

UNIVERSIDADE FEDERAL DE SANTA CATARINA

**PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO**

**IDENTIFICAÇÃO DIFUSA DE SISTEMAS:
PROPOSTA DE UM MODELO ADAPTATIVO**



0.279.229-1



UFSC-BU

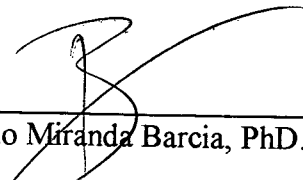
ARAN BEY TCHOLAKIAN MORALES

FLORIANÓPOLIS, ABRIL DE 1997

Autor : ARAN BEY TCHOLAKIAN MORALES

IDENTIFICAÇÃO DIFUSA DE SISTEMAS: PROPOSTA DE UM MODELO ADAPTATIVO

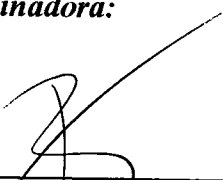
Esta tese foi julgada adequada para a obtenção do título de “Doutor em Engenharia”, especialidade Engenharia de Produção e aprovada em sua forma final pelo Programa de Pós-graduação.



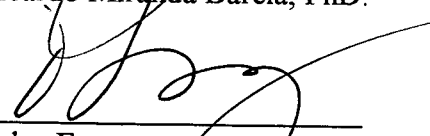
Ricardo Miranda Barcia, PhD.

Coordenador do Programa de Pós-Graduação em Engenharia de Produção

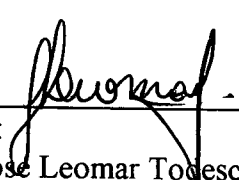
Banca Examinadora:



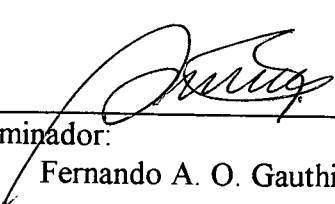
Presidente:
Ricardo Miranda Barcia, PhD.




Examinador Externo:
Tamio Shimizu, Dr.



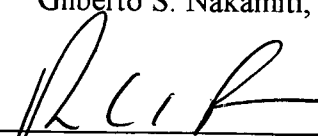
Examinador:
Jose Leomar Todesco, Dr.



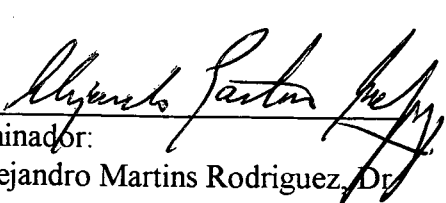
Examinador:
Fernando A. O. Gauthier, Dr.




Examinador Externo:
Gilberto S. Nakamiti, Dr.



Examinador:
Roberto C. S. Pacheco, Dr.



Examinador:
Alejandro Martins Rodriguez, Dr.



Moderador:
Luiz F. Maia, Dr.

**Este trabajo está dedicado a mi esposa Glória,
por el tiempo que estuvo entre nosotros.**

AGRADECIMENTOS

Ao Capes pelo apoio financeiro no decorrer do trabalho.

À Universidade Federal de Santa Catarina e ao Departamento de Engenharia de Produção, por oferecerme a suas instalações para o desenvolvimento deste trabalho.

Ao Professor Plínio Stange, por sua orientação e dedicação.

Ao Coordenador do Curso de Pós-Graduação em Engenharia de Produção, e Co-Orientador deste trabalho, Prof. Ricardo Miranda Barcia.

Aos professores Alejandro Martins e Roberto Pacheco, pelas sugestões e correções e dedicação no decorrer do trabalho.

Aos professores José Leomar Todesco, Fernando Gauthier, Gilberto Nakamiti e Tamio Shimizu pelas sugestões para melhorar o trabalho.

Aos colegas de curso pelas idéias desenvolvidas em conjunto e discussões sobre os diferentes temas abordados durante todo o curso.

E a todos os que, de forma direta ou indireta, colaboraram para a realização deste trabalho.

RESUMO

Identificação de sistemas é um processo iterativo, que facilita obter novo conhecimento sobre a natureza do sistema observado a cada nova iteração. A identificação de sistemas está vinculada à invenção e avaliação de teorias científicas.

O propósito deste trabalho é explorar métodos alternativos para o processo da identificação de sistemas. O modelo proposto é um modelo baseado em regras, que representa as relações entre os agentes do sistema. A incerteza associada aos sistemas é incorporada no modelo via teoria dos conjuntos difusos.

A técnica de busca utilizada para descobrir as regras e escolher as funções de pertinência dos conjuntos difusos que otimizam a resposta do modelo, são os algoritmos genéticos. A natureza robusta e os mecanismos simples dos algoritmos genéticos fazem deles uma ferramenta adequada para este propósito.

Os algoritmos genéticos são uma técnica baseada nos princípios evolutivos de Darwin. No entanto, biologicamente sempre foi discutido de que forma as adaptações adquiridas por aprendizado durante o tempo de vida de um indivíduo são passadas para seus descendentes. Neste trabalho consideramos estas teorias biológicas e propomos um modelo onde evolução e aprendizado interagem.

ABSTRACT

System identification is an interactive process that facilitates the acquisition of new knowledge about the nature of an observed system at every *iteration cycle*. It has a direct relationship with the invention and evaluation of scientific theories.

The general purpose of the present work is to explore alternative methods for the identification of systems. The proposed model is based on production rules, which represent the relationship between different system's agents.

The uncertainty inherent to a system was incorporate in the model through the use of Fuzzy Sets Theory. Genetic Algorithms were used as the search technique for the processes of rule discovery and making a choice between several possible membership functions, and, at the same time, to optimize the performance of a model. The robust nature and simple mechanisms of the genetic algorithms make them a suitable tool for such tasks.

Genetics Algorithms is a mathematical technique based on Darwin's evolutionary principles. But, an idea that has always had attractions is that adaptations acquired during an individual's lifetime, by learning or in other way, are passed on to its offspring. This work propose and uses a model that incorporates evolution and learning.

SUMÁRIO

CAPÍTULO I - INTRODUÇÃO

1.1 Introdução	1
1.2 Sistemas e Modelos.....	2
1.3 Identificação de Sistemas.....	4
1.4 Objetivos do Trabalho.....	7

CAPÍTULO II - IDENTIFICAÇÃO DIFUSA DE SISTEMAS

2.1 Introdução	10
2.2 Sistemas Especialistas.....	11
2.3 Sistemas Difusos.....	12
2.4 Identificação Difusa de Sistemas.....	16
2.5 Algoritmos Genéticos	19
2.6 Os Algoritmos Genéticos na Identificação Difusa de Sistemas.....	22
2.7 Processos Adaptativos: Evolução e Aprendizado	25

CAPÍTULO III - MODELO PROPOSTO : FUNDAMENTAÇÃO E IMPLEMENTAÇÃO

3.1 Introdução	36
----------------------	----

3.2 Modelo Proposto.....	37
3.3 Implementação do Modelo Proposto	42
3.3.1 Representação dos Indivíduos	43
3.3.2 Nível Local	44
3.3.3 Nível Global	49

CAPÍTULO IV - SIMULAÇÕES REALIZADAS E RESULTADOS OBTIDOS

4.1 Introdução	51
4.2 Funções Testadas	52
4.3 Sistemas Simulados.....	68

CAPÍTULO V - CONCLUSÕES E PROBLEMAS EM ABERTO

5.1 Introdução	71
5.2 Resultados, Objetivos Alcançados e Limitações.....	71
5.3 Trabalhos Futuros	74

REFERÊNCIAS BIBLIOGRÁFICAS.....	76
--	-----------

LISTA DE FIGURAS

Figura 1: Estrutura do Sistema Difuso	13
Figura 2: Funções de Pertinência Pertencentes à Variável Lingüística Temperatura do Secador da Massa Cerâmica	14
Figura 3: Processo de Inferência Max-Mín	15
Figura 4: Ciclo Básico dos Algoritmos Genéticos.....	21
Figura 5: Uma Visão mais Completa de Evolução	31
Figura 6: Regiões Cobrindo o Comportamento do Sistema.....	39
Figura 7: Regiões Difusas Cobrindo o Comportamento do Sistema	40
Figura 8: Modelo Adaptativo Proposto.....	42
Figura 9: Representação de um Indivíduo com Duas Características e Cinco Lócus.....	43
Figura 10: Funções de Pertinência de Uma Característica de um Indivíduo	44
Figura 11: <i>Fuzzification</i> Aplicado sobre a Característica de Um Indivíduo.....	45
Figura 12: Cálculo da Área Sombreada.....	47
Figura 13: Método de <i>Defuzzification</i>	47
Figura 14: Comparação entre <i>Fuzzification</i> e o Método de Inferência Aplicado sobre a Variável de Saída	48
Figura 15: Resultado da Função <i>f1</i> com 7 Valores Lingüísticos e 18 Regras.....	53
Figura 16: Resultado da Função <i>f1</i> com 5 Valores Lingüísticos e 6 Regras.....	54
Figura 17: Resultado da Função <i>f2</i> com 7 Valores Lingüísticos e 16 Regras.....	54
Figura 18: Resultado da Função <i>f2</i> com 5 Valores Lingüísticos e 6 Regras.....	55
Figura 19: Resultado da Função <i>f3</i> com 5 Valores Lingüísticos e 17 Regras.....	57

Figura 20: Resultado da Função f_3 com 7 Valores Lingüísticos e 24 Regras.....	57
Figura 21: Resultado da Função f_3 com 9 Valores Lingüísticos e 50 Regras.....	58
Figura 22: Resultado da Função f_4 com 7 Valores Lingüísticos e 36 Regras.....	59
Figura 23: Resultado da Função f_4 com 9 Valores Lingüísticos e 43 Regras.....	60
Figura 24: Resultado da Função f_4 com 12 Valores Lingüísticos e 61 Regras.....	60
Figura 25: Resultado da Função f_5 com 5 Valores Lingüísticos e 16 Regras.....	61
Figura 26: Resultado da Função f_5 com 9 Valores Lingüísticos e 41 Regras.....	62
Figura 27: Resultado da Função f_5 com 12 Valores Lingüísticos e 77 Regras.....	62
Figura 28: Representação gráfica da função f_6	63
Figura 29: Representação gráfica dos erros da função f_6	64
Figura 30: Representação Gráfica da Função f_7	65
Figura 31: Representação gráfica dos erros da função f_7	66
Figura 32: Representação Gráfica da Função f_8	67
Figura 33: Representação Gráfica dos Erros da Função f_8	67

LISTA DE QUADROS

Quadro 1: Comparação dos Erros da Função f_1 , com Aprendizado e sem Aprendizado	55
Quadro 2: Comparação dos Erros da Função f_2 , com Aprendizado e sem Aprendizado	56
Quadro 3: Comparação de Resultados das Funções f_1 e f_3	58
Quadro 4: Comparação de Resultados da Função f_4	61
Quadro 5: Resultados da função f_5	63
Quadro 6: Comparação dos Resultados da Função f_6	64
Quadro 7: Comparação dos Resultados da Função f_7	65
Quadro 8: Comparação dos Resultados da Função f_8	66
Quadro 9: Comparação dos Erros de Diferentes Modelos Aplicados no Processo de Limpeza da Água.....	69

CAPÍTULO I

INTRODUÇÃO

1.1 Introdução

O trabalho que ora apresentamos constitui-se do relato de pesquisa desenvolvida sobre a proposta de um modelo adaptativo para a identificação difusa de sistemas, e tem a seguinte estrutura:

Neste capítulo, fazemos uma introdução sobre a importância da identificação de sistemas como método científico, as motivações que nos levaram a realizar a presente pesquisa e por consequência os objetivos da mesma.

No capítulo 2, enumeramos as ferramentas e trabalhos encontrados na literatura, sobre identificação difusa de sistemas. Fazemos algumas considerações sobre estes trabalhos e ressaltamos as principais diferenças com o modelo proposto. O estado da arte dos principais ferramentas utilizadas no modelo é abordado.

No capítulo 3, é apresentada a fundamentação do modelo proposto, assim como a descrição do algoritmo proposto para a implementação do modelo. Ambos os itens são defendidos e amplamente explanados. A contribuição científica e originalidade deste trabalho são encontradas nesse capítulo.

No capítulo 4, várias funções e sistemas são simulados com o algoritmo implementado, e seus resultados expostos, discutidos e comparados com os resultados obtidos por outros trabalhos.

No capítulo 5, são compiladas as conclusões da presente pesquisa, discutidos os objetivos alcançados, as principais limitações do algoritmo proposto, e sugeridos futuros trabalhos de pesquisa na área.

1.2 Sistemas e Modelos

A tarefa de observar os fenômenos que acontecem no ambiente em que vivemos, com o fim de construir modelos que permitam fazer previsões, controlar ou afetar tais fenômenos foi e é uma das principais funções da ciência. A importância da construção de modelos de fenômenos que acontecem ao nosso redor, baseia-se no fato de que a construção destes modelos está diretamente vinculada à avaliação de teorias científicas (Fogel, 1991), objetivo primordial da ciência.

Durante muito tempo, a ciência desenvolveu métodos que, a seu ver, levavam a uma descrição exata dos processos físicos e naturais. A ciência baseava-se em leis deterministas e era associada à certeza. Assim, considerava-se que a matemática e as relações numéricas nos levaram a um conhecimento mais seguro do que nos diziam nossos sentidos. Estes métodos são geralmente processos iterativos que nos conduzem a modelos que representam os fenômenos que estamos identificando.

Para especificar os passos que levarão a obter os modelos de fenômenos que acontecem em nosso ambiente, é necessário ter primeiramente uma definição formal de sistemas, como também citar os mecanismos e ferramentas disponíveis que permitirão obter tal modelo. Como se verá adiante, as ferramentas utilizadas e as características destas são de fundamental importância para que o modelo obtido responda de forma satisfatória ao sistema observado.

Então, antes de estabelecer qual é a tarefa do construtor do modelo, é necessário definir o que se entende por sistema. Neste trabalho, utilizam-se algumas das idéias desenvolvidas por Caswell (1972) sobre sistemas. Pode-se definir um sistema como uma coleção de agentes, onde cada um comporta-se de um modo, isto é, possui um comportamento próprio, capaz de interagir com os outros, para satisfazer um conjunto específico de exigências e manter um comportamento coerente com seu ambiente.

Os agentes referidos poderiam ser desde simples células, genes, formigas, variáveis numéricas, *string* de *bits* ou até um organismo humano. Estes agentes, segundo a definição de sistemas utilizada e conforme dissemos anteriormente, possuem um comportamento

próprio e interagem com os outros agentes, influenciando em seus comportamentos e sobre o ambiente. Tal fato fornece um indicativo claro da existência de relações agente-agente e agente-ambiente.

Para poder afirmar que os agentes exibem comportamento, é necessário haver um conjunto de estruturas, escolhidas pelo observador, onde é representado o conhecimento dos agentes, e através das quais são adquiridas as estratégias de adaptação de cada agente, para responder às exigências ambientais. Às exigências ambientais podem mudar com o tempo, o que implica que as interações entre agentes do sistema e destes com o ambiente, não são estáticas. Assim, as estruturas comportamentais do sistema devem aceitar as adaptações que o comportamento dos agentes deve ter para responder de forma adequada às exigências do ambiente e às exigências dos outros agentes.

O conhecimento representado nestas estruturas (Caswell, 1972) servirá para obter a descrição do sistema por meio do comportamento de seus agentes através do tempo. Isto é, em cada instante de tempo, o comportamento dos agentes será determinado por uma instância dessas estruturas. Assim, a descrição comportamental do sistema é uma série temporal determinada pelas instâncias de cada agente.

Uma vez identificados os agentes do sistema e as estruturas que contêm o comportamento de cada agente, para especificar completamente o sistema, necessita-se do estabelecimento das relações agente-agente e agente-ambiente. O estabelecimento das relações entre os agentes expressa as limitações de recursos disponíveis no ambiente e como a distribuição dos recursos entre os agentes do sistema influi em seus comportamentos.

Uma vez definido o que se entende por sistemas, pode-se estabelecer a tarefa do construtor do modelo. De maneira formal, a tarefa do construtor consiste na identificação dos agentes e de seus comportamentos, assim como das relações significativas entre os agentes do sistema, e deles com o ambiente. Para realizar esta identificação, o construtor tem à disposição várias ferramentas e procedimentos de análise.

O primeiro passo na construção de um modelo é a divisão do universo em um

sistema e seu ambiente. Isto significa delimitar o sistema e identificar os agentes que pertencem ao mesmo e os que não pertencem.

O segundo passo é representar o comportamento dos agentes, que é percebido pelo observador como uma relação causa-efeito. Assim, as estruturas comportamentais podem ser representadas como uma estrutura de estímulos e respostas. O critério usual da escolha dos estímulos e respostas está baseado no conhecimento que o observador possui sobre o sistema.

Por estes motivos, a modelagem de sistemas tem muito de arte, além de ciência, já que seu estudo está baseado na premissa de modelar o observado pelo construtor do modelo que possui, além do conhecimento objetivo obtido das observações e do acúmulo de informações, muito conhecimento subjetivo, intuitivo e qualitativo do sistema, obtido da sua experiência no convívio com os fenômenos observados.

Uma vez identificados os agentes, comportamentos e relações, poderá ser construído um modelo do sistema que, testado com o sistema real, poderá modificar-se e ajustar-se. É através deste processo iterativo que o observador forma o modelo final do sistema. Assim, este processo de identificação pode ser visto como um processo iterativo que obtém, a cada iteração, novo conhecimento sobre o sistema observado.

1.3 Identificação de Sistemas

Existem diferentes maneiras de obter um modelo para um determinado sistema (Ljung, 1987), incluindo modelos matemáticos, intuitivos e gráficos, sendo os primeiros os mais utilizados. Neste sentido, são muito comuns as técnicas de análise quantitativa, baseadas na modelagem matemática e no conhecimento objetivo, que, por meio do uso de símbolos, equações e outras sentenças matemáticas, representam uma certa realidade, (Zadeh, 1973). Isto significa que os agentes, seus comportamentos e relações são expressos através de um modelo matemático.

Para construir um modelo matemático de um sistema, pode-se adotar duas formas

(Ljung, 1987):

- a) baseando-se em leis físicas e relações que supostamente governam o comportamento do sistema, ou
- b) utilizando procedimentos de identificação, para inferir um modelo baseado em dados observados do sistema. Isto ocorre quando a modelagem direta é impossível, devido a que o conhecimento do mecanismo do sistema é incompleto, ou as suas propriedades exibidas mudam de forma imprevisível.

Em Ljung (1987), são descritos vários procedimentos para obter modelos matemáticos derivados a partir de dados de entrada-saída de um fenômeno observado. No entanto, alguns sistemas cujo comportamento é classificado como de difícil identificação, resistem a tais procedimentos matemáticos. Nestes casos, o observador tem basicamente dois problemas:

- a) a identificação das partes componentes do sistema (agentes, comportamentos e relações) e
- b) a dificuldade de representar o conhecimento e a intuição que ele possui do sistema com as técnicas de análise quantitativa.

Por tal motivo, para poder obter um modelo do sistema com as ferramentas disponíveis, o observador tende a simplificar a realidade observada. Esta simplificação observa-se basicamente nas relações entre agentes e entre eles e seu ambiente, já que os procedimentos quantitativos são limitados no tratamento de relações altamente não lineares e dinâmicas. Como consequência, o modelo obtido não reflete a realidade tal como ela é.

Entre os sistemas que são classificados como de difícil identificação, podem ser citados: os processos de tomada de decisão, os sistemas econômicos, os sistemas físicos e químicos, e os processos biológicos, entre outros (Crosby, 1987; Rosen, 1987).

Os processos de tomada de decisão, por exemplo, podem ser considerados como sistemas de difícil identificação, devido às influências dos fatores humanos sobre eles. A

habilidade na manipulação de conceitos vagos, mal definidos, ambíguos; a tolerância à imprecisão e, conseqüentemente, a inferência de conhecimento destes conceitos, são habilidades humanas. Assim, os fatores humanos são muito complexos para serem descritos com todas as suas propriedades, limites, tolerâncias, por intermédio de uma ferramenta matemática, que seja capaz de descrever e integrar todas as medidas mencionadas anteriormente (Zadeh, 1973).

De fato, as relações matemáticas podem só descrever certos aspectos do mundo real, e sua utilização se deve muitas vezes à falta de outras ferramentas mais apropriadas, como explica Zadeh (1973):

“ ... as técnicas quantitativas convencionais de análise de sistemas são intrinsecamente inadequadas para delinear sistemas fortemente influenciados pelo homem ...” e afirmou, *“ ... para ser capaz de fazer inserções significativas a respeito do comportamento de sistemas humanistas, é necessário abandonar o alto rigor e precisão de nossa análise matemática e ser mais tolerante, pois tal comportamento é aproximado por natureza ...”*.

A complexidade dos sistemas anteriormente citados inclui multidimensionalidade, estruturas hierárquicas, interações mútuas (Yager e Filev, 1994). Estas características podem ser assim resumidas:

- a) os sistemas exibem um rico e variável comportamento, sendo governados por complexas relações causa-efeito;
- b) o comportamento de sistemas inteiros é, em algum sentido, mais que a soma de suas partes, o que, para alguns autores (Farmer e Packard, 1986a; Forrest, 1990) é um indicativo de não linearidade;
- c) os sistemas mencionados se desviam do comportamento previsto sobre a base de qualquer aproximação dinâmica, e o desvio varia com o tempo (Rosen, 1987).

O motivo da dificuldade em identificar sistemas com as características citadas é explicado por alguns pesquisadores, de maneira muito clara, como se descreve a seguir.

Para Ruthen (1993) o enfoque de resolução, que tenta identificar todas as partes e agentes do sistema, se não é impossível, é muitas vezes impraticável neste tipo de sistemas. Isto é devido à dificuldade de descrever todas as interações entre as diferentes partes e agentes do sistema, já que estes contêm muitas partes e agentes.

Back e Chen (1991) descrevem o problema com os seguintes argumentos:

“ ... tradicionalmente, tem-se analisado grandes sistemas do mesmo modo que os pequenos, principalmente porque os métodos desenvolvidos para Sistemas Simples têm tido sucesso. Eles tentam prever o comportamento de grandes sistemas, estudando seus elementos separadamente e analisando seus mecanismos individualmente ...” e concluem, “ ... tais sistemas contêm muitos componentes e são governados por muitas interações, o que torna impossível construir modelos matemáticos que sejam realistas e teoricamente manejáveis ... ”.

Neste sentido, para resolver o problema da identificação de sistemas, as pesquisas se direcionam ao campo da Inteligência Artificial. Assim, surge a idéia da identificação difusa de sistemas, onde um conjunto de dados de entrada-saída do sistema é utilizado para estimar o comportamento de regras lógicas, que representam um modelo do sistema a ser identificado.

1.4 Objetivos do Trabalho

Considerando a análise exposta nos itens anteriores e devido à importância da identificação de sistemas para a ciência e dos paradigmas e ferramentas utilizados na modelagem de um sistema, o objetivo principal deste trabalho é:

- 1º) desenvolver um modelo adequado para o tratamento do problema da identificação de sistemas. O modelo desenvolvido deverá fornecer um modelo formal do sistema identificado.

Para o alcance deste objetivo, o modelo proposto deverá cumprir os seguintes requisitos básicos:

- a) ser capaz de processar as entradas ambientais e armazenar o conhecimento obtido a partir da experiência. Isto significa que o modelo utiliza dados numéricos do sistema real;
- b) organizar esse conhecimento para reconhecer situações anteriores;
- c) adaptar-se aos requerimentos e mudanças ambientais e
- d) descobrir os agentes relevantes do sistema em questão, assim como as relações existentes entre os diferentes agentes do sistema.

Para cumprir os requisitos citados, o modelo baseia-se em várias idéias assim consideradas:

- A utilização de regras do tipo Se-Então para representar as relações entre os agentes do sistema e seu meio ambiente. Um modelo, formado por um conjunto de regras, representa o comportamento do sistema, como se verá no capítulo 3.
- A idéia de representar as relações entre os agentes do sistema com regras do tipo Se-Então leva a considerar a possibilidade da modelagem aproximada em lugar de exata. Assim, consideramos ferramentas de aproximação ao comportamento de sistemas, as quais são mais adequadas para serem utilizadas na adaptação às exigências ambientais e dos outros agentes do sistema.
- A utilização de processos adaptativos para formar o modelo do sistema. Isto significa que a identificação da estrutura do sistema em termos de regras pode ser realizada pelos processos adaptativos. Nesse sentido, distinguimos dois processos adaptativos diferentes:
 - a evolução, baseada nas informações genéticas obtidas dos ancestrais, onde se processam as mudanças mais gerais, a longo prazo e
 - o aprendizado, baseado na experiência, onde se processam as mudanças com metas imediatas.

As ferramentas comumente utilizadas para simular os processos adaptativos são fundamentadas principalmente na teoria evolutiva de Darwin. No entanto, biologicamente sempre foi discutido de que forma as adaptações adquiridas durante o tempo de vida do

indivíduo, por aprendizagem ou outros caminhos, são passadas a seus descendentes. Neste trabalho, consideramos estas teorias biológicas e propomos um modelo onde o aprendizado e a evolução interagem. Como consequência, um outro objetivo primordial deste trabalho é:

- 2º) incorporar a evolução e o aprendizado em um modelo, e estudar as suas interações.

Um outro objetivo deste trabalho, como consequência dos anteriores, é:

- 3º) formalizar o modelo proposto em um algoritmo que deverá implementar os requisitos citados anteriormente com as hipóteses em que se baseia o modelo proposto. O algoritmo deverá ter como parâmetro de entrada um conjunto de dados de entrada-saída do sistema real. O algoritmo será capaz de gerar uma base de regras que aproxime o comportamento dos dados de entrada-saída ao comportamento do sistema real. O algoritmo proposto é testado na identificação de várias funções e sistemas.

CAPÍTULO II

IDENTIFICAÇÃO DIFUSA DE SISTEMAS

2.1 Introdução

Como foi enunciado no capítulo anterior, os paradigmas e ferramentas utilizados na modelagem de um sistema são de fundamental importância para que o modelo obtido reflita o fenômeno observado.

Para resolver o problema da identificação de sistemas, as pesquisas direcionam-se ao campo da Inteligência Artificial, objetivando de procurar métodos que permitam implementar modelos alternativos aos métodos matemáticos, e resolver algumas das dificuldades que o construtor de modelos enfrenta, como a representação do conhecimento intuitivo e subjetivo. A meta comum desses métodos é resolver problemas que resistam aos métodos convencionais de análise e identificação de sistemas, como foi colocado no capítulo anterior. Esta família de métodos considera ferramentas que representam de forma mais adequada as características dos sistemas anteriormente mencionados no item 1.2.

Assim, algumas das ferramentas utilizadas tentam emular os especialistas humanos na execução de uma tarefa específica (Sistemas Especialistas). Outras ferramentas preocupam-se com o tratamento das incertezas e imprecisões inerentes ao raciocínio humano e dos fenômenos observados (Sistemas Difusos), enquanto outras, ainda, tentam reproduzir algumas das virtudes dos seres vivos, como a sua capacidade de aprender pela experiência (Redes Neurais) e de adaptar-se às mudanças ambientais (Algoritmos Evolutivos).

Neste capítulo são apresentadas algumas ferramentas com as características enunciadas no parágrafo anterior, e a maneira como elas têm sido utilizadas em vários trabalhos no tratamento do problema da identificação de sistemas. Os trabalhos encontrados na literatura, que estão na direção deste trabalho, são citados, e as diferenças com o modelo e o algoritmo proposto são ressaltadas.

2.2 Sistemas Especialistas

Um Sistema Especialista (Hayes-Roth, Waterman e Lenat, 1983; Buchanan e Shortliffe, 1984; Waterman, 1986) é um programa de computador que usa conhecimento sobre um domínio específico e executa uma tarefa simulando as capacidades que especialistas humanos têm de resolver problemas. Assim, a meta mais importante de um Sistema Especialista é trabalhar como um especialista humano que executa uma determinada tarefa.

Os Sistemas Especialistas representam de forma computacional os processos decisórios dos tomadores de decisões em domínios específicos do conhecimento humano. Porém, muito do conhecimento dos especialistas é heurístico, o que pode ser definido como uma estratégia, um método, uma simplificação ou uma regra de decisão, geralmente empírica, usada pelo especialista para limitar o espaço de soluções do problema (Barr e Feigenbaum, 1982). Assim, pode-se dizer que os Sistemas Especialistas ancoram seu poder nas informações que armazenam em suas bases de conhecimento. Portanto, a qualidade dos Sistemas Especialistas está diretamente vinculada à base de conhecimento e às inferências que se pode fazer sobre tais bases.

Como consequência, extrair o conhecimento dos especialistas passou a ser um dos principais alvos das pesquisas da Inteligência Artificial. Este processo de aquisição do conhecimento, quando possível, pode ser muito difícil, devido às características subjetivas e incertezas atribuídas ao conhecimento dos especialistas humanos.

Além da dificuldade na aquisição de conhecimento, os Sistemas Especialistas apresentam vários outros problemas quando são enfocados a partir da ótica da modelagem de sistemas, dentre os quais podemos citar:

- a) as incertezas da base de conhecimento, atribuídas à imprecisão do conhecimento e fatores humanos (Zadeh, 1985; Bouchon, 1985) e
- b) a base de conhecimento, uma vez obtida, permanece estática, portanto, de difícil aplicação em sistemas dinâmicos.

A dificuldade atribuída à imprecisão do conhecimento humano é tratada dentre outras ferramentas, pela Lógica Difusa, ferramenta adequada para o tratamento dos dados com subjetividade, vagueza e imprecisão dos tomadores de decisão, implementados no que chamaremos de Sistemas Difusos.

A aquisição do conhecimento dos especialistas humanos e as causas referentes às características dinâmicas do conhecimento foram tratadas com o desenvolvimento, principalmente, do aprendizado de máquina. Entre os paradigmas do aprendizado de máquina encontram-se os paradigmas evolucionário e conexionista.

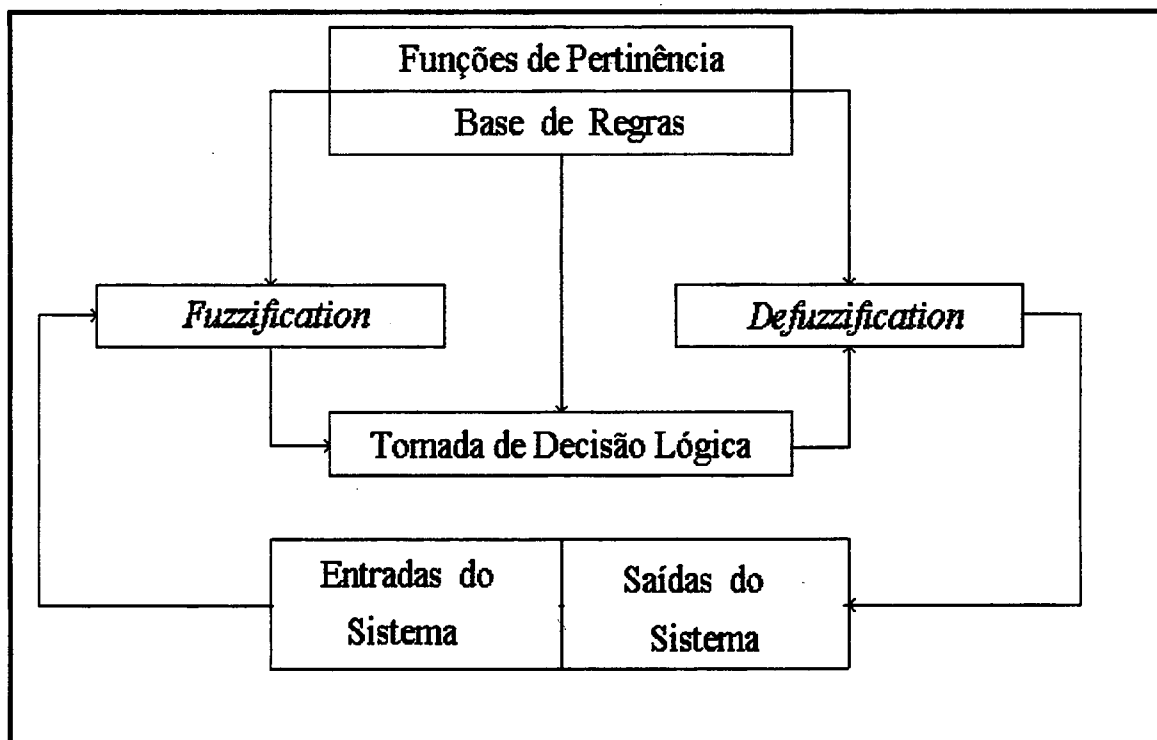
2.3 Sistemas Difusos

Os Sistemas Difusos estimam funções e sistemas com descrição parcial do comportamento do sistema, onde os especialistas “ *podem* ” prover o conhecimento heurístico, ou esse conhecimento pode ser inferido a partir de dados de entrada-saída do sistema.

Assim, os Sistemas Difusos estimam funções de entrada-saída, sem um modelo matemático, simplesmente através de como as saídas dependem das entradas. Por tal motivo, os Sistemas Difusos podem ser vistos como um sistema estímulo-resposta, onde o sistema tem uma determinada resposta, para um determinado estímulo (Kosko, 1991).

Neste sentido, pode-se dizer que os Sistemas Difusos são sistemas baseados em regras que utilizam variáveis lingüísticas difusas (Zadeh, 1975a e 1975b) para executar o processo de tomada de decisão. A figura 1 mostra o fluxo de dados através do Sistema Difuso, onde se pode observar os vários módulos que formam parte do modelo e que serão descritos a seguir.

Figura 1: Estrutura do Sistema Difuso



Fonte: Lee (1990b)

Um Sistema Difuso, como já foi enunciado, é um sistema baseado em regras, portanto um de seus módulos será a base de regras. Estas regras são do tipo Se-Então, e as variáveis do antecedente e conseqüente utilizadas são variáveis lingüísticas.

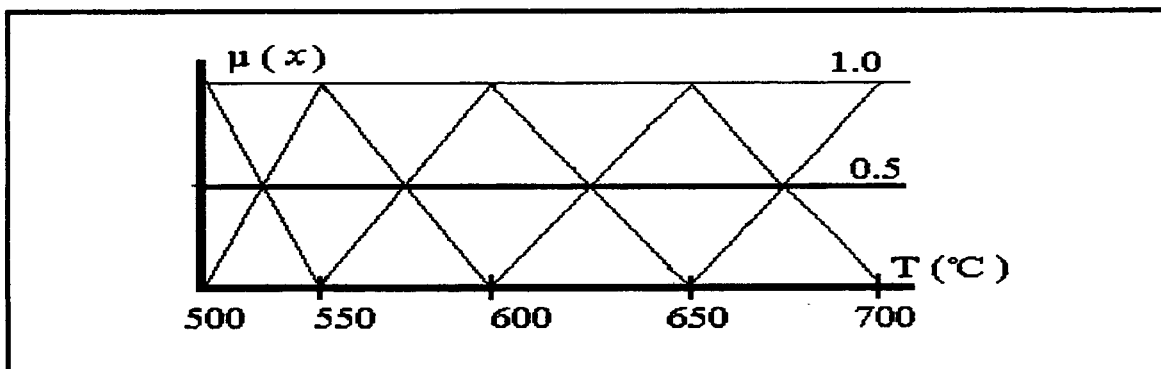
Os possíveis valores de uma variável lingüística são representados por conjuntos difusos (Zadeh, 1975a e 1975b). A forma de caracterizar um conjunto difuso é através de uma função que associa a cada elemento do domínio da variável, um valor que indica o grau de pertinência do elemento ao conjunto difuso. Tais funções são chamadas de funções de pertinência (figura 2). As funções de pertinência dos conjuntos difusos, assim como as regras do sistema, são definidas, *a priori*, pelo especialista humano.

O processo de *fuzzification* (Lee, 1990a) que utiliza as funções de pertinência predefinidas, mapeia cada variável de entrada do sistema em graus de pertinência de algum conjunto difuso que representa a variável em questão.

A Tomada de Decisão Lógica é de vital importância no Sistema Difuso, já que nela

se realiza o processo de inferência para obter, a partir das regras do sistema e da entrada do ambiente, uma ação a ser realizada pelo sistema. A Tomada de Decisão Lógica é realizada através da chamada regra de inferência composicional (Zadeh, 1975c; Mandami, 1977).

Figura 2: Funções de Pertinência Pertencentes à Variável Lingüística Temperatura do Secador da Massa Cerâmica



Fonte: Tcholakian (1992)

Assim, no processo de inferência, temos três passos (Yager e Filev, 1994):

- 1º) Encontrar o nível de contribuição de cada uma das regras. O grau de contribuição é dado pelo valor de pertinência obtido da interseção entre o valor lingüístico do antecedente das regras e o valor de entrada obtido do ambiente. Por exemplo, na regra:

$$\text{Se } x_1 = A_i \text{ e } x_2 = B_i \text{ Então } y = C_i$$

o grau de contribuição é dado por:

$$R_i(x) = A_i(x_1) \wedge B_i(x_2), \quad (1)$$

onde \wedge representa a interseção, modelada por um determinado operador, por exemplo, o operador Min. Em (Mizumoto e Zimmermann, 1982), descrevem-se e comparam-se vários operadores que podem ser utilizados no processo de inferência.

- 2º) A saída de cada regra individual é dada por :

$$F_i(y) = (R_i(x) \wedge C_i(y)) \quad (2)$$

- 3º) A agregação das saídas das regras individuais para obter um subconjunto difuso de F sobre y é

$$F(y) = \bigvee_i F_i(y) = \bigvee_i (R_i(x) \wedge C_i(y)), \quad (3)$$

onde \bigvee representa a união, modelada por outro operador, como por exemplo o operador Max.

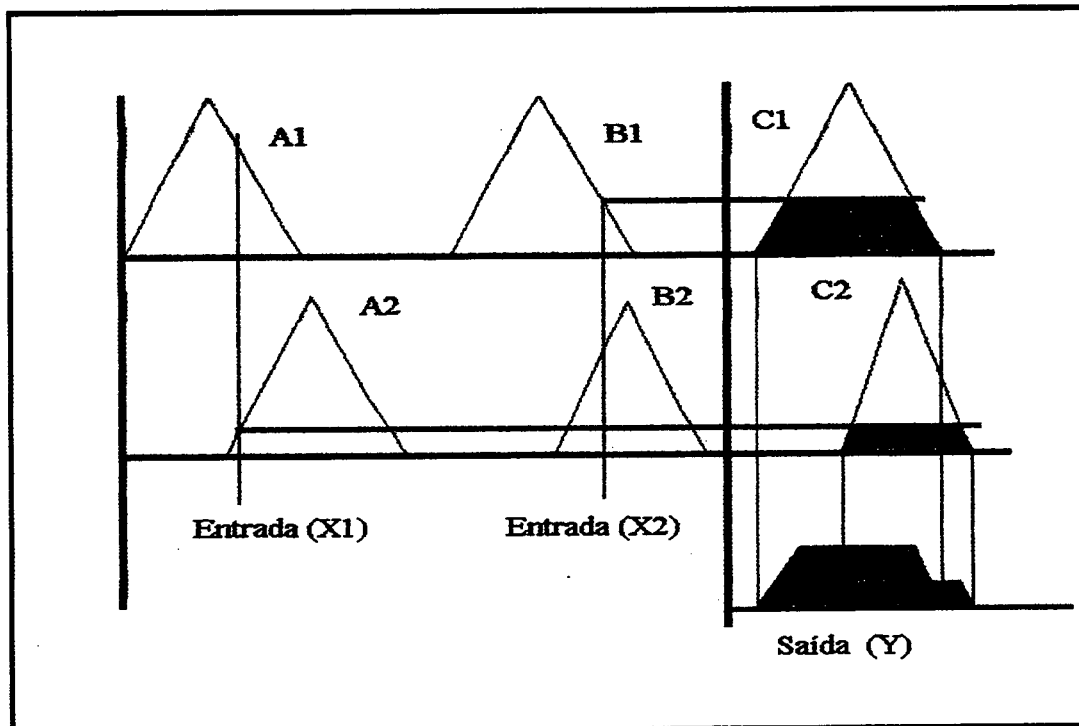
Na figura 3, temos um exemplo do método de inferência Max-Mín, com duas variáveis de entrada e uma de saída, e as seguintes regras:

Se $x_1 = A_1$ e $x_2 = B_1$ Então $y = C_1$

Se $x_1 = A_2$ e $x_2 = B_2$ Então $y = C_2$

Para terminar de descrever os módulos que formam o Sistema Difuso, temos o processo de *defuzzification* que calcula a saída com base na inferência obtida no módulo de tomada de decisão lógica, com as funções de pertinência das variáveis linguísticas da parte consequente das regras (Lee, 1990b), para obter uma saída não difusa.

Figura 3: Processo de Inferência Max-Mín



Fonte: do Autor

Os Sistemas Difusos têm sido utilizados com grande êxito no desenvolvimento de sistemas de controle e alguns exemplos são dados em Kickert e Van Nauta Lanke (1976),

Pappis e Mandami (1977), Sasaki e Akiyama (1988), Yu, Cao e Kandel (1990) e Tcholakian (1992). Os Sistemas Difusos foram propostos basicamente para resolver problemas como a vagueza, subjetividade e imprecisão do conhecimento dos especialistas.

No entanto, os Sistemas Difusos, apesar de solucionarem alguns dos problemas citados, atribuídos aos Sistemas Especialistas, também trazem outros problemas, como a divisão do espaço das variáveis do sistema em subespaços difusos, a escolha dos valores lingüísticos representados por conjuntos difusos, e as funções de pertinência que caracterizam os conjuntos difusos. Todos estes itens são projetados com ajuda do especialista humano, o que, em vários casos, é muito difícil de obter. Além disso, a obtenção da base de regras continua sendo fundamentada nos especialistas humanos.

2.4 Identificação Difusa de Sistemas

Existem várias situações nas quais podemos derivar um Modelo Difuso de um fenômeno observado. Segundo Dubois e Prade (1980), estas situações podem ser:

- a) as descrições lingüísticas, obtidas de um especialista humano, que refletem o conhecimento qualitativo de um processo e permitem construir um conjunto de regras lógicas difusas, como, por exemplo, em Tcholakian (1992);
- b) o caso em que temos equações conhecidas que descrevem o comportamento de um processo, mas os parâmetros não podem ser precisamente identificados e, portanto, são interpretados de uma forma difusa;
- c) equações conhecidas para o processo são demasiado complexas e são interpretadas em um caminho difuso para construir um modelo;
- d) os dados da entrada e saída são utilizados para estimar o comportamento de regras lógicas difusas. Este procedimento é conhecido como identificação difusa de sistemas, e consiste basicamente de duas partes:

– a identificação da estrutura do modelo e

- a identificação dos parâmetros do modelo.

Os passos mais importantes do processo de identificação difusa podem ser descritos como segue:

- 1º) determinar as variáveis que constituem os antecedentes e os conseqüentes das regras;
- 2º) determinar as funções de pertinência dos conjuntos difusos, isto é, a partição do espaço das variáveis do antecedente e das variáveis conseqüentes em subespaços difusos;
- 3º) gerar as relações entre as variáveis do antecedente e as variáveis do conseqüente das regras, a partir de dados de entrada-saída.

Os primeiros trabalhos nesta direção encontram-se em Takagi e Sugeno (1985), Sugeno e Kang (1988), e Sugeno e Tanaka (1991). Em todos estes trabalhos utilizam-se regras lógicas, cuja parte antecedente são variáveis difusas, e a parte conseqüente, uma combinação linear das variáveis do antecedente.

O algoritmo proposto por Takagi e Sugeno (1985), basicamente, é um algoritmo iterativo que podemos resumir nos seguintes passos.

- 1º) Considerando um sistema de múltiplas entradas e uma saída, por exemplo k variáveis de entrada, formamos k modelos compostos cada um por uma única variável, dividindo o espaço de entrada em dois valores lingüísticos, e calculamos as funções de pertinência através de um problema de programação não linear.
- 2º) O conseqüente é uma combinação linear das variáveis da premissa, e os parâmetros são calculados utilizando métodos estatísticos como o de mínimos quadrados.
- 3º) Os k modelos são avaliados, utilizando como índice de desempenho a diferença quadrática entre os valores de saída obtidos pelo modelo proposto e os dados de saída reais do sistema.

- 4º) Pegando o melhor modelo, por exemplo, o modelo formado pela variável x_i , novamente montamos k modelos, agora com a variável do modelo anterior considerado melhor, mais uma variável dividida em dois valores lingüísticos. No caso do modelo i , a variável x_i é dividida em quatro valores lingüísticos.
- 5º) Continua-se com o mesmo procedimento, até que o índice de desempenho fique estável.

Várias críticas podem ser feitas ao algoritmo anteriormente descrito. As regras deste tipo representam corretamente o comportamento global do sistema, no entanto, o comportamento local não é descrito corretamente, (Yen e Gillespie, 1995).

Outro problema está na forma de calcular as variáveis do antecedente das regras, o que é um problema evidentemente combinatório, e limita muito sua aplicação. Um terceiro problema é o algoritmo que calcula as funções de pertinência das variáveis antecedentes, que foi reduzido a um problema de programação matemática não linear.

Outros trabalhos, que tratam da identificação difusa de sistemas, preocupam-se com a geração das regras difusas (Wang e Mendel, 1992). No trabalho de Wang e Mendel, as funções de pertinência são predefinidas, como sendo todas iguais e simétricas, isto é, o domínio de cada variável é dividido em espaços de entrada-saída iguais. O algoritmo consiste basicamente em:

- a) dividir o espaço de entrada-saída em regiões difusas;
- b) gerar as regras difusas desde dados de entrada-saída do sistema;
- c) proporcionar um grau a cada regra, para posteriormente resolver possíveis conflitos de regras;
- d) formar uma base de regras difusa;
- e) utilizar um método de *defuzzification* para obter uma saída a partir da base de regra formada.

As críticas a este algoritmo são a simetria das funções de pertinência e a divisão do

domínio de cada variável em espaços de entrada-saída iguais. O problema de proporcionar graus às regras é resolvido por outros autores por redes neuronais, como, por exemplo, em Kosko (1991), ou utilizando o modelo de Dempster-Shafer, como é citado em Yager e Filev (1994).

No trabalho de Delgado e Gonzalez (1993), trata-se do problema da identificação dos parâmetros do sistema, isto é, encontrar os valores antecedentes e conseqüentes de cada regra. Já outros trabalhos, por exemplo, Higgins e Goodman (1994), utilizam redes neuronais como método de aprendizado de funções de pertinência e regras.

2.5 Algoritmos Genéticos

Desde a década de 60, alguns cientistas em computação começavam a pesquisar sistemas evolucionários, com a idéia de que a evolução poderia ser utilizada como ferramenta de otimização para resolver problemas de engenharia e problemas matemáticos. Em Golberg (1989), encontramos uma breve história da evolução de tais sistemas e os primeiros trabalhos nesta área. Mas, os Algoritmos Genéticos como hoje são conhecidos, foram propostos por John Holland em 1975 (Holland, 1992).¹

Os Algoritmos Genéticos (Goldberg, 1989; Davis, 1991; Rawlins, 1991; Beasley, Bull e Martin, 1993a e 1993b; Forrest, 1996) são uma técnica de busca baseada na teoria evolucionária. Basicamente, os Algoritmos Genéticos diferem de outras técnicas de busca nos seguintes pontos:

- a) são independentes do domínio do problema;
- b) trabalham com *string* de caracteres, geralmente em código binário, para representar um conjunto de parâmetros. A representação dos indivíduos também pode ser feita com valores inteiros ou reais;
- c) usam regras probabilistas para guiar a busca, e não regras deterministas;
- d) são um método indutivo, e não dedutivo, porque procuram soluções através da justaposição de hipóteses;

- e) são relativamente imunes à alta dimensionalidade, mínimos locais e a ruídos;
- f) só necessitam de informação concernente à qualidade da solução produzida por cada conjunto de parâmetros (informação da função objetivo) e não requerem, como nos métodos de otimização, informação derivada, ou completo conhecimento dos parâmetros e da estrutura do problema;
- g) consideram uma população de pontos, e não um só. Aqui, encontra-se implícito o paralelismo dos algoritmos genéticos, já que esses pontos vão evoluir juntos, partilhar e disputar recursos, e aqueles que melhor se adaptarem aos requisitos ambientais gerarão descendentes.

Nos Algoritmos Genéticos, cada ponto do espaço de solução é considerado como um cromossomo ou um indivíduo. O conjunto de cromossomos forma o que é chamado de população. Os indivíduos da população evoluem, de geração em geração, por meio de operações entre os cromossomos, como cruzamento e mutação. O ciclo básico dos Algoritmos Genéticos está mostrado na figura 4.

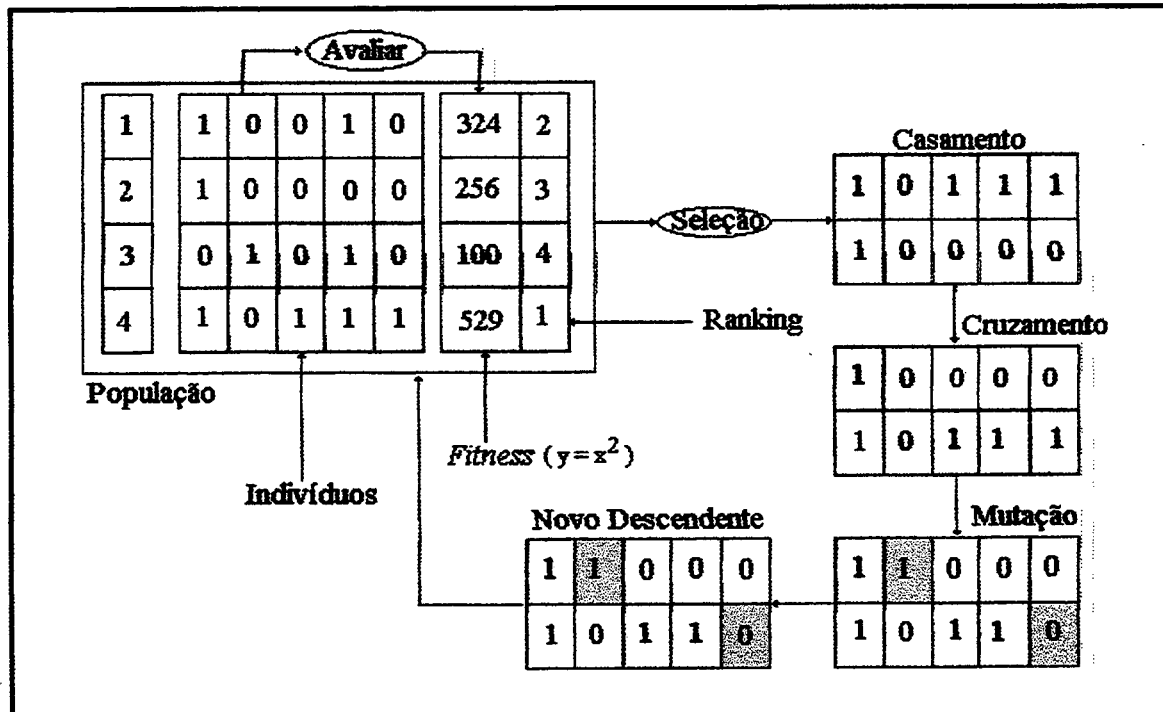
A obtenção da população da geração $t + 1$, a partir da população da geração t , é realizada da seguinte maneira:

- 1º) Uma vez representados os indivíduos, calcula-se o valor que representa a capacidade do cromossomo, com respeito ao problema a resolver, isto é, a adequação do indivíduo ao meio ambiente. Esta medida de capacidade é chamada de *fitness*. Esta medida é importante, porque cromossomos com *fitnesses* baixos, terão menos possibilidades no processo evolutivo. Assim, a escolha dos cromossomos mais adequados está em função do *fitness*, e é realizada pelo operador de seleção.
- 2º) Os cromossomos são selecionados de forma aleatória: aqueles mais aptos às exigências ambientais têm maior probabilidade de serem selecionados. O cruzamento é realizado entre dois cromossomos. A decisão se os cromossomos vão ser cruzados ou não está determinada por uma probabilidade de cruzamento. Quando o cruzamento é realizado, uma posição física do cromossomo é determinada, de

¹ A referência bibliográfica corresponde à segunda edição do livro.

forma aleatória, como ponto de corte do cromossomo. Assim, os dois cromossomos intercambiam suas partes cortadas, formando dois descendentes. Na figura 4, vê-se como é realizada esta operação.

Figura 4: Ciclo Básico dos Algoritmos Genéticos



Fonte: Adaptado de Sudjianto, Wasserman e Sudabo (1996)

3º) Para os descendentes obtidos, pode-se aplicar a operação de mutação, em função de uma probabilidade de mutação. Esta operação é aplicada a cada posição física do cromossomo. Em cromossomos codificados em binário, o operador muda zeros por uns e vice-versa (figura 4). Assim, são gerados novos indivíduos, que vão passar para a próxima geração.

Os Algoritmos Genéticos têm sido utilizados em uma variedade de problemas, como problemas de otimização (Goldberg e Kuo, 1987), otimização multimodal (Golberg e Deb, 1989; Tcholakian, 1994a), otimização multiobjetivo (Fonseca e Fleming, 1993), problemas de otimização combinatória (Khurl, Bäck e Heitkötter, 1993), tais como o problema do caixeiro viajante (Homaifar, Guan e Liepins, 1993), problema de fluxo máximo (Munakata e Hashier, 1993), *Job shop scheduling* (Rahmani e Ono, 1993; Bruns, 1993; Kidwell, 1993).

Também têm sido utilizados na identificação difusa de sistemas como será descrito no próximo item.

2.6 Os Algoritmos Genéticos na Identificação Difusa de Sistemas

Devido às vantagens dos Algoritmos Genéticos descritas anteriormente, alguns trabalhos têm surgido para tentar resolver o problema da identificação difusa de sistemas. Em Cordón e Herrera (1996a e 1996b) distinguem-se três grupos de trabalho na identificação difusa de sistemas utilizando algoritmos genéticos:

- a) os que descobrem a Base de Dados, isto é, os parâmetros do sistema, os limites das variáveis de entrada-saída do sistema, as funções de pertinência das variáveis do sistema;
- b) os que descobrem a Base de Regras, isto é, as regras de comportamento do sistema;
- c) por último, os que descobrem a Base de Conhecimento, que agrupa a Base de Dados e a Base de Regras.

Os trabalhos encontrados na literatura, têm utilizado os Algoritmos Genéticos para melhorar a eficiência dos Sistemas Difusos, descobrir funções de pertinência e as regras do sistema. Estes trabalhos são de autoria de Karr (1991a, 1991b, 1992, 1996), Espy, Vombrack e Aldridge (1992), Tcholakian (1993 e 1994b), Yen e Gillespe (1995), Vergara e Moraga (1995), Herrera, Lozano e Verdagay (1995), Cordón, Herrera, Herrera-Viedma e Lozano (1995), Cordón e Herrera (1996a e 1996b).

Os trabalhos de Karr (1991a, 1991b, 1992, 1996), na modelagem de um sistema de controle, consistem de um modelo que consta de três módulos: um módulo de controle, que consiste de um controlador lógico difuso; um módulo de análise, que descobre as alterações que se produzem no ambiente; e um módulo de aprendizado que, baseado em algoritmos genéticos, altera as funções de pertinência. Karr (1991a) reconhece a dificuldade de selecionar funções de pertinência para que os controladores sejam eficientes. O algoritmo genético apresentado é baseado em código binário e calcula as funções de pertinência do sistema de controle. Em Karr (1991a, 1991b e 1992), o modelo é aplicado sobre um sistema de pH; em Karr (1996), o modelo é aplicado a um sistema caótico.

Em Espy, Vombrack e Aldridge (1992), utilizam-se os algoritmos genéticos para calcular funções de pertinência, baseadas em uma codificação dos cromossomos da população como valores inteiros. Em Tcholakian (1993 e 1994b), também são utilizados os algoritmos genéticos baseados em uma codificação inteira, para calcular funções de pertinência de um sistema difuso baseado em regras. Através do algoritmo genético, são obtidas as funções de pertinência das variáveis difusas que produzem os melhores resultados quando comparado o valor do sistema real com o valor do modelo difuso.

O trabalho de Yen e Gillespe (1995) pode ser destacado pela utilização de uma função *fitness* que combina uma medida de *fitness* local e uma medida de *fitness* global. Esta proposta foi fundamentada na afirmação de que as regras do tipo Se-Então propostas por Takagi e Sugeno (1985) não descrevem corretamente o comportamento local.

Os trabalhos citados acima utilizam o algoritmo genético apenas como método de busca dos parâmetros das funções de pertinência; estes trabalhos podem ser incluídos no primeiro grupo na classificação realizada por Cordón e Herrera (1996a e 1996b). Nestes exemplos, as regras difusas do sistema já são conhecidas, o que estabelece uma diferença essencial com o trabalho aqui apresentado, onde as regras que representam o modelo do sistema deverão ser descobertas.

Os trabalhos de Vergara e Moraga (1995), Herrera, Lozano e Verdagay (1995), Cordón, Herrera, Herrera-Viedma e Lozano (1995), Cordón e Herrera (1996a e 1996b), utilizam um algoritmo dividido em três processos evolutivos diferentes:

- 1º) Gerar um conjunto de regras desde dados de entrada-saída, por meio de um processo evolutivo, baseados em uma partição difusa preestabelecida. O Algoritmo Genético aqui utilizado é codificado com valores reais, e um indivíduo representa uma regra. O algoritmo consiste em um processo iterativo, que permite que um conjunto de regras difusas seja obtido para cobrir o conjunto de exemplo de dados de entrada-saída.
- 2º) Um processo evolutivo para simplificar a Base de Regras. Aqui, utiliza-se um outro

Algoritmo Genético binário, onde um indivíduo representa uma Base de Regras completa. O algoritmo consiste em retirar da Base de Regras possíveis sobreposições de regras.

- 3º) Um processo evolutivo que ajusta as funções de pertinência das regras difusas, baseadas na Base de Regras já descoberta. Aqui utiliza-se um Algoritmo Genético de código real. Este algoritmo pode ser incluído no primeiro grupo da classificação de Cordón e Herrera (1996a e 1996b).

Em outro trabalho, Valenzuela-Rendón (1991), propõem os Sistemas Classificadores Difusos, utilizando o potencial dos Sistemas Classificadores como paradigma de aprendizado, e o potencial da Lógica Difusa como representação da incerteza e do raciocínio aproximado. Sistemas Classificadores representam um aprendizado de máquina, baseado no princípio adaptativo extraído da genética de populações biológicas e da economia (Booker, Golberg e Holland, 1989; Holland, Holyoak, Nisbett e Thagand, 1993).

Em Parodi e Bonelli (1993), encontra-se um Sistema Classificador Difuso que utiliza um esquema de distribuição de créditos para as regras, diferente do *Bucket Brigade*, baseado em aprendizado supervisionado. O *Bucket Brigade* é o algoritmo de avaliação das regras utilizado pelos Sistemas Classificadores (Booker, Golberg e Holland, 1989; Holland, Holyoak, Nisbett e Thagand, 1993).

No modelo de Parodi e Bonelli (1993), os cromossomos do Algoritmo Genético representam uma regra do tipo Se-Então, onde cada variável do sistema é representada por dois parâmetros que formam uma função de pertinência. Cada variável tem associado um conjunto difuso, isto é, cada variável é descrita por uma função de pertinência. Cada classificador contém a descrição das funções de pertinência que correspondem a cada variável de entrada-saída, a qual consiste de parâmetros que definem os conjuntos difusos associados. O problema com esta abordagem é o fato de que a solução vai depender do número de regras que estão diretamente associadas aos indivíduos da população do Algoritmo Genético.

Uma diferença substancial dos trabalhos citados com o proposto aqui está na distinção que fazemos entre os processos adaptativos e a função de cada processo na identificação do modelo do sistema, assim como a interação entre a evolução e o aprendizado. Outra diferença a ser destacada é que os trabalhos anteriormente citados não fornecem um modelo descritivo do sistema identificado.

Outra diferença é que em nossa representação do indivíduo incluímos as funções de pertinência e as regras do sistema juntos. Esta e a distinção entre processos adaptativos são as diferenças essenciais do trabalho proposto com os trabalhos do grupo da Universidade de Granada (Vergara e Moraga, 1995; Herrera, Lozano e Verdagay, 1995; Cordón, Herrera, Herrera-Viedma e Lozano, 1995; Cordón e Herrera, 1996a e 1996b). Enquanto estes trabalhos utilizam três processos evolutivos, para identificar a estrutura e parâmetros do sistema, neste trabalho utilizamos um único processo. Isto implica diferenças na estrutura do modelo, na representação dos indivíduos e suas regras, e na avaliação das estruturas e parâmetros do sistema.

Com respeito aos trabalhos de Valenzuela-Rendón (1991) e Parodi e Bonelli (1993), além das diferenças enunciadas nos parágrafos anteriores, pode ser acrescentada a construção da base de regras que representa o modelo do sistema.

2.7 Processos Adaptativos: Evolução e Aprendizado

Visualizando um sistema como um conjunto de agentes que se adaptam às exigências ambientais e dos demais agentes do sistema, pode-se afirmar que os sistemas têm características dinâmicas. As características dinâmicas destes sistemas surgem devido ao fato de que qualquer agente que interage em um ambiente que sofre mudanças não pode antecipar todas as possíveis situações, mas pode e deve ser capaz de efetuar uma adaptação em suas estratégias, para se ajustar às mudanças desse ambiente.

É difícil apresentar uma definição rigorosa de adaptação. Para alguns pesquisadores (Farmer e Packard, 1986b; Forrest, 1990), o comportamento adaptativo é uma propriedade emergente, conhecida como comportamento emergente, a qual aparece a partir da interação

de simples componentes. Podem-se distinguir basicamente dois processos adaptativos, com diferentes características e funções sobre os agentes do sistema: a evolução e o aprendizado.

As principais hipóteses sobre a teoria da evolução, discutidas até hoje, começaram a surgir na primeira metade do século XIX. Foi Jean Baptista Lamarck que apontou a adaptação dos indivíduos ao ambiente como o problema central da evolução.

De acordo com Lamarck, os órgãos que desenvolvem intensa atividade crescem e se tornam mais eficazes, e aqueles pouco utilizados se atrofiam e degeneram (lei do uso e desuso). Estas alterações poderiam ser transmitidas aos descendentes (herança de caracteres adquiridos). Ao longo de muitas gerações, o acúmulo dessas alterações poderia levar ao aparecimento de novas espécies (Soncini, 1993).

Portanto, para Lamarck, as diferentes espécies de animais tinham desenvolvido exatamente aquilo de que precisavam, e a herança das características adquiridas durante o tempo de vida de um indivíduo foi a principal causa da evolução. Isto é, as características adquiridas pelo indivíduo por meio de seu próprio esforço eram herdadas depois por seus descendentes.

Quando Charles Darwin iniciou sua viagem de exploração científica a partir da qual elaborou as principais idéias de sua teoria científica, a idéia de evolução de Lamarck já era conhecida. Apesar das vastas argumentações em contrário, muitos estudiosos faziam sua defesa. Assim, mais do que comprovar que a evolução ocorre, o empenho de Darwin recaiu em explicar como ela se processa e, no centro de suas preocupações, estava a adaptação dos indivíduos (Soncini, 1993).

Darwin apontou dois fatores como determinantes do processo evolutivo: a seleção natural e as variações hereditárias. A seleção natural é o mecanismo da evolução biológica, que faz com que as espécies melhor adaptadas ao ambiente tendem a sobreviver e deixar descendência.

Em qualquer população ocorrem variações hereditárias. Algumas delas tornam seus

portadores mais aptos a explorar seu ambiente. Estes sobrevivem e se reproduzem deixando uma descendência portadora dessas variações. Por outro lado, indivíduos portadores de variações que os tornam pouco hábeis a explorar seu ambiente, acabam por desaparecer, não conseguindo, portanto, deixar descendência (Soncini, 1993).

Assim, as variações hereditárias que ocorrem em uma população podem ser vantajosas ou não para os indivíduos que as experimentam. Aqueles indivíduos mais aptos às exigências ambientais sobrevivem e se reproduzem, deixando descendentes, e os menos aptos acabam desaparecendo. Com o passar das gerações, os indivíduos melhores dominam a espécie. É desta maneira que age a seleção natural, atuando sobre as variações hereditárias, eliminando os indivíduos portadores de variações que dificultam ou impedem a sobrevivência e mantendo aqueles indivíduos cujas variações permitem melhor explorar o ambiente (Soncini, 1993). A grande contribuição deixada por Darwin foi a idéia da seleção natural atuando sobre as variações hereditárias, como mecanismo fundamental da evolução.

Pode-se resumir a diferença básica entre as idéias de Darwin e de Lamarck, da seguinte maneira: para Lamarck, existia um mapeamento de condições ambientais e fenótipos que determinam o genótipo, e para Darwin, as condições ambientais e o genótipo determinavam o fenótipo. Entende-se por fenótipo as características físicas e morfológicas de um indivíduo, e por genótipo o conjunto de genes de um indivíduo.

Mas, tanto Lamarck como Darwin, deixaram muitas perguntas ainda sem respostas. Entre estas perguntas, podemos apontar: qual é a causa dessas variações hereditárias, e como elas ocorrem ?

Assim, o início do século XX foi marcado pela polêmica entre os neolamarckistas e neodarwinistas, ambos tentando solucionar os pontos obscuros da teoria da evolução.

Os neolamarckistas negavam a seleção natural e apontavam, como pilar da evolução, os efeitos hereditários do uso e desuso, e os efeitos hereditários da ação direta do meio sobre os organismos. Os neodarwinistas rejeitavam tudo da teoria de Darwin, menos a seleção natural, a qual é considerada como o mecanismo central da evolução (Soncini, 1993).

Foi a genética que mostrou a ausência de base do neolamarckismo, ao explicar que as características dos indivíduos são determinadas por genes e que os genes, e não os caracteres, é que são transmitidos de uma geração a outra. Portanto, caracteres adquiridos não são herdados.

A genética também apontou que o código genético de uma população pode mudar devido a três fatores principais: as migrações, a recombinação genética e as mutações (Soncini, 1993).

As migrações introduzem novos genes numa população, na medida em que os migrantes levam consigo seu patrimônio genético e, ao se reproduzirem, transferem-no aos seus descendentes. Desta forma, novas variações de características surgem na população.

Já a recombinação genética é consequência da reprodução sexuada. Embora não introduza novos genes na população, garante novas combinações genéticas nos indivíduos que compõem essa população.

As mutações, fator mais importante de alterações no código genético, produzem variações nos genes que podem ser transmitidos à geração seguinte, isto é, as mutações são capazes de alterar as características hereditárias de um organismo. As mutações parecem ocorrer sem causa aparente, mas o fato é que surgem novos genes nessa população e, com eles, variações de características.

Para a teoria evolucionista, estas pequenas alterações, se consideradas em grandes intervalos de tempo, são capazes de causar profundas transformações em uma determinada espécie.

Biologicamente, as mutações são mudanças no código genético de um ser vivo e responsáveis pelo surgimento de novas características que podem ou não ser benéficas. O que define se uma mutação é vantajosa ou não a um indivíduo, ou a sua espécie, é a interação com o meio ambiente. Estas diferenças dão aos indivíduos vantagens ou desvantagens sobre os outros.

Desta forma, a genética confirmou que o desenvolvimento de um indivíduo, desde seu nascimento até sua morte, é um mapeamento do genótipo e o ambiente para o fenótipo. O mapeamento inverso é impossível biologicamente. No entanto, uma idéia que sempre tem sido atrativa é que adaptações adquiridas durante o tempo de vida do indivíduo, por aprendizado ou outros caminhos, são passadas a seus descendentes (Smith, 1987).

Uma das perguntas que se deve responder é como esse aprendizado ou experiências do indivíduo são passadas para seus descendentes. Em Smith (1987), afirma-se que com o desenvolvimento da biologia molecular, temos como rejeitar a idéia de que as adaptações individuais podem alterar a informação genética. No entanto, é possível para um indivíduo aprender para facilitar a evolução. A idéia é que o aprendizado auxilia a sobrevivência e, então, os organismos poderiam variar geneticamente em sua capacidade de aprender, incrementando a frequência dos genes responsáveis pelo aprendizado.

Definir o termo aprendizado é muito difícil. No entanto, um denominador comum de muitos sistemas de aprendizado é a capacidade que possui o indivíduo de fazer mudanças ao longo do tempo, com a intenção de melhorar o desempenho de tarefas definidas por seu ambiente. Isto permite que o desempenho dos agentes no sistema melhore.

As idéias de que o aprendizado poderia auxiliar a evolução surgiram em 1896, e foram propostas por James Mark Baldwin (1896). Estas idéias, subseqüentemente conhecidas como efeito Baldwin, foram consideradas como um “novo fator” da teoria evolutiva.

Baldwin considerava que a aprendizagem individual pode explicar fenômenos evolucionários que parecem requer herança lamarckista. Este autor explica que a aprendizagem inicialmente facilita a evolução de comportamentos complexos, posteriormente, os comportamentos aprendidos são substituídos pelo comportamento instintivo. Assim, comportamentos aprendidos poderiam ser comportamentos instintivos em gerações subseqüentes, sem apelar à herança lamarckista. Portanto, o efeito Baldwin não requer o mapeamento de fenótipo e ambiente em genótipo, como requer a teoria de Lamarck, o que biologicamente não é aceitável.

Depois de quase meio século, em que as discussões sobre as causas das variações hereditárias pareciam ter sido encerradas, o trabalho de Cairns e colegas (Cairns, Overbaugh e Miller, 1988) reabriu tais discussões (Goodman, 1992).

O trabalho de Cairns sugere que algumas bactérias têm a capacidade de “*escolher*” mutações vantajosas para se adaptarem ao ambiente. A hipótese de Cairns é uma heresia em termos darwinianos, já que entra em aparente contradição com a idéia convencional de que as mutações são aleatórias. Esta hipótese lembra Lamarck. As bactérias “*lamarckistas*” de Cairns desenvolveram mutações para poder aproveitar um tipo de açúcar abundante no ambiente.

Este experimento levou Cairns a afirmar : “ ... *o experimento sugere que populações de bactérias tenham algum modo de produzir (ou reter seletivamente) apenas as mutações mais apropriadas ...* ”. Assim, as mutações vantajosas só ocorreram depois que as bactérias tiveram contato com o meio, “*aprendendo*” a comer açúcar. Estas afirmações contradizem a idéia de que as mutações são totalmente aleatórias; ou como indica Lenski e Mittler (1993), “ ... *alguns experimentos sugerem que as células poderiam ter mecanismos para escolher qual mutação ocorre ...* ”.

O que Cairns e seus colegas afirmam é que as mutações são mais frequentes quando o organismo está sob pressão seletiva para que a mutação apareça. Quando populações de simples células estão sujeitas a certas formas de pressão seletiva, emergem variantes que mudam o DNA, que poderiam trazer apropriadas mudanças no fenótipo. Assim, o ambiente não só seleciona entre variantes preexistentes, como na visão convencional da evolução (Thaler, 1994), mas também interatua com o organismo para gerar a variação na qual a seleção atua. Desta forma, o genótipo original da população poderia ser suplementado pelas variantes.

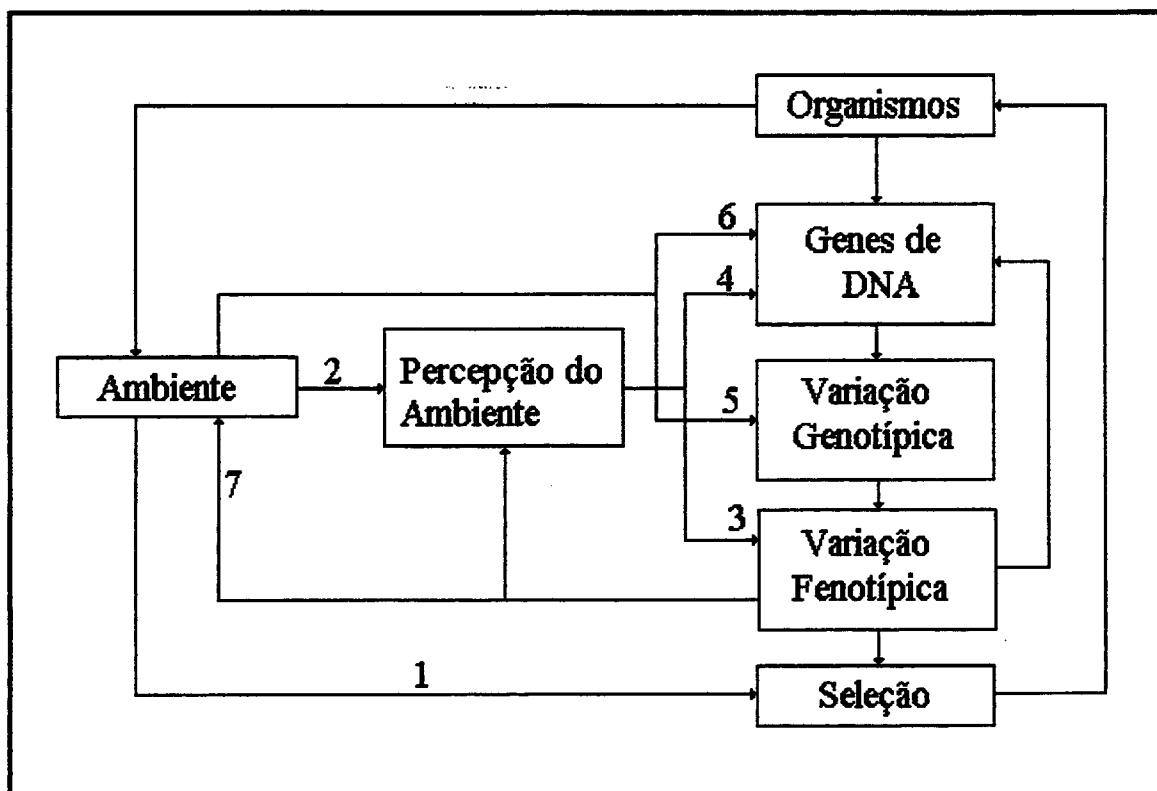
Este fenômeno de mutação sob seleção é chamado de “*seleção induzida*” ou de “*mutação direcionada*”. A idéia deste tipo de mutação sugere que ocorre quando o resultado fenotípico é vantajoso; portanto deve existir uma relação muito específica entre mutação e ambiente. Mutação direcionada requer que as mutações vantajosas ocorram para responder

especificamente a condições seletivas particulares (Lenski e Mittler, 1993).

O problema que existe é que, biologicamente, não se conhece um vínculo genético claro, uma relação causa-efeito entre pôr uma bactéria em um meio com açúcar e o surgimento das mutações que permitem que ela utilize esse alimento. Uma explicação neodarwinista para este tipo de mutação é que algumas mutações são favorecidas quando estão sob pressão seletiva. Isto é, a probabilidade de ocorrer uma mutação em certos indivíduos que estão sob pressão seletiva é maior que em aqueles progenitores onde não existe esta pressão (Lenski e Mittler, 1993).

Baseado nas idéias de mutação direcionada, Thaler (1994), apresenta uma visão mais completa de evolução, onde os conceitos acima expostos são considerados e ampliados (figura 5). Thaler (1994), afirma que a seleção natural atua em alelos particulares e que o metabolismo genético gera alelos com uma alta probabilidade de passar os testes de seleção ambiental. Se o ambiente influi afetando a geração de variações, pode subseqüentemente afetar a seleção.

Figura 5: Uma Visão mais Completa de Evolução



Fonte: Thaler (1994)

Os itens enumerados na figura 5 podem ser interpretados da seguinte maneira:

- (1) O ambiente tem influência na seleção.
- (2) O ambiente é percebido pelo indivíduo.
- (3) Os indivíduos utilizam essa percepção para modificar sua fisionomia (fenótipo).
- (4) Os organismos utilizam a percepção do ambiente para modificar seu metabolismo genético.
- (5) O ambiente atua nas variações genotípicas.
- (6) O ambiente atua diretamente no DNA dos genes.
- (7) O organismo modifica sua interação com o ambiente com seu novo fenótipo.

Em resumo, pode-se afirmar que adaptação envolve um processo de otimização no qual um sistema “aprende” através de sua experiência, em uma escala de tempo rápida, onde acontecem os problemas locais, e “evoluem” em uma escala de tempo mais lenta, onde se processam as mudanças mais gerais, não propriamente correspondentes ao agente como um indivíduo, mas sim à espécie do indivíduo. Por exemplo, em um jogo, as jogadas acontecem em uma escala de tempo rápida, o ajuste das estratégias aparece e acontece em uma escala de tempo lenta.

Portanto, os agentes se adaptam em dois níveis: individual e espécies. Adaptação individual é considerada uma forma de aprendizado, e a adaptação de espécies é considerada uma forma de evolução. Ou, de outra forma, entendemos por aprendizado as mudanças que um agente sofre em seu tempo de vida, e por evolução, as mudanças sofridas pela espécie no tempo de existência da espécie.

Para que os termos aprendizado e evolução fiquem mais claros, podem ser citadas as suas diferenças expostas anteriormente como:

- a) Ambos os processos operam em diferentes escalas de tempo: o aprendizado no tempo de vida de um indivíduo, a evolução sobre a espécie do indivíduo.
- b) aprendizado está baseado na experiência do indivíduo e a evolução na adaptação deste aos requerimentos do ambiente.
- c) aprendizado realiza um tipo de busca local, com metas imediatas, a curto prazo. A evolução realiza um tipo de busca global, a longo prazo, já que sua meta é

assegurar a vida a seus descendentes.

Recentemente, vários pesquisadores têm estudado a interação entre evolução e o aprendizado (Hinton e Nowlan, 1987; Belew, 1989; Ackley e Littman, 1991; Gruau e Whitley, 1993; French e Messinger, 1994; Whitley, Gordon e Mathias, 1994). Estes trabalhos têm em comum mostrar as vantagens, em termos de desempenho do aprendizado, quando este interage com a evolução.

Davis (1991) denomina Algoritmos Genéticos Híbridos à combinação de algoritmos genéticos convencionais (item 2.5) com métodos de aprendizado ou otimização local. A forma utilizada comumente em Algoritmos Genéticos Híbridos é equivalente à idéia de evolução de Lamarck, isto é, um aprendizado que modifica o código genético do indivíduo. Outro caminho onde aprendizado e evolução interagem, é permitir que o aprendizado mude o valor da função *fitness*, o que representa um aprendizado conhecido como efeito Baldwin, aprendizado que não afeta o código genético do indivíduo. Ambas as estratégias de aprendizado são mais eficientes e efetivas que o Algoritmo Genético convencional, segundo Whitley, Gordon e Mathias (1994).

No trabalho de Hinton e Nowlan (1987), implementa-se o efeito Baldwin, utilizando os Algoritmos Genéticos e um processo de aprendizado randômico que desenvolve uma rede neuronal. A codificação genética especifica a topologia da rede, indicando qual das 20 conexões potenciais poderá ser utilizada para conectar um conjunto de neurônios. A função objetivo é tal que a rede “correta” incrementa o *fitness*, que em outro caso diminui. O código genético utilizado consiste de um alfabeto de 3 caracteres (0, 1, ?), o qual especifica ausência de conexão, presença de conexão ou conexão não especificada. Se a conexão é não especificada, ela poderia ser feita randomicamente ou por aprendizado. Assim, o aprendizado é uma busca randômica de possíveis conexões não especificadas, denotadas pelo caractere “?”.

No trabalho de Ackley e Littman (1991), é utilizado um algoritmo de aprendizado para otimizar a função *fitness* durante o tempo de vida do indivíduo. No material genético do indivíduo, é representada uma função de “boa forma” que representa o aprendizado durante

o tempo de vida do indivíduo, sob a assunção que esta função é a representação exata da função *fitness*.

No trabalho de Gruau e Whitley (1993), adiciona-se aprendizado ao processo de desenvolvimento e evolução de árvores gramaticais. Aqui são explorados o efeito Baldwin e a estratégia de aprendizado de Lamarck.

Em Whitley, Gordon e Mathias (1994), também utilizam-se as duas formas de aprendizado (Baldwin e Lamarck) e mostra-se que Lamarck converge ao ótimo local enquanto que Baldwin consegue convergir ao ótimo global. Uma afirmação importante deste trabalho é que a alteração do *fitness* sem alterar o código genético é uma forma de simular o efeito Baldwin.

Mayley (1996), destaca os benefícios e custos do aprendizado. Entre os benefícios, pode-se destacar a adaptabilidade temporal, adaptabilidade espacial e manutenção da diversidade genética, o que garante que indivíduos geneticamente diferentes são capazes de aprender comportamentos semelhantes. Isto implica que a evolução de uma população pode ser guiada por tratamento fenotípico adquirido por membros de uma população em seu tempo de vida.

Dos trabalhos anteriormente mencionados, pode-se concluir que são utilizados dois tipos de algoritmos, um baseado nos algoritmos genéticos utilizados como uma busca global; outro, em algoritmos de aprendizados utilizados como busca local. A busca local é realizada ativamente, via herança lamarckista, ou passivamente, via efeito Baldwin.

Em nosso trabalho, a idéia de aprendizado utilizada é diferente à apresentada nos trabalhos mencionados anteriormente. Em lugar de modificar diretamente o código genético dos indivíduos (aprendizado tipo Lamarck), ou em lugar de alterar a função *fitness* (utilizado como aprendizado do tipo efeito Baldwin), o método proposto baseia-se na idéia de que os indivíduos percebem o ambiente, portanto aprendem durante seu tempo de vida, fazendo com que as mutações mais benéficas estejam mais propensas a ocorrer. Estas mutações mais benéficas não ocorrem no próprio indivíduo, mas este aprendizado é

transmitido para seus descendentes de forma passiva. Desta forma, é possível que aquelas mutações mais apropriadas tenham uma maior possibilidade de manifestar-se nas gerações posteriores.

A idéia de aprendizado utilizado está fortemente influenciada pelas idéias de mutação direcionada, no entanto é utilizada a explicação neodarwinista para o aumento das frequências de genes que formam as características que são responsáveis por uma boa resposta ao ambiente.

CAPÍTULO III

MODELO PROPOSTO : FUNDAMENTAÇÃO E IMPLEMENTAÇÃO

3.1 Introdução

Devido à importância da modelagem de sistemas como método científico, contrastada com a dificuldade e até a impossibilidade de se construir modelos matemáticos que descrevam o comportamento de sistemas, é prudente a consideração de um modelo que auxilie na tarefa do observador do sistema.

A identificação da estrutura do sistema e de seus parâmetros é a meta do modelo proposto. Para alcançar este propósito, o modelo a ser proposto neste capítulo baseia-se nas seguintes idéias:

- a) utilização de regras do tipo Se-Então para representar as relações entre os agentes do sistema e destes com seu meio ambiente;
- b) a lógica difusa, utilizada em duas funções diferentes neste trabalho:
 - como ferramenta de aproximação ao comportamento dos agentes e
 - como modelo para inclusão de graus nos processos adaptativos, sendo utilizada no tratamento das imprecisões e inexatidão inerentes aos próprios processos adaptativos;
- c) os processos adaptativos utilizados para encontrar o modelo do sistema podem ser distinguidos como:
 - a evolução para processar as mudanças de longo prazo e
 - o aprendizado, onde se processam as mudanças de curto prazo.

Neste capítulo, serão expostas as principais idéias utilizadas para o desenvolvimento do modelo proposto, assim como sua implementação, em um algoritmo.

3.2 Modelo Proposto

No capítulo 1, foram expostas a importância da modelagem de sistemas como método científico e a necessidade de ferramentas que permitam a obtenção de um modelo do sistema observado. Na definição de sistemas, apresentada no item 1.1, foi colocado que os agentes que formam o sistema exibem comportamento próprio e, portanto, é necessária a existência de um conjunto de estruturas que servirão para descrever o comportamento dos agentes do sistema. Na continuação especificamos formalmente a definição de sistemas utilizada no item 1.1, tomando como ponto de partida a especificação de Caswell (1972).

As estruturas comportamentais dos agentes de um sistema podem ser representadas como $(c1, c2, \dots, cn)$, com domínio em $C1, C2, \dots, Cn$. Assim, em qualquer instante de tempo t , o comportamento dos agentes pode ser representado pelos valores destas estruturas:

$$[c1(t), c2(t), \dots, cn(t)] \in (C1 \times C2 \times \dots \times Cn) \quad (4)$$

Assim, o comportamento dos agentes é uma série temporal de instâncias que pode ser representada através de uma variável comportamental (δ) e sua equação :

$$\delta : (t0, t1) \rightarrow (C1 \times C2 \times \dots \times Cn) \quad (5)$$

O domínio da variável comportamental δ é Δ , o conjunto de todas as séries temporais em $(t0, t1)$ e domínio $(C1 \times C2 \times \dots \times Cn)$. Portanto Δ inclui todos os possíveis comportamentos dos agentes. Uma descrição comportamental completa é obtida quando é possível especificar o conjunto de comportamentos que ocorrem em um instante de tempo, e que pode ser considerado um subconjunto de Δ (Δ^*), isto é:

$$\delta \in \Delta^* \subset \Delta \quad (6)$$

Obter o subconjunto Δ^* é a meta da modelagem de sistemas.

No entanto, na prática, as estruturas comportamentais são impostas pelo observador, como uma estrutura estímulo-resposta. A idéia é dividir a estrutura comportamental dos agentes em um conjunto de estímulos e respostas. Assim, em lugar das estruturas comportamentais $(c1, c2, \dots, cn)$, tem-se um conjunto de estímulos $(e1, e2, \dots, em)$ com domínio $(E1, E2, \dots, Em)$ e um conjunto de respostas $(r1, r2, \dots, rp)$ com domínio $(R1, R2,$

... , Rp). Assim, a variável comportamental (δ) fica:

$$\delta = (\varepsilon, \lambda) \quad (7)$$

onde

$$\varepsilon: (t0, t1) \rightarrow (E1 \times E2 \times \dots \times Em), \text{ e} \quad (8)$$

$$\lambda: (t0, t1) \rightarrow (R1 \times R2 \times \dots \times Rp), \quad (9)$$

então de (6) deduzimos que $\delta = (\varepsilon, \lambda) \in \Delta^* \subset \Delta$ (10)

O maior esforço na modelagem de um sistema está em selecionar e representar os estímulos-respostas, já que os comportamentos e estratégias dos agentes do sistema são representados aqui (Holland, 1995). O critério usual da escolha dos estímulos e das respostas é o que o observador percebe, isto é, uma relação causa-efeito. A relação causa-efeito pode ser descrita como:

$$\lambda = \Gamma(\varepsilon) \quad (11)$$

esta relação Γ ocorre entre valores instantâneos de estímulos e respostas, isto é:

$$r(t) = \gamma(e(t)) \quad (12)$$

A equação (12) significa que, para determinados valores de estímulos, tem-se uma única resposta. Assim, a função γ é uma função $n \rightarrow 1$.

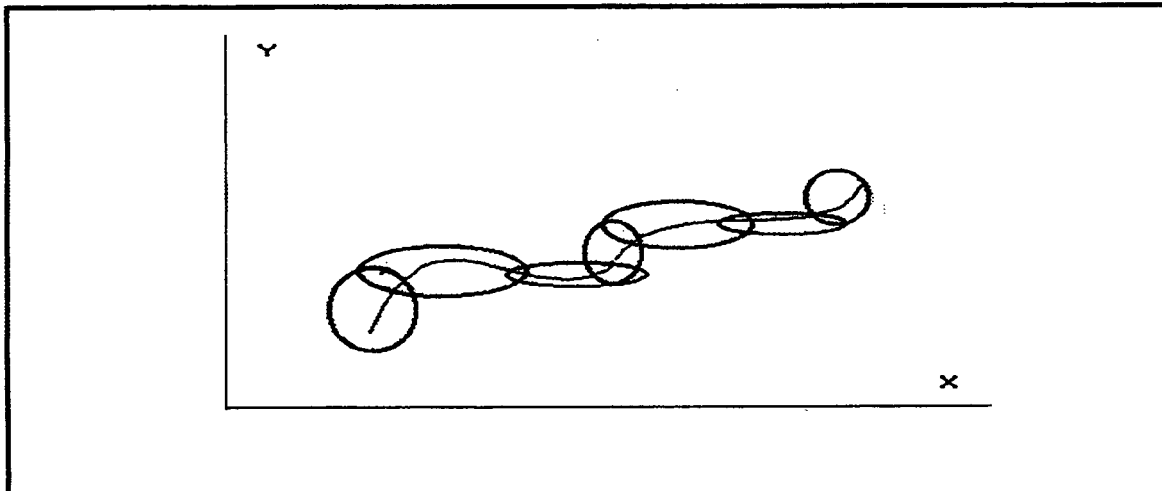
Uma possível maneira de representar uma relação causa-efeito é através de regras do tipo Se-Então, as quais são um caminho para descrever o comportamento dos agentes e as interações entre os agentes e seu ambiente. Isto significa que a função γ está representada por uma regra do tipo Se-Então. As regras do tipo Se-Então trabalham como regiões que identificam o comportamento do sistema, como se pode observar na figura 6.

Por exemplo a regra “Se $x = A$ Então $y = B$ “, onde os conjuntos A e B representam uma região do domínio de x e y , respectivamente, cobrem uma região do espaço do sistema. Na nomenclatura anterior, podemos dizer que os conjuntos A e B são os domínios de determinados estímulos-resposta (x, y) associados aos agentes do sistema.

Por outro lado, o critério usual da escolha dos estímulos e das respostas é o que o observador percebe como uma relação causa-efeito. Devido ao fato de que a descrição dos estímulos e respostas está baseada no conhecimento do observador, este conhecimento é

impreciso e incerto por natureza. Esta é a principal motivação para a utilização da lógica difusa.

Figura 6: Regiões Cobrindo o Comportamento do Sistema



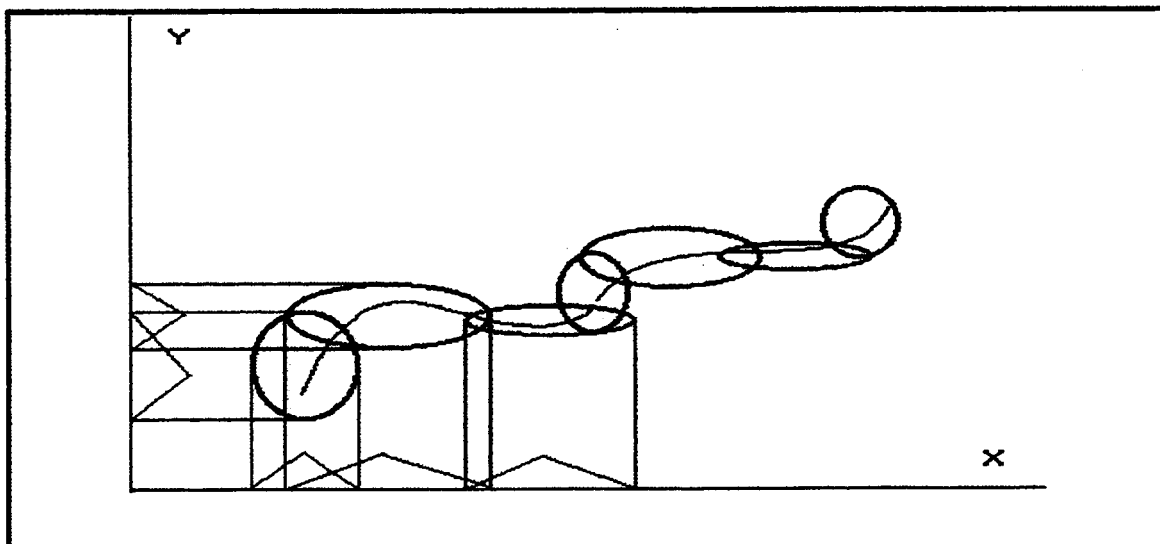
Fonte: Adaptado de Kosko (1992)

Se os conjuntos A e B da regra anterior fossem difusos, a regra poderia ser considerada uma regra difusa. Neste sentido, o domínio das variáveis x e y seria dividido em regiões difusas representando diferentes valores lingüísticos das variáveis, como é mostrado pela figura 7.

Para descrever o comportamento do sistema, devemos encontrar o subconjunto Δ^* , ou, de outra forma, encontrar o conjunto de relações Γ (ou regras) que representam o modelo. Se geramos um conjunto de regras para o sistema, é necessário estabelecer uma condição sobre este conjunto. Esta condição consiste em cobrir com as regras geradas todos os estímulos-respostas disponíveis, o que garante que inclui todas os comportamentos possíveis do sistema. O algoritmo que encontra estas regras deve garantir esta condição.

Estas regras que constituem o modelo do sistema poderiam começar com um conjunto de regras aleatórias, ou com regras obtidas de algum conhecimento prévio que se tenha do sistema. Entretanto, serão os processos adaptativos que, considerando grandes intervalos de tempo, realizarão o ajuste das regras e, portanto, a obtenção final do modelo do sistema.

Figura 7: Regiões Difusas Cobrindo o Comportamento do Sistema



Fonte: Adaptado de Kosko (1992)

Os processos adaptativos são um método válido para a avaliação e invenção de teorias científicas (Fogel, 1991), se considerados como mecanismos de busca iterativos, que tentam aproximar o modelo obtido ao sistema observado. Os indivíduos da população são hipóteses do comportamento do sistema. Isto significa que os comportamentos e relações dos agentes são representados no código genético dos indivíduos da população. A validade destas hipóteses fica demonstrada pela sobrevivência de seus descendentes.

O processo evolutivo atua sobre os indivíduos da população, alterando as características genéticas dos mesmos, o que garante a diversidade genética da população. Por outro lado, o aprendizado é representado pelo tempo de vida dos indivíduos onde, através da experiência, eles incrementam ou diminuem suas possibilidades de sobrevivência futura.

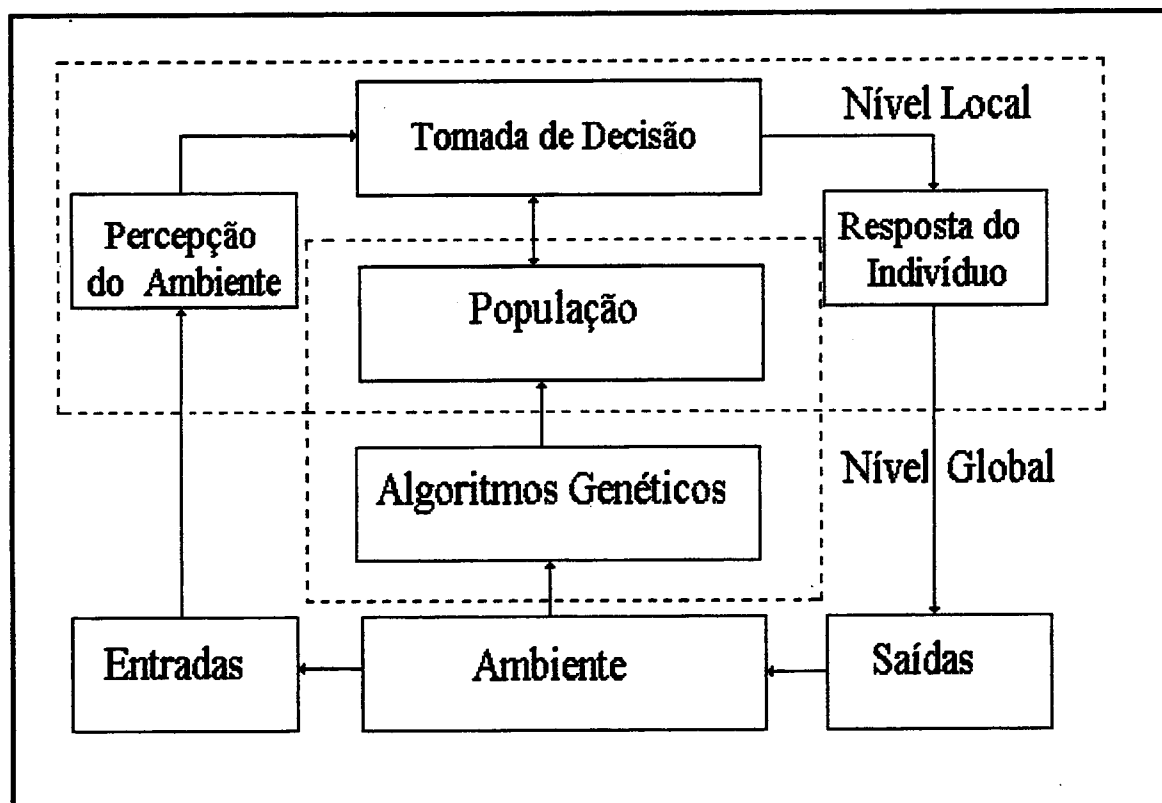
Assim, o modelo adaptativo proposto (figura 8) é formado de dois níveis: um nível evolutivo ou global, que atua, através do código genético, sobre as gerações da população; e um nível local ou de aprendizado, que atua no tempo de vida dos agentes, pela resposta do indivíduo aos estímulos do ambiente. Assim, no nível local, os indivíduos percebem o ambiente na forma de valores lingüísticos e respondem a ele através de suas estratégias que são representadas por regras do tipo Se-Então e estão codificadas no código genético do próprio indivíduo.

No aprendizado ou nível local, a Lógica Difusa é utilizada como ferramenta de aproximação entre instâncias do comportamento do sistema e as regras do modelo. Na evolução ou nível global, onde as mudanças ocorrem nas estruturas comportamentais, a Lógica Difusa é utilizada na representação do código genético dos indivíduos.

As funções principais do indivíduo no nível local são responder aos estímulos ambientais, aprender da experiência e modificar sua avaliação do ambiente. Para que este aprendizado seja transmitido a seus descendentes, propomos que cada gene do indivíduo tenha uma probabilidade de mutação própria. Desta forma, o indivíduo altera as probabilidades de seus genes em função de seu aprendizado no tempo de vida. Quando o indivíduo responde satisfatoriamente aos estímulos ambientais, as probabilidades se alteram para que o indivíduo, ao cruzar-se com outros indivíduos, transmita seu código genético sem alterações. Desta forma, o indivíduo estaria passando aos descendentes, de forma indireta (passiva), o aprendizado em seu tempo de vida, o que seria uma maneira do aprendizado auxiliar à evolução, como foi colocado no item 2.7.

A figura 8, mostra que as entradas são percebidas pelo indivíduo como valores de pertinência de seus valores lingüísticos, para tomar uma decisão em função das estratégias que possui e que estão codificadas em seu genótipo. Quando a decisão é tomada, o indivíduo responde ao ambiente. Isto é realizado com todos os dados de treinamento, o que representa o tempo de vida do indivíduo. As respostas que o indivíduo dá ao ambiente são acumuladas para obter o *fitness* do indivíduo. Esta função de *fitness* é que deve garantir que o conjunto de regras encontradas cubra todos os conjuntos de estímulos-respostas. Na geração subsequente, o Algoritmo Genético convencional é aplicado, alterando a diversidade genética da população.

Figura 8: Modelo Adaptativo Proposto



Fonte: do Autor

3.3 Implementação do Modelo Proposto

Propomos neste item um algoritmo para o modelo sugerido. O algoritmo possui três módulos principais, dois correspondentes à fase de treinamento do algoritmo e um terceiro módulo que trata da construção final da base de regras. A fase de treinamento possui dois sub-módulos principais, correspondentes à evolução das gerações das populações de indivíduos e ao tempo de vida de cada indivíduo da população.

O algoritmo possui, além dos módulos referidos acima, alguns parâmetros de entrada necessários no começo do algoritmo, que devem ser fornecidos pelo usuário, além do banco de dados com informações de dados de Entrada-Saída sobre o sistema que se deseja identificar. Estes parâmetros são correspondentes ao Algoritmo Genético, como o número de indivíduos da população e as probabilidades iniciais dos diferentes operadores.

Outro parâmetro importante é o número de valores lingüísticos de cada variável do

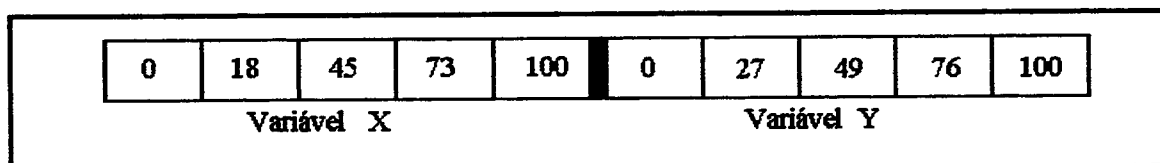
sistema, isto é, a divisão do espaço das variáveis em regiões difusas, o que está diretamente vinculado à representação dos indivíduos da população do Algoritmo Genético. Esta divisão é muito importante, já que está diretamente vinculada ao número potencial de regras do sistema, devido ao fato de este número ser resultado da combinação dos valores lingüísticos das variáveis do sistema (Ishibuchi, Nozaki, Yamamoto e Tanaka, 1995).

3.3.1 Representação dos Indivíduos

A representação dos indivíduos foi escolhida de forma que facilite tanto a representação das funções de pertinência dos valores lingüísticos, como a obtenção da base de regras que representa o modelo formal do sistema identificado.

Para se conseguir tais propósitos, a representação dos indivíduos escolhida foi tal que cada indivíduo possui todos os valores lingüísticos codificados nos lócus dos cromossomos como valores inteiros. Por exemplo, se o sistema que se deseja identificar tem duas variáveis, cada uma com cinco valores lingüísticos, sua representação pode ser visualizada na figura 9.

Figura 9: Representação de um Indivíduo com Duas Características e Cinco Lócus



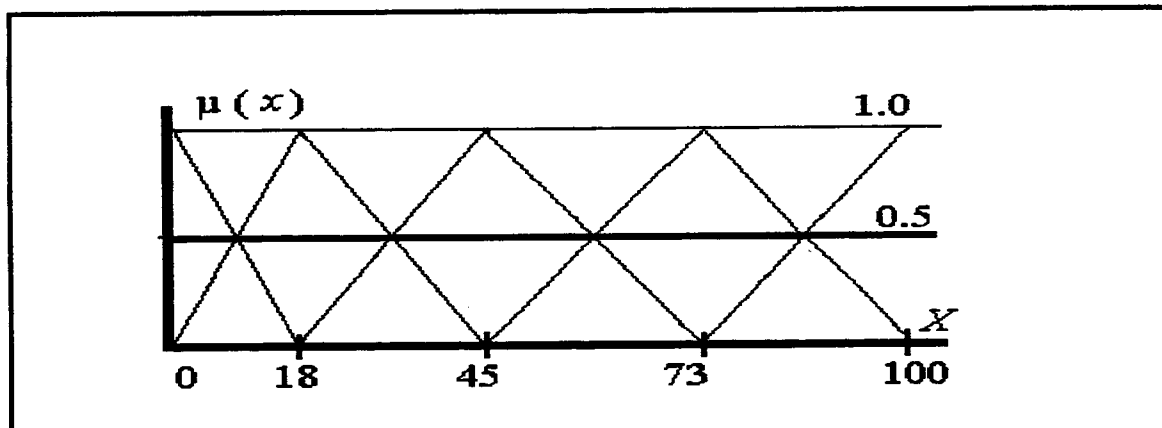
Fonte: do Autor

Em nossa nomenclatura, dizemos que o indivíduo i tem duas características x e y , e que cada característica tem 5 genes que representam os picos das funções de pertinência, que são consideradas como triangulares. Esta representação tem a vantagem de o ponto de corte entre dois valores lingüísticos ocorrer no ponto 0.5 (ver figura 10); isto facilita o cômputo dos futuros cálculos. Os extremos de cada função de pertinência estão representados pelos lócus vizinhos do lócus que representa o pico. Na figura 10, observa-se a representação da característica x do cromossomo da figura 9.

Uma observação importante nesta representação é que um indivíduo representa um

modelo do sistema, isto é, um conjunto de regras difusas que aproxima os dados de entrada-saída aos valores do sistema real. Esta é uma diferença essencial deste modelo, comparativamente aos trabalhos citados no item 2.6. Desta forma, quando o processo evolutivo entra em ação, está-se simulando uma população de modelos de regras, e não uma população de regras.

Figura 10: Funções de Pertinência de Uma Característica de um Indivíduo



Fonte: do Autor

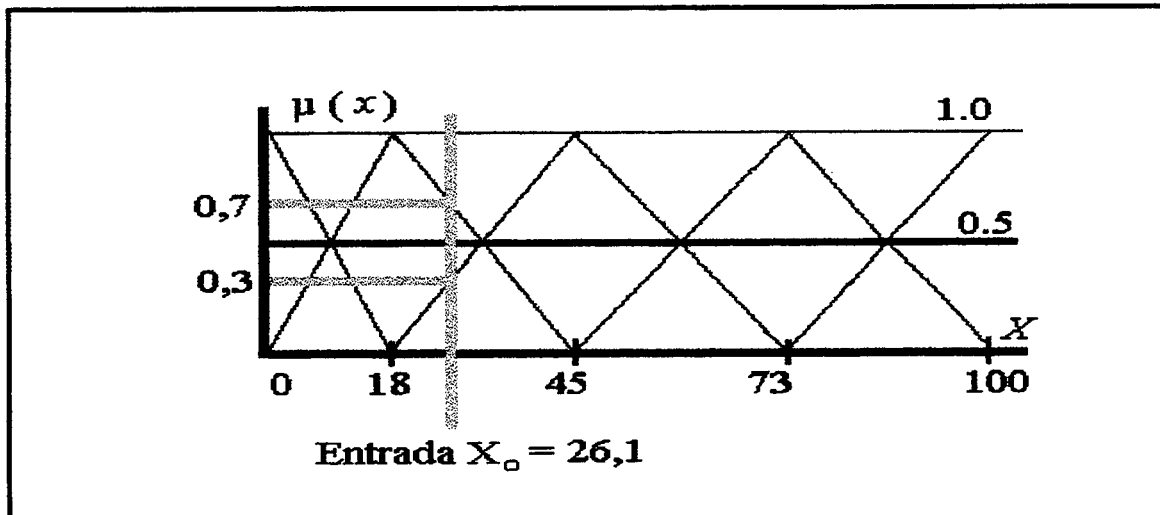
3.3.2 Nível Local

Uma vez definidos todos os parâmetros de entrada do algoritmo, podemos executá-lo. Começando por uma primeira geração aleatória de indivíduos, cada indivíduo da população é testado com os dados de entrada-saída do sistema que estamos identificando. Aqui, o algoritmo possui os seguintes sub-módulos: a *fuzzification*, a *defuzzification* e a escolha das regras que serão utilizadas com o método de inferência, que se encontram implícitas no modelo, como veremos nos parágrafos a seguir. Estes sub-módulos modelam a percepção do indivíduo do ambiente, para poder tomar uma decisão e dar uma resposta.

O módulo de *fuzzification*, como já foi explicado no capítulo 2, mapeia uma variável de entrada em graus de pertinência de algum conjunto difuso. Assim, os dados de entrada (x_i , y_i) são projetados sobre as funções de pertinência formadas pelo indivíduo testado, e assim são obtidos os valores de pertinência de cada dado de entrada e saída, como mostra a figura

Pode-se observar que o método de *fuzzification* corta no máximo duas funções de pertinência. Isto se dá devido à maneira como foram definidas estas funções. Isto também implica que o número potencial de modelos que podem ser aplicados, para um mesmo conjunto de dados de entrada-saída, é 2^{n-1} conjuntos de duas regras, onde n é o número de variáveis do sistema.

Figura 11: *Fuzzification* Aplicado sobre a Característica de Um Indivíduo.



Fonte: do Autor

A decisão que deve ser tomada é saber qual modelo (conjunto de regras) deve ser aplicado para um conjunto específico de dados. Por exemplo, tendo duas variáveis de entrada (A e B) e uma de saída (C), com cinco valores lingüísticos para cada variável, dados os valores de entrada-saída (x_1, x_2, y), e aplicando o método de *fuzzification*, consideramos que os valores lingüísticos que os valores (x_1, x_2, y) cortam, são $A_1, A_2, B_4, B_5, C_2, C_3$. Os modelos que podem ser aplicados, são por exemplo:

- Se A_1 e B_4 Então C_2 , Se A_2 e B_5 Então C_3 , Modelo 1;
- Se A_1 e B_4 Então C_3 , Se A_2 e B_5 Então C_2 , Modelo 2;
- Se A_1 e B_5 Então C_2 , Se A_2 e B_4 Então C_3 , Modelo 3;
- Se A_1 e B_5 Então C_3 , Se A_2 e B_4 Então C_2 , Modelo 4.

É importante notar que de todos estes possíveis modelos que podem ser aplicados para um conjunto de dados específicos, só um deles deve ser considerado: aquele que obtém

o melhor resultado. Este fato é importante, já que o módulo de treinamento do algoritmo poderá escolher diferentes modelos de regras para dados similares.

Tendo as regras que devem ser aplicadas, calculamos o valor obtido pelo modelo, através do módulo de *defuzzification*. Para cada regra aplicada na inferência, é calculado o valor de saída. Este valor está dado pela área sombreada da figura 12, que é calculada como será mostrado no próximo parágrafo:

Seja AT a área total da função de pertinência, B a base dessa função, a altura do triângulo (h) sabemos que é igual a 1, então:

$$AT = (B * h) / 2 = B / 2. \quad (13)$$

Seja AS a área sombreada, e AB a área do triângulo formado acima da área sombreada, então temos que:

$$AS = AT - AB. \quad (14)$$

A altura deste triângulo AB é dada por $(1 - H)$, sendo H o valor obtido pela *fuzzification* das variáveis de entrada e obtido pelo operador do método de inferência.

Assim, a área branca têm o valor:

$$AB = (B * (1 - H) * (1 - H)) / 2. \quad (15)$$

A área sombreada fica como:

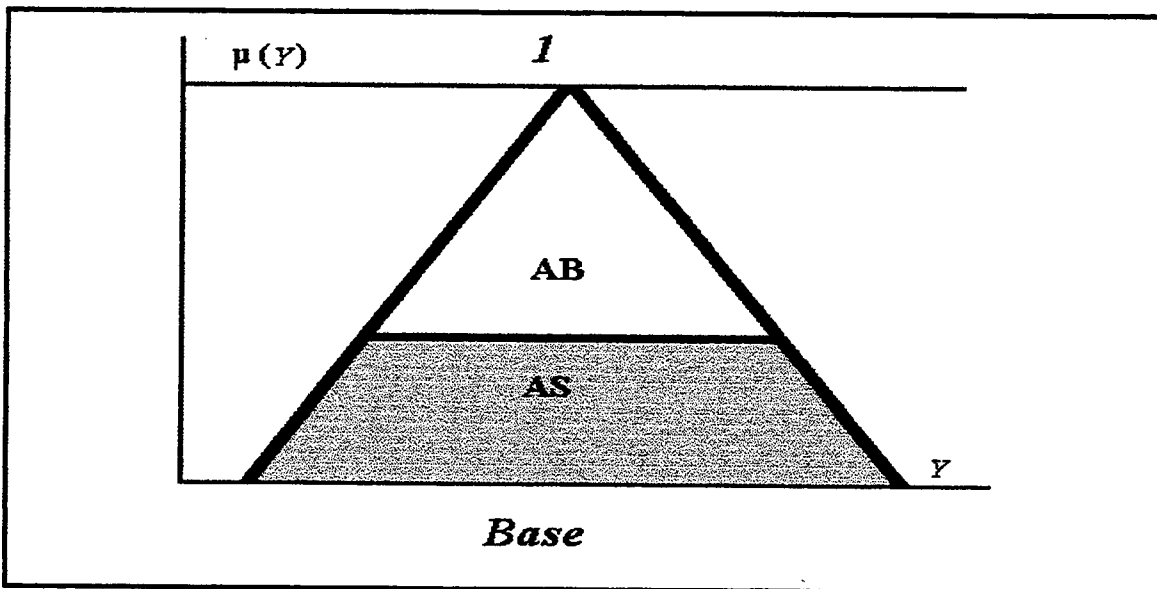
$$AS = (B * H * (2 - H)) / 2. \quad (16)$$

Uma vez calculada esta área para cada regra do modelo aplicado, combinamo-las para obter o valor final (figura 13), resultando seguinte cálculo:

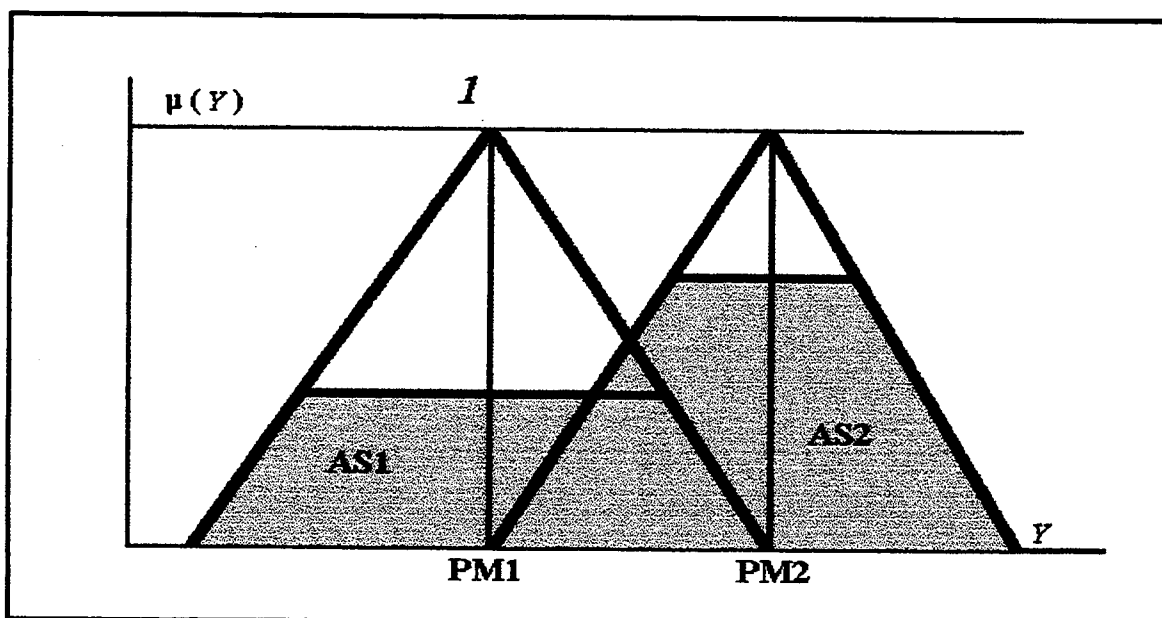
$$((AS1 * PM1) + (AS2 * PM2)) / (AS1 + AS2), \quad (17)$$

onde $AS1$ e $AS2$ são as áreas sombreadas das duas funções de pertinência que contribuem para a obtenção do valor final; $PM1$ e $PM2$ são os picos dessas funções de pertinência.

Figura 12: Cálculo da Área Sombreada.



Fonte: do Autor

Figura 13: Método de *Defuzzification*

Fonte: do Autor

Neste nível local é onde implementamos o esquema de aprendizado. Cada gene do indivíduo tem associada uma probabilidade de mutação. Esta probabilidade é alterada para cada gene em função da resposta do indivíduo às exigências ambientais, como é comentado a seguir:

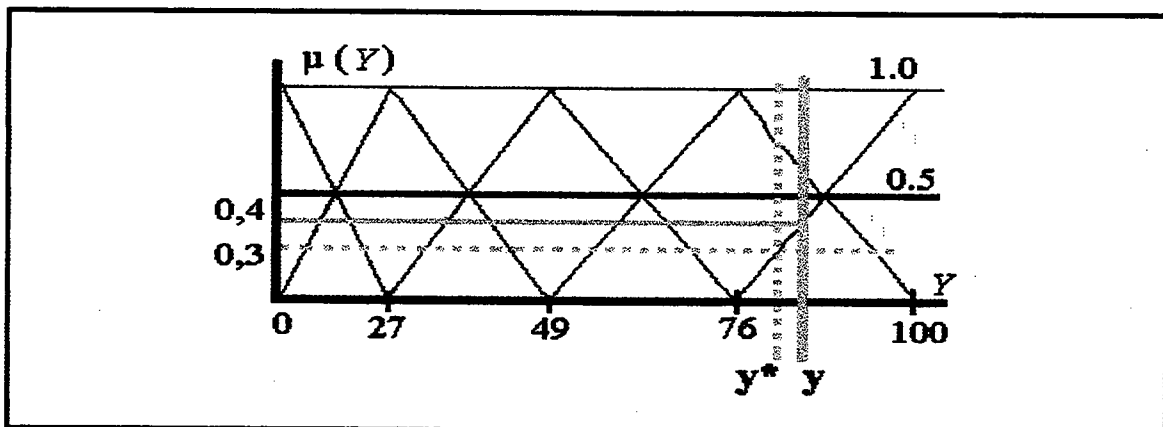
Para cada dado ambiental (x_k, y_k) , quando é aplicado o método de *fuzzification* sobre as funções de pertinência da variável de saída, observa-se uma diferença com o valor projetado obtido pela *fuzzification* na variável de entrada (figura 14). Para cada dado ambiental, temos um erro R_k , dado pela diferenças entre os graus de pertinência. Na fórmula (18), está colocado o erro R_k , onde P_{ik} é o grau de pertinência do termo lingüístico, A_i correspondente à variável x e P_{mk} é o grau de pertinência do termo lingüístico B_m correspondente à variável y .

$$R_k = | P_{ik} - P_{mk} | \quad (18)$$

Ao mesmo tempo, existe um erro entre o valor real de sistema (y) e o valor encontrado pelo modelo (y^*) (figura 14).

$$E_k = | y_k - y_k^* | \quad (19)$$

Figura 14: Comparação entre *Fuzzification* e o Método de Inferência Aplicado sobre a Variável de Saída



Fonte: do Autor

O próximo passo é normalizar as medidas das fórmulas (18) e (19):

$$V_{nk}(i) = R_{ik} / \left(\sum_j R_{jk} \right) \quad (20)$$

onde j representa todos os termos lingüísticos que foram afetados pelas regras utilizadas pelos dados (x_k, y_k) , e i um determinado termo lingüístico de uma variável do sistema. Também normalizamos a fórmula (19), obtendo a fórmula (21).

$$E_{nk} = | y_k - y_k^* | / | y_{max} - y_{min} | \quad (21)$$

onde y_{max} e y_{min} podem ser duas constantes que representam os valores máximo e mínimo do sistema, respectivamente.

Uma vez normalizadas as medidas, distribuímos o erro Enk entre os termos lingüísticos que contribuíram com a saída (y^*):

$$Dk(i) = Enk * Vnk(i) \quad (22)$$

onde i representa um termo lingüístico de uma variável lingüística.

Contabilizamos para cada termo lingüístico, quantas vezes contribuiu positivamente com uma saída:

$$N(i) = N(i) + 1 \quad (23)$$

Este método é aplicado para todos os dados do ambiente (x, y). Quando este processo está terminado, a probabilidade de mutação de cada termo lingüístico (gene) $Pm(i)$ é dada por:

$$Pm(i) = \left(\sum_j Dj(i) \right) / N(i) \quad (24)$$

3.3.3 Nível Global

Uma vez que todos os indivíduos da população foram avaliados e passaram pelo nível local, o segundo módulo de treinamento começa a ser aplicado. Trata-se do algoritmo genético propriamente dito, como foi explicado no item 2.5. Aqui, o código genético dos indivíduos da população pode ser alterado, o que permite o aparecimento de novos indivíduos e garante que o processo evolutivo seja realizado.

O *fitness* de cada indivíduo é dado por:

$$ET = (1 / (2 * n)) * \left(\sum_i (ymax - ymin)^2 - (yi^* - yi)^2 \right), \quad (25)$$

onde yi é o valor real do sistema, e yi^* é o valor obtido pelo algoritmo e $ymax$ e $ymin$ são os valores reais máximo e mínimo do sistema respectivamente; a soma é realizada sobre todos os dados de treinamento, n . Observa-se que ET é máximo, quando o segundo termo da somatória tende a zero, e isto acontece quando $yi \rightarrow yi^*$.

Além dos operadores de mutação e cruzamento utilizados no Algoritmo Genético, como foi descrito no item 2.5 (como a representação nos locos dos cromossomos foi com

valores inteiros), são utilizados outros dois operadores, *creep* e *merge*, que ajudam a manter a diversidade genética da população (Espy, Vombrack e Aldridge, 1992). Em nosso algoritmo, estes operadores são implementados de maneira diferente à encontrada na bibliografia já citada.

O operador *creep* realiza, para cada locus do indivíduo onde o operador é aplicado, uma “pequena” alteração no valor do locus. Esta alteração é realizada em função de uma probabilidade de *creep*, que determina se o operador é aplicado ou não. Quando o operador é aplicado, a “pequena” alteração é realizada da seguinte maneira: dado o valor do locus, sorteia-se um valor randomicamente entre os valores dos locus vizinhos. O valor sorteado ficará no lugar do valor atual.

O operador *merge* pega dois indivíduos e gera um terceiro que representa a mistura entre os dois indivíduos escolhidos. Esta alteração é realizada em função de uma probabilidade de *merge*. Esta mistura se dá em nível de locus, isto é, o locus i do indivíduo $x1$ se combina com o locus i do indivíduo $x2$. Esta combinação entre os locus é realizada através de uma combinação linear, onde os coeficientes são sorteados randomicamente para cada locus.

CAPÍTULO IV

SIMULAÇÕES REALIZADAS E RESULTADOS OBTIDOS

4.1 Introdução

Neste capítulo, é testado o algoritmo descrito no capítulo anterior, com diferentes funções e sistemas. São simulados diferentes tipos de funções, com a intenção de obter uma idéia do funcionamento do algoritmo em identificar diferentes tipos de comportamentos. Algumas funções são testadas utilizando o algoritmo proposto com e sem o esquema de aprendizado, para comparar os resultados obtidos.

Os resultados obtidos são apresentados e comparados com os resultados obtidos em outros trabalhos, como os de Valenzuela-Rendón (1991), Parodi e Bonelli (1993), Cordón e Herrera (1996a, 1996b e 1996c) e Takagi e Sugeno (1985).

Para realizar as simulações, foi gerada uma série aleatória de dados de treinamento de 100 dados, para as cinco primeiras funções, e uma série de 100 dados para teste. Foram realizadas simulações com 100, 500 e 1000 gerações e com diferentes números de valores lingüísticos. Para as funções *f6* e *f7*, mostradas a seguir, foi gerada uma série de 1681 dados aleatórios (41 dados para a variável *x1* e 41 para a variável *x2*), e 168 dados de teste, como foi utilizado em Cordón e Herrera (1996a e 1996b). Para a função *f8*, foi gerada uma série de 26 dados para cada variável respectivamente, como dados de treinamento e para teste foram utilizados 67 dados com em Cordón e Herrera (1996c).

Os parâmetros utilizados no algoritmo genético são:

- a) população de 50 a 100 indivíduos,
- b) probabilidade de cruzamento e de *merge* de 0.8,
- c) probabilidade de mutação inicial de 0.01 e 0.005,
- d) probabilidade de *creep* 0.025.

4.2 Funções Testadas

As funções testadas são:

$$- f1: y = x; \text{ com } x \in [0; 1]; \quad (26)$$

$$- f2: y = x^2; \text{ com } x \in [0; 1]; \quad (27)$$

$$- f3: y = 4 * (x - 0.5)^2; \text{ com } x \in [0; 1]; \quad (28)$$

$$- f4: y = \text{sen}(20 * x); \text{ com } x \in [0; 1]; \quad (29)$$

$$- f5: y = \text{sen}^6(5 * \text{Pi} * x) * e^{-2 * \text{Ln}(2) * (x - 0.1 / 0.8)^2}; \quad (30)$$

$$\text{com } x \in [0; 1];$$

$$- f6: y = (x_1^2 + x_2^2); \text{ com } x_1, x_2 \in [-5; 5]; \quad (31)$$

$$- f7: y = (x_1^2 + x_2^2 - \cos(18 * x_1) - \cos(18 * x_2)) \quad (32)$$

$$\text{com } x_1, x_2 \in [-1; 1];$$

$$- f8: y = e^{x_1} * \text{sen}^2 x_2 + e^{x_2} * \text{sen}^2 x_1; \quad (33)$$

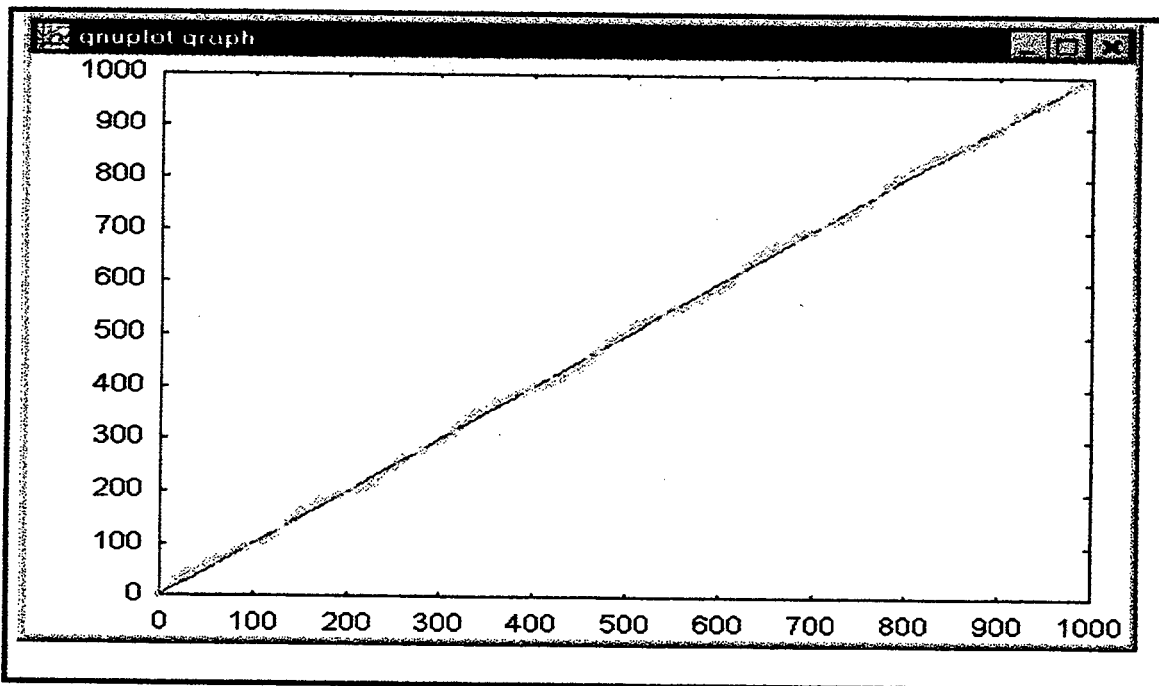
$$\text{com } x_1, x_2 \in [-8; 8];$$

As funções *f1* e *f3* foram utilizadas nos trabalhos de Valenzuela-Rendón (1991) e Parodi e Bonelli (1993). A função *f4* foi simulada por Parodi e Bonelli (1993); já as funções *f6* e *f7* encontram-se em Córdón e Herrera (1996a e 1996b); a função *f8* encontra-se em Córdón e Herrera (1996c). A função *f5* é uma função multimodal, e foi utilizada para ver o comportamento do modelo neste tipo de funções.

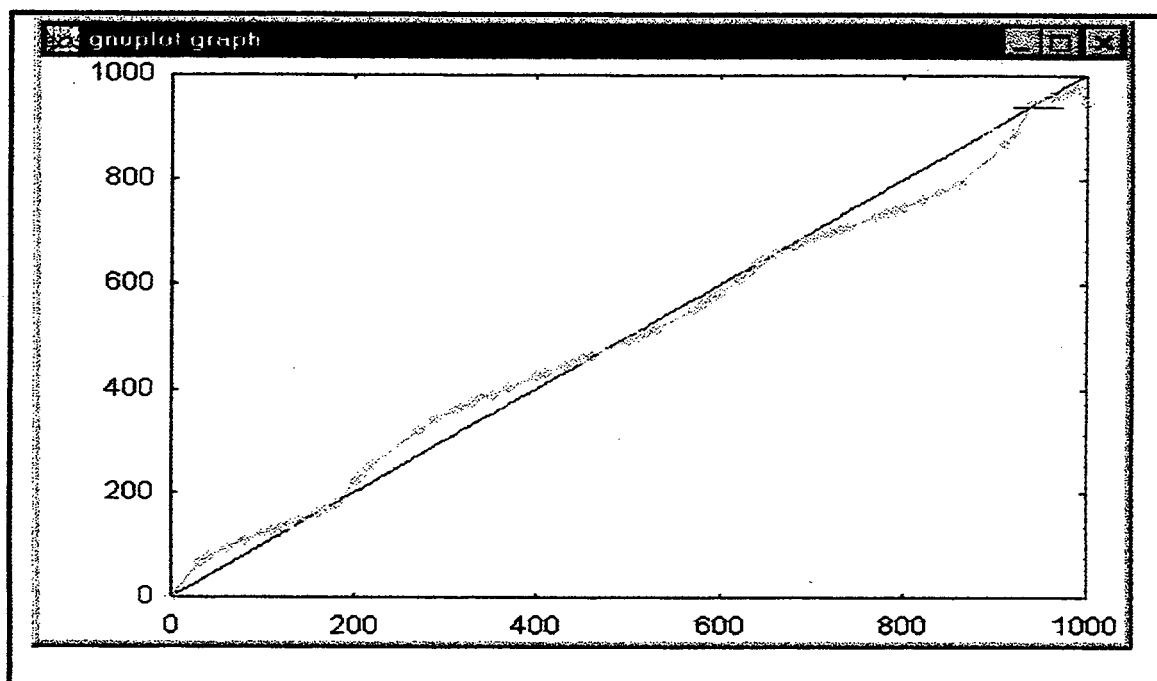
Na seqüência são apresentadas algumas figuras, mostrando diferentes simulações realizadas para as funções anteriormente apresentadas, assim como quadros comparativos das simulações realizadas com o algoritmo implementado utilizando o esquema de aprendizado e sem utilizar o esquema proposto. Também são apresentados quadros comparativos com os resultados encontrados por outros trabalhos. Os resultados são discutidos, e são apresentadas algumas conclusões sobre os quadros mostrados.

Na apresentação das figuras representamos com linha lisa (de cor azul) o comportamento real da função, e a linha com pontos (de cor verde) representa o comportamento encontrado com o modelo de regras que o algoritmo encontrou para cada simulação.

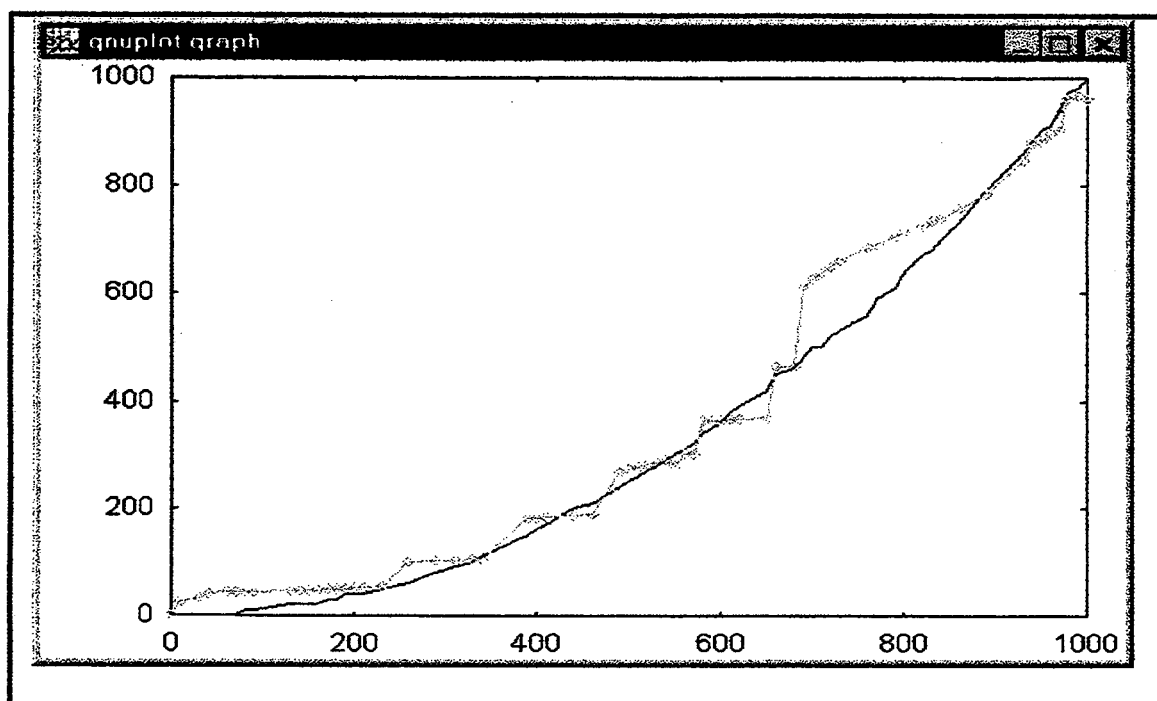
Figura 15: Resultado da Função f_I com 7 Valores Lingüísticos e 18 Regras



Fonte: do Autor

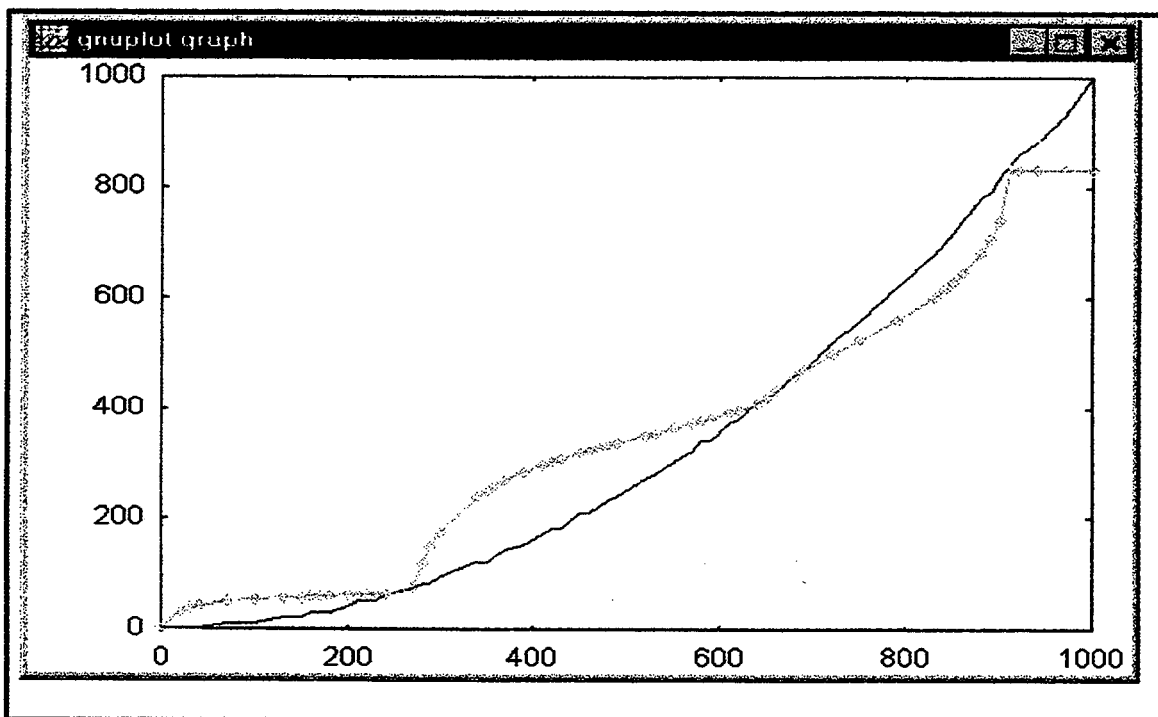
Figura 16: Resultado da Função $f1$ com 5 Valores Lingüísticos e 6 Regras

Fonte: do Autor

Figura 17: Resultado da Função $f2$ com 7 Valores Lingüísticos e 16 Regras

Fonte: do Autor

Figura 18: Resultado da Função f_2 com 5 Valores Lingüísticos e 6 Regras



Fonte: do Autor

No quadro 1 e 2, apresenta-se um resumo dos resultados das funções f_1 e f_2 . Ambos os quadros comparam os resultados encontrados com o algoritmo proposto, utilizando o esquema de aprendizado, indicado por (CA), e sem utilizar o esquema de aprendizado (SA). São comparados também simulações para diferentes números de iterações (100, 500 e 1000) e diferentes valores lingüísticos (5 e 7) para as variáveis difusas.

O erro utilizado é o erro absoluto (34), expresso em porcentagem (Valenzuela-Rendóm, 1991; Parodi e Bonelli, 1993):

$$EA = (1/b - a) * \int |y^* - y| dy. \quad (34)$$

Valores Lingüísticos	$f_1 (5, CA)$	$f_1 (5, SA)$	$f_1 (7, CA)$	$f_1 (7, SA)$
Gerações	EA	EA	EA	EA
$f_1 (100)$	4.474	5.789	2.681	3.298
$f_1 (500)$	2.831	2.856	1.508	1.879
$f_1 (1000)$	2.299	2.350	0.963	1.341

Quadro 1: Comparação dos Erros da Função f_1 , com Aprendizado e sem Aprendizado

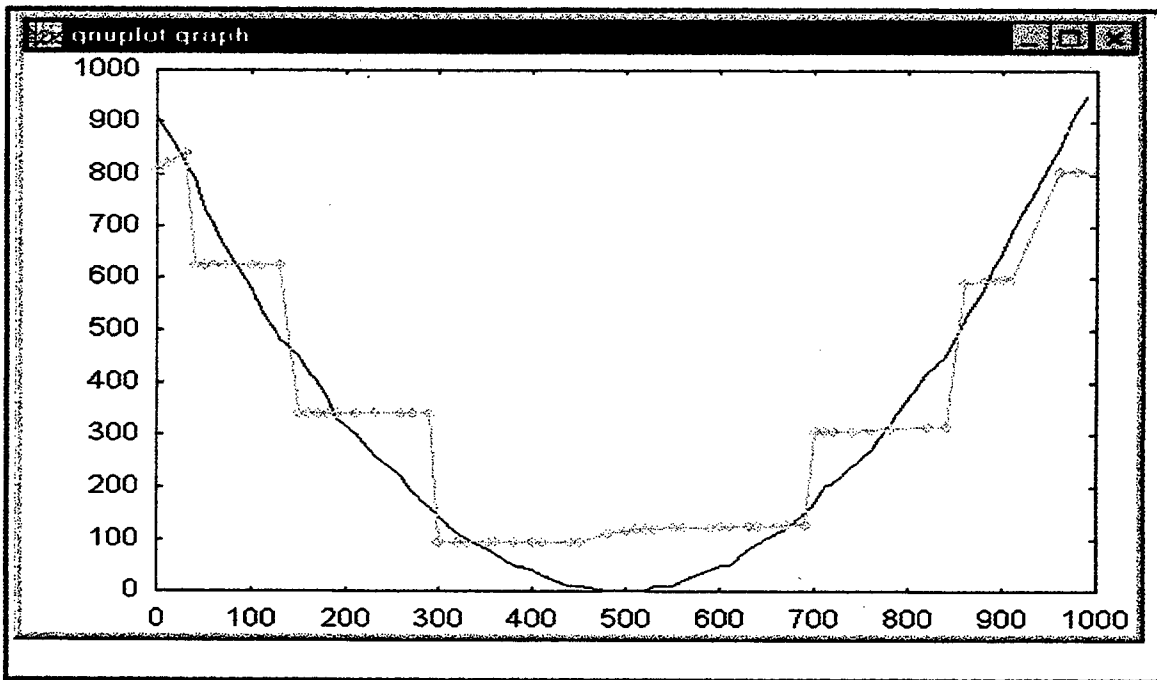
Valores Lingüísticos	$f2 (5, CA)$	$f2 (5, SA)$	$f2 (7, CA)$	$f2 (7, SA)$
Gerações	<i>EA</i>	<i>EA</i>	<i>EA</i>	<i>EA</i>
$f2 (100)$	4.956	6.317	3.377	4.637
$f2 (500)$	3.276	3.388	2.402	2.700
$f2 (1000)$	2.538	2.581	1.984	2.150

Quadro 2: Comparação dos Erros da Função $f2$, com Aprendizado e sem Aprendizado

A primeira conclusão que se pode tirar dos dados expressos nestes quadros é que no algoritmo em que se utiliza o esquema de aprendizado, os resultados são notoriamente melhores que no algoritmo sem tal esquema. Este fato valida a utilização de um esquema de aprendizado interagindo com o processo evolutivo. Valida também o esquema de aprendizado apresentado no item 3.3.2.

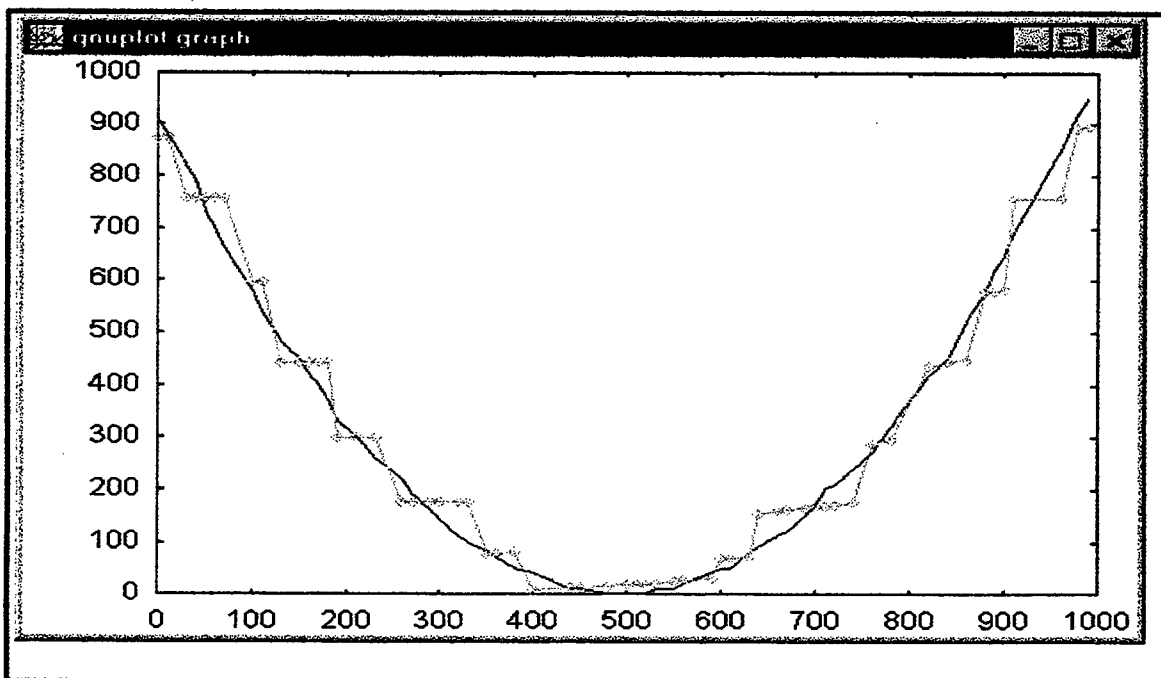
Também pode-se observar que com o aumento das gerações, as diferenças encontradas por ambos os algoritmos diminui. Isto parece ser um indicativo de que o aprendizado acelera a convergência, encontrando melhores resultados com menor número de iterações. Com o aumento das iterações, a vantagem do aprendizado tende a diminuir. Uma terceira conclusão esperada dos quadros 1 e 2 é que, com o aumento das gerações e o aumento da partição do domínio das funções em valores lingüísticos, os resultados também melhoram.

Figura 19: Resultado da Função f_3 com 5 Valores Lingüísticos e 17 Regras



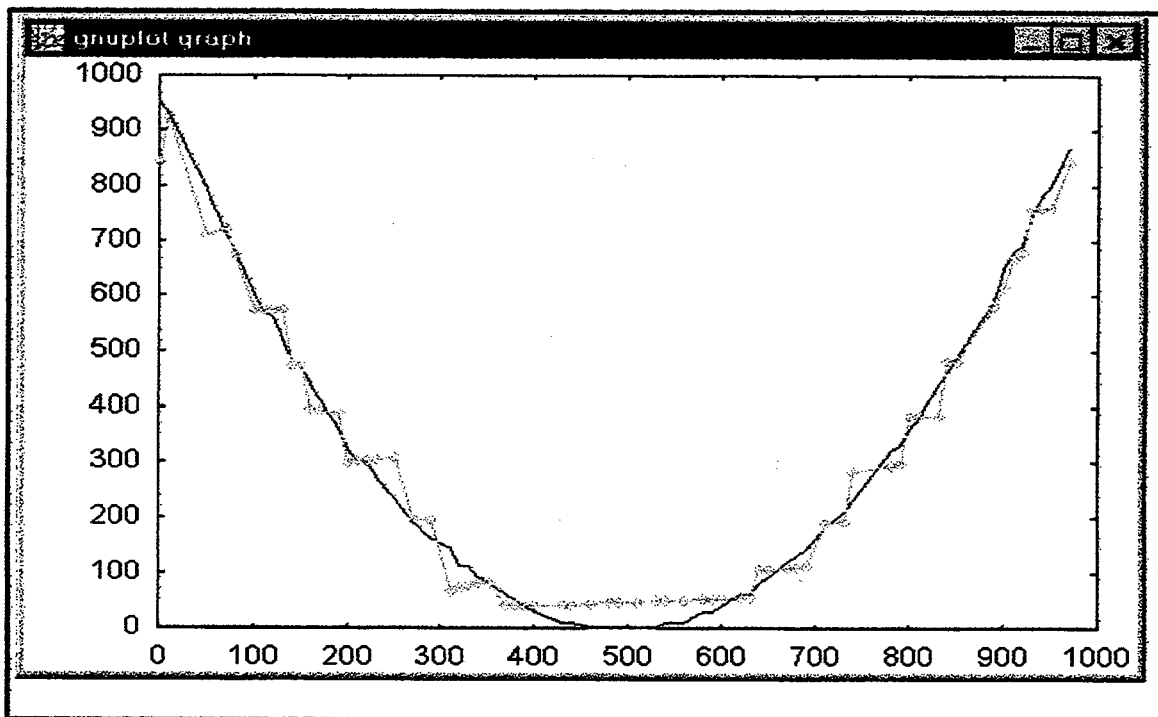
Fonte: do Autor

Figura 20: Resultado da Função f_3 com 7 Valores Lingüísticos e 24 Regras



Fonte: do Autor

Figura 21: Resultado da Função f_3 com 9 Valores Lingüísticos e 50 Regras



Fonte: do Autor

O quadro 3 compara os resultados obtidos com o modelo proposto, e os resultados obtidos por outros pesquisadores utilizando diversos algoritmos. Os algoritmos comparados são:

- (1) Valenzuela-Rendón (1991);
- (2) Parodi e Bonelli (1993);
- (3) Modelo Proposto neste trabalho, com 5 valores lingüísticos;
- (4) Modelo Proposto com 7 valores lingüísticos ;
- (5) Modelo Proposto com 9 valores lingüísticos;
- (6) Modelo Proposto com 10 valores lingüísticos;
- (7) Modelo Proposto com 12 valores lingüísticos.

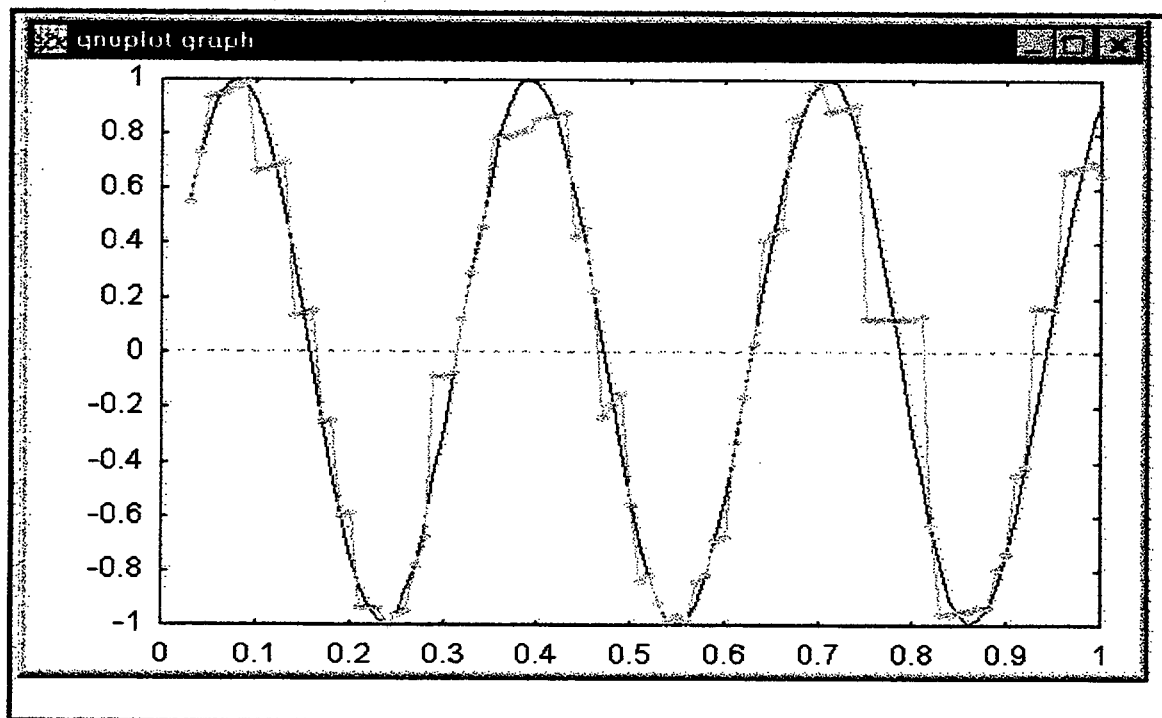
Algoritmo	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Função	<i>EA</i>	<i>EA</i>	<i>EA</i>	<i>EA</i>	<i>EA</i>	<i>EA</i>	<i>EA</i>
f_1	1.79	1.19	2.37	1.59	0.96	-	-
f_3	3.76	3.48	7.59	4.66	3.42	2.97	1.29

Quadro 3: Comparação de Resultados das Funções f_1 e f_3

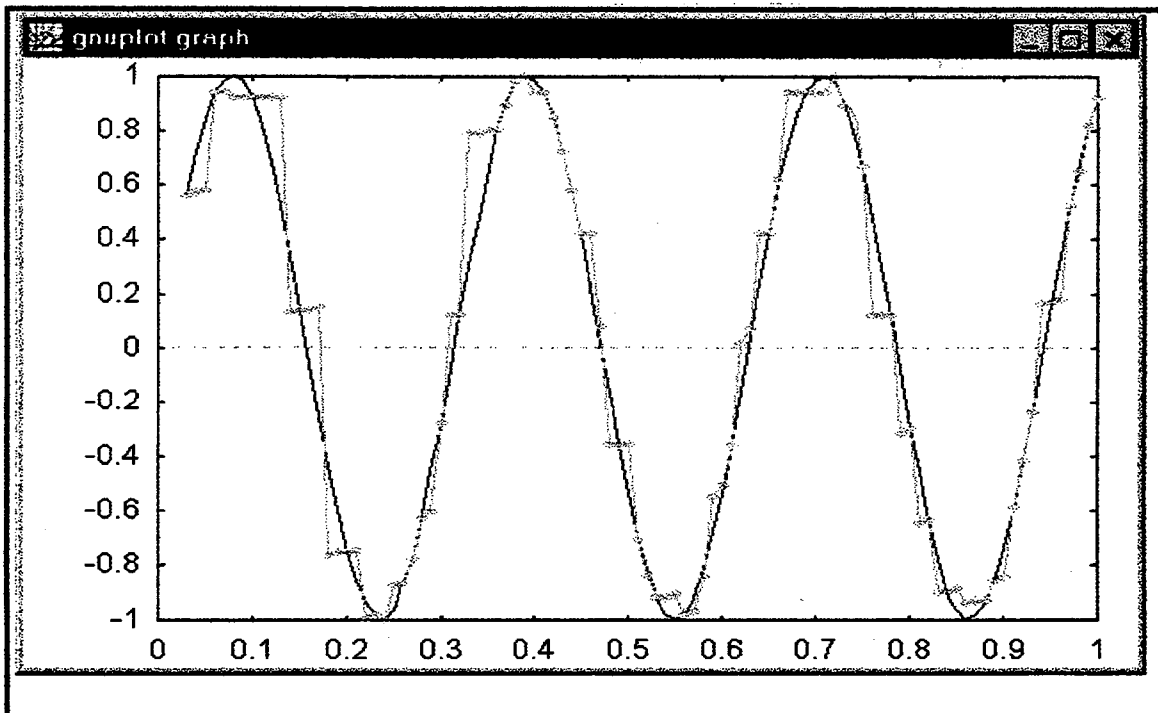
Observa-se que o modelo aqui proposto apresenta resultados semelhantes e até melhores que os resultados encontrados por outros autores. Deve ser ressaltado que o modelo proposto por Valenzuela-Rendón (1991), utiliza 40 regras para obter o resultado apresentado na quadro 3, enquanto o modelo proposto utiliza 5, 7, 9, 10 e 12 valores lingüísticos respectivamente. O trabalho de Parodi e Bonelli (1993), não indica o número de regras utilizadas nas simulações apresentadas.

Como é mostrado no quadro 3, dependendo da função, o algoritmo proposto encontra resultados semelhantes aos encontrados por outros algoritmos, com 5 ou 7 valores lingüísticos. Semelhantemente às constatações feitas através da leitura dos quadros 1 e 2, no quadro 3 pode-se observar como o algoritmo se comporta com o aumento do número de valores lingüísticos. Os resultados mostram que o comportamento do algoritmo melhora, o que era esperado, já que aumenta o número de regras do modelo, o que permite uma maior exatidão do modelo.

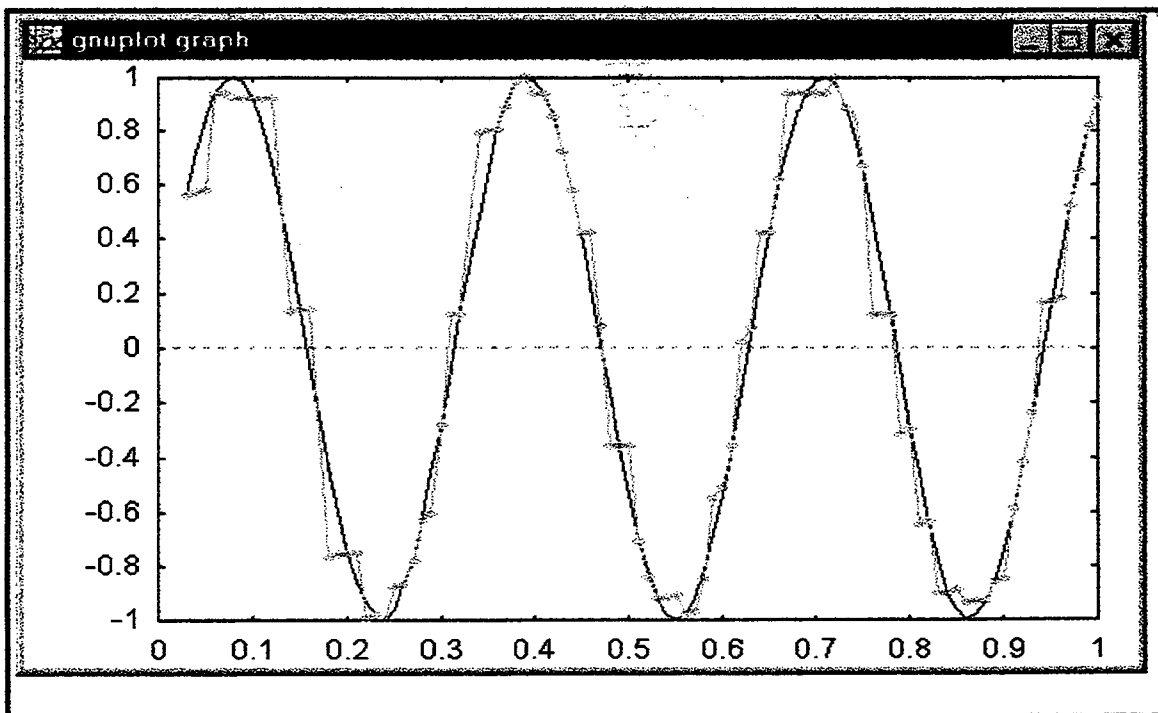
Figura 22: Resultado da Função f_4 com 7 Valores Lingüísticos e 36 Regras



Fonte: do Autor

Figura 23: Resultado da Função f_4 com 9 Valores Lingüísticos e 43 Regras

Fonte: do Autor

Figura 24: Resultado da Função f_4 com 12 Valores Lingüísticos e 61 Regras

Fonte: do Autor

O quadro 4 mostra um resumo dos resultados da função f_4 , comparados com os resultados de Parodi e Bonelli (1993).

Algoritmo	(1)	(2)	(3)	(4)
Função	<i>EA</i>	<i>EA</i>	<i>EA</i>	<i>EA</i>
f_4	4.67	4.81	4.07	3.51

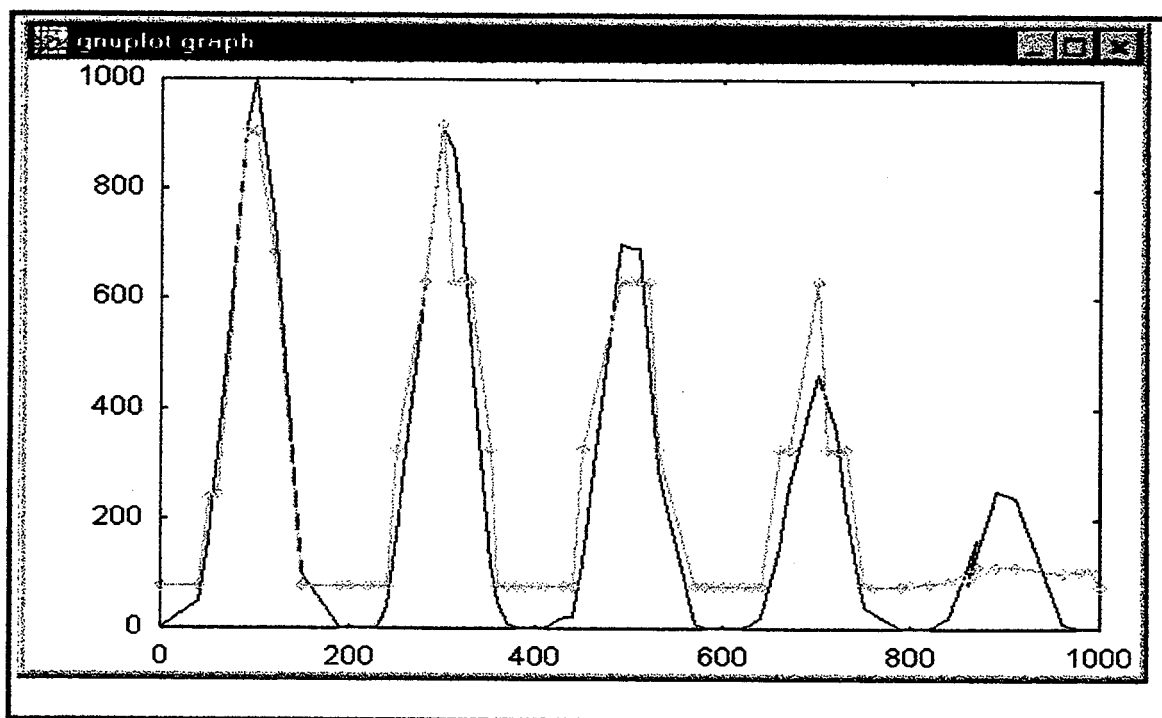
Quadro 4: Comparação de Resultados da Função f_4

Legenda:

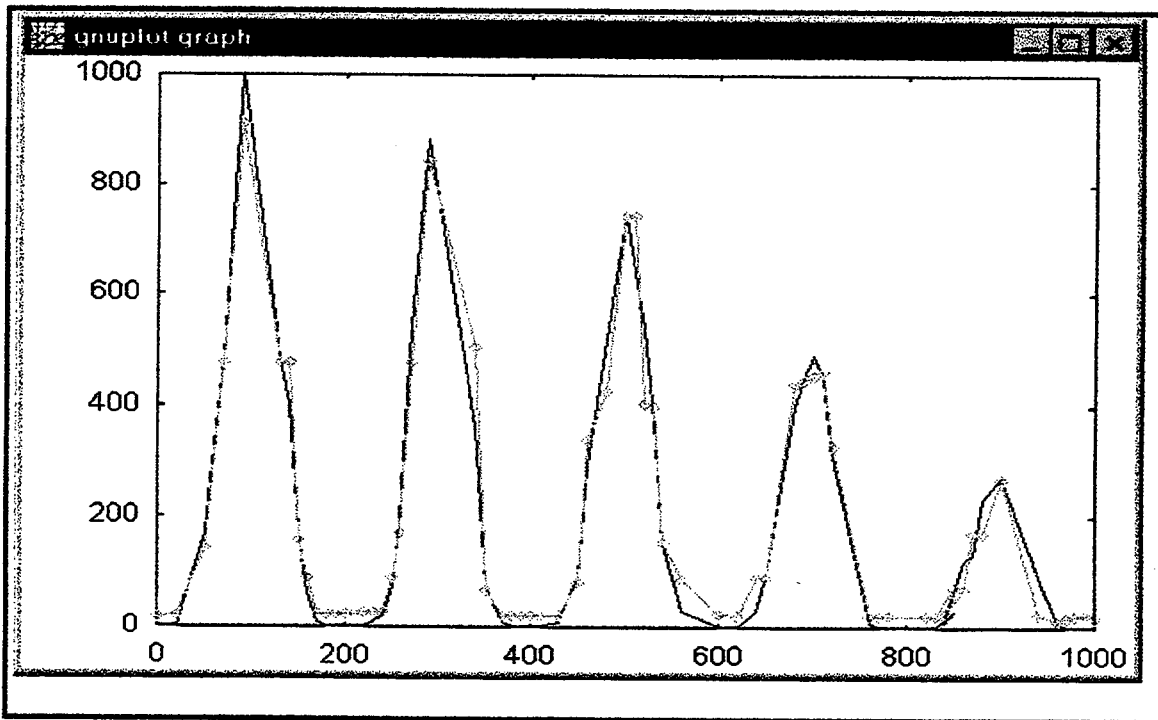
- (1) Parodi e Bonelli (1993);
- (2) Modelo Proposto neste trabalho, com 7 valores lingüísticos;
- (3) Modelo Proposto com 9 valores lingüísticos;
- (4) Modelo Proposto com 12 valores lingüísticos.

No quadro 4 pode-se observar que os resultados apresentados pelo modelo proposto são similares e até melhores aos resultados encontrados por Parodi e Bonelli (1993). As conclusões sobre o aumento no número de valores lingüísticos das variáveis da função são similares as conclusões do quadro 3.

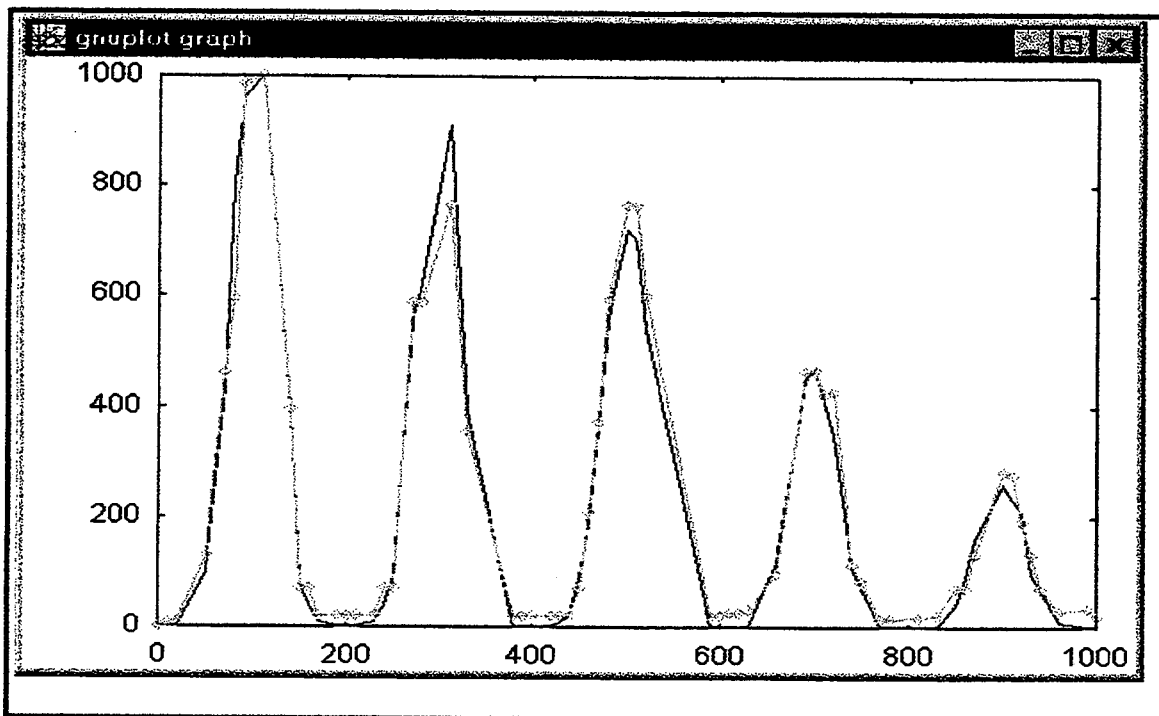
Figura 25: Resultado da Função f_5 com 5 Valores Lingüísticos e 16 Regras



Fonte: do Autor

Figura 26: Resultado da Função f_5 com 9 Valores Linguísticos e 41 Regras

Fonte: do Autor

Figura 27: Resultado da Função f_5 com 12 Valores Linguísticos e 77 Regras

Fonte: do Autor

O quadro 5 mostra os resultados obtidos pela simulação da função f_5 .

Algoritmo	(1)	(2)	(3)
Função	<i>EA</i>	<i>EA</i>	<i>EA</i>
f_5	7.39	3.34	2.71

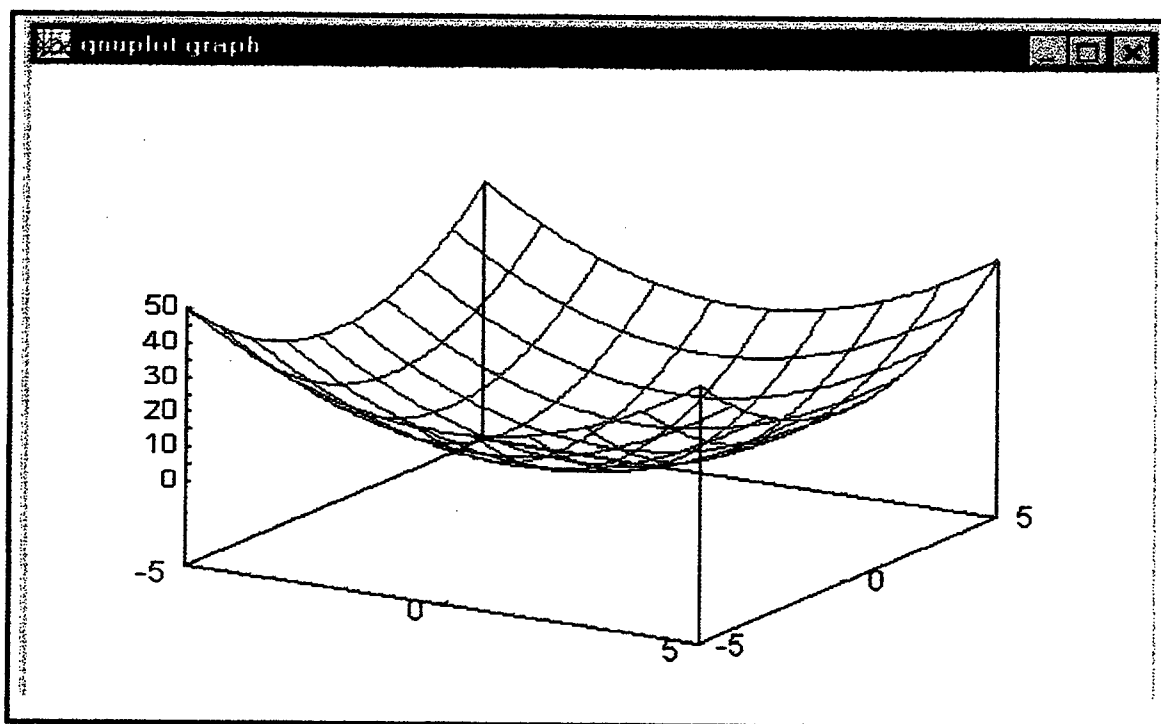
Quadro 5: Resultados da função f_5 .

Legenda:

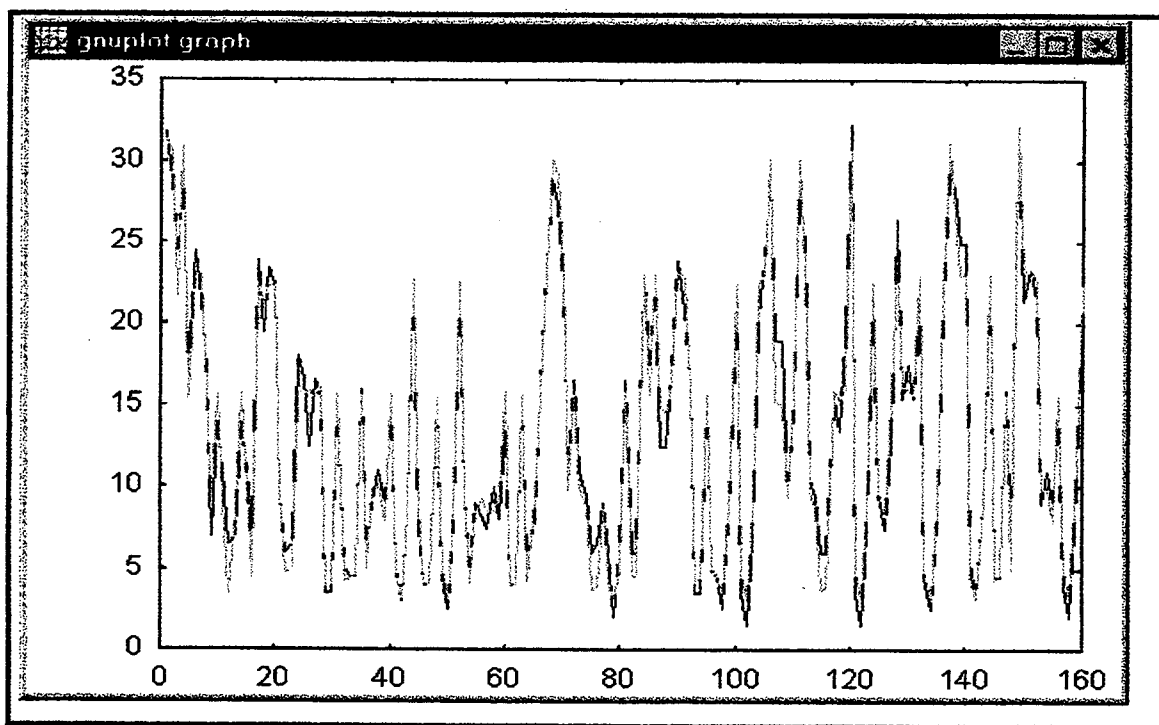
- (1) Modelo Proposto neste trabalho, com 5 valores lingüísticos;
- (2) Modelo Proposto com 9 valores lingüísticos;
- (3) Modelo Proposto com 12 valores lingüísticos.

A intenção de simular esta função (f_5) é mostrar a habilidade do algoritmo em identificar funções não monótonas e multimodais. Como aconteceu nas funções anteriores, o algoritmo foi capaz de identificar o comportamento da função; a aproximação encontrada depende novamente do número de regras encontradas e do número de valores lingüísticos das variáveis do sistema.

Figura 28: Representação gráfica da função f_6 .



Fonte : do Autor

Figura 29: Representação gráfica dos erros da função f_6 .

Fonte: do Autor

O quadro 6 mostra os resultados da função f_6 (ver figura 28). Os resultados são comparados com os resultados de Cordón e Herrera (1996b), que apresentam além de seus resultados, os obtidos por Wang e Mendel, pelo algoritmo por eles proposto. O erro utilizado aqui é o erro médio quadrático (35). Os resultados apresentados por Cordón e Herrera (1996b) são os das fases um, dois e três de algoritmo, como foi discutido no capítulo 2. A figura 29 mostra graficamente os erros da f_6 , o que permite obter uma idéia visual da aproximação encontrada.

$$EQ = (1/(2*n)) * \sum_i (|y_i^* - y_i|)^2 \quad (35)$$

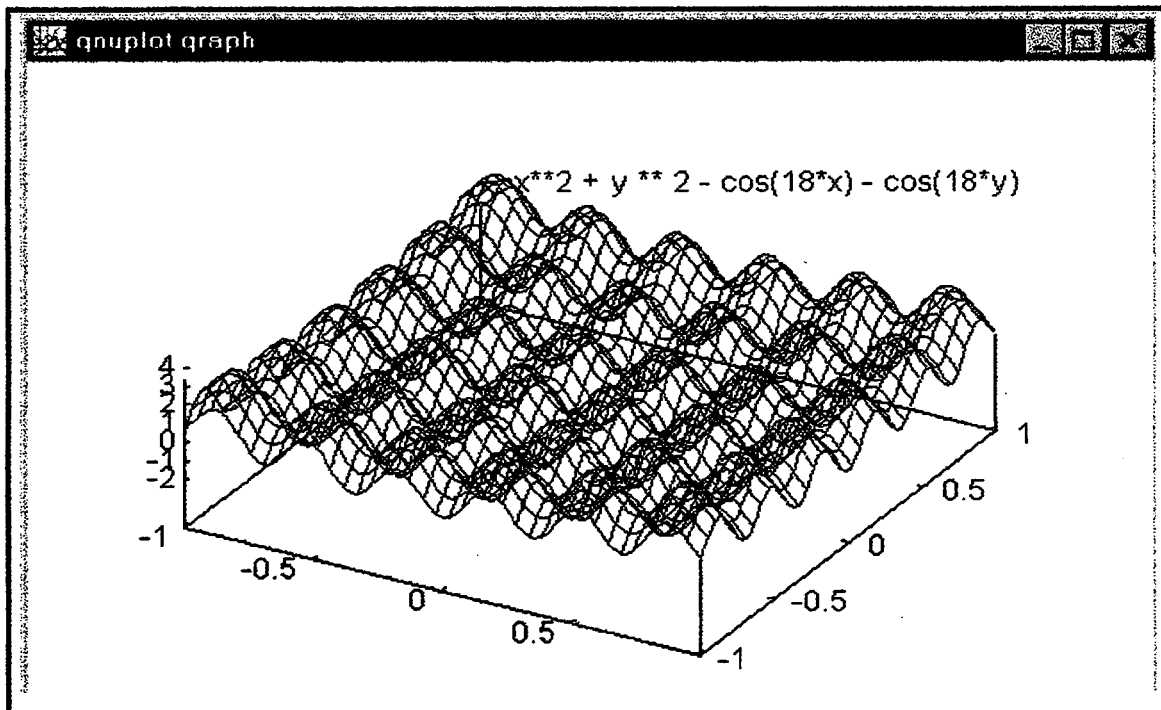
(1)	(1)	(2)	(2)	(3)	(4)	(4)	(5)	(5)
EQ	#R	EQ	#R	EQ	EQ	#R	EQ	#R
1.99247	70	1.63051	58		4.65181	49	3.26233	70
1.99247	70	1.77093	63				2.94546	81
2.41140	98	1.77914	67	0.69687			2.58822	102
2.41140	98	2.13019	73	1.11825			1.89826	132

Quadro 6: Comparação dos Resultados da Função f_6

Legenda:

- (1) Cordón e Herrera (1996b), fase 1;
- (2) Cordón e Herrera (1996b), fase 2;
- (3) Cordón e Herrera (1996b), fase 3;
- (4) Wang e Mendel (1992);
- (5) Modelo Proposto, com 7 valores lingüísticos.

Figura 30: Representação Gráfica da Função $f7$

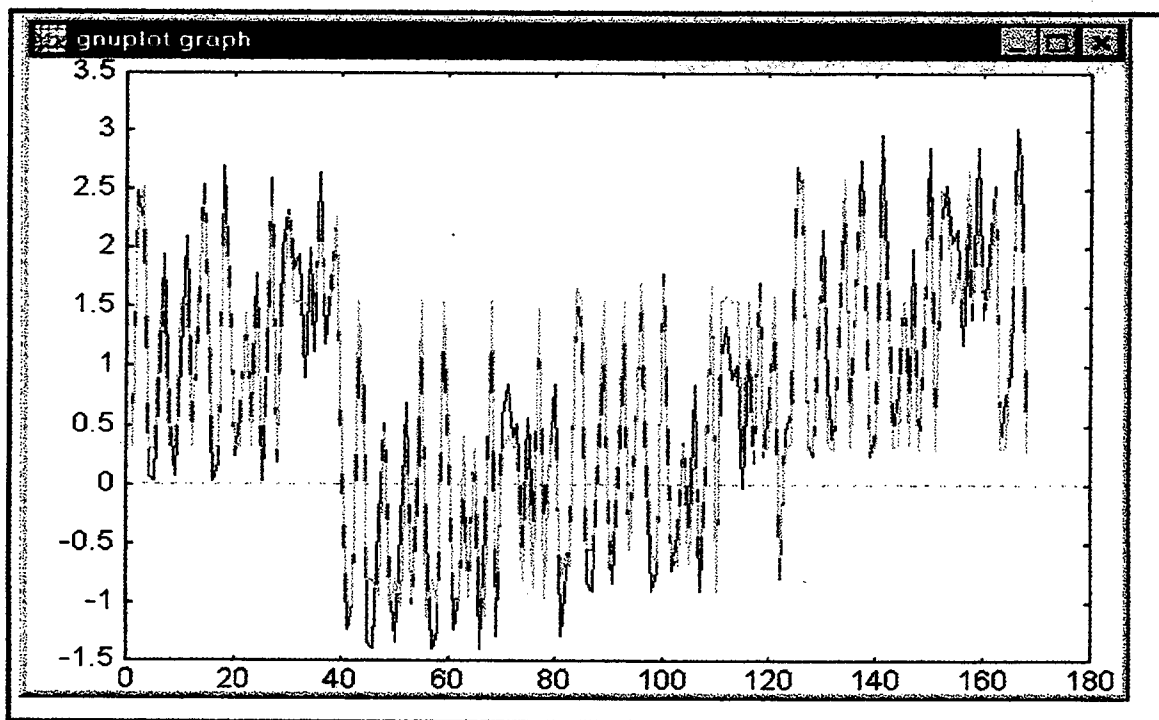


Fonte: do Autor

O quadro 7 mostra os resultados da função $f7$ (ver figura 30), iguais às do quadro 6. Para obter os dois últimos resultados do modelo proposto, o sistema foi simulado com 9 e 12 valores lingüísticos, respectivamente. A figura 31 mostra graficamente os erros da função $f7$.

(1)	(1)	(2)	(2)	(3)	(4)	(4)	(5)	(5)
EQ	#R	EQ	#R	EQ	EQ	#R	EQ	#R
0.34682	251	0.25605	170	0.21841	2.09409	49	0.41804	193
0.34682	251	0.26800	167	0.23429			0.36745	217
0.26803	346	0.21396	232	0.19523			0.29401	259
0.26803	346	0.232196	253	0.21017			0.09593	580

Quadro 7: Comparação dos Resultados da Função $f7$

Figura 31: Representação gráfica dos erros da função f_7 

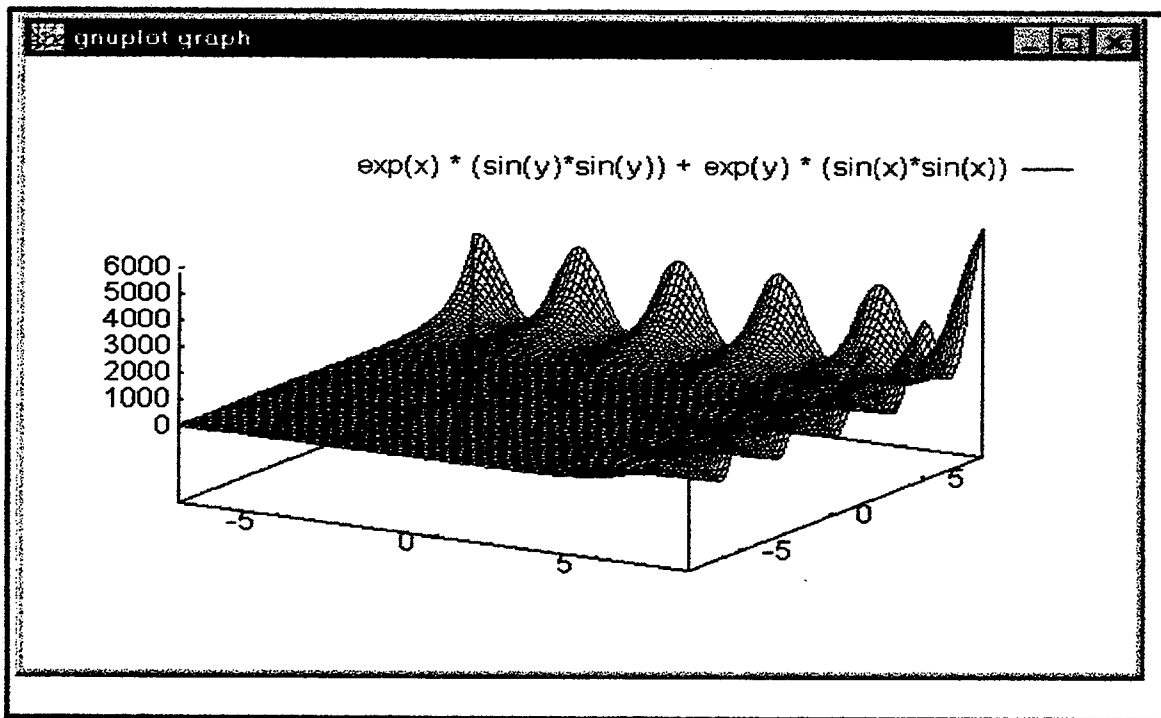
Fonte: do Autor

O quadro 8 mostra os resultados da função f_8 (ver figura 32) onde:

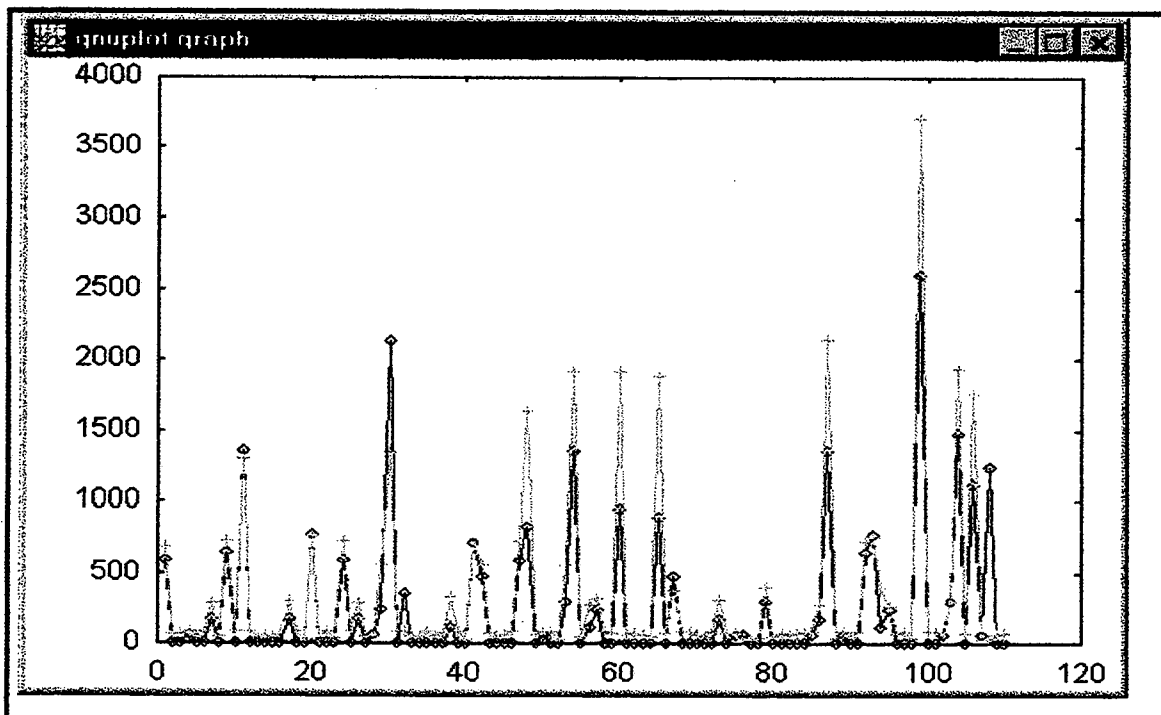
- (1) Cordón e Herrera (1996c), fase 1;
- (2) Cordón e Herrera (1996c), fase 2;
- (3) Cordón e Herrera (1996c), fase 3;
- (4) Wang e Mendel (1992);
- (5) Modelo Proposto, com 7 valores lingüísticos.

(1)	(1)	(2)	(2)	(3)	(4)	(4)	(5)	(5)
EQ	#R	EQ	#R	EQ	EQ	#R	EQ	#R
48970.9	84	22817.4	39	12083.8	159576.9	49	8341.7	130
		22893.7	40	8876.9			17951.1	73

Quadro 8: Comparação dos Resultados da Função f_8

Figura 32: Representação Gráfica da Função f_8 

Fonte: do Autor

Figura 33: Representação Gráfica dos Erros da Função f_8 

Fonte: do Autor

Os quadros 6, 7 e 8 mostram o resultado das funções $f6$, $f7$ e $f8$, respectivamente. A intenção destas funções é mostrar o comportamento do algoritmo com funções de dimensão maior do que nas simulações anteriores. Os resultados encontrados são notoriamente melhores do que os resultados encontrados pelo algoritmo proposto por Wang e Mendel (1992). Já quando comparamos com os resultados encontrados pelos algoritmos propostos por Cordón e Herrera (1996a, 1996b e 1996c), o algoritmo proposto apresenta resultados semelhantes em alguns casos, melhores e piores em outros.

Estes últimos resultados requerem uma análise mais minuciosa. Os algoritmos proposto por Cordón e Herrera (1996a, 1996b e 1996c), têm a vantagem considerável de ter três fases (especificamente são três processos evolutivos diferentes): uma primeira fase para encontrar a base de regras, uma segunda para simplificar a base de regras, e a terceira para ajustar as funções de pertinência. Cada fase têm uma representação diferente dos indivíduos da população e do objetivo do processo evolutivo, como foi especificado no item 2.6. Esta é uma vantagem considerável, a propor um só processo evolutivo que se encarrega de descobrir as funções de pertinência e as regras difusas.

Nas simulações realizadas neste trabalho, observou-se que o número de regras varia de simulação para simulação em função dos dados aleatórios utilizados no treinamento do algoritmo. Este fato fica claramente visualizado nos quadros 6, 7 e 8, onde se apresentam várias simulações, e o número de regras descobertas varia de forma notória, utilizando em alguns casos, o mesmo número de valores lingüísticos.

4.3 Sistemas Simulados

O sistema simulado é o de um processo de limpeza da água, testado por Takagi e Sugeno (1985). Os resultados obtidos com o modelo aqui proposto são comparados com os resultados obtidos por esses pesquisadores.

O processo de limpeza da água consiste basicamente da mistura, em um tanque, de água do rio com produtos químicos, chamados *PAC*. A água misturada flui para um tanque de sedimentação, de onde depois de algumas horas passa para outro tanque onde é filtrada

para produzir água limpa. O principal controle neste processo consiste na determinação do *PAC* a ser adicionado à água.

As variáveis que influem no processo e que podem ser medidas são:

TBI, turbidez da água original;

TB2, turbidez da água tratada;

PAC, total de *PAC* adicionado na água;

TE, temperatura da água;

PH, Ph da água;

CL, cloro da água;

Al, alcalinidade da água.

As variáveis consideradas nesta simulação são *PAC*, *TE*, *PH*, *Al*.

As simulações foram realizadas com 3 e 5 valores lingüísticos para cada variável, respectivamente. Foram gerados 600 dados de treino, utilizando o modelo estatístico fornecido por Takagi e Sugeno (1985):

$$PAC = 9.11 * (TBI)^{1/2} - 79.8 * PH + 12.7 * CL + 1255.6 \quad (36)$$

Os dados utilizados na geração dos novos dados são os fornecidos na Tabela II do trabalho de Takagi e Sugeno. A medida de erro utilizada é a média dos erros absolutos:

$$EAb = (1/n) * \sum_i |y^* - y| \quad (37)$$

No quadro 9, podem-se observar os resultados obtidos.

Modelo Estatístico	Modelo Difuso de Tagaki e Sugeno	Modelo Proposto (3 val. lingüísticos)	Modelo Proposto (5 val. lingüísticos)
128.0	48.5	74.95	39.66

Quadro 9: Comparação dos Erros de Diferentes Modelos Aplicados no Processo de Limpeza da Água.

O indivíduo escolhido pelo algoritmo, no caso de 3 valores lingüísticos para cada variável, encontrou 16 regras, e os seguintes valores, para cada variável do sistema :

PH: 6.90 - 7.05 - 7.30

TE: 16.60 - 19.17 - 23.60

AL: 35.00 - 52.64 - 55.00

PAC: 900.00 - 938.50 - 1400.00

O indivíduo escolhido pelo algoritmo, no caso de 5 valores lingüísticos para cada variável, encontrou 28 regras e os seguintes valores, para cada variável do sistema :

PH: 6.90 - 7.03 - 7.07 - 7.20 - 7.30

TE: 16.60 - 17.08 - 19.08 - 21.20 - 23.60

AL: 35.00 - 40.60 - 51.60 - 52.66 - 55.00

PAC: 900.00 - 920.50 - 1172.50 - 1326.50 - 1400.00.

As conclusões que podem ser obtidas do quadro 9 podem ser enumeradas da seguinte maneira. Foram realizadas dois tipos de simulações, com 3 e 5 valores lingüísticos para cada variável do sistema, respectivamente. O quadro compara os resultados nas duas simulações realizadas com o trabalho de Takagi e Sugeno. Neste caso, a comparação de resultados numéricos deve ser observada com cuidado, por vários motivos: os dados utilizados nas simulações são diferentes, Takagi e Sugeno utilizam dados reais, o modelo proposto gera dados em função da fórmula utilizada na solução estatística fornecida em Takagi e Sugeno (1985).

Outra diferença importante é que as regras utilizadas pelos dois algoritmos têm estruturas diferentes. Enquanto aqui utilizamos uma estrutura de regras do tipo Mandami (1977), o trabalho de Takagi e Sugeno (1985) utilizou um tipo de regras onde o conseqüente é uma combinação linear das variáveis do sistema. A simulação deste sistema é válida para conferir como se comporta o algoritmo em casos de sistemas de dimensões maiores.

CAPÍTULO V

CONCLUSÕES E PROBLEMAS EM ABERTO

5.1 Introdução

Neste capítulo, ressaltam-se os itens mais importantes do trabalho aqui apresentado. Destacam-se a contribuição científica do mesmo, assim como sua originalidade e os objetivos alcançados. Discutem-se os resultados obtidos nas simulações realizadas, apresentados no capítulo 4. As limitações do modelo e do algoritmo apresentado no capítulo 3 são enumerados, assim como os trabalhos complementares, que podem aperfeiçoar este trabalho.

Este capítulo foi dividido em duas partes: numa primeira, discutem-se de forma geral os resultados obtidos nas diferentes simulações realizadas, os objetivos alcançados descritos no item 1.3, e enumeram-se as limitações do algoritmo. Numa segunda parte, enumeram-se os possíveis trabalhos futuros.

5.2 Resultados, Objetivos Alcançados e Limitações

Como conclusão final dos resultados obtidos, válida para todas as simulações mostradas nos quadros do capítulo anterior, pode-se afirmar que o algoritmo proposto foi satisfatório em seus resultados, quando testado em diferentes tipos de comportamentos. Foram simuladas funções lineares, não lineares, senóides, unimodais e multimodais, para observar como o algoritmo consegue responder aos diferentes comportamentos das funções.

Os quadros apresentados no capítulo anterior também mostram que os resultados melhoram com o aumento dos valores lingüísticos, aumentando o número de regras descobertas pelo algoritmo, melhorando notoriamente a precisão da identificação encontrada pelo algoritmo.

A discussão e comparação de resultados com outros métodos, como foi apresentado no capítulo 4, embora fundamental do ponto de vista da análise de performance do modelo e do algoritmo, não deveria esconder o fato de que sua abordagem apresenta não uma forma competitiva de identificar sistemas, mas sim um novo paradigma de identificação.

As comparações dos resultados do algoritmo proposto com outros trabalhos foram feitas com a intenção de mostrar que o modelo proposto é outra alternativa viável a ser utilizada no problema da identificação difusa de sistemas. Os resultados numéricos das diferentes simulações confirmam este propósito, já que eles se apresentaram satisfatórios e semelhantes aos resultados apresentados por outros algoritmos.

Isto dá lugar a afirmar que o modelo de regras que descobre o modelo proposto é válido. Desta maneira, o objetivo principal do trabalho, isto é, obter um modelo formal, foi alcançado, baseado em regras de um sistema como alternativa a um modelo matemático.

Isto mostra que a utilização dos processos adaptativos e a lógica difusa, para encontrar as relações entre os agentes do sistema e o comportamento dos mesmos, de uma forma aproximada é válida. Esta validade é mostrada através dos resultados obtidos e da comparação destes resultados com os obtidos por outros trabalhos.

Os resultados do capítulo anterior mostram também que a idéia da utilização do processo de aprendizado interagindo com o processo evolutivo ajuda na convergência do algoritmo. A contribuição deste trabalho neste ponto, fundamenta-se no método de aprendizado utilizado, diferente dos métodos apresentados em outros trabalhos, como foi mostrado no item 2.7. Desta forma foi alcançado outro objetivo proposto de forma primordial.

O algoritmo proposto pode ser utilizado em diversos tipos de sistemas, como foi observado pelos exemplos testados, que englobam funções com diferentes tipos de comportamento, funções de várias variáveis e outras; no entanto, o algoritmo tem várias restrições como é descrito a seguir. Observe-se que as restrições são referentes ao algoritmo implementado, e não ao modelo proposto.

Para ser utilizado, o algoritmo necessita de um banco de dados com informações de entrada-saída sobre o sistema a identificar. O algoritmo encontra as relações entre as variáveis do sistema, a partir do banco de dados, o que pressupõe que tais relações entre as variáveis já existe. Isto significa que o algoritmo não identifica variáveis não representativas ou supérfluas. O banco de dados deve ser representativo do domínio das variáveis do sistema, já que as regras geradas devem cobrir todo o domínio das variáveis.

A divisão do domínio das variáveis em valores lingüísticos deve ser fornecida pelo usuário, como parâmetro de entrada. O desempenho do algoritmo está diretamente vinculado a esta divisão, devido ao esquema de geração das regras. Por este motivo, este parâmetro pode ser considerado como uma restrição do algoritmo. Isto não significa que o usuário deva ter um conhecimento prévio do sistema, significa que o modelo não descobre o número ótimo de valores lingüísticos para cada sistema identificado.

A necessidade de gerar todas as regras possíveis para cada dado do ambiente e de escolher aquele modelo de melhor resultado, é uma restrição importante do algoritmo, já que este número de regras cresce de forma exponencial, com o aumento de variáveis do sistema.

A representação do código genético dos indivíduos, a utilização dos métodos de inferência, do método de *fuzzification* e *defuzzification* não invalidam o modelo proposto, já que é possível utilizar outros métodos para estes módulos sem modificar a estrutura do modelo.

Como conclusão final, é possível afirmar que o algoritmo proposto encontra um modelo formal do sistema identificado. Esta aproximação fornecida pelo modelo de regras difusas pode ser considerada aceitável, como mostram os resultados apresentados. Assim, o modelo consegue auxiliar no problema da representação e aquisição do comportamento de agentes e relações que atuam sobre o sistema.

5.3 Trabalhos Futuros

São vários os trabalhos que podem ser realizados no sentido de melhorar o algoritmo proposto. Estas melhoras vão desde o aperfeiçoamento de alguns módulos do algoritmo, até a simulação de sistemas com outras características que não as dos sistemas aqui apresentados.

O processo de aprendizado utilizado foi simulado através das probabilidades de operadores de mutação individualizadas para cada característica do indivíduo. No entanto, existem outras formas de implementar o processo de aprendizado, como foi descrito no item 2.7, que podem ser testadas. Além disso, as novas idéias sobre as mutações e novos modelos de evolução que surgem na área biológica, certamente poderão induzir a novos algoritmos de aprendizado.

Foi utilizado um esquema de representação do indivíduo no cromossomo, de tal forma que as funções de pertinência têm certas particularidades, como serem triangulares e com seu ponto de corte no valor 0.5. Outras representações do indivíduo podem ser implementadas e testadas.

Além disso, poderão ser realizadas simulações de sistemas que possuam outras características, que os classificam como de difícil identificação (segundo foi descrito no item 1.2), como, por exemplo, sistemas que possuem competição e cooperação de indivíduos em um ambiente. Nestes casos, a experiência e o aprendizado obtidos no tempo de vida de um indivíduo podem ser muito importantes para as futuras gerações. A simulação de sistemas que tenham estas características deve ser realizada para se obter conclusões mais precisas sobre os métodos de aprendizado e sua interação com a evolução.

No item 3.2, foi descrita uma especificação formal de sistemas, onde se afirmou que, para determinados valores de estímulos, tem-se uma única resposta (função $n \rightarrow I$, equação 12). No entanto, a geração de múltiplas respostas para o mesmo estímulo é resultado do fato de que o comportamento do agente não só depende dos valores dos estímulos, mas também de sua história. A teoria de estados é capaz de expressar a dependência do comportamento dos agentes através da sua história em um caminho que poderia gerar uma determinada

relação estímulo-resposta, e retém o caráter dinâmico do comportamento do agente. Assim, constrói-se uma descrição espaço-estado do agente. Esta descrição tem duas partes básicas, construídas pelo observador, uma variável de estado e um estado relação estímulo-resposta. Esta alternativa poderia ser abordada em trabalho futuro.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ACKLEY, D., LITTMAN, M. Interactions between learning and evolution. In: LANGTON, C.G., TAYLOR, C., FARMER, J.D., RASMUSSEN, S. (Ed.). **Artificial Life II**. Reading, MA: Addison-Wesley, 1990. p. 487-507.
2. BAK, P., CHEN, K. Self-Organized criticality. **Scientific American**, p. 26-33, Jan. 1991.
3. BALDWIN, J.M. A new factor in evolution. **American Naturalist**, v. 30, p. 441-451, 1896.
4. BARR, A., FEIGENBAUM, E.A. **The Handbook of Artificial Intelligence**. Reading, MA: Addison-Wesley, 1982. v. 2.
5. BEASLEY, D., BULL, D.R., MARTIN, R.R. An overview of genetic algorithms: Part 1, Fundamentals. A genetic algorithm tutorial. **Technical Report CS-93-103**, Colorado: Colorado State University, Nov. 1993.
6. BEASLEY, D., BULL, D.R., MARTIN, R.R. An overview of genetic algorithms: Part 2, Research Topics. A genetic algorithm tutorial. **Technical Report CS-93-103**, Colorado: Colorado State University, Nov. 1993.
7. BELEW, R. K. When both individuals and populations search: Adding simple learning to the genetic algorithm. In: INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 3. **Proceedings...** Washington D.C.: George Mason University, Morgan Kaufmann Publishers, 1989. p. 34-41.
8. BOOKER, L.B., GOLBERG, D.E., HOLLAND, J.H. Classifier systems and genetic algorithms. **Artificial Intelligence**, n. 40, p. 235-282, 1989.
9. BOUCHON, B. On the forms of reasoning in expert systems. **Approximate Reasoning in Expert Systems**, Netherlands: North-Holland, p. 341-354, 1985.
10. BRUNS, R. Direct chromosome representation and advanced genetic operators for production scheduling. INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 5. **Proceedings...** Pittsburgh, PA: Lawrence Erlbaum, 1993. p. 352-359.
11. BUCHANAN, B.G., SHORTLIFFE, E.H. **Rule based expert systems**. Reading, MA: Addison-Wesley, 1984.
12. CAIRNS, J., OVERBAUGH, J., MILLER, S. The origin of mutants. **Nature**, v. 335, p. 142-145, 1988.

13. CASWELL, H. Systems analysis and simulation in ecology. In: PATTERN, BERNARD C. **An introduction to systems science for ecologists**. London: Academic Press, 1972. v. 2, p. 1- 78.
14. CORDÓN, O., HERRERA, F., HERRERA-VIDEIRA, E., LOZANO, M. Genetic algorithms and fuzzy logic in control processes. **Technical Report #DECSAI-96109**, Granada, Espanha: E.T.S. de Ingeniería Informática, Universidad de Granada, Mar. 1995.
15. CORDÓN, O., HERRERA, F. Fuzzy identification by means of genetic algorithms. **Technical Report #DECSAI-96107**, Granada, Espanha: E.T.S. de Ingeniería Informática: Universidad de Granada, Mar. 1996
16. CORDÓN, O., HERRERA, F. A hybrid genetic algorithm: evolution strategy process for learning fuzzy logic controller knowledge bases. In: HERRERA, F., VERDEGAY J.L. (Ed.). **Genetic Algorithms and Soft Computing**. Berlin: Physica-Verlag, 1996. p. 251-278.
17. CORDÓN, O., HERRERA, F. Hybridizing genetic algorithms with sharing scheme and evolution strategies for designing approximate fuzzy logic controllers. **Technical Report #DECSAI-96126**, Granada, Espanha: E.T.S. de Ingeniería Informática, Universidad de Granada, Dec. 1996.
18. CROSBY, R.W. Toward a classification of complex systems. **European Journal of Operacional Research**, n. 30, p. 291-293, 1987.
19. DAVIS, L. **HandBook of genetic algorithms**. New York: Van Nostrand Reinhold, 1991.
20. DELGADO, M., GONZALEZ, A. An inductive learning procedure to identify fuzzy systems. **Fuzzy Sets and Systems**, n. 55, p. 121-132, 1993.
21. DUBOIS, D., PRADE, H. **Fuzzy sets and systems: theory and applications**. London: Academic Press, 1980.
22. ESPY, T., VOMBRACK, E., ALDRIDGE, J. Application of genetic algorithms to tuning fuzzy control systems. In: INTERNATIONAL WORKSHOP ON NEURAL NETWORKS AND FUZZY LOGIC, 3. **Proceedings...** Houston, Texas: NASA, 1992. p. 237-248 LOCAL
23. FARMER, J. D., PACKARD, N. H. The immune system, adaptation, and machine learning. **Physica D**, v. 22 , p. 187-204, 1986.
24. FARMER, J.D., PACKARD, N.H. Evolution, games and learning: models for adaptation in machines and nature. **Physica D**, v. 22, p. 7-12, 1986.

25. FOGEL, D.B. **System identification through simulated evolution: a machine learning approach to modeling.** Needham Heights, MA: Ginn Press, 1991.
26. FONSECA, C.M., FLEMING, P.J. Genetic algorithms for multiobjective optimization: formulation, discussion and generalization. In: INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 5. **Proceedings...** Pittsburgh, PA: Lawrence Erlbaum, Morgan Kaufmann Publishers, 1993. p. 416-423.
27. FORREST, S. Emergent computation: self-organizing, collective and cooperative phenomena in natural and artificial computing networks. **Physica D**, v. 42, p. 1-11, 1990.
28. FRENCH, R., MESSINGER, A. Genes, phenes and the Baldwin effect. In: BROOKS, R., MAES, P. (Ed.). **Artificial Life IV.** Cambridge, MA: MIT Press, 1994.
29. GOLDBERG, D.E. **Genetic algorithms in search, optimization, and machine learning.** Reading, MA: Addison-Wesley, 1989.
30. GOLDBERG, D.E., DEB, K. An investigation of niche and species formation in genetic function optimization. In: INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 3. **Proceedings...** Washington D.C.: George Mason University, 1989. p. 42-50.
31. GOLDBERG, D. E., KUO, C. H. Genetic algorithms in pipeline optimization. **Journal of Computing in Civil Engineering**, v. 1, n. 2, p.128-141, 1987.
32. GOODMAN, B. Directed mutations: heredity made to order. **Mosaic**, v. 23, n. 1, p. 24-33, 1992.
33. GRUAU, F., WHITLEY, D. Adding learning to the cellular development of neural networks: evolution and the Baldwin effect. **Computation**, v. 1, p. 213-233, 1993.
34. HAYES-ROTH, F., WATERMAN, D.A., LENAT, D.B. **Building Expert Systems.** Reading, MA: Addison-Wesley, 1983.
35. HERRERA, F., LOZANO, M., VERDEGAY, J.L. A learning process for fuzzy control rules using genetic algorithms. **Technical Report #DECSAI-95108**, Granada, ES: E.T.S. de Ingeniería Informática: Universidad de Granada, Feb. 1995.
36. HIGGINS, C.M., GOODMAN, R.M. Fuzzy rule-based networks for control. **IEEE Transactions on Fuzzy Systems**, v. 2, n. 1, p.82-88, Feb. 1994.

37. HINTON, G.E., NOWLAN, S.J. How learning can guide evolution. **Complex Systems**, v. 1, p. 495-502, 1987.
38. HOLLAND, J.H. **Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control and artificial intelligence.** Cambridge, MA: MIT Press, Bradford Books, 1992.
39. HOLLAND, J.H., HOLYOAK, K.J., NISBETT, R.E. THAGARD, P.R. **Induction: processes of inference, learning, and discovery.** Cambridge MA: MIT Press: Bradford Books, 1993.
40. HOLLAND, J.H. **Hidden orden: how adaptation builds complexity.** Reading, MA: Helix Books : Addison-Wesley, 1995.
41. HOMAIFAR, A., GUAN, S., LIEPINS, G. A new approach on the traveling salesman problem by genetic algorithms. In: **INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 5. Proceedings...** Pittsburgh, PA: Lawrence Erlbaum, Morgan Kaufmann Publishers, 1993. p. 460-466.
42. ISHIBUCHI, H., NOZAKI, K., YAMAMOTO, N., TANAKA, H. Selecting fuzzy: if then rules for classification problems using genetic algorithms. **IEEE Transactions on Fuzzy Systems**, v. 3, n. 3, p. 260-270, Aug. 1995.
43. KARR, C.L. Genetics algorithms for fuzzy controllers. **A.I. Expert**, v.6, n.2, p.26-33, Feb. 1991.
44. KARR, C.L. Appling genetics algorithms to fuzzy logic. **A.I. Expert**, v.6, n.3, p.38-43, Mar. 1991.
45. KARR, C.L. Adaptive process control using fuzzy logic and genetic algorithms. **INTERNATIONAL WORKSHOP ON NEURAL NETWORKS AND FUZZY LOGIC, 3. Proceedings...** Houston, Texas: NASA, 1992. p. 186-193.
46. KARR, C.L. Designing precise fuzzy systems with genetic algorithms. In: HERRERA, F., VERDEGAY J.L. (Ed.). **Genetic algorithms and soft computing.** Berlin: Physica-Verlag, 1996. p. 331-348.
47. KAUFFMAN, S.A. Antichaos and Adaptation. **Scientific American**, p. 64-70, Aug. 1991.
48. KICKERT, J.M., VAN NAUTA LANKE, H.R. Application of a fuzzy controller in a warm water plant. **Automatica**, v. 12, p. 301-308, 1976.

49. KIDWELL, M.D. Using genetic algorithms to schedule distributed tasks on a bus-based system. In: **INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 5. Proceedings...** Pittsburgh, PA: Lawrence Erlbaum, Morgan Kaufmann Publishers, 1993. p. 368-374.
50. KOSKO, B. **Neural Networks and Fuzzy Systems**. Englewood Cliffs, New Jersey: Prentice Hall, 1991.
51. KOSKO, B. **Fuzzy thinking: the new science of fuzzy logic**, hyperion. New York, NY: Hiperion, 1992.
52. KHURL, S., BÄCK, T., HEITKÖTTER, J. An evolutionary approach combinatorial optimization problems. A genetic algorithm tutorial. **Technical Report CS-93-103**, Colorado: Colorado State University, Nov. 1993.
53. LEE, C.C. Fuzzy logic in control systems: fuzzy logic controller, Part I. **IEEE Transactions on Systems, Man and Cybernetics**, v.20, n.2, p. 404-418, Mar./Apr. 1990.
54. LEE, C.C. Fuzzy logic in control systems: Fuzzy logic controller, Part II. **IEEE Transactions on Systems, Man and Cybernetics**, v.20, n.2, p. 419-432, Mar./Apr. 1990.
55. LENSKI, R.E., MITTLER, J.E. The directed mutation controversy and neo-darwinism. **Science**, v. 259, p. 188-194, 1993.
56. LJUNG, L. **System identification: Theory for the user**. Englewood Cliffs, New Jersey: Prentice Hall, 1987.
57. MANDAMI, E.H. Applications of fuzzy logic to approximate reasoning using linguistic synthesis. **IEEE Transactions on Systems, Man and Cybernetics**, v. c-26, n. 12, p. 1182-1191, 1977.
58. MAYLEY, G. Landscapes, learning costs and genetic assimilation. **Evolutionary Computation**, v.4, n.3, p. 213-234, 1996.
59. MIZUMOTO, M., ZIMMERMANN, H.J. Comparison of fuzzy reasoning methods. **Fuzzy Sets and Systems**, n. 8, p. 253-283, 1982.
60. MUNAKATA, T., HASHIER, D.J. A genetic algorithms applied to the maximum flow problem. In: **INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 5. Proceedings...** Pittsburgh, PA: Lawrence Erlbaum, Morgan Kaufmann Publishers, 1993. p. 488-493.

61. PAPPIS, C.P., MAMDAMI, E.H. A fuzzy logic controller for a traffic junction. **IEEE Transactions on Systems, Man and Cybernetics**, v. 7, n. 10, p. 707-717, 1977.
62. PARODI, A., BONELLI, P. A new approach to fuzzy classifier systems. In: **INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 5. Proceedings...** Pittsburgh, PA: Lawrence Erlbaum, Morgan Kaufmann Publishers, 1993. p. 223-230,
63. RAHMANI, A.T., ONO, N. A genetic algorithm applied to the maximum flow problem. In: **INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 5. Proceedings...** Pittsburgh, PA: Lawrence Erlbaum, Morgan Kaufmann Publishers, 1993. p. 494-498.
64. RAWLINS, G. **Foundations of genetic algorithms**. Los Altos, CA: Morgan Kaufmann Publishers, 1991.
65. ROSEN, R. On complex systems. **European Journal of Operational Research**, n. 30, p. 129-134, 1987.
66. RUTHEN, R. Trends in nonlinear dynamics: adapting to complexity. **Scientific American**, p.110-117, Jan. 1993.
67. SASAKI, T., AKIYAMA, T. Traffic control process of expressway by fuzzy logic. **Fuzzy Sets and Systems**, n. 26, p. 165-178, 1988.
68. SMITH, J.M. When learning guides evolution. **Nature**, v. 329, p. 761-762, 1987.
69. SONCINI, M.I. A evolução das idéias evolucionistas. **Revista de Ensino de Ciências**, n.24, p. 4-12, 1993.
70. SUDJIANTO, A., WASSERMAN, G.S., SUDARBO, H. Genetic subsets regression. **Computers and Industrial Engineering, an International Journal**, v. 30, n. 4, p. 839-849, Sep. 1996.
71. SUGENO, M., KANG, G.T. Structure identification of fuzzy model. **Fuzzy Sets and Systems**, n.28, p. 15-35, 1988.
72. SUGENO, M., TANAKA, K. Successive identification of a fuzzy model and its applications to prediction of a complex system. **Fuzzy Sets and Systems**, n. 42, p. 315-334, 1991.
73. TAKAGI, T., SUGENO, M. Fuzzy identifications of systems and its applications to modeling and control. **IEEE Transactions on Systems, Man and Cybernetics**, v. 15, n.1, p. 116-132, 1985.

74. TCHOLAKIAN, A.B. **Modelagem lingüística: alternativa na análise de processos complexos.** Florianópolis, 1992. Dissertação (Mestrado em Engenharia da Produção) - Universidade Federal de Santa Catarina, 1992.
75. TCHOLAKIAN, A.B. Construção de funções de pertinência para controladores lógicos difusos via algoritmos genéticos. In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, 1., 1993, Rio Claro, S.P. **Anais ...** São Paulo: UNESP, Sociedade Brasileira de Automática, 1993. p. 273-281.
76. TCHOLAKIAN, A.B. Otimização de funções multimodais via algoritmos genéticos. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 14., 1994, Florianópolis, S.C. **Anais...** Florianópolis: Ed. da UFSC, 1994. p. 727-732.
77. TCHOLAKIAN, A.B. Um algoritmo para a construção de funções de pertinência via algoritmos genéticos. In: SIMPOSIO BRASILEIRO DE PESQUISA OPERACIONAL, 1994, Florianópolis, SC. **Anais...** Rio de Janeiro: Sociedade Brasileira de Pesquisa Operacional, 1994. p. 111-116.
78. THALER, D.S. The evolution of genetic intelligence. **Science**, v. 64, p. 224-225, 1994.
79. VALENZUELA-RENDÓN, M. The fuzzy classifier system: a classifier system for continuously varying variables. INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 4. **Proceedings...** San Diego: Morgan Kaufmann Publishers, 1991. p. 346-353.
80. VERGARA, V., MORAGA, C. Optimal fuzzy identification models. In: INTERNATIONAL FUZZY SYSTEMS ASSOCIATION WORLD CONGRESS, 6. 1995, São Paulo, SP. **Anais...** [S.l.]: International Fuzzy System Association, 1995. v. 1, p. 109-112.
81. WANG, L., MENDEL, J.M. Generating fuzzy rules by learning from examples. **IEEE Transactions on Systems, Man and Cybernetics**, v.22, n.6, 1992.
82. WATERMAN, D.A. **A guide to expert systems.** Reading, MA: Addison-Wesley, 1986.
83. WHITLEY, D. A genetic algorithm tutorial. **Technical Report CS-93-103.** Colorado: Colorado State University, 1993.
84. WHITLEY, D., GORDON, V.S., MATHIAS, K Lamarckian evolution, the Baldwin effect and function optimization. In: Davidor, Y., Schwefel, H.P., Manner, R. (Ed.). **Parallel Problem Solving from Nature III.** Berlin: Springer-Verlag, 1984. p. 6-15.

85. YAGER, R.R., FILEV, D.P. **Essentials of fuzzy modeling and control.** [S.l]: Wiley: Interscience, 1994.
86. YEN, J., GILLESPIE, W. Integrating global and local evaluations for fuzzy model identifications using genetic algorithms. In: **INTERNATIONAL FUZZY SYSTEMS ASSOCIATION WORLD CONGRESS, 6., 1995, São Paulo. Anais...** [S.l]: International Fuzzy System Association, 1995. v. 1, p. 109-112.
87. YU, C., CAO, Z., KANDEL, A. Application on fuzzy reasoning to the control of an activated sludge plant. **Fuzzy Sets and Systems**, n.38, p. 1-38, 1990.
88. ZADEH, L.A. Outline of a new approach to the analysis of complex systems and decision processes. **IEEE Transactions on Systems, Man and Cybernetics**, v. 3, n. 1, p. 28-44, 1973.
89. ZADEH, L.A. The concept of a linguistic variable and its application to approximate reasoning, part I. **Information Sciences**, n.8, p.199-249, 1975.
90. ZADEH, L.A. The concept of a linguistic variable and its application to approximate reasoning, part II. **Information Sciences**, n.8, p. 301-357, 1975.
91. ZADEH, L.A. The concept of a linguistic variable and its application to approximate reasoning, part III. **Information Sciences**, n.9, p. 43-80, 1975.
92. ZADEH, L.A. The role of fuzzy logic in the management of uncertainty in expert systems. **Approximate reasoning in expert systems**, Netherlands: North-Holland, p. 3-31, 1985.