Treating utthita and utthāpya ākāṅkṣās independently for efficient parsing

Amba Kulkarni and Pavan Kumar

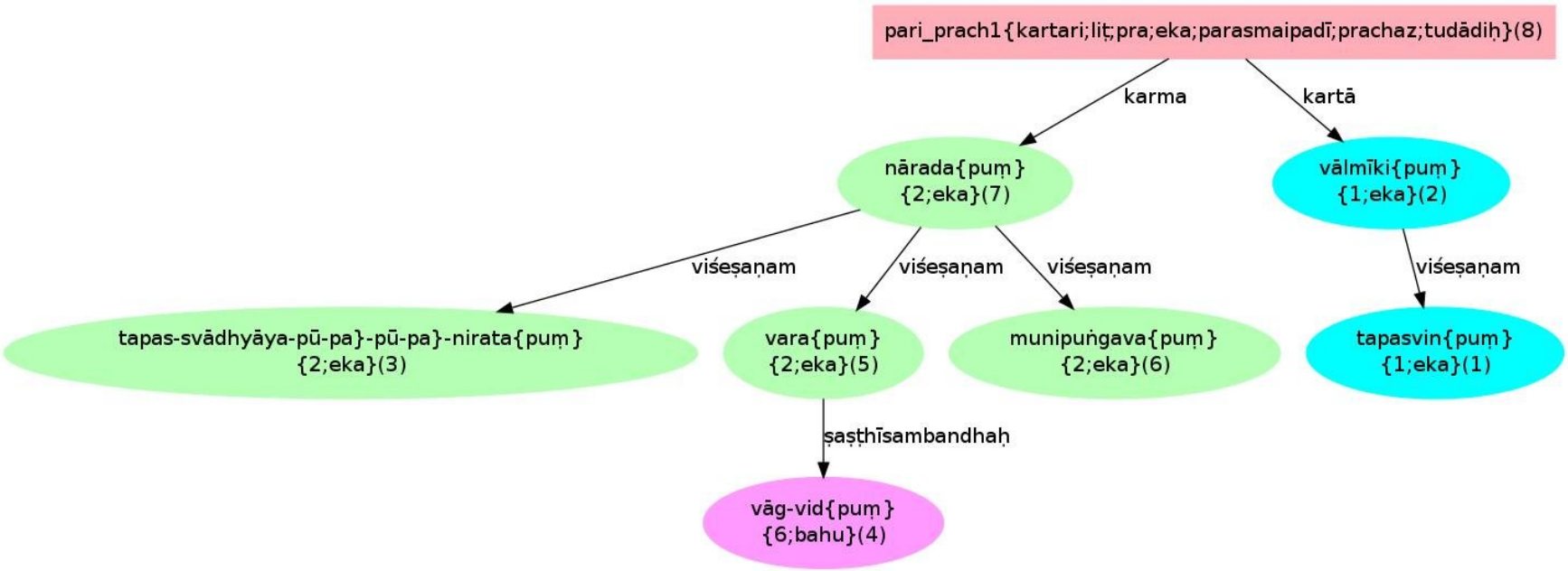
Department of Sanskrit Studies,
University of Hyderabad,
Hyderabad

Motivation

tapas-svādhyāya-niratam tapasvī vāg-vidām varam
nāradam paripapraccha vālmīkiḥ munipuṅgavam

gloss: Valmiki, who performs penance, asked Narada who is continuously engaged in self-study and penance, who is greatest amongst the wise men, and who is greatest among sages.

Parsed Output



Multiple Parses

Machine produces more than 1 parses, since it does not have world knowledge.

The total number of parses for this śloka are 128.

Permutations of (tapasvī, vālmīki) = 2

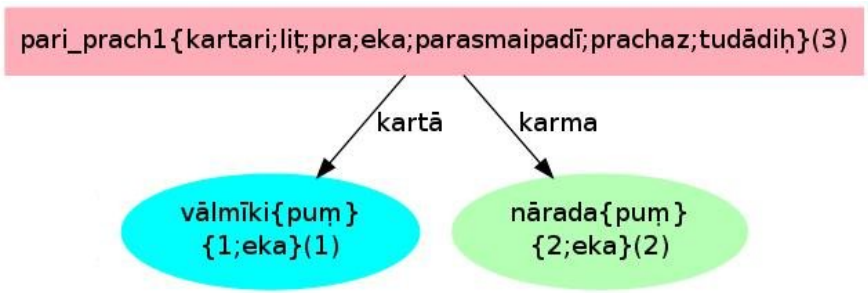
Permutations of (tapas-svādhyāya-niratam, varam, munipuṅgavam, nāradam) = 64

Complexity of parses

If there were no adjectives

for example:

vālmīkiḥ nāradam paripapraccha.



There is only one parse.

contd ...

Adding an adjective => multiple parses.

for example:

tapasvī vālmīkiḥ nāradam paripapraccha.

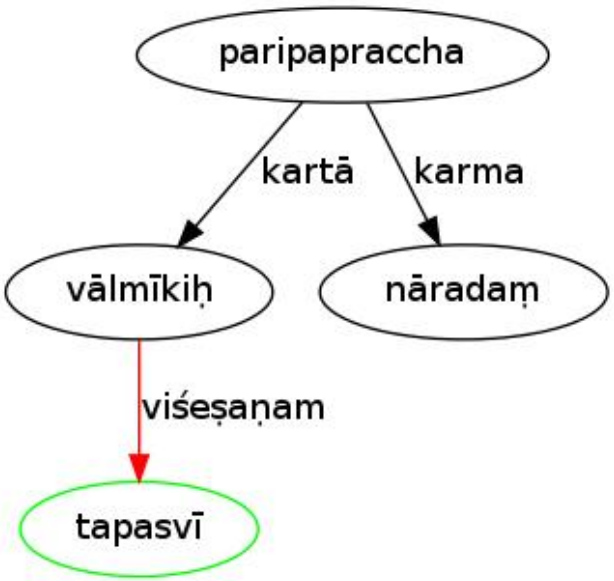


Fig.1

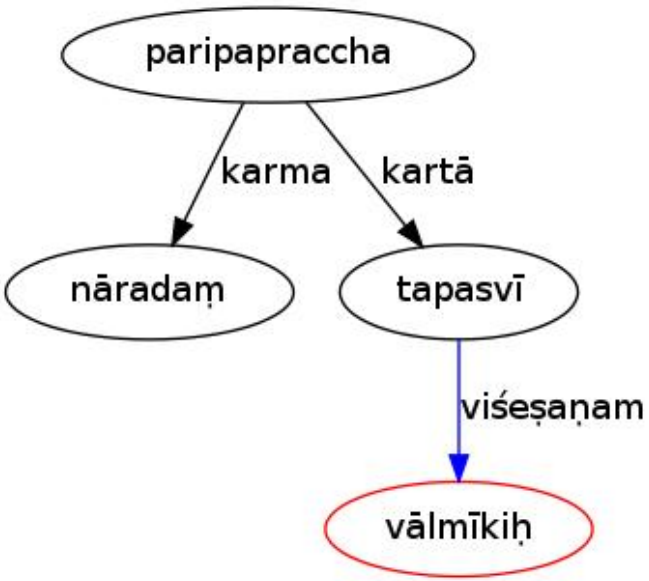


Fig.2

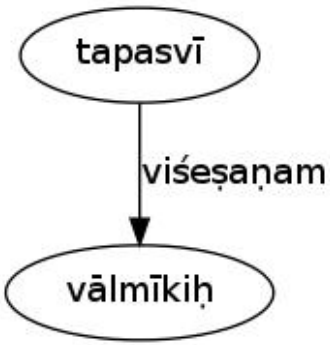
contd ...

Reason : Lack of Semantic Information.

vālmīkiḥ : jāti vācakaḥ
tapasvī : guṇa vācakaḥ.

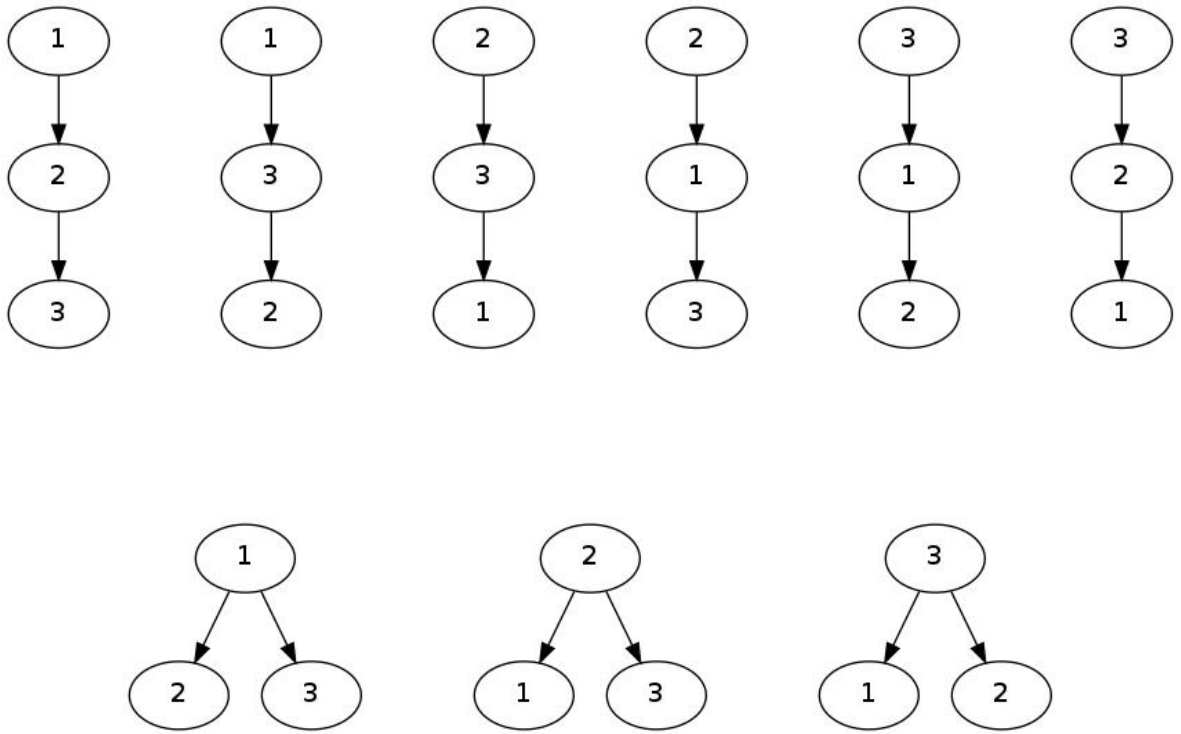
One adjective

One adjective - 2 parses



Two adjectives

Two adjectives - 9 parses



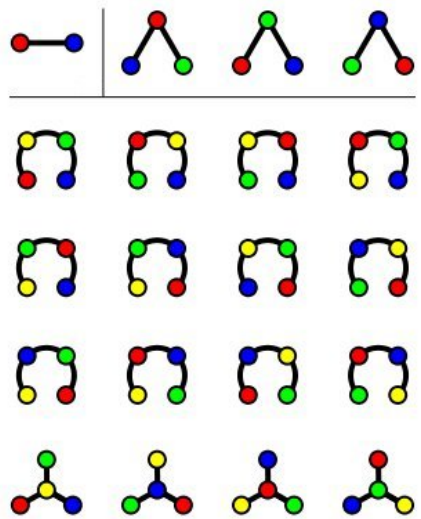
Spanning Tree

G : Complete Graph with 'n' nodes.

=> Every node is connected to every other node by a direct edge.

t(G) = no of Spanning Trees

t(G) = n^{n-2} (Cayley's formula) where the edges are undirected.



Contd...

t(G) = n*n^{n-2}, directed edges

adjective	nodes	no of spanning trees
1	2	2 * $2^0 = 2$
2	3	3 * $3^1 = 9$
3	4	4 * $4^2 = 64$
4	5	5 * $5^3 = 625$

Conclusion

Adjectives increase the parses.

Do adjectives occur frequently?

If Yes, how frequent are they?.

Contd...

A small Empirical Evidence

Viśeṣaṇā's occur as frequently as frequent kāraka relations

Relation	gītā	stories
kartā	1277	4683
karma	1270	4399
adhikaraṇam	876	3465
viśeṣaṇam	1280	3735
Total Relations	8279	18823

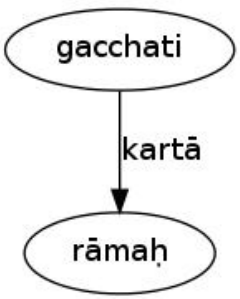
kāraka

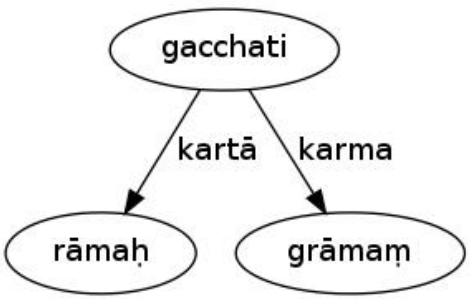
kāraka - adds new dimensions

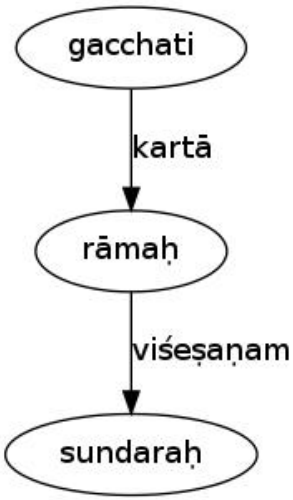
viśeṣaṇa - Qualifies the existing dimensions

Bone vs Flesh

Information point of view







Contd ...

Need to treat viśeṣaṇa's differently from the kāraka relations.

ākāṅkṣā

There are two kinds of ākāṅkṣā's mentioned in Tradition.

- utthitākāṅkṣā
- utthāpyākāṅkṣā

contd....

नियताकाङ्क्षा यथा क्रियाकारकपदानां परस्पराकाङ्क्षा। यथा वा पुत्रादिपदानां पित्रादिरूपप्रतियोगिवाचकपदाकाङ्क्षा नियता (रामरुद्री १ पृ० २)। नियतत्वं च अत्र अवश्यंभावित्वम्। अत्र च पुत्र इत्युक्ते कस्य इत्याकाङ्क्षोदयात् देवदत्तस्य इत्याद्यपेक्षत इति देवदत्तस्येति षष्ठ्यन्तपदपुत्रपदयोर्नियताकाङ्क्षेति ज्ञेयम्। - Nyāyakośaḥ

utthitākāṅkṣā

- Mutual expenctancy between the verb and kārakās and Relational Words
- Relational words such as putraḥ, bhāryā etc.
- This Expectancy is regular/restricted/controlled.

  for example:-

  putraḥ - kasya?
  pidhehi - kim?
  dvāram - kim karaṇīyam?

utthāpyākāṅkṣā

अनियताकाङ्क्षा (रामरुद्री पृ० २)। यथा उदयति चन्द्रः कुमुदबान्धव इत्यादौ चन्द्रपदकुमुदबान्धवपदयोरनियताकाङ्क्षा। - Nyāyakośaḥ

utthāpyākāṅkṣā

- Unidirectional Expectancy.
- This Expectancy is irregular/unrestricted/uncontrolled.

  for example:-

  sundaraḥ - kaḥ
  rāmaḥ - X

How the Parser works?

Consider the sentence:

rāmaḥ vanaṃ gacchati

1. rāmaḥ = rāma {masc.} {sg.} {nom.}
2. rāmaḥ = rā {pr.} {1p} {pl.}
3. vanaṃ = vana {neu.} {sg.} {nom.}
4. vanaṃ = vana {neu.} {sg.} {acc.}
5. gacchati = gam {pr.} {3p.} {sg.}
6. gacchati = gam {pr. part.} {masc.} {sg.} {loc.}
7. gacchati = gam {pr. part.} {neu.} {sg.} {loc.}

How the Parser works?

1. rāmaḥ = rāma {masc.} {sg.} {nom.}
2. rāmaḥ = rā {pr.} {1p} {pl.}
3. vanaṃ = vana {neu.} {sg.} {nom.}
4. vanaṃ = vana {neu.} {sg.} {acc.}
5. gacchati = gam {pr.} {3p.} {sg.}
6. gacchati = gam {pr. part.} {masc.} {sg.} {loc.}
7. gacchati = gam {pr. part.} {neu.} {sg.} {loc.}

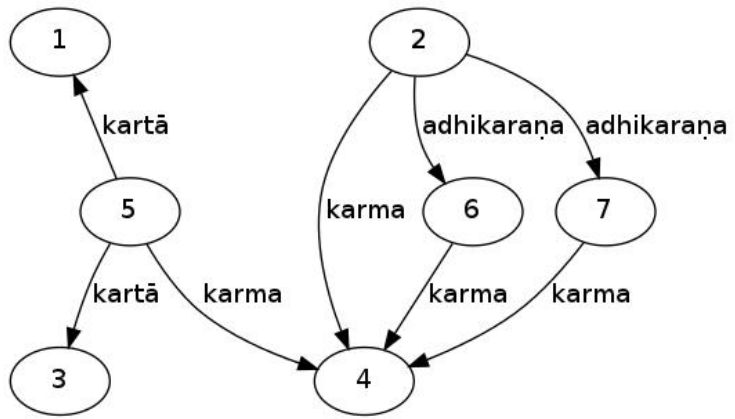


Figure: Possible Relations

How the Parser works?

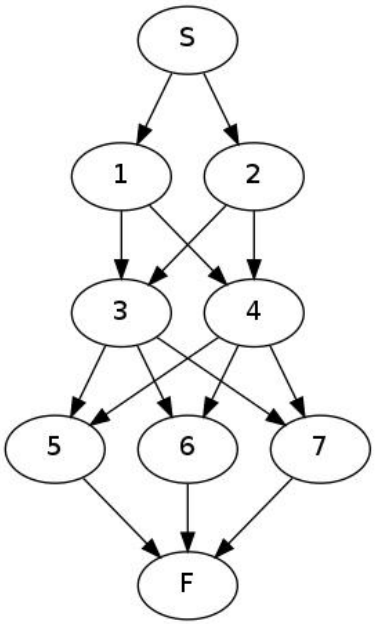


Figure: Adjacency and Possible paths

1. rāmaḥ = rāma {masc.} {sg.} {nom.}
2. rāmaḥ = rā {pr.} {1p} {pl.}
3. vanaṃ = vana {neu.} {sg.} {nom.}
4. vanaṃ = vana {neu.} {sg.} {acc.}
5. gacchati = gam {pr.} {3p.} {sg.}
6. gacchati = gam {pr. part.} {masc.} {sg.} {loc.}
7. gacchati = gam {pr. part.} {neu.} {sg.} {loc.}

A **path** P is a sequence of edges which connects the nodes from 'S' to 'F'.

How the Parser works?

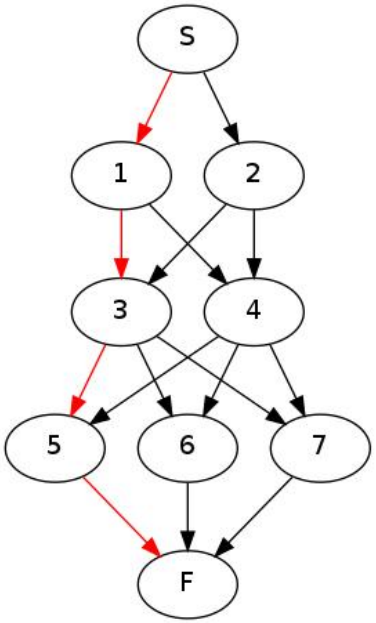


Figure: A sample Path

For example, S-1-3-5-F is a path.

How the Parser works?

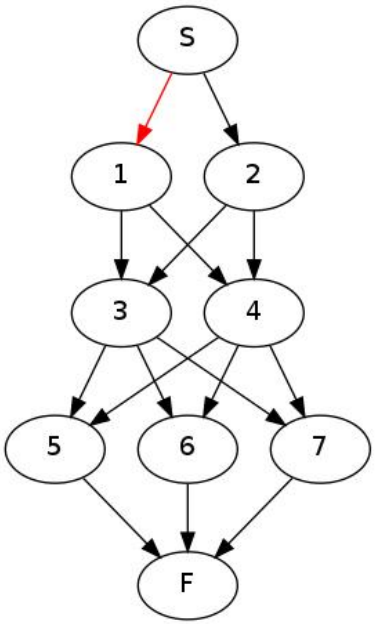


Figure: Adjacency and Possible paths

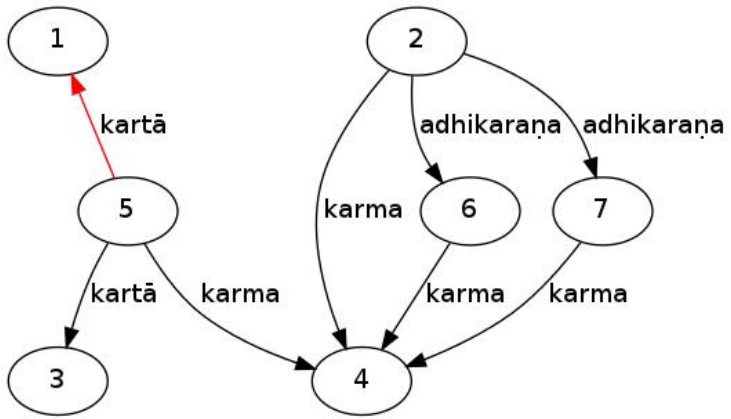


Figure: Relations

How the Parser works?

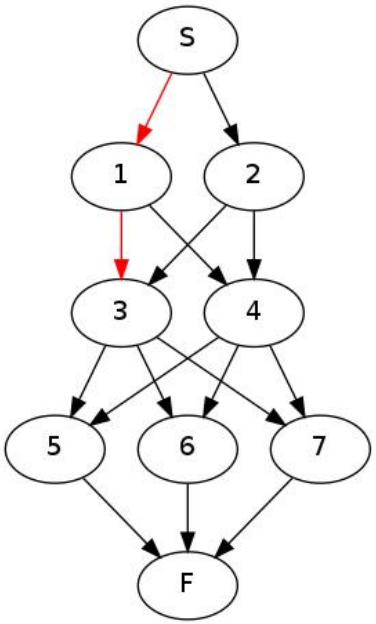


Figure: Adjacency and Possible paths

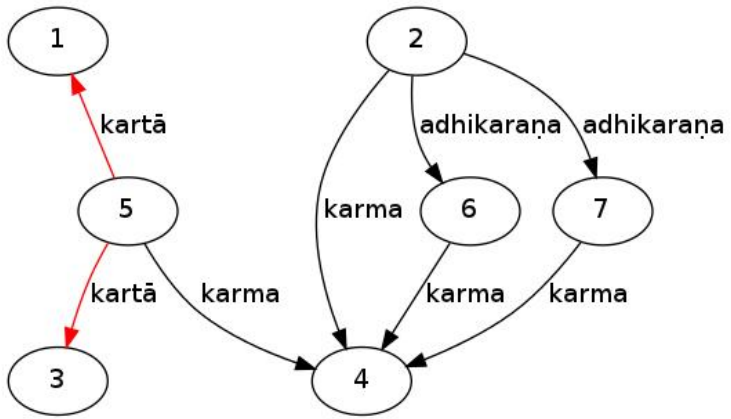


Figure: Relations

How the Parser works?

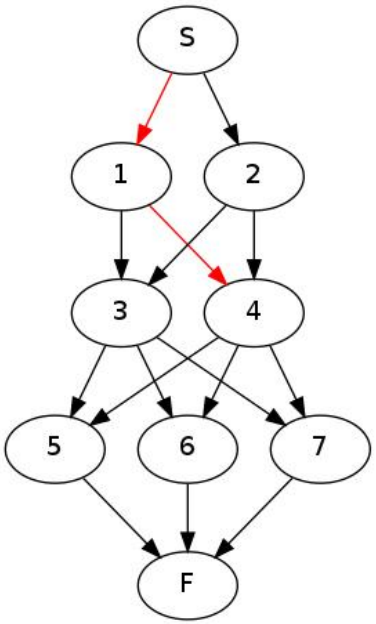


Figure: Adjacency and Possible paths

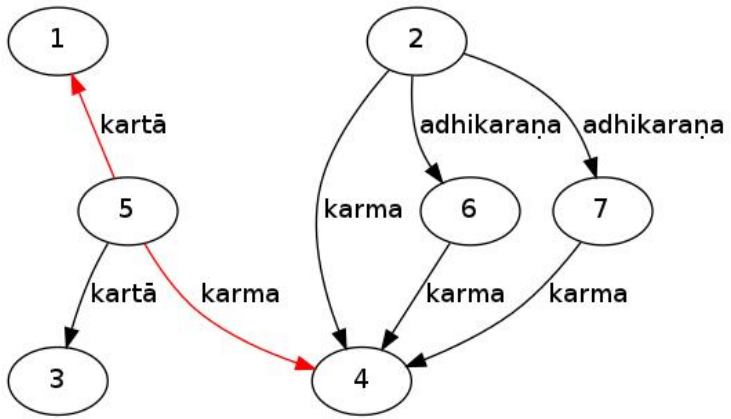


Figure: Relations

How the Parser works?

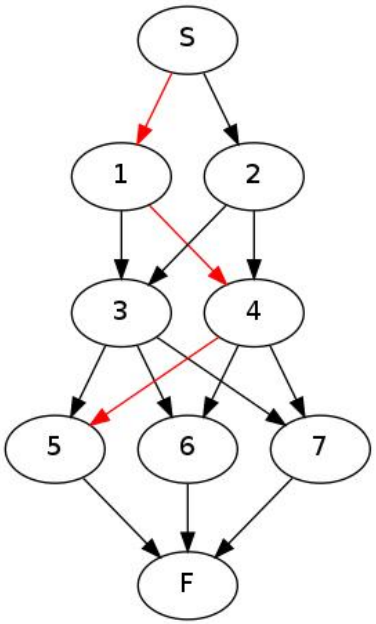


Figure: Adjacency and Possible paths

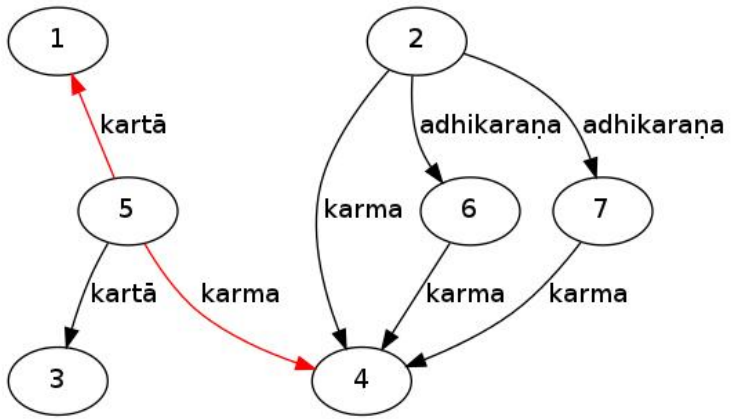


Figure: Relations

How the Parser works?

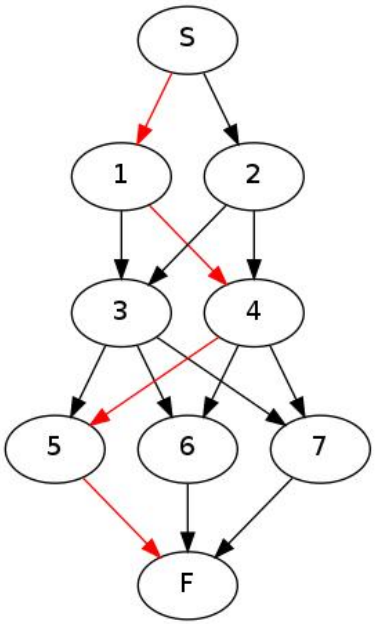


Figure: Adjacency and Possible paths

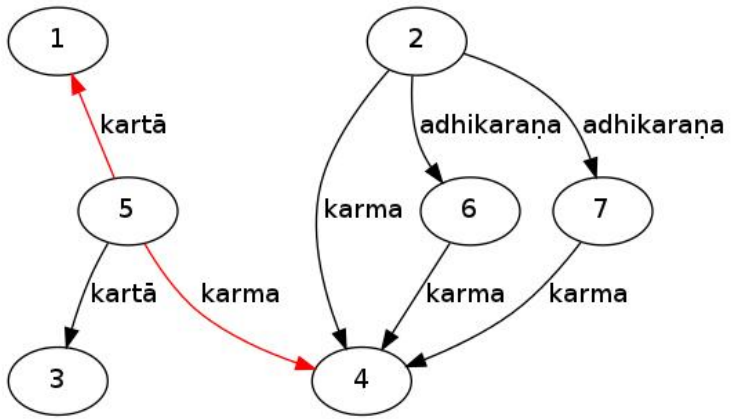
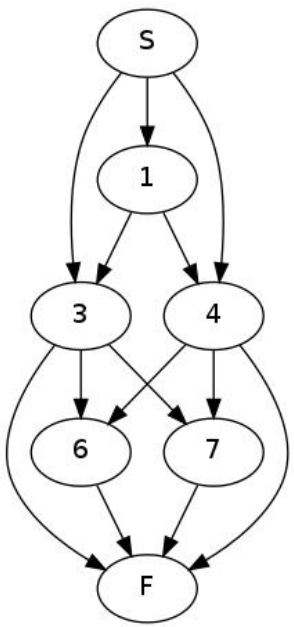
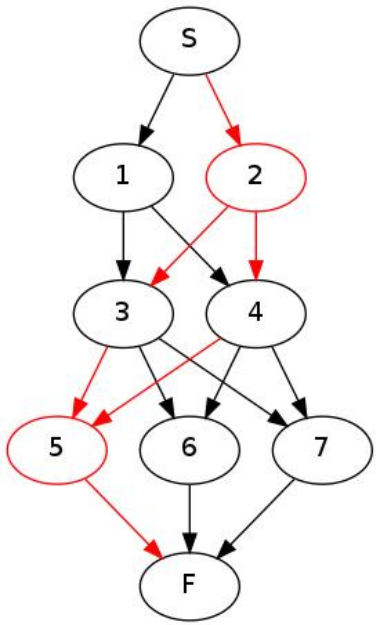


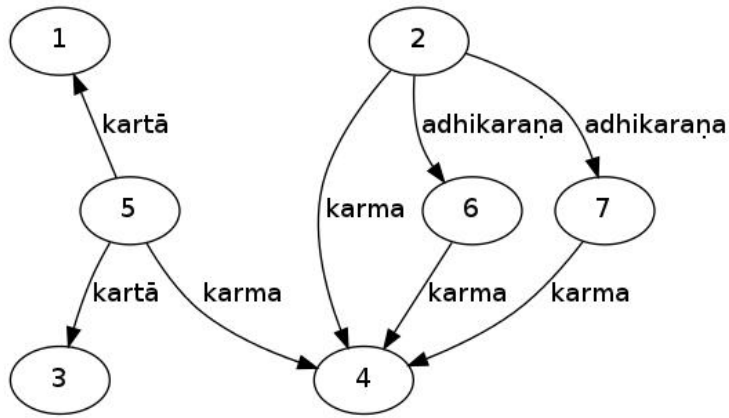
Figure: Relations

Modifications in the Algorithm:

- Bypass nodes corresponding to empty relations
  No incoming arrows corresponding to Nodes 2 and 5.
  Hence remove them.
  Modified graph showing paths is

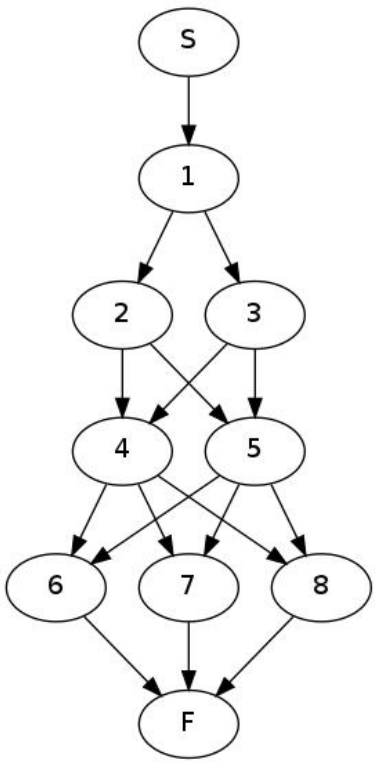


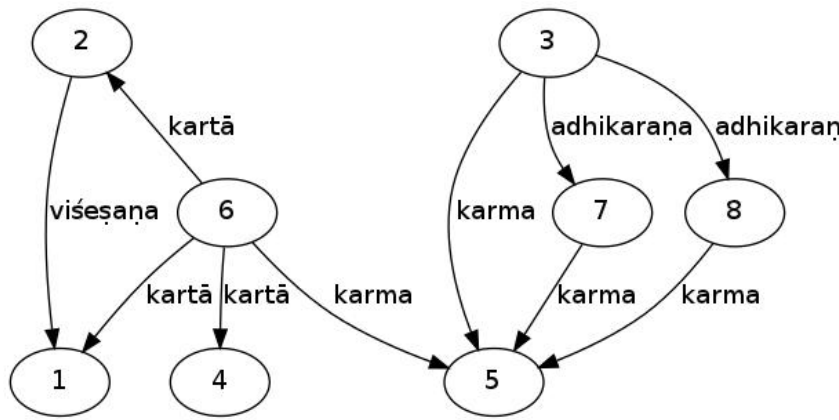




Modifications in the Algorithm:

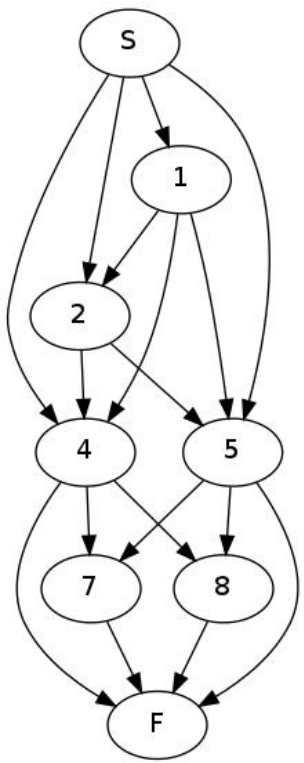
- Remove all the viśeṣaṇa relations, and put them in a viśeṣaṇa bag.
  Sentence: śūraḥ rāmaḥ vanaṃ gacchati
  Original graph with viśeṣaṇa relation:

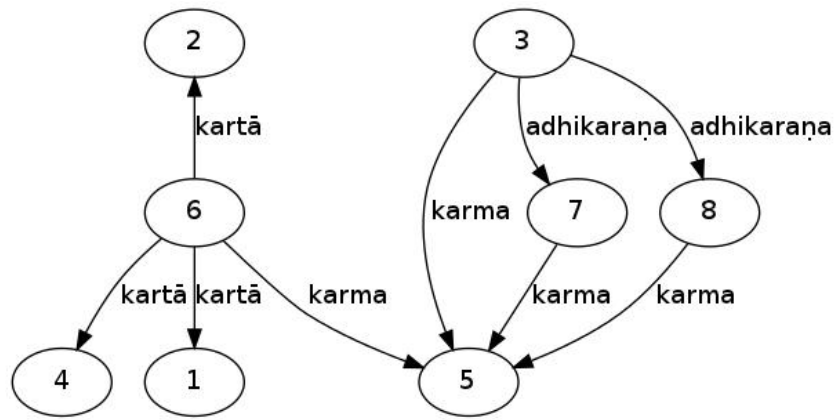




Modifications in the Algorithm:

- Remove all the viśeṣaṇa relations, and put them in a viśeṣaṇa bag.
  Sentence: śūraḥ rāmaḥ vanaṃ gacchati
  Original graph with viśeṣaṇa relation:

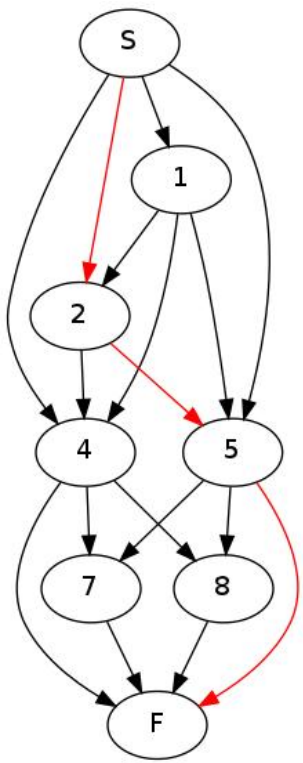


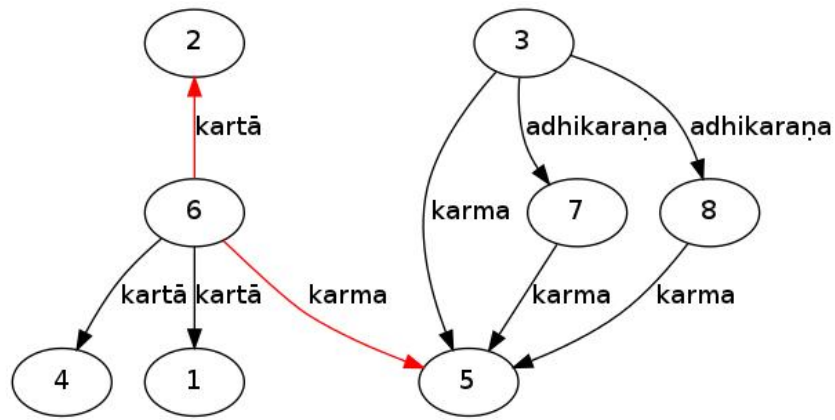


Viśeṣaṇa bag: | Node=2 | Node=1 |

Modifications in the Algorithm:

- Remove all the viśeṣaṇa relations, and put them in a viśeṣaṇa bag.
  Sentence: śūraḥ rāmaḥ vanaṃ gacchati
  Original graph with viśeṣaṇa relation:

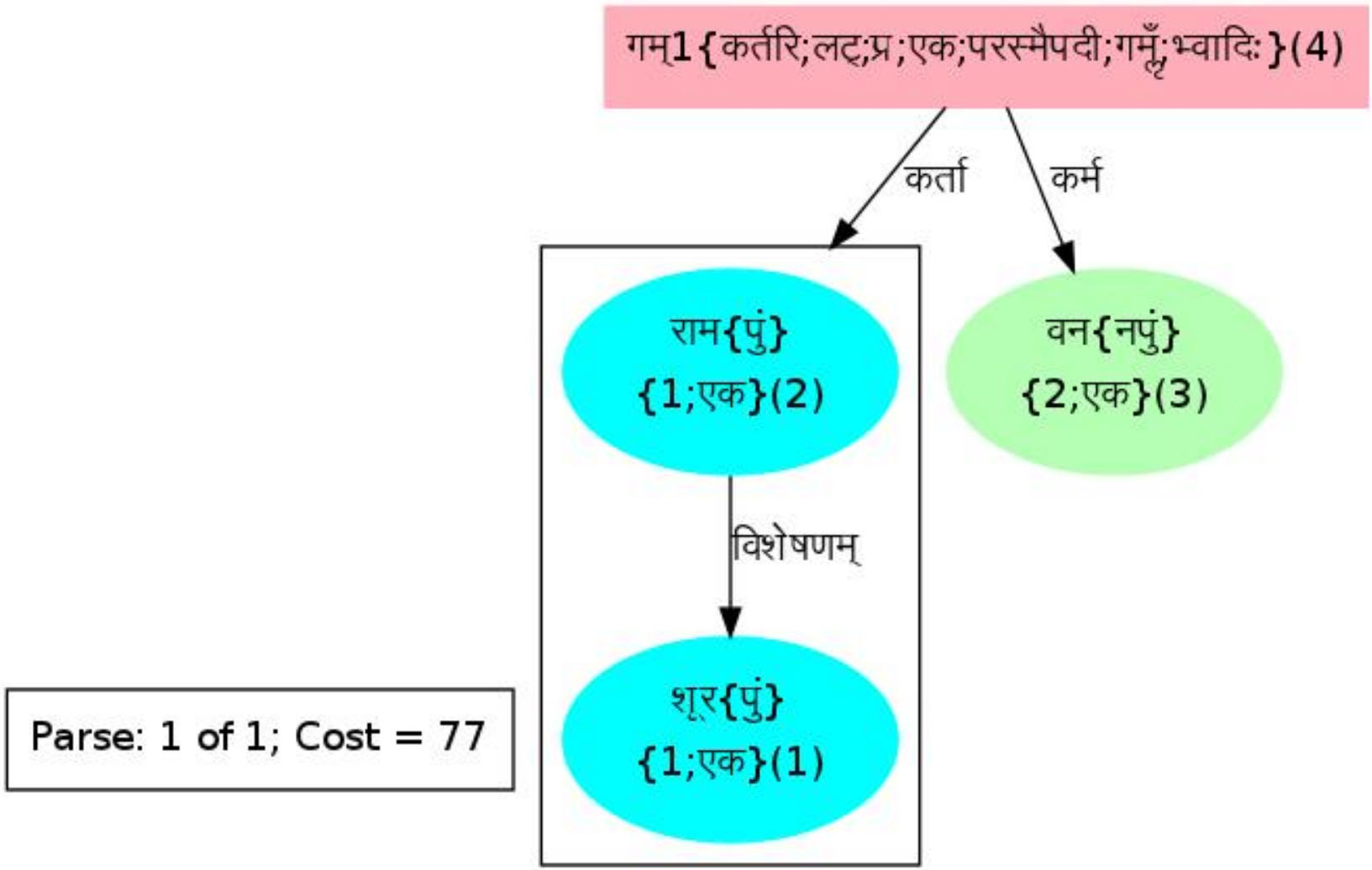




Viśeṣaṇa bag: | Node=2 | Node=1 |

Parsed output for

śūraḥ rāmaḥ vanaṃ gacchati



Parsed output for

tapasvī vālmīkiḥ tapas-svādhyāya-niratam vāg-vidām varam munipuṅgavam nāradam paripapraccha