

Vaiyyākaraṇaḥ: A Sanskrit Grammar Bot for Telegram

Hrishikesh Terdalkar Mahesh A V S D S
hrishirt@cse.iitk.ac.in maheshak@cse.iitk.ac.in
Shubhangi Agarwal Arnab Bhattacharya
sagarwal@cse.iitk.ac.in arnabb@cse.iitk.ac.in
Dept. of Computer Science and Engineering,
Indian Institute of Technology Kanpur,
India.

Rise of social media opens novel opportunities of learning. Telegram¹ is an instant messaging service available as a cross-platform, freemium software. Telegram bots are instances of Telegram clients that are capable of performing actions in response to various user actions. Vaiyyākaraṇaḥ is a telegram bot aimed towards helping the learners of Sanskrit grammar (vyākaraṇa).

Vaiyyākaraṇaḥ is made using *Telethon*², a Telegram client library in Python 3. It makes use of data (Dhātupāṭhaḥ) from <https://ashtadhyayi.com> for conjugations. The bot also uses some of the state-of-the-art Sanskrit computational linguistic tools. The Heritage Platform (Goyal et al., 2012) is used for tasks related to declensions. Sanskrit Sandhi and Compound Splitter (Hellwig and Nehrlich, 2018) is used for the word segmentation task. The user of the bot may type a set of provided keywords as commands followed by the appropriate input text in their Telegram client to obtain the output.

The salient features of the bot are:

- Stem finder (Prātipadikam)
The keyword ‘\shabda’ can be used, followed by a word form (subanta) to obtain the stem (prātipadikam) and morphological information.
- Declension generation (Subantāḥ)
Upon searching a word form, the bot also provides an option to show all morphological forms (declensions) of the provided word.
- Root finder (Dhātuḥ)
The keyword ‘\dhatu’ can be used, followed by a verb form (tiñanta) to obtain the root (dhātu) and morphological information.
- Conjugation generation (Tiñantāḥ)
Upon searching a verb form, the bot also provides an option to display all morphological forms (conjugations) of the provided verb.
- Word segmentation (Sandhisamāsau)
The keyword ‘\vighraha’ can be used, followed by Sanskrit text can be provided to obtain the word segmentation (splitting both sandhi and samāsa).

Figure 1 and Figure 2 illustrate various capabilities of Vaiyyākaraṇaḥ. A video demonstration can be found at <https://sanskrit.iitk.ac.in/jnanasangraha/vaiyyaakarana>. Vaiyyākaraṇaḥ is available for use at https://t.me/vyakarana_bot. The source code can be found at <https://github.com/hrishikeshrt/vaiyyakarana/>³.

References

- [Goyal et al.2012] Pawan Goyal, Gérard Huet, Amba Kulkarni, Peter Scharf, and Ralph Bunker. 2012. A distributed platform for Sanskrit processing. In *COLING*, pages 1011–1028.
- [Hellwig and Nehrlich2018] Oliver Hellwig and Sebastian Nehrlich. 2018. Sanskrit word segmentation using character-level recurrent and convolutional neural networks. In *EMNLP*, pages 2754–2763.

¹<https://telegram.org/>

²<https://docs.telethon.dev/en/latest/>

³Please refer to `INSTALL.md` for installation instructions.

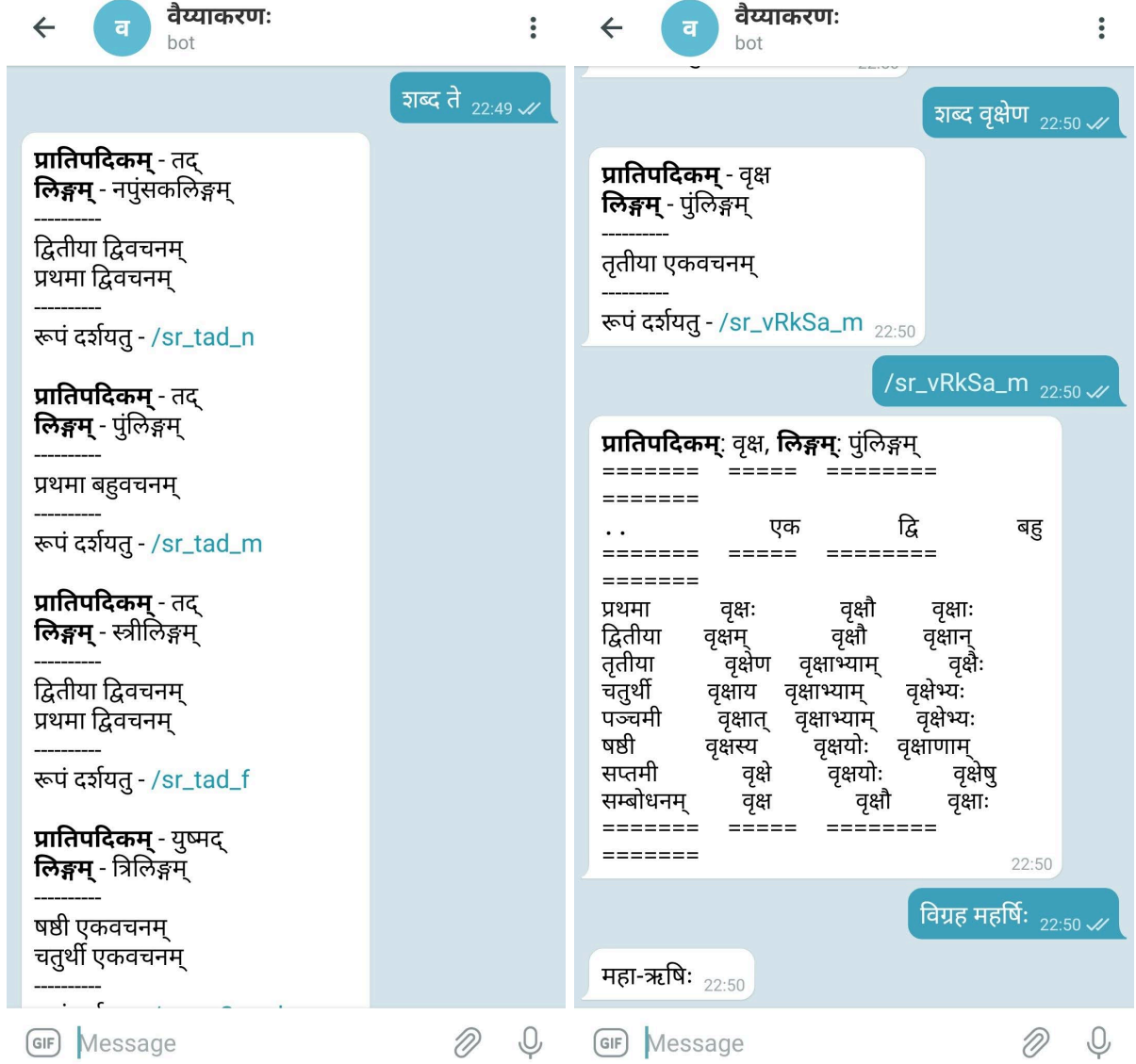


Figure 1: Stem Finder and Declension Generator

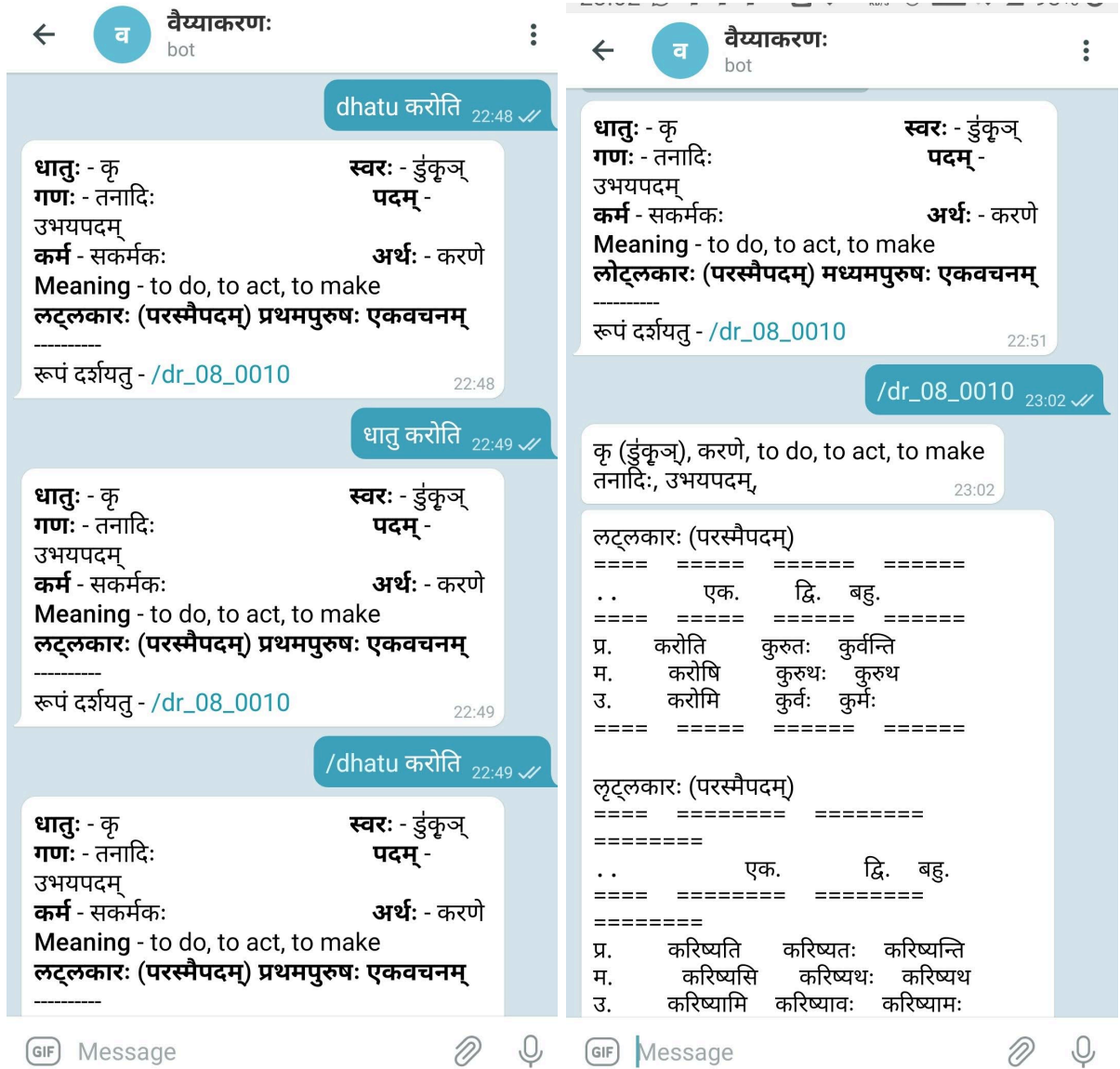


Figure 2: Root Finder and Conjugation Generator