Software Requirements Specification

for

AI Meets the Classics

Version 1.0 approved

Prepared by Sidhraj Solanki

TuanAnh Nguyen

Rawan Alghofaili

University of Massachusetts-Boston

11/15/2018

Table of Contents

Revision History

Name	Date	Reason For Changes	Version

1. Introduction

1.1 Purpose

A chatbot that allows querying Topos Lab's Edison's backend. It is the first release of the product. Moreover, as the SRS document describes, the scope includes the chatbot, an ontology and a relationship model which specifies the queries to the chatbot. In addition to querying Edison backend there is another option to query knowledge base. The knowledge base has questions and answers related to the book.

1.2 Document Conventions

Every requirement statement will be marked according to its priority:

- HIGH: a core requirement and must be completed by the due date
- MED: a medium priority requirement. The absence of this requirement will not compromise the product's core function, but may improve its usability.
- LOW: a bonus requirement that may be completed if time permits.

1.3 Intended Audience and Reading Suggestions

This document is intended to be read by developers and project managers.

1.4 Product Scope

The product's scope is limited to a relationship model created by developing an ontology via Edison on the first chapter of Lord of the Rings, i.e Fellowship of the Ring. The scope also includes a chatbot which allows querying the relationship model. This product will be used as a learning tool to understand the context of Lord of the Rings, its characters and their interactions.

2. Overall Description

2.1 Product Perspective

Although the product is self-contained, it relies on Edison's relationship model as a backend for queries. Topos labs provides a python API for that purpose.

2.2 Product Functions

- Create a dictionary by defining lexemes from the first three chapters of Fellowship of the Ring.
- Create a set of questions that can be utilized to query the relationship model.

- Query the relationship model in the form of a chatbot.
- Feed the chapters to Edison, paragraph by paragraph.
- Create a knowledge base using popular questions based on Fellowship of the Ring

2.3 User Classes and Characteristics

This product is intended to be used by literature enthusiasts. The product's usability could be extend to include average users as well.

2.4 Operating Environment

The relationship model will operate within Edison's backend. The chatbot will query Edison using Topos lab's python API. The chat bot's interface will be built using python's TKinter GUI toolkit. To query the knowledge base, there will be a test processor, which will find a closest match with existing questions and return the response from knowledge base.

2.5 Design and Implementation Constraints

The developers need access to Topos Lab's python API to interact with Edison. The developer must have an account with Topos Labs. The knowledge base has only a small number of questions at present, but it can be extended.

2.6 Assumptions and Dependencies

We assume that the user using the dictionary has access to Edison. The user must have an account with Topos Labs which will enable him/her to connect to Edison's backend.

3. External Interface Requirements

3.1 User Interfaces

The user interface will be simple. There are two options to query the chatbot. For dropdown select if user wants to query Edison backend(Feeds) or the knowledge base (kb).
Anyone can open the chatbot application, select the option(kb or feeds) and type their queries. The chatbot will respond based on the knowledge base or the relationship model it has been trained with, depending on which option was selected.'
We will use standard design elements from TKinter to allow users to query the chatbot. There will be a form in the bottom of the screen. The upper portion of the screen will be reserved for a frame designed to display the user queries along with the chatbot response. This frame will allow users to scroll up and down to read the chat.

3.2 Hardware Interfaces

There is no designated hardware so there is no hardware interfaces.

3.3 Software Interfaces

The chatbot provide two options for users to query. One will be to query the knowledge base. The queries from the chat bot will be matched with closest possible query in the question knowledge base using Jaccard similarity, and the corresponding response will be generated. The query knowledge base will be a simple JSON dictionary.
Another option will be to query Edison backend. The query will be broken down into terms, and these terms will be sent to Edison backend and all the paragraphs containing those terms will be returned. The paragraph containing highest score will be displayed. The communications are handled by underlying operating system.

4. System Features

4.1 Dictionary for Lord of the Rings

4.1.1 Description and Priority

- **Priority**: HIGH
- Develop a dictionary for Lord of the Rings by defining the lexemes from the first three Chapters.
- Feed the complete book to Edison paragraph by paragraph.

4.1.2 Functional Requirements

REQ-1: When inputting text into Edison, it should define all the lexemes present.

4.2 Relationship model for Lord of the Ring's lexemes

4.2.1 Description and Priority

- **Priority:** HIGH
- Produce the relationship model for Lord of the Rings

4.2.2 Functional Requirements

REQ-1: Edison should build a relationship model, structured using the Lexemes and their definitions.

4.3 Chatbot to query the relationship model

4.3.1 Description and Priority

- **Priority:** HIGH
- Users should be able to ask the chatbot a pre-specified set of questions about Lord of the Rings.

4.3.2 Stimulus/Response Sequences

The user will start the chatbot application. When the user types text into the text box, and then presses Enter, the chatbot will respond.

4.3.3 Functional Requirements

REQ-1: The user will open the chatbot application, the chatbot will be displayed.

REQ-2: The user will select the option to query the chatbot. If the user wants to query the knowledge base, then select KB. If the user want to query Edison, select Feeds.

REQ-3: The user will type text into the input text box, the text will be displayed as the user types.

REQ-4: The user will press the enter button, and that would initiate parsing the user question.

REQ-5: The chatbot parser will parse the user question into tokens.

REQ-6: If the user selected KB, the chatbot will match the tokens with the question knowledge base. The question with the largest number of matches with the token list will be displayed.

REQ-7: Even if the tokens have no matches in the question knowledge base, the closest possible question will be found(based on the jaccard distance between the two) and the corresponding answer will be displayed.

REQ-8: If user selected Feeds, then the user query will be broken into terms and passed as json query to Edison. All the paragraphs will be matched against the query terms and the paragraph with highest score is returned..

REQ-9: The chatbot will display the query result.

5. Other Nonfunctional Requirements

5.1 Performance Requirements

The performance requirements are as follows:
a). The tool Edison should identify any keyword (character, race, geolocations) that appears in the text to be analysed. For that the ontology should be complete and should be able to identify any lexemes from the first three chapters of Lord of the Ring.
b). Given a query to the chatbot it should return a answer.

5.2 Security Requirements

The chatbot should not save a log of its conversations with the user, to preserve the user's privacy. Nor will the chatbot rely on data stored in the user's computer other than the username and password to login to Tobos Labs; the question knowledge base; and any information extracted from the relationship model explicitly.

5.3 Software Quality Attributes

The chatbot UI should be easy to use. It should not have any complicated design elements. It should contain elements found in standard chat software only.

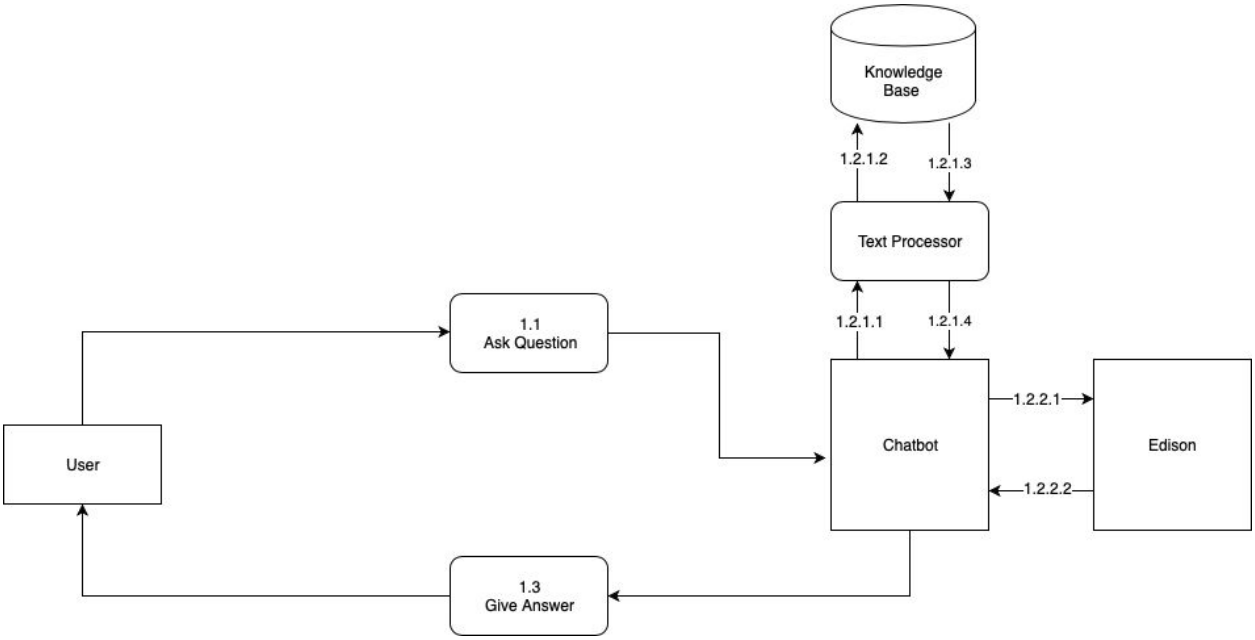
5.4 Business Rules

This chatbot is created only for Lord of The Rings, part 1. So it may not be useful to use it for queries which are related to any text other than it.

Appendix A: Analysis Models

Class diagram

State Machine Diagram



Data Flow Diagram