Sri Lanka Institute of Information Technology

SLIIT

COMPUTING | BUSINESS | ENGINEERING

Software Requirements Specification -

Arithmetic Problem Solving System for Grade 5 Students using neural networks

(මහෞෂධ2)

B.Sc. Special (Hons) Degree in Information Technology

Specialized in Information Technology

Date of Submission: 2015-11-09

**Title : **Arithmetic Problem Solving System for Grade 5 Students using neural networks

Project ID: 15J-004

Author:

Student ID	Name	Signature
IT 12143214	De Silva K.C.E	

……………………………………

Mr.Yashas Mallawarachchi

Supervisor

……………………………………

Mr.Anupiya Nugaliyadde

Co-Supervisor

Date of Submission:

2015-11-09

Declaration

I hereby declare that this proposal document or part of it is not a copy of a document done by any organization, university, institute or a previous student project group at SLIIT and is not copied from the Internet or other sources except where stated.

Author:

Student ID	Name	Signature
IT 12143214	De Silva K.C.E	

…………………………………….

Mr.Yashas Mallawarachchi

Supervisor

Table of content

List of figures

List of tables

1Introduction

The introduction serves to orient the reader. It describes both the system and the SRS itself by giving purpose and scope of the document. System Overview contains a description of the system. This is essentially a brief summary of the vision and scope of the project.

1.1Purpose

This document is intended to provide all of the requirements for the product Arithmetic problem solving system for grade 5 students using neural network. This document addresses different areas of the software under different sections and provides comprehensive information regarding the product.

This Software Requirement Specification (SRS) will act as the contract for the development of the system. Furthermore, it will offer the team of developers a clear and precise overview of the system and its requirements. It will bring about a reduction in the development effort of the system, provide a basis for cost estimation and schedules , provide a baseline for the plan of testers, serve as a basis for enhancement and support the system designers and implementers with pertinent information to release the most accurate system.

The functionality of the system in focus the mode of interaction between the system and the user hardware, other software identified as external interfaces, the attributes pertaining to the system in terms of accuracy, maintainability, security, the design constraints imposed upon the implementation of the software are the basic minutiae discuss in this Requirement Specification.
The document will describe the interfaces of the system, what the system will do, the under which the system should operate.

The intended audiences of this System Requirement Specification are, the members of the research group, project supervisor, Mr.Yashas Mallawarachchi , project co – supervisor Mr.Anupiya Nugaliyadde

1.2Scope

The requirements for release “Arithmetic problem solving system for grade 5 student using neural network” are documented. The approach to accomplish expected goals is researched in advance and conversed meticulously right through the document. The main characteristics and behavior of the system is firmed to attest a user-friendly, effective and efficient system.

This system is intended to be used by grade 5 scholarship students who are not interesting to solve arithmetic problems while studying mathematics at the school. Proposed mathematical problem solving system that is capable of understanding a detailed sentence mathematical problem in Sinhala language.

1.3Definitions, Acronyms, and Abbreviations

Term	**Definition**
NLP	Natural Language Processing
PPMI	Positive Point Wise Mutual Information
QA	Question Answering
APSS	Arithmetic Problem Solving System
PMI	Point Wise Mutual Information
NN	Neural Network

Table 1 definition , acronyms , abbreviation

1.4Overview

The main goal of APSS is automatically learning to solve arithmetic problems. From the literature survey which was conducted to identified that most of the QA systems are designed only for the information retrieval. The functionality provided by this system it is intend to guide the user to work efficiently and reduce the time consumption to solve arithmetic problems.

The system is a combination of four components.

1. Key word identification
2. Mathematical Question identification
3. Identifying Mathematical Operation and generating an equation
4. Neural Network

Objectives to be fulfilled while implementing the system are stated below:

Get the arithmetic question as a text .understand the question and generate the equation and solve it.

Provide user friendly interfaces and maintain better interaction with user and the system. Especially in here we try to cover all the areas in the scholarship syllabus that deal with addition and subtraction operation. Understand correctly the given mathematical problem is very complicated thing. Because we are working with Sinhala language increase the accuracy of the system is very important.so we use neural network to achieve this task. Neural network will check the accuracy and choose the most suitable outcomes of the key word identification and mathematical operation identification

2Overall Description

Question answering is an important approach in NLP. In this proposed mathematical word problem solving system we encourage grade 5 students to learn mathematics in their native language. Since the system is to be created with appropriate interfaces that attract children we hope the system would be a productive one for students in grade 4-5.

Normally Sri Lankan education is a teacher centered paradigm. Within that frame Knowledge is transmitted from professor to students and students are not actively involved. Under that situation a better method is needed to get the student out of the ordinary frame and let the students experience the students-centered paradigm. That leads the student to practice active learning and learn from their own mistakes. To achieve that goal this system will provide simulations for the students on solving the given mathematical problem. Therefor students can do self-learning through the system.

Most of the time mathematics is considered to be a difficult subject. Grade 5 mathematics is the foundation for the higher level mathematical theories. Majority of the Sri Lankan students are learning in the 'Sinhala' language. So it is important to make it available for them to study such a difficult to learn subject in their mother tongue. On that purpose 'Mahoshadha 2' is to be developed to understand and solve mathematical problems in Sinhala language.

We expect that the system will be up to quality standards which will make the system preform the given task efficiently and correctly.

2.1Product perspective

According to literature survey done previously no software designed for arithmetic problem solving system for Sinhala language. Most of the QA systems are designed only for the information retrieval only. This research cover up main functionalities that are handling arithmetic problems. System that is capable of understanding detailed sentence arithmetic problem in Sinhala language. In this action research study, a classroom of 5th grade mathematics students was used to investigate how students solve word problems and how they determine which mathematical approach to use to solve a problem. It was discovered that many of the students read and re-read a question before they try to find an answer.

Most students will check their answer to determine if it is correct and makes sense. Most students agree that mastering basic math facts is very important for problem solving and prefer mathematics that does not focus on problem solving. As a result of this research, it will be emphasized to the building principal and staff the need for a unified and focused curriculum with a scope and sequence for delivery that is consistently followed. . To achieve that goal this system will provide simulations for the students on solving the given mathematical problem. Therefor students can do self-learning through the system.

2.1.1System Interfaces

In order to do the communication between user and the system we have to use interfaces. Using these interfaces data transfers within the system and other modules of APSS can be done in user friendly and trusted manner.

2.1.2User Interfaces

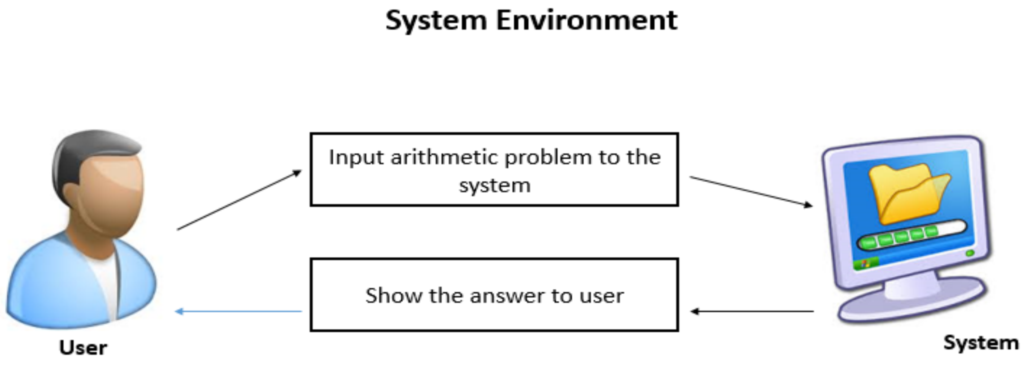


Figure 1

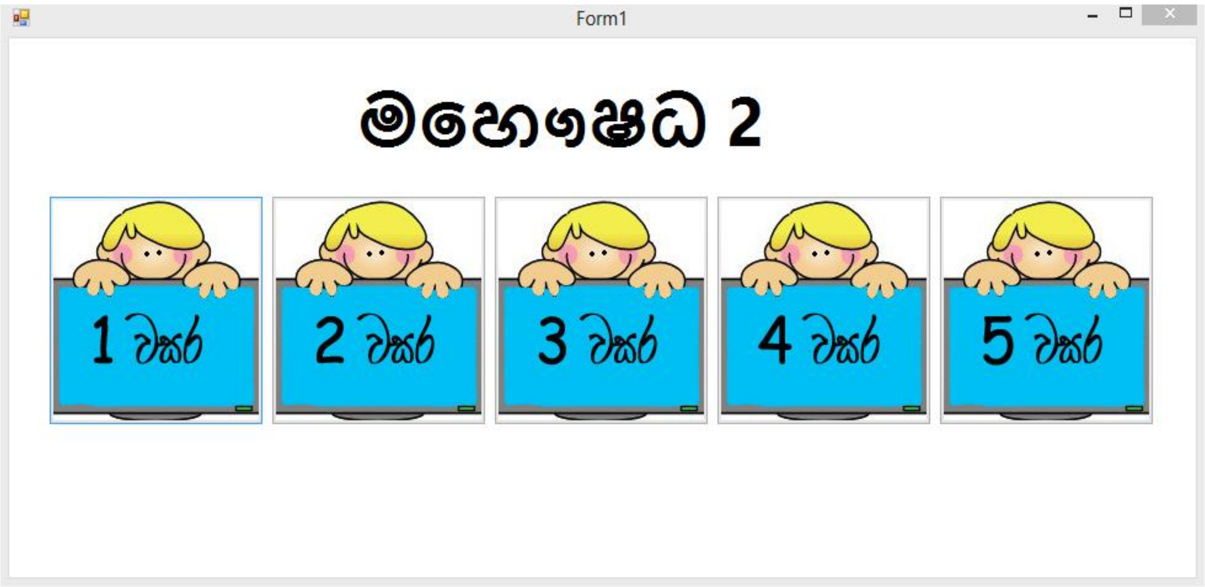
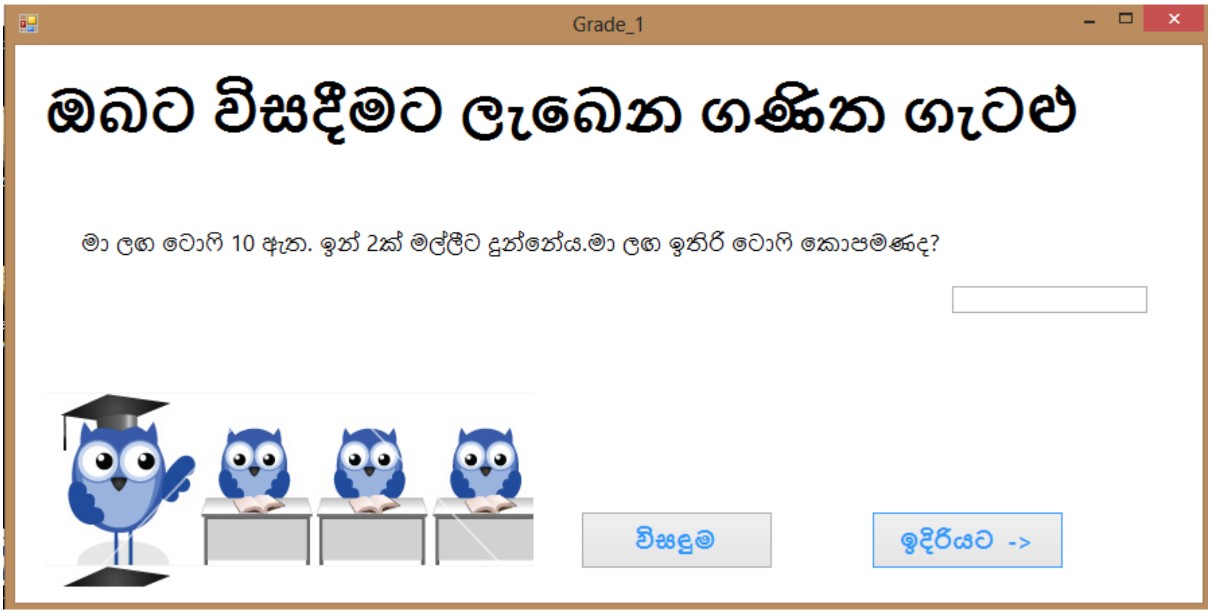


Figure 2 main page



ඔබට විසඳීමට ලැබෙන ගණිත ගැටළු

මා ළඟ ටොෆි 10 ඇත. ඉන් 2ක් මල්ලීට දුන්නේය.මා ළඟ ඉතිරි ටොෆි කොපමණද?

Figure 3 solving problems

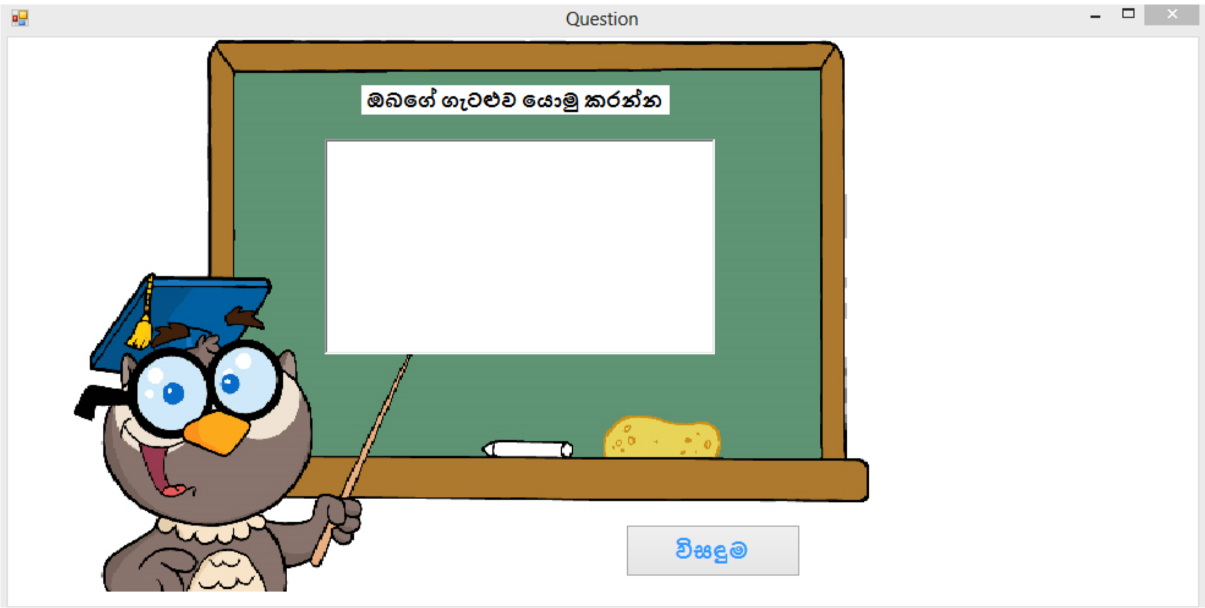


Figure 4 user enter their questions

Available functionalities of keyword extraction are listed bellow

1. User can enter the question here that need to solve.
    a. Only Sinhala String characters allowed
    b. User can enter maximally 100 characters at a time and it can be set of words or set of sentences.
2. If the program is currently run and if any words are currently selecting as keyword then progressing text are loading to next function and corresponding visual output is displaying in terminal.
    a. It is read only text.
3. Visual output of input text is display in this window as a 3D animation.

2.1.2.1Nonfunctional user interfaces requirements

- Less number of controls so that anyone can use easily.
- User friendly tools and environment.
- User can interacting with the system using Sinhala language.

2.1.3Hardware interfaces

The following hardware specification are required by the APSS in order to operate.

- Dual core processer or above
- Hard Disk 5GB or above
- Keyboard
- Mouse
- Display unit (LCD or CRT monitor)
- Networking Equipment
- I GB above RAM

According to the drivers software products that are used to develop explicit components of the system,the prerequisite of Processing power and memory (RAM) may be at variance.

2.1.4Software interfaces

The following software interfaces are required for the development of APSS

- Install Linux OS

2.1.5Memory constraints

Following memory capacities are need by APSS to operate

- Internal Memory: 512 MB
- Disk Space: 1024 MB of free disk space

2.1.6Operations

1. Key word identification
2. Mathematical Question identification
3. Identifying Mathematical Operation and generating an equation
4. Neural Network

2.2Product functions

2.2.1Identify Keywords

Develop a new keyword extraction technique .Identify verbs and entities (candidates) in the question by using morphological parser and name entity recognition analysis. Parser is a program, usually part of a compiler that receives input in the form of sequential source program instructions, every question is parsed to parser and breaks them up into nouns, verbs, and their attributes or options are extracted.

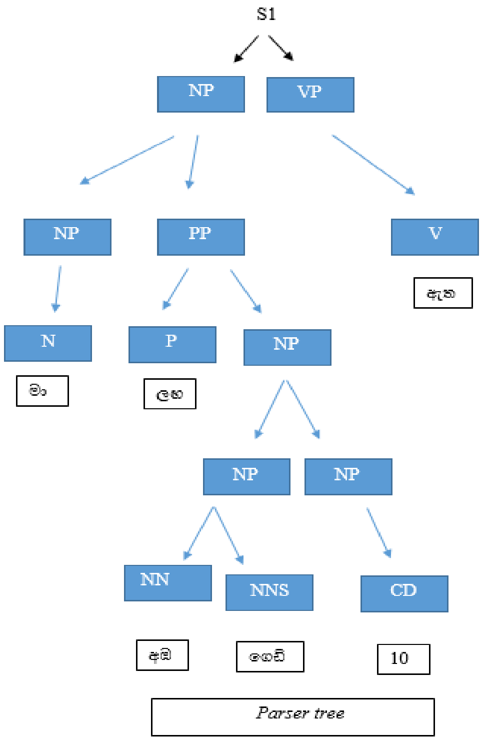


Figure 5 : parser tree example

For our project we hope to use morphological parser [7] capable of analyzing and generating Sinhala verbs. Morphological analysis and generation plays a vital role in many applications related to natural language processing. Morphological parser is not like Stanford dependency parser which can only identify the verbs in Sinhala sentence so to identify the entities we have to use name entity recognition system.

The named entity recognition output is used to identify numbers and people etc.. Because entities are references to some object whose quantity is observed or changing throughout the problem .

Properties calculation

After identifying each candidate, we need to calculate properties that indicate that it may be a keyword. For that we use PPMI to identify the word similarity.

Word similarity is a measure of how semantically similar a pair of words is, with synonyms having the highest value. It is widely used for applications in (NLP), information retrieval, and artificial intelligence.

Calculating PMI only requires simple statistics about two words, their marginal frequencies and their co-occurrence frequency in a corpus. More frequent words tend to have more senses, the synonyms under-weighted by PMI are often those with high frequency.so to identify more suitable keywords we hope to develop a new PMI metric.

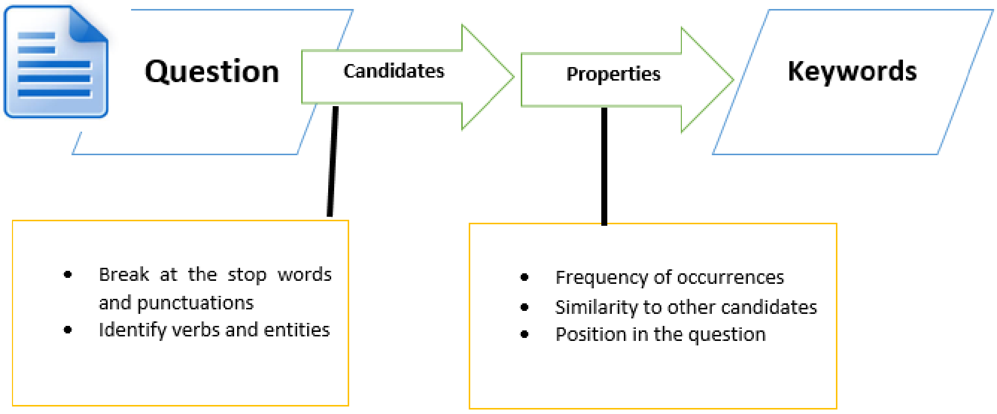

Question
Candidates
Properties
Keywords
Break at the stop words and punctuations
Identify verbs and entities
Frequency of occurrences
Similarity to other candidates
Position in the question

Figure 6 candidate selection

2.2.2Question identification

Create algorithm to identify the mathematical problem. This step involves training a model to identify verb categories for sentences. For that Conditional random field algorithms will be used.

2.2.3Identify and perform mathematical operation

Develop a method to find mathematical operation using relevant key words and verb categories. After finding the consisting of entities, Containers, Attributes, relations among the entities and quantities, mapping the algebraic problem in to an equation. Solving the equation and get the final answer.

2.2.4Neural Network

Develop the NN for the proposed system. This is a large scale neural network which consists of 2 hidden layers, input layer and output layer. Outputs from other 3 functions are inserted as inputs to the neural network in separate layers. Final answer will be generated at the output layer.

3Specific requirements

3.1Use case scenarios

3.1.1Use case scenario for keyword extraction

Use Case name	Key word extraction
Actor	User , system
Pre-Condition	System is running, arithmetic question is received from the user
Success Scenario	1. The use case starts when the system received a arithmetic question 2. Token extraction from given question. 3. Extracting nouns, verbs, adjectives …. etc. 4. Calculate the properties of each candidate. 5. Select possible keywords.
Exceptions	4a. if system get different values for the calculation get the highest value among them. 4b. if system get negative values for the calculation omit that words.
Post-Conditions	System has successfully extracted required keywords

Table 2 keyword extraction

3.1.2Use case scenario for Identifying Mathematical Operation

Use Case name	Identifying Mathematical Operation and Generate the equation
Actor	User , system
Pre-Condition	System is running, key words are extract from the question, Question identification is done
Success Scenario	The use case starts when the system received extracted key words and question identification is done. 2. Forming the states. 3. Match the entities and categories. 4. Initializes or updates the values of the containers in the state. 5. Quantities of entities update or observe. 6. Using the learned verb categories find the relevant mathematical operation. 7. Forms an equation by comparing the quantities for containers matched between the two states.
Exceptions	3a. An entity/category is matched if it has the same head word and same set of attributes as an existing entity/category. 3b. If an entity or category cannot be matching, then a new state is created. 4a. For the matched entities, initializes or updates the values of the containers in the state.
Post-Conditions	

Table 3 Identify mathematical operation

3.1.3 Use case scenario for question identification

Use Case name	Question Identification
Actor	User , system
Pre-Condition	System is running, arithmetic question is received from the user
Success Scenario	1. The use case starts when the system received an arithmetic question 2. Word chunking for the received question 3. Tagging the words 4. Identifying the question words(what, how many , how much etc..) 5. Identify the question
Exceptions	
Post-Conditions	System has successfully identified the question

3.1.4Use case scenario for Neural network

Use Case name	Handle inputs in hidden layers
Actor	User , system
Pre-Condition	System is working, outputs have been generated by other functions
Success Scenario	1. Use case scenario starts when the neural network receives keywords to the first hidden layer. 2. Weights are assigned. 3. Sigmoid activation function is applied to weights. 4. Threshold the results. 5. Get the values that exceed the threshold. 6. Outputs are sent to next hidden layer.
Exceptions	5a. if no input value pass through the threshold, layer takes next set of inputs
Post-Conditions	Output must be sent to the next layer.

Table 4 Neural network

Use Case name	Insert a mathematical problem to the database
Actor	User , system database
Pre-Condition	System is working properly
Success Scenario	1. Use case scenario starts when the user press “add new question” button 2. User is given a text area to insert new question. 3. User enters the word problem. 4. User press enter button.
Exceptions	-
Post-Conditions	Update the database.

Table 5Insert a mathematical problem to database

3.2Performance requirements

The key areas needed to be taken into consideration to enhance the performance of the system can be categorized as follows. The system should run on 32 bit (x86) or 64 bit (x64) Dual-core 2.66-GHZ or faster processor. The speed at which a gesture is recognized shall not exceed 10ms.

- **Efficient and reliable keyword extraction**

At the runtime system has to extract necessary Sinhala words in order to recognize keyword accurately. Averagely system should be able to extract all Sinhala words within one or two second so that give more reliable output. System should be able to continuously calculate PPMI value of each sentence as they are submitted to the system while sending data values to the next step.

3.3Design Constraint

C++ is the programming language that will be used in the software together. In order to achieve availability, accuracy and performance and efficiency aspirations, a high quality design process is needed. Effective processes can help developers to measure how well design goals are being met and regulate functional modules. The systems design should acclimatize the efficiency, data analyze and cost benefit tradeoffs. As most of the users are disabled it is necessary to have simple interfaces which can be used with no prior computer skill.

3.4Performance Requirements

The only way in which systems will meet their performance targets is for them to be specified clearly and unambiguously. It is a simple fact that if performance is not a stated criterion of the system requirements then the system designers will generally not consider performance issues. The performance of the product must be to the highest level, since lot of mathematical questions are handled through the system. Since the system consists of lot sub modules coming up with a suitable design specifically concentrating on performance wise is very important.

3.5Safety Requirements

Safety consistent with mission requirements, is designed into the software in a timely, cost effective manner. The system will always be confirmed to run properly and give the correct answer effectively, therefore there is no risk of any data losses. Further data retrieval is also efficient.

3.6Security Requirements

Security is often an afterthought during software development. Realizing security early, especially in the requirement phase, is important so that security problems can be tackled early enough before going further in the process and avoid rework. According to the user privileges of the system, user can perform his /her relevant tasks which are assigned to them by the system

.

4References

[1] Improving Word Similarity by Augmenting PMI with Estimates of Word Polysemy Lushan Han1, Tim Finin1;2, Paul McNamee2, Anupam Joshi1 and Yelena Yesha1 1 Computer Science and Electrical Engineering ,University of Maryland, Baltimore County , 2 Human Language Technology Center of Excellence , Johns Hopkins University
29 December 2011

[2] “Learning to Solve Arithmetic Word Problems with Verb Categorization” (by Mohammad Javad Hosseini (University of Washington), Hannaneh Hajishirzi(University of Washington), Oren Etzioni(Allen Institute for AI), and Nate Kushman(Massachusetts Institute of Technology)

[3] Automatic Key phrase Extraction based on NLP and Statistical Methods
Martin Dostal and Karel Je_zek
Department of Computer Science and Engineering, Faculty of Applied Sciences
University of West Bohemia, Plze_n, Czech Republic, fmadostal, jezek kag@kiv.zcu.cz

[4] Automatic Keyword Extraction From Any Text Document Using N-gram Rigid Collocation
Bidyut Das, Subhajit Pal, Suman Kr. Mondal, Dipankar Dalui, Saikat Kumar Shome
International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-2

[5] Sinhala Ontology Generator for English to Sinhala Machine Translation
CONFERENCE PAPER · JANUARY 2014 Budditha Hettige General Sir John Kotelawala Defence Unive...George Rzevski The Open University (UK)

[6] http://www.cplusplus.com/forum/general/

[7] http://texlexan.sourceforge.net/

5 Appendices

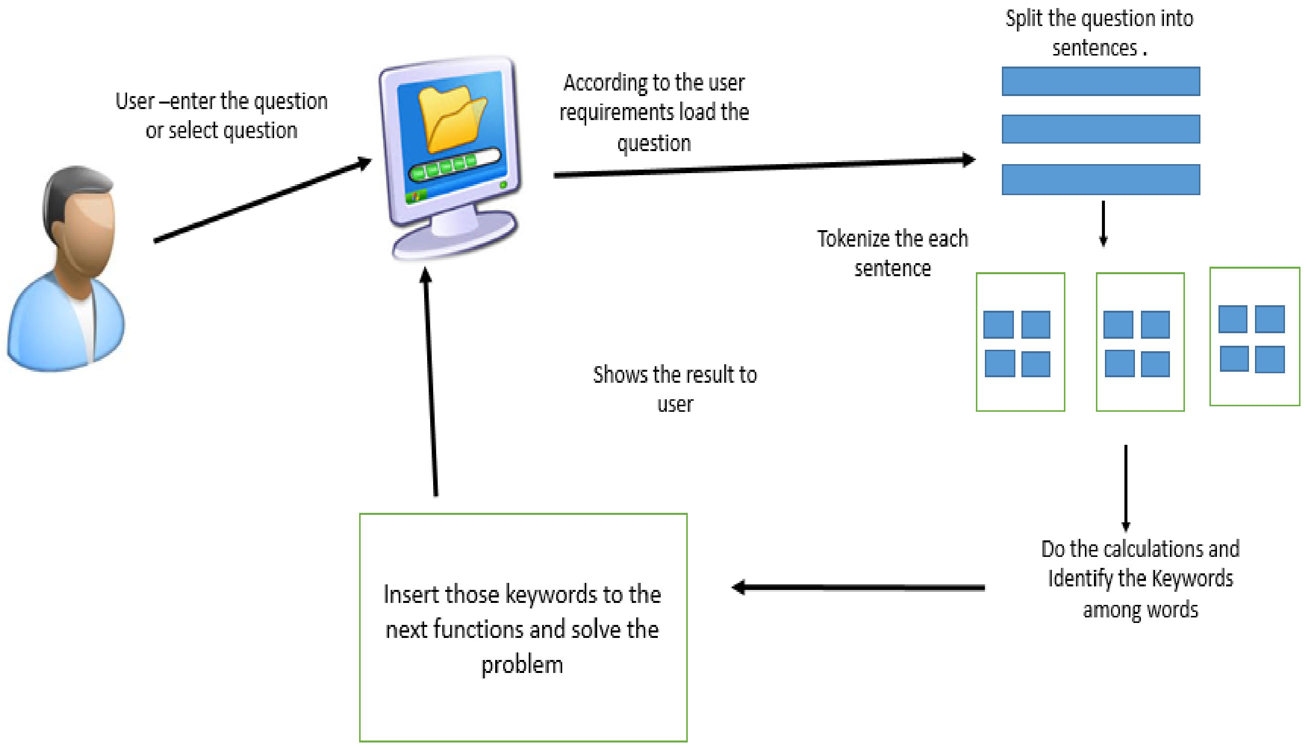

User –enter the question or select question
According to the user requirements load the question
Split the question into sentences .
Tokenize the each sentence
Shows the result to user
Do the calculations and Identify the Keywords among words
Insert those keywords to the next functions and solve the problem

Figure 7Purposed Process for keyword extraction at runtime