and technology in the 21st Century.

Conclusion

A hybrid synthesis has been proposed between large scale high fidelity molecular dynamics simulations and distributed cameo simulations acting as an aggressive discovery system for the *genes of emergence* for some of life's building blocks. The EvoGrid is a framework under construction to support such distributed cameo simulations. Early results from a prototype implementation indicate that our search tree with temporal back-tracking optimization is performing as predicted as a stochastic hill climbing system. The EvoGrid software architecture has been shown to operate successfully with a large number of small, naïve chemical simulations run with the support of an industry standard MD engine. A listing of the current system's shortcomings and a roadmap for future development of the EvoGrid was presented. The authors concluded with a look at a few of the open questions applicable to the emerging field of computational origins of life (COoL) which is dedicated to "achieve the transition to life in an artificial chemistry *in silico*" (Bedau, et al. 2000).

Acknowledgements

We would like to thank Galen R. Brandt for a critical reading and Elixir Technologies Corporation for funding support of this work.

References

Adami, C., Brown, C.T. (1994). Evolutionary learning in the 2D artificial life system "Avida" In R. Brooks & P. Maes (Eds.), *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, Cambridge, MA: MIT Press, 377–381.

Anderson, D.P. (2004), BOINC: A system for public-resource computing and storage, *Proceedings of the 5th IEEE/ACM international Workshop on Grid Computing* (November 08 - 08, 2004), International Conference on Grid Computing. IEEE Computer Society, Washington, DC, 4-10.

Anderson J.A., Lorenz C.D., Travesset A., (2008), General purpose molecular dynamics simulations fully implemented on graphics processing units, *Journal of Computational Physics*, 227 (10), 5342-5359.

Barbalet, T., Daigle J.P. (2008)
Biota Podcast 19, Retrieved April 8, 2010 from http://www.archive.org/download/biotapodcasts/biota_052308.mp3

Barbalet, T., Damer, B., Gordon, R. (2009)
Biota Podcast 44, Retrieved April 8, 2010 from http://www.archive.org/download/biotapodcasts/biota_032909.mp3

Bedau, M.A., McCaskill, J.S., Packard, N.H., Rasmussen, S., Adami, C., Green, D.G., Ikegami, T., Kaneko, K., & Ray, T.S. (2000). Open problems in artificial life. *Artificial Life*, 6, 363–376.

Bartel, D., Szostak, J., (1993), Isolation of new ribozymes from a large pool of random sequences, *Science*, 261, 1411-1418.

Damer, B., (2008), The God detector, In: *Divine Action and Natural Selection: Science, Faith and Evolution*, Eds.: R. Gordon & J. Seckbach. Singapore, World Scientific: 67-82.

Damer, B., (2010), SETI Institute Colloquium Lecture, retrieved June 10, 2010 from http://www.seti.org/csc/lecture/archive/2010

Dittrich, P., Ziegler, J., Banzhaf, W., (2001). Artificial chemistries - a review, *Artificial Life*, 7, 225–275.

Fellermann, H. (2009). Spatially resolved artificial chemistry, In: A. Adamatzky and M. Komosinski (eds.), *Artificial Life Models in Software 2nd Edition*, Springer, 2009: 343-370.

Gordon, R. (2008). Hoyle's tornado origin of artificial life, a computer programming challenge. In: *Divine Action and Natural Selection: Science, Faith and Evolution*. Eds.: R. Gordon & J. Seckbach. Singapore, World Scientific: 354-367.

Gordon, R., Hoover, R.B., (2007). Could there have been a single origin of life in a big bang universe? Proc. SPIE 6694, doi:10.1117/1112.737041.

Hoyle, F. (1984). *The Intelligent Universe*. New York, Holt Rinehart and Winston.

Hutton, T.J., (2009) The Organic Builder: A public experiment in artificial chemistries and self-replication, *Artificial Life*, 15(1), 21-28

Kalos, M.H. & P.A. Whitlock (2008). *Monte Carlo Methods*. Weinheim, Wiley-VCH Verlag GmbH & Co. KGaA, 2nd Edition.

Langton, C.G. (1986). Studying artificial life with cellular automata. *Physica D*, 22, 120–149.

Maruyama, K., Nori, F., Vedral, V. (2009). Colloquium: The physics of Maxwell's demon and information. *Reviews of Modern Physics* 81(1), 1-23.

Pande, V.S., Baker, I., Chapman, J., Elmer, S.P., Khaliq, S., Larson, S.M., Rhee, Y.M., Shirts, M.R., Snow, C.D., Sorin, E.J., and Zagrovic, B. (2003), Atomistic protein folding simulations on the submillisecond time scale using worldwide distributed computing. *Biopolymers*, 68(1), 91-109.

Phillips, J.C., Braun, R., Wang, W., et al., (2005), Scalable molecular dynamics with NAMD, *J. Comput. Chem*, 26(16): 1781-1802.

Plimpton, S.J., (1995), Fast parallel algorithms for short-range molecular dynamics, *J Comp Phys*, 117, 1-19.

Rajagopalan, R. (2001). Simulation of self-assembly systems, *Current Opinion in Colloid and Interface Science*, 6, 357-365.

Rasmussen S., Chen L., Nilsson M., Abe S., (2003), Bridging nonliving and living matter, *Artificial Life*, v.9 n.3, p.269-316, Summer 2003

Russell, S.J., Norvig, P. (2003), *Artificial Intelligence: A Modern Approach (2nd ed.)*, Upper Saddle River, New Jersey: Prentice Hall, 111-114.

Segre, D., Lancet, D. (1999), A statistical chemistry approach to the origin of life, *Chemtracts – Biochemistry and Molecular Biology* 12, 382–397.

Segre, D., Lancet, D. (2000), Composing life, *MBO Reports* 1, 217–222.

Shaw, D.E., Dror, R. et al. (2009), Millisecond-scale molecular dynamics simulations on Anton, *Proceedings of the ACM/IEEE Conference on Supercomputing (SC09)*, Portland, Oregon, November 14–20, 2009

Shaw, D.E., Martin M. et al. (2008), Anton, a special-purpose machine for molecular dynamics simulation, *Communications of the ACM*, 51 (7), 91–97.

Shenhav B., Lancet D. (2004), Prospects of a computational origin of life endeavor, *Origins of Life and Evolution of Biospheres*, 34(1-2), 181-94.

Stewart, J.P., MOPAC2009, Stewart Computational Chemistry, Colorado Springs, CO, USA, HTTP://OpenMOPAC.net (2008).

Van der Spoel, D., Lindahl, E., Hess, B., et al. (2005), GROMACS: Fast, flexible, and free, *Journal of Computational Chemistry*, 26(16), 1701-1718.