Dependency Relations for Sanskrit Parsing and
Treebank

Amba Kulkarni\textsuperscript{1}  Pavankumar Satuluri\textsuperscript{2}  Sanjeev Panchal\textsuperscript{1}
Malay Maity\textsuperscript{1}  Amruta Malvade\textsuperscript{1}

\textsuperscript{1}Department of Sanskrit Studies
University of Hyderabad

\textsuperscript{2}School of Linguistics & Literary Studies
Chinmaya Vishwavidyapeeth
Ernakulam

19\textsuperscript{th} International Workshop on Treebanks and Linguistic
Theories 2020
Pānini’s Grammar
Introduction

1. Pānini’s Grammar
2. Saṁsādhanī Tagset
Introduction

1. Pānini’s Grammar
2. Saṁsādhanī Tagset
3. Enhanced Tagset
Introduction

1. Pānini’s Grammar
2. Saṁsādhanī Tagset
3. Enhanced Tagset
4. Saṁsādhanī Parser
Introduction

1. Pānini’s Grammar
2. Saṁsādhanī Tagset
3. Enhanced Tagset
4. Saṁsādhanī Parser
5. Sanskrit Treebank
Introduction

1. Pānini’s Grammar
2. Saṁsādhanī Tagset
3. Enhanced Tagset
4. Saṁsādhanī Parser
5. Sanskrit Treebank
6. Evaluation
Introduction

1. Pānini’s Grammar
2. Saṁsādhanī Tagset
3. Enhanced Tagset
4. Saṁsādhanī Parser
5. Sanskrit Treebank
6. Evaluation
7. Conclusion
Introduction

1. Pānini’s Grammar
2. Saṁsādhanī Tagset
3. Enhanced Tagset
4. Saṁsādhanī Parser
5. Sanskrit Treebank
6. Evaluation
7. Conclusion
Pāṇini’s Grammar

- Around 400 BCE
- Full-fledged grammar of the then prevalent Sanskrit
- Covering both Vedic and Non-Vedic texts
- Only full-fledged, manually written, computational grammar for any natural language
• The first tagset for Sanskrit (Ramakrishnamacharyulu 2009)
• Approximately 100 relations
• Inter-sentential relations
• Intra-sentential relations
Examined from computational view point
Found to be fine-grained
Fine-grain relations - 100
Kartā

- anubhavī (experiencer)
- amūrtaḥ (abstract)
- prayojakaḥ (causative agent)
- prayojyaḥ (causee)
- abhiprerakaḥ (cause of temptation)
- karmakartṛ
- karaṇakartṛ
- ṣaṣṭīkartā
Skt: Rāmaḥ(Kartā) pacati
Eng: Rama cooks

Skt: Ghaṭaḥ(anubhavī kartā) naśyati
Eng: Pot is experiencing the destruction

Coarse-grain relations - 31
Lacunae in the Samsādhanī Tagset

Relations without any associated semantics
- Relations due to morphological requirement
- Complementiser

Inconsistent with Pāṇinian Grammar
- Coordinating conjuncts
Relations due to morphological requirement

Figure: Old annotation

g्रामम् village  paritah surrounding  vṛkṣāḥ trees  santi. are.

Figure: New annotation

g्रामम् village  paritah surrounding  vṛkṣāḥ trees  santi. are.
Complementiser

Figure: Complementiser: old version

Figure: Complementiser: new version
Coordinating Conjuncts

Figure: Conjunctions: Old version

Figure: Conjunctions: New version
Saṁsādhanī Enhanced Tagset

The current version has 54 relations and classified into following categories.

- Predicate-argument relations
- Non-Predicate argument relations
  - Verb-verb relations
  - Verb-noun relations
  - Noun-noun relations
- Relations due to special words
- Conjunctions and Disjunctions
- Miscellaneous
Samśādhanī Parser

1. First full-fledged Parser for Sanskrit.
2. Follows Pāṇinian Grammar and the Theories of Verbal cognition.
   - ākāṅkṣā (expectancy)
   - yogyatā (meaning congruity)
   - sannidhi (proximity)
3. Implemented as an edge-centric binary join to build a dependency tree, in bottom-up approach, with local and global constraints on the edges and the edge labels.
3. Treebank of Vedic Sanskrit (Hellwig et al., 2020)
Samśādhanī Platform

1. Saṅkṣepa Rāmāyaṇam
2. Śrīmad-Bhagavad-Gītā
3. Śiśupālavadham
Grammar Books

2 NCERT (National Council for Education, Research and Training) 9th grade - 284 sentences

3 Independent sentences from various Sanskrit books.

4 130 Short stories
Skt: mārgāḥ avaruddhāḥ bhavanti.
Gloss: Road{pl,nom} blocked{pl,nom} be{pres,pl,3p}.
Eng: The roads are blocked.

Figure: Inflectional Information
Skt: mitrāṇi kathayanti.

Gloss: friend{pl,nom} / friend{pl,acc} tell{pres,pl,3p } / tell{pres,pl,3p [causative]}

Eng: Friends tell / (They) tell friends / Friends make (somebody) tell / (They) make (somebody) tell friends.
Skt: rāmaḥ pustakaṃ kṛītvā paṭhati.
Gloss: Rama{sg,nom} book{sg,acc} purchase{abs} read{pres,sg,3p}.
Eng: Rama reads a book after purchasing it.
## Evaluation

<table>
<thead>
<tr>
<th>Source</th>
<th>Sentences</th>
<th>Exact Match</th>
<th>Failed</th>
<th>Partial Match</th>
<th>LAS</th>
<th>UAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar 9th grade</td>
<td>468</td>
<td>343</td>
<td>2(.4%)</td>
<td>123</td>
<td>89%</td>
<td>97%</td>
</tr>
<tr>
<td>Skt Learner</td>
<td>1070</td>
<td>817</td>
<td>15(.6%)</td>
<td>87</td>
<td>82%</td>
<td>89%</td>
</tr>
<tr>
<td>BhG sample</td>
<td>36</td>
<td>7</td>
<td>66(6%)</td>
<td>181</td>
<td>88%</td>
<td>92%</td>
</tr>
<tr>
<td>Average</td>
<td>1858</td>
<td>1350</td>
<td>86 (4.6%)</td>
<td>417</td>
<td>85.5%</td>
<td>91.5%</td>
</tr>
</tbody>
</table>

Table: Performance of Parser
<table>
<thead>
<tr>
<th>machine →</th>
<th>kartā (agent)</th>
<th>karma (goal)</th>
<th>adjective</th>
<th>pred adj</th>
<th>..</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>kartā (agent)</td>
<td>1322</td>
<td>14</td>
<td>10</td>
<td>26</td>
<td>..</td>
<td>1523</td>
</tr>
<tr>
<td>karma (goal)</td>
<td>31</td>
<td>883</td>
<td>7</td>
<td>..</td>
<td>1069</td>
<td></td>
</tr>
<tr>
<td>adjective</td>
<td>29</td>
<td>12</td>
<td>260</td>
<td>..</td>
<td>406</td>
<td></td>
</tr>
<tr>
<td>pred adj.</td>
<td>23</td>
<td>..</td>
<td>..</td>
<td>114</td>
<td>..</td>
<td>162</td>
</tr>
<tr>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>Total</td>
<td>1460</td>
<td>952</td>
<td>306</td>
<td>140</td>
<td>..</td>
<td>6226</td>
</tr>
</tbody>
</table>

Table: Confusion Matrix
Conclusion

- Improvised version of dependency relations
- Significant improvements
Thank You