

# As Perspectivas e Grandes Desafios da Computação no Brasil - 2006 a 2016 -

Título: *Computação na Biologia Molecular e Bionanotecnologia*

Proponente: *Leila Ribeiro*

## Descrição:

Quando falamos em *Computação*, geralmente associamos a noção de programa à máquina concreta na qual esses programas executam, os computadores baseados em chips de silício que temos hoje em dia nas nossas casas, laboratórios de pesquisa, etc. Mas as bases teóricas da Ciência da Computação são independentes de máquinas concretas. As máquinas que chamamos de computadores são uma possível implementação de máquina capaz de *computar*. Existem várias propostas de outras máquinas que podem executar esse trabalho, mas até agora elas não passaram de propostas acadêmicas. Os grandes avanços na área de Biologia Molecular da última década podem mudar esse panorama, dando à Ciência da Computação uma máquina completamente diferente: uma *máquina biológica*. Por outro lado, um modelo teórico dessa máquina biológica pode servir de base para a área emergente de Bionanotecnologia, provendo técnicas para estruturar e construir *programas biológicos*, e para ajudar a entender como a natureza *computa*.

Os organismos vivos são sistemas reativos: dependendo de estímulos (condições ambientais e presença ou ausência de certas substâncias), a célula produz uma reação (produzindo as proteínas/enzimas adequadas para tratar o estímulo). A informação sobre quais proteínas/enzimas devem ser geradas para cada situação está guardada no núcleo, no DNA. A área de Biologia Molecular começou a se desenvolver mais intensamente desde o descobrimento da estrutura do DNA, como uma dupla hélice de seqüências de bases (Adenina, Guanina, Citosina e Timina) e do chamado *Dogma Central da Biologia Molecular*, que explica como a célula transforma informações contidas no DNA em proteínas ou enzimas. O DNA contido em nossas células contém informações precisas sobre as seqüências de bases que geram cada proteína, e como e quando essas proteínas devem ser produzidas. Fazendo uma analogia, poderíamos ver o DNA como o software (englobando dados e programa), e a célula com o hardware biológico. Com o sequenciamento dos genomas de várias espécies, temos as listas de bases que estão guardadas nos núcleos das células de indivíduos dessas espécies. Conhecendo o DNA (software) e o funcionamento da célula (hardware), poderíamos entender/prever o comportamento fisiológico dos seres vivos. Isso possibilitaria, por exemplo, o projeto de remédios e terapias muito mais eficazes. Porém, ainda há um grande caminho a ser percorrido, e a Computação pode ter um papel importante nessa área.

O estágio atual das pesquisas pode ser comparado a termos uma listagem do conteúdo de um disco rígido, sem conhecer o sistema operacional nem o sistema de arquivos utilizado. Conhecemos a máquina, mas não a organização dos dados. Como descobrir, olhando um seqüência imensa de 0s e 1s que programas estão guardados em um computador? Aproximadamente 3% das  $3 \times 10^9$  bases que compõem o DNA humano são os chamados genes, ou regiões que codificam proteínas (ou seja, os dados). O restante do DNA, chamado DNA não-codificante, guarda a informação sobre quando e como as proteínas devem ser produzidas, ou seja, o programa que regula a expressão gênica.

Nesse contexto, existem muitos desafios científicos a serem solucionados, e a Computação pode dar muitas contribuições:

- 1) **Descobrimto da estrutura tridimensional do DNA:** ver o DNA como uma seqüência de (pares de ) bases é uma abstração. Na realidade, ele é um emaranhado de moléculas, obtidas através das ligações das bases que o compõe. Quando vemos a estrutura tridimensional de uma seqüência de bases, podemos ter idéias sobre o impacto que essa estrutura pode ter no seu meio, ou seja, sobre o seu funcionamento. Isso ocorre porque somente quando vemos a imagem 3D entendemos em quais tipos de outras moléculas esta pode se ligar, o que dá dicas sobre sua função. Conhecer a função de cada proteína (seja ela natural ou projetada artificialmente) é de fundamental importância para a modelagem biomolecular (pois as proteínas são os agentes dos

processos). Já existem programas para prever a estrutura 3D de sequências, mas ainda há muitas limitações no seu uso. A área de Computação Gráfica pode contribuir muito neste tópico.

- 2) **Simulação de interações entre moléculas:** além de visualizar tridimensionalmente as moléculas de DNA, é preciso poder prever se e como diferentes moléculas colocadas em um mesmo meio interagiriam. Essas interações, que são a base dos processos na biologia molecular, ocorrem se existir a possibilidade das moléculas se unirem (ou seja, se as formas 3D delas forem complementares). As áreas de Computação Gráfica, Inteligência Artificial e Simulação Computacional podem prover métodos mais eficazes para simular, analisar e prever essas interações.
- 3) **Modelagem de projetos biomoleculares:** na Computação, a área de software evoluiu de sistemas muito simples, implementações de funções simples, para sistemas bastante complexos, com milhares de linhas de código em linguagens de alto nível. Estamos acostumados a desenvolver software em fases, usando técnicas de refinamentos e modularização. Na biologia molecular, os processos ainda são descritos ou através de conjuntos de equações matemáticas (o que dá uma visão extremamente limitada, pois somente processos muito pequenos podem ser modelados nesse nível tão concreto), ou através de métodos ad hoc. Existem idéias de como modularizar ou descrever refinamentos, mas as técnicas de engenharia de software, principalmente as baseadas em métodos formais, poderiam contribuir muito nessa área. Já há trabalhos iniciais usando álgebras de processos ou redes de Petri para modelagem de vias metabólicas e/ou regulatórias, mas nos próximos anos, as áreas de Engenharia de Software e Métodos Formais podem oferecer várias outras técnicas de modelagem, análise e verificação para sistemas biológicos.
- 4) **Bionanotecnologia:** em 1996, Leonard Adleman mostrou que pode-se usar DNA e operações sobre DNA para resolver o problema do caixeiro viajante. Desde então, vários pesquisadores tem investigado como computar usando DNA. Normalmente, os experimentos são feitos in vitro, ou seja, as operações sobre as sequências são executadas pelo homem em laboratório. Mas, e se conseguíssemos usar a máquina que já existe, ou seja, a célula para executar programas com DNA? Esse é um dos objetivos da Biologia Molecular, e em especial, da Bionanotecnologia. Essa área se preocupa em produzir organismos (ou moléculas) muito pequenos, mas que carregam um código (ou programa) capaz de influenciar no comportamento das células nas quais eles são inseridos. Atualmente, a Bionanotecnologia tem se concentrado mais nos problemas físico-químicos da síntese dessas moléculas tão pequenas (nem toda sequência tem a estabilidade necessária). Mas as áreas de Teoria da Computação e Linguagens de Programação poderiam contribuir com técnicas de programação, já que hoje em dia a “programação biológica” é realizada de maneira ad hoc. Um novo modelo de computação poderia servir como base da Biologia Molecular e Bionanotecnologia. Note-se que nessa área, a maioria dos sistemas são críticos, ou seja, erros em programas podem ter consequências catastróficas. Por isso, técnicas formais de prova de correção seriam altamente desejáveis.

#### CV resumido de Leila Ribeiro:

- **Formação:** Doutora em Ciência da Computação (Universidade Técnica de Berlim, Alemanha/1996), Mestre em Ciência da Computação (UFRGS, 1991), Bacharel em Ciência da Computação (UFRGS, 1988).
- **Cargo:** Professor adjunto no Departamento de Informática Teórica da UFRGS (desde 1998), bolsista do CNPq – nível 2.
- **Áreas de interesse:** Métodos formais, Bioinformática, Engenharia de Software

Leila Ribeiro tem trabalhado na especificação, semântica e análise formal de sistemas reativos há vários anos. Nos últimos anos, tem usado técnicas de engenharia de software para a modelagem e análise de processos biológicos. Coordenou projetos nacionais (Laboratório de Bioinformática, PLATUS) e internacionais (Graphit, DACHIA, IQ-Mobile). De 1997 a 2002 foi consultora do MEC para avaliação de cursos de graduação na área de Informática. De 2001 a 2002 foi Coordenadora da Comissão de Pesquisa do Instituto de Informática da UFRGS. De 2003 a 2004 foi Chefe do Departamento de Informática Teórica da UFRGS. Foi coordenadora do comitê de programa do SBES (2002), SBMF (2000,2006) e International Conference on Graph Transformation (2006), além de se membro de comitês de vários outros eventos nacionais e internacionais. Por suas contribuições à área de Informática, ganhou em 1999 o *Prêmio Santista em Informática* – categoria Juventude.