

Universidade de São Paulo
Instituto de Matemática e Estatística
Bachalerado em Ciência da Computação

Leonardo Pereira Macedo

Desenvolvimento de um módulo
para a *game engine* Godot

São Paulo
3 de abril de 2017

Desenvolvimento de um módulo para a *game engine* Godot

Monografia final da disciplina
MAC0499 – Trabalho de Formatura Supervisionado

Supervisor: Prof. Dr. Marco Dimas Gubitoso

São Paulo
3 de abril de 2017

Agradecimentos

Resumo

A área de *games* evoluiu muito desde o início da década da 70, quando começaram a ser comercializados. As principais causas estão relacionadas aos avanços em diferentes áreas da Computação.

Com o passar do tempo, surgiram as *game engines: frameworks* voltados especificamente para a criação de jogos, visando a facilitar o desenvolvimento e/ou algumas de suas etapas.

Focaremos em uma *game engine* em particular, *Godot*. Por possuir código aberto, este *software* permite a extensão de suas funcionalidades através da criação de novos módulos.

Este projeto busca implementar um módulo de reconhecimento de voz para *Godot*, depois demonstrando a nova capacidade em um jogo simples desenvolvido na própria plataforma.

Palavras-chave: *software, game engine, Godot*, desenvolvimento de módulo, extensão de funcionalidade.

Abstract

Video games have evolved considerably since the beginning of the 70's, when they started to be commercialized. The main reasons are related to several advances in different fields of Computer Science.

Over time, *game engines* started appearing: *frameworks* designed specifically to assist on game creation, simplifying the process and/or some of its steps.

We will focus on a specific game engine, *Godot*. Since it is an open source project, it is possible to extend its functionalities by creating new modules.

This project's goal is to implement a speech recognition module for *Godot*, then showing the new feature in a simple game developed on the engine itself.

Keywords: software, game engine, *Godot*, module development, functionality extension.

Sumário

Agradecimentos	i
Resumo	iii
Abstract	v
Sumário	vii
Lista de Figuras	ix
Lista de Tabelas	xi
1 Introdução	1
1.1 Motivação e objetivo	1
1.2 Organização do trabalho	1
2 Bibliotecas para Reconhecimento de Voz	3
2.1 Considerações iniciais	3
Referências Bibliográficas	5

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

1.1 Motivação e objetivo

Hoje em dia, não há como negar que o mercado de *games* é um fenômeno mundial, gerando mais de US\$ 91 bilhões em 2016 (SuperData Research, 2016). Comparado aos primeiros jogos, comercializados no início da década de 1970 (Wikipedia, 2017a), a evolução em diversas áreas da computação permitiu grandes avanços nos jogos criados. Inclui-se nisso a evolução dos computadores por conta da *Lei de Moore* (Wikipedia, 2017b), permitindo processamento mais rápido; *games* em 3D e gráficos cada vez mais sofisticados e realistas devido à Computação Gráfica; e adversários sofisticados e de raciocínio rápido com a Inteligência Artificial.

Junto aos próprios jogos, as tecnologias usadas para desenvolvê-los também tiveram progressos. Em especial, temos as *game engines*, que podem ser descritas como “*frameworks* voltados especificamente para a criação de jogos” (Enger, 2013). Elas oferecem diversas ferramentas para acelerar o desenvolvimento de um jogo, como maior facilidade na manipulação gráfica e bibliotecas prontas para tratar colisões entre objetos. Além disso, como eficiência é um fator essencial para manter um bom valor de FPS (*Frames per Second*), as *engines* costumam ter sua base construída em linguagens rápidas e compiladas, como C e C++.

Focaremos em uma *game engine* em particular, *Godot* (Juan Linietsky, Ariel Manzur, 2017). O principal motivo de ter sido escolhida é por ser um *software* de código aberto, o que permite a qualquer pessoa baixar seu código fonte e fazer modificações. Em especial, a *engine* permite a criação de novos módulos para adicionar a ele novas funcionalidades.

Este trabalho visa a criar um novo módulo para *Godot*. Tal extensão adicionará funções simples de reconhecimento de voz, algo ainda inexistente no *software*. Feito isso, a nova funcionalidade será demonstrada em um jogo simples criado nessa *engine*.

1.2 Organização do trabalho

O capítulo 2 contém pesquisas e buscas por uma biblioteca de código aberto que faça reconhecimento de voz eficientemente. O capítulo 3 consiste em integrar a biblioteca encontrada ao *Godot*, expandindo suas funcionalidades. No capítulo 4, mostram-se os passos realizados para criar um jogo que demonstre a capacidade do novo módulo. O capítulo 5 apresenta as conclusões do trabalho.

Por fim, há uma parte subjetiva contendo a apreciação pessoal do TCC e uma descrição das matérias que mais ajudaram no desenvolvimento do projeto.

Capítulo 2

Bibliotecas para Reconhecimento de Voz

2.1 Considerações iniciais

Se o objetivo do projeto é criar um módulo de reconhecimento de voz, primeiramente é necessário chegar a uma biblioteca que implemente tal funcionalidade. Uma implementação do zero fugiria do tema deste trabalho, pois seria necessário aprender sobre reconhecimento de padrões voltado a sons, assunto relacionado a Inteligência Artificial.

A outra opção existente, e a que seguiremos, é procurar por uma biblioteca existente e aprender a manejá-la. Idealmente, gostaríamos que a biblioteca possuísse as seguintes características:

1. *Ter código aberto e licença permissiva:* Provavelmente o único aspecto essencial, uma vez que integraremos a biblioteca em uma *game engine*. A importância é ainda maior se levarmos em conta que jogos com fins comerciais podem ser produzidos em *Godot*.
2. *Ser implementada em C/C++:* Toda a base do *Godot* é construída em C++. A criação de módulos utiliza essa mesma linguagem, como será visto no capítulo 3.
3. *Reconhecer inglês:* Escolheremos inglês como uma linguagem obrigatória a se ter na biblioteca pelo caráter universal dessa linguagem.
4. *Ser multiplataforma:* Ao menos, oferecer suporte para sistemas Windows, MacOS e Linux.
5. *Implementar eficientemente reconhecimento de voz:* A biblioteca será usada em jogos, onde eficiência (velocidade) é uma questão central.
6. *Não ser pesada:* Embora tenha menos importância que os aspectos acima, não é desejável ter uma biblioteca que ocupe muito espaço.

Após uma pesquisa, que pode ser resumida pelo artigo em ([NeoSpeech, 2016](#)), cinco bibliotecas de reconhecimento de voz em geral se destacam por seu uso:

- *Kaldi* ([Kaldi, 2017](#)): É a biblioteca mais recente da lista, com seu código publicado em 2011. Escrita em C++.
- *CMUSphinx* ([CMUSphinx, 2015](#)): Desenvolvida pela *Carnegie Mellon University*, possui diversos pacotes para diferentes tarefas e aplicações. O pacote principal é escrito em Java. Existe também a variante *Pocketsphinx*: escrita em C e com suporte para sistemas embutidos, é uma opção mais interessante a ser considerada para este trabalho.

- *HTK* (HTK, 2016): Desenvolvida pela *Cambridge University Engineering Department*, HTK é uma sigla para *Hidden Markov Model Toolkit*. É escrita em C, com novas versões sendo lançadas consistentemente.
- *Simon* (Simon, 2017): Popular para Linux e escrita em C++, Simon utiliza *CMUSphinx*, *HTK* e *Julius* internamente. Não havia suporte para *MacOS* até 3 de abril de 2017.
- *Julius* (Julius, 2014): Desenvolvida pela *Interactive Speech Technology Consortium* e escrita em C. Infelizmente, o suporte para inglês é limitado e não pode ser usado para propósitos comerciais, o que nos força a descartar esta biblioteca como uma possível opção.

Das cinco bibliotecas descritas, as mais viáveis são as três primeiras (*Kaldi*, *CMUSphinx* e *HTK*). *Simon* está em seus primeiros passos para MacOS, por isso a relutância em seu uso.

Um artigo de 2014 comparou *Kaldi*, *CMUSphinx* e *HTK* em relação a precisão e tempo gasto (Gaida, 2014). *Kaldi* obteve resultados vastamente superiores; *CMUSphinx* obteve bons resultados em pouco tempo; *HTK* precisou de muito mais tempo e treino para conseguir resultados na ordem dos outros dois.

Referências Bibliográficas

- CMUSphinx(2015)** CMUSphinx. *About the CMUSphinx*. <http://cmusphinx.sourceforge.net/wiki/about>, Fevereiro 2015. Acessado: 2017-04-03. 3
- Enger(2013)** Michael Enger. *Game Engines: How do they work?* <https://www.giantbomb.com/profile/michaelenger/blog/game-engines-how-do-they-work/101529/>, Junho 2013. Acessado: 2017-04-03. 1
- Gaida(2014)** Christian Gaida. *Comparing Open-Source Speech Recognition Toolkits*. <http://suendermann.com/su/pdf/oasis2014.pdf>, 2014. Acessado: 2017-04-03. 4
- HTK(2016)** HTK. *What is HTK?* htk.eng.cam.ac.uk, 2016. Acessado: 2017-04-03. 4
- Juan Linietsky, Ariel Manzur(2017)** Juan Linietsky, Ariel Manzur. *Godot Engine*. <https://godotengine.org>, 2017. Acessado: 2017-04-03. 1
- Julius(2014)** Julius. *Open-Source Large Vocabulary CSR Engine Julius*. http://julius.osdn.jp/en_index.php, 2014. Acessado: 2017-04-03. 4
- Kaldi(2017)** Kaldi. *About the Kaldi project*. <http://kaldi-asr.org/doc/about.html>, Março 2017. Acessado: 2017-04-03. 3
- NeoSpeech(2016)** NeoSpeech. *Top 5 Open Source Speech Recognition Toolkits*. <http://blog.neospeech.com/top-5-open-source-speech-recognition-toolkits/>, 2016. Acessado: 2017-04-03. 3
- Simon(2017)** Simon. *About Simon*. <https://simon.kde.org/>, 2017. Acessado: 2017-04-03. 4
- SuperData Research(2016)** SuperData Research. *Worldwide game industry hits \$91 billion in revenues in 2016, with mobile the clear leader*. <https://venturebeat.com/2016/12/21/worldwide-game-industry-hits-91-billion-in-revenues-in-2016-with-mobile-the-clear-leader>, 2016. Acessado: 2017-04-02. 1
- Wikipedia(2017a)** Wikipedia. *The commercialization of video games*. https://en.wikipedia.org/wiki/History_of_video_games#The_commercialization_of_video_games, 2017a. Acessado: 2017-04-03. 1
- Wikipedia(2017b)** Wikipedia. *Moore's Law*. https://en.wikipedia.org/wiki/Moore%27s_law, 2017b. Acessado: 2017-04-03. 1