

Constructive Approach to Teach Inflections in Marathi Language

Abhijit R. Joshi
D.J.Sanghvi College of Engineering, Mumbai.
Email-id: abhijit@svkm-kingston.com

M. Sasikumar
CDAC, Mumbai.
Email-id:sasi@cdacmumbai.in

Abstract

Advancement in information and communication technologies influences the learning and teaching processes in class room; e-learning is no exception. E-learning has been an active field of research over the last couple of decades or so, covering aspects such as instruction design and delivery, assessment and collaboration [3]. Among the various domains, language learning is an important area with high potential for e-learning. Through this paper, we present here some of the results obtained within a research project which is aimed at developing an ILEIL (e-marathi). ILEIL (e-marathi) is an intelligent tutoring system which supports student in learning Indian language specifically Marathi. This paper describes constructive approach to teach inflections in Marathi language. The system presents word to student and asks for appropriate inflection of it. The system evaluates student's response and gives intelligent feedback and guidance to the student in their language learning process. Mean time system also records student's progress. Research is in progress and it is hypothesized that the fully developed system will be sufficient to provide an intelligent - rich learning environment for students.

Keywords Constructivist Learning Environments (CLE); Computer Assisted Language Learning (CALL); Intelligent Learning Environment for Indian Languages (ILEIL); Intelligent Computer Assisted Language Learning (ICALL); Intelligent Tutoring System (ITS).

1. Introduction

Marathi is the language spoken by the native people of Maharashtra. It is also spoken and understood in parts of other Indian states also. Marathi belongs to the group of Indo-Aryan languages. All of the Indo-Aryan languages originated from Sanskrit. Marathi is also evolved from Sanskrit through Prakrit and Apabhramsha [1]. Its grammar and syntax derive from Pali and Prakrit. In ancient times, Marathi was called *Maharashtri*, *Marhatti*, *Mahratti* etc. The script currently used in Marathi is called 'baaLbodha' which is a modified version of Devnaagari script [1]. Since Marathi has been mainly derived from Sanskrit so majority of words found in Marathi are Sanskrit-based. Other than Sanskrit, Marathi has also been influenced by the languages of its neighboring states. Marathi has absorbed words from the languages of different people who ruled India at different times.

It is the official language of the state of Maharashtra. There are 90 million fluent speakers worldwide. Marathi is the 4th most spoken language in India and the 15th most spoken language in world [5,6].

The need of ILEIL for Marathi is come up due to 1) the growing interest in language learning system; 2) to know the rich culture embedded in Marathi; 3) Marathi language student 4) tourist who are visiting Maharashtra and willing to talk in Marathi and 5) NRI who want to keep in touch with their families.

In the design of ILEIL, we made an attempt to integrate the potentials of three areas namely Constructive Learning Environment (CLE), Computer Assisted Language Learning (CALL) and Intelligent Tutoring System (ITS) [2].

CLE is a student centric, constructive and collaborative framework that provides a powerful environment where students can play around as they wish and while playing, they learn the language.

CALL is an approach to language teaching and learning in which computer technology is used as an aid to the presentation, reinforcement and assessment of material to be learned, usually including a substantial interactive element.

ITS attempts to simulate such a “teacher”, who is capable of providing individualized tutoring or guidance in a given subject using pedagogical methods suitable to a student. It also monitors progress of an individual one, in an online setup based on his or her level of understanding in the subject.

A full fledged ITS for language learning consists of: (1) language expertise model, (2) a model that keeps track on student’s current knowledge about the language in order to identify the reason(s) of his/her mistakes, (3) a

model that helps or guide the student in learning and teaching process and makes it more effective by applying appropriate tutoring strategies and pedagogical principles of language learning.

The goal of the ongoing research is to integrate the recent developments in the field of CLE, CALL, ITS and ITC to construct an effective learning environment that helps student in learning Marathi language in its Colloquial form. The system make the user familiar with expressions in everyday conversations and typical dialogue in Marathi language that can take place in public places like railway station, restaurant, post office, bank etc in a systematic way.

The ILEIL is divided into various modules. These modules can be for the word level operations (inflectional operations) or the sentence level operations. This paper focuses on the inflectional operations. The word level modules will have the capability to operate on individual words for operations like plural to singular, vibhakti and verb formation etc. These modules are then be used to teach sentence formation.

The paper is organized as follows: section 2 discuss about inflectional operations in Marathi Section 3 gives idea about our approach. The paper concludes with discussion of our ongoing work.

2. Inflectional Operations in Marathi

A sentence is the well arranged group of words. These words have different roles and as per their roles they are given a name like noun, pronoun, adjective, verb, etc. In Marathi there are eight different types of words present. Sentence is constructed using these eight types of

words but the words used in any sentence are not necessarily in their original form. To form a correct sentence sometimes you need to change the form of these words. This change is required due to the change in gender or number. Such change does not occur in all types of words and are called as non inflectional words. The words in Marathi are broadly categorized into two types namely inflectional and non inflectional. Noun, pronoun, adjective and verb are inflectional words. In subsequent section we will see the implementation of it.

3. The ILEIL/e-marathi system

The ILEIL system is an ITS for any one who is willing to learn Marathi.

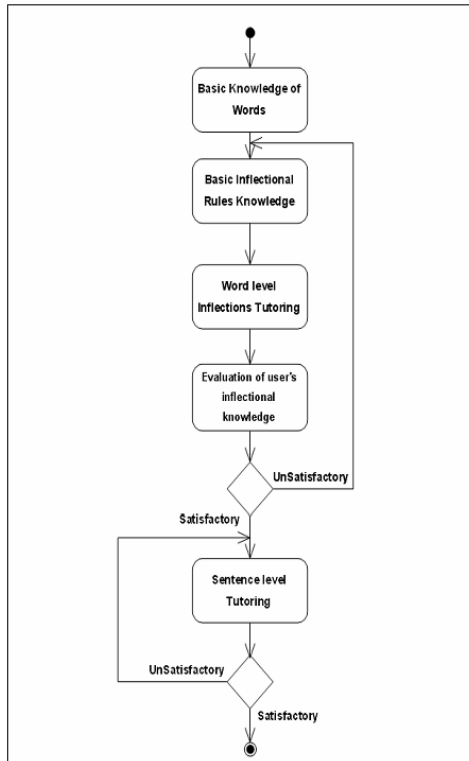


Figure 1: Tutoring Strategy of ILEIL

The system follows the classical architecture of an ITS where the language knowledge is stored in

domain model. The tutor model consists of the strategies of the teacher. The tutor model decides the necessity of an intervention at any given point of time. It also decides the kind of question to be asked to the user. The student model will keep track of the student's current state of knowledge and thus will help the tutor model to decide on the future course of the teaching. The tutoring strategy of ILEIL is shown in figure 1.

3.1 Noun Inflections

Noun form changes due to the change in gender or number or vibhakti and called as noun inflection. Here we are concentrating on singular to plural operation for noun and vice versa. Noun inflection uses two knowledge repositories viz noun-domain and noun-rules which form the domain knowledge of the system. The noun-domain consists of noun along with its gender. These two information are stored as - noun:gender. e.g. aamba:M. This arrangement is made in order to fetch both the data together in one file read operation. The noun-rules consists of the rules of conversion from singular to plural. The rules are stored in the following format -

*gender: ends_with_character/s:
replace_by_character/s: rule_id*

Gender (लिंग)	Words ends in	Rule (नियम)
Masculine पुल्लिंगी	a अ	R1: a → a अ → अ
	aa आ	R2: aa → e आ → ए

Table 1: Example - Rules for singular to plural [4]

In noun inflectional, the ITS selects a word/gender pair from the noun-domain at random. The word is used to ask student the plural without stating the gender, which is left to the

student to decide. The correct plural is thus generated by searching the rules and matching the gender as well as the ending string.

The student is asked for the plural of the word randomly picked up from the noun-domain. The answer given by the student is analyzed by applying reverse of each rule to the word. The analysis is done keeping in mind the type of mistakes that the student can make. These include the mistakes in identifying the gender or application of the wrong rule.

The answer of the student, if wrong, is operated upon by the reverse of each rule in the knowledge domain. Any such operation, if results in the original word, is noted. If the student answer satisfies more than one rule then the student is told the gender and asked the plural again. This intervention addresses the issue of constructivism where the student is not given the correct answer, instead is made to arrive at the answer. If only one rule matches the student answer, the student is informed about the mistake he has made, be it the gender, the rule applied or the application of a non-existent rule.

The module to convert a noun from plural to singular implements the similar technique but the rules are applied in reverse. Thus, both the modules can use the same set of rules.

3.2 Vibhakti Infections

The relation between noun and pronoun with verb or any word in sentence is shown by inflection called as vibhakti. To know the right vibhakti of a noun you need to know first saamaanyaroop of that noun. The rules for noun to saamaanyaroop are stored in following format –

*gender: ends_with_character/s:
replace_by_character/s: rule_id*

Gender (लिंग)	Words ends in	Rule (नियम)
Masculine पुल्लिगी	A अ	R1: a → aa अ → आ
	aa आ	R2: aa → yaa आ → या

Table 2: Example - Rules for noun to saamaanyaroop [4]

The saamaanyaroop is obtained by applying the appropriate rule. Then vibhakti form (prathamaa, dwiteeyaa,...) of the saamaanyaroopa is obtained for the plural or singular by adding appropriate pratyaya from table 3. The process of obtaining vibhakti from noun is shown in figure 2.



Figure 2: noun to vibhakti

Vibhakti	Singular	Plural
Prathamaa	---	---
Dwiteetyaa	Sa, laa	sa, naa
Truteeyaa	ne, shee	nee, shee
Chaturthee	sa, laa	sa, naa
Panchamee	hoona	hoona
Shashthee	chaa, chee, che	chaa,chee, che
Saptamee	ta	ta
Sambodhan	---	no

Table 3: Types of vibhakti [4]

The rules for saamaanyaroop to vibhati are stored in following format –
vibhakti:number(plural/singular): pratyaya1 @ pratyaya2 @..... @ pratyayan

As you can see the rule is decomposed into three parts vibhakti, number, and the possibilities of characters ‘pratyaya’ from table 3 to be added to the ‘saamaanyaroopa’ to

convert it into the vibhakti of the given number. These possibilities are then further decomposed to obtain all the possibilities of the vibhakti. Vibhakti inflection follows the similar analysis technique as used in case of noun inflection. The ITS select noun from domain randomly and ask for the saamaanyaroop of it. If it is correct then it ask further to tell the vibhakti form of the saamaanyroop. The student's input is then evaluated against the ITS solution which is evaluated using the rules for 'vibhakti' given in table 3. If the student's answer is correct, it is indicated to him along with the other possibilities of vibhakti. Otherwise, he is given all the possibilities of vibhakti of the samanyaroop.

3.3 Verb Inflections

Verb is the word which tells something about the person or thing in the state of doing, being, being acted upon. Thus it is the doing or being word. Verb also completes the meaning of sentence. The verb formations depend on four major grammar constituents viz tense, person, gender and number. The basic word of the verb is called as 'dhaatoo'. We add pratyaya in dhaatoo to get the verb. E.g. 'kheLa' is dhaatoo in which if we add pratyaya 'to' which leads to a verb 'kheLato'. In similar way we can get the different forms of verb for any person, gender, number and tense by adding appropriate pratyaya from table 4. All dhaatoo are stored in a domain model. Whereas the rules are stored in the following format –

Ruleid : person type: gender : vachan : tense : characters to be added

पुरुष (Person)	पुंसलिंग (Masculine)						स्त्रीलिंग (Feminine)						नपुंसक (Neutral)							
	अन्यवचन			अन्यवचन			अन्यवचन			अन्यवचन			पुरुषवचन			पुरुषवचन				
	Singular		Plural		Singular		Plural		Singular		Plural		Singular		Plural		Singular		Plural	
	काळ (Tense)	Pre	Pt	F	काळ (Tense)	Pre	Pt	F	काळ (Tense)	Pre	Pt	F	काळ (Tense)	Pre	Pt	F	काळ (Tense)	Pre	Pt	F
प्रथम (First)	तो	ला	ला	ला	ते	ला	ला	ला	ते	ते	ला	ला	ते	ला	ला	ला	ते	ते	ला	ला
	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला
	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला
द्वितीय (Second)	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला
	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला
	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला
तृतीय (Third)	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला
	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला
	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला	ला

Table 4: rules for dhaatoo to verb [1]

Here also we follow the similar analysis technique as mentioned in the previous sections. The student is asked to give the verb of the mentioned person, gender, number and tense of the dhaatoo. This dhaatoo is selected at random. The student response is then evaluated against the ITS solution, which is evaluated using the rules for verb given in table 4. If the student's answer is correct, it is indicated to him along with the other possibilities of verb if any. In case the student answers it wrong then it identifies the mistakes, analyses the student's response, provides a suitable intervention and gives the student a second chance and if still unsuccessful it provides the correct conversion and the answer.

4. Conclusion

In this paper we have presented various issues involved in the design of modules for inflectional words in Marathi language. This work is a part

of a research which is aimed at developing a generic architecture – an ITS that helps students in learning Marathi language. A prototype of inflectional operations in Marathi is implemented using a Python so that system can run on heterogeneous platforms. We started with the traditional rules used for generating inflections which are stored in domain model and these rules are applied on words from domain model to get the ITS solution. We feel that teaching a second language can constitute a good benchmark for evaluating the soundness, completeness and legacy of framework that integrates the areas like CLE, ICALL and ITS. Ongoing work includes reinforcement of the present models and integrating them with simple sentence module. Special attention is given to the adaptability of the models to produce web-based programs to be introduced in existing e-learning tools.

5. Acknowledgement

The authors would like to thanks Buch Rudri, Busa Nayan and Siddiqui Sadaf for extending their help in implementation and carrying out this research.

6. References

- [1] Gowilkar, Leela, Marathiche Vyakaran, Mehta publishing House, 2006.
- [2] Joshi, A. R. and Sasikumar M., Intelligent Learning Environment for Learning Indian Languages. National conference on New Horizons in Technology and Applications, Mumbai, Maharashtra, March 2007.
- [3] Sasikumar, M. and Joshi, A. R. Intelligent Instructional Systems for Language Learning. Instructional conference on Research Trends in Information

Technology, Kolhapur, Maharashtra, March 2007.

- [4] Walambe, M. R., Sugam Marathi Vyakaran Lekhan, Nitin Prakashan, 2005.
- [5] http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/Statement1.htm
- [6] http://encarta.msn.com/media_701500404/Languages_Spoken_by_More_Than_10_Million_People.html