

11-712: NLP Lab Report

Pradeep Prabakar Ravindran

April 26, 2013 (Final Due date)

Abstract

First draft of a report documenting my experience developing a dependency parser for Tamil.

TamilDep is a tool that performs dependency parsing of the Tamil Language. This tool would soon be open-sourced and become publicly available at <https://github.com/pradeep-gnr/Tamil-dependency-parser>. Along with the code for a dependency parser, a corpora of tamil sentences annotated with their dependency parse trees is also expected to be released. I hope that this resource would prove useful to researchers working on Tamil natural language processing.

1 Basic Information about Tamil

2 Past Work on the Syntax and NLP tools for Tamil

3 Available Resources

[include discussion of your corpora –NAS]

4 Survey of Phenomena in [Your Language –NAS]

5 Initial Design

6 System Analysis on Corpus A

7 Lessons Learned and Revised Design

8 System Analysis on Corpus B

9 Final Revisions

10 Future Work

References