

Detecting and correcting grammatical mistakes due to subject-verb inconformity and conflicts in tense aspects in Tamil sentences

Sakuntharaj R.^{1*} and Mahesan S.²

¹*Centre for Information and Communication Technology, Eastern University, Sri Lanka*

²*Department of Computer Science, Faculty of Science, University of Jaffna, Sri Lanka*

A grammar checker checks each sentence in a text whether it conforms to the grammar, and in case it finds conflicts in the conformity, it would suggest alternatives to the word(s) that causes the conflict. An approach to detecting and correcting grammatical mistakes due to subject and finite-verb disagreement with regard to person, number and/or gender and due to disagreement in tense aspects in Tamil sentences is proposed in this paper. Hierarchical POS tags of words are used to detect the grammatical mistakes and two sets of Tamil grammar rules are used to generate suggestions. First set of Tamil grammar rules is used to generate suggestions to correct mistakes due to subject and finite-verb disagreement, and the other one is used to generate suggestions for tense aspect disagreement. These grammar rules have been defined in the form of Python conditional statements based on standard Tamil grammar. Alternative suggestions for both subject and finite-verb are given for personal pronouns of third person and alternative suggestion is given for finite-verb for all types of subjects, not just for pronouns. Moreover, a list of tense-aspects words mapped to tenses is prepared and used for giving suggestions in case of tense aspect disagreement. Test results show that the proposed grammatical mistake detection and correction system performs well. For 100 testing sentences with all kinds of grammatical mistakes mentioned above, it gives 136 suggestions, a Scholar in Tamil language approved 132 suggestions as most suited and the remaining 4 are as acceptable.

Keywords: Tamil, Grammar Checker, Subject and Finite-verb, Tense, Error Correction

*Corresponding Author: sakuntharaj@esn.ac.lk