

A Digital Aid for Hindi Language Learning and Teaching

By

Shivani Thakkar (13CEUOS090)

Yash Pradhan (13ICUOG002)

A project submitted

In

partial fulfillment of the requirements
for the degree of

BACHELOR OF TECHNOLOGY

in

Computer Engineering

Internal Guide

Dr. C. K. Bhensdadia
Head of Department
Dept. of Comp. Engg.

External Guide

Prof. Pushpak Bhattacharyya
FNAE, Director, IIT Patna.
Professor, Dept. of CSE, IIT Bombay



Faculty of Technology
Department of Computer Engineering
Dharmsinh Desai University
April 2017

CERTIFICATE

This is to certify that the project work titled

A Digital Aid for Hindi Language Learning and Teaching

is the bonafide work of

**Shivani Thakkar (13CEUOS090)
Yash Pradhan (13ICUOG002)**

carried out in the partial fulfillment of the degree of Bachelor of Technology in Computer Engineering at Dharmsinh Desai University in the academic session

December 2016 to April 2017

Dr. C. K. Bhensdadia
Guide & Head
Dept. of Computer Engg.



**Faculty of Technology
Department of Computer Engineering
Dharmsinh Desai University
April 2017**

Acknowledgements

We would like to heartily thank our guide **Dr. Pushpak Bhattacharyya** for allowing us to work as research interns in CFILT Lab at IIT Bombay. We would like to thank him for his endless support and motivation. His perfection and meticulous nature has inspired us much.

We would like to convey our heartfelt gratitude to **Dr. C. K. Bhensdadia** (Head of Department, Computer Engineering Department, DDU) for granting us an opportunity to do internship in one of the best Research/Engineering institute of India, IIT Bombay. We are grateful for his constant guidance and encouragement.

We would like to thank entire Tata Center team members: Technical as well as Lexical team. We are highly indebted to the entire team for their guidance and constant supervision. We would also like to thank our mentors **Sandhya Singh** and **Hanumant Redkar** for their guidance.

We would like to thank entire **CFILT** (Center for Indian Language Technology), IIT Bombay Lab members for sharing their pearls of wisdom and being an integral part of our internship period here.

We are also thankful to all our **family-members** for their unconditional support and motivation throughout.

Abstract

Education domain is gaining great attention as it serves as the base of a successful career. It has become imperative to educate the students in an environment which liberates them from burden and inspires them to learn more effectively.

To fulfill the above purpose, digital education plays a major role. Today's children are born into a technology based society and are exposed to laptops and smart devices a lot. In such scenario, traditional book based education seems little monotonous to them. To overcome this issue, digital learning should be emphasized right from the school days. There are multiple benefits of digital learning in transforming a child's life like visual learning, cultural awareness, improved academic performance, inventiveness and many more.

Language impacts the daily lives of members of any race, creed, and region of the world. The main agenda of the project is to remodel Hindi language WordNet data to suit school teaching and learning environment and provide digital aid for students. It has been observed by psychologists that school children have a tendency to learn efficiently if they could relate the material with real life objects. The learning process will be aided through audio visual inputs by simplifying definitions of words and their grammatical usage along with audio pronunciations.

The major task for us under this project was to develop an android application and handle database synchronization. The android application encapsulates features like learning words by associating with images, audio pronunciations, concept definitions, illustrations and grammatical usage. We have completed database synchronization task of our project's database with Hindi WordNet database with the aid of scripts.

Contents

List of Figures	vii
List of Tables	viii
1 Introduction	1
1.1 Introduction to Natural Language Processing	2
1.1.1 Stages of Natural Language Processing and its associated ambiguities	2
1.1.2 Applications of Natural Language Processing	6
1.2 Introduction to Machine Learning	6
1.2.1 What is Machine Learning?	6
1.2.2 Classification of Machine Learning Algorithms	7
1.2.3 Applications of Machine Learning	8
1.3 Introduction to A Digital Aid for Hindi Language Teaching and Learning	9
1.3.1 Aim and Motivation	9
2 Introduction to WordNet	11
2.1 Origin and Evolution of WordNet	12
2.2 About Hindi WordNet	14
2.3 Relations in Hindi WordNet	15
3 A Digital Aid for Hindi Language Learning and Teaching	17
3.1 Description of the digital aid	18
3.2 Software Requirement Specifications	20
3.2.1 Purpose	20
3.2.2 Scope	20
3.2.3 Definitions, Acronyms and Abbreviations	21
3.2.4 Hardware Requirements	21
3.2.5 User Characteristics	21
3.2.6 Software Process Model	21
3.2.7 Non-functional Requirements	22
3.2.8 Installation Requirements	22
3.2.9 Functional Requirements	22
3.3 Analysis and Design	24

3.3.1	Database	24
3.3.2	Use-case diagram	26
3.3.3	Entity Relationship Diagram	27
3.3.4	Class diagram	29
3.3.5	Sequence Diagram	30
3.3.6	Activity diagram	32
3.3.7	Implementation Details	33
3.3.8	Testing	39
3.3.9	Screen shots	40
4	Prototype Field Testing of Digital Aid at Schools	46
4.1	Response from the students	47
4.1.1	IIT Bombay Campus School	47
4.1.2	IIT Bombay Kendriya Vidyalaya	49
4.2	Feedback from the teachers	50
5	Database Synchronization	51
5.1	Introduction to Database Synchronization	52
5.1.1	Challenges of Database Synchronization	54
5.2	Syncing of tatacenter's database with Hindi WordNet	56
5.2.1	Structure of Hindi WordNet database	56
5.2.2	Structure of tatacenter_hindi Database	57
5.2.3	Challenges faced during tatacenter database syncing	61
5.2.4	Implementation details	63
5.2.5	Cron Job for scheduling synchronization	68
5.2.6	Screenshots	69
6	Conclusion and Future Extension	70
6.1	Conclusion	71
6.2	Future Extension	71
	Bibliography	72

List of Figures

1.1	Structure Detection	4
2.1	Relationship between Concepts	15
2.2	Web Interface of Hindi WordNet	16
3.1	Database - Abstract	24
3.2	Use Case Diagram	26
3.3	ER diagram for tatacenter_hindi database	27
3.4	ER diagram for tatacenter_master database	28
3.5	Class Diagram	29
3.6	Sequence Diagram Word Found	30
3.7	Sequence Diagram Word Not Found	31
3.8	Activity Diagram	32
3.9	Select Expertise Level Screen	40
3.10	Search Interface Screen	41
3.11	Varnamala search for letter	42
3.12	Display details for the the varnamala search	43
3.13	Display details for the the word search of Beginner level	44
3.14	Display details for the the word search of Intermediate level	45
4.1	IIT Bombay Campus School	48
4.2	IIT Bombay Kendriya Vidyalaya	49
4.3	IIT Bombay Kendriya Vidyalaya	49
5.1	source_table	52
5.2	destination_table_beforesync	53
5.3	destination_table_aftersync	53
5.4	cron_job_scheduling	69
5.5	cron_job_logging_details	69

List of Tables

3.1	Test Cases	39
5.1	tbl_all_gender	56
5.2	tbl_all_synset	57
5.3	tbl_all_words	57
5.4	tcp_synset	58
5.5	tcp_synset_example	58
5.6	tcp_word_collection	59
5.7	tcp_word	59
5.8	tcp_synser_words	60

Chapter 1

Introduction

1.1 Introduction to Natural Language Processing

Computer science has grown, by leaps and bounds in recent times and one of the major reasons has been Artificial Intelligence. Artificial Intelligence has always pushed in the boundary of computer science and engineering by demanding more and more from the machine.

In today's internet world, we have enormous amount of text, in form of a number of documents. If a human is asked to make sense of all those data, how much of data can human being possibly see? Natural Language Processing is a field of Artificial Intelligence that deals with analyzing, understanding and generating the languages that humans use naturally in order to interface with computers in both written and spoken contexts using natural human languages instead of computer languages. NLP researchers mainly focus on gathering knowledge that how human interpret and use language so that tools and techniques can be developed to achieve desired goals. This is a very challenging problem as human language in itself is ambiguous.

1.1.1 Stages of Natural Language Processing and its associated ambiguities

Natural Language Processing is divided into following stages:

1. Phonetics and Phonology
2. Morphology
3. Lexical Analysis
4. Syntactic Analysis
5. Semantic Analysis
6. Pragmatics
7. Discourse

Phonetics and Phonology

This stage of NLP is concerned with processing of speech which includes speech characteristics like accent, pause, amplitude and tone.

Challenges in this phase

1. Homophones - Homophones are the words which are nearly identical in pronunciation but different in meaning.

For example, ate and eight are the words which sound similar but have different meanings, ‘ate’ is the past tense of eat while ‘eight’ represents number 8.

2. Word Boundary – It becomes difficult to infer the meaning of a statement due to the different word boundary interpretation.

For example, I got [ula]plate has two interpretations:

- (a) I got up late.
- (b) I got a plate.

Morphology

This stage of NLP is concerned with word formation rules from root words. A smallest meaningful unit in the grammar of a language is known as a morpheme.

For example, ‘playing’ has morph ‘play’ and ‘stretched’ has morph ‘stretch’.

This is the crucial first stage of NLP. The languages with rich morphology have the advantage of easier processing at higher stages of processing.

Lexical Analysis

This stage of NLP deals with dictionary access and obtaining the properties of the word. For example, ‘dog’ has following properties:

1. It is a noun. (Lexical Property)
2. It takes ‘s’ in plural form. (Morphological Property)
3. It is an animate object. (Semantic Property)
4. It is a four legged animal. (Semantic Property)
5. It is a carnivore. (Semantic Property)

This stage of NLP holds great significance in the fields of research like Information Extraction, Information Retrieval, Question Answering and many more.

Challenges

1. Part of Speech Disambiguation

The word ‘dog’ has two part of speech categories- noun and verb. ‘dog’ as a noun means the four legged animal and ‘dog’ as a verb means ‘to pursue’.

2. Sense Disambiguation

If we consider ‘dog’ as a noun, it can also have two interpretations: one as an animal and other as a detestable person.

Syntactic Analysis

This stage of NLP is concerned with detection of structure of the sentence. In this stage, any sentence is broken down into noun phrases and verb phrases. They are further broken down into different entities. This is known as structure detection and the whole problem of converting a sentence in a tree like structure like this, is called parsing.

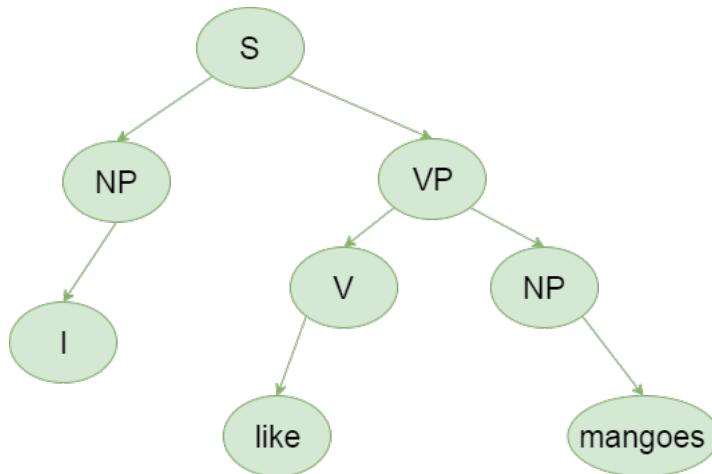


Figure 1.1 Structure Detection

Challenges

1. Scope Ambiguity

Consider the sentence ‘The old men and women were taken to safe locations’, the word ‘old’ has a scope ambiguity. It is difficult to infer that whether old men and old women were taken to safe locations or old men and all women were taken to safe locations.

2. Preposition Phrase Attachment Ambiguity

For example, in the sentence ‘I saw a boy with a telescope’ has Preposition Phrase Attachment Ambiguity. There are two possible interpretations of this sentence.

- (a) I saw the boy and he had the telescope.
- (b) I saw the boy with the help of telescope.

Semantic Analysis

In this stage of NLP, the structures created by the syntactic analyzer are assigned contextual meaning. After deriving the words and their syntactic structures, sentence processing moves towards meaning extraction. The sentence is represented in one of the unambiguous forms like predicate calculus, semantic net, frame, conceptual dependency, conceptual structure, etc.

The ambiguity at this stage arises out of the ambiguities of semantic roles and representation. For example, in the sentence “Visiting aunts can be trying”, are the aunts visitors (agent role) or are they being visited (object role)?

Pragmatics

In this stage of NLP, knowledge regarding a sentence is derived from external commonsense information.

For example, in the sentence ‘The old men and women were taken to safe locations’, it is inferred that along with old men, both young and old women are likely to be taken to safe location.

Discourse

The meaning of an individual sentence may depend on the sentence preceding it and may influence the meaning of the sentence they follow it.

For instance, ‘John is sleeping. He is a boy.’ Here the word ‘He’ refers to John.

1.1.2 Applications of Natural Language Processing

Any application that utilizes text is to be derived from NLP. The most frequent applications utilizing NLP include the following:

1. Sentiment Analysis
2. Information Retrieval
3. Information Extraction
4. Question - Answering
5. Text Summarization
6. Machine Translation

1.2 Introduction to Machine Learning

Each time when a web search on Google or Bing is performed, the reason that it works so well is because their machine learning software has figured out how to rank those pages. When Facebook or Apple's photo application automatically recognizes your friends in your pictures, that's also machine learning.

1.2.1 What is Machine Learning?

Machine Learning is a sub-domain of Artificial Intelligence that provides computers with the ability to learn without being explicitly programmed. Machine learning focuses on the development of computer programs that can change when exposed to new data. A Machine Learning system searches through data to look for patterns, extracts and learns from the patterns and adjusts actions of the program accordingly.

Two definitions of Machine Learning are offered:

1. Arthur Samuel described it as: "The field of study that gives computers the ability to learn without being explicitly programmed." This is an older, informal definition.
2. Tom Mitchell provides a more modern definition: "A computer program is said to learn from Experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

Examples:

1. Playing Checkers.

E: The experience of playing many games of checkers

T: The task of playing checkers.

P: The probability that the program wins the next game of checkers.

2. Email Spam classifier.

E: The experience of classifying emails.

T: The task of classification of emails.

P: The probability that the program correctly classifies next email.

1.2.2 Classification of Machine Learning Algorithms

In general, any of the machine learning algorithm fall under the following two categories:

1. Supervised Learning
2. Unsupervised Learning

Supervised Learning

Supervised Learning is the machine learning task of learning a model from labeled set of training data. In supervised learning, each training example is a pair of an input object (typically a vector) and a desired output value. A supervised learning algorithm analyzes the training data and learns a model, which can be used for mapping new examples. In simple words, we are given a training data set and we already know what our correct output should look like, having the idea that there is a relationship between the input and the output.

Supervised Learning problems are further categorized into following:

1. Regression: In a regression problem, we are trying to predict results within a continuous output, meaning that we are trying to map input variables to some continuous function.

Example: Given data about the size of houses on the real estate market, we try to predict their price. Price as a function of size is a continuous output.

2. Classification: In a classification problem, we are instead trying to predict results in a discrete output. In other words, we are trying to map input variables into discrete categories.

Example: Given the information about breast tumor, we can classify the breast cancer as benign or malign.

Unsupervised Learning

Unsupervised Learning, on the other hand, allows us to approach problems with little or no idea what our results should look like. We can derive structure from data where we don't necessarily know the effect of the variables. We can derive this structure by using clustering algorithms to group the data based on relationships among the variables in the data.

Example: Google News, groups news articles into cohesive groups based upon similarity between them.

1.2.3 Applications of Machine Learning

There are numerous applications of Machine Learning around us. A few of them are listed below

1. Predicting prices of houses based upon available information of houses.
2. Determining whether the breast cancer is malignant or benign based on factors such as size of tumor, etc.
3. News articles are grouped together using machine learning.
4. Machine Learning algorithms are extensively used in the field of computational biology.
5. Automatic photo tagging.
6. Spam filters uses machine learning algorithms.
7. Grouping medical records based on the similarity between patients, helps to treat them in a better way.
8. Handwriting recognition and Computer Vision.
9. In Natural Language Processing for sentiment analysis, language generation, statistical machine translation etc.
10. Youtube uses machine learning to recommend new videos to users.

1.3 Introduction to A Digital Aid for Hindi Language Teaching and Learning

Education domain is gaining great attention as it serves as the base of a successful career. Different ways and methods of education are explored and applied to improve the quality of education. It has become imperative to educate the students in an environment which liberates them from burden and inspires them to learn more effectively.

To fulfill the above purpose, digital education plays a major role. Today's children are born into a technology based society and are exposed to laptops and smart devices a lot. They are inclined towards games and social media. In such scenario, traditional book based education seems little boring to them. To overcome this issue, digital learning should be emphasized right from the school days. There are multiple benefits of digital learning in transforming a child's life like visual learning, cultural awareness, improved academic performance, inventiveness and many more.

Language impacts the daily lives of members of any race, creed, and region of the world. It helps express our feelings, desires, and queries to the world around us. The communication skills that a child learns early in life will be the foundation for his or her communication abilities for the future. Thus language learning holds great significance in a child's life.

1.3.1 Aim and Motivation

To foster language learning process with the help of digital aid in schools is the main aim of Tata Center for Technology and Design at IIT Bombay. The digital aid addresses this aim of providing education through vocabulary teaching. Education is one of the primary challenges facing India. With the population growth rate of 1.5%, there is tremendous pressure on the education system to provide quality education at affordable price and improve the literacy rate.

The rapid digitization of education has created a gap due to the unavailability of digital resources in the field of language learning. There is a scarcity of digital teaching aids in languages which can be used by both teachers and students. This limited resource trend can be addressed by the product by making available resources in a digital form synced with the school curriculum of all boards.

With the ever increasing size of the classes, there is a need to device ways and means for better classroom management and learner engagement. There is enough study and

evidence to support the claim that technology can be used to improve the learning environment for teachers and students at schools. The goal is to create learning environments where students are actively engaged in the learning process. Student engagement is one of most important factors that affect teaching and student motivation to learn. When students are apathetic toward learning, a barrier to learning is created. The aim of the product addresses this need in the most effective way by providing an audio visual enhanced teaching aid for language acquisition.

Burden of the school bag is a long standing problem of the school going children in India. Various strategies are being thought of to reduce this burden, which becomes one of the major causes of the musculoskeletal pain. The aim of product is to reduce the burden of carrying bulky dictionaries and thesauri and the children can access information through mobiles and other such devices.

The major task for us under this project was to develop an android application and handle database synchronization. The android application encapsulates features like learning words by associating with images, audio pronunciations, concept definitions, illustrations, list of other related words and grammatical usage. We need to write scripts for database syncing of Hindi WordNet database with our project's database.

Chapter 2

Introduction to WordNet

WordNet is an on-line lexical reference system whose design is inspired by current psycholinguistic theories of human lexical memory. WordNet has been used for a number of different purposes in information systems, including word sense disambiguation, information retrieval, automatic text classification, automatic text summarization, machine translation and even automatic crossword puzzle generation.

2.1 Origin and Evolution of WordNet

Standard alphabetical procedures organize lexical information and words that are spelled alike together. A frequent objection to this solution is that finding things on an alphabetical list can be tedious and time-consuming. Many people who would like to refer to a dictionary decide not to bother with it because finding the information would interrupt their work and break their train of thought. [9]

In 1985 a group of psychologists and linguists at Princeton University undertook to develop a lexical database. The initial idea was to provide an aid to use in searching dictionaries conceptually, rather than merely alphabetically—it was to be used in close conjunction with an on-line dictionary of the conventional type. As the work proceeded, however, it demanded a more ambitious formulation of its own principles and goals. WordNet is the result. WordNet can be said to be a dictionary based on psycholinguistic principles.

WordNet includes the lexical categories nouns, verbs, adjectives and adverbs but ignores prepositions, determiners and other function words. Words from the same lexical category that are roughly synonymous are grouped into synsets. The different senses of a polysemous word form are assigned to different synsets. The meaning of a synset is further clarified with a short defining gloss and one or more usage examples. An example adjective synset is: good, right, ripe – (most suitable or right for a particular purpose; "a good time to plant tomatoes"; "the right time to act"; "the time is ripe for great sociological changes").

The price of imposing this syntactic categorization on WordNet is a certain amount of redundancy that conventional dictionaries avoid—words like back, for example, turn up in more than one category. But the advantage is that fundamental differences in the semantic organization of these syntactic categories can be clearly seen and systematically exploited.

All synsets are connected to other synsets by means of semantic relations. These relations, which are not all shared by all lexical categories, include:

1. Nouns

- Hypernyms: Y is a hypernym of X if every X is a (kind of) Y (canine is a hypernym of dog)
- Hyponyms: Y is a hyponym of X if every Y is a (kind of) X (dog is a hyponym of canine)
- Meronyms: Y is a meronym of X if Y is a part of X (window is a meronym of building)
- Holonyms: Y is a holonym of X if X is a part of Y (building is a holonym of window)

2. Verbs

- Hypernyms: the verb Y is a hypernym of the verb X if the activity X is a (kind of) Y (to perceive is an hypernym of to listen)
- Troponyms: the verb Y is a troponym of the verb X if the activity Y is doing X in some manner (to lisp is a troponym of to talk)
- Entailment: the verb Y is entailed by X if by doing X you must be doing Y (to sleep is entailed by to snore)

WordNet is sometimes called ontology, a persistent claim that its creators do not make. The hypernym/hyponym relationships among the noun synsets can be interpreted as specialization relations among conceptual categories. In other words, WordNet can be interpreted and used as a lexical ontology in the computer science sense.

The Open Multilingual WordNet provides access to open licensed wordnets in a variety of languages, all linked to the Princeton Wordnet of English (PWN). The goal is to make it easy to use wordnets in multiple languages.

2.2 About Hindi WordNet

The Hindi WordNet is a system for bringing together different lexical and semantic relations between the Hindi words. It organizes the lexical information in terms of word meanings and can be termed as a lexicon based on psycholinguistic principles. The design of the Hindi WordNet is inspired by the famous English WordNet. In the Hindi WordNet the words are grouped together according to their similarity of meanings, two words that can be interchanged in a context are synonymous in that context. For each word there is a synonym set, or synset, in the Hindi WordNet, representing one lexical concept. This is done to remove ambiguity in cases where a single word has multiple meanings. Synsets are the basic building blocks of WordNet. The Hindi WordNet deals with the content words, or open class category of words. Thus, the HindiWordNet contains the following category of words- Noun, Verb, Adjective and Adverb.

Each entry in the Hindi WordNet consists of following elements.

1. Synset:

It is a set of synonymous words. For example, “कहानी, किस्सा, कथा, दास्तान” represents the concept of narrative, story, tale. The words in the synset are arranged according to the frequency of usage.

2. Gloss:

It describes the concept. It consists of two parts: Text definition: It explains the concept denoted by the synset. For example, “मन से गढ़ा हुआ या किसी वास्तविक घटना के आधार पर प्रस्तुत किया हुआ मौखिक या लिखित विवरण जिसका मुख्य उद्देश्य पाठकों का मनोरंजन करना, उन्हें कोई शिक्षा देना अथवा किसी वस्तु-स्थिति से परिचित कराना होता है” precisely explains the concept of story, narrative. Example sentence: It gives the usage of the words in the sentence. Generally, the words in a synset are replaceable in the sentence. For example, “मुंशी प्रेमचंद की कहानियाँ ग्रामीण परिवेश को अच्छी तरह से दर्शाती हैं।” gives the usage for the words in the synset representing the concept of story, narrative.

3. Position in Ontology:

An ontology is a hierarchical organization of concepts, more specifically, a categorization of entities and actions. For each syntactic category namely noun, verb, adjective and adverb, a separate ontological hierarchy is present. Each synset is mapped into some place in the ontology. A synset may have multiple parents. The ontology for the synset representing the concept school is shown in figure 1.

2.3 Relations in Hindi WordNet

A WordNet is a word sense network. A word sense node in this network is a synset which is regarded as a basic object in the WordNet. Each synset in the Hindi WordNet is linked with other synsets through the well-known lexical and semantic relations of hypernymy, hyponymy, meronymy, troponymy, antonymy, entailment etc. Semantic relations are between synsets and lexical relations are between words. These relations serve to organize the lexical knowledge base.

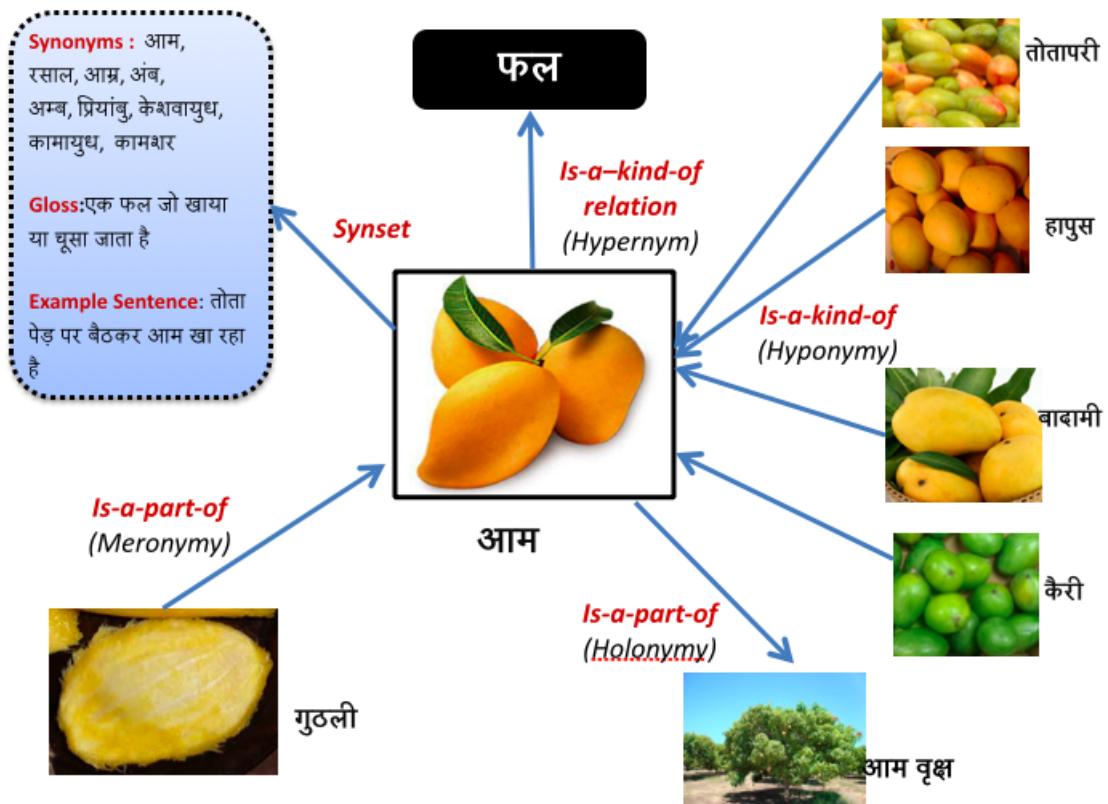


Figure 2.1 Relationship between Concepts



Figure 2.2 Web Interface of Hindi WordNet

The following relations are captured in the Hindi Wordnet.

1. Hyponymy
2. Hypernymy
3. Meronymy
4. Holonymy
5. Entailment
6. Troponymy
7. Antonymy
8. Gradation
9. Causative

Chapter 3

A Digital Aid for Hindi Language Learning and Teaching

The main agenda of the project is to remodel Hindi language Wordnet data to suit school teaching and learning environment and provide digital aid for students. It has been observed by psychologists that school children have a tendency to learn efficiently if they could relate the material with real life objects. The learning process will be aided through audio visual inputs by simplifying definitions of words and their grammatical usage along with audio pronunciations. This kind of resource assisted language learning will address the long standing need of the country of well groomed youth who communicate well, express themselves aptly and foster better understanding in society.

3.1 Description of the digital aid

Mastering vocabulary is a major part of language learning and a prudent understanding of the mental lexicon helps in the development of better instructional strategies. Understanding a word involves myriad of tasks like analyzing different forms of the words, their concept definitions and analogies with other words. Having proper insight of interconnections of words and their lexical relations holds prime significance to define various relations among words like hyponymy/hypernymy, meronymy/holonymy, antonymy, synonymy, polysemy, entailments and many more.

The semantic structure of a language poses the problem of incorrect usage for the learner. Such problems cannot be resolved by merely consulting conventional dictionaries as much of the structural information is omitted from them. However, the relations that are found in the WordNet are superior to the information found in conventional dictionaries and help in solving the problem of incorrect usage.

The main feature of the WordNet is the cognitive synonym set (synset) which is interlinked to other such sets by means of conceptual semantic and lexical relations. This enables the learners to know all possible connotations of a target word and gives them exposure to its various possible contexts. A classroom teacher in a traditional classroom setting will certainly not be able to provide such data to students. This can contribute effectively to both context based and corpus based methods of teaching and learning vocabulary.

The project is divided into 5 levels by allocating different classes of schools in different levels. The aim behind dividing into disparate levels is to introduce different level of features and interface keeping into account the environment suitable for different school standards.

Technologies used:

- Development Environment:

Android Studio is the official Integrated Development Environment (IDE) for Android app development, based on IntelliJ IDEA. On top of IntelliJ's powerful code editor and developer tools, Android Studio offers even more features that enhance your productivity when building Android apps, such as: A fast and feature-rich emulator: a unified environment where you can develop for all Android devices, instant run to push changes to your running app without building a new APK, lint tools to catch performance, usability, version compatibility, and other problems.

- Diagram Tools :

Microsoft Visio, Google's draw.io

- Language:

Android Studio uses Java as a programming language, web services have been developed in PHP.

3.2 Software Requirement Specifications

3.2.1 Purpose

‘Hindi Shabdamitra’, the purpose of this android application is to suit school teaching and learning Hindi language. The learning process will be aided through audio visual inputs by simplifying definitions of words and their grammatical usage along with audio pronunciations.

This SRS document is meant to describe the features of ‘Hindi Shabdamitra’ Application, so as to serve as a guide to designers and developers of the application.

3.2.2 Scope

The task that is going to be carried out by this application is to aid the user in learning Hindi language with the help of images, audios and grammar features as per school curriculum. This android application will be divided into five levels viz. Beginner, Intermediate, Expert, Proficient and Advanced. This version of the application supports language learning for the beginner level i.e. classes I and II. The words are collected from the textbooks of CBSE and ICSE boards for class I and II, which covers around 1500 words. In this version, the following information will be displayed for each searched word:

1. Definition
2. Part of Speech category
3. Usage through example
4. Synonyms
5. Antonyms
6. Gender
7. Plural form
8. Image
9. Audio pronunciation carried out by native speaker of Hindi language

All of this information will be linked with the Tata centre’s database from Hindi Wordnet. For Beginner level students, instead of using definition and example from Hindi Wordnet simplified gloss and example will be included for their ease of understanding. This gloss and example simplification task will be carried out by linguist team

at IIT Bombay CFILT Lab. For understanding of concepts in a better way, images designed by experts from Industrial Design Centre will be included. These images will also capture abstract concepts which are not easily understandable by young minds.

3.2.3 Definitions, Acronyms and Abbreviations

- User: Anyone who uses this app, is a User.
- UI: User Interface
- UX: User Experience
- DB: Database

3.2.4 Hardware Requirements

The application will be deployed on server and the database will also reside on server. The android application will be available on the play store after full development.

3.2.5 User Characteristics

The user should know how to navigate through an easy to use android interface.

3.2.6 Software Process Model

The Hindi Shabdamitra Application is implemented using the incremental model as the application is divided in modules and so they can be easily managed. Each of the modules can be implemented and tested separately. Each of the subsequent release of the new module will add functionality to the previous release of the application. The separately developed modules are needed to be integrated.

As each of the module is loosely coupled and autonomous we can deliver the application first with limited functionalities and then later on can add the various functionalities on incremental basis.

Advantages:

- New features can be suggested with each release.
- Easy to implement and test the application as it is divided into a number of modules.
- Helps in creating the basic working application quickly.
- Re-usability is achieved as we have loosely coupled and autonomous modules.

3.2.7 Non-functional Requirements

- Ease of use: The Hindi Shabdamitra App should have a user friendly as well as a well-designed GUI for a better UX.
- Error Handling: The error handling mechanism of the application should be developed so that proper messages are displayed.

3.2.8 Installation Requirements

This application requires an android device with Android Version 5.0 or higher.

3.2.9 Functional Requirements

R1: Application should allow the user to select his level of Expertise from 5 available levels.

R1.1 Application should provide facility to select level.

Input: Desired Level

Output: Interface pertaining to selected level

R2: Application should allow the user to search words based on Varnamala by just clicking any single vowel/consonant from the keyboard.

R2.1 Application should provide facility to click any letter from keyboard.

Input: Desired Level

Output: Two most frequently used words starting with the selected letter.

R3: Application should allow the user to search words by typing in Hindi with the help of keyboard.

R3.1 Application should provide facility to search any word.

Input: Desired word

Output: Interface showing details related to searched word

R4: Application should allow the user to search words by typing in English using transliteration feature with the help of keyboard.

R4.1 Application should provide facility to search any word.

Input: Desired word

Output: Interface showing details related to searched word

R5: Application should allow the user to listen to audio pronunciation associated with a searched word.

R5.1 Application should provide facility to listen to audio.

Input: Click on audio button

Output: Audio

R6: Application should allow the user to navigate through the interface screen for viewing details along with an associated image for multiple senses of the word searched.

R6.1 Application should provide facility to view details pertaining to a word.

Input: Desired word

Output: Details of the searched word along with an image

R7: Application should allow the user to get appropriate error messages if the word searched by him is not found.

R7.1 Application should provide facility to receive appropriate error messages.

Input: Desired word

Output: Appropriate error message

R8: Application should allow the user to get access to user manual for guidelines to use the interface.

R8.1 Application should provide facility to receive appropriate guidelines.

Input: User Manual Request

Output: Guidelines

3.3 Analysis and Design

3.3.1 Database

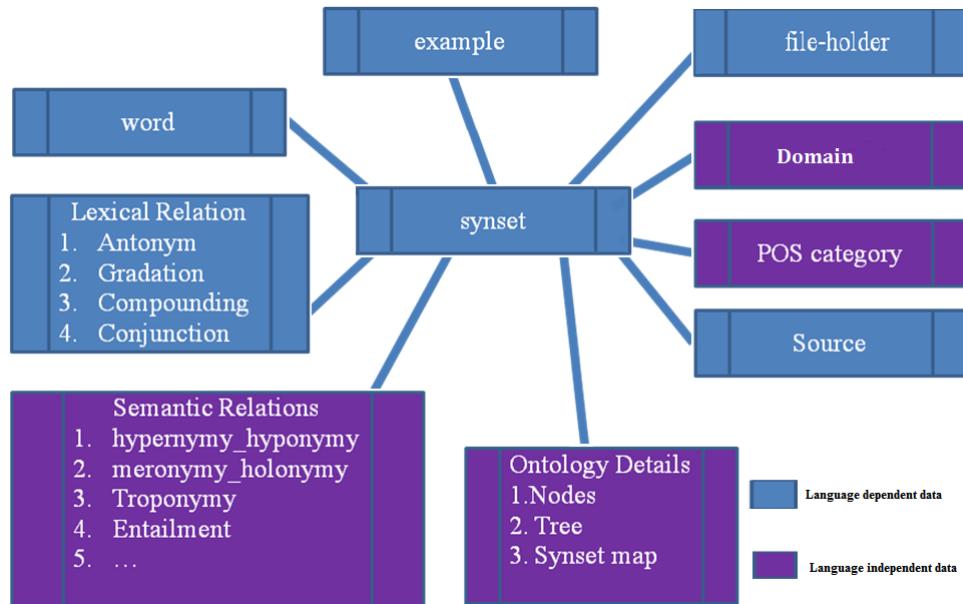


Figure 3.1 Database - Abstract

The tatacenter database stores the data for a project in two or more databases. Here tables from original IndoWordNet database structure are added/deleted/modified as per the requirement of the Tata Center Project. In this design, the common data such as semantic relations and ontology details, which is common to all languages are maintained in a master database called `tatacenter_master`, which is an extension of `wordnet_master` database from IndoWordNet database structure. The synset data, language specific data and lexical relations for Hindi language is maintained in a separate database called `tatacenter_hindi`.

Database 1

Name: `tatacenter_master`

Purpose: To maintain the data shared by all the languages. This database will keep tables having data which is common to all the languages. It includes tables having ontology details, semantic relations, and data common across all the languages.

Database 2

Name: tatacenter_hindi

Purpose: To maintain the data which is not shared by all the languages. This database keeps tables which will have information specific to the Hindi language. It will include tables to keep synset details, words in the language, examples, simplified gloss, grammatical information, etc.

Database 3

Name: tatacenter_admin

Purpose: In addition to the above mentioned databases, there can be another database to keep administration related information such as user details, user feedback, website administration, and statistics and so on.

Database 4

Name: tatacenter_english

Purpose: To maintain the English language data. This database will contain the princeton WordNet data. This database will have same database schema as wordnet_hindi database.

3.3.2 Use-case diagram

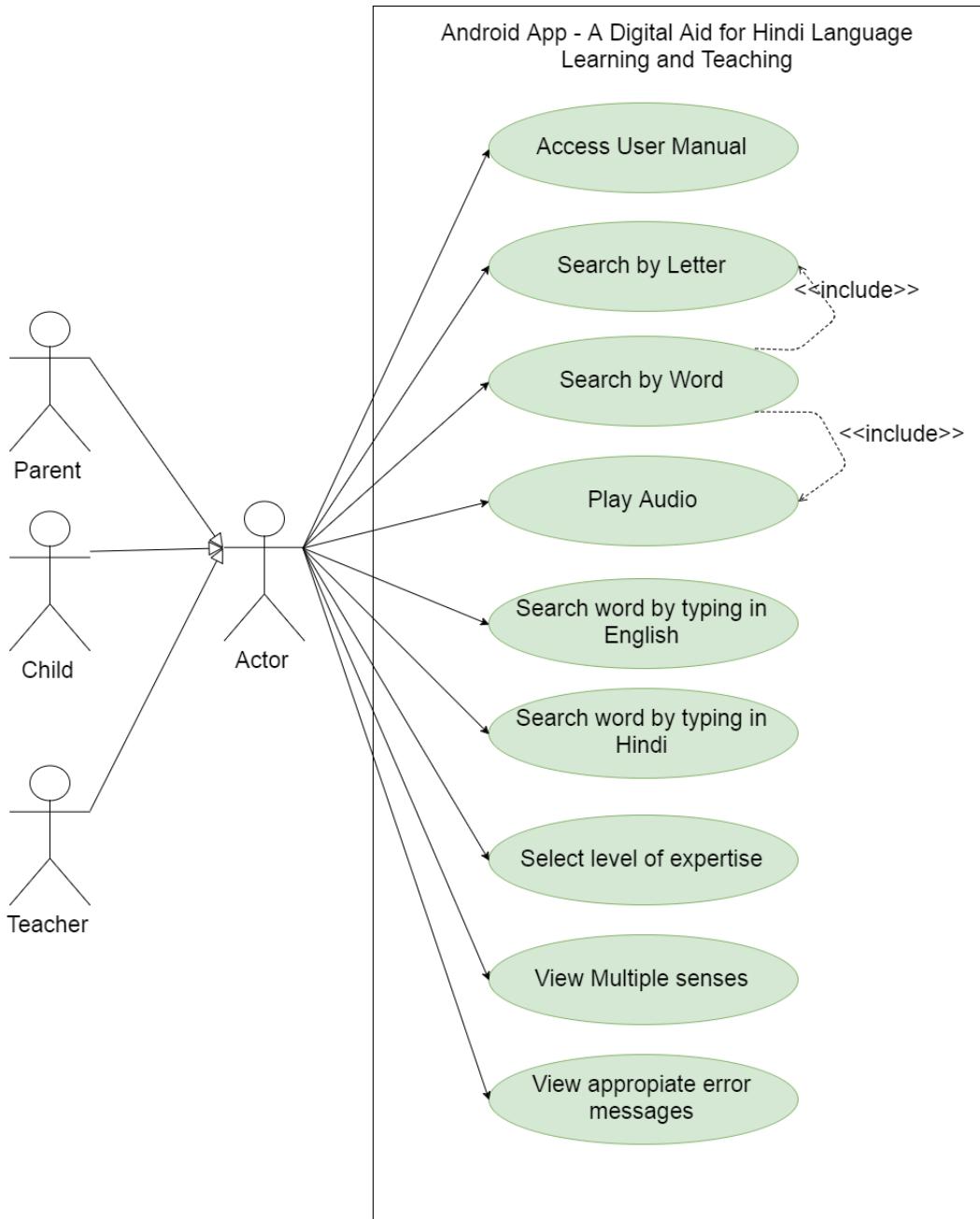


Figure 3.2 Use Case Diagram

3.3.3 Entity Relationship Diagram

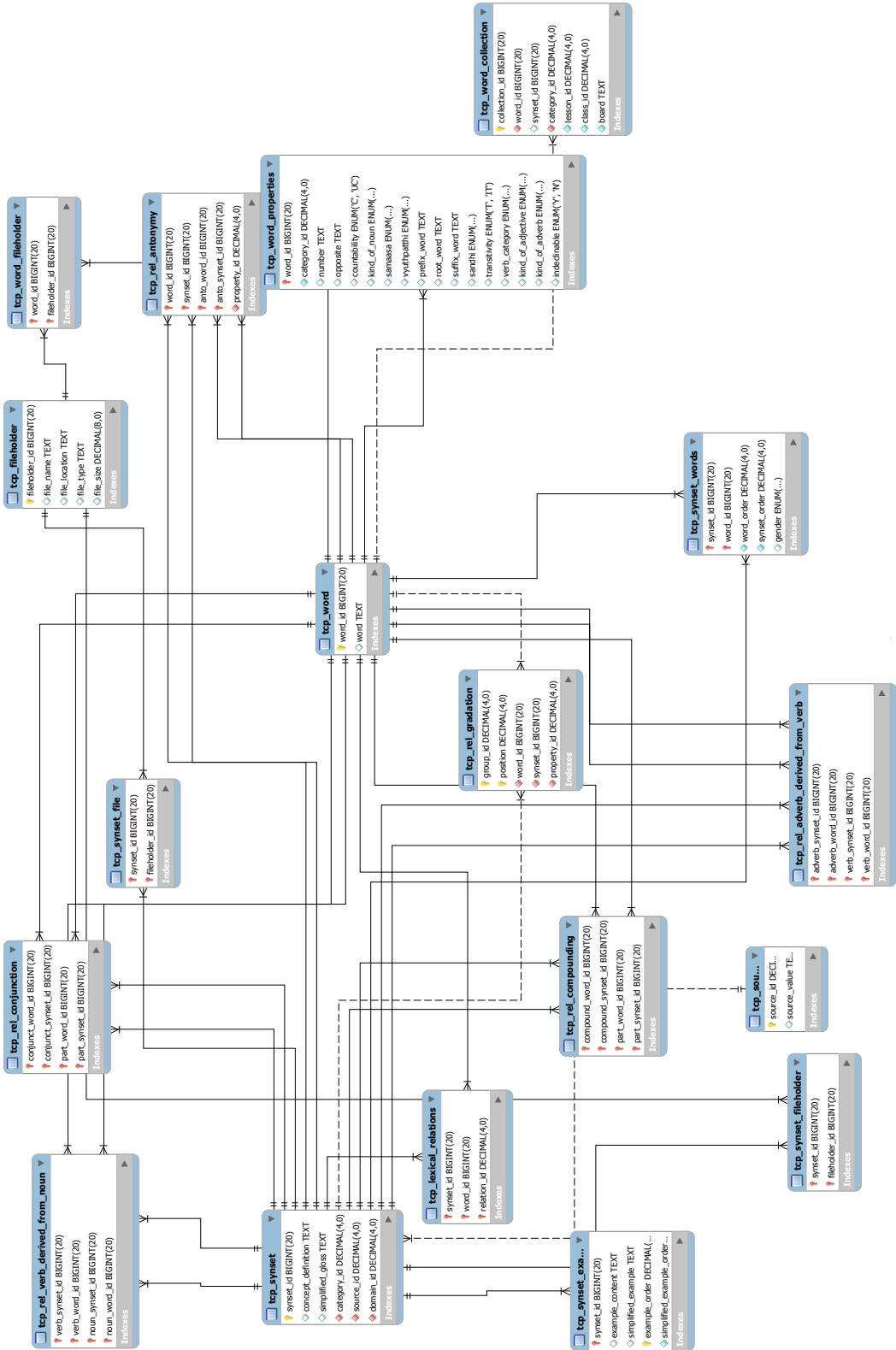


Figure 3.3 ER diagram for tatacenter_hindi database

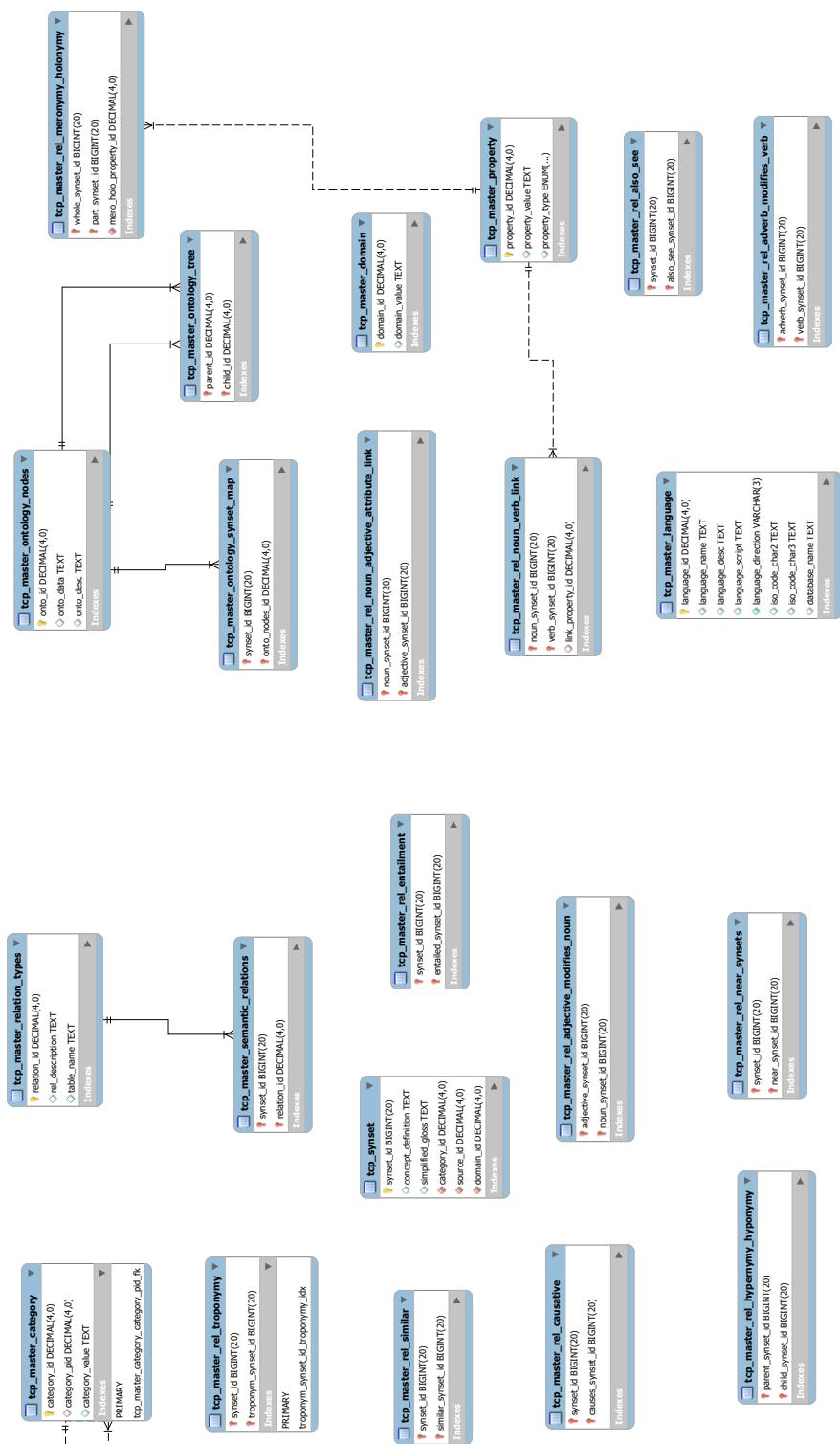


Figure 3.4 ER diagram for tatacenter_master database

3.3.4 Class diagram

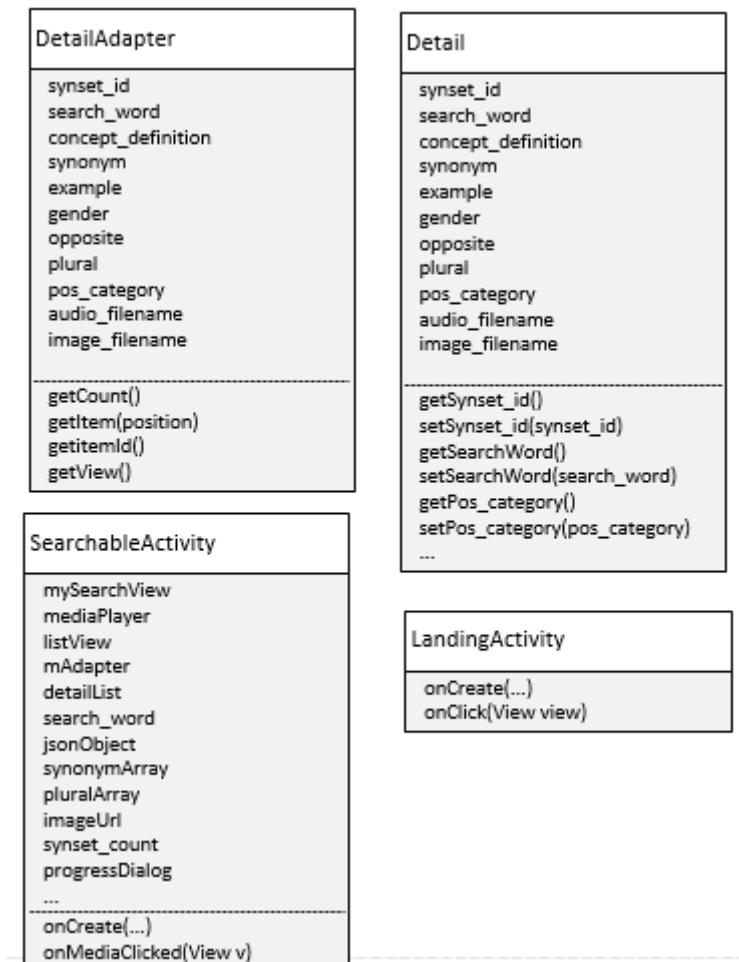


Figure 3.5 Class Diagram

3.3.5 Sequence Diagram

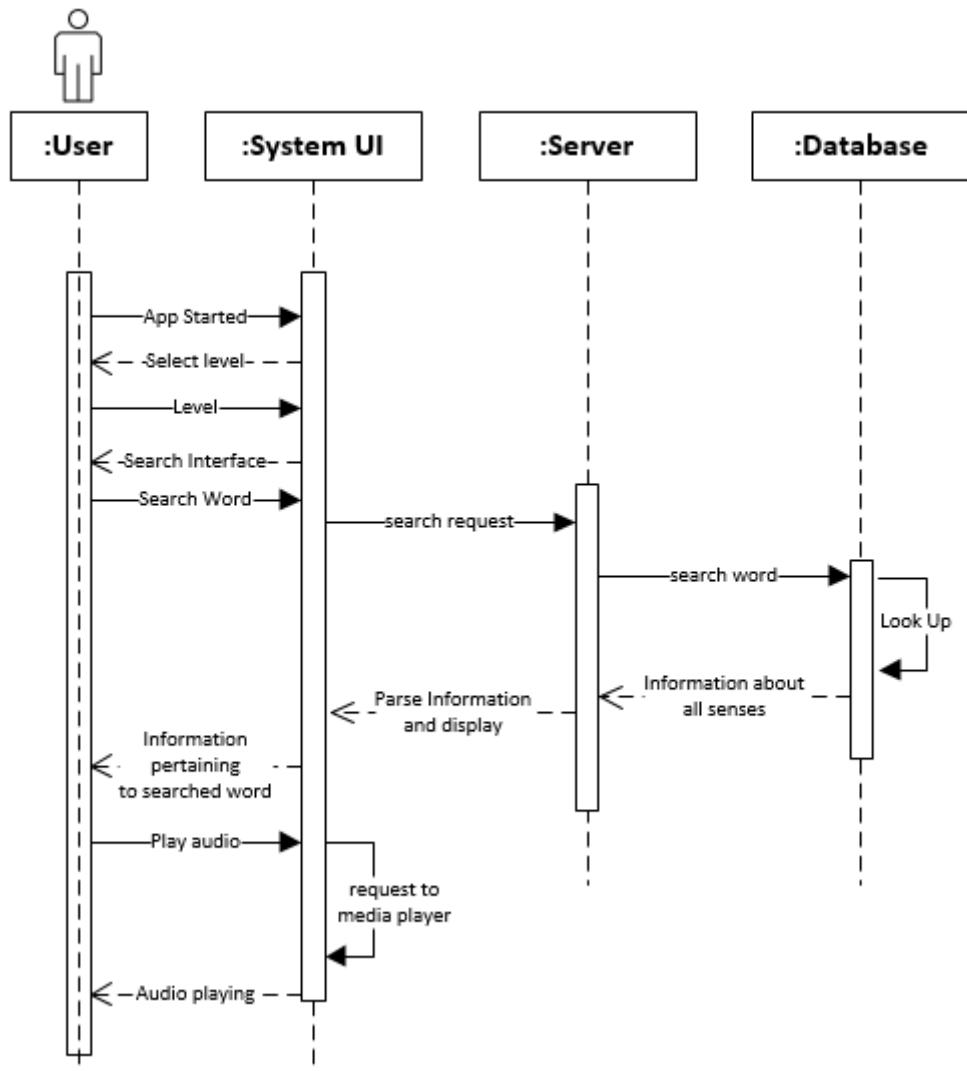


Figure 3.6 Sequence Diagram Word Found

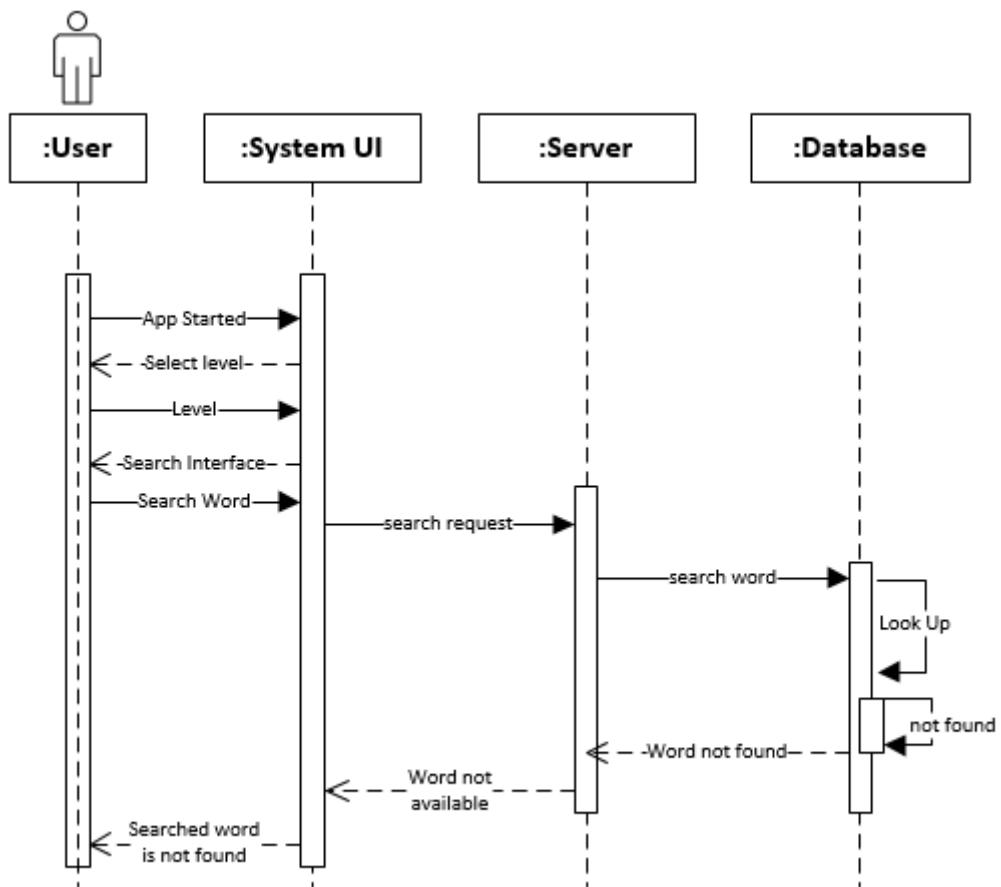


Figure 3.7 Sequence Diagram Word Not Found

3.3.6 Activity diagram

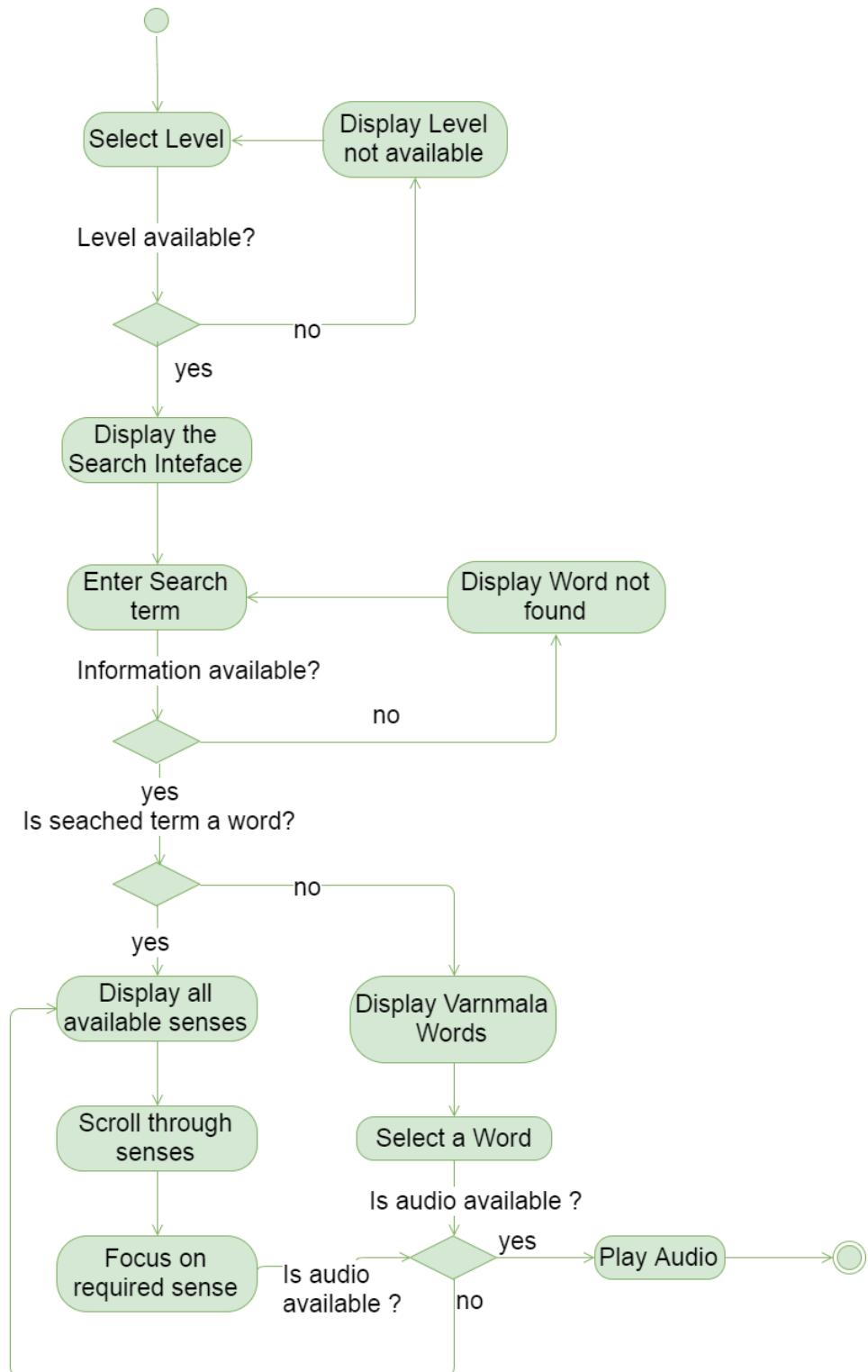


Figure 3.8 Activity Diagram

3.3.7 Implementation Details

There are two major divisions in implementation of an android application involving live databases. First one is Web Service and the other is developing the Android application.

Importance of Web Services in Android Applications Nowadays, existing Web applications are in a requirement of creating mobile applications to show their presence in mobile platform as well. Almost all web applications are having their Mobile applications created in Android, iOS or Windows platform. Exposing the existing functionalities of the applications is bit tough as all the functionalities have to re-implemented in the respective platforms. But this purpose can be easily fulfilled with much ease by creating Web Service and expose the existing functionalities as web methods to Mobile platforms.

There are few advantages of Web Services like:

- Re-usage of existing functionalities:

By designing the web service, we could also get significant benefits by reusing the existing functionalities by exposing them as web methods. This web services can be reused for other platforms as well.

- Remote DB hit made simple:

Databases residing remotely can also be hit from inside Android applications through Web Service calls. Making Web Service call from Android applications allows us to add functionality outside the scope of a DB like caching data, applying business rules over the data etc.

- Flexibility of Programming Language:

Web Services can be written in numerous programming languages like PHP, Java, Python, Ruby, Node.js and many more. This flexibility relieves the programmer from the overhead of expertizing multiple languages for developing different products.

Web Services for tatacenter project

In the android application for 'A Digital Aid for Language (Hindi) Teaching and Learning', there are two major functionalities:

1. Search by word

The user of the application searches word in the search bar and gets all relevant details based on the expertise level selected by the user.

2. Search by letter

The user of the application clicks any letter on the keyboard and two most frequently used words starting with the clicked letter appears. Clicking on any one of them shows relevant details of that word.

For implementation of these two functionalities, two web services **Search_word** and **Search_letter** are developed by us in PHP language.

Search_word web service

The basic idea of the project is to develop an interface which aids in classroom teaching environment. The motive is that the teacher teaches any chapter of their course curriculum and searches word in the aid that seems difficult to understand for students. We have a table `tcp_word_collection` in our database that has `word_id` and corresponding `synset_id` of the collected word. So the logic of the web service is designed such that the interface displays the relevant synset information for the word appearing in `word_collection` first and then the other synset information for that word is displayed. The input parameter to this web service is the searched word.

Following methods are implemented to get details for any word.

1. `getGloss(synset_id)`
2. `getExample(synset_id)`
3. `getSynonym(synset_id)`
4. `getCategory(synset_id)`
5. `getGender(synset_id,word_id)`
6. `getOpposite(synset_id,word_id)`
7. `getPlural(synset_id,word_id)`
8. `getImage(synset_id)`
9. `getAudio(word_id)`

10. getCountability(synset_id,word_id)

11. getOntologicalDetails(synset_id)

Description of each method

1. getGloss(synset_id)

This method returns gloss for any synset. For beginner level students of class one to five, it is agreed upon to have simpler gloss and examples for any concept. If simplified_gloss exists for any concept, then it is returned by this method otherwise concept_definition mentioned in Hindi Wordnet is returned.

2. getExample(synset_id)

This method returns example sentences for any synset. For beginner level students of class one to five, it is agreed upon to have simpler gloss and examples for any concept. If simplified_example exists for any concept, then it is returned by this method otherwise example_content mentioned in Hindi Wordnet is returned. For multiple examples, the example sentence having the searched word as a substring either in the exact form, plural form or any inflected form is returned for better discernment of students. If any of the example does not have any inflected form of the searched word as a substring, then all examples are returned. For instance, if the search word is समझ, following examples exists

- (a) लता की बुद्धि बहुत तेज है।
- (b) कुछ समझने के लिए हम अपनी अकल का उपयोग करते हैं।
- (c) लता की बुद्धि बहुत तेज है।

Out of these three examples, following examples are returned.

- (a) कुछ समझने के लिए हम अपनी अकल का उपयोग करते हैं।
- (b) लता की बुद्धि बहुत तेज है।

3. getSynonym(synset_id) This method returns set of synonyms for any synset. Five most frequently used words for that synset are returned. For instance, if the search word is समझ, few of synonyms are

अकल, दिमाग, प्रज्ञा, विवेक, धी, शक्ति, मति, मनीषा, मेधा, मस्तिष्क, बूझ, अकल, अकल, अकल, समझ, जिहन, ज़िहन, जेहन, ज़ेहन, जहन, ज़हन, संज्ञा, मनीषिका, प्राज्ञता, प्राज्ञत्व, अभिबुद्धि, आत्मसमुद्घवा, आत्मोद्घवा, इडा, प्रतिभान

Out of these, अकल, दिमाग, प्रज्ञा, विवेक, धी are returned.

4. getCategory(synset_id)

This method returns Part Of Speech Category for the searched word. For instance, if the search word is समझ, then NOUN is returned.

5. getGender(synset_id,word_id)

This method returns Gender for the searched word. For instance, if the search word is समझ, then F is returned.

6. getOpposite(synset_id,word_id)

This method returns Opposite for the searched word. For instance, if the search word is अच्छा, then बुरा is returned.

7. getPlural(synset_id,word_id)

This method returns plural for the searched word. For instance, if the search word is बिल्ली, then बिल्लियाँ is returned.

8. getImage(synset_id)

This method returns image filename for the searched word.

9. getAudio(word_id)

This method returns audio filename for the searched word.

10. getCountability(synset_id,word_id)

This method returns countability for the searched word. For instance, if the search word is छत्ते, then Countable is returned.

11. getOntologicalDetails(synset_id)

This method returns ontological details for the searched word.

Search_letter web service

The user of the application clicks any letter on the keyboard and two most frequently used words starting with the clicked letter appears. There is a table tcp_varnamala from which two words starting with the clicked letter are fetched and above methods are used to fetch details for them.

Android Application for Digital Aid

A web interface for Digital Aid is already being developed by the team at CFILT, IIT Bombay. But because of cell phones becoming ubiquitous, the need for developing an android app was felt and hence this app's development began.

For the prototype release, we were required to implement basic functionalities as already listed in the SRS document. The Android App makes use of PHP web services in order to communicate with the server.

The application provides a simple easy to use interface to search for meaning of words in hindi language. The first screen is designed using Buttons and Linear Layout where user needs to choose his level of proficiency after which user is taken to the search interface. The user can search for meaning of word using this interface.

Layout Files

1. activity_landing.xml
Contains the layout for first page of the android app, user has to select his level of expertise over here.
2. activity_searchable.xml
This layout contains the search interface for the user to search for meaning of words.
3. detail_item.xml
A word can have multiple meanings, and each of the possible sense is displayed in the interface. This layout files contains the structure for displaying information for a single sense. All senses of the words are displayed in a ListView in the activity_searchable layout.
4. alert.xml
This layout file contains the message to be displayed in case the searched word is not found in the database.
5. future.xml
As the application is under development here at CFILT, not all levels have been developed. This file contains the message to be displayed if the selected level is not developed yet.

Java Files

1. LandingActivity.java

Redirects the user as per his selected level of expertise.

2. SearchableActivity.java

This is the most important file for the application. When a Query word is typed, a request is made to the web service which is hosted on the web servers at IIT Bombay. The web service returns a JSON Object containing information about the searched word. This is because SearchView has a listener associated with it. Information is retrieved from the JSON object and displayed in the ListView .

3. Detail.java

This is a java bean. It contains variables to store all the information for a particular sense of word. It has appropriate constructor for initializing the instance of this class. This also contains getters and setter methods to get and set values of the variables, which contain information pertaining to sense of a word.

4. DetailAdapter.java

As mentioned earlier, a word may contain multiple senses and all needs to be shown to the user. To do this ListView has been used wherein each item will contain information pertaining to a single sense. And to bind this information to the objects of the list view an Adapter is required. The DetailAdapter class helps to bind the information and hence display the same in the interface.

Libraries used

1. Picasso

For displaying images of the searched concepts, first we tried using the normal way of setting the url of image on server to the image holder. But that lead to flickering of images because android tries to optimize performance. Picasso is a powerful image downloading and caching library for Android. It handles ImageView recycling, automatic memory and disk caching.

Following is the code snippet.

```
Picasso.with(context)
.load(url)
.resize(50, 50)
.centerCrop()
.into(imageView)
```

2. Volley

Volley is an HTTP library that makes networking for Android apps easier and most importantly, faster. Volley offers following benefits:

- (a) Automatic scheduling of network requests.
- (b) Multiple concurrent network connections.
- (c) Multiple concurrent network connections.
- (d) Ease of customization, for example, for retry and backoff.
- (e) Strong ordering that makes it easy to correctly populate your UI with data fetched asynchronously from the network.

3.3.8 Testing

Black Box testing

Black-box testing is a method of software testing that examines the functionality of an application without peering into its internal structures or workings. This method of test can be applied to virtually every level of software testing: unit, integration, system and acceptance. It typically comprises most if not all higher level testing, but can also dominate unit testing as well.

Table 3.1 Test Cases

Test Case Id	Test Scenario	Test Data	Expected Result	Actual Result
T01	Select expertise level	Beginner	Navigation to a new page with search interface	As expected
T02	Select expertise level	Advanced	Display message that the feature will be added soon	As expected
T03	Search word	কবন্না	Display error message that word is not found	As expected
T04	Search word	কবিতা	Display details of the word	As expected
T05	Search letter	অ	Display two most frequently used words starting with অ	As expected

3.3.9 Screen shots

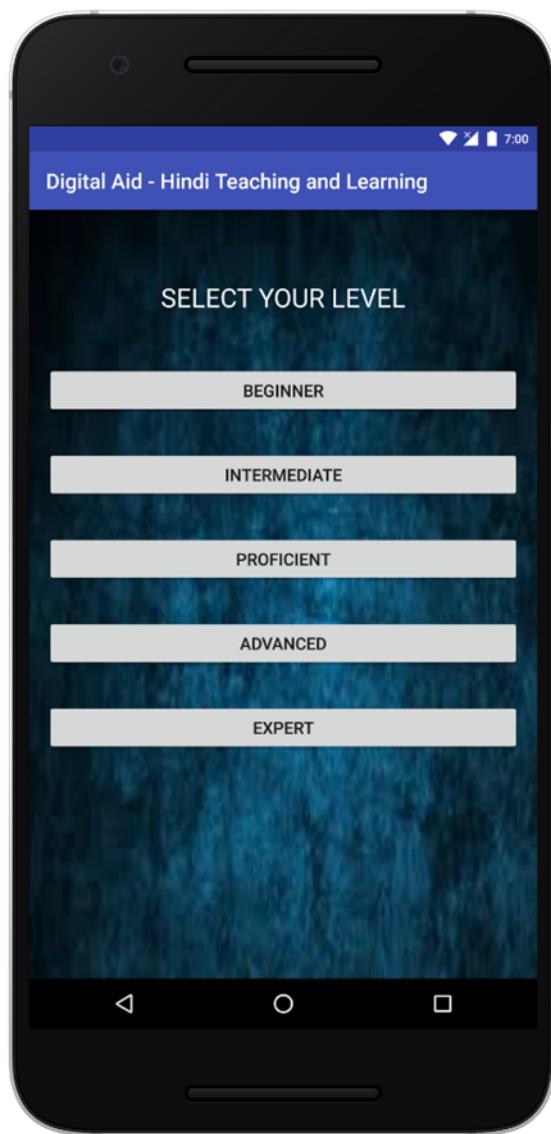


Figure 3.9 Select Expertise Level Screen



Figure 3.10 Search Interface Screen

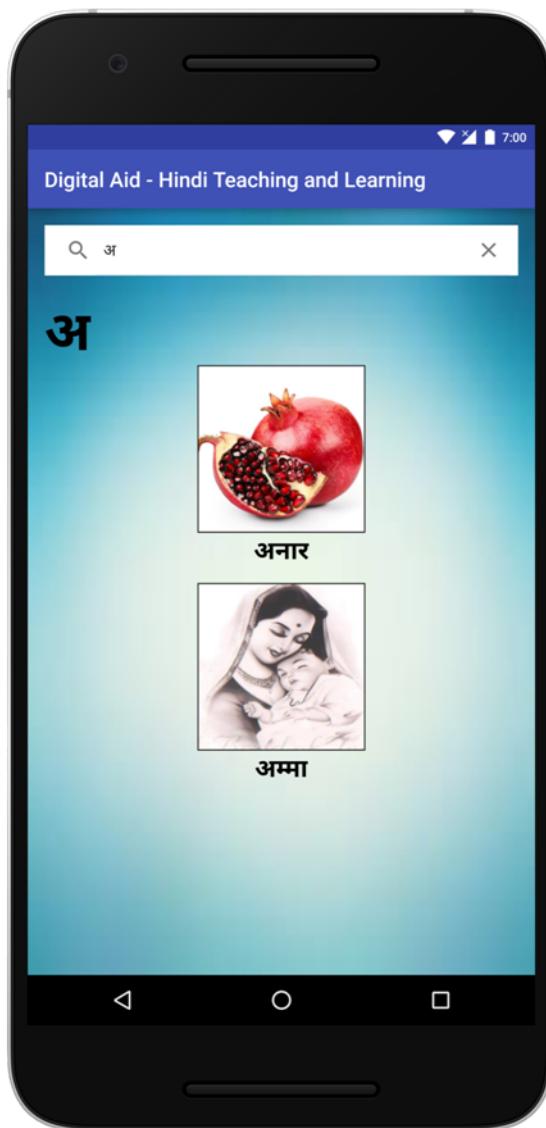


Figure 3.11 Varnamala search for letter

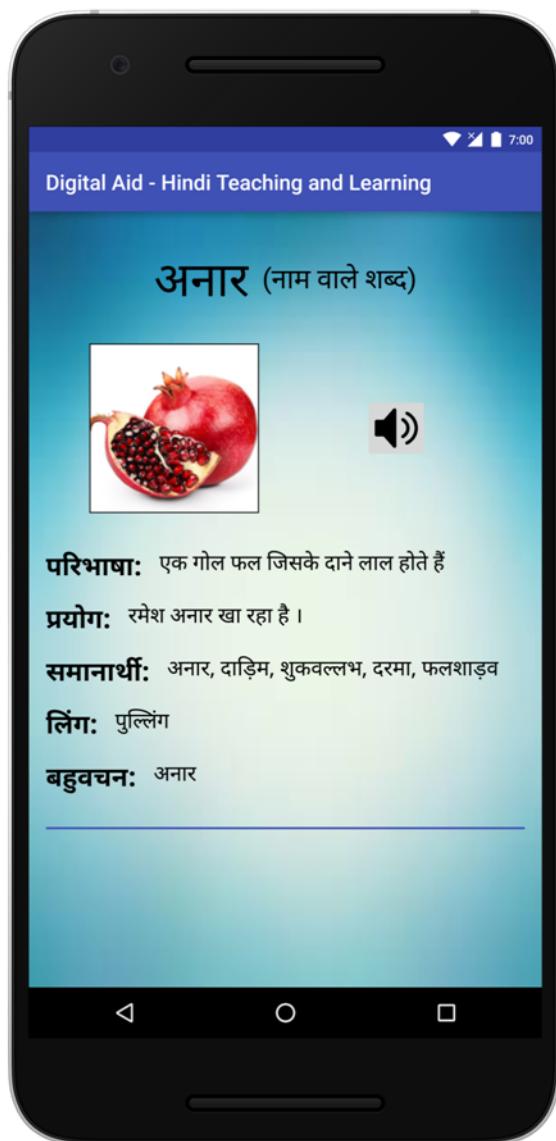


Figure 3.12 Display details for the varnamala search

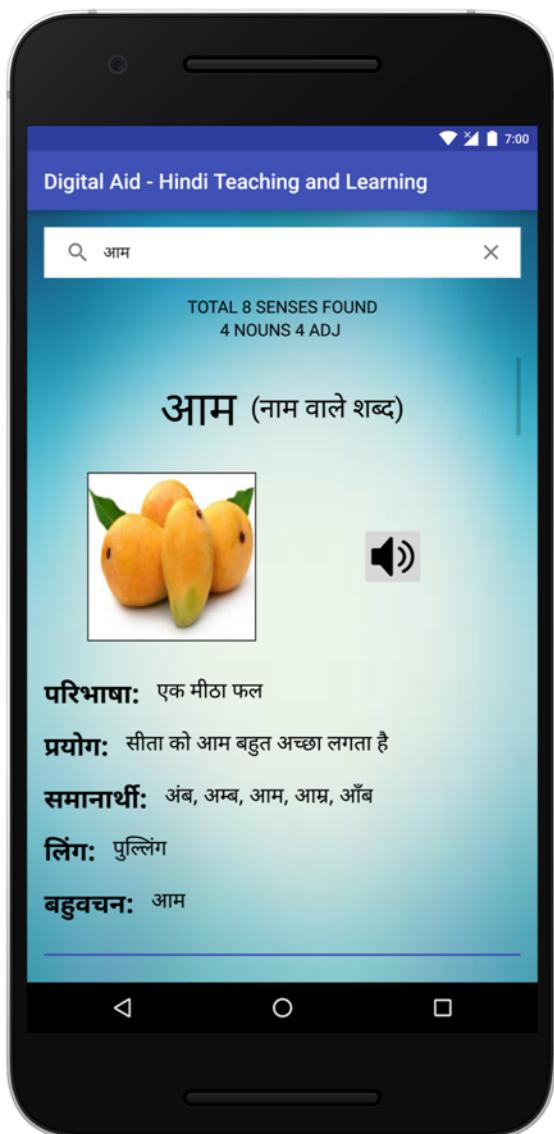


Figure 3.13 Display details for the the word search of Beginner level

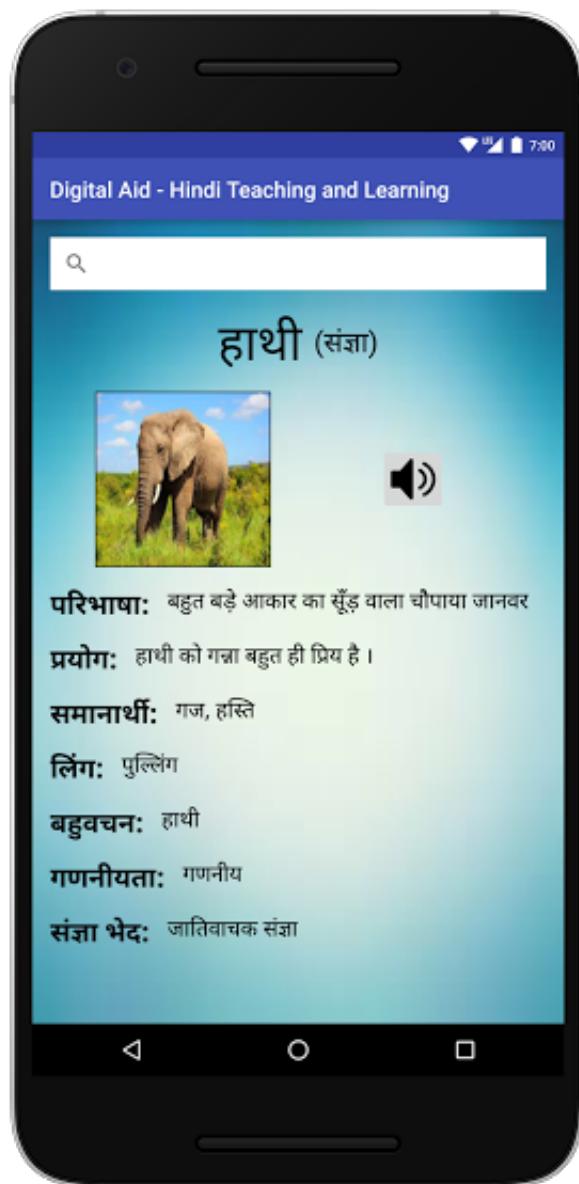


Figure 3.14 Display details for the the word search of Intermediate level

Chapter 4

Prototype Field Testing of Digital Aid at Schools

The value of our efforts and ideas can be worth if we can convince the people for whom we have developed something. After development of android application for Beginner Level students, our team from the project '**A Digital Aid for(Hindi) Language Teaching and Learning'** led by **Prof Malhar Kulkarni** from Department of Humanities and Social Science went to local **IIT Bombay Campus School** for our first field testing of the prototype on 8th February. The application was demonstrated to teacher and students of 5th standard students as their curriculum starts teaching Hindi language from class 5. Then we went to **IIT Bombay Kendriya Vidyalaya** for field testing of our prototype on 16th Feb and 17th February. The application was demonstrated to teacher and students of 1st standard on 16th February and students of 2nd standard on 17th February. We also went to **Kendriya Vidyalaya Bhandup** branch on 1st March. Students and teachers were quite receptive and they appreciated this approach of learning.

4.1 Response from the students

4.1.1 IIT Bombay Campus School

The Hindi class was not conducted the regular way- it had a supplementary aid, the Hindi WordNet. Narrating a new poem to the students on our soldiers, the Hindi teacher started her session. All eyes were glued to the screen for reading. The teacher explained each difficult word with its meaning, a reference image, its synonym, its usage with a sentence - all features included in the new interface initiated at Tata Centre.

There was an added element wherein the students could listen to the exact pronunciation of that specific word, which somehow raised the interest levels of the students. They were, in turn, asked to pronounce the difficult ones and repeat the meanings, as per their understanding. Pictures of freedom fighters were shown and while the students were asked to identify them, they enjoyed the interaction with the images in an otherwise regular Hindi class.

Apart from the content projected on the screen, a different application showcased the same on a mobile device developed by us. We volunteered to demonstrate the application and few students were using this to relate to what was being taught. They enjoyed learning through this method and mobile application was really an exciting experience for them.

Bringing languages closer

The project 'A digital aid for language and learning' in the Education domain has its first sights of the web interface and mobile application interface ready.

EDUCATION

This platform in Hindi language learning attempts to change the classroom environment. As a resource-assisted automated system, the Hindi WordNet data has been remodelled to suit the school teaching and learning environment. With a single platform to teach and learn the language, it will simultaneously serve as a dictionary and thesaurus, with audio-visual input.

"Developed and put to practice hardly anywhere in India, the core idea of Hindi WordNet is to establish semantic and lexical relations between words, concepts and parts of concepts. All this is available under one big umbrella that provides information and linkages between languages," says Prof Malhar Kulkarni, PI, from Humanities and Social Sciences.

The need for this project was first pointed out when the first sight of WordNet was still not ready to be used in an academic setup. Schools and educational institutions apparently wanted more applications with Word Net after its launch at events over the past two years in IIT Bombay.

The project team funded by Tata Centre has gone a few steps ahead by working on the web browser and as a mobile application. The design interface gets support from Co-PI, Prof Anirudha Joshi, IDC, while the audio aspect is being looked into by Co-PI, Prof Preethi Jyoti, Department of Computer Science and Engineering. The prototype in language learning holds its own with about 40,000 synsets. It is now being tested in local schools for Level 1- Classes 1 and 2.

The data of many languages and their transliteration are advantages that the Hindi WordNet platform speaks of. To make it useful to schools and have the teachers adapt to it quickly, games, interactive exercises and multiple question papers have been designed to evaluate the students with just a click.

Building such an automated system has had its share of challenges. Experts in different languages have not been an easy resource, and what's more, there are different sources giving diverse information on the same sets of words in the Hindi language. The aim is to utilize this tool to its optimum potentiality with the added features of sound, image and meaning. "Given the mandate, if this works well for the Hindi language, the model can be translated to inspire a pan-India activity that will bring the country's languages closer," says Prof Kulkarni.




- Gayathri Thakoor
Project Manager

Editor: Gayathri Thakoor
Design: Mrunal Phansalkar
Photo Credit: Romit Patil

Tata Centre Office, 2nd Floor, L2, Lecture Hall Complex(LHC),
Opp.KReSIT, IIT Bombay, Powai, Mumbai-400 076, India
www.tatacentre.iitb.ac.in

8 | TURNING POINT FEBRUARY 2017

Figure 4.1 IIT Bombay Campus School

4.1.2 IIT Bombay Kendriya Vidyalaya

We went to IIT Bombay Kendriya Vidyalaya on 16th February and 17th February for demonstration purpose. In the lecture, a poem from their textbook was taught to students with the help of the web interface and mobile application. We observed a remarkable transformation in learning pattern of students. They were found to be more disciplined and attentive than in regular teaching. This type of learning made them inclined towards language importance. Images and audio pronunciations helped them in a significant way.



Figure 4.2 IIT Bombay Kendriya Vidyalaya



Figure 4.3 IIT Bombay Kendriya Vidyalaya

4.2 Feedback from the teachers

The teachers of both the schools were receptive to this new approach of teaching. They praised this unique idea of using Hindi WordNet data in a different way for school education. Teachers were creative and open to accept the fact that traditional education system requires some remodeling to attract students. According to them, audio pronunciations and images helped a lot to clarify many abstract concepts which are generally difficult to explain to students. They appreciated the mobile application as it would also help the parents to teach their children at home. They were willing to use this aid in future for higher classes too.

Chapter 5

Database Synchronization

5.1 Introduction to Database Synchronization

Database Synchronization is the methodology of maintaining data consistency across two databases. Maintaining consistency includes reflecting all changes in the target database from the source database. There should be uniformity between two databases over time and this process should be repeated periodically. Following are various types of synchronization:

1. Insert Synchronization
2. Update Synchronization
3. Drop Synchronization
4. Mixed Synchronization

1. Insert Synchronization

In Insert Synchronization, new records from the source table are reflected to the target database. It is ensured that if data with unique primary key is not available in target database, new database entry is made for that.

The screenshot shows a MySQL Workbench interface. At the top, there is a SQL query editor window containing the following code:

```
SELECT * FROM `source_table`
```

Below the query editor are several control buttons: a checkbox for "Show all", a dropdown for "Number of rows" set to 25, and a search bar labeled "Filter rows: Search this table".

Underneath these controls is a table titled "source_table". The table has the following structure and data:

	+ Options	← T →	employee_id	name	designation	salary
<input type="checkbox"/>	Edit Copy Delete		1	Rahul	Application Engineer	52000
<input type="checkbox"/>	Edit Copy Delete		2	Shreya	Project Staff	39000
<input type="checkbox"/>	Edit Copy Delete		3	Hiren	Project Manager	90000

Figure 5.1 source_table

	<input type="checkbox"/>	<input checked="" type="checkbox"/> Edit	<input type="checkbox"/> Copy	<input type="checkbox"/> Delete	employee_id	name	designation	salary
	<input type="checkbox"/>	<input checked="" type="checkbox"/> Edit	<input type="checkbox"/> Copy	<input type="checkbox"/> Delete	1	Rahul	Application Engineer	52000
	<input type="checkbox"/>	<input checked="" type="checkbox"/> Edit	<input type="checkbox"/> Copy	<input type="checkbox"/> Delete	2	Shreya	Project Staff	39000

Figure 5.2 destination_table_beforesync

	<input type="checkbox"/>	<input checked="" type="checkbox"/> Edit	<input type="checkbox"/> Copy	<input type="checkbox"/> Delete	employee_id	name	designation	salary
	<input type="checkbox"/>	<input checked="" type="checkbox"/> Edit	<input type="checkbox"/> Copy	<input type="checkbox"/> Delete	1	Rahul	Application Engineer	52000
	<input type="checkbox"/>	<input checked="" type="checkbox"/> Edit	<input type="checkbox"/> Copy	<input type="checkbox"/> Delete	2	Shreya	Project Staff	39000
	<input type="checkbox"/>	<input checked="" type="checkbox"/> Edit	<input type="checkbox"/> Copy	<input type="checkbox"/> Delete	3	Hiren	Project Manager	90000

Figure 5.3 destination_table_aftersync

2. Update Synchronization

Whenever any modification in source database occur, appropriate changes in target database needs to be made. Synchronizer compares record's values at first. Then altered records are replaced in the destination tables in order to establish consistency between two tables. All data keeps updated in case of update synchronization.

3. Drop Synchronization

If some records have been deleted from source database, corresponding records have to be deleted from destination as well. All obsolete records will be removed from target database if they do not exist in source database.

4. Mixed Synchronization

If any change is made in source table like addition of completely new records, removal of obsolete rows or modifications of existing rows, in order to keep target database in relevance we need to add, delete and update these correspondent records in destination database.

5.1.1 Challenges of Database Synchronization

There are some challenges while dealing with Database Synchronization like:

- Similarity between structures of two databases. If their structures are dissimilar, database synchronization becomes more tedious and complex
- Type of synchronization - one way or two way
- Time duration at which database synchronization needs to be performed
- Choosing method to carry out database synchronization

Brief about Challenges

(a) Structure of two databases

Structure of two databases considered for synchronization purpose holds a great importance. If the schema of two databases are similar, replication or migration is a suitable approach. Replication will take new or modified data from a source database and copy it to a target database. If two systems have different schemas and different purposes, most of the standard answers like replication, log shipping and others are off the table.

(b) One way or two way synchronization

Complexity of database synchronization also varies based on whether synchronization is required one way or two way. If the schema of databases are same, replication works well, and if it is two way, it might be only viable option, since conflict resolution is built in. But if the synchronization is one way, triggers can be used. It has no built-in mechanisms for tracking changes so we will have to add fields to your tables to track changes. We can add a timestamp column to our source table and update the timestamp column whenever we change a row. This can be done in our SQL statement or through triggers. At sync time, the target database checks timestamp and if it is changed, then appropriate changes are reflected for those records. We can also use a flag column in the source table which sets the flag true across the records for which changes are done in the source table.

(c) Time duration for performing synchronization

Time is a major factor for data synchronization. The total time consumed for completion of sync and the time interval at which the process needs to be performed again are matters of concern. Complexity involved in completion of sync process should be efficient enough in terms of both time and efforts. If the sync process is required to schedule less frequently, then total time duration for execution can be little high. There should be balance between both the factors depending on the database type and business requirements.

(d) Selection of methods to do synchronization

There are manifold approaches to carry out data sync process.

- **Use a hosted system:**

This approach is not really data syncing. If all the users needing to use the database are on the same network, they can access the database server via Internet. We can use File system data sharing features for this purpose.

- **Do not make changes to more than one copy of your database**

This approach is also not really syncing. This is the most simplest way but limiting method. It implies that one of our database copies will always be the 'master' copy, and that all other copies should only be used in read-only mode. Offline database cannot be used for data entry or updates. This method is not viable if both the databases are live.

- **Update data in existing records during imports**

This approach is as simple as doing an import. When any modification is made to existing data during an import, data in the existing table is replaced with the data of the table which is imported. This approach is useful when we are certain that data from the same records has not been modified in both tables since the last time they were synced. The prime drawback of this method is that it lacks any concept of a 'winner' when a given tuple has been modified in both copies of database. The imported record data will always win over the data in the existing record that is being imported into.

- **Use a third party tool dedicated to database syncing**

Using a third party tool can greatly simplify the need of a full-fledged sync system. Much of the complex work is handled automatically. The disadvantage of this approach is that there is an additional expense of purchasing / licensing a third party sync tool. There is also an extra overhead on the developer's side to learn to use the tool and detecting its limitations based on the application requirements.

- **Creating system for synchronization**

A custom system can be completely optimized for our database solution and work environment. Creating own sync system is both technically and semantically complex. This method is likely to be the most complex one to implement. It demands professional developer-level technical skills. It also involves cost in terms of time and efforts required to write own scripts.

5.2 Syncing of tatacenter's database with Hindi WordNet

The project 'A digital aid for language (Hindi) Teaching and Learning' has a separate database based on Hindi WordNet database. It has different database design and schema compared to WordNet database. The database design is prepared which suits the requirements of the project.

5.2.1 Structure of Hindi WordNet database

Major tables of Hindi WordNet database

1. `tbl_all_gender`
2. `tbl_all_synset`
3. `tbl_all_words`

1. **`tbl_all_gender`**

Purpose:

To maintain the gender information for a word with respect to the concept / synset.

Table 5.1 `tbl_all_gender`

Sr. No.	Field Name	Data Type	Purpose
1	<code>synset_id</code>	<code>int(11)</code>	Foreign key from <code>tbl_all_synset</code> table.
2	<code>word</code>	<code>varchar(200)</code>	Foreign key from <code>tbl_all_words</code> table.
3	<code>gender</code>	<code>varchar(20)</code>	Stores gender information of a word w.r.t a synset

2. **tbl_all_synset**

Purpose:

To maintain the details of a synset (i.e a concept in a language). A synset (or concept) has a gloss, example sentences and synonym word set.

Table 5.2 tbl_all_synset

Sr. No.	Field Name	Data Type	Purpose
1	synset_id	int(11)	Uniquely identifies a concept/synset in the language
2	head	char(100)	Contains first word from synset blob file
3	synset	blob	Contains a blob file having words with same synset_id
4	gloss	blob	Contains a blob file having concept definition and examples

3. **tbl_all_words**

Purpose:

To maintain the unique words of the language. Holds the vocabulary of the language. It also contains mapping with the synset_ids.

Table 5.3 tbl_all_words

Sr. No.	Field Name	Data Type	Purpose
1	synset_id	int(11)	Foreign key from tbl_all_synset table
2	word	char(100)	The word of the language.
3	pos	char(10)	POS Category of the word for mapped synset specifying whether the concept is a noun, verb, adjective or adverb
4	sense_num	int(11)	Contains a blob file having concept definition and examples

5.2.2 Structure of tatacenter_hindi Database

Following are few major tables of tatacenter_hindi database:

1. tcp_synset
2. tcp_synset_example
3. tcp_word_collection

4. tcp_word
5. tcp_synset_words

1. **tcp_synset**

Purpose:

To maintain the details of a synset (i.e a concept in a language). A synset (or concept) has a gloss, example sentences and synonym word set.

Table 5.4 tcp_synset

Sr. No.	Field Name	Data Type	Purpose
1	synset_id	Bigint(20)	Uniquely identifies a concept/synset in the language
2	concept_definition	Text	The gloss / concept definition of a synset
3	simplified_gloss	Text	The simplified concept definition for level 1
4	category_id	Decimal(4,0)	Foreign key from tcp_master_category table. Specifying whether the concept is a noun, verb, adjective or adverb
5	source_id	Decimal(4,0)	Foreign key from tcp_source table. Specifies the source from where the concept is taken.
6	domain_id	Decimal(4,0)	Foreign key from tcp_master_domain table. Specifies the domain in which the synset belongs.

2. **tcp_synset_example**

Purpose: To maintain the example sentences for a concept / synset. A synset may have more than one example sentences. It also captures simplified example.

Table 5.5 tcp_synset_example

Sr. No.	Field Name	Data Type	Purpose
1	synset_id	Bigint(20)	Foreign key from tcp_synset table.
2	example_content	Text	The simplified example.
3	simplified_example	Text	The simplified example for level 1
4	example_order	Decimal(4,0)	Gives the order for the example sentences. The order in which the examples are to be displayed. Here 1 is used for highest order.
5	simplified_example_order	Decimal(4,0)	Gives the order for the simplified example sentences. The order in which the examples are to be used.

3. **tcp_word_collection**

Purpose: To maintain the example sentences for a concept / synset. A synset may have more than one example sentences. It also captures simplified example.

Table 5.6 tcp_word_collection

Sr. No.	Field Name	Data Type	Purpose
1	collection_id	Bigint(20)	Primary key of tcp_word_collection
2	word_id	Bigint(20)	Foreign key from tcp_word table. Points to a word.
3	synset_id	Bigint(20)	Provision to store a synset id belonging to a word. (Not a FK of tcp_synset table)
4	category_id	Decimal(4,0)	Foreign key to tcp_master_category. Points to a POS category
5	lesson_id	Decimal(4,0)	Lesson number of a class
6	class_id	Decimal(4,0)	Class number or division number in school
7	board	Text	Board of a textbook from which word is collected
8	page_no	Decimal(4,0)	Page no. of the word from where the word has been collected.

4. **tcp_word**

Purpose: To maintain the unique words of the language. Holds the vocabulary of the language.

Table 5.7 tcp_word

Sr. No.	Field Name	Data Type	Purpose
1	word_id	Bigint(20)	Uniquely identifies a word of the,language.
2	word	Text	The word of the language.

5. tcp_synset_words

Purpose: To maintain the synonymous words in a synset which are used to describe a concept in a language by maintaining the principle of coverage and minimality

Table 5.8 tcp_synset_words

Sr. No.	Field Name	Data Type	Purpose
1	synset_word_id	Bigint(20)	Unique id for identifying a synset id, word id combination
2	synset_id	Bigint(20)	Foreign key from tcp_synset table.
3	word_id	Bigint(20)	Foreign key from tcp_word table.
4	word_order	Decimal(4,0)	Gives the order for the synonymous words. Where the highest order is given to the most commonly used word for a concept. Here 1 is used for highest order
5	simplified_word_order	Decimal(4,0)	Gives the simplified word order for, the synonymous words.
6	synset_order	Decimal(4,0)	Gives the ranking of synset in which, they are ranked/to be ranked

5.2.3 Challenges faced during tatacenter database syncing

From the structure of Hindi WordNet database and tatacenter_hindi database, it can be inferred that they have disparate schemas. Both the databases are live and gets modified almost everyday. Only one way synchronization is required. Web interface tools are developed for linguists to make entries in both the databases. New synsets and new words which are in frequent usage of Hindi language and have a potential meaning are added into WordNet database along with many grammar features. The base of tatacenter database design is upon Hindi WordNet design but it is modified to suit the objective of the project. Some of the additional data are required apart from WordNet data. This project is primarily meant for classroom education and will aid teachers, students and parents to have an ease in Hindi language learning and teaching areas. For beginner level classes upto fifth standard, the meaning and examples related to any concept are simplified and those entries are added to tatacenter database. This task is carried out by language experts in weekdays and the tool is on throughout.

We faced some challenges while working on synchronization process. Following are some challenges:

1. Structure of databases

Structure of both databases differ, so replication or migration methods cannot be used. There should be some alternative approach to proceed with the task.

2. One way or two way synchronization

We require only one way synchronization from WordNet to tatacenter database. For one way synchronization, the source database should have a time stamp column or a flag column in each of its table. WordNet database does not have any such column and it is not possible to modify its structure by us. Its design and working is predefined by higher authority. Thus, using SQL triggers is not feasible to accomplish our task.

3. Time duration for performing synchronization

The requirement is to perform the sync process on every Friday night after working hours. The time required for completion of sync process depends on method used for it.

4. Selection of methods to do synchronization

This is the prime challenge faced by us. We had following alternatives:

- **Use a hosted system**

Using a hosted system is possible only if there is one live database and multiple users need access to it. This approach is not suitable in our case.

- **Do not make changes to more than one copy of database**

This method is simple but is applicable only if there is one live database and we want to restrict right of modification to a single copy. In our case, we have two live databases and we require read-write access for both. Moreover the structures are also different.

- **Update data in existing records during imports**

This method is partially useful in our sync process. Apart from some major tables described in structure of both databases, there are around 20 tables in tatacenter database that can be synced with Hindi WordNet database with the aid of this approach. There are some tables in Hindi WordNet for storing lexical, semantic and ontological relations between concepts. Data in these tables are not so frequently updated and their schema also matches with tatacenter database. So this method can be useful.

- **Use a third party tool dedicated to database syncing**

This method is not attractive due to involvement of additional expense and efforts to learn the tool. Moreover, it is practically not possible to get a fully fledged tool that satisfy all our requirements.

- **Creating system for synchronization**

This method is the most appropriate one as it gives us the liberty to develop custom features and can take care of project specific requirements.

From the above mentioned description, we finally decided to develop scripts to sync major tables and use import approach for other tables which are not frequently updated in source database.

5.2.4 Implementation details

The database synchronization mainly involves four tables from tatacenter_hindi database and three tables from Hindi WordNet database. They are

1. tbl_all_synset
2. tbl_all_words
3. tbl_all_gender
4. tcp_synset
5. tcp_synset_example
6. tcp_word
7. tcp_synset_words

Mixed synchronization is required so there is a need to handle all cases - insertion, modification/updation and deletion of records.

1. **tcp_word**

This table stores word along with a corresponding word id. The Hindi WordNet is expanding everyday, new words are added to it, also new words which are collected by word collectors needs to go in tcp_word table. A live tool has already been developed for the word collectors to add new words on the fly. But to add the newly added words of Hindi WordNet synchronization is required.

If there were a word_id field in the Hindi WordNet database, the task would have been simpler, but as WordNet being based upon solely on synonymous sets and relations, such field does not exists in the Hindi WordNet database.

- Insert sync

To find out the new words, we grouped the words of Hindi WordNet into sets 1000 synsets each, and compared which words was not already present in the tata center database and consecutively, these new words are inserted into the tata center database and a new word_id is assigned to it. We used the following query to achieve the same.

```
SELECT word  
FROM tbl_all_words  
WHERE word  
NOT IN ( SELECT word FROM tcp_word)  
AND tbl_all_words.synset_id BETWEEN lower_limit AND upper_limit
```

2. tcp_synset

For tcp_synset table, synset_id, concept_definition and category_id are the fields that gets affected due to Hindi WordNet.

- Insert sync

synset_id is the primary key in both tcp_synset and tbl_all_synset tables. But in tbl_all_synset, they are not added in incremental order. Some of the synsets which are not valid synsets are deleted. So new synsets also gets inserted at those empty positions. We used following query to find synset_ids which are newly added.

```
SELECT synset_id
FROM tbl_all_synset
WHERE synset_id NOT IN
(SELECT synset_id FROM tcp_synset)
```

To get concept_definition and category_id, gloss and category fields of tbl_all_synset are used. gloss field is a blob file which has data in format-
concept_definition:"example1\example2...\examplen"

We used string tokenizer to extract concept_definition and example_content.

```
token = strtok(string, ":" )  
concept_def_wordnet=token
```

Category field of tbl_all_synset gives category value. From this category value, corresponding category_id is obtained from tcp_master_category and added to tcp_synset table.

- Update sync

If there is any modification in concept_definition and category_id of any existing synset, it is handled.

3. **tcp_synset_example**

- Insert sync

We used the same query to find synset_ids which are newly added.

```
SELECT synset_id
FROM tbl_all_synset
WHERE synset_id NOT IN
(SELECT synset_id FROM tcp_synset)
```

To get example_content and example_order, gloss field of tbl_all_synset is used. gloss field is a blob file which has data in format-concept_definition:"example1\example2...\examplen"

We used string tokenizer to get concept_definition.

```
token = strtok ( string , ":" )  
concept_def_wordnet=token
```

To get example_content, we removed quotes, extracted multiple examples and kept track of count for getting example_order.

```
while ( token )  
    token = strtok ( "/" )  
    if ( token )  
        token=trim ( token , "\'"')  
    _____end_if  
end_while
```

- **Update sync**

First of all, number of examples in tcp_synset_example and tbl_all_synset are compared. If the examples have same count, then example_content of tbl_all_synset and tcp_synset_example are compared. If there is any change, we used the following query:

```
UPDATE tcp_synset_example
SET example_content=example_content_wordnet[current],
example_order=example_order_wordnet[current]
WHERE synset_id=synset
AND example_order = example_order_synset[current];
```

If number of examples in `tbl_all_synset` is more than in `tcp_synset_example`, then new example entry is added in `tcp_synset_example`. We used the following query:

```
INSERT into tcp_synset_example (synset_id,example_content,example_order)
values(synset_id, example_content_wordnet, example_order_wordnet)
```

If number of examples in `tbl_all_synset` is less than in `tcp_synset_example`, then old example entry is replaced by `NULL` in `tcp_synset_example`. We used the following query:

```
UPDATE tcp_synset_example
SET example_content=null,example_order=0
WHERE synset_id=synset
AND example_order=example_order_synset[current];
```

4. `tcp_synset_words`

- Insert sync

`tcp_synset_words` table is meant for storing the mapping of words with their corresponding synsets, gender information and `word_order` associated with that word. For each synset/concept, multiple words are associated. All the words are ranked according to the frequency of their use. This ranking is `word_order`.

In `tbl_all_synset` table, `synset` field is a blob file which has data in format-
`word1,word2,word3,....,wordn`

Whenever any new word is added in `tcp_word` after sync process, mapping should be done in `tcp_synset_words` table. For each newly added word, corresponding `synset_id` is obtained from `tbl_all_words`. For that `synset_id`, newly added word is searched in the `synset` blob field and its order is noted. This order is stored in `word_order` field of `tcp_synset_word` table. For the newly added word and its corresponding `synset_id` there is a gender field in `tbl_all_gender` table which is stored in `gender` field of `tcp_synset_words` table.

```
INSERT into tcp_synset_words
(synset_id, word_id, gender, word_order)
values (synset_id_for_new_word, new_word_id, gender, word_order)
```

Whenever any new synset is added in tcp_synset after sync process, mapping should be done in tcp_synset_words table. For each newly added synset, multiple words corresponding to that synset can be obtained from tbl_all_words table. For each pair of word and synset, if an entry is already present in tcp_synset_words table, nothing is done. Otherwise, a new entry is added to tcp_synset_words table. For that new synset_id, word not having mapping in tcp_synset_words is searched in the synset blob field and its order is noted. This order is stored in word_order field of tcp_synset_word table. For the newly added word and its corresponding synset_id there is a gender field in tbl_all_gender table which is stored in gender field of tcp_synset_words table.

```
INSERT into tcp_synset_words
(synset_id, word_id, gender, word_order)
values (new_synset_id, word_id_for_new_synset, gender, word_order)
```

- Update sync The tcp_synset_words table contains mappings of words with their corresponding synsets, gender information and word_order associated with that word. The word_order is determined according to frequency of usage.

Due to the dynamics of natural languages it is quite possible that these order is modified in the WordNet. Hence to reflect this changes, we used following query to get the task done. After checking whether the synset order is same or not, following query is executed.

```
UPDATE tcp_synset_words
SET word_order = "word_order"
WHERE synset_id = "synset_id"
AND word_id = "word_id"
```

Surprisingly, same thing applies for gender of an word belonging to a synset too. Upon discussion with language experts, we came to know that the gender is also modified if required. So to sync this changes with the tatacenter_hindi database, we used the following query when the gender values were different.

```
UPDATE tcp_synset_words
SET gender = "gender"
```

```
WHERE synset_id="synset_id"
AND word_id="word_id"
```

5.2.5 Cron Job for scheduling synchronization

Cron is a time-based job scheduler in Unix based operating systems. It is used by people who maintain software environment, databases etc to schedule jobs(commands or shell scripts) to run periodically at fixed times, dates or intervals. It automates the process of system maintenance and administration. For the task of database synchronization we scheduled a cron job that automated the process, executing the sync code every week.

Cron Job Schedule

The schedule for the cron job is created in a text file which is known as the crontab file. The crontab file contains list of jobs and other instructions for the cron daemon to execute.

Example of a crontab schedule

```
m h dom mon dow command_to_execute
m : (0 -59)
h : (0 - 23)
dom : (1 - 31)
mon : (1 - 12)
dow : (0 - 6) Sunday to Saturday; 7 is also sunday
```

To define the time you can provide concrete values for minute (m), hour (h), day of month (dom), month (mon), and day of week (dow) or use '*' in these fields (for 'any'). Note that tasks will be started based on the cron's system daemon's notion of time and timezones.

Schedule for Synchronizing Tata center's database

For the project, we need to synchronize the database every week with Hindi WordNet. The syncing scripts execute at 21:55 every friday. The schedule for same as in the crontab file is as follows.

```
55 21 * * 5 curl www.cfilt.iitb.ac.in/~tatacenterproject/
dbsync/home.php > log_ddmmyy.txt
```

To schedule this job, use the crontab command.

```
crontab crontab.txt
```

5.2.6 Screenshots

```
tatacenterproject@tukaram:~$ cd public_html/
tatacenterproject@tukaram:~/public_html$ cd dbsync/
tatacenterproject@tukaram:~/public_html/dbsync$ crontab crontab.txt
tatacenterproject@tukaram:~/public_html/dbsync$ crontab -l
MAILTO="cfilt_admin@gmail.com"
55 23 * * 5 curl www.cfilt.iitb.ac.in/~tatacenterproject/dbsync/home.php
tatacenterproject@tukaram:~/public_html/dbsync$
```

Figure 5.4 cron_job_scheduling

SELECT *
FROM `sync_details`
LIMIT 0 , 30

The screenshot shows a MySQL database table named 'sync_details'. The table has columns: sync_id, filename, message, data, and time_stamp. The data is ordered by sync_id. The table contains 18 rows of log entries. The first few rows show initial files being processed ('initial.php'). Subsequent rows show various stages of the sync process, including 'tcp_word_sync_1.php' and multiple entries for 'tcp_synset_sync_file1.php' through 'tcp_synset_sync_file3.php'. The log ends with 'home.php' and 'SYNC END'.

	sync_id	filename	message	data	time_stamp
<input type="checkbox"/>	1	home.php	SYNC START		2017/03/15 00:27:01
<input type="checkbox"/>	2	initial.php	SYNC IN PROGRESS		2017/03/15 00:27:01
<input type="checkbox"/>	3	initial.php	Last Word ID	104773	2017/03/15 00:27:01
<input type="checkbox"/>	4	initial.php	SYNC IN PROGRESS		2017/03/15 00:27:02
<input type="checkbox"/>	5	tcp_word_sync_1.php	SYNC IN PROGRESS		2017/03/15 00:27:02
<input type="checkbox"/>	6	tcp_word_sync_1.php	SYNC IN PROGRESS		2017/03/15 01:51:16
<input type="checkbox"/>	7	tcp_synset_sync_file1.php	SYNC IN PROGRESS		2017/03/15 01:51:16
<input type="checkbox"/>	8	tcp_synset_sync_file1.php	NOT IN Query done	9	2017/03/15 01:51:16
<input type="checkbox"/>	9	tcp_synset_sync_file1.php	SYNC IN PROGRESS		2017/03/15 01:51:23
<input type="checkbox"/>	10	tcp_synset_sync_file2.php	SYNC IN PROGRESS		2017/03/15 01:51:23
<input type="checkbox"/>	11	tcp_synset_sync_file2.php	SYNC IN PROGRESS		2017/03/15 02:30:11
<input type="checkbox"/>	12	tcp_synset_sync_file3.php	SYNC IN PROGRESS		2017/03/15 02:30:11
<input type="checkbox"/>	13	tcp_synset_sync_file3.php	SYNC IN PROGRESS		2017/03/15 03:07:41
<input type="checkbox"/>	14	tcp_synset_words_sync_1.php	SYNC IN PROGRESS		2017/03/15 03:07:41
<input type="checkbox"/>	15	tcp_synset_words_sync_1.php	SYNC IN PROGRESS		2017/03/15 04:12:06
<input type="checkbox"/>	16	tcp_remaining_words_mapping.php	SYNC IN PROGRESS		2017/03/15 04:12:06
<input type="checkbox"/>	17	tcp_remaining_words_mapping.php	SYNC IN PROGRESS		2017/03/15 04:12:09
<input type="checkbox"/>	18	home.php	SYNC END		2017/03/15 04:12:09

Figure 5.5 cron_job_logging_details

Chapter 6

Conclusion and Future Extension

6.1 Conclusion

Hereby we conclude that we developed the project by understanding all functionalities of this project. We checked the feasibility and requirement for this System. Then we defined overall look and flow of control among modules in paper. All the functionalities of the android application are developed separately. Then all of them were integrated in the main application. After coding and integration of all the functionalities, we tested all functionalities separately which is basically Unit Testing of all modules after completion of Unit Testing, whole system was then tested once again (i.e. Integration Testing). Thus, the android application allows the user to learn Hindi language with the aid of illustrations, audios, grammatical features for Beginner and Intermediate level students.

We have completed development of a database synchronization subsystem by developing scripts. We have scheduled time based cron job which runs these scripts every weekend to perform synchronization process. We have tested the subsystem and works perfectly.

6.2 Future Extension

- Completion of Android application development for proficient, advanced and expert level of expertise.
- Completion of ongoing theme based Hindi Shabdamitra visualizer

Bibliography

- [1] Android developers. <https://developer.android.com/develop/index.html>.
- [2] Picasso. <http://square.github.io/picasso/>.
- [3] StackOverFlow Syncing Databases. <http://stackoverflow.com/questions/3234974/synchronizing-2-database-with-different-schemas>, .
- [4] Syncing Approaches. [http://help.filemaker.com/app/answers/detail/a_id/7720/~database-synchronization---an-overview-of-approaches\(mostimp\)](http://help.filemaker.com/app/answers/detail/a_id/7720/~database-synchronization---an-overview-of-approaches(mostimp)), .
- [5] Android Volley Intro. <https://developer.android.com/training/volley/index.html>, .
- [6] Android Volley Simple Request. <https://developer.android.com/training/volley/simple.html>, .
- [7] Android Volley Standard Request. <https://developer.android.com/training/volley/request.html>, .
- [8] Pushpak Bhattacharyya. Indowordnet. In *The WordNet in Indian Languages*, pages 1–18. Springer, 2017.
- [9] George A Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J Miller. Introduction to wordnet: An on-line lexical database. *International journal of lexicography*, 3(4):235–244, 1990.
- [10] Venkatesh Prabhu, Shilpa Desai, Hanumant Redkar, NR Prabhugaonkar, Apurva Nagvenkar, and Ramdas Karmali. An effcient database design for indowordnet development using hybrid approach. 2012.