Escalonamento de Operações em Dispositivos IoT para Assistentes Virtuais*

Felipe R. de C. Marques¹, Otávio H. Higa¹, Alfredo Goldman²

¹Escola de Artes Ciências e Humanidades - EACH/USP

²Instituto de Matemática e Estatística – IME/USP

{feliperanmarques, otaviohiga}@usp.br, gold@ime.usp.br

Abstract. *In order to interact with the real world, the A.D.A (Advanced Distributed Assistant) Project aims to develop an open source personal assistant able to interact with users from an ecosystem of devices through voice commands in brazilian portuguese. Here, we are developing the interface between the assistant and the IoT devices, such as sensors and actuators, so that the user can control their own equipments. Our goal is to research and integrate open standards of smart home device communication with the assistant and build a system for correctly ordering the execution of commands given by the user.*

Resumo. *Para interagir com o mundo real, o projeto A.D.A (Assistente Distribuída Avançada)¹ tem como objetivo desenvolver uma assistente pessoal de código aberto que consiga interagir com os usuários por meio de comandos de voz e executá-los em diferentes dispositivos IoT. Neste trabalho, estamos desenvolvendo a interface da assistente com os dispositivos embarcados, como sensores e atuadores, para que o usuário possa controlar seus próprios equipamentos. Nosso objetivo é pesquisar e integrar padrões de comunicação de dispositivos de smart home à assistente e construir um escalonador de comandos para conseguir executar o desejo do usuário corretamente.*

1. Introdução

Com o avanço de técnicas de aprendizado de máquina na última década e, em particular, com o sucesso de arquiteturas de aprendizado ponta-a-ponta, as assistentes virtuais comandadas por voz deixaram de ser um problema em aberto, para se transformar em um produto. Mas a despeito do sucesso comercial, ainda apresentam limitações importantes que não foram devidamente exploradas.

Podemos definir uma assistente comandada por voz como um agente artificial virtual que é capaz de reconhecer comandos verbais, extrair seu significado e, a partir de uma representação de conhecimento expressa por ações pré-programadas, executar a tarefa ou tarefas pretendidas pelo usuário [Mitchell et al. 1994]. Dessa forma, a assistente desempenha, na prática, o papel de uma camada intermediária simplificadora e compatível com o modelo mental do usuário, desobrigando-o de interagir com inúmeras interfaces distintas e com a constante mudança de contexto envolvida no processo. Tal papel pode

*Este projeto foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pelo grupo de extensão CodeLab uclab.xyz/site

¹uclab.xyz/ada

ser entendido também como uma manifestação da “transparência” propagada por Mark Weiser no contexto da computação ubíqua [Weiser 1993]. Isso resulta do fato que a assistente atua como uma ponte para a realização da tarefa solicitada, ocultando do usuário os mecanismos computacionais que a viabilizam.

Por outro lado, essa conveniência tem um custo de infraestrutura não desprezível, uma vez que o agente deve gerenciar e interagir adequadamente com os dispositivos que integram o ecossistema do usuário, o que demanda, entre outros aspectos, interoperabilidade de interfaces de programação de aplicações (APIs, do inglês *Application Program Interfaces*) e escalabilidade de processos. Todo o processamento que deve ocorrer para transformar comandos de voz em operações a serem executadas por dispositivos precisa ser modularizado e distribuído, para que seja possível lidar com a crescente demanda de usuários por sistemas desse tipo.

A possibilidade de conectar pessoas e objetos inteligentes através de uma infraestrutura comum – a Internet –, tem sido o foco de muitas pesquisas recentes na área da computação. Essa nova abordagem recebe o nome de Internet das Coisas (*Internet of Things* – IoT). A ideia principal se refere à uma rede unificada para interconectar pessoas a qualquer tipo de objeto, que pode ser algo real ou virtual, um dispositivo de hardware ou um *web service*.

Nesse contexto, está sendo pesquisado o núcleo de interpretação de programas, já que, com a expectativa do grande número de dispositivos que poderão se conectar à internet, assume-se que a escalabilidade será um grande desafio. Também estão sendo analisados alguns protocolos de comunicação, entendendo suas principais características para buscar a melhor execução de ações pelos dispositivos IoT dentro do âmbito de uma assistente virtual. Ao final, será criada uma prova de conceito, de código aberto, que irá demonstrar as capacidades da assistente e aplicações em domínios de interesse público e privado.

2. Metodologia

Antes de qualquer aprofundamento em relação à comunicação de uma assistente virtual, é importante estudar conceitos básicos de rede para o entendimento da relação entre assistente e dispositivos. Conhecer o *Internet Protocol* (IP) é essencial, pois este define a comunicação utilizada entre máquinas dentro de uma rede. Além disso, é necessário escolher entre os protocolos TCP e UDP, pois através destes que se realiza a transmissão de dados entre dois dispositivos distintos, com a diferença de que o primeiro possui checagem de erro de transmissão e procedimentos para evitar perda de dados, enquanto o segundo descarta essas funcionalidades para aumentar a velocidade de transmissão.

Pela importância de se receber todos os dados transmitidos para o entendimento integral do comando, foi escolhido o protocolo TCP no projeto. Assim como usado pela assistente virtual da Apple, a Siri [Caviglione 2015], e pelo Spotify [Kreitz and Niemela 2010], a confiabilidade na transmissão é fundamental, mesmo que a velocidade seja afetada.

Alguns exemplos de comandos que uma assistente virtual pode realizar atualmente são [Hoy 2018]:

- ler e enviar mensagens de texto e e-mails, fazer chamadas telefônicas;

- responder perguntas simples (“Que horas são? Qual a previsão do tempo?”);
- definir alarmes, cronômetros e eventos no calendário;
- definir lembretes, criar listas, fazer cálculos de matemática simples;
- controlar a reprodução de mídia;
- contar piadas e histórias;
- controlar dispositivos IoT como termostatos, luzes, alarmes e fechaduras.

O foco nesse projeto, na parte de IoT, foi no último item com o desenvolvimento de um software que permitiria que um componente fosse invocado pela assistente virtual. Ou seja, no caso do usuário falar para acender uma lâmpada, a assistente teria que verificar se a lâmpada foi registrada corretamente, se existe uma conexão entre assistente e dispositivo, realizar a ação de acender a lâmpada e devolver um *feedback* sobre o sucesso ou insucesso do comando.

Essas etapas de execução são organizadas pelo escalonador, que irá ordenar as operações para dispositivos conectados em uma instância A.D.A a partir de programas deduzidos pelo serviço de análise de intenções. O interpretador escolherá em que momento disparar as operações e reportará para a interface caso haja falha na execução. Isto é, para acender a lâmpada, o interpretador terá que certificar, antes, se existe uma conexão entre eles, para depois verificar se a função pedida pelo usuário existe e está ativa, e não o contrário.

3. Próximos passos

Relacionado à IoT, pretendemos construir um totem para a A.D.A, com microfone e alto falante a partir de um *Raspberry Pi*, tendo como exemplo o *Google Assistant* [Dempsey 2017]. Além disso, pretendemos concluir a interface de comandos para os dispositivos que a assistente pode acessar.

Relacionado à interface dos dispositivos com a A.D.A, pretendemos pesquisar sobre técnicas de implementação de interpretadores [Appel and Palsberg 2002] e modelos de computação distribuída relacionados à arquitetura de microsserviços [Reichert and Weber 2012], para encontrar o melhor algoritmo de escalonamento para os processos da assistente. Além disso, devemos investigar ferramentas para representação de linguagens de programação visuais.

Ao final do projeto, pretende-se construir uma assistente virtual capaz de entender comandos em português brasileiro, transcrevê-los para linguagem de máquina e executar a ação requisitada com uma infraestrutura modularizada e distribuída.

Referências

- Appel, A. W. and Palsberg, J. (2002). *Modern compiler implementation in Java*. Cambridge University Press.
- Cavaglione, L. (2015). A first look at traffic patterns of siri. *Transactions on Emerging Telecommunications Technologies*, 26(4):664–669.
- Dempsey, P. (2017). The teardown: Google home personal assistant. *Engineering Technology*, 12(3):80–81.
- Hoy, M. B. (2018). Alexa, siri, cortana, and more: An introduction to voice assistants. *Medical Reference Services Quarterly*, 37(1):81–88. PMID: 29327988.

- Kreitz, G. and Niemela, F. (2010). Spotify – large scale, low latency, p2p music-on-demand streaming. In *2010 IEEE Tenth International Conference on Peer-to-Peer Computing (P2P)*, pages 1–10.
- Mitchell, T. M., Caruana, R., Freitag, D., McDermott, J., and Zabowski, D. (1994). Experience with a learning personal assistant. *Communications of the ACM*, 37(7):80–91.
- Reichert, M. and Weber, B. (2012). *Enabling Flexibility in Process-Aware Information Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Weiser, M. (1993). Some computer science issues in ubiquitous computing. *Communications of the ACM*, 36(7):75–84.