

Apresentação

[Identificação](#)

[Introdução](#)

[Objetivos](#)

Fala

[Som, Voz Humana, Fala](#)

Rec. de Fala

[Introdução](#)

[Fases](#)

[Problemas](#)

Redes Neurais

[Introdução](#)

Pré-Proc.

[Introdução](#)

[Média](#)

[FFT](#)

Rbf-Fuzzy Artmap

[Introdução](#)

[Estrutura](#)

[Funcionamento](#)

[Vantagens X Desvantagens](#)

Automation OLE

[Introdução](#)

[Tipos](#)

[Modos de Acesso](#)

Protótipo

[Introdução](#)

[Classes Automation OLE](#)

[Ferramenta RnaTools 1.0v](#)

[Aplicação Exemplo](#)

Resultados

[Demonstração](#)

Conclusões

[Conclusões](#)

Universidade Regional de Blumenau Centro de Ciências Exatas e Naturais Curso de Ciências da computação

Classes Automation OLE para reconhecimento de fala através da Rede Neural Artificial RBF-Fuzzy Artmap

Arnoldo Uber Junior

M. Eng. Marcel Hugo (Orientador)

Blumenau, Julho/2001

[◀ Voltar](#)

[Avançar ▶](#)

- **Palavras “ouvidas” por computador**
- **Síntese da fala**
- **Cotidiano dos desenvolvedores**

Facilitar a utilização de RNA aplicada ao reconhecimento de fala no desenvolvimento de software e utilização

- **Desenvolver classes que fazem o interfaceamento entre a base de conhecimento e a aplicação do usuário;**
- **Desenvolver uma ferramenta para treinamento da RNA;**
- **Desenvolver um protótipo demonstração do uso das classes e da ferramenta de treinamento;**
- **Apresentar um estudo referente à técnica implementada na ferramenta com vantagens e desvantagens da RNA;**
- **Possibilitar a um desenvolvedor leigo em RNA e reconhecimento de fala, adicionar estes recursos em suas aplicações.**

O som é a forma de energia que se apresenta como resultado de um movimento vibratório rápido, transmitida através de ondas em um meio elástico, e caracterizando-se principalmente por uma sensação sonora. Possui três qualidades fundamentais: tom ou altura, intensidade e timbre.

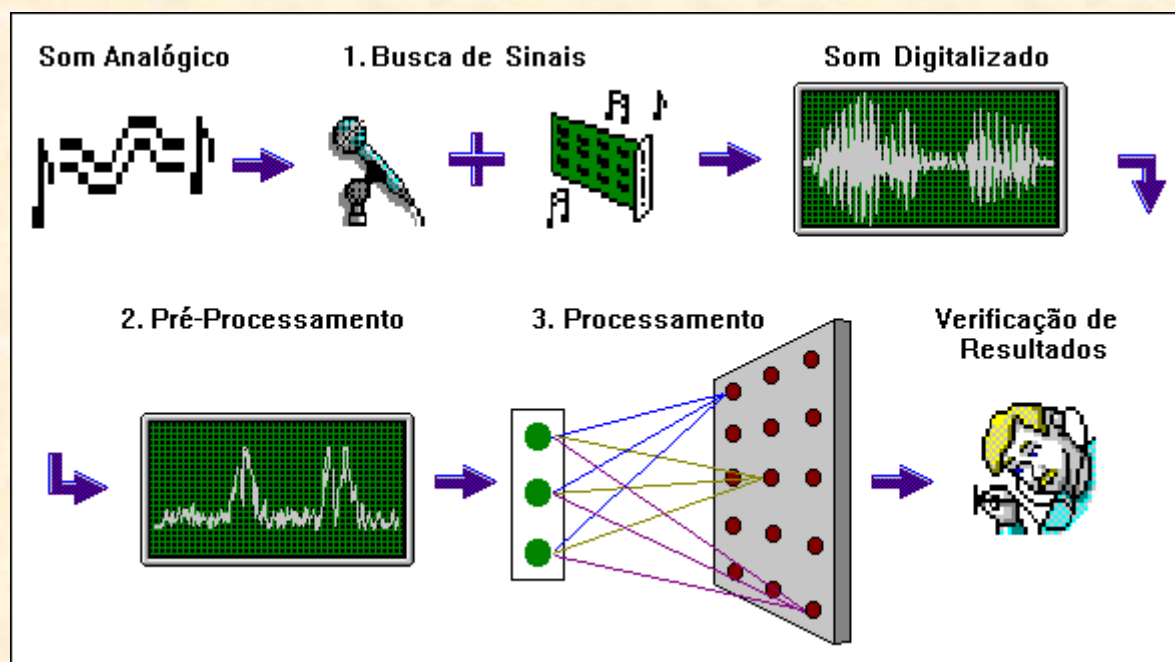
A voz é um elemento da linguagem; é a produção que o ser humano faz de sons através das cordas vocais. É o elemento sonoro da comunicação

A fala é a produção dos fonemas com articulação na palavra, onde articulação é o contato que os órgãos fonatórios tem, resultando em sons diferentes que se combinam formando a comunicação verbal

O reconhecimento de fala é o processo de extração da informação lingüística do sinal de voz

- **Início em 1950**
- **Interfaces**
- **Utilização da fala**
- **1990 primeiro sistema de rec. de fala com RNA**

Fases adotadas



Fonte: Hugo (1995)

- **Diferentes formas de onda de uma mesma palavra;**
- **Segmentação da fala;**
- **Tamanho do dicionário;**
- **Diferenças acústicas não lineares no tempo (ritmo), em frequência (timbre) e em amplitude (intensidade);**
- **Insuficiência de toda a informação lingüística (sotaque, erros de português) ;**
- **Dependência X Independência de locutor.**

“Aptidão ou capacidade de um dispositivo para desempenhar funções que são normalmente associadas à inteligência humana, tais como raciocínio, aprendizagem e auto-aperfeiçoamento”. Observando e procurando reproduzir a natureza, a IA alcança êxito, paradoxalmente, sendo natural (Xerez, 1996).

RNAs são sistemas computacionais, de implementação em hardware ou software, que imitam as habilidades computacionais do sistema nervoso biológico, usando um grande número de simples neurônios artificiais interconectados (Loesch, 1996).

O pré-processamento de sinais é uma forma de organizar ou ajustar os sinais captados procurando torná-los passíveis de processamento na fase posterior do processo de reconhecimento de fala. Ele é responsável por gerar o vetor característico do padrão a ser analisado (Hugo, 1995).

Propostas:

- Identificar os sinais mais significativos (envoltória);
- Diminuir a quantidade de sinais para compor a entrada da RNA;
- Melhorar a representatividade do sinal em relação ao sinal original.

Etapas:

- Eliminação do ciclo negativo do sinal amostrado;
- Redução do sinal amostrado detectando forma de onda;
- Mediação do sinal reduzido;
- Normalização do sinal mediado.



A FFT (Fast Fourier Transform) tem condições de identificar ou distinguir as senóides de diferentes frequências e suas respectivas amplitudes que são combinadas para formar a onda sonora (Hugo, 1995).

Moser (1997) considera os resultados provindos da utilização de FFT insatisfatórios por não revelarem características marcantes para cada sinal individualmente, não justificando portanto a insistência de sua aplicação em um sistema de reconhecimento de fala independente de locutor.

A RNA RBF-Fuzzy Artmap apresentada por Tontini (1995), é composta pelas RNAs: RBF e Fuzzy Artmap

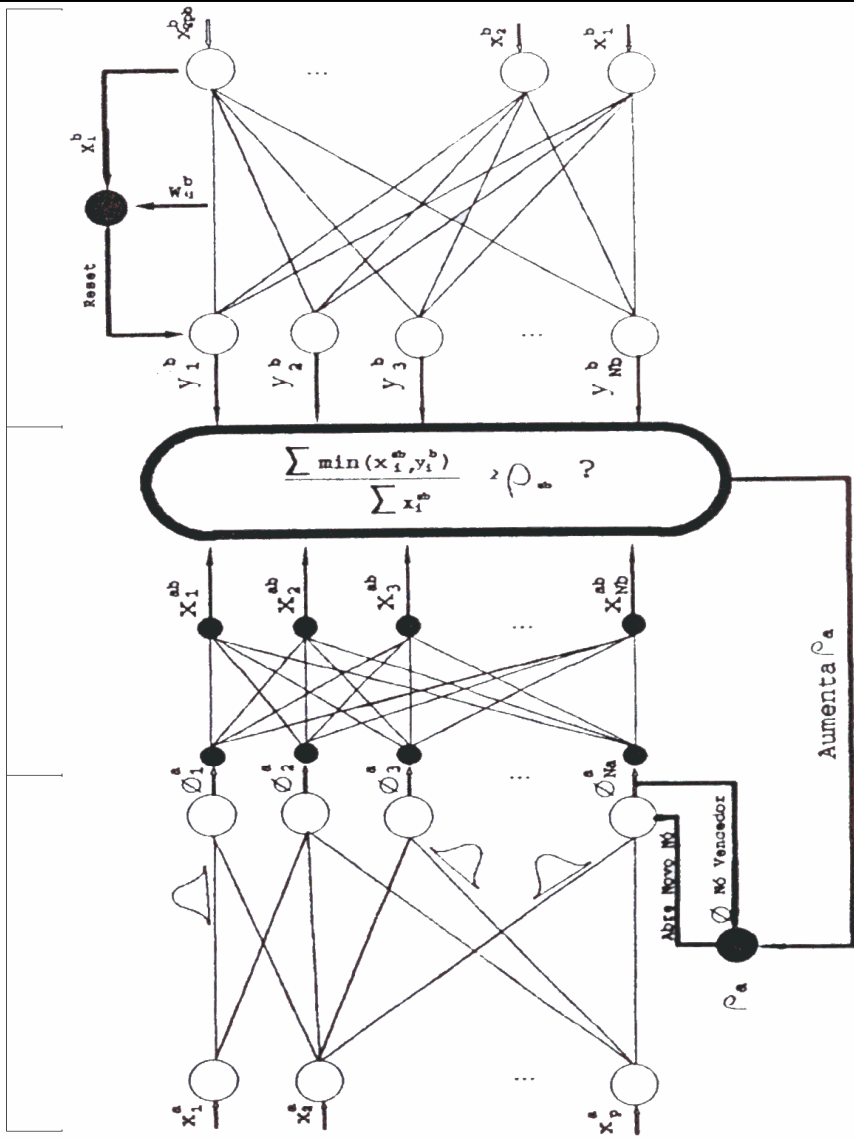
Elimina as limitações apresentadas pelas RNAs RBF e Fuzzy Artmap

A capacidade de aprendizado instantâneo e incremental, apresentada pela RNA Fuzzy-Artmap, tornando-a uma forte candidata à aplicação em sistemas de reconhecimento de padrões (Exemplo : reconhecimento de fala).

FUZZY - ARTb

MODULO DE LIGAÇÃO

REDE RBF



Vantagens:

- Sua arquitetura permite a incorporação das qualidades da rede RBF na rede Fuzzy Artmap, diminuindo a sensibilidade a ordem dos padrões de treinamento;
- Capacidade de aprendizado instantâneo.

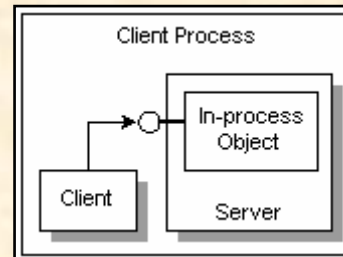
Desvantagem:

- Necessidade de um número maior de nós, na camada RBF, do que a rede Fuzzy-Artmap necessita na rede ARTa.

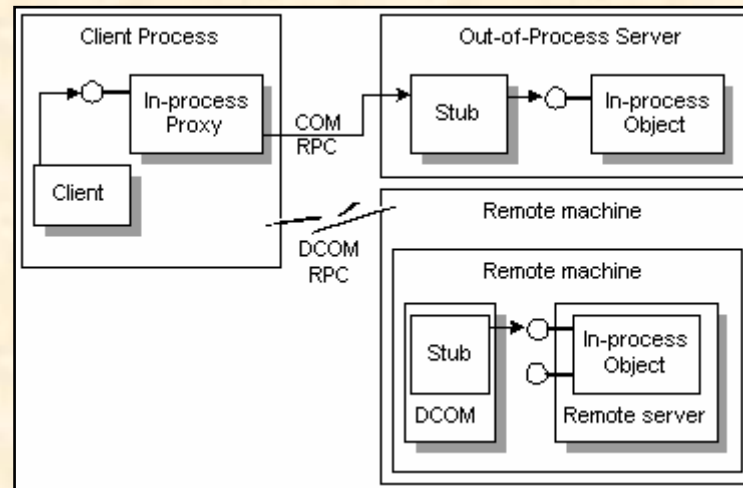
São aplicações ou bibliotecas que expõem suas funcionalidades para serem chamadas por aplicações clientes (*Automation Controllers*). As aplicações clientes podem ser qualquer aplicação que suporte *Automation OLE*, podendo ser escritas por exemplo, em ambiente Delphi, Visual Basic ou C++ Builder (Borland, 1999).

Existem três tipos de Automation OLE:

- **No processo;**



- **Fora do processo;**
- **Remota.**

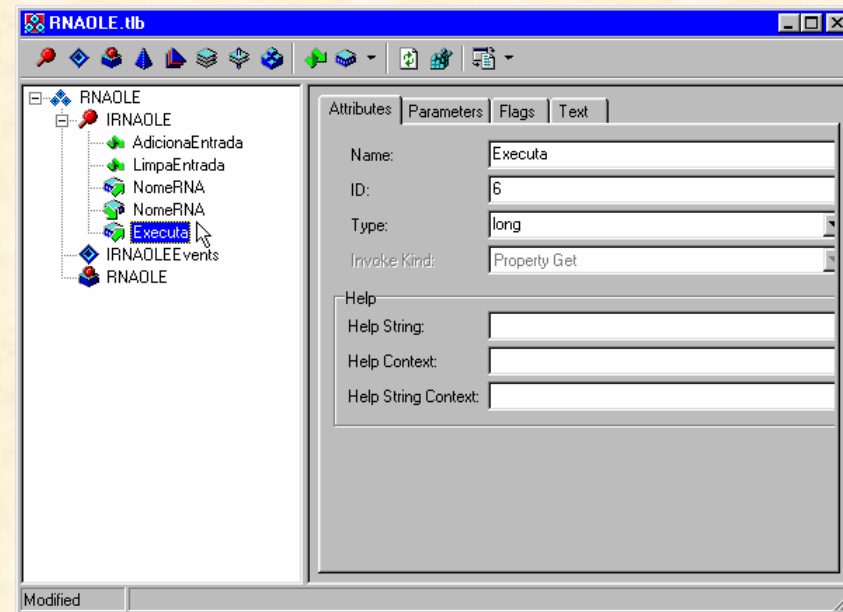


Existem duas maneiras de acessar um servidor Automation OLE:

- Tipo de dados variant;

```
var  
    Varw: Variant;  
Begin  
    Varw := CreateOleObject('PrototipoRF.Create');  
    Varw.Inicializa;  
    Varw.Reconhece;  
End;
```

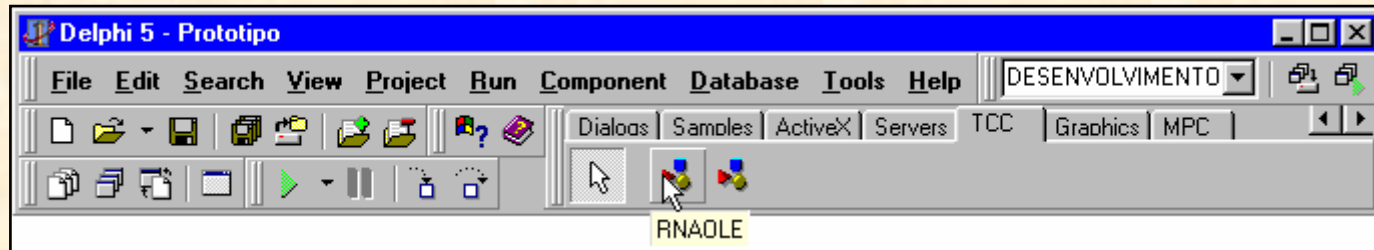
- Interface (Type Library);



O protótipo é constituído de três etapas:

- **Ferramenta de treinamento;**
- **Classes *Automation OLE*;**
- **Aplicação exemplo.**

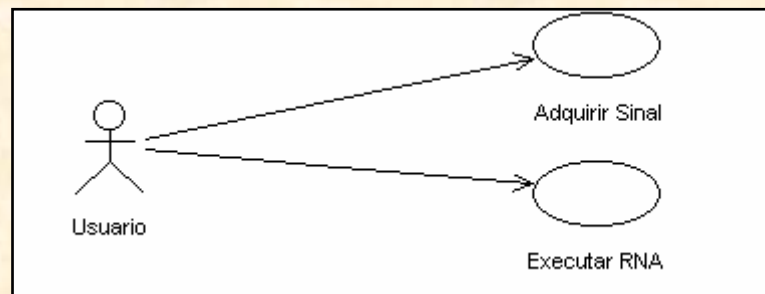
Apresentação:



Especificação:

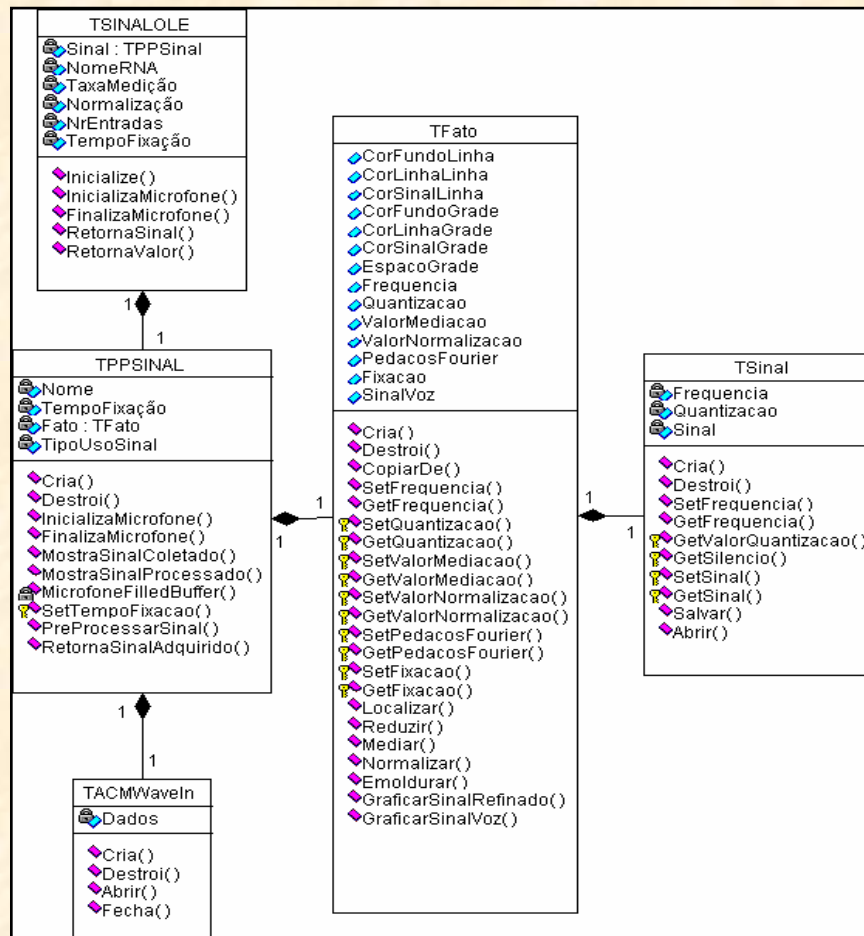
Foram criadas as classes Automation OLE TRNAOLE e TSINALOLE.

- **Caso de uso**

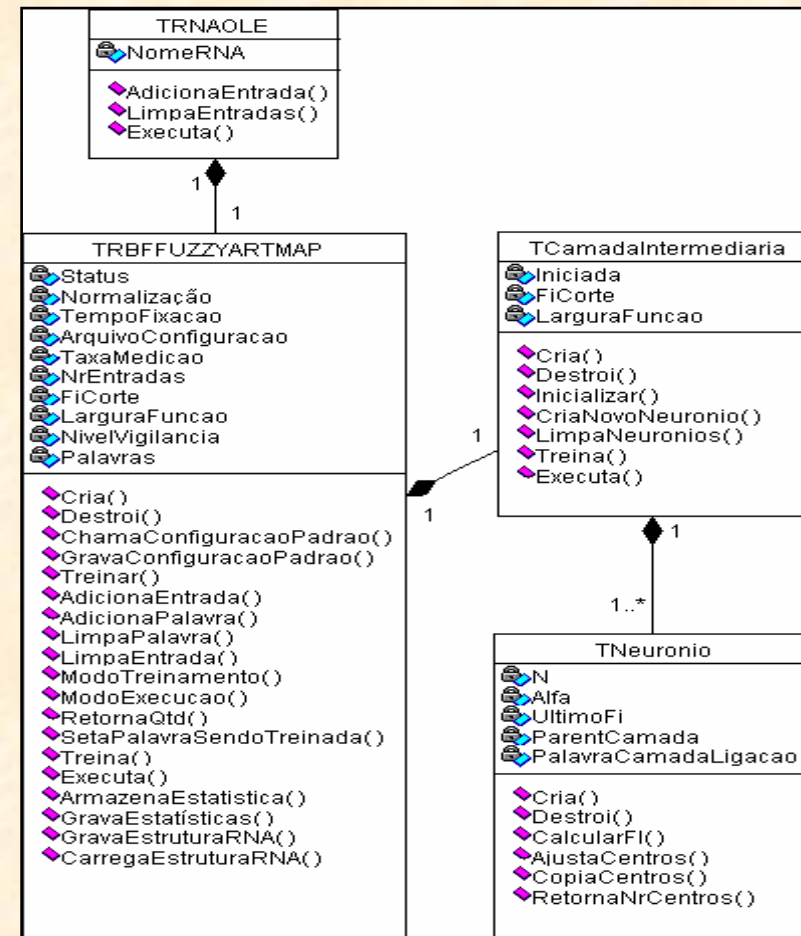


• Diagramas de Classe

TSinalOLE

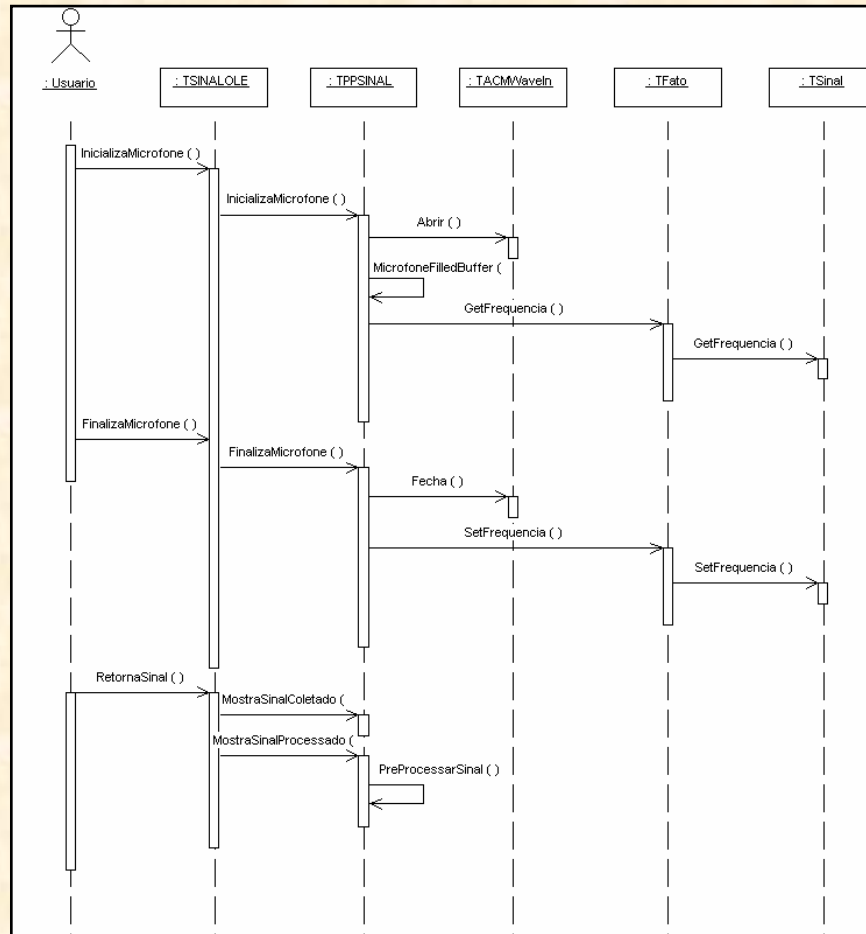


TRnaOLE

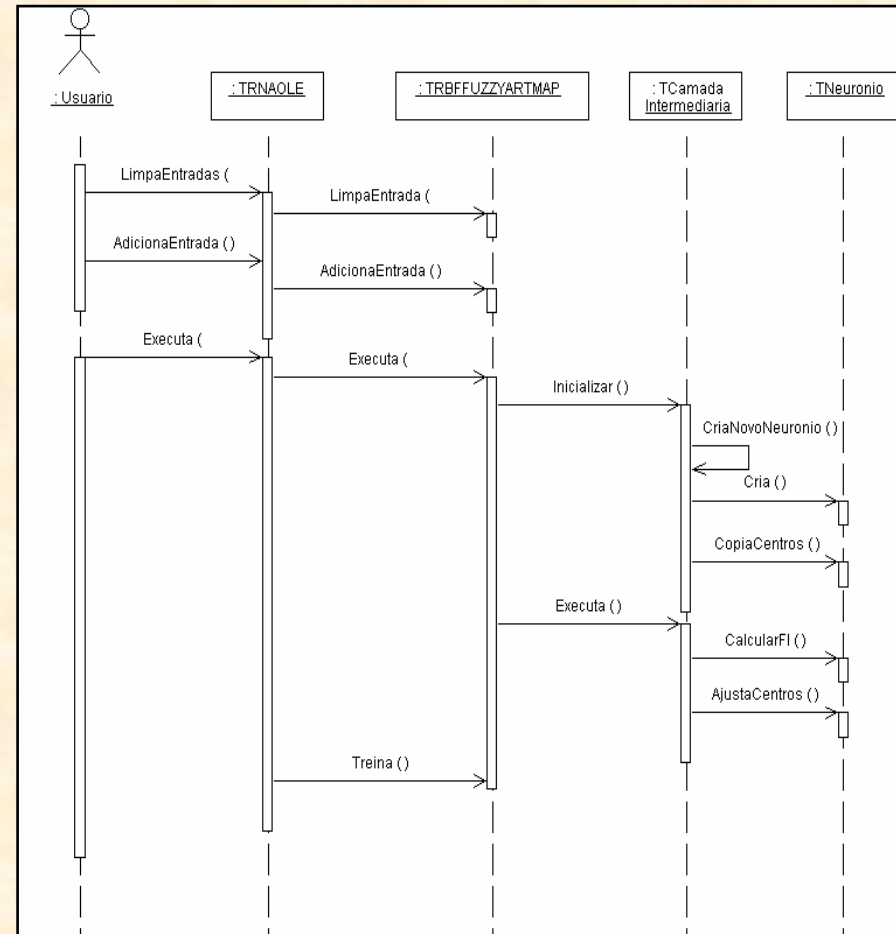


- Diagramas de Sequência

TSinalOLE



TRnaOLE

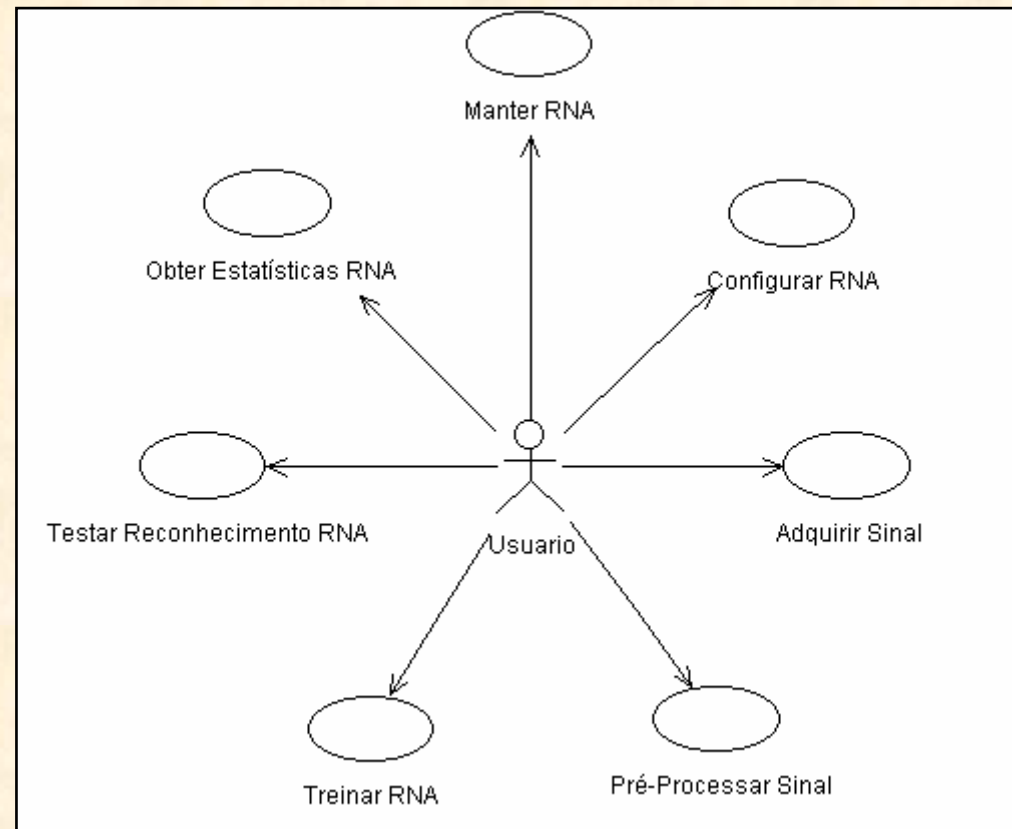


Apresentação:

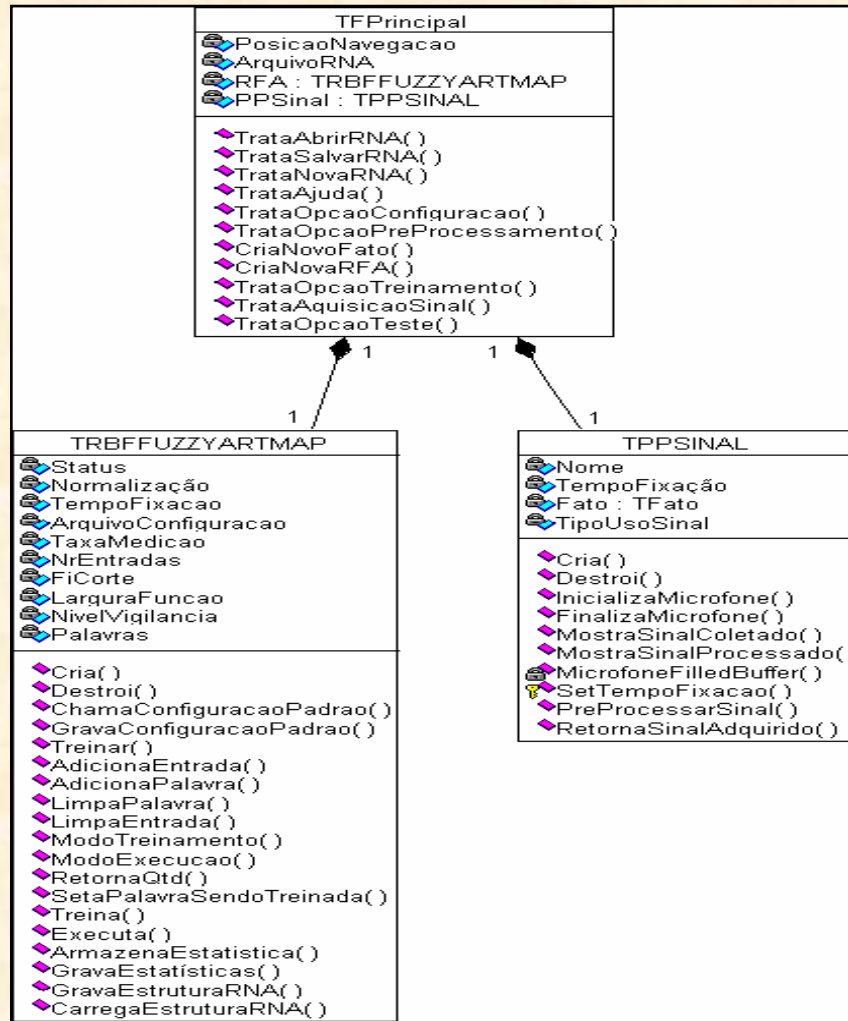


Especificação:

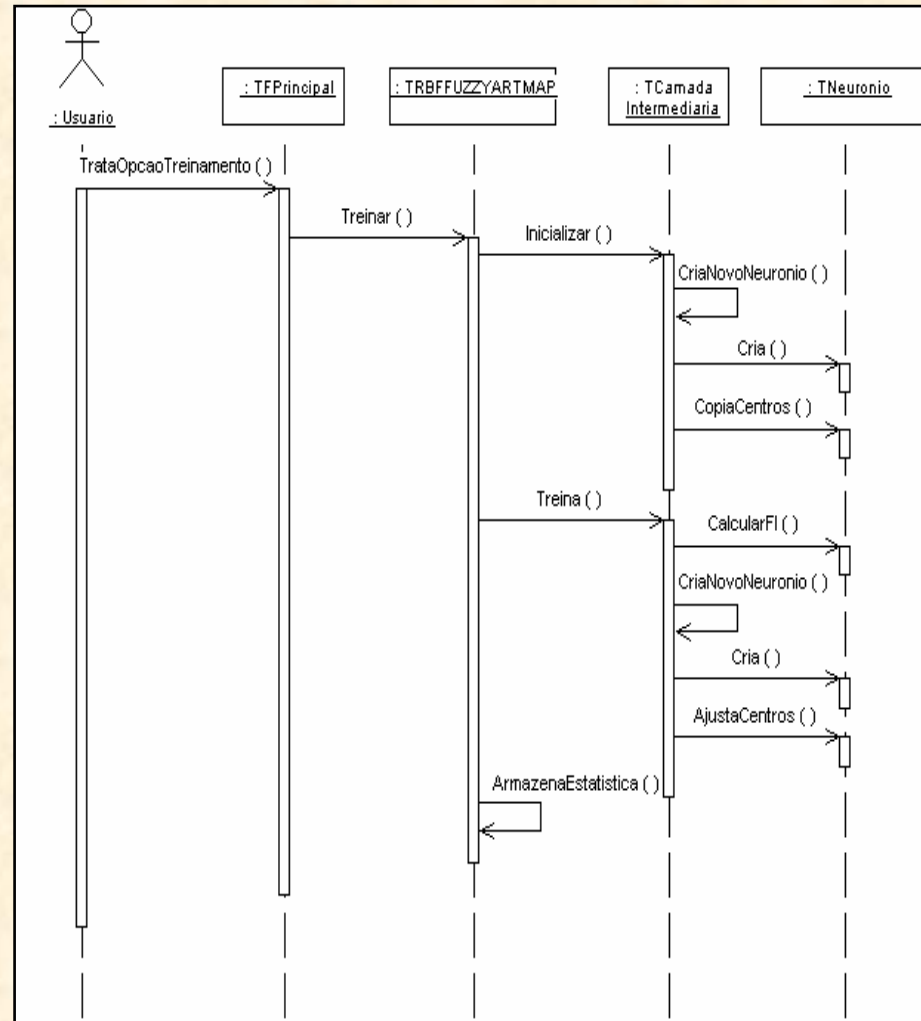
- Caso de Uso



• Diagrama de Classe



• Diagrama de Sequência

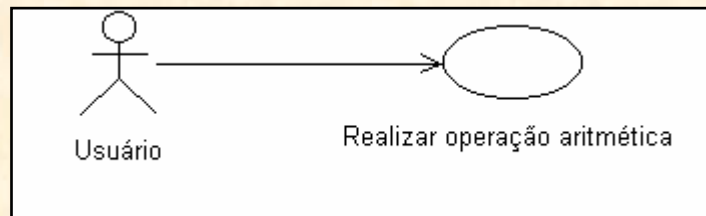


Apresentação:



Especificação:

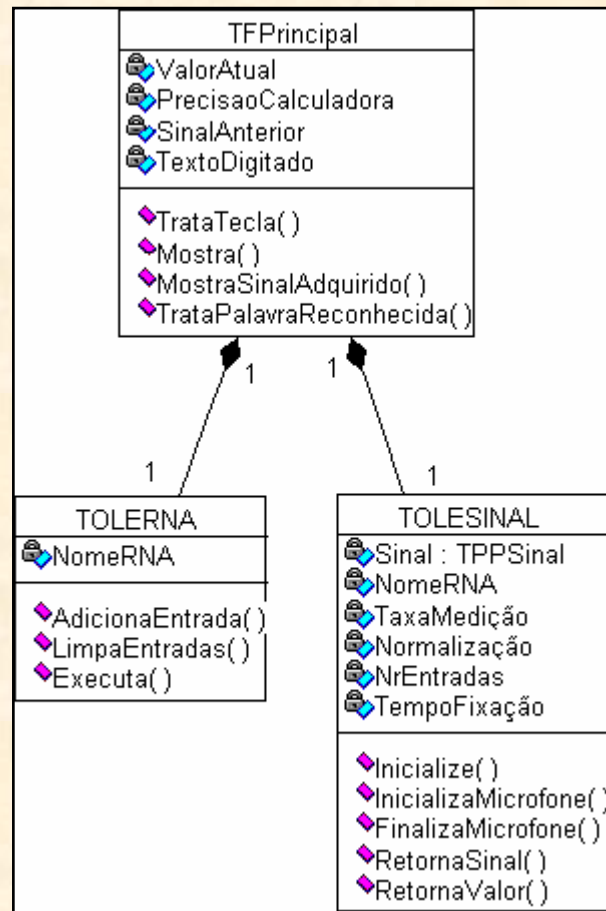
- Caso de Uso



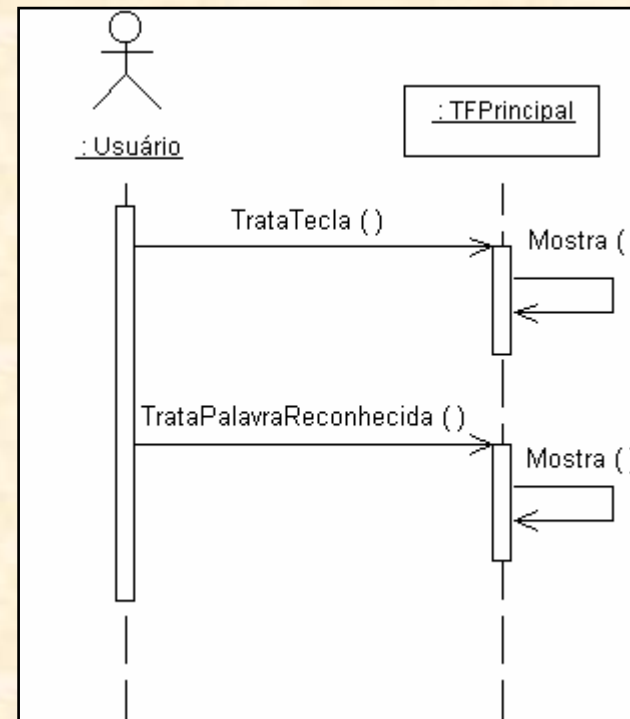
Protótipo

Aplicação Exemplo (Calculadora)

• Diagrama de Classe



• Diagrama de Sequência



Dos testes realizados com a ferramenta de treinamento obteve-se as seguintes tabelas de porcentagens:

- Apresentação do locutor 2 no treinamento do locutor 1;

Palavra	zero	um	dois	três	quatro	cinco	seis	sete	oito	nove	Total
Indefinida	0%	10%	10%	0%	0%	20%	10%	10%	10%	20%	8%
Incorreta	10%	20%	10%	60%	30%	0%	10%	50%	10%	10%	21%
Certa	90%	70%	80%	40%	70%	80%	80%	40%	80%	70%	70%

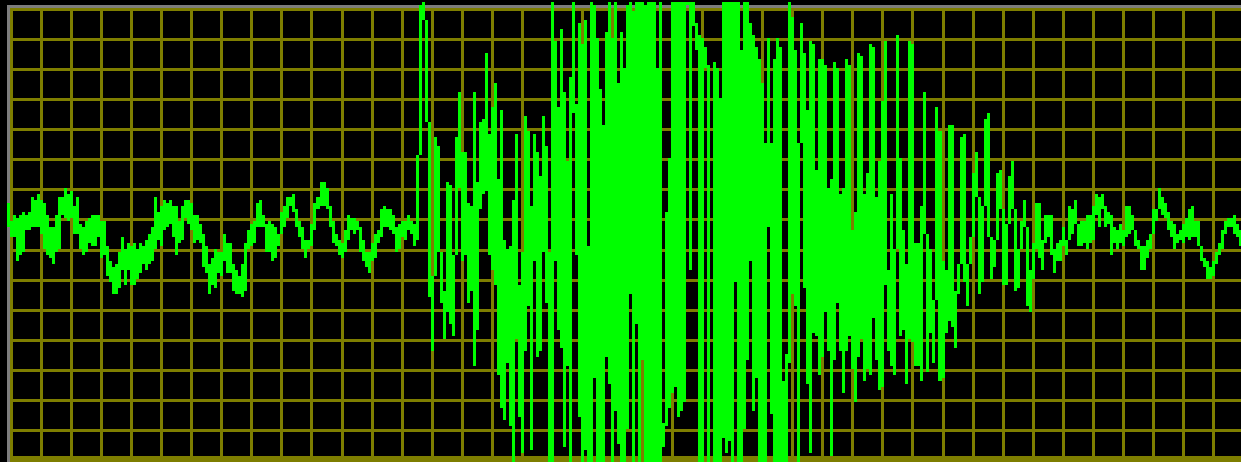
- Apresentação do locutor 1 no treinamento do locutor 2.

Palavra	zero	um	dois	três	quatro	cinco	seis	sete	oito	nove	Total
Indefinida	10%	0%	30%	30%	20%	10%	20%	0%	10%	0%	13%
Incorreta	10%	30%	10%	30%	0%	0%	20%	10%	30%	10%	15%
Certa	80%	70%	60%	40%	80%	90%	60%	90%	60%	90%	72%

Conclusões

Demonstrou-se através da aplicação exemplo (calculadora) a funcionalidade das classes *Automation OLE* desenvolvidas neste trabalho de pesquisa, e também a utilização da ferramenta de treinamento RNA Tools, que possibilita além de gerar a base de conhecimento para uso posterior, testar os padrões adquiridos, com a finalidade de buscar a melhor parametrização da RNA RBF-Fuzzy Artmap, visando a melhor performance no reconhecimento.

Através do protótipo apresentado, composto pela ferramenta, classes *Automation OLE* e aplicação exemplo, atingiu-se o objetivo de oferecer ao usuário final, uma interface por comandos de fala, não trazendo grande tempo de desenvolvimento e pesquisa ao desenvolvedor de software.



“Fim”

◀ Voltar

Avançar ▶

AMORIM, Antônio. **Fundamentos científicos da fonoaudiologia**. São Paulo: Ciências Humanas, 1981. 123p.

AMORIM, Antônio. **Fonoaudiologia geral**. São Paulo: Livraria Pioneira, 1972. 116p.

BAEKER, Ronald M.; Grudin, Jonathan. **Human-computer interaction** : toward the year 2000. San Francisco: Morgan Kaufmann Publishers, 2000. 949p.

BARSA Enciclopédia. São Paulo: Enciclopédia Britânica do Brasil Publicações Ltda, 1998. 16v.

BETEMPS, Rafael da Silva; Silveira, Liliana Nieto. **Redes neurais: conceitos e métodos**. Pelotas, RS, n. 1, p.31-42, 1997.

BORLAND Delphi, versão 5.02. Borland Corporation, 1999. Ambiente de desenvolvimento/Linguagem de programação Object Pascal. 1 CD-ROM.

BRUNS, Fábio Augusto. **Protótipo para o reconhecimento de palavras através da fala**. 1995. 67 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

CAMPESTRINI, Márcio Rodrigo. **Protótipo de um sistema de reconhecimento de caracteres baseado em redes neurais**. 2000. 36 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

CANTÚ, Marco. **Dominando o Delphi 5** – “a bíblia”. São Paulo: Pearson Education do Brasil, 2000. 860p.

COELHO, Helder. **Inteligência artificial: o balanço da década de 80**. Lisboa: Editora Caminho, 1990. 276p.

COLLING, Suzete Teresinha. **Utilização da tecnologia ActiveX Data Objects (ADO) em um sistema com objetos distribuídos**. 2000. 77 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

DALABRIDA, Alexandre Daniel; Itaborai, Gilvani Kisner. **Utilização da Inteligência artificial em reconhecimento de voz**. Universidade Regional de Blumenau, Blumenau, set. 2000. Disponível em: <>. Acesso em: 09 maio 2001.

FERRARO, Nicolau Gilberto; Junior, Francisco Ramalho; Soares, Paulo Antônio de. **Os fundamentos da física : Termologia, Óptica Geométrica e Ondas**. São Paulo: Moderna, 1988. 419p.

FOLMER-JOHNSON, Tore Nils Olof. **Acústica**. São Paulo: Moderna, 1968. 157p.

FURLAN, José Davi. Modelagem de objetos através de UML – the unified modeling language. São Paulo: Makron Books, 1998.

HECKMANN, Jacques Robert. **Sistematização das atuais técnicas de inteligência artificial e análise de sua aplicabilidade**. 1998. 69 f. Monografia (Curso de Pós Graduação), Universidade Regional de Blumenau, Blumenau.

HONORATO JUNIOR, Paulo Francioli. **COM - Component Object Model**. São Paulo. Universidade Regional de Blumenau, Blumenau, set. 2000. Disponível em: <>. Acesso em: 09 maio 2001.

HUGO, Marcel. **Construção de um protótipo de software comandado por voz**. 1992. 41 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

HUGO, Marcel. **Uma interface de reconhecimento de voz para o sistema de gerenciamento de central de informação de fretes.** 1995. 60f. Dissertação (Mestrado em Engenharia da Produção) – Universidade Federal de Santa Catarina, Florianópolis.

KITANO, Hiroaki. **Speech-to-speech translation.** Boston: Kluwer Academic Publishers, 1994. 193p.

KLABUNDE, Charles Cristiano. **Aplicação do modelo de rede neural RBF-fuzzy artmap para o reconhecimento de fala.** 1996. 64 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

LFM programming's page. Curitiba. Desenvolvido pela Universidade Federal do Paraná, 1999. Apresenta conceitos e técnicas sobre redes neurais e pré-processamento de sinais. Disponível em <>. Acesso em: 30 abr. 2001.

LOESCH, Cláudio e SARI, Solange T. **Redes neurais artificiais : fundamentos e modelos.** Blumenau: FURB, 1996.

MATRAS, Jean-Jacques. **O som.** São Paulo: Universidade Hoje, 1995. 116p.

MOSER, Bóris. **Utilização da transformada de Fourier no tratamento de sinal para a utilização em uma rede neural.** 1997. 59 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

SANTOS, Roosevelt dos. **Reconhecimento de fala por fonemas.** 1996. 64 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

SCHMITZ, Eber; SILVEIRA, Denis. **Desenvolvimento de software orientado a objetos** – utilizando UML e Delphi 5. Rio de Janeiro: Brasport, 2000.

STARKE, Júlio Frederico. **Comparativo entre modelos de redes neurais aplicadas ao reconhecimento da fala**. 1996. 69 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

STEVE, Teixeira; PACHECO, Xavier. **Delphi 5 - developers guide**. London: Pearson, 1999. 1556p.

REVOX Sistema de Reconhecimento de Voz. Caxias do Sul. Desenvolvido pela Universidade Federal do Rio Grande do Sul (UFRGS), 1997. Apresenta informações sobre um sistema para reconhecimento de fala. Disponível em <>. Acesso em: 20 abr. 2001.

TAFNER, Malcon Anderson. **O reconhecimento de palavras faladas isoladas usando redes neurais artificiais**. 1996. 89f. Dissertação (Mestrado em Engenharia da Produção) – Universidade Federal de Santa Catarina, Florianópolis.

TONTINI, Gérson. **Automatização da identificação de padrões em gráficos de controle estatístico de processos (CEP) através de redes neurais com lógica difusa**. 1995. 137 f. Tese (Doutorado em Engenharia Mecânica) – Universidade Federal de Santa Catarina, Florianópolis.

VIEIRA, Eduardo. **Pagamento sem neura...** Revista Info Exame. São Paulo, ano 16, n. 182, p.92-93, maio 2001.

XEREZ, Marcos de; Filho, Ilson W. Rodrigues; Tafner, Malcon Anderson. **Redes neurais artificiais** – introdução e princípios de neurocomputação. Blumenau: Editora da Furb, 1996. 199p.

Quando um padrão é apresentado a RNA RBF-Fuzzy Artmap:

- Calcula-se o valor do F_i ;
- Acha-se o valor que tenha resposta máxima;
- Compara-se a saída do neurônio com resposta máxima ao nível de vigilância;
 - Se a resposta máxima é superior ao nível de vigilância:
 - O neurônio é ativado passando para o módulo de ligação.
 - Caso contrário em modo de:
 - Execução: o padrão de entrada não é identificado;
 - Treinamento: a criação de um novo neurônio com centros iguais ao padrão de entrada;
 - Se é inferior, cria-se um novo neurônio com centros iguais ao padrão de entrada;
- No módulo de ligação em modo de:
 - Treinamento: é feita uma checagem para verificar se o neurônio vencedor identifica o padrão apresentado, se o padrão apresentado for:
 - Identificado: a RNA resulta a identificação do padrão;
 - Não identificado: pode retornar a camada RBF e criar um novo neurônio com centros iguais ao padrão de entrada.
 - Execução: não há a possibilidade de se executar checagem, onde o neurônio ativado pela camada RBF identifica o padrão de entrada, podendo ocorrer a identificação do padrão correta ou não.