

UNIVERSIDADE REGIONAL DE BLUMENAU
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
CURSO DE CIÊNCIA DA COMPUTAÇÃO – BACHARELADO

**VIKI: INTELIGÊNCIA VIRTUAL COM INTERAÇÃO POR
VOZ APLICADA À INTERNET DAS COISAS**

LUÍS HENRIQUE GOLL

BLUMENAU
2016

LUÍS HENRIQUE GOLL

**VIKI: INTELIGÊNCIA VIRTUAL COM INTERAÇÃO POR
VOZ APLICADA À INTERNET DAS COISAS**

Trabalho de Conclusão de Curso apresentado ao curso de graduação em Ciência da Computação do Centro de Ciências Exatas e Naturais da Universidade Regional de Blumenau como requisito parcial para a obtenção do grau de Bacharel em Ciência da Computação.

Profa. Joyce Martins, Mestre - Orientadora

**BLUMENAU
2016**

VIKI: INTELIGÊNCIA VIRTUAL COM INTERAÇÃO POR VOZ APLICADA À INTERNET DAS COISAS

Por

LUÍS HENRIQUE GOLL

Trabalho de Conclusão de Curso aprovado para obtenção dos créditos na disciplina de Trabalho de Conclusão de Curso II pela banca examinadora formada por:

Presidente: _____
Profa. Joyce Martins, Mestre – Orientadora, FURB

Membro: _____
Prof. Miguel Alexandre Wisintainer, Mestre – FURB

Membro: _____
Prof. Roberto Heinzle, Doutor – FURB

Blumenau, 05 de dezembro de 2016

Dedico este trabalho à minha mãe, por ter sempre me apoiado e incentivado em todos os objetivos que eu já tive na vida e, em especial, neste trabalho.

AGRADECIMENTOS

A Deus, por me permitir que cada dia possa ser um novo começo na minha jornada da vida.

Aos meus amigos e familiares, Camila Viviani Martins, Carolina Kovaleski De Souza, Daniel Gielow Junior, Douglas Eduardo dos Santos, Jean Carlos Kruger, Juliano Mueloschat, Fabiola Telles Goll, Filipe Rodrigo Miguel, Johnny Ricardo Manke, Luan Christian Froehlich, Luana Bachmann, Maria Ilésia Marquato, Thiago Henrique Goll, William Christopher Schepers, Yan Schreiber, que se dispuseram a testar e avaliar o desenvolvimento do trabalho.

Ao meu amigo Rodrigo de Sousa, pela ajuda com o circuito infravermelho para controle da televisão.

Aos meus professores Alexander Roberto Valdameri, Aurélio Faustino Hoppe, Mauro Marcelo Mattos, Miguel Alexandre Wisintainer, que souberam do meu projeto e contribuíram ou incentivaram para a realização dele.

À minha orientadora Joyce Martins, pelo apoio e dedicação, incentivo a realizar este trabalho como TCC e pela amizade.

Felicidade é alguém para amar, algo para fazer
e algo para aspirar.

Joseph Addison

RESUMO

Este trabalho apresenta o desenvolvimento de uma aplicação de inteligência virtual com conversação por voz voltada para Internet das Coisas. A aplicação, denominada VIKI, realiza o reconhecimento da voz do usuário e transforma um comando recebido em uma ação. Essas ações podem ser, por exemplo, ligar uma luz, ligar uma televisão, trocar de canal ou uma simples resposta da VIKI dependendo do comando ou do contexto. Para a realização do reconhecimento e síntese de voz foi utilizada a API Microsoft Speech Platform. Destaca-se também o uso de dois modelos do microcontrolador ESP8266 necessários para o controle dos dispositivos físicos (luz e televisão). Foram feitos experimentos no intuito de validar o funcionamento e a usabilidade da VIKI. Os resultados obtidos demonstram que a aplicação desenvolvida é capaz de dialogar com o usuário, embora com algumas limitações, e realizar as ações correspondentes aos comandos de forma satisfatória.

Palavras-chave: Reconhecimento de voz. Síntese de voz. Internet das coisas. Inteligência virtual.

ABSTRACT

This work presents the development of a virtual intelligence application with speech recognition applied to the Internet of Things. The application, denominated VIKI, performs speech recognition from the user's voice and transform a received command into an action. These actions can be, for instance, turn on some light, turn on a television, switching channel or a simply answer from VIKI depending on the command or context in which the application stands. In order to achieve the speech recognition and synthesis was used the API Microsoft Speech Platform. The use of two ESP8266 microcontrollers were also required to control physical devices (light and television). Experiments were made in order to validate VIKI's operation and usability. The results obtained show that the developed application is able to dialogue with the user, although with some limitations, and perform the correspondent command actions satisfactorily.

Key-words: Speech recognition. Speech synthetizer. Internet of things. Virtual intelligence.

LISTA DE FIGURAS

Figura 1 – Principais áreas da inteligência artificial.....	19
Figura 2 – Assistente pessoal Lili.....	20
Figura 3 – Assistente pessoal Cortana.....	21
Figura 4 – Programa de dicção e reconhecimento de voz Dragon Naturally Speaking	22
Figura 5 – Cenários de conversação e ações da VIKI	24
Figura 6 – Diagrama de fluxo da interação com a VIKI	26
Figura 7 – Diagrama de casos de uso	27
Figura 8 – Diagrama de pacotes	28
Figura 9 – Placa AFSmart 031604 e microcontrolador ESP8266-01.....	30
Figura 10 – Fonte HLK-PM01	30
Figura 11 – Dispositivos para controle da luz	31
Figura 12 – Dispositivos para controle da televisão.....	32
Figura 13 – Diagrama de fluxo para controle da luz	33
Figura 14 – Diagrama de fluxo para controle a televisão.....	34
Figura 15 – Tela principal da VIKI	37
Figura 16 – Notificação de frase reconhecida	37
Figura 17 – Configuração de IP na VIKI	55
Figura 18 – Configuração de IP do microcontrolador da luz	55
Figura 19 – Esquemática da placa	58

LISTA DE QUADROS

Quadro 1 – Pacotes e classes da aplicação	28
Quadro 2 – Perfil dos voluntários do experimento 01	39
Quadro 3 – Avaliação de usabilidade do experimento 01	39
Quadro 4 – Perfil dos voluntários do experimento 02.....	40
Quadro 5 – Avaliação de usabilidade do experimento 02	41
Quadro 6 – Comparativo entre o trabalho desenvolvido e os trabalhos correlatos	42
Quadro 7 – Definição em XML do comando para ligar a luz	47
Quadro 8 – Definição em XML do comando para trocar de canal por número	48
Quadro 9 – Propriedades da estrutura XML.....	48
Quadro 10 – Respostas internas da VIKI	49
Quadro 11 – UC01: Ação VIKI	51
Quadro 12 – UC02: Ação stand-by	51
Quadro 13 – UC03: Ação agradecer	51
Quadro 14 – UC04: Ação cancelar	51
Quadro 15 – UC05: Ação repetir fala	51
Quadro 16 – UC06: Ação ignorar	51
Quadro 17 – UC07: Ação pedir ajuda	52
Quadro 18 – UC08: Ação conversar	52
Quadro 19 – UC09: Ação ligar luz	52
Quadro 20 – UC10: Ação desligar luz	52
Quadro 21 – UC11: Ação ligar televisão	52
Quadro 22 – UC12: Ação desligar televisão	52
Quadro 23 – UC13: Ação aumentar volume	53
Quadro 24 – UC14: Ação diminuir volume	53
Quadro 25 – UC15: Ação colocar no mudo	53
Quadro 26 – UC16: Ação tirar do mudo	53
Quadro 27 – UC17: Ação trocar de canal por número	53
Quadro 28 – UC18: Ação trocar de canal nome	53
Quadro 29 – UC19: Ação trocar para o último canal	54
Quadro 30 – Questões de perfil do usuário	56

Quadro 31 – Questões de usabilidade e opinião.....	56
--	----

LISTA DE ABREVIATURAS E SIGLAS

API – Application Programming Interface

IA – Inteligência Artificial

IO – Input and Output

IDE – Integrated Development Environment

RF – Requisito Funcional

RNF – Requisito Não Funcional

TI – Tecnologia da Informação

USB – Universal Serial Bus

XML – eXtensible Markup Language

SUMÁRIO

1 INTRODUÇÃO.....	14
1.1 OBJETIVOS.....	15
1.2 ESTRUTURA.....	15
2 FUNDAMENTAÇÃO TEÓRICA	16
2.1 RECONHECIMENTO E SÍNTESE DE VOZ.....	16
2.2 INTERNET DAS COISAS	17
2.3 INTELIGÊNCIA VIRTUAL.....	18
2.4 TRABALHOS CORRELATOS.....	19
2.4.1 Fibaro	19
2.4.2 Cortana	20
2.4.3 Dragon Naturally Speaking.....	22
3 DESENVOLVIMENTO DA VIKI	23
3.1 REQUISITOS.....	23
3.2 ESPECIFICAÇÃO	23
3.2.1 Cenários de interação	24
3.2.2 Fluxo da interação com a VIKI.....	25
3.2.3 Diagramas de casos de uso e de pacotes	26
3.3 CONTROLE DOS DISPOSITIVOS FÍSICOS	29
3.3.1 Definição e montagem dos microcontroladores e placas.....	29
3.3.2 Implementação dos microcontroladores	32
3.4 IMPLEMENTAÇÃO	34
3.4.1 Técnicas e ferramentas utilizadas.....	34
3.4.2 Operacionalidade da implementação	36
3.5 ANÁLISE DOS RESULTADOS	38
3.5.1 Experimento 01: uso da VIKI em um ambiente real.....	38
3.5.2 Experimento 02: uso da VIKI em um ambiente simulado.....	40
3.5.3 Comparativo entre os trabalhos correlatos.....	41
4 CONCLUSÕES	43
4.1 EXTENSÕES	44
REFERÊNCIAS	45
APÊNDICE A - DESCRIÇÃO DA ESTRUTURA DO COMANDO E DO XML	47

APÊNDICE B - DESCRIÇÃO DOS CASOS DE USO.....	50
APÊNDICE C – INSTRUÇÕES DE CONFIGURAÇÃO DA APLICAÇÃO	55
APÊNDICE D – QUESTIONÁRIO DE AVALIAÇÃO DE USABILIDADE	56
ANEXO A – ESQUEMA ELÉTRICO DA PLACA UTILIZADA	58

1 INTRODUÇÃO

Hoje a tecnologia está presente no dia a dia. Existem diversos recursos e aplicativos que auxiliam no cotidiano de praticamente todas as pessoas. Há também nos últimos anos uma vasta expansão da Internet das Coisas (em inglês Internet of Things - IoT), que é a integração de dispositivos do mundo físico com o mundo virtual. No entanto, segundo Goll (2012), observa-se que toda esta tecnologia tem uma curva de aprendizado associada à árdua necessidade do público geral ter que se adaptar a ela diariamente. Considerando esse cenário, pergunta-se: por que as pessoas têm que se adaptar à tecnologia? Não deveria ser ao contrário, a tecnologia se adaptar aos indivíduos? (GOLL, 2012). Nesse sentido, é indispensável que a interação seja facilitada e simplificada antes que fique impossível gerenciar e controlar tantos dispositivos e aplicativos.

Na comunicação diária, a linguagem natural é a linguagem utilizada pela maioria das pessoas. Indivíduos facilmente desenvolvem domínio do seu primeiro idioma durante os primeiros anos de vida. Assim, segundo Lee, Chang e Hsieh (2014), a flexibilidade sintática e semântica de uma linguagem natural permite que este tipo de comunicação seja natural para seres humanos. Por isso, usar a linguagem natural para responder e tirar dúvidas, solucionar problemas, enfim, interagir de forma rápida e efetiva, tem levado “a uma crescente demanda por softwares de assistentes virtuais inteligentes” (DINO VISIBILIDADE ONLINE, 2016).

Diante do problema exposto, desenvolveu-se uma aplicação denominada de VIKI. VIKI usa inteligência virtual para interagir com as pessoas pelo reconhecimento e síntese de voz e para analisar contextos e diálogos a partir do processamento da linguagem natural, em conjunto com o conceito de informação integrada para permitir a comunicação e o controle de dispositivos do mundo físico. Em suma, VIKI é uma inteligência virtual que se comunica pela voz e é capaz de unificar tecnologias de forma simples e natural, controlando luzes e aparelhos eletrodomésticos. Com ela o usuário pode controlar seu ambiente físico de forma natural utilizando apenas a voz e obtendo feedback da aplicação conforme o uso. A aplicação também distingue diferentes contextos dentro deste ambiente e questiona o usuário em caso de dúvidas ou ambiguidades, deixando toda interação fluída e intuitiva, similar à comunicação entre pessoas.

VIKI visa à integração e acesso das pessoas através do rompimento da dificuldade de uso de sistemas e equipamentos tecnológicos. Ainda neste mesmo aspecto, espera-se que a ferramenta desenvolvida possa ajudar indivíduos a utilizarem sistemas e soluções tecnológicas de forma natural e intuitiva. O emprego de uma aplicação capaz de conversar e

tomar ações somente pela linguagem natural também visa gerar uma maior aceitação e aderência a soluções tecnológicas que antes poderiam ser descartadas pela sua curva de aprendizado do consumidor final. Usuários comuns e de áreas não relacionadas à informática muitas vezes passam dificuldades para interagir e entender novas tecnologias que exigem um grande processo de aprendizado e adaptação.

1.1 OBJETIVOS

Este trabalho tem como objetivo desenvolver uma ferramenta de inteligência virtual com conversação por voz aplicada à Internet das Coisas.

Os objetivos específicos deste trabalho são:

- a) reconhecer e sintetizar voz, permitindo toda interação do usuário ser exclusivamente por voz, em português brasileiro;
- b) ser capaz de tomar ações a partir de um diálogo com o usuário, como por exemplo questionar o usuário em caso de dúvidas e processar a resposta, ou ainda processar sentenças que façam sentido diante do contexto da última sentença reconhecida, dando a impressão de uma continuidade no diálogo;
- c) ser integrada com automação de alguns eletrodomésticos, tais como televisão e luzes, em um determinado ambiente.

1.2 ESTRUTURA

Este trabalho está dividido em quatro capítulos. O primeiro capítulo apresenta a introdução e os objetivos do trabalho. O segundo capítulo apresenta a fundamentação teórica sobre reconhecimento e síntese de voz, Internet das Coisas e inteligência virtual. Descreve também ferramentas similares à VIKI. O terceiro capítulo mostra o desenvolvimento do trabalho com requisitos, especificação, implementação, operacionalidade da aplicação e resultados obtidos. Por último, o quarto capítulo relata as conclusões e extensões que poderiam ser desenvolvidas.

2 FUNDAMENTAÇÃO TEÓRICA

A seção 2.1 expõe uma visão geral sobre reconhecimento e síntese de voz. A seção 2.2 aborda o termo Internet das Coisas e justifica sua importância neste trabalho. A seção 2.3 oferece uma visão sobre inteligência virtual e sua crescente demanda. Por fim, a seção 2.4 traz um estudo de três trabalhos correlatos com aspectos relevantes para o trabalho proposto.

2.1 RECONHECIMENTO E SÍNTESE DE VOZ

Reconhecimento de voz é, conforme descreve Adams (1990 apud BAUMANN, 1993), a tecnologia pela qual sons, palavras ou frases faladas por humanos são convertidos em sinais elétricos, os quais são transformados em padrões de código com um significado. Já a síntese de voz é basicamente o processo contrário, que transforma qualquer texto do mundo virtual em sons de fala com elevados graus de inteligibilidade e naturalidade (VOICE INTERACTION DO BRASIL, 2016).

Tanto o reconhecimento quanto a síntese de voz são excelentes formas de imersão e acessibilidade dos usuários de forma geral, pois permitem que a interação com computadores seja feita de uma maneira mais natural e intuitiva. Também é um meio de permitir que pessoas com deficiência visual possam utilizar computadores de forma muito melhor. Nos dias de hoje já existem muitos produtos e soluções que oferecem reconhecimento de voz de qualidade para controlar sistemas operacionais e aplicativos, além de utilizar e navegar na internet de forma geral. É uma área em demanda também para usuários comuns, que possuem dificuldades de operar sistemas ou tecnologias complexas. Utilizando a voz como forma de entrada e saída, sistemas ou tecnologias podem se tornar mais fáceis de serem operados.

No entanto, devido à alta complexidade computacional para processar e interpretar a voz humana, tornou-se comum a criação e a reutilização de programas, ou Application Programming Interface (APIs), como, por exemplo, a Microsoft Speech Platform (MICROSOFT CORPORATION, 2016b), que é utilizada no trabalho desenvolvido. A API em questão é capaz de reconhecer e sintetizar voz em vinte e seis idiomas, incluindo o português brasileiro e pode ser utilizada livremente por qualquer programa desenvolvido no *framework* .NET ou na linguagem de programação C++ dentro da plataforma Windows (MICROSOFT CORPORATION, 2016b).

A Microsoft Speech Platform reconhece somente sentenças pré-definidas, isto é, todas as palavras ou frases que serão reconhecidas devem ser determinadas explicitamente. Porém, a gramática com as sentenças pode ser definida em tempo de execução, utilizando palavras ou frases alternativas, podendo também ser utilizado, opcionalmente, em conjunto com coringas.

Coringas indicam para a API que qualquer coisa falada será aceita, mas o que foi falado antes do coringa deve corresponder à sentença definida. Também é possível atribuir semântica para as frases, o que permite que várias sentenças possam ser vinculadas a um mesmo significado ou comando. Quando é feito o reconhecimento de voz, a API dispara um evento informando qual foi a frase reconhecida, sua semântica e precisão estimada do reconhecimento. Quanto à síntese de voz, a API permite controlar de forma simplificada entonação, volume e velocidade das palavras a serem sintetizadas.

2.2 INTERNET DAS COISAS

Internet das Coisas é uma rede de objetos do cotidiano que podem compartilhar informações e realizar tarefas em conjunto do mundo físico com o mundo virtual (SAS INSTITUTE INC, 2016). Conceitualmente, Lira (2013) afirma que é a possibilidade de conectar o mundo físico com o mundo digital por meio da web. Esta conectividade do mundo físico inclui desde eletrodomésticos até veículos e máquinas industriais. A Internet das Coisas é então a conexão de todos estes objetos para automatizar e melhorar a vida das pessoas e corporações (SANTAELLA et al., 2013).

O acesso à internet por dispositivos móveis cresce e empresas estão cada vez mais online (CÂMARA DE COMÉRCIO E INDÚSTRIA BRASIL-ALEMANHA DE PORTO ALEGRE, 2012). Assim, é natural que tarefas repetitivas, tanto de indivíduos quanto de corporações, comecem a vislumbrar a automação e praticidade que a Internet das Coisas pode oferecer. Já é possível estimar uma grande demanda de desenvolvedores para esta área, pois em 2014 o número de desenvolvedores para Internet das Coisas era de 300 mil, e estima-se que este segmento aumente tão drasticamente até 2020 que chegará a ser necessário cerca de 4.5 milhões de desenvolvedores na área (ASAY, 2014). A Internet das Coisas está em rápida expansão no momento e é perceptível que o mundo físico está evoluindo e ficará cada vez mais conectado com o mundo virtual, mudando a forma que as pessoas interagem com aparelhos do cotidiano.

Existem diversos modelos e placas de microcontroladores e dispositivos para Internet das Coisas, sendo que neste trabalho foram escolhidos dois modelos do microcontrolador ESP8266 (ESP8266 COMMUNITY FORUM, 2016), um para controlar a televisão e outro para a luz (LIMA, 2016). O ESP8266 NodeMCU possui entrada mini Universal Serial Bus (USB) para alimentação e gravação do código fonte, o que facilita o desenvolvimento pois basta conectar no USB para enviar um código novo. Como já conta com um conversor USB embutido e trabalha com entrada na tensão de 5V, pode ser utilizado no emissor

infravermelho para simular um controle remoto e emitir os sinais para a televisão. Já o ESP8266-01 é uma das versões mais compactas do ESP8266. Este modelo tem tamanho reduzido, sendo facilmente adaptado em um interruptor comum para controlar uma luz.

2.3 INTELIGÊNCIA VIRTUAL

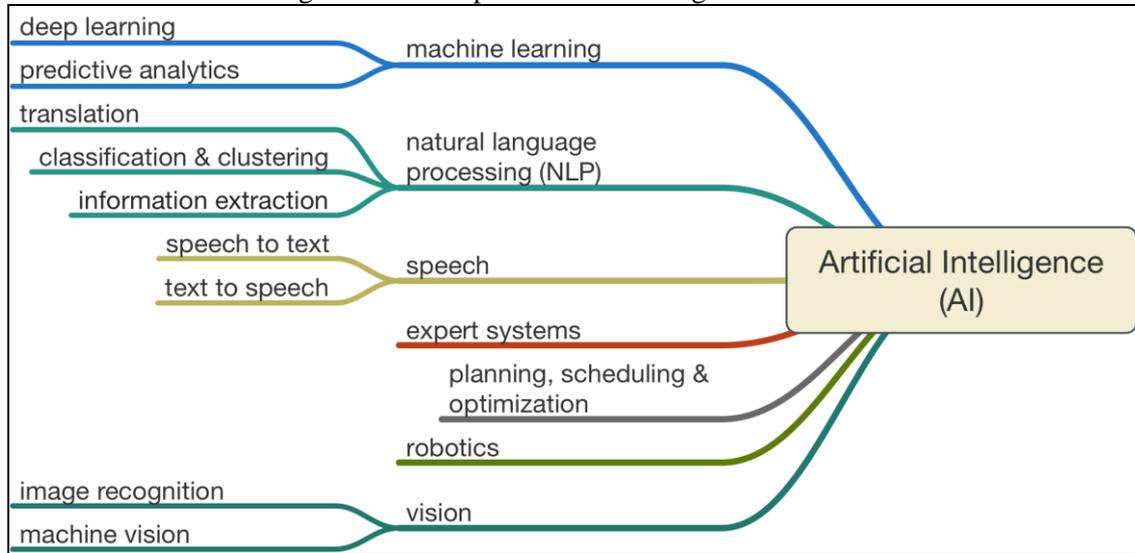
A inteligência virtual pode ser considerada um ramo da Inteligência Artificial (IA). No entanto, Sato (2009) afirma que não existe uma definição ou um conceito exato para IA, mas é basicamente fazer com que computadores pensem ou sejam tão inteligentes quanto os seres humanos. Assim, um dos objetivos desta área é que possa substituir os humanos em determinadas atividades que podem ser consideradas repetitivas ou até mesmo perigosas.

A inteligência artificial, segundo Computer Science Degree HUB (2015), tem uma ampla aplicação, incluindo área comercial, governamental e até mesmo usuário final. Cita-se, como exemplo, a utilização da IA para criar sistemas especialistas capazes de coletar dados, processá-los e tornar estes dados úteis para tomar decisões em determinada aplicação. Também pode ser aplicada na agricultura, em sistemas de saúde ou militares. A área de robótica utiliza-se da IA para construir máquinas capazes de trabalhar em áreas que não são adequadas para os seres humanos, como no fundo do oceano ou no espaço, além de também ser utilizada em entidades industriais de grande porte. Usuários se beneficiam da IA quando há necessidade de aprender ou compreender um outro idioma sem o acompanhamento de um tutor. Programas de reconhecimento de voz permitem que alunos estrangeiros pratiquem um segundo idioma por meio do computador e recebam correções conforme sua pronúncia. Sendo assim, a inteligência artificial também é fortemente utilizada em programas de tradução que permite que usuários convertam texto de uma linguagem para a outra. Na Figura 1 é possível ver algumas das principais áreas e ramificações da inteligência artificial, muitas delas interligadas com outras áreas, tais como robótica, aprendizado de máquina, processamento de linguagem natural e processamento da fala.

Nesse contexto, a inteligência virtual em específico é geralmente um programa que toma ou fornece decisões inteligentes, baseado em uma situação previamente programada. A inteligência virtual por si só não é inteligente, pois não é capaz de evoluir sozinha e tomar decisões independentes, mas aparenta ser e tenta interagir de forma natural com um humano já que foi programada para agir assim. Portanto, utiliza-se do processamento de linguagem natural e da fala. Hoje há uma grande necessidade de responder, tirar dúvidas e interagir com usuários de forma rápida e eficiente, somada a uma grande quantidade de informações e dados recebidos constantemente. Por isso, empresas vêm procurando por soluções que otimizem este

processo (DINO VISIBILIDADE ONLINE, 2016), tornando a inteligência virtual uma área promissora e importante no mercado e também no campo computacional.

Figura 1 – Principais áreas da inteligência artificial



Fonte: Mills (2015).

2.4 TRABALHOS CORRELATOS

São descritas três aplicações correlatas ao trabalho proposto: Fibaro (FIBAR GROUP, 2016), com foco na automação residencial, incluindo desde itens de lazer até segurança; Cortana (MICROSOFT CORPORATION, 2016a), aplicativo de assistência virtual, possuindo reconhecimento e sintetizador de voz que permite uma maior interação com o usuário; Dragon Naturally Speaking (NUANCE COMMUNICATIONS, 2016), que fornece controle completo do computador somente por voz.

2.4.1 Fibaro

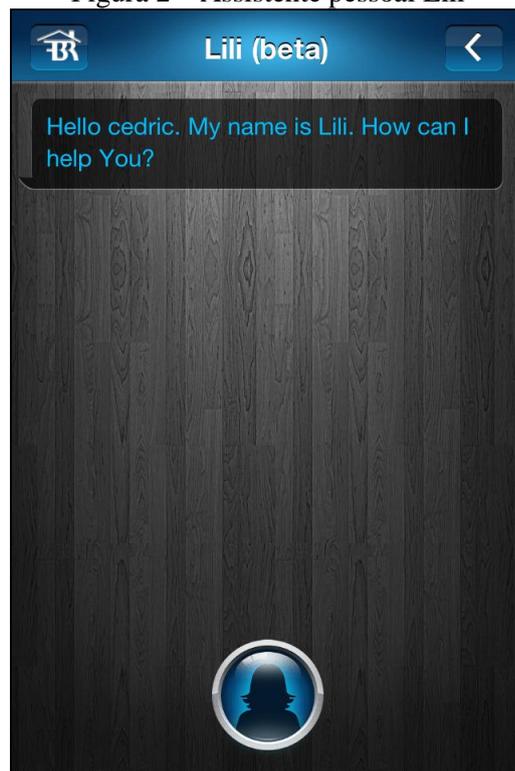
O Fibaro, desenvolvido pelo Fibar Group (2016), é um sistema completo de automação residencial sem fio. Possui uma ampla gama de dispositivos que permitem controlar a maioria dos aparelhos domésticos de uma residência, tais como: luzes, televisões, cortinas, alarmes, câmeras, entre outros. Todo o controle destes dispositivos é feito a partir de uma central que pode ser acessada remotamente. Também contém uma pequena inteligência virtual, denominada Lili, que permite comandar ações por voz. No entanto, ela não dialoga com o usuário e todos os comandos são formados por sentenças simples e pré-definidas, como por exemplo *Turn on the light*, a única sentença utilizada para ligar a luz.

Para utilizar o sistema Fibaro é preciso ter a central do sistema instalada e um módulo para cada dispositivo que será controlado. O produto Fibaro oferece na sua central o controle de todos os dispositivos adquiridos e instalados na residência, também possui aplicativos para

Android e iOS que permitem fazer estas mesmas ações remotamente. Por meio destes aplicativos, é possível interagir com Lili, que permite controlar os principais dispositivos utilizando somente a voz (FIBAR GROUP, 2016).

Embora seja um sistema completo e funcional, a parte de automação residencial não é compatível de fábrica com os padrões de tomadas do Brasil. E apesar de existirem empresas e produtos à disposição que realizem estas modificações e adaptações no Brasil, isto acaba gerando um impacto no preço final do produto. Além disso, a inteligência virtual Lili oferece somente controle das ações por palavras e sentenças estáticas, sem nenhum controle de contexto ou diálogo, o que deixa toda interação um tanto robótica. Por fim, toda a solução, desde o aplicativo do sistema até a inteligência virtual, não está disponível em português, o que gera uma maior dificuldade e barreira de uso para o público brasileiro. A Figura 2 mostra a tela inicial da Lili.

Figura 2 – Assistente pessoal Lili



Fonte: Fibar Group (2016).

2.4.2 Cortana

Cortana é uma assistente pessoal inteligente desenvolvida pela Microsoft Corporation (2016a) e disponibilizada para plataforma Windows. Esta inteligência virtual funciona com reconhecimento e síntese de voz. Seu processamento de voz permite realizar tarefas e comandos pré-definidos a partir de um reconhecimento de voz livre, com base em uma análise semântica do que foi falado pelo usuário. A partir da frase ou comando identificado, a

assistente também é capaz de responder e interagir com o usuário por voz, incluindo respostas com diferentes entonações, dando a impressão de uma interação muito natural e similar à de um humano.

A assistente está disponível nas plataformas mais recentes da Microsoft, incluindo computadores, smartphones e Xbox One, sendo possível interagir com os usuários por reconhecimento de voz, respondendo às perguntas ou dando informações utilizando síntese de voz. Cortana possui diferentes comandos e funcionalidades dependendo da plataforma em que está sendo utilizada, mas de forma geral permite realizar as ações mais básicas e comuns exclusivamente por voz, como por exemplo gerenciamento de agenda, pesquisas na internet e navegação pelo sistema.

A inteligência virtual possui uma excelente síntese de voz e possui todas as funcionalidades no idioma português brasileiro. No entanto, seu foco é puramente ser um programa de computador que auxilia e interage com o usuário neste mundo virtual que a plataforma oferece, deixando de fora a Internet das Coisas e o mundo físico. A Figura 3 mostra a interface inicial da Cortana.

Figura 3 – Assistente pessoal Cortana



Fonte: elaborado pelo autor.

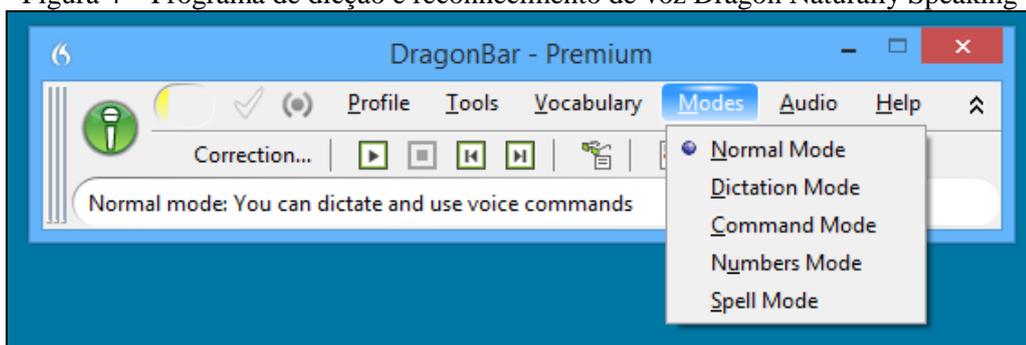
2.4.3 Dragon Naturally Speaking

Dragon Naturally Speaking é um programa de computador de reconhecimento e controle por voz desenvolvido pela Nuance Communications (2016). É disponibilizado para diversas plataformas e tem como foco um reconhecimento de voz rápido e preciso, sendo recomendado para pessoas com deficiência visual para navegação no sistema operacional e editores de texto. Também pode ser usado por escritores que queiram digitalizar textos e livros somente usando a voz.

O aplicativo pode ser utilizado nos sistemas operacionais Windows, Mac OS X e Linux, além das versões para Android e iOS para dispositivos móveis. Para utilizar a aplicação, a mesma exige um treinamento da voz do usuário na primeira vez em que for executada, o qual permite que o algoritmo da aplicação consiga se adaptar a diferentes sotaques de acordo com o treinamento realizado. A ferramenta permite controlar alguns aspectos básicos do sistema, como, por exemplo, abrir aplicativos e pesquisar na internet, além de permitir definir macros, ou seja, definir uma série de ações que devem ser executadas no sistema quando determinada frase ou comando for reconhecido.

A aplicação não trabalha com o idioma português, mas oferece suporte e adaptação para diferentes sotaques de todo o mundo no idioma inglês. Esta aplicação tem como foco o reconhecimento de voz livre e não possui uma inteligência virtual. Embora possua uma síntese de voz, a mesma é bastante robótica, limita-se apenas a repetir e confirmar comandos ou frases ditadas pelo usuário. A Figura 4 mostra a tela inicial do Dragon Naturally Speaking.

Figura 4 – Programa de ditação e reconhecimento de voz Dragon Naturally Speaking



Fonte: Anderson (2014).

3 DESENVOLVIMENTO DA VIKI

Este capítulo apresenta o desenvolvimento da VIKI. A seção 3.1 apresenta os requisitos da aplicação. A seção 3.2 exibe a especificação do trabalho. A seção 3.3 mostra a definição e implementação dos microcontroladores para controle dos dispositivos físicos. A seção 3.4 descreve as ferramentas utilizadas e a operacionalidade da aplicação. Por fim, a seção 3.5 discute os resultados obtidos.

3.1 REQUISITOS

Os requisitos da aplicação são:

- a) permitir o controle da luz a partir da aplicação (Requisito Funcional - RF);
- b) permitir o controle da televisão a partir da aplicação (RF);
- c) ser capaz de questionar o usuário em caso de dúvidas e processar a resposta do usuário (RF);
- d) ser capaz de reconhecer por voz sentenças pré-definidas (RF);
- e) ser capaz de sintetizar respostas diferentes (RF);
- f) a aplicação deve ser desenvolvida na linguagem C# (Requisito Não Funcional - RNF);
- g) a aplicação deve utilizar a API Microsoft Speech Platform para realizar o reconhecimento e a síntese de voz (RNF);
- h) a aplicação deve ser totalmente funcional por conversação por voz (RNF);
- i) a aplicação deve ser de fácil utilização (RNF);
- j) a aplicação deve se comunicar com microcontroladores via wi-fi para efetuar o controle dos dispositivos físicos (RNF);
- k) a programação do microcontrolador deve ser desenvolvida na linguagem C++ (RNF);
- l) o controle da luz deve ser realizado por relé através de um ESP8266-01 (RNF);
- m) o controle da televisão deve ser realizado por led emissor infravermelho através de um ESP8266 NodeMCU (RNF).

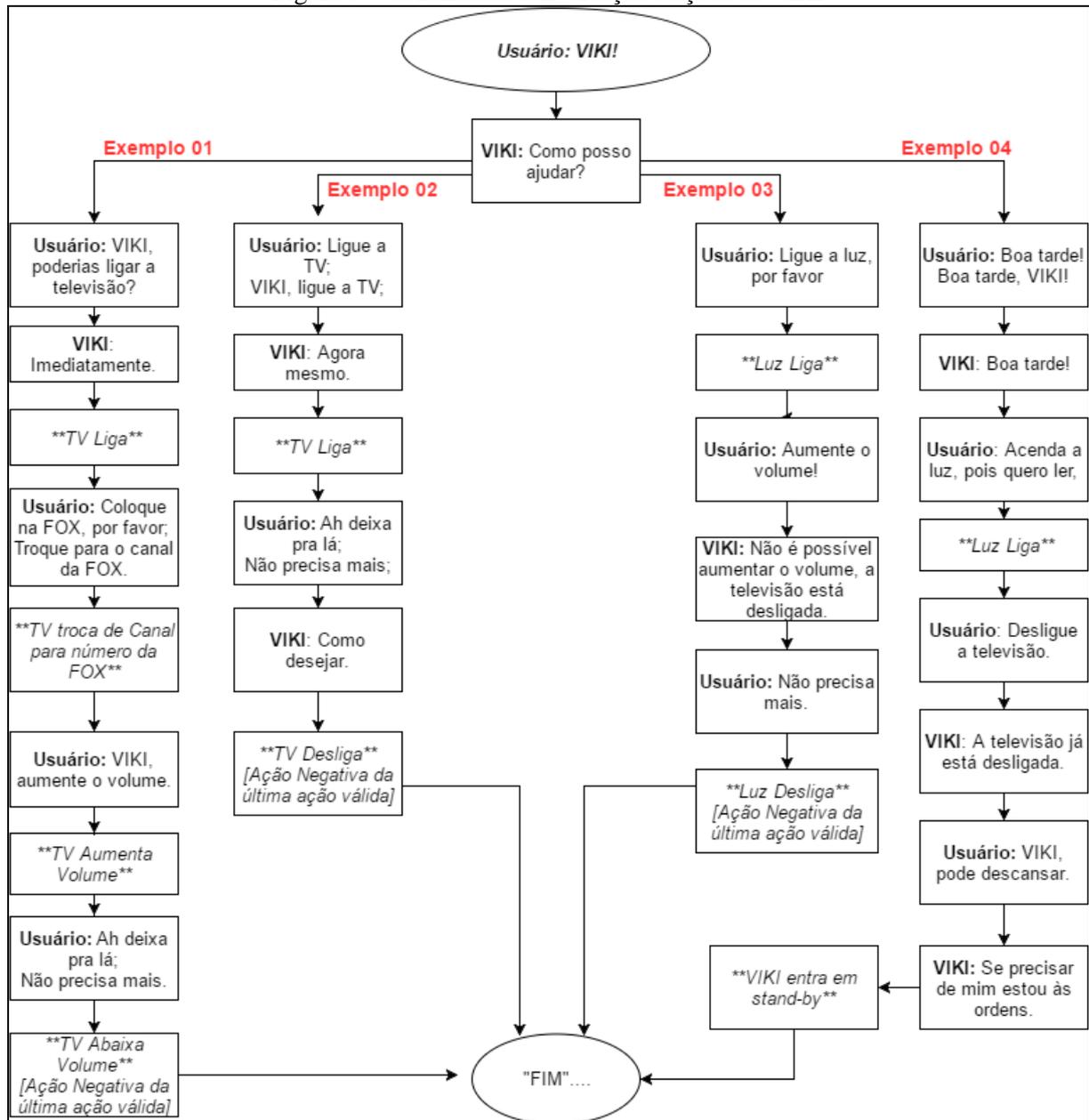
3.2 ESPECIFICAÇÃO

Nesta seção são apresentados os cenários que a aplicação deverá ser capaz de atender, bem como a lógica especificada para fazer o reconhecimento de voz. Também são apresentados os diagramas casos de uso e de pacotes.

3.2.1 Cenários de interação

A Figura 5 traz quatro exemplos de cenários que foram definidos, os dois primeiros com ações para controlar a televisão e os dois últimos para ligar / desligar a luz.

Figura 5 – Cenários de conversação e ações da VIKI



Fonte: elaborada pelo autor.

O primeiro cenário exemplificado tem quatro comandos. Inicia-se com o usuário pedindo para a VIKI ligar a televisão, em seguida há uma resposta falada da VIKI e, por fim, a ação efetiva de ligar a televisão. Logo após, o usuário solicita para colocar no canal da FOX, sendo apresentadas duas maneiras distintas, ‘Coloque na FOX, por favor’ e ‘Troque para o canal da FOX’. Essas duas sentenças são algumas das possíveis definidas para este comando. Adiante o usuário pede para a VIKI aumentar o volume e o volume da televisão é aumentado.

Imediatamente, o usuário pede o cancelamento do comando executado, sendo a ação correspondente abaixar o volume da televisão.

Nos outros três cenários são exemplificados comandos ilustrando as diferentes respostas e formas de interagir com a VIKI. Vale destacar que é efetuado o controle do estado dos dispositivos físicos para informar ao usuário quando não é possível executar alguma ação em virtude do estado atual do dispositivo. Exemplifica essa situação o terceiro cenário, onde, após a luz ser ligada, o usuário fala ‘Aumente o volume!’, ação que não é possível, como informado pela VIKI, porque a televisão está desligada. Já no quarto cenário tem-se o comando stand-by que faz com que a VIKI pare de reconhecer qualquer comando. Para tirá-la do stand-by, é necessário chamá-la diretamente.

A partir dos cenários especificados, pode-se definir as seguintes sequências de interação:

- a) para controlar os dispositivos físicos:
 - comando do usuário, resposta da VIKI (opcional), efetivação da ação,
 - comando do usuário, resposta da VIKI (opcional), efetivação da ação negando a última ação executada,
 - comando do usuário, resposta da VIKI informando não ser possível executar a ação em função do estado atual do dispositivo físico ou por não conseguir se comunicar com o dispositivo físico;
- b) para dialogar com a VIKI:
 - comando do usuário, resposta da VIKI, efetivação da ação para entrar ou sair do modo do stand-by,
 - comando do usuário (agradecer, pedir ajuda, conversar sobre assuntos casuais, entre outros), resposta da VIKI.

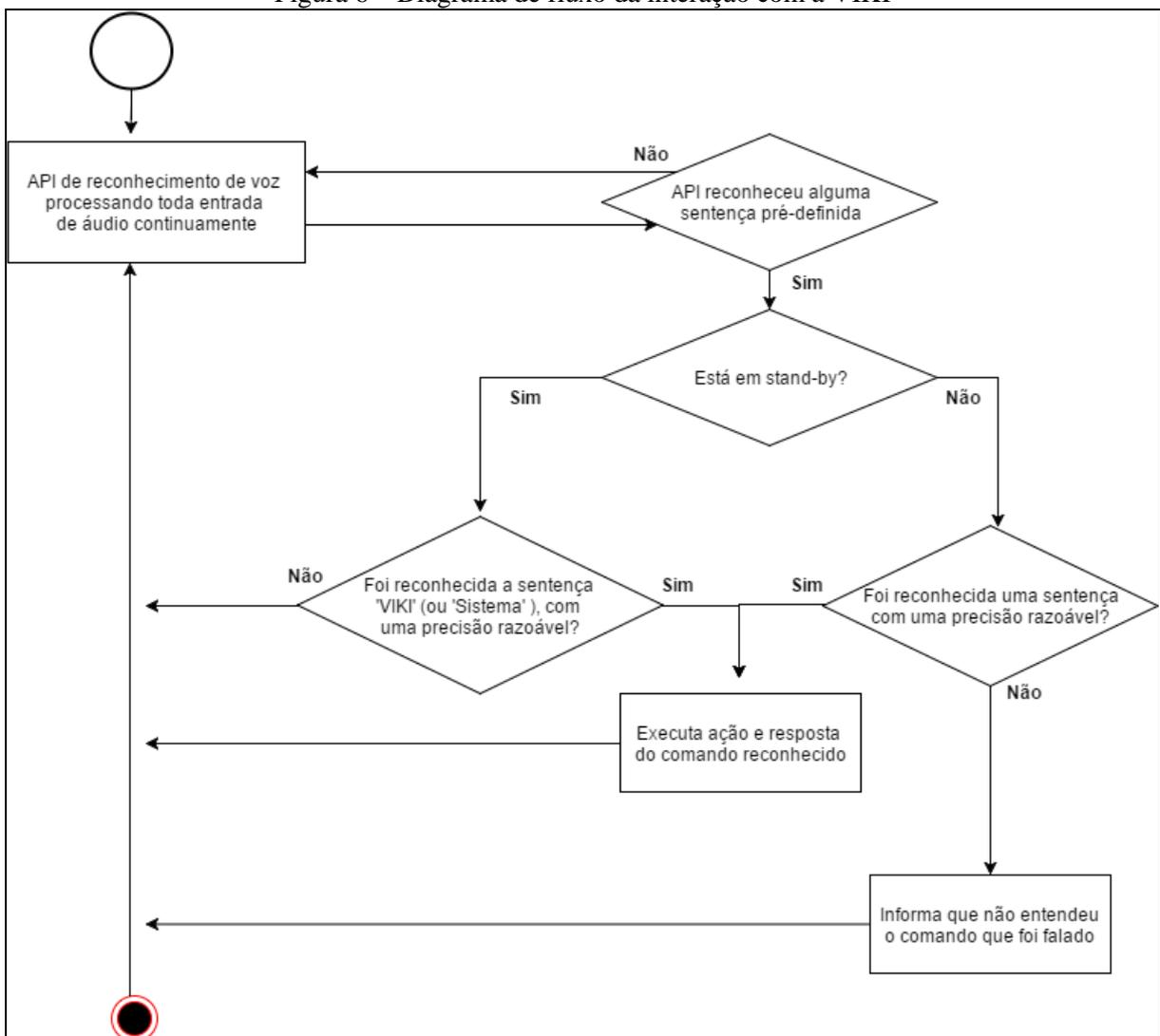
Observa-se que, como exemplificado com o comando para troca de canal, um comando pode possuir diferentes maneiras de ser verbalizado, de acordo com a definição feita. Os comandos da VIKI podem ser configurados em arquivos eXtensible Markup Language (XML), de forma que as sentenças reconhecidas e as respostas correspondentes possam ser facilmente modificadas e adaptadas se necessário, conforme descrito no Apêndice A.

3.2.2 Fluxo da interação com a VIKI

O reconhecimento dos comandos ocorre por meio da API Microsoft Speech Platform e funciona conforme ilustrado na Figura 6. Inicialmente, a API fica monitorando

ininterruptamente tudo que for recebido pela entrada de áudio. Assim que reconhecer alguma sentença pré-definida, ou seja, um comando, dispara um evento. A aplicação então verifica o estado da VIKI. Se estiver em stand-by, somente serão aceitas as sentenças ‘VIKI’ ou ‘Sistema’ com uma precisão razoável, ignorando qualquer outra sentença reconhecida. Caso não esteja em stand-by, é verificado se a sentença reconhecida possui uma precisão razoável. Em caso positivo, a VIKI dá uma resposta (opcional) e executa a ação correspondente ao comando reconhecido, conforme explicando na seção anterior. Ainda, se não está em stand-by, mas também não obteve uma precisão razoável no reconhecimento da sentença, a VIKI responde que não entendeu o que foi falado pelo usuário.

Figura 6 – Diagrama de fluxo da interação com a VIKI



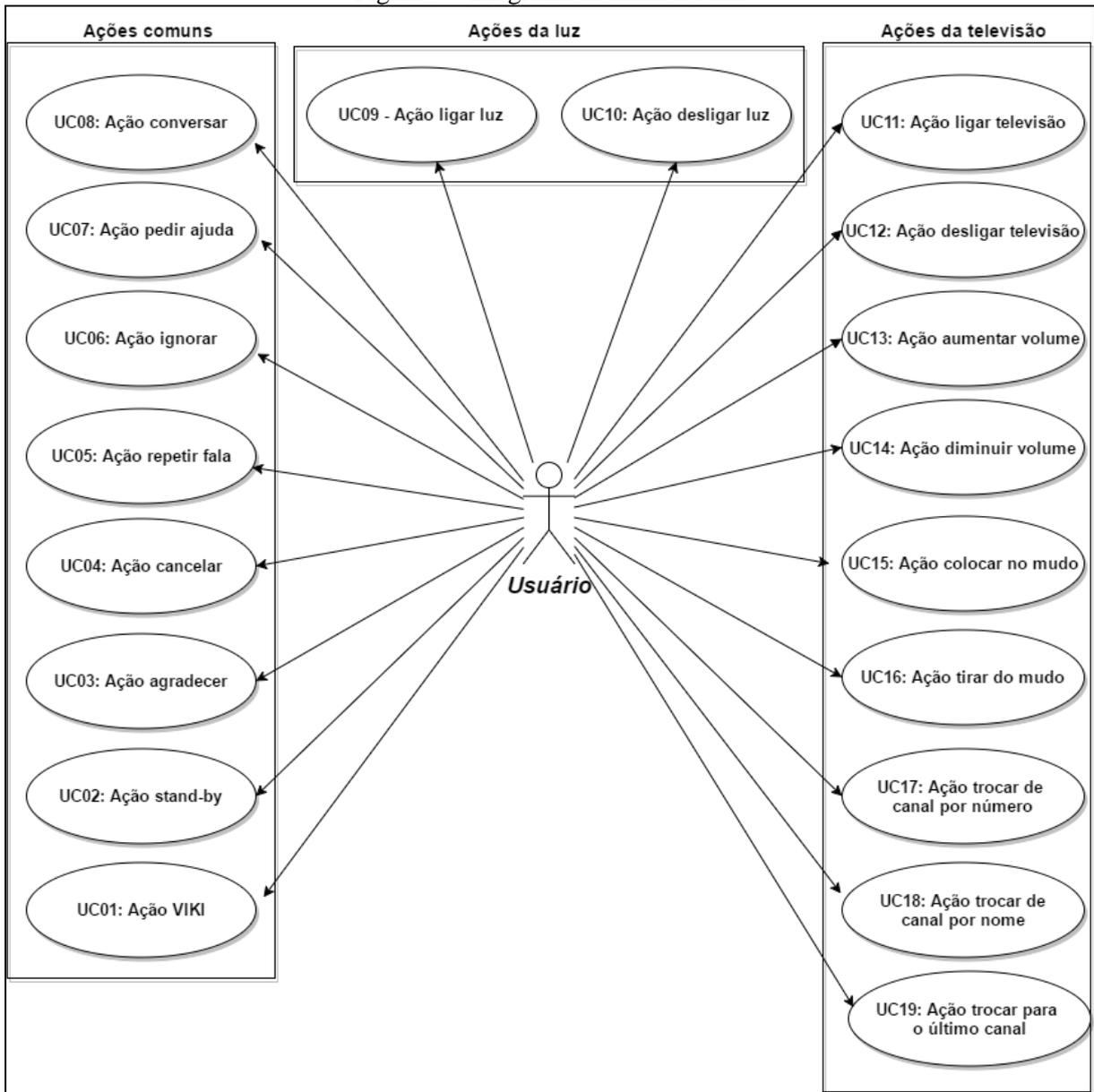
Fonte: elaborada pelo autor.

3.2.3 Diagramas de casos de uso e de pacotes

Os casos de uso que representam as funcionalidades da aplicação são exibidos na Figura 7. O detalhamento pode ser visto no Apêndice B. De acordo com a Figura 7, a

aplicação possui apenas o Usuário como ator, que é responsável por realizar todas as dezenove ações da aplicação. Cada ação resulta em uma interação diferente com a VIKI. Os casos de uso 01 a 08 refletem ações comuns da VIKI. Os casos de uso 09 e 10 são ações para ligar ou desligar a luz, enquanto os casos de uso 11 a 19 são relacionados às ações da televisão.

Figura 7 – Diagrama de casos de uso

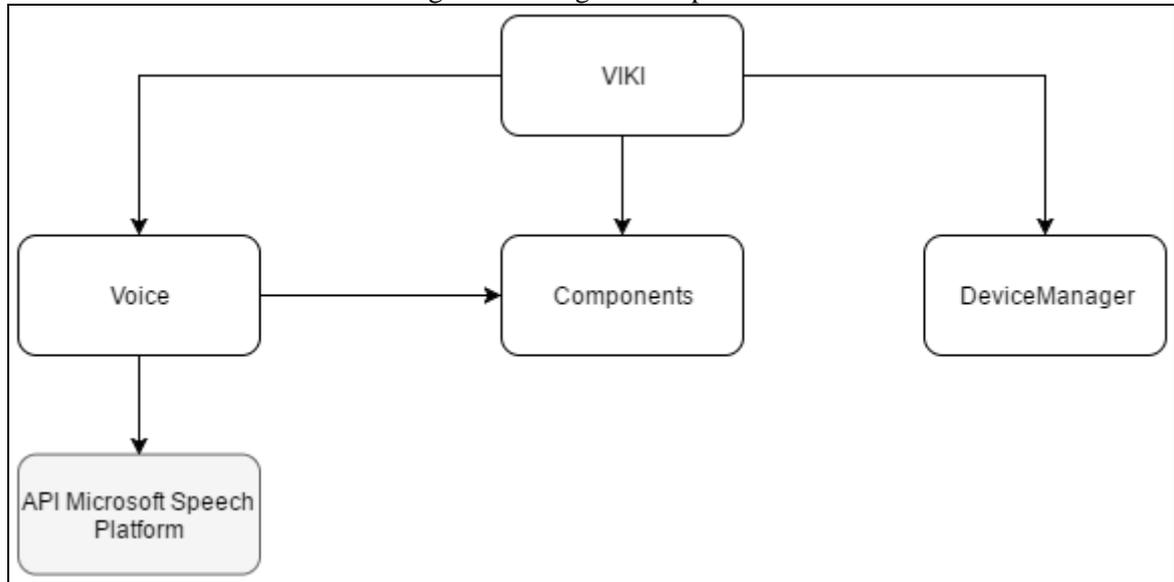


Fonte: elaborado pelo autor.

Os pacotes que compõem a VIKI, conforme ilustrado na Figura 8, são: (a) *VIKI*, responsável por toda a lógica e inteligência da aplicação, os demais pacotes são utilizados por ela; (b) *DeviceManager*, pacote responsável por tratar da comunicação com microcontroladores e manter o controle do estado de cada dispositivo; (c) *Components*, que mantém a especificação e a estrutura dos comandos existentes na VIKI e sua serialização por

XML; (d) `Voice`, pacote responsável por interagir com a API Microsoft Speech Platform e tratar os comandos de reconhecimento e síntese de voz, abstraindo a utilização da API e utilizando as estruturas da `VIKI`, definidas no pacote `Components`. A composição de cada pacote e suas principais classes estão descritas no Quadro 1.

Figura 8 – Diagrama de pacotes



Fonte: elaborada pelo autor.

Quadro 1 – Pacotes e classes da aplicação

pacote	classe	descrição
VIKI	MainWindow	Principal classe da VIKI, onde fica contida toda a lógica da aplicação. Recebe do pacote <code>Voice</code> os comandos reconhecidos, trata-os através das classes do pacote <code>Components</code> e efetua a comunicação com os dispositivos físicos com o auxílio do pacote <code>DeviceManager</code> . Sintetiza a voz da VIKI por meio do pacote <code>Voice</code> .
DeviceManager	Device	Classe abstrata que contém as propriedades em comum dos dispositivos físicos.
	Light	Classe que realiza o controle da luz por meio de comunicação com o microcontrolador, além de manter salvo os estados do dispositivo.
	TV	Classe que realiza o controle da televisão por meio de comunicação com o microcontrolador, além de manter salvo os estados do dispositivo.
Components	Command	Classe que contém as especificações de um comando, tais como: sentenças através das quais o comando pode ser reconhecido, respostas que devem ser dadas quando reconhecer o comando, se o comando deve ser sempre reconhecido ou não.
	Component	Classe abstrata que contém as propriedades em comum de cada componente, sendo que um componente representa uma categoria de ação

pacote	classe	descrição
		da VIKI.
	VIKIBase	Subclasse de <code>Component</code> que define todos os comandos essenciais e básicos de ações da VIKI, tais como: desfazer uma ação, entrar e sair do stand-by, agradecer, entre outros.
	VIKICHatty	Subclasse de <code>Component</code> que define todos os comandos de conversação casual com a VIKI.
	Physical	Subclasse de <code>Component</code> que define todos os comandos referentes à controle da luz e da televisão.
	Serialization	Classe responsável por transformar os componentes e comandos em XML e vice-versa.
Voice	Hear	Classe responsável por toda a lógica e tratamento do reconhecimento de voz. Mantém comunicação com a API Microsoft Speech Platform e trata os eventos e informações recebidas, tais como precisão, estado de stand-by e controle dos comandos temporários. Abstrai ainda mais toda parte de reconhecimento de voz, devolvendo para VIKI os comandos relevantes devidamente tratados por meio da classe <code>VoiceRecogEventArgs</code> .
	Speak	Classe responsável por todo e qualquer comando de síntese de voz. É possível enviar uma <code>string</code> a ser sintetizada, inclusive contendo controle de entonação se necessário.
	VoiceRecogEventArgs	Classe que contém as principais informações de um comando reconhecido, tais como: comando, frase e argumento, data e hora do reconhecimento.

Fonte: elaborado pelo autor.

3.3 CONTROLE DOS DISPOSITIVOS FÍSICOS

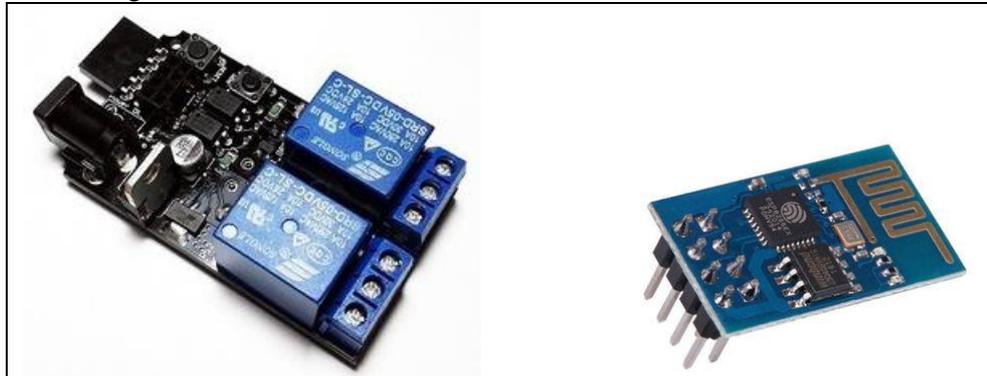
Nesta seção são descritas a definição e a montagem dos microcontroladores e placas, bem como as suas implementações para controle da luz e da televisão.

3.3.1 Definição e montagem dos microcontroladores e placas

Para o controle da luz foi utilizada uma placa AFSmart 031604 (AF ELETRÔNICA, 2016), ilustrada na Figura 9 (à esquerda), que contém dois relés embutidos, alimentação de 5V ou 9V, e entrada para conectar um microcontrolador ESP8266-01, mostrado na Figura 9 (à direita), que pode ser removido para gravação de um novo programa. O relé da placa está conectado ao fio de fase da luz. Assim, quando este relé for fechado, via programação do

microcontrolador, a luz será acesa, e quando for aberto, a luz será apagada. A esquemática da placa disponibilizada pelo fabricante pode ser encontrada no Anexo A.

Figura 9 – Placa AFSmart 031604 e microcontrolador ESP8266-01



Fonte: AF Eletrônica (2016) e Santos (2016).

Como a placa AFSmart 031604 é alimentada somente por corrente contínua de 5V ou 9V, foi necessário utilizar uma fonte adaptadora de tensão. Foi utilizada a fonte HLK-PM01, apresentada na Figura 10. A mesma é capaz de converter de 100V a 240V alternado para 5V de corrente contínua. Esta fonte também conta com proteção de sobrecarga e curto circuito.

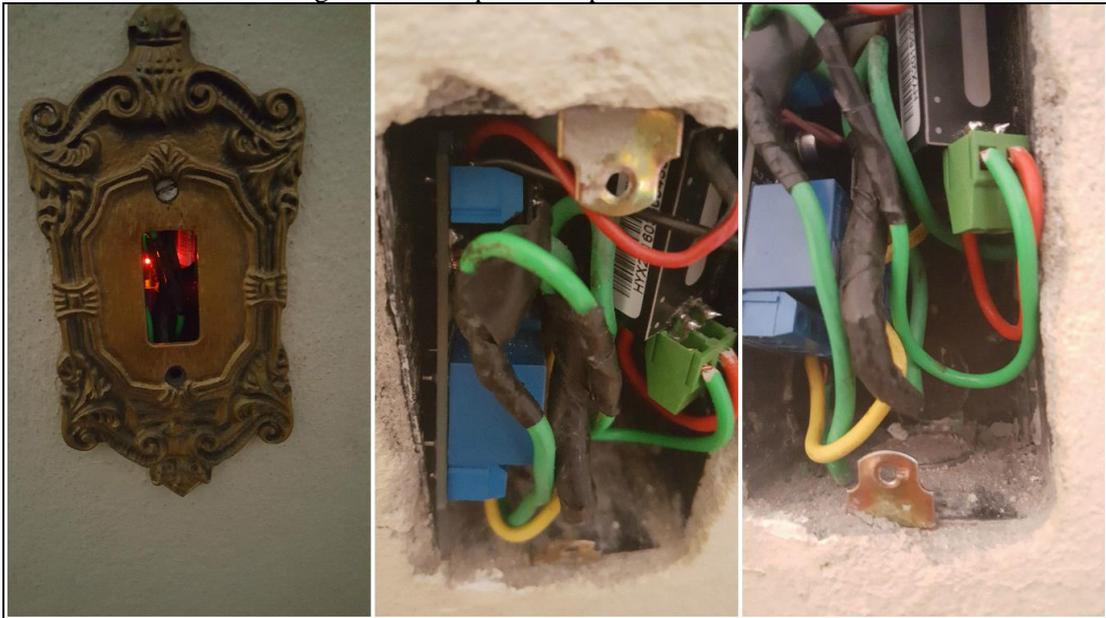
Figura 10 – Fonte HLK-PM01



Fonte: HI-Link (2016).

Observa-se que a placa, o microcontrolador e a fonte são pequenos o suficiente para caberem dentro de um interruptor de luz. Suas dimensões são (em milímetros), respectivamente, 63x37x20, 14.3x24.8 e 5x29.4x15.4. Na Figura 11 é possível visualizar a placa AFSmart, devidamente instalada em um interruptor comum de luz, com um ESP8266 programado embutido na própria placa e com a fonte ligada diretamente no 220V da casa.

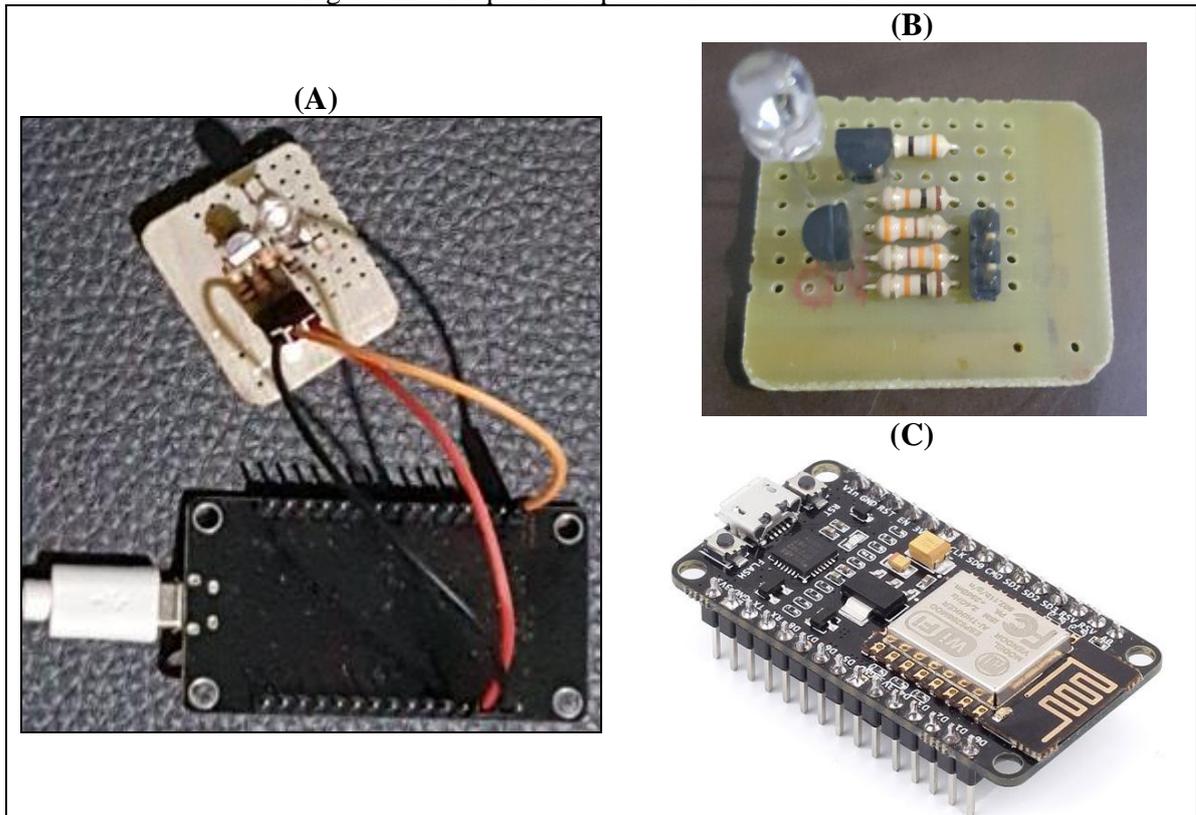
Figura 11 – Dispositivos para controle da luz



Fonte: elaborada pelo autor.

Quanto ao controle da televisão (Figura 12A), foi utilizado um pequeno circuito (Figura 12B) conectado ao microcontrolador ESP8266 NodeMCU (Figura 12C). O circuito possui um led emissor infravermelho, similar ao de um controle remoto, que passa por transistores para aumentar sua potência, ou seja, o alcance da emissão do sinal infravermelho, em torno de 4 metros. Assim, a aplicação é capaz de simular um controle remoto e enviar para a televisão os comandos desejados. O circuito possui ainda três conectores, são eles: GND, IO (em inglês Input and Output – IO) e VIN. Tanto o circuito como o microcontrolador tem alimentação de 5V. Como a alimentação do microcontrolador é via mini USB, podendo ser utilizado USB de um computador, bateria ou um carregador de celular comum com saída mini USB, pode ser usada para fornecer os 5V deste circuito.

Figura 12 – Dispositivos para controle da televisão

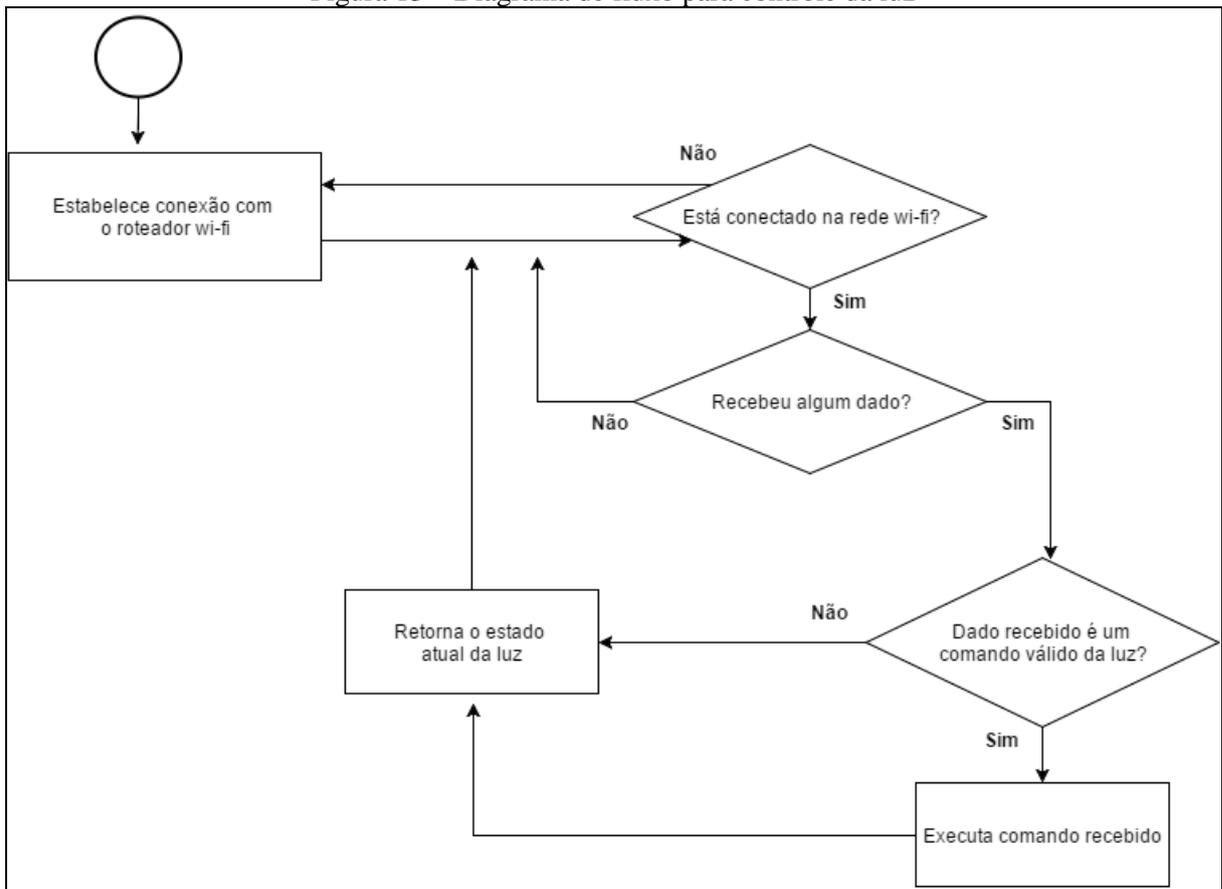


Fonte: elaborada pelo autor.

3.3.2 Implementação dos microcontroladores

O controle da luz é relativamente simples, pois, tendo as placas devidamente conectadas e instaladas, para ligar ou desligar a luz basta controlar um relé, que nada mais é que uma IO digital do ESP8266. Conforme mostra a Figura 13, o microcontrolador inicia tentando estabelecer comunicação com uma determinada rede wi-fi. Então verifica se está devidamente conectado. Em caso negativo e sempre que o microcontrolador não estiver conectado, tenta-se estabelecer essa comunicação. Em seguida fica em loop aguardando receber algum dado. Caso o dado recebido seja um comando válido (ligar ou desligar), executa este comando, ligando ou desligando o relé que controla a luz. Independentemente de ter recebido um comando válido, retorna o estado atual da luz, isto é, se o relé está ou não acionado. Essa informação é útil quando não se quer alterar o estado do relé e sim apenas ler o estado do mesmo para saber o estado da luz.

Figura 13 – Diagrama de fluxo para controle da luz



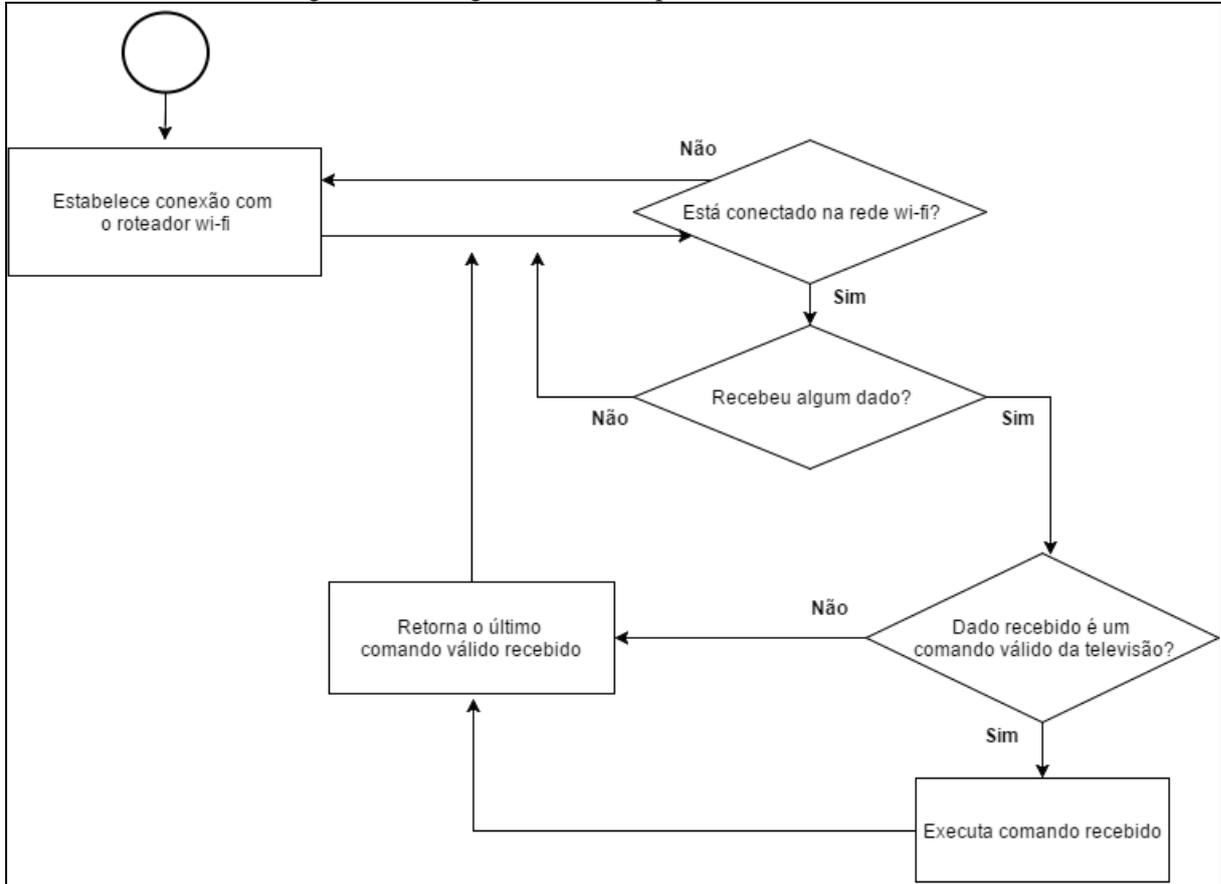
Fonte: elaborada pelo autor.

Quanto ao controle da televisão, foi utilizada a biblioteca `IRremote` que possui métodos para enviar comandos e protocolos de controle remoto de diferentes dispositivos. Cada botão de um controle remoto tem um código que é enviado para a televisão por meio do seu emissor infravermelho. Este código varia de fabricante para fabricante. Uma vez sabendo estes códigos, a biblioteca faz a conversão dos mesmos para os pulsos necessários emitidos pelo led emissor infravermelho.

Conforme apresentado na Figura 14, da mesma forma como explicado anteriormente, o microcontrolador inicia tentando estabelecer comunicação com uma determinada rede wi-fi. Então verifica se está devidamente conectado. Em caso negativo e sempre que não estiver conectado, o microcontrolador tenta estabelecer essa comunicação. Em seguida fica em loop aguardando receber algum dado. O dado recebido é processado, podendo ser um dos comandos válidos para televisão, os quais simulam um controle remoto, tais como botões de 0 a 9, aumentar canal, diminuir canal, aumentar volume, diminuir volume, ligar/desligar e mudo. Caso seja um comando válido, efetua o comando acionando o led emissor infravermelho por meio da biblioteca `IRremote`. Independentemente de ter recebido um comando válido, retorna o último comando executado pelo microcontrolador. Nota-se que

diferente da luz, não é possível enviar o estado da televisão, pois o estado da mesma é abstrato e não está vinculado a nenhum relé ou algo do gênero. Isto é, não é possível determinar de forma real o estado da televisão. Portanto, fica a cargo da VIKI, em nível de software, manter o histórico e o controle do estado da televisão.

Figura 14 – Diagrama de fluxo para controle a televisão



Fonte: elaborada pelo autor.

3.4 IMPLEMENTAÇÃO

A seguir são mostradas as técnicas e ferramentas utilizadas para o desenvolvimento deste trabalho. São apresentadas também as principais rotinas de reconhecimento e síntese de voz em conjunto com as tomadas de ações da aplicação.

3.4.1 Técnicas e ferramentas utilizadas

A aplicação VIKI foi desenvolvida utilizando a linguagem de programação C# através da IDE Visual Studio. A programação dos microcontroladores foi desenvolvida em C++ utilizando o Visual Studio. Já para realizar a compilação e a implantação do código nos microcontroladores ESP8266 NodeMCU e ESP8266-01 foi necessário instalar a IDE do Arduino (BANZI, 2016) em conjunto com a biblioteca do ESP8266. Por meio das bibliotecas

do ESP8266 e da `IRremote`, foi possível utilizar das funções do microcontrolador de forma fácil.

Para o reconhecimento e síntese de voz foi utilizada a biblioteca Microsoft Speech Platform, que é composta por duas principais classes, `Microsoft.Speech.Recognition` e `Microsoft.Speech.Synthesis`. A classe `Microsoft.Speech.Recognition`, responsável pelo reconhecimento de voz, permite definir uma entrada de áudio, sendo elas: por arquivos *wav*, pelo dispositivo de entrada de áudio padrão do sistema (usado pela VIKI) ou por um dispositivo de entrada de áudio específico. Para que seja efetuado o reconhecimento, é preciso definir as gramáticas, por meio da classe `Microsoft.Speech.Grammar`, contendo as sentenças que a aplicação deve ser capaz de reconhecer. Uma gramática deve conter um nome e ao menos uma sentença a ser reconhecida, as quais podem conter coringas ou palavras opcionais. Assim, para cada comando da VIKI foi definida uma gramática correspondente. Cita-se como exemplo o comando para agradecer a VIKI, para o qual foi definida a gramática `VIKI_Thanks` contendo as seguintes sentenças ‘Valeu’, ‘Obrigada’, ‘Obrigado’ e ‘Agradecido’, ou seja, formas como o usuário pode agradecer a VIKI por alguma ação efetuada. Uma vez definida a entrada de áudio e as gramáticas que a API deve reconhecer, deve-se iniciar o processo de reconhecimento de voz. A API monitora e processa todo áudio recebido pela entrada especificada em busca de um reconhecimento válido. Assim que o mesmo for identificado, a API dispara um evento informando qual gramática foi reconhecida, qual sentença e qual a precisão estimada, além de informar também outras sentenças que são similares e podem ter sido confundidas pelo algoritmo de reconhecimento.

As informações recebidas da API são tratadas pela VIKI. A API possui um filtro de precisão interno e não dispara eventos para reconhecimentos com precisão muito baixa. Dispõe também de uma precisão estimada quando reconhece alguma sentença para que a aplicação possa fazer uma verificação mais refinada se necessário. Assim, definiu-se que sentenças com menos de 50% de precisão estimada são ignoradas pela VIKI, de 50% a 60% é uma sentença que gera dúvida e acima de 60% considera-se que possui uma precisão razoável. A partir deste ponto, identifica-se qual comando está associado com a sentença reconhecida. Antes de executar efetivamente a ação correspondente ao comando, é verificado se é um comando temporário. Um comando temporário só é reconhecido em determinadas situações. Por exemplo, as sentenças usadas para agradecer só são reconhecidas e tratadas quando a VIKI determinar que executou uma ação que faça sentido ser agradecida. Portanto, a VIKI é que insere ou remove os comandos temporários da relação de comandos que podem

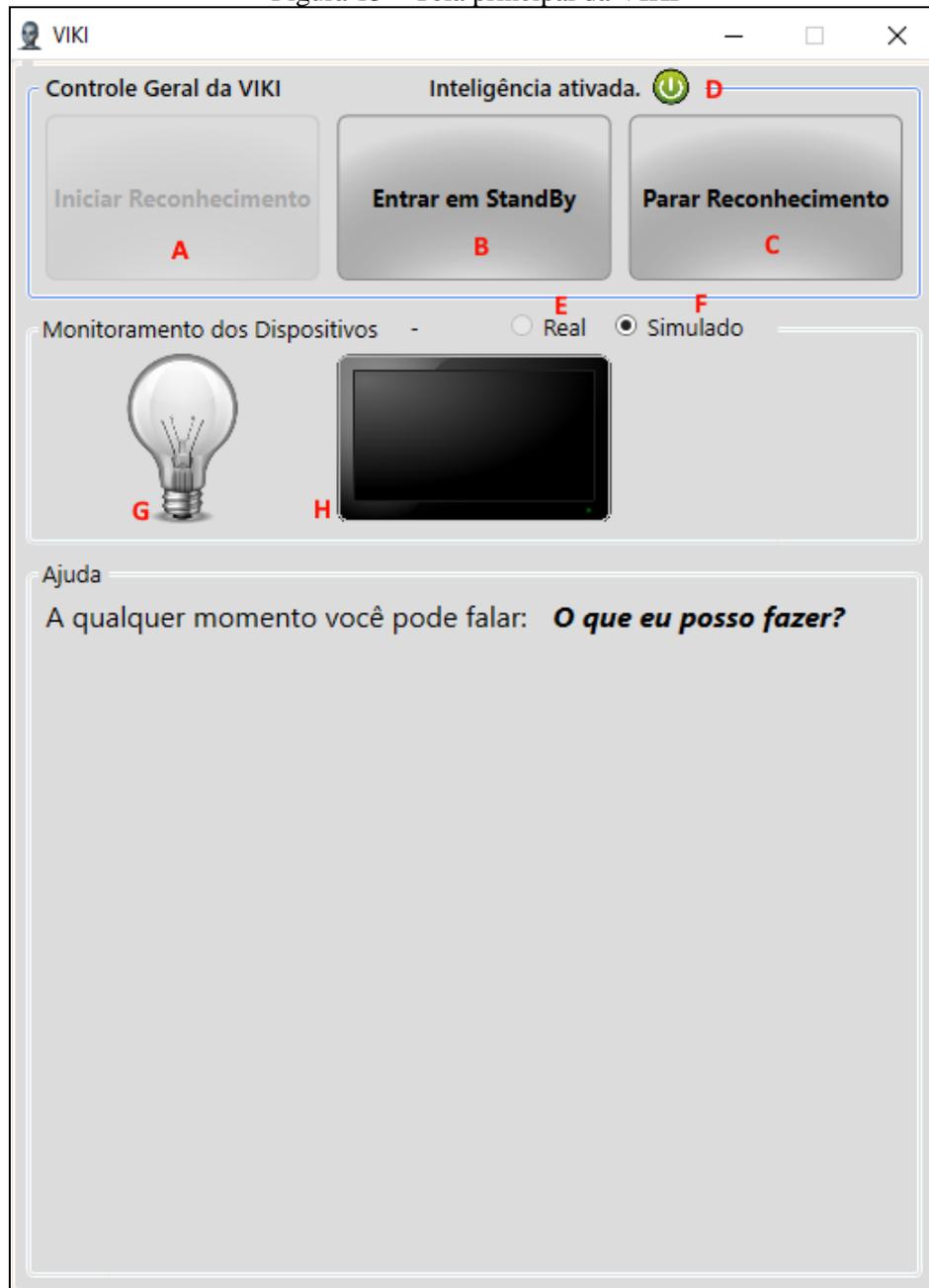
ser executados. Feito isso, a VIKI determina o que deve falar em resposta ao comando reconhecido e qual ação deve ser executada.

O uso da classe `Microsoft.Speech.Synthesis`, responsável pela síntese de voz, é extremamente simples e direto, basta apenas passar como parâmetro um literal contendo palavras gramaticalmente corretas da língua portuguesa, que o método `SpeakAsync` executa o áudio correspondente. Há possibilidade de adicionar entonação, isto é, deixar determinada palavra ou sílaba mais grave ou mais aguda, alterar seu volume e sua velocidade de fala. A partir disso e visando ter uma interação mais natural, a VIKI dá diferentes respostas para cada comando especificado e tem diferentes formas de agradecer, concordar, discordar ou negar as ações, todos com um controle de aleatoriedade.

3.4.2 Operacionalidade da implementação

A interface da VIKI, apresentada na Figura 15, conta com: (a) botão para iniciar o reconhecimento de voz, que faz com que a VIKI fique ouvindo e processando tudo que for falado; (b) botão para entrar em stand-by, que, quando pressionado, faz com que a VIKI permaneça ouvindo, mas só responda ao ser chamada pelo nome, saindo desse estado; (c) botão para parar o reconhecimento, nesta situação a VIKI não ouve nem processa mais nenhuma entrada de áudio; (d) descrição do estado atual (ativa, em stand-by ou desativada), seguido de uma imagem (verde, amarela ou vermelha, respectivamente); (e) e (f) opção para indicar o tipo de monitoramento: dispositivos reais, neste modo, a VIKI tentará enviar os comandos de luz e televisão para os microcontroladores; simulado, neste caso, a VIKI está trabalhando sem dispositivos, ou seja, apenas simulando o estado deles, portanto, nunca tentará enviar comandos para os microcontroladores; (g) uma imagem para indicar o estado da luz, podendo ser acesa ou apagada; (h) uma imagem para indicar o estado da televisão, podendo ser ligada ou desligada, bem como mostrar o nome do canal e o volume da televisão quando a mesma estiver ligada.

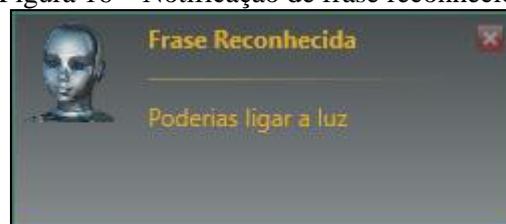
Figura 15 – Tela principal da VIKI



Fonte: elaborada pelo autor.

Ao ocorrer o reconhecimento de alguma frase, é apresentada uma pequena notificação da VIKI, conforme apresentado na Figura 16. Neste caso em questão, foi reconhecida a frase ‘Poderias ligar a luz’.

Figura 16 – Notificação de frase reconhecida



Fonte: elaborada pelo autor.

Assim que uma frase for reconhecida, além da VIKI notificar o usuário, ela executa o comando relacionado a esta frase. A notificação pode ser uma resposta direta a uma pergunta, uma indicação se a ação será executada com sucesso ou se não será executada e o porquê. Por exemplo, ao tentar trocar de canal com a televisão desligada, a aplicação falará que não é possível mudar de canal no momento, pois a televisão está desligada.

Destaca-se que a interface aqui descrita não é necessária para a utilização da VIKI. Ela é útil especialmente no modo simulado para que seja possível testar o reconhecimento e a lógica da aplicação, além de poder visualizar o estado simulado da luz e da televisão sem estar de fato conectado aos microcontroladores e demais dispositivos eletrônicos. Em qualquer momento, a interface pode ser minimizada ou nunca exibida e todas as operações continuam sendo realizadas somente por voz. Por fim, as instruções para configuração da VIKI encontram-se no Apêndice C.

3.5 ANÁLISE DOS RESULTADOS

Nesta seção são apresentados os testes realizados com a aplicação. Foram realizados dois experimentos. A seção 3.5.1 apresenta o experimento com pessoas em um ambiente real com luz e televisão sendo controladas por voz pela VIKI, detalhando a metodologia utilizada e a análise dos resultados. A seção 3.5.2 relata o experimento com pessoas tendo a VIKI instalada em seus próprios computadores, a fim de testar melhor a usabilidade, o reconhecimento e a síntese de voz. Nesse último caso, tem-se um ambiente simulado. Por fim, a seção 3.5.3 traz um comparativo entre os trabalhos correlatos e o trabalho desenvolvido.

3.5.1 Experimento 01: uso da VIKI em um ambiente real

O experimento foi realizado com quatro pessoas em ambiente real. Ocorreu durante o mês de outubro através de um teste individual com os voluntários. Foi feito em um ambiente com uma luz e uma televisão devidamente preparadas com os microcontroladores e conectadas à VIKI. A interação ocorreu exclusivamente por voz, sem nenhum contato com a interface gráfica da VIKI. Para realização dos testes foi disponibilizado um *earset bluetooth* de propriedade do autor. Inicialmente, foram fornecidas instruções prévias pelo autor deste trabalho, que estava acompanhando o experimento. Cada voluntário foi instruído sobre algumas das ações que a VIKI poderia fazer e o que ela poderia controlar.

Após o experimento, os voluntários receberam um questionário contendo perguntas de perfil e de usabilidade da aplicação que está disponível no Apêndice D. Os voluntários foram orientados a responder o questionário através da ferramenta Google Forms. O questionário

possui três questões objetivas de perfil, oito questões objetivas e cinco questões descritivas opcionais de usabilidade e feedback. As perguntas procuravam obter informações sobre o uso da aplicação e as impressões dos usuários.

O Quadro 2 mostra o perfil dos voluntários envolvidos no experimento 01. É possível observar que o perfil dos voluntários é diversificado, permitindo avaliar as impressões de pessoas com pouco contato (50%) ou profissionais da área de Tecnologia da Informação (TI) (50%). Os voluntários também tinham diferentes vozes por serem de idades e sexos diferentes.

Quadro 2 – Perfil dos voluntários do experimento 01

Qual o seu sexo?	50% masculino; 50% feminino
Qual a sua idade?	50% 18 a 30 anos; 50% 31 a 50 anos
Como você avalia seu nível de conhecimento com informática e tecnologia de forma geral?	25% apenas o básico 25% regular 50% profissional de TI

Fonte: elaborado pelo autor.

No Quadro 3 tem-se as questões objetivas ligadas ao uso e impressões a respeito da VIKI, onde é possível perceber que a aplicação desenvolvida obteve resultados satisfatórios quanto à usabilidade. Nesse quesito, o único ponto que teve um resultado intermediário foi se a VIKI entendeu tudo o que a pessoa disse. Isso se deve ao fato, conforme observado, que os voluntários, no início do experimento, tenderam a falar de forma diferente do que quando conversam com uma pessoa, isto é, falaram extremamente devagar ou até mesmo berrando. Outro ponto é que alguns dos voluntários tentaram emendar frases em uma conversação mais dinâmica, como no seguinte cenário: (1) voluntário: ‘Desligue a televisão’; (2) VIKI: ‘A televisão já está desligada’; (3) voluntário replica: ‘Então liga ela pra mim’. Neste exemplo, o comando não foi reconhecido e a VIKI ignorou a réplica do usuário, pois a mesma não faz parte da lista de sentenças reconhecidas. Quanto às impressões e opiniões, todos acharam interessante e recomendariam o uso da aplicação.

Quadro 3 – Avaliação de usabilidade do experimento 01

Você foi previamente explicado sobre o que é e qual a função da VIKI?	100% sim
De forma geral, você conseguiu entender e utilizar a VIKI?	100% sim
Você conseguiu entender todas as falas e respostas da VIKI?	100% - 5
A VIKI entendeu tudo que você quis dizer, dentro dos comandos especificados?	25% - 4 75% - 5
A VIKI travou ou apresentou problema alguma vez?	100% nunca
O quanto você gostou e achou interessante a VIKI?	100% - 5
Em um momento em que a VIKI esteja mais desenvolvida e disponível para o público, qual a chance de você recomendá-la para outra pessoa?	25% - 4 75% - 5

Fonte: elaborado pelo autor.

3.5.2 Experimento 02: uso da VIKI em um ambiente simulado

O experimento foi realizado com oito pessoas em seus próprios computadores com a VIKI no modo simulado. Neste modo, a VIKI dá feedback visual indicando o que reconheceu e qual ação foi tomada na tela do computador. O experimento ocorreu durante o mês de novembro através de testes remotos e independentes. Para realização dos testes, a VIKI foi disponibilizada em modo simulador para que cada usuário instalasse e testasse a vontade em seu próprio computador. Cada voluntário recebeu: (a) a aplicação para instalação; (b) uma breve instrução textual sobre o que é a VIKI e o que ela pode fazer, bem como algumas orientações sobre uso do microfone e do ambiente; (c) um link do Google Forms para o questionário de perfil e usabilidade. Neste experimento a conversação ocorreu exclusivamente por voz sem a presença ou conhecimento direto do autor.

O Quadro 4 mostra o perfil dos voluntários envolvidos no experimento 02, onde é possível observar que também é razoavelmente diversificado como no experimento 01, com exceção que a maioria dos usuários possuía um nível de conhecimento elevado em informática. O experimento também contou com voluntários com diferentes vozes, de sexos e idades diferentes.

Quadro 4 – Perfil dos voluntários do experimento 02

Qual o seu sexo?	62,5% masculino; 37,5% feminino
Qual a sua idade?	75% 18 a 30 anos; 25% 31 a 50 anos
Como você avalia seu nível de conhecimento com informática e tecnologia de forma geral?	12,5% apenas o básico 37,5% usuário avançado 50% Profissional de TI

Fonte: elaborado pelo autor.

No Quadro 5 estão as questões diretamente ligadas ao uso e impressões dos voluntários. A aplicação desenvolvida obteve resultados satisfatórios quanto à usabilidade e foram feitas críticas positivas sobre a aplicação, de forma que 100% dos voluntários gostaram e acharam a aplicação interessante, assim como recomendariam seu uso em um estado mais desenvolvido. Foram relatados problemas de travamento durante o uso da VIKI uma vez ou outra (12,5%), assim como foram relatados outros problemas (25%) relacionados ao uso do instalador. Inicialmente o instalador foi disponibilizado somente para ambientes 64 bits e a API Microsoft Speech Platform eventualmente não era instalada como esperado, de tal forma que a VIKI não funcionava devidamente por falta da API. Nesses casos, foi necessária uma orientação do autor para a reinstalação da mesma. Quanto à questão ‘A VIKI entendeu tudo que você quis dizer, dentro dos comandos especificados?’, foram informados alguns comandos que eventualmente não foram corretamente reconhecidos, como por exemplo:

‘ligue a televisão’ e ‘desligue a televisão’. Estas sentenças são muito parecidas existindo apenas uma pequena diferença, que se não forem faladas de forma bem declarada, podem ocasionar no reconhecimento incorreto do comando. Por fim, os quesitos interesse pela VIKI e possibilidade de recomendá-la foram altos, embora tenha algumas limitações e problemas a serem melhorados e tratados.

Quadro 5 – Avaliação de usabilidade do experimento 02

Você foi previamente explicado sobre o que é, e qual a função da VIKI?	37,5% sim 62,5% mais ou menos
De forma geral, você conseguiu entender e utilizar a VIKI?	100% sim
Você conseguiu entender todas as falas e respostas da VIKI?	100% - 5
A VIKI entendeu tudo que você quis dizer, dentro dos comandos especificados?	62,5% - 4 37,5% - 5
A VIKI travou ou apresentou problema alguma vez?	62,5% nunca 12,5% uma vez ou outra 25% outros
O quanto você gostou e achou interessante a VIKI?	25% - 4 75% - 5
Em um momento em que a VIKI esteja mais desenvolvida e disponível para o público, qual a chance de você recomendá-la para outra pessoa?	25% - 4 75% - 5

Fonte: elaborado pelo autor.

3.5.3 Comparativo entre os trabalhos correlatos

Para avaliar e comparar os trabalhos correlatos foram levadas em consideração as características mais importantes de cada um deles e do trabalho desenvolvido, como o reconhecimento de voz exclusivamente por comandos e sentenças específicas que é a forma como a VIKI trabalha, sendo também a forma em que produto Fibaro trabalha atualmente. Ainda assim, há produtos, como o Dragon Naturally Speaking e a Cortana, que permitem também um reconhecimento de voz livre, ou seja, que em determinado modo ou situação aceita e reconhece qualquer palavra ou texto falado pelo usuário, além das sentenças de comandos específicos, mas neste caso são produtos focados quase que exclusivamente no reconhecimento ou síntese de voz. Outra característica essencial é o idioma, o Fibaro e o Dragon Naturally Speaking não processam português brasileiro, somente a VIKI e a Cortana possuem esta característica. Indo além, o produto Fibaro tem como principal objetivo a automação residencial e embora possua reconhecimento de voz por comandos, a mesma é bem simples, além de possuir síntese de voz limitada, ou seja, sem capacidade de entonação ou com áudios pré-gravados para dar impressão de síntese de voz em determinadas situações.

A VIKI não aborda o reconhecimento de voz livre, mas somente o reconhecimento de voz por comandos e sentenças específicas, os quais julga-se serem suficientes para uma inteligência virtual dentro de um contexto específico e bem definido. Também conta com síntese de voz que permite a fala de qualquer palavra ou frase com controle de entonação no idioma português brasileiro. VIKI tem a característica, até então única nestes trabalhos correlatos, de ser uma inteligência virtual com capacidade básica de diálogo. Observa-se que esta capacidade permite que a VIKI possa questionar o usuário em caso de dúvidas e o mesmo possa respondê-la de acordo, dando continuidade ao processamento da aplicação. Além disso, o usuário também pode efetuar algumas ações de acordo com o contexto e histórico da conversa. Por exemplo, ao solicitar ‘Aumente mais’, a VIKI conseguirá entender implicitamente que o usuário deseja repetir a ação de aumentar o volume, se o último comando processado for referente a aumentar o volume. O Quadro 6 traz um comparativo entre os trabalhos correlatos descritos anteriormente e a VIKI.

Quadro 6 – Comparativo entre o trabalho desenvolvido e os trabalhos correlatos

características	VIKI	Dragon Naturally Speaking	Fibaro Home System	Cortana
reconhecimento de voz por comandos e sentenças específicas	X	X	X	X
reconhecimento de voz livre		X		X
reconhecimento e síntese do idioma português brasileiro	X			X
síntese de voz	X	limitada	limitada	X
capacidade de diálogo básico	X			
integração com Internet das Coisas	X		X	

Fonte: elaborado pelo autor.

4 CONCLUSÕES

Um dos principais objetivos da tecnologia é tornar a vida do homem mais fácil. As pessoas utilizam no dia-a-dia diversos equipamentos e programas feitos para melhorar e dar mais conforto para as suas vidas e atividades cotidianas. No entanto, novas tecnologias costumam apresentar uma grande curva de aprendizado e um árduo tempo de adaptação do público que irá utilizá-la, deixando inclusive uma boa parte de pessoas de fora de todas as vantagens que os sistemas e ferramentas possam ofertar. Visto isto, este trabalho buscou oferecer uma forma de utilizar uma aplicação integrada à automação residencial, permitindo controlar luz e televisão, de forma natural e sem grandes dificuldades para a maior parte dos usuários, por meio da interação com uma inteligência virtual utilizando somente a linguagem natural. A aplicação foi desenvolvida utilizando a linguagem C# no ambiente Visual Studio e a API Microsoft Speech Platform, em conjunto com a linguagem C++ e a IDE do Arduino especificamente para a programação do microcontrolador. Todas as ferramentas se mostraram funcionais, não apresentando problemas durante o desenvolvimento do trabalho.

No que diz respeito aos experimentos realizados, os resultados foram satisfatórios. Os objetivos de realizar o reconhecimento de voz, fornecer resposta ao usuário por meio de um sintetizador e acionar dispositivos físicos, como a luz e a televisão por meio de microcontroladores, foram alcançados. Também foram incluídos alguns comandos que não estavam previstos inicialmente, como conversação e frases casuais não relacionadas à automação, com objetivo de passar a impressão da VIKI ter uma personalidade própria e ser mais descontraída nos diálogos com o usuário.

As limitações encontradas foram principalmente relacionadas à síntese e ao reconhecimento de voz. A API Microsoft Speech Platform mostrou-se extremamente difícil para criar uma entonação de pergunta, fazendo com que a VIKI evite elaborar perguntas ou sempre precise deixar explícito na sentença que se trata de uma pergunta. Ainda, a API não reconhece adequadamente vozes muito agudas, em especial de crianças.

A partir dos resultados dos experimentos realizados, pode-se concluir que a aplicação desenvolvida, apesar das limitações, é uma opção interessante tanto para entretenimento quanto para ser a porta de entrada para pessoas que possuem dificuldades para interagir com novas tecnologias, visto que todos os voluntários que testaram a VIKI conseguiram utilizá-la e avaliaram bem o uso da aplicação. Por fim, de acordo também com a pesquisa, o trabalho desenvolvido pode servir como base para outras aplicações como controle e automação de aparelhos e ambientes para deficientes.

4.1 EXTENSÕES

Algumas possíveis extensões para este trabalho são:

- a) desenvolver ou usar um sintetizador capaz de elaborar perguntas, em especial dando ênfase e entonação para as frases sintetizadas;
- b) processar diálogos e contextos mais aprofundados, utilizando alguma estrutura de dados como árvores;
- c) adaptar para outros idiomas, como espanhol ou inglês, que já são suportados pela mesma API utilizada;
- d) ampliar a parte de Internet das Coisas permitindo que também possa acionar a luz fisicamente por meio de um interruptor que se comunique com o microcontrolador e com a aplicação.

REFERÊNCIAS

- AF ELETRÔNICA. **Esp8266 Wifi 802.11 b/g/n kit desenvolvimento**. [S.l.], [2016?]. Disponível em: <<http://www.afeletronica.com.br/pd-33eeb4-esp8266-wifi-802-11-b-g-n-kit-desenvolvimento.html?ct=&p=1&s=1>>. Acesso em 17 nov. 2016.
- ANDERSON, Tim. **Review: Nuance Dragon Naturally Speaking 13**. [S.l.], 2014. Disponível em: <<http://gadgets.itwriting.com/3158-review-nuance-dragon-naturally-speaking-13.html>>. Acesso em: 11 nov. 2016.
- ASAY, Matt. **The Internet Of Things will need millions of developers by 2020**. [S.l.], 2014. Disponível em: <<http://readwrite.com/2014/06/27/internet-of-things-developers-jobs-opportunity>>. Acesso em: 18 mar. 2016.
- BANZI, Massimo et al. Arduino Software. [S.l.], 2016. Disponível em: <<https://www.arduino.cc/en/Main/Software>>. Acesso em: 11 out. 2016.
- BAUMANN, Jim. **Voice recognition**. [S.l.], 1993. Disponível em: <http://www.hitl.washington.edu/research/knowledge_base/virtual-worlds/EVE/I.D.2.d.VoiceRecognition.html>. Acesso em: 31 mar. 2016.
- CÂMARA DE COMÉRCIO E INDÚSTRIA BRASIL-ALEMANHA DE PORTO ALEGRE. **Acesso à internet por dispositivos móveis cresce e empresas estão cada vez mais online**. [Porto Alegre], 2012. Disponível em: <<http://ahkemfoco.com.br/noticias/acesso-a-internet-por-dispositivos-moveis-cresce-e-empresas-estao-cada-vez-mais-online>>. Acesso em: 24 mar. 2016.
- COMPUTER SCIENCE DEGREE HUB. **Is artificial intelligence a growing field?** [S.l.], [2015]. Disponível em: <<http://www.computersciencedegreehub.com/faq/artificial-intelligence-growing-field/>>. Acesso em: 16 nov. 2016.
- DINO VISIBILIDADE ONLINE. **Empresas passam a adotar inteligência virtual como ferramenta para o atendimento de clientes**. [S.l.], 2016. Disponível em: <<http://exame.abril.com.br/negocios/dino/noticias/empresas-passam-a-adotar-inteligencia-virtual-como-ferramenta-para-o-atendimento-de-clientes.shtml>>. Acesso em: 31 mar. 2016.
- ESP8266 COMMUNITY FORUM. **Everything ESP8266**. [S.l.], 2016. Disponível em: <<http://www.esp8266.com>>. Acesso em: 11 out. 2016.
- FIBAR GROUP. **Fibaro**. [S.l.], 2016. Disponível em: <<http://www.fibaro.com/en>>. Acesso em: 18 mar. 2016.
- GOLL, Luís H. A tecnologia está realmente evoluindo e tornando as nossas vidas mais simples? **Profissionais TI: pra quem respira inspiração**. [S.l.], 2012. Disponível em: <<http://www.profissionaisiti.com.br/2012/10/a-tecnologia-esta-realmente-evoluindo-e-tornando-as-nossas-vidas-mais-simples>>. Acesso em: 18 mar. 2016.
- HI-LINK. **Ultra-compact power module HLK-PM01**. [S.l.], [2016?]. Disponível em: <http://www.hlktech.net/product_detail.php?ProId=54>. Acesso em: 17 nov. 2016.
- LEE, Ming C.; CHANG, Jia W.; HSIEH, Tung C. A grammar-based semantic similarity algorithm for natural language sentences. **The Scientific World Journal**, [S.l.], v. 2014, 2014. Não paginado. Disponível em: <<http://www.hindawi.com/journals/tswj/2014/437162/>>. Acesso em: 31 mar. 2016.

LIMA, Irving S. **NodeMCU (ESP8266) o módulo que desbanca o Arduino e facilitará a Internet das Coisas...** [S.l.], 2016. Disponível em: <<http://irving.com.br/esp8266/nodemcu-esp8266-o-modulo-que-desbanca-o-arduino-e-facilitara-a-internet-das-coisas>>. Acesso em: 15 nov. 2016.

LIRA, Davi. **Internet das Coisas**. [S.l.], 2013. Disponível em: <<http://porvir.org/wiki/internet-das-coisas>>. Acesso em: 31 mar. 2016.

MICROSOFT CORPORATION. **Cortana e seu aplicativo: ajudando as pessoas a realizar tarefas**. [S.l.], 2016a. Disponível em: <<https://dev.windows.com/pt-br/cortana>>. Acesso em 18 mar. 2016.

_____. **Microsoft speech platform**. [S.l.], 2016b. Disponível em: <<https://msdn.microsoft.com/en-us/library/jj127857.aspx>>. Acesso em: 30 mar. 2016.

MILLS, Michael. **Artificial intelligence in law: the state of play in 2015?** [S.l.], 2015. Disponível em: <<http://www.legaltechnology.com/latest-news/artificial-intelligence-in-law-the-state-of-play-in-2015/>>. Acesso em: 16 nov. 2016.

NUANCE COMMUNICATIONS. **Dragon speech recognition software**. [S.l.], 2016. Disponível em: <<http://www.nuance.com/dragon/index.htm>>. Acesso em: 24 mar. 2016.

SANTAELLA, Lucia et al. Desvelando a Internet das Coisas. **Revista GEMInIS**, São Carlos, v. 1, n. 2, p.19–32, 2013. Disponível em <<http://www.revistageminis.ufscar.br/index.php/geminis/article/view/141/pdf>>. Acesso em: 24 mar. 2016.

SANTOS, Nuno. **ESP8266**. [S.l.], 2016. Disponível em: <<https://techtutorialsx.wordpress.com/2016/02/14/esp8266/>>. Acesso em: 20 nov. 2016.

SAS INSTITUTE INC. **O que é a Internet das Coisas (IoT)?** [S.l.], 2016. Disponível em: <http://www.sas.com/pt_br/insights/big-data/internet-das-coisas.html>. Acesso em 31 mar. 2016.

SATO, Paula. O que é inteligência artificial? Onde ela é aplicada? **Revista Nova Escola**, [São Paulo], Jun. 2009. Disponível em: <<http://revistaescola.abril.com.br/ciencias/fundamentos/inteligencia-artificial-onde-ela-aplicada-476528.shtml>>. Acesso em: 03 abr. 2016.

VOICE INTERACTION DO BRASIL. **Síntese de voz: DIXI.SERVER**. [S.l.], 2016. Disponível em: <http://www.voiceinteraction.com.br/?page_id=367>. Acesso em: 31 mar. 2016.

APÊNDICE A - Descrição da estrutura do comando e do XML

Neste apêndice é descrita a estrutura dos comandos da VIKI e sua definição em XML. No Quadro 7 é possível visualizar a estrutura do comando `VIKI_Light_TurnOn`, que aceita algumas formas de ligar a luz utilizando palavras opcionais e alternativas de palavras obrigatórias, permitindo que o reconhecimento do comando pode ser feito a partir de várias sentenças. Por exemplo, o comando exige obrigatoriamente que a sentença inicie com ‘acenda’, ‘acender’, ‘ligue’ ou ‘ligar’, seguido opcionalmente do artigo ‘a’ e, por fim, a palavra ‘luz’. Aceita também ‘tens como’ ou ‘poderias’ seguido de ‘ligar a luz’. Este comando não possui argumentos.

Quadro 7 – Definição em XML do comando para ligar a luz

```
<Command>
  <CommandName>ligar luz</CommandName>
  <CommandValue>Light_TurnOn</CommandValue>
  <CommandExample>Ligue a luz</CommandExample>
  <CommandPhrases>
    <string>[acenda&acender&ligue&ligar] [a] luz</string>
    <string>[tens como&poderias] ligar a luz</string>
  </CommandPhrases>
  <CommandArguments />
  <VocativePrefix>true</VocativePrefix>
  <WildcardSuffix>true</WildcardSuffix>
  <IsBaseCommand>true</IsBaseCommand>
  <ResponsePhraseId>0</ResponsePhraseId>
  <ResponsePercentage>0</ResponsePercentage>
</Command>
```

Fonte: elaborado pelo autor.

Para permitir a especificação de frases alternativas foram definidas algumas regras. Uma sentença pode ser colocada entre colchetes, significando que esta sentença é opcional. Quando houver mais de uma sentença entre colchetes, as sentenças devem ser separadas por | ou &. As sentenças entre colchetes separadas por | são todas opcionais. Para sentenças entre colchetes separadas por & é obrigatório o uso de uma delas para que a frase seja reconhecida. Caso sejam usados | e & entre colchetes, por padrão, todas as sentenças são opcionais.

Conforme apresentado no Quadro 8, a definição do comando para trocar de canal por número é composta pela sentença ‘coloque no’, seguida, opcionalmente, de ‘canal’, e, como este comando tem argumentos, deverá ser informado obrigatoriamente um dos argumentos, que neste caso são os números de canais disponíveis. Quando este comando for reconhecido, a VIKI receberá devidamente tratado qual foi o comando e o argumento em questão, podendo então tomar a ação adequada.

Quadro 8 – Definição em XML do comando para trocar de canal por número

```

<Command>
  <CommandName>Coloque no Canal Número</CommandName>
  <CommandValue>TV_ChannelNumber</CommandValue>
  <CommandPhrases>
    <string>coloque no [canal]</string>
  </CommandPhrases>
  <CommandArguments>
    <string>10</string>
    <string>20</string>
    <string>30</string>
    <string>40</string>
    <string>41</string>
    <string>42</string>
    <string>43</string>
    <string>102</string>
    <string>120</string>
  </CommandArguments>
  <VocativePrefix>>true</VocativePrefix>
  <WildcardSuffix>>true</WildcardSuffix>
  <IsBaseCommand>>true</IsBaseCommand>
  <ResponsePhraseId>0</ResponsePhraseId>
  <ResponsePercentage>0</ResponsePercentage>
</Command>

```

Fonte: elaborado pelo autor.

No Quadro 9 são relacionadas as propriedades do XML, bem como apresentadas suas descrições detalhadas.

Quadro 9 – Propriedades da estrutura XML

propriedade	descrição
CommandArguments	argumentos do comando, propriedade opcional.
CommandName	nome do comando.
CommandValue	identificador interno do comando.
CommandPhrases	todas as frases pelas quais o comando pode ser reconhecido
CommandExample	frase a ser sintetizada pela VIKI quando o usuário pedir ajuda. Se deixada em branco, o exemplo e, conseqüentemente, o comando, nunca será informado pela VIKI.
IsBaseCommand	indica se o comando é um comando base, ou seja, se ele está sempre habilitado ou se só será habilitado e reconhecido quando a VIKI julgar necessário.
VocativePrefix	indica se o comando aceita e reconhece o vocativo 'VIKI' ou 'Sistema' antes das sentenças definidas em CommandPhrases.
WildcardSuffix	indica se o comando aceita um coringa após as sentenças compostas por CommandPhrases e CommandArguments.
ResponsePercentage	porcentagem de chance da VIKI retornar alguma resposta em conjunto com a execução do comando.
ResponsePhrases	frases a serem sintetizadas pela VIKI quando o comando for reconhecido, a seleção é aleatória.
ResponsePhraseId	respostas internas a serem sintetizadas pela VIKI quando o comando for reconhecido. Caso seja utilizado em conjunto com o ResponsePhrases, a VIKI sintetiza uma das respostas internas junto com uma das ResponsePhrases aleatórias.

Fonte: elaborado pelo autor.

O `ResponsePhraseId` remete a alguma das frases internas da VIKI, as quais foram agrupadas por tipo, conforme apresentado no Quadro 10.

Quadro 10 – Respostas internas da VIKI

Phrase	descrição	exemplos
0 - None	Nenhuma resposta	
1 - Accept	A VIKI retorna alguma frase de aceitação e subordinação.	Como desejar. Agora mesmo.
2 - Deny	A VIKI retorna alguma frase de que o pedido não é possível.	Não é possível. Ação indisponível.
3 - Dismissal	A VIKI retorna alguma frase indicando que foi dispensada e entrou em stand-by.	Qualquer coisa me chame. Ficarei no aguardo caso precise de mim.
4 - Salute	A VIKI retorna alguma frase de saudação.	Oi. Precisa de algo.
5 - PermissionDenied	A VIKI retorna alguma frase de permissão negada.	Acesso negado. Acesso não autorizado.
6 - PermissionGranted	A VIKI retorna alguma frase de permissão concedida.	Acesso permitido. Acesso concedido.
7 - Self	A VIKI retorna alguma frase respondendo a ter sido chamada.	As suas ordens. O que deseja.
8 - Processing	A VIKI retorna alguma frase de que está carregando ou processando algo.	Aguarde. Carregando.
9 - Agree	A VIKI retorna alguma frase de que concorda com algo.	Está certo. Tudo bem então.
10 - PositiveResponse	A VIKI retorna alguma frase contendo uma resposta positiva de agradecimento.	Não há de que. Disponha.
11 - Greetings	A VIKI retorna alguma frase de que ela foi recém ativada.	Aguardando ordens. Inteligência virtual ativada.
12 - GoodMorning	A VIKI retorna alguma frase de bom dia.	Bom dia. Uma manhã maravilhosa.
13 - GoodAfternoon	A VIKI retorna alguma frase de boa tarde.	Olá. Boa tarde
14 - GoodNight	A VIKI retorna alguma frase de boa noite.	Boa noite. Uma bela noite.

Fonte: elaborado pelo autor.

APÊNDICE B - Descrição dos casos de uso

Os casos de uso da aplicação estão descritos a seguir:

- a) UC01: Ação VIKI: descreve como o usuário chama a VIKI e a tira do modo stand-by;
- b) UC02: Ação stand-by: descreve como o usuário coloca a VKI em modo stand-by;
- c) UC03: Ação agradecer: descreve como o usuário agradece a VIKI por algum comando realizado;
- d) UC04: Ação cancelar: descreve como o usuário cancela a última ação realizada;
- e) UC05: Ação repetir fala: descreve como o usuário pede que a VIKI repita a última frase falada;
- f) UC06: Ação ignorar: descreve como o usuário informa a VIKI que não estava falando com a aplicação;
- g) UC07: Ação pedir ajuda: descreve como o usuário pode pedir ajuda da VIKI sobre quais comandos e frases ele pode falar no momento;
- h) UC08: Ação conversar: descreve como o usuário pode falar sobre conversas e assuntos casuais com a VIKI;
- i) UC09: Ação ligar luz: descreve como o usuário liga a luz;
- j) UC10: Ação desligar luz: descreve como o usuário desliga a luz;
- k) UC11: Ação ligar televisão: descreve como o usuário liga a televisão;
- l) UC12: Ação desligar televisão: descreve como o usuário desliga a televisão;
- m) UC13: Ação aumentar volume: descreve como o usuário aumenta o volume da televisão;
- n) UC14: Ação diminuir volume: descreve como o usuário diminui o volume da televisão;
- o) UC15: Ação colocar no mudo: descreve como o usuário coloca a televisão no mudo;
- p) UC16: Ação tirar do mudo: descreve como o usuário tira a televisão do mudo;
- q) UC17: Ação trocar de canal por número: descreve como o usuário troca de canal especificando o número dele;
- r) UC18: Ação trocar de canal por nome: descreve como o usuário troca de canal especificando o nome dele;
- s) UC19: Ação voltar para o último canal: descreve como o usuário volta para

o último canal.

A partir do Quadro 11 até o Quadro 29 são mostradas as condições, cenários e exceções de cada caso de uso.

Quadro 11 – UC01: Ação VIKI

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo ou em stand-by.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação VIKI. 2. A aplicação responde, saindo do modo stand-by, se for o caso.

Fonte: elaborado pelo autor.

Quadro 12 – UC02: Ação stand-by

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação stand-by. 2. A aplicação entra no modo stand-by.

Fonte: elaborado pelo autor.

Quadro 13 – UC03: Ação agradecer

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo. 2. O último comando executado é passível de agradecimento.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação agradecer. 2. A aplicação responde ao agradecimento.

Fonte: elaborado pelo autor.

Quadro 14 – UC04: Ação cancelar

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo. 2. O último comando executado possui uma ação negativa vinculada.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação cancelar. 2. A aplicação busca pelo último comando executado e executa a ação negativa correspondente.

Fonte: elaborado pelo autor.

Quadro 15 – UC05: Ação repetir fala

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo
Cenário principal	1. O usuário fala uma das sentenças referentes à ação repetir fala. 2. A aplicação sintetiza novamente a última frase falada.

Fonte: elaborado pelo autor.

Quadro 16 – UC06: Ação ignorar

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo. 2. O usuário fala algum comando ou frase não reconhecida pela VIKI ou de forma diferente, resultando em uma precisão baixa.
Cenário principal	1. O usuário fala alguma das sentenças referentes à ação ignorar. 2. A VIKI irá ignorar a última sentença que foi reconhecida com baixa precisão e apagará o contexto do último comando válido.

Fonte: elaborado pelo autor.

Quadro 17 – UC07: Ação pedir ajuda

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação pedir ajuda. 2. A aplicação busca os comandos que possuem sentenças de exemplos definidas. 3. A aplicação filtra os comandos que podem ser executados no contexto atual. 4. A aplicação seleciona aleatoriamente um dos comandos e sintetiza a sentença de exemplo especificada para o comando.
Exceção	1. Não ter nenhum comando com sentença de exemplo definida.

Fonte: elaborado pelo autor.

Quadro 18 – UC08: Ação conversar

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação conversar. 2. Baseado em qual frase de conversa casual o usuário falou, a aplicação responde aleatoriamente alguma de suas respostas vinculadas à frase em questão.

Fonte: elaborado pelo autor.

Quadro 19 – UC09: Ação ligar luz

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação ligar luz. 2. A aplicação liga a luz.
Exceção	1. A luz já está ligada.

Fonte: elaborado pelo autor.

Quadro 20 – UC10: Ação desligar luz

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação desligar luz. 2. A aplicação liga a luz.
Exceção	1. A luz não está ligada.

Fonte: elaborado pelo autor.

Quadro 21 – UC11: Ação ligar televisão

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação ligar televisão. 2. A aplicação liga a televisão.
Exceção	1. A televisão já está ligada.

Fonte: elaborado pelo autor.

Quadro 22 – UC12: Ação desligar televisão

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação desligar televisão. 2. A aplicação desliga a televisão.
Exceção	1. A televisão não está ligada.

Fonte: elaborado pelo autor.

Quadro 23 – UC13: Ação aumentar volume

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação aumentar volume. 2. A aplicação aumenta o volume da televisão.
Exceção	1. A televisão não está ligada.

Fonte: elaborado pelo autor.

Quadro 24 – UC14: Ação diminuir volume

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação diminuir volume. 2. A aplicação diminui o volume da televisão.
Exceção	1. A televisão não está ligada.

Fonte: elaborado pelo autor.

Quadro 25 – UC15: Ação colocar no mudo

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação colocar no mudo. 2. A aplicação coloca a televisão no mudo.
Exceção	1. A televisão não está ligada. 2. A televisão já está no mudo.

Fonte: elaborado pelo autor.

Quadro 26 – UC16: Ação tirar do mudo

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação tirar do mudo. 2. A aplicação tira a televisão do mudo.
Exceção	1. A televisão não está ligada. 2. A televisão não está no mudo.

Fonte: elaborado pelo autor.

Quadro 27 – UC17: Ação trocar de canal por número

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação trocar de canal por número. 2. A aplicação troca para o canal em questão.
Exceção	1. A televisão não está ligada.

Fonte: elaborado pelo autor.

Quadro 28 – UC18: Ação trocar de canal nome

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação trocar de canal por nome. 2. A aplicação busca internamente pelo número do canal referente ao nome que o usuário falou. 3. A aplicação troca para o canal em questão.
Exceção	1. A televisão não está ligada.

Fonte: elaborado pelo autor.

Quadro 29 – UC19: Ação trocar para o último canal

Ator	Usuário
Pré-condições	1. Estado da VIKI ativo.
Cenário principal	1. O usuário fala uma das sentenças referentes à ação trocar para o último canal. 2. A aplicação troca para o último canal válido da televisão.
Exceção	1. A televisão não ter trocado para nenhum canal ainda.

Fonte: elaborado pelo autor.

APÊNDICE C – Instruções de configuração da aplicação

Para a utilização da VIKI, deve-se primeiramente instalar o *framework* .NET versão 4.0 *full*. Em seguida, instala-se as APIs de reconhecimento e síntese de voz da Microsoft Speech Platform versão 11. Os arquivos a serem instalados são: (a) *SpeechPlatformRuntime* (x86 ou x64 de acordo com o sistema operacional), (b) *MMSpeech_SR_pt-BR_TELE*, (c) *MSSpeech_TTS_pt-BR_Heloisa*. Os requisitos das APIs utilizadas e da aplicação são: (a) processador de 1GHz ou superior, (b) 1GB RAM, (c) *framework* .NET versão 4.0 *full*. O sistema operacional necessário é o Windows 2003 server ou mais recente, que seja compatível com o *framework* .NET versão 4.0 *full*.

Tendo o ambiente devidamente instalado, basta iniciar a VIKI por meio do executável 'VIKI.exe'. Com a aplicação aberta, é possível alterar o IP dos microcontroladores da luz e da televisão, se necessário, pressionando a tecla 'F7', conforme apresentado na Figura 17.

Figura 17 – Configuração de IP na VIKI

Fonte: elaborada pelo autor.

Quanto à configuração dos microcontroladores, tanto do ESP8266 NodeMCU para controle da televisão quanto o ESP8266-01 para o controle da luz, deve-se configurar no seu código-fonte: (a) o nome da rede wi-fi com a qual ele deve se comunicar, (b) a senha desta rede wi-fi, (c) as informações de IP. É necessário que o IP definido nos códigos-fonte reflita os configurados na VIKI anteriormente. A configuração do microcontrolador da luz está ilustrada na Figura 18. A configuração para o microcontrolador da televisão é igual ao da luz, somente devendo ter um IP diferente para não haver conflito.

Figura 18 – Configuração de IP do microcontrolador da luz

```
const char* ssid = "MinhaRedeWiFi"; //Nome da rede wi-fi
const char* password = "minhasenha"; //Senha da rede wi-fi
IPAddress ip(192, 168, 25, 50); //IP que deseja na rede wi-fi
IPAddress gateway(192, 168, 25, 1); //IP do gateway da rede wi-fi
IPAddress subnet(255, 255, 255, 0); //Máscara de sub-rede da rede wi-fi
```

Fonte: elaborada pelo autor.

APÊNDICE D – Questionário de avaliação de usabilidade

Neste apêndice constam as questões de perfil (Quadro 30) e usabilidade (Quadro 31) que os usuários responderam no final dos experimentos.

Quadro 30 – Questões de perfil do usuário

PERFIL DO USUÁRIO
<p>Qual é o seu sexo?</p> <p><input type="checkbox"/> Masculino</p> <p><input type="checkbox"/> Feminino</p> <p>Qual a sua idade?</p> <p><input type="checkbox"/> Menos de 18 anos</p> <p><input type="checkbox"/> Entre 18 e 30 anos</p> <p><input type="checkbox"/> Entre 31 e 50 anos</p> <p><input type="checkbox"/> Mais de 50 anos</p> <p>Como você avalia seu nível de conhecimento com informática e tecnologia de forma geral?</p> <p><input type="checkbox"/> Não conheço nada</p> <p><input type="checkbox"/> Apenas o básico</p> <p><input type="checkbox"/> Regular</p> <p><input type="checkbox"/> Usuário avançado</p> <p><input type="checkbox"/> Profissional de TI</p>

Fonte: elaborado pelo autor.

Quadro 31 – Questões de usabilidade e opinião

AVALIAÇÃO DA USABILIDADE E OPINIÃO SOBRE A APLICAÇÃO
<p>Você foi previamente explicado sobre o que é e qual a função da VIKI?</p> <p><input type="checkbox"/> Sim <input type="checkbox"/> Mais ou menos <input type="checkbox"/> Não</p> <p>De forma geral, você conseguiu entender e utilizar a VIKI?</p> <p><input type="checkbox"/> Sim <input type="checkbox"/> Mais ou menos <input type="checkbox"/> Não</p> <p>Você conseguiu entender todas as falas e respostas da VIKI?</p> <p><input type="checkbox"/> 1 – Não entendi nada</p> <p><input type="checkbox"/> 2</p> <p><input type="checkbox"/> 3</p> <p><input type="checkbox"/> 4</p> <p><input type="checkbox"/> 5 – Entendi tudo</p> <p>A VIKI falou alguma frase ou resposta de forma estranha, robótica ou incompreensível? (Se sim, por favor diga quais)</p> <hr/> <p>A VIKI entendeu tudo que você quis dizer, dentro dos comandos especificados?</p> <p><input type="checkbox"/> 1 – Não entendeu nada</p> <p><input type="checkbox"/> 2</p> <p><input type="checkbox"/> 3</p> <p><input type="checkbox"/> 4</p> <p><input type="checkbox"/> 5 – Entendeu tudo</p>

A VIKI entendeu errado algum comando que você falou? (Se sim, por favor diga quais)

A VIKI travou ou apresentou problema alguma vez?

- Nunca Uma vez ou outra Várias vezes

Por favor dê sua opinião sobre quais aplicações e locais você acha que a VIKI poderia ser utilizada

	não serve	pode ser utilizada	muito útil
automação residencial (luzes, TV, ar condicionado, música, cortinas, entre outros)			
asilos e/ou hospitais para auxiliar pessoas com alguma deficiência física (controlar cama, TV, aparelhos, entre outros)			
embutida no carro para controle de funções, multimídia e GPS			
como assistente virtual para orientar e ajudar novos funcionários a usar algum sistema ou programa			
como assistente pessoal para auxiliar, monitorar e gerenciar e-mails, compromissos ou aplicativos			

O quanto você gostou e achou interessante a VIKI?

- 1 – Não gostei
 2
 3
 4
 5 – Gostei muito

Em um momento em que a VIKI esteja mais desenvolvida e disponível para o público, qual a chance de você recomendá-la para outra pessoa?

- 1 – Pouco provável
 2
 3
 4
 5 – Muito provável

Existe alguma forma de conversar ou palavras diferentes, que você ache interessante da VIKI reconhecer e conversar?

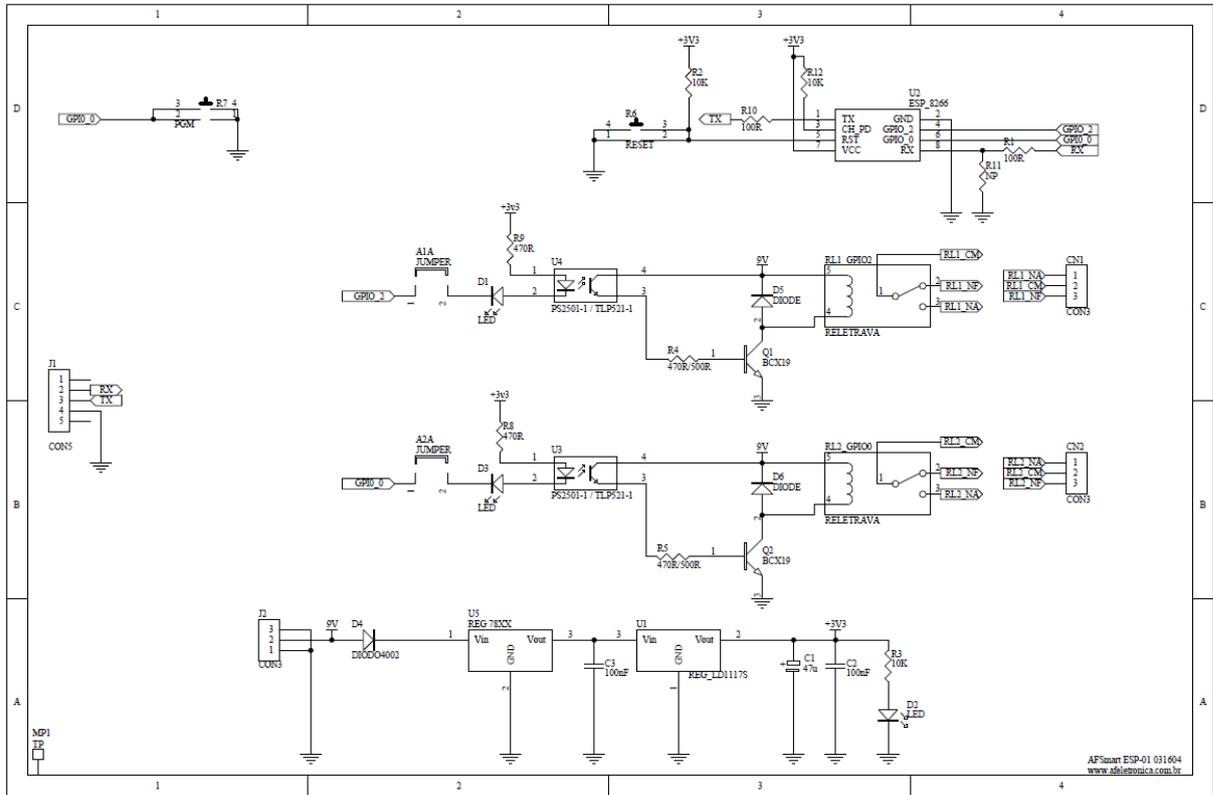
Por favor, diga em suas próprias palavras, quais aspectos da VIKI poderiam ser melhores:

Por favor, relate quais problemas ou dificuldades você teve ao utilizar a VIKI:

ANEXO A – Esquema elétrico da placa utilizada

Na Figura 19 é possível visualizar a esquemática da placa AFSmart 031604 utilizada no trabalho para controlar a luz.

Figura 19 – Esquemática da placa



Fonte: AF Eletrônica (2016).