Um Estudo de Métricas de Similaridade em Sistemas Baseados em Casos Aplicados à Área da Saúde

Márcia Regina Ferro Móss Júlio

Trabalho Final de Mestrado Profissional

BIBLIOTECA CENTRAL
DESENVOLVIMENTO
COLEÇÃO
UNICAMP

UNIDADE BC
Nº CHAMADA
T/UNICAMP
V EX
TOMBO BC/ 65561
PROC. 16-86-05
C D
PREÇO 11,00
DATA 14-9-05
Nº CPD

Bib. ID. 364450

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Júlio, Márcia Regina Ferro Móss

J945e Um estudo de métricas de similaridade em sistemas baseados em casos aplicados à Área de Saúde / Márcia Regina Ferro Móss -- Campinas, [S.P. :s.n.], 2005.

Orientador : Gilberto Shigueo Nakamiti; Heloísa Vieira da Rocha

Trabalho final (mestrado profissional) - Universidade Estadual de Campinas, Instituto de Computação.

1. Raciocínio baseado em casos. 2. Sistemas especialistas (Inteligência Artificial). 3. Inteligência Artificial (Aplicações médicas). I. Nakamiti, Gilberto Shigueo. II. Rocha, Heloísa Vieira da. III. Universidade Estadual de Campinas. Instituto de Computação. IV. Título.

Instituto de Computação

Universidade Estadual de Campinas

Um Estudo de Métricas de Similaridade em Sistemas Baseados em Casos Aplicados à Área da Saúde

Márcia Regina Ferro Móss Júlio

Fevereiro de 2005

Banca Examinadora:

- Prof. Dr. Gilberto Shigueo Nakamiti (Orientador)
  Instituto de Computação - UNICAMP
- Prof. Dr. Ricardo Luís de Freitas
  Pontifícia Católica de Campinas
- Profª. Drª. Ariadne Maria Brito Rizzoni Carvalho
  Instituto de Computação - UNICAMP
- Prof. Dr. Jacques Wainer (Suplente)
  Instituto de Computação - UNICAMP

UNICAMP 326

BIBLIOTECA CENTRAL
DESENVOLVIMENTO
COLEÇÃO
UNICAMP

Um Estudo de Métricas de Similaridade em Sistemas Baseados em Casos Aplicados à Área da Saúde

Este exemplar corresponde à redação final do Trabalho Final devidamente corrigida e defendida por Márcia Regina Ferro Móss Júlio e aprovada pela Banca Examinadora.

Campinas, 18 de fevereiro de 2005.

Prof. Dr. Gilberto Shigueo Nakamiti
(Orientador)

Profª. Drª. Heloísa Vieira Rocha
(Co-Orientadora)

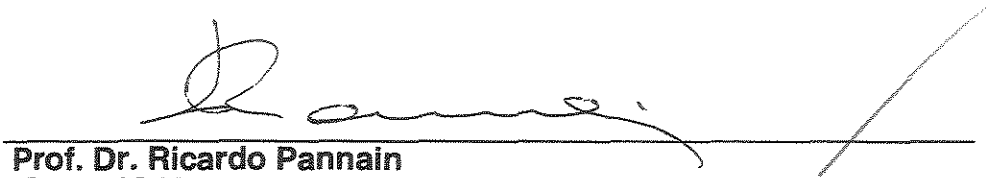
Trabalho Final apresentado ao Instituto de Computação, UNICAMP, como requisito parcial para a obtenção do título de Mestre em Computação na área de Engenharia de Software.

Prof. Dr. ALEXANDRE XAVIER FALCÃO
Coordenador do Mestrado Profissional
Instituto de Computação - UNICAMP
Matr. 27.180-2

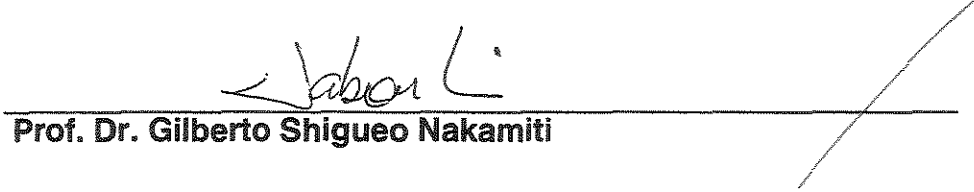
TERMO DE APROVAÇÃO

Trabalho Final Escrito defendido e aprovado em 18 de fevereiro de 2005, pela Banca Examinadora composta pelos Professores Doutores:

Prof. Dr. Ricardo Luis de Freitas
PUCCAMP



Prof. Dr. Ricardo Pannain
IC - UNICAMP



Prof. Dr. Gilberto Shigueo Nakamiti

DEDICATÓRIA

Dedico esse trabalho ao meu filho, pequeno Caio e ao meu marido Luís Antônio, que sempre me deram total apoio e carinho.

AGRADECIMENTOS

Agradeço a Deus e à minha família pelo apoio e compreensão, sem os quais não conseguiria concluir este trabalho.

Ao apoio logístico e financeiro do Programa Institucional de Capacitação Docente – PICD do Centro Universitário Anhanguera – Campus Leme – 2002/2004.

Ao Prof. Dr. Gilberto S. Nakamiti, orientador deste trabalho, pelo interesse e dedicação, desde a concepção do projeto inicial até sua conclusão. À Profa. Dra. Heloísa Vieira Rocha pela co-orientação.

Aos amigos Letícia, Paulo, Sérgio, Giovana, Igor, Christiane, Márcio, Ingleds, Gesiane, e Adriano pela contribuição dada a este trabalho.

Enfim, agradeço a todas as pessoas que, de forma direta ou indireta, contribuíram para o desenvolvimento deste trabalho.

RESUMO

No momento da escolha da solução para um problema, muitas vezes o ser humano se vale de experiências passadas, ocorridas com problemas semelhantes e que, portanto, podem prever soluções de sucesso ou não.

Sistemas Baseados em Casos (SBC) podem utilizar soluções anteriores para interpretar uma nova situação, ou criar uma solução apropriada para um novo problema.

Este trabalho apresenta um estudo de métricas de similaridade em sistemas baseados em casos, aplicados à área da saúde, mais especificamente sobre epicondilite lateral, uma tendinite do cotovelo.

O estudo sobre métricas de similaridade em sistemas baseados em casos foi realizado a partir de levantamentos bibliográficos sobre Raciocínio Baseados em Casos e sobretudo com o estudo e aprendizado obtido por meio da aplicação de RBC na área da Saúde.

A aplicação foi desenvolvida com a participação de profissionais da área da saúde que muito colaboraram na construção da aplicação, bem como com o fornecimento de casos reais para os cadastros na base de casos e aplicação de testes de validação.

PALAVRAS - CHAVE: Sistemas Baseados em Casos, Métricas de Similaridade, Aplicação na Área da Saúde.

ABSTRACT

When solving a problem, humans often use past experiences with similar situations, which can help the prediction of failure or success.

Case-Based Systems use past experiences to interpret a new situation, or to create an appropiate solution for a new problem.

For work presents a study on similarity metrics in case-based systems, and an application concerning the health area, more specifically about Lateral Epiconditis, an elbow tendinitis.

The study on similarity metrics in case-based systems was conducted from bibliographic research and more importantly, with the study and learning abtained with the health area application development.

Health area professionals took part and helped the application development, as well as provided real cases to configure and validate the system.

KEY-WORDS: Case-Based Systems, Similarity Metrics, Health Area Application.

LISTA DE FIGURAS

Figura 01	Ciclo de RBC	16
Figura 02	Mapeamento de um Caso Real	48
Figura 03	Formulário do FisiSmart representando um Caso Real	49
Figura 04	Diagrama de Contexto do FisioSmart	66
Figura 05	Diagrama de Fluxo de Dados de Cadastro	67
Figura 06	Diagrama de Fluxo de Dados de Recuperação de Casos	68
Figura 07	Modelo do Banco de Dados do FisioSmart	72
Figura 08	Tela de Entrada do FisioSmart	79
Figura 09	Tela de Cadastro de Pacientes	80
Figura 10	Tela de Cadastro do Histórico de Pacientes	81
Figura 11	Tela de Cadastro de Descritores	82
Figura 12	Tela de Cadastro de Medicamentos	83
Figura 13	Tela de Cadastro de Casos	84
Figura 14	Tela de Cadastro dos Atributos dos Casos	85
Figura 15	Tela de Similaridade dos Casos	86
Figura 16	Tela de Cadastro do Tratamento	87

Figura 17 Tela de Cadastro dos Pesos dos Atributos 88

Figura 18 Gráfico de Similaridade 89

Figura 19 Listagem dos Casos 90

Figura 20 Métrica City Block (Manhattan) 92

Figura 21 Métrica Contagem de Características 93

Figura 22 Métrica Euclidiana 94

Figura 23 Métrica Euclidiana Quadrada 95

Figura 24 Métrica Euclidiana Quadrada Ponderada 96

Figura 25 Métrica Vizinho mais Próximo – Função Inclusão de Consulta 97

Figura 26 Métrica Vizinho mais Próximo – Função Interseção 98

Figura 27 Métrica Vizinho mais Próximo – Função Linear 99

LISTA DE TABELAS

Tabela 01	Primeiros Sistemas RBC	07
Tabela 02	Exemplos de Atributo e Valor	28
Tabela 03	Cálculo de Similaridade para um Atributo	51
Tabela 04	Ilustração do Cálculo da Função Linear	53
Tabela 05	Ilustração do Cálculo da Função Interseção	54
Tabela 06	Ilustração do Cálculo da Função Inclusão de Consulta	55
Tabela 07	Ilustração do Cálculo da Similaridade Global	56
Tabela 08	Pesos dos Atributos	59
Tabela 09	Representação do Ajuste 1	62
Tabela 10	Representação do Ajuste 2	63
Tabela 11	Descrição das Tabelas de Dados utilizadas no Fisiosmart	73
Tabela 12	Dicionário de Dados da Tabela de Atributos	74
Tabela 13	Dicionário de Dados da Tabela de Casos	75
Tabela 14	Dicionário de Dados da Tabela de Descritores	75
Tabela 15	Dicionário de Dados da Tabela de Relacionamento entre Atributos e Descritores	76

Tabela 16 Dicionário de Dados da Tabela de Relacionamento entre Casos, Atributos e Descritores 76

Tabela 17 Dicionário de Dados da Tabela de Medicamentos 76

Tabela 18 Dicionário de Dados da Tabela de Pacientes 77

Tabela 19 Dicionário de Dados a Tabela de Relacionamento entre Medicamentos e Pacientes 77

Tabela 20 Dicionário de Dados da Tabela de Parâmetros 78

Tabela 21 Métrica City-Block (Manhattan) 92

Tabela 22 Métrica Contagem de Características 92

Tabela 23 Métrica Euclidiana 93

Tabela 24 Métrica Euclidiana Quadrada 94

Tabela 25 Métrica Euclidiana Quadrada Ponderada 95

Tabela 26 Métrica Vizinho mais Próximo – Função Inclusão de Consulta 95

Tabela 27 Métrica Vizinho mais Próximo – Função Interseção 96

Tabela 28 Métrica Vizinho mais Próximo – Função Linear 97

Tabela 29 Casos Retornados pela Aplicação 131

SUMÁRIO

1. INTRODUÇÃO 01

2. RACIOCÍNIO BASEADO EM CASOS 03
   2.1 Definição 04
   2.2 Histórico 06
   2.3 Casos 11
      2.3.1 Componentes de um Caso 11
      2.3.2 Organização dos Casos 14
   2.4 Ciclo de Desenvolvimento em RBC 15
   2.5 Etapas para a Construção de um Sistema Baseado em Casos 16
      2.5.1 Aquisição do Conhecimento 18
         2.5.1.1 Obtenção de Casos para a Base de Casos 19
      2.5.2 Definição dos Atributos 20
      2.5.3 Seleção dos Índices ou Problema de Indexação 20
         2.5.3.1 Processo de Indexação dos Casos 23
         2.5.3.2 Vocabulário de Indexação dos Casos 25
      2.5.4 Recuperação dos Casos 26
      2.5.5 Adaptação dos Casos (Reutilização) 28
      2.5.6 Revisão dos Casos 30
      2.5.7 Definição do Processo de Aprendizado sobre a Solução (Retenção) 31

3. MÉTRICAS DE SIMILARIDADE 34
   3.1 Classificação 35
      3.1.1 Similaridade Local e Global 35
      3.1.2 Similaridade Sintática e Semântica 36
      3.1.3 Similaridade Estrutural, Organizacional e Pragmática 36
      3.1.4 Medidas Cardinais e Medidas Ordinais 37
   3.2 Exemplos de Métricas 37

3.2.1 Distância através do método do Vizinho mais Próximo (Nearest Neighbour) ... 37
3.2.2 Distância Euclidiana ... 41
3.2.3 Distância Euclidiana Quadrada ... 41
3.2.4 Distância Euclidiana Quadrada Ponderada ... 41
3.2.5 Distância City-block (Manhattan) ... 42
3.2.6 Distância Chebychev ... 42
3.2.7 Distância para Pesos Diferentes ... 42
3.2.8 Contagem de Características (Features Count) ... 43
3.2.9 Métrica Booleana ... 43
3.3 Métricas de Ordenação e Agrupamento ... 44
3.3.1 Threshold (Limite) ... 44
3.3.2 Matching e Ranking (Comparação e Ordenação) ... 44
3.3.3 Recuperação Indutiva (Inductive Retrieval) ... 44
3.4 Formas de Estruturação ... 45
3.4.1 Flat Memory, Serial Search (Memória Plana, busca Serial) ... 45
3.4.2 Discrimination Networks (Redes Discriminatórias) ... 45
3.4.3 Shared Feature Networks (Redes de Características Compartilhadas ... 46
3.4.4 Redundant Discrimination Networks (Redes Discriminatórias Redundantes) ... 46

4. DESENVOLVIMENTO DE UMA APLICAÇÃO ... 47
4.1 Representação dos Casos ... 47
4.2 Armazenamento e Recuperação dos Casos ... 50
4.3 Métricas Implementadas ... 50
4.4 Aplicação na Área da Saúde ... 58
4.4.1 Representação dos Casos na Aplicação ... 59
4.4.2 Tratamento dos Atributos Monovalorados e Multivalorados ... 60
4.5 Modelagem do Sistema ... 64
4.5.1 Lista de Eventos ... 65
4.5.2 Diagrama de Contexto ... 65

4.5.3 Diagramas de Fluxo de Dados (DFD's) 66
4.5.4 Descrição dos Processos 69
4.5.4.1 Processo 1 - Manipular Registro de Paciente 69
4.5.4.2 Processo 2 - Manipular Medicamentos 69
4.5.4.3 Processo 3 - Manipular Histórico do Paciente 69
4.5.4.4 Processo 4 - Configurar Pesos 70
4.5.4.5 Processo 5 - Manipular Descritores 70
4.5.4.6 Processo 6 - Manipular Casos 70
4.5.4.7 Processo 7 - Armazenar Novo Caso 70
4.5.4.8 Processo 8 - Selecionar Métrica 71
4.5.4.9 Processo 9 - Recuperar Casos Similares 71
4.5.4.10 Processo 10 – Reter Caso Similar Revisado 71
4.5.5 Modelo Entidade Relacionamento (MER) 71
4.5.6 Dicionário de Dados 73
4.5.6.1 Descrição das Tabelas de Dados 73
4.5.6.2 Descrição dos Atributos das Tabelas 74
4.6 Telas 79
4.6.1 Cadastro de Pacientes 80
4.6.2 Cadastro de Descritores 81
4.6.3 Cadastro de Medicamentos 82
4.6.4 Cadastro de Casos 83
4.6.5 Cadastro de Pesos 87
4.7 Gráficos e Relatórios 88
4.8 Resultados Obtidos 91
4.8.1 Análise dos Resultados Obtidos 100

5. CONSIDERAÇÕES FINAIS 105
5.1 Contribuições 106
5.2 Trabalhos Futuros 107

BIBLIOGRAFIA 109
ANEXO 01 - Manual de Utilização do FisioSmart 1.0 117

ANEXO 02 - Questionário utilizado na Entrevista dos Pacientes com Epicondilite Lateral 132
ANEXO 03 - Casos Retornados pela Aplicação 133

Capítulo 1 Introdução

Quando se busca uma solução ou uma explicação diante de um problema, faz-se rapidamente uma recordação de situações passadas, onde se encontra a mesma ou uma situação problema similar com o problema em questão.

Raciocínio Baseado em Casos (RBC) pode ser entendido como um método de soluções de problemas usando adaptações de soluções anteriores similares a estes problemas.

Sistemas Baseados em Casos (SBC) podem adaptar velhas soluções para encontrar novas; usar velhos casos para explicar novas situações e criticar novas soluções; utilizar raciocínios anteriores para interpretar uma nova situação; ou criar uma solução apropriada para um novo problema.

A motivação de RBC é a utilização de um conhecimento específico para a solução de problemas verídicos, que são os casos passados. Propõe a solução de um problema buscando um caso ocorrido anteriormente e que seja semelhante ao atual. Uma característica do RBC é o aprendizado contínuo, uma vez que a experiência nova pode ser retida no momento da solução de um novo problema, sendo incorporada à base de conhecimento e ficando disponível para solucionar novos problemas **[Aamodt, 1994]**.

Esse trabalho se propõe a fazer um estudo de métricas de similaridade em sistemas baseados em casos, aplicados à área da saúde, mais especificamente sobre epicondilite lateral, uma tendinite do cotovelo.

A pesquisa foi realizada por meio de levantamentos bibliográficos e desenvolvimento de uma aplicação para atender casos de epicondilite lateral, utilizando RBC.

A realização deste trabalho vem sendo executada a partir de 2002 e apresenta dois aspectos: o estudo de métricas de similaridade em SBC e a aplicação com RBC na área da saúde.

Este trabalho faz a comparação do uso de métricas de similaridade para a melhor solução dos casos de epicondilite previstos na aplicação com RBC.

A aplicação será disponibilizada para o uso na rede pública e privada da área da saúde a fim de melhor indicar o tratamento para os diversos casos de epicondilite de forma segura e dinâmica.

O capítulo 2 apresenta uma revisão bibliográfica sobre RBC e mostra o ciclo de desenvolvimento para um SBC, com o propósito de subsidiar conceitos básicos para a evolução do assunto nos capítulos seguintes.

O capítulo 3 apresenta as medidas de similaridades e ilustra sua forma de cálculo em aplicações RBC.

No capítulo 4 foram descritos os processos de análise e desenvolvimento da aplicação, bem como, mostradas as métricas de similaridade utilizadas.

Finalmente, foram apresentadas as considerações finais deste trabalho.

Capítulo 2 Raciocínio Baseado Em Casos

Neste capítulo, são apresentados os principais conceitos sobre Raciocínio Baseado em Casos e o Ciclo de Desenvolvimento para aplicações que se propõe utilizar RBC. Não se pretende esgotar o assunto, mas prover os subsídios necessários para o entendimento do trabalho.

Raciocínio Baseado em Casos é uma técnica de IA que reproduz aspectos da cognição humana para resolver problemas especialistas, ou seja, problemas que são normalmente solucionados por especialistas humanos em determinada área do conhecimento. Os sistemas de RBC simulam o ato humano de relembrar um problema anterior para resolver um problema atual, em função da identificação de afinidades entre os mesmos **[Darpa, 1989, Lee, 1998, Ramalho, 1999, Sbrocco, 2003].**

Em um sistema RBC, um caso representa uma descrição completa de um problema do domínio com uma solução já aplicada. Um conjunto de casos usualmente é indexado por atributos, a fim de agilizar a busca e recuperação de casos similares. Depois da recuperação de um caso similar ao novo problema, é feita sua adaptação para que se ajuste ao problema e ofereça uma melhor solução. Assim, um caso é a abstração de uma experiência descrita por meio de atributos que foram devidamente valorados. Estes atributos descrevem o conteúdo da experiência e o seu contexto. As características de um caso são referentes aos conjuntos de pares atributo-valor **[Camargo, 1999, Darpa, 1989].**

Para recorrer às experiências passadas, o ser humano utiliza uma base de conhecimento fazendo hipóteses sobre essa base, que é, na verdade, seu conjunto de experiências acerca de determinado assunto. O conjunto de casos extraídos dessas experiências representa a base de casos. Portanto, o RBC é utilizado para a resolução de um problema quando após a identificação de um problema atual, utilizando uma base de casos, faz-se a busca da experiência mais semelhante no conjunto de casos disponível, e aplica-se o conhecimento dessa

experiência passada na resolução do problema atual **[Lagemann, 1997, Mello, 2002]**.

2.1 Definição

Dentro de Inteligência Artificial (IA), tem-se uma técnica conhecida como Raciocínio Baseado em Casos (RBC), que se propõe a resolver novos problemas, por meio da recuperação e adaptação de soluções dadas a problemas anteriores semelhantes.

Uma das motivações do RBC é a utilização de conhecimento específico para a solução de problemas verídicos, que são os casos passados. Propõe a solução de um problema buscando um caso ocorrido anteriormente e que seja semelhante ao atual. Uma característica do RBC é o aprendizado contínuo, uma vez que a experiência nova pode ser retida no momento da solução de um novo problema, sendo incorporada à base de conhecimento e ficando disponível para solucionar novos problemas **[Aamodt, 1994]**.

Quando o RBC faz a busca, ou seja, pesquisas em sua base de conhecimento, verifica se há casos semelhantes às características apresentadas pelo problema atual. Assim, pode encontrar um ou vários casos que poderão ser utilizados na solução do problema. Se for preciso, pode-se modificar partes dos casos que não se adequem ao problema atual, criando um novo caso para posterior utilização **[Lee, 1998]**.

O conhecimento em RBC é representado por meio da base de casos. A representação do conhecimento trata da descrição dos casos e de como organizá-los.

Um caso é composto normalmente por dois componentes:

- descrição do problema e

- descrição da solução.

Um terceiro componente pode ser considerado:

- o resultado da aplicação da solução ao problema **[Abel, 1996]**.

A escolha do tipo de organização da base de casos deve ser feita com muita cautela. Deve-se considerar a estrutura e a eficiência. Para tanto, alguns aspectos devem ser observados:

- o tipo de representação do sistema, que pode ser, por exemplo:
  - *redes semânticas:* são grafos direcionados ligados por nós para representar objetos e conexões que representam a relação entre objetos
  - *memória episódica:* memória que procura recuperar os casos por seus episódios, ou
  - *memória dinâmica:* que pode usar, por exemplo, uma estrutura hierárquica denominada de Pacotes de Organização de Memória (POM), que agrupam um conjunto de casos com características similares. A modelagem de memória refere-se ao tipo de modelo de memória utilizada para a representação do conhecimento;

- a forma de estruturação da memória de casos. O estilo de organização dos casos na memória é escolhido de acordo com a natureza dos dados, com sua forma de representação e com as outras etapas de desenvolvimento do sistema de RBC.

2.2 Histórico

A origem do RBC ocorreu em trabalhos sobre Memória Dinâmica de Roger Scahnk e Abelson em 1977 **[Lee, 1998].**

O princípio de Memória Dinâmica originou da tentativa de se criar modelos cognitivos na solução de problemas e no aprendizado de situações, baseado em memórias episódicas **[Robertson, 1977, [Wangenheim, 2003].**

No trabalho de Schank e Abelson **[Schank, 1977]**, os autores afirmam que o conhecimento humano é armazenado no cérebro em forma de roteiros ("scripts"). Os roteiros compõem a memória conceitual e descrevem eventos típicos, como a ida a um restaurante ou consulta médica **[Lagemann, 1997].**

Diversas aplicações que utilizam RBC já foram desenvolvidas, em várias áreas, como suporte ao consumidor, controle de qualidade, planejamento de processos, apoio à decisão, manutenção de aeronaves, roteamento de mensagens, help desks computadorizados, predições, pesquisas jurisprudenciais, entre outras **[Schank, 1982].**

A tabela 01 mostra alguns dos primeiros sistemas criados utilizando RBC **[Wangenheim, 2003].**

Tabela 01 – Primeiros Sistemas RBC

Sistema em RBC	Aplicação	Autor	Local / Ano
CYRUS	Encontros Diplomáticos	Janet Kolodner	Yale University 1983
MEDIATOR	Soluções para Conflitos	Janet Kolodner	Yale University 1989
PERSUADER	Disputas Trabalhistas	Sycara	Atlanta 1987
CHEF	Criar Receitas Culinárias	Hammond	Philadelphia 1986
JULIA	Planejar Menus de Refeições	Hinrichs	Erlabum 1988
HYPO	Assessorar Violação de Patentes	Ashley	Cambridge 1991

O sistema CYRUS **[Kolodner, 1983a]** foi o primeiro sistema considerado raciocinador baseado em casos. Sua autora foi Janet Kolodner, da Universidade de Yale, no ano de 1983. CYRUS funciona com perguntas e respostas a respeito do conhecimento obtido pelas várias reuniões de um ex-secretário de estado americano, Cyrus Vance. É um sistema baseado no modelo de memória dinâmica e na teoria de POMs. O objetivo do sistema CYRUS é encontrar uma situação diplomática ocorrida no passado semelhante a uma situação atual e propor uma solução que Cyrus Vance adotaria.

O sistema MEDIATOR **[Kolodner, 1989]**, também desenvolvido por Janet Kolodner na Universidade de Yale, no ano de 1989, propõe-se a solucionar conflitos.

O sistema PERSUADER **[Sycara, 1987]**, desenvolvido por Sycara, em Atlanta, no ano de 1987, propõe-se a solucionar disputas trabalhistas entre patrão e empregado. Também foi um dos primeiros sistemas de RBC.

O sistema CHEF **[Hammond, 1986b]**, desenvolvido por Hammond, na Philadelphia, no ano de 1986, se propõe a criar novas receitas culinárias. Trata-se de um planejador baseado em casos, que a partir de receitas existentes, cria novas receitas culinárias. Uma receita é tida como um plano que contém passos de execução para se criar um determinado prato.

O sistema JULIA **[Hinrichs, 1988]**, desenvolvido por Hinrichs, em Erlabum, no ano de 1988, propõe-se a planejar menus de refeições. Planeja refeições a partir de um conjunto de restrições, por exemplo, custo do prato, se vegetariano ou não, entre outras características. Dessa forma, usa casos para propor soluções, a partir da decomposição de problemas, se for preciso, e usa as restrições para sintetizar as respostas.

O sistema HYPO **[Rissland, 1991, Smith, 1991]**, desenvolvido por Ashley, em Cambridge, no ano de 1991, propõe-se a assessorar a violação de patentes, utilizando RBC na área da legislação de marcas e patentes. A idéia central é usar casos passados para que o sistema HYPO possa estruturar uma argumentação a fim de conduzir o réu à vitória, prevendo as contra-argumentações que a parte contrária possa fazer **[Wangenheim, 2003]**.

Muitos outros sistemas utilizando RBC têm sido desenvolvidos nos últimos anos. Tem havido o desenvolvimento de aplicações práticas.

Para a área de Central de Atendimento é possível desenvolver aplicações utilizando técnicas de Raciocínio Baseado em Casos, para atender funções usuais de controle e supervisão do fluxo de atendimento aos chamados, além de outras. É possível facilitar a resolução técnica do problema por meio da localização de casos correspondentes, por similaridade, em uma base de casos com problemas resolvidos no passado. Para o desenvolvimento deste tipo de aplicação já existem empresas especializadas, que comercializam essas soluções para empresas que utilizam central de atendimento **[Lagemann, 1997]**.

Para a área de Comércio Eletrônico, ou seja, a troca de informações, produtos ou serviços, por meio de redes de computadores, é possível desenvolver aplicações utilizando técnicas de RBC, para oferecer suporte aos clientes com melhores serviços e facilidades de seleção.

Atualmente, alguns desenvolvimentos de aplicações em RBC estão relacionados à solução de problemas relativos ao suporte aos clientes, na etapa de pós-venda. Alguns exemplos desse tipo de aplicação são os acessos a base de dados, por meio das FAQ ("Frequently Asked Questions" – Perguntas realizadas com freqüência), que proporcionam aos clientes informações após a compra de algum produto. Pode ser uma simples pergunta, rapidamente respondida, até um processo complexo com várias interações entre cliente e atendentes do "call center", mesmo que sobre a Internet **[Turner, 1989b]**.

Para a área de Recuperação Textual, é possível desenvolver aplicações utilizando técnicas de RBC. Em RBC, grande parte das informações pode ser armazenada em forma textual, por meio de relatórios, manuais, notas informativas ou comentários sobre determinado assunto. Os sistemas de recuperação textual não retornam como resposta a informação solicitada pelo usuário, mas retornam os documentos que possuem uma informação relacionada com a questão feita pelo usuário.

Alguns exemplos de aplicações que utilizam recuperação textual com RBC:

- CBR-Answers **[Lenz, 1997]:** ferramenta destinada ao manuseio de conhecimento contido em documentos textuais. Foi projetado para situações de hotline, serviço de atendimento ao consumidor por telefone.
- FAQFinder **[Lenz, 1998]:** tem por objetivo responder questões em linguagem natural, por meio da recuperação de questões a partir de arquivos de FAQ. Essa aplicação se difere de outros projetos de recuperação textual com RBC, por não focalizar um domínio específico.
- Spire **[Lenz, 1998]:** utiliza uma aproximação bastante diferente das usuais para manusear os casos textuais. Combina técnicas de RBC com técnicas de Recuperação de Informação. Resolve problemas de paráfrases, pois sugere frases alternativas feitas pelo usuário, embora os problemas de ambigüidade não são resolvidos.

- FALLQ **[Lenz, 1997]**: foi projetado para ser um sistema de informações para o suporte de uma hotline, manuseando FAQs, contudo tem sido utilizado para a manutenção de diversos tipos de informação dentro da companhia que o desenvolveu.
- SIMATIC **[Lenz, 1997]**: utiliza muitos dos princípios do CBR-Answers. Foi desenvolvido para utilização como uma hotline via Internet.

Para a área de Medicina, também é possível desenvolver aplicações utilizando técnicas de RBC. O processo decisório que um médico faz é único para um determinado domínio médico e está associado com o objeto da decisão, ou seja, o paciente. Esse tipo de aplicação pode ser usado para diagnose, classificação, planejamento, prognose ou tutoriamento.

Independente da área de aplicação, sistemas que utilizam RBC possuem o mesmo princípio, ou seja, comparam um novo caso com casos da base de casos, realizando uma pesquisa dos casos mais semelhantes por meio de métricas de similaridade que determinam o grau de similaridade entre o novo caso e a base de casos. Encontrado um ou mais casos similares, segue a fase de adaptação que pode ser feita pelo usuário ou próprio sistema a fim de resolver o problema em questão. Se for apropriado, os novos casos são acrescidos à base de casos **[Mello, 2002]**.

2.3 Casos

Segundo **[Lee, 1998]**, Um caso pode ser entendido como a abstração de uma experiência que pode ser descrita por meio de atributos valorados. Estes atributos procuram descrever o conteúdo e o contexto da experiência.

Um caso também pode ser entendido como uma lista de objetivos, estados e suas explicações, e suas características podem levar a um resultado particular

de uma determinada situação. De forma mais abrangente, um caso pode ser conectado a um conjunto de sub-casos que, quando integrados, formam a estrutura do espaço de resolução de problemas **[Abel, 1996].**

Um caso ainda pode ser entendido como a forma de conhecimento contextualizado, que representa uma experiência a fim de ensinar uma lição útil. Uma lição útil, pode ser entendida como aquela que possui o potencial de ajudar o raciocinador a alcançar uma meta ou um conjunto delas, e podem advertir sobre a possibilidade de falha ou mesmo apontar para um futuro problema **[Abel, 1997]** .

2.3.1 Componentes de um Caso

Um caso é composto normalmente por dois componentes:

- descrição do problema e
- descrição da solução.

Um terceiro componente pode ser considerado:

- o resultado da aplicação da solução ao problema **[Abel, 1996].**

Descrição do problema – realizada por meio da atribuição de características que descrevem o problema de entrada. Essas características podem assumir a forma de nomes, números, funções ou textos, que representem características, objetivos, metas, restrições e condições para a distinção de um caso do outro.

A descrição do problema observa o estado da "situação" de quando o caso ocorreu, e faz a representação do problema. Contempla:

- Metas a serem alcançadas na solução do problema.
- Restrições contidas nestas metas.
- Características e relações entre as partes da situação.

No momento de decisão da escolha da descrição do problema, algumas diretrizes podem ser consideradas:

- Incluir todas as características que possam contribuir para atingir a meta (resolver o problema);
- Incluir as características que são normalmente utilizadas para descrever casos típicos;
- Incluir, além das anteriores, outras classes de características que descrevam o problema, como características de adaptação.

As características de adaptação têm influência sobre as características da solução. Entretanto, são características que podem fazer parte da descrição do problema, pois, ao invés de determinar a similaridade entre os casos, determinam os ajustes que deverão ser realizados.

Solução do problema - é a solução do problema identificado na descrição do mesmo, composto por:

- A solução propriamente dita.
- Passos necessários para a resolução do problema.
- Justificativas do por que das decisões tomadas na resolução do problema.
- Soluções possíveis que não foram escolhidas (soluções alternativas) ou soluções que foram excluídas por não serem aceitáveis.
- Expectativas dos resultados.

A solução do caso é a solução que o RBC sugere como saída e inclui as tarefas que o sistema deve executar. Por exemplo: um diagnóstico ou uma interpretação assume a forma de solução do problema em função dos objetivos. A solução deve ser descrita através de uma metodologia que aponte qual a melhor solução.

Depois de várias e sucessivas adaptações de soluções, é recomendado que seja registrada a avaliação de cada adaptação, as impressões quanto às novas implementações dessa solução e sua comparação com as expectativas inicialmente propostas. Esses dados podem ajudar a enriquecer e aprimorar a qualidade dos casos e, conseqüentemente, da base de conhecimento.

Resultado da aplicação - composto por:

- O próprio resultado.
- Se o resultado preencheu ou violou as expectativas.
- O resultado foi sucesso ou falha.
- Explicação da violação da expectativa ou da falha.
- Estratégia para reparo que poderia evitar um problema.
- Apontar para uma provável próxima solução.

É necessário identificar mecanismos de avaliação dos resultados da implementação da solução adotada na base de casos. Estes mecanismos podem ser qualitativos ou quantitativos.

Uma experiência vivenciada é um caso. Após isso, as outras execuções dessa mesma tarefa (é intuitivo que se use aquilo que foi aprendido na primeira vez) pode também tornar se uma rotina e deixar de ser um caso **[Abel, 1997].**

Uma definição de caso pode ser mais complexa do que a delimitação do tipo de problema a ser utilizado em RBC. É possível tentar buscar o mesmo tipo de solução por diferentes formulações do problema. Por exemplo, se o problema fosse tentar buscar solução para acidentes de trânsito, o ponto de partida seria a identificação das causas dos acidentes. A modelagem poderia ser orientada pelos agentes envolvidos no acidente, como carros e pessoas, e pelo local, se foi numa esquina ou em uma estrada **[Lee, 1998].**

Dessa forma, é possível observar que a representação dos casos é uma tarefa complexa e muito importante para a determinação do sucesso ou do fracasso do sistema de RBC. Talvez a maior dificuldade esteja no momento de

decidir o que será armazenado e encontrar a estrutura ideal para descrever o conteúdo do caso **[Abel, 1996]**.

Portanto, a descrição dos casos, que é tratada pela representação do conhecimento, assume formas diferenciadas dependendo do domínio da aplicação do sistema. Os componentes serão compostos de acordo com a necessidade. Para tanto, deve-se considerar a viabilidade de representação do conteúdo, do contexto das experiências na modelagem dos casos e da escolha do tipo de base de casos [Kolodner, 1983b].

2.3.2 Organização dos Casos

Para se organizar os casos na memória, é importante escolher os casos conforme a natureza dos dados, forma de representação dos casos e etapas de desenvolvimento do sistema de RBC.

A armazenagem dos casos, na base de casos, pode ser feita em uma estrutura hierárquica, estrutura plana de dados, ou em estruturas intermediárias.

Algumas formas de organização dos casos podem ser utilizadas, como: memória plana, memória hierárquica, banco de dados relacional, redes discriminatórias e redes de características compartilhadas **[Lee, 1998]**.

2.4 Ciclo de Desenvolvimento em RBC

O ciclo de RBC proposto por Aamondt e Plaza **[Aamodt, 1994]** é um modelo bastante aceito para o processo de RBC. É também chamado de 4R, pois engloba um ciclo de raciocínio contínuo composto por 4 tarefas:

- Recuperar – o(s) caso(s) mais similar(es) da base de casos;

- Reutilizar (adaptar) – esse(s) caso(s) para resolver o problema;
- Revisar - a solução proposta, e
- Reter (aprender) – a experiência representando o caso atual (ou partes desta experiência) para reutilização futura.

A figura 01 mostra o ciclo de RBC.

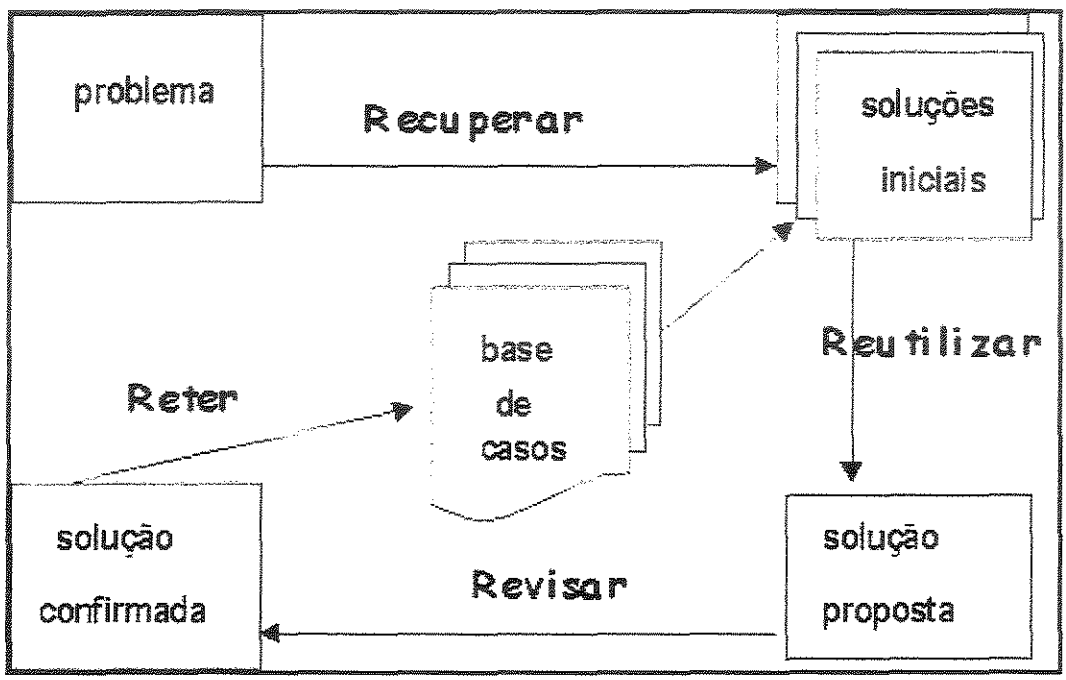


Figura 01 – Ciclo de RBC **[Aamodt, 1994]**

A representação e a indexação dos casos não são explicitadas no ciclo, porque são executadas na criação da base de casos, ainda na construção do sistema.

2.5 Etapas para a Construção de um Sistema Baseado em Casos

O sucesso de Sistemas Baseados em Casos tem sido observado principalmente quando utilizados em domínios bem específicos.

Os sistemas especialistas (SE) baseiam-se em módulos de conhecimento sobre um domínio específico, sendo direcionados para solucionar problemas que ocorram nesse domínio. Um SE possui um motor de inferência (que representa a forma de manipular o conhecimento já representado na base, a fim de resolver o problema) capaz de tomar algumas decisões, fazendo um "raciocínio" sobre as informações de entrada do sistema e extraindo informações previamente armazenadas **[Turner, 1989a]**.

A utilização de SE que aplicam RBC apresentam um processo diferente dos sistemas especialistas tradicionais. As diferenças podem ser notadas em diversas etapas do processo. Assim, a aquisição de conhecimento, por exemplo, não exige a reunião de todas as informações necessárias ao domínio do problema. O "conhecimento" é determinado pelos casos armazenados em sua base de casos, ou base de conhecimento. Os casos são armazenados conforme a sua importância, para isso se determinam pesos. Pode-se observar que sistemas RBC atendem a resolução de problemas, onde o conhecimento é incompleto ou com poucas evidências, portanto RBC processa experiências e não faz inferências baseadas no conhecimento **[Kolodner, 1983b]**.

Basicamente a construção de RBC deve seguir as seguintes etapas:

Aquisição do Conhecimento – é a seleção da base de informações do domínio da aplicação. Pode ocorrer por meio de documentos, banco de dados, fichas, ou similares. Se não houver informações disponíveis, é preciso consultar um especialista da área. A confiabilidade dos sistemas que utilizam RBC é diretamente proporcional a uma correta e completa base de casos.

Definição dos Atributos – para serem utilizados na solução do problema.

Seleção dos Índices ou Problema de Indexação - para o acesso dos casos armazenados na biblioteca de casos. Deve-se procurar uma estrutura apropriada para a descrição das informações contidas nos casos.

Definição dos Métodos de Recuperação – para a verificação da similaridade entre o novo caso e o conjunto de casos na base. As medidas de similaridade são escolhidas por meio de entrevistas com o especialista de acordo com a aplicação.

Adaptação de Casos – a partir de uma solução associada a um caso na base de casos poderá haver uma adaptação para solucionar o novo problema. A adaptação de casos envolve operações de inclusão, exclusão e substituição de componentes de uma solução recuperada.

Definição do Processo de Aprendizado sobre a Solução – Ocorrendo sucesso entre a solução apresentada e adaptada para o novo problema, o sistema deve ser capaz de avaliar se deve armazenar ou descartar essa solução. Se ocorrerem falhas na solução, o sistema deverá explicar e aprender com essas falhas, para que não ocorram novamente **[Althoff, 1995, Leake, 2002]**.

2.5.1 Aquisição do Conhecimento

A aquisição do conhecimento consiste na forma de acúmulo, transferência e transformação de uma perícia em solução para um dado problema em um programa computadorizado. Ocorre por meio de fontes de conhecimento como um especialista humano ou documentos (textos, documentos multimídia, bases de dados, relatórios de pesquisa, informações disponíveis na web e outras). O cenário para a aquisição do conhecimento é composto pela armazenagem organizada das informações para o motor de inferência e por parte dos computadores pela necessidade explícita e em detalhes de informações, utilizada pelos humanos. O ser humano nem sempre se recorda de todos os passos

BIBLIOTECA CENTRAL
DESENVOLVIMENTO
COLEÇÃO

necessários para a resolução de um problema, na hora de transferir ou processar o conhecimento **[Wangenheim, 2003].**

A Aquisição de Conhecimento (AC) pode ser desenvolvida em 6 etapas:

Identificação e Estudo do Problema - Delimitação do problema a ser resolvido e o conhecimento necessário para isso. Normalmente, o especialista humano define qual o problema deve ser resolvido pelo sistema.

Entrevistas - realizadas entrevistas com um especialista humano a fim de detalhar a forma de utilização e o conhecimento aplicado à solução de um problema.

Análise do Conhecimento – a partir das informações obtidas, é necessário analisá-las e elucidar sua consistência, fazendo a verificação se são suficientes e completas para solucionar o problema.

Proposta de Representação do Conhecimento - faz-se a estruturação do conhecimento em uma das formas computacionais.

Implementação e Teste do Sistema - a implementação do sistema é feita por meio de uma linguagem de computação tradicional (Delphi, por exemplo) ou utilizando-se uma ferramenta específica para o desenvolvimento de sistemas RBC. Em seguida o sistema é testado em termos funcionais.

Validação junto ao Especialista e Usuários - finalmente nessa etapa do desenvolvimento, o sistema é validado pelo especialista, em termos do conhecimento e também quanto à facilidade de uso e adequação à solução esperada pelos usuários **[Abel, 1997].**

2.5.1.1 Obtenção de Casos para a Base de Casos

A obtenção dos casos para a base de casos pode correr de diversas formas:

Casos não disponíveis de uma forma externa – quando não há disponíveis registros descritivos dos casos do domínio em questão. Somente o especialista humano se recorda dos casos. Nesta situação, para se obter os casos, são necessários os mesmos processos e limitações da aquisição do conhecimento para a construção de modelos.

Casos semidisponíveis em uma fonte externa - quando há registro dos casos, porém de forma incompleta, não sendo possível seu uso imediato. Nesta situação os casos podem ser utilizados como ponto inicial para construção do banco de conhecimento, mas precisam ser complementados com informações obtidas por meio da engenharia do conhecimento (uma subdisciplina da inteligência artificial que trata de desenvolver esse tipo de sistema para mais de uma aplicação).

Casos disponíveis e contém erros - quando os casos estão registrados de forma completa em uma fonte externa, porém contêm informações conflitantes ou erros. Nesta situação há necessidade de se fazer interações com o especialista humano, para se garantir a integridade das informações contidas na base de casos.

Casos disponíveis e corretos - esta situação ocorre em domínios do conhecimento onde a solução pode ser descrita por meio de poucos casos. Mesmo assim, a engenharia do conhecimento deve ser utilizada para definir a melhor forma de representação dos casos, quais características serão utilizadas para indexação e como será realizada a adaptação das soluções recuperadas **[Abel, 1996]**.

2.5.2 Definição dos Atributos

Atributos são os dados encapsulados em um objeto gerenciado. Cada atributo corresponde a uma das características do recurso que o objeto representa. O atributo é formado por um nome, um tipo e um ou mais valores que refletem o estado corrente do recurso gerenciado. Os valores dos atributos podem ser lidos ou modificados **[Gresse, 1999].**

2.5.3 Seleção dos Índices ou Problema de Indexação

Em sistemas de RBC, os casos são representados por uma coleção de casos. Para que se obtenham casos relevantes e semelhantes a fim de propor soluções a um novo caso, por meio da base de casos, torna-se necessária a criação de índices de acessos a essa base, os quais devem ser escolhidos de acordo com suas características.

Uma definição para indexação de casos pode ser um instrumento destinado à orientação da similaridade. A indexação permite organizar os casos baseados em uma ou mais características. Faz-se necessária a determinação do que em um determinado caso é relevante para comparação com outro caso, a fim de recuperar o caso que obtiver a similaridade necessária para viabilizar sua solução **[Lee, 1998].**

A indexação poderá determinar o que se comparar entre os casos, para determinar sua similaridade. Dessa forma, RBC utiliza os índices para facilitar e agilizar a recuperação, bem como, para atribuir com eficiência a similaridade entre os casos.

Por meio dessa definição de indexação observa-se que as informações de um caso podem ser de dois tipos:

- Informações indexadas - para utilização na recuperação.

- Informações não indexadas - para fornecimento de dados não necessários para recuperação **[Lagemann, 1997].**

É possível observar, que os índices de um caso nada mais são que combinações de descritores, capazes de distinguir entre os casos existentes **[Abel, 1996].**

Para se indexar um caso, algumas questões referenciais podem ocorrer:

- Como é que as pessoas se lembram das coisas corretas na hora certa?
- Os computadores podem reproduzir essa situação?
- Será que quando solucionado o problema de indexação, os problemas de busca também estão solucionados?

As questões acima inerentes a indexação podem ser respondidas com a explicação sobre o problema de indexação.

O problema de indexação trata da certeza do acesso a um caso no momento certo. Dessa forma, primeiramente, se faz a atribuição de rótulos ao caso quando é inserido na base de casos, a fim de assegurar sua recuperação quando oportuno. Os rótulos normalmente definem os problemas e as circunstâncias onde uma tarefa foi aprendida para ser ensinada posteriormente.

Para sucesso do processo de indexação dos casos, é necessária uma boa definição de índices, que pode ser obtida por meio dos levantamentos efetuados com o especialista da área. Bons índices proporcionarão boas respostas na busca. Num segundo momento, trata-se da organização dos casos para que a busca seja feita de forma eficiente e aperfeiçoada **[Kolodner, 1983b].**

Algumas características para um bom trabalho de indexação devem contemplar:

- Antecipar o vocabulário a ser usado na recuperação.

- Utilizar os conceitos usados normalmente para descrever os casos, por meio das características mais abstratas.
- Antecipar as circunstâncias de realização da recuperação.
- Ser suficientemente concreto para facilitar a reorganização em futuras situações.
- Realizar a indexação proporcionando uma recuperação eficiente de casos, a fim de atender as expectativas do usuário **[Darpa, 1989]**.

Alguns fatores como o tamanho da base de casos, coleção dos casos, é relevante quanto ao problema de indexação. Se a base for pequena, a busca não é uma questão muito importante, contudo se a base se tornar grande, é necessário considerar o tempo para se diferenciar entre os casos da base de conhecimento, o caso ou os casos relevantes para propor uma solução ao caso em análise. Daí a importância dos rótulos, para diferenciar ou discriminar diferentes situações e utilidades. Nesse caso o que se chama rótulo pode receber também a denominação de índice **[Dillon, 1982, Fagan, 1987, Maher, 1995a]**.

2.5.3.1 Processo de Indexação dos Casos

Para se realizar a indexação dos casos, alguns aspectos devem ser considerados:

- É necessário que a indexação contenha conceitos utilizados normalmente para descrever os itens que estão sendo indexados; mesmo que sejam aspectos superficiais ou abstratos;
- É preciso, antecipar o vocabulário que o usuário da aplicação possa utilizar. Dessa forma, a indexação deve

prover uma descrição correta para que o usuário possa encontrar o caso ou casos similares na base de casos;

- Também é preciso antecipar as circunstâncias em que o usuário da aplicação deseja recuperar alguma coisa. Dessa forma, antecipar por exemplo, que quem pensa em dor do epicôndilo lateral pode pressupor tratamento com infra-vermelho ou seções de fisioterapia.

Após várias experiências adquiridas na construção de sistemas de RBC, a comunidade científica elaborou um roteiro para a escolha de índices para casos particulares **[Kolodner, 1983b]**:

- Índices devem ser preditivos. Considerando que um caso é a descrição de um problema, ou seja, a tentativa de uma solução e o resultado da solução posta em prática, pode-se considerar também que os aspectos de um caso são responsáveis pelo seu resultado, então pode-se afirmar que eles predizem o resultado. As características preditivas são combinações de descritores de um caso que foram responsáveis pela solução do caso, observando quais dessas combinações influenciaram no seu resultado. Por exemplo, um caso de tratamento de epicondilite lateral que teve um fracasso porque a paciente era gestante e foi prescrito tratamento com ondas curtas. Gestantes não podem receber tratamento com ondas curtas. A combinação dos descritores "mulheres gestantes" e "tratamento com ondas curtas" formam uma característica preditiva (prevêem um fracasso).
- As predições que podem ser feitas devem ser úteis. Devem descrever os propósitos de utilização do caso;
- Índices devem ser abstratos o suficiente para permitirem

> que um caso seja útil em uma variedade de situações futuras. Dessa forma, os descritores de um índice devem ser mais abstratos do que os descritores particulares de um caso. Contudo, os índices devem ser concretos o suficiente para que não ofereçam dificuldades no reconhecimento de situações futuras. Se o índice for muito abstrato há o risco de ocorrer muitas inferências sobre seus descritores. Ser concreto o suficiente significa que devem ser reconhecíveis com pouca inferência. Portanto, quanto mais inferência for usada, menos concreto é o índice.

Ainda para a obtenção de melhor desempenho de um sistema de RBC é recomendada a escolha de índices manualmente, embora haja métodos automáticos para isso **[Kolodner, 1983b].**

Algumas técnicas podem ser escolhidas para a indexação:

Análise Matemática – análise numérica de todos os elementos do domínio e suas dimensões, a fim de identificar quais características influenciam ou determinam as conclusões.

Índices baseados nas diferenças entre os casos – indexação feita especificamente com as características que diferencia casos similares.

Técnicas baseadas em explicação – indexação feita com as características que influenciam na solução dos casos. Para tanto, os casos são analisados individualmente e são determinadas as características do problema utilizadas para construir a solução **[Kolodner, 1983b].**

2.5.3.2 Vocabulário de Indexação dos Casos

É importante a escolha do vocabulário a ser utilizado na indexação. Existem vocabulários específicos para determinados domínios, como também existem vocabulários que transpassam esses domínios. Para que se possa contemplar os movimentos do raciocínio humano, a forma como os humanos organizam o conteúdo de sua memória para algum tipo de analogia, é preciso incluir o mesmo conteúdo nos índices dos casos. Conteúdos de memória são, portanto, itens do vocabulário que descrevem relações abstratas por meio dos domínios.

Um vocabulário de indexação é um subconjunto do vocabulário utilizado para a representação simbólica completa dos casos. As características são utilizadas como rótulos (índices), que são atribuídos aos casos como descritores. Esses índices descrevem o vocabulário de indexação.

Quando um vocabulário é escolhido para a representação de casos deve conter duas partes:

- O conjunto das dimensões cobertas pelo vocabulário e
- O conjunto de símbolos representando os valores que incluem a descrição de cada dimensão.

2.5.4 Recuperação dos Casos

A recuperação é uma etapa no ciclo RBC que tem por objetivo auxiliar o raciocínio produzido do próprio ciclo. Sua realização é feita por meio da utilização das características do novo caso, que poderão ser relevantes na solução de um problema **[Abel, 1996].**

A recuperação também pode ser entendida como a etapa em que é disparada uma função para recuperar os casos similares. Para tanto, pode-se utilizar uma métrica ou métodos de classificação, contudo devem ser

estabelecidos limites para a recuperação dos casos, ou um limite de casos que possam ser recuperados **[Lee, 1998].**

A recuperação de um caso inicia na descrição de um problema e termina quando o melhor caso é encontrado. Para isso algumas etapas devem ser executadas:

- *Identificação das características:* fornece ao sistema as características do caso em questão;
- *Casamento inicial:* faz a recuperação de um conjunto de casos candidatos possíveis;
- *Busca:* etapa mais elaborada de seleção do melhor candidato entre os casos recuperados na etapa do casamento inicial;
- *Seleção:* essa etapa gera conseqüências e expectativas de cada caso recuperado e faz a tentativa da avaliação das conseqüências e justificação das expectativas. Nessa etapa ainda, há uma tentativa de ordenação dos casos, conforme a métrica ou critério de classificação usado, a fim de se realizar a escolha do caso, da base de casos, mais similar possível ao novo problema **[Maher, 1995b].**

A tarefa recuperação é um processo de retorno de um ou mais casos, extraídos da base de casos, que são resultado da comparação entre o caso novo e a base. Essa comparação se dá por meio da avaliação de similaridade, ou seja, a seleção de um caso ou uma combinação de vários casos que propõe uma solução para o caso novo.

Alguns fatores são considerados importantes na etapa de recuperação:

- Eficiência: velocidade de recuperação dos casos pelo sistema.
- Precisão: grau de casos recuperados que podem ser utilizados para alcançar o objetivo proposto.

- Flexibilidade: grau de recuperação de casos para raciocínios inesperados **[Abel, 1996].**

Para um melhor entendimento sobre a recuperação de casos, algumas definições sobre RBC devem ser consideradas:

- Descritor: atributo utilizado na descrição de um caso. Podem descrever aspectos do problema ou da situação, solução ou resultado. Ex.: Cor dos Olhos
- Valor: refere-se à parte atributo do descritor. Ex.: Verde (1). O campo Valor pode ser mapeado para outro tipo. Ex.: Inteiro.
- Característica: usando tanto como descritor (a maior parte das vezes) como dimensão. Ex.: Cor dos Olhos – Verde ou 1. A tabela 02 pode ilustrar os exemplos de descritor e Atributo e Valor **[Darpa, 1989].**

Tabela 02 – Exemplos de Atributo e Valor

Cor dos Olhos (Atributo)	**Valor**
Verde	1
Azul	2
Castanho	3
Preto	4

2.5.5 Adaptação de Casos (Reutilização)

A tarefa de reutilização de casos é caracterizada pela adaptação da solução armazenada de um caso recuperado às necessidades de um novo caso proposto **[Hammond, 1989a, Kambhampati, 1989].**

A reutilização de uma solução de casos recuperados, diante de um novo caso, possui dois aspectos:

- Diferenças entre o caso recuperado e o novo caso e
- Partes do caso recuperado que podem ser transferidas para o novo caso **[Aamodt, 1994].**

Diante da raridade de um problema ser exatamente igual a outro, soluções dadas a um problema passado normalmente devem ser adaptadas para solucionar novos problemas. No processo de adaptação pode haver a substituição de um atributo da solução por outro, ou mesmo uma complexa e total modificação da solução **[Abel, 1996].**

A adaptação pode ser entendida como tendo a função de alterar um caso, quando necessário, a fim de solucionar um problema inicial, avaliando-se as diferenças entre os casos recuperados e o problema inicial **[Lee, 1998].**

O processo de adaptação pode ser realizado de diversas formas:

- *Inclusão* de um novo comportamento à solução recuperada.
- *Eliminação* de um comportamento da solução recuperada.
- *Substituição* de parte de um comportamento **[Lagemann, 1997].**

Alguma formas de adaptação:

- *Adaptação Estrutural:* a adaptação de regras é realizada diretamente na solução armazenada no caso.

- *Adaptação Derivacional:* as regras geradas para a solução original são executadas novamente para gerar uma solução nova. Ao recuperar-se um caso, há uma verificação por parte do sistema para observar se as diferenças entre o novo caso e o caso passado afetam fatores básicos para a solução armazenada no caso. A solução armazenada é adaptada pela re-execução das partes do processo da solução original e não a alterando completamente. A principal vantagem desse tipo de adaptação é a adaptação da solução de um caso em problemas dentro de seu domínio, e poder transferir estes a outros domínios **[Abel, 1996].**

- *Adaptação nula:* não faz nada, somente aplica a solução de um caso recuperado para uma nova situação.

- *Solução parametrizada:* quando há a recuperação de um caso para sua aplicação em uma situação determinada. As descrições dos problemas de ambos os casos são comparadas por parâmetros específicos, sendo as diferenças utilizadas para modificar as soluções dos parâmetros na solução apropriada.

- *Abstração e Especialização:* técnica estrutural usada para realizar adaptações simples de forma complexa para gerar novas soluções.

- *Reinstalação:* a reinstalação não opera na solução original, mas nos métodos utilizados na geração da solução **[Abel, 1996].**

A tarefa adaptação ou de ajuste de situação se faz necessária quando há dificuldades em apresentar o problema de entrada nos moldes dos casos da base. Ocorrendo esse problema, são necessários ajustes a fim de se comparar o caso de entrada com os casos da base. A complexidade dessa tarefa é variável de acordo com o tipo de situação tratada, podendo ser uma simples substituição de um componente do caso recuperado ou uma complexa modificação na estrutura geral da solução **[Lee, 1998]**.

2.5.6 Revisão dos Casos

Quando uma solução para um caso é gerada de forma incorreta, na etapa de reutilização (adaptação), surge a oportunidade do aprendizado com a falha ocorrida. Conseqüentemente ocorre a revisão do caso, que pode ser composta por:

- Avaliar a solução do caso gerada pela reutilização e, se for positiva, aprender (retenção do caso);
- Reparar a solução do caso, utilizando reconhecimento específico de domínio **[Aamodt, 1994]**.

2.5.7 Definição do Processo de Aprendizado sobre a Solução (Retenção)

A retenção pode ser entendida como a absorção da parte útil do problema resolvido. Após a avaliação e possíveis reparos, realiza-se o aprendizado do sucesso ou insucesso da solução do problema analisado **[Hammond, 1989a, Kambhampati, 1989]**.

A aprendizagem ocorre principalmente com o acúmulo de novas experiências na memória de um sistema RBC e da correta indexação dos problemas. A eficiência de um sistema RBC ocorre justamente quando consegue aprender por meio de experiências passadas e correta indexação dos problemas **[Lagemann, 1997]**.

Para que ocorra a aprendizagem, dois importantes fatores podem considerados em um sistema RBC:

- Ter habilidade na identificação do que deve conhecer e
- Processar logo que esteja disponível, a informação adquirida na fase anterior **[Abel, 1996]**.

A aprendizagem, ainda pode ser entendida como a incorporação do que é útil do problema resolvido. Faz-se a avaliação e identifica-se os possíveis reparos necessários e daí ocorre o aprendizado com a observação do sucesso ou falhas da solução proposta **[Maher, 1995b]**.

É possível observar que um sistema de RBC consegue aprender por suas experiências, o que pode ser muito útil em experiências onde a aprendizagem é ponto crucial para a solução. O aproveitamento de soluções passadas para resolução de novos problemas, antecipando e evitando erros cometidos melhora muito a qualidade da solução. O aprendizado ocorre como subproduto do raciocínio, armazenando novas experiências na memória de casos e disponibilizando-as para casos futuros, ou na informação e análise das falhas, armazenando-as para evitar novas falhas no futuro. Esses casos devem ser indexados de tal forma que facilitem o aproveitamento deste conhecimento **[Lagemann, 1997]**.

A aprendizagem na base de casos ocorre com a adição de uma solução adaptada ao problema de entrada, quando sua qualidade for considerada satisfatória **[Lee, 1998]**.

O ciclo de RBC pesquisa na base de casos tarefas resolvidas anteriormente com descrição similar à do problema atual. Essa pesquisa é realizada em uma ordem de preferência onde se obtém casos mais similares e menos similares.

Dessa forma, a solução de um caso suficientemente similar ao atual pode ser reutilizada. A métrica de similaridade determina a medida de similaridade, isto é, define sua forma de cálculo. Assim, aplica-se repetidamente, par a par (situação atual comparada à situação da base de casos), o cálculo (métrica) de similaridade. Após o cálculo, os casos na base de casos são ordenados de acordo com o valor de similaridade e os mais similares são propostos como solução potencial. Contudo, essa solução potencial pode ainda não satisfazer completamente os requisitos da situação presente. Nesses casos é necessário fazer uma reutilização com adaptações, podendo ser feitas desde simples cópias de soluções até adaptações complexas de acordo com regras que refletem um modelo do domínio de aplicação do sistema RBC em questão.

Finalmente, é feita a validação da solução adaptada, por meio da confirmação de uma solução plausível para um problema. Sua verificação ocorre normalmente durante a aplicação prática ou por meio de uma simulação. Em seguida, a nova solução poderá ser incorporada à base de casos, tornando assim os sistemas RBC continuamente atualizados. Sob esta ótica, pode-se dizer que informações sobre o processo de recuperação e adaptação do novo caso podem ser capturadas e oferecer informações sobre o desempenho do sistema RBC, podendo ser continuamente melhorado. Assim, genericamente, pode-se dizer que a recuperação, reuso e revisão podem ficar sob a responsabilidade do usuário e a retenção de novos casos para o engenheiro do conhecimento responsável pela manutenção do sistema **[Turner, 1989b]**.

A seguir serão apresentadas, as métricas de similaridade que determinam o grau de semelhança entre os casos.

Capítulo 3 Métricas de Similaridade

Similaridade pode ser vista como um conceito intuitivo para descrever a percepção de um observador humano da semelhança existente entre dois objetos **[Turner, 1989b]**.

A formalização mais aceita do conceito de similaridade está na definição de uma medida numérica de distância ou similaridade. Pode-se entender também como medida de similaridade a formalização de uma determinada filosofia de julgamento de semelhança por meio de um modelo matemático concreto **[Gresse, 1999]**.

Métrica pode ser entendida como um atributo mensurável de uma determinada entidade. Se o atributo mensurável pode ser lido diretamente, então tem-se uma métrica primitiva, ou seja, que pode ser obtida da própria entidade. Dessa forma, define-se como métrica derivada aquela obtida por meio de cálculos executados com as métricas primitivas, sendo que os atributos mensuráveis, nesse caso, não podem ser obtidos diretamente **[Martins, 2000]**.

Uma métrica de similaridade sintetiza a similaridade por meio de cada atributo através de uma medida da importância que é usada para modelar a relevância de cada atributo na avaliação da similaridade. Conseqüentemente, qualquer cálculo matemático orientado para avaliações sintáticas é válido, tais como a média ponderada e as integrais difusas. Contudo, ainda é possível apresentar a avaliação de similaridade como um problema de reconhecimento de padrões dentro do espaço dos casos. Desta forma, algoritmos que realizam estes tipos de avaliação também podem ser usados, como os algoritmos de vizinho mais próximo, que serão apresentados nas próximas sub-seções **[Lee, 1998]**.

Dessa forma, entende-se que as métricas levam a uma medida de similaridade. Consideram-se aqui métricas como métodos que levam a uma medida quantitativa ou qualitativa referente a similaridade entre casos, ou seja, o seu grau de similaridade.

A escolha da métrica de similaridade é uma das questões essenciais em sistemas de RBC, já que a base fundamental do paradigma de RBC é a solução de um problema atual a partir da reutilização de uma solução dada a um problema anterior similar e as métricas são fundamentais para a recuperação do problema anterior similar **[Lee, 1998].**

3.1 Classificação

As métricas de similaridade podem ser classificadas de diferentes formas, de acordo com diferentes autores. Dessa forma, algumas classificações são possíveis:

3.1.1 Similaridade Local e Global

Segundo [Gresse, 1999] a similaridade entre dois objetos pode ocorrer de forma intuitiva e então é classificada como *similaridade global*. Nesse caso determina-se, por meio da utilidade de um caso em relação a um novo problema, considerando-se todos os seus atributos.

Quando se deseja uma avaliação mais compreensiva da similaridade entre um caso e um novo problema, pode-se usar a *similaridade local*, ou seja, a similaridade entre atributos específicos. Dessa forma, é permitida a utilização das similaridades entre atributos isolados para o cálculo da similaridade, possibilitando uma maior especificidade **[Gresse, 1999].**

Assim similaridade global pode ser entendida como a junção de todos os atributos do caso da base de casos com o novo caso em questão, e a similaridade local pode ser entendida como a comparação de atributo a atributo do caso da

base com o caso novo. Similaridade Local é calculada atributo a atributo e a Global é a junção dos resultados obtidos na Local.

3.1.2 Similaridade Sintática e Semântica

A *Similaridade Semântica* pode ser vista como uma forma de avaliação onde os atributos são comparados em termos de semelhança sintática. Alguns exemplos de categorias são: sinônimos, ordinais, análise de perfil, clusterização e quantificadores.

A *Similaridade Sintática* pode ser vista como uma forma de avaliação que não considera fatores contextuais, se referindo aos atributos que são sintaticamente idênticos em duas situações **[Aamodt, 1994]**.

3.1.3 Similaridade Estrutural, Organizacional e Pragmática

A *Similaridade Estrutural* exige uma consistência estrutural entre os casos e um dos nós conceitos. É mais complexa que a Similaridade Semântica. Os casos devem estar ligados por meio de estruturas isomórficas. O isomorfismo depende da consistência da ligação de proposições, predicados ou argumentos. Por exemplo, dois casos são similares se as suas estruturas de relações e argumentos forem semelhantes.

A *Similaridade Organizacional* tem relação com a similaridade imposta aos casos armazenados em lugares próximos na base de casos.

A *Similaridade Pragmática* ocorre quando duas partes ocupam papéis similares em suas respectivas situações **[Hammond, 1989a, Kambhampati, 1989]**.

3.1.4 Medidas Cardinais e Medidas Ordinais

As medidas Cardinais levam a resultados numéricos que expressam, por exemplo, uma medida de distância e as medidas Ordinais expressam uma ordem, um "ranking".

Algumas das características dos casos são responsáveis pela aproximação cardinal. Por exemplo a similaridade de meses pode ser medida por sua proximidade, isto é, a similaridade de novembro e dezembro é 1, enquanto que a de agosto e novembro é 3.

3.2 Exemplos de Métricas

Esta seção apresenta uma série de métricas de similaridade, em particular relativas a medidas de distância, as mais utilizadas no contexto dos sistemas baseados em casos.

3.2.1 Distância através do método do Vizinho mais Próximo (*Nearest Neighbor*)

Este método avalia as distâncias entre o novo caso e os casos armazenados na base, com o objetivo de encontrar o mais similar.

É uma técnica simples, que pode ser usada para buscar casos similares na base de casos, a partir de um novo caso, não requer muitos cálculos para sua compreensão. É necessário identificar as características essenciais para a resolução do problema. Essas características devem ser representadas por um sistema de coordenadas, de forma que a medida de distância entre o novo

problema e os casos da base possa ser medida. Pode também haver pesos diferenciados para cada atributo, dando mais importância a alguns e menos a outros, que precisam ser determinados pelo especialista **[Piva, 2004a, Piva, 2004b]**.

Esse método é classificado como uma similaridade global. A partir dele, há a necessidade de utilizar uma função local para complementá-la, dessa forma foram utilizadas as funções locais (linear, interseção, inclusão) para os atributos multivalorados e a função booleana para os atributos monovalorados **[Hammond, 1989a]**.

Para uma recuperação de casos eficaz, os aspectos de definição e identificação dos índices constituem fator fundamental. A partir desses aspectos, a técnica de busca deverá indicar em qual região do espaço de busca o problema em questão está inserido, localizando então os casos mais parecidos por meio de comparação e valorização **[Lagemann, 1997]**.

Para cada atributo, deve ser determinada a similaridade entre o caso novo e o existente. Essa medida de similaridade deve ser multiplicada por um fator peso e então calcula-se a somatória de todos os valores. Dessa forma, fica estabelecida a medida de similaridade entre os casos da base de casos e o novo caso **[Abel, 1996]**.

A fórmula da similaridade do Vizinho mais Próximo pode ser representada por:

$$Similaridade\ (N,C) = \left[\sum_{i=1}^{n} f(N_i, C_i) * w_i \right] \Big/ n$$

onde:

N = Novo caso

C = Casos existentes na memória de casos

n = número de atributos

i = atributo individual, variando de 1 a n

f = função de similaridade para o atributo i nos casos N e C

w = peso do atributo i

Para se obter um *ranking* (ordenação) dos casos na base de casos, o cálculo do Vizinho mais Próximo deverá ser repetido para toda a base de casos. Normalmente, as similaridades variam em um intervalo entre 0 e 1 (0 para casos totalmente diferentes e 1 para casos idênticos). A dificuldade consiste em determinar os pesos relativos das características **[Abel, 1996].**

A similaridade local depende altamente do domínio de aplicação. Essa medida deve ser cuidadosamente definida e revisada a fim de refletir o conceito de similaridade entre os objetos do domínio da aplicação.

Para se obter similaridade local a partir da métrica do vizinho mais próximo, pode-se utilizar alguns dos tipos básicos de representação de casos (número, símbolo binário, conjunto, intervalo, string e outros). O método do Vizinho mais Próximo corresponde a soma de todos os atributos calculados por uma função local, como:

- Tipo de Representação de Caso: Numérico – Função: Linear

  Os atributos podem ser descritos por números, por ex. R$278,00 no atributo Preço. A medida de distância local genérica entre dois números determina o ponto de partida para a similaridade entre dois valores quaisquer e pode ser expressa pelo módulo da diferença entre valores $|v_i - v_j|$.

  A função linear é bastante utilizada para o tipo de representação numérica. Nessa função a similaridade cresce com o decréscimo da distância entre os dois valores, ponderada pelo tamanho do intervalo assumido pelo domínio dos valores do atributo, assim:

$$f(v_i, v_j) = \begin{cases} 1 & \text{,sendo } v^i = v^j \\ 1 - [\,(v_i - v_j) / (ub - lb)\,] & \text{, caso contrário} \end{cases}$$

onde ***ub*** é o limite superior da faixa de valores do atributo e ***lb*** o limite inferior.

- Tipo de Representação de Caso: Conjunto –

  Pode haver a necessidade de se representar as entidades de informação como conjunto de valores. Dessa forma, diferentes estratégias podem ser usadas como as funções Interseção e a Inclusão de Consulta.

Função Interseção: Calcula-se a proporção da interseção de dois conjuntos para a sua união. Como na fórmula abaixo, onde v_i representa o valor da consulta e v_j o valor do caso.

$$f(v_i, v_j) = \begin{cases} (v_i \cap v_j) / (v_i \cup v_j) \text{ , sendo } v_i \neq \{\} \text{ e } v_j \neq \{\} \\ 0 \text{ , caso contrário} \end{cases}$$

Função Inclusão de Consulta: A similaridade é 1.0, quando a consulta está incluída no caso. Se houver entradas na consulta inexistentes no caso, diminuem o valor. A fórmula abaixo mostra a medida de similaridade.

$$f(v_i, v_j) = \begin{cases} (v_i\,' / v_i) \text{ , sendo } v_i\,' \subset v_j \text{ e } v_i \neq \{\} \text{ e } v_j \neq \{\} \\ 0 \text{ , caso contrário} \end{cases}$$

3.2.2 Distância Euclidiana

Essa forma de cálculo da distância é muito utilizada, e trata da distância geométrica em um espaço multidimensional.

$$Distância\ (x,y) = \left\{ \sum_{i} (x_i - y_i)^2 \right\}^{1/2}$$

A distância euclidiana é a real distância entre dois pontos em um espaço qualquer, onde a função é a raiz quadrada do somatório da diferença dos atributos ao quadrado.

3.2.3 Distância Euclidiana Quadrada

A distância euclidiana quadrada é a real distância entre dois pontos em um espaço qualquer, onde a função é o somatório da diferença dos atributos ao quadrado.

É o quadrado da distância Euclidiana padrão.

$$Distância\ (x,y) = \sum_{i} (x_i - y_i)^2$$

3.2.4 Distância Euclidiana Quadrada Ponderada

Com a introdução de pesos w_i, é possível estender a métrica distância Euclidiana para distância Euclidiana Ponderada.

A distância euclidiana quadrada ponderada é a real distância entre dois pontos em um espaço qualquer, onde a função é o somatório da diferença dos atributos ao quadrado, multiplicado pelo peso do atributo.

$$Distância\ (x,y) = \sum_{i} w_i\ (x_i - y_i)^2$$

3.2.5 Distância *City-Block* (*Manhattan*)

A distância city block é média entre as dimensões, onde a função é o somatório da diferença dos atributos, dividido pelo número de atributos do caso. Essa medida, se utilizada em uma série de casos, gera resultados parecidos com a distância Euclidiana, minimizando o efeito de grandes diferenças.

$$Distância\ (x,y) = \left[\sum_{i} |x_i - y_i| \right] \Big/ n$$

3.2.6 Distância *Chebychev*

Recomendada para casos onde é necessária a definição do grau máximo de diferença entre dois objetos, em qualquer dimensão.

$$Distância\ (x,y) = Máximo\ |\ x_i - y_i\ |$$

3.2.7 Distância para Pesos Diferentes

Quando os pesos são diferentes para as características.
onde:

N = Novo caso

C = caso existente

$$Distância\ (x,y) = (x_C * w_x) + (y_n * w_y)$$

A técnica do Vizinho mais Próximo, além de ser a mais utilizada, talvez seja a técnica mais indicada para resolver problemas que tenham uma base de casos pequena e com poucos atributos indexados, devido ao volume de cálculos necessários para a determinação de cada atributo indexado em cada um dos casos **[Hammond, 1989a, Kambhampati, 1989].**

3.2.8 Contagem de Características (Features Count)

Neste método, o cálculo da similaridade é feito por meio da contagem dos atributos do caso armazenado na base que sejam iguais aos atributos do novo caso. A cada atributo i do caso armazenado na base que seja igual ao atributo i do novo caso, é somado 1 ao número de atributos coincidentes entre os dois casos. Se os atributos forem iguais recebe similaridade 1, caso contrário 0.

O grau de similaridade entre os dois é dado pelo percentual do número de atributos iguais entre os dois casos, levando-se em conta o número de atributos considerados na comparação destes.

Este método, também possibilita a indicação de um limiar de similaridade, como será apresentado na seção 3.3.1. Este limiar é expresso por meio de um percentual mínimo de similaridade **[Aamodt, 1994]**.

3.2.9 Métrica Booleana

Em algumas aplicações, a similaridade de certos atributos booleanos pode ser mapeada para os valores 0 e 1. Outra possibilidade é utilizar a similaridade de forma eliminatória, isto é, casos que não satisfazem certa condição (ou não possuem determinados atributos) não são considerados.

3.3 Métricas de Ordenação e Agrupamento

Classificam, ordenam ou agrupam os casos, baseados em algum critério. Podem servir para a seleção de casos ou diminuição do espaço de busca.

3.3.1 *Threshold* (Limite)

Threshold pode ser usado como uma medida de limite da quantidade de casos que serão recuperados ou um valor mínimo para consideração dos casos. Esse limite pode ser estabelecido pelo número absoluto de casos que se pretende recuperar ou pelo intervalo de medida de similaridade que se entenda necessária e suficiente. Por exemplo, desconsiderar os casos que possuem similaridade abaixo de 50% e selecionar os três casos mais similares **[Mantaras, 1995].**

3.3.2 *Matching* e *Ranking (Comparação e Ordenação)*

Para uma procura direta em locais na memória na busca por casos potencialmente úteis, os algoritmos de recuperação utilizam-se de heurísticas de *Matching* e *Ranking* para a escolha na base de casos.

Matching é um processo de comparação estrita entre dois casos para a determinação de seu grau de similaridade. Possui a possibilidade de produção de uma contagem (*score*), que represente o grau de similaridade ou determine se um caso é suficientemente similar, como ocorre com a contagem de características.

Ranking é um processo de ordenação de casos parcialmente emparelhados (*partially-matching*) de acordo com a utilidade, ou seja, a identificação de qual caso é melhor que os outros. Esse processo pode utilizar, por exemplo, a saída do procedimento de *match*, a fim fazer a ordenação dos casos parcialmente emparelhados **[Kolodner, 1983a, Mantaras, 1995, Rich, 1988].**

3.3.3 Recuperação Indutiva (*Inductive Retrieval*)

A técnica de Recuperação Indutiva pode ser utilizada em sistemas que necessitem de classificação de padrões entre os casos, separando-os por conjuntos (*clusters*). Um conjunto possui casos similares e um requisito de indução, que é a própria definição da característica de um novo caso. Os casos de um mesmo conjunto possuem:

- descrição similar;
- referem-se a problemas similares, ou
- buscam soluções similares **[Abel, 1996].**

3.4 Formas de Estruturação

Agrupam e armazenam os casos, baseado em alguma técnica.

3.4.1 *Flat Memory, Serial Search* (Memória Plana, Busca Serial)

A técnica da memória plana com busca seqüencial armazena os casos seqüencialmente, em uma lista simples, vetor ou arquivo. A recuperação é realizada por uma função de *match* para cada caso da base de casos, de forma seqüencial.

Trata-se de uma técnica mais apropriada para bases de casos pequenas, na medida em que a base cresce a técnica se torna ineficiente e se confunde na busca dos melhores casos similares **[Kolodner, 1983a].**

3.4.2 *Discrimination Networks* (Redes Discriminatórias)

O método de redes de características compartilhadas agrupa os casos similares. Portanto, agrupar é principal e discriminar é secundário. Contudo, na

técnica de Redes Discriminatórias, ocorre o contrário: discriminar é principal e agrupar é secundário.

A busca é realizada por meio dos nós. Cada nó interno tem uma questão abaixo dele. Cada nó filho possui uma resposta diferente para a questão proposta pelo pai. A questão é respondida e o nó de sua resposta é selecionado, e assim sucessivamente, até o encontro de um único item ou conjunto de itens **[Kolodner, 1983a].**

3.4.3 *Shared Feature Networks* (Redes de Características Compartilhadas)

Essa técnica trata de bases de casos grandes. Para tanto, organiza os casos hierarquicamente e considera apenas pequenos conjuntos para o processo de recuperação.

Possíveis formas de agrupamento:

- agrupar por conjuntos que são partes de muitos itens;
- agrupar por características individuais, que poderão dividir o conjunto em grupos de igual tamanho; ou
- agrupar por características individuais, com pequenos grupos de itens diferentes de outros **[Hammond, 1989a].**

3.4.4 Redundant Discrimination Networks (Redes Discriminatórias Redundantes)

Essa técnica procura tratar o problema das informações perdidas. Para isso, organiza os itens por meio de diferentes redes discriminatórias, cada uma com uma ordem diferente das questões **[Hammond, 1989a, Kambhampati, 1989].**

Capítulo 4 Desenvolvimento de uma Aplicação

Para se desenvolver um sistema inteligente, existe a necessidade do conhecimento especialista, e normalmente esse especialista é um profissional da área em questão. Desta forma, por meio da utilização da análise do conhecimento junto àqueles que o detêm, faz-se a extração do conhecimento.

O processo de extração do conhecimento foi realizado por meio da identificação dos itens a serem avaliados em cada questão (atributos do caso) e como deveriam ser armazenados na base de casos, suas correlações com os possíveis diagnósticos e similaridades entre os atributos e casos diferentes **[Júlio, 2004]**.

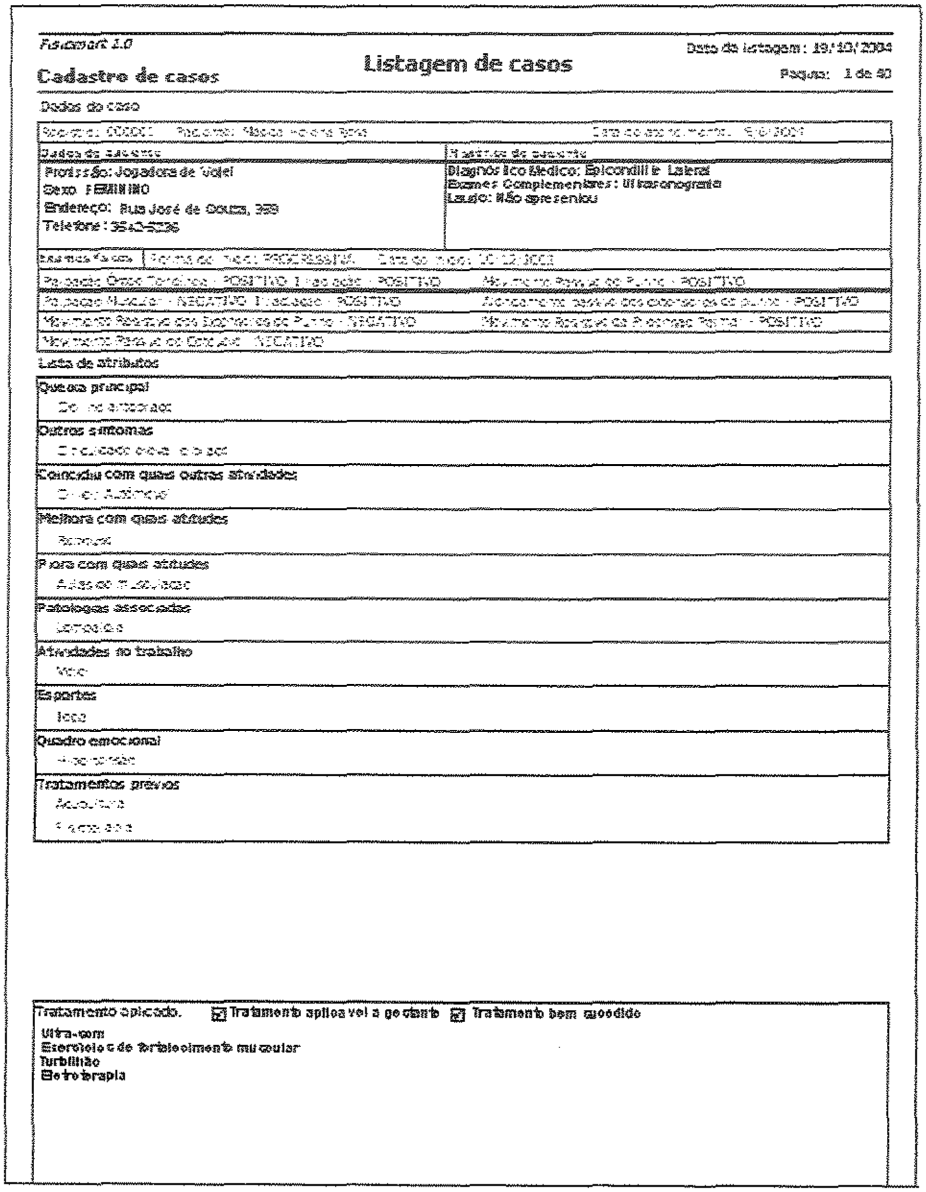
O FisioSmart é uma aplicação na área da saúde que trata de casos de epicondilite lateral e que utiliza conceitos de RBC. Este sistema se propõe a realizar o cadastro de casos de epicondilite lateral e a partir das métricas de similaridade utilizadas fazer a proposta de tratamentos para um novo caso depois de analisados os casos da base de casos. Pode ser utilizado em redes de saúde pública, bem como, em clínicas de fisioterapia. Para se compor esse trabalho foi necessária a participação de uma especialista em Fisioterapia que também exerce atualmente a função de professora e coordenadora de curso de uma faculdade da rede privada. Essa área de atuação foi escolhida por se conseguir delimitar o escopo da aplicação das métricas de distância e para se propor uma ferramenta de trabalho gratuita para a rede pública de saúde à população mais carente.

4.1 Representação dos Casos

A representação dos casos pode ser feita por meio de um mapeamento de um caso real. As figuras 02 e 03 mostram um exemplo de mapeamento de um caso real.

Registro: 000001
Nome: Magda Helena Rosa
Data de Nascimento:
Sexo: Feminino
Cor: Branca
Estado Civil:
Profissão: Jogadora de Vôlei
Endereço: Rua José de Souza, 359 – Araras – SP
Data da 1ª avaliação: 09/.06/2004
Exames complementares (vêm junto com o encaminhamento e diagnóstico médico)
História clínica
- Queixa principal: Dor no Antebraço
- Outros sintomas: Dificuldade em elevar o braço
- Data de início dos sintomas: 10/12/2003
- Forma de início dos sintomas: Progressiva
- Se há melhora dos sintomas com alguma posição ou atitude: Repouso
- Se há piora dos sintomas com alguma posição ou atitude:Aulas de Musculação
- Antecedentes familiares:
- Antecedentes pessoais (patologias associadas):
- Tratamentos prévios e resultados: acupuntura e fisioterapia
- Descrição das atividades no trabalho: Vôlei
- Descrição de atividades no lar (relevantes):
- Prática de esportes ou hobby: Yoga
- Quadro emocional: hipertensão
- Afastamento do trabalho ou função por incapacidade:
- Uso de medicamentos:
Exame físico (resultado + se apresentar dor no epicôndilo lateral com irradiação ou não)
- Palpação (ósteo-tendínea/muscular): Positivo/ Negativo
- Dor ao movimento passivo: Positivo
- Dor ao movimento resistido: Negativo
- Dor ao alongamento dos extensores do punho: Positivo
- Dor à preensão palmar: Positivo
Tratamentos:
- eletroterapia (laser, TENS, ultra-som, sonoforese)
- cinesioterapia (alongamentos, exercícios resistidos)
- crioterapia (criomassagem)
- massoterapia
- reeducação postural
- ergonomia
- splints e braces

Figura 02 - Mapeamento de um Caso Real


FisioSmart 1.0

Data da listagem: 19/10/2004

Cadastro de casos

Listagem de casos

Página: 1 de 40

Dados do caso

Dados da paciente

Profissão: Jogadora de Volei
Sexo: FEMININO
Endereço: Rua José de Souza, 359
Telefone: 3542-5236

Diagnóstico Médico: Epicondilite Lateral
Exames Complementares: Ultrassonografia
Laudo: Não apresentou

Lista de atributos

Queixa principal

Outros sintomas

Coincidiu com quais outras atividades

Melhora com quais atitudes

Piora com quais atitudes

Patologias associadas

Atividades no trabalho

Esportes

Quadro emocional

Tratamentos previos

Tratamento aplicado. ☑ Tratamento aplicavel a paciente ☑ Tratamento bem sucedido

Ultra-som
Exercicios de fortalecimento muscular
Turbilhão
Eletroterapia

Figura 03 – Formulário do FisioSmart representando um Caso Real

4.2 Armazenamento e Recuperação dos Casos

Dando seqüência ao ciclo do RBC, existe a necessidade de identificação da melhor técnica para a representação dos casos, ou seja, a modelagem da memória de casos.

Neste trabalho, a estrutura utilizada para o armazenamento dos casos foi a estrutura de memória plana. A estrutura linear é indicada para métodos de recuperação puramente associativos. Os casos são armazenados seqüencialmente em uma lista-vetor. A recuperação de casos é feita de forma seqüencial, pela aplicação da medida de similaridade, onde há um casamento entre um caso novo com um caso da base de casos, um por vez **[Hinrichs, 1992]**.

Essa forma de recuperação dos casos é vantajosa pela simplicidade do algoritmo, onde a avaliação é feita com todos os casos. Contudo, pode apresentar um baixo desempenho conforme cresce o número de casos na base de casos. Para amenizar esse problema, pode-se utilizar algoritmos que otimizam a recuperação por meio de filtros que são aplicados sobre a base de casos para a escolha de grupos específicos, de acordo com o conhecimento do especialista da área. Após o resgate de um grupo de casos em que se tem interesse é que se aplica o casamento (medida de similaridade) em cada um deles **[Kambhampati, 1989]**.

Neste trabalho, como o foco não é a recuperação de casos, foi implementada somente a técnica de memória plana, visto que não foi utilizado um número muito grande de casos, no total a base chegou a 40 casos, já com a inserção dos novos casos.

4.3 Métricas Implementadas

Existem diversas funções de cálculo de distância, como descrito

anteriormente no capítulo 3.

Um exemplo simples de cálculo de similaridade pode ser entendido conforme a tabela 03.

Tabela 03 – Cálculo de Similaridade para um Atributo

Novo Caso	Valor	Caso da Base	Valor	Distância
Atributo_1	0,4	Atributo_1	0,7	|0,4 – 0,7| = 0,3
Atributo_2	0,1	Atributo_2	0,9	|0,1 – 0,9| = 0,8

Dessa forma, os valores do Atributo_1 possuem uma diferença menor que a diferença de valores Atributo_2 do novo caso. Quanto menor o valor, menor a distância. Portanto, o Atributo_1 do novo caso é mais similar que o Atributo_2 do novo caso.

O valor da similaridade pode ser obtido subtraindo 1 da diferença para obtenção da similaridade de cada atributo.

Esse cálculo é exemplificado a seguir:

Atributo_1 = 1 - |0,4 – 0,7| = 0,7

Atributo_2 = 1 - |0,1 – 0,9| = 0,2

Nesse caso, o Atributo_1 do novo caso é mais similar que o Atributo_2.

Na aplicação das métricas de similaridade são propostas algumas técnicas de ajuste, cujos detalhes serão descritos em 4.4.2.

As métricas implementadas no FisioSmart foram:

- City-Block (Manhattan)
- Contagem de Características;
- Euclidiana;
- Euclidiana Quadrada;
- Euclidiana Quadrada Ponderada;
- Vizinho mais Próximo – Função Inclusão de Consulta;

- Vizinho mais Próximo – Função Interseção;
- Vizinho mais Próximo – Função Linear.

Para a implementação das métricas da distância city block, euclidiana, euclidiana quadrada e euclidiana quadrada ponderada foi necessário aplicar alguns ajustes, pois o cálculo dessa distância não é indicado para atributos multivalorados.

Para a implementação da métrica de distância através do método do vizinho mais próximo em suas funções no FisiSmart foram utilizadas:

- **Função Linear** – utilização da distância dos descritores para subtrair o valor do atributo do caso novo com o atributo do caso da base, após utilizar o ajuste 1, descrito na seção 4.4.2, obtendo-se assim, o cálculo da função linear. Essa função utiliza limites superior e inferior da faixa de valores do atributo. Para o FisioSmart os limites dos valores dos descritores foram 0 e 1 para os limites inferior e superior respectivamente.
  Função Linear:
  $1 - [(v_i - v_j) / (ub - lb)]$ onde, (ub – lb) => ub = 1 (limite superior) e lb = 0 (limite inferior), portanto
  (ub – lb) => (1 – 0) = 1
  v_i => atributo do caso Novo e
  v_j => atributo do caso da Base
  A tabela 04 mostra um exemplo de cálculo da função linear.

Tabela 04 – Ilustração do Cálculo da Função Linear

Linha	Caso Novo	Caso da Base
	Atributo - Queixa Principal	**Atributo – Queixa Principal**
01	Dor no Cotovelo (0,7)	Dor no Cotovelo (0,7)
02	Formigamento (0,8)	Formigamento (0,8)
03		Dor no Punho (0,6)

$1 - [(v_i - v_j) / (ub - lb)]$

1-(L1N – L1B) → 1- |0,7 – 0,7| /1 = 1
1-(L1N – L2B) → 1- |0,7 – 0,8| /1 = 0,9
1-(L1N – L3B) → 1- |0,7 – 0,6| /1 = 0,9
1-(L2N – L1B) → 1- |0,8 – 0,7| /1 = 0,9
1-(L2N – L2B) → 1- |0,8 – 0,8| /1 = 1
1-(L2N – L3B) → 1- |0,8 – 0,6| /1 = 0,8

Média → (1+0,9+0,9+0,9+1+0,8) /6 = 0,91
Máxima → 1

Onde, na tabela 04:
L1N → Linha 01 do Novo Caso → v_i
L2N → Linha 02 do Novo Caso → v_i
L1B → Linha 01 do Caso da Base → v_j
L2B → Linha 02 do Caso da Base → v_j
L3B → Linha 03 do Caso da Base → v_j

- **Função Interseção** – foi feita a interseção entre os conjuntos (novo caso com o caso da base) dividindo-se

pela união dos dois conjuntos. Não houve necessidade da utilização dos ajustes 1 e 2, pois essa função já é definida para conjuntos (atributos multivalorados). Também não houve a necessidade da utilização dos valores dos descritores, por se tratar de conjuntos, onde foram utilizados o número de elementos desses conjuntos. Função Interseção:

$(v_i \cap v_j) / (v_i \cup v_j)$, sendo $v_i \neq \{\}$ e $v_j \neq \{\}$

A tabela 05 mostra um exemplo de cálculo da função interseção.

Tabela 05 – Ilustração do Cálculo da Função Interseção

Caso Novo	Caso da Base
Atributo - Queixa Principal (v_i)	**Atributo – Queixa Principal (v_j)**
Dor no Cotovelo	Dor no Cotovelo
Formigamento	Formigamento
	Dor no Punho

$(v_i \cap v_j) / (v_i \cup v_j)$ onde:

$(v_i \cap v_j)$ = {Dor no Cotovelo, Formigamento}

$(v_i \cup v_j)$ = {Dor no Cotovelo, Formigamento, Dor no Punho}

$(v_i \cap v_j) / (v_i \cup v_j) = 2/3 = 0,66$ (divide-se o n°. de elementos da interseção pela união)

- **Função Inclusão de Consulta** - quando o caso da base contiver o caso novo a similaridade será 1, contudo ocorrendo diferença entre os conjuntos (alguns dos elementos do caso da base não contiver os elementos do caso novo) a similaridade irá diminuir.

Não houve necessidade da utilização dos ajustes 1 e 2, pois essa função já é definida para conjuntos (atributos multivalorados). Também não houve a necessidade da utilização dos valores dos descritores, por se tratar de conjuntos, onde foram utilizados a quantidade de elementos desses conjuntos. Função Inclusão de Consulta:

(v_i' / v_i) , sendo v_i' v_j e $v_i \neq \{\}$ e $v_j \neq \{\}$

A tabela 06 mostra um exemplo de cálculo da função inclusão de consulta

Tabela 06 – Ilustração do Cálculo da Função Inclusão de Consulta

Caso Novo	**Caso da Base**
Atributo - Queixa Principal (v_i)	**Atributo – Queixa Principal (v_j)**
Dor no Cotovelo	Dor no Cotovelo (v_i')
Formigamento	Formigamento (v_i')
	Dor no Punho

(v_i' / v_i) onde:

v_i' = {Dor no Cotovelo, Formigamento}

v_i = {Dor no Cotovelo, Formigamento}

$(v_i' / v_i) = 2/2 = 1$ (divide-se o nº. de elementos de v_i' por v_i)

A utilização da função booleana atribui valor 1 para os atributos monovalorados quando o novo caso for igual ao caso da base e 0 para quando for diferente.

A tabela 07 ilustra o cálculo da similaridade global.

Tabela 07 – Ilustração do Cálculo da Similaridade Global

Atributos	Peso do Atributo	Caso Novo (6)	Caso da Base (17)	Valor da Similaridade Local
Queixa Principal	10	Alteração de Sensibilidade do membro superior (0,4) Dor no Membro superior (0,5)	Dor no Cotovelo (0,7)	0,50
Outros Sintomas	1	Diminuição de força no movimento extensor	Diminuição de força no movimento extensor	1,00
Forma de Início	2	Repentina	Progressiva	0,00
Coincidiu com quais outras atividades	4	Prática de Tênis (1)	Prática de Esportes (0,5)	0,25
Melhora com quais atitudes	1	Imobilizar o membro superior (0,6)	Repouso durante o trabalho (1)	0,30
Piora com quais atitudes	1	Exercícios Físicos (0,7)	Carregar Peso (0,9)	0,40
Antecedentes Familiares	5	Não	Não	1,00
Antecedentes Pessoais	1	Não	Não	1,00
Patologias Associadas	6	Não	Não	1,00
Quais tratamentos prévios	7	Não	Fisioterapia (1)	0,00
Atividades no Trabalho	5	Professor de Tênis (0,8) Professor de Natação (0,1)	Movimento Repetitivo (1)	0,30
Atividades no Lar	1	Não	Não	1,00
Esportes	1	Natação (0,3) Tênis (1)	Natação (0,3)	0,43
Quadro Emocional	1	Não	Não	1,00
Palpação Ósteo Tendínea	10	Positivo	Positivo com irradiação	0,50
Movimento Resistivo dos Extensores do Punho	5	Positivo	Positivo	1,00

Movimento Passivo do Cotovelo	10	Negativo	Positivo	0,00
Movimento Passivo do Punho	9	Negativo	Positivo	0,00
Alongamento Passivo dos Extensores do Punho	6	Positivo	Positivo	1,00
Palpação Muscular	8	Positivo	Positivo	1,00
Movimento Resistivo da Preensão Palmar	10	Positivo	Positivo	1,00
Valor da similaridade global: (10*0.50 + 1*1.00 + 2*0.00 + 4*0.25 + 1*0.30 + 1*0.40 + 5*1.00 + 1*1.00 + 6*1.00 + 7*0.00 + 5*0.45 + 1*1.00 + 1*0.43 + 1*1.00 + 10*0.50 + 5*1.00 + 10*0.00 + 9*0.00 + 6*1.00 + 8*1.00 + 10*1.00) / (10 + 1 + 2 + 4 + 1 + 1 + 5 + 1 + 6 + 7 + 5 + 1 + 1 + 1 + 10 + 5 + 10 + 9 + 6 + 8 + 10) = 58,38/104 = 0,56				

A função chebychev não foi implementada, pois o ajuste 1 já obtém o valor máximo de similaridade de cada atributo.

Não foi possível implementar distância para pesos diferentes na aplicação Fisiosmart, pois cada atributo do sistema tem um peso único. Essa distância prevê valores diferentes para os atributos diante de uma outra condição que lhe possa ser atribuído. Essa definição de peso único para os atributos, no caso do FisioSmart foi dada pelo especialista da área.

A implementação da contagem de características foi efetuada sem a necessidade de ajustes.

Assim, os atributos multivalorados foram implementados pelas métricas (city-block, contagem de características, euclidiana, euclidiana quadrada, euclidiana quadrada ponderada, vizinho mais próximo funções – linear, inclusão de consulta e interseção) e a métrica booleana foi aplicada para todos os atributos monovalorados utilizados no FisioSmart.

Na seção seguinte poderá ser observada uma aplicação na área da saúde com a utilização de RBC.

4.4 Aplicação na Área de Saúde

O FisioSmart é um Sistema para Avaliação de Diagnósticos de Casos de Epicondilite Lateral, a tendinite do cotovelo. Essa aplicação utiliza RBC e serve de base para o teste de algumas métricas de similaridade, em particular as baseadas em distância, além de apresentar um tratamento para atributos multivalorados.

Para o desenvolvimento dessa aplicação não foi utilizada nenhuma *Shell* de RBC. Todos os programas necessários foram desenvolvidos e implementados na ferramenta de programação Delphi, utilizando o banco de dados FireBird. O anexo 01 apresenta um manual de utilização dessa aplicação.

Quando um fisioterapeuta recebe um paciente já com o diagnóstico de Epicondilite Lateral, ele inicia uma entrevista onde o paciente responde a uma série de perguntas que poderão auxiliar o fisioterapeuta na escolha da melhor conduta e procedimentos em fisioterapia para realizar o tratamento do paciente. No anexo 02 pode-se encontrar o questionário utilizado na entrevista.

Assim como faz o fisioterapeuta, ao receber um novo paciente (e, portanto, um novo caso) busca em sua memória as experiências de tratamentos prescritos anteriormente, utilizados em situações semelhantes e que obtiveram sucesso no tratamento. Para resolver o problema atual, o sistema FisioSmart por meio dos métodos de comparação, também busca na memória de casos, utilizando RBC, casos anteriores .

Dessa forma, este trabalho foi realizado em conjunto com profissionais da área de fisioterapia. O verdadeiro nome dos pacientes foram omitidos por motivos éticos e foram utilizados nomes fictícios.

4.4.1 Representação dos Casos na Aplicação

As figuras 02 e 03 mostraram o exemplo de mapeamento de um caso real para a aplicação FisioSmart.

A tabela 08 mostra os pesos dos atributos, propostos pelo especialista.

Tabela 08 – Pesos dos Atributos

Atributos	Peso do Atributo
Queixa Principal	10
Outros Sintomas	6
Forma de Início	10
Coincidiu com quais outras atividades	10
Melhora com quais atitudes	5
Piora com quais atitudes	5
Antecedentes Familiares	6
Antecedentes Pessoais	8
Patologias Associadas	8
Quais tratamentos prévios	7
Atividades no Trabalho	8
Atividades no Lar	8
Esportes	10
Quadro Emocional	10
Palpação Ósteo Tendínea	10
Movimento Resistivo dos Extensores do Punho	10
Movimento Passivo do Cotovelo	10
Movimento Passivo do Punho	9
Alongamento Passivo dos Extensores do Punho	6
Palpação Muscular	10
Movimento Resistivo da Preensão Palmar	6

4.4.2 Tratamento dos Atributos Monovalorados e Multivalorados

A partir dessas métricas foram realizados testes a partir de uma base de casos contendo 40 casos colhidos em postos de saúde e com especialistas da área.

Foram cadastrados 25 casos na base de casos e 15 casos foram utilizados para a realização dos testes.

Cada bateria de casos foi testada com cada uma das seguintes métricas:

- City-Block (Manhattan);
- Contagem de Características;
- Euclidiana;
- Euclidiana Quadrada;
- Euclidiana Quadrada Ponderada;
- Vizinho mais Próximo – Função Inclusão de Consulta;
- Vizinho mais Próximo – Função Interseção;
- Vizinho mais Próximo – Função Linear.

Por meio desses testes, foram obtidos os casos mais similares, de acordo com os limites de similaridade pré-definidos (50%), definição dada pelo especialista da área, ou seja, os casos que não atingiram esse percentual foram descartados. Estes casos mais similares ficam disponíveis para o especialista (fisioterapeuta) decidir qual tratamento deverá ser aplicado para o caso em questão.

Observando-se as características das métricas de similaridades, pode-se fazer algumas considerações:

1. Alguns atributos monovalorados foram classificados como booleanos, portanto utilizam a métrica booleana, independente da métrica que está sendo utilizada.

Os atributos monovalorados booleanos do FisioSmart são:

- Irradiação da Palpação Ósteo Tendínea;
- Movimento Passivo do Punho;
- Irradiação da Palpação Muscular;
- Alongamento Passivo dos Extensores do Punho;
- Movimento Resistivo dos Extensores do Punho;
- Movimento Resistivo da Preensão Palmar;
- Movimento Passivo do Cotovelo;
- Forma de Início (pode ser progressiva ou repentina e uma forma exclui a outra).

2. Os atributos multivalorados podem utilizar os valores definidos em outros descritores, que foram realizados no processo de indexação.

Os atributos multivalorados do FisioSmart são:

- Queixa Principal;
- Outros Sintomas;
- Coincidiu com quais outras atividades;
- Melhora com quais atitudes;
- Piora com quais atitudes;
- Patologias Associadas;
- Atividades no trabalho;
- Esportes;
- Quadro Emocional;
- Quais tratamentos prévios;
- Antecedentes Familiares;
- Antecedentes Pessoais;
- Atividades no Lar.

Para a aplicação das métricas de similaridade são propostas algumas técnicas de ajuste:

1. Usualmente, as métricas utilizam a distância dos descritores dos casos para calcular o valor as similaridade dos atributos. Essa forma de cálculo é muito útil para atributos monovalorados, mas não utilizável da forma usual com atributos multivalorados. A adaptação consiste em analisar os conjuntos de valores, como explicado abaixo com os atributos presentes na Tabela 09.

Tabela 09 – Representação do Ajuste 1

Linha	Caso Novo		Caso da Base	
	Atributo - Queixa Principal	Valor do Descritor	Atributo – Queixa Principal	Valor do Descritor
01	Dor no Cotovelo	0,7	Dor no Cotovelo	0,7
02	Formigamento	0,8	Formigamento	0,8
03			Dor no Punho	0,6

No sistema proposto, faz-se uma comparação de todos os valores do atributo do caso novo com todos os valores dos atributos do caso da base (produto cartesiano). O 1º valor do atributo do caso novo é comparado com o 1º valor do atributo do caso da base, o 1º valor do atributo do caso novo é comparado com o 2º valor do atributo do caso da base, e assim por diante. Após todas as comparações, são calculadas a média e o valor máximo desse conjunto de valores de atributos, com cada uma das métricas de similaridades usadas na aplicação. A função a seguir ilustra esse cálculo:

1-(L1N – L1B) → 1- | 0,7 – 0,7| = 1

1-(L1N – L2B) → 1- | 0,7 – 0,8| = 0,9

1-(L1N – L3B) → 1- |0,7 – 0,6| = 0,9

1-(L2N – L1B) → 1- |0,8 – 0,7| = 0,9

1-(L2N – L2B) → 1- |0,8 – 0,8| = 1

1-(L2N – L3B) → 1- |0,8 – 0,6| = 0,8

Média → (1+0,9+0,9+0,9+1+0,8) /6 = 0,91

Máxima → 1

Onde, na tabela 09:

L1N → Linha 01 do Novo Caso

L2N → Linha 02 do Novo Caso

L1B → Linha 01 do Caso da Base

L2B → Linha 02 do Caso da Base

L3B → Linha 03 do Caso da Base

2. Outro ajuste também realizado, refere-se aos atributos multivalorados do item anterior. Quando se faz a comparação de atributos multivalorados, com valores idênticos nos dois casos, pode ocorrer um valor resultante inferior a 1. A tabela 10 mostra esse ajuste.

Tabela 10 – Representação do Ajuste 2

Linha	Caso Novo		Caso da Base	
	Atributo - Queixa Principal	**Valor do Descritor**	**Atributo - Queixa Principal**	**Valor do Descritor**
01	Dor no Cotovelo	0,7	Dor no Cotovelo	0,7
02	Formigamento	0,8	Formigamento	0,8

É possível observar que os valores dos casos analisados são idênticos. Portanto a similaridade deveria ser igual a 1. No entanto, se os valores forem inseridos na função o resultado poderá ser inferior a 1. Dessa forma, o ajuste foi realizado para garantir a similaridade igual a 1. A função a seguir ilustra esse cálculo:

1-(L1N – L1B) → 1- |0,7 – 0,7| = 1

1-(L1N – L2B) → 1- |0,7 – 0,8| = 0,9

1-(L2N – L1B) → 1- |0,8 – 0,7| = 0,9

1-(L2N – L2B) → 1- |0,8 – 0,8| = 1

Média → (1+0,9+0,9+1) /4 = 0,95 que deveria resultar 1

Máxima → 1

Onde, na tabela 10:

L1N → Linha 01 do Novo Caso

L2N → Linha 02 do Novo Caso

L1B → Linha 01 do Caso da Base

L2B → Linha 02 do Caso da Base

É importante observar os multivalores que podem ser assumidos pelo atributo Queixa Principal:

- Formigamento;
- Dor no cotovelo;
- Dor no membro superior;
- Dor no punho;
- Alteração de sensibilidade no membro superior.

A seguir será apresentada a modelagem do sistema.

4.5 Modelagem do Sistema

Esta seção apresenta a modelagem do sistema FisioSmart, utilizando a Análise Essencial **[Yourdon, 1992].** A opção pelo uso dessa metodologia, deve-se ao fato que esta abordagem proporciona simplicidade na tarefa de modelagem. Dessa forma utilizou-se as seguintes ferramentas:

- Lista de Eventos,
- Diagrama de Contexto,
- Diagramas de Fluxo de Dados (DFD's),

- Descrição dos Processos,
- Modelo Entidade-Relacionamento (MER) e
- Dicionário de Dados.

4.5.1 Lista de Eventos

Eventos são acontecimentos externos ao sistema que desencadeiam alguma ação de resposta. A lista de eventos é uma lista de estímulos do ambiente que o sistema deve responder.

Após a definição dos objetivos do sistema, iniciou-se a construção da Lista de Eventos, a seguir:

1. Usuário cadastra Paciente
2. Usuário cadastra Medicamentos
3. Usuário informa histórico do Paciente
4. Usuário configura pesos
5. Usuário cadastra vocabulário
6. Usuário armazena casos na base de casos
7. Usuário informa dados do Novo Caso a ser pesquisado
8. Usuário escolhe métrica
9. Usuário obtém casos similares

4.5.2 Diagrama de Contexto

O Diagrama de contexto trata da descrição gráfica que representa a troca de informações entre o sistema e o ambiente.

A figura 04 mostra o Diagrama de Contexto, baseado na Lista de Eventos.

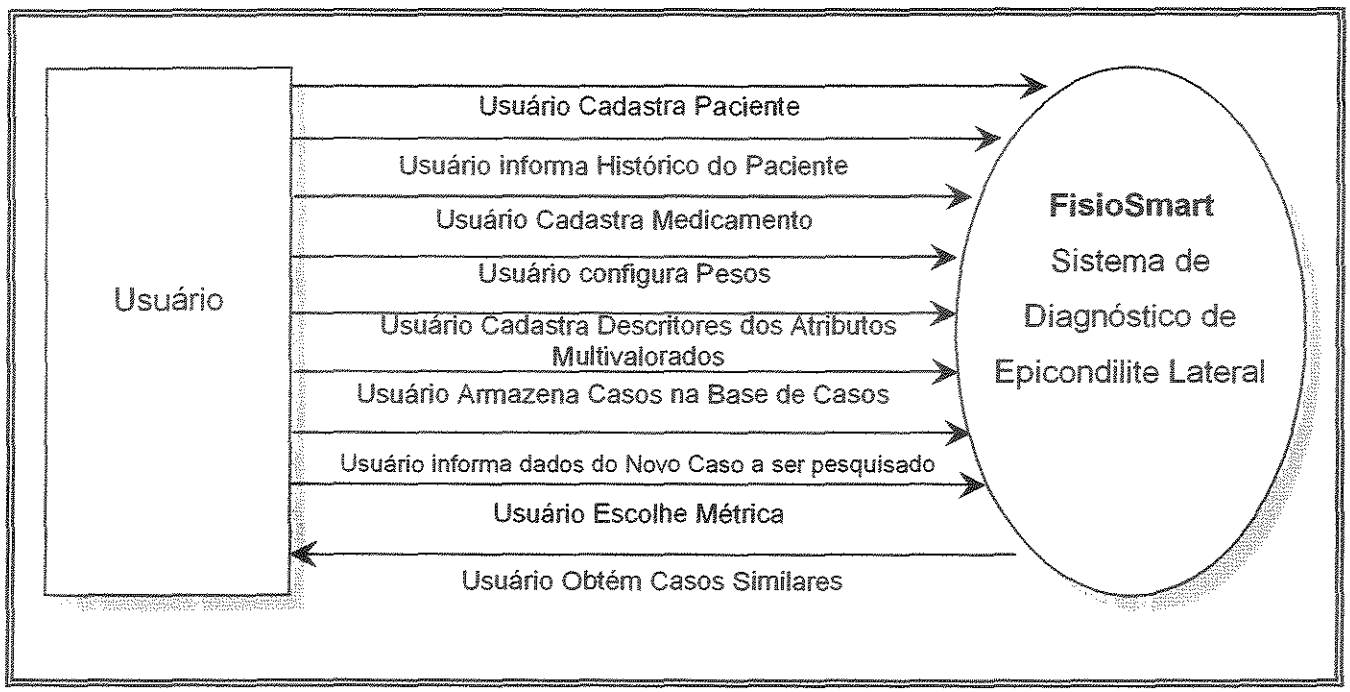


Figura 04 – Diagrama de Contexto do FisioSmart

4.5.3 Diagramas de Fluxo de Dados (DFD's)

O Diagrama de Fluxo de Dados é uma forma de representação gráfica de um sistema, e mostra os componentes funcionais do sistema e as interfaces de dados entre eles, representando assim seu funcionamento.

As figuras 05 e 06 mostram os DFD's de Cadastro e Recuperação dos Casos, respectivamente.

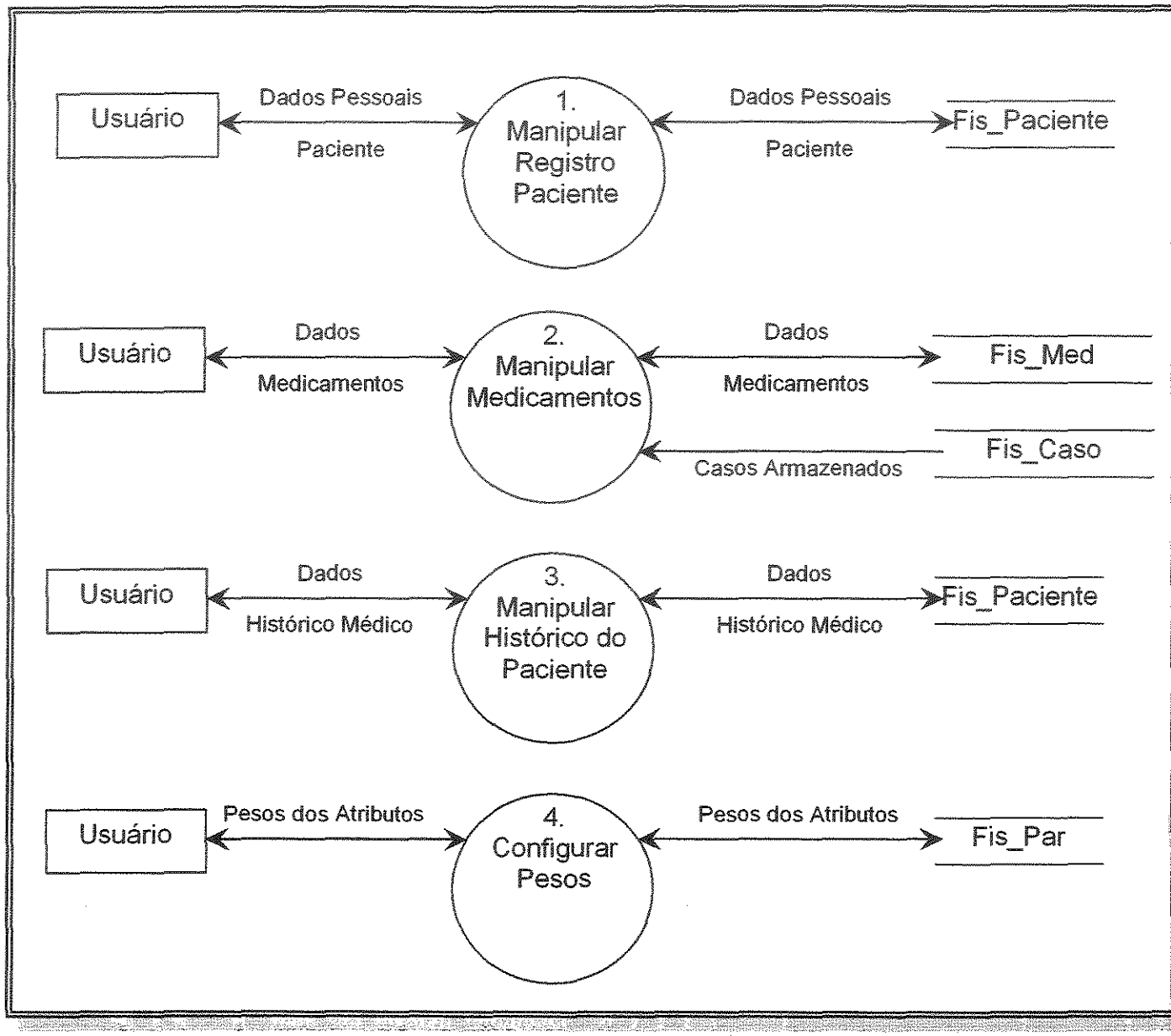


Figura 05 – Diagrama de Fluxo de Dados de Cadastro

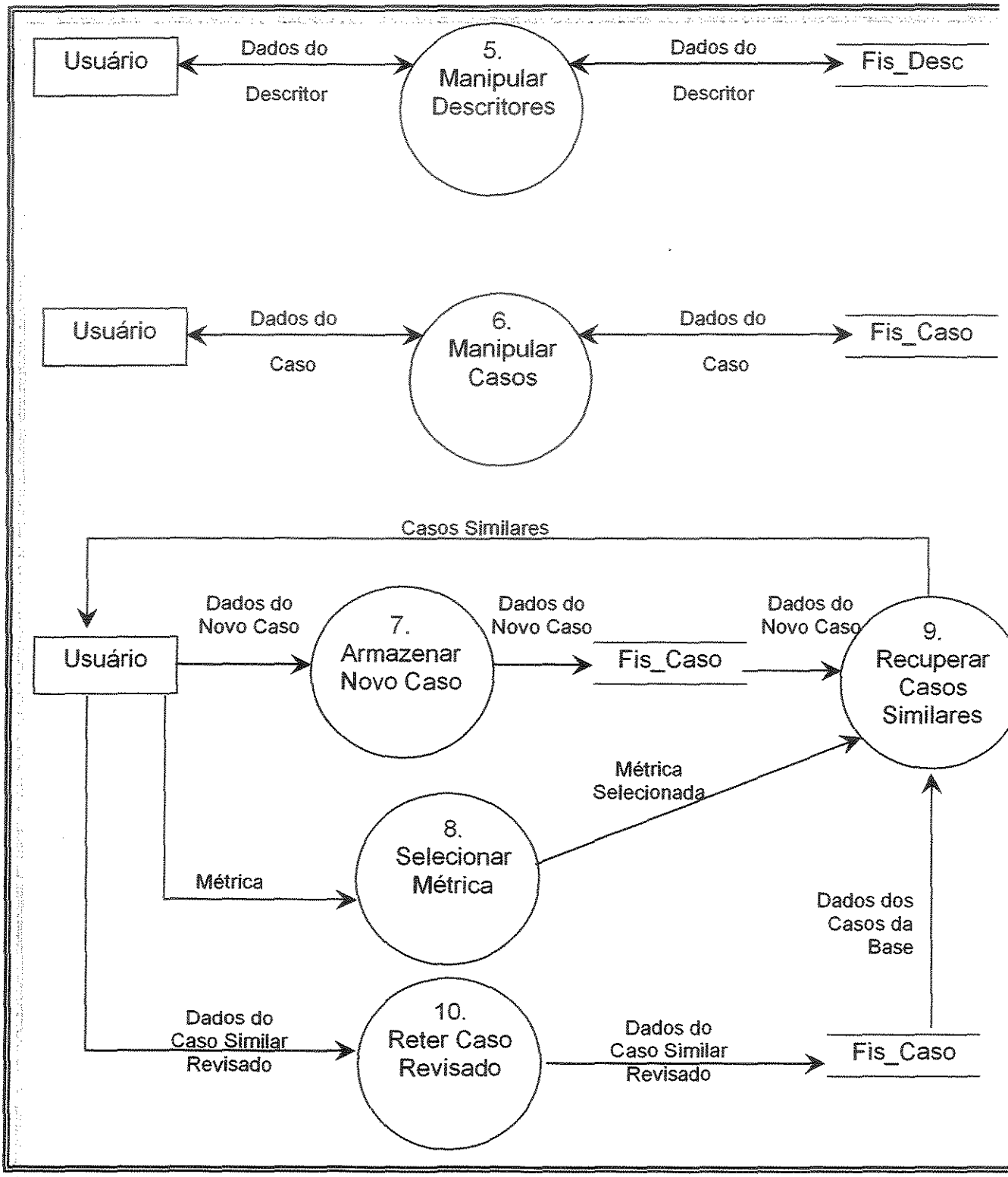


Figura 06 – Diagrama de Fluxo de Dados da Recuperação de Casos

4.5.4 Descrição dos Processos

A seguir, os processos são descritos, através de uma descrição textual de cada um deles.

4.5.4.1 Processo 1 – Manipular Registro Paciente

Este processo faz a manutenção do Registro do Paciente, possibilitando inclusão, consulta, alteração e exclusão. A exclusão só será permitida caso o paciente não possua casos associados.

4.5.4.2 Processo 2 – Manipular Medicamentos

Este processo faz a manutenção da Base de Dados de Medicamentos, possibilitando inclusão, consulta, alteração e exclusão. A exclusão só será permitida, caso o medicamento não esteja relacionado a um paciente.

4.5.4.3 Processo 3 – Manipular Histórico do Paciente

Este processo faz a manutenção do Histórico do Paciente, possibilitando inclusão, consulta, alteração e exclusão de informações.

4.5.4.4 Processo 4 – Configurar Pesos

Este processo permite a atribuição de pesos de consulta para cada um dos atributos analisados na pesquisa de casos similares.

4.5.4.5 Processo 5 – Manipular Descritores

Este processo permite a manutenção da base de Descritores do Atributo Multivalorado, onde é possível incluir, consultar, alterar e excluir descritores. Neste processo é possível definir os valores de similaridade entre os descritores contidos em um mesmo Atributo Multivalorado.

4.5.4.6 Processo 6 – Manipular Casos

Neste processo é possível manipular a base de casos, incluindo, consultando, alterando e excluindo casos da base. Para cada caso há diversos atributos que serão armazenados e posteriormente utilizados para o cálculo da similaridade. Este processo é utilizado para definir a base inicial de casos.

4.5.4.7 Processo 7 – Armazenar Novo Caso

Este processo coleta informações sobre o novo caso que vai ser consultado, armazenando as informações temporariamente para comparação com os casos na base de casos.

4.5.4.8 Processo 8 – Selecionar Métrica

Este processo permite ao usuário a escolha da métrica, com a qual serão realizados os cálculos de similaridade para encontrar os casos mais similares na base de casos.

4.5.4.9 Processo 9 – Recuperar Casos Similares

Este processo é responsável pelo cálculo das métricas escolhidas, analisando o novo caso armazenado temporariamente com os demais casos da base, recuperando os casos mais similares, estes casos serão revisados pelo usuário.

4.5.4.10 Processo 10 – Reter Caso Similar Revisado

Este processo é responsável pelo armazenamento definitivo do novo caso na base de casos, após a revisão do usuário.

4.5.5 Modelo Entidade Relacionamento (MER)

A abordagem Entidade Relacionamento é uma técnica de modelagem conceitual de propriedades estáticas de sistemas.

O Modelo ER (MER), proposto por CHEN em 1977, especifica a estrutura lógica geral do BD baseado na percepção do mundo real. É montado sobre três

conceitos básicos: a entidade, o relacionamento e o atributo. Mais tarde, com a difusão do modelo, diversas extensões foram propostas, tendo outros conceitos como a entidade associativa, e a generalização/especialização sido incorporados ao modelo.

Para a modelagem do banco de dados do sistema, utilizou-se a ferramenta Modelo Entidade Relacionamento, conforme a figura 07. Utilizou-se, para a sua construção, a ferramenta *Case Power Designer Data Architect* 6.1.

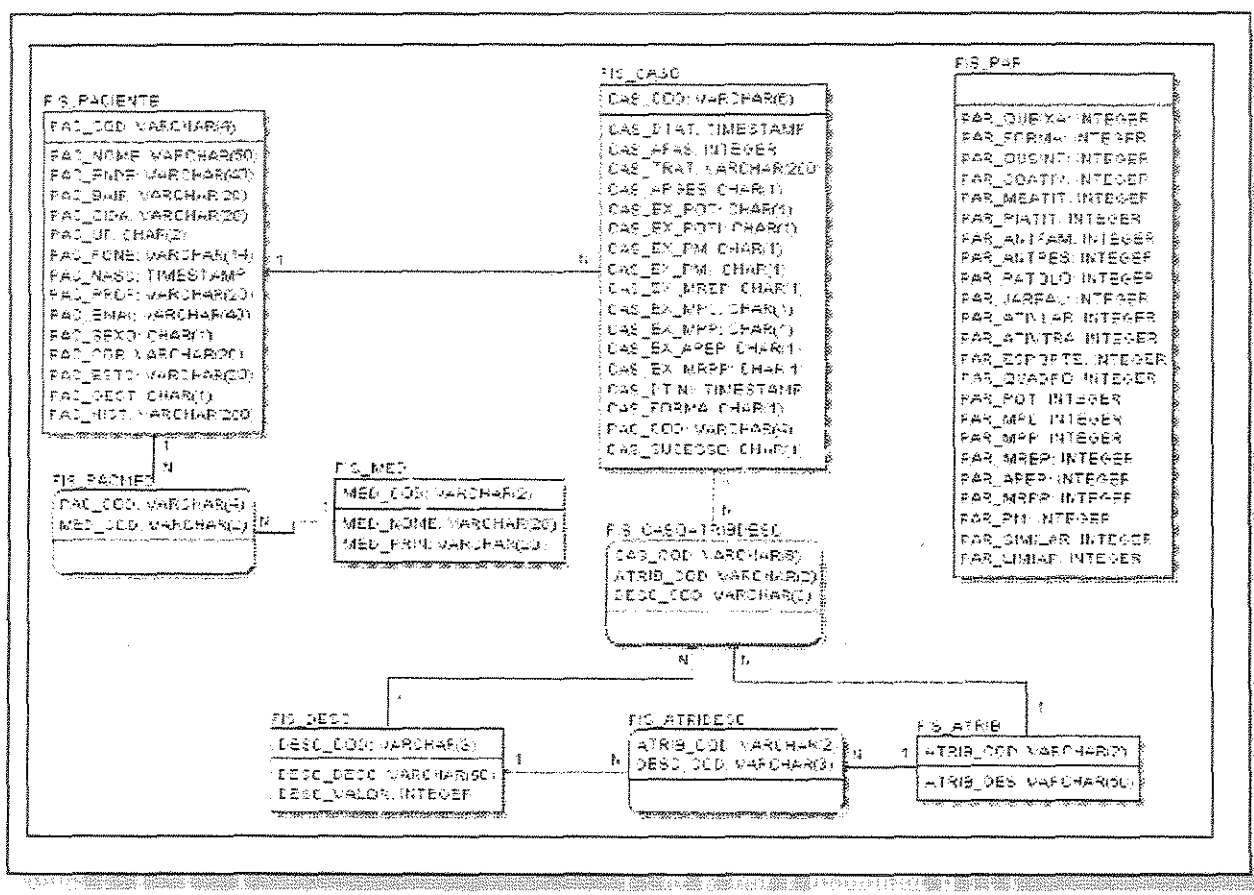


Figura 07 – Modelo do Banco de Dados do FisioSmart

4.5.6 Dicionário de Dados

O dicionário de dados consiste na descrição de todas as tabelas componentes da base de dados, assim como seus itens de dados e fluxos dos processos modelados nos DFD's, definindo, assim, a composição/estrutura dos mesmos.

4.5.6.1 Descrição das Tabelas de Dados

A tabela 11 apresenta a descrição das tabelas de dados do sistema.

Tabela 11 – Descrição das Tabelas de Dados utilizadas no FisioSmart

Nome das Tabelas	Descrição
FIS_ATRIB	Atributos cadastrados
FIS_CASO	Casos da base
FIS_DESC	Armazena os descritores do(s) atributo(s) multivalorado(s)
FIS_ATRIBDESC	Permite o relacionamento entre as tabelas de Atributos e Descritores, normalização de um relacionamento muitos-para-muitos
FIS_CASOATRIBDESC	Permite o relacionamento entre as tabelas de Casos, Atributos e Descritores, normalização de relacionamentos muitos-para-muitos
FIS_MED	Cadastro dos medicamentos
FIS_PACIENTE	Informações pessoais dos pacientes
FIS_PACMED	Permite o relacionamento entre as tabelas Pacientes e Medicamentos, normalização de um relacionamento muitos-para-muitos
FIS_PAR	Tabela de Parâmetros. Armazena os pesos dos atributos. Os pesos são utilizados em algumas métricas na busca por casos mais similares.

4.5.6.2 Descrição dos Atributos das Tabelas

As tabelas 12 a 20 apresentam a descrição dos atributos do banco de dados, organizadas em quatro colunas. A primeira possui o nome do atributo, a segunda coluna representa se o atributo é uma chave sendo convencionados os seguintes valores: "PK" – chave primária e "FK" – chave estrangeira, a terceira coluna determina o tipo do atributo, e a última coluna descreve o atributo.

Tabela 12 – Dicionário de Dados da Tabela de Atributos – FIS_ATRIB

Atributo	Chave	Tipo	Descrição
ATRIB_COD	PK	VARCHAR(2)	Atributo inequívoco que identifica cada atributo armazenado na tabela
ATRIB_DES	-	VARCHAR (50)	Descrição do Atributo armazenado

Tabela 13 – Dicionário de Dados da Tabela de Casos – FIS_CASO

Atributo	Chave	Tipo	Descrição
CAS_COD	PK	VARCHAR (6)	Atributo inequívoco que identifica um caso
CAS_DTAT	-	DATE	Data do caso
CAS_AFAS	-	INTEGER	Duração do afastamento do paciente
CAS_TRAT	-	VARCHAR (200)	
CAS_APGES	-	CHAR (1)	Apenas gestante
CAS_EX_POT	-	CHAR (1)	Exame Físico Palpação Ósteo Tendínea
CAS_EX_POTI	-	CHAR (1)	Irradiação Palpação Ósteo Tendínea
CAS_EX_PM	-	CHAR (1)	Exame Físico Palpação Muscular
CAS_EX_PMI	-	CHAR (1)	Irradiação Palpação Muscular
CAS_EX_MREP	-	CHAR (1)	Exame Físico Movimento Resistivo dos Extensores do Punho
CAS_EX_MPC	-	CHAR (1)	Exame Físico Movimento Passivo do Cotovelo
CAS_EX_MPP	-	CHAR (1)	Exame Físico Movimento Passivo do Punho
CAS_EX_APEP	-	CHAR (1)	Exame Físico Alongamento passivo dos extensores do punho
CAS_EX_MRPP	-	CHAR (1)	Exame Físico Movimento Resistivo da preensão palmar
CAS_DTINI	-	DATE	Data de início do sintoma
CAS_FORMA	-	CHAR (1)	Forma de início
PAC_COD	-	VARCHAR (4)	Código do paciente
CAS_SUCESSO	-	CHAR (1)	Tratamento obteve sucesso

Tabela 14 – Dicionário de Dados da Tabela de Descritores - FIS_DESC

Atributo	Chave	Tipo	Descrição
DESC_COD	FK	VARCHAR (3)	Código do Descritor
DESC_DESC	-	VARCHAR (50)	Descrição do Descritor
DESC_VALOR	-	INTEGER	Determina distância entre um descritor e outro

Tabela 15 – Dicionário de Dados da Tabela de Relacionamento entre Atributos e Descritores - FIS_ATRIBDESC

Atributo	Chave	Tipo	Descrição
ATRIB_COD	FK, PK	VARCHAR (2)	Código do Atributo
DESC_COD	FK, PK	VARCHAR (3)	Código do Descritor

Tabela 16 – Dicionário de Dados da Tabela de Relacionamento entre Casos, Atributos e Descritores - FIS_CASOATRIBDESC

Atributo	Chave	Tipo	Descrição
CAS_COD	FK, PK	VARCHAR (6)	Código do Caso
ATRIB_COD	FK, PK	VARCHAR (2)	Código do Atributo
DESC_COD	FK, PK	VARCHAR (3)	Código do Descritor

Tabela 17 – Dicionário de Dados da Tabela de Medicamentos – FIS_MED

Atributo	Chave	Tipo	Descrição
MED_COD	PK	VARCHAR (2)	Código do Medicamento
MED_NOME	-	VARCHAR (20)	Nome do Medicamento
MED_PRINC	-	VARCHAR (20)	Princípio Ativo do Medicamento

Tabela 18 – Dicionário de Dados da Tabela de Pacientes – FIS_PAC

Atributo	Chave	Tipo	Descrição
PAC_COD	PK	VARCHAR (4)	Código do Paciente
PAC_NOME	-	VARCHAR (50)	Nome do Paciente
PAC_ENDE	-	VARCHAR (40)	Endereço do Paciente
PAC_BAIR	-	VARCHAR (20)	Bairro do Paciente
PAC_CIDA	-	VARCHAR (20)	Cidade do Paciente
PAC_UF	-	CHAR (2)	Estado da cidade do paciente
PAC_FONE	-	VARCHAR (14)	Telefone do Paciente
PAC_NASC	-	TIMESTAMP	Data de Nascimento do paciente
PAC_PROF	-	VARCHAR (20)	Profissão do Paciente
PAC_EMAIL	-	VARCHAR (40)	E-mail do paciente
PAC_SEXO	-	CHAR (1)	Sexo do Paciente
PAC_COR	-	VARCHAR (20)	Cor da Pele do paciente
PAC_ESTC	-	VARCHAR (20)	Estado Civil do Paciente
PAC_GEST	-	CHAR (1)	Identifica se paciente é gestante
PAC_HIST	-	VARCHAR (200)	Histórico do Paciente

Tabela 19 – Dicionário de Dados da Tabela de Relacionamento entre Medicamentos e Pacientes – FIS_PACMED

Atributo	Chave	Tipo	Descrição
PAC_COD	FK, PK	VARCHAR (4)	Código do Paciente
MED_COD	FK, PK	VARCHAR (2)	Código do Medicamento

Tabela 20 – Dicionário de Dados da Tabela de Parâmetros – FIS_PAR

Atributo	Chave	Tipo	Descrição
PAR_QUEIXA	-	INTEGER	Peso do atributo – Queixa principal
PAR_FORMA	-	INTEGER	Peso do atributo - Forma de inicio
PAR_OUSINT	-	INTEGER	Peso do atributo - Outros sintomas
PAR_COATIV	-	INTEGER	Peso do atributo - Coincidiu com outras atividades
PAR_MEATIT	-	INTEGER	Peso do atributo - Melhora com alguma atitude
PAR_PIATIT	-	INTEGER	Peso do atributo - Piora com alguma atitude
PAR_ANTFAM	-	INTEGER	Peso do atributo – Antecedentes familiares
PAR_ANTPES	-	INTEGER	Peso do atributo – Antecedentes pessoais
PAR_PATOLO	-	INTEGER	Peso do atributo – Patologia Associada
PAR_JAREAL	-	INTEGER	Peso do atributo – Já realizou algum tratamento prévio
PAR_ATIVLAR	-	INTEGER	Peso do atributo – Atividades no lar
PAR_ATIVTRA	-	INTEGER	Peso do atributo – Atividade no trabalho
PAR_EXPORTE	-	INTEGER	Peso do atributo – Pratica algum exporte
PAR_QUADRO	-	INTEGER	Peso do atributo – Quadro emocional
PAR_POT	-	INTEGER	Peso do atributo – Palpação ósteo tendínia
PAR_MPC	-	INTEGER	Peso do atributo – Movimento passivo do cotovelo
PAR_MPP	-	INTEGER	Peso do atributo – Movimento passivo do punho
PAR_MREP	-	INTEGER	Peso do atributo – Movimento resistivo dos extensores do punho
PAR_APEP	-	INTEGER	Peso do atributo – Alongamento passivo dos extensores do punho
PAR_MRPP	-	INTEGER	Peso do atributo – Movimento resistivo da preensão palmar
PAR_PM	-	INTEGER	Peso do atributo – Palpação muscular
PAR_SIMILAR	-	INTEGER	Mínimo de similaridade para retorno de um caso
PAR_LIMIAR	-	INTEGER	Limiar de retorno.

4.6 Telas

A tela apresentada pela figura 8, permite a entrada em outros módulos do sistema, através do menu principal o usuário pode acessar o cadastro de pacientes, medicamentos, descritor, casos e configurar pesos.

A figura 08 mostra a tela de entrada do sistema FisioSmart.

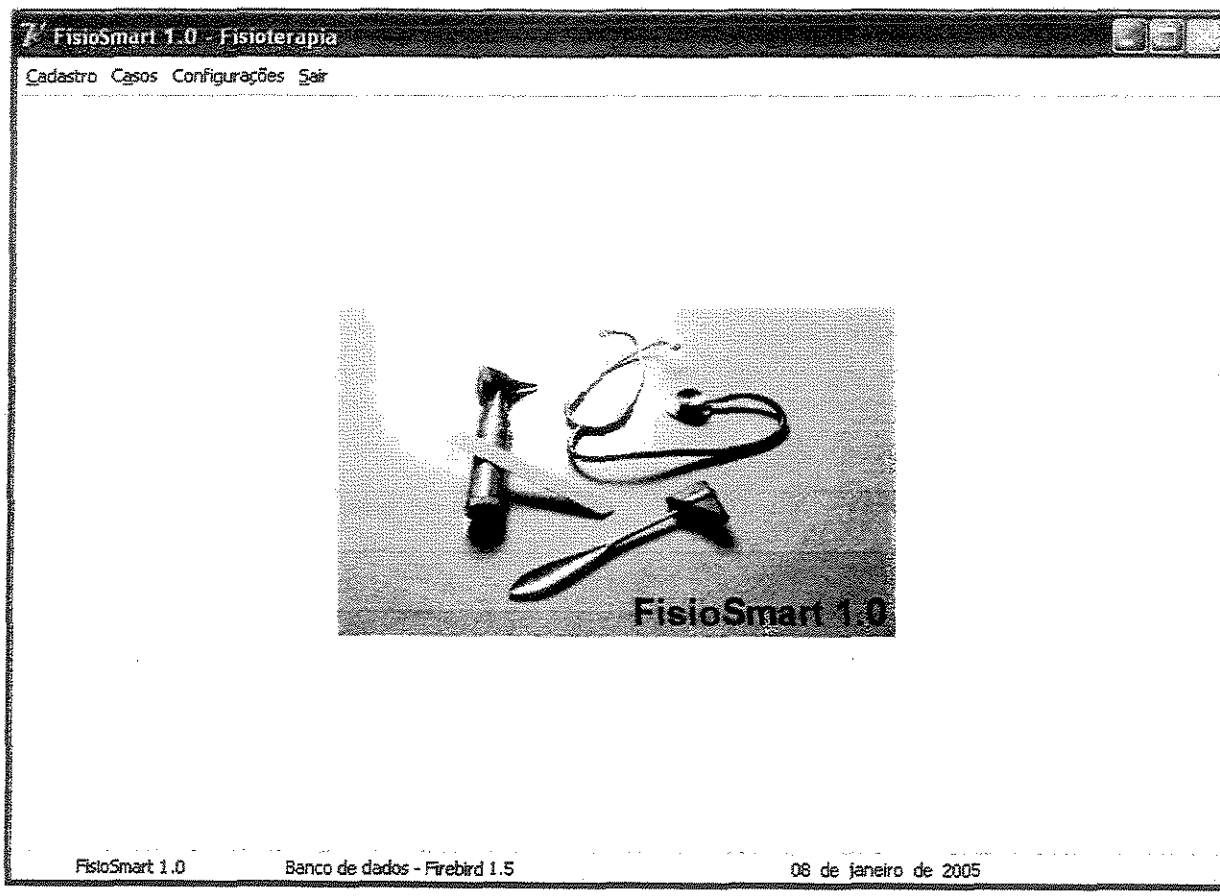


Figura 08 – Tela de Entrada do Fisiosmart

4.6.1 Cadastro de Pacientes

O usuário inclui os dados do paciente conforme os descritores de cada campo, após informar os dados o usuário clica na aba "histórico" para incluir seu histórico médico e os possíveis medicamentos que o mesmo pode utilizar.

A figura 09 mostra a tela de cadastros de pacientes do FisioSmart e a figura 10 mostra a tela de cadastros de histórico dos pacientes do FisioSmart.

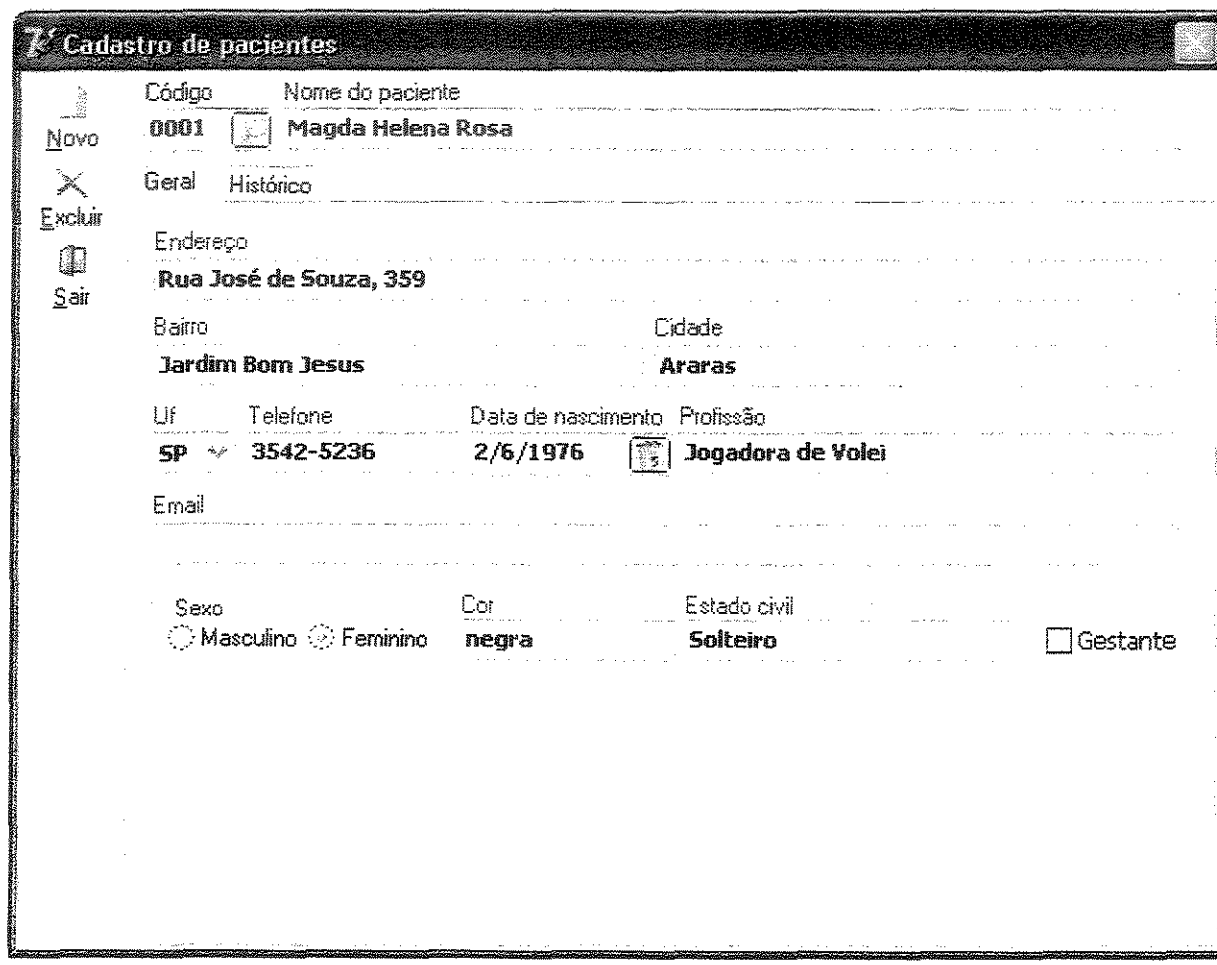


Figura 09 – Tela de Cadastro de Pacientes

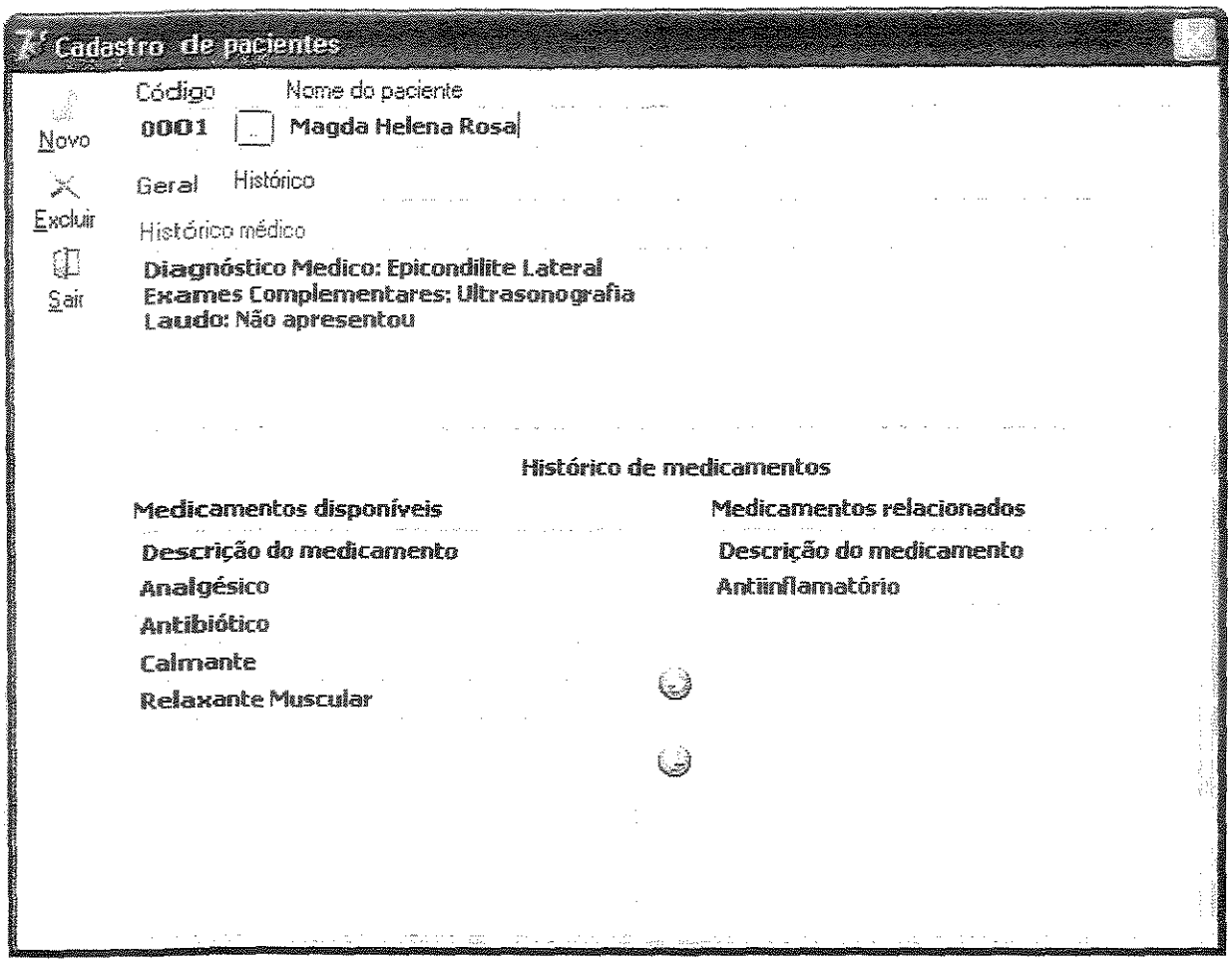


Figura 10 – Tela de Cadastro do Histórico de Pacientes

4.6.2 Cadastro de Descritores

O cadastro dos descritores é onde o especialista define a distância entre os descritores de cada atributo, o especialista define uma distância de 0 a 10 entre os descritores, é através destes valores que as métricas podem distinguir distância entre os descritores de cada atributo.

A figura 11 mostra a tela de cadastros de descritores do FisioSmart.

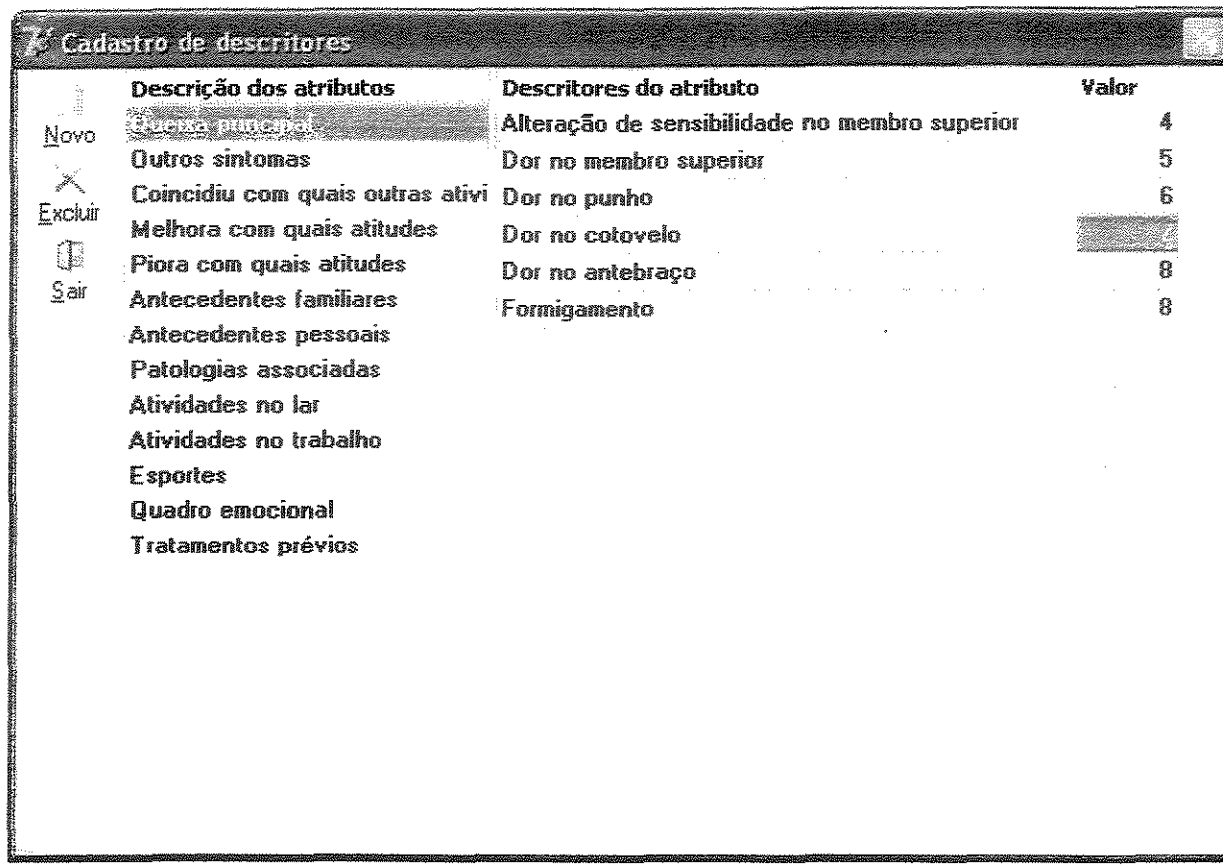


Figura 11 – Tela de Cadastro de Descritor

4.6.3 Cadastro de Medicamentos

O usuário cadastra os medicamentos a serem relacionados no cadastro de paciente.

A figura 12 mostra a tela de cadastros de medicamentos do FisioSmart.

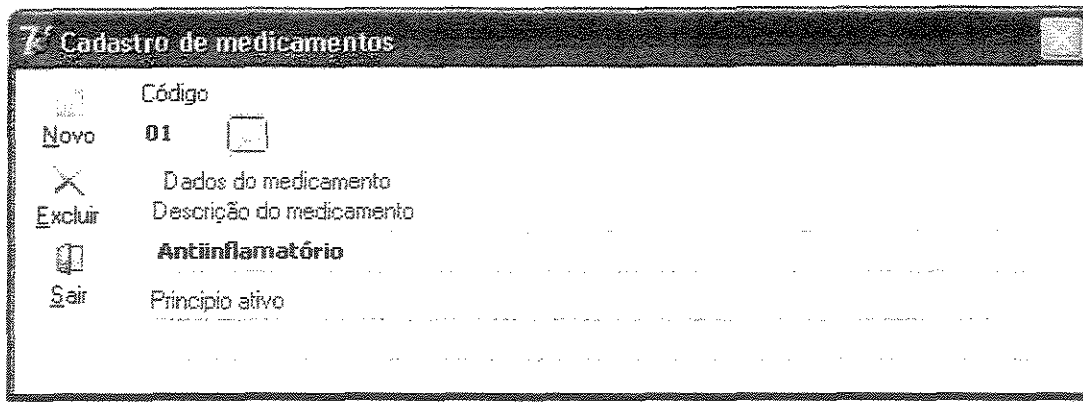


Figura 12 – Tela de Cadastro de Medicamentos

4.6.4 Cadastro de Casos

O usuário insere um novo caso, para que o sistema possa recuperar o caso mais similar ao caso atual. O usuário informa o paciente que está sendo atendido e o resultado de cada exame que o mesmo já realizou. Após informar o resultado dos exames, o usuário clica na aba "atributos" e informa os sintomas do paciente através de cada atributo que esta disponível no sistema.

A figura 13 mostra a tela de cadastros de casos do FisioSmart.

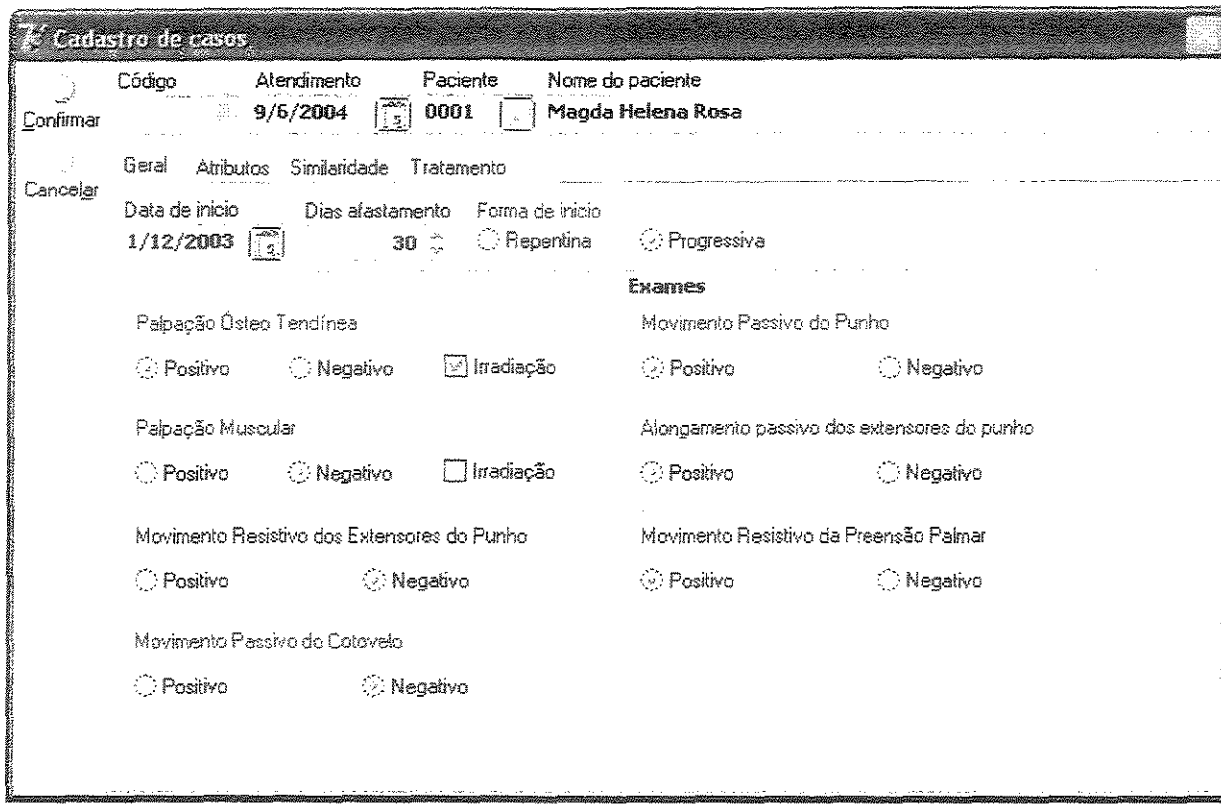


Figura 13 – Tela de Cadastro de Casos

A figura 14 mostra a tela de cadastros dos atributos dos casos do FisioSmart.

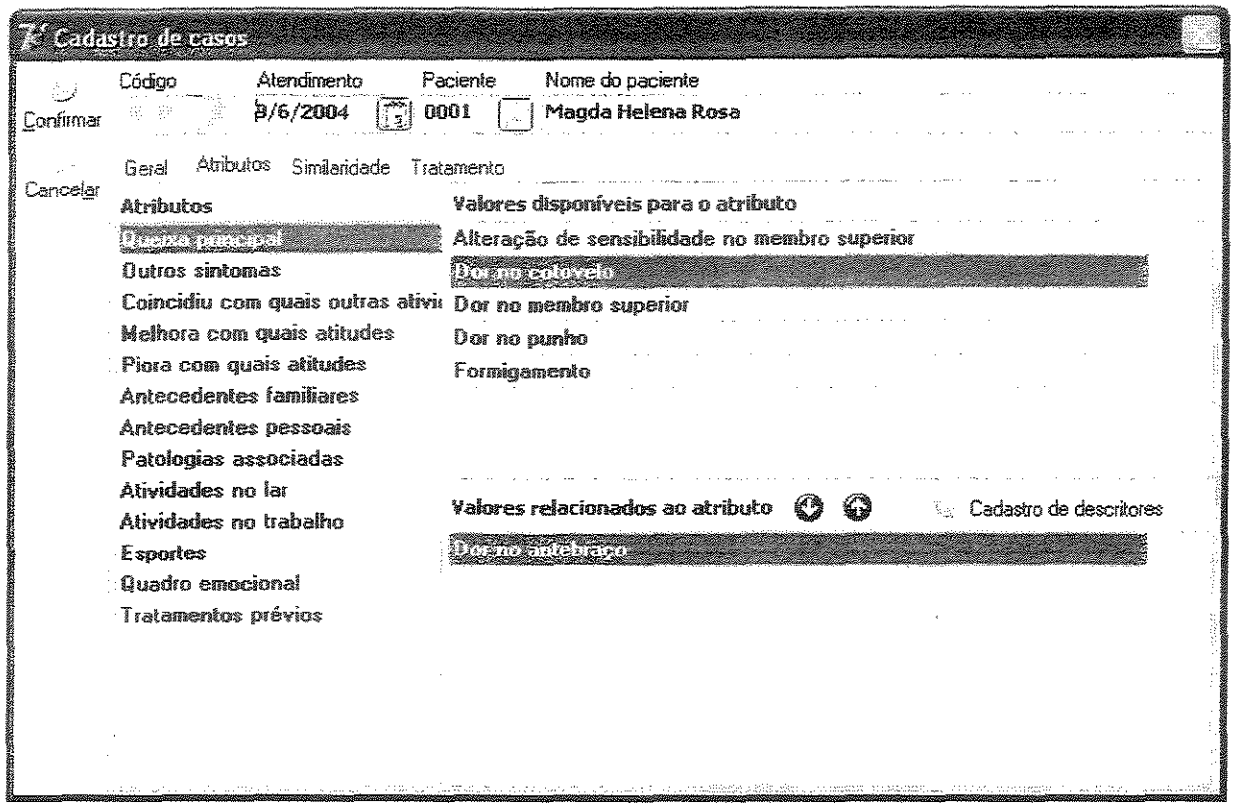


Figura 14 – Tela de Cadastro dos Atributos dos Casos

Após informar os dados apresentados o usuário clica na aba "similaridade" para obter o caso mais similar ao atual. Para isto, o sistema disponibiliza várias métricas de similaridade.

A figura 15 mostra a tela dos cálculos de similaridade dos casos do FisioSmart.

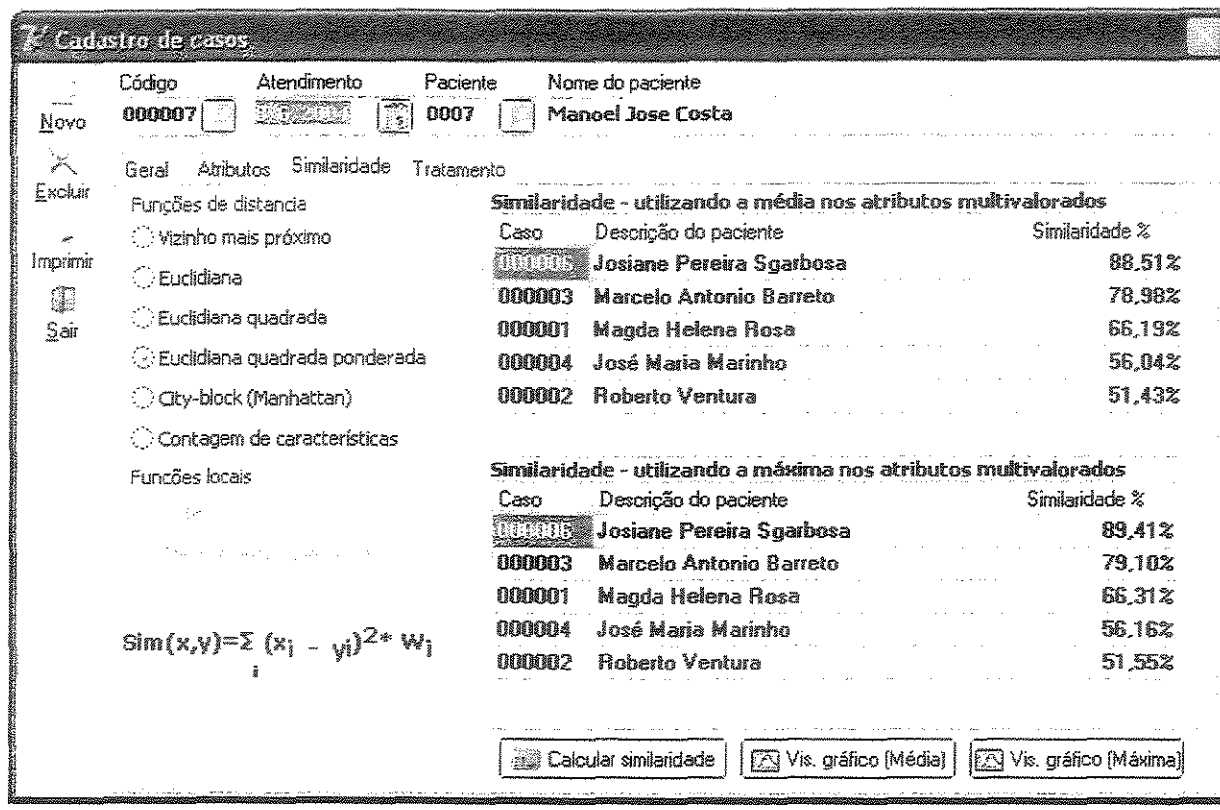


Figura 15 – Tela de Similaridade dos Casos

Na aba "tratamento", o especialista pode visualizar o tratamento aplicado para o primeiro e segundo caso mais similar. Através disso, o especialista pode adotar o tratamento aplicado ou pode adaptá-lo para o caso em questão, tornando o mesmo uma nova fonte de referência para novos casos.

A figura 16 mostra a tela de cadastro do tratamento dos casos do FisioSmart.

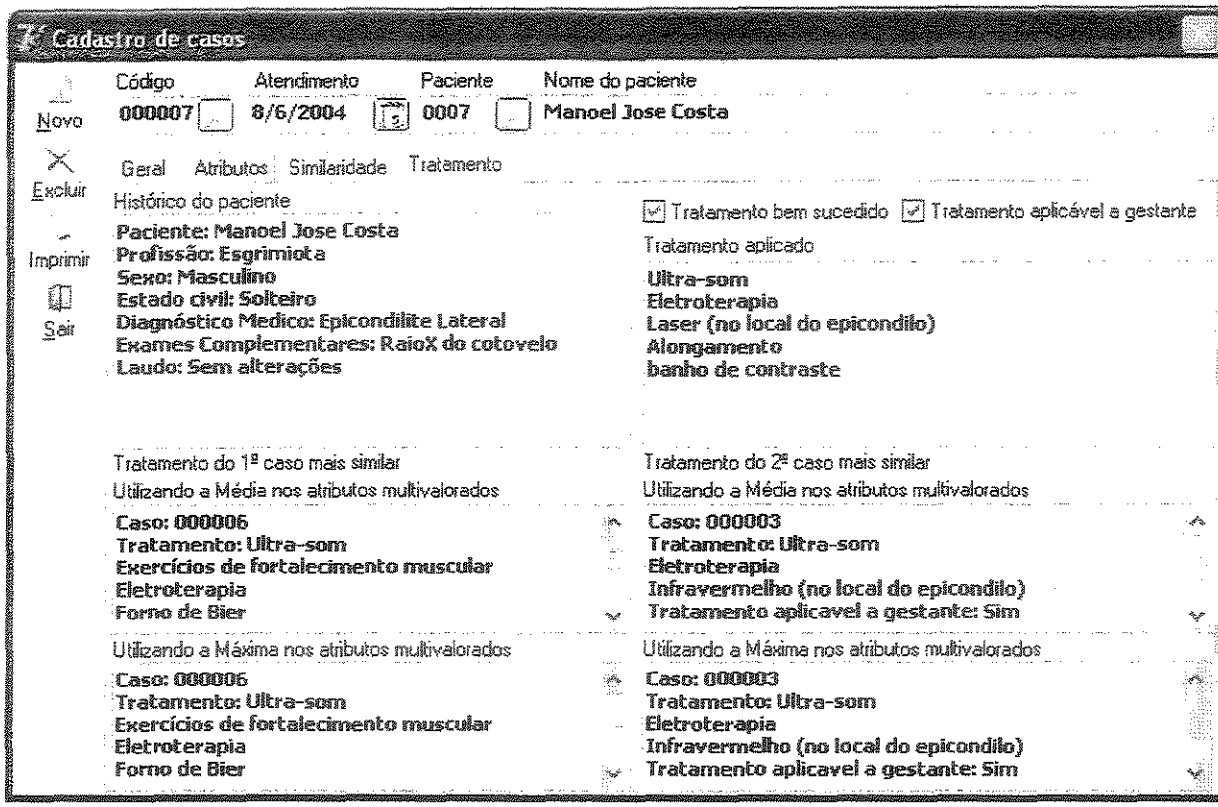


Figura 16 – Tela de Cadastro do Tratamento

4.6.5 Cadastro de Pesos

O especialista atribui um peso para cada atributo do sistema de acordo com o seu grau de importância, sendo que o mesmo pode atribuir um limiar, ou seja, o número máximo de registros que se deseja como retorno e um valor mínimo de similaridade de registros que serão retornados.

A figura 17 mostra a tela de cadastro dos pesos dos atributos do FisioSmart.

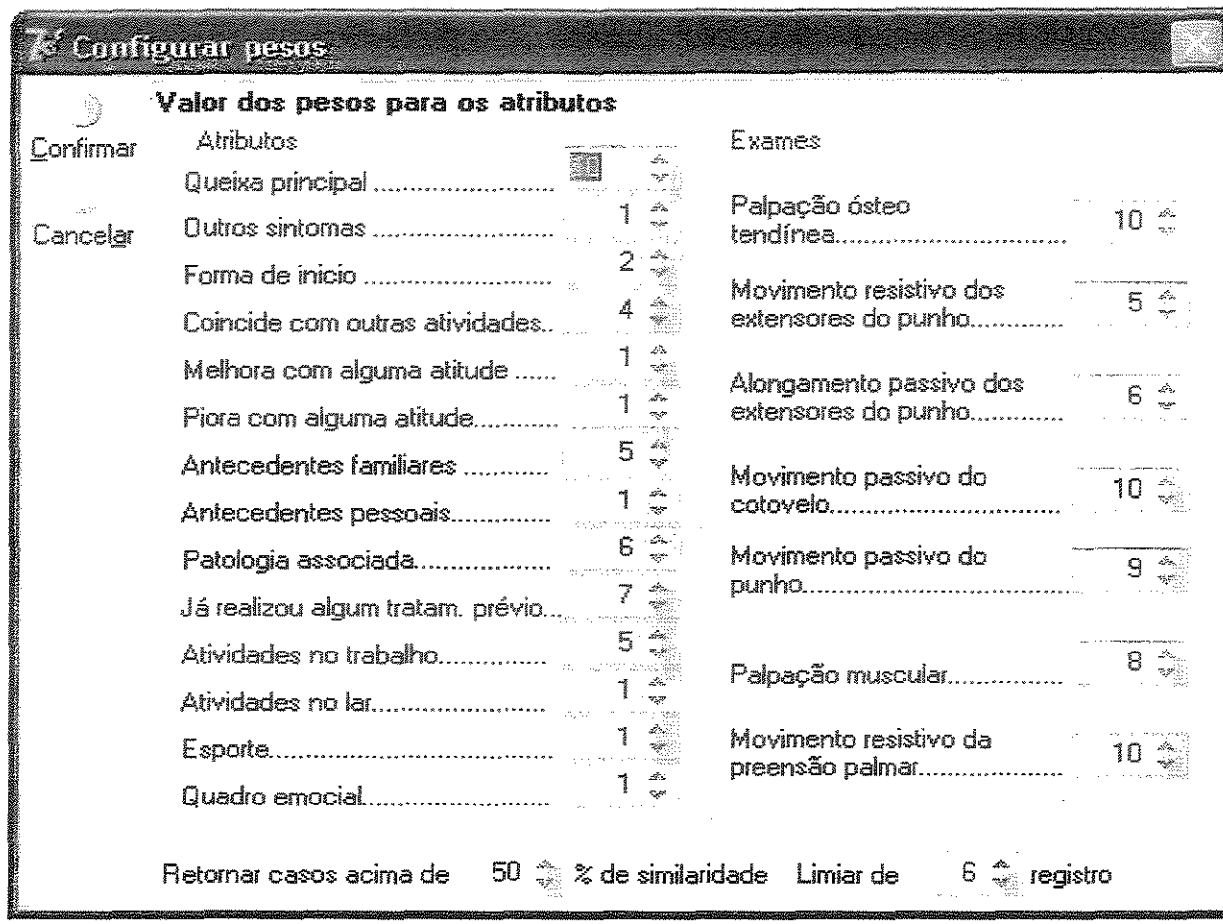


Figura 17 – Tela de Cadastro dos Pesos dos Atributos

4.7 Gráficos e Relatórios

Após obter a lista dos casos mais similares, o especialista pode clicar no botão "visualizar gráfico". O gráfico gerado exibe o grau de similaridade de cada atributo do sistema.

A figura 18 mostra a gráfico de similaridade dos atributos do FisioSmart.

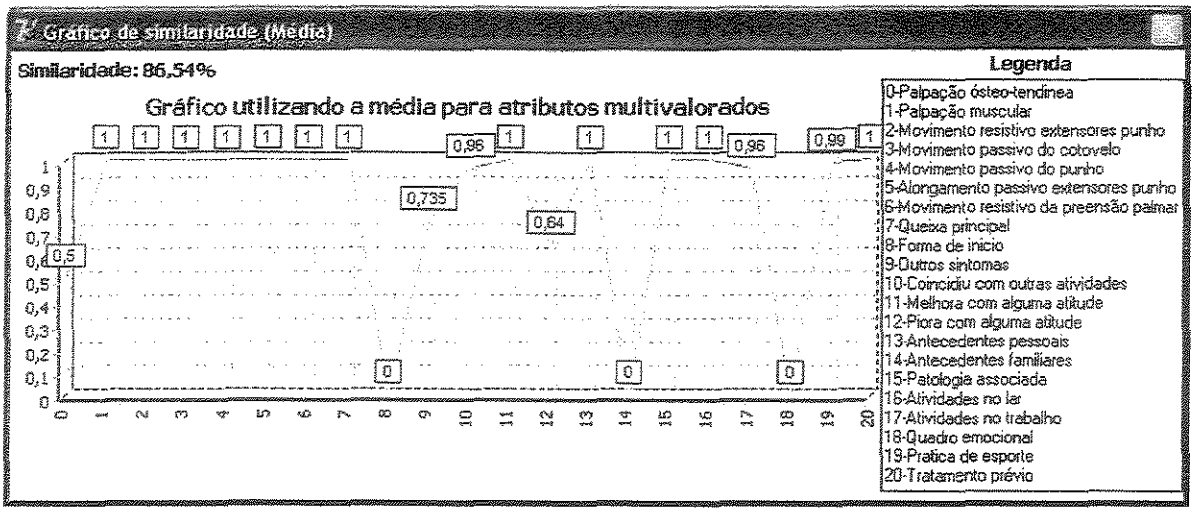


Figura 18 – Gráfico de Similaridade

O usuário pode imprimir a descrição de um caso para análise posterior. A figura 19 mostra a listagem dos casos cadastrados no FisioSmart.

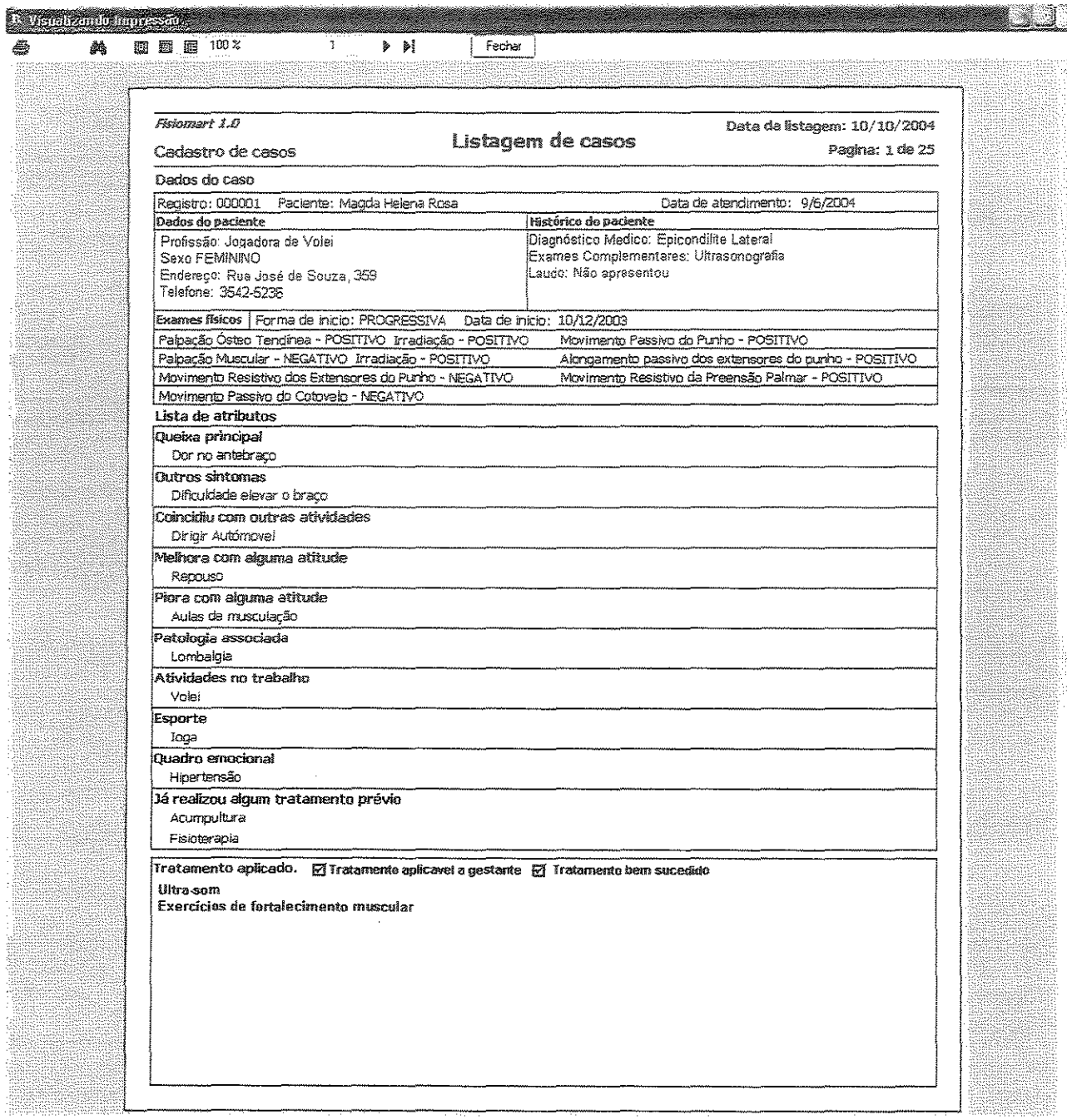


Figura 19 – Listagem dos Casos

4.8 Resultados Obtidos

A partir dos testes efetuados com o FisioSmart e com uma base de casos contendo 25 casos e 15 casos novos para teste, foi possível observar alguns comportamentos diferentes de resultados diante das métricas utilizadas. Dessa forma, o 1º caso de teste foi comparado com os 25 casos da base, o 2º caso de teste foi comparado com 26 casos da base, uma vez que o 1º caso de teste já foi inserido na base de casos e assim por diante. No final se obteve 40 casos na base, 25 casos iniciais da base mais os 15 novos casos que foram inseridos e comparados com todos os casos da base no momento da busca pela similaridade.

Grande parte dos atributos são multivalorados. Dessa forma as métricas utilizadas em sua forma original, sem serem submetidas aos ajustes, foram a do vizinho mais próximo utilizando a função interseção e a contagem de características. As demais sofreram esses ajustes.

A partir dos testes realizados foi gerada uma tabela para cada métrica utilizada, mostrando o caso mais similar retornado, o valor máximo e médio em percentual, a classificação dos casos mais similares feita pelo especialista e a comparação entre os casos retornados pela aplicação e pelo especialista. Utilizou-se o valor máximo e o valor médio para se realizar um melhor comparativo entre os resultados obtidos a partir do ajuste 1, no cálculo dos atributos multivalorados, pois algumas métricas não são implementadas diretamente em sua forma original.

A tabela 21 mostra a Métrica City Block (Manhattan).

Tabela 21 – Métrica City Block (Manhattan)

Caso de Teste	Casos Retornados pela Média	Valor Média %	Casos Retornados pela Máxima	Valor Máxima %	Classificação do Especialista	Classificação do Caso Recuperado
26	1-21-2	63,00 - 59,50 - 59,00	1-21-2	63,00 - 59,50 - 59,00	1, 5, 21	1º
27	12-9-24	93,65 - 81,73 - 80,77	12-9-24	93,65 - 81,73 - 80,77	12, 17, 25	1º
28	12-27-7	89,50 - 87,00 -84,25	12-27-7	90,00 - 87,50 - 85,50	17, 8, 12	3º
29	27-10-3	75,00-74,75-73,50	27-10-3	75,00-75,00-73,50	3, 9, 20	3º
30	17-27-25	84,50-82,00-81,00	17-27-25	84,50-82,00-81,00	15, 17, 20	2º
31	27-13-12	66,50-66,00-65,50	27-13-12	66,50-66,00-65,50	9, 27, 13	2º
32	19-2-27	82,50-65,25-63,75	19-2-27	83,50-65,50-64,20	19, 11, 22	1º
33	20-11-25-23-10	72,00-68,50-68,00-67-50-66,50	20-11-25-23-10	72,50-69,00-68,50-68-00-67,00	10, 9, 13	5º
34	12-27-17	93,00-84,00-81,00	12-27-17	95,00-86,00-83,00	12, 17, 3	1º
35	12-27-25	83,75-82,25-79,25	12-27-25	85,00-83,50-80,50	34, 3, 12	3º
36	25-8-4	83,00-80,50-78,50	25-8-4	83,00-80,50-78,50	8, 9, 25	2º
37	34-35-12	85,50-85,00-84-50	34-35-12	87,00-87,00-85,00	17, 3, 4, 34	4º
38	35-9-4	88,00-83,25-75,75	35-9-4	88,50-83,50-76,00	9, 10, 35	2º
39	4-25-37	89,00-85,00-83,00	4-25-37	90,00-85,50-84,00	4, 9, 17	1º
40	25-4-37	90,50-89,00-85,00	25-4-37	91,00-90,00-85,50	17, 4, 39	2º

A Figura 20 mostra os melhores resultados da Métrica City Block (Manhattan).

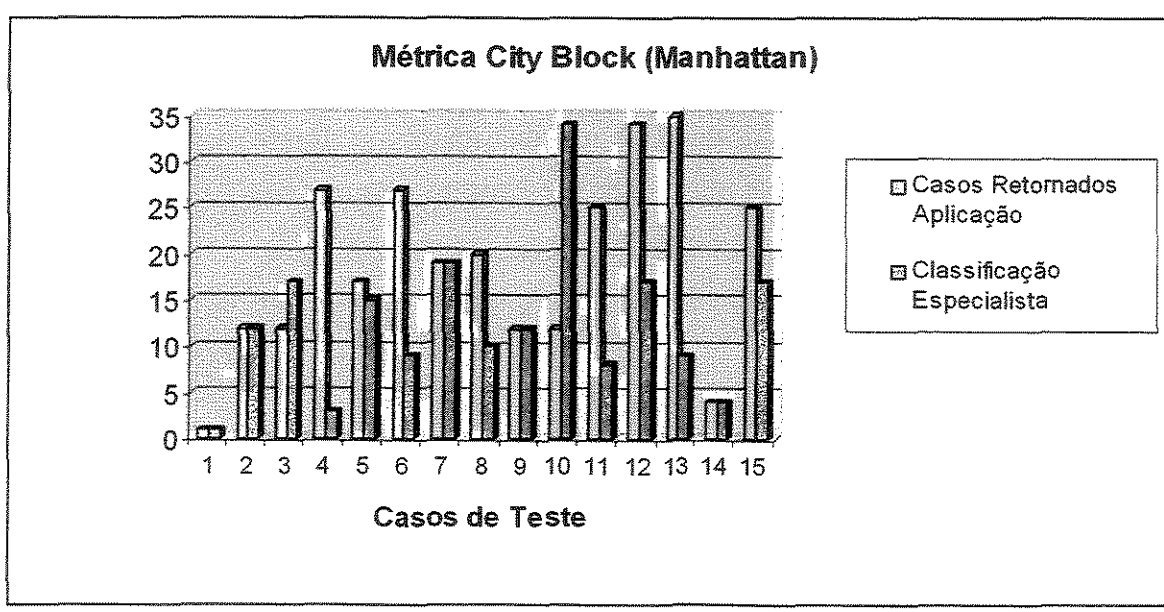


Figura 20 - Métrica City Block (Manhattan)

A tabela 22 mostra a Métrica Contagem de Características.

Tabela 22 – Métrica Contagem de Características

Caso de Teste	Casos Retornados pela Média	Valor Média %	Classificação do Especialista	Classificação do Caso Recuperado
26	21	50,00	1, 5, 21	3º
27	12-25-17	77,50-70,00-67,50	12, 17, 25	1º
28	27-12-17	75,00-72,50-72,50	17, 8, 27, 12	3º
29	3-4-12	57,50-57,50-57,50	3, 9, 20	1º
30	17-3-27	75,00-72,50-72,50	15, 17, 20	2º
31	12-27-23	60,00-55,00-50,00	9, 27, 13, 23	2º
32	19-2-11	65,00-57,50-52,50	19, 11, 22	1º
33	25-10-9	60,00-55,00-55,00	10, 9, 13	2º
34	12-27-3	90,00-72,50-70,00	12, 17, 3	1º
35	34-12-3	75,00-75,00-70,00	34, 3, 12, 17	1º
36	25-8-9	70,00-65,00-60,00	8, 9, 25	2º
37	3-25-34	75,00-75,00-75,00	17, 3, 4, 30, 34	2º
38	35-9-4	75,00-65,00-65,00	9, 10, 35	2º
39	4-25-37	70,00-65,00-65,00	4, 9, 17	1º
40	25-4-37	80,00-72,50-70,00	17, 4, 39	2º

Obs.: Não tem valor máximo para esta métrica

A figura 21 mostra os melhores resultados da Métrica Contagem de Características.

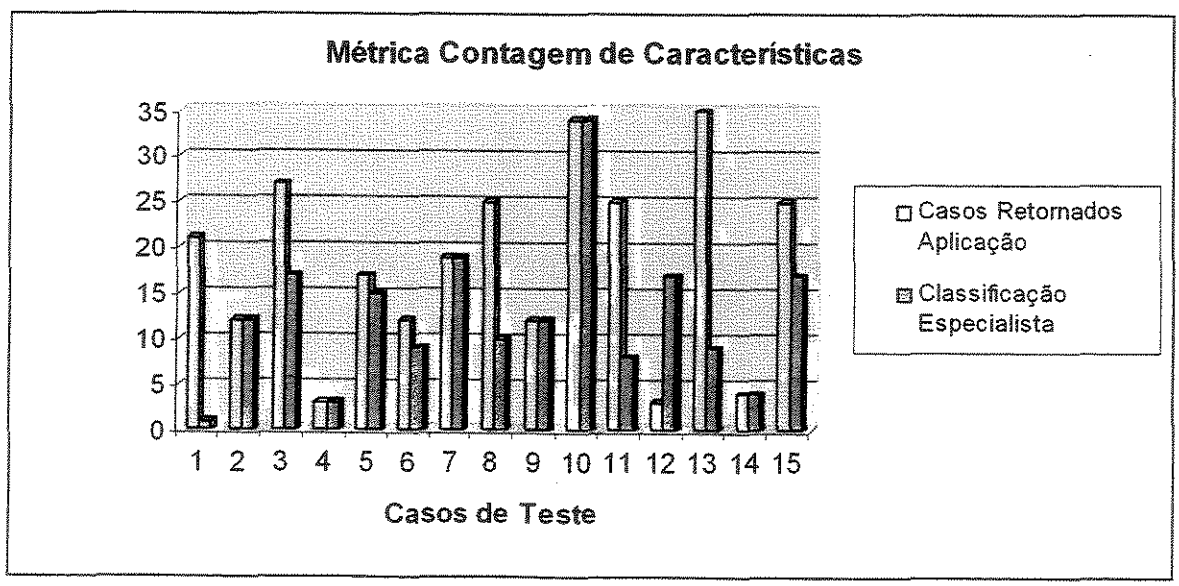


Figura 21 - Métrica Contagem de Características

A tabela 23 mostra a Métrica Euclidiana.

Tabela 23 – Métrica Euclidiana

Caso de Teste	Casos Retornados pela Média	Valor Média %	Casos Retornados pela Máxima	Valor Máxima %	Classificação do Especialista	Classificação do Caso Recuperado
26	1-21-2	79,40-76,94-76,63	1-21-2	79,46-76,94-76,63	1, 5, 21	1°
27	12-23-25	96,02-90,53-88,77	12-23-25	96,02-90,53-88,77	12, 17, 25	1°
28	12-7-27	93,49-92,47-92,20	12-7-27	93,53-92,71-92,25	17, 8, 27, 12	3°
29	27-21-25-10-3	86,46-86,38-86,22-85,62-85,30	27-21-25-10-3	86,46-86,38-86,22-85,62-85,30	3, 9, 20	5°
30	17-15-25	91,52-90,42-89,95	17-15-25	91,52-90,42-89,95	15, 17, 20	2°
31	13-27-9	82,40-82,32-80,83	13-27-9	82,40-82,32-80,83	9, 27, 13, 23	2°
32	19-27-20	91,26-80,52-79,81	19-27-20	91,47-80,53-79,88	19, 11, 22	1°
33	20-11-23-10	85,02-82,87-82,58-81,97	20-11-23-10	85,08-82,92-82,64-81,97	10, 9, 13	4°
34	12-27-17	94,31-90,32-89,68	12-27-17	95,12-91,16-90,53	12, 17, 3	1°
35	27-12-34	90,44-89,80-89,16	27-34-12	90,61-89,97-89,97	34, 3, 12, 17	3°-2°
36	25-8-4	89,76-89,68-88,05	25-8-4	89,76-89,68-88,05	8, 9, 25	2°
37	35-12-25-34	91,90-91,65-91,26-91,23	34-35-12	92,07-92,07-91,65	17, 3, 4, 30, 34	5°
38	35-9-10	93,83-91,09-86,52	35-9-10	93,99-91,10-86,52	9, 10, 35	2°
39	4-25-38	94,06-91,68-89,93	4-25-38	94,26-91,73-90,13	4, 9, 17	1°
40	25-4-37	94,26-93,09-91,55	25-4-37	94,37-93,22-91,68	17, 4, 39	2°

A figura 22 mostra os melhores resultados da Métrica Euclidiana.

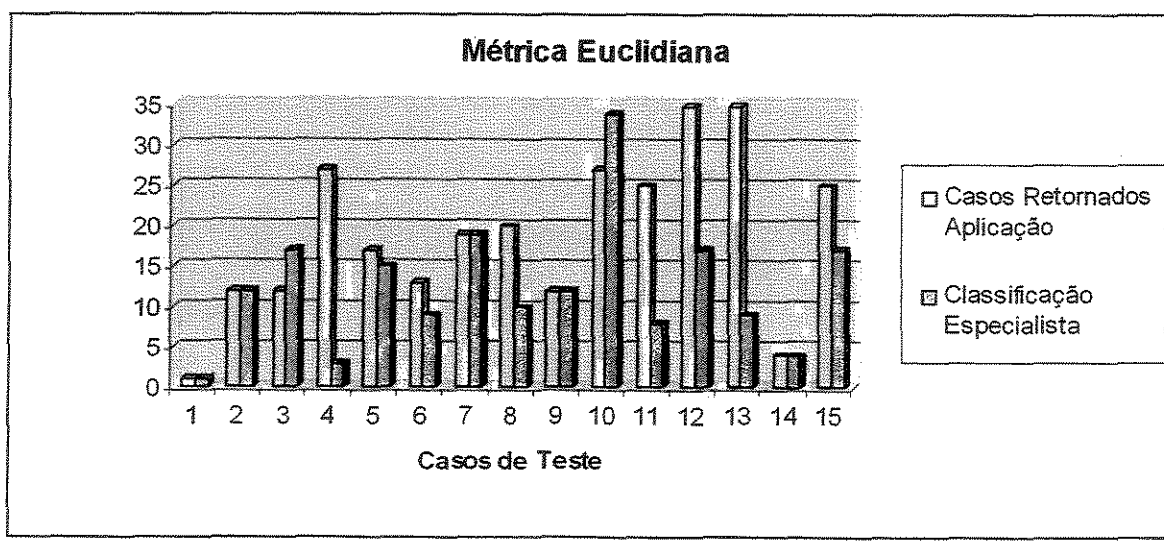


Figura 22 - Métrica Euclidiana

A tabela 24 mostra a Métrica Euclidiana Quadrada.

Tabela 24 – Métrica Euclidiana Quadrada

Caso de Teste	Casos Retornados pela Média	Valor Média %	Casos Retornados pela Máxima	Valor Máxima %	Classificação do Especialista	Classificação do Caso Recuperado
26	1-21-2	63,05-59,19-58,71	1-21-2	63,14-59,19-58,71	1, 5, 21	1º
27	12-23-25	92,19-81,95-78,81	12-23-25	92,19-81,95-78,81	12, 17, 25	1º
28	12-7-27	87,40-85,52-85,02	12-7-27	87,48-85,95-85,10	17, 8, 27, 12	3º
29	27-21-25-10-3	74,76-74,62-74,33-73,38-72,76	27-21-25-10-3	74,76-74,62-74,33-73,38-72,76	3, 9, 20	5º
30	17-15-25	83,76-81,76-80,90	17-15-25	83,76-81,76-80,90	15, 17, 20	2º
31	13-27-9	67,90-67,76-65,33	13-27-9	67,90-67,76-65,33	9, 27, 13, 23	2º
32	19-27-20	83,29-64,83-63,69	19-27-20	83,67-64,86-63,81	19, 11, 22	1º
33	20-11-23-10	72,29-68,67-68,19-67,19	20-11-23-10	72,38-68,76-68,29-67,19	10, 9, 13	4º
34	12-27-17	88,95-81,57-80,43	12-27-17	90,48-83,10-81,95	12, 17, 3	1º
35	27-12-34	81,79-80,64-79,50	27-34-12	82,10-80,95-80,95	34, 3, 12, 17	3º-2º
36	25-8-4	80,57-80,43-77,52	25-8-4	80,57-80,43-77,52	8, 9, 25	2º
37	35-12-25-34	84,45-84,00-83,29-83,24	35-34-12	84,76-84,76-84,00	17, 3, 4, 30, 34	5º
38	35-9-10	88,05-82,98-74,86	35-9-10	88,33-83,00-74,86	9, 10, 35	2º
39	4-25-38	88,48-84,05-80,88	4-25-38	88,86-84,14-81,24	4, 9, 17	1º
40	25-4-37	88,86-86,67-83,81	25-4-37	89,05-86,90-84,05	17, 4, 39	2º

A figura 23 mostra os melhores resultados da Métrica Euclidiana Quadrada.

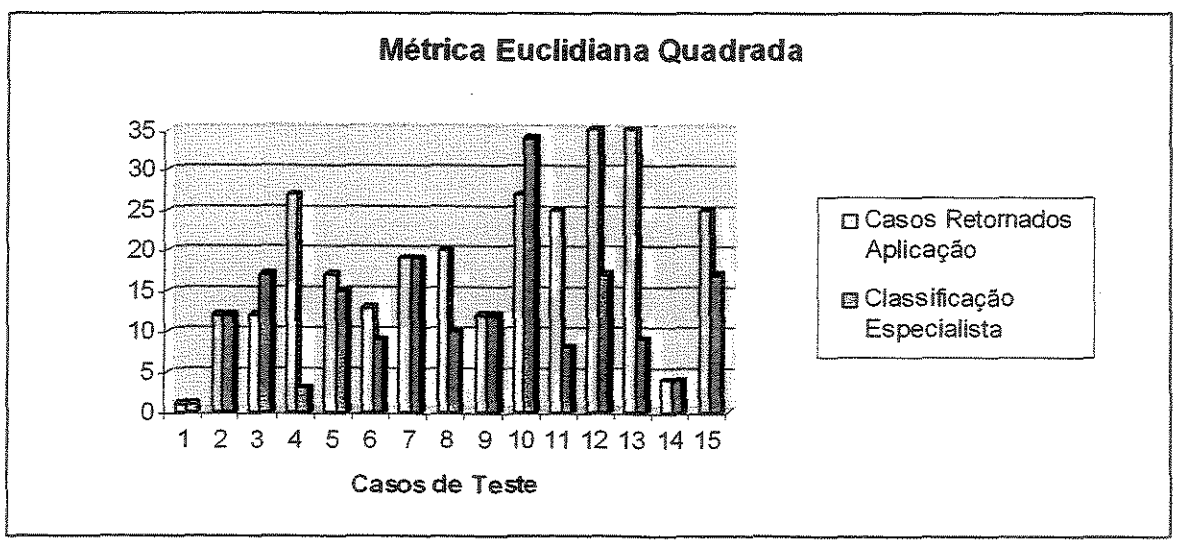


Figura 23 - Métrica Euclidiana Quadrada

A tabela 25 mostra a Métrica Euclidiana Quadrada Ponderada.

Tabela 25 – Métrica Euclidiana Quadrada Ponderada

Caso de Teste	Casos Retornados pela Média	Valor Média %	Casos Retornados pela Máxima	Valor Máxima %	Classificação do Especialista	Classificação do Caso Recuperado
26	21-1-2	74,84-70,79-70,13	21-1-2	74,84-70,92-70,13	1, 5, 21	2º
27	12-9-25	94,10-83,13-82,26	12-9-25	94,10-83,13-82,26	12, 17, 25	1º
28	27-12-23	92,98-89,14-80,51	27-12-23	93,14-89,30-80,67	17, 8, 27, 12	3º
29	3-27-26	76,47-76,39-76,11	3-27-26	76,47-76,39-76,11	3, 9, 20	1º
30	15-3-17	90,73-89,15-84,66	15-3-17	90,73-89,15-84,66	15, 17, 20	1º
31	28-27-12	78,30-78,11-77,49	28-27-12	78,46-78,11-77,49	9, 27, 13, 23	2º
32	19-22-11	91,69-68,94-61,30	19-22-11	92,08-68,99-61,35	19, 11, 22	1º
33	18-6-11	78,05-77,21-76,12	18-6-11	78,41-78,24-76,31	10, 9, 13	—
34	12-27-11	89,81-83,99-83,94	12-27-11	91,35-85,53-85,48	12, 17, 3	1º
35	34-12-27	94,47-88,89-84,80	34-12-11	96,15-89,42-85,76	34, 3, 12, 17	1º
36	8-25-9	89,04-86,46-84,11	8-25-9	89,04-86,46-84,11	8, 9, 25	1º
37	35-34-12	96,39-95,38-90,04	35-34-12	96,92-96,92-90,04	17, 3, 4, 30, 34	5º
38	35-37-34	91,01-89,17-88,52	35-37-34	91,30-89,22-88,57	9, 10, 35	3º
39	4-38-7	90,09-86,95-85,43	4-38-7	90,21-87,02-85,52	4, 9, 17	1º
40	9-38-25	93,34-87,29-86,00	9-38-25	93,52-87,63-86,16	17, 4, 39	—

A figura 24 mostra os melhores resultados da Métrica Euclidiana Quadrada Ponderada.

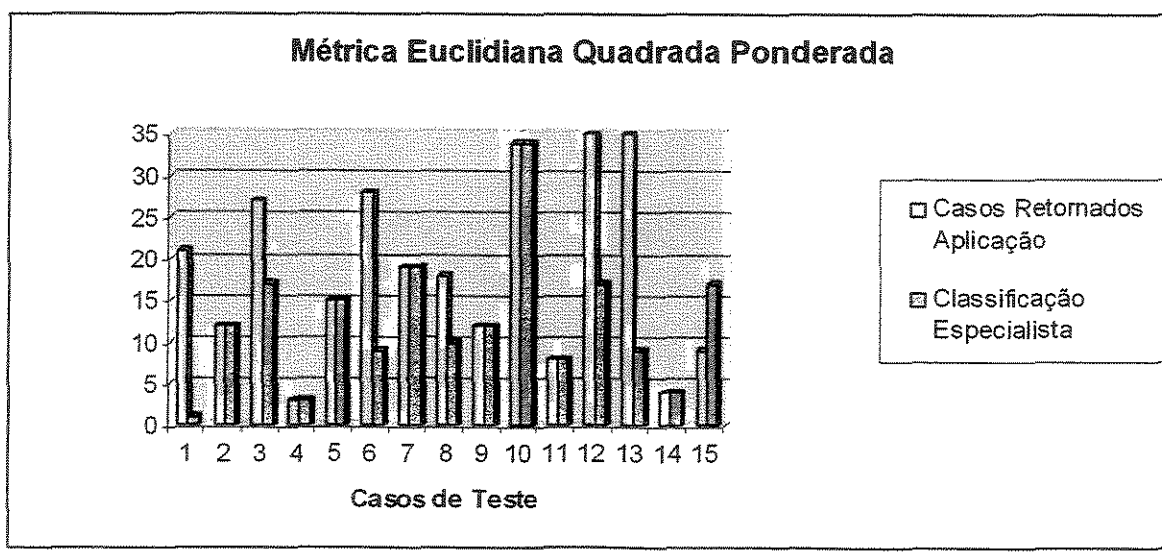


Figura 24 - Métrica Euclidiana Quadrada Ponderada

A tabela 26 mostra a Métrica Vizinho mais Próximo – Função Inclusão de Consulta.

Tabela 26 – Métrica Vizinho mais Próximo – Função Inclusão de Consulta

Caso de Teste	Casos Retornados pela Média	Valor Média %	Classificação do Especialista	Classificação do Caso Recuperado
26	21-1	58,65-53,73	1, 5, 21	2º
27	12-25-9	91,35-76,92-75,96	12, 17, 25	1º
28	27-12-17	84,94-81,09-67,63	17, 8, 27, 12	3º
29	1-26-3	66,35-61,54-60,58	3, 9, 20	–
30	3-17-20	81,73-76,92-75,00	15, 17, 20	2º
31	12-17-28	65,96-75,00-74,04	9, 27, 13, 23	–
32	19-22-11	75,96-58,65-53,85	19, 11, 22	1º
33	18-11-28	72,12-67,31-64,42	10, 9, 13	–
34	12-27-11	88,94-81,25-81,25	12, 17, 3	1º
35	34-12-11	86,54-79,81-75,96	34, 3, 12, 17	1º
36	8-25-9	78,85-77,88-75,96	8, 9, 25	1º
37	34-35-12	95,19-90,38-87,50	17, 3, 4, 30, 34	5º
38	37-34-35	78,85-78,85-76,92	9, 10, 35	–
39	38-4-37	77,88-74,04-74,04	4, 9, 17	2º.
40	9-38-25	75,00-72,60-72,12	17, 4, 39	–

Obs.: Não tem valor máximo para esta métrica

A figura 25 mostra os melhores resultados da Métrica Vizinho mais Próximo – Função Inclusão de Consulta.

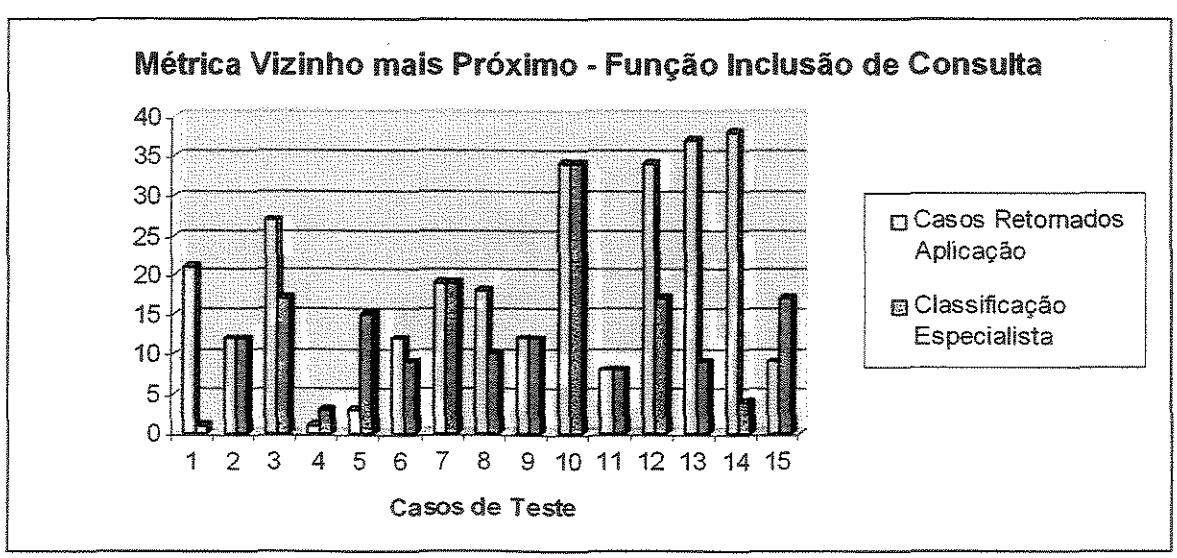


Figura 25 - Métrica Vizinho mais Próximo – Função Inclusão de Consulta

A tabela 27 mostra a Métrica Vizinho mais Próximo – Função Interseção.

Tabela 27 – Métrica Vizinho mais Próximo – Função Interseção

Caso de Teste	Casos Retornados pela Média	Valor Média %	Classificação do Especialista	Classificação do Caso Recuperado
26	21-1	58,65-53,37	1, 5, 21	2º
27	12-25-9	88,94-76,92-75,96	12, 17, 25	1º
28	27-12-17	84,94-78,69-67,63	17, 8, 27, 12	3º
29	1-26-3	62,98-61,54-60,58	3, 9, 20	3º
30	3-17-15	81,73-76,92-75,00	15, 17, 20	2º
31	12-27-28	75,96-75,00-67,63	9, 27, 13, 23	2º
32	19-22-11	75,96-58,65-53,85	19, 11, 22	1º
33	18-11-9	72,12-67,31-63,46	10, 9, 13	2º
34	12-27-11	88,14-81,25-81,25	12, 17, 3	1º
35	34-12-11	86,54-79,81-75,96	34, 3, 12, 17	1º
36	8-25-9	78,85-77,88-75,96	8, 9, 25	1º
37	34-35-12	92,79-85,58-85,10	17, 3, 4, 30, 34	5º
38	37-34-35	78,85-78,85-76,92	9, 10, 35	3º
39	38-4-37	77,88-74,04-74,04	4, 9, 17	2º
40	9-25-38	75,00-72,12-70,99	17, 4, 39	–

Obs.: Não tem valor máximo para esta métrica

A figura 26 mostra os melhores resultados da Métrica Vizinho mais Próximo – Função Interseção.

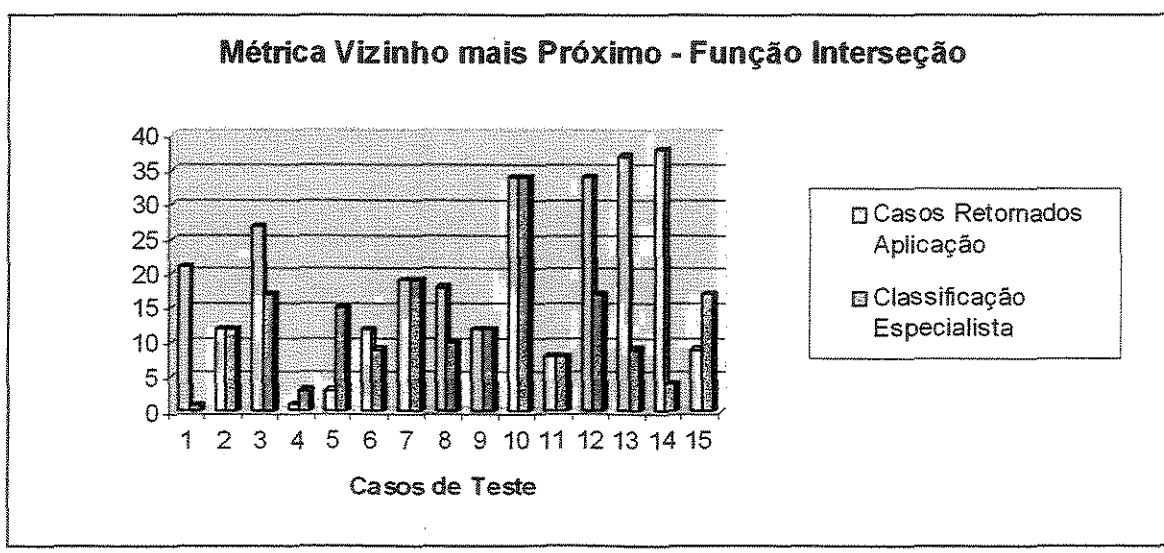


Figura 26 - Métrica Vizinho mais Próximo – Função Interseção

A tabela 28 mostra a Métrica Vizinho mais Próximo – Função Linear.

Tabela 28 – Métrica Vizinho mais Próximo – Função Linear

Caso de Teste	Casos Retornados pela Média	Valor Média %	Casos Retornados pela Máxima	Valor Máxima %	Classificação do Especialista	Classificação do Caso Recuperado
26	21-1-2	72,02-67,40-67,02	21-1-2	72,02-68,08-67,02	1, 5, 21	2º
27	12-9-24	93,65-81,73-80,77	12-9-24	93,65-81,73-80,77	12, 17, 25	1º
28	27-12-13	91,83-87,98-77,31	27-12-13	91,83-87,98-77,31	17, 8, 27, 12	3º
29	3-27-26	62,98-61,54-60,58	3-27-26	74,52-74,33-73,56	3, 9, 20	1º
30	3-15-17	87,31-86,73-82,79	3-15-17	87,31-86,73-82,79	15, 17, 20	2º
31	27-12-18	77,21-77,02-76,15	27-12-18	77,21-77,02-76,15	9, 27, 13, 23	2º
32	19-22-11	88,46-68,27-60,10	19-22-11	89,42-68,75-60,58	19, 11, 22	1º
33	18-6-11	77,16-74,18-74,04	18-6-11	77,40-75,87-75,00	10, 9, 13	—
34	12-27-11	89,42-83,37-83,27	12-27-11	91,35-85,29-85,19	12, 17, 3	1º
35	34-12-27	92,79-87,50-82,88	34-12-27	96,15-89,42-84,81	34, 3, 12, 17	1º
36	8-25-9	86,92-86,15-82,60	8-25-9	86,92-86,15-82,60	8, 9, 25	1º
37	35-34-12	94,62-94,62-89,42	35-34-12	96,54-96,54-89,42	17, 3, 4, 30, 34	5º
38	35-37-34	87,31-87,12-86,73	35-37-34	87,79-87,60-87,21	9, 10, 35	3º
39	4-38-37	87,45-84,47-84,38	4-38-37	88,08-85,00-84,62	4, 9, 17	1º
40	9-38-25	89,47-84,47-83,41	9-38-25	90,00-85,96-83,94	17, 4, 39	–

A figura 27 mostra os melhores resultados da Métrica Vizinho mais Próximo – Função Linear.

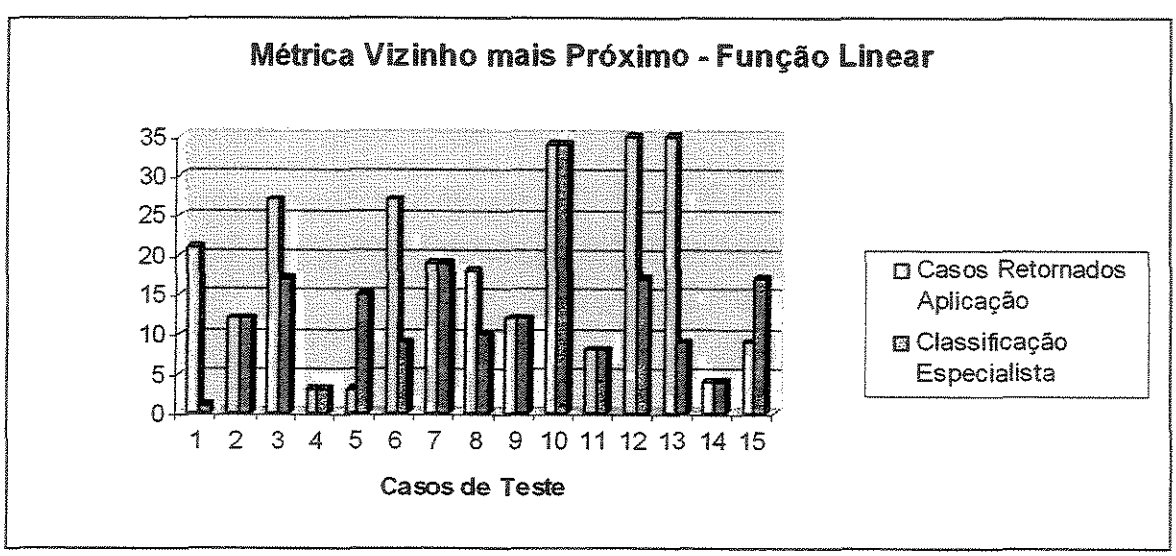


Figura 27 - Métrica Vizinho mais Próximo – Função Linear

4.8.1 Análise dos Resultados Obtidos

As tabelas 29 a 36 mostram os resultados dos testes, considerando o uso do ajuste 1, vista no item 4.4.2, obtendo-se em percentual os resultados médio e máximo de cada métrica utilizada e o seu referido caso. A partir desses resultados, foi possível observar que o resultado médio obteve uma melhor similaridade que o resultado máximo. Essas tabelas mostram também a classificação do especialista em relação aos casos mais similares encontrados pela aplicação.

No anexo 03, podem ser observados os casos que foram retornados pela aplicação (1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 34, 35, 37, 38, 39) e os casos que não foram (5, 14, 29, 30, 31, 32, 33, 36, 40).

Os casos de maior similaridade indicados pelo especialista foram 1, 3, 4, 5, 8, 9, 10, 11, 12, 13, 15, 17, 19, 20, 21, 22, 23, 25, 27, 30, 34, 35, e 39 e os casos não indicados pelo especialista foram 2, 6, 7, 14, 16, 18, 24, 26, 28, 29, 31, 32, 33, 36, 37, 38 e 40.

A escolha do melhor caso fica para a tomada de decisão do especialista, depois de observar os valores retornados pela aplicação. Dessa forma se houver alguma necessidade de adaptação para a indicação do tratamento aos casos de epicondilite, o especialista deverá fazê-la.

A partir dos resultados obtidos pode-se observar que algumas métricas retornaram os casos mais similares: A métrica Euclidiana Quadrada Ponderada retornou 8 casos mais similares dos 15 testes efetuados, seguida do Vizinho mais Próximo – Função Linear que retornou 7 casos.

Também foi possível observar a proximidade dos resultados obtidos pela aplicação com os resultados atribuídos pelo especialista, dessa forma obteve-se para cada métrica os seguintes desempenhos:

- City-Block (Manhattan)

05 casos coincidiram com a 1ª classificação

04 casos coincidiram com a 2ª classificação

03 casos coincidiram com a 3ª classificação

03 casos não coincidiram

- Contagem de Características

06 casos coincidiram com a 1ª classificação

07 casos coincidiram com a 2ª classificação

02 casos coincidiram com a 3ª classificação

- Euclidiana

05 casos coincidiram com a 1ª classificação

06 casos coincidiram com a 2ª classificação

01 caso coincidiu com a 3ª classificação

03 casos não coincidiram

- Euclidiana Quadrada

05 casos coincidiram com a 1ª classificação

06 casos coincidiram com a 2ª classificação

01 caso coincidiu com a 3ª classificação

03 casos não coincidiram

- Euclidiana Quadrada Ponderada

08 casos coincidiram com a 1ª classificação

02 casos coincidiram com a 2ª classificação

02 casos coincidiram com a 3ª classificação

03 casos não coincidiram

- Vizinho mais Próximo – Função Inclusão de Consulta

05 casos coincidiram com a 1ª classificação

03 casos coincidiram com a 2ª classificação

01 casos coincidiram com a 3ª classificação

06 casos não coincidiram

- Vizinho mais Próximo – Função Interseção

05 casos coincidiram com a 1ª classificação

05 casos coincidiram com a 2ª classificação

03 casos coincidiram com a 3ª classificação

02 casos não coincidiram

- Vizinho mais Próximo – Função Linear

07 casos coincidiram com a 1ª classificação

03 casos coincidiram com a 2ª classificação

02 casos coincidiram com a 3ª classificação

03 casos não coincidiram

Os resultados obtidos pelas melhores métricas que a aplicação selecionou foram:

- Euclidiana Quadrada Ponderada retornou 8 casos e
- Vizinho mais Próximo – Função Linear retornou 7 casos mais similares dos 15 testes efetuados.

Esses resultados relacionados com as classificações do especialista da área:

- Euclidiana Quadrada Ponderada 8 casos e
- Vizinho mais Próximo – Função Linear 7 casos coincidentes com a 1ª. Classificação

É possível observar que os melhores resultados foram obtidos pelas métricas Euclidiana Quadrada Ponderada e Vizinho mais Próximo – Função Linear para os casos de Epicondilite Lateral analisados.

O entendimento desses resultados obtidos pode ser observado a partir das características das métricas. A métrica Euclidiana representa a menor distância

entre dois pontos, portanto uma reta. A métrica Euclidiana Quadrada é utilizada com freqüência na análise classificatória nos casos em que a variáveis são quantitativas. Já a métrica Euclidiana Quadrada Ponderada é dada pela soma dos quadrados das diferenças dos valores de todas as variáveis, considerando-se seus pesos. A métrica do Vizinho mais Próximo é a mais simples, pois não requer muitos cálculos para o seu entendimento. Primeiramente identifica-se os atributos essenciais para a solução do problema, sendo que estes podem ser representados através de um sistema de coordenadas, possibilitando a medição da distância entre caso novo e os casos da base **[Lee, 1998]**.

O desempenho atingido pelas métricas Contagem de Características e City-Block (Manhattan) também podem ser entendido a partir de suas características. A métrica Contagem de Características é calculada por meio da contagem dos atributos coincidentes (dos casos da base com o novo caso). Dessa forma encontra-se o percentual sobre o número dos atributos iguais entre os dois casos, considerando-se os atributos usados na comparação. Permite ainda a indicação de um limiar mínimo de similaridade. Já a métrica City-Block, conhecida como distância do quarteirão, é uma medida de distância neutra, pois pondera todas as diferenças de forma idêntica. Por exemplo em um modelo urbano com ruas perpendiculares, corresponde ao menor caminho **[Wangenheim, 2003]**.

Uma justificativa para os resultados obtidos por essas métricas diante dos atributos do FisioSmart pode ser o fato que essas duas métricas realizam a ponderação pelo peso dos atributos e utilizam os valores dos descritores, enquanto as demais métricas, não combinam esses dois fatores. Embora a métrica do Vizinho mais Próximo em suas funções Interseção e Inclusão de Consulta, também realizem a ponderação dos pesos, não utilizam os valores dos descritores, por se tratar de conjuntos, onde são utilizados o número de elementos desses conjuntos.

Na próxima seção serão apresentadas as Considerações Finais deste trabalho.

Capítulo 5 Considerações Finais

A Inteligência Artificial vem propiciando avanços na área de informática, estudando maneiras para que a máquina possa realizar tarefas inteligentes próprias do ser humano. Uma das técnicas é Raciocínio Baseado em Casos, que visa encontrar uma solução para o problema atual comparando-o com casos anteriormente armazenados.

O Raciocínio Baseado em Casos (RBC) constitui um paradigma de aprendizado de máquina em IA, pois é capaz de simular o raciocínio de um especialista, utilizando a experiência de casos passados na análise e resolução de uma situação atual.

A principal fonte de conhecimento é a memória de casos e o raciocínio é realizado através da recuperação de casos na base de conhecimento.

As tendências de desenvolvimento dos métodos RBC permitem o seu agrupamento em torno de quatro tópicos principais: integração com outros métodos de aprendizado, integração com outros componentes de raciocínio, incorporação com processamento paralelo e métodos avançados focando novos aspectos cognitivos.

Este trabalho apresentou um estudo sobre métricas de similaridade em Sistemas Baseados em Casos, aplicados à área da saúde, mais especificamente sobre epicondilite lateral, uma tendinite do cotovelo. Para esse área de aplicação e as atributos definidos para o FisioSmart as medidas de distância utilizadas foram as que mais se ajustaram ao problema.

Para apresentar esse estudo sobre métricas de similaridade em SBC foi necessário realizar uma revisão bibliográfica sobre as métricas que possibilitam as medidas de distância.

Algumas métricas de similaridade, em particular relativas às medidas de distância foram utilizadas nesse trabalho, sendo elas:

- Distância através do método do vizinho mais próximo;
- Distância Euclidiana;

BIBLIOTECA CENTRAL

- Distância Euclidiana Quadrada;
- Distância Euclidiana Quadrada Ponderada;
- Distância City_Block;
- Distância ChebyChev;
- Distância Percentual;
- Distância para Pesos diferentes;
- Contagem de Características e;
- Booleana.

A partir das métricas utilizadas nessa aplicação pode-se obter uma análise dos resultados obtidos, sendo que, em relação aos casos estudados, o maior sucesso obtido foi por meio das métricas Euclidiana Quadrada Ponderada e o Vizinho mais Próximo – Função Linear. Para uma melhor adequação das métricas aos atributos utilizados fez-se necessário alguns ajustes na aplicação da métrica a fim de se solucionar as diferenças obtidas se aplicadas as métricas em sua forma original.

5.1 Contribuições

As principais contribuições deste trabalho foram:

1. Um estudo comparativo entre algumas métricas de similaridade.

2. Proposta de técnicas de manipulação de atributos multivalorados, para as métricas de similaridade que não os manipulam diretamente.

3. O desenvolvimento de uma aplicação na área da saúde com RBC, propondo tratamentos para casos de epicondilite lateral.

5.2 Trabalhos Futuros

Alguns tópicos a serem abordados em trabalhos futuros:

1. A implementação de outros algoritmos de recuperação de casos, de casamento inicial, de aprendizado e retenção. Esses novos algoritmos poderiam ser estruturados de maneira genérica, de modo que fiquem armazenados em uma biblioteca, podendo ser utilizado para qualquer domínio de aplicação, servindo como uma ferramenta *shell* para desenvolvimento de aplicações baseadas em casos.
2. O desenvolvimento de sistema híbridos trabalhando com Raciocínio Baseado em Casos e Algoritmos Genéticos, Raciocínio Baseado em Casos e Redes Neurais, de tal forma a combinar o uso de várias técnicas.
3. A comparação com outras métricas de similaridades.
4. Implementação de outros algoritmos de recuperação de casos que melhorem o desempenho do sistema.

Referências Bibliográficas

[Aamodt, 1994] AAMODT, A., E. Plaza. *Case-Based Reasoning: Foundational Issues, methodological Variations, and System Approaches*, AICOM, 7(1), March 1994.

[Abel, 1996] ABEL, M. *Raciocínio Baseado em Casos.* Porto Alegre, Instituto de Informática da UFRGS, 120p,1996.

[Abel, 1997] ABEL, M. *Engenharia de conhecimento.* Porto Alegre, Instituto de Informática a UFRGS,130p, junho, 1997.

[Althoff, 1995] ALTHOFF, Klaus. *Dieter Evaluating Case-Based Reasoning Systems. Proc. Unicom Workshop on CBR*, London, 270p, 1995.

[Barreto, 1997] BARRETO, Jorge Muniz. *Inteligência Artificial no limiar do século XXI.* Florianópolis: J.M. Barreto, 150p, 1997.

[Camargo, 1999] CAMARGO, Kátia Gavranich. *Inteligência Artificial Aplicada à Nutrição a Prescrição de Planos Alimentares.* Dissertação (Mestrado em Engenharia de Produção) – Centro Tecnológico, Universidade Federal de Santa Catarina. Florianópolis, 1999.

[Chen, 1995] CHEN, Peter. *Modelagem De Dados - A Abordagem Entidade.* São Paulo. Makron Books, 256p, 1995.

[Darpa, 1989] DEFENSE ADVANCED REASEARCH PROJECTS AGENCY (DARPA), *Case-Based Reasoning, Proceedings of a Workshop on Case-Based Reasoning*, Flórida, Jun-1989.

[Dillon, 1982] DILLON, M. e Gray, S. *FASIT: A Fully Automatic Syntactically Based Indexing System*. Journal of the American Society of Information Science, 34. 1982.

[Fagan, 1987] FAGAN, J. *Experiments in Automatic Phrase Indexing for Document Retrieval: A Comparison of Sintactic e Non-Syntactic Methods*. Ph.D. Thesis, Technical Report 87-868, Cornell University, Computer Science Department. 1987.

[Freitas, 2004] FREITAS, R. e NAKAMITI, G. S. e JÚLIO, M.M., *Um Estudo sobre Métricas de Similaridade em Sistema de RBC*. Network Tecnologies, ano 3n°. 3. ISN 1677-778. , 27-36 p. 2004.

[Gelbart, 1991] GELBART, D. e JC Smith. *Beyond Boolean Search: FLEXICON, a legal text-based intelligent system*. Third International Conference on Artificial Intelligence e Law, ACM, Oxford, UK, 225-234. 1991.

[Gey, 1989] GEY, F. e Chan, W. *Comparing Vector Space Retrieval with the RUBRIC Expert System*. SIGIR FORUM, 23, 1. 1989.

[Gresse, 1999] GRESSE, C. *REMEX - A case based approach for reuse of software measurement experienceware*. In Proceedings of 3 rd Int. Conference on Case-Based Reasoning, Germany, 1999.

[Gupta, 1987] GUPTA, P. *Boolean Interpretation of Conjunctions for Document Retrieval*. Journal of the American Society for Information Science, 38. 1987.

[Hammond, 1989a] HAMMOND, C., *Adaptation of Cases, Proceedings of Workshop on Cased-Based Reasoning*, 88-89p, Flórida, Jun-1989.

[Hammond, 1989b] HAMMOND, C., *Planejamento Baseado em Casos.* Academic Press, 257p, 1989.

[Hinrichs, 1992] HINRICHS, T., *Problem solving in open worlds.* Lawrence Eribaum Associates, 1992.

[Júlio, 2004] JÚLIO, M.M. e FRUTUOSO, L. *Uma definição sobre RBC.* Revista de Informática, Unifian, 34-54p, 2004.

[Kambhampati, 1989] KAMBHAMPATI, S., *Representational Requirements for Plan Reuse*, Proceedings of a Workshop on Case-Based Reasoning, 20-23p, Flórida, Jun-1989.

[Kolodner, 1983a] KOLODNER, J. *Reconstructive memory, a computer model. Cognitive Science*, Vol.7. s.281-328, 1983.

[Kolodner, 1983b] KOLODNER, J. *Maintaining organization in a dinamic long-term memory.* Cognitive ScieNCe, Vol.7, s.243-280, 1983.

[Kolodner, 1988] KOLODNER, J & Simpson, R., *The Mediator: A Case Study of a Case-Based Problem Solver*, School of Information and Computer Science, Georgia Institute of Technology, Atlanta, 1-53p, Feb-1988.

[Lagemann, 1997] LAGEMANN, Gerson Volney. *RBC para o Problema de Suporte ao Cliente nas Empresas de Prestação de Serviço de Software: O Caso Datasul.* Dissertação (Mestrado em Engenharia de Produção). Centro Tecnológico, universidade federal de Santa Catarina. Florianópolis, 1997.

[Leake, 1996] LEAKE, D. B., *Case-Based Reasoning Experiences, Lessons & Future Directions*, AAAI Press/The MIT Press, Menlo Park, CA 94025, 1996.

[Leake, 2002] LEAKE, David. *Experience, Introspection, and Expertise: Learning to Refine the Case-Based Reasoning Process*. [online]. Disponível na Internet via www. URL:http://citeseer.nj.nec.com/1595.html. Consultado em janeiro de 2002.

[Lee, 1998] LEE, Rosina Weber. *Pesquisa Jurisprudencial Inteligente*. Tese (Doutorado em Engenharia de Produção) – Centro Tecnológico, Universidade Federal de Santa Catarina. Florianópolis, 1998.

[Lenz, 1997] LENZ, M., Burkhard, H. D.. *CBR for Document Retrieval: The FallQ Project*. Lecture Notes in Artificial Intelligence: 2nd Int. Conference on CBR, ICCBR97. David Leake, Enric Plaza (ed.) Berlin: Springer Verlag, 1997.

[Lenz, 1998] LENZ , M., Hübner A., Kunze M. *Textual CBR*. M. Lenz. B. Boutsh-Sporl, H.D. Burkhard, S. Wess (eds.). Case Based Reasoning Technology. Springer Verlag, 1998.

[Maher, 1995a] MAHER, Mary Lou, M. Bala Balachandran and Dong Mei Zhang. *Case Based Reasoning in Design*. New Jersey: Lawrence Erlbaum Associates, 1995.

[Maher, 1995b] MAHER, Mary Lou, M. Bala Balachandran and Dong Mei Zhang. *Case-Based Reasoning in Design*. New Jersey. Lawrence Erlbaum Associates, 1995.

[Mantaras, 1995] MANTARAS, R. & Plaza E., *Case-Based Reasoning, State of the Art in Machine Learning*, 29-37p, Setembro-1995.

[Martins, 2000] MARTINS, Agenor S. *Computação Baseado em Casos: Contribuições Metodológicas aos Modelos de Indexação, de avaliação, de*

Ranking e de Similaridade de casos. Tese (Doutorado) – COPELE, Universidade Federal da Paraíba, Campina Grande, 2000.

[Mello, 2002] MELLO, Moacyr C. F°. *Aplicando as Sete Métricas de Controle de Projeto*,. Rational Software Latin América. White Paper. 260p, São Paulo. 2002.

[Osborne, 1997a] OSBORNE, Hugh & Bridge, Derek G.: *Similarity Metrics: A Formal Unification of Cardinal and Non-Cardinal Similarity Measures*, pp.235-244 . In David B. Leake, Enric Plaza (Eds.): Case-Based Reasoning Research and Development, Second International Conference, ICCBR-97, Proceedings. Lecture Notes in Computer Science 1266 Springer 1997, ISBN 3-540-63233-6. Providence, Rhode Island, USA, July 25-27, 1997.

[Osborne, 1997b] OSBORNE, Hugh & Bridge, Derek G.: *We're all going on a summer holiday: An Exercise in Non-Cardinal Case Base Retrieval*, pp.211-220. In Gösta Grahne (Ed.): SCAI'97 - Frontiers in Artificial Intelligence and Applications, Vol. 40, IOS Press, Amsterdam, The Netherlands, ISBN 90-5199-354-4. Proceedings of the Sixth Scandinavian Conference on Artificial Intelligence, Helsinki, Finland, 18-20 August, 1997.

[Piva, 2004a] PIVA, D., FREITAS, R.L., MISKULIN, M.S., NAKAMITI, G.S. and TOBAR, C.M.. Auxiliar: *A Case-Based System to Assist Online Courses*, Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04), IEEE, 2004.

[Piva, 2004b] PIVA, D., MISKULIN, M.S., FREITAS, R.L. and NAKAMITI, G.S.. Auxiliar Reasoner: *The core of a case-based system to aid engineering online courses*, World Congress on Engineering and Technology Education, 356-360p, São Paulo, 2004.

[Ramalho, 1999] RAMALHO, Benício José Almeida, *Sapiens: Sistema de apoio ao ensino de graduação baseado na Internet*. Dissertação de Mestrado em Engenharia de Produção. Universidade Federal de Santa Catarina, Florianópolis, 1999.

[Rich, 1988] RICH, E., *Inteligência Artificial*, McGraw-Hill, 280p, São Paulo, 1988.

[Rissland, 1983] RISSLAND, E., *Examples in legal reasoning: Legal hypotheticals*. In: Proceedings of the Eighth International Joint ConfereNCe on Artificial IntelligeNCe, IJCAI, Karlsruhe, 1983.

[Robertson, 1977] ROBERTSON, S.E. *The Probability Ranking Principle in IR*. Journal of Documentation, 33, 294-304.1977.

[Salton, 1984] SALTON, Gerard. *The Use of Extended Boolean Logic in Information Retrieval*. SIGMOD Conference, 277-285. 1984.

[Sbrocco, 2003] SBROCCO, J.H.T.C., FREITAS, R.L. and NAKAMITI, G.S., *Detecção de Falhas em Servidores de Dados usando a Tecnologia de Agentes Inteligentes*. VI Simpósio Brasileiro de Automação Inteligente, Bauru-SP, 2003.

[Schank, 1977] SCHANK, R., Abelson, R. *Scripts, Plans, Goals and Understanding*. Hillsdale NJ, Erlbaum , 1977.

[Schank, 1982] SCHANK, R., *Dynamic memory. A Theory of Reminding and Learning in Computers and People*. Cambridge University Press, 1982.

[Skalak, 1992] SKALAK, C.B. and RISSLAND, E (1992) *Arguments and cases: An inevitable twining. Artificial IntelligeNCe and Law*. An International Journal, 1(1), pp. 3-48. 1992.

[Smith, 1978] SMITH, E.E., *Adams N. Fact Retrieval And The Pardox of Interference*. Cognitive Psychology 10, 1978.

[Sycara, 1987] SYCARA, K. Using *Raciocínio Baseado em Casos for plan adaptation and repair*. Proceedings Raciocínio Baseado em Casos Workshop, DARPA. Clearwater beach, Florida, Morgan Kaufman, pp. 425-434, 1987

[Turner, 1989a] TURNER, R., *Case-Based and Schema-Based Reasoning of Probleming Solving*, Proceedings of a Workshop on Case-Based Reasoning, 341-344p, Flórida, Jun-1989.

[Turner, 1989b] TURBAN, Efrain and Jay E. Aronson. *Ecision Suport Systems and Intelligent Systems*, 5th ed. Prentice-Hall, Inc., 380p, New Jersey, 1998.

[Wangenheim, 2003] WANGENHEIM, C.G.V., Wangenheim, A.V.. *Raciocínio Baseado em Casos*. Ed. Manole Ltda. 294p, Barueri, SP, 2003.

[Watson, 1997] WATSON, I., *Applying Case-Based Reasoning: Techniques for Enterprise Systems*, Morgan Kaufmann Pub. Inc., San Francisco, CA, 1997.

[Yourdon, 1992] YOURDON, Edward; e ALENCAR, Dalton Conde de; *Análise estruturada moderna*. Rio de Janeiro: Campus, 1992.

Anexo A – Manual de utilização do FisioSmart 1.0

O FisioSmart – é um Sistema para Avaliação de Diagnósticos de Casos de Epicondilite Lateral, a tendinite do cotovelo. Essa aplicação utiliza RBC (Raciocínio Baseado em Casos) e algumas métricas de similaridade.

O FisioSmart foi desenvolvido e implementado utilizando a ferramenta de programação Delphi e Banco de Dados FireBird.

Quando um fisioterapeuta recebe um paciente, já com o diagnóstico de Epicondilite Lateral, ele inicia uma entrevista onde o paciente responde a uma série de perguntas que poderão auxiliar o fisioterapeuta na escolha da melhor conduta e procedimentos em fisioterapia para realizar o tratamento do paciente.

O objetivo do FisioSmart é buscar casos semelhantes ao novo caso fazendo consultas a Base de Casos armazenada, propondo condutas e procedimentos em fisioterapia para realizar o tratamento do paciente. Assim como faz o fisioterapeuta ao receber um novo paciente.

O FisioSmart possibilita que o especialista possa escolher uma métrica de similaridade para o retorno dos casos mais similares, ficando a critério do especialista qual métrica obteve um melhor resultado. As métricas de similaridade disponíveis no FisioSmart são:

- Vizinho mais Próximo (Nearest Neighbor)
  - Linear;
  - Inserção;
  - Inclusão de Consulta.

- Euclidiana;
- Euclidiana Quadrada;
- Euclidiana Quadrada Ponderada;
- City Block;
- Método da Contagem de Características;

Apresentação do sistema

A figura 08 apresenta a tela de entrada do sistema.

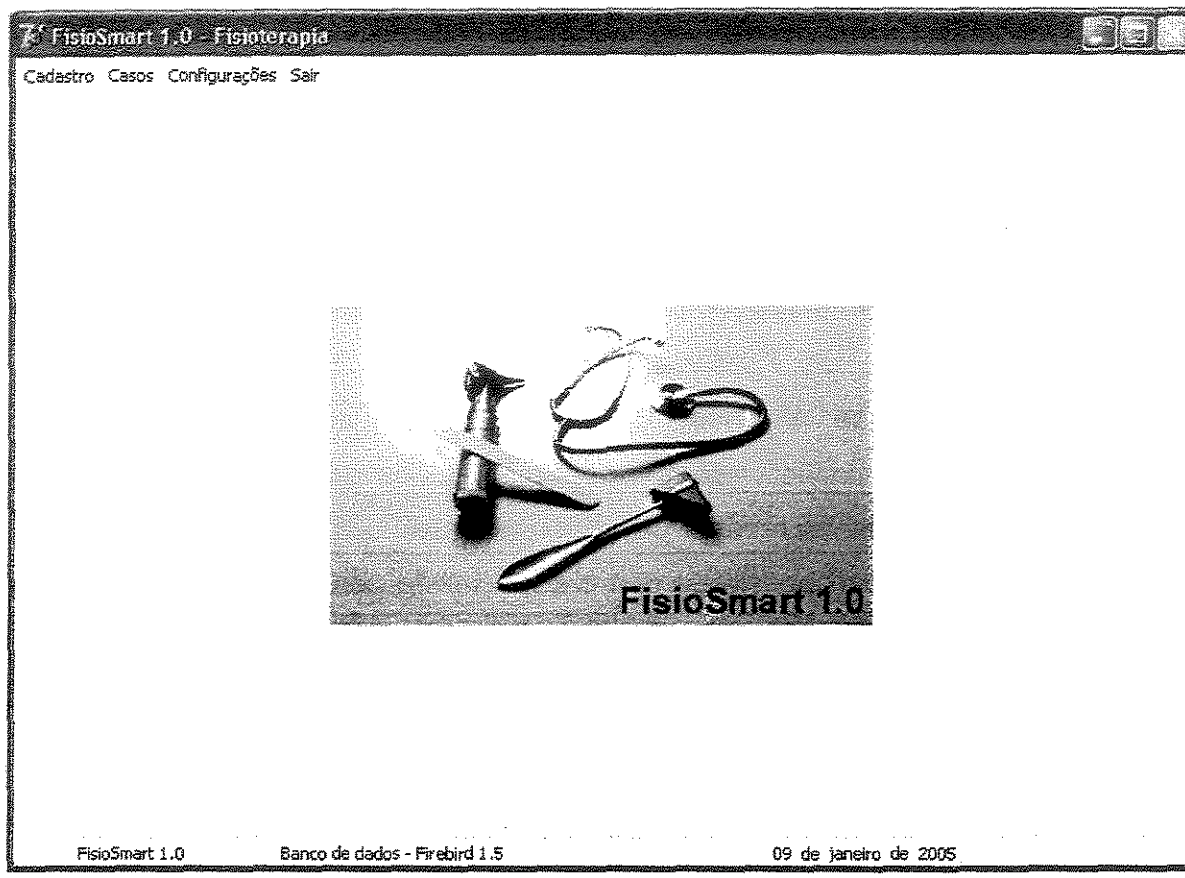


Figura 08 - Tela de Entrada do FisioSmart

A partir da tela representada pela Figura 08 é possível acessar todos os módulos do sistema. No menu principal apresentam-se as seguintes opções:

Cadastro:

- **Pacientes** → Módulo responsável pela manutenção do cadastro de paciente.

- **Descritores** → Módulo responsável pelos valores dos descritores dos atributos.
- **Medicamentos** → Módulo responsável pela manutenção do cadastro de medicamentos.

Casos:

Cadastrar casos → Módulo responsável pela manutenção da base de casos do sistema e pela análise de um novo caso.

Configurações: Módulo responsável pela atribuição de pesos para os atributos do sistema e também para definir um limiar e uma porcentagem de similaridade mínima de retorno.

A seguir serão apresentados detalhes sobre os módulos acima mencionados.

Cadastro de Pacientes

A figura 09 representa a tela Geral de Cadastro de Pacientes.

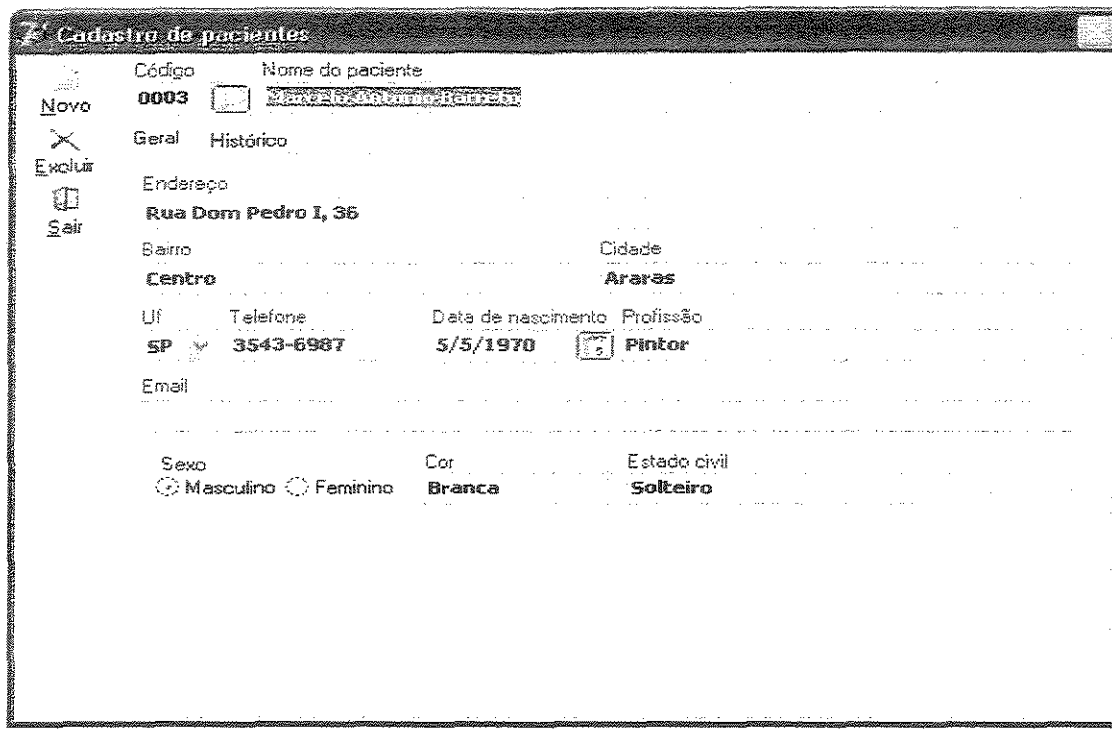


Figura 09 – Tela de Cadastro de Pacientes

O módulo de Cadastro de Pacientes é responsável pela manutenção do registro do paciente. As informações contidas neste módulo são:

- Código do paciente: O sistema gera um código seqüencial automaticamente. Sendo este código um número que identifica o paciente entre os demais.
- Nome do paciente: O usuário deve informar o nome do paciente.
- Endereço: O usuário deve informar neste campo o endereço onde o paciente reside.
- Bairro: O usuário deve informar neste campo o bairro em que o paciente reside.
- Cidade: O usuário deve informar neste campo a cidade em que o paciente reside.
- UF: O usuário deve informar neste campo o estado em que o paciente reside.

Cadastro de descritores

A figura 11 representa a tela de Cadastro de Descritores

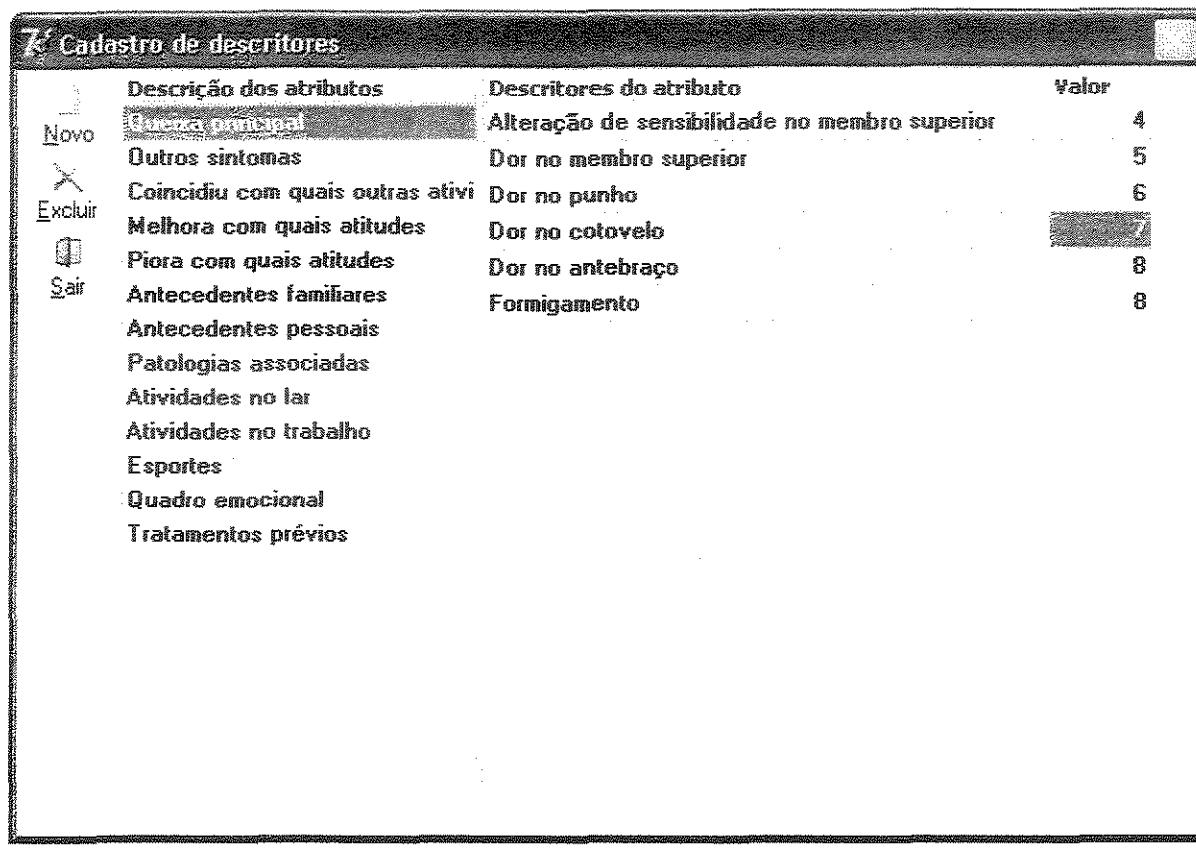


Figura 11 – Tela de Cadastro de Descritores

O módulo Cadastro de Descritores é responsável pela valorização dos descritores dos atributos. O especialista deve cadastrar um descritor e valorizar o mesmo em relação aos já cadastros na base, os valores estão definidos de 1 a 10, estes valores são usados para definir o grau de distância entre um descritor e outro, para que os mesmos possam ser implementados nas métricas de similaridade. Este módulo exibe os atributos definidos no sistema. Para visualizar os seus descritores, basta selecionar o atributo desejado.

As operações possíveis neste módulo são:

Inserir um novo descritor: Para inserir um novo descritor basta clicar no botão "Novo" na barra de ferramentas à esquerda, preencher os campos descritor e o campo valor e clicar no botão "Confirmar" para efetivar a inclusão.

Alterar um paciente: O especialista pode alterar os dados diretamente em seus respectivos campos e clicar no botão "Confirmar" para efetivar a alteração.

Excluir um paciente: Para excluir, é necessário clicar no botão "Excluir" na barra de ferramentas à esquerda, uma mensagem de confirmação ira aparecer para confirmar a exclusão basta clicar em "Ok". Obs.: um descritor só poderá ser excluído se não houver casos relacionados ao mesmo.

Cadastro de Medicamentos

A figura 12 representa a tela de Cadastro de Medicamentos.

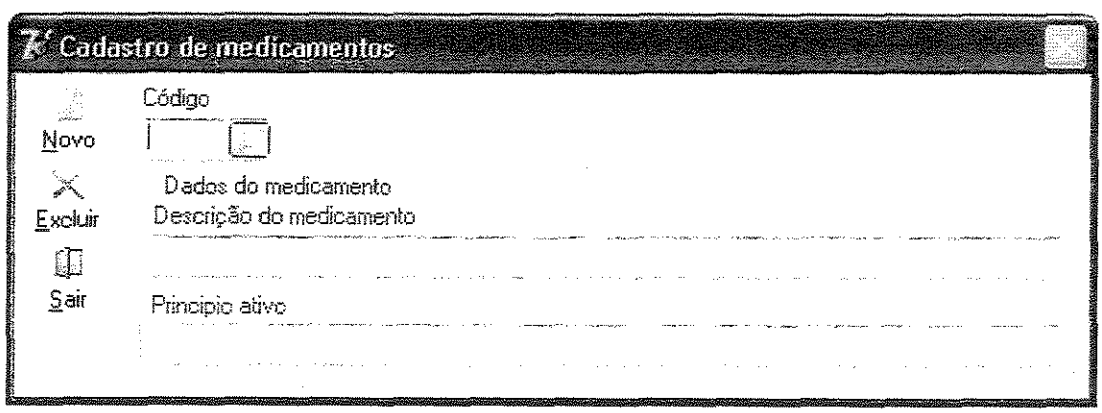


Figura 12 – Tela Cadastro de Medicamentos

Este módulo é responsável pela manutenção dos registros dos medicamentos que são relacionados ao(s) paciente(s). As informações contidas neste módulo são:

- Código: O sistema gera um código seqüencial automaticamente. Sendo este código um número que identifica o medicamento entre os demais.
- Descrição do medicamento: O usuário deve informar a descrição do medicamento.
- Principio ativo: O usuário deve informar neste campo o principio ativo do medicamento.

Inserir um novo medicamento: Clique no botão "Novo" na barra de ferramentas à esquerda, preencha os campos acima citados e clique no botão "Confirmar" para efetivar a inclusão.

Alterar um medicamento: O usuário pode alterar os dados diretamente em seus respectivos campos acima citado e clicar no botão "Confirmar" para efetivar a alteração.

Excluir um medicamento: Clique no botão "Excluir" na barra de ferramentas à esquerda, Uma mensagem de confirmação ira aparecer e clique em "Ok". Obs.: um medicamento só poderá ser excluído se não houver pacientes relacionados ao mesmo.

Configurar pesos

A figura 17 monstra a tela de configuração dos pesos.

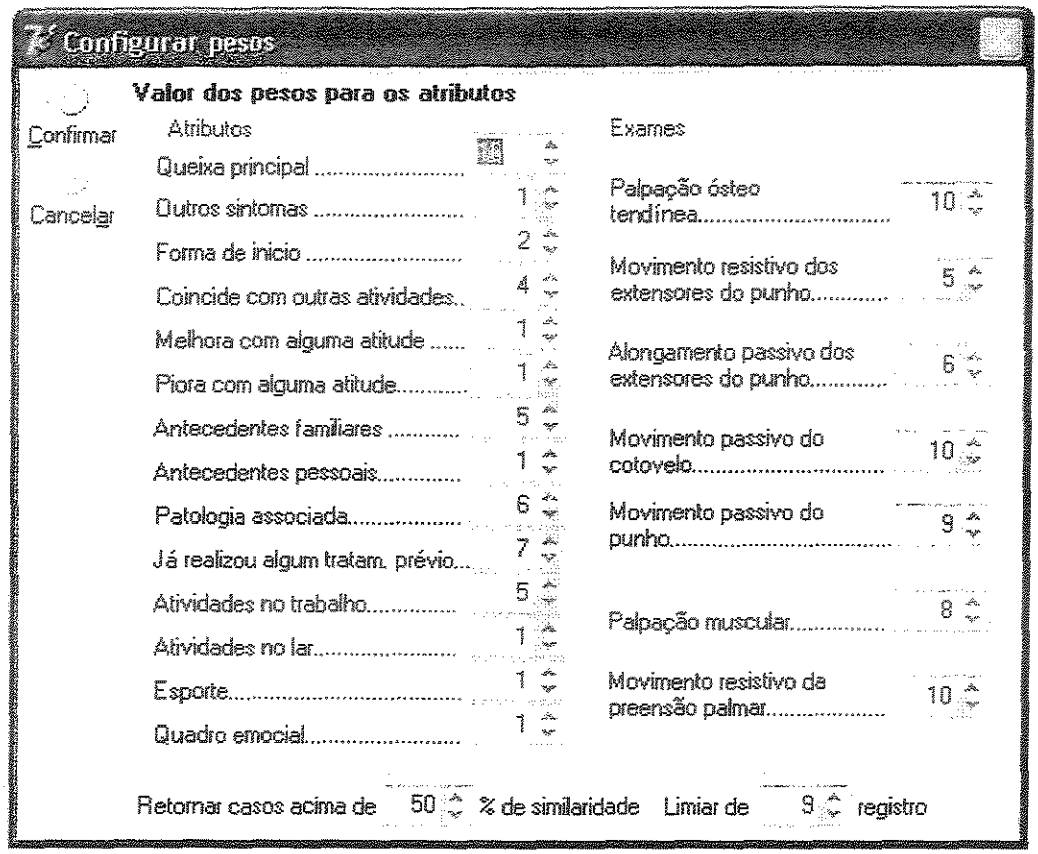


Figura 17 – Tela de Cadastro dos Pesos dos Atributos

Este módulo configura os pesos para cada atributo do sistema, os pesos podem variar de 1 a 10, para modificar o peso de um atributo basta informá-lo em seu respectivo campo, e clicar no botão "Confirmar" na barra de ferramentas à esquerda. O especialista pode definir um limite de similaridade e um limiar de registros a serem retornados.

Cadastro de casos

A figura 13 representa a tela Geral de Cadastro dos Casos.

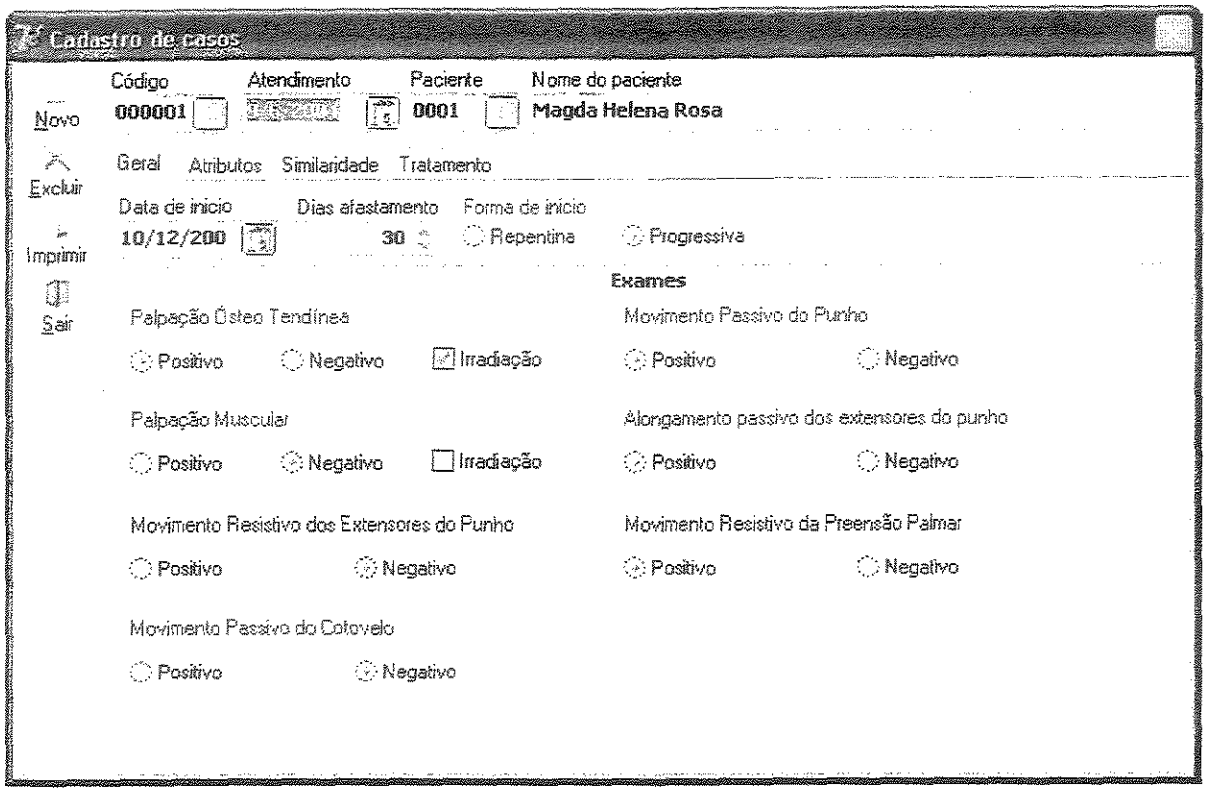


Figura 13 – Tela de Cadastro de Casos

Este módulo é responsável pela manutenção do registro do caso. É neste módulo que é realizado o cálculo entre os casos mais similares. As informações contidas neste módulo são:

- Código do caso: O sistema gera um código seqüencial automaticamente. Sendo este código um número que identifica o caso entre os demais.
- Atendimento: O especialista deve informar a data de atendimento do paciente, o sistema ira informa a data do sistema automaticamente, sendo que a mesma pode ser alterada.
- Código do paciente: O especialista deve informa o código do paciente que esta sendo atendido, Caso não saiba o código do paciente o especialista pode clicar no botão ao lado do campo para localizá-lo pelo nome.
- Nome do paciente: Campo informativo não pode ser alterado, apenas identifica o paciente.

- Data de inicio: O especialista deve informar neste campo a data de inicio dos sintomas.
- Dias afastamento: O especialista deve informar neste campo quantos dias o paciente ficou afastado.
- Forma de inicio: O especialista deve informar se a forma de inicio do sintoma do paciente foi repentina ou progressiva.
- Exames: O especialista dever informar os resultados dos exames apresentados se positivou ou negativo.

Na segunda aba **Atributos**, o especialista deve relacionar os descritores do atributo em questão de acordo com as informações passadas pelo paciente. A figura 14 representa esta tela.

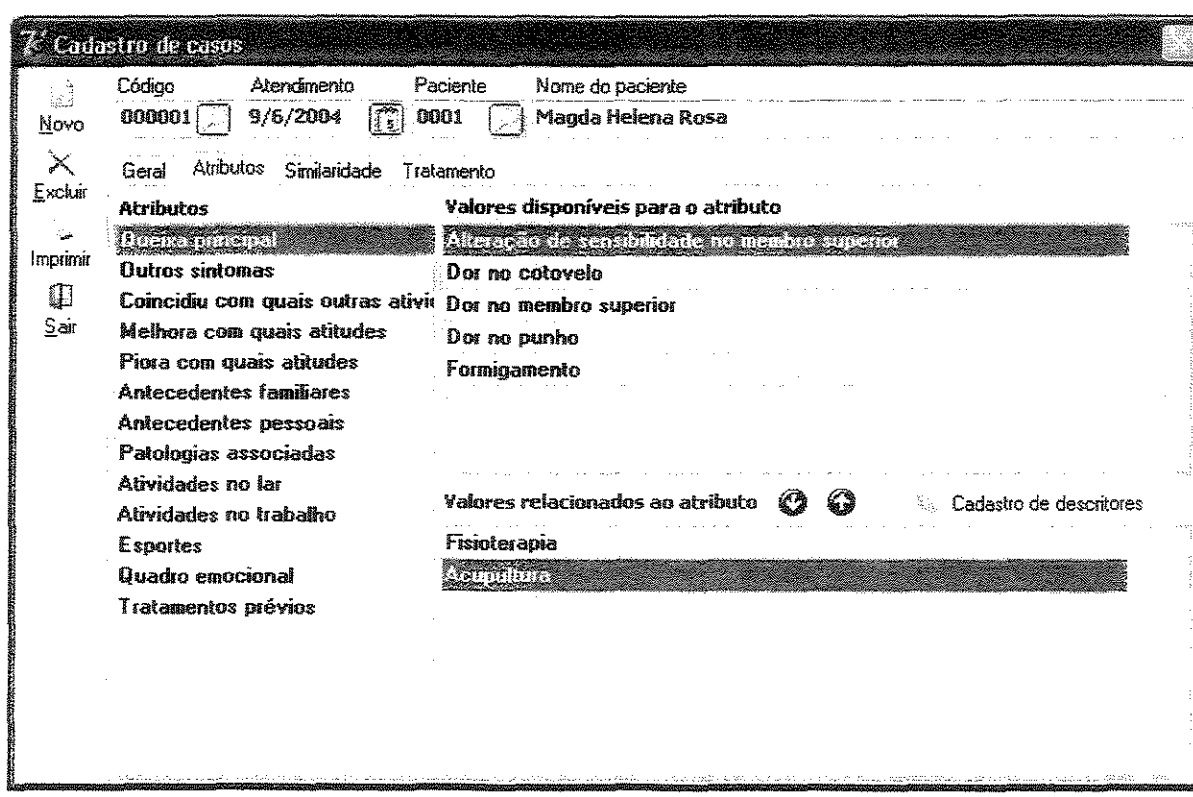


Figura 14 – Tela de Cadastro dos Atributos dos Casos

A terceira aba, **Similaridade,** é responsável pela escolha e cálculo das métricas, após informar os dados acima citados o especialista deve escolher uma métrica para calcular a similaridade do caso em questão, o especialista indica uma métrica e clica no botão "Calcular similaridade", o sistema irá retornar uma lista com o código do caso, descrição do paciente e a porcentagem de similaridade em relação ao caso atual. A figura 15 apresenta os casos retornados pelo FisioSmart como similares ao novo caso.

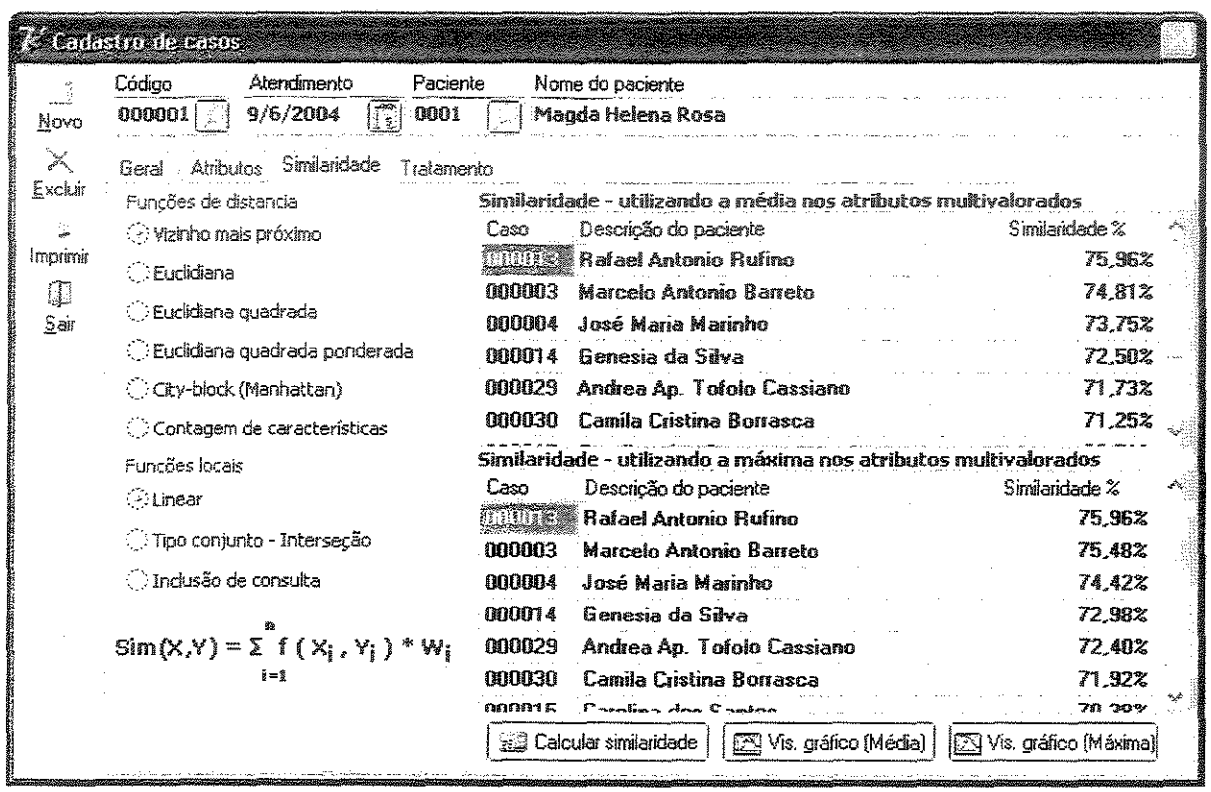


Figura 15 – Tela de Similaridade dos Casos

O especialista pode clicar no botão "Vis. Gráfico", para visualizar a similaridade de cada atributo em relação ao caso atual. A figura 18 apresenta um gráfico de similaridade.

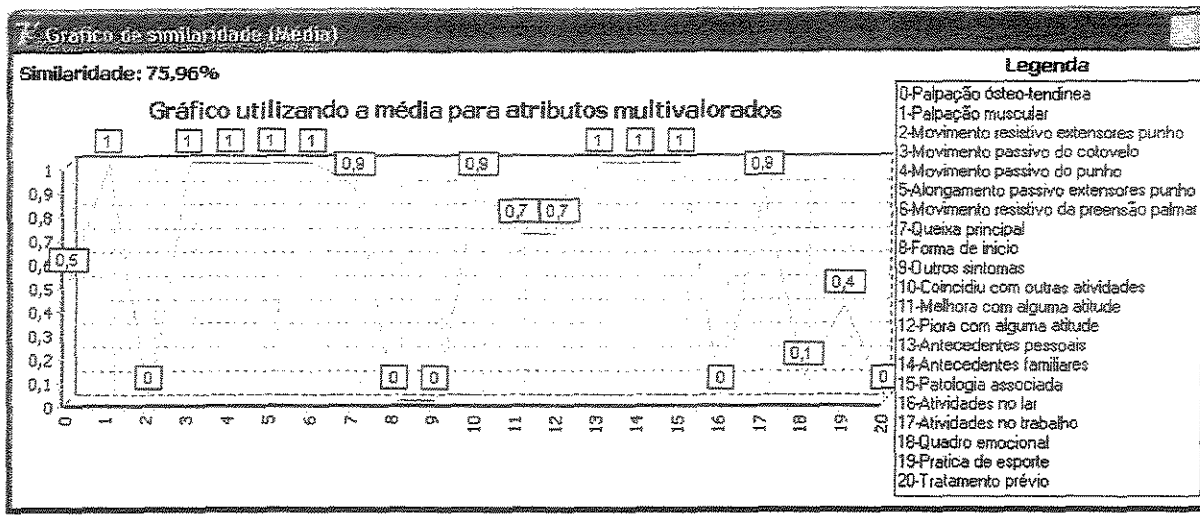


Figura 18 – Gráfico de Similaridade

Ao clicar na aba **Tratamento** o especialista terá a sua disposição o histórico do paciente e o tratamento aplicado ao 1º e 2º caso mais similar ficando a critério do especialista adotar ou não o tratamento sugerido pelo sistema, sendo que o especialista pode também aplicar o tratamento do caso mais similar com algumas modificações criando um novo tratamento e gerando mais uma informação válida para uma futura consulta. Caso um tratamento não obtiver sucesso o campo "tratamento bem sucedido" deve ser desmarcado, valido também para o campo "Tratamento aplicável a gestante". A figura 16 apresenta a tela de Tratamento.

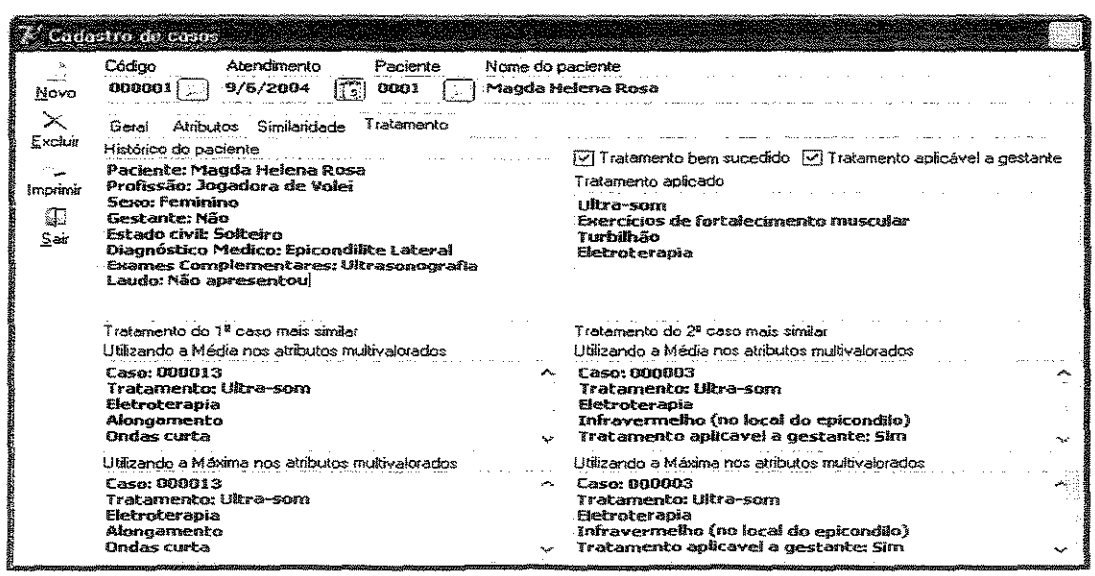


Figura 16 – Tela de Cadastro de Tratamento

É possível realizar algumas operações referentes ao cadastro de casos, como:

Inserir um novo caso: Basta clicar no botão "Novo" na barra de ferramentas à esquerda, preencher os campos acima citados e clique no botão "Confirmar" para efetivar a inclusão.

Alterar um caso: O especialista pode alterar os dados diretamente em seus respectivos campos acima citado e clicar no botão "Confirmar" para efetivar a alteração.

Excluir um caso: Para excluir um caso é preciso clicar no botão "Excluir" na barra de ferramentas à esquerda, uma mensagem de confirmação irá aparecer, para confirmar a atualização e clicar em "Ok".

Imprimir casos: Para imprimir os casos, é necessário clicar no botão "Imprimir" na barra de ferramentas à esquerda. Após clicar um relatório irá aparecer para imprimir os casos.

ANEXO 02 – Questionário Utilizado na Entrevista dos Pacientes com Epicondilite Lateral

Paciente			
Registro			
1ª Consulta			
Nome			
Nascimento			
Endereço			
Bairro			
Cidade			
UF			
Profissão			
Sexo			
Estado Civil			
Cor			
Diagnóstico Médico			
Exames Complementares			
Laudo			
Medicamentos			
Histórico			
Queixa Principal			
Outros Sintomas			
Data Início do Sintoma			
Forma do Sintoma	Repentina ou Progressiva		
Coincidiu com outras atividades			
Melhora com alguma atitude			
Piora com alguma atitude			
Antecedentes familiares			
Antecedentes Pessoais			
Patologia Associada			
Já realizou algum tratamento prévio referente a queixa principal (Qual)			
Descrição das atividades no trabalho			
Descrição das atividades no lar			
Prática de esporte / hobby			
Quadro emocional			
Duração Afastamento			
Exames Físicos			
Palpação ósteo-tendínea	S/N	Irradiação	S/N
Palpação Muscular	S/N	Irradiação	S/N
Movimento Passivo do cotovelo	S/N		
Movimento Passivo do punho	S/N		
Dor ou movimento resistido - Extensores do punho	S/N		
Dor ou movimento resistido - preensão palmar	S/N		
Dor ao alongamento passivo dos extensores do punho	S/N		

ANEXO 03 – Casos Retornados pela Aplicação

A tabela 29 mostra os casos de testes com os resultados obtidos em cada uma das métricas de similaridade utilizadas. Dessa forma, pode-se observar o comportamento das métricas a cada caso novo inserido na base de casos.

Tabela 29 – Casos Retornados pela Aplicação

Métrica de Similaridade	Casos Retornados pela Média	Valor Média %	Casos Retornados pela Máxima	Valor Máxima %	Classificação do Especialista	Classificação do Caso Recuperado
Caso - 26						
City Block - Manhattan	1-21-2	63,00 - 59,50 - 59,00	1-21-2	63,00 - 59,50 - 59,00	1, 5, 21	1º
Contagem de Características	21	50,00	—	Não tem máxima	1, 5, 21	3º
Euclidiana	1-21-2	79,40-76,94-76,63	1-21-2	79,46-76,94-76,63	1, 5, 21	1º
Euclidiana Quadrada	1-21-2	63,05-59,19-58,71	1-21-2	63,14-59,19-58,71	1, 5, 21	1º
Euclidiana Quadrada Ponderada	21-1-2	74,84-70,79-70,13	21-1-2	74,84-70,92-70,13	1, 5, 21	2º
Inclusão de Consulta	21-1	58,65-53,73	—	Não tem máxima	1, 5, 21	2º
Interseção	21-1	58,65-53,37	—	Não tem máxima	1, 5, 21	2º
Linear	21-1-2	72,02-67,40-67,02	21-1-2	72,02-68,08-67,02	1, 5, 21	2º
Caso – 27						
City Block - Manhattan	12-9-24	93,65 - 81,73 - 80,77	12-9-24	93,65 - 81,73 - 80,77	12, 17, 25	1º
Contagem de Características	12-25-17	77,50-70,00-67,50	—	Não tem máxima	12, 17, 25	1º
Euclidiana	12-23-25	96,02-90,53-88,77	12-23-25	96,02-90,53-88,77	12, 17, 25	1º
Euclidiana Quadrada	12-23-25	92,19-81,95-78,81	12-23-25	92,19-81,95-78,81	12, 17, 25	1º
Euclidiana Quadrada Ponderada	12-9-25	94,10-83,13-82,26	12-9-25	94,10-83,13-82,26	12, 17, 25	1º
Inclusão de Consulta	12-25-9	91,35-76,92-75,96	—	Não tem máxima	12, 17, 25	1º
Interseção	12-9-25	88,94-76,92-75,96	—	Não tem máxima	12, 17, 25	1º
Linear	12-9-24	93,65-81,73-80,77	12-9-24	93,65-81,73-80,77	12, 17, 25	1º
Caso – 28						
City Block - Manhattan	12-27-7	89,50 - 87,00 -84,25	12-27-7	90,00 - 87,50 - 85,50	17, 8, 12	3º
Contagem de Características	27-12-17	75,00-72,50-72,50	—	Não tem máxima	17, 8, 27, 12	3º
Euclidiana	12-7-27	93,49-92,47-92,20	12-7-27	93,53-92,71-92,25	17, 8, 27, 12	3º
Euclidiana Quadrada	12-7-27	87,40-85,52-85,02	12-7-27	87,48-85,95-85,10	17, 8, 27, 12	3º
Euclidiana Quadrada Ponderada	27-12-23	92,98-89,14-80,51	27-12-23	93,14-89,30-80,67	17, 8, 27, 12	3º
Inclusão de Consulta	27-12-17	84,94-81,09-67,63	—	Não tem máxima	17, 8, 27, 12	3º
Interseção	27-12-17	84,94-78,69-67,63	—	Não tem máxima	17, 8, 27, 12	3º
Linear	27-12-13	91,83-87,98-77,31	27-12-13	91,83-87,98-77,31	17, 8, 27, 12	3º
Caso – 29						
City Block - Manhattan	27-10-3	75,00-74,75-73,50	27-10-3	75,00-75,00-73,50	3, 9, 20	3º
Contagem de Características	3-4-12	57,50-57,50-57,50	—	Não tem máxima	3, 9, 20	1º
Euclidiana	27-21-25-10-3	86,46-86,38-86,22-85,62-85,30	27-21-25-10-3	86,46-86,38-86,22-85,62-85,30	3, 9, 20	5º
Euclidiana Quadrada	27-21-25-10-3	74,76-74,62-74,33-73,38-72,76	27-21-25-10-3	74,76-74,62-74,33-73,38-72,76	3, 9, 20	5º

Euclidiana Quadrada Ponderada	3-27-26	76,47-76,39-76,11	3-27-26	76,47-76,39-76,11	3, 9, 20	1°
Inclusão de Consulta	1-26-3	66,35-61,54-60,58	—	Não tem máxima	3, 9, 20	2°
Interseção	1-26-3	62,98-61,54-60,58	—	Não tem máxima	3, 9, 20	3°
Linear	3-27-26	62,98-61,54-60,58	3-27-26	74,52-74,33-73,56	3, 9, 20	1°
Caso – 30						
City Block - Manhattan	17-27-25	84,50-82,00-81,00	17-27-25	84,50-82,00-81,00	15, 17, 20	2°
Contagem de Características	17-3-27	75,00-72,50-72,50	—	Não tem máxima	15, 17, 20	2°
Euclidiana	17-15-25	91,52-90,42-89,95	17-15-25	91,52-90,42-89,95	15, 17, 20	2°
Euclidiana Quadrada	17-15-25	83,76-81,76-80,90	17-15-25	83,76-81,76-80,90	15, 17, 20	2°
Euclidiana Quadrada Ponderada	15-3-17	90,73-89,15-84,66	15-3-17	90,73-89,15-84,66	15, 17, 20	1°
Inclusão de Consulta	3-17-20	81,73-76,92-75,00	—	Não tem máxima	15, 17, 20	1°
Interseção	3-17-15	81,73-76,92-75,00	—	Não tem máxima	15, 17, 20	2°
Linear	3-15-17	87,31-86,73-82,79	3-15-17	87,31-86,73-82,79	15, 17, 20	2°
Caso – 31						
City Block - Manhattan	27-13-12	66,50-66,00-65,50	27-13-12	66,50-66,00-65,50	9, 27, 13	2°
Contagem de Características	12-27-23	60,00-55,00-50,00	—	Não tem máxima	9, 27, 13, 23	2°
Euclidiana	13-27-9	82,40-82,32-80,83	13-27-9	82,40-82,32-80,83	9, 27, 13, 23	2°
Euclidiana Quadrada	13-27-9	67,90-67,76-65,33	13-27-9	67,90-67,76-65,33	9, 27, 13, 23	2°
Euclidiana Quadrada Ponderada	28-27-12	78,30-78,11-77,49	28-27-12	78,46-78,11-77,49	9, 27, 13, 23	2°
Inclusão de Consulta	12-17-28	65,96-75,00-74,04	—	Não tem máxima	9, 27, 13, 23	2°
Interseção	12-27-28	75,96-75,00-67,63	—	Não tem máxima	9, 27, 13, 23	2°
Linear	27-12-18	77,21-77,02-76,15	27-12-18	77,21-77,02-76,15	9, 27, 13, 23	2°
Caso – 32						
City Block - Manhattan	19-2-27	82,50-65,25-63,75	19-2-27	83,50-65,50-64,20	19, 11, 22	1°
Contagem de Características	19-2-11	65,00-57,50-52,50	—	Não tem máxima	19, 11, 22	1°
Euclidiana	19-27-20	91,26-80,52-79,81	19-27-20	91,47-80,53-79,88	19, 11, 22	1°
Euclidiana Quadrada	19-27-20	83,29-64,83-63,69	19-27-20	83,67-64,86-63,81	19, 11, 22	1°
Euclidiana Quadrada Ponderada	19-22-11	91,69-68,94-61,30	19-22-11	92,08-68,99-61,35	19, 11, 22	1°
Inclusão de Consulta	19-22-11	75,96-58,65-53,85	—	Não tem máxima	19, 11, 22	1°
Interseção	19-22-11	75,96-58,65-53,85	—	Não tem máxima	19, 11, 22	1°
Linear	19-22-11	88,46-68,27-60,10	19-22-11	89,42-68,75-60,58	19, 11, 22	1°
Caso – 33						
City Block - Manhattan	20-11-25-23-10	72,00-68,50-68,00-67-50-66,50	20-11-25-23-10	72,50-69,00-68,50-68-00-67,00	10, 9, 13	5°
Contagem de Características	25-10-9	60,00-55,00-55,00	—	Não tem máxima	10, 9, 13	2°
Euclidiana	20-11-23-10	85,02-82,87-82,58-81,97	20-11-23-10	85,08-82,92-82,64-81,97	10, 9, 13	4°
Euclidiana Quadrada	20-11-23-10	72,29-68,67-68,19-67,19	20-11-23-10	72,38-68,76-68,29-67,19	10, 9, 13	4°
Euclidiana Quadrada Ponderada	18-6-11	78,05-77,21-76,12	18-6-11	78,41-78,24-76,31	10, 9, 13	—
Inclusão de Consulta	18-11-28	72,12-67,31-64,42	—	Não tem máxima	10, 9, 13	—
Interseção	18-11-9	72,12-67,31-63,46	—	Não tem máxima	10, 9, 13	2°
Linear	18-6-11	77,16-74,18-74,04	18-6-11	77,40-75,87-75,00	10, 9, 13	—
Caso – 34						
City Block - Manhattan	12-27-17	93,00-84,00-81,00	12-27-17	95,00-86,00-83,00	12, 17, 3	1°
Contagem de Características	12-27-3	90,00-72,50-70,00	—	Não tem máxima	12, 17, 3	1°
Euclidiana	12-27-17	94,31-90,32-89,68	12-27-17	95,12-91,16-90,53	12, 17, 3	1°

Euclidiana Quadrada	12-27-17	88,95-81,57-80,43	12-27-17	90,48-83,10-81,95	12, 17, 3	1°
Euclidiana Quadrada Ponderada	12-27-11	89,81-83,99-83,94	12-27-11	91,35-85,53-85,48	12, 17, 3	1°
Inclusão de Consulta	12-27-11	88,94-81,25-81,25	—	Não tem máxima	12, 17, 3	1°
Interseção	12-27-11	88,14-81,25-81,25	—	Não tem máxima	12, 17, 3	1°
Linear	12-27-11	89,42-83,37-83,27	12-27-11	91,35-85,29-85,19	12, 17, 3	1°
Caso – 35						
City Block - Manhattan	12-27-25	83,75-82,25-79,25	12-27-25	85,00-83,50-80,50	34, 3, 12	3°
Contagem de Caracteristicas	34-12-3	75,00-75,00-70,00	—	Não tem máxima	34, 3, 12, 17	1°
Euclidiana	27-12-34	90,44-89,80-89,16	27-34-12	90,61-89,97-89,97	34, 3, 12, 17	3°-2°
Euclidiana Quadrada	27-12-34	81,79-80,64-79,50	27-34-12	82,10-80,95-80,95	34, 3, 12, 17	3°-2°
Euclidiana Quadrada Ponderada	34-12-27	94,47-88,89-84,80	34-12-11	96,15-89,42-85,76	34, 3, 12, 17	1°
Inclusão de Consulta	34-12-11	86,54-79,81-75,96	—	Não tem máxima	34, 3, 12, 17	1°
Interseção	34-12-11	86,54-79,81-75,96	—	Não tem máxima	34, 3, 12, 17	1°
Linear	34-12-27	92,79-87,50-82,88	34-12-27	96,15-89,42-84,81	34, 3, 12, 17	1°
Caso – 36						
City Block - Manhattan	25-8-4	83,00-80,50-78,50	25-8-4	83,00-80,50-78,50	8, 9, 25	2°
Contagem de Caracteristicas	25-8-9	70,00-65,00-60,00	—	Não tem máxima	8, 9, 25	2°
Euclidiana	25-8-4	89,76-89,68-88,05	25-8-4	89,76-89,68-88,05	8, 9, 25	2°
Euclidiana Quadrada	25-8-4	80,57-80,43-77,52	25-8-4	80,57-80,43-77,52	8, 9, 25	2°
Euclidiana Quadrada Ponderada	8-25-9	89,04-86,46-84,11	8-25-9	89,04-86,46-84,11	8, 9, 25	1°
Inclusão de Consulta	8-25-9	78,85-77,88-75,96	—	Não tem máxima	8, 9, 25	1°
Interseção	8-25-9	78,85-77,88-75,96	—	Não tem máxima	8, 9, 25	1°
Linear	8-25-9	86,92-86,15-82,60	8-25-9	86,92-86,15-82,60	8, 9, 25	1°
Caso – 37						
City Block - Manhattan	34-35-12	85,50-85,00-84-50	34-35-12	87,00-87,00-85,00	17, 3, 4, 34	4°
Contagem de Caracteristicas	3-25-34	75,00-75,00-75,00	—	Não tem máxima	17, 3, 4, 30, 34	2°
Euclidiana	35-12-25-34	91,90-91,65-91,26-91,23	35-35-12	92,07-92,07-91,65	17, 3, 4, 30, 34	5°
Euclidiana Quadrada	35-12-25-34	84,45-84,00-83,29-83,24	35-34-12	84,76-84,76-84,00	17, 3, 4, 30, 34	5°
Euclidiana Quadrada Ponderada	35-34-12	96,39-95,38-90,04	35-34-12	96,92-96,92-90,04	17, 3, 4, 30, 34	5°
Inclusão de Consulta	34-35-12	95,19-90,38-87,50	—	Não tem máxima	17, 3, 4, 30, 34	5°
Interseção	34-35-12	92,79-85,58-85,10	—	Não tem máxima	17, 3, 4, 30, 34	5°
Linear	35-34-12	94,62-94,62-89,42	35-34-12	96,54-96,54-89,42	17, 3, 4, 30, 34	5°
Caso – 38						
City Block - Manhattan	35-9-4	88,00-83,25-75,75	35-9-4	88,50-83,50-76,00	9, 10, 35	2°
Contagem de Caracteristicas	35-9-4	75,00-65,00-65,00	—	Não tem máxima	9, 10, 35	2°
Euclidiana	35-9-10	93,83-91,09-86,52	35-9-10	93,99-91,10-86,52	9, 10, 35	2°
Euclidiana Quadrada	35-9-10	88,05-82,98-74,86	35-9-10	88,33-83,00-74,86	9, 10, 35	2°
Euclidiana Quadrada Ponderada	35-37-34	91,01-89,17-88,52	35-37-34	91,30-89,22-88,57	9, 10, 35	3°
Inclusão de Consulta	37-34-35	78,85-78,85-76,92	—	Não tem máxima	9, 10, 35	3°
Interseção	37-34-35	78,85-78,85-76,92	—	Não tem máxima	9, 10, 35	3°
Linear	35-37-34	87,31-87,12-86,73	35-37-34	87,79-87,60-87,21	9, 10, 35	3°

Caso – 39						
City Block - Manhattan	4-25-37	89,00-85,00-83,00	4-25-37	90,00-85,50-84,00	4, 9, 17	1°
Contagem de Características	4-25-37	70,00-65,00-65,00	—	Não tem máxima	4, 9, 17	1°
Euclidiana	4-25-38	94,06-91,68-89,93	4-25-38	94,26-91,73-90,13	4, 9, 17	1°
Euclidiana Quadrada	4-25-38	88,48-84,05-80,88	4-25-38	88,86-84,14-81,24	4, 9, 17	1°
Euclidiana Quadrada Ponderada	4-38-7	90,09-86,95-85,43	4-38-7	90,21-87,02-85,52	4, 9, 17	1°
Inclusão de Consulta	38-4-37	77,88-74,04-74,04	—	Não tem máxima	4, 9, 17	2°
Interseção	38-4-37	77,88-74,04-74,04	—	Não tem máxima	4, 9, 17	2°
Linear	4-38-37	87,45-84,47-84,38	4-38-37	88,08-85,00-84,62	4, 9, 17	1°
Caso – 40						
City Block - Manhattan	25-4-37	90,50-89,00-85,00	25-4-37	91,00-90,00-85,50	17, 4, 39	2°
Contagem de Características	25-4-37	80,00-72,50-70,00	—	Não tem máxima	17, 4, 39	2°
Euclidiana	25-4-37	94,26-93,09-91,55	25-4-37	94,37-93,22-91,68	17, 4, 39	2°
Euclidiana Quadrada	25-4-37	88,86-86,67-83,81	25-4-37	89,05-86,90-84,05	17, 4, 39	2°
Euclidiana Quadrada Ponderada	9-38-25	93,34-87,29-86,00	9-38-25	93,52-87,63-86,16	17, 4, 39	—
Inclusão de Consulta	9-38-25	75,00-72,60-72,12	—	Não tem máxima	17, 4, 39	2°
Interseção	9-25-38	75,00-72,12-70,99	—	Não tem máxima	17, 4, 39	—
Linear	9-38-25	89,47-84,47-83,41	9-38-25	90,00-85,96-83,94	17, 4, 39	—