**Software Requirements Specification**

**for**

**Syntactical Parsing Of Marathi Text**

**Prepared by**

# Pranav Pradip Ranaware 19121056

# Kalyani Mahadev Kadam 19121017

# Vaishnavi Nilkamal Swami 20221082

**GCOEARA30.09.2022**

Copyright © 1999 by Karl E. Wiegers. Permission is granted to use, modify, and distribute this document.

Software Requirements Specification for < Syntactical Parsing Of Marathi Text > **Page 2**

**Table of Contents**

**Table of Contents ii**

**Revision History ii**

1. **Introduction 1**
   1. Purpose 1
   2. Document Conventions 1
   3. Intended Audience and Reading Suggestions 1
   4. Product Scope 1
   5. References 1
2. **Overall Description 2**
   1. Product Perspective 2
   2. Product Functions 2
   3. User Classes and Characteristics 2
   4. Operating Environment 2
   5. Design and Implementation Constraints 2
   6. User Documentation 2
   7. Assumptions and Dependencies 3
3. **External Interface Requirements 3**
   1. User Interfaces 3
   2. Hardware Interfaces 3
   3. Software Interfaces 3
   4. Communications Interfaces 3
4. **System Features 4**
   1. System Feature 1 4
   2. System Feature 2 (and so on) 4
5. **Other Nonfunctional Requirements 4**
   1. Performance Requirements 4
   2. Safety Requirements 5
   3. Security Requirements 5
   4. Software Quality Attributes 5
   5. Business Rules 5
6. **Other Requirements 5**

**Appendix A: Glossary 5**

**Appendix B: Analysis Models 5**

**Appendix C: To Be Determined List 6**

# Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Date** | **Reason For Changes** | **Version** |
|  |  |  |  |
|  |  |  |  |

**Software Requirements Specification for <** Syntactical Parsing Of Marathi Text **> Page 3**

# Introduction

## Purpose

We are working on to develop parser for marathi language, as there is no parsing technique has been developed for marathi language.

A very little attempt has been made to develop a syntactical parser on South Indian

languages including Kannada. Kannada is a spoken as well as written language in

Karnataka state and it is basically very agglutinative in nature. We made an effort

to develop a systematic syntax analyzer to parse all types of Kannada texts, and the

same has been presented in the paper

A very little attempt has been made to develop a syntactical parser on South Indian

languages including Kannada. Kannada is a spoken as well as written language in

Karnataka state and it is basically very agglutinative in nature. We made an effort

to develop a systematic syntax analyzer to parse all types of Kannada texts, and the

same has been presented in the paper

Introduction

A very little attempt has been made to develop a syntactical parser on South Indian

languages including Kannada. Kannada is a spoken as well as written language in

Karnataka state and it is basically very agglutinative in nature. We made an effort

to develop a systematic syntax analyzer to parse all types of Kannada texts, and the

same has been presented in the paper

## Document Conventions

Font: Times new roman

## Intended Audience and Reading Suggestions

Marathi is a spoken as well as written language in Maharashtra state. The audience for this project are who have knowledge of marathi language.

## Product Scope

* To analyze the sequence of tokens to determine their grammatical structure.
* To Compute the parse tree corresponding to the input.
* To separates a series of text based on Marathi grammar rules NLP is used.
* To display the parser tree for every valid Marathi sentence NLTK tree module is used.

## References

Harish BS, Rangan RK, “A comprehensive survey on Indian regional language processing”, Proceedings of the international conference [ACCTA-2010] on Special Issue of IJCCT, June 2020.

M. Rajani Shree, “Syntactic Parsing in Kannada Text/Natural Language Processing”, Third International Conference on Intelligent computing information and control systems ICICCS, March 2022.

Das BR, Singh D, Bhoi PC, Priyadarshini P “Parsing for Natural Language in Odia: a novel study” , International Conference on Computer Science, Engineering and Applications (ICCSEA), March 2020.

Teodora Dordevic and Suzana Stojkovic , “Different Approaches in Serbian Language Parsing Using Context free Grammers” , Proceedings of the second ınternational conference on data science, E-learning and ınformation systems, Sep2020.

Denis Eka Cahyani, Langlang Gumilar, Ajie pangestu “Indonesian Parsing Using Probabilistic Context-Free Grammer and CYK”, Third International Seminar On Research Of Technology and Intelligent System, Dec 2020.

Deepa Yogish, T.N.Manjunath, Ravindra S.Hegadi “Review on Natural Language Processing Trends and Techniques Using NLTK”, Part of the [Communications in Computer and Information Science](https://link.springer.com/bookseries/7899) book series (CCIS,volume 1037),July 2017.

# Overall Description

## Product Perspective

Self-contained product.

## Product Functions

* To Parse string of symbols, according to the rules of a Marathi grammar.
* To analyse the grammatical structure of natural language.
* To separate units like subject, verb, and object and determines the relations between these units.
* when there is a need to represent input data from source code abstractly as a data structure so that it can be checked for the correct syntax

## User Classes and Characteristics

The user will be those who have knowledge of Marathi language, they can use this project.

## Operating Environment

**Software Requirement**

* Python
* NLP

**Hardware Requirement**

* Windows 7 or higher
* I3 processor system or higher
* 4 GB RAM or higher
* 100 GB ROM or higher

## Design and Implementation Constraints

Hardware requirement

* Windows 7 or higher
* I3 processor system or higher
* 4 GB RAM or higher
* 100 GB ROM or higher

**Technoloogy**

Python, NLTK, NLP

**Tool**

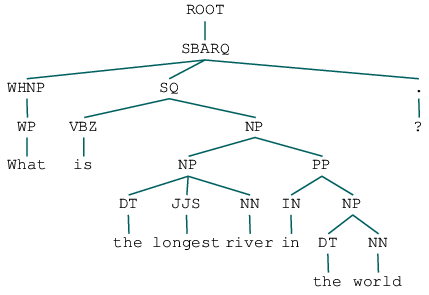
Jupyter notebook

## Assumptions and Dependencies

Our project is dependent on Jupyter notebook tool.

# External Interface Requirements

## User Interfaces



## Hardware Interfaces

* Windows 7 or higher
* I3 processor system or higher
* 4 GB RAM or higher
* 100 GB ROM or higher

## Software Interfaces

* Jupyter notebook
* NLTK
* NLP

# System Features

* To Parse string of symbols, according to the rules of a Marathi grammar.
* To analyse the grammatical structure of natural language.
* To separate units like subject, verb, and object and determines the relations between these units.
* when there is a need to represent input data from source code abstractly as a data structure so that it can be checked for the correct syntax

**5. Other NonFunctional Requirement**

**5.1 Performance Requirement**

The performance of the system lies in the way it is handled. Every API user and end user must be given proper guidance regarding how to use the system.

**5.2 Safety Requirement**

To ensure the safety of the system, perform regular monitoring of the system so as to perform the proper working of the parse.

**5.3 Security Requirement**

A data flow on the network should be handled with necessary encryption of data. Any unauthorized API user should be prevented from accessing the system.

**5.4 Software Quality Attributes**.

Reliability - The system is fault tolerant and the user can check syntax easily.

Portability - The system will be an NLP system and can be installed on any system.

Usability - The system is easy to use, understand, and highly secure.

**5.5 Business Rules**

The parser module provides an interface to Python's internal parser and byte-code compiler. The primary purpose for this interface is to allow Python code to edit the parse tree of a Python expression and create executable code

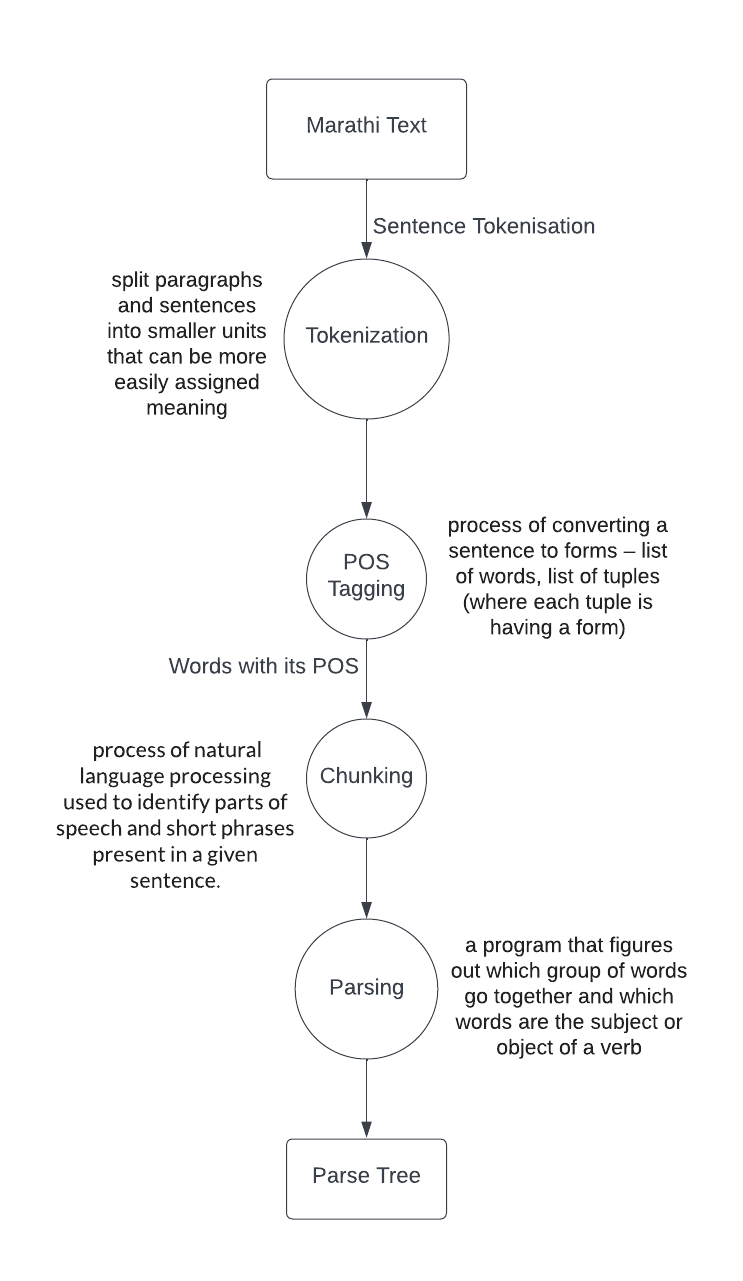
**Appendix A: Glossary**

NLP: Natural Language Processing

NLTK: Natural Language Tool Kit

**Appendix B:**

**Data flow Diagram**



:

Sequence Diagram

