A Digital Aid for Hindi Language Teaching and Learning

Shivani Thakkar
Yash Pradhan

Internal Guide: Dr. C. K. Bhensdadia
External Guide: Prof. Pushpak Bhattacharyya

April 15, 2017

Outline

- Introduction
- WordNet, Hindi Wordnet, IndoWordNet
- Android Application Development
- Database Synchronization
- Ongoing: Hindi Shabdamitra Visualizer
- Project Prototype Demonstrations
- Summary and Future work

About Project

- Worked on the project titled 'A Digital Aid for Hindi Language Teaching and Learning'
- Sponsored by Tata Center for Technology and Design, IIT Bombay
- Executed at CFILT Research Lab

Introduction

- Challenges in Education
  - Rapid digitization
  - Better learner engagement in classroom
  - The burden of the school bag
- Challenges in Language Learning
  - Learning vocabulary and grammar features
  - Understanding of meaning, usage and pronunciation
  - Understanding the related concepts
  - Visualizing the concepts

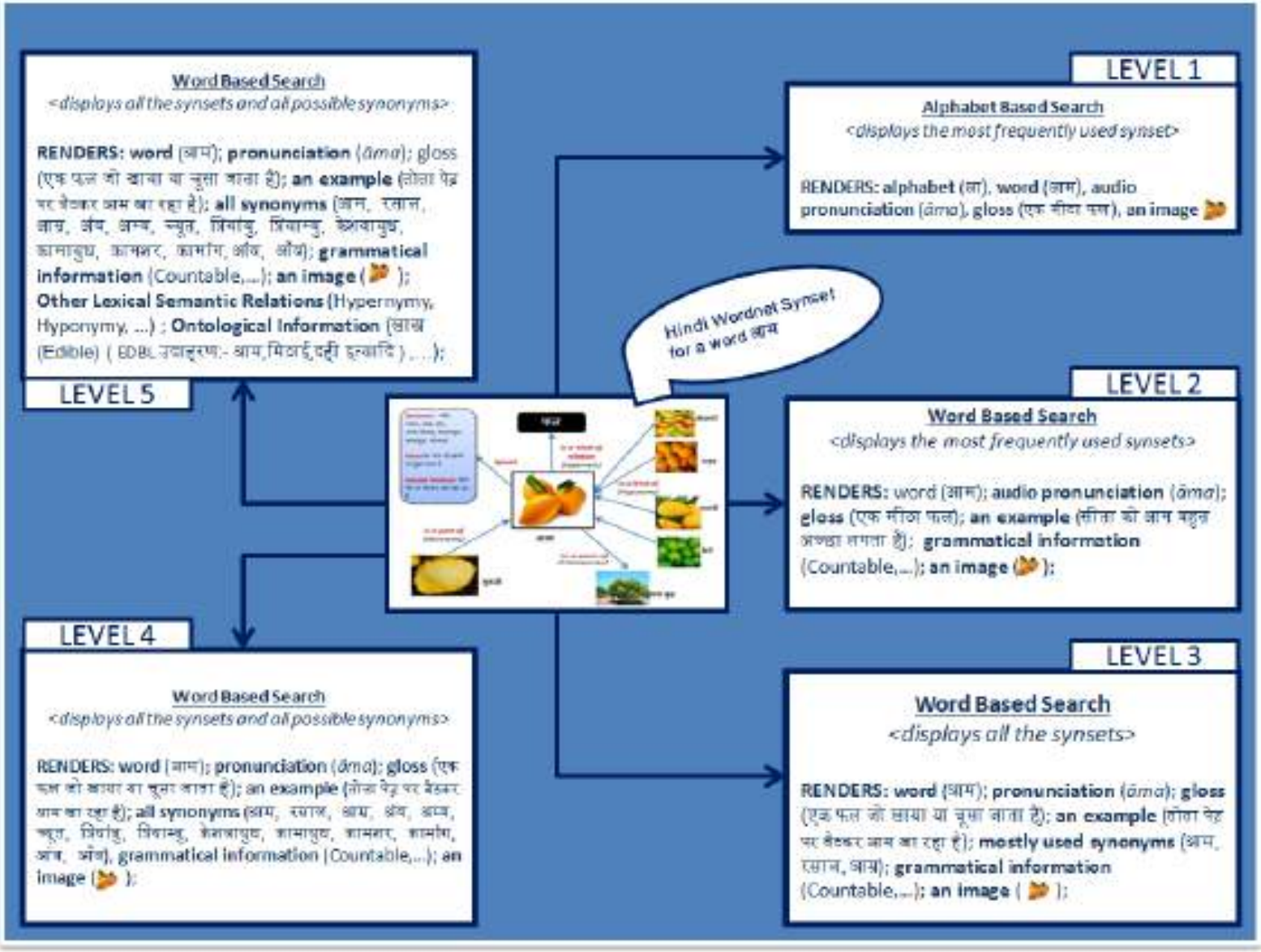
Introduction

Solution

A Digital Aid for Hindi Language Teaching and Learning: Hindi Shabdamitra

Using **Hindi Wordnet** as a resource

Hindi Shabdamitra - Project at a Glance



WordNet (English WordNet[1])

- WordNet developed at Princeton University, is an online lexical resource based on psycholinguistic principles.
- Words from the same lexical category that are synonymous are grouped into synsets.
- WordNet has different lexical relations like antonymy, gradation, etc. and semantic relations like hypernymy (is-a), meronymy (part of), etc.

[1] http://wordnetweb.princeton.edu/perl/webwn.

Hindi Wordnet

Hindi Wordnet[2]

- Hindi Wordnet inspired by Princeton WordNet is an online lexical resource created at CFILT Lab, IIT Bombay and was released in 2010.

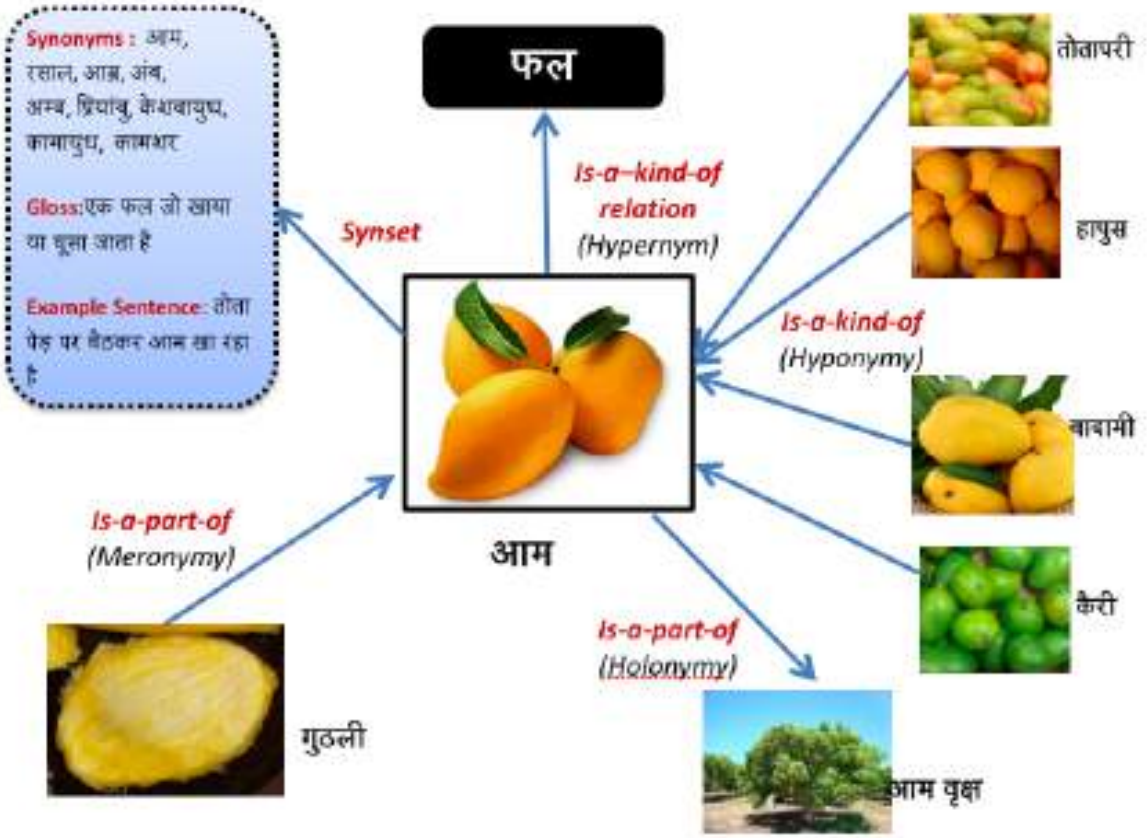


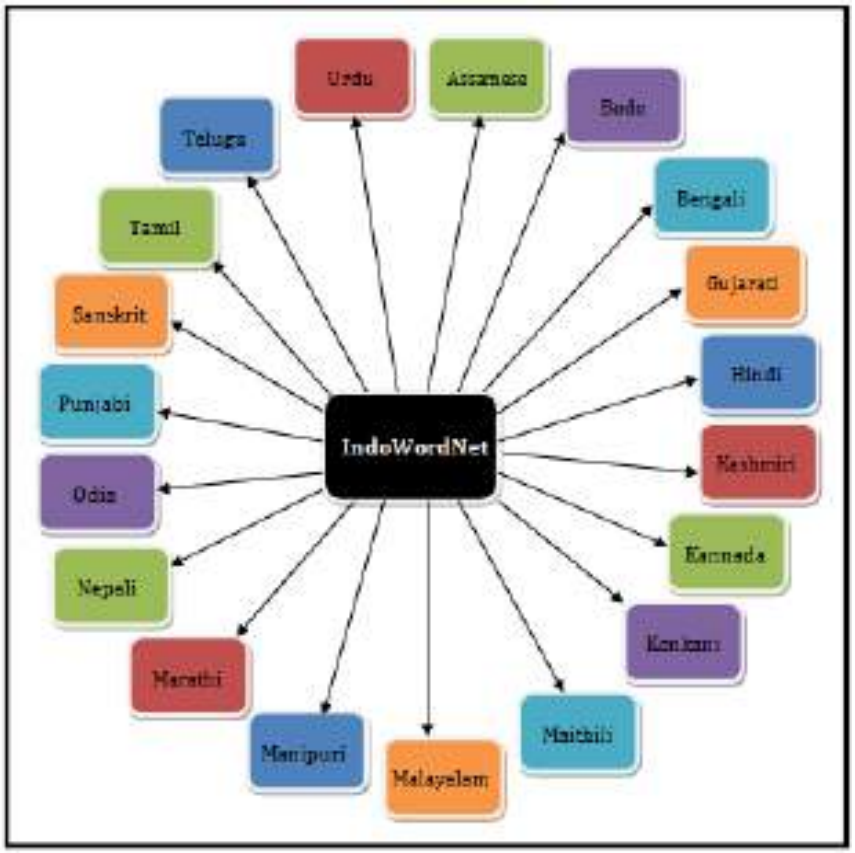
Figure: Relationship between Concepts

[2] http://www.cfilt.iitb.ac.in/wordnet/webhwn/wn.php

IndoWordNet

IndoWordNet[3]

- A multilingual integrated Wordnet for 18 Indian languages, including Hindi.

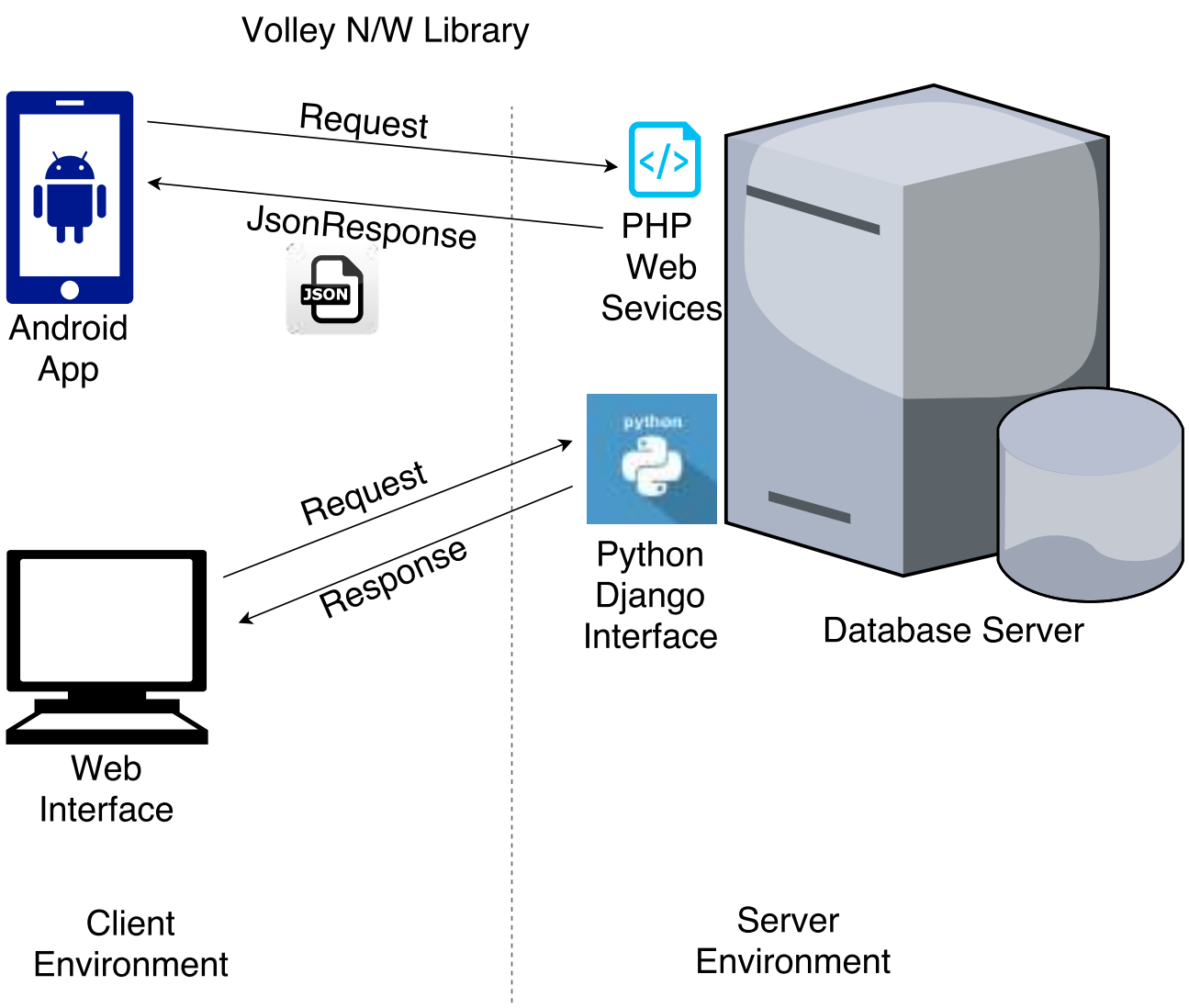


[3] http://www.cfilt.iitb.ac.in/indowordnet/

Features

- Layered Interface
  - The interface is divided into five levels depending on the expertise level of the learner
- Picture Depiction
  - Illustrations to demonstrate concepts in a better way
- Audio Pronunciation
  - To learn actual and proper pronunciation of words
- Grammatical Features
  - Word level features such as gender, plurality, case markers, etc.

Architecture of the Android and Web Application



Android Application Development

Implementation details

- Designed and developed layered interface using Android Studio
- Picasso library is utilized for picture depiction which handles ImageView recycling
- MediaPlayer Class is integrated to test and run the audio pronunciation files
- PHP web services are developed to extract layer wise information, including grammatical features from database server.
- JSON based parsing is performed to render information in the Android application.
- Client server communication is effectively handled using Volley (HTTP library)

Importance of Volley

- Volley utilizes power of android to access network efficiently, hence minimum effort on developers side.
- Benefits: Faster network operations, cleaner code.
- Overcomes limitations of AsyncTask (standard way, one needs to write code to access network, use HttpURLConnection etc.)

Select Expertise Level

- User can select the level of expertise

Search by Letter

- User will be shown two most frequently used words starting from the entered letter

Search by Letter (Contd.)

- User can select any of the two words and can view all grammatical information related to that word.

Search by Word for Beginner level

- The user can search any word and all grammatical information for all senses of the searched word will be displayed.

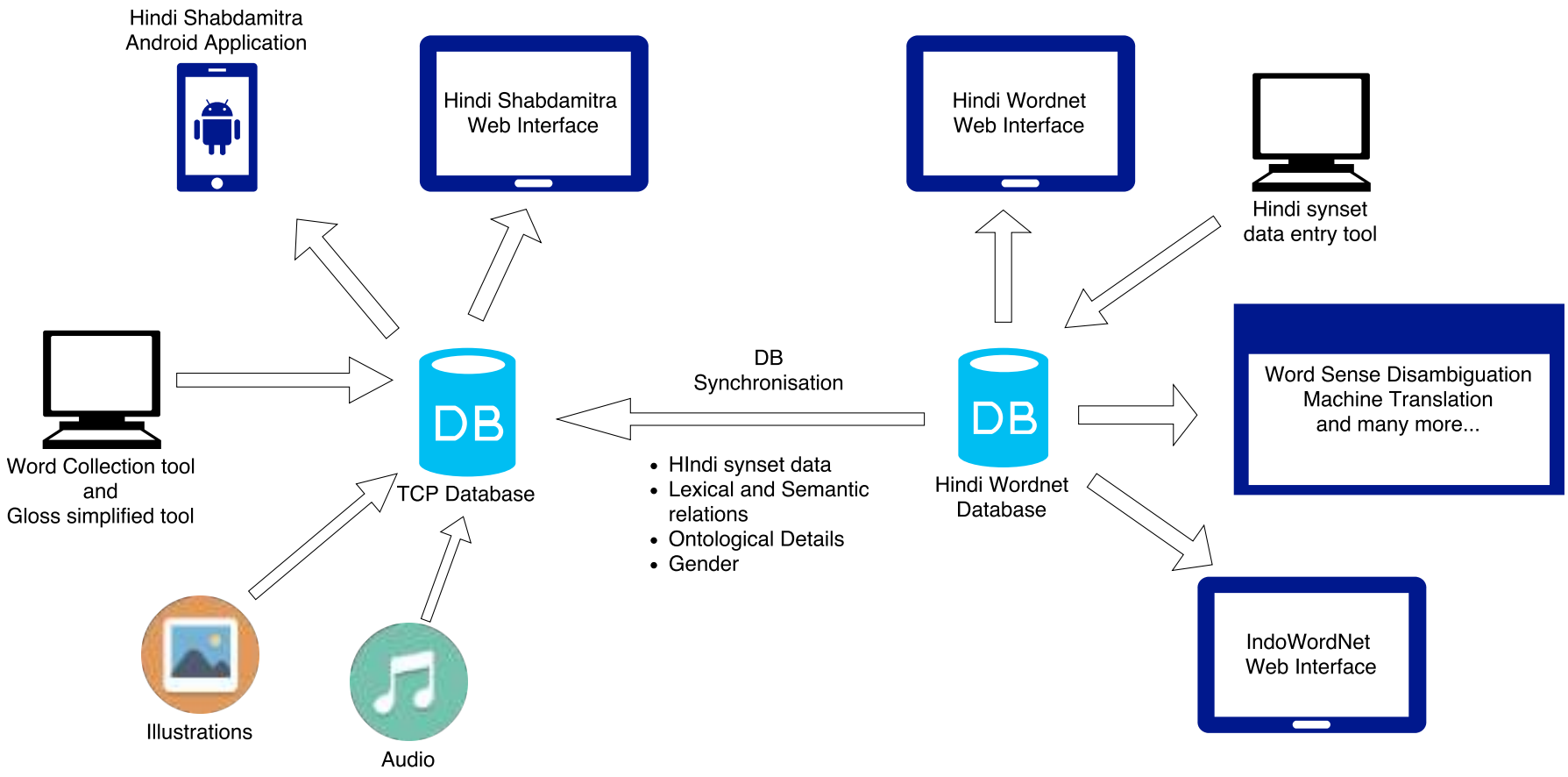
Search by Word for Intermediate level

- The user can search any word and all grammatical information for all senses of the searched word will be displayed.

Database Synchronization

- Methodology of maintaining consistency across databases over time and repeating the process periodically
- Types of synchronization
  - Insert synchronization
  - Update synchronization
  - Drop synchronization
  - Mixed synchronization

Database synchronization work flow diagram



Challenges of Database synchronization

- Structure of two databases
- One way or two way synchronization
- Time duration for performing synchronization
- Selection of methods to do synchronization

Database Synchronization

Methods of database synchronization

- Use a hosted system
- Master slave approach
- Update data in existing records during imports
- Use a third party tool dedicated to database syncing
- Creating system for synchronization

Database Synchronization

Need of Database synchronization

- Hindi Wordnet database
  - Devised for research purpose
  - Generalized
- TCP Database
  - Devised for research as well as academic purpose (in sync with school curriculum)
  - Well organized
  - Normalized
  - Specialized (audio, illustration, grammatical features)

Cron Job

- Time based job scheduler in UNIX based environment
- The schedule for the cron is created in a text file

```
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 > log06042017.txt
tatacenterproject@tukaram:~/public_html/dbsync$
```

Figure: Scheduling cron job

Database Synchronization

```
SELECT *
FROM `sync_details`
LIMIT 0 , 30
```

Page number: 1 ▾ > >>

Show: 30 row(s) starting from row # 30 in horizontal ▾ mode and repeat headers after 100 cells

Sort by key: None ▾

+ Options

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

Figure: Synchronization Progress Logging details

Ongoing: Hindi Shabdamitra Visualizer

Hindi Shabdamitra Visualizer

- The motive of Hindi Shabdamitra Visualizer is to serve as a pictorial aid for graphically depicting various relations and associations between concepts
- Principle - Theme Based Learning

Theme Based Learning

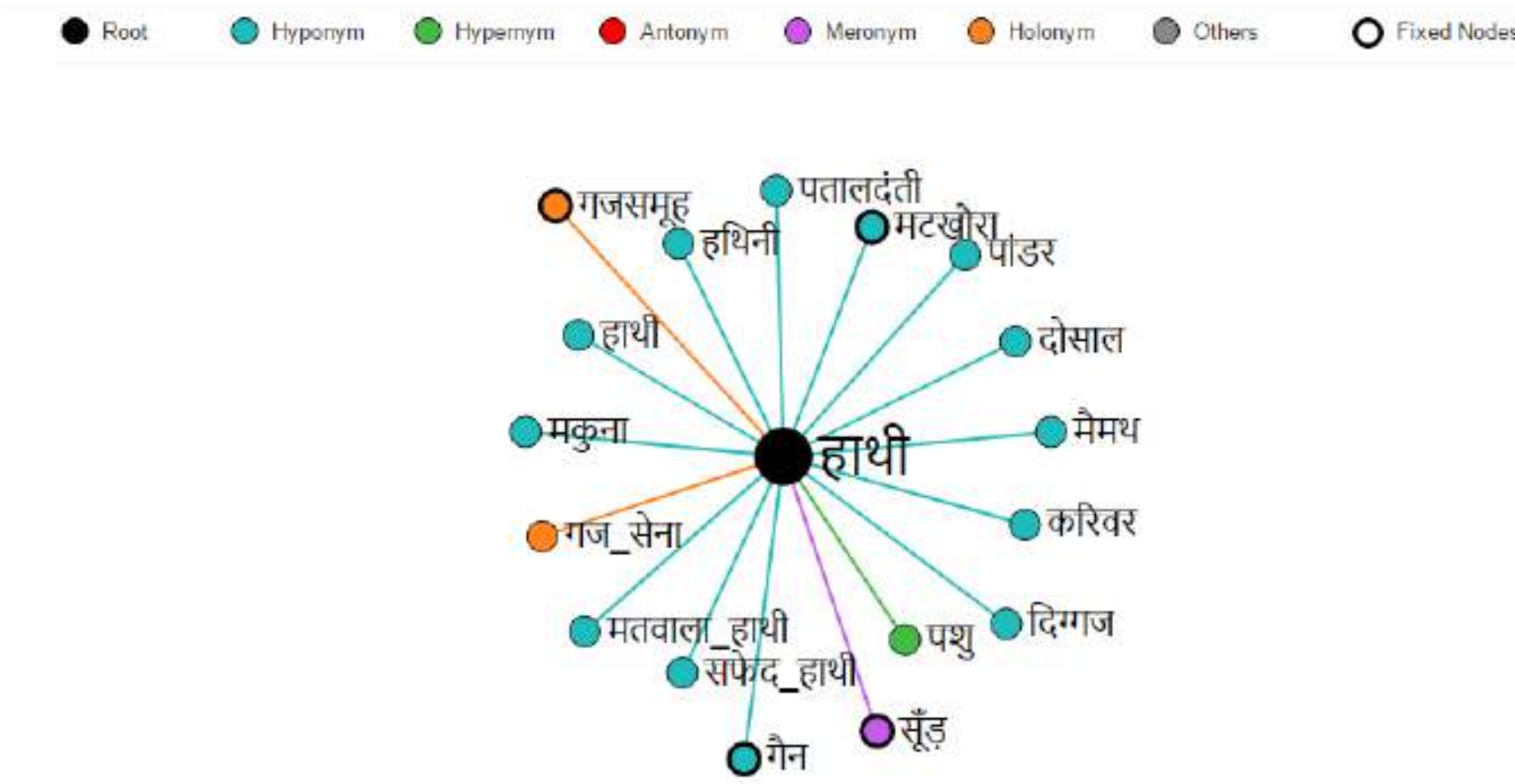
- Emphasizes on association and relation between concepts
- Supported by child psychology researchers

Ongoing: Hindi Shabdamitra Visualizer

Relations

- Hypernymy/Hyponymy (is-a)
- Meronymy/Holonymy (part of)
- Synonymy and Antonymy
- Ontological relations
- Related concepts based on theme/scene/setup
- Applications of concepts
- Onomatopoeic words

Ongoing: Hindi Shabdamitra Visualizer



Project Prototype Demonstrations

Prototype demonstration of web and app interfaces done at:

- Campus school, IIT Bombay
- Kendriya Vidyalaya, IIT Bombay
- Kendriya Vidyalaya, Bhandup
- Workshop organized by the funding institute (Tata Center) at IIT Bombay

Feedback

- App interface well appreciated by both teachers and students
- Improved learner engagement
- Students could learn and understand concepts in a better way using audio pronunciations and illustrations

Project Prototype Demonstrations

4

[4] http://tctd-iitb.tumblr.com/post/156973978834/supplementing-language-teaching-a-field-test

Summary and Future work

Summary

- Designed and developed Android application serving as an audio visual dictionary
- Developed database synchronization subsystem
- Ongoing: Hindi Shabdamitra Visualizer for 'Theme Based Learning'

Future work

- Hindi Shabdamitra visualizer
- Android application for higher levels

References

- Beckwith, Richard, Christiane Fellbaum, Derek Gross, K. Miller, G. A. Miller, and R. Tengi. "Five papers on WordNet." Special Issue of Journal of Lexicography 3, no. 4 (1990): 235-312.
- Bhattacharyya, Pushpak. "IndoWordNet." In The WordNet in Indian Languages, pp. 1-18. Springer Singapore, 2017.
- Prabhu, Venkatesh, Shilpa Desai, Hanumant Redkar, N. R. Prabhugaonkar, Apurva Nagvenkar, and Ramdas Karmali. "An effcient database design for IndoWordNet development using hybrid approach." (2012).
- Chaplot, Devendra Singh, Sudha Bhingardive, and Pushpak Bhattacharyya. "IndoWordnet visualizer: A graphical user interface for browsing and exploring wordnets of Indian languages." In Global WordNet Conference 2014 (GWC 2014). 2014.

Acknowledgement

Our sincere thanks to

- Dr. C. K. Bhensdadia, Dharmsinh Desai University
- Prof. Pushpak Bhattacharyya, IIT Bombay

Thank you