Issues in the computational processing of Upamā alaṅkāra

Bhakti Jadhav[1], Amruta Barbadikar[2], Amba Kulkarni[2], and Malhar Kulkarni[1]

[1]Department of Humanities and Social Sciences, Indian Institute of Technology Bombay.
[2]Department of Sanskrit Studies, University of Hyderabad, Hyderabad.
bhaktij96@gmail.com, amruta.barbadikar@gmail.com,
ambakulkarni@uohyd.ac.in, malharku@gmail.com

Abstract

Processing and understanding of figurative speech is a challenging task for computers as well as humans. In this paper, we present a case of Upamā alaṅkāra (simile). The verbal cognition of the Upamā alaṅkāra by a human is presented as a dependency tree, which involves the identification of various components such as upamāna (vehicle), upameya (topic), sādhāraṇa-dharma (common property) and upamādyotaka (word indicating similitude). This involves the repetition of elliptical elements. Further, we show, how the same dependency tree may be represented without any loss of information, even without repetition of elliptical elements. Such a representation would be useful for the computational processing of the alaṅkāras.

1 Introduction

Sanskrit is one of the oldest natural languages with rich variety of literature ranging from *Vyākaraṇa* (grammar), *Śikṣā* (phonetics), *Nirukta* (etymology), etc. to *Mīmāmsā* (exegesis), *Nyāya* (logic), *Kāvyaśāstra* (poetics), *sāhitya* (literature), *nāṭyaśāstra* (dramaturgy), *dharmaśāstra* (jurisprudence), etc. *Kāvyaśāstra* (poetics) is one of the subjects that has been dealt with elaborately in the Sanskrit tradition. The study of *alaṅkāras* (that can loosely be translated as 'Figures of speech') constitutes an important part of the study of this branch. Figures of speech are an integral part of language. They are constantly used in common parlance as well as in literature. They contribute to the richness of the language and increase its effectiveness. In Sanskrit Poetics, the figures of speech have been categorized into 3 types:

1. Śabdālaṅkāra- of words.

2. Arthālaṅkāra- of senses.

3. Ubhayālaṅkāra- of both.

In arthālaṅkāras, Upamā is one of the oldest figures of speech mentioned in the earliest text of Kāvyaśāstra, that is Nāṭyaśāstra of Bharata. Upamā serves as the basis for several other figures of speech such as *apahnuti*, *sasandeha*, *utprekṣā*, *rūpaka*, etc. that are based on similarity. In this figure of speech, the central idea is to compare one object with another due to both possessing a common characteristic. Herein, a comparison takes place by measuring the object of comparison closely with the standard of comparison. Such a comparison may intend to exalt or degrade an object.

1. To exalt- e.g., *mukham candraḥ iva ramaṇīyam*- 'The face is as pleasant as the moon.'

2. To degrade- e.g., *durjanavachaḥ halāhalam iva dussaham*- 'The speech of a wicked person is as unbearable as the poison called Halāhala.[1]'

The notion of the English figure of speech- Simile (Qadir et al., 2016), is similar to the concept of *Upamā*. *Upamā* consists of four components, viz.

1. ***Upameya*** – The object of comparison, which is known as 'topic' in simile.

2. ***Upamāna***- The standard of comparison, that is known as 'vehicle' in simile.

3. ***Sādhāraṇadharma***- common property on grounds of which two objects are compared, which is known as 'event' or 'state' in simile.

4. ***Upamādyotaka*** (that also is known as *dyotaka* or *vādi*) is the word that expresses similarity or *Upamā*, which is known as 'comparator' in simile.

---

[1]Halāhala was a deadly poison that emerged from the ocean due to its churning. This is a reference from *puraṇas*

Processing of figurative language is a challenging domain in Natural Language Processing (NLP). Here are a few instances of computational work in this field in other languages. Mpouli (2017) proposed a framework for annotating similes in a literary text in the English language. Audichya and Saini (2021) proposed a specific methodology for the identification of the three Hindi 'Alankaaras'. Wöckener et al. (2021) designed an end-to-end model for poetry generation based on conditioned recurrent neural networks (RNN). Niculae (2013) worked on automatic recognition and classification of comparisons and similes. Kesarwani et al. (2017) focuses on metaphor detection in a poetry corpus. Naaz and Singh (2022) mapped the aesthetic properties of poetry onto a numeral scale for the Hindi language. For the Sanskrit language, Barbadikar and Kulkarni (2023, 2024) designed a tool to identify and classify Yamaka and Anuprāsa, which are the types of *Śabdālaṅkāra*. However, no work has been done in the field of *arthālaṅkāras* in Sanskrit language from the computational point of view.

The cognitive ability of the human mind is way more complex than a machine. *Upamā* is a poetic device where a creative analogy is presented. Unlike other simple sentences, sentences where *Upamā* is employed, the tagging scheme is expected to provide certain more specifications than the thematic roles. This level of parsing is higher than the *thematic roles* level. We aim to parse sentences with *Upamā* in a manner similar to how human beings understand it and then represent it within a parse tree.The current parser does not handle constructions with *Upamā alaṅkāra*. We propose an extension to this parser to handle *Upamā alaṅkāra* as well. First we present a diagrammatic representation of the human cognition of *Upamā*. Then, taking into account the limitation of the machine to deal with only words and not their meanings, we propose another representation that is feasible to arrive at automatically, yet not leading to any loss of information.

2 Background

Understanding a sentence with *Upamā alaṅkāra* is not easy. One needs to first identify the word expressing similitude. This is followed by the identification of the common property and then the *Upamāna* and upameya. Once these are identified, then in order to understand the overall meaning of

the sentence, typically the reader will construct two parallel sentences, one with *Upamāna* and another with *upameya* by providing the elliptic words if any. For example, in the sentence *mukham candraḥ iva ramaṇīyam asti* (The face is as pleasant as the moon.), the word *ramaṇīyam* (pleasant) is identified as the common property and then the *Upamāna* and *upameya* are identified as '*candra*' (the moon) and '*mukham*' (the face) respectively, with '*iva*' as the marker. Now, the understanding that takes place in the mind of a reader is as follows: *candraḥ ramaṇīyaḥ asti* (The moon is pleasant). On similar lines, *mukham ramaṇīyam asti* (The face is pleasant). Thus, the reader provides the elliptic words and constructs two separate independent sentences and understands them.

In order to develop a computational module for automatic handling of *Upamā alaṅkāra*, since the identification of elliptic words and reconstruction of two sentences with their repetition is not feasible, we propose to analyse the sentence, taking the advantage that the *Upamāna* and *upameya* in Sanskrit have the same syntactic behaviour as adjective-substantive. Further, the marker of *Upamā alaṅkāra* provides a clue for identifying the *Upamāna-upameya* relation. Taking these clues, we propose a module to enhance the existing Sanskrit parser.[2] In what follows, we first provide the classification of six types of *Upamā*. Each of these presentations includes a graphical representation of human understanding followed by the representation that we expect from the machine. In the fifth section, we discuss the potential problems in the implementation. This is followed by the conclusion and future scope for analysis of figures of speech.

3 About *Upamā alaṅkāra*

The literature in the Sanskrit language has dealt with rhetorics in quite a detail. The literature available in the tradition of *Kāvyaśāstra* on *Upamā alaṅkāra* discusses various definitions of this figure of speech, followed by its types and examples. Further, a few texts also explain the *śābdabodha* of *Upamā* instances in *Navya-Nyāya* style from the logicians' perspective. Here, we refer to *Mammaṭācārya*'s *Kāvyaprakāśa* (KP, henceforth) for classification of *Upamā alaṅkāra* into various types. We have chosen this text from *kāvyaśāstra* because the classification is easy and clear as it is

---

[2] https://sanskrit.uohyd.ac.in/scl/MT/index.html

done on syntactical grounds. This figure of speech has been divided into two types:

1. **Pūrnā** (Complete) – The type which consists of all four components of Upamā.

2. **Luptā** (*Elliptical*) – The type which has either one or two or even three components of *Upamā* missing.

This being a pilot study, we have considered the *Pūrnā* variety where all the factors are present for better analysis at the initial stage. The *Pūrnā* variety is further divided into 6 sub-types on the following two criteria:

- Whether the similarity is explicit or implicit:

    1. Where the similarity is explicit the variety is called 'Ṡrautī'.
    2. Where the similarity is implicit the variety is called 'Ārthī'.

- The morpho-syntactic form of expression of similarity:

    1. **Vākyagā**- Where the similarity is expressed in the form of a sentence.
    2. **Samāsagā**- Where the similarity is expressed in the form of a compound.
    3. **Taddhitagā** - Where the similarity is expressed in the form of *taddhita* formation using the *taddhita* affix vat.

4 Graphical representation

Herein, we give graphs for every variety of *pūrnopamā* in two ways. In the first type, we show how human intellect understands a sentence having *Upamā alaṅkāra*. In this pattern, we divide the expression into two sentences. One is the *Upamāna* sentence (sentence standing as topic) and another is the *Upameya* sentence (sentence standing as vehicle). And these two sentences are connected by the word indicating similitude.

While deriving these two different sentences from one sentence expressing *'Upamā alaṅkāra'*, the words denoting common property i.e., the state or event in Simile (*samāna-dharma*) have to be repeated in both. Such a word indicating 'state' can be an adjective, a headword (verb) or a noun. Also, there are a few words that are considered in both sentences but are not the 'state'. For example in the verse discussed in the section 4.3, the words '*bhuvanam*' and '*babhāra*' are repeated in both topic

and vehicle sentences. But they cannot be termed as 'state', for they do not contribute to the beauty of expression of 'Upamā'. On the contrary, they have been taken twice to complete the sentence and give the context. Human intellect can make such distinction by understanding the role and the relation of each word in the sentence based on meaning.

In pattern 2, the words are not repeated. A whole sentence is considered to be a single sentence. The reason for the non-repetition of the common property is as follows: The common property used for simile can be an action or a property, and it is only the context that tells us which one is being referred to. It may also involve extra linguistic information. For the Machine learning techniques to learn this, we need a huge corpus with such instances, preferably annotated. In the absence of such a corpus, it is not possible even for the Machine Learning algorithms to mark this common ground for comparison.

In the graphs the pattern code is as follows: Essentially the four main components of *Upamā* are shown in various shapes in both the patterns. The *Upameya* is given in dashed oval. The *Upamāna* is shown in dashed box. The *Dyotaka* is shown in a pentagon shape. The common property is given in a simple box in the human understanding graph, whereas in machine graphs just like other components of the sentence, it is given in a simple oval shape like other sentential relations. Common property has not been given a different shape in the machine graph because the machine cannot identify it. The dotted line in pattern 2 denotes the supplement of ellipsis ('*adhyāhāra*') of the *Dyotaka* (comparator) when there are more than one pair of *Upameya* and *Upamāna* and only one comparator is provided. The first pattern does not require the node of *Dyotaka*. We now discuss every variety of *pūrnopamā* and provide its diagrammatic representation. In this diagram, the main verb is connected to the *kārakas* and other indeclinables such as 'na'. Further, the *viśeṣya* is connected to *viśeṣaṇa*.

4.1 Śrauti vākyagā pūrṇā Upamā

In this type, all four components are present and the similitude is expressed by words such as *yathā*, *iva*, *vā* and *va* explicitly in the form of the complete sentence. The example from KP is as follows:

svapne'pi samareṣu tvāṃ
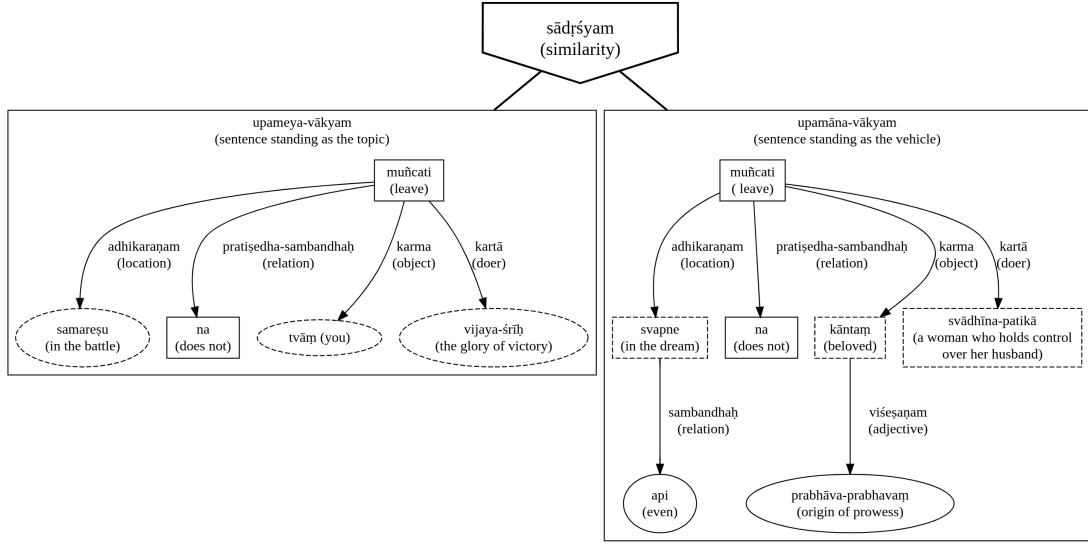*vijayaśrīrna muñcati |*
prabhāvaprabhavaṃ kāntaṃ

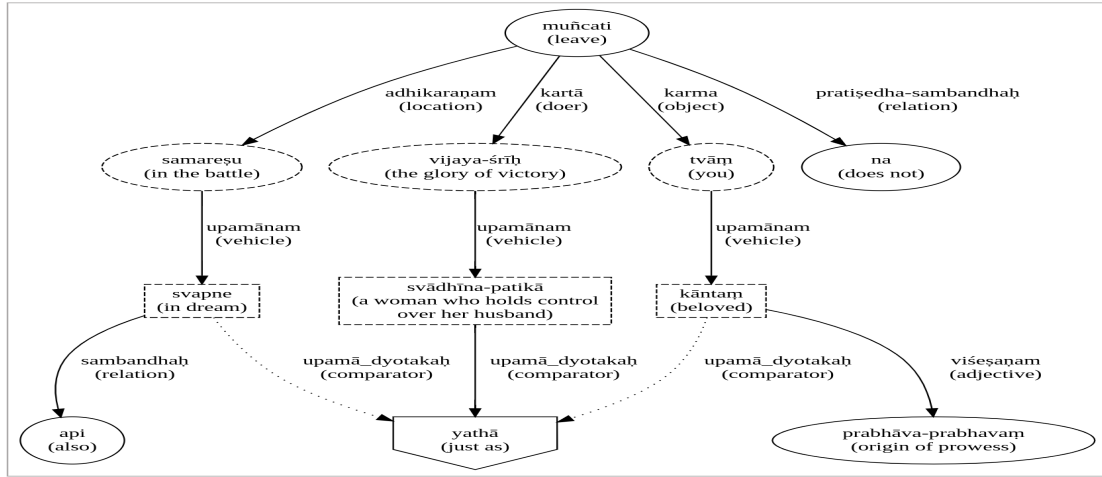Figure 1: Śrauti vākyagā pūrṇā Upamā: Pattern 1
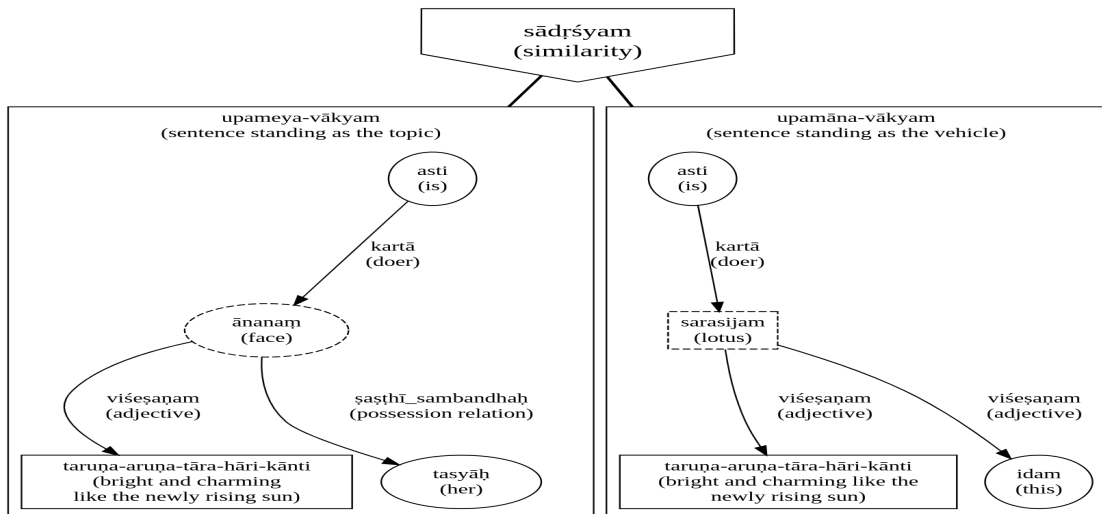

Figure 2: Śrauti vākyagā pūrṇā Upamā: Pattern 2
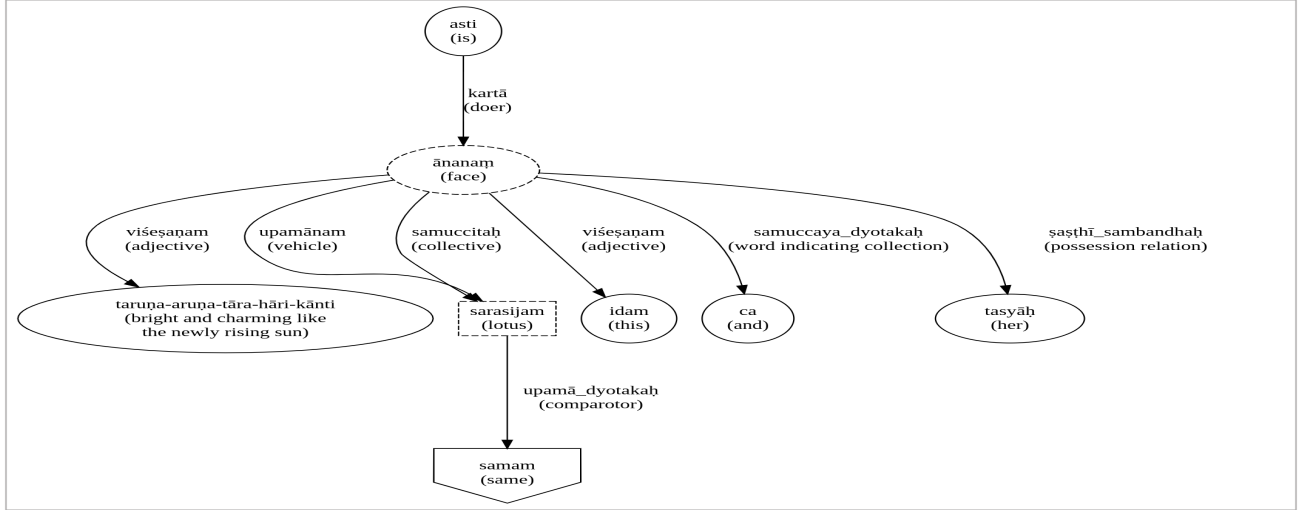

Figure 3: Ārthī vākyagā pūrṇā Upamā: Pattern 1

Figure 4: Ārthī vākyagā pūrṇā Upamā: Pattern 2

svādhīnapatikā yathā ॥

Meaning- Even in the dream (of your enemies), the glory of Triumph does not forsake you, the fountain-head of prowess in battles, just as a woman who holds her husband under her control, does not forsake him, her beloved, the fountain-head of ardent love, even in her dream.

Topic- vijayaśrīḥ (the glory of victory), Vehicle- svādhīnapatikā (a woman who holds control over her husband), State- amuñcana (to not forsake), Comparator- yathā (like)

- **Pattern 1**: In figure1, we have divided the Upamā instance into two sentences as perceived by the human intellect. The first sentence becomes the object of comparison and the second becomes the standard of comparison. These two sentences possess the similarity that is indicated by '*yathā*'. All components in the sentence are connected to the main verb '*muñcati*' by various relations such as kartā, karma, adhikaraṇa, sambandha, etc. 1. Sentence that stands as the topic- *vijayaśrīḥ samareṣu tvāṃ na muñcati* . Meaning- The glory of victory doesn't leave you in the battle-field.) 2. Sentence that stands as the vehicle- *yathā prabhāvaprabhavaṃ kāntaṃ svādhīnapatikā svapne api na muñcati* . Meaning- Just as a woman who holds control over her husband doesn't leave her beloved even in the dream.

- **Pattern 2**: In figure 2, the representation is given in the format that is convenient for the machine, where the verb is taken only once by avoiding the '*adhyāhāra* (supplying an ellipsis)'.

4.2 *Ārthī vākyagā pūrṇā Upamā*

In this variety of Upamā words such as *sadṛśam*, *tulyam* are used to indicate similarity. (See figure 3 and figure 4)

> *cakitahariṇalolalocanāyāḥ krudhi*
> *taruṇāruṇatārahārikānti |*
> *sarasijamidamānanaṃ ca tasyāḥ*
> *samamiti cetasi saṃmadaṃ vidhatte ॥*

Meaning- The face (of that heroine) of her, whose eyes are tremulous like that of a startled deer, when in anger, has the bright and charming lustre like a newly rising sun (new dawn). Such a face resembles the lotus which too has a bright and charming lustre like a new dawn. This very phenomenon of a face resembling a lotus creates immense joy in the mind of the hero.

Topic- *ānanam* (face), Vehicle- *sarasijam* (lotus), State- *taruṇāruṇatārahārakānti* (the lustre that is bright like the new dawn and charming), Comparator- *samam* (are equals).

- From the perspective of human understanding the two sentences are:

  1.   Sentence that stands as a topic - *tasyāḥ ānanam taruṇāruṇatārahārakānti asti*. Meaning- Her face has the lustre that is bright and charming like the new dawn.
  2.   Sentence that stands as a vehicle- *idam sarasijam taruṇāruṇatārahārakānti asti*.

Meaning - This lotus has a lustre that is bright and charming like the new dawn.

- Sentence for machine understanding is: *Tasyāḥ taruṇāruṇatārahārakānti ānanam, idam taruṇāruṇatārahārakānti sarasijam cha samam asti.*

  Meaning- Her face and this lotus, both of which have a shared attribute of having the lustre that is bright and charming like the new dawn, are similar.

4.3 *Śrautī samāsagā pūrṇā Upamā*

This variety of Upamā is possible because of a particular grammatical technicality.[3] (see figure 5 and figure 6) Here, *iva* is compounded with the preceding Upamāna word. In a normally compounded word, the case termination of the previous word is dropped, but in the case of compounding with iva, the previous word retains its case termination e.g., Rāmasyeva (Rāmasya + iva). Example-

> *atyāyatairniyamakāribhiruddhatānāṃ divyaiḥ*
> *prabhābhiranapāyamayairupāyaiḥ|*
> *śaurirbhujairiva caturbhiradaḥ sadā yo lakṣmīvilāsabhavanairbhuvanaṃ babhāra ॥*

Meaning- The king sustained this world with the four expediments (of statecraft), that were far-reaching (in their consequences) and curbed the arrogant (men) and were exalted and glorious, that could not be obstructed and were an abode of the free play of wealth; just as *Viṣṇu* sustains the world with his four arms, that are very long and curb the unruly (demons), that are divine and glorious, that cannot be harmed and are the place for the free play of Lakṣmī.

Topic- *upāyāḥ*; Vehicle- *bhuja*; State- *divyaiḥ*, *caturbhiḥ*; Comparator- *iva*

- From the perspective of human understanding the two sentences are:
  1. Sentence that stands as a topic - *yāḥ divyaiḥ, caturbhiḥ upāyaiḥ bhuvanaṃ babhāra.* Meaning- He sustained the world with four divine expediments.
  2. Sentence that stands as a vehicle- *śauriḥ divyaiḥ,caturbhiḥ bhujaiḥ bhuvanaṃ*

³By the *vārtika* of Kātyāyana muni- '*ivena nityasamāsaḥ vibhaktyalopaḥ pūrvapadaprakṛtisvaratvaṃ ca |*'

babhāra. Meaning - Sentence that stands as a vehicle- Lord Krishna sustained the world with (his)four divine arms.

- Sentence for machine understanding is: *śauriḥ divyaiḥ,caturbhiḥ bhujaiḥ iva yāḥ divyaiḥ,caturbhiḥ upāyaiḥ bhuvanaṃ babhāra.* Meaning- Just as Lord Krishna sustained the world with his four divine arms, the king sustained the world with his four divine expediments.

4.4 *Ārthī samāsagā pūrṇā Upamā*

avitathamanorathapathaprathaneṣu praguṇagarimagītaśrīḥ |
suratarusadṛśaḥ sa bhavānabhilaṣaṇīyaḥ
kṣitīśvara na kasya ॥

Meaning- O king, you who are like the divine tree (i.e., kalpataru), you the one whose glories about enlarging the path of fulfilment the desires (of people) are sung (who constantly keeps fulfilling the desires of people), for whom would you not be desirable? (See figure 7 and figure 8) Topic- *bhavan* (you), Vehicle- *surataru* (the divine tree), State - *Abhilaṣaṇīyatva* (desirability) Comparator- *sadṛśaḥ* (similar)

- From the perspective of human understanding the two sentences are:
  1. *Avitathamanorathaprathaneṣu praguṇagarimagītaśrīḥ bhavan abhilaṣaṇīyaḥ.* Meaning- You, whose glories are sung regarding enlarging the path of fulfilment of the desires (of people), are desirable.
  2. *Avitathamanorathaprathaneṣu praguṇagarimagītaśrīḥ surataruḥ abhilaṣaṇīyaḥ.* Meaning-The divine tree whose glories are sung regarding enlarging the path of fulfilment of the desires (of people), is desirable.

- Sentence for machine understanding is: *Avitathamanorathaprathaneṣu praguṇagarimgītaśrīḥ Suratarusadṛśaḥ bhavan abhilaṣaṇīyah.* Meaning- You, whose glories are sung regarding enlarging the path of fulfilment the desires (of people), like the divine tree, are desirable.

4.5 *Śrautī taddhitagā pūrṇā Upamā*

Taddhita affixes are special kind of secondary affixes that are attached to substantives and thus
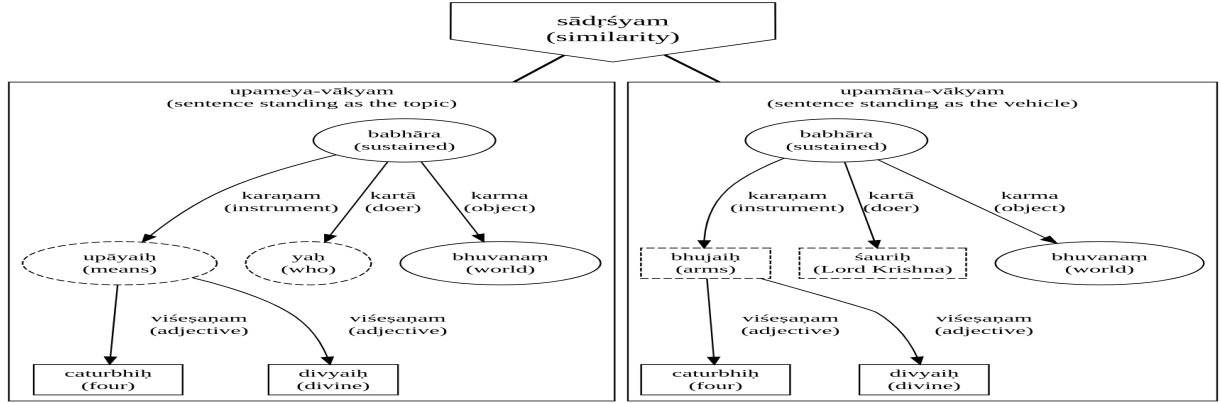
Figure 5: Śrautī samāsagā pūrṇā Upamā: Pattern 1

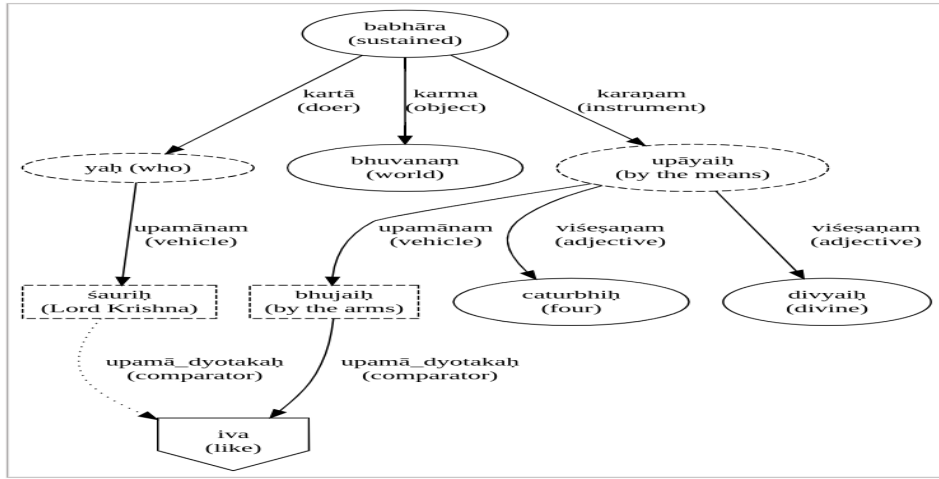Figure 6: Śrautī samāsagā pūrṇā Upamā: Pattern 2

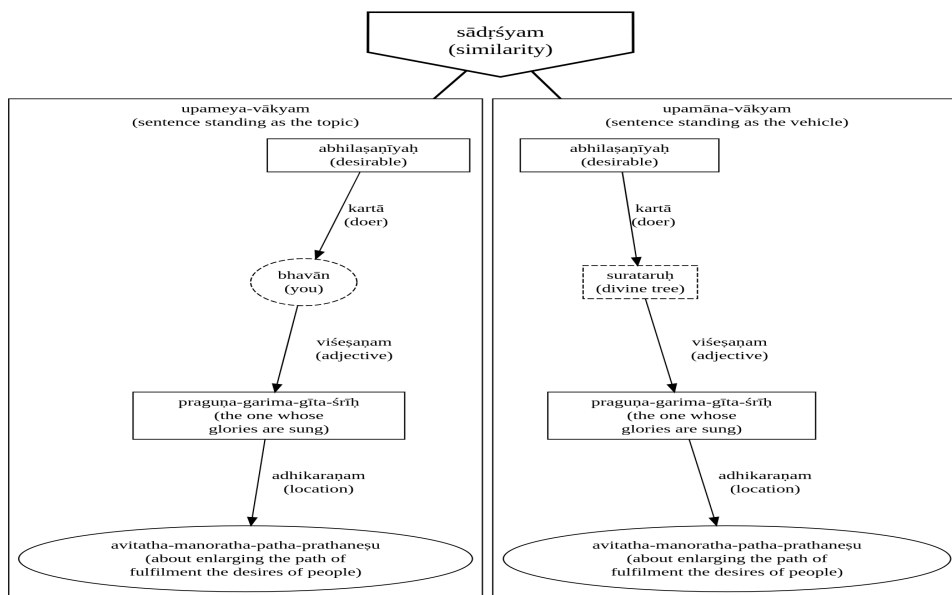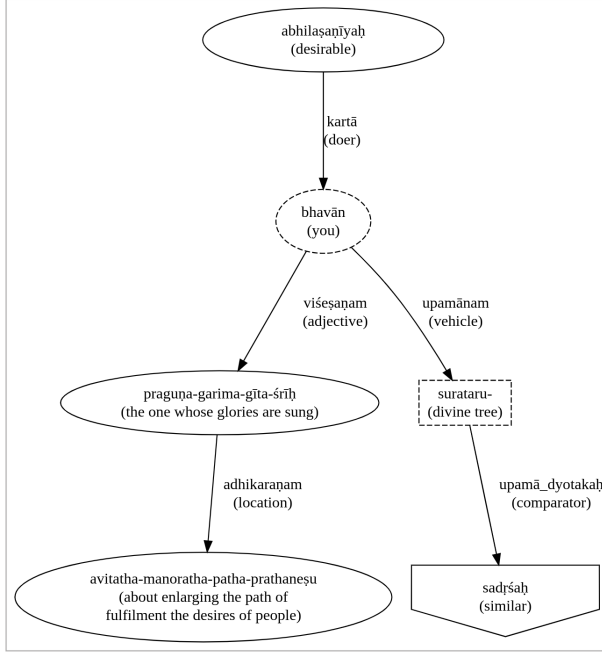Figure 7: ārthī samāsagā pūrṇā Upamā: Pattern 1

Figure 8: ārthī samāsagā pūrṇā
Upamā: Pattern 2

Figure 9: Śrautī taddhitagā pūrṇā
Upamā: Pattern 1

other words like *rāmavat* (like of Rāma), *ayo-dhyāvat* (like in Ayodhyā), *paurastya* (Eastern), *vaiyākaraṇa* (One who studies or has learnt grammar), etc. are derived. (See figure 9 and figure 10)
Example:

gāmbhīryagarimā tasya satyaṃ gaṅgābhujaṅgavat |

Meaning- The greatness of unfathomableness (of his mind) is indeed like the lover (Upapati) of Gaṅgā i.e., ocean (his mind is as deep as the ocean i.e. it cannot be known- durjñeya). Topic- (*saḥ*) *tasya* (his), Vehicle- *Gaṅgābhujaṅga* (Sea), State- *gambhīryagarimā* (unfathomableness), Comparator - *vat*

- From the perspective of human understanding the two sentences are:
  1. Sentence that stands as a topic: *Tasya gambhīryagarimā asti* Meaning- His mind is greatly unfathomable.
  2. Sentence that stands as a vehicle: *Gaṅgābhujaṅgasya gambhīryagarimā asti.* Meaning- The sea is greatly unfathomable.

- Sentence for machine understanding is: *Tasya gambhīryagarimā gaṅgābhujaṅgavat asti.* Meaning- His mind is as greatly unfathomable as the sea.
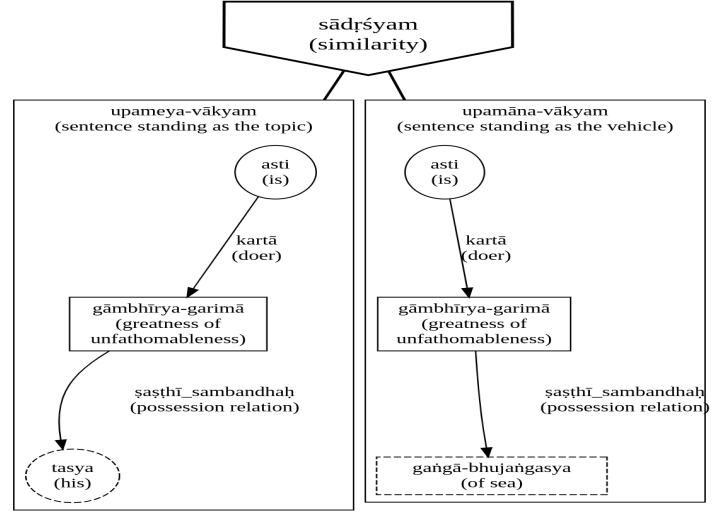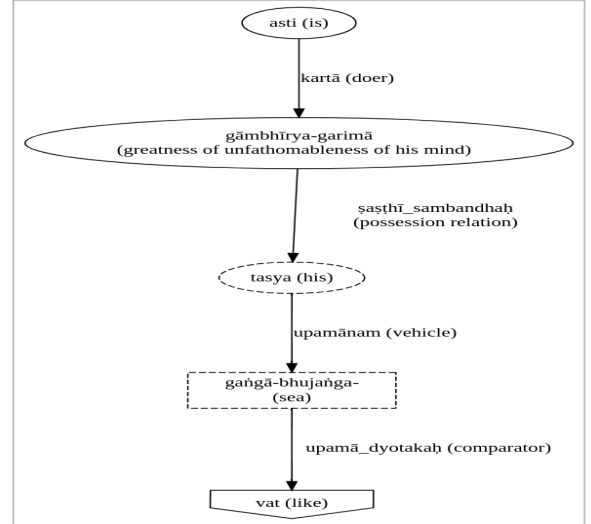

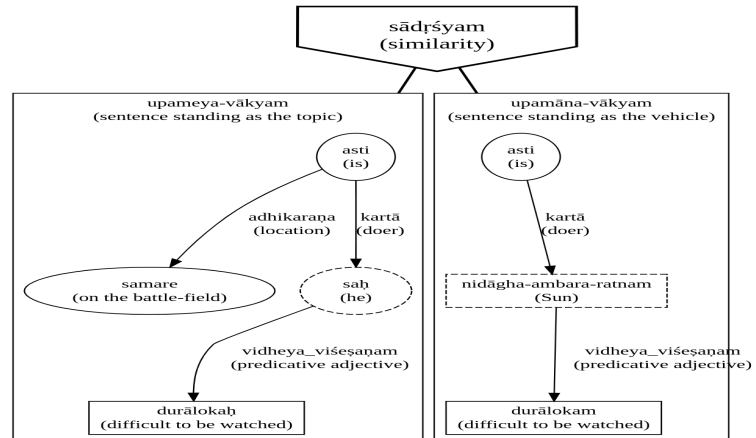Figure 10: Śrautī taddhitagā pūrṇā
Upamā: Pattern 2

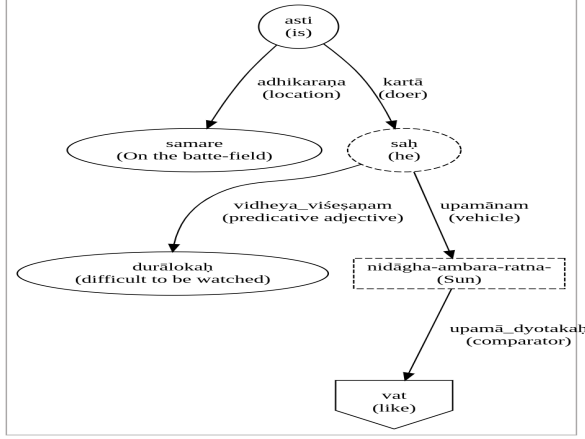Figure 11: Ārthī taddhitagā pūrṇā
Upamā: Pattern 1

Figure 12: Ārthī taddhitagā pūrṇā
Upamā: Pattern 2

Note- *Yathā gaṅgābhujaṅgasya 'gāmbhīryamahimā' tathā tasya rājñaḥ manasaḥ 'gāmbhīryamahimā' durjñeyāntakaraṇatvam iti bhāvaḥ |*- the similarity in the relation of genitive case is employed. In this case, *vat* is attached to the Upamāna.[4]

4.6 *Ārthī taddhitagā pūrṇā Upamā*

When the similarity between the object of comparison and standard of comparison is of action and it is denoted by a taddhita suffix, it is Ārthi Taddhitagā Pūrṇopamā. See figure 11 and figure 12, the Upamāna here is in Nominative case and the then affix *vat* is attached to it. It occurs when the affix *vat* is added to nouns according to the sūtra from Aṣṭādhyāyi.[5] Here, the affix (*vat*) comes after a word in the Instrumental case in construction, in the sense of 'like that' when the meaning is 'similarity of action'. Example-

> *durālokaḥ sa samare nidāghāmbararatnavat |*

Meaning- On the battlefield, he is very difficult to watch at (his sight is unbearable) like that of the sun in summer. The components here are: Topic- *saḥ* (he), Vehicle - *Nidāghāmbararatna* (Sun), State- *durālokatva*, Comparator – *vat*.

- From the perspective of human understanding the two sentences are:
  1. Sentence that stands as a topic: *Saḥ samare durālokaḥ asti.*
  Meaning- He is difficult to be watched in the battlefield. 2. Sentence that stands as a vehicle: *Nidāghāmbararatnam durālokam asti.* Meaning- The Sun is difficult to watch in Summer.

- Sentence for machine understanding is: Saḥ samare nidāghāmbararatnavat duralokaḥ asti. Meaning- On the battlefield, he is very difficult to be watched like the sun in summer.

5 Potential problems in the implementation of the proposed representation

In this context, we anticipate a few challenges that a machine would face during implementation:

- One of the primary challenges for a machine is to identify that, of the given pairs of words, which word is the *upameya* and which word is *Upamāna*. This is because, in most of the cases there is an agreement of case and number between two words. This problem is more specific to the *Ārthī vākyagā pūrṇā Upamā* variety since the comparator is disjoined from the *Upamāna* (vehicle) as in the variety 4.2.

- In the case of the two varieties viz. 4.5 *Śrautī taddhitagā pūrṇā Upamā* and 4.6 *Ārthī taddhitagā pūrṇā Upamā*, the comparator is the same i.e., *vat*. Hence, when an instance of pūrṇā Upamā having *vat* is given to the machine it would recognize it as *taddhitagā* but would not be able to decide whether it is a *Śrautī* or *Ārthī* variety. Further, in *Śrautī taddhitagā pūrṇā Upamā* whether the relation is possession or location also has to be decided. A human intellect decides this based on several other factors such as context, intuitive understanding, logic, etc. Developing this skill of discernment in a machine would be a challenge. It is possible to decide this only when the meaning or definition for the suffix is provided to the machine.

- In multiple examples we see that both *upameya* and *upamāna* have not only the case agreement but also the gender and number are the same. In such a situation, if a machine encounters an adjective within the sentence or verse, it faces a challenge in determining whether it should be associated with the topic

---

[4]tatra tasyeva |5.1.116, Aṣṭādhyāyi of Pāṇini, The affix *vat* comes in the sense of 'like what is therein or thereof' by this sūtra

[5]*tena tulyam kriya chet vatiḥ |*, 5.1.115, Aṣṭādhyāyi of Pāṇini

or the vehicle. Furthermore, if an adjective itself is a common property (ground), then it must be linked with both the topic and the vehicle, as in the variety 4.2

Finding solutions to these problems is a real challenge. The answers to these problems can be sought with the help of of ontological classification of such objects. The Word Net[6] or the Concept Net or the Knowledge Net of Amarakośa[7] can also come to aid for this. Similarly, the creation of a data set of possible pairs of *Upamāna* and *Upameya* with common ground can help the machine classify a word as a topic or vehicle in a given instance. However, such a data set will not be useful for the sentences where *Upamāna* and Upameya are not listed in this data set.

6 Conclusions and future scope

While analysing any complex topic that has several components interconnected to each other, the graph is a suitable format for its cognition. We realize this prominently when we see a graphical representation of parsed sentences in Samsaadhanii, Wordnet, concept net, representation of the relations between the objects or concepts in scientific texts, etc. The inherent nature of a figure of speech like Upamā is to have an interconnection between two objects, and there could be more than one such pair in a poetic instance. In such a situation, this graphical representation could be of great aid to the student or reader, since it eliminates ambiguity and room for doubt that could possibly arise while reading a complex statement. Since this concept of Upamā is found in other languages as well, a similar pattern of representation can be adopted by them while explaining the instances of this figure of speech in their respective languages.

Identifying the state in an instance of Upamā, is a task that seems unachievable in the present scenario. However, this is an open field for researchers to develop a technique to identify it. Further, we aim to construct such parse trees for another variety of *Upamā alaṅkāra* known as *luptopamā* that is sub-classified into 19 different varieties in KP. We also aim to analyse the phenomena called *camatkṛti* i.e., the poetic astonishment by analysing these dependency structures.

---

[6]https://www.cfilt.iitb.ac.in/wordnet/webswn/english_version.php
[7]https://sanskrit.uohyd.ac.in/scl/amarakosha/index.html

References

Ācārya Viśveśvara (Ed.&Tr.). 2017. *Kāvyaprakāśa.* jñānamaṇḍala Limited, Varanasi.

Milind Kumar Audichya and Jatinderkumar R. Saini. 2021. Towards Natural Language Processing with figures of speech in Hindi poetry. In *International Journal of Advanced Computer Science and Applications.*

Amruta Barbadikar and Amba Kulkarni. 2023. Yamaka Identifier and Classifier: A computational tool for the analysis of Sanskrit figure of sound. *The Annals of the Bhandarkar Oriental Research Institute (forthcoming)*, 102.

Amruta Barbadikar and Amba Kulkarni. 2024. Anuprāsa Identifier and Classifier: A computational tool to analyze Sanskrit figure of sound. In *Proceedings of the 7th International Sanskrit Computational Linguistics Symposium (forthcoming)*, Pondicherry, India. Association for Computational Linguistics.

Sharma Hari Dutta (Ed.&Tr.). 1935. *Kāvyaprakāśa Ullāsa X.* Oriental Book Agency, Pune.

Edwin Gerow. 1971. *A Glossary of Indian Figures of Speech.* Mouton & Co. N. V. Publishers, Hague.

Vaibhav Kesarwani, Diana Inkpen, Stan Szpakowicz, and Chris Tanasescu. 2017. Metaphor detection in a poetry corpus. In *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 1–9, Vancouver, Canada. Association for Computational Linguistics.

Amba Kulkarni. 2021a. *Sanskrit Parsing: Based on the Theories of Śābdabodha.* DK Printworld (P) Ltd.

Amba Kulkarni. 2021b. Sanskrit parsing following Indian Theories of Verbal Cognition. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, 20(2).

Amba Kulkarni, Pavankumar Satuluri, Sanjeev Panchal, Malay Maity, and Amruta Malvade. 2020. Dependency relations for Sanskrit parsing and treebank. In *Proceedings of the 19th International Workshop on Treebanks and Linguistic Theories*, pages 135–150, Düsseldorf, Germany. Association for Computational Linguistics.

M. Kulkarni, C. Dangarikar, I. Kulkarni, A. Nanda, and Bhattacharya. 2010. Introducing Sanskrit Wordnet. In *5th Global Wordnet Conference, IIT Mumbai.*

Malhar Kulkarni. 2017. Sanskrit Word Net at Indian Institute of Technology (iitb) mumbai. *The WordNet in Indian Languages*, pages 231–241.

Malhar Kulkarni, Nilesh Joshi, Sayali Khare, Hanumant Redkar, and Pushpak Bhattacharyya. 2019. Introduction to Sanskrit shabdamitra: An educational application of Sanskrit Wordnet. In *Proceedings of the 6th International Sanskrit Computational Linguistics Symposium*, pages 117–133, IIT Kharagpur, India. Association for Computational Linguistics.

Revati Kulkarni and Malhar Kulkarni. 2023. Morph-dependency model for Upamā alaṅkāra with special reference to Raghuvaṁśam. In *To be published in the proceedings of 43rd International Conference of the Linguistic Society of India*.

Suzanne Mpouli. 2017. Annotating similes in literary texts. In *Proceedings of the 13th Joint ISO-ACL Workshop on Interoperable Semantic Annotation (ISA-13)*.

Komal Naaz and Niraj Kumar Singh. 2022. Design and development of computational tools for analyzing elements of Hindi poetry. In *IEEE Access*.

Sivaja S Nair and Amba Kulkarni. 2010. The knowledge structure in amarakośa. In *Sanskrit Computational Linguistics: 4th International Symposium, New Delhi, India, December 10-12, 2010. Proceedings*, pages 173–189. Springer.

Vlad Niculae. 2013. Comparison pattern matching and creative simile recognition. In *Proceedings of the Joint Symposium on Semantic Processing. Textual Inference and Structures in Corpora*, pages 110–114, Trento, Italy.

Ashequl Qadir, Ellen Riloff, and Marilyn A. Walker. 2016. Automatically inferring implicit properties in similes. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1223–1232, San Diego, California. Association for Computational Linguistics.

Da Ren, Pengfei Zhang, Qing Li, Xiaohui Tao, Junying Chen, and Yi Cai. 2020. A hybrid representation-based simile component extraction. In *Neural Computing and Applications volume*.

Robyn Speer and Catherine Havasi. 2013. Conceptnet 5: A large semantic network for relational knowledge. In *The People's Web Meets NLP*.

Arjunwādkar K S (Ed. & Tr.). 1962. *Kāvyaprakāśa Ullāsa X*. Deshmukh Prakashan, Pune.

Sukthankar S.S. (Ed. & Tr.). 1941. *Kāvyaprakāśa Ullāsa X*. Karnataka Publishing House, Mumbai.

Jörg Wöckener, Thomas Haider, Tristan Miller, The-Khang Nguyen, Thanh Tung Linh Nguyen, Minh Vu Pham, Jonas Belouadi, and Steffen Eger. 2021. End-to-end style-conditioned poetry generation: What does it take to learn from examples alone? In *Proceedings of the 5th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 57–66, Punta Cana, Dominican Republic (online). Association for Computational Linguistics.