

Bioinformática/Informática Médica

198

UM AMBIENTE BASEADO EM AGENTES PARA ANOTAÇÃO AUTOMÁTICA DE GENOMAS. *Abner Novak Pitinga, Ana Lucia Cetertich Bazzan (orient.)* (Departamento de Informática Teórica, Instituto de Informática, UFRGS).

Uma das grandes necessidades nos projetos genoma de hoje é a de analisar e anotar uma grande quantidade de seqüências de proteínas de forma rápida e eficaz. Entretanto, os pesquisadores envolvidos nessas pesquisas não possuem tempo suficiente para fazer a anotação completa dessas seqüências, o que torna óbvia a necessidade de se criar ferramentas automáticas para essa tarefa. Este trabalho propõe a criação de um ambiente baseado em agentes que auxilie em todo o processo de anotação de genomas. Esse ambiente usa técnicas de aprendizado de máquina para que, a partir de um conjunto de dados já anotados, sejam geradas regras para anotação de keywords para novas seqüências de proteínas. O conjunto de dados usados para geração dessas regras é proveniente do Swiss-Prot, um banco de dados de proteínas. As seqüências de proteínas usadas para geração de cada conjunto de regras estão relacionadas com a família do organismo com que se está trabalhando, ou seja, são escolhidas seqüências de proteínas de organismos da mesma família. Basicamente o trabalho resume-se em se obter os dados do Swiss-Prot, formatar esses dados para a sua utilização com algum algoritmo de mineração de dados, geração das regras e armazenamento dessas regras em uma base dados. (CNPq-Proj. Integrado).