Extra Linguistic Information needed for Automatic Generation of Sanskrit Compounds: A study

Pavankumar and Amba Kulkarni

Department of Sanskrit Studies,
University of Hyderabad,
Hyderabad

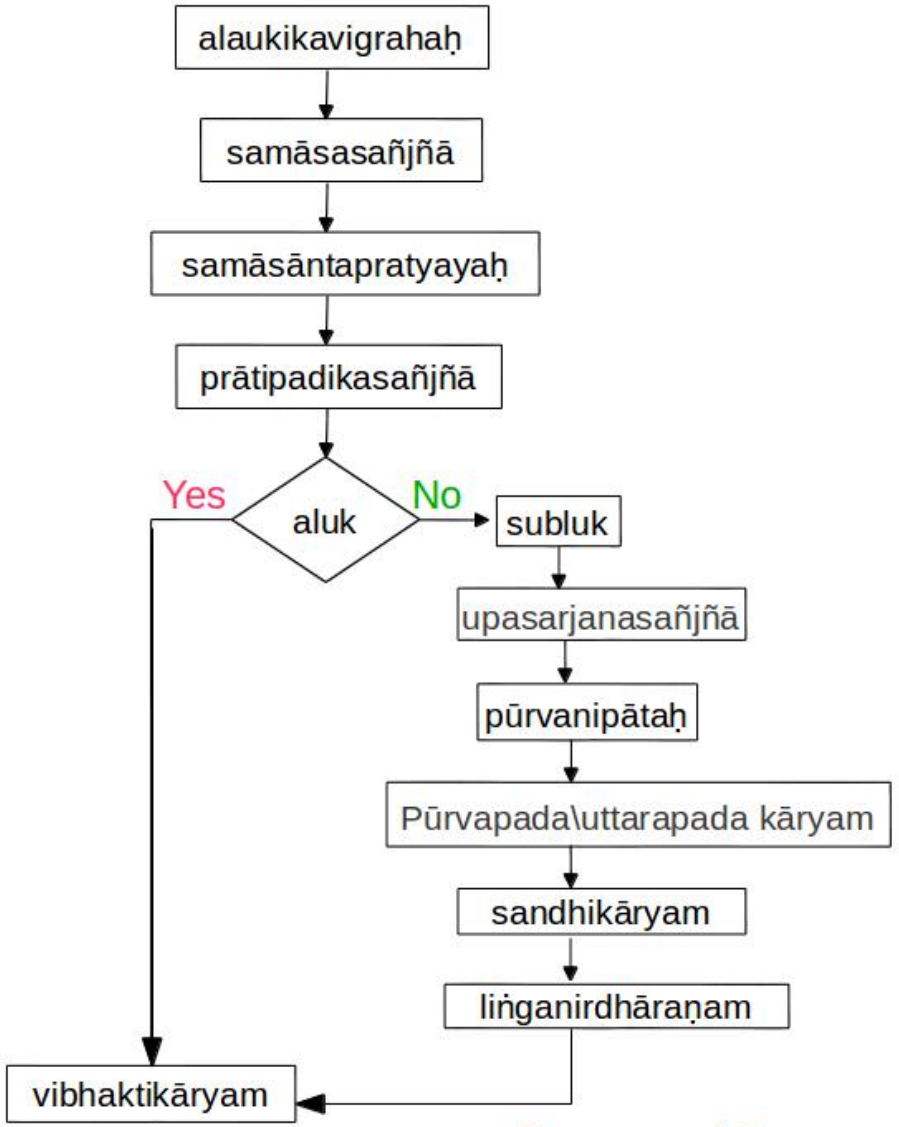
Generation of Sanskrit Compounds

Input: a) Concept
Blue Lotus
nīlam utpalaṃ

b) Intermediate Linguistic Expression (alaukikavigraha)
nīla + su utpala + su type : karmadhāraya

Output: nīlotpalam

Process of Compound Generation



(B1+S1 B2+S2)

(B1+S1 B2+S2) samāsa

(B1+S1 B2+S2 S3)

(B1+S1 B2+S2 S3) prātipadikam

(B1+[S1] B2+[S2] S3)

Decide the order

B1 ⟶ B3 /_ B2

B2 ⟶ B4 / B1_

B1' + B2' = B'

Decide the Gender

B = Base
S = Suffix

Process of Compound Generation

Grammar for Compound Generation

padam:	subantam	
|	tiṅantam	suptiṅantam padam
;		
subantam:	stem sup	
;		
stem:	$stem_{NT}$	
|	$stem_{T}$	
;		
$stem_{NT}$:	kṛt	
|	taddhitaḥ	
|	samāsaḥ	kṛtaddhitasamāsāśca
;		
samāsa:	alaukikavigrahaḥ	
;		
alaukikavigrahaḥ:	stem '+' sup stem '+' sup	
|	avyaya stem '+' sup	
|	stem '+' sup avyaya	
;		

Lexer for Compound Generator

```
sup:    su|au |jas
    |   am |auṭ |śas
    |   tā |bhyām |bhis
    |   ṅe |bhyām |bhyas
    |   ṅasi |bhyām |bhyas
    |   ṅas |os |ām
    |   ṅi |os |sup
    ;
avyaya  su
    |   upa
    |   prati
    ;
stem_T  [a-zA-Z]+arthavadadhāturapratyayaḥ prātipadikam
    ;
```

Vidhisūtrās

Vidhisūtrās are further classified into three types:

1. Rule with Right context
2. Rule with Left context
3. Rule with Extended right context

Rule with Right Context

$W_1\ W_2 \rightarrow W_3\ W_2$.

pādasya padājyātigopahateṣu P6.3.52

pāda → *pada*/_*āji*

Rule with Left Context

$W_1\ W_2 \rightarrow W_1\ W_3$.

jyotirāyuṣaḥ stomaḥ P8.3.83

stoma → *ṣṭoma*/*jyotis*_

Rule with Extended Right Context

$W_1\ W_2 \rightarrow W_1\ W_2\ W_3$.

antarbahirbhyām ca lomnaḥ P5.4.117

antar $|$*bahir loman* $\rightarrow$ *antar* $|$*bahir loman ap*

Contd ...

$$B_1 \rightarrow B_3 \ / \ \_ \ B_2$$

$$B_2 \rightarrow B_4 \ / \ B_1 \ \_$$

Most of the sūtrās require sequence of phonemes but there are some sūtrās which require a rich structure involving following features:

- The morphological analysis of a word,
- The semantic category of a word,
- The meaning of a word, and
- Cases where user intervention is needed

Morphological analysis of a word

svayam kṛta + su → svayaṁkṛtam
svayam ktena P2.1.25

Semantic category of a word

go+ su garbhiṇī + su → gogarbhiṇī
catuṣpādo garbhiṇyā P2.1.71

Meaning of a word

antika + ṅasi āgata + su → antikādagataḥ
stokāntikadūrārthakṛcchrāṇi ktena P2.1.39

Cases of User Intervention

- Words with multiple meanings
- Relations between two components
- The property of the referent of the compound
- Property of the referent of the second component

Words with multiple meanings

Following words with multiple meanings can form compounds only in certain meanings.

Words	Meanings
yathā	asādṛśye
yāvat	avadhāraṇe
prati	mātrārthe
āṅ	maryāda and abhividhi
abhi and prati	ābhimukhye
anu	samayā
anu	āyāma
kim	kṣepe
akṣi	adarśana
uras	agrākhya

Relation between two words

- viśeṣaṇaviśeṣyabhāvaḥ[1](modifier-modified relation)
  eg: nīlotpalam
- upamānopameyabhāvaḥ[2](Relation of Resemblance)
  eg: ghanaśyāmaḥ
- kriyakarākabhāvaḥ[3](Verb Agreement)
  eg: śaṅkulākhaṇḍaḥ
- atyantasaṃyogaḥ[4](Total contact)
  eg: muhūrtasukham

[1] *viśeṣaṇaṃ viśeṣyeṇa bahulam 2.1.57*
[2] *upamānani samānyavacanaiḥ 2.1.55*
[3] *tṛitīyā tatkṛtārṭhena guṇavacanena 2.1.30*
[4] *atyantasaṃyoge ca 2.1.29*

Property of the Referent of the compound

- pūjyamāna[5](Respect)
  eg : sadvaidyaḥ
- jāti and saṃjñā[6](class and Proper name)
  eg : upānasam
- jānapadākhya[7](Name of the inhabitant of a country)
  eg : surāṣṭrabrahmaḥ

[5] *sanmahatparamottamotkṛṣṭāḥ pūjyamānaiḥ 2.1.61*
[6] *ano'śmāyassarasāṃ jātisaṃjñayoḥ 5.4.94*
[7] *brahmaṇo jānapadākhyāyām 5.4.104*

Property of the Referent of the second component

vandita[8](Praised)
eg: praśastabhrātā, subhrātā

[8] *vandite bhrātuḥ 5.4.157*

Statistics

We examined 400 sūtrās out of 5% sūtrās needed user intervention.

Problems

- jāti paripraśna[9]
- kāravācī[10]

[9] *katarakatamau jātipariprašne 2.1.63*
[10] *kāranāmni ca prācāṃ halādau 6.3.10*

Thank you