



UNIVERSIDADE FEDERAL DE PERNAMBUCO

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

SISTEMAS DE APOIO À DECISÃO NO DIAGNÓSTICO MÉDICO DA HIPERTENSÃO ARTERIAL E DAS ARRITMIAS CARDÍACAS

por

Domingos Vanderlei Filho

Tese submetida ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do grau de Doutor em Engenharia Elétrica

Orientador:

Professor. Fernando Menezes Campello de Souza, PhD

Departamento de Eletrônica e Sistemas

Recife, setembro / 2005

Vanderlei Filho, Domingos

Sistemas de apoio à decisão no diagnóstico médico da hipertensão arterial e das arritmias cardíacas / Domingos Vanderlei Filho. – Recife : O Autor, 2005.

xiv, 174 folhas : il., fig., tab.

Tese (doutorado) – Universidade Federal de Pernambuco. CTG. Engenharia Elétrica, 2005.

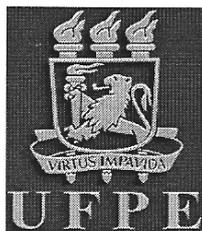
Inclui bibliografia, apêndices e anexo.

1. Engenharia elétrica – Sistemas. 2. Modelos matemáticos – Redes neurais artificiais e estatística. 3. Hipertensão arterial e arritmias cardíacas – Diagnóstico médico – Apoio à decisão. I. Título.

**621.3 : 616
621.3**

**CDU (2.ed.)
CDD (22.ed.)**

**UFPE
BC2005-587**



Universidade Federal de Pernambuco

Pós-Graduação em Engenharia Elétrica

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE TESE DE DOUTORADO

DOMINGOS VANDERLEI FILHO

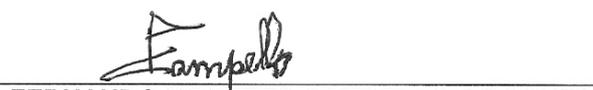
TÍTULO

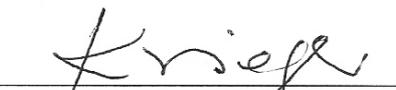
“SISTEMAS DE APOIO À DECISÃO NO DIAGNÓSTICO MÉDICO DA HIPERTROFIA VENTRICULAR ESQUERDA E DAS ARRITMIAS CARDÍACAS”

A comissão examinadora composta pelos professores: FERNANDO MENEZES CAMPELLO DE SOUZA, DES/UFPE, FERNANDA MARIA RIBEIRO DE ALENCAR, DES/UFPE, EDUARDO MOACYR KRIEGER, INCOR/USP, HILTON DE CASTRO CHAVES JÚNIOR, DMC/UFPE e MÊUSER JORGE SILVA VALENÇA, UNIVERSO, sob a presidência do primeiro, consideram o candidato **DOMINGOS VANDERLEI FILHO APROVADO.**

Recife, 30 de setembro de 2005.

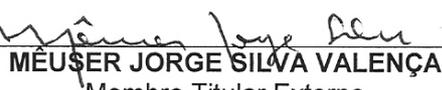

JOAQUIM FERREIRA MARTINS FILHO
Coordenador do PPGE


FERNANDO MENEZES CAMPELLO DE SOUZA
Orientador e Membro Titular Interno


EDUARDO MOACYR KRIEGER
Membro Titular Externo


FERNANDA MARIA RIBEIRO DE ALENCAR
Membro Titular Interno


HILTON DE CASTRO CHAVES JÚNIOR
Membro Titular Externo


MÊUSER JORGE SILVA VALENÇA
Membro Titular Externo

Dedico este trabalho a: Domingos e Mere, meus Pais

Telma, minha Esposa

Milena, minha Filha

AGRADECIMENTOS

A DEUS inteligência suprema causa primária de todas as coisa, por esta existência e pela oportunidade de estar aprendendo um pouco a respeito da grandiosidade do mundo que nos cerca.

Ao meu Orientador o Professor Fernando Campello Menezes de Souza pela sua competência, atenção e paciência em que nos soube conduzir de forma competente nos vários desafios durante o curso.

A Francisco Lyra pela sua capacidade de gerenciar pessoas associada a sua visão humana, que nos apoiou nos momentos decisivos e assim possibilitou compatibilizar tarefas profissionais com a conclusão deste trabalho.

À minha família e pelo apoio e por tolerarem a minha falta de atenção durante minhas horas de estudo e pesquisas.

À Coordenação e Secretaria do PPGEE, em especial a Andréa Tenório pelo apoio e atenção dispensada, e resolução dos problemas de controle administrativo e acadêmico.

A todos os professores da UFPE, que tiveram alguma participação direta ou indireta neste trabalho, pela dedicação e entusiasmo no cumprimento da árdua tarefa de fazer a ciência evoluir.

Aos meus colegas de pós-graduação pelo espírito de equipe, momentos de estudo e descontração, em especial os colegas do grupo de engenharia de sistemas.

Aos colegas Milde Lira, André Marques e André Leite pelas trocas de informações e pelo apoio recebido nos momentos que se fizeram necessário.

Aos amigos do NEIL e todas as pessoas que de alguma forma contribuíram para o desenvolvimento e conclusão desta tese, meu muito obrigado.

Resumo da Tese apresentada à UFPE como parte dos requisitos necessários para a obtenção do grau de Doutor em Engenharia Elétrica.

SISTEMAS DE APOIO À DECISÃO NO DIAGNÓSTICO MÉDICO DA HIPERTROFIA VENTRICULAR ESQUERDA E DAS ARRITMIAS CARDÍACAS

Domingos Vanderlei Filho

Setembro/2005

Orientador: Fernando Menezes Campello de Souza, Ph.D.

Área de Concentração: Processamento de Energia.

Palavras Chaves: Apóio a Decisão, Redes Neurais Artificiais, Engenharia de Sistemas, Classificação, Cardiologia, Hipertrofia Ventricular Esquerda.

Número de Páginas: 174

RESUMO: Neste trabalho é desenvolvida uma nova ferramenta matemática para análise de dados baseada em redes neurais artificiais. Mais especificamente buscou-se extrair informações relevantes sobre o sistema cardio-circulatório de uma série de indivíduos para ajudar e avançar no diagnóstico, prognóstico e tratamento de doenças cardiovasculares. Este sistema foi escolhido em função de sua complexidade que enseja a introdução de novas abordagens e técnicas para entender, explicar e analisar de forma mais eficiente e eficaz o seu funcionamento. O conjunto universo dos dados utilizados nesta pesquisa incluem: antropométricos, biográficos, pressões de consultório e exames complementares da série bioquímica, hormônios, do eletrocardiograma, do ecocardiograma, registro da atividade elétrica do coração pelo método HOLTER-24h e da MAPA-24h, além de parâmetros do sistema cardiovascular como RC, IPPA, IPPARC e Harmonia. O desenvolvimento da pesquisa abrangeu a prospecção de ferramenta para análise de dados com redes neurais tipo MLP e Construtivas do tipo NSRBN. Para fins de avaliação, os resultados obtidos por meio da ferramenta desenvolvida neste trabalho, foram confrontados com desempenhos alcançados com outras técnicas estatísticas tradicionalmente utilizadas, como a regressão logística e múltipla, por meio de medidas de sensibilidade, especificidade, taxa de acertos, coeficiente de determinação e área da curva ROC. Os resultados encontrados superam os das técnicas tradicionais ou utilizam variáveis de mais fácil obtenção ou menor custo financeiro, isto permitirá reduzir ou justificar melhor a prescrição de exames e procedimentos de custo elevado.

Abstract of Thesis presented to UFPE as a partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering.

MEDICAL DECISION SUPPORT SYSTEMS TO ASSIST LEFT VENTRICULAR HYPERTROPHY AND CARDIAC ARRHYTHMIAS DIAGNOSIS

Domingos Vanderlei Filho

September/2005

Supervisor: Fernando Menezes Campello de Souza, Ph.D.

Concentration Area: Electric Power.

Keywords: Decision Support, Artificial Neural Networks, Systems Engineering, Classification, Cardiology, Left Ventricular Hypertrophy, Arrhythmias.

Number of Pages: 174

ABSTRACT: In this work, a new mathematical tool based on artificial neural networks and aimed to data analysis is developed. More specifically it was tried to obtain more relevant information about the cardiocirculatory system of a group of individuals to assist and improve the diagnosis, prognosis and a cardiovascular diseases medical treatment. The cardiovascular system has been chosen because of its complexity, which gives the opportunity to introduce new approaches and techniques to understand, explain and analyze efficaciously and efficiently how it works. The whole data set used in this research includes: anthropometric, biographical, blood pressure measures, the biochemistry series, hormones, results of the cardiologic tests like ABPM-24h, HOLTER-24h, electrocardiography, echocardiogram and cardiovascular system parameters like RC, BPPI, BPPIRC and Harmony. The development of this research included the investigation of a data analysis tool using artificial neural networks types: MLP and constructive NSRBN. The results obtained using the artificial neural networks were evaluated and compared with a classical approach used in medical area, logistic and multiple regressions, in terms of sensitivity, specificity, correct classification rate, coefficient of determination and the area under ROC curve. The results obtained overcome those from traditional techniques or the variables they use are low cost or easily obtained, therefore the system proposed has the potential to reduce or to justify the use of costly invasive procedures in the cardiology area.

SUMÁRIO

LISTA DE FIGURAS	IX
LISTA DE TABELAS.....	XI
LISTA DE SIGLAS E ABREVIATURAS	XIII
1. INTRODUÇÃO	1
1.1 OBJETIVOS.....	5
1.2 SOBRE O TRABALHO.....	5
1.3 ORGANIZAÇÃO DO TRABALHO.....	7
2. O SISTEMA CARDIOVASCULAR.....	9
2.1 INTRODUÇÃO	9
2.1.1 Sistema Vascular	10
2.1.2 O Sangue	11
2.1.3 O Coração.....	14
2.1.4 Funções do Sistema Cardiovascular	19
2.2 DISFUNÇÕES	22
2.2.1 Hipertensão	22
2.2.2 Hipertrofia Ventricular Esquerda	25
2.2.3 Arritmia	25
2.2.4 Infarto do Miocárdio	27
2.2.5 Insuficiência Cardíaca	28
2.2.6 Diabetes.....	29
2.3 ANAMNESE, EXAME FÍSICO, EXAMES COMPLEMENTARES E TECNOLOGIAS PARA O DIAGNÓSTICO.....	30
2.3.1 Avaliação Clínica e Laboratorial.....	30
2.3.2 Equipamentos e Exames para a Medida da Pressão Arterial	33
2.3.3 Equipamentos e Exames para Medida de Outras Variáveis.....	39
2.3.4 Mecanismos Probabilísticos e Erros de Medidas.....	42
3. SISTEMAS NÃO LINEARES.....	44
3.1 SISTEMAS LINEARES.....	44
3.2 SISTEMAS NÃO LINEARES HOMOGÊNEOS.....	47
3.3 SISTEMAS POLINOMIAIS E SISTEMAS DE VOLTERRA.....	57
3.4 ASPECTOS HEURÍSTICOS E MATEMÁTICOS	58
3.5 AS TÉCNICAS INDUTIVAS E AS REDES NEURAIS ARTIFICIAIS.....	60
3.6 MECANISMOS DE AUTO ORGANIZAÇÃO NAS REDES NEURAIS ARTIFICIAIS	62
4. MODELOS DE REDES NEURAIS ARTIFICIAIS E DE REGRESSÃO	64
4.1 MODELAGEM DE SISTEMAS COMPLEXOS.....	64
4.1.1 Classificação de Padrões	66
4.1.2 Aprendizagem e as Teorias de Piaget e Behaviorista.....	70
4.1.3 O Neurônio Biológico.....	73
4.2 MODELOS DE REDES NEURAIS ARTIFICIAIS	75
4.2.1 O Neurônio Artificial.....	78
4.2.2 Perceptron Multicamadas	80
4.2.3 Limitações dos Modelos Tradicionais	86
4.3 MODELOS DE REDES NEURAIS CONSTRUTIVAS	88
4.3.1 Características Gerais das Redes Neurais Construtivas.....	88
4.3.2 Redes Neurais Compostas por Blocos de Regressões Sigmóides Não-lineares	90
4.4 MODELOS DE REGRESSÃO LOGÍSTICA E MÚLTIPLA.....	97
4.4.1 Regressão Logística.....	97
4.4.2 Regressão Múltipla.....	98
4.5 SIMULADORES UTILIZADOS	100
4.5.1 Neural Networks Toolbox - MatLab.....	100

4.5.2	<i>Neuro Hidro</i>	101
4.5.3	<i>Statistica</i>	102
5. O APOIO AO DIAGNÓSTICO DA HIPERTROFIA VENTRICULAR ESQUERDA.....		103
5.1	DESCRIÇÃO DA BASE DE DADOS	104
5.2	EXPERIMENTOS PARA CLASSIFICAÇÃO DE HVE	106
5.2.1	<i>Experimentos com Redes Neurais MLP para HVE</i>	109
5.2.2	<i>Experimentos com Redes Neurais NSRBN para HVE</i>	110
5.2.3	<i>Experimentos com Regressão Logística para HVE</i>	114
5.3	EXPERIMENTOS PARA CLASSIFICAÇÃO DE IMVE	116
5.3.1	<i>Experimentos com Redes Neurais NSRBN para IMVE</i>	117
5.3.2	<i>Experimentos com Regressão Múltipla para IMVE</i>	120
6. O APOIO AO DIAGNÓSTICO DA ARRITMIA CARDÍACA.....		122
6.1	DESCRIÇÃO DA BASE DE DADOS	122
6.2	EXPERIMENTOS COM REDES NSRBN PARA CLASSIFICAÇÃO DE ARRITMIAS ...	123
7. UM SISTEMA DE INFORMAÇÕES PARA UMA CLÍNICA DE CARDIOLOGIA		129
8. CONCLUSÕES, COMENTÁRIOS E SUGESTÕES.....		135
BIBLIOGRAFIA.....		138
ANEXO A SISTEMAS DE CONTROLE AUTO-ADAPTATIVOS.....		146
APÊNDICE A TEOREMAS		160
A.1	<i>DEFINIÇÃO 1</i>	160
A.2	<i>DEFINIÇÃO 2</i>	160
A.3	<i>TEOREMA 1</i>	160
A.4	<i>TEOREMA 2</i>	161
APÊNDICE B ALGORITMO REDES NSRBN		162
APÊNDICE C CRITÉRIOS DE SELEÇÃO NO ALGORITMO CONSTRUTIVO PARA REDES NSRBN		163
C.1	<i>CRITÉRIO DE REGULARIDADE (CR)</i>	163
C.2	<i>CRITÉRIO DE MÍNIMA TENDÊNCIA OU CRITÉRIO CONSISTENTE (CMT)</i>	164
C.3	<i>CRITÉRIO DE PREDIÇÃO (CP)</i>	165
C.4	<i>CRITÉRIO COMBINADO (CC)</i>	165
C.5	<i>CRITÉRIO DE BALANÇO DE VARIÁVEIS (CBV)</i>	166
APÊNDICE D ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS		167
LISTA DE PUBLICAÇÕES.....		174

LISTA DE FIGURAS

Figura 1.1 - Mortalidade proporcional por causas selecionadas no Brasil 1980-1999.....	3
Figura 1.2 – Componentes interdisciplinares do trabalho	7
Figura 2.1 - Representação parcial do Sistema Cardiovascular Humano.....	10
Figura 2.2 - Detalhe do sistema vascular.....	11
Figura 2.3 - Componentes do sangue	14
Figura 2.4 - Câmaras, vasos e válvulas do coração	15
Figura 2.5 - Curvas de Pressão, Volume Ventricular, Eletrocardiograma e ocorrência das Bulhas durante ciclo cardíaco.	17
Figura 2.6 - Ciclo cardíaco.	17
Figura 2.7 - Representação simplificada do sistema elétrico do coração.....	18
Figura 2.8 - Forma de onda do Eletrocardiograma.....	19
Figura 2.9 - Representação da circulação pulmonar e sistêmica e funções do SCV.....	21
Figura 2.10 - Medida Indireta da Pressão pelo Método Auscultatório.....	34
Figura 2.11 - Esfigmomanômetros de coluna de mercúrio	35
Figura 2.12 - Esfigmomanômetro aneróide.....	35
Figura 2.13 - Tensiômetros digitais de dedo (E) e punho (D).....	36
Figura 2.14 - Paciente monitorado por equipamento Holter	42
Figura 3.1 - Representação em Bloco de um Sistema Dinâmico.	44
Figura 3.2 - Diagrama em Bloco Sistema Linear Estacionário	47
Figura 3.3 - Diagrama em Bloco do Sistema Não Linear	49
Figura 3.4 - Um sistema estruturado por interconexão.	49
Figura 3.5 - Um sistema estruturado por interconexão mais complicada.	52
Figura 3.6 - Calculador da integral de uma quadrática.....	52
Figura 3.7 - Calculador do quadrado de uma integral.	52
Figura 4.1 - Formas de folhas.....	67
Figura 4.2 - Neurônio Biológico	74
Figura 4.3 - Arquitetura de uma RNA de uma camada; os círculos representam neurônios e as setas as interligações ponderadas.	77
Figura 4.4 - Neurônio de McCulloch e Pitts.....	79
Figura 4.5 - Representação de uma Rede MLP	82
Figura 4.6 - Arquitetura combinatorial.....	91

Figura 4.7 - Arquitetura de um bloco polinomial.....	93
Figura 4.8 - Tela do Simulador de redes MLP	101
Figura 4.9 - Tela do Simulador Neuro Hidro	102
Figura 5.1 - Curva ROC para o Modelo 6 variáveis.....	112
Figura 5.2 - Curva ROC para o Modelo Regressão Logística.....	115
Figura 5.3 - Curvas dos valores esperados vs. obtidos – modelo NSRBN	119
Figura 5.4 - Curvas dos valores esperados vs. obtidos – modelo Regressão Múltipla.	120
Figura 7.1 - Fases de desenvolvimento do sistema de apoio a decisão	132
Figura 7.2 - Sistema de apoio à decisão em uma clínica de cardiológica	134

LISTA DE TABELAS

Tabela 2.1 - Composição Plasmática.....	14
Tabela 2.2 - Classificação Diagnóstica da Hipertensão Arterial em Adultos (maiores de 18 anos).	24
Tabela 2.3 - Classificação da Pressão Sangüínea de acordo com o JNC 7.	24
Tabela 4.1 - Atributos das Folhas.....	68
Tabela 4.2 - Variáveis e atributos de pacientes.	69
Tabela 4.3 - Semelhança entre termos utilizados em RNAs e em Estatística	98
Tabela 5.1 - Resultados para Rede MLP com 5 entradas.....	109
Tabela 5.2 - Resultados para Rede com 2 entradas	110
Tabela 5.3 - Resultados para Rede com 3 entradas	111
Tabela 5.4 - Resultados para Rede com 6 entradas	111
Tabela 5.5 - Resultados para Rede com 2 entradas do Eletrocardiograma	112
Tabela 5.6 - Resultados para Rede com 3 entradas do Eletrocardiograma	113
Tabela 5.7 - Resultados para Rede com 4 entradas do Eletrocardiograma	113
Tabela 5.8 - Desempenho Modelo de Regressão Logística.....	114
Tabela 5.9 - Resultados das Redes Neurais e da Regressão Logística	115
Tabela 5.10 - Resultados para Rede com 2 entradas	117
Tabela 5.11 - Resultados para Rede com 3 entradas	118
Tabela 5.12 - Resultados para Rede com 4 entradas	118
Tabela 5.13 - Resultados para Rede com 6 entradas	119
Tabela 5.14 - Resultados das Redes Neurais NSRBN e da Regressão Múltipla.....	121
Tabela 6.1 - Resultados para Rede – Caso 1	124
Tabela 6.2 - Resultados para Rede – Caso 2	124
Tabela 6.3 - Resultados para Rede – Caso 3	125
Tabela 6.4 - Resultados para Rede – Caso 4	125
Tabela 6.5 - Resultados para Rede – Caso 5	126
Tabela 6.6 - Resultados para Rede – Caso 6	126
Tabela 6.7 - Resultados para Rede – Caso 7	127
Tabela 6.8 - Resultados para Rede – Caso 8	127
Tabela 6.9 - Resultados para Rede – Caso 9	128
Tabela 6.10 – Resultado Geral das Redes NSRBN para Classificação de Arritmias...	128

Tabela D.1 - Estatística Descritiva da Amostra - Caso HVE	167
Tabela D.2 - Estatística Descritiva da Amostra - Caso Arritmias	171

LISTA DE SIGLAS E ABREVIATURAS

ACC – *American College of Cardiology*
ADALINE – *ADaptive LINear Element ou ADaptive Linear NEuron*
AHA – *American Heart Association*
ART – *Adaptative Ressonance Theory*
AV – *Átrio Ventriculares*
AVC – *Acidente Vascular Cerebral*
BHS – *British Hypertension Society*
bpm - *Batimentos por Minuto*
C/Q – *Relação Cintura Quadril*
CCR – *Taxa de Padrões Classificados Corretamente*
CENEPI – *Centro Nacional de Epidemiologia*
ECG – *Eletrocardiograma*
EUA – *Estados Unidos da América*
Evisolbig – *Extrasístoles Ventriculares Isoladas e/ou Bigeminadas/24h*
FCC – *Frequência Cardíaca Medida em Consultório*
FCMEDIA24H – *Média da Frequência Cardíaca nas 24h*
FCMEDIANOITE – *Média da Frequência Cardíaca durante à Noite*
FUNASA – *Fundação Nacional de Saúde*
GMDH – *Group Method of Data Handling*
HA – *Hipertensão Arterial*
HDL – *High Density Lipoprotein*
HMM – *Modelos de Markov Oculto*
HVE – *Hipertrofia Ventricular Esquerda*
IAM – *Infarto Agudo do Miocárdio*
IEEE – *Institute of Electrical and Electronic Engineers*
IMC – *Índice de Massa Corporal*
IMIP - *Instituto Materno Infantil de Pernambuco*
IPPA – *Índice Pulsátil da Pressão Arterial*
IRE – *Institute of Radio Eletronics*
JNC 7 – *The Seventh Report of Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure*

LDL – *Low Density Lipoprotein*
MAPA – Monitoração Ambulatorial da Pressão Arterial
MAPE – Erro Percentual Absoluto Médio
MLP – *Multi-layer Perceptron*
MRPA – Monitoração Residencial da Pressão Arterial
MS – Ministério da Saúde
MSE – Erro Médio Quadrático
NSRBN – Redes Compostas por Blocos de Regressões Sigmóides Não-lineares
NYHA – *New York Heart Association*
OMS – Organização Mundial de Saúde
PAM – Pressão Arterial Média
PDC – Pressão Diastólica Medida em Consultório
PSC – Pressão Sistólica Medida em Consultório
QRS – Duração do Complexo QRS
RNA – Redes Neurais Artificiais
RNAC – Redes Neurais Construtivas
ROC – *Receiver Operating Characteristic*
RRMedioDia – Média de Intervalos RR durante o Dia
SBH – Sociedade Brasileira de Hipertensão
SBPO – Simpósio Brasileiro de Pesquisa Operacional
SCV – Sistema Cardiovascular
SDNNDia – Desvio Padrão dos Intervalos RR normais durante o Dia
SE – Erro Padrão
SOM – *Self-Organizin Maps*
TaquicVentricular – Taquicardia supraventricular/24h
TSH – *Thyroid Stimulating Hormone*
UFPE – Universidade Federal de Pernambuco
USP – Universidade de São Paulo

1. INTRODUÇÃO

“Every noble work is at first impossible.”

Thomas Carlyle (1795-1881), ensaísta e historiador escocês.

“Todo trabalho nobre parece impossível ao início.” (Tradução livre)

“O Homem é do tamanho do seu sonho.”

Fernando Pessoa (1888-1935), poeta português.

Um dos assuntos mais importantes e abrangentes relacionados à área de saúde é o estudo do sistema cardiovascular (SCV), pois ele tem relação com o funcionamento de todo o organismo humano. Portanto, quadros clínicos onde se observa disfunções no SCV podem conduzir a diversas complicações e em casos mais extremos à própria morte do organismo.

As funções de assimilação e eliminação são importantes para a sobrevivência em todos os organismos vivos. Em organismos mais elementares, como os protozoários, observa-se que o mecanismo para a execução destas funções é realizado diretamente por intermédio da superfície do corpo por uma difusão. Entretanto, em organismos mais complexos, como o humano, com milhões de células, torna-se necessário uma forma mais rápida para a troca de substâncias. Nos organismos pluricelulares existe um sistema interno de transporte para prover com os materiais necessários, oxigênio e outros, todas as células de forma rápida e segura. No caso do organismo humano o sangue percorre o corpo por intermédio de um sistema de tubos fechados e intercomunicantes, mantendo-se a corrente por meio da ação de uma bomba que o aspira e o impele a qual é chamada de coração.

Existem relatos feitos por Hipócrates, que viveu por volta de 400 a.C., nos quais ele descreveu condições muito provavelmente provocadas por doenças relacionadas ao aparelho circulatório. Em torno de 1500, Antonio Benivieni (1443-1502) descreveu em seu livro *De abditis nonnullis ac mirandis morborum et sanationum causis* (Das causas ocultas e maravilhosas das enfermidades e de suas curas) pela primeira vez condições

pós-infarto do miocárdio (Tamayo, 2005). O inglês Samuel Johnson fez o seguinte relato em 1783:

“I waked and sat up...when I felt a confusion and indistinctness in my head which lasted, I suppose about half a minute. Soon after I perceived that I had suffered a paralytick stroke, and that my Speech was taken from me.” (Mackay e Mensah, 2004)

“Eu acordei e fiquei de pé...quando eu senti uma confusão e uma um vazio na minha cabeça que perdurou, eu suponho por cerca de meio minuto. Logo após eu percebi que eu havia sofrido um derrame cerebral, e que minha voz tinha sido subtraída de mim.” (tradução livre)

Em 1880 o patologista Carl Weigert interpretou as cicatrizes do miocárdio como conseqüências de oclusões das coronárias.

Dentre as doenças relacionadas com alguma disfunção do SCV, pode-se citar: infarto do miocárdio, isquemia e derrame cerebral, insuficiência cardíaca, aterosclerose arterial, aneurismas da aorta, *angina pectoris*, valvulopatias, hipertensão arterial e arritmias cardíacas.

Segundo dados da Organização Mundial de Saúde (OMS) as doenças coronarianas levam a morte no mundo, a cada ano, mais de 7 milhões de pessoas (3,8 milhões de homens e 3,4 milhões de mulheres) (Mackay e Mensah, 2004). Dentre os 15 milhões que sofrem derrames cerebrais, aproximadamente 5 milhões morrem e outros 5 milhões tornam-se deficientes que os torna incapacitados a desenvolver suas funções anteriores na família e na sociedade (Mackay e Mensah, 2004). Outro dado relevante sobre este assunto é que a maioria destas mortes ocorre em países em desenvolvimento como o Brasil. A hipertensão é apontada como uma das mais importantes causas de mortes prematuras no mundo que podem ser prevenidas e representa cerca de metade de todas as doenças cardiovasculares.

De acordo com o relatório *The Seventh Report of Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure* (JNC 7) a hipertensão afeta aproximadamente 50 milhões de pessoas nos Estados Unidos da América (EUA) e 1 bilhão no mundo. Em seu relatório de fevereiro de 2005 a *Division*

of *Vital Statistics* apresenta dados de 2003 nos quais as doenças do coração e cerebrovasculares estão entre as três maiores causas de mortes nos EUA (Hoyert; Kung e Smith, 2005, p.3-4).

Conforme pode ser observado na Figura 1.1, entre os anos de 1980 e 1999 houve no Brasil um aumento na proporção das doenças do aparelho circulatório e diminuição das infecciosas e parasitárias – de 9% para 6% – e das afecções originadas no período perinatal – de 7% para 5%. O Ministério da Saúde (MS), no anuário estatístico de 2001 informa que as doenças do aparelho circulatório representam a principal causa de óbitos no país (32%) e em todas as regiões, seguidas pelas causas externas (15%), neoplasias (15%) e doenças do aparelho respiratório (11%) (Brasil.Ministério da Saúde, 2001 e 2003). Ao serem estratificados os 263.240 casos de óbitos por doenças do aparelho circulatório registrados no período de 2001, observa-se que 33% foram de doenças cerebrovasculares, 30% de doenças isquêmicas do coração e 37% para as demais doenças do aparelho circulatório (Brasil.Ministério da Saúde, 2003).

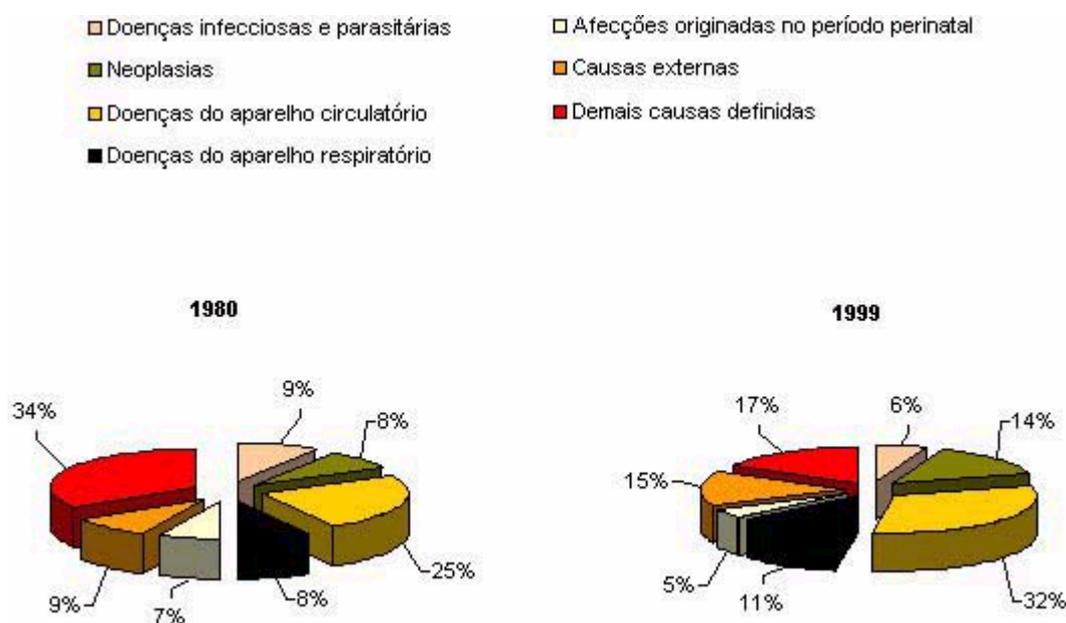


Figura 1.1 - Mortalidade proporcional por causas selecionadas no Brasil 1980-1999

(Fonte: MS/FUNASA/CENEPI)

Adicionalmente, as IV Diretrizes Brasileiras de Hipertensão Arterial trazem a informação de que as doenças cardiovasculares superam as outras causas de morte

desde 1963. É acrescentado ainda que as tendências de risco de morte por doenças cardiovasculares são diferentes nas diversas regiões com queda no Sudeste e Sul, aumento no Centro-Oeste e Nordeste e estabilidade no Norte. Outros trabalhos publicados também mostram e reforçam as conseqüências das doenças relacionadas ao SCV na sociedade brasileira (Mansur *et. al.*, 2001), (De Souza *et al.*, 2001), (Mansur *et. al.*, 2002) e (Lessa, 2003).

Outro aspecto que pode ser mencionado sobre o tema das doenças cardiovasculares refere-se aos custos associados a elas. Entre estes custos podem ser citados: aqueles relativos aos cuidados médicos assumidos pelo indivíduo e sua família, as ausências e licenças médicas no trabalho, os pertinentes ao sistema público de saúde e as perdas de produtividade para o país, sem mencionar o mais importante e difícil de ser quantificado que é o da perda da vida humana.

Nos EUA estimativas da *American Heart Association* (AHA) apontavam para um custo total em 2004 de US\$53,6 bilhões com derrames cerebrais, enquanto custos diretos com cuidados médicos e terapias de US\$33 bilhões e custos indiretos com perda de produtividade de US\$20,6 bilhões (Mackay e Mensah, 2004).

Ao observar os dados apresentados nos parágrafos anteriores os quais provêm de fontes distintas, percebe-se que eles apontam na mesma direção, sendo portanto possível perceber-se as repercussões e alto custo social para a população, bem como o impacto negativo que representam as doenças cardiovasculares na produtividade da economia brasileira. Assim, fica evidente a importância social da área de cardiologia e abre-se a possibilidade de conceber e desenvolver modelos, técnicas e ferramentas de engenharia de sistemas para cooperar na busca de soluções para os problemas de diagnóstico e tratamento de tais doenças.

1.1 OBJETIVOS

O objetivo geral desta tese é desenvolver novas ferramentas de engenharia de sistemas e métodos quantitativos que permitam analisar, a partir de observações, o sistema cardiovascular bem como propor ações que tragam melhores resultados em relação aos disponíveis na atualidade para a sociedade em termos de saúde pública na área de cardiologia.

O conjunto de objetivos específicos compreende:

- Desenvolver sistemática de apoio ao diagnóstico médico de hipertensão e arritmias, mais eficiente e eficaz do que as disponíveis na área de cardiologia;
- Desenvolver uma metodologia para um melhor aproveitamento dos dados do exame MAPA-24h (Monitoração Ambulatorial da Pressão Arterial de 24 horas) extraindo-se informações além daquelas que atualmente são obtidas, para melhorar o diagnóstico, o prognóstico e tratamento da hipertensão arterial;
- Desenvolver uma metodologia para um melhor aproveitamento dos dados do exame Holter 24 horas (monitorização eletrocardiográfica que observa a variabilidade da frequência cardíaca durante um período de 24 horas) extraindo-se informações além daquelas que atualmente são obtidas, para melhorar o diagnóstico, o prognóstico e tratamento das arritmias.

1.2 SOBRE O TRABALHO

Esta tese foi elaborada dentro do paradigma da interdisciplinaridade (ou transdisciplinaridade). Ela insere-se dentro do Projeto Vesalius, que inclui: o Projeto Dinâmica do Sistema Cardiovascular; a Engenharia de operação e manutenção de equipamentos médicos; a Teoria da decisão em medicina; o Apoio inteligente ao diagnóstico; a Análise estatística de dados; o Planejamento de pesquisas; o Acompanhamento de pacientes; a qualidade da imagem médica e os Modelos matemáticos do sistema cardiovascular. Cabe mencionar que esta tese foi desenvolvida em estreito contato com pesquisadores da área médica, em particular com o prof. Dr.

Hilton Chaves, do Departamento de Medicina Clínica do Centro de Ciências da Saúde da Universidade Federal de Pernambuco.

Optou-se pela metodologia das redes neurais artificiais, na forma de uma proposta recente nessa área (Valença, 1999); (Valença e Ludermir, 2000a); (Valença e Ludermir, 2000b); (Valença e Ludermir, 2001); (Valença e Ludermir, 2002). O objetivo desta tese não foi o de desenvolver as técnicas de redes neurais artificiais, e sim elaborar metodologias de apoio ao diagnóstico médico em cardiologia, que fosse de custo mais baixo com relação às práticas vigentes. Foi necessário portanto estudar diversos aspectos do sistema cardiovascular e fazer uso das bases de dados existentes e disponíveis no grupo de pesquisa do orientador.

A escolha do modelo de rede neural artificial para abordar os problemas de apoio ao diagnóstico de hipertrofia ventricular esquerda e de arritmias cardíacas, levou em conta a pequena dimensão da base de dados disponível. Os resultados mostram que a escolha foi acertada, haja vista a comparação favorável com os métodos existentes e resultados de pesquisas anteriores. A escolha de variáveis pertinentes exigiu o estudo de elementos de fisiologia do sistema cardiovascular. A base de dados foi pequena, mas com muitas variáveis.

A redação, haja vista a interdisciplinaridade da pesquisa buscou contemplar tanto leitores da área de engenharia quanto de medicina, de forma a contribuir com o estreitamento cada vez maior entre essas duas áreas aparentemente díspares. Fez-se então uma exposição sobre algumas questões relativas ao sistema cardiovascular, bem como às redes neurais artificiais, procurando-se mediações cognitivas. Esse *cross breeding* tem se mostrado frutífero, e a área de saúde pode se beneficiar com isso. Para a engenharia, o desafio é o desenvolvimento de técnicas de modelagem, a partir de metodologias mais modernas, como é o caso das redes neurais artificiais.

Na Figura 1.2 pode-se observar as principais áreas envolvidas nesta pesquisa bem como a interseção entre elas, representada pela área hachurada, onde esteve concentrada boa parte da pesquisa.

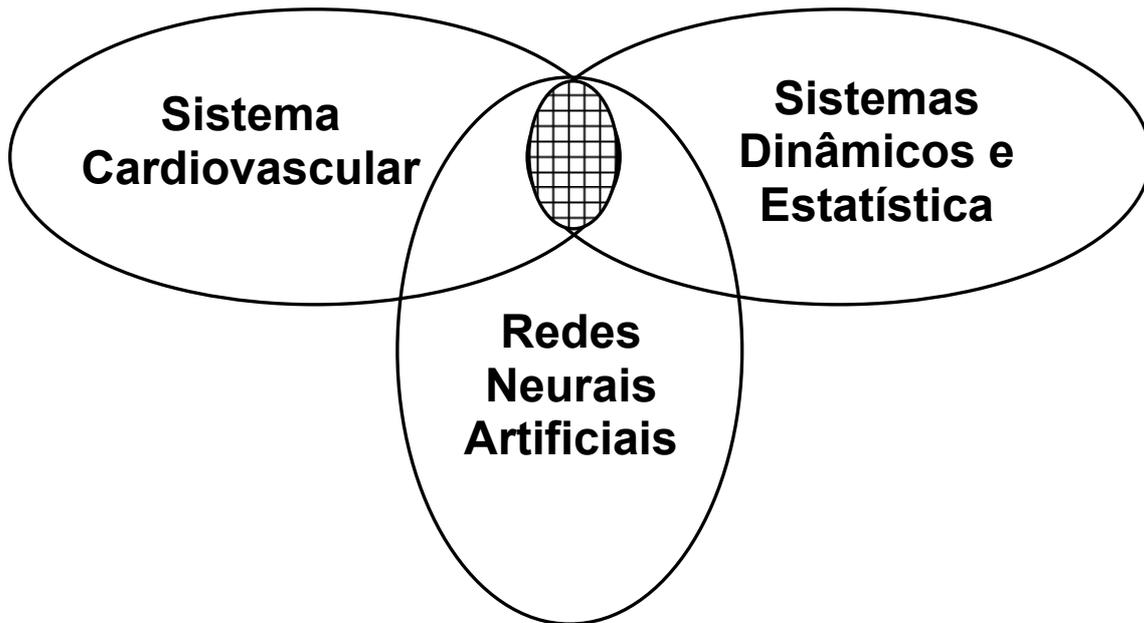


Figura 1.2 – Componentes interdisciplinares do trabalho

1.3 ORGANIZAÇÃO DO TRABALHO

Esta tese está dividida em oito capítulos. No primeiro capítulo apresenta-se uma breve introdução ao tema proposto no trabalho, de forma a destacar a sua importância e abrangência, bem como algumas repercussões sociais e também econômicas, tanto do ponto de vista mundial como do Brasil. O capítulo 1 também mostra os objetivos gerais e específicos e outras informações sobre o trabalho e sua organização.

O segundo capítulo trata das questões relativas ao sistema cardiovascular. No terceiro capítulo são apresentados aspectos dos sistemas não lineares. O quarto capítulo trata da modelagem de sistemas complexos, o problema de classificação e modelos de redes neurais artificiais (RNA) e de regressão. No capítulo quarto ainda são brevemente comentados os simuladores utilizados para os modelos desenvolvidos nesta tese. A base de dados utilizada na tese é descrita no quinto capítulo onde também são apresentados os resultados obtidos com os modelos de apoio ao diagnóstico da hipertrofia ventricular esquerda. No capítulo sexto estão mostrados os resultados obtidos com os modelos de apoio ao diagnóstico de arritmias cardíacas, também baseados em redes neurais artificiais construtivas (RNAC). O capítulo sétimo traz uma proposta para um sistema

de informações para uma clínica de cardiologia. As conclusões, comentários e sugestões para trabalhos futuros são apresentados no capítulo oitavo.

2. O SISTEMA CARDIOVASCULAR

“All the knowledge I possess everyone else can acquire, but my heart is all my own.”

Johann Wolfgang von Goethe, *The Sorrows of Young Werther* 1774.

“Todo o conhecimento que eu tenho qualquer pessoa também pode ter, mas o meu coração é todo somente meu.” (Tradução livre)

“In nature there are neither rewards nor punishments there are consequences.”

Robert Greene Ingersoll (1833-1899)

Lectures and Essays.

“Na natureza não há recompensas nem punições, há conseqüências.” (Tradução livre)

Neste capítulo são mostrados alguns conceitos sobre o sistema cardiovascular, de forma a apoiar o leitor da área de engenharia na compreensão do contexto no qual a engenharia de sistemas foi aplicada. Os tópicos apresentados nas seções deste capítulo seguem basicamente Bronzino (1995), Guyton e Hall (2002), Braunwald; Zipes e Libby (2003v1) e Braunwald; Zipes e Libby (2003v2). Estes tópicos abordam desde a descrição do SCV humano, seus componentes principais e as suas importantes funções para o organismo como um todo. São abordadas também algumas disfunções, principalmente aquelas relacionadas com este trabalho, bem como a anamnese, o exame físico e exames complementares mais comumente utilizados para o diagnóstico médico. O capítulo é finalizado com uma breve exposição sobre os mecanismos probabilísticos e erros de medidas.

2.1 INTRODUÇÃO

O sistema circulatório humano é composto pelo sangue, pelos vasos do sistema vascular (artérias, veias, arteríolas, vênulas e capilares) e pelo coração formado por quatro câmaras (os átrios e os ventrículos direito e esquerdo). No caso humano o sistema circulatório é dito fechado, pois nele o sangue está contido no sistema vascular. Na Figura 2.1 está representado de forma simplificada o sistema cardiovascular humano,

onde está mostrada a circulação pulmonar e a sistêmica, e também o coração, os vasos sanguíneos e o sangue.

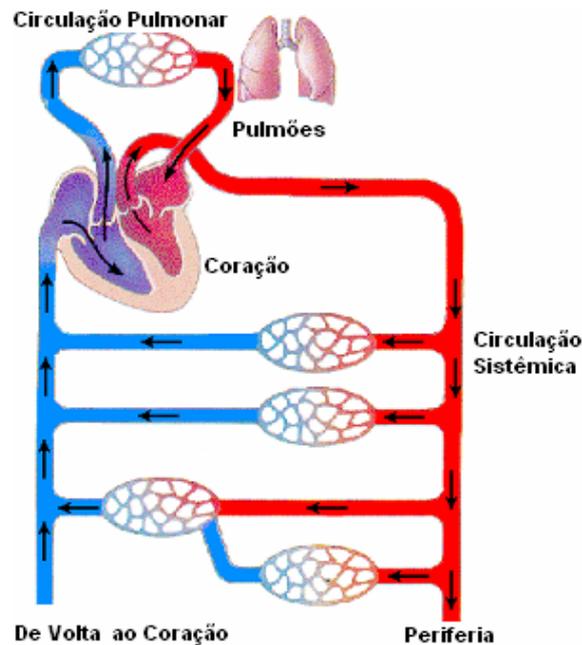


Figura 2.1 - Representação parcial do Sistema Cardiovascular Humano

2.1.1 Sistema Vascular

Uma explicação do papel desempenhado por cada um dos elementos componentes do sistema vascular é fornecida a seguir (Guyton e Hall, 2002):

- Artérias – Elas desempenham a função de transportar sangue sob alta pressão aos tecidos. As paredes das artérias são resistentes de forma que elas podem suportar a alta pressão.
- Arteriolas – Consistem nas últimas e pequenas ramificações do sistema arterial. Elas têm a capacidade de atuar como válvulas de controle e por meio delas o sangue é liberado para os capilares. Esta capacidade é função das paredes musculares fortes que são capazes de fechar as arteriolas totalmente ou e permitir que elas se dilatem.
- Capilares – Estes elementos desempenham a função de meio de troca de líquidos, nutrientes, eletrólitos, hormônios e outras substâncias entre o sangue e os espaços intersticiais. O fato das paredes capilares serem muito delgadas e permeáveis a

substâncias de pequeno tamanho molecular, torna possível para os capilares desempenharem este papel.

- Vênulas – Elas coletam o sangue dos capilares e aglutinam-se gradualmente em veias de calibre progressivamente maiores.
- Veias – Elas funcionam como elementos condutores de sangue que retorna dos tecidos pra o coração. As paredes venosas são delgadas em relação as paredes arteriais, isto em princípio não traz maiores problemas considerando-se que a pressão do sistema venoso é bem mais baixa em relação ao arterial. Apesar deste aspecto mencionado no período anterior estas paredes são musculares e isso permite que elas se contraíam ou se expandam, e portanto atuem como um reservatório de sangue suplementar para quantidades grandes ou pequenas, em função da necessidade do organismo.

A Figura 2.2 mostra o detalhe de uma transição entre a circulação do sangue rico em oxigênio em uma artéria e depois nas arteríolas até chegar aos capilares, onde ocorre a troca gasosa (oxigênio-dióxido de carbono), e na seqüência o sangue pobre em oxigênio é recolhido pelas vênulas quando então é conduzido até uma veia.

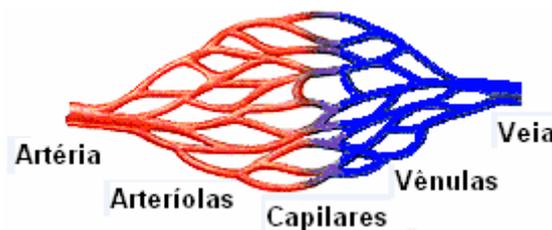


Figura 2.2 - Detalhe do sistema vascular

2.1.2 O Sangue

O sangue pode ser descrito como um tecido composto por uma mistura de células especializadas (eritrócitos, leucócitos) e plaquetas, que juntamente com determinadas substâncias estão em suspensão em um fluido chamado de plasma. O sangue circula no sistema vascular e desempenha as funções de: transportar oxigênio e nutrientes

(proteínas, açúcar, gordura, água e sais minerais) para os tecidos do corpo; remover o dióxido de carbono e outros resíduos metabólicos para serem eliminados por meio da respiração, do suor, da urina e das fezes; transferir mensagens hormonais entre órgãos; prevenir hemorragias e transportar anticorpos e células de defesa para regiões onde haja infecções. O controle da temperatura do corpo, o equilíbrio da distribuição de água e o processo de absorção celular também estão diretamente ligados ao sangue.

A sua composição típica do sangue é de mais de 99% de células chamadas eritrócitos (glóbulos vermelhos ou hemácias), portanto as células denominadas leucócitos que também o compõem não influenciam a determinação das suas características físicas (Guyton e Hall, 2002). O hematócrito é um termo que expressa o percentual de células no volume sanguíneo total de um indivíduo. Estes valores podem variar em função de vários aspectos como: o sexo; o grau de atividade orgânica; altitude na qual reside ou também se o indivíduo apresenta um quadro de anemia.

Os eritrócitos têm a função de transportar hemoglobina, elemento este que participa na condução do oxigênio dos pulmões para os tecidos. Quando a hemoglobina está saturada de oxigênio assume uma coloração avermelhada viva (sangue arterial), quando saturada de gás carbônico, torna-se escuro (sangue venoso). Além do transporte de hemoglobina as hemácias desenvolvem outras funções como a de servir como excelente tampão ácido-básico e conter grande quantidade da enzima anidrase carbônica, para catalizar a reação entre a água e o dióxido de carbono. Isto permite o transporte do dióxido de carbono na forma de íon bicarbonato (HCO_3^-), dos tecidos para os pulmões.

Os leucócitos têm papel vital na defesa do organismo contra agressores externos. Uma outra função destas células é na limpeza de células mortas e fragmentos de tecidos que possam acumular e vir a causar problemas para o organismo. Existem algumas classes de leucócitos como: basófilo, linfócito, monócito, neutrófilo e eosinófilo. Embora não seja verdade para todos os casos, em muitas infecções é observado um aumento no número de leucócitos na corrente sanguínea. Este aumento pode ser detectado por um tipo de teste hematológico de forma a colaborar no diagnóstico médico.

Um terceiro elemento de importância fundamental no sangue são as plaquetas. Sua importância é fundamental no mecanismo da hemóstase e coagulação do sangue. As

plaquetas não são células, mas apenas fragmentos de megacariócitos (células especiais nativas da medula óssea) liberados na circulação. Uma diminuição acentuada no número de plaquetas aumenta a possibilidade de ocorrência de hemorragias espontâneas pela pele ou mucosa.

A parte líquida é chamada de plasma sangüíneo. Cerca de 90% do plasma constituem-se de água pura, na qual estão dissolvidas as numerosas substâncias existentes no sangue. Entre essas substâncias estão sais como: sódio, cloro, fósforo, potássio, magnésio, cálcio e outros. Importância fundamental cabe as proteínas, que também estão dissolvidas no plasma.

Entre as proteínas plasmáticas, encontram-se:

- Albumina - que representa a maior parte e tem como função primária participar na produção da pressão osmótica no nível da membrana capilar. Esta pressão osmótica impede que o líquido do plasma extravase dos capilares para os espaços intersticiais;
- Fibrinogênio e a protombina - que participam no processo de coagulação; e
- Globulinas - estas são divididas em três grupos (alfa, beta e gama). As alfas e betas desempenham funções na circulação como a de agir como substrato para a formação de outras substâncias e transportar a própria proteína de uma parte do corpo para outra. O papel de proteger o organismo contra infecções é realizado pelas gamaglobulinas e em menor extensão as beta globulinas, pois os anticorpos são constituídos principalmente dessas globulinas. Isto permite ao organismo resistência às infecções e à toxidade, e portanto imunidade ao organismo.

Adicionalmente, existem gases dissolvidos e substâncias como a uréia, ácido úrico, creatinina, glicose, gorduras e ácidos graxos que também se encontram presentes neste sistema de alimentação e defesa do corpo humano.

Na Figura 2.3 estão mostrados os componentes do sangue e na Tabela 2.1 está apresentada a composição do plasma sangüíneo.

Tabela 2.1 - Composição Plasmática

CLASSE DA SUBSTÂNCIA	NOME
Proteínas especiais	Albuminas, Globulinas (anticorpos), Fibrinogênio, Protombina, Aglutininas
Outras substâncias orgânicas	Enzimas, Anticorpos, Hormônios, Vitaminas
Lipídios	Colesterol, Triglicérides
Glucídios	Glicose
Substâncias nitrogenadas	Uréia, Ácido úrico, Creatinina
Sais inorgânicos	Sódio, Cloro, Potássio, Cálcio, Fosfatos

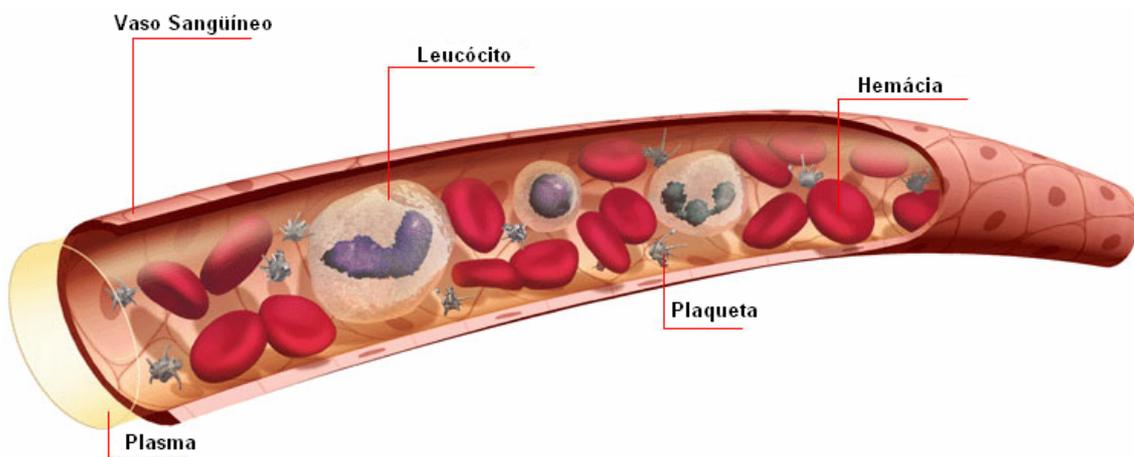


Figura 2.3 - Componentes do sangue

2.1.3 O Coração

O coração pode ser visto tecnicamente, como um conjunto de duas bombas do tipo aspirante e premente as quais são equipadas com válvulas anti-refluxo. Há nestas bombas uma alimentação elétrica automática que atua de forma que o enchimento e a ejeção, realizados nos dois corpos das bombas, ocorram praticamente de forma

simultânea. Para realizar a tarefa de bombear algo em torno de 10 mil litros/dia de sangue esta bomba tem como motor um músculo muito forte, que consome uma grande quantidade de oxigênio e de glicose. O suprimento destas substâncias se faz por intermédio de duas artérias que se distribuem na forma de duas coroas ao redor do coração, chamadas de artérias coronárias.

Na Figura 2.4 podem ser vistas as quatro câmaras, as válvulas e os vasos que se conectam ao coração e permitem a chegada e saída do sangue por ele bombeado.

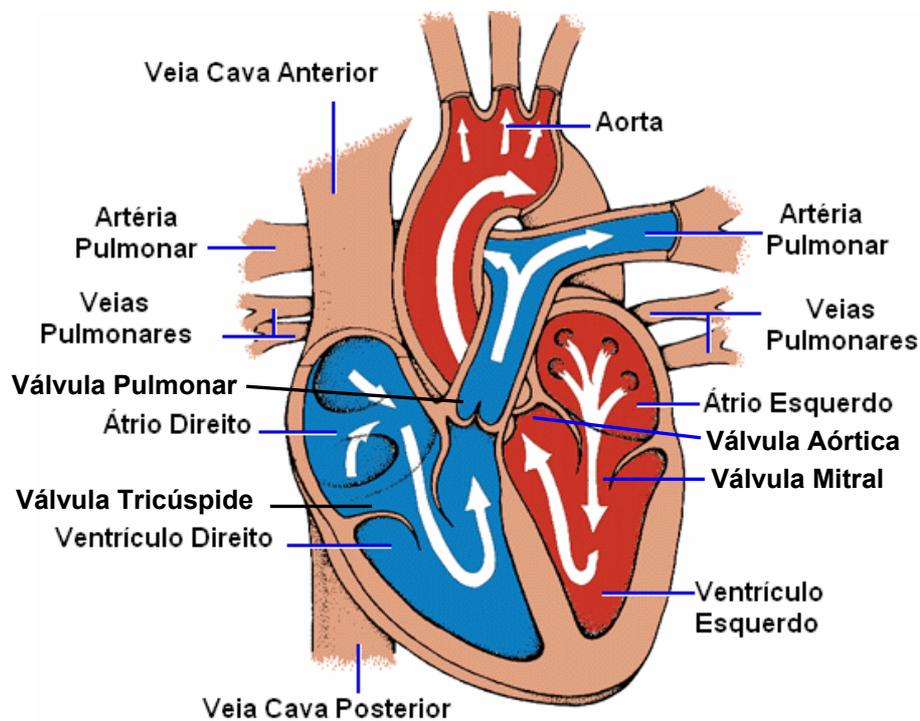


Figura 2.4 - Câmaras, vasos e válvulas do coração

O coração executa uma seqüência cíclica de movimentos de contração e relaxamento. A esta seqüência de eventos é dado o nome de ciclo cardíaco. De uma forma simplificada e resumida pode-se dividir o ciclo em três estágios:

- Sístole atrial – o qual consiste da contração dos átrios e a correspondente entrada de sangue nos ventrículos com a abertura das válvulas átrio-ventriculares (mitral e tricúspide). Após o sangue ter deixado os átrios, as válvulas átrio-ventriculares (AV), que estão localizadas entre as câmaras dos átrios e dos ventrículos, se fecham.

Este movimento evita que ocorra de retorno de sangue para os átrios. Vale ressaltar que é o fechamento das válvulas que gera o som associado ao batimento cardíaco.

- **Sístole Ventricular** – Após a sístole atrial os ventrículos estão com sua capacidade máxima de volume e com sua maior pressão diastólica. As válvulas aórtica e pulmonar encontram-se fechadas, pois as pressões diastólicas ainda são bem maiores que a pressão dos ventrículos. A ativação elétrica chega ao ventrículo e inicia a contração muscular. Neste instante a pressão no interior do ventrículo sobe e as válvulas AV se fecham de forma completa. Surge então a primeira bulha, linha inferior na Figura 2.5. Ao ser atingido um determinado valor de pressão interna que supere a pressão diastólica das grandes artérias as válvulas aórtica e pulmonar se abrem. Esta abertura permite a expulsão do volume de sangue nos ventrículos para as grandes artérias. O pico da curva de pressão ventricular corresponde ao máximo da contração ventricular.
- **Diástole Cardíaca** – Com o fim da contração muscular há o início do relaxamento dos átrios e ventrículos como preparação para a entrada de um novo volume de sangue no coração. A pressão se reduz, parte descendente da curva de pressão ventricular Figura 2.5. Em um determinado instante de tempo o valor da pressão arterial supera a ventricular, o que provoca o fechamento das válvulas, primeiro a aórtica e em seguida a pulmonar. A segunda bulha (vide Figura 2.5) ocorre neste ponto do ciclo cardíaco. Este é o momento que tem início a diástole cardíaca. Quando a pressão ventricular se torna inferior a pressão atrial, conforme indicado na Figura 2.5, as válvulas AV se abrem e um grande fluxo de sangue segue para o ventrículo, aproximadamente 70% do enchimento ventricular ocorre neste intervalo do ciclo. Após o enchimento do ventrículo as válvulas AV tendem a se fechar de forma passiva.

Na Figura 2.6 (a) está representada a diástole atrial e ventricular, em (b) a sístole atrial e em (c) a sístole ventricular.

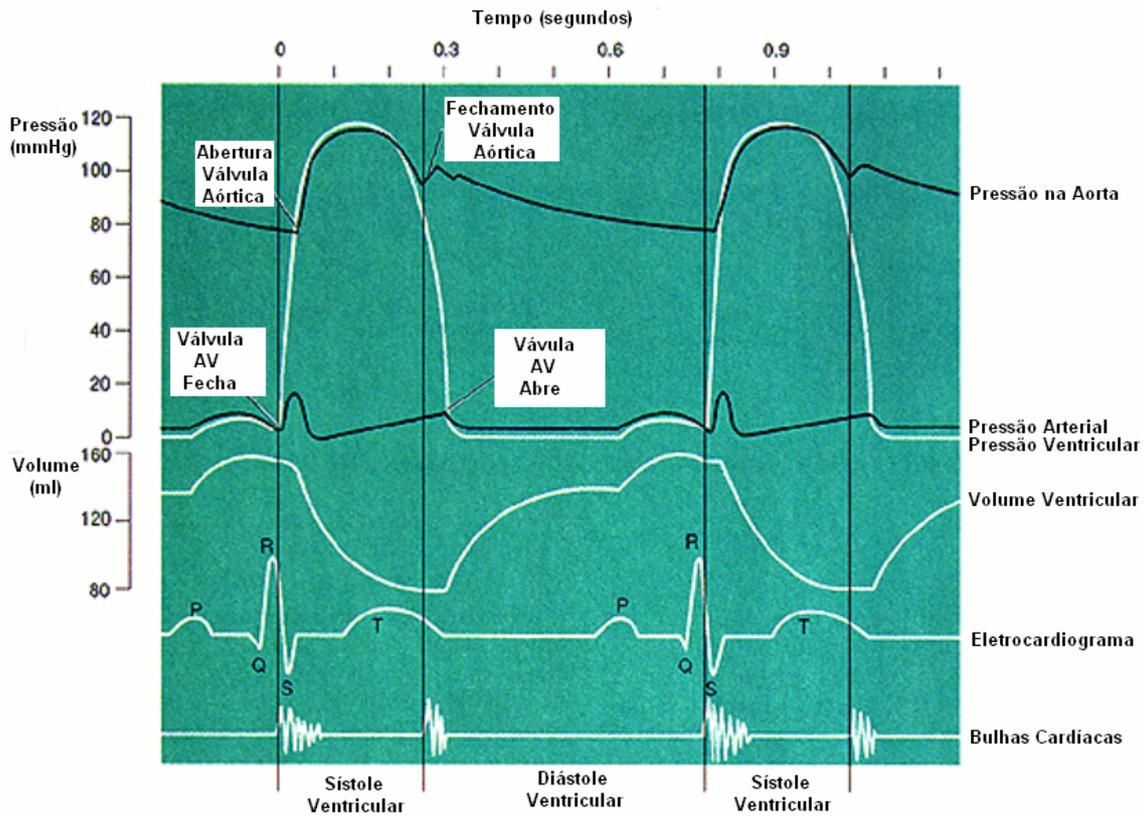


Figura 2.5 - Curvas de Pressão, Volume Ventricular, Eletrocardiograma e ocorrência das Bulhas durante ciclo cardíaco.

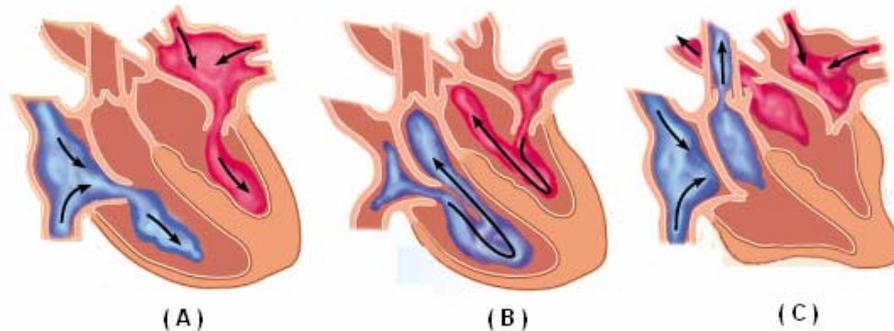


Figura 2.6 - Ciclo cardíaco.

Considerando-se o que foi apresentado nos parágrafos anteriores percebe-se que as contrações que ocorrem em um coração, em funcionamento normal, estão sincronizadas. Primeiro os átrios se contraem juntos e em seguida os ventrículos se contraem juntos. Na população em geral de indivíduos adultos e saudáveis, a frequência cardíaca encontra-se na faixa de 60 a 100 batidas por minuto. Frequências acima de 100 batidas por minuto podem ser observadas como resposta ao estresse físico ou

emocional, como também abaixo de 60 batidas por minuto quando em estado de repouso. Comportamento diferenciado pode ser observado em atletas e pessoas com disfunções.

A contração dos músculos cardíacos que proporcionam o bombeamento de sangue está associada à existência de uma pequena corrente elétrica, a qual é responsável por gerar os estímulos necessários para ativação destes músculos. Esta corrente elétrica é gerada por uma pequena massa de tecido especializado, formada por células marca-passo denominada de nódulo sinusal (também é chamada de nódulo sinoatrial, atriosinusal ou nódulo SA), localizada no coração conforme a Figura 2.7. Este sinal faz com que os átrios ao serem estimulados se contraíam durante um breve período de tempo e, portanto o sangue em seu interior seja bombeado para os ventrículos. Os estímulos elétricos gerados pelo nódulo sinoatrial viajam por meio de fibras de condução (feixes internodais) até o nódulo atrioventricular (AV) onde a passagem do estímulo se faz mais lentamente (retardo fisiológico), permitindo, assim, que haja tempo para se fazer a sístole atrial mecânica. Em seguida o sinal prossegue por intermédio do feixe de Hiss, e finalmente pelas fibras de Purkinje até o miocárdio ventricular.

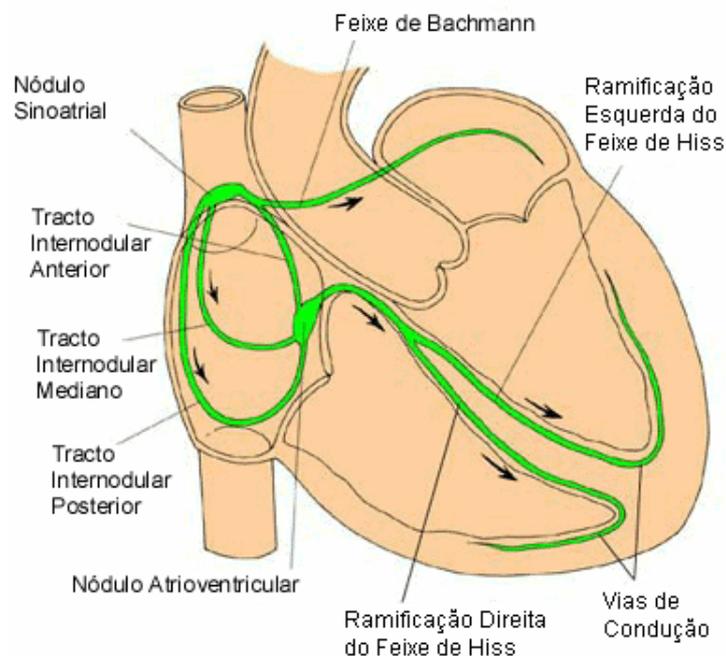


Figura 2.7 - Representação simplificada do sistema elétrico do coração

Em condições normais de funcionamento, a contração dos átrios ocorre um pouco antes que a dos ventrículos, isto permite que o sangue encha os ventrículos antes da sua contração.

Portanto, o sistema especializado de geração de impulsos e condução dos estímulos no coração, em condições normais, garante um ritmo para a contração cardíaca, bem como ativação das diferentes partes do miocárdio de uma forma ordenada e em uma seqüência bem definida, o que garante o funcionamento do coração como uma bomba eficiente.

Na Figura 2.8 pode ser vista a representação gráfica de uma onda típica correspondente ao sinal elétrico do ECG, bem como as diversas denominações associadas a regiões específicas. A correspondência temporal desta onda com as curvas de pressão está mostrada na Figura 2.5.

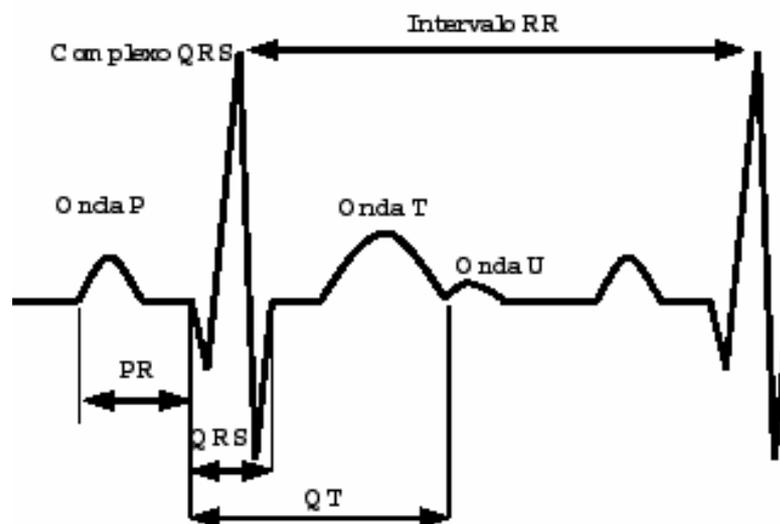


Figura 2.8 - Forma de onda do Eletrocardiograma

2.1.4 Funções do Sistema Cardiovascular

A importância do sistema cardiovascular para a sobrevivência do organismo se manifesta desde o estágio embrionário do ser humano, pois o embrião depende da

circulação do sangue para a manutenção de um ambiente celular favorável e da homeostase. Como resposta a esta necessidade do organismo o SCV se desenvolve e atinge um estado funcional antes que outros importantes sistemas. O coração surge no feto e começa a bater regularmente algo em torno de quatro semanas após a fertilização.

Podem ser citadas as seguintes funções associadas ao sistema cardiovascular: Transporte de Gases (oxigênio e dióxido de carbono) entre os pulmões e os demais tecidos do corpo; Transporte de Nutrientes resultantes da digestão que alcançam o sangue e então são levados aos tecidos do corpo; Transporte de Resíduos metabólicos provenientes das células até os órgãos de excreção; Transporte de Hormônios secretados por certos órgãos que capazes de modificar o funcionamento de outros órgãos do corpo; Intercâmbio de Materiais na forma de algumas substâncias produzidas ou armazenadas em uma parte do corpo e utilizadas em outra parte; Transporte de Calor com sua distribuição homogênea pelas diversas partes do organismo para manutenção de uma temperatura adequada em todas as regiões; Distribuição de Mecanismos de Defesa contra agentes infecciosos na forma de anticorpos e células fagocitárias; Coagulação Sangüínea, pois pelo sangue circulam as plaquetas com função na coagulação sangüínea. O sangue contém ainda fatores de coagulação, capazes de bloquear eventuais vazamentos em caso de rompimento de um vaso sangüíneo.

Considerando-se o que foi apresentado no parágrafo anterior, além da nutrição e purificação dos órgãos e tecidos percebe-se que o sangue também sofre vários tipos de processamentos. Eles ocorrem na forma de oxigenação na circulação pulmonar; filtração na circulação renal; aquecimento na circulação cutânea e provimento de nutrientes na circulação intestinal.

A Figura 2.9 apresenta um diagrama das circulações sistêmica e pulmonar, bem como as funções associadas ao SCV e os demais órgãos envolvidos no desenvolvimento destas funções.

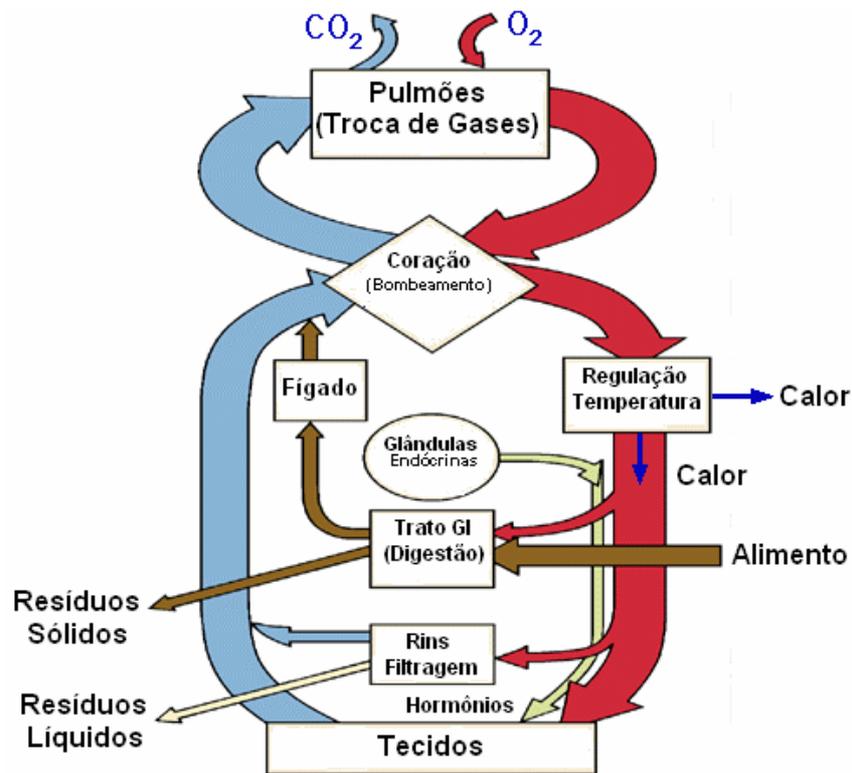


Figura 2.9 - Representação da circulação pulmonar e sistêmica e funções do SCV

As prioridades na distribuição sanguínea seguem regras que refletem uma otimização do sistema com relação a utilização deste recurso escasso que é o sangue. Observa-se que o comportamento natural do SCV é de tal forma que a prioridade ocorre segundo os critérios:

- Importância da região perfundida para a manutenção da vida;
- Importância da região a ser perfundida para ajudar ao organismo escapar de situações que tragam ameaça a vida;
- A capacidade da região perfundida de funcionar ou sobreviver com um déficit no suprimento de sangue.

O sistema de controle cardiovascular se faz na forma de um controle intrínseco, realizado por mecanismos baseados em propriedades físico-químicas dos tecidos e órgãos, ou pelo controle externo, fruto de respostas sobre os tecidos cardiovasculares de órgãos sistêmicos, destacadamente o sistema endócrino e o sistema nervoso autônomo.

Assim, pode-se citar quatro classes de processos de regulação dos elementos coração, vasos e sangue:

- Controle intrínseco, dependente apenas de propriedades do elemento;
- Controle local, resultante de variações físico-químicas do meio interno;
- Controle nervoso, proveniente da função do sistema nervoso central e periférico; e
- Controle humoral, função da atividade das glândulas de secreção interna.

2.2 DISFUNÇÕES

2.2.1 Hipertensão

A força com a qual o sangue é empurrado contra as paredes das artérias é chamada de pressão arterial. Hipertensão Arterial (HA) é o termo médico para uma pressão arterial persistentemente alta, pois é natural que a pressão possa variar como uma forma de resposta do organismo frente às diversas situações a que ele é submetido, como por exemplo, diminua durante o repouso e se eleve nos casos em que ele é exigido fisicamente ou emocionalmente.

Esta disfunção pode se manifestar de maneira silenciosa, ou seja, sem sintomas físicos, e também pode não haver uma causa conhecida (hipertensão primária, idiopática ou essencial) ou estar associada a outras doenças primárias (hipertensão secundária). São exemplos de causas da HA secundária: problemas renais, problemas na artéria aorta, tumores (feocromocitoma) e algumas doenças endocrinológicas.

Geralmente não existem sintomas que sejam relacionados apenas à hipertensão arterial. Alguns sintomas possíveis como dores de cabeça, hemorragia nasal, tonturas ou cansaço excessivo tipicamente não ocorrem até que a pressão sanguínea tenha atingido valores elevados.

A hipertensão é considerada um fator de risco para se desenvolver uma série de problemas de saúde, como por exemplo, romper os vasos sanguíneos cerebrais (acidente vascular cerebral¹, AVC ou derrame cerebral), renais (insuficiência renal) ou de outros órgãos vitais. Pode também provocar diminuição ou perda na visão por lesões na retina e outras conseqüências como sobrecarga excessiva sobre o coração, causando a falência daquele órgão.

Um indivíduo hipertenso tem aumentada a chance de desenvolver: cardiopatia isquêmica (doença das artérias coronárias como angina² ou infarto), claudicação intermitente (dor em membros inferiores ao caminhar, secundária a obstrução de alguma artéria), e insuficiência cardíaca (falha do coração como bomba, levando a falta de ar) do que o indivíduo normotenso.

Segundo a Sociedade Brasileira de Hipertensão: “ ... A hipertensão, ou pressão alta, existe quando a pressão, medida várias vezes em consultório médico, é igual a 14 por 9 ou maior.” (Sociedade Brasileira de Hipertensão, 2001)

“É uma doença crônica, não transmissível, de natureza multifatorial, assintomática (na grande maioria dos casos) que compromete fundamentalmente o equilíbrio dos mecanismos vasodilatadores e vasoconstritores, levando a um aumento da tensão sanguínea nos vasos, capaz de comprometer a irrigação tecidual e provocar danos aos órgãos por eles irrigados.” (SOCIEDADE BRASILEIRA DE HIPERTENSÃO, 2003)

A classificação da hipertensão arterial segundo IV Diretrizes Brasileiras de Hipertensão Arterial está mostrada na Tabela 2.2.

1 Acidente Vascular Cerebral também chamado de AVC ou Derrame Cerebral, corresponde obstrução ou ruptura de uma artéria cerebral que causa uma necrose no cérebro semelhante ao infarto no coração.

2 Angina é o termo médico em latim que significa dor no peito, causada pela pouca irrigação de sangue do músculo do coração. Nome dado à dor que antecede ao infarto. Nem todas as vezes que surgem dores no peito é um prenuncio de infarto. A dor no peito suspeita é aquela que surge depois de esforços físicos; dura poucos minutos e vai embora.

Tabela 2.2 - Classificação Diagnóstica da Hipertensão Arterial em Adultos (maiores de 18 anos).

Pressão Arterial Diastólica (mmHg)	Pressão Arterial Sistólica (mmHg)	Classificação
<80	<120	Ótimo
<85	<130	Normal
85-89	130-139	Limítrofe
90-99	140-159	Hipertensão Leve (Estágio 1)
100-109	160-179	Hipertensão Moderada (Estágio 2)
≥ 110	≥ 180	Hipertensão Grave (Estágio 3)
<90	≥ 140	Hipertensão Sistólica Isolada

Fonte: IV Diretrizes Brasileiras de Hipertensão Arterial, 2002.

As diretrizes para classificar a presença e o grau de hipertensão segundo *the Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure* (JNC VII), estão apresentados na Tabela 2.3.

Tabela 2.3 - Classificação da Pressão Sangüínea de acordo com o JNC 7.

Categoria	PAS (mmHg)		PAD (mmHg)
Normal	<120	e	<80
Pré-Hipertensão	120-139	ou	80-89
Hipertensão, Estágio 1	140-159	ou	90-99
Hipertensão, Estágio 2	≥ 160	ou	≥ 100

Fonte: JNC VII.

Diversos fatores tanto endógenos como exógenos podem influenciar no aparecimento da hipertensão arterial. Como fatores endógenos podem ser citados: a hereditariedade, a idade, a raça do indivíduo além do sobrepeso. No caso dos fatores exógenos tem-se: a ingestão de bebida alcoólica, a falta de exercícios físicos regulares (sedentarismo), o tabagismo, dieta com excesso de sal e o estresse.

2.2.2 Hipertrofia Ventricular Esquerda

Quando ocorre uma carga excessiva de pressão que o músculo cardíaco deve vencer na contração ou por aumento do débito cardíaco que deve ser bombeado, o remodelamento miocárdico ocorre, promovendo um dos mecanismos compensatórios fundamentais que permite o ventrículo sustente essa carga. Contudo, quando o ventrículo é submetido a uma carga anormal e elevada por um longo período ele pode vir a falhar na função de manter a compensação necessária e pode vir a apresentar insuficiência da bomba. Assim, a hipertrofia ventricular esquerda (HVE) pode ser resumidamente explicada como o aumento anormal da massa muscular do ventrículo esquerdo do coração, e tem como consequência uma piora na sua capacidade de bombeamento de sangue. Este tipo alteração do músculo cardíaco está frequentemente associado com hipertensão e é um fator de risco independente para morte súbita cardíaca.

2.2.3 Arritmia

As arritmias podem ser explicadas como problemas que afetam o sistema elétrico do músculo cardíaco e produzem distúrbios no ritmo regular dos batimentos cardíacos. Estes batimentos irregulares podem ocorrer em um coração saudável e gerar consequências mínimas. Contudo, as arritmias causam um bombeamento menos efetivo do coração e também podem causar indicar problemas sérios e levar o indivíduo a óbito.

Conforme descrito na seção 2.1.3 o coração possui quatro câmaras e é excitado por sinais elétricos originários de células especiais que formam marca-passos naturais. Os sinais são conduzidos por fibras (Sistema de His-Purkinje) e por sua vez promovem excitação nas diversas regiões do músculo cardíaco de forma a realizar as contrações necessárias para o bombeamento apropriado do sangue para o organismo. As contrações são de tal forma que os átrios se contraem ligeiramente antes dos ventrículos, o que permite a transferência do conteúdo dos átrios para os ventrículos antes deles se contraírem. Para uma pessoa em repouso os batimentos ocorrem normalmente de 60 a 80 vezes por minuto.

Sob certas condições outras regiões do coração podem dar origem a sinais elétricos comportando-se também como marca-passos. Situações como estas descritas na frase anterior são um dos tipos que podem levar a ocorrência de arritmias. Outras situações são: quando o marca-passo natural produz sinais com ritmo anormal ou quando há interrupção na condução do sinal.

Embora muitas arritmias tenham causa desconhecida, pode-se citar alguns fatores que contribuem com o seu aparecimento, como: doença arterial coronária, hipertensão arterial, diabetes, tabagismo, uso excessivo de álcool, uso de drogas e estresse. Algumas substâncias como: medicamentos com prescrição médica e sem a necessidade de prescrição médica, suplementos alimentares e remédios fitoterápicos podem causar arritmias em algumas pessoas.

A bradicardia ou bradiarritmia é tipificada por um ritmo anormal e baixo dos batimentos cardíacos. Nos adultos considera-se bradicardia quando a frequência cardíaca cai abaixo de 60 batimentos por minuto (bpm). Atletas e pessoas bem condicionadas fisicamente podem apresentar batimentos cardíacos, em repouso, abaixo de 60bpm. Isto pode ser considerado normal caso destas pessoas não apresentem sintomas associados, descritos mais adiante no texto. O critério também é diferente para o caso de recém nascidos, os quais têm um ritmo cardíaco naturalmente mais elevado que os adultos.

A bradicardia usualmente não requer tratamento a não ser que os sintomas como: fadiga, tonturas, sensação de desfalecimento ou desmaios se apresentem de forma repetitiva ou prolongada. Estes sintomas são resultantes de suprimento insuficiente de sangue para o cérebro. Pessoas idosas são mais propensas a ter problemas de batimentos cardíacos mais lentos. Os sintomas podem ser tratados e em alguns casos é utilizado o implante de marca-passos artificiais.

A taquicardia ou taquiarritmia é uma categoria de distúrbio do ritmo cardíaco quando ele se apresenta anormal e rápido e como consequência o bombeio de sangue se torna ineficiente. No caso de indivíduos adultos em repouso o termo é usualmente aplicado para valores de frequência cardíaca acima de 100 bpm. Da mesma forma que as bradicardias as taquicardias assumem várias denominações, geralmente associadas com

a sua causa ou origem, como: atriais, ventriculares, Wolff-Parkinson-White e outras. A taquicardia pode gerar um desconforto no peito ou desmaio, palpitações e em casos extremos à morte súbita. As causas destas arritmias podem ser de origem congênita ou adquirida, neste segundo caso pode-se citar a taquicardia ventricular pós-infarto agudo do miocárdio ou pós-doença de chagas.

A taquicardia ventricular surge quando batimentos rápidos do coração começam nos ventrículos, como já mencionado, isto pode interferir na quantidade de sangue bombeada para o cérebro e outros órgãos vitais. Uma situação mais grave surge se esta forma de arritmia muda para uma forma mais grave chamada de fibrilação ventricular. Na fibrilação ventricular os ventrículos vibram e o coração não bombeia sangue algum. Nesta situação caso não haja providências médicas imediatas o indivíduo é levado à óbito.

Uma forma de tratar este quadro é por meio de cardioversão elétrica, de forma a restabelecer o ritmo normal dos batimentos cardíacos. Outras formas de tratamento das taquicardias são por meio de medicamentos ou identificando-se e destruindo-se os pontos de origem dos ritmos anormais (pontos ectópicos) ou ainda implantando-se dispositivos eletrônicos chamados desfibriladores.

2.2.4 Infarto do Miocárdio

O infarto agudo do miocárdio (IAM), conhecido popularmente como ataque cardíaco, pode apresentar vários sintomas como: vários graus de dor ou desconforto no peito; fraqueza; sudorese, náusea, vômito e muitas vezes com a perda da consciência. Ele ocorre quando uma ou mais partes do músculo cardíaco (miocárdio) são necrosadas, morte do tecido, pela falta ou redução severa no suprimento de sangue oxigenado como resultado de uma obstrução ou ainda estreitamento em um ou mais vasos sanguíneos que irriga o músculo. A falta ou privação de oxigênio para o músculo (isquemia) ou a morte (necrose) de uma porção do músculo ocorre usualmente pelo aparecimento de placas de gordura que se depositam na parede das artérias (aterosclerose).

A aterosclerose é um distúrbio que se caracteriza pela deposição de placas na parede da artéria (túnica íntima). A lesão gordurosa mistura-se aos poucos a diversas células e substâncias existentes no sangue e torna-se cada vez mais e mais densa, elevada e aderente. Nesta fase essas placas recebem o nome de ateromas, e a partir deste momento elas aumentam de forma progressiva salientando-se para a luz do vaso até a obstrução do fluxo sanguíneo ou ocasionando o desprendimento de trombos (pedaços do material aderido a parede do vaso), induzindo a doenças coronarianas.

Este problema pode ocorrer em qualquer artéria, sendo as mais freqüentes as coronarianas e cerebrais. Caso o suprimento não seja restabelecido dentro de um tempo adequado e a depender da extensão do dano causado, isto pode trazer sérias conseqüências para o indivíduo e até mesmo levar o a óbito.

Um estado que pode anteceder o IAM é o aparecimento da *Angina Pectoris*. Este é o termo médico em latim que significa dor no peito, causada pela pouca irrigação de sangue do músculo do coração. Nome dado à dor que antecede ao infarto. Contudo, nem todas as vezes que surgem dores no peito é um prenuncio de infarto. A dor no peito suspeita é aquela que surge depois de esforços físicos; dura poucos minutos e vai embora.

2.2.5 Insuficiência Cardíaca

Pode ser simplifadamente compreendida como a incapacidade do coração em bombear o sangue em quantidade suficiente para as necessidades do organismo, devido a alguma disfunção causada, por exemplo, pela:

- Redução da contratilidade do miocárdio devido a um fluxo sanguíneo coronariano menor;
- Dano nas válvulas cardíacas;
- Pressão externa ao redor do coração;
- Deficiência vitamínica;
- Doença primária do músculo cardíaco;
- Anormalidade que transforme o coração em uma bomba hipofetiva;

2.2.6 Diabetes

A diabetes é uma doença caracterizada pela incapacidade do organismo de transformar o açúcar do sangue, usando a insulina. É também chamado de diabete ou diabetes melitus. Ela pode ser classificada em dois tipos: Tipo 1 e Tipo 2. A diabetes do tipo 1 é uma doença crônica que dura por toda a vida da pessoa. Embora possa ocorrer em qualquer idade ela é usualmente diagnosticada em crianças e adultos jovens. Ela também é chamada de diabetes insulino-dependente, pois as pessoas com esta disfunção necessitam tomar insulina para manter-se vivos.

A insulina é um hormônio produzido no pâncreas e é utilizado pelo organismo para transportar a glicose, proveniente dos alimentos, do sangue para dentro das células. No caso da diabetes tipo 1, por razões desconhecidas o sistema imunológico do organismo atacam as células betas do pâncreas. Após a destruição destas células o pâncreas passa a produzir muito pouca ou nenhuma insulina, o que então faz com que a glicose permaneça no sangue. Este excesso de glicose no sangue por longos períodos de tempo produz danos aos órgãos do corpo. Portanto, pessoas com este tipo de diabetes necessitam tomar doses diárias de insulina, além de adotar hábitos adequados alimentares com dieta balanceada para evitar hipoglicemia e hiperglicemia. São sinais típicos de diabetes: Muita sede; Vontade de urinar diversas vezes; Perda de peso (mesmo sentindo mais fome e comendo mais do que o habitual); Fadiga (cansaço inexplicável) e Visão embaçada. Algumas complicações resultantes da diabetes são: doença do coração, cegueira e complicações renais.

A forma mais comum de diabetes é a do tipo 2 e ela freqüentemente aparece em adultos de meia idade. Este tipo de diabetes é uma doença progressiva que ocorre quando o organismo deixa de produzir insulina em quantidade suficiente e ainda não utiliza de forma eficiente a insulina produzida. Ela está correlacionada com obesidade e pode ser controlado ou retardado o seu aparecimento por meio de exercícios físicos e dieta alimentar adequada. Na sua forma leve pode ficar oculta por muito tempo. Contudo, quando não tratada pode levar ao aparecimento de muitos problemas de saúde,

incluindo-se doenças cardiovasculares. A sua ocorrência é maior em pessoas com mais de 45 anos, com sobrepeso e que tenham histórico familiar da doença.

Indivíduos com diabetes enfrentam um desafio diário de controlar o nível de glicose no sangue de forma a prevenir ou retardar uma série de complicações de saúde. Pessoas com diabetes são mais propensas a desenvolver doenças cardiovasculares.

2.3 ANAMNESE, EXAME FÍSICO, EXAMES COMPLEMENTARES E TECNOLOGIAS PARA O DIAGNÓSTICO

Nas subseções seguintes, 2.3.1, 2.3.2 e 2.3.3, procura-se apresentar alguns tópicos para os leitores que não são da área médica. Estes pontos buscam familiarizar estes leitores com procedimentos e testes comuns para os médicos. O texto traz informações sobre: a anamnese, o exame físico e exames complementares, bem como as tecnologias hoje utilizadas para o diagnóstico médico na área de cardiologia. As tecnologias mostradas incluem tanto aquelas para avaliação da pressão arterial como de outras variáveis. Na subseção 2.3.4 são indicados os mecanismos probabilísticos e erros de medidas.

2.3.1 Avaliação Clínica e Laboratorial

Os diversos métodos e procedimentos hoje disponíveis que fazem uso de modernos equipamentos são um avanço no suporte à área médica. Entretanto, esta gama de testes requer um poder de decisão mais acurado que o necessário na época em que só era disponível o eletrocardiograma e a radiografia de tórax. A boa anamnese e o exame físico cuidadoso podem ser suficientes no estabelecimento do diagnóstico sem que haja a necessidade de realizar procedimentos onerosos, ocasionalmente desconfortáveis e até arriscados. Portanto, a utilização apropriada destes recursos deve servir de suplemento ao exame clínico cuidadoso.

Este exame do paciente deve abranger a história clínica e o exame físico do paciente com suspeita de doença cardiovascular ou ainda com a doença já diagnosticada. A

anamnese clínica pode ser obtida por várias abordagens, um primeiro passo consiste em facultar ao paciente a oportunidade de narrar suas experiências e queixas. Um passo posterior consiste em obter informações sobre o aparecimento e a duração dos sintomas; sua localização; fatores precipitantes, agravantes e de melhora; situações em que ocorrem e quaisquer sintomas associados, seu ritmo de progressão e a resposta à terapêutica.

Uma anamnese completa inclui antecedentes pessoais, ocupacionais, nutricionais e revisão de diversos sintomas. Devem ser investigados outros aspectos como os hábitos relativos a atividade física, tabagismo, etilismo ou uso de drogas, ilícitas ou não, assim como medicações em uso e suas indicações. Outros pontos a serem pesquisados são os fatores de risco para coronariopatias e a natureza do trabalho do paciente, estresses físicos e emocionais.

Segundo Braunwald; Zipes e Libby (2003v1), os principais sintomas das cardiopatias são: Dispnéia (sensação desconfortável da respiração), Desconforto ou dor torácica, Cianose (coloração azulada de pele e membranas resultante de aumento da concentração de hemoglobina reduzida ou de pigmentos de hemoglobina anormal do sangue que perfunde determinada área), Síncope (perda de consciência), Palpitação e Edema.

No exame físico geral é feita uma avaliação do quadro geral do paciente onde podem ser destacados:

- cor da pele, palidez, inspeção da face e de aspectos sugestivos de hipertensão secundária;
- Pescoço: palpação e ausculta das artérias carótidas, verificação da presença de estase venosa e palpação da tireóide;
- Sinais Vitais: Medidas de pressão arterial separadas no tempo e em posição deitada e sentada e medida da frequência cardíaca.
- Medida de peso e altura o que também possibilita o cálculo do índice de massa corporal;
- Exame do Pulmão: dificuldades respiratórias, a frequência respiratória, regularidade, esforço, ausculta dos estertores, roncos e sibilos;

- Exame do abdome: verificar massas abdominais indicativas de rins policísticos, hidroftose, tumores e aneurismas. Identificação de sopros abdominais (aorta e artérias renais);
- Exame do precórdio: icto sugestivo de hipertrofia ou dilatação do ventrículo esquerdo, arritmias e bulhas.
- Extremidades: palpação de pulsos braquiais, radiais, femorais, tibiais posteriores e pediosos. Verificação de edema.
- Exame de fundo de olho: identificar estreitamento arteriolar, cruzamentos arteriovenosos patológicos, hemorragias, exsudatos e papiledema e
- Exame neurológico sumário.

A composição da avaliação laboratorial inclui uma série de exames conforme indicação clínica como: Exame de urina (bioquímica e sedimento); Creatinina; Potássio; Glicemia; Colesterol Total, *High Density Lipoprotein* (HDL), *Low Density Lipoprotein* (LDL); Triglicerídeos; Ácido úrico; Hematócrito e Hemoglobina; Cálcio; *Thyroid Stimulating Hormone* (TSH); Proteinúria de 24 horas; marcadores de necrose miocárdica e outros.

Nas seções a seguir estão apresentados alguns testes aplicáveis na área de cardiologia. Como já mencionado a prescrição de cada tipo de teste deve levar em conta: os fatores de risco associados com cada paciente, o histórico de problemas cardíacos, os sintomas atuais e a interpretação médica de todos estes fatores. A seqüência natural é de se prescrever inicialmente testes mais simples e então parte-se para testes mais complexos caso seja necessário. Os testes podem ser classificados em não invasivo, quando não envolvem a colocação de instrumentos ou fluidos no interior do corpo.

Entre os não invasivos tem-se:

- Eletrocardiograma de Repouso;
- Raio-X do Tórax;
- Holter;
- Ecocardiograma;
- Teste de esforço;
- Tomografia Computadorizada do coração e vasos;
- Ressonância Magnética do coração e vasos;

- MAPA;
- Prova de mesa inclinada ou *Tilt Test*.

Entre os testes invasivos:

- Ecografia transesofágica;
- Cintilografia miocárdica;
- Estudo eletrofisiológico intracardíaco;
- Cateterismo Cardíaco.

2.3.2 Equipamentos e Exames para a Medida da Pressão Arterial

São duas as formas de medida de pressão arterial no ser humano: a direta e a indireta. Na medida direta é fornecida a pressão intra-arterial e é o padrão-ouro. Esta forma de medida além de ser um método invasivo e apresenta inconvenientes como ser: dolorosa, incômoda, de pouca praticidade e exigir equipamentos e técnicas mais sofisticadas. Ela é mais comumente utilizada em situações de pesquisa.

A forma indireta utiliza um manômetro conectado a um manguito e é mais utilizada em função de sua praticidade e comodidade técnica. Quando a medida é realizada por meio do método auscultatório (utiliza os sons de Korotkoff), o uso do estetoscópio se faz necessário. Na Figura 2.10 estão mostrados os equipamentos e suas respectivas posições para a medida indireta da pressão pelo método auscultatório.

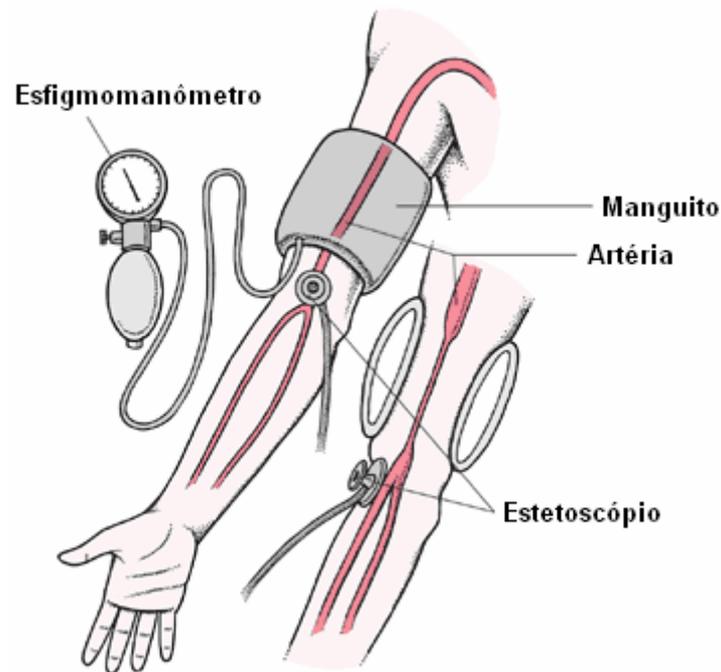


Figura 2.10 - Medida Indireta da Pressão pelo Método Auscultatório

No caso do método oscilométrico a medida é baseada nas oscilações de amplitude do pulso e são determinadas a pressão sistólica e a pressão arterial média, neste caso a pressão diastólica é estimada matematicamente.

O esfigmomanômetro de coluna de mercúrio representa o padrão-ouro para o registro indireto da pressão sanguínea, e portanto os demais equipamentos são aferido a partir dele. Este tipo de equipamento em geral se apresenta boas características de precisão, fácil manutenção e dificilmente perde a calibração. Um dos principais cuidados necessários é o de evitar a perda do mercúrio e quando necessário deve-se acrescentar mercúrio no reservatório até que o menisco esteja na marca do zero. Outros cuidados para a obtenção de leituras corretas é que a coluna de mercúrio esteja na posição vertical e que o tubo onde fica contido o mercúrio esteja livre de sujeira e limpo. Um inconveniente da utilização do esfigmomanômetro de coluna de mercúrio é ele não ser de fácil transporte e requerer cuidados na manipulação, o que talvez o leve a não ser tão adotado na rotina como deveria. Na Figura 2.11 são mostrados exemplos deste equipamento nos modelos de mesa e com rodízio.



Figura 2.11 - Esfigmomanômetros de coluna de mercúrio

O esfigmomanômetro aneróide é o que tem o uso mais difundido. Uma vantagem desse tipo de aparelho está no fato de ele ser mais fácil de transportar e acondicionar. Neste equipamento são utilizadas estruturas metálicas que sofrem expansão com a aplicação de pressão e tem uma agulha indicadora a qual é movimentada pelo amplificador mecânico. A posição de repouso da agulha deve ser o ponto zero. É recomendado que este tipo de equipamento tenha sua calibração verificada. Este teste deve ser realizado contra um manômetro de coluna de mercúrio, o qual por sua vez deve estar devidamente calibrado. Na Figura 2.12 está representado um esfigmomanômetro de aneróide durante uma medição de pressão.



Figura 2.12 - Esfigmomanômetro aneróide

Um comentário sobre os esfigmomanômetros, tanto de coluna de mercúrio como aneróide, é que apesar de simples eles exigem bom treinamento do observador para que os erros de medidas se tornem menos freqüentes.

Um outro tipo de equipamento que tem a utilização cada vez mais freqüente é o eletrônico. Ele dispensa o uso de estetoscópio e fornece medidas de forma automática, em geral por mostrador digital. Estas características que resultam em fácil manipulação, o torna preferido por pessoas leigas porem bem orientada. O treinamento é de fácil assimilação e fica eliminado o erro dependente do observador. Existem algumas variedades destes aparelhos, pois alguns fazem medidas pelo método auscultatório, outros pelo oscilométrico e alguns fazem uso dos dois simultaneamente. Outra variedade encontrada diz respeito a utilização de manguito na posição tradicional (braço), ou no punho ou ainda existem aparelhos para medida da pressão no dedo, conforme mostrado na Figura 2.13.

Vários fatores podem afetar a sensibilidade da medida realizada, como: som ambiente, danos mecânicos ou elétricos e movimentação e posicionamento incorreto do manguito. Hoje já existem aparelhos deste tipo devidamente testados e validados por organismos internacionais.



Figura 2.13 - Tensiômetros digitais de dedo (E) e punho (D).

É sabido que a pressão arterial apresenta variações ao longo do tempo, devido à interação de fatores neuro-humorais, comportamentais e ambientais e que em pacientes hipertensos essa variabilidade da pressão arterial apresenta uma amplitude maior que

nos indivíduos normais. No intervalo de tempo da vigília, os valores da pressão arterial são maiores do que aqueles obtidos durante o sono.

A medida da pressão arterial realizada no consultório médico, apesar de ser considerada um procedimento padrão para o diagnóstico de hipertensão e para o acompanhamento de pacientes hipertensos, está sujeita a inúmeros fatores de erro, sendo que o mais importante é a influência do observador. Adicionalmente, tal medida propicia um número reduzido de leituras, que não apresentam boa reprodutibilidade ao longo do tempo.

Em função dessas variações novos métodos surgiram para melhor avaliação do comportamento da pressão arterial em períodos predeterminados. Dentre eles destacam-se: medidas repetidas casuais da pressão arterial, monitorização ambulatorial (MAPA) e/ou residencial da pressão arterial (MRPA).

A II Diretriz para uso da MRPA a define como: “o registro da pressão arterial por método indireto, pela manhã e à noite, durante cinco dias, realizado pelo paciente ou outra pessoa devidamente treinados, durante a vigília, no domicílio ou no trabalho” (Sociedade Brasileira de Cardiologia, 2005). A MRPA constitui uma alternativa útil para evitar os inconvenientes comentados nos parágrafos anteriores relativos as medidas de consultório. Dentre as vantagens da MRPA em relação às medidas casuais (de consultório) está a melhoria dos índices de adesão ao tratamento, o que levou vários organismos internacionais a sugerirem o uso rotineiro de tal procedimento.

Segundo a II Diretriz para uso da MRPA (Sociedade Brasileira de Cardiologia, 2005) as principais indicações para o uso deste método, são:

- Identificação e seguimento do hipertenso do avental branco;
- Identificação do efeito do avental branco;
- Identificação de hipertensão mascarada;
- Avaliação terapêutica anti-hipertensiva durante a vigília.

Como vantagens potenciais da MRPA em relação à medida casual, podem ser citadas:

- Maior número de medidas;

- Boa aceitabilidade, inclusive por idosos e muito idosos;
- Maior adesão ao tratamento;
- Boa reprodutibilidade;
- Quantificação do efeito do avental branco;
- Avaliação da pressão sem influência do ambiente do consultório;
- Atenuação de erros e preferências do observador;
- Melhor correlação com lesão de órgãos-alvo;
- Possibilidade de armazenamento, impressão e transmissão a distância das leituras obtidas;
- Diminuição do número de visitas ao consultório;
- Baixo custo dos aparelhos.

A monitoração ambulatorial de pressão arterial (MAPA) é um método iniciado nos anos 50 com Maurice Sokolow (Ernst e Bergus, 2002), voltado para o desenvolvimento de pesquisa onde se buscou correlacionar a curva de pressão com o risco cardiovascular. Os primeiros equipamentos eram incômodos e imprecisos, contudo ao longo do tempo a tecnologia avançou consideravelmente. A MAPA é o método que permite o registro indireto e intermitente da pressão arterial por um período de 24 horas (Sociedade Brasileira de Hipertensão, 2002). Neste intervalo de tempo, que inclui o período de vigília e do sono, o indivíduo realiza as suas atividades habituais e rotineiras.

Segundo a IV Diretriz para uso da MAPA (Sociedade Brasileira de Cardiologia, 2005) as principais indicações para o uso deste método, são:

- Suspeita de hipertensão do avental branco;
- Avaliação da eficácia terapêutica anti-hipertensiva: a) quando a pressão arterial causal permanecer elevada apesar da otimização do tratamento anti-hipertensivo para diagnóstico de hipertensão arterial resistente ou efeito do avental branco; b) quando a pressão arterial casual estiver controlada e houver indícios da persistência ou progressão de lesão de órgãos alvo;
- Avaliação de normotensos com lesão de órgãos-alvo;
- Avaliação de sintomas, principalmente hipotensão.

As principais vantagens para o uso da MAPA são:

- Obtenção de múltiplas medidas nas 24 horas;
- Avaliação da pressão arterial durante as atividades cotidianas;
- Avaliação da pressão arterial durante o sono;
- Avaliação do padrão circadiano da pressão arterial;
- Avaliação das médias, cargas e variabilidade da pressão arterial;
- Identificação da reação de “alarme”;
- Atenuação do efeito placebo;
- Avaliação do efeito anti-hipertensivo nas 24 horas;
- Possibilidade de estratificação do risco cardiovascular.

Algumas limitações da MAPA podem ser citadas:

- Braços que não permitam ajuste adequado do manguito;
- Valores muito elevados de pressão arterial sistólica;
- Situações clínicas associadas a distúrbios de movimento (Parkinsonismo, etc);
- Pulsos muito irregulares (fibrilação e “flutter” atriais);
- Hiato auscultatório quando empregado o método auscultatório.

2.3.3 Equipamentos e Exames para Medida de Outras Variáveis

O eletrocardiógrafo é um equipamento que permite monitorar e gravar, de forma não invasiva, o ritmo e a atividade elétrica do coração (vide seção 2.1.3). No exame chamado de eletrocardiograma (ECG) são colocados eletrodos, na pele na região do peito do indivíduo, para captar os sinais elétricos correspondentes a excitação do músculo cardíaco. Os sinais captados são armazenados e também são geralmente impressos em uma fita de papel. Este procedimento permite ajudar na detecção ou caracterização de anormalidades nos batimentos cardíacos, defeitos de condução, hipertrofia, danos localizados no músculo cardíaco, distúrbios nas válvulas e angina entre outros. O teste pode ser realizado com o paciente em repouso ou submetido a determinado esforço físico.

Se o ECG de repouso for normal e a história do paciente for sugestiva de uma doença do coração, o ECG de esforço pode mostrar alterações não reveladas no ECG de

repouso. Serve para avaliar o coração sob condições de estresse, pois o exercício é um estresse fisiológico comumente utilizado para tornar mais evidente anormalidades cardiovasculares ausentes em repouso e também para determinar a adequação da função cardíaca (Braunwald; Zipes e Libby, 2003v1). As indicações mais frequentes são auxiliar no diagnóstico da doença das artérias coronárias, determinar a capacidade funcional e estimar. A *American Heart Association* (AHA) e o *American College of Cardiology* (ACC) determinaram várias categorias de indicações do teste de esforço, baseando-se em um conjunto representativo de materiais publicados sobre o assunto (Gibbons et al., 1997).

O ecocardiógrafo é um equipamento que permite realizar procedimento não invasivo para obter imagens do coração utilizando-se ondas sonoras de alta frequência (ultra-som). Estas imagens possibilitam avaliar o tamanho do coração e de suas câmaras, a espessura das paredes, as válvulas, o miocárdio e vasos sanguíneos localizados próximo ao coração. O exame é realizado colocando-se e movendo-se um tipo de dispositivo manual de emissão e captação de sinais no peito do paciente. Por meio das imagens obtidas torna-se possível avaliar melhor a estrutura e o funcionamento do coração e fornecer mais informações para o diagnóstico médico como no caso de distúrbios nas válvulas cardíacas, cardiomiopatia e infarto do miocárdio.

Nos estudos realizados com Doppler, a velocidade do fluxo sanguíneo pode ser detectada por meio do deslocamento da frequência do ultra-som refletido. Adicionalmente sabe-se que o ultra-som é muito menos danoso para os tecidos que os raios X, pois até o momento não existe evidência clínica de efeitos adversos, apesar do uso da técnica durante mais de duas décadas.

A Ecografia transesofágica é um outro exame que também faz uso da técnica de ultra-som, porém neste caso o teste passa a ser invasivo. Ela é uma ecografia semelhante as não invasivas, mas neste caso o aparelho é colocado dentro do esôfago, o que então permite pela proximidade do coração uma melhor avaliação dos detalhes de algumas válvulas cardíacas. É particularmente útil no diagnóstico de lesões em válvulas causadas pela endocardite.

Um outro conjunto de procedimentos envolve várias técnicas utilizadas nas investigações rádio-nucleares. Estas técnicas utilizam a injeção de materiais radioativos especialmente preparados, que podem ser detectados na corrente sanguínea ou nos tecidos. A radioatividade é detectada por meio de uma câmara nuclear (gama) e é convertida em sinais elétricos por tubos de fotomultiplicação. A principal utilização destas técnicas é fornecer informações sobre a uniformidade da perfusão pulmonar e miocárdica, informações estas ainda não obtidas por outros meios. Neste grupo está a Cintilografia miocárdica na qual a captação de um radioisótopo pelo músculo cardíaco é proporcional à sua perfusão. Pode-se fazer o exame em repouso ou sob esforço físico ou farmacológico. Zonas isquêmicas, menos bem perfundidas serão reveladas.

Na avaliação global da área cardíaca é utilizada a radiografia do tórax. Uma regra geral leva em consideração que a porção mais larga do coração deverá ser menor que a metade do diâmetro do tórax. Embora o ecocardiograma seja mais apurado que a radiografia de tórax na identificação do tamanho cardíaco, uma radiografia pode revelar se há crescimento das cavidades ou de todo o coração. Pode mostrar alterações vasculares ou pulmonares como causa ou repercussão de uma doença cardíaca.

A eletrocardiografia dinâmica de 24 horas, ou simplesmente Holter, em homenagem ao inventor norte-americano Norman Jefferies Holter, é um exame principalmente indicado na investigação de sintomas possivelmente decorrentes de arritmias transitórias, bem como na avaliação de eficácia terapêutica antiarrítmica (Holter e Gengerelli, 1949) (Holter, 1961).

O Holter 24 horas é a monitorização eletrocardiográfica, que observa a variabilidade da frequência cardíaca por um longo período por meio de um método não invasivo. Os monitores Holter atuais são, em geral, equipamentos portáteis que possuem internamente um dispositivo de armazenamento no qual são registrados dados da atividade elétrica do coração do paciente. O eletrocardiograma resultante desta atividade elétrica, captada por intermédio de eletrodos conectados à pele do paciente, é gravado de forma continuada.

Em um período de 24 horas são armazenadas cerca de 100.000 medidas referentes aos batimentos cardíacos. Paralelamente ao registro automático feito pelo equipamento, é

solicitado ao paciente que anote em um diário as atividades realizadas ao longo das 24 horas bem como os eventuais sintomas, para que sejam posteriormente correlacionados à presença ou não de arritmias armazenadas no Holter. Assim, as anormalidades detectadas do ritmo cardíaco e gravadas no dispositivo de armazenamento de dados são analisadas posteriormente em computador por meio de um software dedicado.

A Figura 2.14 mostra um paciente monitorado por um modelo de equipamento Holter.

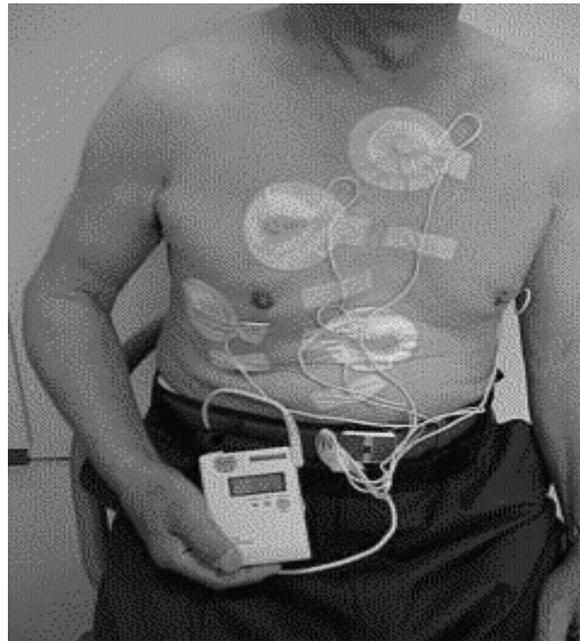


Figura 2.14 - Paciente monitorado por equipamento Holter

2.3.4 Mecanismos Probabilísticos e Erros de Medidas

Um primeiro aspecto a ser tratado são os vários tipos possíveis de erros encontrados, como:

- Erro de escrita, ao fazer a anotação da medida realizada por outra pessoa;
- Erro do instrumento por estar descalibrado ou defeituoso;
- Erro de leitura da medição realizada pelos instrumentos;
- Erro causado pela má utilização da técnica de medida.

Outro mecanismo presente é a preferência por números. É possível perceber que algumas pessoas têm a preferência por determinados números como 125 no lugar de 128. Embora isto não seja um procedimento correto este tipo de situação existe entre os profissionais de saúde, inclusive quando realizam a medida da pressão arterial.

Estar na presença de um profissional da área de saúde, principalmente o médico, pode também causar alterações na medida da pressão arterial. Isto pode se manifestar pelo efeito do avental branco ou da hipertensão mascarada.

Uma outra situação que pode ocorrer é quando um único profissional realiza várias medidas de pressão arterial em um mesmo indivíduo, neste caso pode ocorrer que inconscientemente ele coloque um viés na medição.

Neste capítulo foram tratados aspectos básicos, porém relevantes e necessários para uma melhor compreensão do SCV e também deste trabalho. O próximo capítulo aborda a área de sistemas não lineares a qual está intimamente relacionada com as redes neurais artificiais, que por sua vez são modelos matemáticos. Busca-se aqui tornar clara esta relação, pois estas redes neurais artificiais são fortemente utilizadas neste trabalho no apoio à decisão voltada para o contexto do diagnóstico médico.

3. SISTEMAS NÃO LINEARES

Quando confrontados com problemas de sistemas não lineares, a primeira abordagem que se tenta é a linearização, tipicamente em torno de algum ponto de equilíbrio. Noutras palavras, tenta-se evitar os aspectos não lineares do problema. Se as não linearidades forem fracas ou moderadas, em geral consegue-se algum resultado satisfatório. Quando não, é muito difícil chegar-se a algum resultado útil.

Técnicas não lineares especiais têm sido desenvolvidas para classes de problemas. Entre estas estão a técnica do plano de fase, o segundo método de Liapunov, a linearização harmônica, e outras. Algumas técnicas são oriundas da teoria das equações diferenciais, da análise funcional e da teoria dos operadores. Observa-se a ausência de uma teoria intermediária que ocupe esse espaço em generalidade e aplicabilidade.

3.1 SISTEMAS LINEARES

As explicações dadas a seguir baseiam-se no que está exposto em Rugh (1981). A representação de um sistema dinâmico em termos de uma relação entre a entrada e a saída é dada por:

$$y(t) = \int_{-\infty}^{\infty} h(\tau)u(t - \tau)d\tau, \quad (3.1)$$

onde u é a entrada e y é a saída do sistema, como ilustrado na Figura 3.1.

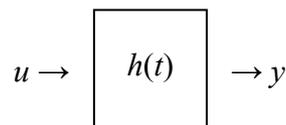


Figura 3.1 - Representação em Bloco de um Sistema Dinâmico.

Supõe-se que a resposta impulsional $h(t)$, tipicamente chamada de *kernel*, satisfaz à condição $h(t)=0$ para $t<0$.

Várias hipóteses técnicas são feitas com relação à expressão 3.1:

- h é uma função real definida para $t = (-\infty, \infty)$, contínua por partes, exceto, possivelmente, em $t = 0$, onde pode ocorrer uma função (generalizada) impulso (a função delta de Dirac);
- u é uma função real definida em $t = (-\infty, \infty)$, tipicamente contínua por partes, mas podendo conter impulsos;
- Desconsiderando a questão dos impulsos, estas condições implicam em que o sinal de saída é uma função real contínua definida em $t = (-\infty, \infty)$.

Hipóteses mais fracas podem ser adotadas, mas essas são suficientes para os propósitos deste capítulo.

É óbvio que o sistema descrito por meio da expressão 3.1 é linear, haja vista as propriedades da integral. Ou melhor, que a descrição do sistema é uma descrição, representação, linear. A hipótese $h(t)=0$ para $t<0$ quanto à h corresponde à causalidade; a propriedade de que a saída do sistema num dado instante não pode depender dos valores futuros da entrada (unilateralidade). Por uma simples inspeção vê-se que a resposta a uma versão retardada de u é a versão retardada de uma resposta a u , e portanto o sistema representado por 3.1 é estacionário. Mais precisamente, se a resposta a $u(t)$ é $y(t)$, então a resposta a $u(t-\tau)$ é $y(t-\tau)$, para qualquer $\tau>0$.

A hipótese $h(t)=0$ para $t<0$ implica que o limite inferior de $-\infty$ na integral na expressão 3.1 pode ser substituído por 0. Considerando-se apenas sinais de entrada que são nulos antes de $t=0$, como é freqüentemente o caso, possibilita que o limite superior (∞) da integral na expressão 3.1 seja substituído por t . A vantagem em manter-se os limites infinitos, manifesta-se na hora de se fazer muitas trocas de variáveis de integração; raramente haverá necessidade de se trocar os limites. Uma das desvantagens é que, apenas no caso dos limites finitos é que se pode proceder impunemente com a troca da ordem de integração em integrais múltiplas, teorema de Fubini vide (Rudin, 1966).

Por uma troca da variável de integração, a expressão 3.1 pode ser escrita como:

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)d\tau \quad (3.2)$$

Nesta forma, a hipótese $h(t)=0$ para $t \leq 0$ implica que o limite superior da integral pode ser diminuído para t , enquanto que a hipótese $u(t)=0$ para $t < 0$ permite que o limite inferior seja aumentado para 0. Para sistemas estacionários, a representação favorita é a expressão 3.1, principalmente porque o *kernel* aparece com um argumento que não é molestado, contrariamente à forma apresentada na expressão 3.2.

No diagrama de blocos da Figura 3.1, o sistema é denotado pelo seu *kernel*. Se o *kernel* for desconhecido, a Figura 3.1 corresponderá à famosa caixa preta linear.

Na sua *Theory of Functionals* (Teoria dos Funcionais), o matemático italiano Vito Volterra (1860–1940) introduziu a primeira noção do que é agora conhecido como uma série de Volterra. A primeira aplicação importante do trabalho de Volterra à análise de circuitos não lineares foi feita pelo matemático americano Norbert Wiener (1894–1964), que as usou de uma maneira geral para analisar vários problemas, incluindo o espectro de um sistema de FM com uma entrada sendo um ruído gaussiano. Desde então as séries de Volterra têm sido largamente usadas no cálculo de termos de distorção, pequenos mas inoportunos, perturbadores, em amplificadores a transistor e outros sistemas (Wiener, 1942), (Volterra, 1959), (Schetzen, 1980).

Se se remove a hipótese de estacionaridade, ter-se-á:

$$y(t) = \int_{-\infty}^{\infty} h(t, \tau)u(\tau)d\tau \quad (3.3)$$

onde h é uma função real definida para $t \in (-\infty, \infty)$, $\tau \in (-\infty, \infty)$ e $h(t, \tau)=0$ se $\tau > t$ (causalidade). Os comentários anteriores sobre a redução dos limites de integração aplicam-se aqui. Apenas a propriedade da invariância aos retardos puros, que corresponde à estacionaridade, foi abandonada. Permite-se, tipicamente, que h contenha impulsos em $\tau-t$, mas fora disso, supõe-se que ela é contínua por partes para $0 \leq \tau \leq t$.

Comparando-se a expressão 3.2 com a 3.3, vê-se que um sistema linear estacionário pode ser considerado como um caso particular de um sistema linear não estacionário. É

conveniente, pois, chamar o *kernel* $h(t, \tau)$ apresentado na expressão 3.3 estacionário se existir um *kernel* $g(t)$ tal que

$$g(t-\tau)=h(t, \tau). \quad (3.4)$$

Uma maneira fácil de verificar a estacionaridade de $h(t, \tau)$ é verificar a condição:

$$h(0, \tau-t)=h(t, \tau).$$

Se esta condição for satisfeita, então, fazendo-se

$$g(t) = h(0, -t)$$

verifica a expressão 3.4, pois

$$g(t-\tau)= h(0, \tau-t)=h(t, \tau).$$

Um sistema linear estacionário (possivelmente) é diagramado usando-se a representação da expressão 3.3 como mostrado na Figura 3.2.

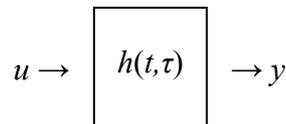


Figura 3.2 - Diagrama em Bloco Sistema Linear Estacionário

3.2 SISTEMAS NÃO LINEARES HOMOGÊNEOS

O enfoque a ser adotado na representação por relações entrada/saída de sistemas não lineares, envolve uma simples generalização das representações discutidas anteriormente na seção 3.1. Não serão discutidas aqui algumas questões não resolvidas, de caráter, num certo sentido, filosófico, quanto à generalidade e utilidade deste enfoque. O objetivo aqui é mostrar a representação, discutir algumas de suas propriedades e dar alguns exemplos, de forma a que se possa apreciar o seu interesse.

Seja h uma função real de n variáveis $t_1, t_2, \dots, t_n, t_i \in (-\infty, \infty), i=1, 2, \dots, n$, com valor $h_n(t_1, t_2, \dots, t_n)$, sendo $h_n(t_1, t_2, \dots, t_n)=0$ se qualquer $t_i < 0$. Define-se, neste caso, a relação entrada saída por:

$$y(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h_n(\tau_1, \dots, \tau_n) u(t - \tau_1) u(t - \tau_2) \dots u(t - \tau_n) d\tau_1 \dots d\tau_n \quad (3.5)$$

É visível a semelhança com as representações de sistemas lineares mostradas anteriormente. Além disso, todas as hipóteses técnicas feitas anteriormente são aplicáveis nesta representação da expressão 3.5, que é algumas vezes chamada de convolução generalizada.

Note-se que a unilateralidade de h com respeito a cada variável corresponde à causalidade. O sistema aqui não é linear, mas pode-se ver facilmente que é estacionário.

Um sistema representado pela expressão 3.5 é chamado de sistema homogêneo de grau n . Esta terminologia deriva do fato de que a aplicação da entrada $\alpha u(t)$, onde α é um escalar, produzirá uma saída $\alpha^n y(t)$, onde $y(t)$ é a resposta a $u(t)$. Observe-se que esta terminologia inclui o caso de um sistema linear como um sistema homogêneo de grau 1. Como anteriormente, $h_n(t_1, t_2, \dots, t_n)$ será chamado o *kernel* do sistema.

Como no caso linear, também, os limites inferiores podem ser substituídos por 0, devido à hipótese de unilateralidade do *kernel*. Se se assume, ademais, que o sinal de entrada é unilateral, então todos os limites superiores podem ser substituídos por t . Finalmente, uma mudança de cada variável de integração mostra que a expressão 3.5 pode ser reescrita como:

$$y(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h_n(t - \tau_1, \dots, t - \tau_n) u(\tau_1) u(\tau_2) \dots u(\tau_n) d\tau_1 \dots d\tau_n \quad (3.6)$$

A representação por diagrama de blocos, Figura 3.3, é semelhante aos casos anteriores:

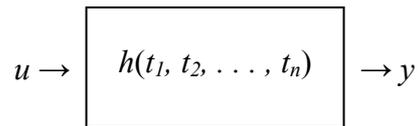


Figura 3.3 - Diagrama em Bloco do Sistema Não Linear

Pode-se tipificar duas situações onde os sistemas homogêneos podem ocorrer em aplicações de engenharia. A primeira envolve sistemas físicos que naturalmente são estruturados em termos de interconexões de subsistemas lineares e não linearidades simples. De interesse particular, pela sua simplicidade e praticidade, com razoável grau de aderência a muitas situações reais, são os casos que envolvem subsistemas lineares estacionários e não linearidades que podem ser representadas em termos de multiplicadores. Para sistemas estruturados por interconexões, como estes, é freqüentemente fácil estabelecer o *kernel* do sistema como um todo, a partir dos *kernels* dos subsistemas, simplesmente rastreando o sinal de entrada ao longo do diagrama de blocos do sistema. Um exemplo de conexão multiplicativa é mostrado na Figura 3.4.

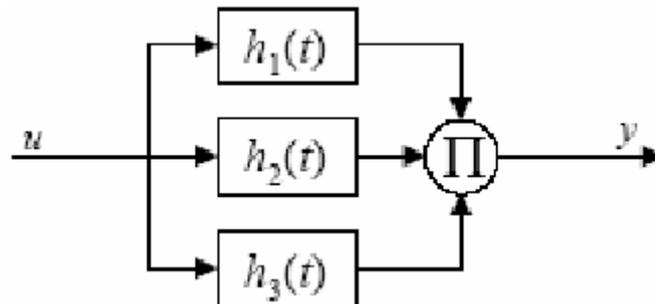


Figura 3.4 - Um sistema estruturado por interconexão.

Trata-se de uma conexão multiplicativa de três subsistemas lineares, cada um deles representado por:

$$y_i(t) = \int_{-\infty}^{\infty} h_i(t)u(t-\tau)d\tau, \quad i = 1, 2, 3,$$

e então o sistema global é descrito por:

$$y(t) = y_1(t)y_2(t)y_3(t)$$

$$y(t) = \int_{-\infty}^{\infty} h_1(t)u(t-\tau)d\tau \int_{-\infty}^{\infty} h_2(t)u(t-\tau)d\tau \int_{-\infty}^{\infty} h_3(t)u(t-\tau)d\tau$$

$$y(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(\tau_1)h_2(\tau_2)h_3(\tau_3)u(t-\tau_1)u(t-\tau_2)u(t-\tau_3)d\tau_1d\tau_2d\tau_3$$

É claro que um *kernel* para este sistema homogêneo de grau 3 é dado por:

$$h(t_1, t_2, t_3) = h_1(t_1)h_2(t_2)h_3(t_3)$$

Um segundo tipo de situação aparece quando se tem uma descrição de um sistema não linear por uma equação de estado. Considere-se, para ilustrar, um modelo de compartimento no qual cada variável $x_i(t)$ representa uma população, uma concentração, ou outra quantidade de interesse. Se a taxa de variação no tempo de $x_i(t)$ depende linearmente das outras $x_j(t)$'s, mas com um sinal de controle paramétrico escalar, então $\dot{x}(t)$ conterà termos da forma $du(t)x_j(t)$. Modelos compartimentais não lineares deste tipo levaram ao estudo das chamadas equações de estado bilineares.

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Dx(t)u(t) + bu(t) \\ y(t) &= cx(t), \quad t \geq 0, \quad x(0) = x_0\end{aligned}$$

Considere-se, por exemplo, um sistema não linear descrito pela equação diferencial:

$$\begin{aligned}\dot{x}(t) &= Dx(t)u(t) + bu(t) \\ y(t) &= cx(t), \quad t \geq 0, \quad x(0) = 0,\end{aligned}$$

onde $x(t)$ é um vetor 2×1 , $u(t)$ e $y(t)$ são escalares e

$$D = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad c = [0 \quad 1]$$

Sabe-se que para um sinal de entrada contínuo por partes, para $t \geq 0$, esta equação terá uma solução única dada por:

$$x(t) = \int_0^t e^{D \int_{\tau_2}^t u(\tau_1) d\tau_1} b u(\tau_2) d\tau_2$$

onde a matriz exponencial é dada por:

$$e^{D \int_{\tau_2}^t u(\tau_1) d\tau_1} = I + D \int_{\tau_2}^t u(\tau_1) d\tau_1 + \frac{1}{2!} \left[D \int_{\tau_2}^t u(\tau_1) d\tau_1 \right]^2 + \dots$$

Para o caso particular em questão, $D^2 = 0$ e então

$$e^{D \int_{\tau_2}^t u(\tau_1) d\tau_1} = \begin{bmatrix} 1 & 0 \\ \int_{\tau_2}^t u(\tau_1) d\tau_1 & 1 \end{bmatrix}$$

A relação entrada/saída pode então ser colocada na forma:

$$\begin{aligned} y(t) &= \int_0^t c e^{D \int_{\tau_2}^t u(\tau_1) d\tau_1} b u(\tau_2) d\tau_2 \\ &= \int_0^t \int_{\tau_2}^t u(\tau_1) u(\tau_2) d\tau_1 d\tau_2 \end{aligned}$$

Vê-se claramente por esta expressão que o sistema é homogêneo e de grau 2. Para colocar a representação da relação entrada/saída numa forma mais familiar, use-se a função degrau unitário:

$$\delta_{-1}(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}$$

Ter-se-á então:

$$y(t) = \int_0^t \int_0^t \delta_1(\tau_1 - \tau_2)$$

Logo, um *kernel* para o sistema é:

$$h(t_1, t_2) = \delta_{-1}(\tau_1 - \tau_2)$$

Haverá ocasiões em que os sistemas homogêneos podem ser não estacionários. Tais sistemas são representados pela relação entrada saída:

$$y(t) = h(t, \tau_1, \dots, \tau_n)u(\tau_1) \cdots u(\tau_n) d\tau_1 \cdots d\tau_n \tag{3.7}$$

Pode-se pensar em muitas configurações de interconexões, como ilustrado nas Figuras 3.5, 3.6 e 3.7.

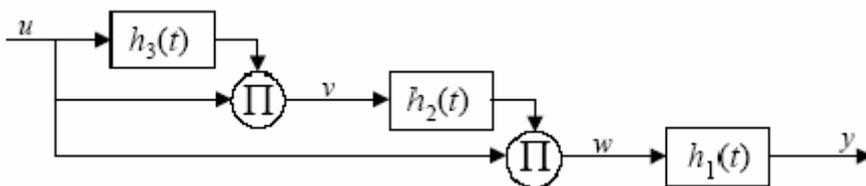


Figura 3.5 - Um sistema estruturado por interconexão mais complicada.

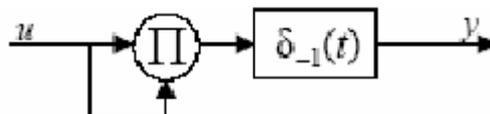


Figura 3.6 - Calculador da integral de uma quadrática.

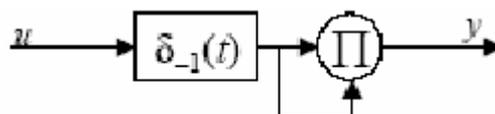


Figura 3.7 - Calculador do quadrado de uma integral.

Um *kernel* $h(t, \tau_1, \dots, \tau_n)$ é estacionário se e somente se existe um *kernel* $g(t_1, t_2, \dots, t_n)$ tal que a relação:

$$g(t - \tau_1, \dots, t - \tau_n) = h(t, \tau_1, \dots, \tau_n) \quad (3.8)$$

seja válida para todo t, τ_1, \dots, τ_n . É conveniente, usualmente, verificar a estacionaridade verificando a relação funcional:

$$h(0, \tau_1 - t, \dots, \tau_n - t) = h(t, \tau_1, \dots, \tau_n) \quad (3.9)$$

pois se esta for satisfeita, então a expressão 3.8 é obtida fazendo-se

$$g(t_1, \dots, t_n) = h(0, -t_1, \dots, -t_n) \quad (3.10)$$

Logo, quando a expressão 3.9 é satisfeita, pode-se escrever:

$$y(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(t - \tau_1, \dots, t - \tau_n) = h(t, \tau_1, \dots, \tau_n) u(\tau_1) u(\tau_2) \dots u(\tau_n) d\tau_1 d\tau_2 \dots d\tau_n \quad (3.11)$$

no lugar da expressão 3.7.

Um *kernel* que descreve um sistema homogêneo de grau n é dito ser separável se puder ser escrito na forma:

$$h(t_1, \dots, t_n) = \sum_{i=1}^m v_{1i}(t_1) v_{2i}(t_2) \dots v_{ni}(t_n) \quad (3.12)$$

ou

$$h(t, \tau_1, \dots, \tau_n) = \sum_{i=1}^m v_{0i}(t) v_{1i}(\tau_1) v_{2i}(\tau_2) \dots v_{ni}(\tau_n) \quad (3.13)$$

onde cada $v_{ji}(\cdot)$ é uma função contínua. Será chamado separável diferenciável se cada $v_{ji}(\cdot)$ for diferenciável. A separabilidade é uma propriedade freqüente dos *kernels*.

Mais de um *kernel* pode ser usado para se descrever um dado sistema. Isto leva a dificuldades quando se deseja descrever as propriedades do sistema em termos das propriedades do *kernel*. É importante, portanto, em muitas situações, impor unicidade, trabalhando-se com formas especiais, restritas, do *kernel*. Dentre elas, três formatos especiais são:

- Simétrico;
- Triangular;
- Regular.

Um *kernel* simétrico no caso estacionário satisfaz à:

$$h_{sym}(t_1, \dots, t_n) = h_{sym}(t_{\pi(1)}, \dots, t_{\pi(n)}) \quad (3.14)$$

ou, no caso estacionário,

$$h_{sym}(t, \tau_1, \dots, \tau_n) = h_{sym}(t, \tau_{\pi(1)}, \dots, \tau_{\pi(n)}) \quad (3.15)$$

onde $\pi(\cdot)$ denota qualquer permutação dos inteiros $1, 2, \dots, n$. É fácil mostrar que, sem perda de generalidade, o *kernel* de um sistema homogêneo pode ser suposto simétrico. De fato, qualquer *kernel* dado, $h(t_1, \dots, t_n)$, pode ser substituído por um *kernel* simétrico fazendo-se

$$h_{sym}(t_1, \dots, t_n) = \frac{1}{n!} \sum_{\pi(\cdot)} h_{sym}(t_{\pi(1)}, \dots, t_{\pi(n)}) \quad (3.16)$$

onde a soma indicada é feita sobre todas as possíveis $n!$ permutações dos inteiros de 1 até n .

Para ver que esta substituição não afeta a relação entrada/saída, considere-se a expressão:

$$\int_{-\infty}^{\infty} h(\tau_1, \dots, \tau_n) u(t - \tau_1) \cdots u(t - \tau_n) d\tau_1 \cdots d\tau_n =$$

$$= \frac{1}{n!} \sum_{\pi(\cdot)} \int_{-\infty}^{\infty} h(\tau_{\pi(1)}, \dots, \tau_{\pi(n)}) u(t - \tau_{\pi(1)}) \cdots u(t - \tau_{\pi(n)}) d\tau_{\pi(1)} \cdots d\tau_{\pi(n)} \quad (3.17)$$

Introduzindo-se a mudança de variáveis (de fato apenas uma re-rotulação) $\zeta_i = \tau_{\pi(i)}$, $i = 1, \dots, n$ em todos os termos da soma na expressão 3.17, mostra que todos os termos são idênticos. Logo, somando-se todos os $n!$ termos idênticos do lado direito mostra que os dois *kernels* produzem o mesmo comportamento entrada/saída.

É freqüente que um *kernel* de interesse seja parcialmente simétrico no sentido de que nem todos os termos da soma na expressão 3.16 são distintos. Numa situação destas, uma versão simétrica do *kernel* pode ser obtida somando-se naquelas permutações que dão somandos distintos e substituindo $n!$ pelo número de tais perturbações. Obtém-se, freqüentemente, uma redução significativa do número de termos.

Em suma, os *kernels* de Volterra são sinônimos de efetiva conectividade.

Um *kernel* $h(t, \tau_1, \tau_2, \dots, \tau_n)$ é triangular se e somente se satisfaz a propriedade adicional que $h(t, \tau_1, \tau_2, \dots, \tau_n) = 0$ quando $\tau_{i+j} > \tau_j$ para inteiros positivos i, j . Para um tal *kernel*, a representação pode ser escrita como:

$$y(t) = \int_{-\infty}^t \int_{-\infty}^{\tau_1} \cdots \int_{-\infty}^{\tau_{n-1}} h_{tri}(t, \tau_1, \dots, \tau_n) u(\tau_1) u(\tau_2) \cdots u(\tau_n) d\tau_n \cdots d\tau_1 \quad (3.18)$$

ou

$$y(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_{tri}(t, \tau_1, \dots, \tau_n) \delta_{-1}(\tau_1 - \tau_2) \delta_{-1}(\tau_2 - \tau_3) \cdots \cdots \delta_{-1}(\tau_{n-1} - \tau_n) u(\tau_1) u(\tau_2) \cdots u(\tau_n) d\tau_n \cdots d\tau_1 \quad (3.19)$$

A terceira forma especial para o *kernel*, o chamado *kernel* regular, envolve uma forma especial da representação como um todo. Fica mais fácil defini-la baseando-se no *kernel* triangular.

De um ponto de vista intuitivo, ele desloca a descontinuidade do *kernel* triangular, tirando-a d “visão” e produz um *kernel* suave sobre todo o primeiro octante. Ele é usado apenas para o caso de sistemas estacionários, e apenas para os casos onde os sinais de entrada são unilaterais.

A representação correspondente pode ser escrita da forma:

$$y(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_{tri}(\tau_1, \dots, \tau_n) u(t - \tau_1) u(t - \tau_2) \cdots u(t - \tau_n) d\tau_1 \cdots d\tau_n \quad (3.20)$$

onde as funções degrau unitário foram suprimidas e os limites infinito mantidos apenas para manter a notação mais simples.

Como exemplo de aplicação pode-se supor a relação causal entre as saídas e a história recente das entradas de um sistema dinâmico não linear, para uma região do cérebro, ou voxel. As séries de Volterra têm sido descritas como uma “série de potência sem memória” e são geralmente concebidas como uma convolução de ordem alta, ou “convolução não linear” das entradas para fornecer a saída. Os *kernels* de Volterra são úteis na caracterização de efetiva conectividade ou influências que um sistema neural exerce sobre outro porque eles representam as características causais do sistema em questão.

Neurobiologicamente eles têm uma interpretação, pois, simples e convincente - eles são sinônimos de efetiva conectividade. É evidente que o *kernel* de primeira ordem incorpora a resposta evocada por uma mudança na entrada em $t-\tau_1$. Noutras palavras, é uma medida, dependente do tempo, da eficácia da excitação. Similarmente, o *kernel* de segunda ordem reflete a influência modulatória da entrada em $t-\tau_1$ na resposta evocada em $t-\tau_2$. E assim por diante para as ordens mais altas.

3.3 SISTEMAS POLINOMIAIS E SISTEMAS DE VOLTERRA

Um sistema descrito por uma soma finita de termos homogêneos da forma:

$$y(t) = \sum_{n=1}^N \int_{-\infty}^{\infty} h_n(\tau_1, \dots, \tau_n) u(t - \tau_1) \cdots u(t - \tau_n) d\tau_1 \cdots d\tau_n \quad (3.21)$$

é chamado um sistema polinomial de grau N , supondo-se $h_N(\tau_1, \dots, \tau_N) \neq 0$. Se um sistema for definido por uma soma infinita de termos homogêneos, ele será chamado um sistema de Volterra. A mesma terminologia é usada, naturalmente, para o caso não estacionário. Adicionando-se um termo de grau 0, diga-se $y_0(t)$, sistemas que têm respostas não nulas a entradas identicamente nulas podem ser representados.

Note-se que como casos especiais estão incluídos os sistemas não lineares estáticos descritos por um polinômio ou uma série infinita na entrada:

$$y(t) = a_1 u(t) + \cdots + a_N u^N(t)$$

$$y(t) = \sum_{n=1}^{\infty} a_n u^n(t) \quad (3.22)$$

Para se ver isto, basta tomar $h_n(t_1, \dots, t_n) = a_n \delta_0(t_1) \dots \delta_0(t_n)$ na expressão geral (3.21) dos sistemas polinomiais. Exemplos de sistemas polinomiais são fáceis de gerar a partir de sistemas estruturados por interconexões. O caso mais simples é uma conexão em cascata de um sistema linear seguido de uma não linearidade polinomial. Se a não linearidade for descrita por uma série infinita de potência, ter-se-á um sistema de Volterra.

Dado que a representação de um sistema de Volterra é por uma série infinita, há que se garantir a sua convergência, para que a representação tenha sentido. A garantia, de uma maneira geral, da convergência de um polinômio de grau n , quando $n \rightarrow \infty$, para qualquer função contínua, é dada pelo teorema de Weierstrass (Rudin, 1964).

As condições para a convergência de séries de Volterra envolvem tipicamente uma cota superior no intervalo de tempo e uma cota para $u(t)$ neste intervalo. Estas cotas dependem tipicamente umas das outras de uma maneira grosseiramente inversa. Isto é, à medida que se aumenta o intervalo de tempo, a cota superior da entrada deve diminuir, e vice-versa. Os cálculos para se encontrar cotas são, em geral, difíceis.

Essa possibilidade de trabalhar com polinômios é que permite o uso no estudo, por exemplo, de modelos de regressão múltipla, ou regressão logística, em estatística.

3.4 ASPECTOS HEURÍSTICOS E MATEMÁTICOS

Uma justificativa para a representação por séries de Volterra é baseada num enfoque bastante intuitivo da descrição de sistemas não lineares. É natural enxergar a saída $y(t)$ de um sistema não linear num instante particular de tempo t como dependente (de uma maneira não linear) de todos os valores da entrada antes do instante t . Isto é, $y(t)$ depende de $u(t - \tau)$ para todo $\tau \geq 0$. É conveniente, embora não necessário, considerar t como sendo o instante presente, e então re-enunciar o conceito como: a saída presente depende dos valores de todas as entradas passadas. Isto leva à seguinte idéia. Se $u(t - \tau)$, para todo $\tau \geq 0$, puder ser caracterizado por um conjunto de quantidades $u_1(t), u_2(t), \dots$, então a saída $y(t)$ pode ser representada como uma função não linear destas quantidades.

$$y(t) = f(u_1(t), u_2(t), \dots). \quad (3.23)$$

O primeiro passo nessa linha de raciocínio é encontrar uma caracterização para o passado de um sinal de entrada. Suponha-se então que t é fixo, e que a entrada $u(t - \tau)$, $0 \leq \tau \leq \infty$, é um elemento do espaço de Hilbert das funções quadrático-integráveis $L^2(0, \infty)$. Isto é:

$$\int_0^{\infty} u^2(t - \tau) d\tau < \infty$$

Ademais, suponha-se que $\phi_1(t), \phi_2(t), \dots$, seja uma base ortonormal para esse espaço:

$$\int_0^{\infty} \phi_i(\tau)\phi_j(\tau) d\tau = \begin{cases} 1, & i \neq j \\ 0, & i = j \end{cases}$$

Então o valor do sinal de entrada em qualquer tempo no passado pode ser escrito na forma:

$$u(t - \tau) = \sum_{i=1}^{\infty} u_i(t)\phi_i(\tau) \quad (3.24)$$

onde

$$u_i(t) = \int_0^{\infty} u(t - \tau)\phi_i(\tau) d\tau \quad (3.25)$$

Embora t seja considerado fixo, este desenvolvimento produz uma caracterização do passado de $u(t)$ em termos de $u_1(t), u_2(t), \dots$, independentemente de t .

Com esta caracterização em mãos, expanda-se a função $f(u_1(t), u_2(t), \dots)$ numa série de potências de maneira que a saída em qualquer instante de tempo seja dada por:

$$y(t) = a + \sum_{i=1}^{\infty} a_i u_i(t) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} a_{i_1 i_2} u_{i_1}(t) u_{i_2}(t) + \dots \quad (3.26)$$

É claro que, na prática, todas as somas infinitas são truncadas para se obter uma representação aproximada. Para que se veja o que tudo isso tem a ver com a representação de Volterra/Wiener, simplesmente substitua-se 3.25 em 3.26 para obter:

$$\begin{aligned} y(t) = & a + \int_0^{\infty} \sum_{i=1}^{\infty} a_i \phi_i(\tau_1) u(t - \tau_1) d\tau_1 + \\ & + \int_0^{\infty} \int_0^{\infty} \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} a_{i_1 i_2} \phi_{i_1}(\tau_1) \phi_{i_2}(\tau_2) u(t - \tau_1) u(t - \tau_2) d\tau_1 d\tau_2 + \dots \end{aligned} \quad (3.27)$$

Com a óbvia definição dos *kernels* em termos das funções ortonormadas $\phi_i(\tau)$, este é precisamente o tipo de representação que se tem discutido até agora.

Além das referências já citadas mais informações sobre sistemas não lineares, séries de Volterra e tópicos correlatos podem ser obtidas em: (Brilliant, 1958), (George, 1959), (Liusternik e Sobolev, 1961), (Barrett, 1963), (Zames, 1963), (Zames, 1964), (Prenter, 1970), (Willems, 1971), (Halme e Orava, 1972), (Isidori e Ruberti, 1973), (Desoer e Vidyasagar, 1975), (Porter e Clark, 1975), (Root, 1975), (Gallman, 1976), (Porter, 1976), (Schetzen, 1976), (Marmarelis e Marmarelis, 1978), (Porter, 1978) e (Mitzel et al., 1979).

3.5 AS TÉCNICAS INDUTIVAS E AS REDES NEURAIS ARTIFICIAIS

Suponha-se que se dispõe de uma amostra de N observações e um conjunto de pares entrada/saída. Trata-se de identificar a relação entrada/saída por um processo de treinamento e validação.

Nos enfoques indutivos, uma forma geral de função considerada é o polinômio de Kolmogorov-Gabor, que é uma versão discreta da série funcional de Volterra:

$$\begin{aligned}\hat{y} &= a_0 + \sum_{i=1}^m a_i u_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} u_i u_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} u_i u_j u_k + \dots \\ &= a_0 + a_i u_1 + a_2 u_2 + \dots + a_{11} u_1^2 + a_{12} u_1 u_2 + \dots + a_{111} u_1^3 + a_{112} u_1^2 u_2 + \dots \\ &= a_0 + a_i u_1 + a_2 u_2 + \dots + a_{11} u_{m+1} + a_{12} u_{m+2} + \dots + a_{mm} u_{ml},\end{aligned}\quad (3.28)$$

onde a saída estimada é denotada por \hat{y} , o vetor externo de entrada u por $(u_1, u_2, \dots, u_{ml})$ e os a_i 's são os pesos, ou coeficientes. Esta expressão é linear nos coeficientes e não linear em u .

A literatura revela que já há muito tempo o número de estruturas, conceitos, métodos e aplicações de redes neurais têm sido bem conhecidos, na modelagem neural. Começou com o trabalho de McCulloch e Pitts (1943), que consideraram o cérebro como sendo um computador consistindo de elementos de cálculo bem definidos, os neurônios

(McCulloch e Pitts, 1943). Os enfoques de teoria dos sistemas ao funcionamento do cérebro são discutidos em várias disciplinas como cibernética, reconhecimento de padrões, biofísica, biologia teórica, psicologia matemática, sistemas de controle e outras.

O “conceito” de redes neurais tem sido adotado no estudo de soluções de problemas relacionados a várias ciências aplicadas e a estudos de implementações de hardware de computador para processamento distribuído paralelo e estruturas de projeto não von Neumann. Em 1958 Rosenblatt introduziu o Perceptron, baseado no funcionamento do neurônio (Rosenblatt, 1958). O elemento neural linear adaptativo (ADALINE), o qual é baseado no *Perceptron*, foi desenvolvido por Widrow e Hoff para reconhecimento de padrões no começo da década de 1960 (Widrow e Hoff, 1960). É popular por seu uso em várias aplicações em processamento de sinais e comunicações. A técnica de aprendizado indutivo chamada de método de grupo para manuseio de dados (*Group Method of Data Handling* (GMDH)), a qual é baseada no *Perceptron*, foi desenvolvida por Alexey Grigorevich Ivakhnenko, na década de 1960, para identificação de sistemas, modelagem e predição em sistemas complexos. Versões modificadas desses algoritmos têm sido usadas em várias aplicações de modelagem (Ivakhnenko, 1970) e (Ivakhnenko, 1971).

Observa-se um rápido desenvolvimento de modelagem por intermédio de redes neurais artificiais, principalmente na direção do conexionismo entre unidades neurais em estruturas de rede e em adaptações de mecanismos de “aprendizagem”.

As técnicas diferem de acordo com os mecanismos adaptados nas redes. Elas são distinguidas por fazerem ajustes sucessivos em força de conexão até que a rede execute uma computação desejada com uma certa precisão. A técnica dos mínimos quadráticos que é usada no ADALINE é uma das importantes contribuições ao desenvolvimento do *Perceptron*. A técnica de aprendizagem por retro-propagação tornou-se bem conhecida mais recentemente. Tornou-se popular pelos trabalhos em redes multicamadas *feedforward* para a solução de vários problemas.

3.6 MECANISMOS DE AUTO ORGANIZAÇÃO NAS REDES NEURAIS ARTIFICIAIS

Qualquer rede neural artificial consiste de unidades de processamento. Estas podem ser de três tipos: entrada, saída e escondida ou associativa. As unidades associativas são os enlaces de comunicação entre as unidades de entrada e as de saída. A maior tarefa da rede é fazer, estabelecer, um conjunto de associações dos padrões da entrada u com a saída y . Quando um novo padrão de entrada é adicionado à configuração, a associação deve ser apta a identificar o seu padrão de saída. As unidades são conectadas a cada outra por intermédio de pesos de conexão; usualmente os valores negativos são chamados inibitórios, e os positivos, excitatórios.

Um processo é dito passar por uma auto-organização quando categorias de identificação ou reconhecimento emergem por intermédio do ambiente do sistema. A auto-organização do “conhecimento” é formada primordialmente na adaptação do mecanismo de aprendizagem na estrutura da rede. A auto-organização na rede é considerada enquanto ele está construindo as conexões entre as unidades de processamento nas camadas para representar os itens de entrada e saída discretos. Processos adaptativos (interações entre variáveis de estado) são considerados dentro das unidades. Funções de limiar lineares ou não lineares são aplicadas nas unidades para uma ativação adicional das suas saídas. Uma função limiar padrão é uma função de transferência linear que é usada para a categorização binária de padrões de características.

Funções de transferência não lineares (não confundir com as clássicas funções de transferência dos sistemas dinâmicos lineares) como as funções sigmóide são usadas para transformar as saídas da unidade. Funções objetivo limiars são usadas nas redes indutivas como um caso especial para medir a objetividade da unidade e decidir se se vai fazer a unidade “ligar” ou “desligar”.

A estratégia é que as unidades competem entre si e ganham a corrida. No caso anterior a saída da unidade é transformada de acordo com a função limiar e alimentada para frente (*feedforward*); enquanto que no caso seguinte, a saída da unidade é alimentada

diretamente se ela está “ligada” de acordo com a função objetivo limiar. Uma função do estado é usada para computar a capacidade de cada unidade. Cada unidade é analisada independentemente das outras. O próximo nível de interação é proveniente das conexões mútuas entre as unidades; o fenômeno coletivo é considerado a partir de malhas da rede. Devido a tais conexões, cada unidade depende do estado de muitas outras unidades. Uma tal estrutura de rede pode ser comutada para um modo de auto-organização usando-se uma lei estatística de aprendizado.

Uma lei de aprendizagem é usada para conectar uma forma específica de mudança adquirida por intermédio dos pesos das sinapses — uma forma que conecte o comportamento presente ao comportamento passado de uma maneira adaptativa, de forma que resultados positivos ou negativos de eventos sirvam como sinais para alguma outra coisa. Esta lei pode ser uma função matemática, tal como uma função energia que dissipa energia na rede, ou uma função erro que mede o erro residual da saída.

Um método de aprendizagem segue um procedimento que avalia esta função para fazer mudanças pseudo-aleatórias nos valores dos pesos, retendo aquelas mudanças que resultam em melhorias para se obter a resposta de saída ótima. Vários procedimentos diferentes têm sido desenvolvidos baseados na minimização do erro quadrático médio da unidade de saída (a técnica dos mínimos quadráticos é a mais simples e a mais popular). Existem também técnicas baseadas em lógica booleana de auto-organização.

Neste capítulo apresentou-se conceitos e fundamentos sobre os sistemas não lineares presentes em várias situações do mundo real. Na seqüência do capítulo foram introduzidas as séries de Volterra e a possibilidade de estabelecer uma correspondência entre uma dada representação e a estrutura de interconexões do sistema. Mostrou-se que os *kernels* de Volterra são sinônimos de efetiva conectividade. Nas seções 3.5 e 3.6 procurou-se mostrar a relação entre as técnicas indutivas e as redes neurais artificiais e os mecanismos de auto organização nestas redes. Um maior detalhamento sobre as redes neurais artificiais é apresentado no próximo capítulo, bem como modelos de regressão logística e de regressão múltipla. A abordagem tradicional com modelos de regressão é utilizada neste trabalho na comparação com as redes neurais artificiais.

4. MODELOS DE REDES NEURAIIS ARTIFICIAIS E DE REGRESSÃO

“All models are wrong – but some models are useful.”

(Box, 1979, p. 202)

“Todos modelos estão errados – mas alguns modelos são úteis” (Tradução livre)

Neste capítulo é feita inicialmente uma breve exposição sobre: a modelagem de sistemas complexos, classificação de padrões e teorias da aprendizagem. Em seguida é mostrado o neurônio biológico e são introduzidos os modelos tradicionais de Redes Neurais Artificiais, bem como algumas limitações associadas a estes modelos. Posteriormente, são apresentadas as redes *feedforward* construtivas do tipo NSRBN, Regressão Logística e Múltipla, além dos Simuladores utilizados neste trabalho.

4.1 MODELAGEM DE SISTEMAS COMPLEXOS

O estudo e a modelagem de objetos complexos, ou sistemas complexos, constitui um permanente desafio. Entre estes pode-se citar:

- Aproximação e extrapolação de funções;
- Identificação de sistemas;
- Reconhecimento de padrões;
- Previsão de processos estocásticos;
- Classificação.

Os métodos usados para o estudo e solução desses problemas dividem-se em dois tipos principais: hipotético-dedutivos (lógica-matemática) e indutivos. Como afirmado em Campello de Souza (2005), “O estudo da natureza do conhecimento é complexo e indigesto”. Segundo Campello de Souza (2005), as noções filosóficas introduzidas pelo filósofo americano Charles Sanders Pierce podem ajudar nessa tarefa. Segundo ele:

“No sistema lógico de Peirce, a abdução e a dedução contribuem ao entendimento conceitual de um fenômeno, enquanto que a indução adiciona detalhes quantitativos (mensuração) ao conhecimento conceitual. Uma dedução é um silogismo, que parte de duas premissas, uma maior e uma menor, e chega a uma conclusão (consequência lógica, necessária). Uma abdução é um silogismo cuja premissa menor é apenas provável. A indução é um raciocínio em que, de fatos particulares, se tira uma conclusão genérica. Peirce justificou a validade da indução como um processo capaz de auto-correção, mas afirmou também que nem a indução nem a dedução podem ajudar a desvendar a estrutura interna do conhecimento. Da mesma maneira que a análise exploratória de dados desempenha um papel de ‘construtora de modelos’ para a análise confirmatória de dados (quem trabalha com massas de dados e constrói modelos conhece esse processo), a abdução desempenha o papel de ‘exploradora de caminhos viáveis para futuras indagações. Vê-se então que a lógica da abdução ajusta-se bem ao paradigma da análise exploratória de dados.

Em resumo, para encerrar esta digressão filosófica:

- *no estágio da abdução o objetivo é explorar os dados, encontrar padrões e sugerir uma hipótese plausível;*
- *na dedução refina-se a hipótese, baseado em outras premissas plausíveis; e*
- *na indução se faz a substanciação empírica.*

Um pensador deve portanto usar as três coisas, pois a abdução cria, a dedução explica e a indução verifica. Esse é o pensamento de Peirce. No estudo e aplicação da teoria da decisão, freqüentemente esses aspectos estão próximos uns dos outros.” (Campello de Souza, 2005, p.249)

Este é o caso das redes neurais artificiais (RNA). Por um lado há todo um aparato matemático que dá estrutura aos modelos conceituais básicos (entrada, processamento, saída, realimentação, arquitetura, otimização, adaptação, estimação e outros; funciona como se fossem células tronco), e por outro um enfoque indutivo ao fenômeno que se está estudando, ou ao problema que se está procurando resolver (treinamento a partir de uma base de dados, estimação, validação, calibração, e outros). Assim, não há teoremas (resultados científicos) sobre o objeto do estudo. A fase de abdução é essencialmente a mesma que é usada no procedimento científico normal. A fase dedutiva, no caso da abordagem por RNA, não existe propriamente. Ela fica muito mais embutida na própria

rede, em conceitos mais fundamentais, mais básicos, como mencionado. No caso do procedimento científico usual, depois dos teoremas, dos silogismos, resultados hipotéticos-dedutivos, procede-se à indução para a vindicação de uma teoria. Isto é feito tipicamente por intermédio de técnicas de observação, com instrumentos de estatística.

No caso das RNA há uma espécie de antecipação dessa fase de vindicação, precedida de uma fase de treinamento. As três fases do sistema filosófico de Peirce estão, portanto, mais próximas e mais organicamente ligadas. Existe abdução, quando se escolhe as variáveis pertinentes e se delinea o problema, “dedução” quando se escolhe a arquitetura, por assim dizer, da rede, e indução quando se faz o treinamento e a calibração e validação da rede. As três fases são, por conseguinte, mais amalgamadas. A fase “dedutiva”, por assim dizer, pois a rigor não existe, é mais recôndita. Assim, as RNA não concorrem com a ciência e sim colaboram com ela. No estudo e operação de sistemas complexos, para os quais não houver teoria, oferece um meio efetivo de se avançar.

O desafio é maior em áreas como finanças, meteorologia, medicina, entre outras. Em muitos casos os problemas são mal definidos, a informação *a priori* é escassa, existe um número grande de variáveis e fatores condicionantes não mensuráveis, ruído (variância) e, para piorar a situação, amostras pequenas. Essa é, freqüentemente, a situação prática. É nesse ambiente que as RNA têm algo a oferecer. Pode-se ter uma compreensão mais detalhada desses conceitos no Anexo A.

4.1.1 Classificação de Padrões

De uma forma geral, classificação pode ser entendida como: um processo cognitivo básico de distribuir de forma organizada coisas ou pessoas em classes ou categorias do mesmo tipo. No desenvolvimento deste processo, grandes categorias podem ser divididas em outras menores até que cada coisa ou pessoa esteja em seu grupo próprio. A classificação está implícita em várias tarefas do dia a dia das pessoas e das organizações. Um caso de classificação muito utilizado na área de ciências biológicas é aquele relativo às plantas. A taxonomia utilizada nesta área é antiga e utiliza

características morfológicas (formatos de folhas e frutos, e outros) das plantas para separá-las em grupos similares. É nestes casos é muito comum usar as próprias características como parte dos nomes das plantas. Isto pode ser percebido na classificação de folhas, que tem um dos primeiros registros atribuído a Theophrastus (300 aC), considerado pai da botânica.

No processo decisório, no caso específico tratado aqui utilizado para classificar, variáveis, atributos e aspectos são componentes importantes (Campello de Souza, 2005). Estes componentes permitem caracterizar melhor as situações e assim obter as informações desejadas, dado que as palavras não são suficientes. Isto sem dúvida é algo aderente ao caso do diagnóstico médico. Para se entender melhor o problema da classificação, serão considerados os conceitos fornecidos por Campello de Souza (2005) para Variável, Atributo e Aspecto:

Variável: quantidade que pode assumir qualquer valor dentro de um conjunto de valores;

Atributo: combinação de variáveis segundo uma certa regra, uma certa fórmula (é um indicador);

Aspecto: é uma síntese dos atributos, uma agregação de atributos.

Tomando-se, por exemplo, o caso de classificar folhas segundo as diversas formas que elas apresentam. Um conjunto de oito possíveis formas e a correspondente denominação morfológica está apresentado na Figura 4.1.

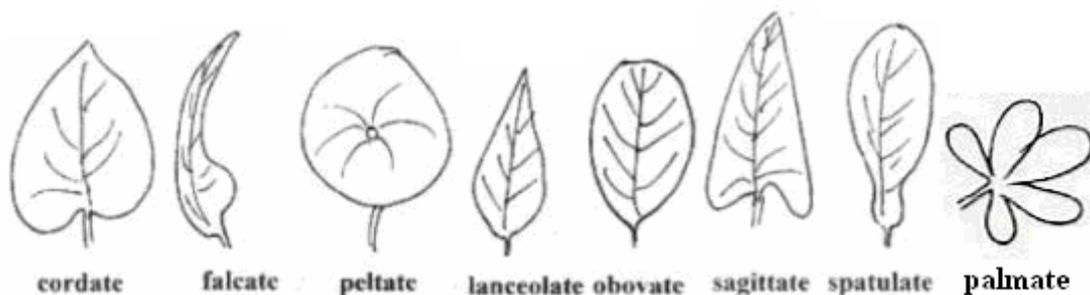


Figura 4.1 - Formas de folhas

Podem-se definir as seguintes variáveis:

- Área da folha;

- Perímetro da folha;
- Área da envoltória convexa da folha (menor figura convexa que contém a folha dada; e figura convexa deve ser entendida como aquela que um segmento de reta unindo quaisquer dois pontos dentro da figura, tem que estar totalmente contido na figura);
- Diâmetro maior da folha (maior distância obtida entre duas retas colocadas de forma paralela aproximando-as externamente de um lado e de outro da folha até tocá-la);
- Diâmetro menor da folha (menor distância obtida entre duas retas colocadas de forma paralela aproximando-as externamente de um lado e de outro da folha até tocá-la);

Utilizando-se as variáveis alguns indicadores ou atributos podem ser definidos, pelas Equações 4.1, 4.2 e 4.3 a seguir:

$$\text{Atributo 1} = \frac{\text{Área da Folha}}{(\text{Perímetro da Folha})^2} \quad (4.1)$$

$$\text{Atributo 2} = \frac{\text{Diâmetro Menor da Folha}}{\text{Diâmetro Maior da Folha}} \quad (4.2)$$

$$\text{Atributo 3} = \frac{\text{Área da Folha}}{\text{Área da Envoltória Convexa}} \quad (4.3)$$

Ao serem calculados os atributos para uma amostra de folhas dos tipos mostrados na Figura 4.1, poderiam ser obtidos resultados como aqueles apresentados na Tabela 4.1.

Tabela 4.1 - Atributos das Folhas

AMOSTRA	ATRIBUTO 1	ATRIBUTO 2	ATRIBUTO 3
1	0,2	0,25	0,82
2	0,88	0,93	0,95
3	0,31	0,87	0,6

É interessante perceber que é possível classificar os tipos de folhas a partir dos resultados obtidos. A amostra (1) poderia pertencer ao tipo falcate, a (2) ao peltate e a (3) ao palmate.

De forma análoga pode ser realizada a classificação de pacientes quanto ao acometimento de alguma disfunção em seu organismo. Para tornar mais clara esta possibilidade, apresenta-se um exemplo na área de médica a seguir. A Tabela 4.2 trás variáveis e atributos (parâmetros) relativos a uma amostra de pacientes. Estes pacientes (1, 2, 3 e 4) poderiam ser então classificados individualmente em função das variáveis e atributos como hígidos ou portadores de alguma disfunção. Assim, da mesma maneira que a classificação das folhas e plantas poderia apoiar o agrônomo nos cuidados e medidas a serem tomadas para plantar, a correta classificação do doente pode ajudar o médico a tratar os pacientes por meio de melhores diagnósticos, o que conduzirá a condutas terapêuticas mais eficientes e eficazes.

Tabela 4.2 - Variáveis e atributos de pacientes.

PACIENTE	HVE	IMC	PSC	PDC	FCC	RENINA	ESPESSURA PAREDE VE	C/Q	IDADE	COLEST. TOTAL
1	1	22,9	204	116	76	1,0	1,1	0,94	57	134
2	0	30,8	150	108	84	2,5	0,8	1,02	40	209
3	1	27,4	146	102	97	1,1	1,1	1,00	71	154
4	1	27,1	166	86	52	0,0	1,1	0,91	67	110

As variáveis e parâmetros apresentados na Tabela 4.2 formam apenas um subconjunto das várias combinações possíveis, pois poderiam ser adicionados ou substituídos por outros como: Pressão de Pulso, Produto da Resistência Periférica pela Complacência dos Grandes Vasos Arteriais (RC), Índice Pulsátil da Pressão Arterial (IPPA), IPPARC, Harmonia, Pressão Arterial Média (PAM), Desvio Padrão dos Intervalos RR normais durante o dia (SDNNDia), Média de intervalos RR durante o dia (RRMedioDia), Média da frequência cardíaca nas 24h (FCMEDIA24H), Extrasístoles ventriculares isoladas e/ou bigeminadas/24h (EVisolbig), Fração de Ejeção, Duração do complexo QRS (QRS), Taquicardia supraventricular/24h (TaquicVentri), Média da frequência cardíaca durante à noite (FCMEDIANOITE).

Outras ferramentas são também utilizadas em tarefas de classificação, além das redes Neurais Artificiais. Alguns exemplos destas ferramentas são: Regressão Logística, Discriminante Linear de Fisher, Árvores de Decisão e Modelos de Markov Oculto (HMM) (Hosmer e Lemeshow, 1989), (Beale, 1990), (Krogh, 1998), (Rabiner, 1989).

4.1.2 Aprendizagem e as Teorias de Piaget e Behaviorista

“Knowledge, then, is a system of transformations that become progressively adequate.”

(O conhecimento, portanto, é um sistema de transformações que se torna progressivamente adequado – tradução livre).

“From this time on, the universe is built up into an aggregate of permanent objects connected by causal relations that are independent of the subject and are placed in objective space and time.”

(De agora em diante, o universo é construído em uma agregação de objetos permanentemente conectados por relações causais que são independentes do assunto e são colocados de forma objetiva no espaço e no tempo – tradução livre).

“Logic and mathematics are nothing but specialised linguistic structures.”

(Lógica e matemática não são nada mais do que estruturas lingüísticas especializadas – tradução livre).

Jean Piaget (1896-1980) psicólogo e filósofo suíço foi um importante teórico do processo do conhecimento humano (epistemologia). Ele desenvolveu uma teoria na busca de explicação sobre a gênese do conhecimento. Por conhecimento deve ser entendido uma explicação não empírica e não sensorial da realidade objetiva, e uma progressiva e refinada adaptação e alteração do meio que envolve o indivíduo.

Piaget identificou quatro estágios no desenvolvimento da capacidade de raciocínio de um indivíduo: sensorio motor, pré-operatório, operacional concreto e das operações formais (Piaget, 1983). Na sua teoria estes estágios se sucedem até o início da

adolescência e têm correspondência com as sucessivas fases de crescimento físico. Assim, a criança, no seu desenvolvimento, constrói estruturas cognitivas sofisticadas que vão dos poucos e primitivos reflexos do recém-nascido até às mais complexas atividades mentais do jovem adulto. Segundo a teoria de Piaget, por estrutura cognitiva entende-se como um "mapa" mental interno, um "esquema" ou uma "rede" de conceitos construídos pelo indivíduo para compreender e responder às experiências que decorrem dentro do seu meio envolvente. O terceiro estágio, que corresponde ao período dos sete aos doze anos, é onde ocorrem as primeiras operações lógicas e o indivíduo é capaz de classificar objetos conforme suas semelhanças ou diferenças.

Os estágios têm caráter integrativo, portanto as estruturas construídas em um dado nível são integradas nas estruturas do nível seguinte. Por exemplo, um "esquema de reunião" para comportamento como de um bebê que empilha pequenas peças em formato de cubo, permanece na criança mais velha que junta objetos procurando classificá-los e mesmo em operações lógicas como a reunião de duas classes (os pais mais as mães igual a todos os pais).

A psicologia da aprendizagem utiliza pontos importantes da teoria de Piaget para estruturar situações de aprendizagem:

- os estágios de desenvolvimento cognitivo fornecem indicadores para a definição da complexidade da situação, ou seja, deve-se propor situações de aprendizagem compatíveis com o estágio atual de desenvolvimento cognitivo do indivíduo;
- para Piaget aprender é atuar sobre o objeto da aprendizagem para compreendê-lo e modificá-lo. Surge então o outro conceito chave - a aprendizagem ativa;
- como aprender é uma contínua adaptação ao meio externo, aprende-se quando se entra em conflito cognitivo, ou seja, quando o indivíduo se defronta com uma situação que não sabe resolver. O organismo se desequilibra frente ao novo, mas como todo organismo vivo procura o equilíbrio. Então, para encontrar o equilíbrio, ele lança mão de um complexo processo de adaptação. Adaptação deve ser entendida neste contexto, como o processo pelo qual o sujeito adquire um equilíbrio entre assimilação e acomodação. A assimilação refere-se à introjeção de conhecimentos sobre o meio e a incorporação ao conjunto de conhecimentos já existentes. Por meio da incorporação, a estrutura de conhecimento existente se

modifica de modo a acomodar-se a novos elementos - esta modificação é denominada acomodação. Equilíbrio é o processo de organização das estruturas cognitivas num sistema coerente, interdependente, que possibilita ao indivíduo a adaptação à realidade. É a partir deste entendimento que as situações de aprendizagem baseiam-se em jogos e desafios, nos quais o sujeito é defrontado com um problema novo para resolver.

A psicologia possui de um lado as aproximações cognitivas e de outro o estudo do comportamento (behaviorista). Há contribuições das duas visões para as redes neurais artificiais. Tomando-se o caso do treinamento de uma RNA, neste procedimento um teste é realizado em geral por uma pessoa e o sistema artificial responde semelhantemente ao que acontece com um animal quando treinado pela tecnologia behaviorista. Os behavioristas utilizam uma modelagem (técnica) pra ensinar de forma progressiva um animal. Da mesma forma as RNA podem emular respostas adequadas por intermédio de treinamento progressivo.

Uma definição possível para aprendizado é a ocorrência de mudança no comportamento em função de procedimentos de treinamento. Esta mudança é verificada utilizando-se uma medida de desempenho antes e após o treinamento. Condicionamento é uma forma freqüente entre os behavioristas de se reportar ao aprendizado. São descritas duas formas ou tipos de condicionamento:

- Clássico: Este tipo está amplamente exemplificado pelos conhecidos experimentos com a salivação de cães em resposta a um sino, do fisiologista e psicólogo experimental russo Ivan Petrovich Pavlov (1849-1936). Durante as medições das secreções salivares Pavlov observou que os cães salivavam não somente em resposta à comida (estímulo incondicionado), mas também em relação a um estímulo externo, no caso do experimento, o toque de um sino (estímulo condicionado). Para o segundo caso a salivação para o toque do sino é a resposta condicionada. Uma de suas obras mais importantes sobre este tema foi *Conditioned Reflexes* em 1926. O aprendizado é pertinente tanto para estímulos condicionados como incondicionados e uma associação de dois estímulos (oferta de alimento e toque do sino) pode ser utilizada para definir o aprendizado.
- Operante: Ele é também chamado de instrumental e teve em Burrhus Frederic Skinner (1904–1990) um de seus fortes proponentes. Ele envolve três componentes:

feedback, treinamento baseado em respostas previamente conhecidas e amostragem ou treinamento progressivo. Este tipo de treinamento depende de *feedback*, dessa forma quando por exemplo, um animal faz algo correto ele é recompensado (reforçado) e quando faz algo incorreto, ele é punido. Reforço e punição são os estímulos que aumentam a frequência de um comportamento. Nas seções a seguir é mostrado que as redes neurais artificiais também utilizam *feedback* como instrumento de aprendizagem. Adicionalmente, é importante mencionar que o conceito de reforço também está presente nas RNA, pois ele é utilizado para adaptação das conexões entre neurônios, em alguns tipos de RNA.

4.1.3 O Neurônio Biológico

Em conjunto os sistemas nervoso e endócrino produzem a maior parte das funções do corpo humano. O sistema nervoso, em geral controla as atividades rápidas do corpo enquanto o sistema endócrino as funções do metabolismo. Uma quantidade muito grande de informações provenientes de diversos tipos de receptores sensoriais são recebidas pelo sistema nervoso. Assim, a resposta produzida que será dada pelo corpo é função da combinação das informações recebidas pelos sensores com outras obtidas pela hereditariedade ou por mecanismos de aprendizagem. A parte do sistema nervoso chamada de cérebro é que provê uma quantidade grande de controle sobre as funções do corpo. Os neurônios são um tipo de célula existente no cérebro especializada à atividade nervosa. O cérebro humano é composto por algo em torno de 10^{11} neurônios. A estrutura de um neurônio pode ser observada na Figura 4.2.

O neurônio possui ramificações chamadas de dendritos que conduzem sinais de uma extremidade para o corpo da célula, chamada de soma. A partir de um ponto da célula surge um filamento comprido que é o axônio. Na extremidade do axônio existem estruturas chamadas terminais do neurônio, os quais são conectados aos dendritos dos outros neurônios. Esta conexão é chamada de sinapse. O fluxo de informações caminha sempre dos dendritos para o axônio. Ocorre um processo de soma dos estímulos de entrada e como resultado pode ser gerado um impulso elétrico que se propaga até os terminais do axônio.

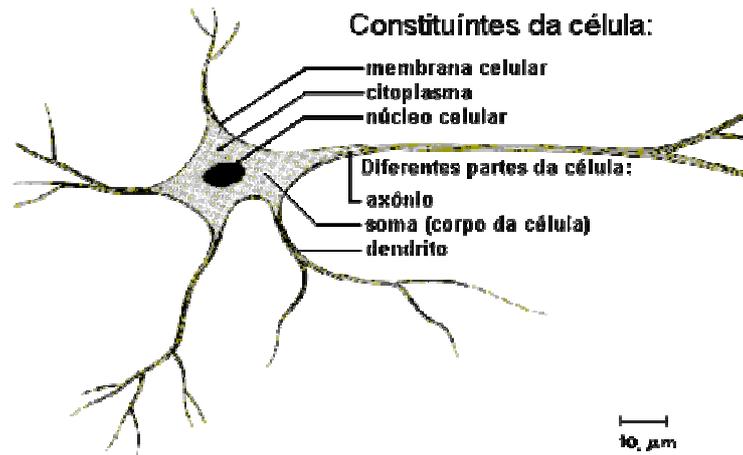


Figura 4.2 - Neurônio Biológico

O processo que controla a produção do impulso é normalmente aceito como uma função de limiar. Esta função produz uma saída, caso a soma das entradas seja maior ou igual a um dado limiar. Este tipo de mecanismo define uma característica chamada “tudo ou nada”, ou seja, o neurônio dispara ou não (a sinapse de um único neurônio é sempre toda excitatória ou inibitória). Uma vez que o neurônio esteja conectado a vários outros o processo ocorre a partir de uma mistura de entrada positivas e negativas. Portanto, a frequência de geração de impulsos elétricos pelo neurônio depende do processo de integração das suas entradas e seu padrão de disparo pode ser modificado pelo mecanismo de aprendizagem (adaptação) (Loesch e Sari, 1995).

Convém acrescentar mais algumas considerações básicas sobre o funcionamento dos neurônios. Uma delas é sobre a existência de um período refratário no intervalo entre dois impulsos, durante o qual um impulso não pode ser gerado nem propagado. A propagação só será possível após a restauração química dentro do neurônio. Uma segunda consideração trata do espaço entre duas células chamado de *gap* sináptico (50 a 200×10^{-8} cm). Um neurônio comunica-se com outro pelo *gap* sináptico por intermédio de produtos químicos específicos chamados de neurotransmissores. Existem muitos tipos de neurotransmissores, dos quais somente um deles é liberado em todos os terminais durante uma dada ativação do neurônio. Um papel fundamental na estocagem de informações é atribuído às sinapses, onde a força da conexão é responsável pela quantidade de informação estocada. Este mecanismo pode ser mais bem compreendido

pela lei de Hebb: a força de uma conexão sináptica é incrementada se esta estiver entre dois neurônios excitados simultaneamente (Hebb, 1949).

Na seção 4.2.1 é apresentado um modelo de neurônio artificial, onde é possível comparar o neurônio biológico com este modelo simplificado.

4.2 MODELOS DE REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais são modelos matemáticos inspirados na estrutura do cérebro humano. Elas foram desenvolvidas inicialmente em uma tentativa de modelar processos de aprendizado do cérebro, e diferentemente dos neurônios biológicos, são implementadas por meio de *hardware* ou de *software* (McCulloch e Pitts, 1943). A palavra neural é utilizada hoje em dia por razões históricas ligadas às primeiras pesquisas realizadas (Braga; Ludermir e Carvalho, 2000, p.5), (Nascimento e Yoneiama, 2000, p.112). Portanto, é importante entender que nem todas as redes neurais artificiais têm como motivação a modelagem do sistema nervoso.

Observa-se que não há uma definição universalmente aceita de rede neural artificial. Contudo, pode-se dizer de uma forma geral, que as RNAs são sistemas distribuídos de forma paralela, constituídos por unidades de processamento simples chamadas também de: neurônios artificiais, nodos, nós ou simplesmente neurônios, em alusão aos neurônios biológicos. Estes neurônios efetuam cálculos de determinadas funções matemáticas e estão dispostos em camadas. As camadas são interligadas por meio de conexões, geralmente unidirecionais, que por sua vez têm pesos associados a elas, para os modelos de RNAs com pesos. Estes pesos são a forma pela qual ocorre o armazenamento do conhecimento pela RNA, os quais ponderam cada entrada recebida por um determinado neurônio que compõe a rede, e desta forma definem a correspondente influência na sua saída (Braga; Ludermir e Carvalho, 2000, p.1).

Aplicações de RNA na área de reconhecimento de padrões, incluindo-se problemas de classificação, vão desde a análise de imagens médicas, o reconhecimento óptico de caracteres, de voz, de escrita, a identificação de características biométricas, a

classificação de documentos, dispositivos de busca na internet e classificação de crédito. Contudo, o campo de aplicações é vasto e permeia várias áreas e vai além das tarefas de classificação como: Financeira (bolsa de valores e mercado de ações); Predição de séries temporais (clima, tempo); Sistemas de Controle (robôs autônomos); Redução de Ruído (ECG); Genética (seqüenciamento de DNA) entre outras.

A capacidade das RNAs para fornecer soluções a problemas requer uma etapa de aprendizagem, na qual um conjunto de dados de treinamento, que funciona como exemplo, é apresentado à rede. Durante a fase de treinamento um conjunto de procedimentos bem definidos, chamados algoritmos de aprendizado, permitem que haja a adaptação dos parâmetros da rede, de forma que sejam extraídas características relevantes dos dados de entrada da rede. Os diversos tipos de algoritmos de aprendizado existentes variam pela forma como o ajuste dos pesos das conexões é realizado. Durante a fase de aprendizado das RNAs as interconexões entre os neurônios podem sofrer três tipos de modificações: estabelecimento de novas conexões, eliminação de conexões existentes ou alteração dos pesos das conexões.

Os métodos para treinamento de RNA podem ser classificados em dois paradigmas principais: aprendizado supervisionado e não supervisionado. O aprendizado supervisionado é o mais comumente utilizado, nele os valores de entrada e saída desejada são fornecidos para a RNA e em seguida é calculada a resposta da rede. Então, é feita uma comparação entre a ação ótima esperada para a rede, resposta desejada, e a calculada por ela. A soma dos erros quadráticos, Equação 4.1, é um exemplo de função de “custo” utilizada, a qual é minimizada pelo algoritmo de treinamento para definição dos ajustes a serem realizados nos pesos das conexões. Estes ajustes conduzem a rede para uma solução, caso ela exista. O método de aprendizado supervisionado pode ser realizado na forma *on-line* ou *off-line*, nesta última os dados do conjunto de treinamento não se alteram e, portanto a solução encontrada permanece fixa. No *on-line* há uma contínua adaptação da rede ao conjunto de dados, portanto caso haja alguma mudança no conjunto, a rede será capaz de ajustar-se a esta mudança e dessa forma continuar a fornecer uma resposta adequada.

$$E(p) = \frac{1}{2} \sum_{i=1}^k (d_i - o_i)^2 \quad (4.4)$$

onde: p é o padrão de entrada considerado, k é o número de neurônios de saída, d_i a saída desejada correspondente ao neurônio i e o_i a saída atual da rede.

O tipo de problema a ser solucionado e a maneira como os neurônios encontram-se organizados em uma RNA estão intrinsecamente vinculados. A definição da arquitetura é feita em função de alguns parâmetros da RNA como: número de camadas; número de neurônios por camada; forma como os neurônios estão conectados e topologia da rede. São citados a seguir alguns tipos de classificações possíveis.

Quanto ao número de camadas, as RNAs podem ser:

- Redes com uma única camada (só existe um nó entre qualquer entrada e qualquer saída);
- Redes com múltiplas camadas (existe mais de um neurônio entre uma entrada e uma saída).

Quanto aos tipos de conexões entre os neurônios, as RNAs podem ser:

- *feedforward* ou acíclica, neste caso a saída de algum neurônio em uma camada de ordem i não pode ser usada como entrada de neurônios em camadas de ordem inferior ou igual a i .
- *feedback* ou cíclica, contrariamente ao caso anterior nestas redes a saída de algum neurônio na camada de ordem i é usada como entrada para neurônios em camadas de índices menor ou igual a i .

Um outra possível forma de classificar as RNAs é quanto a sua conectividade. Elas podem ser completamente conectadas ou parcialmente conectadas.

Na Figura 4.3 pode ser visto um exemplo de uma RNA *feedforward* com uma camada completamente conectada.

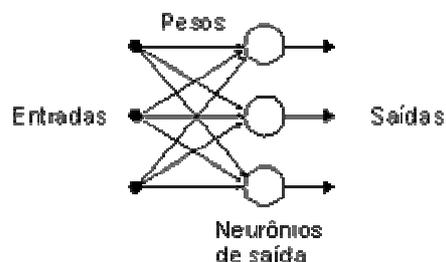


Figura 4.3 - Arquitetura de uma RNA de uma camada; os círculos representam neurônios e as setas as interligações ponderadas.

Uma característica importante das redes neurais é sua capacidade de generalizar; isto significa responder de forma coerente a dados não apresentados à rede anteriormente, utilizando-se para tanto as características extraídas dos exemplos de treinamento. De uma forma geral pode-se citar outras características principais das redes neurais além da generalização, como:

- a aptidão para resolver problemas não linearmente separáveis;
- a habilidade de extrair conhecimentos a partir de exemplo (aprendizagem);
- a capacidade de testar simultaneamente todas as possibilidades associadas às diversas informações recebidas (processamento paralelo);
- gerar um desempenho satisfatório quando há perda de neurônios (tolerância a falhas);
- reaprender adequando-se a novos conjunto de exemplos (adaptabilidade) além de trabalhar com uma grande quantidade e diversidade de dados (Valença, 2005).

4.2.1 O Neurônio Artificial

Dentre os primeiros pesquisadores na área de RNAs estão Warren McCulloch e Walter Pitts, os quais desenvolveram e publicaram em 1943 um modelo de neurônio artificial (McCulloch e Pitts, 1943). A representação matemática deste modelo apresenta um neurônio como uma unidade de n entradas (x_1, x_2, \dots, x_n) , equivalentes aos dendritos, e uma única saída y a qual foi associada ao axônio. As sinapses são representadas no modelo por meio dos pesos (w_1, w_2, \dots, w_n) correspondentes a cada entrada. De forma análoga ao neurônio biológico o artificial dispara quando a soma ponderada $(\sum_{i=1}^n x_i w_i)$ de suas entradas excede o limiar de excitação ou *threshold* (θ). A aplicação do valor correspondente a soma ponderada das entradas na função de ativação, permitirá definir se a saída do neurônio de McCulloch e Pitts estará ativa (1) ou não (0). Na forma original do modelo de McCulloch e Pitts a função de ativação do neurônio é dada pela função de limiar Equação 4.5.

$$\sum_{i=1}^n x_i w_i \geq \theta \quad (4.5)$$

Ou ainda de forma equivalente:

$$\sum_{i=1}^n x_i w_i - \theta \geq 0 \tag{4.6}$$

Na Figura 4.4 está ilustrado um neurônio de McCulloch e Pitts.

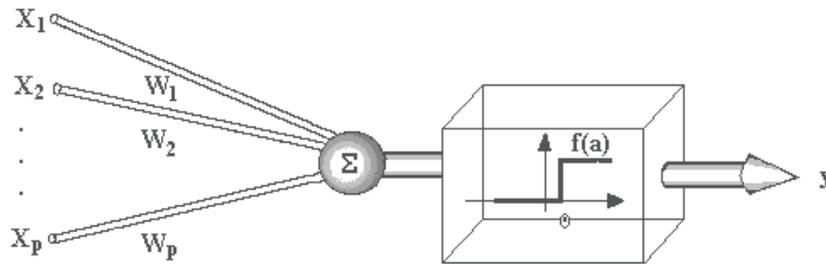


Figura 4.4 - Neurônio de McCulloch e Pitts

O modelo de McCulloch e Pitts apresenta algumas limitações como o fato de que redes com apenas uma camada só conseguem implementar funções linearmente separáveis e que o modelo foi proposto com pesos fixos, não ajustáveis. Portanto, surgiram outros modelos com diferentes funções de ativação e que permitem produzir qualquer tipo de saída. Dentre as outras possíveis funções de ativação, ou funções de transferência como também são chamadas, tem-se:

Função Linear:

$$y = \alpha x \tag{4.7}$$

Função Rampa:

$$y = \begin{cases} +\delta & \text{se } x > +\delta \\ x & \text{se } |x| < +\delta \\ -\delta & \text{se } x \leq -\delta \end{cases} \tag{4.8}$$

Função Degrau:

$$y = \begin{cases} +\delta & \text{se } x > 0 \\ -\delta & \text{se } x \leq 0 \end{cases} \tag{4.9}$$

Função Sigmoidal - Logística:

$$y = \frac{1}{1 + e^{-x/T}} \tag{4.10}$$

Onde T define a suavidade da curva.

4.2.2 *Perceptron* Multicamadas

O *Perceptron* foi proposto pelo psicólogo Frank Roseblatt (Roseblatt, 1958) e é uma das primeiras e mais simples formas de rede *feedforward*. O reconhecimento de padrões foi uma das tarefas para qual ele foi pensado. Na sua estrutura mais básica ele é constituído por uma camada de entrada, uma de conexões e uma de saída composta por neurônios do tipo McCulloch e Pitts (seção 4.2.1). Somente o nível de saída possui propriedades adaptativas. O *perceptron* incorporou um algoritmo de aprendizagem, no qual os pesos na camada de saída ou resposta são adaptados de forma a obter uma saída desejada. Após este modelo ter surgido houve um grande desenvolvimento nas pesquisas na área de RNA, na tentativa de superar as limitações deste modelo em relação ao cálculo de funções mais complexas.

Na década de 60 houve a proposição do teorema de convergência do *perceptron*. Este teorema mostra que um neurônio do tipo McCulloch e Pitts treinado com algoritmo do *perceptron* sempre converge, caso o problema seja linearmente separável (Roseblatt, 1962).

Em 1960 Bernard Widrow e Ted Hoff propuseram um modelo chamado de ADALINE (*ADaptive LINear Element ou ADaptive Linear NEuron*), a descrição deste modelo foi dentro do contexto de processamento de sinais no então *Institute of Radio Eletronics* (IRE), hoje *Institute of Electrical and Electronic Engineers* (IEEE) (Widrow e Hoff, 1960). O algoritmo *Least Mean Squares* de treinamento ficou também conhecido como regra delta e deu origem ao algoritmo *back-propagation* para treinamento de *perceptrons* multicamadas. Esta regra de aprendizado minimiza a soma dos quadrados dos erros lineares sobre o conjunto de treinamento. A distinção chave entre o ADALINE e uma *perceptron* com um neurônio de saída está no modo como as leis de aprendizado são aplicadas. No *perceptron* o delta é calculado usando-se a saída binária do neurônio e a saída desejada, e no ADALINE o delta é calculado usando a ativação neural antes que a função transferência binária seja aplicada.

Na década de 60 as escolas simbolista e conexionista (relativo às redes neurais artificiais) competiam por prestígio e recursos para as suas pesquisas. Em 1969 Marvin

Minsky e Seymour Papert mostraram matematicamente e chamaram a atenção para algumas tarefas que o *perceptron* não era capaz de realizar, pois que ele só resolve problemas linearmente separáveis (Minsky e Papert, 1969). Embora fosse lógico que redes *perceptron* com mais de uma camada de neurônios, chamadas *multi-layer perceptrons*, tivessem a capacidade de aprender padrões não-linearmente separáveis era desconhecido um algoritmo que pudesse realizar esta tarefa. Assim, por mais de uma década a abordagem conexionista ficou em segundo plano.

No início da década de 80 Jonh Hopfield publicou um artigo que chamou a atenção para as abordagens associativas das redes neurais artificiais (Hopfield, 1982). Este artigo foi responsável por parte da retomada das pesquisas na área. Hopfield mostrou a relação entre redes recorrentes auto-associativas e sistemas físicos, isto permitiu a utilização de teorias então correntes da física para estudar esses modelos. Apesar da apatia em relação as pesquisas na área, em grande parte devido à repercussão do trabalho de Minsky e Papert, além de Hopfield alguns poucos pesquisadores continuaram trabalhos na área, entre eles: Steven Grossberg em 1980 estabeleceu um novo princípio de auto-organização chamado de *Adaptive Resonance Theory* – ART e Teuvo Kohonen em 1982 introduziu as *Self-Organizing Maps* – SOM.

Paul Werbos em 1974 demonstrou um método para treinar *multilayer perceptrons*, essencialmente idêntico ao *backpropagation* (Werbos, 1974), que só se tornou amplamente conhecido após a edição do livro *Parallel Distributed Processing* em 1986 (Rumelhart; Hinton e Williams, 1986). Neste livro Rumelhart, Hinton e Williams anunciaram a descoberta de um método que permitia uma rede *perceptron* multicamadas aprender a discriminar classes não linearmente separáveis. Este método, a Regra Delta Generalizada, ficou mais conhecido como algoritmo de correção de erros *backpropagation*. O *backpropagation* foi na verdade inventado e reinventado por várias pessoas, pois além de Paul Werbos, David Parker em 1982 e Yann Le Cun em 1985 publicaram descobertas similares (Parker, 1982).

O modelo de rede com múltiplas camadas chamada *multi-layer perceptron* (MLP) constitui-se de uma camada de entrada, uma ou mais camadas intermediárias ou escondidas e uma camada de saída de *perceptrons*. A arquitetura de uma rede MLP está mostrada na Figura 4.5.

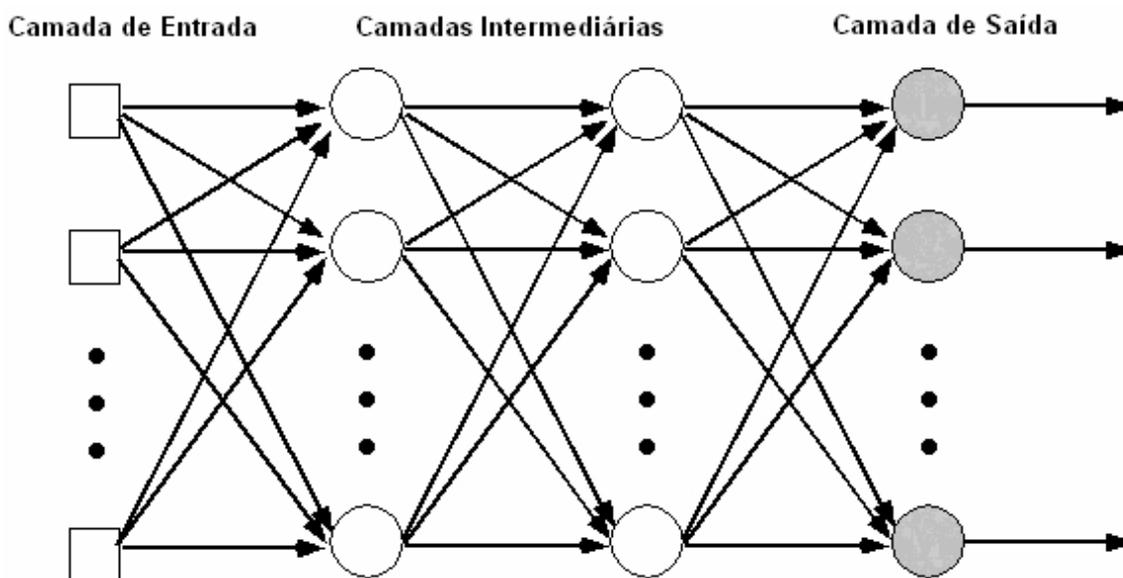


Figura 4.5 - Representação de uma Rede MLP

As redes MLP comparativamente as redes sem camadas intermediárias apresentam um poder computacional muito maior, pois redes com uma camada intermediária podem implementar qualquer função contínua (Cybenko, 1989) e redes com duas camadas intermediárias podem implementar qualquer função matemática (Cybenko, 1988). Por isto se diz que este modelo é um aproximador universal de funções.

Embora a utilização de duas ou até mais camadas possa facilitar o treinamento da rede, um número elevado de camadas não é recomendado, pois o erro medido durante o treinamento o qual é propagado da camada de saída para a última intermediária e desta para as demais se torna a cada passagem menos preciso. Ao passo que o erro é propagado cada camada intermediária recebe assim apenas uma estimativa da estimativa do erro cometido.

Braga; Ludermir e Carvalho (2000, p.55) citam alguns fatores relacionados com o número de neurônios adequados na camada intermediária:

- Números de exemplos de treinamento;
- Quantidade de ruído presente nos exemplos;
- Complexidade da função a ser aprendida e
- Distribuição estatística dos dados de treinamento.

No caso de problemas de reconhecimento de padrões é recomendado não utilizar neurônios demais para evitar a memorização dos padrões (*overfitting*), nem uma quantidade muito pequena que pode levar a rede a gastar um tempo elevado para encontrar a representação ótima.

O processo de treinamento da rede MLP-*backpropagation* apresenta dois passos. O primeiro deles é o passo para frente ou *forward pass*. Neste passo um padrão é apresentado à camada de entrada, então os neurônios desta primeira camada geram saídas que são aplicadas às entradas da camada imediatamente seguinte. Os neurônios da camada seguinte computam as suas saídas e a atividade da rede é de tal forma que se repete o processo, ou seja, flui de camada em camada até que a resposta final seja produzida pela camada de saída.

No segundo passo também chamado de passo para trás ou *backward pass*, a saída (resposta) obtida da rede é comparada com a resposta desejada, gerada por esse padrão particular que foi aplicado na entrada da rede. Os erros são propagados a partir da camada de saída até a camada de entrada, camada por camada. Este erro é então utilizado para ajustar os pesos da última camada de modo a reduzir seus erros. Em seguida, são ajustados os pesos da penúltima camada, a depender dos erros da última camada ponderados pelos pesos das conexões entre a penúltima e a última camada. Este processo se repete para as demais camadas da rede. Portanto, conforme o erro é retro-propagado e os pesos das conexões das unidades da camada de saída e das camadas intermediárias vão sendo atualizados, até que os pesos da primeira camada intermediária sejam atualizados (Rumelhart; Hinton e Williams, 1986).

O algoritmo *backpropagation* implementa o gradiente descendente no espaço dos pesos para o treinamento de uma rede *feedforward*. Neste método a rede efetua o cálculo das derivadas parciais da função $F(\vec{w}; \vec{x}_p)$ em relação ao vetor peso \vec{w} para um dado vetor de entrada \vec{x}_p . A garantia da minimização do erro na camada de saída é realizada pelo emprego destas derivadas no ajuste dos pesos.

A forma mais comum para a regra de propagação é a soma ponderada (produto interno) dos valores fornecidos como entrada e utiliza-se como pesos os valores associados com as conexões do neurônio conforme a Equação 4.11.

$$net_j = \sum_i w_{ji} x_i \quad (4.11)$$

onde w_1, w_2, \dots, w_n correspondem aos pesos das conexões, e x_1, x_2, \dots, x_n correspondem aos valores de entrada para o neurônio j .

A função de ativação é obtida a partir do valor calculado de net_j e de outra variável chamada de limiar de ativação ou *threshold*. Este limiar representa um valor o qual é adicionado negativamente às entradas e contribui no estado de ativação do neurônio. A saída gerada pela função de ativação está ilustrada pela Equação 4.12.

$$o_j = f_j(net_j - \theta_j) \quad (4.12)$$

onde θ_j é o limiar de ativação.

O algoritmo *backpropagation* é baseado no método do gradiente descendente, portanto as funções de ativação utilizadas devem ser contínuas e diferenciáveis. A mais utilizada é a função de ativação *sigmoidal* (Prechelt, 1994).

Um resumo dos passos realizados pelo algoritmo *backpropagation* estão apresentados a seguir:

Passo I: O vetor de entrada \vec{x}_p é apresentado à entrada da rede por meio de um esquema de apresentação previamente definido. Em geral a apresentação é feita seguindo-se uma ordem aleatória do conjunto de padrões de treinamento. A resposta desejada t_p correspondente a cada estímulo \vec{x}_p também é apresentada;

Passo II: São calculadas as respostas obtidas pelos neurônios da rede, camada após camada, por meio da função de ativação;

Passo III: É calculado o erro correspondente à diferença entre o vetor de saída desejada e o vetor de saída obtido por meio da função custo E_p . Caso o valor de E_p seja menor

que um limiar ε previamente estabelecido, o processo de aprendizagem termina então se vai ao passo VI. Em caso contrário, segue-se ao passo IV. Existem vários tipos de função objetivo ou de custo, contudo uma das mais utilizadas é o erro médio quadrático dado pela expressão 4.13.

$$E_p = \frac{1}{2} \sum_{j=1}^n (t_{pj} - o_{pj})^2 \quad (4.13)$$

onde o_{pj} é a resposta fornecida pela neurônio u_j e t_{pj} em correspondência a resposta desejada;

Passo IV: É realizada a atualização dos pesos das conexões. Este processo ocorre camada após camada, no sentido da camada de saída para a camada de entrada. O ajuste dos pesos é segundo a expressão 4.14.

$$\Delta w_{ij} = \eta \delta_j o_i \quad (4.14)$$

onde η representa a taxa de aprendizagem, o_i a resposta do neurônio i o qual se conecta ao neurônio j e δ_{pj} corresponde ao erro ocorrido entre a resposta desejada e a resposta fornecida pelo neurônio j . Para os neurônio da camada de saída, o valor de δ_{pj} é dado pela Equação 4.15.

$$\delta_{pj} = f'_j(\text{net}_{pj})(t_{pj} - o_{pj}) \quad (4.15)$$

Passo V: No caso dos neurônios das camadas intermediárias não há como calcular diretamente os ajustes em função do erro $(t_{pj} - o_{pj})$, portanto estes pesos são calculados em função dos erros δ_k observados nos neurônios das camadas subseqüentes da forma mostrada pela Equação 4.16.

$$\delta_{pj} = f'_j(\text{net}_{pj}) \sum_k w_{jk} \delta_{pj} \quad (4.16)$$

Após os ajustes retorna-se ao passo I.

Passo VI: Fim do algoritmo.

4.2.3 Limitações dos Modelos Tradicionais

O desenvolvimento de redes neurais construtivas (RNAC) tem como sua principal motivação a tentativa de se construir modelos para resolver problemas presentes nas redes *feedforward* tradicionais, como a rede MLP. São exemplos de limitações encontradas nas redes tradicionais o problema da topologia fixa e a lentidão no processo de aprendizagem.

O entendimento da limitação imposta pela topologia fixa está diretamente associado a como se definir, antes que ocorra o processo de treinamento, o número de neurônios e camadas intermediárias para resolver um problema específico. Estes valores estão diretamente relacionados ao modo como os neurônios realizam o mapeamento desejado entrada-saída da RNA.

Dentre as razões que levam as redes MLP-*backpropagation* a um processo de aprendizagem lento, dois podem ser considerados mais significativos: o problema da largura de passo e da mudança de objetivos. A ocorrência do problema da largura de passo está vinculada ao fato do algoritmo *backpropagation* padrão computar apenas a derivada parcial da função de erro em relação a cada peso na rede. Assim, dadas estas derivadas, o erro é minimizado a cada passo, através do método chamado de gradiente descendente. Neste método ao serem tomados passos infinitesimais a função erro eventualmente atingirá um mínimo local. Contudo, investigações práticas têm apresentado em vários casos que este mínimo local ou será o mínimo global, solução ótima, ou apenas aproxima-se de uma solução satisfatória (Fahlman e Lebiere, 1990).

Na prática não é desejável uma infinidade de passos, pois isto conduz a um processo de treinamento extremamente lento. Assim, a agilização do treinamento, em princípio, seria conseguida por meio de uma largura de passo maior, embora a obtenção de uma boa convergência não seja garantida por valores de largura de passo muito grandes. As informações necessárias para se escolher uma largura de passo razoável como, por exemplo, sobre derivadas de alta ordem da função de erro, não estão disponíveis no algoritmo *backpropagation* padrão.

É possível observar várias tentativas de otimização do algoritmo *backpropagation*. O algoritmo *Quickprop* proposto por Fahlman, é um dos mais bem sucedidos que trata do problema da largura de passo (Fahlman, 1988). Este algoritmo calcula as derivadas primeiras parciais da função de erro em relação aos pesos de maneira semelhante ao *backpropagation* padrão, contudo o *Quickprop* utiliza na atualização dos pesos um método de segunda ordem relacionado ao método de Newton no lugar do gradiente descendente simples.

Com o intuito de contribuir com melhorias no algoritmo de treinamento das RNA hoje existentes, e considerando-se as experiências vivenciadas no decorrer deste trabalho coloca-se como proposta para trabalhos futuros, capítulo 8, a utilização de um novo algoritmo baseado em distribuição de probabilidade, chamado de TALUS (Cavalcante, 1996), (Campello de Souza, 1997), (Lins, 1999), (Guerra Belém, 1999), (Cavalcanti, 2004) (Cavalcanti et al., 2004). Este algoritmo tem demonstrado ser capaz de obter de forma eficiente a direção da busca de ótimo para problemas de otimização global sem restrição.

Um fato relacionado ao problema da mudança de objetivos é que durante o processo de aprendizagem, cada neurônio da rede busca tornar-se um detector de características. Estes neurônios tomam para si alguma função importante no processo como um todo. A complexidade do processo aumenta ao passo em que todos os neurônios estão se modificando ao mesmo tempo. Considerando que os neurônios das camadas escondidas não se intercomunicam, e portanto recebem apenas os sinais de suas entradas e o de erro, o qual por sua vez vem retropropagado pelas unidades de saída da RNA. O problema que cada neurônio tenta solucionar está vinculado com este sinal de erro retropropagado, contudo o mais comum é que ele mude constantemente. Assim, o que se observa é um cenário complexo o qual requer muito tempo de processamento para estabilizar-se, o que não ocorreria se cada neurônio assumisse um papel específico, o que conduziria a uma solução do problema de forma rápida e objetiva.

Experimentos realizados com o algoritmo *backpropagation* têm revelado que, na medida em que cresce o número de neurônios nas camadas escondidas da RNA o tempo de treinamento sofre um grande retardo (Fahlman e Lebiere, 1990). Parte desse retardo

é atribuída a certa atenuação ocorrida no sinal de erro, quando ele é propagado pelas camadas da rede. Pode-se adicionalmente considerar nesta análise o efeito da mudança constante de objetivo entre as unidades. Fahlman (1990) menciona o chamado efeito “rebanho” como uma manifestação comum do problema de mudança de objetivos.

Um caso exemplo do efeito “rebanho” ocorre quando os neurônios da rede necessitam resolver dois problemas distintos A e B. Considerando-se que todos os neurônios têm capacidade semelhante para lidar com cada um dos problemas, se A gerar um sinal de erro mais coerente que B, observa-se uma tendência de todos os neurônios direcionarem a solução para A, desconsiderando o problema B. Assim, uma vez que o problema A tenha sido solucionado, os neurônios concentram-se na outra fonte de erro, o problema B, e desta forma o problema A reaparece. Pode ocorrer que todos os neurônios tratem ambos os problemas ao mesmo tempo, todavia até que isto venha a ocorrer leva-se um tempo considerável.

Uma forma de se enfrentar o problema da mudança de objetivos é permitir que somente uma quantidade reduzida de neurônios da RNA mude simultaneamente, permanecendo os demais neurônios constantes (Fahlman e Lebiere, 1990).

4.3 MODELOS DE REDES NEURAIAS CONSTRUTIVAS

Apesar de ser possível identificar várias aplicações de sucesso onde são empregadas RNAs com topologia fixa, na seção anterior 4.2.3 foram citadas algumas limitações associadas a elas. Assim, as RNAC têm surgido como uma forma alternativa aos modelos tradicionais de arquiteturas fixas. Na sessão a seguir 4.3.1 são apresentadas as características das redes construtivas e alguns fatores motivadores.

4.3.1 Características Gerais das Redes Neurais Construtivas

A arquitetura é uma característica importante de uma RNA, pois ela pode influenciar de forma significativa o desempenho final da rede para uma dada aplicação. Uma forma de

classificar a arquitetura de uma RNA é de acordo com a possibilidade de mudanças no seu tamanho durante o processo de treinamento. Assim, ela pode ser classificada em estática ou dinâmica. Uma rede é dita de arquitetura estática quando o número de camadas e unidades de processamento permanece constante desde a concepção da RNA até a finalização do seu treinamento. Por outro lado, uma RNA de arquitetura dinâmica tem como característica principal um processo de mudanças no tamanho durante o treinamento.

Uma das abordagens utilizadas na otimização da arquitetura de uma rede utiliza as chamadas redes neurais construtivas. Nas RNAC a otimização da topologia ocorre, pois existe a possibilidade de introduzir e retirar elementos no decorrer do processo de treinamento. Nos algoritmos de aprendizado construtivos, também chamados de *growing*, a arquitetura final da rede é definida com a utilização de uma rede inicialmente pequena, na qual são adicionadas novas unidades no decorrer do treinamento (Braga; Ludermir e Carvalho, 2000).

Sabe-se que os algoritmos de ajuste de pesos tipicamente procuram por um vetor de peso solução que satisfaça algum critério de desempenho escolhido. Para que esta solução seja realmente satisfatória, ela deve pertencer ao espaço de pesos no qual ela está sendo realizada e o procedimento de busca utilizado deve ser capaz de atingi-la. O procedimento se resume a um problema de tentativa e erro, a menos que haja um suporte à tarefa, na forma de conhecimento específico sobre o problema, o que conduzirá à escolha de uma topologia adequada. A utilização de algoritmos construtivos pode potencialmente fornecer uma forma de superar este problema por meio de uma extensão da busca de uma solução ao espaço de topologias da rede.

É desejável que a complexidade da rede em termos, por exemplo, do número de neurônios, conexões e outros parâmetros seja proporcional à complexidade intrínseca da aplicação para a qual ela se destina. Assim, torna-se desejável que os algoritmos de aprendizagem construam redes contemplando esta proporcionalidade citada na frase anterior. Os algoritmos construtivos, potencialmente, podem encontrar redes próximas à configuração mínima que dê resposta correta sobre um determinado conjunto de dados.

Os algoritmos construtivos podem ser distinguidos por meio de características como: quando adicionar um novo neurônio na rede; conectividade do neurônio inserido; escolha quanto a representação das entradas (binárias, bipolares ou reais); inicialização da unidade a ser adicionada; como treinar o neurônio adicionado entre outras.

4.3.2 Redes Neurais Compostas por Blocos de Regressões Sigmóides Não-lineares

As Redes Compostas por Blocos de Regressões Sigmóides Não-Lineares (NSRBN) foram desenvolvidas por Valença em 1999 (Valença, 1999). Dentre as características principais das redes NSRBN, duas merecem destaque: a primeira diz respeito à sua estrutura compacta, por possuir apenas uma camada escondida e a segunda por utilizarem um algoritmo construtivo baseado em um método de aprendizado construtivo, o qual é responsável pela definição de sua complexidade ótima.

A principal motivação para o surgimento das redes NSRBN, foi a possibilidade de se empregar um método sistemático por intermédio da abordagem construtiva, de forma a estabelecer um mecanismo natural para o crescimento da rede e simultaneamente selecionar o modelo de complexidade ótima. Somando-se aos aspectos já citados, estas redes apresentam grande capacidade de mapeamento de uma rede neural de alta ordem, com uma simples camada escondida, e ainda limitam o aumento explosivo no número de termos requeridos (Valença, 2005).

A estrutura das redes NSRBN é montada por meio de um algoritmo construtivo. Este algoritmo parte de uma pequena estrutura, até atingir uma de tamanho ótimo, segundo critérios estatísticos de seleção adequados (vide Apêndice C).

A forma de construção no modelo NSRBN é por meio de blocos de polinômios homogêneos, empregando-se uma função de ativação adequada para os neurônios escondidos, como por exemplo, a tangente hiperbólica. Em relação aos neurônios de saída, eles têm uma ativação mais complexa, uma vez que executam uma regressão logística não-linear em relação à saída dos neurônios escondidos. Portanto, as redes

NSRBN podem aproximar uniformemente qualquer função contínua definida em um conjunto compacto limitado (vide Apêndice A), e também trabalhar com a realização de funções booleanas, o que não é possível com os algoritmos tradicionais do *Group Method Of Data Handling* (GMDH), sejam combinatorial ou *multilayer* (Valença,1999).

A construção de um modelo combinatorial tem como primeiro passo dividir a função $f(x)$ em blocos compostos de termos de igual grau, conforme mostrado na Figura 4.6.

$$f(x) = f_1(x) + f_2(x) + \dots + f_d(x) \tag{4.17}$$

O bloco de grau p ($p = 1, 2, \dots, d$), de acordo com o polinômio de Kolmogorov-Gabor (teorema de Weierstrass) (Gabor et al., 1961) (Madala e Ivakhnenko, 1994) (Pinkus, 2000), pode ser expresso como uma soma de produtos das entradas, isto é:

$$f_p(x) = \sum_{i_1=1}^N \sum_{i_2=i_1}^N \dots \sum_{i_N=i_1}^N a_p(i_1, i_2, i_3, \dots, i_N) \cdot x_1^{q_{p1}(i_1)} \cdot x_2^{q_{p2}(i_2)} \dots x_N^{q_{pN}(i_N)} \tag{4.18}$$

em que, $q_{p1}(i_1) + q_{p2}(i_2) + q_{p3}(i_3) + \dots + q_{pN}(i_N) = p$ e $p \geq q_{pm}(\cdot) \geq 0$ para $m = 1, 2, 3, \dots, N$.

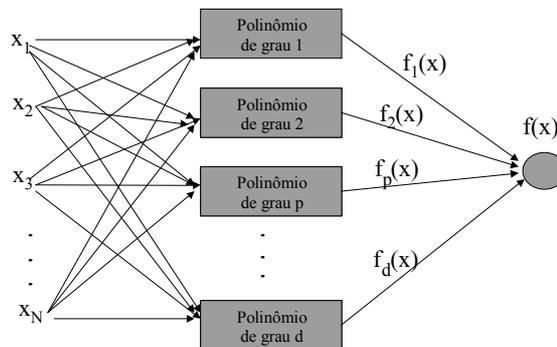


Figura 4.6 - Arquitetura combinatorial

Uma maneira simples de realizar o mapeamento de cada um dos polinômios homogêneos de grau p é fazer todos os produtos das variáveis de entrada, buscando-se aproximar diretamente a hipersuperfície não-linear, contudo isto não se mostra uma solução prática, a menos que d e N sejam relativamente pequenos (modelo combinatorial). Portanto, para se obter um modelo eficiente, torna-se necessário que

este modelo possua funções de ativação que sejam capazes de realizar vários termos dos polinômios homogêneos de forma otimizada, ou seja, sem que ocorra um crescimento explosivo no número de parâmetros necessários para ajustar a função desejada. Para tal objetivo, as NSRBNs têm por base o algoritmo combinatorial (composta pela soma de blocos) com uma estrutura de blocos similar a uma rede *Multilayer Perceptron* (MLP), sendo que, diferentemente destas, os neurônios de saída realizam a técnica estatística de regressão logística não-linear (Hosmer e Lemeshow, 1989), com relação aos neurônios escondidos.

Uma rede NSRBN pode ser definida como uma rede combinatorial composta pela soma de p blocos ($p=1, \dots, d$) com estrutura similar a uma MLP, na qual as unidades escondidas desses blocos realizam uma regressão sigmóide linear das entradas (tal qual uma MLP), e as unidades de saída realizam uma regressão do tipo sigmóide não-linear das unidades escondidas (tal qual uma regressão logística não-linear). Essa nova estrutura criada é um aproximador universal com forte capacidade de mapeamento não-linear, vide teoremas Apêndice A. Emprega-se aqui o termo sigmóide, no sentido de qualquer função sigmoidal, como uma generalização da regressão logística clássica.

Cada bloco é composto de apenas uma camada escondida, onde as Nh unidades escondidas que o compõem têm função de ativação do tipo:

$$f(x) = (\sigma_{net(h)} + \theta_h) \quad (4.19)$$

onde: $net(h)$ é a entrada líquida para cada neurônio escondido, h é um número inteiro que representa as unidades escondidas em cada bloco (variando de 1 até a ordem p do bloco), θ_h é a tendência e $\sigma_{net(h)}$ é, por exemplo, a função tangente hiperbólica dada por:

$$\sigma_{net(h)} = \frac{e^{net(h)} - e^{-net(h)}}{e^{net(h)} + e^{-net(h)}} \quad (4.20)$$

É possível utilizar outras funções, como a função recíproca representada na Equação 4.21, isto é,

$$f_r(x) = \frac{net(h)}{1 + |net(h)|} \quad (4.21)$$

O número de neurônios em cada bloco é definido pelo valor de p , o qual corresponde ao grau do bloco, isto é: o bloco de grau 1 ($p=1$) tem $h=1$, logo um neurônio escondido; o bloco de grau 2 ($p=2$) terá 2 neurônios escondidos ($h=1$ e $h=2$) e assim sucessivamente. Portanto, o bloco de grau p terá $h=p$ e conseqüentemente p neurônios escondidos.

Na Figura 4.7 está apresentada a estrutura proposta para cada bloco, na qual mostra-se o mapeamento para um bloco de grau p (a ordem dos neurônios escondidos h variará de 1 até p), para i variáveis de entrada (onde $i = 1, \dots, N$) e uma única unidade de saída.

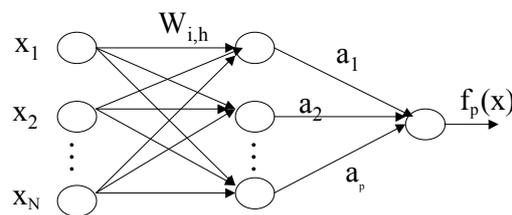


Figura 4.7 - Arquitetura de um bloco polinomial

O bloco apresentado na Figura 4.7 é uma rede *feedforward* completamente conectada com uma única camada escondida. A diferença básica desse bloco de grau p para uma rede MLP é que os neurônios de saída são expressos como uma regressão sigmóide não-linear das unidades escondidas.

Seja $\mathbf{x} = [1, x_1, x_2, x_3, \dots, x_N]^T$ o vetor coluna das variáveis de entrada, no qual o elemento x_i denota um elemento qualquer do vetor \mathbf{x} . As entradas são ponderadas pelos vetores de pesos $\mathbf{w} = [w_{0h}, w_{1h}, w_{2h}, \dots, w_{Nh}]^T$, $h=1, \dots, p$ nos quais p é a ordem do bloco. Portanto, utilizando-se a função de ativação estabelecida na Equação 4.19, a saída de cada unidade escondida, $f(net(h))$, desse bloco de grau p pode ser escrita como:

$$f(net(h)) = (\sigma_{net(h)} + \theta_h) \tag{4.22}$$

onde: $net(h) = \mathbf{w}_h^T \mathbf{x}$, ou seja, $net(h) = \sum_{i=1}^N w_{ih}x_i + w_{0h}$, $h=1,2, \dots, p$.

Portanto, a unidade de saída (que realiza uma regressão logística não-linear) desse bloco

pode ser representada pela Equação 4.23:

$$f_p(x) = a_1(f(\text{net}(1)))^1 + a_2(f(\text{net}(2)))^2 + a_3(f(\text{net}(3)))^3 + \dots + a_p(f(\text{net}(p)))^p \quad (4.23)$$

sendo $\mathbf{a}=[a_1, a_2, a_3, \dots, a_p]^T$ o vetor dos pesos das unidades escondidas para a unidade de saída.

Substituindo-se $f(\text{net}(h))$ Equação 4.22 em $f_p(x)$ Equação (4.23) resulta:

$$f_p(x) = a_1(\sigma_{\text{net}(1)} + \theta_1) + a_2(\sigma_{\text{net}(2)} + \theta_2)^2 + \dots + a_p(\sigma_{\text{net}(p)} + \theta_p)^p \quad (4.24)$$

ou ainda

$$f_p(x) = \sum_{h=1}^p a_h(\sigma_{\text{net}(h)} + \theta_h)^h \quad (4.25)$$

onde h representa a ordem dos neurônios na camada escondida, ($h=1, \dots, p$), do bloco de grau p . A arquitetura de uma rede NSRBN vista na Figura 4.6 é composta pela soma de blocos com grau variando de 1 a d . Porém, nesta descrição está sendo utilizado o bloco p para representar qualquer um desses blocos, conforme mostrado na Figura 4.7. Portanto, usando-se a arquitetura descrita nos parágrafos anteriores, uma função desconhecida f em R^N pode ser aproximada por uma rede NSRBN, conforme Equação 4.26:

$$f(x) = \sigma_{\text{net}(0)} \left(\sum_{p=1}^d f_p(x) \right) \quad (4.26)$$

onde: $\sigma_{\text{net}(0)}$ é uma função de ativação adequada para a unidade de saída.

As redes NSRBN foram concebidas de tal forma que sua construção é feita de forma incremental, de modo que se pode representá-la em uma forma similar à de um bloco (Equação 4.24), conforme apresenta-se a seguir. Considere-se o passo do algoritmo para o qual é adicionado à rede o bloco de ordem d , f_d , tendo-se anteriormente adicionado os blocos f_0, f_1, \dots, f_{d-1} . Fazendo-se $f_0 \equiv 0$, uma função f pode ser representada por:

$$f(x) = \sigma_{net(o)} \left(\sum_{p=0}^{d-1} f_p(x) + f_d(x) \right) \quad (4.27)$$

na qual os pesos dos blocos f_0, f_1, \dots, f_{d-1} são fixados no momento imediatamente anterior a adição do bloco de ordem d, f_d .

Portanto, f_x pode ser reescrita na forma da Equação 4.28:

$$f(x) = \sigma_{net(o)} \{ c_0 + c_1(\sigma_{net(i)} + \theta_1) + c_2(\sigma_{net(2)} + \theta_2)^2 + \dots + c_d(\sigma_{net(d)} + \theta_d)^d \} \quad (4.28)$$

na qual o vetor $\mathbf{c} = [c_1, c_2, \dots, c_d]$ representa os pesos das unidades escondidas para as unidades de saída e a constante c_0 representa todos os blocos congelados até a ordem $d-1$. A Equação 4.25 é, portanto, semelhante a uma regressão logística não-linear quando a função de ativação utilizada é a sigmóide logística tradicional.

A potencialidade dessas redes está fundamentada nas características impostas pela regressão sigmóide não-linear na unidade de saída.

O algoritmo construtivo das NSRBNs tem como base um método de aprendizagem construtiva (também chamado de auto-organização na teoria das redes GMDH). O princípio construtivo pode ser formulado da seguinte forma: quando um modelo incrementa gradualmente sua complexidade, certos critérios, que são chamados de critérios de seleção ou função objetivo e que têm a propriedade de complemento externo, passam por um valor mínimo global. O encontro deste mínimo indicará a existência de um modelo de complexidade ótima.

Assim, usando-se essa arquitetura proposta uma função qualquer $f(x)$ pode ser aproximada pelo uso direto de uma rede NSRBN por dividir $f(x)$ em blocos homogêneos de ordem 1 até o bloco de ordem d . Desse modo, a função total $f(x)$ (polinômio completo) pode ser aproximada pela soma das funções parciais $f_p(x)$, $p=1, \dots, d$ (polinômios homogêneos), de forma incremental o que permite ser representada matematicamente por:

$$f(x) = g(\hat{f}_1(x) + \hat{f}_2(x) + \dots + \hat{f}_p(x) + \dots + f_d(x)) \quad (4.29)$$

É necessário lembrar que o modelo de rede ora apresentado tem como princípio manter fixos os pesos de todos os blocos anteriores ao atualmente otimizado, de tal forma que o número de pesos a ser otimizado a cada grau de complexidade não aumente de forma explosiva. Este problema ocorre no algoritmo combinatorial tradicional.

A forma como atua o algoritmo de abordagem construtiva está descrita a seguir.

Seja p um passo do algoritmo para qual o bloco $f_p(x)$ é adicionado à rede, então uma função qualquer $f(x)$ é sucessivamente aproximada por:

$$f(x) = g(f_1(x)) \quad (4.30)$$

$$f(x) = g(\hat{f}_1(x) + f_2(x)) \quad (4.31)$$

...

...

...

$$f(x) = g(\hat{f}_1(x) + \hat{f}_2(x) + \dots + f_p(x)) \quad (4.32)$$

em que os pesos dos blocos $\hat{f}_{p-1}(x)$ são fixados quando o bloco de ordem p é adicionado à rede e g representa uma função de ativação adequada para a unidade de saída.

Ao observar a Equação 4.28, percebe-se que o problema consiste em resolver um sistema de equações não-lineares. Portanto, técnicas de segunda ordem tradicionais como gradientes conjugados (Bishop, 1995) e Levenberg-Marquardt (Bishop, 1995) podem ser utilizadas com grande ganho de eficiência. Contudo, testes realizados com gradiente descendente mostraram que essa técnica pode ser utilizada. As aplicações realizadas com ela forneceram resultados satisfatórios em todos os casos analisados (Valença, 1999).

Embora o algoritmo construtivo das NSRBN fixe os pesos dos blocos anteriormente adicionados na estrutura em construção, ele introduz uma forma de compensação. Isto é realizado com a colocação de uma tendência adicionada a função

tangente hiperbólica, de tal forma que cada bloco de sigmóide realize não apenas valores do bloco atual como também termos dos blocos anteriores. Este procedimento permite uma correção nos pesos e no número de unidades escondidas. Uma descrição na forma de passos do algoritmo das redes NSRBN está apresentado no Apêndice B.

4.4 MODELOS DE REGRESSÃO LOGÍSTICA E MÚLTIPLA

4.4.1 Regressão Logística

Pode-se dizer de forma geral que os modelos de regressão consistem de uma variável dependente, chamada de variável de saída ou resposta, e de uma ou mais variáveis independentes, as quais são chamadas de projeto ou explicativas, além de um termo chamado de erro. Uma forma de expressar a relação entre estes termos pode ser dada por:

$$Y = f(X_1, X_2, \dots, X_p) + \text{Erro} \quad (4.33)$$

A regressão logística é uma técnica que permite prever uma resposta discreta a partir de um conjunto de variáveis que podem ser contínuas, discretas, dicotômicas ou uma mistura destas categorias, pois não é feita suposição sobre a distribuição das variáveis independentes (Hosmer e Lemeshow, 1989) e (Agresti, 1996). Na regressão logística a variável de saída é dicotômica, ou seja, ela pode assumir o valor 1 com probabilidade de sucesso θ , ou o valor 0 com probabilidade de falha $1-\theta$. Este tipo de variável também é chamada de Bernoulli ou binária.

O relacionamento entre a variável resposta e as variáveis explicativas na regressão logística é não linear. A expressão matemática da função pode ser observada a seguir:

$$\theta = \frac{e^{g(x)}}{1 + e^{g(x)}} \quad (4.34)$$

É comum usar $g(x)$ uma função linear dos parâmetros, contudo $g(x)$ pode ser qualquer função.

Fazendo-se $g(x) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i$ temos:

$$\theta = \frac{e^{(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i)}}{1 + e^{(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i)}} \quad (4.35)$$

onde:

α = constante da equação

β = o coeficiente da variável explicativa

Uma outra forma alternativa de expressar pode ser vista em (4.36).

$$\text{logit}[\theta(x)] = \ln \left[\frac{\theta(x)}{1 - \theta(x)} \right] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i \quad (4.36)$$

Na Tabela 4.3 está apresentada uma comparação entre termos utilizados na abordagem estatística e com redes neurais (Tu, 1996).

Tabela 4.3 - Semelhança entre termos utilizados em RNAs e em Estatística

Termo em Redes Neurais	Termo em Estatística
Entrada da Rede	Variável Independente, Explicativa, Auxiliar, Regressora
Saída da Rede	Variável Dependente, Resposta
Pesos das Conexões	Coefficientes de Regressão
Peso do Viés	Intercepto-y
Erro	Resíduos
Aprendizado, Treinamento	Estimação de Parâmetros
Padrão de Treinamento	Observação

4.4.2 Regressão Múltipla

No modelo linear de regressão múltipla (Draper, 1981), a relação entre a variável dependente e as variáveis explicativas assume a forma:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (4.37)$$

Onde a variável resposta Y está numa escala contínua (de razão). Embora variáveis explicativas categóricas possam ser usadas também na regressão múltipla, tipicamente,

as variáveis independentes são também contínuas. O método de estimação dos parâmetros utilizado é o dos mínimos quadráticos. Podem ser usadas transformações nas variáveis explicativas, e o modelo apresenta a forma:

$$Y = \beta_0 + \beta_1 f_1(X_1) + \beta_2 f_2(X_2) + \dots + \beta_p f_p(X_p) + \varepsilon \quad (4.38)$$

As transformações introduzidas servem para garantir a linearidade. Para um modelo clássico, é imperioso que a variável Y seja gaussiana, isto cria a necessidade em alguns casos de também fazer-se uma transformação na variável resposta de forma a ser garantida esta condição. Um tipo de transformação realizada para este fim é por meio da aplicação do logaritmo neperiano.

Este modelo permite fazer a estimação do valor da variável resposta em função dos valores das variáveis explicativas, apresentando-se o resultado em termos de uma média e um intervalo de confiança ao redor deste valor. O modelo é considerado ajustado quando os testes estatísticos sobre os valores estimados dos parâmetros forem significativos (teste t de Student) e a análise de variância também for significativa (teste F). Este modelo é considerado tanto melhor quanto maior for o coeficiente de determinação, o qual expressa o percentual da variabilidade observada na variável resposta que é explicado pela equação da regressão. Assim, uma vez atendida a condição de ser estatisticamente significante, o modelo será tanto mais preciso na estimação do valor da variável resposta para um dado indivíduo, ou seja, mais estreito o intervalo de confiança ao redor do valor estimado, quanto maior for o seu coeficiente de determinação.

O que é desejado com a utilização de modelos de regressão múltipla é identificar o máximo de fatores explicativos, que possam esclarecer a variabilidade da resposta estimada empiricamente, procurando-se deixar o mínimo de variabilidade atribuída ao erro. Esta variabilidade imputada ao erro é chamada variabilidade não-explicada pela equação de regressão ou explicada pela variância.

4.5 SIMULADORES UTILIZADOS

Com o propósito de documentar fielmente a pesquisa realizada são apresentados a seguir informações sobre as ferramentas de simulação utilizadas. No desenvolvimento deste trabalho vários cálculos, testes e simulações foram efetuados empregando-se os aplicativos citados nas subseções 4.5.1, 4.5.2 e 4.5.3 a seguir. Contudo, é preciso destacar que estas tarefas poderiam ter sido executadas por intermédio de outras soluções equivalentes, tanto disponíveis comercialmente como desenvolvidas especificamente para estes fins.

4.5.1 Neural Networks Toolbox - MatLab

O *neural networks toolbox* é uma ferramenta da Math Works para resolver problemas com redes neurais artificiais, e consiste de uma coleção de funções construídas no Matlab, o qual por sua vez, é um ambiente de computação numérica (Demuth e Beale,1996). O objetivo desta ferramenta é oferecer um ambiente para criar e simular redes de vários tipos e arquiteturas. Nos experimentos realizados neste trabalho, o *neural networks toolbox* 4.0 foi utilizado para treinamento de redes *feedforward* MLP.

Um simulador específico para testar redes MLP foi desenvolvido com o *neural networks toolbox* 4.0. A tela com as diversas possibilidades de pré-processamento dos dados, ajustes de parâmetros e configurações da rede, bem como métricas para avaliação da rede estão apresentadas na Figura 4.8.

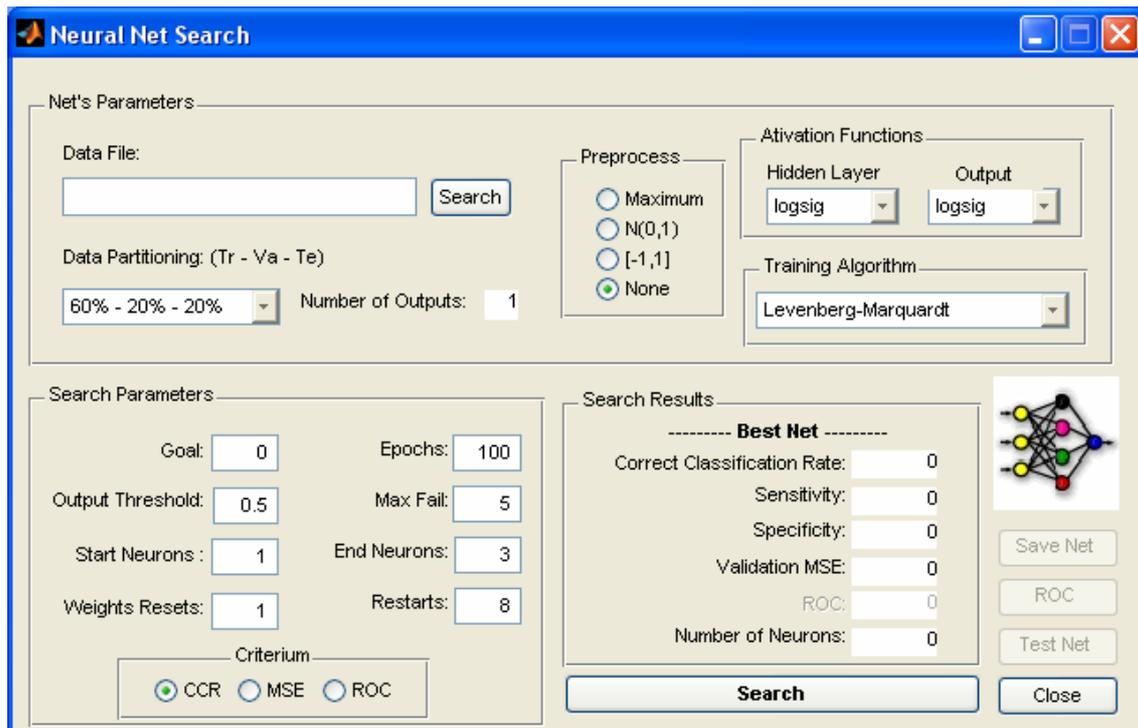


Figura 4.8 - Tela do Simulador de redes MLP

4.5.2 Neuro Hidro

O Neuro Hidro é um simulador de redes neurais construtivas do tipo NSRBN, o qual na sua versão acadêmica permite aos pesquisadores desenvolver, simular e analisar modelos com este tipo de rede (Valença, 2005). O Neuro Hidro possui interface bastante amigável onde é possível controlar vários parâmetros da arquitetura da rede e do processo de treinamento, bem como visualizar os resultados tanto na forma de valores como gráfica. A tela do Neuro Hidro pode ser vista na Figura 4.9.

The image shows the 'TREINAMENTO' (Training) tab of the Neuro Hydro Simulator. It is divided into several sections:

- DADOS DE ENTRADA** (Input Data): Includes fields for 'NÚMERO DE ENTRADAS' (Number of Inputs) and 'NÚMERO DE SAÍDAS' (Number of Outputs).
- CONJUNTOS** (Sets): Includes fields for '(%) TREINAMENTO' (Training), '(%) VALIDAÇÃO CRUZADA' (Cross-validation), and '(%) VERIFICAÇÃO' (Verification).
- PARÂMETROS DO ALGORITMO** (Algorithm Parameters): Includes 'Taxa de Aprendizagem (Alfa)(%) (0 < Alfa < 100)' (Learning Rate), 'Semente para Geração' (Seed), and 'Momentum (beta)(%) (0 < Beta < 100)' (Momentum).
- CRITÉRIOS DE PARADA** (Stopping Criteria): Includes 'EPOCASMIN' and 'EPOCASMAX'.
- DADOS DE SAÍDA** (Output Data): Includes radio buttons for 'Sem Transformação' (No Transformation) and 'Com Transformação' (With Transformation).
- Erros** (Errors): Includes 'Erro Percentual Médio Absoluto' (Mean Absolute Percentage Error) and 'Erro Padrão' (Standard Error).
- Erro Mínimo** (Minimum Error): A field at the bottom left.
- Gráficos** (Plots): Two plots are shown: 'Conjunto de Validação Cruzada' (Cross-validation Set) and 'Conjunto Total' (Total Set). Both plots show 'Erro Médio Quadrático' (Mean Squared Error) on the y-axis and have a legend for 'Ocorrida' (Observed) and 'Calculada' (Calculated).

Figura 4.9 - Tela do Simulador Neuro Hidro

4.5.3 Statistica

O Statistica é uma ferramenta para análise estatística de dados da StatSoft que permite realizar desde estatística descritiva, análise não-paramétrica, análise de cluster, análise de componentes principais e diversos tipos de regressão lineares e não lineares entre outros. Adicionalmente, o Statistica incorpora interface de dados amigável, recursos gráficos e facilidades de importação e exportação de dados. Na maior parte das análises estatísticas realizadas neste trabalho foi utilizado o Statistica 6.

5. O APOIO AO DIAGNÓSTICO DA HIPERTROFIA VENTRICULAR ESQUERDA

**“If you can look into the seeds of time,
And say which grain will grow and which will not,
Speak then to me,”** Shakespeare (1564-1616), Macbeth
“Se puderdes olhando as sementes, e dizer qual irá ou não germinar, então, falai
comigo,...” (Tradução livre)

As RNA têm sido utilizadas na área de classificação em cardiologia e vêm mostrando resultados promissores. Uma grande parte das abordagens propostas faz uso da conhecida rede *Multi-Layer Perceptron* (MLP) treinada com o algoritmo *Backpropagation* (Rumelhart; Hinton e Williams, 1986). Contudo, existem outros tipos de redes neurais que podem ser testados para o mesmo fim, em particular as redes construtivas do tipo NSRBN, com resultados que podem ser comparativamente superiores dos obtidos por abordagens tradicionais aplicadas aos mesmos padrões de entrada.

Durante a etapa de prospecção de modelos foi desenvolvido um simulador na plataforma do MatLab para redes MLP. Contudo, todos os resultados obtidos foram inferiores ao desempenho da regressão, portanto partiu-se na busca de modelos com redes construtivas do tipo NSRBN.

Neste capítulo são apresentados os resultados de um sistema de apoio a decisão para diagnóstico na área de cardiologia, particularmente no diagnóstico da Hipertrofia Ventricular Esquerda e na estimação do Índice de Massa Ventricular Esquerda. Os resultados foram comparados com aqueles obtidos por regressões Logística e Múltipla respectivamente.

5.1 DESCRIÇÃO DA BASE DE DADOS

A base de dados utilizada neste capítulo foi gerada por Chaves Júnior (2000). A coleta de dados foi realizada em uma amostra de indivíduos normotensos e hipertensos dentro de um estudo de natureza observacional e constituído de uma série de casos. Este estudo atendeu aos padrões éticos com protocolo de estudo e termo de consentimento aprovados pelos comitês de ética em pesquisas da Universidade Federal de Pernambuco (UFPE) e Universidade de São Paulo (USP).

A população estudada consistiu de 119 homens e mulheres, todos matriculados no local de realização do estudo, que foi a Clínica de Hipertensão do Serviço de Cardiologia do Hospital das Clínicas da UFPE e voluntários normotensos. A amostra buscou abranger o maior número possível de perfis de indivíduos. Portanto, ela foi subdividida em termos de sexo, idade, status de hipertenso e presença ou não de hipertrofia ventricular esquerda (HVE) ao ecocardiograma. Como característica adicional da amostra todos os hipertensos usavam droga anti-hipertensiva e 50% dos indivíduos eram portadores de HVE ao ecocardiograma.

No estudo realizado que gerou a base de dados foram realizados três exames de Monitoração Ambulatorial da Pressão Arterial de 24h (MAPA) em intervalos curtos de 8 a 15 dias. Para determinação do tamanho da amostra levou-se em consideração os seguintes valores:

- Diferença a ser detectada maior ou igual a 6 mmHg;
- Valor do risco (α) = 0,05;
- Poder do teste ($1-\beta$) = 0,80 e
- Variabilidade da pressão arterial sistólica dada pelo coeficiente de variação=11%.

Considerando-se estes valores especificados e também que as pressões têm uma distribuição gaussiana, foi calculado o tamanho da amostra (Bowker e Lieberman, 1972). O valor calculado resultou em um n de 100 para um teste bicaudal. Neste caso o poder do teste é de 80%. De forma a serem absorvidas perdas de pontos na amostra, adicionou-se uma margem de segurança de 20% o que resultou em 119 indivíduos para amostra. Contudo, dos 119 indivíduos selecionados 101 ficaram na amostra final.

Portanto, na amostra final havia 73% de hipertensos (leves, moderados e graves em uso de droga anti-hipertensiva).

Os critérios de inclusão adotados no estúdio em questão que gerou a amostra foram:

- Homens e mulheres entre 21 e 81 anos de idade;
- Indivíduos normotensos e hipertensos em uso de droga anti-hipertensiva;
- 50% dos indivíduos deveriam ser portadores de algum grau de hipertrofia ventricular esquerda concêntrica ao ecocardiograma;
- Os indivíduos deveriam manifestar de forma espontânea o desejo em participar da pesquisa, após o convite verbal e minuciosa explicação do estudo;
- O termo de consentimento pós-informado devia ser lido e assinado por cada indivíduo ou seu responsável direto autorizando sua participação voluntária no estudo.

Como critérios de exclusão no estudo foram utilizados:

- Portadores de hipertensão arterial maligna e secundária;
- Portadores de valvopatia aórtica, miocardiopatia dilatada, cardiopatia hipertrófica idiopática e/ou qualquer doença representada por obstrução da via de saída do ventrículo esquerdo;
- Portadores de insuficiência coronariana aguda ou subaguda caracterizada por angina do peito ou infarto do miocárdio nos últimos seis meses;
- Portadores de insuficiência cardíaca congestiva, classes funcionais III e IV (*New York Heart Association* – NYHA);
- Doença encefálica (AVC) aguda ou subaguda nos últimos seis meses;
- Arritmia cardíaca tipo *flutter* ou fibrilação atriais e extrassístolia ventricular freqüente (superior a 10 extrassístoles/minuto);
- Portadores de qualquer doença crônica invalidante.

Na consulta médica cada indivíduo foi entrevistado pelo médico e respondeu a um questionário e foi submetido a um exame físico, o que permitiu coletar: dados biográficos (idade, sexo, raça, procedência e outros); antropométricos (peso, altura, medida da cintura e do quadril, índice de massa corpórea); do interrogatório sintomatológico; sobre antecedentes pessoais e familiares e hábitos de vida (ingestão

informada de bebida alcoólica, de sal e tabagismo); Eletrocardiograma convencional de repouso (12 derivações); pressões arteriais com o paciente na posição sentada por meio do esfigmomanômetro de mercúrio (segundo a *British Hypertension Society* - BHS).

Adicionalmente também foram realizados exames de ecocardiograma, MAPA-24h além da série bioquímica e hormônios. Particularmente, a reprodutibilidade da MAPA-24h já foi verificada o que vem incentivar a utilização dos dados com ela obtidos (Chaves Jr, 2005). No Apêndice D Tabela D.1 e glossário da Tabela está mostrada a estatística descritiva das variáveis contidas na amostra.

5.2 EXPERIMENTOS PARA CLASSIFICAÇÃO DE HVE

Nos experimentos procurou-se abranger e testar as diversas variáveis existentes na base de dados, tanto de forma isolada como combinadas. O conjunto total de variáveis possíveis de serem exploradas ultrapassa 2000. Foi utilizado como critério de seleção dos modelos investigados aqueles que produzissem resultados satisfatórios em termos de: classificação correta de padrões (CCR), sensibilidade, especificidade, *odds ratio*, área sob a curva ROC (*Receiver Operating Characteristic*) bem como utilizassem variáveis de fácil obtenção e baixo custo financeiro.

Outro objetivo perseguido foi selecionar modelos que, sem negligenciar o desempenho final, apresentassem menor custo computacional e complexidade, em termos do número de blocos na configuração após a fase de treinamento da NSRBN bem como utilizaram um menor número de variáveis de entrada. Nos resultados obtidos com as redes MLP também buscou-se menor custo computacional e complexidade, em termos do número de neurônios presentes na rede.

Os resultados obtidos com estes modelos conexionistas foram comparados com outros gerados com a abordagem clássica estatística (Regressão Logística). Em todos os casos as redes tiveram como variável de saída os valores correspondentes a ter ou não ter hipertrofia ventricular esquerda, HVE=1 e HVE=0 respectivamente. Neste estudo as medidas obtidas com o ecocardiograma foram utilizadas como padrão ouro e os valores

da variável de saída HVE foram obtidos de acordo com o critério recomendado pela *American Society of Echocardiography* que é de 134g/m^2 para homens ou 110g/m^2 para mulheres (Hammond et al., 1986).

A análise de desempenho foi realizada utilizando-se tanto o simulador de redes MLP mostrado na subseção 4.5.1, como também por meio de um *software* simulador de redes NSRBN mostrado na subseção 4.5.2 (Valença, 2005), os quais permitiram escolher e testar uma grande variedade de:

- Tamanhos dos conjuntos de treinamento, validação e teste;
- Valores dos parâmetros como taxa de aprendizado, momento e semente;
- Critério de parada para o treinamento.

Em função da característica de cada simulador estava disponível o mínimo erro médio quadrático (MSE) no conjunto de validação, o erro percentual absoluto médio (MAPE) e o erro padrão (SE) no conjunto de teste. A taxa de padrões classificados corretamente (CCR), sensibilidade, especificidade, a razão de chances (*odds ratio*) e a curva ROC foram calculados por meio de um procedimento externo ao simulador de redes NSRBN, e excetuando-se o *odds ratio* os demais foram calculados internamente pelo simulador de redes MLP.

A curva ROC expressa a relação entre a sensibilidade e o falso positivo (1-Especificidade). Para se construir a curva ROC foram utilizados 20 valores de pontos de corte distintos. Foi considerado o ponto com maior eficiência de operação aquele que apresentou a maior média entre sensibilidade e a especificidade.

$$CCR = \frac{(Vp + Vn)}{(Vp + Fn + Vn + Fp)} \quad (5.1)$$

$$Sensibilidade = \frac{Vp}{(Vp + Fn)} \quad (5.2)$$

$$Especificidade = \frac{Vn}{(Vn + Fp)} \quad (5.3)$$

$$\text{Odds Ratio} = \frac{(Vp.Vn)}{(Fn.Fp)} \quad (5.4)$$

onde:

Verdadeiro Positivo (Vp) - o teste é positivo e o paciente tem a doença;

Falso Positivo (Fp) - o teste é positivo, mas o paciente não tem a doença;

Falso Negativo (Fn) - o teste é negativo e o paciente tem a doença;

Verdadeiro Negativo (Vn) - o teste é negativo e o paciente não tem a doença.

$$MSE = \frac{1}{n} \sum_{i=1}^n (O_d - O_o)^2 \quad (5.5)$$

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{O_d - O_o}{O_d} \right|}{n} \quad (5.6)$$

$$SE = \frac{DP}{\sqrt{n}} \quad (5.7)$$

onde:

O_d é a saída desejada ou alvo;

O_o é a saída obtida por intermédio do modelo;

DP é o desvio padrão;

n é o número de padrões considerados.

Os experimentos envolvendo a regressão logística foram conduzidos como descrito em Chaves Júnior (2000), onde se estudou os seguintes aspectos: As variáveis que fortalecem o modelo; o comportamento das variáveis selecionadas; a importância relativa das variáveis isoladas e combinadas e os acertos e erros do modelo.

Na construção dos modelos foram seguidos cuidados técnicos, incluindo a análise de resíduos, a verificação da estabilidade dos parâmetros em sub-amostras e as

significâncias dos parâmetros estimados. O método de estimação é o máximo de verossimilhança, o qual requer o uso de algoritmos de programação não-linear, calculando-se deste modo, a razão de chances (*odds ratio*) do evento de interesse ocorrer.

5.2.1 Experimentos com Redes Neurais MLP para HVE

Nas simulações realizadas com redes MLP a melhor rede encontrada correspondeu a seguinte configuração no simulador de redes MLP:

- Algoritmo de Treinamento: Levenberg-Marquardt;
- Funções de ativação camada escondida e de saída: logsig;
- Número de *restarts*: 8;
- Número de *reset* nos pesos: 10;
- Teste de até 18 neurônios na camada escondida;
- Particionamento dos Dados: 60% Conjunto de Treinamento, 20% Conjunto de Validação, 20% Conjunto de Teste.
- Critério para parar o treinamento: MSE validação;
- Valor do Limiar para classificar a saída: 0,5.

A arquitetura final da rede MLP para o melhor modelo com 5 entradas, apresentou os resultados indicados na Tabela 5.1.

Tabela 5.1 - Resultados para Rede MLP com 5 entradas

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,0686
Número de neurônios na Camada Escondida	3
CCR no teste (%)	71,4
Sensibilidade no teste (%)	75
Especificidade no teste (%)	69,2

5.2.2 Experimentos com Redes Neurais NSRBN para HVE

Nos experimentos com duas variáveis de entrada o modelo que apresentou o menor erro no conjunto de validação foi composto pelas variáveis: Idade e Colesterol Total. Os valores obtidos estão apresentados na Tabela 5.2.

Tabela 5.2 - Resultados para Rede com 2 entradas

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,736
MAPE no Conjunto de Teste	2,783
SE no Conjunto de Teste	0,438
Número de Blocos da Rede NSRBN	7
CCR (%)	76,24
Sensibilidade (%)	70,83
Especificidade (%)	81,13
Eficiência (Média Sens. e Espec.) (%)	75,98
Odds Ratio	10:1

Nos modelos de redes experimentadas com três variáveis de entrada aquele que apresentou o menor erro no conjunto de validação foi composto pelas variáveis: Idade, PAM do sono na MAPA 1 (MASMEDM1) e Altura. Os valores obtidos estão apresentados na Tabela 5.3.

Tabela 5.3 - Resultados para Rede com 3 entradas

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,297
MAPE no Conjunto de Teste	3,498
SE no Conjunto de Teste	0,453
Número de Blocos da Rede NSRBN	11
CCR (%)	82,18
Sensibilidade (%)	85,42
Especificidade (%)	79,25
Eficiência (Média Sens. e Espec.) (%)	83,80
<i>Odds Ratio</i>	22:1

Para os experimentos realizados com modelos de redes escolhendo-se seis variáveis de entrada aquele que apresentou o menor erro no conjunto de validação foi composto pelas variáveis: Idade, Sexo, Relação Cintura-Quadril, Colesterol Total, PAM e Frequência Cardíaca. A Figura 5.1 mostra a curva ROC para o modelo e na Tabela 5.4 pode ser observado o desempenho.

Tabela 5.4 - Resultados para Rede com 6 entradas

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,849
MAPE no Conjunto de Teste	2,55
SE no Conjunto de Teste	0,428
Número de Blocos da Rede NSRBN	7
CCR (%)	89,10
Sensibilidade (%)	87,50
Especificidade (%)	90,60
Eficiência (Média Sens. e Espec.) (%)	89,05
<i>Odds Ratio</i>	67:1
Área Curva ROC	0,776

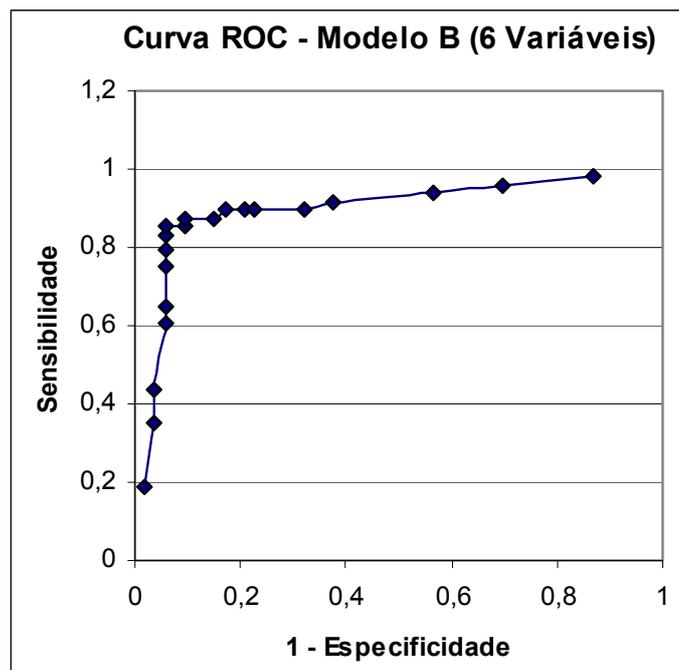


Figura 5.1 - Curva ROC para o Modelo 6 variáveis

É apresentado a seguir um conjunto de resultados para classificação de HVE utilizando-se apenas variáveis obtidas por meio do eletrocardiograma. O melhor resultado para um modelo com duas variáveis resultou na seleção de: D2S e AVFS. Os resultados correspondentes estão na Tabela 5.5.

Tabela 5.5 - Resultados para Rede com 2 entradas do Eletrocardiograma

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	1,1
MAPE no Conjunto de Teste	4,12
SE no Conjunto de Teste	0,5
Número de Blocos da Rede NSRBN	10
CCR (%)	69,39
Sensibilidade (%)	58,33
Especificidade (%)	80
Eficiência (Média Sens. e Espec.) (%)	69,16
Odds Ratio	5,6

Para um modelo com três variáveis o melhor na seleção resultou: D2S, AVFS e AVRS. Os resultados correspondentes estão na Tabela 5.6.

Tabela 5.6 - Resultados para Rede com 3 entradas do Eletrocardiograma

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,68
MAPE no Conjunto de Teste	4,18
SE no Conjunto de Teste	0,544
Número de Blocos da Rede NSRBN	10
CCR (%)	73,47
Sensibilidade (%)	72,92
Especificidade (%)	74
Eficiência (Média Sens. e Espec.) (%)	73,46
Odds Ratio	7,66

A seleção entre os modelos com quatro variáveis resultou na seleção: V1R, D2S, AVFS e QRS. Os resultados correspondentes estão na Tabela 5.7.

Tabela 5.7 - Resultados para Rede com 4 entradas do Eletrocardiograma

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,846
MAPE no Conjunto de Teste	3,99
SE no Conjunto de Teste	0,529
Número de Blocos da Rede NSRBN	12
CCR (%)	73,47
Sensibilidade (%)	60,42
Especificidade (%)	86
Eficiência (Média Sens. e Espec.) (%)	73,21
Odds Ratio	9,37

5.2.3 Experimentos com Regressão Logística para HVE

Na investigação de modelos para estimar a probabilidade de HVE por meio de regressão logística, foram identificadas as variáveis Idade, Pressão Média no Sono (sistólica, de pulso e arterial média) e os valores do índice eletrocardiográfico de Cornell para HVE como aquelas que mais fortalecem o modelo (Chaves Júnior, 2000) e (Campello de Souza; Chaves Júnior e Krieger, 2000).

Quando foi realizada a análise da importância de variáveis combinadas Chaves Júnior (2000) encontrou a melhor resposta para o modelo composto pelas variáveis: Idade, índice eletrocardiográfico de Cornell e os valores abaixo e acima da mediana da pressão arterial média no sono. Para este modelo obteve-se o desempenho apresentado na Tabela 5.8 e a curva ROC Figura 5.2.

Tabela 5.8 - Desempenho Modelo de Regressão Logística

MÉTRICA	VALOR OBTIDO
Taxa de Acertos (CCR)	82,00%
Sensibilidade	81,00%
Especificidade	83,00%
Área Curva ROC	0,609
Odds Ratio	21:1

A equação da regressão logística está apresentada a seguir:

$$\text{logit}(p) = -6,743 + 0,097IDADE + 1,782CORNELL + 2,047MAPSME01 \quad (5.8)$$

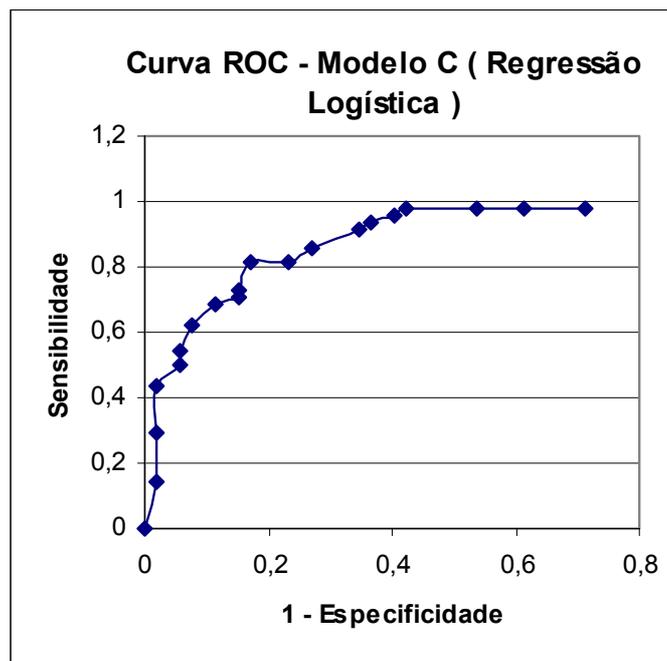


Figura 5.2 - Curva ROC para o Modelo Regressão Logística

Na Tabela 5.9 estão apresentados os resultados das redes neurais NSRBN e MLP, bem como da regressão logística de forma a facilitar uma comparação entre eles.

Tabela 5.9 - Resultados das Redes Neurais e da Regressão Logística

ASPECTO OBSERVADO	MODELO UTILIZADO							
	MLP	REDES NSRBN						LOG.
	Tab. 5.1	Tab. 5.2	Tab. 5.3	Tab. 5.4	Tab. 5.5	Tab. 5.6	Tab. 5.7	Tab. 5.8
Nº Variáveis Ent. / Aux.	5	2	3	6	2	3	4	3
MSE (Validação)	0,0686	0,736	0,297	0,849	1,1	0,68	0,846	-
MAPE (Teste)	-	2,783	3,498	2,55	4,12	4,18	3,99	-
SE (Teste)	-	0,438	0,453	0,428	0,5	0,544	0,529	-
Nº neurônios na Camada Escondida	3	-	-	-	-	-	-	-
Nº de Blocos da Rede NSRBN	-	7	11	7	10	10	12	-
CCR (%)	71,40	76,24	82,18	89,10	69,39	73,47	73,47	82,00
Sensibilidade (%)	75,00	70,83	85,42	87,50	58,33	72,92	60,42	81,00
Especificidade (%)	69,20	81,13	79,25	90,60	80	74	86,00	83,00
Eficiência (%)	72,10	75,98	83,80	89,05	69,16	73,46	73,21	82,00
Odds Ratio	-	10:1	22:1	67:1	5,6:1	7,7:1	9,4:1	21:1
Área Curva ROC	-	-	-	0,776	-	-	-	0,609

5.3 EXPERIMENTOS PARA CLASSIFICAÇÃO DE IMVE

Nos experimentos também procurou-se abranger e testar as diversas variáveis existentes na base de dados, tanto de forma isolada como combinadas. O conjunto total de variáveis possíveis de serem exploradas ultrapassa 2000. Foi utilizado como critério de seleção dos modelos investigados aqueles que produzissem resultados satisfatórios em termos de: coeficiente de determinação (r^2) bem como utilizassem variáveis de fácil obtenção e baixo custo financeiro. Outro objetivo perseguido foi selecionar modelos que, sem negligenciar o desempenho final, apresentassem menor custo computacional e complexidade, em termos do número de blocos na configuração após a fase de treinamento da NSRBN bem como utilizaram um menor número de variáveis de entrada.

Os resultados obtidos com estes modelos conexionistas foram comparados com outros gerados com a abordagem clássica estatística (Regressão Múltipla). Em todos os casos as redes tiveram como variável de saída os valores correspondentes ao índice de massa do ventrículo esquerdo (IMVE). Neste estudo as medidas obtidas com o ecocardiograma foram utilizadas como padrão ouro. A análise de desempenho foi realizada utilizando-se um *software* simulador de redes NSRBN (Valença, 2005), o qual permitiu escolher e testar uma grande variedade de:

- Tamanhos dos conjuntos de treinamento, validação e teste;
- Valores dos parâmetros como taxa de aprendizado, momento e semente;
- Critério de parada para o treinamento.

Para cada simulação realizada estava disponível o mínimo erro médio quadrático (MSE) no conjunto de validação, o erro percentual absoluto médio (MAPE) e o erro padrão (SE) no conjunto de teste. O coeficiente de determinação foi calculado por meio de um procedimento externo ao simulador e informa que fração da variabilidade de uma característica é explicada pela outra variável.

$$r^2 = \frac{\sum_{i=1}^n (O_o - \overline{O_d})^2 / n}{\sum_{i=1}^n (O_d - \overline{O_d})^2 / n} \quad (5.8)$$

5.3.1 Experimentos com Redes Neurais NSRBN para IMVE

Nos experimentos com duas variáveis de entrada o modelo que apresentou o menor erro no conjunto de validação foi composto pelas variáveis: Idade e BPTMDM1. Os valores obtidos estão apresentados na Tabela 5.10.

Tabela 5.10 - Resultados para Rede com 2 entradas

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,15
MAPE no Conjunto de Teste	27,03
SE no Conjunto de Teste	35,65
Número de Blocos da Rede NSRBN	9
Coefficiente de Determinação	38,37

Nos modelos de redes experimentadas com três variáveis de entrada aquele que apresentou o menor erro no conjunto de validação foi composto pelas variáveis: Idade, Colesterol Total e BPSMDM1. Os valores obtidos estão apresentados na Tabela 5.11.

Tabela 5.11 - Resultados para Rede com 3 entradas

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,127
MAPE no Conjunto de Teste	27,88
SE no Conjunto de Teste	35,13
Número de Blocos da Rede NSRBN	11
Coefficiente de Determinação	56,41

Para os experimentos realizados com modelos de redes escolhendo-se quatro variáveis de entrada aquele que apresentou o menor erro no conjunto de validação foi composto pelas variáveis: Idade, Relação Cintura-Quadril, Duração do Intervalo QT (eletrocardiograma) e BPSMDM1. Os valores obtidos estão apresentados na Tabela 5.12 e a Figura 5.3 apresenta as curvas com os valores esperados e os obtidos com o este modelo.

Tabela 5.12 - Resultados para Rede com 4 entradas

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,066
MAPE no Conjunto de Teste	20,76
SE no Conjunto de Teste	31,02
Número de Blocos da Rede NSRBN	7
Coefficiente de Determinação	85,82

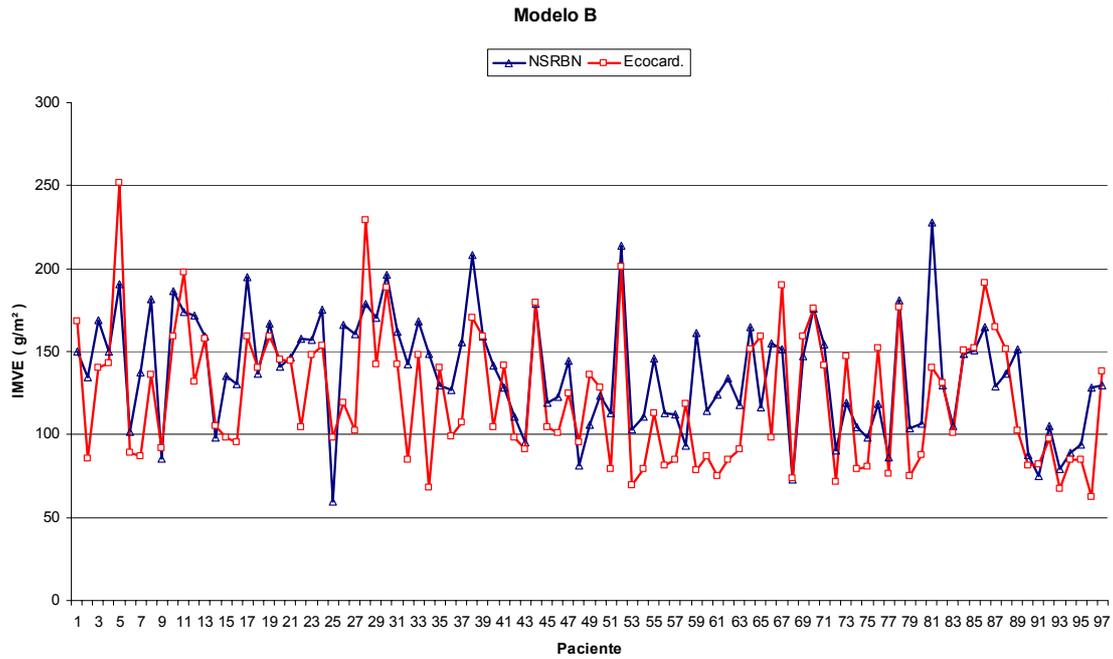


Figura 5.3 - Curvas dos valores esperados vs. obtidos – modelo NSRBN

Nos modelos de redes experimentadas com seis variáveis de entrada aquele que apresentou o menor erro no conjunto de validação foi composto pelas variáveis: Idade, Sexo, Relação Cintura-Quadril, MAPTDPM1 e FCTMEDM1. Os valores obtidos estão apresentados na Tabela 5.13.

Tabela 5.13 - Resultados para Rede com 6 entradas

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,067
MAPE no Conjunto de Teste	29,42
SE no Conjunto de Teste	45,8
Número de Blocos da Rede NSRBN	9
Coefficiente de Determinação	68,45

5.3.2 Experimentos com Regressão Múltipla para IMVE

As variáveis que compuseram o melhor modelo com a regressão múltipla foram: Idade, Relação cintura/quadril (Cintquad), valores contínuos do índice eletro cardiográfico de Cornell (Cornellc) e a pressão sistólica média do sono (Pssmed) (Chaves Júnior, 2000). Elas explicaram 47% da variabilidade do logaritmo neperiano do índice de massa ventricular esquerda (coeficiente de determinação igual a 0,47). A expressão analítica do modelo está na Equação 5.9.

$$\text{IMVE} = \exp[2,6535 + 0,0065(\text{Idade}) + 0,8247(\text{Cintquad}) + 0,0080(\text{Cornellc}) + 0,0065(\text{Pssmed})] \quad (5.9)$$

A Figura 5.4 apresenta as curvas com os valores esperados e os obtidos com o modelo de regressão múltipla.

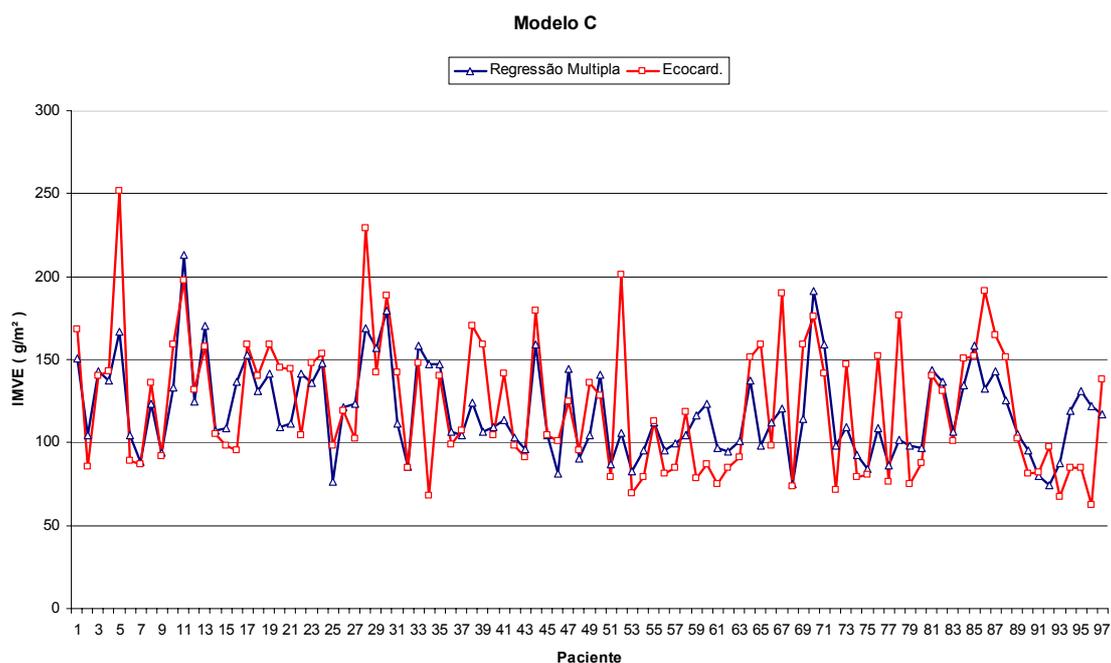


Figura 5.4 - Curvas dos valores esperados vs. obtidos – modelo Regressão Múltipla

Na Tabela 5.14 estão apresentados os resultados das redes neurais NSRBN e da regressão múltipla de forma a facilitar uma comparação entre eles.

Tabela 5.14 - Resultados das Redes Neurais NSRBN e da Regressão Múltipla

ASPECTO OBSERVADO	MODELO UTILIZADO				
	REDES NSRBN				REG. MÚLTIPLA
	Tab. 5.10	Tab. 5.11	Tab. 5.12	Tab. 5.13	-
Nº Variáveis Ent. / Aux.	2	3	4	6	4
MSE (Validação)	0,15	0,127	0,066	0,067	-
MAPE (Teste)	27,03	27,88	20,76	29,42	-
SE (Teste)	35,65	35,13	31,02	45,8	-
Nº de Blocos da Rede NSRBN	9	11	7	9	-
Coef. Determinação r^2 (%)	38,37	56,41	85,82	68,45	47,0

6. O APOIO AO DIAGNÓSTICO DA ARRITMIA CARDÍACA

“Salus Populi Suprema Est Lex.”

Cícero (106 aC–43 aC)

“O bem estar do povo é a lei suprema” (Tradução livre)

Neste capítulo são apresentados os resultados de um sistema de apoio a decisão para diagnóstico de arritmias cardíacas.

6.1 DESCRIÇÃO DA BASE DE DADOS

A base de dados utilizada neste capítulo foi gerada por Silva Filho (2003). A amostra foi resultante de um estudo de natureza observacional com 108 indivíduos hipertensos leves, moderados ou graves, de ambos os sexos, com idade média igual a 55,2 anos (23 a 77 anos) e todos em uso de medicação anti-hipertensiva, exceto as classes de betabloqueadores e bloqueadores dos canais de cálcio dos grupos diltiazem e verapamil. Dois terços destes indivíduos eram portadores de algum grau de HVE concêntrica. O padrão-ouro para a tipificação e contagem das arritmias cardíacas utilizado foram os resultados do exame de Holter-24 h. O tamanho dimensionado da amostra para esta pesquisa preconizava 106 o que foi atendido.

Além da classificação e contagem das arritmias pelo método de Holter-24 h, foram anotados os valores das estatísticas referentes ao domínio do tempo. Os valores dessas estatísticas foram calculados levando-se em consideração todos os registros dos valores do intervalo RR no período de 24 h. No referido estudo o padrão ouro para o diagnóstico de HVE foi o ecocardiograma.

Este estudo atendeu aos padrões éticos com protocolo de estudo e termo de consentimento aprovados pelos comitês de ética em pesquisas do Instituto Materno Infantil de Pernambuco (IMIP). As variáveis que compõem a amostra estão apresentadas no Apêndice D Tabela D.2.

6.2 EXPERIMENTOS COM REDES NSRBN PARA CLASSIFICAÇÃO DE ARRITMIAS

Nos experimentos procurou-se abranger e testar as diversas variáveis existentes na base de dados, tanto de forma isolada como combinadas. Foi utilizado como critério de seleção dos modelos investigados aqueles que produzissem resultados satisfatórios em termos de: classificação correta de padrões (CCR), sensibilidade, especificidade, bem como utilizassem variáveis de fácil obtenção e baixo custo financeiro. Outro objetivo perseguido foi selecionar modelos que, sem negligenciar o desempenho final, apresentassem menor custo computacional e complexidade, em termos do número de blocos na configuração após a fase de treinamento da NSRBN bem como utilizaram um menor número de variáveis de entrada.

A análise de desempenho foi realizada utilizando-se um *software* simulador de redes NSRBN (Valença, 2005), o qual permitiu escolher e testar uma grande variedade de:

- Tamanhos dos conjuntos de treinamento, validação e teste;
- Valores dos parâmetros como taxa de aprendizado, momento e semente;
- Critério de parada para o treinamento.

Para cada simulação realizada estava disponível o mínimo erro médio quadrático (MSE) no conjunto de validação, o erro percentual absoluto médio (MAPE) e o erro padrão (SE) no conjunto de teste. A taxa de padrões classificados corretamente (CCR), sensibilidade, especificidade.

Para uma primeira seleção de variáveis procurou-se estudar as variáveis separadamente. Dentre as variáveis que apresentaram algum poder de classificação nesta base de dados, são apresentados a seguir os resultados de alguns casos:

- **Caso 1** (Variável de Entrada: HISTIFAMI; Variável de Saída: ESV01)

Os valores obtidos estão apresentados na Tabela 6.1.

Tabela 6.1 - Resultados para Rede – Caso 1

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,076
MAPE no Conjunto de Teste	2,16
SE no Conjunto de Teste	0,44
Número de Blocos da Rede NSRBN	7
CCR (%)	78,70
Sensibilidade (%)	53,33
Especificidade (%)	82,80
Odds Ratio	5,5

- **Caso 2** (Variável de Entrada: IDADE; Variável de Saída: ATRIAL01)

Os valores obtidos estão apresentados na Tabela 6.2.

Tabela 6.2 - Resultados para Rede – Caso 2

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	0,42
MAPE no Conjunto de Teste	1,74
SE no Conjunto de Teste	0,34
Número de Blocos da Rede NSRBN	3
CCR (%)	87,85
Sensibilidade (%)	96,47
Especificidade (%)	54,55
Odds Ratio	32,8

- **Caso 3** (Variável de Entrada: SEXO; Variável de Saída: ATRIAL01)

Os valores obtidos estão apresentados na Tabela 6.3.

Tabela 6.3 - Resultados para Rede – Caso 3

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	1,56
MAPE no Conjunto de Teste	2,94
SE no Conjunto de Teste	0,44
Número de Blocos da Rede NSRBN	2
CCR (%)	46,73
Sensibilidade (%)	35,29
Especificidade (%)	90,91
<i>Odds Ratio</i>	5,45

- **Caso 4** (Variável de Entrada: IDADE; Variável de Saída: VENTR01)

Os valores obtidos estão apresentados na Tabela 6.4.

Tabela 6.4 - Resultados para Rede – Caso 4

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	1,37
MAPE no Conjunto de Teste	4,17
SE no Conjunto de Teste	0,505
Número de Blocos da Rede NSRBN	3
CCR (%)	69,16
Sensibilidade (%)	66,67
Especificidade (%)	75,86
<i>Odds Ratio</i>	6,28

- **Caso 5** (Variável de Entrada: ETILISMO; Variável de Saída: VENTR01)

Os valores obtidos estão apresentados na Tabela 6.5.

Tabela 6.5 - Resultados para Rede – Caso 5

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	1,86
MAPE no Conjunto de Teste	4,59
SE no Conjunto de Teste	0,55
Número de Blocos da Rede NSRBN	2
CCR (%)	70,09
Sensibilidade (%)	88,46
Especificidade (%)	20,69
<i>Odds Ratio</i>	2,0

- **Caso 6** (Variável de Entrada: PESO; Variável de Saída: VENTR01)

Os valores obtidos estão apresentados na Tabela 6.6.

Tabela 6.6 - Resultados para Rede – Caso 6

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	1,75
MAPE no Conjunto de Teste	4,67
SE no Conjunto de Teste	0,54
Número de Blocos da Rede NSRBN	3
CCR (%)	64,49
Sensibilidade (%)	71,79
Especificidade (%)	44,83
<i>Odds Ratio</i>	2,06

- **Caso 7** (Variável de Entrada: IMC; Variável de Saída: VENTR01)

Os valores obtidos estão apresentados na Tabela 6.7.

Tabela 6.7 - Resultados para Rede – Caso 7

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	1,85
MAPE no Conjunto de Teste	4,67
SE no Conjunto de Teste	0,54
Número de Blocos da Rede NSRBN	2
CCR (%)	62,62
Sensibilidade (%)	65,38
Especificidade (%)	55,17
<i>Odds Ratio</i>	2,32

- **Caso 8** (Variável de Entrada: PSC; Variável de Saída: VENTR01)

Os valores obtidos estão apresentados na Tabela 6.8.

Tabela 6.8 - Resultados para Rede – Caso 8

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	1,57
MAPE no Conjunto de Teste	2,94
SE no Conjunto de Teste	0,43
Número de Blocos da Rede NSRBN	2
CCR (%)	42,99
Sensibilidade (%)	30,59
Especificidade (%)	90,91
<i>Odds Ratio</i>	4,4

- **Caso 9** (Variável de Entrada: PPULSOC; Variável de Saída: VENTR01)

Os valores obtidos estão apresentados na Tabela 6.9.

Tabela 6.9 - Resultados para Rede – Caso 9

ASPECTO OBSERVADO	VALOR ENCONTRADO
MSE no Conjunto de Validação	1,53
MAPE no Conjunto de Teste	2,81
SE no Conjunto de Teste	0,43
Número de Blocos da Rede NSRBN	2
CCR (%)	57,01
Sensibilidade (%)	48,24
Especificidade (%)	90,91
Odds Ratio	9,31

Na Tabela 6.10 estão apresentados os resultados das redes neurais NSRBN para a classificação de arritmias de forma a facilitar uma comparação entre eles.

Tabela 6.10 – Resultado Geral das Redes NSRBN para Classificação de Arritmias

ASPECTO OBSERVADO	MODELO UTILIZADO								
	Caso 1	Caso 4	Caso 5	Caso 6	Caso 7	Caso 8	Caso 9	Caso 2	Caso 3
Variável de Saída	ESV01	VENTR01						ATRIAL01	
MSE (Validação)	0,076	1,37	1,86	1,75	1,85	1,57	1,53	0,42	1,56
MAPE (Teste)	2,16	4,17	4,59	4,67	4,67	2,94	2,81	1,74	2,94
SE (Teste)	0,44	0,51	0,55	0,54	0,54	0,43	0,43	0,34	0,44
Nº de Blocos da Rede NSRBN	7	3	2	3	2	2	2	3	2
CCR (%)	78,70	69,16	70,09	64,49	62,62	42,99	57,01	87,85	46,73
Sensibilidade (%)	53,33	66,67	88,46	71,79	65,38	30,59	48,24	96,47	35,29
Especificidade (%)	82,80	75,86	20,69	44,83	55,17	90,91	90,91	54,55	90,91
Odds Ratio	5,5	6,28	2,0	2,06	2,32	4,4	9,31	32,8	5,45

7. UM SISTEMA DE INFORMAÇÕES PARA UMA CLÍNICA DE CARDIOLOGIA

“In a moment of decision the best thing you can do is the right thing. The worst thing you can do is nothing.”

Theodore Roosevelt, (1858-1919)

“Em qualquer momento de decisão, a melhor coisa que você pode fazer é a coisa certa.

A pior coisa que você pode fazer é não fazer nada.” (Tradução livre)

“Conhecimento não é aquilo que você sabe, mas o que você faz com aquilo que sabe.”

Aldous Huxley (1894-1963)

Este capítulo busca aproximar as áreas de apoio a decisão com a de redes neurais artificiais. Portanto, procura-se propor uma aplicação em problema real na área de cardiologia de forma a contribuir e disseminar esta modelagem na área médica. Observa-se que diferentemente do cérebro humano como padrão original, até o momento atual o padrão na área de computação é fortemente o de von-Neumann, o qual é bastante diferente daquele. Assim, a proposta de utilizar redes neurais artificiais para apoio a decisão vai ao encontro do paradigma original.

A decisão é um tema que se faz presente e já era estudado desde a antigüidade, pois na Grécia Aristóteles (384-322 aC), propôs a Teoria das Razões em seu *“De Motu Animalium”* (Schick, 1997). Toda razão segundo a teoria de Aristóteles é um estado mental com duas partes: uma crença e um desejo. Uma decisão é fruto do exercício da vontade do decisor em escolher uma dentre todas as opções existentes e elegíveis. Portanto, uma decisão implica em uma escolha, que por sua vez implica em opções e para que haja opções implica em haver liberdade de escolha. Assim, pode-se dizer que uma decisão é uma escolha de opções tomada livremente. Ela requer um posicionamento frente ao problema pelo decisor. O entendimento do problema tratado,

do seu contexto e os fatos a ele relacionados são pontos importantes para uma boa decisão.

No contexto da área médica a decisão muita das vezes é complexa e tem consequência direta na qualidade de vida e risco de morte de um indivíduo (paciente). O ambiente de decisão típico observado na área de cardiologia apresenta: um elevado número de fatores internos e externos a serem considerados relativos ao paciente; a existência de possíveis respostas adversas para uma mesma indicação de fármaco e também as causas e efeitos do mecanismo de satisfação não estão completamente compreendidos.

Na década de 40 do século 20, o biólogo Ludwig von Bertalanffy (1901-1972), quando de seus estudos de organismos vivos, chegou a conclusão que há uma clara interação entre as partes componentes destes organismos e o seu comprometimento com os objetivo ou finalidades primordiais. Como consequência destes estudos e observações surgiu a Teoria Geral dos Sistemas (Cautela, 1986). Uma definição possível para sistema é um conjunto em interação, inter-relacionado ou ainda elementos independentes ou partes que funcionam juntas como um todo para atingir um objetivo.

Um sistema de informação é um conjunto de elementos independentes (subsistemas), logicamente associados, para que de sua interação sejam geradas informações necessárias para a tomada de decisões (Cautela, 1986). Sistemas de suporte ou apoio a decisão são sistemas especiais a fim de melhorar a qualidade das decisões em problemas semi-estruturados (Turban, 1995).

De forma genérica podem ser listadas algumas classes de problemas que com relativa facilidade ocorrem, isoladamente ou simultaneamente, e que demandam apoio a decisão. Estas classes e suas explicações seguem: Predição (capacidade de antecipar fatos e eventos futuros); Diagnose (inferir causas de problemas por observação de sintomas); Monitoração (comparar observações de fenômenos e ventos com padrões previamente estabelecidos); Controle (interagir de forma adaptativa com um dado sistema e exercer sobre ele as correções necessárias para atingir as metas estabelecidas) entre outras.

As razões para utilizar um sistema de apoio à decisão baseado neural decorre principalmente das características dos problemas tratados e das próprias redes neurais. Como razões mais relevantes podem-se citar: a natureza complexa em função de uma alta cardinalidade de entradas; não haver o conhecimento completo da iteração entre os fatores de entrada e saída; por demandarem uma habilidade geralmente não trivial para sua solução; pela criticidade e relevância de emprego; grande capacidade de paralelismo e robustez das RNA; capacidade de lidar com ambigüidade entre outros (Lima Neto e Ludermir, 1997) (Kasabov, 1996).

A seguir são propostas algumas fases a serem utilizadas no desenvolvimento do sistema de decisão, baseado em redes neurais artificiais.

Concepção: Estudar o problema e seus condicionantes.

Aquisição e Preparação dos Dados: Fazer o levantamento de fatores relevantes e ordená-los pelo seu grau de relevância, bem como verificar a disponibilidade, viabilidade e qualidade destes fatores em termos de serem representativos do problema tratado. Construir um filtro para a aquisição dos dados existentes na base de dados que serão utilizados. Remover os *outliers* (observação que é estatisticamente diferente dos demais valores da distribuição de dados considerada). Proceder ao ajustamento dimensional das entradas e saídas, além de normalizar e codificar os valores. Fazer a preparação dos arquivos de treinamento e teste. Possibilitar que o conjunto de dados de treinamento e teste represente da forma mais fiel possível, o problema a ser resolvido pelo sistema de apoio a decisão.

Modelagem Neural: Classificar o tipo de problema e analisar possíveis arquiteturas. Realizar simulações aleatórias iniciais, identificar relevâncias entrada/saída e montar um conjunto de simulações. Treinar as RNA e tabular os resultados. Validar as RNA submetidas ao treinamento e tabular estes resultado obtidos. Aprovar junto ao usuário os valores obtidos de desempenho, generalização e erros admissíveis. Obter uma solução específica de *hardware* e *software*, a ser integrada ao parque de sistemas.

Diálogo com o Usuário: Definir uma forma simples e objetiva de comunicação entre o sistema e o usuário. Permitir facilidade na entrada e saída de dados com interfaces

amigáveis. Obter a aprovação do usuário na adequação da solução final oferecida pelo sistema de apoio à decisão.

As quatro fases descritas nos parágrafos anteriores estão representadas na Figura 7.1.

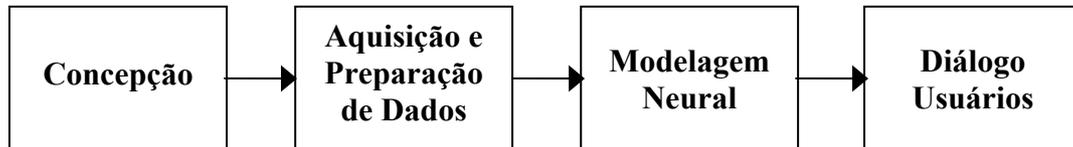


Figura 7.1 - Fases de desenvolvimento do sistema de apoio a decisão

Como seqüência a estas etapas de implementação sugere-se mais uma etapa complementar:

Definições Gerais de Utilização: Esta etapa deve abranger os questionamentos finais que antecedem a utilização do sistema na rotina médica. Assim, deverá ser definido o conjunto de usuários e também devem ser observados aspectos como navegabilidade, acessibilidade e definição de acesso a dados e funções (quem acessa o quê?). Deve haver segurança e confidencialidade no acesso aos dados de pacientes. Deve ainda ser definido um calendário de retreinamento dos modelos neurais e de revisão conceitual do sistema.

Pode-se listar um conjunto de vantagens gerais associadas com a introdução de um sistema de apoio a decisão:

- Redução no intervalo de tempo requerido para a tomada de decisões pelos médicos;
- O conhecimento uma vez armazenado na rede neural permite que ele permaneça para o grupo ou equipe médica, mesmo quando houver rotatividade de pessoal;
- O sistema pode ficar disponível para ser utilizado vinte e quatro horas por dia inclusive por intermédio de uma rede de computadores;
- Atenuação de erros grosseiros nas decisões, pois o sistema provê uma referência em termos de decisão possível, que sugere ao decisor um balizamento para sua decisão;
- Diminuição em termos de discrepâncias entre decisores diferentes;

- Redução de decisões viciosas ou induzidas, isto também é favorecido pela característica do sistema em ser adaptável as novas situações do ambiente de decisão;

Um outro uso ainda pouco convencional, embora possível, é a utilização do sistema como ferramenta de auditoria, ou seja, verificar se as decisões tomadas eram “aceitáveis” no contexto.

Na Figura 7.2 estão apresentados os elementos relacionados com o contexto de um sistema de apoio à decisão voltado para uma clínica de cardiologia. Os dados que servirão de entrada para o sistema provêm de diversas formas como: a anamnese, o exame físico e os exames complementares. Estas são informações sobre o estado da natureza.

O sistema de apoio à decisão baseado em redes neurais permite realizar a classificação em relação a um dado aspecto desejado como exemplo: ter ou não ter HVE, ter ou não ter um tipo específico de arritmia, ter ou não ter diabetes ou outros tipos de classificação. É necessário salientar que para que esta classificação ocorra de forma satisfatória, deverá ter havido um treinamento prévio do modelo neural utilizando-se para isto uma base de dados representativa de indivíduos doentes e hígidos. Este treinamento prévio serve para formar a base de conhecimento do sistema neural.

Uma vez que o sistema já tenha sido treinado a classificação correspondente a um novo paciente será então rapidamente realizada. O resultado desta classificação poderá gerar um conjunto de alternativas para o decisor na área médica. Baseando-se nesta classificação e na particularidade de cada caso, o médico poderá fazer a opção pela melhor alternativa e sua correspondente conduta terapêutica.

Considerando-se que para cada alternativa escolhida há um conjunto associado de conseqüências possíveis, a melhoria que pode ser obtida com um sistema de apoio à decisão baseado em modelos de redes neurais, pode resultar em ganhos para o profissional da área médica e para a sociedade como um todo.

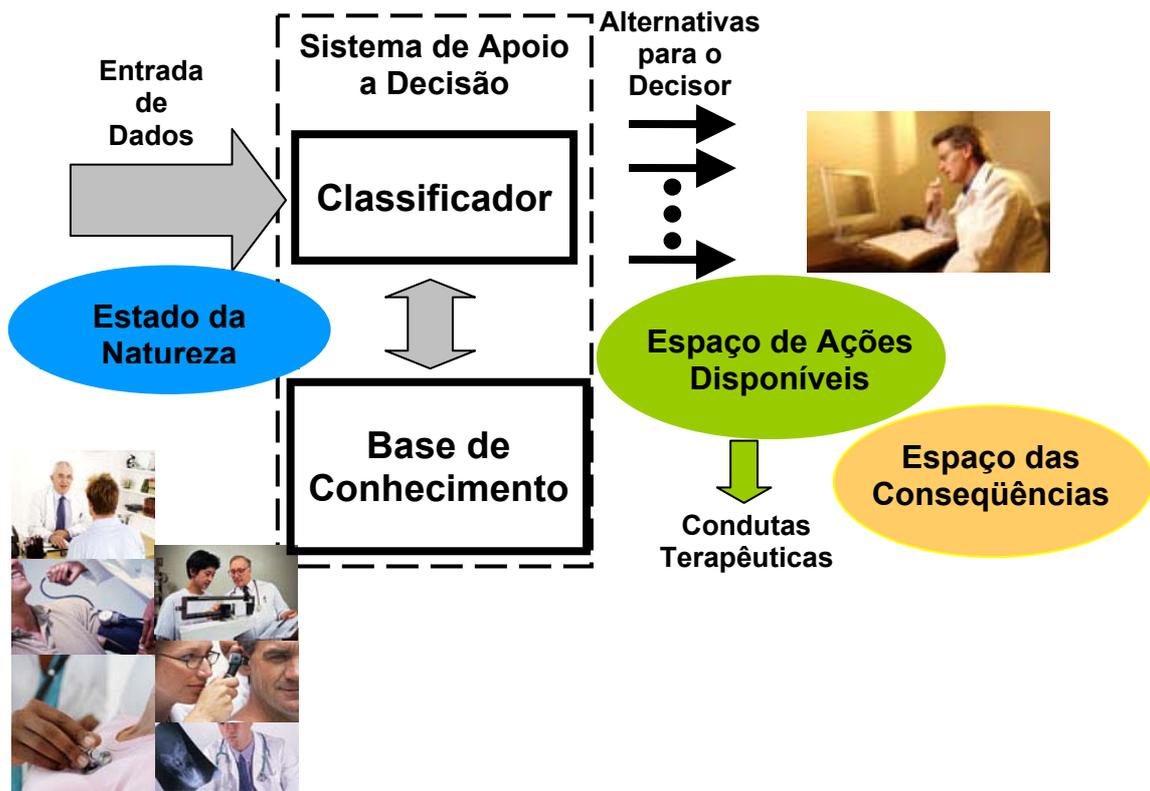


Figura 7.2 - Sistema de apoio à decisão em uma clínica de cardiológica

8. CONCLUSÕES, COMENTÁRIOS E SUGESTÕES

"I never think of the future – it comes soon enough."

Albert Einstein (1879–1955)

“Eu nunca penso no futuro, ele chega rápido demais.” (Tradução livre)

“Life is the art of drawing sufficient conclusions from insufficient premises”

Samuel Butler, (1612-1680)

“A vida é a arte de tirar conclusões suficientes a partir de premissas insuficientes”

(Tradução livre)

O estudo do sistema cardiovascular é uma área em permanente expansão, pela complexidade do objeto tratado. Muitas hipóteses já foram confirmadas por experimentos práticos, entretanto outros ainda carecem de esclarecimento o que demanda mais estudos a fim de possibilitar um melhor entendimento da complexa interação entre as diversas partes constituintes do SCV.

Neste trabalho foi realizada uma extensiva investigação que resultou no desenvolvimento de ferramentas de engenharia de sistemas e métodos quantitativos, no sentido de propor ações que tragam melhores resultados em relação aos disponíveis hoje para a área de saúde (cardiologia).

Diversas dificuldades foram encontradas durante a realização deste trabalho. Vale ressaltar a dificuldade em adquirir dados confiáveis e adequados para fazer os experimentos, uma vez que a intenção inicial era investigar a área de arritmias cardíacas. O poder de classificação encontrado nos diversos experimentos realizados com a base de dados informada no capítulo 6, provavelmente deve-se ao fato de só possuir informações de indivíduos doentes na sua composição. Uma comprovação disto pode ser feita partindo-se de outras técnicas. Assim, sem perder o foco na área de cardiologia houve um redirecionamento da pesquisa, ampliando-a para o apoio ao diagnóstico de HVE.

É importante salientar que alguns dos modelos propostos fazem uso de variáveis de baixo custo financeiro e que podem ser obtidas facilmente durante uma consulta médica. Contribuiu-se também com uma melhor utilização das medidas fornecidas pelo exame MAPA-24h, fazendo-se uso de variáveis obtidas por este exame.

Considerando a pequena amostra disponível para treinamento da rede neural, 101 indivíduos, os resultados obtidos (capítulo 6), tanto para classificação de HVE quanto de estimação de IMVE, até agora indicam que os sistemas propostos têm potencial para reduzir ou ainda justificar melhor a prescrição de exames e procedimentos de custo elevado na área de cardiologia. Isto é reforçado pelas características de das redes de detectar relações não lineares de alta ordem entre as variáveis, além de sua robustez para tratar dados incompletos.

Contudo, entende-se que o modelo aqui apresentado seja validado em outras amostras, de forma a testar seu poder preditivo obtido a partir desta população com outras. Entende-se que estes novos experimentos são necessários antes que estes modelos de redes possam ser aceitos para apoio à decisão médica.

Em função dos estudos realizados foi possível propor uma modelagem para as arritmias cardíacas. Esta modelagem foi aceita na forma de um trabalho no Simpósio Brasileiro de Pesquisa Operacional (SBPO) e foi demonstrada a sua aplicação por meio dos dados de um exame Holter-24h (vide lista de publicações).

Conclui-se que este trabalho acrescentou melhorias que podem apontar na direção de se construir sistemas de apoio a decisão médica, dotados de flexibilidade e robustez no provimento de diagnósticos médicos.

Entre as perspectivas futuras, está a possibilidade, já mencionada no capítulo 4, de desenvolver uma nova rede construtiva que utilize o algoritmo TALUS no seu treinamento. Este algoritmo de otimização global já mostrou sua potencialidade em fornecer bons resultados e a união das redes NSRBN com este algoritmo poderá resultar

em uma nova rede de maior desempenho tanto para tarefas de classificação como outras.

Outra possibilidade é a introdução de um bloco para pré-processamento das variáveis de entrada com análise de componentes principais. Para este fim será necessário coletar dados contendo a dinâmica do SCV.

BIBLIOGRAFIA

“If I have seen further, it is by standing on the shoulders of giants”

Sir Issac Newton (1642-1727)

Letter to Robert Hooke.

AGRESTI, A. *An introduction to categorical data analysis*. New York: John Willey, 1996.

BARRETT, J. The use of functionals in the analysis of nonlinear physical systems. *Journal of Electronics and Control*, v. 15, p. 567-615, 1963.

BEALE, R.; JACKSON, T. *Neural computing: an introduction*. Bristol: IOP Publishing, 1990.

BISHOP, C. M. *Neural networks for pattern recognition*. Oxford: Oxford University Press, 1995.

BOWKER, A. H.; LIEBERMAN, G. J. *Engineering statistics*. 2nd edition. Englewood Cliffs, New Jersey, 1972. p.232.

BOX, G. E. P. Robustness in scientific model building. In: LAUNER, R. L.; WILKINSON, G. N. (Eds.), *Robustness in statistics*. New York: Academic Press, 1979. p. 201-236.

BRAGA, A. P.; LUDERMIR, T. B.; CARVALHO, A. C. P. de L. F. *Redes Neurais Artificiais: Teoria e Aplicações*. Rio de Janeiro: Livro Técnico e Científico, 2000.

BRASIL.Ministério da Saúde. *Anuário estatístico de saúde do Brasil 2001*. Disponível em <http://portal.saude.gov.br/saude/aplicacoes/anuario2001/index.cfm>. Acesso em março de 2003.

_____. *Indicadores e Dados Básicos – Brasil – 2003*. Disponível em <http://tabnet.datasus.gov.br/cgi/idb2003/matriz.htm>. Acesso em abril de 2005.

BRAUNWALD, E.; ZIPES, D. P.; LIBBY, P. *Tratado de Medicina Cardiovascular*, 6.ed. São Paulo: ROCA, 2003. 2v.

BRILLIANT, M. B. *Theory of the analysis of nonlinear systems*. Cambridge: Massachusetts Institute of Technology, 1958. (Massachusetts Institute of Technology Research Laboratory of Electronics Technical Report No. 345).

BRONZINO, Joseph D. *The Biomedical Engineering Handbook*. Boca Raton: CRC Press & IEEE Press, 1995.

CAMPELLO DE SOUZA, F. M. *Sistemas de controle auto-adaptativos*. Recife, 1974. (Notas de Aula de Sistemas de Controle, Departamento de Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Pernambuco).

_____. A global optimization algorithm for solving linear programming problems. In: CONFERENCE ON MANAGEMENT AND CONTROL OF PRODUCTION AND LOGISTICS, MCPL, 1997, Campinas, SP. Proceedings... Campinas, 1997.

_____. *Decisões racionais em situações de incerteza*. 2 ed. Recife, 2005.

_____.; CHAVES JÚNIOR, H. de C.; KRIEGER, E. M. Modelo de regressão logística para a hipertrofia ventricular esquerda (HVE) pela monitorização ambulatorial da pressão arterial de 24 horas (MAPA-24h): poster apresentado. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE HIPERTENSÃO, 9., Salvador, BA. 2000.

CAUTELA, A. L.; POLLONI, E. G. F. *Sistemas de informação na administração de empresas*. São Paulo: Atlas, 1986.

CAVALCANTE, J. R. R. *Algoritmo probabilístico paralelo para otimização global*. 1996. 516f. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal de Pernambuco, Recife, 1996.

CAVALCANTI, A. M. *Desenvolvimento de algoritmos probabilísticos de otimização global*. Tese (Doutorado em Engenharia Elétrica) - Universidade Federal de Pernambuco, 2004.

_____.; OLIVEIRA NETO, R. F. de; CAVALCANTE, J. R. R. et al. Algoritmo de Otimização Global Irrestrita Utilizando Nuvens Controladas. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, SBPO, 36. 2004, São João del-Rei, MG., Anais... 2004.

CHAVES JUNIOR, H. de C. *Monitoração ambulatorial da pressão arterial (MAPA-24h): sua reprodutibilidade e importância na construção de modelos para estimar a hipertrofia cardíaca*, 2000. 222f. Tese (Doutorado em Medicina) - Universidade de São Paulo, São Paulo, 2000.

_____.; CAMPELLO DE SOUZA, F. M.; KRIEGER, E. M. The reproducibility of dipping status: beyond the cutoff points. *Blood Pressure Monitoring*, v. 10 n. 4, p. 201-205, July/August, 2005.

CYBENKO, G. *Continuous valued neural networks with two hidden layers are sufficient*. Technical Report, Department of Computer Science, Tufts University, 1988.

_____. Approximation by superpositions of a sigmoid function. In: *Mathematics of Control, Signals and Systems*, 2 p.303-314, 1989.

DE SOUZA, M. de F.; TIMERMAN, A.; SERRANO JÚNIOR, C. V. et al. Tendências do risco de morte por doenças circulatórias nas cinco regiões do Brasil de 1979 a 1996. *Arq Bras Cardiol*, v. 77, n.6, p.562-8, 2001.

DEMUTH, H.; BEALE, M. *Neural network toolbox user's guide*. Natick, MA: The Math Works, 1996.

DESOER, C.; VIDYASAGAR, M. *Feedback systems: input/output properties*. New York: Academic Press, 1975.

DRAPER, S. *Applied regression analysis*. Ed. John Wiley and Sons, 1981.

ERNST, M. E.; BERGUS, G. R. Noninvasive 24-hour ambulatory blood pressure monitoring: overview of technology and clinical applications. *Pharmacotherapy* 22(5) p.597-612, 2002.

FAHLMAN, S. Faster-learning variations on back-propagation: an empirical study. In: *Proceedings of the 1988 Connectionist Models Summer School*, Morgan Kaufmann, 1988.

_____.; LEBIERE, C. The cascadecorrelation learning architecture. In TOURETZKY, D. (Ed.). *Advances in Neural Information Processing Systems*, , Los Altos, CA: Morgan Kaufmann, 1990. v. 2, p. 524-532

GABOR, D; WILDY, W.; WOODCOCK, R. Universal non-linear filter predictor which optimizes itself by learning. *IEE Proc.* v.108, parte B, p.422-438, 1961.

GALLMAN, P. Representation of nonlinear systems via the Stone-Weierstrass theorem. *Automatica*, v. 12, p. 619-622, 1976.

GEORGE, D. A. *Continuous nonlinear systems*. Cambridge: Massachusetts Institute of Technology, 1959. (Massachusetts Institute of Technology Research Laboratory of Electronics Technical Report No. 355).

GIBBONS, R. J.; BALADY, G. J.; BEASLEY, J. W.; et al. ACC/AHA guidelines for exercise testing. A report of the American College of Cardiology/American Heart Association Task Force on practice guidelines (Committee on exercise testing). *J Am Coll Cardiol* 30:260, 1997.

GUERRA BELÉM, José. *Propostas de circuitos eletrônicos para implementação de algoritmos probabilísticos*. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal de Pernambuco, Recife, 1999.

GUYTON, A. C; HALL, J. E. *Tratado de fisiologia médica*. 10.ed. Rio de Janeiro: Guanabara Koogan, 2002.

HALME, A.; ORAVA, J. Generalized polynomial operators for nonlinear systems analysis. *IEEE Transactions on Automatic Control*, v. AC-17, p. 226-228, 1972.

HAMMOND, I. W.; DEVEREUX, R. B.; ALDERMAN, M. H. et al. The prevalence and correlates of echocardiographic left ventricular hypertrophy among employed patients with uncomplicated hypertension. *J Am Coll Cardiol.*; 1986 Mar;7(3):639-50.

HEBB, D.O. *The organization of behaviour*. New York: Wiley, 1949.

HOLTER, N. J. A New Method for Heart Studies. *Science*, 134: p.1214-20, 1961.

_____.; GENGERELLI, J. A. Remote recording of physiological data by radio. *Rocky Mountain Med. J.*, 46 p.747-751, 1949.

HOPFIELD, J. J. *Neural networks and physical systems with emergent collective computational abilities*. Proceedings of National Academy of Sciences, 79, p.2554-2558, 1982.

HOSMER, D. W.; LEMESHOW, S. *Applied logistic regression*, New York: John Wiley, 1989.

HOYERT, Donna L.; KUNG, Hsiang-Ching; SMITH, Betty L. *National vital statistics reports - deaths: preliminary data for 2003*. Division of Vital Statistics, v. 53, n. 15, February 2005.

ISIDORI, A.; RUBERTI, A. Realization theory of bilinear systems. In: MAYNE, D. Q. e BROCKETT, R. W. (Eds.). *Geometric methods in system theory*. Dordrecht, The Netherlands: Reidel Publishing, p. 83-130, 1973.

IVAKHNENKO, A. G. Heuristic self-organization in problems of engineering. *Automatica*, v. 6, p. 207-219, 1970.

_____. Polynomial theory of complex systems. *IEEE Transactions on Systems, Man, and Cybernetics*, v. SMC-1, n. 4, p. 364-378, October 1971.

JOINT NATIONAL COMMITTEE. *The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure (JNC-VII)*. National Institutes of Health of United States. December 2003.

KASABOV, N. K. *Foundations of neural networks, fuzzy systems and knowledge engineering*, MA, USA: Bradford Book – MIT Press, 1996.

KROGH, A. An introduction to hidden markov models for biological sequences – Chapter 4 In: SALZBERG, S. L.; SEARLS, D. B. e KASIF, S. (Eds.). *Computational methods in molecular biology*. Elsevier, 1998. p. 45-63.

LESSA, I. Assistência médica e óbitos por doença arterial coronariana no Brasil, 1980-1999. *Arq Bras Cardiol*, v. 81, n.4, p.329-35, 2003.

LIMA NETO, F. B. de; LUDERMIR, T. B. Suporte à decisão neural para a colheita de cana-de-açúcar. In: SIMPÓSIO BRASILEIRO DE REDES NEURAIAS, 4. 1997, Goiânia, GO, Anais... Goiânia,GO, 1997. p.27-31.

LINS, G. D. *Estudo, melhoria e implementação computacional do Talus*. 1999. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal de Pernambuco, Recife, 1999.

LIUSTERNIK, L.; SOBOLEV, V. *Elements of Functional Analysis*. New York: Frederick Ungar, 1961.

LOESCH, C.; SARI, S. T. *Redes neurais artificiais: fundamentos e modelos*. Blumenau: ed. da FURB, 1996.

MACKAY, J.; MENSAH, G. A. *The Atlas of Heart Disease and Stroke*. Geneva: World Health Organization/Centers for Disease Control and Prevention, 2004.

MANDALA, H. R.; IVAKHNENKO, A. G. *Inductive Learning Algorithms for Complex Systems Modeling*. Boca Raton: CRC Press Inc., 1994.

MANSUR, A. de P.; FAVARATO, D.; DE SOUZA, M. de F. et al. Tendência do risco de morte por doenças circulatórias no Brasil de 1979 a 1996. *Arq Bras Cardiol*, v. 76, n. 6, p.497-503, 2001.

_____.; DE SOUZA, M. de F.; TIMERMAN, A. et al. Tendência do risco de morte por doenças circulatórias, cerebrovasculares e isquêmicas do coração em 11 capitais do Brasil de 1980 a 1998. *Arq Bras Cardiol*, v. 79, n.3, p.269-76, 2002.

MARMARELIS, P.; MARMARELIS, V. *Analysis of physiological systems, plenum*. New York, 1978.

McCULLOCH, W. S; PITTS, W. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v. 5, p. 115-133, 1943.

MINSKY, M.; PAPERT, S. *Perceptrons: an introduction to computational geometry*. The MIT Press, 1969.

MITZEL, G.; CLANCY, S.; RUGH, W. On transfer function representations for homogeneous nonlinear systems. *IEEE Transactions on Automatic Control*, v. AC-24, p. 242-249, 1979.

NASCIMENTO JR., C. L.; YONEYAMA, T. *Inteligência artificial em controle e automação*. São Paulo: Edgard Blücher, 2000.

PARKER, D. B. *Learning-Logic*. Cambridge: MIT Center for Computational Research in Economics and Management Science, 1985. (Technical Report-47).

PIAGET, J. *A Epistemologia genética - Sabedoria e ilusões da filosofia - Problemas da psicologia genética*. Coleção Os Pensadores. 2.ed., São Paulo: Abril Cultural, 1983.

PINKUS, A. Weierstrass and approximation theory. *J. Approx. Theory* 107, p.1-66, 2000.

PORTER, W. An overview of polynomial system theory. *Proceedings of the IEEE*, v. 64, p. 18-23, 1976.

_____. Approximation by Bernstein systems. *Mathematical Systems Theory*, v. 11, p. 259-274, 1978.

_____.; CLARK, T. Causality structure and the Weierstrass theorem. *Journal of Mathematical Analysis and Applications*, v. 52, p.351-363, 1975.

PRECHELT, L. *Proben1 – A set of neural network benchmark problems and benchmarking rules*. Germany, 1994. (Fakultät für Informatik, Universität Karlsruhe Technical Report 21/94).

PRENTER, P. A Weierstrass theorem for real separable hilbert spaces. *Journal of Approximation Theory*, v. 3, p.341-351, 1970.

RABINER, L. R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. of IEEE*, v.77, n.2, p. 257-286, 1989.

ROOT, W. On the modeling of systems for identification, part I: representations of classes of systems. *SIAM Journal on Control*, v. 13, p. 927-944, 1975.

ROSEMBLATT, F. The Perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, v.65 p.386-408, 1958.

_____. *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. New York: Spartan Books, 1962.

RUDIN, W. *Principles of mathematical analysis*. McGraw-Hill, 1964.

_____. *Real and complex analysis*. McGraw-Hill, 1966.

RUGH, W. J. *Nonlinear system theory: the Volterra/Wiener approach*. The Johns Hopkins University Press, 1981.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations through error propagation. In: RUMELHART, D. E.; McCLELLAND, J. L. e The PDP Research Group (Eds.). *Parallel distributed processing: Experiments in the microstructure of cognition*. v. 1, p. 318-362. Cambridge, MA: MIT Press, 1986.

SCHETZEN, M. Theory of pth-order inverses of nonlinear systems. *IEEE Transactions on Circuits and Systems*, v. CAS-23, p. 285-291, 1976.

_____. *The Volterra and Wiener theories of nonlinear systems*. New York: Wiley & Sons, 1980.

SCHICK, F. *Making choices: a recasting of decision theory*. Cambridge, UK: Cambridge University Press, 1997.

SILVA FILHO, J. F. Estudo da *Modulação Autonômica em Função da Variabilidade da Frequência cardíaca em Hipertensos com e sem Hipertrofia Ventricular Esquerda Concêntrica*. Dissertação (Mestrado em Medicina Interna) – Universidade Federal de Pernambuco, Recife, 2003.

SOCIEDADE BRASILEIRA DE CARDIOLOGIA. IV Diretriz para o uso da Monitoração Ambulatorial da Pressão Arterial e II Diretriz para o uso da Monitoração Residencial da Pressão Arterial. *Arq Bras Cardiol*, v.85, suplemento II, julho 2005.

SOCIEDADE BRASILEIRA DE HIPERTENSÃO. IV Diretrizes Brasileiras de Hipertensão Arterial. *Revista Brasileira de Hipertensão*, São Paulo, v.9, n. 4, p. 359-408, out/dez 2002.

_____. *Aula sobre hipertensão - Plano de Reorganização da Atenção à Hipertensão Arterial e ao Diabetes Mellitus*. Documento criado em 09/06/2003. Disponível em <http://www.sbh.org.br/documentos/index.asp>. Acesso em março de 2005.

_____. *Cartilha do Hipertenso Nº 1* 2001. Disponível em <http://www.sbh.org.br/publico/download/cartilhadohipertenso1.pdf>. Acesso em abril 2005.

TAMAYO, R. P. *La medicina en el renacimiento (siglos XV a XVII)*. http://omega.ilce.edu.mx:3000/sites/ciencia/volumen3/ciencia3/154/htm/sec_. Acesso em 02/03/2005.

TU, J. V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin. Epidemiol* v. 49, n. 11, p. 1225-1231, 1996.

TURBAN, E. *Decision support and expert systems: management support systems*. 4th ed. Englewood Cliffs, N.J.: Prentice Hall, 1995.

VALENÇA, M. J. S. *Análise e projeto de redes neurais com abordagem construtiva para modelagem de sistemas complexos*. Tese (Doutorado em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 1999.

_____. *Aplicando redes neurais: um guia completo*. Olinda-PE, 2005.

_____.; LUDERMIR, T. B. Neural Networks vs. PARMA Modelling: Case studies of river flow prediction. SIMPÓSIO BRASILEIRO DE REDES NEURAI, SBRN, 6. 2000, Rio de Janeiro,RJ. Proceeding... Los Alamitos, California: IEEE Computer Society, p. 113-116, November 2000a.

_____.; LUDERMIR, T. B. Monthly streamflow using an neural fuzzy network model. SIMPÓSIO BRASILEIRO DE REDES NEURAI, SBRN, 6. 2000, Rio de Janeiro,RJ. Proceeding... Los Alamitos, California: IEEE Computer Society, p. 117-119, November 2000b.

_____.; LUDERMIR, T. B. Constructive neural networks in forecasting weekly river flows. INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND MULTIMEDIA APPLICATIONS, ICCIMA, 2001, Yokosuka City, Japan. Proceeding... Published by the IEEE Computer Society, Japan, 2001.

_____.; LUDERMIR, T. B. NeuroInflow: The New Model to Forecast Average Monthly Inflow. SIMPÓSIO BRASILEIRO DE REDES NEURAI, SBRN, 7. 2002, Pernambuco,Brasil. Proceeding... Los Alamitos, California: IEEE Computer Society, p. 74-79, November 2002.

VOLTERRA, V., *Theory of functionals and of integral and integro-differential equations*. New York: Dover Publications, 1959.

WERBOS, P. J. *Beyond regression: new tools for prediction and analysis in the behavioral sciences*. PhD thesis, Harvard University, Cambridge, MA, 1974.

WIDROW, B.; HOFF, M. E. Adaptive switching circuits. *Western Electronic Show and Convention*, Institute of Radio Engineers, 1960.

WIENER, N. *Response of a non-linear device to noise*. Cambridge, MA: Radiation Laboratory, M.I.T., Apr. 1942. (Report No. 129).

WILLEMS, J. *The analysis of feedback systems*. Cambridge, Massachusetts: MIT Press, 1971.

ZAMES, G. Functional analysis applied to nonlinear feedback systems. *IEEE Transactions on Circuit Theory*, v. CT-10, p. 392-404, 1963. 39

_____. Realizability conditions for nonlinear feedback systems. *IEEE Transactions on Circuit Theory*, v. CT-11, p. 186-194, 1964.

**Anexo A SISTEMAS DE CONTROLE AUTO-
ADAPTATIVOS**

Notas de Aula de Sistemas de Controle

Professor Fernando Menezes Campello de Souza, PhD
Departamento de Engenharia Elétrica
Escola de Engenharia
Universidade Federal de Pernambuco

Recife

1974

Apresentação

Estas notas de aula foram escritas com base em notas de aula do curso de especialização em Sistemas de Controle (*Section Spéciale d'Automatique*) que o autor frequentou em Toulouse, na França em 1970/1971, na *École Nationale Supérieure d'Electrotechnique, d'Eletronique, d'Informatique et d'Hydraulique de Toulouse* (ENSEEIH). Elas fazem parte do material que é utilizado na disciplina Sistemas de Controle do curso de Engenharia Elétrica da Universidade Federal de Pernambuco.

1. CLASSIFICAÇÃO DOS SISTEMAS AUTO-ADAPTATIVOS

Krasovskii fez uma classificação dos sistemas auto-adaptativos baseada nas noções de informação inicial, informação de trabalho, precisão particular e precisão generalizada.

1.1 INFORMAÇÃO

Um sistema de controle qualquer tem por base a **informação**, isto é, um conjunto de dados obtidos pela observação, experiência, medidas, etc... Por exemplo, no caso da condução de um automóvel, as reações psicológicas do condutor, e em particular sua visão, são fontes de informação sobre o sistema estrada-automóvel.

O dispositivo de controle do processo utiliza essa informação para agir sobre a parte controlada, mantendo-a nos limites de precisão fixados *a priori*. No exemplo do carro, a visão de uma curva (informação) obriga o condutor a controlar a velocidade e a girar o volante (ação de controle) para que o carro (parte controlada) siga o contorno da estrada sem deixá-la (limites de precisão).

É possível distinguir dois tipos de informação nos sistemas de controle: a **informação inicial** ou *a priori*, que é a informação disponível no ante-projeto, e a informação adquirida durante o funcionamento, ou **informação de trabalho**.

1.1 a) A informação inicial ou *a priori* é relativa à parte controlada. Ela é constituída do conjunto de dados, reunidos *a priori*, sobre as propriedades da parte controlada e dos fatores que agem sobre ela. Por exemplo, no caso da regulação da velocidade de um motor, a informação inicial compreende, entre outras coisas: as características estáticas e dinâmicas da máquina, a faixa de variação dos momentos perturbadores, etc....

A informação inicial praticamente não se renova durante o processo de exploração normal do sistema. Nos sistemas reais ocorre sempre uma degradação ou uma perda da informação inicial ao longo do tempo, em consequência de variações aleatórias, variações que ocorrem sempre nas condições reais (envelhecimento, etc...).

1.1 b) A informação de trabalho elabora-se durante o tempo de exploração do sistema. Os vários blocos de medida do dispositivo de controle servem como fontes de informação de trabalho. É o caso, por exemplo, de um taquímetro numa regulação de velocidade, ou de um termopar numa regulação de temperatura. A informação de trabalho se transmite sob a forma de sinais de comando.

1.2 PRECISÃO

A precisão é um dos índices de qualidade principais de um controle. No caso, por exemplo, de um servomecanismo de posição, a precisão do controle caracteriza o desvio entre a grandeza de saída e a grandeza de entrada.

a) A precisão particular do controle é a precisão obtida para características dadas da parte controlada, para perturbações e uma “mise au point” inicial dadas (informação inicial). Em condições reais a informação inicial não é nunca completa e se degrada ao

longo do tempo, o que faz com que a precisão particular não possa caracterizar completamente a qualidade de um controle.

1.2 b) A precisão generalizada serve para uma avaliação mais completa da qualidade do controle no caso onde a informação inicial é incompleta e no caso de variação das propriedades da parte controlada. A gama de variações da precisão particular para variações do ponto de funcionamento inicial e das características da parte controlada, na presença de perturbação, caracteriza a precisão generalizada do controle.

Quanto maior for a precisão particular, para gamas dadas de variação das propriedades do processo, e uma dada quantidade de informação inicial, maior será a precisão generalizada.

Para uma dada precisão generalizada, a precisão particular será tão grande quanto maior for a faixa de variação das propriedades da parte controlada, e pequena a quantidade de informação inicial.

Por outro lado é evidente que um sistema com grande precisão generalizada é um sistema de baixa sensibilidade.

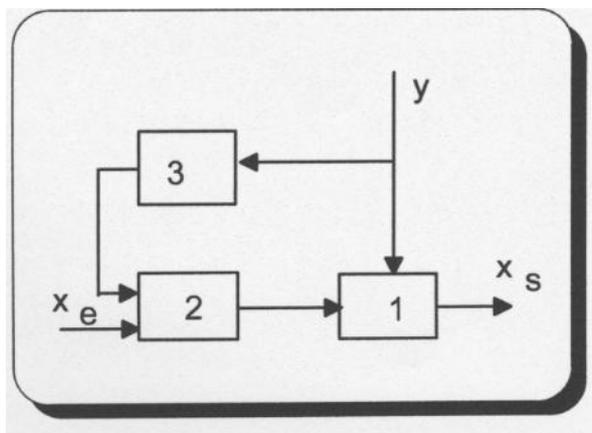
Estas noções de informação (inicial e de trabalho) e de precisão (particular e generalizada) permitem classificar os vários sistemas de controle automático e sistemas adaptativos.

1.3 SISTEMAS DE CONTROLE ORDINÁRIOS

O objetivo principal dos sistemas de controle ordinários é, seja de manter certas grandezas de um processo em níveis fixados a priori (reguladores), seja de forçar uma grandeza a seguir as variações conhecidas de uma outra grandeza (sistemas seguidores). Distingue-se duas classes principais nestes sistemas.

1.3 a) Sistemas de Controle Automático funcionando em Cadeia Aberta

Estes sistemas utilizam o princípio da compensação



O bloco 1 representa a parte controlada sobre a qual agem as perturbações y , causas do desvio entre x_e e x_s . O princípio da compensação consiste em medir y por meio do bloco 3 e em formar com o bloco 2 uma ação de retomo sobre o sistema.

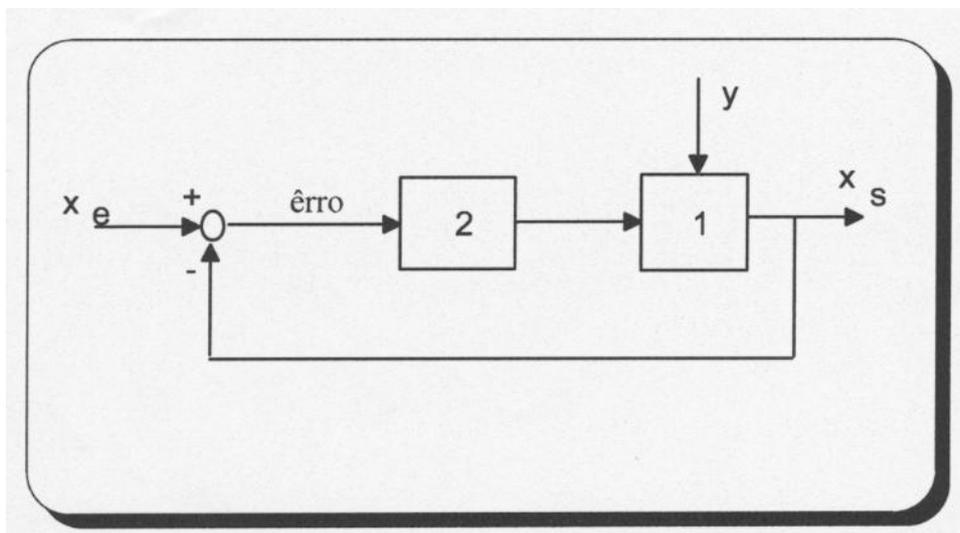
A **informação de trabalho** compõe-se da entrada x_e e da perturbação controlada y .

A informação inicial para este tipo de sistema é sempre muito grande. Para se formar a ação de retomo é necessário ter-se um conhecimento muito preciso das características do sistema, do tipo das perturbações y e da entrada x_e .

A precisão com a qual são conhecidas essas características e medidas essas perturbações, fixa a precisão do controle (precisão particular). Se durante o funcionamento as características do sistema variam, a condição de compensação é violada, ou seja, a precisão generalizada é muito pequena para esse tipo de controle.

1.3 b) Sistemas de Controle Automático funcionando em Cadeia Fechada

Estes sistemas utilizam o erro que existe entre o sinal de saída x_s e o sinal de entrada x_e (informação de trabalho) como reação a uma perturbação y (regulador); ou a uma variação da grandeza de entrada x_e (sistema seguidor).



A informação inicial necessária é, neste caso, menor do que a do caso precedente.

A precisão generalizada pode ser aumentada em grandes proporções, havendo, é claro, limites.

O aumento da precisão generalizada pode ser feito por um aumento do ganho, respeitando as condições de estabilidade do sistema, o que demanda um aumento da quantidade de informação inicial. Com efeito, à medida que se aumenta o ganho, aparece a influência dos pequenos parâmetros, das não-linearidades, desprezados de início no modelo, isto é, a “estrutura fina” do sistema. Sem o conhecimento dessas características, é impossível garantir o aumento do ganho sem prejuízo da estabilidade. Ademais, os pequenos parâmetros são submetidos a variações aleatórias.

Em resumo pode-se dizer que os sistemas ordinários de controle possuem uma precisão particular grande e uma precisão generalizada média, a partir de uma grande quantidade de informação inicial, e de informação de trabalho.

Para se aumentar a precisão generalizada dos sistemas de controle automático é necessário utilizar-se outros tipos de sistemas: os sistemas auto-adaptativos.

1.4 SISTEMAS DE CONTROLE AUTO-ADAPTATIVOS

Estes são sistemas cuja precisão generalizada elevada, para uma informação inicial incompleta, é assegurada a partir de testes e de medidas sobre o processo, por ação contínua ou discreta sobre este processo.

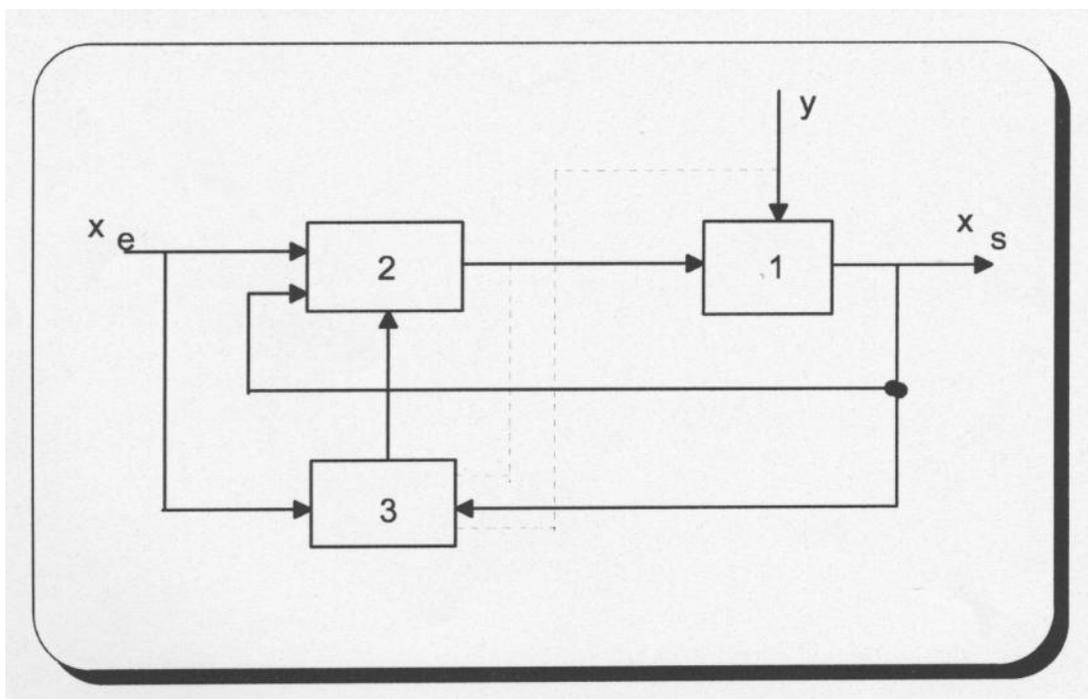
Retome-se o exemplo do carro: para manter seu carro sobre a estrada qualquer que seja o estado desta, o motorista dá pequenos golpes no volante (teste) com o intuito de conhecer as características dinâmicas de seu dispositivo de controle.

Num sistema auto-adaptativo, a compensação da diminuição de informação inicial, e o aumento da precisão generalizada são obtidos por um crescimento da informação de trabalho.

Este aumento da quantidade de informação de trabalho é obtido por intermédio de malhas de retroação sobre o sistema, ou apenas uma parte desse sistema, com ajuda de ações de pesquisa ou de ensaios.

O caráter mais recente dos sistemas auto-adaptativos não permite uma classificação rigorosa dos mesmos. Entretanto pode-se distinguir certos grandes grupos que, graças às noções de informação inicial e de trabalho, podem ser classificados como explicado a seguir.

1.4.1 Os Sistemas Auto-Reguladores, nos quais uma regulagem do elemento de controle é efetuada com o intuito de tornar ótimo um certo critério, a informação sobre este critério sendo obtida a partir de testes e de medidas sobre o sistema. Considere-se, por exemplo, o esquema da figura abaixo.



O bloco de controle 2 age sobre o sistema 1 com o intuito de fazer com que o valor de x_s seja aderente ao valor de x_e da entrada (regulador ou sistema seguidor ordinário).

O bloco de controle 3 age sobre os parâmetros do bloco de comando 2 com o intuito de otimizar um certo critério de qualidade do regulador (1,2), por exemplo, minimização do quadrado do erro entre x_s e x_e , na presença de perturbações y e de modificações das características de 1, não conhecidas a priori.

Aqui a informação inicial pode ser inferior àquela necessária para o regulador (1-2). Não é necessário conhecer as perturbações y , por exemplo. Este tipo de sistema satisfaz portanto à definição de sistema auto-regulador.

Os sistemas auto-reguladores podem ser classificados em três categorias, segundo o seu princípio de funcionamento:

1.4.1 a) Os Sistemas de Regulação Extremal que têm a propriedade de agir seja sobre as variáveis de entrada seja sobre os parâmetros de um sistema, de maneira a tornar ótimo (extremal) um índice de qualidade, seja intrínseco ao sistema (ex: procura da ressonância num circuito oscilante), seja extrínseco ao sistema (ex: mínimo de um custo operatório, erro quadrático mínimo, ...), isto quaisquer que sejam as perturbações exteriores e as variações das características do sistema, desconhecidas a priori.

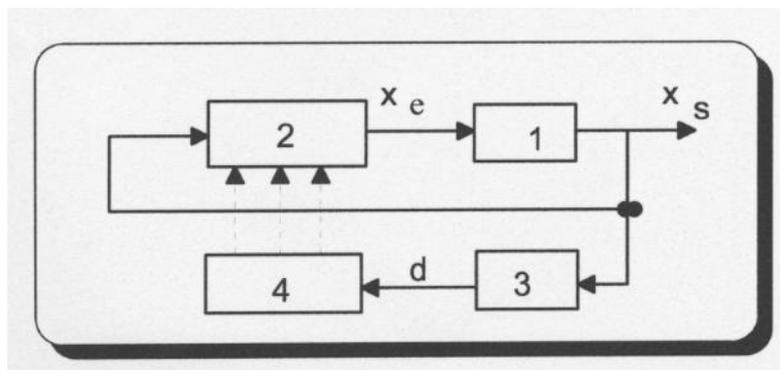
Dois detalhes mostram que esses sistemas pertencem à classe de sistemas adaptativos; não se tem necessidade de um conhecimento preciso da característica extremal, isto é, a informação inicial sobre esses sistemas não precisa ser completa. Estes sistemas utilizam um movimento de pesquisa como meio de obter informação complementar (informação de trabalho).

1.4.1 b) Os Sistemas com Blocos de Correção Adaptativos, que permitem garantir a estabilidade e um índice de qualidade do processo (ver exemplo abaixo).

1.4.1 c) Os Sistemas Auto-Otimizantes, idênticos aos precedentes, mas com uma duração de otimização inferior à duração do regime transitório do sistema.

1.4.2 Os Sistemas a Auto-Aprendizagem: são sistemas que aperfeiçoam o seu algoritmo de controle, analisando, durante um intervalo de tempo, testes e medidas sobre o processo controlado, para levar o índice de ótimo a um valor admissível ou extremal.

Considere-se, por exemplo, um sistema representado pelo diagrama de blocos abaixo. Trata-se de um regulador extremal (2) agindo sobre o sinal de entrada x_e do sistema 1 de maneira a tomar ótima a grandeza de saída x_s .



Pelo seu funcionamento o regulador impõe que a grandeza de saída oscile no entorno de um valor médio x_{se} . Seja d a amplitude dessas oscilações. O bloco 3 mede e coloca na memória o valor d desta amplitude. O bloco 4 tem por função minimizar d , modificando o algoritmo de trabalho do regulador extremal 2.

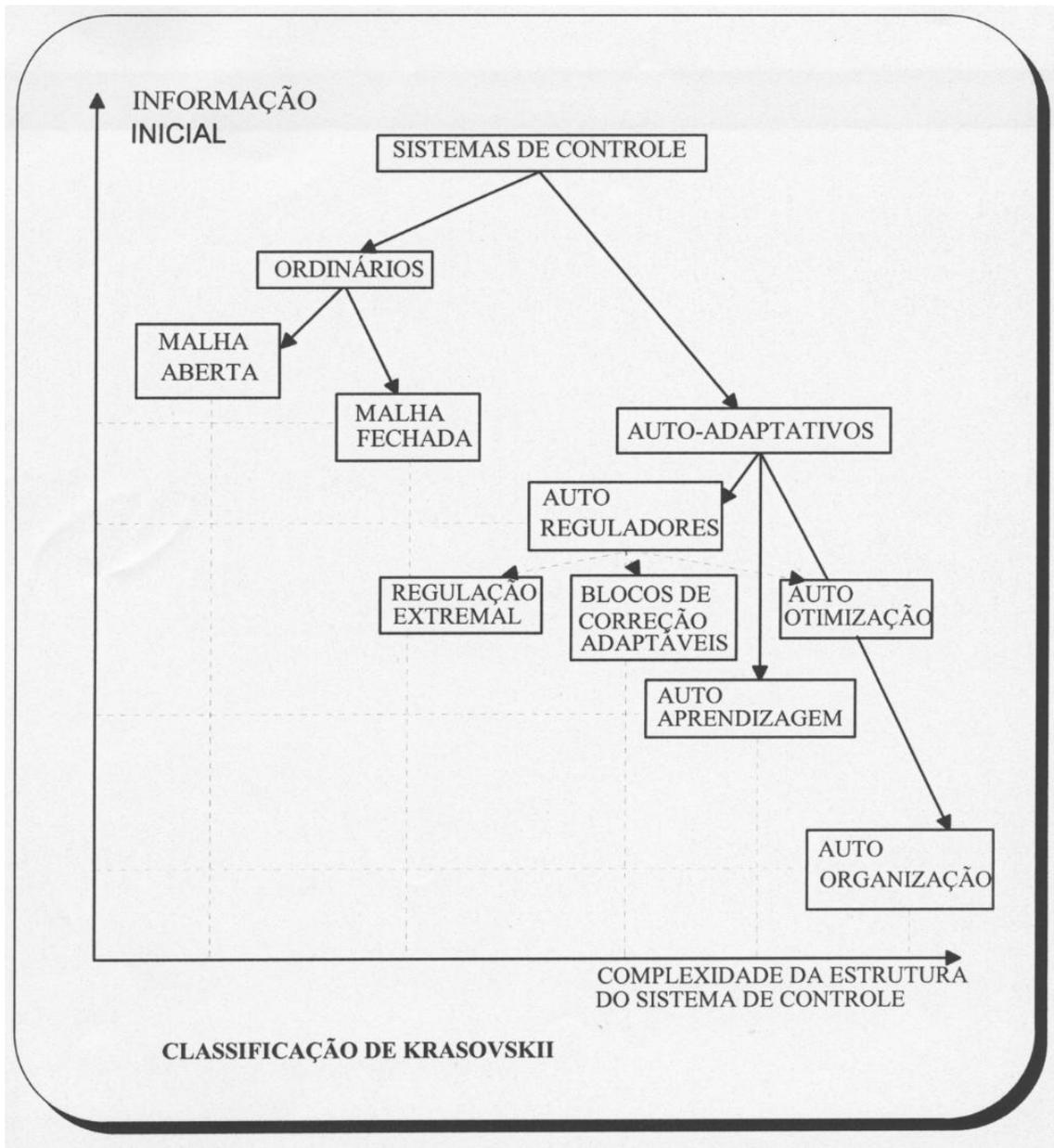
A informação inicial, no caso de um sistema a auto-aprendizagem, é inferior àquela necessária para um sistema auto-regulador, pois o sistema adapta seu algoritmo de trabalho (variação de estrutura do bloco de controle) em função das características do sistema controlado. Em resumo, um tal sistema determina e memoriza sua qualidade de adaptação.

1.4.3 Os Sistemas a Auto-Organização, sistemas capazes de conservar sua capacidade de trabalho sob influência de perturbações bastante grandes para modificar o caráter de funcionamento e a estrutura da parte controlada. Eles são levados então a modificar não apenas os parâmetros de regulagem da parte do controle (controlador) mas também sua estrutura. Frequentemente a influência dessas grandes perturbações obriga a modificar o critério a otimizar.

Acham-se exemplos de auto-organização nos sistemas biológicos.

A informação inicial neste caso é mais reduzida, mas por outro lado a complexidade do sistema é muito grande.

É possível resumir esta classificação sob forma de diagrama, fazendo aparecer a diminuição da formação inicial utilizada assim como o aumento da complexidade nos sistemas.



2. CLASSIFICAÇÃO DOS SISTEMAS DE CONTROLE PROPOSTA POR SOLODOVNIKOV

2.1 ENUNCIADO DO PROBLEMA GERAL DO CONTROLE AUTOMÁTICO

Considere-se um sistema de controle automático que pode ser dividido em bloco de controle e parte controlada, ou controlador e sistema controlado.

A informação disponível sobre o sistema controlado pode ser caracterizada pelo vetor de estado $x(t)$ do sistema ($x(t) \in R^n$). Este vetor de estado tem por componentes as variáveis de estado do sistema,

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

As ações que fazem variar o vetor de estado, bem como as suas derivadas são:

I - As perturbações (influência do meio exterior sobre o sistema controlado) representadas pelo vetor $f(t)$.

II - As variações incontroláveis das propriedades dinâmicas do sistema controlado.

III - O vetor de controle $u(t)$ que varia num certo domínio, limitado pelas restrições impostas.

O objetivo do controle será de tornar ótimo um índice de qualidade (índice de desempenho, funcional objetivo, função custo, etc ...)

$$J(x, f, u, t)$$

fazendo intervir o vetor de estado, o vetor perturbação, o vetor de controle (comando) e o tempo, ou pelo menos garantir a melhor aproximação possível.

É preciso pois achar e realizar (sintetizar) a relação entre u e x (para que o controle seja automático) que, levando em conta f , garanta com uma aproximação suficiente a condição

$$J(x, u, f, t) = \text{Extremo} \quad (1)$$

Seja $u_0(t) = u_0(x, f, t)$ o vetor de controle que realiza esta condição. Este vetor é chamado de controle ótimo, ou algoritmo ideal.

Na prática procura-se uma realização aproximada de $u_0(t)$ sabendo-se que é o limite para o qual se deve tender quando se faz o projeto de um sistema de controle.

A realização de um controle $u(t)$ corresponde a um valor J_r do índice de qualidade diferente do valor ideal $J_0(t)$. O erro

$$\Delta = J_0(t) - J_r(t) \quad (2)$$

é chamado índice de precisão do controle.

O índice que mede a qualidade do controle realizável é um funcional

$$Q = \Phi(\Delta) \quad (3)$$

que se deve otimizar (extremalizar) no estágio de ante-projeto. Entretanto, essa otimização não é, em geral, compatível com as exigências de confiabilidade de funcionamento para o sistema controlado. De fato, quanto maior é a precisão do controle, mais complexo e volumoso é o controlador, e portanto, menos confiável.

Para levar em conta a confiabilidade no estágio de ante-projeto, introduz-se um termo R no índice de qualidade:

$$P=Q+\lambda R \quad (4)$$

Finalmente, o objetivo do controle será tornar mínimo o índice de qualidade:

$$P= \Phi(\Delta)+\lambda R \quad (5)$$

Isto é realizado em três etapas:

1 - Determinação do algoritmo ideal de controle $u_0(t)$ correspondente à $J_0(t)$; o problema correspondente é chamado de otimização primária.

2 - Determinação do regime ótimo, racional, por extremalização de $Q=\Phi(\Delta)$ chamada otimização secundária.

3 - Síntese tecnológica e realização efetiva, para a qual, eventualmente, ter-se-á conta de (5).

2.2 DEFINIÇÃO DA NOÇÃO DE VARIAÇÃO CONTROLADA

Chamam-se “variações controladas” as variações das propriedades do controlador (parâmetros, ações de controle, algoritmos de controle, estrutura, etc ...) que ele produz visando realizar o objetivo do controle.

2.3 CLASSIFICAÇÃO

Entre os sistemas de controle automático pode-se considerar duas grandes classes:

- os sistemas não adaptativos, e
- os sistemas adaptativos.

Os sistemas não adaptativos não produzem variações controladas. Toda a informação inicial é usada para realizar o controle ótimo. Durante o funcionamento as propriedades, escolhidas a priori, do controlador, não podem variar a não ser de uma maneira incontrolável. Este é o caso dos sistemas de controle ordinários, nos quais o esquema e os parâmetros das redes compensadoras (corretoras) são escolhidos em função da formação inicial sobre o sistema, e são fixados no estágio conceitual.

Os sistemas adaptativos podem produzir variações controladas do bloco de controle. São esses sistemas que Solodovnikov classificou, segundo três critérios, levando em consideração:

- a) a natureza das variações controladas;
- b) os meios que permitem realizar essas variações controladas;
- c) os meios pelos quais essas variações controladas garantem o objetivo do controle.

2.3.1 Classificação baseada no Primeiro Critério

- **Sistemas Auto-Reguladores:** as variações controladas das propriedades do bloco de controle (controlador) são variações dos seus parâmetros ou das ações de controle.

- **Sistemas a Auto-Organização:** as variações controladas são variações de estrutura do controlador, além das variações dos parâmetros ou das ações de controle.

- **Sistemas a Auto-Algoritmo:** as variações controladas são variações do algoritmo de trabalho, além das variações de estrutura, de parâmetros, e de ações de controle.

No que conceme os sistemas a auto-organização, nota-se aqui uma diferença com a classificação que se inspira no trabalho de Krasovskii.

2.3.2 Classificação baseada no Segundo Critério

- **Sistemas passivos:** a variação controlada é uma função programada elaborada à partir da informação inicial e das condições de trabalho do sistema. Essa função programada é definida *a priori* para realizar o regime ótimo. Nestes sistemas, como nos sistemas não adaptativos, a resolução do problema de otimização primária ou secundária só se pode fazer no estágio de ante-projeto, o que supõe uma quantidade importante de informação inicial.

- **Sistemas Ativos:** a variação controlada é realizada não somente em função da informação inicial, mas também em função da informação de trabalho. As variações controladas não são conhecidas *a priori*.

Estes sistemas se subdividem, por sua vez, em sistemas a “ciclo aberto de adaptação”, nos quais não há auto-análise dos efeitos das variações controladas, e em sistemas a “ciclo fechado de adaptação” que compreendem malhas fechadas complementares no diagrama de blocos, que servem para analisar o efeito das variações controladas.

Um sistema ativo “a ciclo aberto” dispõe dos meios para resolver o problema da otimização primária, mas não permite resolver o problema da otimização secundária como o faz um sistema “ativo a malha fechada”.

2.3.3 Classificação baseada no Terceiro Critério

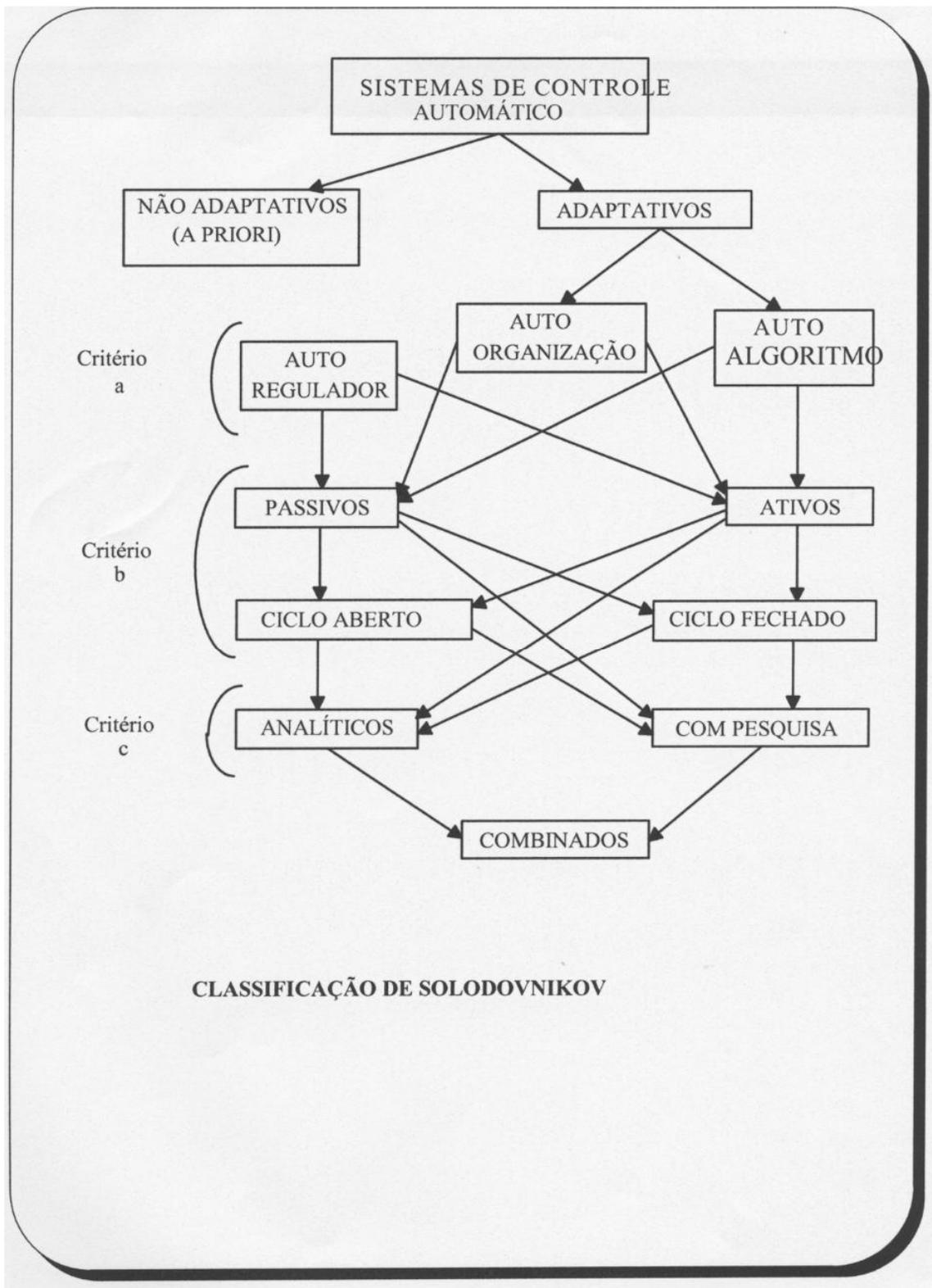
- **Sistemas Auto-Adaptativos Analíticos:** as variações controladas são o resultado de uma determinação analítica das condições de ótimo da relação (1), isto é, de um cálculo.

- **Sistemas utilizando Sinais de Pesquisa:** as variações controladas são o resultado de uma pesquisa, o que supõe um aumento da quantidade de informação de trabalho. É vantajoso utilizar um sistema com pesquisa em todos os casos onde o objetivo do controle é realizar condições ótimas de trabalho sob forma estática, ou ainda

quando as variações das características estáticas são lentas em relação ao tempo de pesquisa (exemplo: regulação extremal).

Os sistemas analíticos são usados principalmente nos casos onde há variações das propriedades dinâmicas do sistema controlado.

Essa classificação de Solodovnikov pode ser representada pelo esquema seguinte.



Apêndice A TEOREMAS

A.1 DEFINIÇÃO 1

$\sigma : R \rightarrow R$ é chamada uma função sigmóide generalizada, se os limites

$$\begin{cases} \lim_{x \rightarrow -\infty} \sigma(x) = 0 \\ \lim_{x \rightarrow \infty} \sigma(x) = 1 \end{cases}$$

Nesse caso, não é necessário que σ seja contínua ou monotônica, de maneira que:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$\sigma(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$

A.2 DEFINIÇÃO 2

Se uma função contínua é definida em R^n , e $\lim_{|x| \rightarrow \infty} f(x)$ existe, então $f(x)$ é dita como sendo uma função contínua no espaço estendido R^n (denotado como \bar{R}^n) e o conjunto de todas as funções contínuas definidas no espaço estendido \bar{R}^n pode ser escrito como $C(\bar{R}^n)$, isto é:

$$C(\bar{R}^n) = \{f \in C(R^n) : \lim_{|x| \rightarrow \infty} f(x), \text{ existe}\}$$

A.3 TEOREMA 1

Teorema de Chen et al.: Se $\sigma(x)$ é uma função sigmóide limitada (generalizada) e $f(x)$ é uma função contínua sobre o intervalo $(-\infty, \infty)$, tal que $\lim_{x \rightarrow -\infty} f(x) = A$ e $\lim_{x \rightarrow \infty} f(x) = B$, em que A e B são constantes, então para algum $\varepsilon > 0$ existe N, c_i , y_i e θ_i , tal que

$$\left| f(x) - \sum_{i=1}^N c_i \sigma(y_i x + \theta_i) \right| < \varepsilon$$

É importante observar que, de acordo com esse teorema, a única restrição imposta à função sigmóide é de que ela seja limitada. Para a arquitetura proposta para as redes NSRBN vide capítulo 4, o uso de funções sigmóides satisfaz, portanto, a esse requisito, uma vez que as funções geradas pelas potências da função sigmóide continuam sendo funções sigmóides limitadas.

A.4 TEOREMA 2

Teorema de Chen et al.: suponha que $g(x) \in C_0(\mathbb{R}^1)$, $g(x) \neq 0$; então o conjunto de $\sum_{i=1}^N c_i g(y_i \cdot x + \theta_i)$ é denso em $C_0(\mathbb{R}^n)$, em que $y_i \in \mathbb{R}^n$, $y_i \neq 0$, $c_i, \theta_i \in \mathbb{R}^1$ ($i = 1, \dots, N$) e $C_0(\overline{\mathbb{R}^n}) = \{f \in C(\overline{\mathbb{R}^n}) : \lim_{|x| \rightarrow \infty} f(x) = 0\}$.

O mérito desse teorema está no fato de permitir explorar outras funções de ativação.

Seja, por exemplo, um *spline* de grau k $B_k(x) = \sum_{i=0}^k c_k^i (-1)^i (x-i)_+^{k-1}$; esta função satisfaz

a todas as condições deste teorema. Logo, $\sum_{i=1}^N c_i B_k(y_i \cdot x + \theta_i)$ é denso em $C_0(\mathbb{R}^n)$.

Similarmente, qualquer *wavelet* também satisfaz este teorema e muitas outras funções, portanto, estas podem também ser utilizadas para aproximar funções contínuas em $C_0(\mathbb{R}^n)$.

Em termos da arquitetura proposta para as redes NSRBN vide capítulo 4, este teorema permite concluir que: qualquer função que o satisfaça e cujas potências continuem satisfazendo este mesmo teorema, podem ser utilizadas como função de ativação da rede.

Apêndice B ALGORITMO REDES NSRBN

O algoritmo pode ser resumido da seguinte forma:

- 1) Passo inicial: escolha dois critérios de seleção (chamados de primeiro e segundo critérios de seleção) adequados ao problema: um para determinação dos pesos e outro para estabelecimento da complexidade (mínima tendência, regularidade, critério combinado (Mandala & Ivakhnenko,1994). Definir critérios de parada adequados como número máximo de épocas, valor mínimo para o critério de seleção e limiar a partir do qual o acréscimo de um bloco não acarreta benefício adicional;
- 2) Adicionar o bloco de grau p (inicialmente $p=1$);
- 3) Utilizando conjuntos de treinamento e validação, para todos os exemplos:
 - Calcular a saída da rede;
 - Otimizar os pesos utilizando-se o primeiro critério de seleção (mínima tendência como sugestão). Técnicas eficientes como gradientes conjugados e Levenberg-Marquardet, podem ser utilizadas, para otimização dos pesos. O uso do tradicional algoritmo *backpropagation* fornece também resultados satisfatórios;
- 4) Ao fim de cada época, calcular os critérios de seleção;
- 5) Verificar os critérios de parada para otimização dos pesos (critério de seleção1). Se os critérios de parada não são satisfeitos, retornar ao passo 3. Caso contrário, continue com o passo 6;
- 6) Em função do segundo critério externo de seleção, definido no início (regularidade, por exemplo), calcular o valor para o critério de parada (valor para o qual o bloco adicionado não acarreta benefício adicional considerado significativo para o segundo critério de seleção);

Caso o ganho obtido seja superior ao limiar fazer $p=p+1$ e retornar ao passo 2 (acrescentando-se outro bloco), repetindo-se todo o processo. Caso contrário, o modelo de complexidade ótima terá sido obtido e o processo estará terminado.

Apêndice C CRITÉRIOS DE SELEÇÃO NO ALGORITMO CONSTRUTIVO PARA REDES NSRBN

Os modelos construtivos procuram resolver: o problema de estimar os parâmetros e a seleção da arquitetura ótima do modelo. De maneira geral, cada modelo utiliza dois critérios: um critério interno com base no qual se estimam os coeficientes do modelo e um critério externo para a seleção da estrutura ótima do modelo. Esses critérios de seleção são também chamados de função objetivo ou função erro. Em função da importância desses critérios de seleção, serão definidos a seguir alguns critérios que podem ser utilizados.

Considerando-se que se dispõe de um conjunto de dados com N pontos que será dividido em três subconjuntos A (N_A valores), B (N_B valores), e C (N_C valores).

C.1 CRITÉRIO DE REGULARIDADE (CR)

Esse critério consiste no cálculo da função objetivo (erro médio quadrático, por exemplo) baseado no conjunto de validação B, para selecionar o modelo ótimo e o critério de parada (Madala & Ivakhnenko, 1994). Entretanto, os pesos são obtidos pelo uso do conjunto de treinamento A:

$$CR = \frac{1}{N_B} \sum_{n=1}^{N_B} (Z_p^B - Z_o^B)^2 \quad (C.1)$$

em que Z_p^B - é o valor previsto para o conjunto B; Z_o^B - são os valores observados no conjunto B, \bar{Z}_o^B é o valor médio dos valores ocorridos para o conjunto B, \bar{Z}_p^B é o valor médio dos valores previstos para o conjunto B e N_B - é o número de valores do conjunto B (conjunto de validação). Outras formas para o critério de regularidade são:

$$CR = \frac{\sum_{n=1}^{N_B} (Z_o^B - Z_p^B)^2}{\sum_{n=1}^{N_B} (Z_o^B - \bar{Z}_o^B)^2} \quad (C.2)$$

$$CR = \frac{1}{N_B} \sum_{n=1}^{N_B} (Z_p^B - Z_o^B)^2 / \sum_{n=1}^{N_B} (Z_o^B)^2 \quad (C.3)$$

C.2 CRITÉRIO DE MÍNIMA TENDÊNCIA OU CRITÉRIO CONSISTENTE (CMT)

Esse critério tem por objetivo minimizar a diferença entre os valores calculados para os conjuntos A e B (Madala & Ivakhnenko, 1994).

$$CMT = \sum_{n=1}^{N_{A+B}} (Z_C^A - Z_C^B)^2 / \sum_{n=1}^{N_{A+B}} (Z_o^{A+B})^2 \quad (C.4)$$

em que Z_C^A - são os valores calculados tendo como conjunto de validação o conjunto A (o conjunto B serve para ajustar os pesos); Z_C^B - são os valores calculados tendo como conjunto de validação o conjunto B (o conjunto A é usado para estabelecer os pesos); Z_o^{A+B} - são os valores ocorridos para os conjuntos A e B. Este critério é bastante útil para dados que apresentam ruídos. Outras formas desse critério são:

$$CMT = \frac{\sum_{n=1}^{N_{A+B}} (Z_C^A - Z_C^B)^2}{\sum_{n=1}^{N_{A+B}} (Z_o^{A+B} - \bar{Z}_o^{A+B})^2} \quad (C.5)$$

$$CMT = \frac{1}{N_{A+B}} \sum_{n=1}^{N_{A+B}} (Z_C^A - Z_C^B)^2 \quad (C.6)$$

onde: \bar{Z}_o^{A+B} é a média dos valores ocorridos para os conjuntos A e B, Z_o^{A+B} são os valores observados para os conjuntos A e B e N_{A+B} - é o número de valores contidos nos conjuntos A e B.

C.3 CRITÉRIO DE PREDIÇÃO (CP)

Quando o conjunto de dados é suficiente para ser dividido em três subconjuntos (A, B e C) este critério é calculado sobre o conjunto C para selecionar o modelo ótimo (Madala&Ivakhnenko, 1994). O conjunto A é utilizado para determinar os pesos e o conjunto B para estabelecer a arquitetura. A função objetivo pode ser a mesma do critério de regularidade sendo que calculada sobre o conjunto C.

$$CP = \frac{1}{N_C} \sum_{n=1}^{N_C} (Z_p^C - Z_o^C)^2 / \sum_{n=1}^{N_C} (Z_o^C)^2 \quad (C.7)$$

C.4 CRITÉRIO COMBINADO (CC)

Em muitas aplicações práticas, pode-se necessitar de um modelo que atenda a mais de um critério (Madala & Ivakhnenko, 1994). O critério combinado, como o próprio nome diz, é o critério que cuja função objetivo é uma função composta de dois ou mais critérios. Suponha que se esteja interessado em desenvolver um modelo que apresente boas características de predição e ao mesmo tempo incorpore características de um sistema com mínima tendência. Nestas circunstâncias, pode-se adotar uma função objetivo combinando estes critérios por um fator de ponderação. Esse critério tem sua representação geral dada pela seguinte expressão:

$$CC = \alpha.CMT + (1 - \alpha)CP \quad (C.8)$$

na qual α é o fator de ponderação.

C.5 CRITÉRIO DE BALANÇO DE VARIÁVEIS (CBV)

Esse critério é útil para previsões de longo prazo. Suponha o caso de uma série temporal com sazonalidade mensal para a qual se deseja realizar previsão de longo prazo, ou seja, doze meses à frente (Madala & Ivakhnenko, 1994). Esse critério tem por objetivo minimizar a diferença entre os valores calculados para o modelo global e para os modelos individuais. Nesse caso esse critério tem sua representação dada pela seguinte expressão:

$$CBV = \sum_{p=1}^{N_C} \left[\left(Y_p^C - \frac{1}{12} (y_{p1}^C - y_{p2}^C - \dots - y_{p12}^C) \right) \right]^2 \quad (C.9)$$

em que Y_p^C - é a previsão feita para a média anual; $y_{p1}^C, y_{p2}^C, \dots, y_{p12}^C$ - são os valores previstos para cada mês do ano.

Portanto, esse critério tem por objetivo minimizar a diferença entre a previsão realizada para a média anual e a média calculada por intermédio das previsões mensais.

Apêndice D ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS

Tabela D.1 - Estatística Descritiva da Amostra - Caso HVE

Variável	N	Média	Mediana	Soma	Mín.	Máx.	Desv.Pad.
SEXO	101	0,48	0	48,00	0,00	1,00	0,502
IDADE	101	53,18	53	5371,00	21,00	81,00	15,096
IDOSO	101	0,28	0	28,00	0,00	1,00	0,450
PESO	101	73,24	73	7397,50	48,00	115,00	14,790
ALTURA	101	1,64	1,64	165,20	1,40	1,85	0,098
IMC	101	27,27	26,99	2754,18	18,07	40,31	4,430
CINTURA	101	93,12	93	9405,00	70,00	121,00	11,820
QUADRIL	101	101,01	99	10202,00	84,00	149,00	10,873
CINT/QUAD	101	0,92	0,92	93,16	0,77	1,07	0,072
HIPERTENSO	101	0,73	1	74,00	0,00	1,00	0,445
PSC1M	101	150,31	146	15181,00	112,00	228,00	23,249
PDC1M	101	91,43	90	9234,00	58,00	132,00	15,143
FCC1M	101	76,54	76	7731,00	48,00	118,00	14,993
PSC2M	101	147,76	146	14924,00	106,00	220,00	21,855
PDC2M	101	90,15	88	9105,00	54,00	130,00	14,479
FCC2M	101	78,43	78	7921,00	49,00	115,00	15,130
PSC3M	101	143,87	140	14531,00	104,00	216,00	21,147
PDC3M	101	87,66	86	8854,00	56,00	132,00	14,081
FCC3M	101	77,24	77	7801,00	35,00	119,00	15,602
PSCmédia-Consult.	101	147,31	144	14878,67	110,00	221,33	21,147
PDCmédia-Consult.	101	89,75	88	9064,33	58,00	131,33	13,875
FCCmédia-Consult.	101	77,40	77,33	7817,67	49,00	113,67	14,367
HVE-SOKOLOW-ECG	100	0,12	0	12,00	0,00	1,00	0,327
HVE-CORNELL-ECG	100	0,19	0	19,00	0,00	1,00	0,394
HVE-CORNELLContínuo	100	19,25	18,5	1924,50	4,00	53,00	8,239
INDMASSA-VE	101	121,08	106,9	12229,40	62,00	251,50	40,140
HVE-Ecocardiógrama	101	0,48	0	48,00	0,00	1,00	0,502
VOLMASSA-VE	101	0,56	0,59	56,80	0,30	0,98	0,143
PSTMEDM1	101	135,06	131,17	13641,39	106,14	176,04	14,834
PDTMEDM1	101	81,89	80,17	8270,48	56,68	113,75	10,889
PULSTM1	101	53,18	50,73	5370,91	32,75	89,66	11,104
MAPTM1	101	100,40	98,13	10140,76	79,97	131,93	11,159
FCTMEDM1	101	75,70	76,19	7645,43	49,92	100,68	12,184
PSTMEDM2	101	132,92	131,06	13424,42	112,02	179,08	13,480
PDTMEDM2	101	80,83	78,73	8163,79	55,44	118,47	10,878
PULSTM2	101	52,09	49,79	5260,62	34,11	92,67	10,598
MAPTM2	101	98,90	98,54	9989,22	78,21	138,78	10,525

Tabela D.1 - Estatística Descritiva da Amostra – Caso HVE(continuação)

Variável	N	Média	Mediana	Soma	Mín.	Máx.	Desv.Pad.
FCTMEDM2	101	76,58	76,15	7734,67	50,11	102,80	12,426
PSTMEDM3	101	133,40	132,08	13473,23	107,61	175,88	13,489
PDTMEDM3	101	80,81	78,73	8161,41	56,86	106,54	10,534
PULSTM3	101	52,59	49,79	5311,82	33,17	87,04	10,498
MAPTM3	101	99,10	97,87	10008,88	80,52	125,37	10,257
FCTMEDM3	101	76,47	75,88	7723,72	51,56	101,68	12,005
PSTMED-3mapas	101	133,79	132,18	13513,01	109,42	176,70	13,313
PDTMED-3mapas	101	81,17	79,48	8198,56	56,32	110,38	10,457
PULTMED-3mapas	101	52,62	50,34	5314,45	33,43	89,79	10,492
MAPTMED-3mapas	101	99,47	98,69	10046,29	81,23	128,85	10,204
FCTMED-3mapas	101	76,25	76,04	7701,28	51,06	101,72	11,694
DCNOPSM1	101	7,27	7,58	734,59	-13,35	23,64	7,128
DCNOPDM1	101	11,69	11,82	1181,18	-8,72	33,42	8,544
DCNOPSM2	101	7,37	7,37	744,06	-14,69	25,21	6,620
DCNOPDM2	101	12,16	11,93	1228,54	-6,09	35,25	8,173
DCNOPSM3	101	6,89	6,67	695,53	-14,54	19,16	6,444
DCNOPDM3	101	11,33	10,60	1144,42	-14,65	30,76	7,741
DCNOPS-Méd 3 MAPAs	101	7,18	7,45	724,73	-14,19	20,98	6,199
DCNOPD-Méd 3 MAPAs	101	11,73	11,07	1184,71	-9,08	29,33	7,497

Segue um glossário relativo as variáveis da Tabela D.1:

- N = número de indivíduos da amostra;
- Mín. = valor mínimo;
- Máx. = valor máximo;
- SEXO = percentual de indivíduos do sexo masculino;
- IDADE = em anos;
- IDOSO = percentual de indivíduos maiores de 65 anos;
- PESO = em quilogramas;
- ALTURA = em centímetros;
- IMC = índice de massa corpórea (Kg/m^2);
- CINT/QUAD = relação cintura quadril;
- HIPERTENSO = percentual de indivíduos hipertensos;
- PSC1M = pressão sistólica medida em consultório na primeira consulta;
- PDC1M = pressão diastólica medida em consultório na primeira consulta;
- FCC1M = frequência cardíaca medida em consultório na primeira consulta;
- PSC2M = pressão sistólica medida em consultório na segunda consulta;
- PDC2M = pressão diastólica medida em consultório na segunda consulta;
- FCC2M = frequência cardíaca medida em consultório na segunda consulta;

- PSC3M = pressão sistólica medida em consultório na terceira consulta;
- PDC3M = pressão diastólica medida em consultório na terceira consulta;
- FCC3M = frequência cardíaca medida em consultório na terceira consulta;
- PSCmédia-Consult. = média das pressões sistólicas medidas em consultório nas três consultas;
- PDCmédia-Consult. = média das pressões diastólicas medidas em consultório nas três consultas;
- FCCmédia-Consult. = média das frequências cardíacas medidas em consultório nas três consultas;
- HVE-SOKOLOW-ECG = critério eletrocardiográfico de Sokolow-Lyon para hipertrofia ventricular esquerda (0=não tem; 1=tem) em milivolts;
- HVE-CORNELL-ECG = critério eletrocardiográfico de Cornell Voltagem para hipertrofia ventricular esquerda (0=não tem; 1=tem) em milivolts;
- HVE-CORNELLContínuo = critério eletrocardiográfico de Cornell Voltagem para hipertrofia ventricular esquerda em valores contínuos (milivolts);
- INDMASSA-VE = índice de massa ventricular esquerda calculado pelo ecocardiograma (g/m^2);
- HVE-Ecocardiograma = percentual de indivíduos com hipertrofia ventricular esquerda, diagnosticada pelo ecocardiograma;
- VOLMASSA-VE = relação volume-massa ventricular esquerda calculada pelo ecocardiograma;
- PSTMEDM1 = pressão sistólica média em 24 horas medida na MAPA 1;
- PDTMEDM1 = pressão diastólica média em 24 horas medida na MAPA 1;
- PULSTM1 = pressão de pulso média em 24 horas medida na MAPA 1;
- MAPTM1 = média da pressão arterial média em 24 horas medida na MAPA 1;
- FCTMEDM1 = frequência cardíaca média em 24 horas medida na MAPA 1;
- PSTMEDM2 = pressão sistólica média em 24 horas medida na MAPA 2;
- PDTMEDM2 = pressão diastólica média em 24 horas medida na MAPA 2;
- PULSTM2 = pressão de pulso média em 24 horas medida na MAPA 2;
- MAPTM2 = média da pressão arterial média em 24 horas medida na MAPA 2;
- FCTMEDM2 = frequência cardíaca média em 24 horas medida na MAPA 2;
- PSTMEDM3 = pressão sistólica média em 24 horas medida na MAPA 3;
- PDTMEDM3 = pressão diastólica média em 24 horas medida na MAPA 3;

-
- PULSTM3 = pressão de pulso média em 24 horas medida na MAPA 3;
 - MAPTM3 = média da pressão arterial média em 24 horas medida na MAPA 3;
 - FCTMEDM3 = frequência cardíaca média em 24 horas medida na MAPA 3;
 - PSTMED-3mapas = média da pressão sistólica média em 24 horas das três MAPAs;
 - PDTMED-3mapas = média da pressão diastólica média em 24 horas das três MAPAs;
 - PULSTMED-3mapas = média da pressão de pulso média em 24 horas das três MAPAs;
 - MAPTMED-3mapas = média da pressão arterial média em 24 horas das três MAPAs;
 - FCTMED-3mapas = média da frequência cardíaca média em 24 horas das três MAPAs;
 - DCNOPSM1 = percentual do descenso sistólico do sono na MAPA 1;
 - DCNOPDM1 = percentual do descenso diastólico do sono na MAPA 1;
 - DCNOPSM2 = percentual do descenso sistólico do sono na MAPA 2;
 - DCNOPDM2 = percentual do descenso diastólico do sono na MAPA 2;
 - DCNOPSM3 = percentual do descenso sistólico do sono na MAPA 3;
 - DCNOPDM3 = percentual do descenso diastólico do sono na MAPA 3;
 - DCNOPS-Méd 3 MAPAs = média percentual dos descensos sistólicos do sono nas 3 MAPAs;
 - DCNOPD-Méd 3 MAPAs = média percentual dos descensos diastólicos do sono nas 3 MAPAs;

Tabela D.2 - Estatística Descritiva da Amostra - Caso Arritmias

	N	Média	Mediana	Soma	Mínimo	Máximo	D.P.
<i>SEXO</i>	108	0,296	0,000	32,0	0,000	1,00	0,459
<i>IDADE</i>	108	51,676	51,500	5581,0	23,000	77,00	12,779
<i>IDADCAT</i>	108	2,704	3,000	292,0	1,000	4,00	1,096
<i>RACA</i>	108	0,380	0,000	41,0	0,000	1,00	0,488
<i>HISTFAMI</i>	108	0,778	1,000	84,0	0,000	1,00	0,418
<i>TABAGISM</i>	108	0,148	0,000	16,0	0,000	1,00	0,357
<i>ETILISMO</i>	108	0,139	0,000	15,0	0,000	1,00	0,347
<i>PESO</i>	108	74,060	73,000	7998,5	45,400	126,00	16,780
<i>ALTURA</i>	108	156,213	155,000	16871,0	133,000	185,00	8,390
<i>IMC</i>	108	30,416	29,078	3284,9	19,068	51,78	6,912
<i>CQ</i>	108	0,891	0,885	96,3	0,701	1,08	0,085
<i>FC</i>	108	77,631	77,300	8384,1	53,000	125,30	12,431
<i>PS</i>	108	157,648	154,450	17026,0	108,000	255,00	26,697
<i>PD</i>	108	100,954	100,100	10903,0	73,300	143,60	14,634
<i>PR</i>	108	142,222	140,000	15360,0	100,000	220,00	23,013
<i>QRS</i>	108	78,611	80,000	8490,0	40,000	150,00	16,375
<i>QT</i>	108	380,926	380,000	41140,0	260,000	520,00	41,027
<i>QTc</i>	108	427,185	424,264	46135,9	320,000	550,00	36,318
<i>RR</i>	108	796,259	800,000	85996,0	80,000	1170,00	159,364
<i>CORNELLC</i>	108	14,602	13,000	1577,0	2,000	35,00	6,834
<i>CORNELL</i>	108	0,157	0,000	17,0	0,000	1,00	0,366
<i>IMVE</i>	108	131,736	128,305	14227,5	72,000	328,60	37,543
<i>LNIMVE</i>	108	4,845	4,854	523,3	4,277	5,79	0,264
<i>HVE</i>	108	0,639	1,000	69,0	0,000	1,00	0,483
<i>AE</i>	108	34,593	35,000	3736,0	27,000	48,00	4,212
<i>FE</i>	108	66,630	67,000	7196,0	53,000	77,00	4,621
<i>ESVisosal</i>	108	271,148	6,000	29284,0	0,000	11934,00	1319,594
<i>TaquiSupra</i>	108	1,278	0,000	138,0	0,000	69,00	6,764
<i>EVisolbig</i>	108	300,000	4,000	32400,0	0,000	8794,00	1064,838

Tabela D.2 - Estatística Descritiva da Amostra – Caso Arritmias (continuação)

	N	Média	Mediana	Soma	Mínimo	Máximo	D.P.
<i>EVparsale</i>	108	5,074	0,000	548,0	0,000	202,00	27,817
<i>TaquicVentri</i>	108	0,074	0,000	8,0	0,000	6,00	0,591
<i>RRMedioDia</i>	108	728,898	722,000	78721,0	536,000	1062,00	110,822
<i>FCMediaDia</i>	108	84,103	83,102	9083,1	56,497	111,94	12,044
<i>SDNN – Dia</i>	108	78,167	71,000	8442,0	32,000	163,00	27,037
<i>SDANN – Dia</i>	108	61,398	59,000	6631,0	19,000	153,00	26,758
<i>SDNNI – Dia</i>	108	43,917	40,500	4743,0	22,000	94,00	13,863
<i>RMSSD – Dia</i>	108	28,185	23,000	3044,0	13,000	132,00	16,720
<i>PNN > 50 – Dia</i>	108	4,695	1,845	507,1	0,060	37,83	6,394
<i>RRmedioNoite</i>	108	917,435	919,000	99083,0	675,000	1195,00	122,221
<i>FCMediaNoite</i>	108	66,598	65,288	7192,6	50,209	88,89	9,183
<i>SDNN – Noite</i>	108	101,157	97,500	10925,0	39,000	245,00	36,122
<i>SDANN – Noite</i>	108	67,269	61,000	7265,0	16,000	192,00	32,608
<i>SDNNI – Noite</i>	108	63,315	60,000	6838,0	24,000	124,00	22,590
<i>RMSSD – Noite</i>	108	44,213	35,000	4775,0	15,000	141,00	25,517
<i>PNN > 50 – Noite</i>	108	15,361	10,890	1658,9	0,210	60,98	14,329
<i>RRMedio – 24h</i>	108	775,213	758,000	83723,0	631,000	1024,00	94,911
<i>FCMedia24h</i>	108	78,523	79,156	8480,5	58,594	95,09	9,353
<i>SDNN – 24h</i>	108	131,991	132,000	14255,0	53,000	206,00	31,784
<i>SDANN – 24h</i>	108	121,315	119,000	13102,0	48,000	196,00	31,854
<i>SDNNI – 24h</i>	108	50,556	48,500	5460,0	21,000	103,00	15,046
<i>RMSSD – 24h</i>	108	33,676	28,000	3637,0	15,000	118,00	17,432
<i>PNN > 50 – 24h</i>	108	7,474	4,770	807,2	0,180	34,70	7,461
<i>ECG – AR</i>	89	0,258	0,000	23,0	0,000	1,00	0,440
<i>LNIDADE</i>	108	3,911	3,942	422,4	3,135	4,34	0,270
<i>LNCQ</i>	108	-0,120	-0,122	-12,9	-0,355	0,07	0,096
<i>Numero</i>	108	2181,844	2158,500	235639,1	1763,000	2623,00	227,594
<i>RLH1</i>	108	5,759	3,109	622,0	0,400	32,72	6,429
<i>RLH2</i>	108	6,408	3,395	692,0	0,136	32,20	7,101

Tabela D.2 - Estatística Descritiva da Amostra – Caso Arritmias (continuação)

	N	Média	Mediana	Soma	Mínimo	Máximo	D.P.
<i>RLH3</i>	108	1,500	0,968	162,0	0,111	11,47	1,850
<i>RLH4</i>	108	1,422	0,932	153,6	0,087	8,34	1,480
<i>R1</i>	108	0,435	0,442	47,0	0,000	0,96	0,253
<i>F1</i>	108	101,065	49,600	10915,1	0,010	1457,54	176,318
<i>R2</i>	108	0,488	0,459	52,7	0,022	0,94	0,245
<i>F2</i>	108	112,303	52,964	12128,7	1,394	926,24	141,225
<i>R3</i>	108	0,266	0,241	28,7	0,010	0,83	0,176
<i>F3</i>	108	31,384	19,819	3389,5	0,612	315,32	42,860
<i>R4</i>	108	0,290	0,280	31,4	0,020	0,86	0,195
<i>F4</i>	108	36,986	24,351	3994,5	1,273	382,21	50,093
<i>R401</i>	108	0,500	0,500	54,0	0,000	1,00	0,502
<i>ESV01</i>	108	0,509	1,000	55,0	0,000	1,00	0,502
<i>TAKSUP01</i>	108	0,222	0,000	24,0	0,000	1,00	0,418
<i>EV01</i>	108	0,139	0,000	15,0	0,000	1,00	0,347
<i>EVPASA01</i>	108	0,194	0,000	21,0	0,000	1,00	0,398
<i>TAKVEN01</i>	108	0,028	0,000	3,0	0,000	1,00	0,165

LISTA DE PUBLICAÇÕES

- VANDERLEI FILHO, D.; VALENÇA, M. J. S.; CAMPELLO DE SOUZA, F. M. Análise Comparativa de Rede Neural Artificial Construtiva e Regressão Logística Aplicadas na Identificação de Hipertrofia Ventricular Esquerda. In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, SBAI, 7. 2005, São Luis,MA. Anais...São Luis-MA, 2005.
- _____.; CAVALCANTI, A. M.; CHAVES JUNIOR, H. C.; CAMPELLO DE SOUZA, F. M. Modelo para Análise de Arritmias Cardíacas Usando Cadeias de Markov. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, SBPO, 37. 2005, Gramado,RS. Anais... 2005.
- _____.; CHAVES JUNIOR, H. C.; VALENÇA, M. J. S.; CAMPELLO DE SOUZA, F. M. Um Estudo Comparativo de Aplicação de Rede Neural Artificial Construtiva e Regressão Múltipla na Estimção do Índice de Massa Ventricular Esquerda. In: SIMPÓSIO DE PESQUISA OPERACIONAL E LOGÍSTICA DA MARINHA, SPOLM, 8. 2005, Rio de Janeiro,RJ. Anais... 2005.
- _____, D.; VALENÇA, M. J. S.; CHAVES JUNIOR, H. C.; CAMPELLO DE SOUZA, F. M. A Constructive Neural Network for Detecting Left Ventricular Hypertrophy. In: THE FIFTH INTERNATIONAL CONFERENCE ON HYBRID INTELLIGENT SYSTEMS, HIS'05, 5. 2005, Rio de Janeiro,RJ. novembro 2005.
- _____.; VALENÇA, M. J. S.; CAMPELLO DE SOUZA, F. M. A Machine-assisted Diagnosis to Left Ventricular Hypertrophy. In: IBERIAN LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGINEERING, CILAMCE, 26. 2005, Guarapari,ES. Anais... 2005.
- CAVALCANTI, A. M.; MARMO, M. C. R. ; VANDERLEI FILHO, D.; CAMPELLO DE SOUZA, F. M. Modelo de Markov Oculito Combinado com Filogenético Aplicado em Análise de Bioseqüências. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, SBPO, 37. 2005, Gramado,RS. Anais... 2005.
- _____.; MARMO, M. C. R. ; VANDERLEI FILHO, D.; CAMPELLO DE SOUZA, F. M. O Reconhecimento de DNA Utilizando o Modelo de Markov Oculito para Análise de Seqüências. In: ENCONTRO NORTE-NORDESTE DE PESQUISA OPERACIONAL E MATEMÁTICA COMPUTACIONAL, EPOMaC, 1. 2005, Maceió,AL. Anais... 2005.
- DA SILVA, A. A.; VANDERLEI FILHO, D.; CAMPELLO DE SOUZA, F. M. Um Marcador de Arritmias Ventriculares em Indivíduos a partir das Medidas de Pressão Sistólica e Diastólica e de Frequência Cardíaca no Consultório. In: CONGRESSO DA SOCIEDADE DE HIPERTENSÃO, 12. 2003, Natal,RN. Anais... 2003.