Parallel Exact Stochastic Simulation of Gene Networks

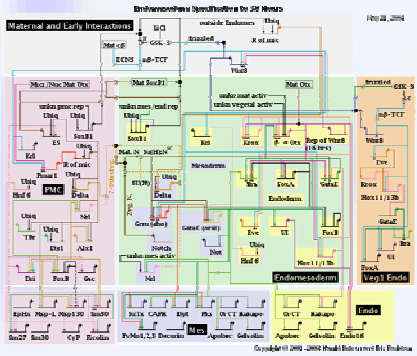
TRAYVON D. LESLIE
ALABAMA A&M UNIVERSITY

Research Alliance in Math and Science
Computer Science and Mathematics Division
Mentor: Dr. Nagiza Samatova

http://www.csm.ornl.gov/Internships/RAMS.html

ABSTRACT

The applicability of computer modeling to whole-cell and multi-cell biochemical models is limited by the accuracy and efficiency of the simulation tools used to model gene regulatory networks. It is widely accepted that exact stochastic simulation algorithms, originally developed by Gillespie, accurately depict the time-evolution of a spatially homogeneous biochemical model, but these algorithms are often abandoned by modelers because their execution time can be on the order of days to months. This work improves the performance of exact stochastic simulation algorithms by developing parallel implementations and measuring their performance for a wide variety of models and emerging computer architectures such as SGI Altix and Cray X1.



$$\frac{dDNA}{dt} = 0$$

$$\frac{dMRNA}{dt} = k_1 DNA - k_2 MRNA$$

$$\frac{dP}{dt} = k_2 MRNA - k_3 P - k_4 P^2 + k_5 Dimer$$

$$\frac{dDimer}{dt} = k_4 P^2 - k_5 Dimer$$

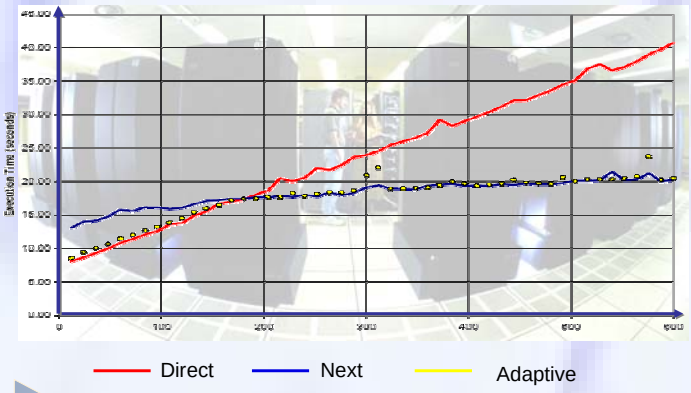
OBJECTIVE AND PLAN

- **Improve ESS performance, by parallelizing Gillespie's algorithms based on parallel discrete event simulation techniques.**
- **Treat the systems as a stochastic process with discrete variables that represent the populations, not concentrations, of chemical species.**
- **Identify performance bottlenecks and areas for improvements**
- **Show a significant performance improvement over a sequential simulation**

Gillespie's First Reaction Algorithm

1. **Initialize M chemical reactions and kinetic constants**
2. **Initialize current time**
3. **Calculate propensity for each reaction**
4. **Generate time for each reaction**
5. **Select reaction with minimal time**
6. **Execute the selected reaction**
7. **Update concentrations of molecular species**
8. **Update time and go to Step 4**

Adaptive vs. Direct Method Performance And Next



CONCLUSION:

Implementing the concepts proposed in the above document will hopefully show a significant performance improvement over using a sequential simulation algorithm. This work will also hopefully serve as another major step in increasing the applicability of exact stochastic simulation to biochemical system modeling.

UT-BATTELLE

OAK RIDGE NATIONAL LABORATORY
U.S. DEPARTMENT OF ENERGY

Office of Science
U.S. DEPARTMENT OF ENERGY

The Research Alliance in Math and Science program is sponsored by the Mathematical, Information, and Computational Sciences Division, Office of Advanced Scientific Computing Research, U.S. Department of Energy. The work was performed at the Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC under Contract No. De-AC05-00OR22725. This work has been authored by a contractor of the U.S. Government, accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

Acknowledgements: I WOULD LIKE TO THANK MY MOTHER, DARLENE LESLIE FOR BEING BEHIND ME 100 PERCENT. I WOULD ALSO LIKE TO THANK DR. Z.T. ZENG (FACULTY ADVISOR; ALABAMA A&M) FOR INTRODUCING THE OPPORTUNITY OF THE RAMS PROGRAM TO ME. MY MENTOR, DR. NAGIZA SAMATOVA, MRS. DEBBIE MCCOY, MS. CHERYL HAMBY AND THE REST OF THE RAMS PARTICIPANTS WHOM HAVE MADE MY EXPERIENCE HERE AT OAK RIDGE A DELIGHTFUL ONE.

This research used resources of the Center for Computational Sciences at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.