The Knowledge Structure in Amarakośa

Sivaja S. Nair and Amba Kulkarni

Department of Sanskrit Studies,
University of Hyderabad,
Hyderabad
sivaja.s.nair@gmail.com and apksh@uohyd.ernet.in

Abstract. *Amarakośa* is the most celebrated and authoritative ancient thesaurus of Sanskrit. It is one of the books which an Indian child learning through Indian traditional educational system memorizes as early as his first year of formal learning. Though it appears as a linear list of words, close inspection of it shows a rich organisation of words expressing various relations a word bears with other words. Thus when a child studies Amarakośa further, the linear list of words unfolds into a knowledge web. In this paper we describe our effort to make the implicit knowledge in Amarakośa explicit. A model for storing such structure is discussed and a web tool is described that answers the queries by reconstructing the links among words from the structured tables dynamically.

Key Words : Amarakośa, Synset, Polysemy, Semantic relations, Ontology

1 Introduction

The Indian tradition of transmitting knowledge orally is on the verge of vanishing. As the oral transmission demands, Indian traditional educational culture was organised to be *formal and intensive* as opposed to the modern culture which is more *informal and extensive* (Wood, 1985). In traditional circumstances, a child would receive his education largely by oral transmission, mainly through rote-learning. The method employed was through recitation and remembering. A child is taught the alphabet (varṇamālā), he would memorise a few verses, subhāṣitas, and then start reciting a dictionary of synonymous words – the Amarakośa – till it is memorised. It typically would take anywhere between 6 months to a year to memorise a list of approximately 10,000 Sanskrit words arranged as a list of synonyms. The close inspection of the structure of the Amarakośa gives much more insight into the way the words are organised. When a student memorises it, though in the beginning it appears as a linear list of words, as he starts understanding the meaning of the words, reads the commentaries on this text and starts using these words, the linear structure

unfolds into a knowledge web with various links.

The Amarakośa printed in the form of a book just shows the linear order, and the index at the end of the book point to various words for easy references. But there is much more to it than just a linear order. The knowledge a student acquires through various commentaries and also its practical use in his own field of expertise – be it Āyurveda, Vyākaraṇa or Sāhitya, is in the form of various links. With the modern education culture that is dominated by the use of computers as a tool, which relies more on the secondary memories such as books, computers, and the World Wide Web, than the human memory, it is necessary to make the *implicit* knowledge in Amarakośa *explicit*. The computers have an advantage over the printed books. Computers can represent multi-dimensional objects, and thus one can navigate through the whole structure and at the same time with the powerful search facilities can search complex queries. In this paper, we illustrate with examples various kinds of links one can 'visualise' in Amarakośa, and provide a database model to store these links in order to facilitate automatic extraction of these links as an answer to a search query.

2 Amarakośa

Amarakośa primarily named as *Nāmaliṅgānuśāsana* (a work that deals with instructions related to the gender of nouns) is authored by Amarasiṃha - 4^{th} century A.D. (Oka, 1981) - and is the most celebrated and authoritative ancient thesaurus of Sanskrit with around 60 commentaries and translations into modern Indian as well as foreign languages such as Chinese, Tibetan, French, etc (Patkar, 1981). It is considered as an essential requisite for a Sanskrit scholar and as such a child is asked to memorise it even before he starts his studies formally. It consists of 1608 verses composed in anuṣṭup meter[1] and are divided into 3 chapters called Kāṇḍas.[2]

Classification Each of the three Kāṇḍas is further subdivided into various *vargas*. The classification of three kāṇḍas into 25 vargas is as below.

- ***prathamakāṇḍam*:**
  svargavargaḥ (heaven)
  vyomavargaḥ (sky)
  digvargaḥ (direction)
  kālavargaḥ (time)
  dhīvargaḥ (cognition)

[1] Śloke ṣaṣṭaṃ guruṃ jñeyaṃ sarvatra laghu pañcamam |
Dvicatuḥpādayorhrasvaṃ saptamaṃ dīrghamanyayoḥ ||

[2] and as such is known as *Trikāṇḍī*

śabdādivargaḥ (sound)
nāṭyavargaḥ (drama)
pātālabhogivargaḥ (nether world)
narakavargaḥ (hell)
vārivargaḥ (water)

– ***dvitīyakāṇḍam*:**
bhūmivargaḥ (earth)
puravargaḥ (towns or cities)
śailavargaḥ (mountains)
vanauṣadhivargaḥ (forests and medicines)
siṃhādivargaḥ (lions and other animals)
manuṣyavargaḥ (mankind)
brahmavargaḥ (priest tribe)
kṣatriyavargaḥ (military tribe)
vaiśyavargaḥ (business tribe)
śūdravargaḥ (mixed classes)

– ***tṛtīyakāṇḍam*:**
viśeṣyanighnavargaḥ (adjective)
saṁkīrṇavargaḥ (miscellaneous)
nānārthavargaḥ (polysemous)
avyayavargaḥ (indeclinables)
liṅgādisaṅgrahavargaḥ (gender)

Amarakośa contains 11,580 content words (tokens). Some of the tokens are repeated either within a kāṇḍa or across the kāṇḍas leading to only 9,031 types. The kāṇḍa-wise distribution of the tokens and types is shown in Table 1.

kāṇḍa	tokens	types
prathamakāṇḍam	2465	2300
dvitīyakāṇḍam	5827	5282
tṛtīyakāṇḍam	3288	2271

Table 1. Tokens and types in each kāṇḍas

Synset A set of synonymous words is termed as a synset. Each synonym may span over one or more verses. The following verse, e.g., provides a synonym for the word *jaṃbuka*.

striyām śivā bhūrimāyagomāyumṛgadhūrtakāḥ |
sṛgālavañcakakroṣṭupherupheravajambukāḥ ||2.5.5 ||

Polysemy Amarakośa has 4,017 synsets. Some of the words fall under more than one synsets, and thus are ambiguous. Most of these polysemous words belong to the nānārthavarga of the third kāṇḍa which lists the polysemous words alphabetically according to their endings. The polysemy distribution in the Amarakośa is summarised in Table 2. There is only one word *hari* in Amarakośa which has as many as 14 senses, the word *antara* belongs to 13 synsets, and the word *go* has 12 synsets. We note that almost 65% words (7459 words) belong to a single synset and thus are not ambiguous.

No. of meanings	No. of words	Words
14	1	*hari*
13	1	*antara*
12	1	*go*
10	2	*kriyā, kūṭa*
9	2	*rasa, vṛṣa*
8	8	*dhātu, dharma, vasu, ariṣṭa...*
7	9	
6	18	
5	49	
4	136	
3	330	
2	1015	
1	7459	

Table 2. Polysemy Distribution

Gender A few verses in the beginning of the Amarakośa describe the meta-language and the techniques employed to indicate the gender of various words. The word *striyām*, for example, in (2.5.5) above is not a token but a word from the meta-language indicating the gender of the following word *śivā* to be feminine. In addition to these general guidelines, in the liṅgādisaṅgrahavarga Amarasiṃha gives certain grammatical and phonological clues for deciding the gender of a word. In the event of absence of any rule, the gender of the remaining words in 2.5.5, constituting two compound words "bhūrimāyagomāyumṛgadhūrtakāḥ" and "sṛgālavañcakakroṣṭupherupheravajambukāḥ" is inferred to be masculine from their compounding-forms.

The avyayavarga lists synsets consisting of indeclinables.

3 Organisation of synsets within a varga

Except the polysemous words (nānārthavarga), all other synsets in a varga show some semantic relation to the varga it belongs to and sometimes even to the preceding or following synsets. These semantic relations indicate various kinds of relations. They may be classified as hierarchical or associative. The hypernym indicating a more general term or the hyponym showing a more specific term are the examples of hierarchical relation. Similarly the holonym-meronym relation marking the whole-part relation is also a hierarchical relation. In addition various other relations are indicated by the adjacency of the synsets. These may be termed as associative relations, which indicate some kind of association of one synset with the other. This association may be the association among human beings, or the association of certain objects with certain other objects. We illustrate below some such relations with examples.

3.1 Example 1: Viṣṇuḥ

The verses from 1.1.18 to 1.1.29 describe various synsets representing *Viṣṇu*, and objects related to/associated with *Viṣṇu*. The relations, as is evident from the following description, are kinship relations such as father, brother, son, grandson, wife, and also associated objects such as conch, discus, sword, vehicle, etc. (See Figure 1).

Viṣṇuḥ (1.1.18 - 1.1.22)[3]
- Kṛṣṇa's father (1.1.22)
- Kṛṣṇa's elder brother (1.1.23 - 1.1.24)
- kāmadevaḥ (1.1.25 - 1.1.26)
  - floral arrows of kāmadevaḥ (1.1.26)
  - physical arrows of kāmadevaḥ (1.1.26)
  - son of kāmadevaḥ - aniruddhaḥ (1.1.27)
- wife of Viṣṇuḥ - lakṣmī (1.1.27)
- Special devices/equipments of Viṣṇuḥ(1.1.28)
- (conch, discus, sword, jewel, bow, horse, mark,etc.)
- Kṛṣṇa's charioteer, minister (1.1.28)
- Kṛṣṇa's younger brother (1.1.28)
- Viṣṇu's vehicle - garuḍaḥ (1.1.29)

3.2 Example 2: samayaḥ

The verses from 1.4.1 to 1.4.9 deal with words related to time, units of measurement, special names of special days, etc.

[3] The English translations of the subheadings, which are given here and in the following examples, describing the ślokas are taken from Colebrooke's commentary on Amarakośa (Colebrooke, 1808).

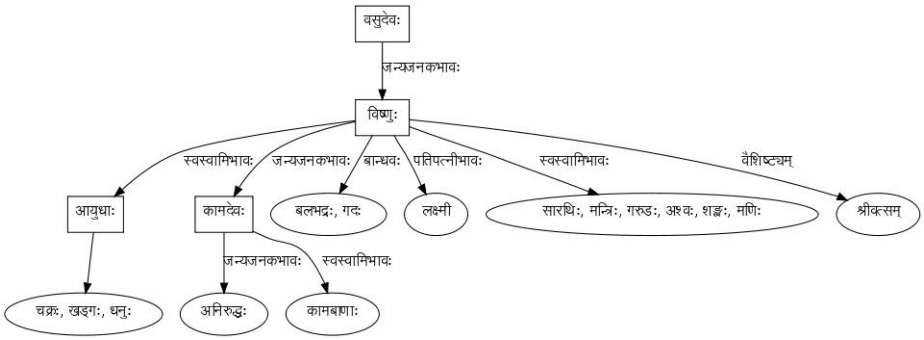

Fig. 1. Relations of Viṣṇu

Time (1.4.1)
- Lunar day (1.4.1)
  - First lunar day (1.4.1)
  - {Day (1.4.2)
    - Morning (1.4.2 - 1.4.3)
    - Twilight (1.4.3)
    - Evening (1.4.3)
    - First four hours of a day (1.4.3)
    - Second four hours of a day (1.4.3)
    - Third four hours of a day (1.4.3)
    - Period of the day (1.4.3)
    - Night (1.4.3 - 1.4.4)
      - A dark night (1.4.5)
      - A moonlight night (1.4.5)
      - A night and two days (1.4.5)
      - First part of night (1.4.6)
      - Midnight (1.4.6)
      - Sequence of nights (1.4.6)
  - Space of three hours (1.4.6) }
    - Last day of the half month (1.4.7)
      - Precise moment of the full or the new moon (1.4.7)
    - Full moon day (1.4.7)
      - Full moon whole day(1.4.8)
      - Full Moon with a little gibbous on part of a day (1.4.8)
    - No moon day (1.4.8)
      - wanning crescent (1.4.9)
      - No moon whole day (1.4.9)

In this example we also see violation of nesting. In between the synsets related to *lunar day* and *last day of the month*, the synsets related to *day* (which refers to the apparent solar motion) are intervened.

3.3 Example 3: kṣatriyaḥ

Here is a group of verses from 2.8.1 to 2.8.10 belonging to the kṣatriyavarga. The words here refer to the king, military, ministers, various category of people engaged in the services of kings, etc.

Man of the military tribe (2.8.1)
- King (2.8.1)
  - Universal monarch (2.8.2)
  - An emperor (2.8.2)
  - King over a country (2.8.2)
  - Paramount sovereign (2.8.3)
- Multitude of kings (2.8.3)
- Multitude of military tribe (2.8.4)
  - Minister (2.8.4)
    - Deputy minister (2.8.4)
  - Priest (2.8.5)
  - Judge (2.8.5)
  - King's companions (2.8.5)
  - Body guards of a king (2.8.6)
  - Warder (2.8.6)
  - Superintendent (2.8.6)
    - Village Superintendent (2.8.7)
    - Superintendent of many villages (2.8.7)
    - Superintendent of Gold (2.8.7)
    - Superintendent of Silver (2.8.7)
    - Superintendent of the womens' appartments (2.8.8)
      - Outside guard of the womens' appartment (2.8.8)
    - attendant of a king (2.8.9)
      - eunuch (2.8.9)
    - Prince whose territories lie on the frontiers of those of the enemy (2.8.9)
    - Neighboring prince (2.8.9)
    - Prince whose territories lie beyond those of the friend (2.8.10)
    - Enemy in the rear (2.8.10)

3.4 Implicit relations

These were three samples from three distinct topics involving totally different kind of relations. All these relations are semantic in nature. A more detailed study of such examples showed that following relations occur more frequently.

- *avayavāvayavī* (part-whole relation)
- *parāparājāti* (is a kind of relation)
- *janyajanaka* (child-parent relation)

– *patipatnī* (husband-wife relation)
– *svasvāmi* (master-possession relation)
– *ājīvikā* (livelihood)

There are a few other relations such as kinship relations, ādhāra-ādheya, vaṃśa-vaṃśīya etc. The extraction of such relations and marking is still ongoing. But the instances of such relations were found to be rare.

4 Amarakośa-jñāna-jāla

In the recent past there have been notable efforts by Sanskrit computational linguists with focus on Amarakośa. Jha et. al. (2010, Online Multilingual Amarakosha) have developed a searchable web interface to Amarakośa which provides the Indian language equivalents of the Amarakośa words in addition to the original sanskrit text. Bharati et. al. (2008) and Nair et. al. (2009) did comparative study of the Amarakośa with the existing Hindi WordNet in order to find the usefulness of Hindi WordNet in augmenting the Amarakośa with relational information. There have been efforts by Kulkarni et. al. (2008, 2010) that describe the development of Sanskrit WordNet. The present effort is altogether an innovative effort that helps reveal the internal structure of the Amarakośa.

The Amarakośa-jñāna-jāla is developed as a web application. The application provides a search result of a query dynamically generated using the structured lexicon of the Amarakośa and the supplementary tables marking the relations.

The structured lexicon as well as the supplementary tables showing the explicit relations are simple ASCII text files. Sanskrit words are stored in a roman transliterated scheme (WX notation).[4] There are two advantages of storing the text in WX notation. The first advantage is, it facilitates a lexicographer to use simple unix tools such as grep, sed, etc. for her day-to-day work of updating the knowledge-base. Unicode for Devanagari mixes the phonemes with the syllables, making it un-natural to write the search expressions. The second advantage, of course, is the size. The size of the tables in UTF-8 for Devanagari is more than 2 times the corresponding files in roman transliteration such as WX notation.

4.1 Structured Lexicon

The main structured lexicon consists of synsets stored in the form of a set of records. Each record corresponds to a word in the Amarakośa (excluding the meta-language words). It consists of 5 fields as described below.

[4] http://sanskrit.uohyd.ernet.in/~anusaaraka/sanskrit/samsaadhanii/wx.html

Stem Amarakośa lists words in nominative cases. However, we decided to go for the nominal stem instead of the nominative case word form. In case of feminine words, this field contains the feminine stem, i.e. the stem after adding the feminine suffix. In case of nānārthavarga (part of the Amarakośa dealing with polysemous words), the polysemous word is entered in this field.

The reason for choosing nominal stem over the nominative case form is the ease in linking the Amarakośa words with the existing computational resources such as morphological analysers and generators and various e-lexicons, which typically expect a prātipadikam and not a prathamānta (ending in nominative case).

Amarakośa index This field contains a reference to an entry in the Amarakośa, as a 5 tuple of numbers, separated by dots. The 5 numbers in the 5 tuple refer to the kāṇḍa, varga, śloka, pāda and the word number respectively. Table 3 shows a sample entry corresponding to the following śloka,

svaravyayaṃ svarganākatridivatridaśālayāḥ |
suraloko dyodivau dve striyāṁ klībe triviṣṭapam ||1.1.6 ||

Word	**Reference**
svar	1.1.6.1.1
svarga	1.1.6.1.2
nāka	1.1.6.1.3
tridiva	1.1.6.1.4
tridaśālaya	1.1.6.1.5
suraloka	1.1.6.2.1
dyo	1.1.6.2.2
div	1.1.6.2.3
triviṣṭapa	1.1.6.2.4

Table 3. Words and references of the *svarga*-synset

Liṅgam (gender) This field contains the gender of the stem. The gender of a word in a śloka is decided with the help of meta-language employed by Amarakośa. These are further cross checked with Devadatta Tiwari's *Devakośa arthāt Amarakośa* (Tiwari, 1989) and Colebrooke's commentary on Amarakośa (Colebrooke, 1808) when in doubt.

Sanskrit has 3 values for gender viz. masculine, feminine and neuter. Thus there are 8 possible combinations (an indeclinable is assigned no gender, and

the adjectives are the ones which take all the three genders). In addition, Amarakośa also provides information about words that are always plural or dual by nature. Following combinations of gender, number information were found in the Amarakośa.

- Indeclinable - (*avya.*)
- Feminine - (*strī.*)
- Masculine - (*puṃ.*)
- Neuter - (*napuṃ.*)
- Masculine and Feminine - (*strī-puṃ.*) [aśani[5]]
- Feminine and Neuter - (*strī-napuṃ.*) [uḍu[6]]
- Feminine dual - (*strī-dvi.*) [dyāvāpṛthvyau[7]]
- Feminine plural - (*strī-bahu.*) [apsaras[8]]
- Masculine and Neuter - (*puṃ-napuṃ.*) [daivatāni[9]]
- Masculine dual - (*puṃ-dvi.*) [nāsatyau[10]]
- Masculine plural - (*puṃ-bahu.*) [gṛhāḥ[11]]
- Neuter and indeclinable - (*napuṃ-avya.*) [apadiśaṃ[12]]
- Adjective - (*vi.*)

Vargaḥ This field contains the name of the varga, as given in the commentaries to which the entry belongs.

Head Word The first four fields cover all the explicit information that can be easily extracted automatically. The important feature of Amarakośa is that it provides synonymous words. The marking of synonymous words is obvious only through the world knowledge or through the commentaries. To provide a handle to each set of synonymous words – called as synset, we created a field termed as *Head Word* which provides a name to each synset. Thus these Head Words are unique and act as a reference ID for a synset. The total number of Head Words give us the total number of synsets in the Amarakośa. We denote the synset corresponding to a Head Word *W* by *Syn(W)*.

The choice of Head-Words is mainly guided by the Bhānuji Dīkṣitā's *Sudhā* commentary on *Amarakośa* (Pandit, 1915). When a better choice was available in the Malayalam commentary *Triveṇī* (Moosath, 1956) or *Pārameśvarī* (Moosath, 1914) or the Hindi commentary *Prabhā*, it was chosen. Table 4 shows an example of a śloka 2.5.5 converted to a structured table, and figure 2 shows the search result of the Amarakośa-jñāna-jāla for the word śṛgāla.

[5] aśanirdvayoḥ (1.1.47)
[6] tārakāpyuḍu vā striyāṃ (1.3.21)
[7] dyāvāpṛthvyau (2.1.19)
[8] striyāṃ bahuṣvapsarasaḥ (1.1.52)
[9] daivatāni puṃsi vā (1.1.9)
[10] nāsatyāvaśvinau dasrāvāśvineyau ca tāvubhau (1.1.51)
[11] gṛhāḥ puṃsi ca bhūmnyeva (2.2.5)
[12] klībāvyayaṃ tvapadiśam (1.3.5)

Token	Reference	Gender	Varga-name	Head-Word
śivā	2.5.5.1.1	strī	siṃhādivargaḥ	jaṃbhūkaḥ
būrimāya	2.5.5.1.2	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ
gomāyu	2.5.5.1.3	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ
mṛgadūrtaka	2.5.5.1.4	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ
śṛgāla	2.5.5.2.1	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ
vañjaka	2.5.5.2.2	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ
kṛoṣṭu	2.5.5.2.3	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ
pheru	2.5.5.2.4	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ
pherava	2.5.5.2.5	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ
jaṃbuka	2.5.5.2.6	puṃ	siṃhādivargaḥ	jaṃbhūkaḥ

Table 4. Example of Head-Word

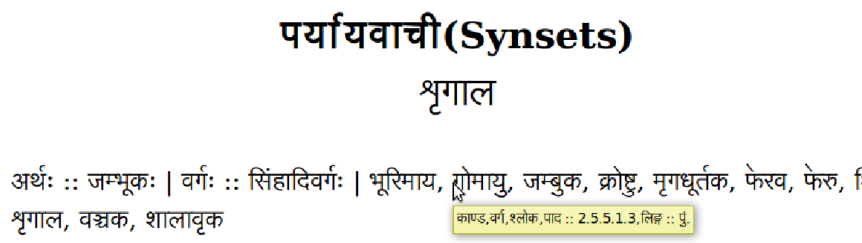

Fig. 2. Example of a synset

4.2 Tables marking various relations

The relations are among various Head Words and are marked as records. Each record corresponds to one synset ID. The first field of each record consists of the synset ID, and remaining six fields correspond to the Head Words that bear a relation of *is_a_part_of (avayavāvayavi)*, *is_a_kind_of(parāparājāti)*, *janya-janaka-bhāva*, *pati-patnī-bhāva*, *sva-svāmi-bhāva*, *ājīvikā* with the synset ID in the first field.

1. Is a part of (avayavāvayavi)
   This field marks *is a part of* relation. Let W be the synset-ID. Then this field will have an entry W' if the member of $SynW$ is a part of member of $SynW'$ (See Table 5).

   For example,
   Syn($rātriḥ$) = *śarvarī, kṣaṇadā, kṣapā, niśā, niśīthinī, rajanī, rātri, vibhāvarī, tamasvinī, tamī, triyāmā, yāminī, naktam, doṣā, vasati, śyāmā.*
   and

Syn(*rātrimadhyaḥ*) = *ardharātra, niśītha.*

Now, *ardharātra, niśītha* are part of *niśā, rajanī, rātri*, etc.. Hence *rātrimadhyaḥ* is marked to be is_a_part_of (*avayava* of) *rātriḥ*
Similarly *pradoṣa, rajanīmukha* (∈ *Syn*(*rātriprāraṃbhaḥ*)) are also part of *niśā, rajanī, rātri*, etc.. Hence *rātriprāraṃbhaḥ*, where Syn (*rātriprāraṃbhaḥ*)= *pradoṣa, rajanīmukha*
also bears a part_of relation with *rātriḥ.*

Head-Word W	**part (*avayava*)-of W**
rātrimadhyaḥ	rātriḥ
rātriprāraṃbhaḥ	rātriḥ

Table 5. Example of is-a-part relation

2. Is a kind of (*parāparājāti*)
   This field marks *is a kind of* relation. The entry contains the Head Word W' such that synset ID W bears a relation of *is a kind of* with W'. The hypernymy and hyponymy relation can be extracted using this field. Here are some entries: (see Table 6.)

Head-Word W	**kind (*parājātiḥ*) of W**
gaṅgā	nadī
yamunā	nadī
narmadā	nadī

Table 6. Example of is_a_kind_of rFelations

3. Janya-janaka-bhāva (parent-child relation)
   This field marks the relation of parent-child (*janya-janaka-bhāva*). (see Table 7.) Where Syn (*jayantaḥ*) = *pākaśāsani, jayanta.*
   and
   Syn (*indraḥ*) = *indra, biḍaujas, maghavan, marutvat, pākaśāsana, sunāsīra, vṛddhaśravas, purandara, puruhūta, jiṣṇu, lekharṣabha, śakra, śatamanyu, divaspati, vṛṣan, vṛtrahan, gotrabhid, sutrāman, vāsava, vajrin, balārāti, śacīpati, surapati, vāstoṣpati, harihaya, jambhabhedin, namucisūdana, svarāj, meghavāhana, saṅkrandana, turāṣā, duścyavana, ākhaṇḍala, ṛbhukṣin, sahasrākṣa, kauṣika, ghanāghana, parjanya, hari.*

   Syn (*sanatkumāraḥ*) = *sanatkumāra, vaidhātra*
   and
   Syn (*brahmā*) = *ātmabhū, brahman, caturānana, hiraṇyagarbha, lokeśa,*

parameṣṭhin, pitāmaha, surajyeṣṭha, svayambhū, abjayoni, aṇḍaja, haṁsavāhana, kamalāsana, kamalodbhava, nābhijanman, nidhana, prajāpati, pūrva, rajomūrtin, satyaka, sadānanda, svaṣṭṛ, vedhas, viriñci, viśvasṛj, vidhātṛ, vidhi, dhātṛ, druhiṇa, ka, ātman, śambhu.

Head-Word W	**Child (*janya*) of W**
indraḥ	jayantaḥ
brahmā	sanatkumāraḥ
śivaḥ	gaṇeśaḥ

Table 7. Example of Janya-janaka relation

4. Pati-patnī-bhāva (husband-wife relation)
   This field marks the husband-wife relation, as shown below. (see Table 8.) Where Syn(*lakṣmī*) = *bhārgavī, haripriyā, indiraā, kamalā, kṣīrasāgarakanyakā, kṣīrodatanayā, lakṣmi, lokajananī, lokamātṛ, mā, padmā, padmālayā, ramā, śrī, vṛṣākapāyī.*
   and
   Syn(*viṣṇuḥ*) = *hṛṣīkeśa, keśava, kṛṣṇa, mādhava, nārāyaṇa, svabhū, vaikuṇḍha, viṣṇu, viṣṭraśravas, dāmodara, acyuta, garuḍadhvaja, govinda, janārdana, pītāmbara, puṇḍarīkākṣa, śārṅgin, viṣvaksena, daityāri, cakrapāṇi, caturbhuja, indrāvaraja, madhuripu, padmanābha, upendra, vāsudeva, trivikrama, adhokṣaja, balidhvaṁsin, kaṁsārāti, puruṣottama, śaurī, śrīpati, vanamālin, xevakīnandana, jalaśāyin, kaiṭabhajit, mukunda, muramardana, narakāntaka, purāṇapuruṣa, śrīvatsalāñchana, viśvambhara, viśvarūpa, vidhu, yajñapuruṣa, lakṣmīpati, murāri, aja, ajita, avyakta, vṛṣākapi, babhru, hari, vedhas.*

Head-Word W	**Husband (*pati*) of W**
lakṣmī	viṣṇuḥ
pārvatī	śivaḥ
lopāmudrā	agastyaḥ

Table 8. Example of Pati-patnī relation

5. Sva-svāmi-bhāva (master-possession relation)
   This field marks the master-possession or sva-svāmi-bhāva relation as shown below: (see Table 9.)
6. Ājīvikā (livelihood)
   This field marks the livelihood relation between two syn-sets. For example,

Head-Word W	**master (*svāmi*) of W**
viṣṇoḥ mantriḥ	viṣṇuḥ
viṣṇoḥ sārathiḥ	viṣṇuḥ
garuḍaḥ	viṣṇuḥ

Table 9. Example of Sva-svāmi relation

the synset with Head Word *matsya* is (*aṇḍaja*, *jhaṣa*, *matsya*, *mīna*, *pṛthuroman*, *śakulī*, *vaisāriṇa*, *visāra*, *animiṣa*) denotes objects which act as a livelihood for the objects expressed through the concept of *dhīvara*, and hence the livelihood for the objects belonging to the synset *dhīvara* is marked as a *matsya*. (see Table 10.)

Head-Word W	**Livelihood (*Ājīvikā*) of W**
dhīvaraḥ	matsyaḥ
nartakī	nṛtyam
nāvikaḥ	naukā
sevakaḥ	sevā

Table 10. Example of Ājīvikā relation

4.3 Quantitative analysis

For every headword, one or more of the relations as specified above are marked. As was expected, the hierarchical relations viz. is_a_kind_of and is_a_part_of appear prominently than the associative relations. The occurrence of various relations in terms of Head-Words and all the words belonging to the synsets denoted by these head words is shown in Table 11.[13]

No.	Relation	Headwords	Words
1	is_a_kind_of	2239	6807
2	is_a_part_of	560	1654
3	janya-janaka	17	193
4	sva-svāmī	36	122
5	ājīvikā	30	106
6	pati-patnī	25	105

Table 11. Relational statistics

[13] Till 16^{th} April 2010

4.4 Implementation

From the structured lexicon table and the table of relations we build data bases using the built-in dbm engines of unix and the programes are written in Perl. These dbm engines use hashing techniques to enable fast retival of the data by key.

Following three hash tables are built from the structured lexicon.

a) Head-word hash
where Key=stem and Value=head-word

b) Synset hash
with Key=head-word and Value=synset

c) Word-info hash
generated by Key=stem and Value=word-index and gender

From the table of relations, corresponding to each relation R, we built a hash table which returns the associates a head-Word W with another head-word W', if W' is related to W by relation R

Amarakośa-jñāna-jāla is presented as a web application developed with 'apache' web server and 'perl' for CGI script. User submits a query a word and a relation, machine produces all the words related to the given word by the chosen relation. The word here may be either a stem or an inflected word form. In the case of inflected word form, machine consults the morphological analyser to get the stem. Figures in appendix - 1 give sample results of queries for different word-relation combinations. When a cursor is placed on a word a tool tip shows its word-index and gender(as shown in Fig. 1.).

5 Conclusion

The study of Amarakośa from a point of view of exploring the relations was undertaken to reveal the implicit knowledge and make it explicit. The resulting computational tool helps a Sanskrit reader to get a feel for various kinds of relations mentioned in the Amarakośa and thereby its richness as a knowledge source. The hierarchical relations such as is_a_part_of and is_a_kind_of will be of help in information extraction, while the associative relations help a reader to get the cultural knowledge.

Sanskrit has a rich tradition of kośas. Most of them are arranged as a list of words with similar meaning (synonymic) or a list of words indicating various shades of a given word (polesemic). *Nāmamālā, Śabdaratnākara, Śabdacandrikā* are a few among the first type and *Nānārthasaṅgraha,*

Anekārthadhvanimañjarī, Viśvaprakāśa are a few examples of the second type. *Amarakośa, Abhidhānaratnamālā* and *Vaijayantīkośa* has both kind of entries.

This implementation may serve as a model to build similar tools for various other kośas mentioned above.

The Amarakośa is now available with various kinds of search facilities as a web service at

`http://sanskrit.uohyd.ernet.in/~anusaaraka/sanskrit/samsaadhanii/amarakosha/home.html.`

References

1. Bharati Akshar, Kulkarni Amba and Nair S. Sivaja (2008) "Use of Amarakosha and Hindi WordNet in Building a Network of Sanskrit Words" International Conference On Natural Language Processing, C-DAC Pune.
2. Colebrooke H.T (1808) "Kosha or Dictionary of the sungskrita language by Umura singha with an English interpretations and annotations" Serampore.
3. Fellbaum, Christiane (1999) "WordNet An Electronic Lexical Database" MIT Press, Massachusetts.
4. Jha Girish Nath, Chandrashekar R, Singh Umesh Kumar, Jha Vibhuti Nath, Pandey Satyendra, Singh Surjit Kumar and Mishra Mukesh Kumar (2010) "Online Multilingual Amarakośa: The Relational Lexical Database" 5th Global Wordnet Conference, IIT Mumbai.
5. Kulkarni, Malhar and Bhattacharya Pushpak (2008) "Verbal roots in the Sanskrit WordNet", 2nd International Sanskrit Computational Linguistics Symposium, Brown University.
6. Kulkarni Malhar, Dangarikar Chaitali, Kulkarni Irawati, Nanda Abhishek and Bhattacharya Pushpak (2010) "Introducing Sanskrit Wordnet" 5th Global Wordnet Conference, IIT Mumbai.
7. Moosath, Parameshvaran T.C. (1914) "Paarameswarii : malayalam commentory of Amarakosha" National Book Stall, Kottayam.
8. Moosath, Parameshvaran T.C.(1956) "Triveni : malayalam commentory of Amarakosha" National Book Stall, Kottayam.
9. Nair S. Sivaja, Swain Pritilaxmi and Kulkarni Amba (2009) "Developing network of Sanskrit words across Part-Of-Speech categories", CSATS'09, Rashtriya Sanskrit Vidhyapeeth, Tirupati.
10. Oka Krishnaji Govind (1981) "The Namalinganushasana : Amarakosa of Amarasimha with the commentary (Amarakoshodghaatana) of Ksheraswamin", Upasana prakashan Delhi, Varanasi.
11. Online Multilingual Amarakosha : http://sanskrit.jnu.ac.in/amara/index.jsp
12. Patkar M.M (1981) "History of Sanskrit Lexicography" Munshiram Manoharlal Publishers Pvt. Ltd., Delhi.
13. Pandit, Sivadatta (1915) "Namalinganusasana (Amarakosha) With the commentary of (Vyakhyasudha or Ramasrami) of Bhanuji Dikshit" Tukaram javaji proprietor of the "Nirnaya Sagar" Press, Bombay.
14. Tiwari, Devadatt (1989) "Devakoṡaḥ ardhāt amarakoṡaḥ" Nag prakashak, Delhi.
15. Wood, Ananda (1985) "Knowledge before printing and after – The Indian tradition in changing Kerala", Oxford University Press, Delhi.

A appendix - 1

आजीविका

धीवर

अर्थः :: धीवरः | वर्गः :: वारिवर्गः | कैवर्त, दाश, धीवर, जालिक

आजीविका

अर्थः :: मत्स्यः | वर्गः :: वारिवर्गः | अण्डज, झष, मत्स्य, मीन, पृथुरोमन्, शकुली, वैसारिण, विसार, अनिमिष

Fig. 3. Example of ājīvikā

अवयवी(Holonymy)

अङ्गुली

> अर्थः :: अङ्गुली | वर्गः :: मनुष्यवर्गः | अङ्गुली, करशाखा, कर्णिका

>अर्थः :: हस्तः | वर्गः :: मनुष्यवर्गः | पञ्चशाख, पाणि, शय, हस्त, कर

>अर्थः :: भुजः | वर्गः :: मनुष्यवर्गः | बाहु, भुज, प्रवेष्ट, दोस्

>अर्थः :: देहः | वर्गः :: मनुष्यवर्गः | गात्र, कलेवर, संहनन, शरीर, वपुस्, वर्ष्मन्, विग्रह, काय, मूर्ति, तनु, तनू, देह, करण, उत्सेध, भूतात्मन्, आत्मन्, धामन्, क्षेत्र, अजिर

Fig. 4. Example of avayvī

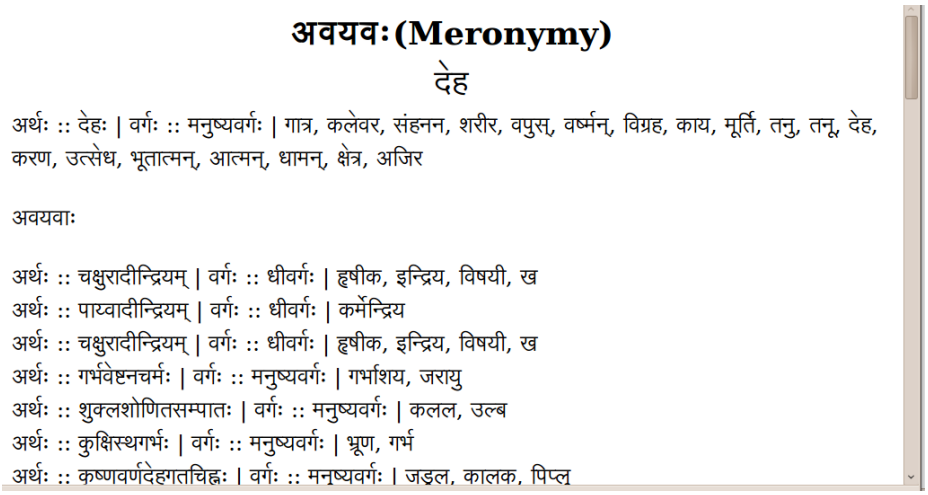

Fig. 5. Example of avayava

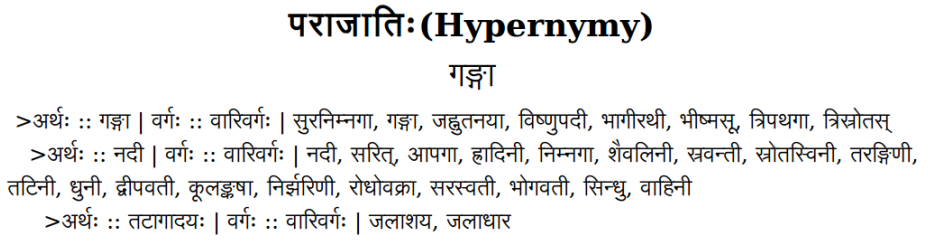

Fig. 6. Example of hypernymy

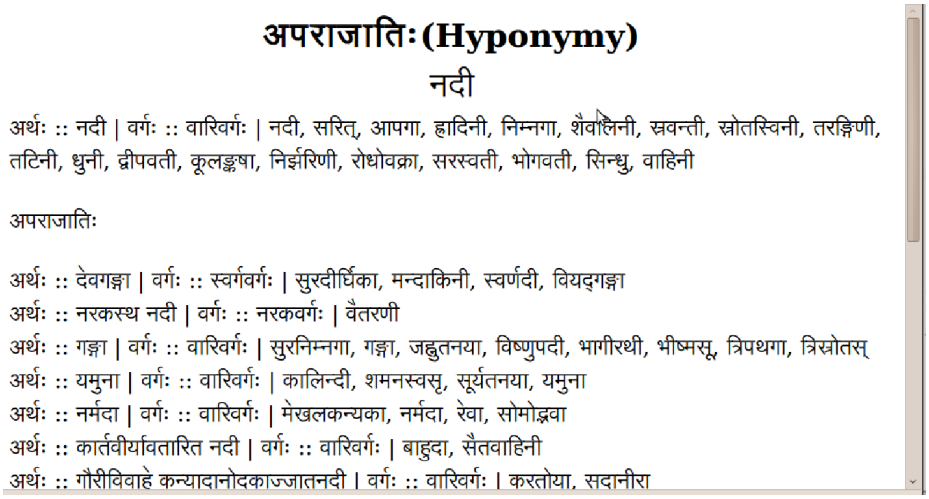

Fig. 7. Example of hyponymy