Grammatical Analysis and Subject Indexing of Mahābhārata and Tātparyanirṇaya

Srinivasa Kumar N Acharya & Arjuna S.R.
Dvaita Philosophy Resource Centre
Department Of European Studies
Manipal University, Manipal
`acharya.srinivasan@manipal.edu, arjunsr1987@gmail.com`

Abstract

Mahābhārata, with its more than 100,000 verses, is the largest epic that deals with several dimensions of human life: physical, psychological and philosophical. This epic text deals with almost all domains of knowledge with multidimensional perspectives; hence, it caters to offer several new outlooks that fosters one's understanding and spiritually elevates to enlightenment. This concept paper discusses about exploring exegetically the epic text using Sanskrit Computational Linguistics tools through analysing syntax and semantics of the literature.

1   Introduction

Traditional texts that hold the wisdom of centuries of human civilization have remained inaccessible to larger and mainstream academic world for being available primarily in Sanskrit, a language that seems to be limited only to scholars in traditional schools. Sanskrit language, one of the world's oldest, houses a huge amount of repository of literature that dates to millennia. The structure of the language remains unchanged in its history due to its strong grammar base. Nevertheless, it is not easy to interpret Sanskrit texts due to 'polysemy' and 'dynamism' that often makes it ambiguous. One should not make out that the language is stagnant and not open to neologism. The grammar and semantics of the language are structured in a way that the words or meanings of the language are highly dynamic. Moreover, more than thousands of traditional scholars and researchers most of whom work in silos with focus on specific schools of thought also have not been able to successfully mainstream the traditional knowledge system into that of today's academia. Due to these, it becomes very pertinent that the power of knowledge hidden in these texts viz., Mahābhārata is unravelled and brought to mainstream discourses in philosophy and social sciences with a focus on contemporary relevance.

1.1   Why Mahābhārata?

Mahābhārata, the magnum opus of Vedavyāsa, is a text with over 100,000 verses in 18 Parvas and it is said to be the world's biggest literature epic that deals with all the aspects of human life4) is one of the biggest literary work in Sanskrit literature with around 100,000 verses in 18 parvas. A verse emphasises importance of Mahābhārata -

> "*Dharme ca arthe ca kāme ca mokṣe ca bharataŗṣabha |*
> *yat iha asti tat anyatra yat na iha asti na tat kvacit ||*".

"On the topics of codes of conduct, wisdom, desire and the final destiny of life, one could find everything in the texts of Mahābhārata and if not, nowhere else."

This vast literature is a source of wisdom, which speaks about the human history, geography, the lineage of many dynasties of India, cultures, traditions, ethical frameworks, art and architecture, education system, trading and economic system, politics, the status of women and almost every aspect of human civilization.

It also deals with the scientific aspects of Vedic system, cosmology, astronomy, warfare and political strategies. Ultimately, it discourses on the Supreme Being, the souls, the philosophical

life and emancipation. Purāṇic literature claims that each verse of Mahābhārata including the Bhagavadgītā and Viṣṇusahasranāma, which are the condensed summaries of this treatise, opens up several dimensions.

1.2 Mahābhārata-tātparyanirṇaya

Ācārya Madhva (Ānandatīrtha), one of the philosophical thinkers in India has written a major commentary on Mahābhārata titled *Mahābhārata-tātparyanirṇaya* and also a short commentary titled *Mahābhāratatātparya (Yamakabhārata)*. The Tātparyanirṇaya comprises of 5180 verses in 32 chapters presents the stories of Rāmāyaṇa, Mahābhārata, Harivamśa, Bhāgavatapurāṇa and other major treatises in a chronological order that are apparently sporadic in these texts. This text emphasises on the philosophical understanding of these texts and addresses several questions raised about different interpretations. It also deals with several questions raised about ethical stances of this epic literature and offers solutions to philosophical conflicts that otherwise arise. The commentary is said to be a realistic perspective of Mahābhārata. The text of Mahābhārata-tātparyanirṇaya edited by Bannanje Govindacharya4) will be referred to the 'e-annotation'.

The details of Mahābhārata-tātparyanirṇaya is shown in Table - 1.

2 Mahābhārata Recensions

There are hundreds of manuscripts available on Mahābhārata with different recensions and retellings based on them, across nations. Amongst them, the most important editions popular(4) (based on northern recension) and critical edition(4) of BORI (Bhandarkar Oriental Research Institute, Pune) and southern recensions published by T R Krishnacharya of Kumbhakonam(4), and P.P.S. Sastry(4) of Chennai.

2.1 BORI and Kumbhakoṇam recension

The team led by Prof Vishnu S Suktankar had prepared the critical edition in BORI. This is widely accepted of 73730 verses4). Prof Muneo Tokunaga was the first person to key the whole text of BORI version into computer. Later it was revised by Prof John Dargavel Smith4).

In 1910, T R Krishnacharya and T R Vyasacharya edited and published a new version based on South-Indian texts4). This recension is also known as Kumbhakoṇam recension. It contains 96645 verse.

2.2 Difference between northern and southern recension

Apparently, there are more than 50 instances in the northern recensions that the narratives miss links to further stories. Nevertheless, the southern recension bridges such gaps. Moreover, southern recension also has provided answers to many queries otherwise bewilder the reader unable to understand details like: who is elder, what is the sequel, philosophical discussions and so on. As a detailed subject indexing under certain classified parameters with the help of required grammatical analysis will lead interpretations and discourses towards a better understanding of Mahābhārata. Table - 2 shows the details of both BORI and Kumbhakoṇam recensions.

Our research focuses on looking into analysis of the northern and southern recensions comparatively.

2.3 Computational Analysis of Mahābhārata

A group of researchers in Tirupati, Hyderabad and Bengaluru under the leadership of Prof Shrinivasa Varakhedi had digitised the whole text of this edition along with the support of Prof K V Ramakrishnamacharyulu, Prof D Prahlada Char and Prof Amba Kulkarni. This group had initiated the work on XML tagging of the subjects that are as indexed in the publication. This work is available[1] at both HTML and XML format at Search And Retrieval of Indic Texts (SARIT) website. The contents of the HTML files will be processed.

---

[1] http://sarit.indology.info/exist/apps/sarit/works/

Chapters	Verses
Sarvaśāstra-tātparyanirṇaya	138
Vakyoddhāra	183
Sargānusargalaya-prādhurbhāvanirṇaya	84
Śrīrāmāvatāre ayodhyāpraveśa	65
Śrīrāmacarite hanumaddarśanam	51
Śrīrāmacarite samudrataraṇaniścaya	59
Śrīrāmacarite hanumatpratiyānam (Sundarakāṇḍa)	50
Śrīrāmacarite śrīrāmadayādānam (Yuddhakāṇḍa)	248
Śrīrāmacarite rāmasvadhāmapraveśa (Uttarakāṇḍa)	143
Vyāsāvatārānuvarṇanam	88
Bhagavadavatārapratijā	237
Pāṇḍavotpatti	136
Kamsavadha	138
Uddhavapratiyānam	112
Pāṇḍava-śastrābhyāsa	66
Sṛgālavadha	33
Haṁsaḍibhakavadha	298
Bhimārjunadigvijaya	169
Pāṇḍavarājyalābha	223
Khāṇḍavadāha	245
Pāṇḍava-vanapraveśa	415
Araṇīprāpti	458
Ajātāvāsasamāpti	59
Yuddhodyoga	88
Bhīṣmapāta	144
Nārāyaṇāstropaśamanam	319
Karṇavadha	192
Pāṇḍavarājyalābha	241
Samastadharmasaṅgraha	63
Yāgasamāpti	179
Dhṛtarāṣṭrādisvargaprāpti	76
Pāṇḍavasvargārohaṇam	180
Total	**5180**

Table 1: Mahābhārata-tātparyanirṇaya content details

	BORI edition		**Kumbhakonam edition**		
Parva	Chapters	Verses/Proses	Upaparva	Chapters	Verses/Proses
Ādi	225	7187	18	260	10889
Sabhā	72	2392	8	103	4367
Vana	299	10284	21	315	12476
Virāṭa	67	1824	5	78	3494
Udyoga	197	6093	10	196	6753
Bhīṣma	117	5416	4	122	5809
Droṇa	173	8141	8	203	10012
Karṇa	69	3871	1	101	4975
Śalya	64	3301	3	66	3596
Sauptika	18	772	2	18	815
Strī	27	730	3	27	807
Śānti	353	12815	3	375	15151
Anuśāsana	154	6526	2	272	11194
Āśvamedhika	96	2743	3	118	4555
Āśramavāsika	47	1062	3	41	1097
Mausala	9	273	1	9	301
Mahāprāsthānika	3	106	1	3	111
Svargārohaṇa	5	194	1	5	232
Total	**1895**	**73730**	**98**	**2312**	**96645**

Table 2: BORI and Kumbhakoṇam recensions' details

Digital Corpus of Sanskrit, developed by Oliver Hellwig[2] is a magnificent e-corpus. Many Sanskrit texts with Morph-anaylsis are available with parallels and headlines. Its advance features for any selected lemma like n-grams, KWIC(Key Word In Context), Collocations, Occurrences, Meanings and a few more are very helpful for the research or to get a particular info from a text. Mahābhārata is also available in DCS with all the above mentioned features.

3   Research tools

A Consortium of Universities and Institutions has worked on development of Sanskrit computational tools and Sanskrit-Hindi Machine Translation System led by University of Hyderabad with the support of Government of India. Thus, the support of such tools will help to unambiguously interpret the Mahābhārata. It should be noted that a couple of institutions have ventured into whole and parts of the text analysis. However, the research focuses on analysis and indexing of the recension that is academically popular in southern parts of India.

There are a handful of tools designed by a group of Sanskrit Computational Linguistics experts for this task using standard models. Among them, the prominent ones are the Saṁsādhanī[3] developed by Prof Amba Kulkarni, Hyderabad and the Sanskrit Heritage[4]tool developed by Prof Gérard Huet, Paris. These tools enhance the understanding of even complex Sanskrit sentences (prose or poetry) and facilitate multi-dimensional interpretations following the Pāṇinian grammar. This helps a reader who is keen on learning Mahābhārata, but finds the grammatical aspects difficult to decipher. Hence, the grammatical analysis will foster the understanding of the syntactical structure of the text with a user-friendly interface.

---

[2]http://kjc-sv013.kjc.uni-heidelberg.de/dcs/index.php?contents=texte
[3]http://sanskrit.uohyd.ac.in/scl/
[4]http://sanskrit.inria.fr/

4 Pertinence of Research

Henceforth, the grammatical analysis and subject indexing of Mahābhārata will foster the understanding of the syntactical structure, semantics and discourses of the text with a user-friendly interface through computational tools employing the best research practices. This also provides the base for knowledge sharing and support building the existing applications. Eventually, it will bridge the gap between traditional and modern academia with the help of technology.

It should be noted that a couple of institutions have ventured into whole and parts of the text analysis as told earlier. However, the research focuses on analysis and indexing of the recension that is academically popular in southern parts of India. The XML files obtained through SARIT contained more than hundreds of verses that are seemingly missing. Furthermore, the present activity also proposes additional parameters that have not been indexed.

5 Methods

Grammatical analysis and subject indexing will consume years of research. This will involve a thorough reading and understanding of verses 101,825[5] of Mahābhārata and Tātparyanirṇaya looking into verifying, analysing and indexing of the whole text.

Grammatical analysis and subject indexing will consume years of research. This will involve a thorough reading and understanding of Mahābhārata and Tātparyanirṇaya looking into verifying, analysing and indexing of the whole text. Therefore, the e-editions of Mahābhārata need to be reviewed completely and in addition to that the print version of Tātparyanirṇaya to be digitised.

The action would be carried out in the following steps -

1. Obtaining the Mahābhārata and Tātparyanirṇaya e-text

2. Verifying (proofreading and editing) the e-recension

3. Grammatical analysis (includes identifying and tagging of grammatical aspects)

4. Subject indexing using XML in different parameters

We will see in detail about these above steps.

5.1 Verifying the e-recension

We are following the same BORI style[6] for verification process and numbering with some modifications.

- Rectifying typos and other errors (spacing, grammatical, etc.) in the e-text

- Inclusion of missing contents (a 'line' or more) with marker '#'

- Denoting the contents (a 'line' or more) that are not available in the published text yet published in the e-edition using marker '∗'

- The change of order in verse/prose or its parts (pada, ardha, etc.) are marked using '$'

- Numbering format PP-UU-AAA-SSS-L (Parva-Upaprva-Chapters-Verses-Line of the pāda)

- 'Tabular' key is used to indicate the gap between numbering and verse/prose

---

[5]96645 verses in Mahābhārata(Kumbhakonam Edition) and 5180 verses of Tātparyanirṇaya
[6]http://bombay.indology.info/mahabharata/text/UD/MBh00.txt

5.2 Grammatical Analysis

Understanding the grammatical aspects viz. Sandhi, Samāsa etc. are the essential to analyse any Sanskrit text. The identification and tagging of these grammatical aspects in the versified/prosified text involve the following processes -

- Word Segmentation (pada-pāṭha)

- Morph Analysis (pada-paricaya)

- Compound Analysis (samasta-pada-paricaya)

- Syntactico-Semantic relations (kāraka-paricaya)

We are using the above mentioned Saṃsādhanī and Sanskrit Heritage tools for this purpose. Presently, we are only concentrating on first three parts and later, we will consider the fourth one - kāraka-paricaya as well. We analysed the morphological aspects of first seven chapters of Ashvamedhika, which contains 169 verses using the Saṃsādhanī tool. We also verified the retrieved data, manually. We can see the performance of the tools in Table - 3.

Number of solutions	Number of cases	Percentage
Correct analysis	1292	57.11
Both correct and incorrect analysis	602	26.61
Incorrect analysis	256	11.31
Doubts	39	1.72
Typo-errors	73	3.22

Table 3: Performance of the tools in Morph-analysis

This performance of the tool shows that grammatical analysis work can be easily done with less effort. This result motivates us to give more attention to subject-indexing part, which is the crucial part of the research project.

5.3 Subject indexing

In this part, we will point out a few important topics and index it using XML (Extensible Markup Language). This part will help in acquiring the notable and precise information available in Mahābhārata. This is also to facilitate unambiguous interpretation of Mahābhārata. For instance, a person wanting to study the status of women during Mahābhārata period can find information about that particular information. Draupadi's dialogue with Satyabhāmā in Vanaparva can be one such example.

It is also important to note the existence of bipolar ideas as reflected in Viduranīti and Kaṇiṅkanīti providing space for debate, discussion and decisions based on consensus. Codes of conduct as listed out by Bhīṣma in Śāntiparva highlight democratic practices as well. Thus, subject-indexing helps us to extract such critical discussions from Mahābhārata. which later could form a theme for a symposium and may even lead to a scholarly publication.

The subjects discussed in the text will be identified and tagged under the classified parameters with the support of grammatical analysis. There shall be inclusions of new parameters or renaming of the identified parameters over time.

−> Definition of tags

−> Sieving of information embedded in the text for subject identification

−> Creation/modification of new/existing tags

−> Indexing of the subjects through XML

XML tagging would help classify the content on the criteria like following list -

- **C**ontent

  Parva-wise: Ādi and Sabhā

  Story-wise: Ākhyāna and Upākhyāna like Rājasūyaprasaṅga and Naadamayantīkathā

  Style-wise: Chandas, Vṛtta, Alaṅkāra and other in detail

According to the northern recensions Abhimanyu entered Cakravyūha. But according to southern recensions, it is Padmavyūha. In other instance, the northern recensions say Arjuna eloped in his true form. But southern recensions say, he eloped in disguise as an ascetic.

Mahābhārata-tātparyanirṇaya4) has extensively played with Chandas and Vṛtta that are used even to communicate the intensity or the story itself. For example, the first half is Vasantatilakā of Śakvarī and second half is Indravajrā of Triṣṭup in chapter 27, verse 11; this is a novel Upajāti. Here Vasantatilakā explains the approach of Bhīmasena in battle, and then Indravajrā illustrates how he leaps on foes. Moreover, the functions of the mainstream grammar has also interpolated some of Mahābhāratas and other such literatures ingenious Sanskrit words like bāhuṣāī, akṣohiṇī, saṅkha and duḥṣanta by modifying them as bāhuśālī, akṣauhiṇī, saṅkhya and duṣyanta.

- **N**amed Entities

  Character-wise: Kṛṣṇa and Bhīma

  Characteristic-wise: Jñāna and Bhakti

  Name-wise: Kṛṣṇa, Acyuta and Janārdana

  Period-wise: Kṛṣṇa's childhood and Rājasūya period

  Place-wise: Mathurā and Hastināpura

A Yakṣa by name Sthūnakarṇa is referred to in Mahabharata who meets Shikhandi. However, there is no such evidence of Sthūnakarṇa in the list of Yakṣas provided by other texts. Commentaries do come in such instances as a handy tool to depict that the Yakṣa Tumburu was also called Sthūnakarṇa.

Many scholars have debated on the interpretation of 'manvādi kecit bruvate' verse of Mahabharata. But Indian traditional scholarship emphasises that this verse claims Mahabharata defers to three striking dimensions referring to the historicity or narration of stories, psyche of an individual and spiritual understanding. Here, Mahābhārata-tātparyanirṇaya throws light on that how Mahabharata illustrates human psychology by depicting a characteristic of a person like Bhīma, Arjuna, Duryodhana and Śakuni, what they stand for which makes the literature relevant beyond times4).

- **S**ubject

  Ritual-wise: Yajñas and Saṃskāras

  Stotra-wise: Viṣṇusahasranāma and others

  Glossary-wise: explaining the meanings and derivations of technical words

  Description-wise: as described about principles in Bhagavadgītā

  Data-wise: Calculations of calendar system and Akṣohiṇī

Mahābhārata describes many kind of saṃskāras, Yajñas and many more rituals. We will tag those all ritual details.

The calculation of the classifications of army is well described during the preparation of battle. Patti refers to an elephant, a chariot, three horses and five infantry warriors. The

multiplication of Patti into three results into Senamukha, Gulma, Gana, Vāhini, Pṛtanā, Camū, Anikini and Akṣohiṇī4).

- **Context**

  Subject-wise: Arthaśāstra and Rājanīti

  Dialogue-wise: Bhagavadgītā and Bhiṣmopadeśa

  Śāstra-wise: Viduranīti and Sanatsujātīya

  Scientific discipline-wise: Description of Universe, Ecology and Social sciences

  Art-wise: (aesthetics domain)

The philosophical discussions such as the idea of Dharma, human agency, freewill in the chapter 32 of Vanaparva as instructed by Draupadi. The ontic debates among Pāṇḍavas on understanding of human pursuits such as codes of life, economy, desire and ultimate destiny in the in the chapter 165 of Śāntiparva.

6   Conclusion

The research activity that shall also employ new models to exegesise interpretations of the discourses of the Sanskrit literatures. Therefore, academically, it will have a greater impact on the field of Sanskrit, main-streaming the Indian traditional scholarship with the help of computational linguistics. This will benefit the global academia allied with such interests. This action will be definitely replicated leading to multiplier effect with many more important works in Sanskrit getting digitized and tagged employing new methods and models also help building technical tools.

Economically, it will create corpus that is free to access with the utility of technical tools and will result integrating traditional knowledge reserve with mainstream academia.

At the end, the activity will contribute a user-friendly interface with a manual guide that help scholars and researchers to study the epic literature and its philosophical annotations like Tātparyanirṇaya. This will increase the understanding of Indian epics and philosophy with a focus on contemporary relevance.

References

F. M. Ferrrari. Gender and Narrative in the Mahābhārata - review. *South Asia Research (London and New York: Routledge)*, 30(2):199–201, 2010.

Bannanje Govindacharya. *Mahabharata-tatparyanirnaya and Mahabharata-tatparya of Anandatirtha.* Īśāvāsya Pratiṣṭhāna, Udupi, 1999.

Bannanje Govindacharya. *Bhimasena: Vyasaru Kandante.* Īśāvāsya Pratiṣṭhāna, Udupi, 2008.

Bannanje Govindacharya. *Mahabharata-tatparyanirnaya of Anandatirtha.* Tattvasamshodana Samsat, 2009.

Bannanje Govindacharya. *Bhagavadgītā in Kannaḍa.* Īśāvāsya Pratiṣṭhāna, Udupi, 2015.

Gérard Huet. Sanskrit Heritage tools `sanskrit.inria.fr`.

A B Keith and MacDonell. *Vedic Index of Names and Subjects.* Motilal Banarasidas, New Delhi, 1912.

T.R. Krishnacharya and T.R. Vyasacarya. *Sriman Mahabharatam: A new edition mainly based on the South Indian texts, with footnotes and readings.* Divine Books, New Delhi, 2011.

Amba Kulkarni. Sanskrit Computational Linguistics (SCL) tools `sanskrit.uohyd.ac.in/scl`.

Amba Kulkarni. Tagging guidelines. 2002-2016 `http://sanskrit.uohyd.ac.in/Corpus/guideline.html`.

Vettam Mani. *Purāṇic Encyclopedia: A Comprehensive Dictionary with Special Reference to the Epic and Purāṇic Literature.* Motilal Banarasidas, New Delhi, 1975.

M. A. Mehendale. *Mahabharata - Cultural Index.* Bhandarkar Oriental Research Institute, Pune, 1993.

P. P. S. Sastri. *The Mahabharata (Southern Recension).* Vavilla Ramaswamy Sastrulu Press, Madras, 1931.

S Sorensen. *An Index to the Names in the Mahabharata with short explanations and a Concordance to the Bombay and Calcutta Editions and P C Roy's translation.* Motilal Banarasidas, New Delhi, 1963.

Vishnu Suktankar. *The Mahābhārata - Critical Edition.* Bhandarkar Oriental Research Institute, Pune, 1942.

Muneo Tokunaga and John Dargavel Smith. *The Mahābhārata.* Bhandarkar Oriental Research Institute, Pune, 1997-2001.