Indian Institute of Technology, Kanpur

Course Project CS689A

Computational Linguistics for Indian Languages

Mrunmayee Amshekar, 22100269
Sunil Dhaka, 17817735

Project Report

TAM Tags in Marathi and Hindi

Introduction

For typological study of languages, tense, aspect and mood are three crucial concepts. Tense is the temporal location of an event. Tense markers in languages relate the time of the described event to the present time (Comrie 1976). Aspect on the other hand relates to different ways to view the internal temporal constitution of a situation (Comrie 1976). That is, whether a situation is viewed from within (imperfective) or as a single point or a whole (perfective). Bhat (1999) describes the characteristics of mood as follows:

> "(i) a speaker's opinion or judgement regarding the actuality of an event,
> (ii) kind of evidence that is available for the speaker to form this judgement, and
> (iii) kind of need or requirement which forces the speaker (or someone else) to get involved   in an event (or to carry out an action)."

These features of tense, aspect and mood are predominantly marked on the verb as they provide specifications about the event being described in the utterance. In this study we have reviewed the TAM tagset being created for Indian languages and made an attempt to adapt it for it to be suitable for Marathi. It must be kept in mind that though all the concerned languages are related, the categories across them may not be identical. A brief critical review of the said tagset has been done. The computational part of the project aims to generate various verb forms in Marathi and Hindi when feature specifications are provided in the input. We have also attempted a task wherein when input verbs are provided in Hindi, an output with identical tense and aspect in Marathi would be given.

TAM tagset for Marathi

TENSE

TAG-Translation	Type	Marathi pattern	Marathi example
Present	Tense	त आहे	मी हसत आहे. मी हसते आहे. मी हसतेय .
Present	Tense	आहे	हे चंदन आहे.
Past	Tense	लो	मी खेळलो

| Past-Aux | Tense | त + होते | मी खेळत होते |
| Future | Tense | -ईल | उद्या राधा येईल. |

ASPECT

perfect	Aspect	(erg) ल	मी गेले
Perfective-past	Aspect	(erg) ल	मी आंबा खाल्ला
Aorist-past	Aspect	लेल	तुम्ही त्या दिवशी जेवण बनवलेलं
Imperfect past	Aspect		मनीषा कपडे धूत होती.
future	Aspect		
Habitual present	Aspect	ते	ती कविता लिहिते
Habitual present	Aspect		
Habitual present	Aspect	यचे	इथे आधी मुलं राहायची
Completive-perfect-past	Aspect		
Perfective present	Aspect	लो आहे	आम्ही पहिला आहे
Progressive-imperfect-present	Aspect	त आहे	ते गात आहेत
Durative Progressive-past	Aspect		
Durative Progressive	Aspect		
Continuous Progressive-present	Aspect		
Perfect-subjunctive	Aspect		

Mood

Indicative-Present	Mood		मी इथे आहे.
Indicative-Future	Mood	ल	इथे पाऊस पडेल
Indicative-Past	Mood	ला असावा	इथे पाऊस पडला असावा
Imperative Polite	Mood	-	-
Imperative semi polite/familiar	Mood	pluralization	तुम्ही घरी जा
Imperative	Mood		तू जा
Imperative, prescriptive	Mood	व	तू घरी जावंस

Hortative	Mood	ऊ	जाऊ, खाऊ
Conditional	Mood	जर-तर	जर मी अभ्यास केला असता तर मला चांगले मार्क मिळाले असते.
Potential-Possibility	Mood	शकतो	राम घरी जाऊ शकतो
Potential-Ability	Mood	शकतो	राम पाच किलोमीटर धावू शकतो
Subjunctive/Conjunctive	Mood		तू औषध घ्यावंस. तिने परत यावं असं मला वाटतं . मुलांनी रस्त्यावर खेळू नये.
Desiderative	Mood	यच	त्याला यायचं आहे
		वो	त्यांचा सत्यानाश होवो
Optative	Mood		
Debitive	Mood	यच	मला लगेच दिल्लीला जायचं आहे
Presumptive	Mood	असेल	तो अभ्यास करत असेल
Contrafactual	Mood	तर - नसतं	मी तुझ्या जागी असते तर मी असं नसतं केलं
Interrogative	Mood	का? ना?	मुलं आली का?
Compulsive/Necessitive/Obligative	Mood	लागेल	रामला खीर खावी लागली

It can be observed in the table above that the three features tense aspect and mood overlap in marking of a sentence. Each verb may carry a certain combination of these three. Additionally mood is different from tense and aspect in the fact that it describes the actuality of the event and may be described by elements other than the main or auxiliary verbs in a sentence. For example

Hindi	English	Features
मैं जाऊंगी	I will go.	Future Tense
मैं शायद जाऊंगी	I might go.	Mood - possibility

It can be observed that a sentence like 'मैं जाऊंगी ' in absence of the word 'शायद' has potential to display both modal and future-tense meaning. Since tense and aspect are more basic to an utterance it is recommended that in the process of analysis tense and aspect be given priority and modal analysis may follow this.

Tense

- Marathi has a feature (perhaps owing to contact with Dravidian languages) wherein it expresses a non-past meaning in its first person verbs e.g.
    जाते - I go OR I will go
  As opposed to Hindi present
    जाती हुं - I go

- The tagset makes two entries - past and past-aux but the hindi examples provided in this category show similar behaviour

  पिछले साल, मैं भारत गया था |

  "मैं खेला था। हम खेले थे। तुम खेले थे। वो खेला था।

  Since both show a perfective+past structure it would not be wise to make two categories.

Aspect

- The difference between perfective and perfect is that perfect is perfective which has relevance to the present. The tag in the tagset wrongly shows perfective as perfect. Marathi and Hindi examples for perfective and perfect are given below:

Marathi	Hindi	Aspect
मी आंबा खाल्ला.	मैने आम खाया|	Perfective
मी आंबा खाल्ला आहे.	मैने आम खाया है|	Perfect

- There are two distinctions within the Imperfective aspect - progressive and habitual hence the tag 'Imperfect past' is incorrect. Also Sanskrit does not make a strict distinction between habitual and progressive hence the examples given seem incorrect.
- The future tense constructions like 'मैं जाऊंगी ।' are only marked for tense as the future event in this construction is neither being viewed as a progressive nor as a whole event.
- In most of the Hindi examples of aspect use of compound verb construction can be seen. Though the different compound verbs do express aspectual meanings, they may have other nuances attached e.g. दिया has a directional meaning towards someone other than the speaker and लिया has a meaning towards self and hence constructions like दे लिया or ले दिया seem unacceptable.

Due to such discrepancies in the Tagset I propose a tagging method with reference to the linguistics literature, for Hindi and Marathi (this same method has been used in our computational part of the project).
- ➢ Given a sentence first it aspect should be analysed as Imperfective or Perfective, with further divisions wherever applicable.
- ➢ Each aspect would be further tagged with respect to the tense related to it.
- ➢ This would give us a tag which is a combination of tense and aspect.
- ➢ It may happen (like in the case of future tense or perfective aspect that one of the two is missing).
- ➢ Compound verbs in a language which afford an aspectual meaning to the utterance should be identified along with the other nuanced meanings that they express. This step would aid in the specification of the aspect (e.g. as being completive or durative).
- ➢ The modal analysis would come as a further step. Sentences may first be divided into realis or irrealis mood (realis is related to past tense and perfective aspect and irrealis is related to events which have not actualized at the time of utterance)

➢ For modal analysis other aspects of the utterance e.g. adverbials have to be taken into consideration.
➢ The distinctions epistemic mood and deontic mood should also be considered as these would help in locating similar modal construction during translation of the utterance into a target language (if an equivalent is not to be found).
➢ After this step the event would get a TAM tag which may be further used to translate the utterance into an equivalent utterance in the target language (consideration of parts of a sentence other than the word was not feasible and hence only tense and aspect have been considered for the computational part of the project).

The code

We performed two tasks:
1. Given a root verb (in Hindi or Marathi) and the tense- aspect and other specifications, the code outputs the desired verb form.
2. Given a verb form in Hindi, from the dictionary given to the code, the program outputs a verb in Marathi having the same tense and aspect.

The outputs of the code are satisfactory although it is not capable of handling irregular verbs.

Works Cited

Comrie, Bernard. *Aspect*. Cambridge, Cambridge University Press, 1976.

Shankara Bhat, D. N. *The Prominence of Tense, Aspect, and Mood*. J. Benjamins Pub.,

      1999. Accessed 15 November 2022.