*dedicado à memória da minha vó,
professora Mercedes,
que sempre soube que eu chegaria até aqui*

“O cara pode até ser 0,1% na estatística... mas para ele... ele é 100%”
— “Véio”, meu pai.

Agradecimentos

Obrigado aos meus dois orientadores, Prof. Dr. Carlos A. de B. Pereira (Carlinhos, para mim) e Prof. Dr. Edson Amaro Jr (tio Eds, para mim), pela orientação e diversão durante toda este trabalho. Sem a ajuda, orientação, financiamento e senso de humor de vocês, nada disso seria possível.

Sem o financiamento do projeto por parte das agência de fomento a pesquisa, este trabalho jamais teria sido realizado. Agradeço às agências: CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e principalmente ao IIEP – Instituto Israelita de Ensino e Pesquisa do Hospital Israelita Albert Einstein.

Quando as aquisições de imagens de ressonância magnética funcional não iam bem, a biomédica Liana Guerra Sanches salvou o projeto, tomando para si a tarefa de contatar os voluntários, organizar os experimentos, etc, etc, além de ter colaborado na realização do experimento. Na prática, quem adquiriu as imagens do banco-de-dados normativo e dos pacientes foi a Liana e por isso agradeço muito. Sem sua ajuda, a metade da tese, que trata de RMf, não teria sido possível. Toda vez que se mencionar “a equipe do Hospital Israelita Albert Einstein”, de fato, está-se referindo à Liana e ao radiologista Marcelo M. Felix, sob orientação próxima do Prof. Edson.

Pelo outro lado, na metade que trata de *microarray*, devo agradecer a todos os pesquisadores colaboradores que cederam seus dados e confiaram nos métodos aqui propostos para a análise destes, tão preciosos para eles. São os pesquisadores: Dr. Oswaldo Keith Okamoto e equipe, do Departamento de Pesquisa Experimental do Instituto Israelita de Ensino e Pesquisa – Hospital Israelita Albert Einstein; Camila M. Egídio e Dr. Eduardo M. Reis, do Departamento de Bioquímica do Instituto de Química – USP e Dr^a. Suely L. Gomes e Tie Koide, também do Departamento de Bioquímica do Instituto de Química – USP. Um agradecimento mais do que especial deve ser feito à Tie Koide, que contribuiu de forma decisiva na elaboração dos métodos aqui propostos, sempre disposta a discutir quaisquer assuntos relacionados à pesquisa com *microarrays*. Por sua força e incentivo, alguns trabalhos citados nesta tese puderam ser publicados como artigos. Sem seu trabalho duro e idéias inteligentes, a metade desta tese que trata de *microarrays* jamais poderia ter sido feita (mas não é só por isso que eu me casei com ela!).

Obrigado a todos os colegas que ajudaram indiretamente dando sugestões, indicando bibliografia, lendo e re-lendo o manuscrito da tese, dos artigos, etc: Maria Angela Barreiros, Maria da Graça M. Martin e toda a equipe do LIM-44 do Hospital das Clínicas da Faculdade de Medicina – USP, colegas do programa de doutorado em Bioinformática, a Patricia secretária do Bioinfo, a Patricia Severino, e ao Paulo A. Zaini (desculpe se esqueci de mencionar alguém). Não poderia deixar de mencionar em separado os manos **djogo**, **tvenancio** e **lvaruzza** pela ajuda e pelas discussões sobre diversos assuntos, em particular sobre Bioinformática.

E finalmente, um agradecimento à minha família, em especial aos meus pais Fausto Vêncio e Elizeth Zorzetto, por terem financiado, apoiado, atrapalhado, vibrado, incentivado e contribuído para minha formação educacional e formação como pessoa.

Resumo

O objetivo deste trabalho é apresentar os métodos originais em Bioinformática desenvolvidos para a análise estatística na interpretação dos dados de duas técnicas baseadas em imagens: a técnica de microarranjos de DNA e a técnica de ressonância magnética funcional. O interesse principal é abordar essas técnicas experimentais quando enfrenta-se uma situação clara de amostras escassas, isto é, quando existem relativamente poucas observações experimentais do fenômeno estudado, sendo a análise individual/personalizada o representante extremo desta situação, que tem que ser resolvida. Para tanto, opta-se pelo uso da Inferência Bayesiana no contexto da Teoria da Decisão sob Incerteza, implementada computacionalmente sob o arcabouço dos Sistemas de Suporte à Decisão. Ambas as tecnologias estudadas produzem dados complexos, baseados na interpretação das diferenças entre imagens obtidas da resposta do sistema a um estímulo e da resposta numa situação controle. O resultado deste trabalho é o desenvolvimento de dois sistemas de suporte à decisão, chamados *HTself* e *Dotslashen*, para a análise de dados de microarranjos e ressonância magnética funcional, respectivamente; e de seus métodos matemáticos/computacionais subjacentes. Os sistemas desenvolvidos extraem conhecimento racional de bancos-de-dados normativos, através de modelos matemáticos específicos, contornando então o problema de amostras escassas. Finalmente, neste trabalho são descritas aplicações a problemas reais, para destacar a utilidade dos sistemas de suporte à decisão desenvolvidos nas áreas de Biologia Molecular e Neuroimagem Funcional.

Abstract

The goal of this work is to present the novel Bioinformatics methods that were developed aiming the statistical analysis of two image-based techniques: DNA microarrays and functional magnetic resonance imaging. The main interest is to approach these experimental techniques in small sample size situations, i.e., when there are relatively few experimental observations of the phenomena of interest, for which the case of single subject/datum analysis is its most extreme. In order to approach these problems we chose to use Bayesian Inference in the context of the Decision Theory under Uncertainty, computationally implemented under the Decision Support Systems framework. Both technologies produce complex data, based on the interpretation of differences between images from the response to a given stimulus and the control situation. The result of this work is the development of two decision support systems, called *HTself* and *Dotslashen*, to analyze microarray and functional magnetic resonance imaging data, respectively; and the underlying mathematical and computational methods. These systems use the rational knowledge from normative databases implemented in specific mathematical models, overcoming the problem of small sample size. Finally, in this work it is described applications to real problems in order to stress the utility for Molecular Biology and Functional Neuroimaging of the developed decision support systems.

Sumário

	Capítulo	pág.
1	Introdução	1
2	Sistemas de Suporte à Decisão	4
2.1	Definições e Classificações	4
3	Microarranjos de DNA	8
3.1	Introdução à Expressão Gênica	8
3.2	Obtenção das Imagens	10
3.3	Pré-processamento das Imagens	13
3.4	Sistema de Suporte à Decisão <i>HTself</i>	19
3.5	Aplicações do Sistema <i>HTself</i>	27
4	Ressonância Magnética Funcional	36
4.1	Introdução à Ativação Neuronal	36
4.2	Obtenção das Imagens	40
4.3	Pré-processamento das Imagens	44
4.4	Sistema de Suporte à Decisão <i>Dotslashen</i>	54
4.5	Aplicações do Sistema <i>Dotslashen</i>	63
5	Conclusões e Perspectivas	71
	Referências Bibliográficas	73