MAC5701 – Tópicos em Ciência da Computação
Plano de Estudos

Aluno: Edgar Szilagyi
edgar@cecm.usp.br

Orientador: Prof. Dr. Marco Dimas Gubitoso
gubi@ime.usp.br

Área: Modelagem e Simulação

IME–USP
São Paulo, 25 de março de 2003

1 Introdução

A área da dissertação de mestrado do aluno será a modelagem e simulação computacionais de sistemas dinâmicos estocásticos discretos bem como de autômatos finitos celulares. Particularmente, o objetivo de tal projeto é construir um software de edição e simulação para o estudo de redes de regulação de expressão gênica. Pretende-se, através deste plano de estudos, aprofundar os conhecimentos envolvidos na modelagem e principalmente na elaboração do software.

2 Motivação

Atualmente, a ciência da bioinformática vem crescendo vertiginosamente. Uma de suas fortes áreas está em congruência de cadeias em diferentes formas, onde é permitido permutas de elementos, taxas de erros controláveis, e uma infinidade de outras características que tornam esse problema bastante particular. Sem essa tecnologia, seriam inviáveis, senão impossíveis, projetos de sequenciamentos de DNA.

Outra vertente da bioinformática é a análise de resultados obtidos a partir de experimentos bioquímicos acerca do funcionamento das células. A tecnologia atual permite que seja gerado um enorme volume de dados a partir de apenas alguns poucos experimentos. Lidar com esse enorme volume de dados sem o apoio computacional também seria inviável. Podemos citar algumas contribuições da computação para este problema como a mineração e agrupamento de dados, análise e extração automática de dados a partir de imagens e inferência de sistemas.

Essa última possui uma peculiaridade muito interessante, pois permite que se automatize a geração de modelos para se descrever o sistema biológico em questão. Entretanto, esses modelos são gerados a partir de hipóteses muitas vezes vagas e não totalmente compreendidas e a partir de dados com grande susceptibilidade a erros. Dessa forma é de suma importância que exista uma forma de se testar e validar esses modelos. É justamente nesse ponto que entra o trabalho do aluno. O que se propõe a ser feito na dissertação de mestrado é um software com funções de edição e simulação que auxiliem no estudo desses modelos.

3 O Simulador

O simulador será parte central da dissertação de mestrado. Será construído em C++ por estabelecer um bom compromisso entre desempenho, facilidade de programação e principalmente com manutenções e expansões futuras. Seu núcleo será fundamentalmente um simulador de sistemas dinâmicos estocástico discreto voltado para o problema das redes de regulação em expressão gênica. Porém sua implementação será de tal forma que permitirá a adição de componentes extras para extração de informações qualitativas do sistema, simulação de autômatos celulares entre outras. Essa estrutura em camadas permite que sejam especializadas diversas ferramentas de edição, simulação e análise, tornando o programa uma ferramenta de uso efetivo.

O software conterá pelo menos dois tipos de interface com usuário, uma em modo texto e outra em modo gráfico. A versão gráfica será voltada inclusive para usuários de outras áreas, como bioquímicos e biólogos, pois são eles que efetivamente farão uso do sistema. O projeto do software será tal que facilmente poderá ser implementada uma nova interface gráfica se necessária. Apesar de que um gene ou uma variável de estado sejam exatamente a mesma coisa, um profissional ou outro pode não se sentir a vontade diante destes termos.

O software possuirá também um analisador léxico/gramatical simples permitindo que se especifique o modelo que se deseja simular/editar através de uma linguagem simples.

4 Objetivos em Tópicos em Ciência da Computação

Diante do desafio de implementar esse sistema, propomos que o aluno aprofunde seu conhecimento nas áreas correlatas, apresentando aspectos importantes de cada uma delas:

- Modelagem
  - Sistemas dinâmicos estocásticos e discretos
  - Autômatos finitos
  - Modelagem e representação do problema bioquímico em termos matemáticos e computacionais
- Tecnologia Computacional
  - Programação orientada a objetos e projeto do sistema
  - Programação distribuída: CORBA
  - A linguagem de especificação de modelos para o simulador
  - A engenharia de software envolvida na elaboração do sistema

5 Cronograma para o primeiro semestre de 2003

Pretende-se abordar todos os tópicos citados de forma concomitante, inclusive com a elaboração de parte do software. Entretanto, o cronograma abaixo sugere as áreas em que o aluno terá maior concentração no decorrer do semestre:

- Abril - Modelagens e a linguagem de especificação
- Maio - Eng. de software e projeto do sistema
- Junho - Elaboração do relatório

6 Bibliografia

- Bioquímica e biologia
  - Lehninger, Albert L.; Nelson, David L.; Cox, Michael M.; *Principles of Biochemistry, 2nd edition*, Worth Publishing, 1993
  - Voet & Voet; *Biochemistry, 2nd edition*, John Wiley & Sons, 1995
  - Stryer, Lubert; *Biochemistry, 4th edition* , W H Freeman and Co, 1995
  - Lewin, Benjamin; *Genes VII, 7th edition*, Oxford University Press, 2000
  - Lawrence, Peter A.; *The Making of a Fly: The Genetics of Animal Design*, Blackwell Science Inc, 1992
- Matemática
  - Preparata, Franco P.; Yeh, Raymond T.; *Introduction to Discrete Structures for Computer Science and Engineering*, Addison-Wesley Pub Co, 1973
  - Desoer, C.A.; *A Second Course on Linear Systems*, Van Nostrand Reinhold Co,1969
  - Nelson, R.J.; *Introduction to Automata*, John Wiley & Sons, 1968
  - Demongeot, J.; Golès, E.; Tchuente, M.; *Dynamical Systems and Cellular Automata*, Academic Press, 1985
  - Lorenz, A.A.; *Stochastic Automata Constructive Theory*, John Wiley & Sons, 1974
  - Bailey, Norman T.J.; *The Elements of Stochastic Processes with Applications to the Natural Sciences*, John Wiley & Sons, 1964.
  - Goodman, Roe; *Introduction to Stochastic Models*, Benjamin/Cummings, 1988
- Computação
  - Sommerville, Ian; *Software Engineering, 6th edition*, Addison-Wesley Pub Co., 2000

– Pressman, Roger S.; *Engenharia de Software, 5a edição*, McGraw Hill, 2002

– Pree, Wolfgang; *Design Patterns for Object-Oriented Software Development*, Addison-Wesley Pub Co, 1995.

– Bushmann, Meunier, Rohnerst, Sommerlad, Stal; *A System of Patterns*, John Wiley & Sons, 1998

– Mowbray Thomas J.; Malveau, Raphael C.; *CORBA Design Patterns*, Addison-Wesley Pub Co, 1997

– Vinoski, M.; Henning, S.; *Advanced CORBA Programming with C++*, Addison-Wesley Pub Co, 1999