The EvoGrid: *in silico* simulation grid for origin of life endeavours

Bruce Damer, Peter Newman, Ryan Norkus, Miroslav Karpis – UC Dublin & DigitalSpace
John Graham - Calit2, University of California, San Diego

1. INTRODUCTION

The Evoluton Grid, or EvoGrid, is an open, distributed artificial chemistry simulation grid now under development (Damer et al., 2010). The EvoGrid is designed to bring industrial scale simulation to bear on problems in origin of life endeavors including the modeling of *de novo* emergence of structures and reaction sequences in a prebiotic chemical milieu (Figure 1). Initial trials of the EvoGrid prototype constructed by a team at California based company DigitalSpace are presented here. The EvoGrid can enable distributed processing of any artificial chemistry simulation engine. The GROMACS molecular dynamics (MD) code was selected for the first tests due to its use in bio-relevant simulations such as Folding@Home (Pande et al., 2003). These *in Silico* experiments be predictive and useful for chemists working *in Vitro* based on the direction set by (Bedau et al, 2000) in the open questions in Artificial Life.

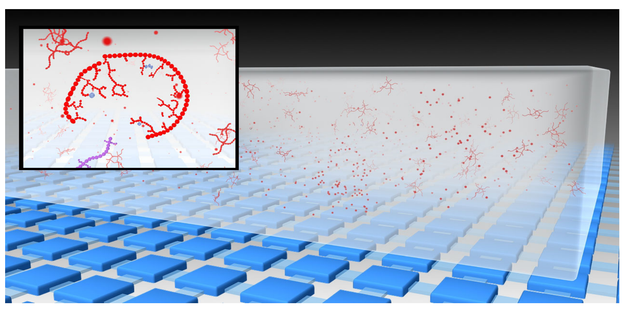

Fig. 1. Vision of the EvoGrid: a cartoon representation of a simulation grid with the observation of the spontaneous formation of a structure.

2. MODELS & METHODS

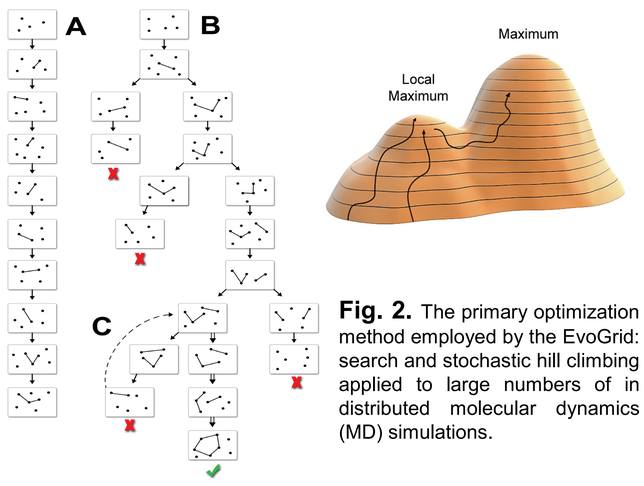

Fig. 2. The primary optimization method employed by the EvoGrid: search and stochastic hill climbing applied to large numbers of in distributed molecular dynamics (MD) simulations.

Due to the extreme computational costs involved in MD simulations, the EvoGrid focuses on simulating large numbers of small volumes containing 1K-10K atoms and employing search functions to select promising simulations for future propagation. Thus a stochastic hill climbing method (Figure 2) is implemented accross the search space. Our hypothesis was that linear simulations of such volumes (A) would produce fewer interesting phenomena than those guided by directed search (B-C). Therefore, the challenge of the first prototype was to show that this optimization can work in a distributed MD simulation supporting the formation of notional molecular bonds.

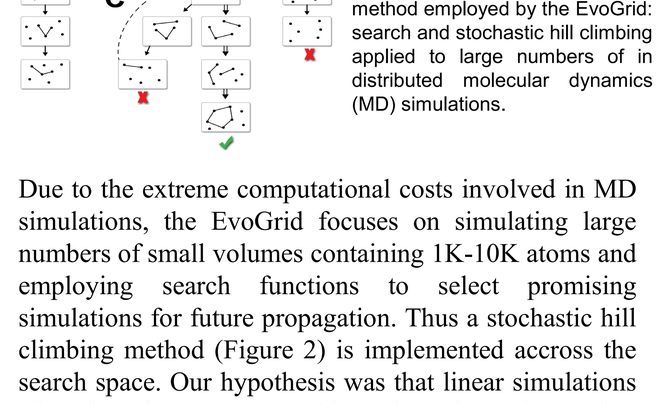

Fig. 3. Conceptual work flow of tasks within the EvoGrid

Figure 3 illustrates the flow of tasks within the first prototype EvoGrid, which implements a continuous cycle of simulation, analysis and search, generating weighted scores which are then reported back to the simulation grid.

3. IMPLEMENTATION

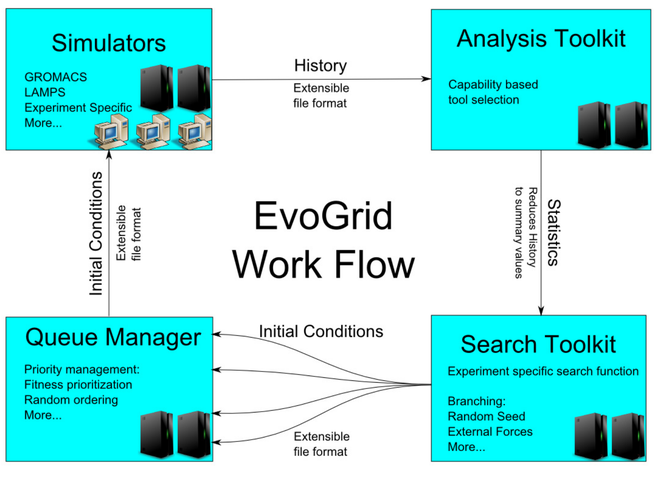

Fig. 4. Logical work flow of tasks within the EvoGrid.

Figure 4 illustrates how simulation volumes are processed first by engines (such as GROMACS) then passed through analysis, searched and then placed in priority queues for continued simulation or abandonment.

4. RESULTS

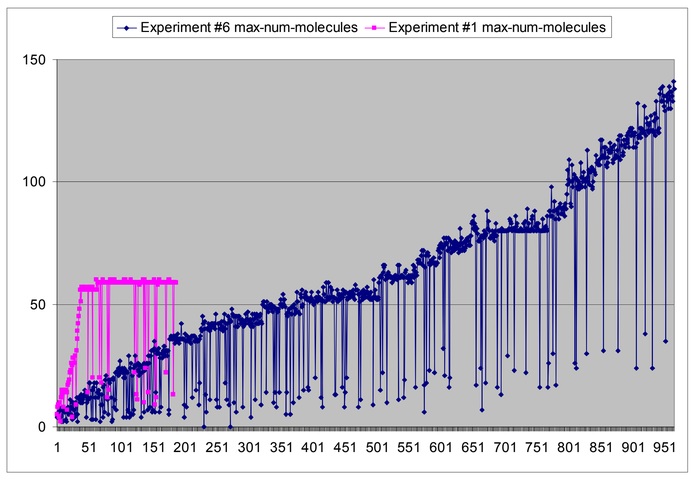

Fig. 5. Growth in the number of molecules for two experiments

Figure 5 illustrates the results of two experiments in which the search function (the teleological goal) was the formation of larger numbers of molecules from an atomistic soup. In Experiment #1 a total of 60 molecules were formed within 1K atom volumes after two months of computing time on a small grid of 4 servers. 187 1K atom simulations were processed through 1,000 time steps, while 11,015 simulations were abandoned. In Experiment #6 with a modified search function permitting the discovery of higher maxima 141 molecules were observed through 966 processed and 61,415 abandoned. In a control experiment, 26 molecules where observed but there was no sustained trend: no ability to sustain a local maximum and search for a higher one.

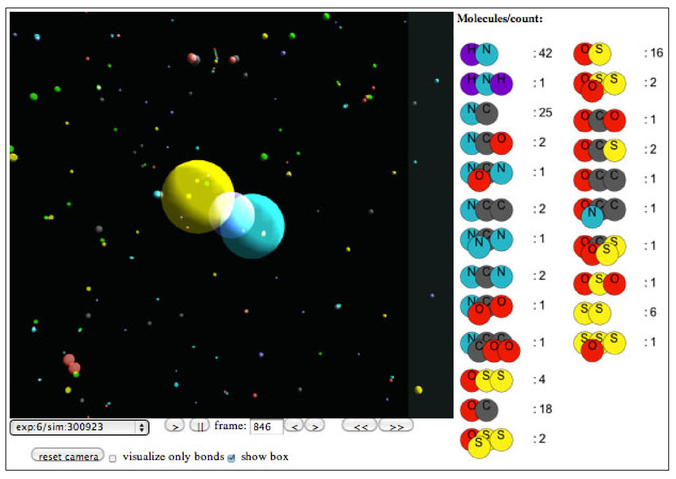

Fig. 6. 3D view of virtual molecular products in high yield volume

Figure 6 shows a 3D view of one of the volumes simulated in Experiment #6 with the notional molecules formed (right side) and one such molecule within the volume (center). It should be noted that the bond formation implemented for the prototype was highly naïve and designed only to test the hill climbing optimization method. Future versions of the EvoGrid will implement more realistic chemistry.

5. CONCLUSIONS

The EvoGrid is designed to bring industrial scale computing to origin of life experiments such as simulating the emergence of lipids and their association into micelles or vesicles, the arising of auto-catalytic sets, and scenarios for polymerization and replication of informational molecules. The following road-map is proposed for future EvoGrid implementations beyond the current prototype (Damer, 2011). These simulation targets are listed in order of perceived difficulty and provide a ten to twenty year time horizon for the ability of computation to handle the suggested volumes over bio-relevant time scales.

Experiment #1: Astrochemistry Model (Allamandola)
Experiment #2: FLiNT Nanocell Model (Fellermann, Rasmussen et al.)
Experiment #3: Riboyzyme Selection Experiment Model (Szostak, Bartel et al.)
Experiment #4: Model of a Hypopopulated Reaction Graph for the Study of Autocatalytic Sets and the Adjacent Possible (Kauffman et al)
Experiment #5: Model for RNA-Making Reactors in Hydrothermal Vents (Russell et al.)
Experiment #6: Model for Encapsulation of Polymers in Multilamellar Structures through Wet/Dry Cycles (Deamer et al.)
Experiment #7: Model of the FLiNT Protocell Life Cycle (Rasmussen et al.)
Experiment #8: Complex Free Encapsulation Origin of Life Model (CREATR-Damer et al.) – *see other poster at this conference*.

Call for collaboration

The authors invite proposals to collaborate on the next phases of implementation of the EvoGrid to incorporate improved chemical bond formation, faster engines and non-trivial experiments such as those listed here. All source code for the framework is available under open source terms and current efforts continue to extend the framework to incorporate other artificial chemistry codes.

Contact

Bruce Damer
Email: bdamer@digitalspace.com
Telephone +1 831 338 9400
DigitalSpace Corporation: www.digitalspace.com

Project web site: www.evogrid.org

LITERATURE CITED

Bedau MA, McCaskill JS, Packard NH, Rasmussen S, Adami C, Green DG, Ikegami T, Kaneko K, & Ray TS (2000). Open problems in artificial life. *Artificial Life*, 6, 363–376.

Damer B, Newman P, Gordon R, Barbalet T, Deamer D, Norkus R (2010). The EvoGrid: A Framework for Distributed Artificial Chemistry Cameo Simulations Supporting Computational Origins of Life Endeavors, *Proceedings of the 12th Conference on Artificial Life* (Aug 18-23, Odense, Denmark), MIT Press, Cambridge, MA.

Damer, B. 2011. The EvoGrid: *An Approach to Computational Origins of Life Endeavours*. PhD, University College Dublin.

Pande VS, Baker I, Chapman J, Elmer SP, Khaliq S, Larson SM, Rhee YM, Shirts MR, Snow CD, Sorin EJ, and Zagrovic B (2003). Atomistic protein folding simulations on the submillisecond time scale using worldwide distributed computing. *Biopolymers*, 68(1), 91-109.

ACKNOWLEDGEMENTS

R. Gordon, Department of Biology, Univ. of Manitoba
D. Deamer, UC Santa Cruz
T. Barbalet, Biota.org
F. J. Dyson, Institute for Advanced Study
NASA and Elixir Technologies Corp for funding support
University College Dublin/SMARTlab