



Projeto Will: comunicação entre o usuário e o computador

Marilei Fátima Kovatli¹

Resumo: este artigo apresenta como tema o projeto de um programa de comunicação entre um usuário e o computador. O objetivo geral do estudo é desenvolver um software que possa estabelecer a comunicação do usuário com a máquina de tal forma que as pessoas tenham a impressão de estar interagindo com outra pessoa e não com um software. Como objetivos específicos definiu-se: Escolher a ferramenta de desenvolvimento; Definir os comandos/recursos do software; Criar o protótipo do software; Testar o software; Verificar conexão com internet, modo off-line; Analisar o retorno (feedback) sobre o tipo de interação. A metodologia deste estudo envolve uma pesquisa bibliográfica e experimental, e está dividido em etapas. Para desenvolver os comandos de reconhecimento de voz, e os comandos específicos que habilitam o software executar as tarefas, utilizou-se algumas ferramentas e aplicativos como: Microsoft speech api, SAPI 5.0, system.speech da Microsoft. A pesquisa está em desenvolvimento e a próxima etapa compreende a criação de um banco de dados que possibilite a aprendizagem do software e também realizar um teste com um maior grupo de usuários. A pesquisa resultou no entendimento acerca dos processos de interação e comunicação que é possível desenvolver através de um software. Através da utilização de conceitos de programação, de IHC - Interface homem-máquina e inteligência artificial foi realizado a simulação da comunicação, após o envio de perguntas em linguagem natural, o programa consulta uma base de conhecimento e em seguida fornece uma resposta que tenta imitar o comportamento humano.

Palavras-chave: interação – software – computador - comunicação.

Abstract: this article presents as a theme the design of a communication program between a user and the computer. The overall goal of the study is to develop software that can establish user communication with the machine so that people have the impression that they are interacting with someone else and not with software. The following specific objectives were defined: Choose the development tool; Define the commands / features of the software; Create the software prototype; Test the software; Check connection to internet, offline mode; Analyze the feedback about the type of interaction. The methodology of this study involves a bibliographical and experimental research, and is divided into stages. To develop the voice recognition commands, and the specific commands that enable the software to perform the tasks, we used some tools and applications such as: Microsoft speech api, SAPI 5.0, system.speech from Microsoft. The research is under development and the next step comprises the creation of a database that enables the learning of the software and also conduct a test with a larger group of users. The research resulted in an understanding of the interaction and communication processes that can be developed through software. Through the use of programming concepts, MMI - Man-machine interface and artificial intelligence was performed the simulation of communication, after sending questions in natural language, the program queries a knowledge base and then provides a response that tries to imitate human behavior.

Keywords: interaction - software - computer - communication.

Introdução

Com o avanço da tecnologia e com o surgimento de processadores com alto poder de processamento, desde os cartões perfurados, para teclados e mouses,

¹ Mestre em Ciência da Computação- UFSC. Professora do Curso de Análise e Desenvolvimento de Sistemas da Faculdade CNEC Santo Ângelo



para telas *touchscreen*², e comandos de reconhecimento de voz em algumas aplicações, permitem uma interação homem-máquina antes nunca imaginado. As aplicações dessas tecnologias variam desde sistemas de ditados, a aplicações de educação, comandos para controlar menus, a automação de ambientes e ao auxílio para portadores de deficiências, onde em sua maioria são destinadas a pessoas com limitação motor ou visual.

O objetivo geral do estudo é desenvolver um software que possa estabelecer a comunicação do usuário com a máquina de tal forma que as pessoas tenham a impressão de estar interagindo com outra pessoa e não com um software. Como objetivos específicos definiu-se: Escolher a ferramenta de desenvolvimento; Definir os comandos/recursos do software; Criar o protótipo do software; Testar o software; Verificar conexão com internet, modo off-line; Analisar o retorno (feedback) sobre o tipo de interação.

Seguindo a ideia do objetivo geral e dos objetivos específicos, é possível justificar que, a pesquisa realizada mostra que as interfaces via voz estão rapidamente se tornando uma necessidade, em um futuro próximo, sistemas interativos irão fornecer fácil acesso a milhares de informações e serviços que irão afetar de forma profunda a vida cotidiana das pessoas. O presente estudo, ainda pode servir como base de fonte de informações para futuros projetos e pesquisas.

Interação entre o homem e a máquina - software de reconhecimento de voz

Estamos numa era onde a interação humana com os computadores está cada vez mais acessível, e com a facilidade que temos em possuir em mãos diariamente equipamentos com sistemas para uso pessoal e até mesmo profissional, onde necessitamos estar com soluções em tempo real.

² A tela sensível ao toque é um display eletrônico visual que pode detectar a presença e localização de um toque dentro da área de exibição, por meio de pressão. <https://www.tecmundo.com.br/multitouch/177-o-que-e-touch-screen-.htm>



Com o avanço da tecnologia, o acesso ao computador tornou-se indispensável para inúmeros setores, causando uma necessidade enorme em facilitar a interação. (MORAIS,LOPER, 2014).

Alguns softwares desenvolvidos para interação entre homem-máquina através do reconhecimento de voz, podemos citar:

DosVox, desenvolvido pelo núcleo de computação eletrônica da Universidade Federal do Rio de Janeiro, no qual vem se dedicando a criação de um sistema para atender as necessidades dos deficientes visuais.

O Sistema MyWay, desenvolvido para automação de ambientes, controlando iluminação, equipamentos eletrônicos, segurança.

O Speech to Text, desenvolvido pela IBM, qual converte o áudio e voz em texto escrito, transcreve em tempo real.

Ferramentas de desenvolvimento que suporta aplicação de reconhecimento de voz: Microsoft speech api, é uma tecnologia de voz API (SAPI) para Windows Embedded CE é um subconjunto da Microsoft Speech Application Programming Interface (SAPI) 5.0. SAPI 5.0 é um padrão de programação que fornece ferramentas e componentes de reconhecimento de fala e aplicativos de software text-to-speech.

Eclipse, é uma ferramenta IDE que compreende vários tipos de linguagem e que aceita a instalação de plugins para emular o desenvolvimento da plataforma.

As interfaces via voz estão rapidamente se tornando uma necessidade. Em um futuro próximo, sistemas interativos irão fornecer fácil acesso a milhares de informações e serviços que irão afetar de forma profunda a vida cotidiana das Pessoas.

Hoje em dia, tais sistemas estão limitados a pessoas que tenham acesso aos computadores, uma parte relativamente pequena da população, mesmo nos países mais desenvolvidos.



São necessários avanços na tecnologia de linguagem humana para que o cidadão médio possa acessar estes sistemas, usando habilidades de comunicação naturais e empregando aparelhos domésticos, tais como o telefone.

Sem avanços fundamentais em interfaces voltadas ao usuário, uma larga fração da sociedade será impedida de participar da era da informação, resultando em uma maior estratificação da sociedade, agravando ainda mais o panorama social dos dias de hoje.

Uma interface via voz, na linguagem do usuário, seria ideal pois é a mais natural, flexível, eficiente, e econômica forma de comunicação Humana. (YNOGUTI, 1999).

Surge à preocupação da interface do usuário com a máquina, para muitos especialistas, não adianta ter um sistema perfeito, se sua interação com o usuário é complicada, de difícil compreensão, sendo assim gera uma grande resistência para sua utilização e sua popularização.

A partir dessas situações surge à necessidade de cada vez mais a interação ser mais próxima de uma humana, seguindo essa linha de pensamento surgiu às plataformas de reconhecimento de voz, onde essa interação torna-se mais “amigável”.

O sistema de reconhecimento de voz, ele digitaliza a fala do comando dado, convertendo ela de analógico para digital, usando um conversor no qual capta as vibrações criadas pela voz do usuário, tornando em ondas dados digitais.

Aplicando medidas para todas as ondas captadas e som digitalizado e filtrando para separa ruídos, efetua uma computação das características espectral contido na voz. Sendo separados os sons fonéticos não maiores que uma sílaba, compara os sons captados com fonemas conhecidos em seu banco de dados. Sendo esse processo uma busca para associar a saída.



O sistema analisa o resultado e faz a comparação com palavras chaves e/ou frases conhecidas, e como resultado ele identifica e converte as palavras ditas pelo usuário em funcionalidades desejadas ou solicitadas.

De acordo com Zukerman e Litman (2001), é possível construir sistemas baseados em comandos de voz interativos e em tempo-real, em que as entradas do usuário sejam capturadas por um reconhecedor automático de voz e as saídas do sistema sejam enviadas através de um sintetizador de voz, ou através de mensagens previamente gravadas. (ZUKERMAN e LITMAN,2001).

Ferramenta escolhida

Utilizou-se como linguagem de programação para o desenvolvimento c sharp, c#, na plataforma de desenvolvimento Microsoft Visual Studio e com ferramentas da Microsoft para plataforma Windows.

A escolha desta linguagem foi pelo motivo dela ser uma linguagem legível, e pela facilidade de acesso as ferramentas, e programação gráfica além de ser essencial para desenvolvimentos de softwares em plataformas Windows.

O System.Speech.Recognition configurado em modo off-line, no qual possibilitou converter áudio falado em texto, nesse modo de configuração necessita carregar dentro do programa um vocabulário, esses métodos de reconhecimento de frases e/ou palavras são realizando internamente, quando o usuário pronuncia uma frase especifica, o computador reconhece e habilita uma ação logo em seguida.

Sendo assim com esta ferramenta pode-se criar um suporte virtual que nos possibilita reconhecer comandos de voz de um usuário, sem uso de internet, para isso deve se criar um mini vocabulário, e carregar no programa, no qual ele busca as palavras e realiza a comparação, e testa as condições pré-determinadas, e retorna a ação resultante do teste.



Uma observação importante à ferramenta pode ser configurada em modo online, sendo que com conexão web, ela pode buscar na rede as palavras para o reconhecimento delas, não sendo necessário adicionar em seu programa um vocabulário, pois ela busca em banco de dados online, mas para essa função a máquina precisa de um alto poder de processamento.

A ferramenta System.Speech.Synthesis foi a qual nos permitiu criar a voz do assistente virtual, sua função é a criação de voz humana artificial, podendo criar fala de pedaços de falas gravadas, vocabulários armazenados, mas sua principal aplicação é converter textos em fala.

Outras ferramentas utilizadas foram bibliotecas próprias da linguagem c#, na qual nos permitiu desenvolver alguns comandos e funções do protótipo tais como: Verificar data atual usando a propriedade Date Time.Now onde retorna um Date.Time que o valor representa a data e hora atuais do computador local.

Projeto will

O projeto tem como objetivo, desenvolver um software que consiga interagir com usuário através de reconhecimento de voz, com comandos específicos, ele consiga executar as tarefas, e dialogar com usuário, passando uma ideia, simulando uma interação mais parecida entre humanos, e não humano e máquina.

Algumas aplicações possíveis de serem realizadas com o sistema: aberturas de páginas na web, e pesquisas, interação com e-mail, facebook, interação com aplicações existente na máquina em que eu esteja instalado, como acesso a bloco de notas, mídia player, realizar cálculos, conectar ele a outros dispositivos via bluetooth, no qual através de comandos de voz manipule outros dispositivos.

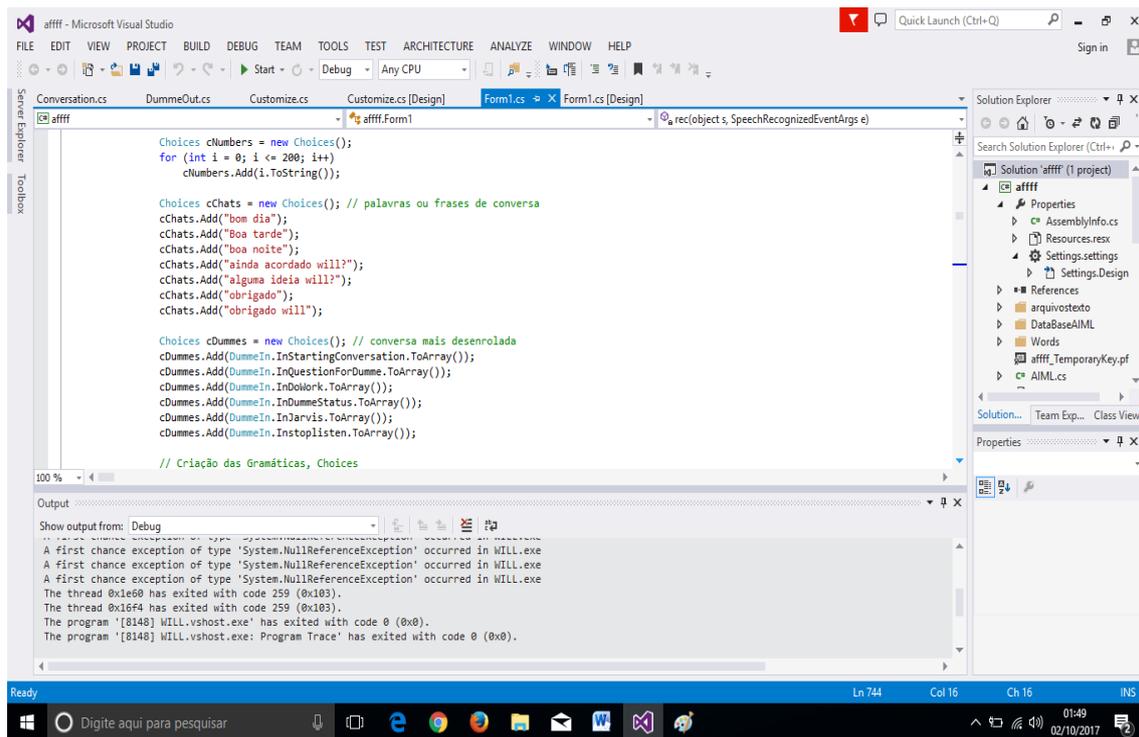


Ilustração 1 – Tela principal da ferramenta de desenvolvimento
Fonte: Software Microsoft Visual Studio

Essa ilustração traz uma parte do código fonte, desenvolvido na ferramenta Visual Studio no qual deu início ao desenvolvimento do projeto. Onde mostra a configuração de uma parte da interação com diálogo, e também as bibliotecas e funções de reconhecimento de voz.

Mas o foco principal desse protótipo é a interação homem e máquina, onde essa seja mais parecida com a interação entre duas pessoas, num diálogo, tanto formal como informal, onde sua interação seja imediata.

Aplicou-se vários testes como diálogos, navegações na internet, pesquisas no google, a interação dele com os softwares como calendário, bloco de notas, navegação por notícias, e sua manipulação com Arduino onde controlou via bluetooth, executando a seguinte tarefa: acender e apagar leds.



Ilustração 2 – Tela da interface do Will em uma interação
Fonte: Produção do pesquisador

A ilustração2 mostra uma interação da interface do Will, ele busca no site de clima, com base na localização do computador onde esta instalado, mostrando tanto na tela escrita como na voz, o clima atual, a temperatura, humidade, nesse caso na cidade de Santo Ângelo-RS.

A Ferramenta de Desenvolvimento

O processo de análise, foram elaborados testes para o desenvolvimento do **Will**, para mostrar o processo será descrito sobre a linguagem no qual o programa está sendo desenvolvido.

C# é uma linguagem elegante, orientada a objeto e fortemente tipada, que permite que os desenvolvedores criem uma variedade de aplicativos robustos e muito mais. A sintaxe de C# é altamente expressiva, mas também é simples e fácil de aprender. A sintaxe de chaves de C# será instantaneamente reconhecível para qualquer pessoa familiarizada com C, C++ ou Java.

Então, com muita pesquisa conseguiu-se uma ferramenta que permitisse manipular eventos de reconhecimento de voz de uma maneira rápida e eficiente.



Também estudos sobre o desenvolvimento de códigos exclusivos em grupos de programadores, pois a ajuda de desenvolvedores que já atuavam com o mesmo assunto foi de grande valia para o estudo, que em tão pouco tempo conseguiu chegar em um resultado surpreendente.

Um dos resultados foi fazer com que o **Will**, reconheça os comandos programados em português brasil ou também inglês-US, sem ter que alterar o idioma da exibição do sistema operacional Windows.

Definição de Comandos/Recursos

Comandos foram desenvolvidos a partir da ferramenta da Microsoft que forneceu o reconhecimento de voz e através dela, claro, na linguagem de programação `c#` foi aperfeiçoado recursos e funções que pudessem realizar a operação de simular uma conversa entre um humano e uma máquina.

Foi utilizado regras lógicas e conjunto de dados disponíveis pela ferramenta para que produzisse uma capacidade de conseguir interagir e aplicar raciocínio nas situações do cotidiano.

Um dos recursos que o Will possui é de que não é necessário estar conectado a internet para que ele reconheça os comandos definidos pelo programador. Destacando-se de softwares similares como Cortana, Siri, Genesys e etc.

Uma vantagem é o fato do Will responder comandos imediatamente, já que não precisa passar por um servidor de internet.

Conclusão

O assistente virtual Will, consiste numa inteligência artificial programada para reconhecer, interpretar e executar comandos dados por voz ou por texto. O objetivo do projeto foi atingido, pois desenvolveu-se e testou-se um software com capacidade de interagir e comunicar-se com o usuário.



Através desta pesquisa, foi possível desenvolver e testar como um software pode simular a conversação entre pessoas, após o envio de perguntas em linguagem natural, o programa consulta uma base de conhecimento e em seguida fornece uma resposta que tenta imitar o comportamento humano, utilizando conceitos de programação, de IHC - Interface homem-máquina e inteligência artificial.

O will é capaz de aprender coisas novas, e compartilhar seus conhecimentos para suas outras cópias, tornando-se, cada vez mais inteligente. Ele sempre estará em evolução, e não existirá maneira correta de se comunicar com ele, você pode falar como se fosse uma pessoa real. Com sua personalidade, suas ações e a maneira de conversar com ele, irá implicar na alteração de seu humor, fazendo com que ele se torne único e suas ações se tornem inesperadas.

Referências

Automação de Ambientes. Disponível em: <<http://sistemamyway.com.br>>. Acesso em: 27 de setembro de 2017.

Como funciona o reconhecimento de voz. Disponível em: <<https://www.tecmundo.com.br/curiosidade/3144-como-funciona-o-reconhecimento-de-voz-.htm> php>. Acesso em: 27 de setembro de 2017.

Desenvolvendo Soluções com Interface Baseada em Voz - Capítulo 2 - Livro de tutoriais do XI Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais. Cuiabá – MT, de 5 a 9 de Novembro de 2012 . SBC, 2012.

Interação humano computador. Disponível em: <<https://www.passeidireto.com/arquivo/21173773/livro-unopar-interacao-humano-computador>>. Acessado em 27 de setembro de 2017.

Introdução a Linguagem C# e ao .net Framework. Disponível em: < <https://docs.microsoft.com/pt-br/dotnet/csharp/getting-started/introduction-to-the-csharp-language-and-the-net-framework>>. Acesso em: 27 de setembro de 2017.

MORAIS, Everson Matias de; LOPER, Adriane Aparecida. **Interação humano-computador**, 2014.

O que é touchscreen. Disponível em: <<https://www.tecmundo.com.br/multitouch/177-o-que-e-touch-screen-.htm>>. Acesso em: 10 de outubro de 2017.

Softwares Deficientes Visuais. Disponível em:



<<http://www.acessibilidadeinclusiva.com.br/programas-para-computadores>>.
Acesso em: 27 de setembro de 2017.

YNOGUTI , Carlos Alberto. **Reconhecimento de Fala Contínua Usando Modelos Ocultos de Markov**. Universidade Estadual de Campinas, 1999.