Pāṇinian Syntactico-Semantic Relation Labels

Amba Kulkarni[1]   Dipti Misra Sharma[2]

[1]University of Hyderabad

[2]International Institute of Information Technology, Hyderabad, India

Depling 2019, August 28th, 2019

Indian Grammatical Tradition
provides a theoretical framework
to understand the two-way communication through Language

The two way communication consists of

- Transforming the thoughts in the minds of a speaker into a language string (Generation)
- Deciphering a language string by the listener (Analysis)

The two way communication process consists of

- Transforming the thoughts in the minds of a speaker into a language string (Generation)
    - Pāṇini's grammar
- Deciphering a language string by the listener (Analysis)
    - Theories of verbal cognition (śābdabodha)

Pāṇini's grammar

- Composed around 500 BC
- Aṣṭādhyāyī (8 chapters, with 4 parts each)
- Around 4000 aphorisms (sūtras), very much similar to mathematical concise formulae
  - minimum number of words
  - devoid of ambiguity
  - contain essence of the topic
  - universal in nature
  - without un-meaningful words
  - without any fault

4-levels in Pāṇini

According to Kiparsky, the grammar analyses sentences at a
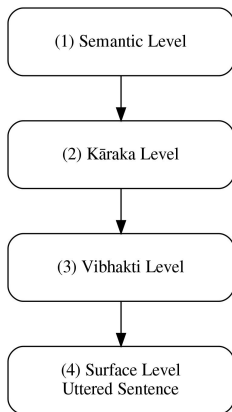hierarchy of 4 levels of description.

Figure: Levels in the generation process in Pāṇini

An activity of going from one place to the other by some person

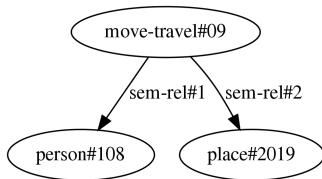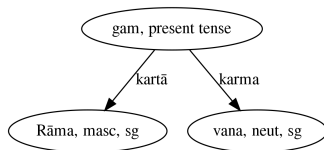Figure: Conceptual representation of a thought

Figure: Representation in abstract grammatical terms

Morphological Spellout rules

word index	stem	morphological features
1	Rāma masc	sg nom
2	vana neut	sg acc
3	gam parasmaipada class-1	laṭ 3p sg

Skt: Rāmaḥ vanam gacchati
Gloss Rama{masc sg nom} forest{neut sg acc} go{pres sg 3p}
Eng: Rama goes to the forest

Main focus: Semantic labels assigned to various participants of the activity
Lables: indicate the role of the participant in the activity.

Pāṇini classifies them into only 6 categories.

Kāraka relations

1. The participant which is the most independent to perform the activity is termed as *kartṛ*.

2. The participant which is the most desired by the *kartṛ* is termed as *karman*.

3. The thing which is most instrumental in bringing the action to accomplishment is called a *karaṇa* (instrument).

4. The participant which the agent wishes to reach through the object is termed *sampradāna* (beneficiary).

5. The participant which is fixed when there is a movement away is termed as an *apādāna* (source).

6. The participant which serves as a locus of an activity is called an *adhikaraṇa* (locus).

Extension of scope of the kāraka assignment rules:

- The associated semantics is totally different
- The extension to the semantics is not obvious

Extension of scope of the kāraka assignment rules:

- The associated semantics is totally different

  sthā (to stand) : locus as an argument
  adhi-sthā (to stand over, as well to govern)

  In the sense of 'to govern', the argument is an object
  (karman), and not a locus.

Extension of scope of the kāraka relations:

- The extension to the semantics is not obvious
  apādāna (source of separation)

  bhī (to be afraid of) : John is afraid of a lion.

  Lion : the source of fear (mental separation) : apādāna

- Kāraka (Predicate-argument) relations
- Non Kāraka relations such as
  - cause/reason (hetu)
  - purpose (prayojana)
  - precedence (pūrvakāla)
  - .
  - .
  - .

Granularity

- Ramakrishnamacharyulu(2009) collected a list of all such relations
  from the texts on the theories of verbal cognition
- Around 100 relations
- Too fine-grained for mechanical processing

Granularity

- A subset of these relations was selected for mechanical processing (Kulkarni)
- The core relations for different Indian languages is common with a few language specific variations.

Salient Features

- The relations are binary.
- All relations are between words denoting concepts.
- Underspecified relations are provided to handle the complexity in processing.
- Most of the relation names are the same as found in the Pāṇinian tradition.
- A few new relations, which were not found in Pāṇinian grammar, are added. These correspond to certain accompanying terms (*upapada*) that govern the case markers of the accompanying word.
- These dependency relations are found to be suitable for automatic parsing with high accuracy.
- The labels are also comprehensible by non-grammarians.
- These relations are also found to be appropriate for both parsing as well as generation.

Purely Syntactic

- duplication of words pervading, several, successive order, series, distributiveness, repetition, and so on (vīpsā)
- Genitive case marker
  part-and-whole, kinship, possession, ...
- Pair of arguments (arg1 and arg2)
  To denote inter-sentential relations

All other relations are purely semantic.

What is the semantics associated with the relation Kartṛ ?

Kartṛ is not a subject

(1) Skt: Rāmaḥ pāṭhaṁ paṭhati
Gloss: Rama{nom.} lesson {acc.} read {pr tense 3p sg}
Eng: Rama reads a lesson.

(2) Skt: Rāmeṇa pāṭhaḥ paṭhyate
Gloss: Rama{ins.} lesson {nom.} read {passive pr tense 3p sg}
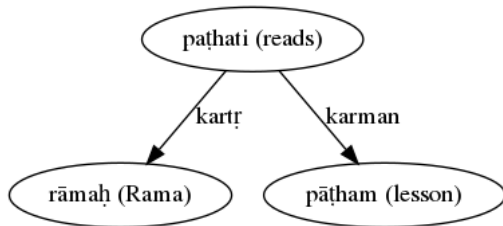Eng: The lesson is read by Rama.
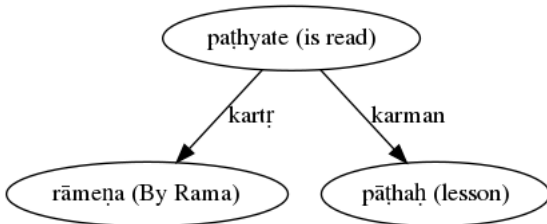
Rama is a Kartṛ in both the sentences.

Figure: analysis of an active sentence

Figure: analysis of a passive sentence

Is kartṛ an agent ?

Kartṛ is not an agent

1) Skt:*rāmaḥ kuñcikayā tālam udghāṭayati.*
Gloss: Rama{nom.} key{ins.} lock{acc} open{pr tense 3p sg}.
Eng: Rama opens the lock with a key.

Thematic: Rāma : Agent
Pāṇinian: Rāma : Kartṛ

Kartṛ is not an agent

2) Skt:*śyāmā kuñcikā tālam udghāṭayati.*
Gloss: Black{nom.} key{nom} lock{acc.} open{pr tense 3p sg}.
Eng: The black key opens the lock.

Thematic: key: Instrument
Pāṇini: key: Kartṛ

Kartṛ is not an agent

3) Skt: *tālaḥ udghāṭyate.*
Gloss: Lock{nom.} open{pr tense 3p sg}.
Eng: The lock opens.

Thematic: lock: Theme
Pāṇini: lock: Kartṛ

What is the semantics associated with Kartṛ?

Pāṇini defines Kartṛ as
The independent participant in an activity

Opening of a lock: three sub-activities

1. the insertion of a key by an agent,
2. pressing of the levers of the lock by an instrument (key), and
3. moving of the latch and opening of the lock.

1-3 : $open_1$
2-3 : $open_2$
3 : $open_3$

Substantive playing the role of kartṛ decides the meaning of the verb.

In Hindi,
$open_1$ and $open_2$ → khola
$open_3$ → khula

2) Skt:*śyāmā kuñcikā tālam udghāṭayati.*
Gloss: Black{nom.} key{nom} lock{acc.} open{pr tense 3p sg}.
Eng: The black key opens the lock.

Thematic: key: Instrument
Pāṇini: key: Kartṛ

- In order to assign the thematic role, one has to appeal to the extra-linguistic information
- Also in doing so, one would miss the underlying semantics associated with the verb in the given context.

- Pāṇini has identified
  - How much information is coded in a language string
  - Gave it a semantic interpretation
- This level is reachable through grammar rules alone
- It puts an upper bound on the analysis without any extra-linguistic information

Sanskrit Parser using Pāṇinian dependencies

Skt: *dṛṣṭvā tu pāṇḍavānīkam vyūḍham duryodhanaḥ tadā |*
ācāryam upasaṅgamya rājā vacanam abravīt || (BhG 1.2)

Gloss: After_seeing[1] the_army_of_the_Pāṇḍavas
arranged_in_military_phalanx Duryodhana at_that_time, teacher
approached King words spoke

Eng: At that time, after seeing the army of the Pāṇḍavas arranged
in military phalanx, King Duryodhana approached (his) teacher
and spoke (these) words.

---

[1] *tu* here is just a filler for metrical purpose

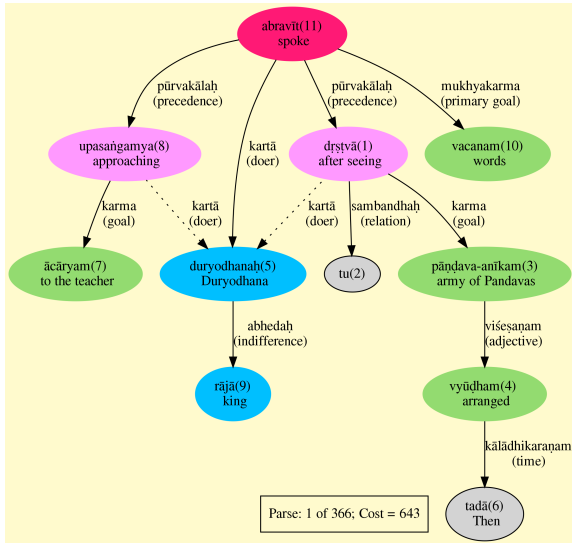# Sanskrit Parser using Pāṇinian dependencies

Figure: Parsed output of the BhG 1.2 verse

- Pāṇinian dependencies offers well-defined semantics for relations that can be extracted purely from a language string
- The same set of relations is useful for both generation and analysis
- Plan eclectic use of rule-based and machine learning approaches for developing better parsers.

Dhanyavādaḥ

Merci

Thank you