



**Universidade de São Paulo**  
Instituto de Matemática e Estatística



# **Parte Subjetiva**

## **Adequabilidade do Kepler para pacotes de Bioinformática do R**

---

**ALUNO:**

André de Melo Folli

**ORIENTADOR**

Prof. Dr. Roberto Hirata Jr.

**2/19/2008**

## Desafios e Frustrações

No começo de 2007 tive a palestra inicial do Trabalho de Formatura e nela descobri que o trabalho não poderia mais ser feito sobre o estágio. A partir daí comecei a buscar um orientador e tema. Apesar do pouco tempo e de algumas adversidades encontradas no caminho, um grupo foi montado (composto por mim, o Gustavo e o Bruno), escolhemos o orientador (Roberto Hirata) e definimos um tema geral pro nosso trabalho (Bioinformática).

Fizemos algumas reuniões para descobrir quais eram as nossas habilidades/facilidades afim de especificar qual seria o trabalho propriamente dito. Tivemos inclusive uma reunião com o irmão do Gustavo (que é biólogo) com o intuito de fazer um trabalho com a participação dele como co-orientador. Essa tentativa não deu certo. Foram feitas mais algumas reuniões e definimos que faríamos a Reengenharia do pacote Maiges para o software R.

A partir daí começamos a estudar um pouco mais sobre Bioinformática (principalmente sobre análise de microarray), o software R (Entender a linguagem R, aprender a instalar pacotes, rodar tutoriais, etc) e o pacote Maiges mais a fundo (Entrando em contato com o código do pacote, estudando mais a fundo sua documentação, além de ler alguns capítulos da tese de Doutorado do Gustavo Henrique Esteves). Durante esse trabalho tivemos muitos empecilhos para instalar os pacotes que queríamos nos computadores do IME e começamos a ficar realmente desanimados. Durante esse tempo percebemos que nós não possuíamos conhecimento suficiente de biologia, ou seja, fazer a reengenharia do pacote não seria viável, pois poderíamos "quebrar" o funcionamento ideal deste e nem perceberíamos. A solução para esse impasse foi mudar o foco de nosso projeto. Após outras reuniões com o professor Hirata mudamos o foco do trabalho (tentamos não fugir do tema proposto), nesse novo foco aproveitamos o estudo do pacote Maiges, estudamos outros dois pacotes para Bioinformática do R (Limma e Marray, fizemos um estudo semelhante ao do pacote Maiges) afim de entender sua organização para podermos sugerir uma reengenharia (já que sabíamos que fazê-la não seria possível). Além disso, entramos em contato com o Kepler. O objetivo era automatizar a execução das funcionalidades do pacote Maiges com o Kepler, montando um workflow. Tivemos muito trabalho para conseguir instalar (nos computadores do IME) o Kepler e as bibliotecas do R que precisamos para rodar o pacote Maiges no Kepler e encontramos outros problemas técnicos para automatizar as funções do Maiges no Kepler (que acabaram motivando a mudança de tema que ocorreu logo depois em nosso projeto). Ainda nessa fase do projeto tentamos fazer um pacote ou função para o R que fizesse a análise de um pacote passado como parâmetro e devolvesse um diagrama de classes desse pacote, essa função seria muito útil na análise de outros pacotes do R. Apesar disso, essa tarefa foi abandonada logo pois descobrimos que ela seria muito trabalhosa e não se encaixava completamente no escopo do projeto. Nos vimos mais uma vez com o projeto empacado e o tempo ficando cada vez menor, percebemos que nesse tema não conseguiríamos terminar o trabalho a tempo.

Em uma nova reunião com o Prof. Hirata foi sugerido uma mudança de tema. Nesse tema novo (Adequabilidade do Kepler para pacotes de Bioinformática do R) pudemos aproveitar grande parte dos assuntos já pesquisados e foi possível realizar um bom trabalho.

## **Disciplinas cursadas no BCC mais relevantes para o trabalho**

Considero que as seguintes disciplinas estudadas no BCC foram as mais importantes pro trabalho:

MAC0122 – Princípios de Desenvolvimento de Algoritmos

MAC0323 – Estruturas de Dados

MAC0446 – Princípios de Interação Homem-Computador

## **Se o aluno fosse continuar atuando na área em que realizou o trabalho, que passos tomaria para aprimorar os conhecimentos relevantes para esta atividade?**

Finalmente, posso dizer que estou bastante contente com os resultados obtidos, que a realização desse trabalho foi mais difícil e trabalhosa do que eu imaginava e que, apesar de eu ter descoberto que não pretendo trabalhar com pesquisa na área de Bioinformática, creio que os problemas enfrentados foram bastante enriquecedores para mim e para o grupo.