

Reconhecimento de Voz para Automação Residencial baseado em Agentes Inteligentes

Marília A. Amaral¹, Rodolfo Barriviera², Eduardo CotrinTeixeira³

^{1,2}Centro de Ciências Exatas e Tecnológicas, Universidade Norte do Paraná, Rua Tietê, 1208, Vila Nova, Londrina, PR

marilia.amaral@unopar.br ; rodolfo.barriviera@unopar.br

³Departamento de Computação, Universidade Estadual de Londrina, Rodovia Celso Garcia Cid, Campus Universitário, Londrina, PR
eduardo@dc.uel.br

Resumo – A utilização de agentes inteligentes é cada vez mais difundida no meio computacional, principalmente nas áreas de interfaces, gerenciamento de redes, recuperação de informação, sistemas distribuídos e automação. Porém, a discussão sobre a definição de agentes inteligentes é muito ampla. Basicamente um agente possui autonomia e executa tarefas, estando sempre vinculado a uma ação disparada automaticamente ou a uma ação externa manual a ser executada. Este artigo relata o desenvolvimento de um Software de Reconhecimento de Voz para Automação Residencial que utiliza arquitetura baseada em agentes inteligentes. O objetivo desta aplicação é auxiliar e proporcionar mais independência a pessoas portadoras de necessidades especiais, como pacientes que possuem um alto grau de lesão, com movimentos apenas acima do pescoço. O sistema desenvolvido é composto de dois programas, um para reconhecimento de voz e o outro para controle do processo de automação residencial. Os programas são interligados por meio de uma interface de software, e o sistema todo possui um circuito acoplado à saída paralela do computador. Este circuito ativa ou desativa oito pontos que estão conectados a qualquer aparelho eletro-eletrônico, através de comando de voz.

Palavras-chave: Agentes inteligentes, Automação Residencial, Portadores de Necessidades Especiais.

Abstract - The use of intelligent agents is being diffuse in the computational area, mainly in the areas of interfaces, network management, information recovery, distributed systems and automation. The discussion on the definition of intelligent agents is very ample. Basically an agent executes tasks, being always linked to an automatically action or a manual external action to be executed. This article discuss the development of a **Software of Recognition of Voice joined to the Residential Automation** that uses an architecture based on intelligent agents. The objective of this application is help and to provide more independence to people with disabilities, with movements only above of the neck. It is made use a computer contend two linked programs to recognition of voice and process controll of residencial automation. This circuit activ or desactiv eight points that are connected to any electro-electronic device, through voice command.

Key-words: Intelligent Agents, Residencial Automation, People with disabilities

Introdução

O processo de desenvolvimento de sistemas de reconhecimento de voz é caracterizado por exigir trabalho árduo de programação e de pesquisa. Para facilitar este processo, o presente estudo propõe a utilização de Agentes Inteligentes aplicados à área de reconhecimento de voz. A pesquisa constitui-se de uma série de etapas com tarefas definidas, que objetivam a obtenção de um software de reconhecimento de voz que utilize a arquitetura de agentes inteligentes.

O trabalho apresentado envolve uma ampla pesquisa sobre Agentes Inteligentes que podem ser utilizados em diversas aplicações como na interação homem-máquina, engenharia de software e sistemas distribuídos. [1]

Antes de descrever a aplicação que é foco deste trabalho é interessante definir e apresentar conceitos relevantes à teoria de agentes inteligentes. A palavra agente se refere a todo ente que possui habilidade, capacidade e permissão para atuar em nome de outro. Em termos usuais, agentes seriam pessoas com um conhecimento ou recurso mais especializado em uma determinada área, mediante o qual auxiliariam outros indivíduos em suas tarefas [2].

Como não existe ainda um consenso quanto à definição de agente, serão destacadas algumas das definições mais relevantes:

Segundo Wooldridge e Jennings [1], agentes são sistemas que apresentam um comportamento que é determinado por um processo de raciocínio baseado na representação de suas atitudes, tais como crenças,

comprometimentos e desejos. Um sistema pode ser visto como um agente se possuir as seguintes propriedades: autonomia, habilidade social, reatividade, pró-atividade.

De acordo com Nwana [3], agente é um componente de software ou hardware capaz de atuar para poder executar uma tarefa em nome de um usuário.

Tem-se ainda a definição de Russel e Norvig [4], segundo a qual a inteligência está intrinsecamente relacionada com as ações racionais. Agir racionalmente significa agir de forma a alcançar as metas definidas por alguém, dadas as crenças deste. Assim, definem de forma genérica que um agente é algo que percebe seu ambiente por meio de sensores e atua nesse ambiente por meio de atuadores. A figura 1 mostra o diagrama de um agente genérico segundo definição de Russel e Norvig [4].

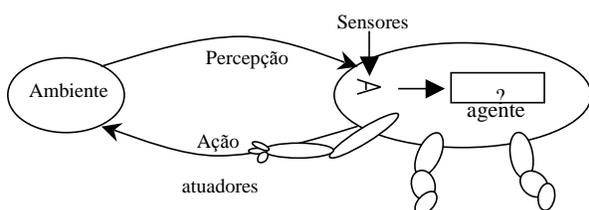


Figura 1 – Diagrama de Agente genérico

A aplicação escolhida para este trabalho está voltada para área de Reconhecimento de Voz para Automação Residencial. Considera-se de grande relevância o uso dos agentes inteligentes, pois estes poderão atuar como catalisadores dos comandos do usuário e poderão interagir com os demais módulos do sistema para efetivamente concretizar a ação solicitada pelo usuário.

Justificativa

O processo de desenvolvimento de software aqui apresentado sugere uma metodologia de desenvolvimento de software de reconhecimento de voz utilizando Agentes Inteligentes que pode facilitar o processo de desenvolvimento desta categoria de software. Este estudo está inserido em um projeto de pesquisa que necessita de uma solução para a área de reconhecimento de voz [5].

A motivação para o desenvolvimento de um software de reconhecimento de voz utilizando a arquitetura de agentes inteligentes surgiu da necessidade de desenvolvimento modular e eficiente de um ambiente que atendesse aos portadores de necessidades especiais, a fim de que estes possam controlar suas residências através de um microcomputador e do recurso de voz.

Como definido anteriormente por Nwana [3] um agente é um componente de software ou hardware capaz de atuar para poder executar uma tarefa em nome de um usuário.

O sistema de reconhecimento de voz neste ambiente denominado Ambiente Computacional de Ajuda a Portadores de Necessidades Especiais, possui as duas características listadas acima: é um componente de software que deve interagir com outros componentes de software e hardware para executar uma tarefa determinada por um usuário.

Objetivos

O objetivo geral deste trabalho é desenvolver um Ambiente Computacional de Ajuda a Portadores de Necessidades Especiais formado por um software de reconhecimento de voz unido à automação residencial dependente do locutor, com o reconhecimento de “comandos únicos” de voz que ativarão determinados pontos de energia conectados a aparelhos eletroeletrônicos e aparelhos domésticos dentro de uma residência. Além disto o trabalho tem como objetivo a contribuição para o processo de inclusão digital dos portadores de necessidades especiais, no caso deste trabalho especificamente, a priori, os portadores de tetraplegia.

Para atingir este objetivo são utilizadas informações da AACD [6], e as tecnologias de Agentes Inteligentes [1], KQML [7], Processamento de Linguagem Natural (PLN), técnicas para processamento de voz e conceitos de Projeto de Software advindos da Engenharia de Software [8].

A pesquisa destas tecnologias forma apenas uma parte do cronograma definido, que rege etapas para cada fase do trabalho. Atualmente o trabalho encontra-se em fase de testes. Foram realizados alguns testes com um portador de necessidade especial e a próxima etapa será a de refinamento do projeto, levando em consideração as opiniões emitidas após o teste.

Os objetivos listados acima se tornam reais somente com o segmento fiel do cronograma definido, que estabelece cada etapa do trabalho a ser desenvolvida com datas e prazos determinados. A criação de um modelo que apresente todos os passos para a criação de um software de reconhecimento de voz, com conceitos de agentes inteligentes, unidos a automação residencial conclui o objetivo geral deste trabalho.

Metodologia

O método adotado para o desenvolvimento do software de reconhecimento de voz para automação residencial é de modularização, onde

cada módulo do software é um módulo independente que executa tarefas distintas, ou seja, um agente inteligente recebendo informações e executando tarefas determinadas, que são de fundamental importância em todo o processo de reconhecimento de voz e execução das tarefas solicitadas pelo usuário. As atividades de responsabilidade dos agentes inteligentes estão descritas como requisitos funcionais na seção *Sistema Experimental* deste documento.

O desenvolvimento do software foi pautado em conceitos de engenharia de software, utilizando ferramentas de programação como Borland C++Builder e Visual Studio (C++). O ciclo de vida básico para desenvolvimento é dividido nas seguintes etapas [8]: levantamento de dados, análise de requisitos, implementação e testes.

A seção *Sistema Experimental* do presente trabalho detalha a arquitetura do software de reconhecimento de voz e como sua integração é feita com o Ambiente Computacional de Ajuda a Portadores de Necessidades Especiais.

Sistema Experimental

Para acessar o Ambiente Computacional de Ajuda a Portadores de Necessidades Especiais, o usuário possui como opção: utilizar o teclado para entrada de comandos via digitação, ou utilizar o “Comando de Voz” para entrada via microfone. Quando a entrada do comando desejado é realizada via voz o “Programa de Reconhecimento de Voz” deve ser ativado.

Analisando estudos na área de reconhecimento de voz já realizados por cientistas e desenvolvedores [9], constatou-se que o desenvolvimento completo de um sistema de digitalização de voz seria inviável devido a sua complexidade, o que desviaria o propósito do projeto em questão. O assunto é muito amplo sendo abordado em dissertações de mestrado, e outros projetos científicos mais específicos [10].

A proposta para este desenvolvimento é um sistema que utiliza técnicas de indexação de voz, que aborda a análise de algumas características da voz como tonalidade e amplitude [11]. O Ambiente Computacional de Ajuda a Portadores de Necessidades especiais se divide em duas etapas, como mostra a figura 2.

Na **Etapa 1** se dá o encontro do usuário com o módulo denominado Agente de Comandos. No **Item 1** o usuário grava, via voz, todos os comandos disponíveis pelo software. No **Item 2** estes comandos, cada um destinado a uma operação do software, ficam gravados num banco de dados. Esta etapa é de suma importância, pois é através dela que o software ‘aprende’ as características da voz do usuário.

Já na **Etapa 2**, após concluir os processos anteriores, o Agente de Execução de Comandos, que caracteriza a etapa 2, está pronto para ser utilizado. No **Item 3** o usuário determina um

comando através da voz, como por exemplo, “acenda luz 1”.

No **Item 4** o sistema grava esta voz em formato “wav” (formato específico para manipulação de som), após este processo, no **Item 5**, é realizada análise para extrair as características da voz, tais como timbre, altura, tonalidade e amplitude.

Com todas as características de entrada analisadas e armazenadas, inicia-se o **Item 6** da arquitetura proposta. Deve ser feito então um processo de comparação das informações de entrada com as informações que já estão no banco de dados.

Já no **Item 7** essas comparações retornarão valores lógicos. Se o resultado for um valor verdadeiro, será executado aquele comando que teve as mesmas características da voz gravada, ou seja, “acenda luz 1”, caso contrário o sistema emite uma mensagem de ‘comando inexistente’, e retorna ao modo de espera do próximo comando.

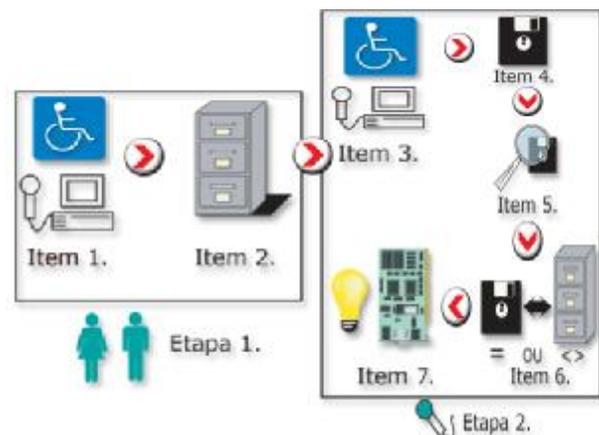


Figura 2 – Diagrama esquemático do processo de reconhecimento de voz

A arquitetura dos agentes inteligentes utilizados no diagrama esquemático da figura 2 é apresentada na figura 3. Como referido anteriormente, são utilizados dois tipos de agentes: Agente de Comandos e Agente de Execução de Comandos.

Na figura 3 estão representados os agentes atuando em seu ambiente, onde cada agente executa uma tarefa determinada.

A entrada de comandos via voz é recebida pelo Agente de Comandos, que faz a gravação dos comandos no banco de dados, no primeiro contato com o usuário. Após este contato o banco de dados possui todos os comandos que podem ser realizados pelo usuário.

A partir do momento que a base de dados é constituída, o usuário pode executar os comandos, que são recebidos pelo Agente de Comandos. Este realiza uma consulta ao banco de dados verificando se o comando é verdadeiro

(consta na base de dados), ou falso (não consta na base de dados). Sendo verdadeiro este agente faz a comunicação com o Agente de Execução de Comandos, que ativa os pontos de energia que conectam aparelhos residenciais eletro-eletrônicos. Se não for verdadeiro, é exibida na tela uma mensagem de erro que permanece somente alguns segundos na tela, para depois retornar ao sistema em “repouso”.

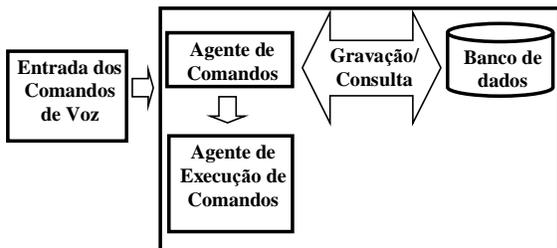


Figura 3 – Arquitetura básica dos agentes utilizados no Programa de Reconhecimento de Voz

Para que as ações determinadas pelo Agente de Execução de Comandos sejam efetivadas, existe uma conexão entre o Programa de Reconhecimento de Voz (PRV) e o Programa Porta Paralela (PPP). Esta conexão é feita através do bloco Interface. É justamente através do PPP que se tem a ativação dos pontos disponíveis para a alimentação das várias tomadas conectadas ao sistema.

A seguir é apresentado, na figura 4, um diagrama de blocos do ambiente de ajuda a portadores de necessidades especiais.

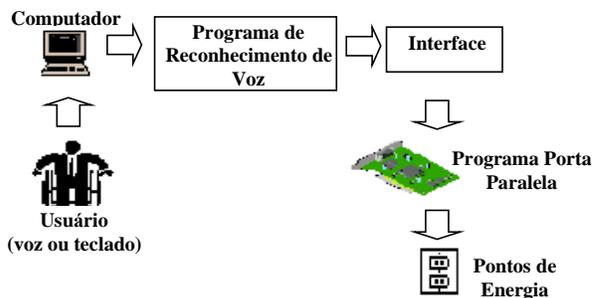


Figura 4 – Diagrama de blocos do Ambiente Proposto

Os elementos apresentados na figura 4 são descritos a seguir.

A porta paralela é o mecanismo de comunicação entre o computador e um periférico. São vários os periféricos que se utilizam desta porta para enviar e receber dados para o computador (exemplo: Scanners, Câmeras de vídeo, Unidade de disco removível). Ela está ligada diretamente à placa mãe do computador e pode ser manipulada através de recursos de várias linguagens de programação, como

C/C++/C++Builder, Pascal/Delphi, Visual Studio, Visual Basic, entre outras.

A Interface é um circuito, dispositivo ou porta que permite que duas ou mais unidades incompatíveis sejam interligadas num sistema padrão de comunicação, permitindo que transfiram dados entre si.

Os pontos de energia são as tomadas onde serão ligados os aparelhos a serem controlados.

A implementação dos agentes inteligentes está vinculada ao Programa de Reconhecimento de Voz.

Na Tabela 1 é apresentada a descrição dos requisitos do Programa de Reconhecimento de Voz [8]. Foram definidos 15 requisitos, cada um deles especificado como Requisito Essencial ou Requisito Desejável.

Um requisito definido como essencial (RE) é imprescindível para o funcionamento do sistema, exemplo: gravar um comando de voz. Um requisito Desejável (RD) é aquele que faz parte do sistema, porém será utilizado eventualmente, exemplo: *Help*.

Estes requisitos foram definidos tomando como base a interação dos agentes inteligentes apresentados na figura 3.

As principais atividades atribuídas aos agentes inteligentes (Agente de Comandos e Agente de Execução de Comandos) estão listadas como requisitos essenciais do Programa de Reconhecimento de Voz, pois estas são importantes para o bom funcionamento do Programa.

Por meio desta descrição dos requisitos foi possível iniciar a implementação do Programa de Reconhecimento de Voz e prever quais seriam as entradas, saídas, origens, destinos e regras de cada ação do Programa.

As entradas são os dados que disparam as execuções. As saídas são os dados produzidos/gerados por estas execuções. As origens e destinos são definições físicas relativas ao *hardware* que está relacionado com as entradas e saídas. As regras definem as condições para que as execuções sejam efetivadas.

Nº	Tipo	Descrição
1	RE	Possuir um Banco de Dados contendo uma tabela denominada DadosUsuário com os seguintes campos: NomeUsuário, Data, Hora, ComandoVoz, Saída e ÁudioVoz.
2	RD	Permitir a execução dos comandos armazenados na tabela DadosUsuário do Banco de Dados via voz ou teclado
3	RE	Possuir um módulo inicial composto pelo Agente de Comandos, que recebe os comandos do usuário via voz, para gravá-los no Banco de Dados.
4	RE	Possuir um módulo de espera composto pelo Agente de Comandos, que recebe os comandos do usuário via voz, para pesquisá-los no Banco de Dados.
5	RE	O Agente de Comandos deve enviar uma informação constando a saída a ser ativada para o Agente de Execução de Comandos, se o resultado da pesquisa for verdadeiro.
6	RE	O sistema deve possuir um Agente de Execução de Comandos que ative através da Placa de Circuito Impresso o aparelho doméstico que foi indicado pelo usuário na informação recebida pelo Agente de Comandos.
7	RE	O Agente de Comandos deve emitir uma mensagem de Erro na tela por 5 segundos, quando não for encontrado na tabela DadosUsuário do banco de dados o comando solicitado pelo usuário.
8	RE	O sistema deve gravar os comandos de vários usuários independentes, via voz (microfone) com o auxílio do teclado, para que estes comandos possam ser utilizados posteriormente.
9	RE	O sistema deve possuir um item que permite escolher qual usuário irá utilizar o software naquele momento.
10	RD	O sistema deve possuir um <i>help</i> on-line
11	RE	O sistema deve permitir a exclusão dos comandos de voz gravados.
13	RD	O sistema deve emitir relatório de todos os usuários com os seus respectivos comandos de voz gravados.
15	RE	O sistema deve conectar-se a uma Placa de Circuito Impresso (PCI) através da Porta Paralela do PC, para ativar os aparelhos eletrodomésticos e eletroeletrônicos via comando de voz ou teclado.

Tabela 1 – Requisitos do Programa de Reconhecimento de Voz

Conclusão

Todo o hardware do Ambiente Computacional de Ajuda a Portadores de Necessidades Especiais já está desenvolvido, assim como o Programa Porta Paralela.

O Programa de Reconhecimento de Voz já está em fase de finalização dos testes, partindo para o refinamento de suas funcionalidades.

A conclusão do projeto proposto vem com a análise realizada por um paciente portador de necessidade especial selecionado pelo grupo de pesquisa que desenvolve o presente trabalho. A princípio foi escolhido um paciente com um bom nível cultural para que seja possível uma análise mais detalhada.

Foram levantados alguns itens propostos como trabalhos futuros, tais como: - reconhecimento de voz de vários usuários durante o mesmo período de execução; - utilização do reconhecimento de voz em rede de computadores e análise de outros meios para a comunicação da porta paralela com o computador.

O projeto proposto mostra-se extremamente viável, pois atualmente é notável que uma maior atenção deveria ser dispensada para o público portador de necessidades especiais.

Como trabalhos futuros pode-se destacar a pesquisa para implementação e inserções de elementos relacionados com Wireless, bem como utilização de tecnologia Java – J2ME.

Neste trabalho os Agentes Inteligentes são tecnologias que tornam o software melhor modelado, para execução de suas tarefas, visto que o objetivo do agente é atuar, através de dados recebidos como entrada, para executar uma determinada tarefa.

Referências

- WOOLDRIDGE, Michael; JENNINGS, Nicholas R. Intelligent Agents: Theory and Practice. Knowledge Engineering Review, 1994.
- BARONE, Dante Augusto Couto; YEPES, Igor. Inteligência Artificial Distribuída: Uma Abordagem ao Comportamento Social Inteligente. Sociedades Artificiais. Editora Bookman, 2002.
- NWANA, H.S. Software agents: an overview. Knowledge Engineering Review. Cambridge University Press, vol 11, 1996.
- RUSSEL, S; NORVIG, P., 1995. Artificial Intelligence: A modern Approach. Prentice Hall, 1995.

5. AMARAL, Marília; FAESARELLA, Annete; BARRIVIERA, Rodolfo; BRAUKO, João II, OLIVEIRA, Dennis Augusto; BARREIRO, Elias P. Júnior. Programa Computacional de Ajuda a Portadores de Necessidades Especiais. Congresso Brasileiro de Computação. Anais... Univali, SC, 2002.
6. AACD 50 anos de medicina e reabilitação – Associação de Assistência à Criança Deficiente. Editora Vida & Consciência. Disponível por: www.aacd.org.br, 2000.
7. FININ, Tim, FRITZON, Richard, MCKAY, Don, MCENTIRE, Robin. KQML as an Agent Communication Language. Conference on Information and Knowledge Management, 1994.
8. PRESSMAN, S. Software Engineering: A Practitioner's Approach. 5ªed. McGraw-Hill series in computer, 2001.
9. Projeto Motrix. Disponível em: <http://intervox.nce.ufrj.br/motrix/>. Acessado em: fev. 2003.
10. GOYA, Denise H., 2002. Ouça e seja ouvido: Casos interessantes. Disponível em: http://pcworld.terra.com.br/pcw/testes/tecno_hard/0048.html. Acessado em mar. 2002.
11. H. P. Nii, "Blackboard Systems: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures," AI Magazine, Summer 1986.