COMPUTATIONAL SANSKRIT

Postgraduate Progamme

Sanskrit Language and Literature

M23SN01AC

Self Learning Material

SREENARAYANAGURU OPEN UNIVERSITY

(The State University for Education, Training and Research in Blended Format, Kerala)

SREENARAYANAGURU OPEN UNIVERSITY

Vision

To increase access of potential learners of all categories to higher education, research and training, and ensure equity through delivery of high quality processes and outcomes fostering inclusive educational empowerment for social advancement.

Mission

To be benchmarked as a model for conservation and dissemination of knowledge and skill on blended and virtual mode in education, training and research for normal, continuing, and adult learners.

Pathway

Access and Quality define Equity.

Computational Sanskrit

Course Code: M23SN01AC

Semester - I

Ability Enhancement Compulsory Course
Postgraduate Programme
Sanskrit Language and Literature
Self Learning Material

വിദ്യകൊണ്ട് സ്വതന്ത്രരാവുക

SREENARAYANAGURU
OPEN UNIVERSITY

SREENARAYANAGURU OPEN UNIVERSITY

The State University for Education, Training and Research in Blended Format, Kerala

M23SN01AC
Computational Sanskrit
MA Sanskrit Language & Literature
Semester - I

SREENARAYANAGURU OPEN UNIVERSITY

Printed and published on behalf of Sreenarayanaguru Open University by Registrar, SGOU, Kollam.
www.sgou.ac.in

ISBN 978-81-966907-9-3
9 788196 690793

DOCUMENTATION

Development of the Content
Dr. Sivaja S. Nair

Review
Content : Dr. Christy D. Ponnan
Format : Dr. I. G. Shibi
Linguistics : Dr. Christy D. Ponnan

Edit
Dr. Christy D. Ponnan

Scrutiny
Dr. Vijayarajan K. U.

Co-ordination
Dr. I. G. Shibi and Team SLM

Design Control
Azeem Babu T. A.

Production
January 2024

YouTube

MESSAGE FROM VICE CHANCELLOR

Dear
I greet all of you with deep delight and great excitement. I welcome you to the Sreenarayanaguru Open University.

Sreenarayanaguru Open University was established in September 2020 as a state initiative for fostering higher education in open and distance mode. We shaped our dreams through a pathway defined by a dictum 'access and quality define equity'. It provides all reasons to us for the celebration of quality in the process of education. I am overwhelmed to let you know that we have resolved not to become ourselves a reason or cause a reason for the dissemination of inferior education. It sets the pace as well as the destination. The name of the University centres around the aura of Sreenarayanaguru, the great renaissance thinker of modern India. His name is a reminder for us to ensure quality in the delivery of all academic endeavours.

Sreenarayanaguru Open University rests on the practical framework of the popularly known "blended format". Learner on distance mode obviously has limitations in getting exposed to the full potential of classroom learning experience. Our pedagogical basket has three entities viz Self Learning Material, Classroom Counselling and Virtual modes. This combination is expected to provide high voltage in learning as well as teaching experiences. Care has been taken to ensure quality endeavours across all the entities.

The university is committed to provide you stimulating learning experience. The postgraduate programme in Sanskrit has a unique blend of language and literature. The focus of the programme is on enhancing the capabilities of the learners to undergo a deeper comprehension of the sociology of the forms in literature, although the required credits are in place to learn other aspects of Sanskrit literature. Care has been taken to expose the students to recent trends in Sanskrit literature. We assure you that the university student support services will closely stay with you for the redressal of your grievances during your studentship.

Feel free to write to us about anything that you feel relevant regarding the academic programme.

Wish you the best.

Regards,
Dr. P. M. Mubarak Pasha

01.01.2024

Contents

Introduction to Sanskrit Computation

BLOCK - 1

Important Sites of Sanskrit Literary Database

Learning Outcomes

- Introducing Sanskrit literary databases
- Introducing different Sanskrit sites
- Enabling to decipher a text with technology
- Enabling to revive traditional methods of learning Sanskrit with technology

Background

Sanskrit is very rich in content and highly rule based in grammar. In this unit we will be discussing Sanskrit literary databases and tools for better understanding and usage of Sanskrit language.

Keywords

Inflected language, Sanskrit language tools, literary database

Discussion

Sanskrit language has a very rich structure and it is a highly inflected language. Sanskrit uses prefixes, suffixes, infixes. It also has a rich system of compounds, which uses commonly in texts.

The nominal system of Sanskrit includes nouns, pronouns, and adjectives. It has three genders as masculine, feminine, and

The nominal system of Sanskrit includes nouns, pronouns, and adjectives

neuter, three numbers as singular, dual, and plural, and seven syntactic cases as nominative, accusative, instrumental, dative, ablative, genitive, and locative. The nouns can generate from a root word (prātipadika) with a suffix from the list of twenty-one suffixes (sup pratyayas). The root word can be from the categories of verb + kṛt-suffix, noun + taddhita-suffix, and compounds.

The verbal system of Sanskrit can be crosslinked in one of three ways: first, second, or third person. There are three numerical forms for verbs: singular, dual, and plural. There are ten tenses and three voices for verbs: active, middle, and passive. The verbs can generate from a verb root (dhātu) which is listed in Pāṇini'sDhātupāṭha or a verb root derived adding sanādi-pratyayas which can be from both noun and verb.

The grammar of Sanskrit is rule based, formula based and logical

The word order of a sentence in Sanskrit can be flexible at most of the cases. Otherwise, the order is SOV (Subject Object Verb). The grammar of Sanskrit is rule based, formula based and logical.

1.1 Importance of Sanskrit tools and Literary Database

We need to have literary database in Sanskrit

To understand the nature of high productivity of words in Sanskrit and the efficient usage of the language can assist by computer tools for Sanskrit. Also, to access and process the texts in Sanskrit in today's digital world we need to have literary database in Sanskrit.

1.1.1 Sanskrit Computational Tools

The different kinds of tools are available for Sanskrit

Sanskrit computational tools assist the user to analyse texts in Sanskrit and helps to understand the grammatical structure of Sanskrit. There are different kinds of tools are available for Sanskrit. They are morphological analyser, morphological generator, Sandhi splitter, Sandhi joiner, Compound analyser, Kāraka analyser, etc.

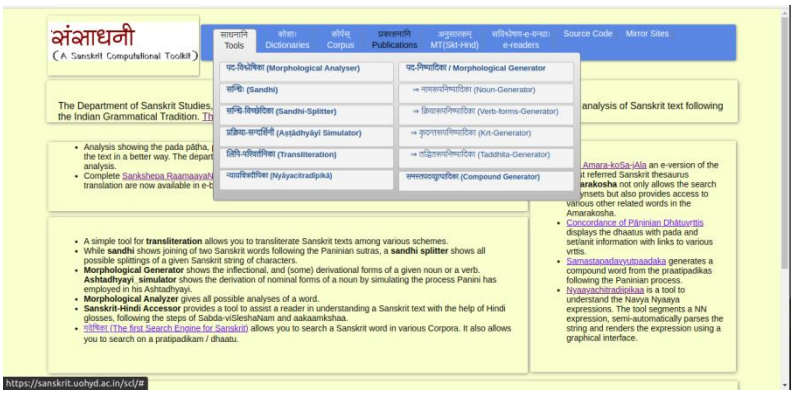
University of Hyderabad, IIT Kanpur, IIIT Hyderabad, Jawaharlal Nehru University, New Delhi, Inria, France etc. are the pioneers in the field of Sanskrit computational Linguistic. They developed different tools which helps to analyse and interpret a Sanskrit text.

1.1.1.1 Samsaadhanii (संसाधनी)

Samsaadhanii is a Sanskrit computational toolkit

Samsaadhanii is a Sanskrit computational toolkit. The Department of Sanskrit Studies, University of Hyderabad, is engaged in developing various computational tools for the analysis of Sanskrit text following the Indian Grammatical Tradition.

Figure – 1 Samsaadhanii home page



Samsaadhanii contains the following tools (as shown in figure 1) related to Sanskrit computations developed under the guidance of Prof. Amba Kulkarni. They are :

Word analyser - gives all possible analyses of a word in Sanskrit.

- Word generator – shows the inflectional, and derivational forms of a given noun or a verb. They have the subdivisions like -
- Noun form generator
- Kridanta form generator
- Tiṅanta form generator
- Taddhita form generator
- Samasta pada vyutpādaka
- Sandhi joiner - shows joining of two Sanskrit words following the Paninian sutras
- Sandhi splitter - shows all possible splittings of a given Sanskrit string of characters.

- Anusaaraka Sanskrit-Hindi Machine Translation system - provides a tool to assist a reader in understanding a Sanskrit text with the help of Hindi glosses, following the steps of śabdaviśleṣaṇam and ākāṅkṣā.
- Transliteration modules - allows you to transliterate Sanskrit texts among various schemes.
- The Amara-kosa-jala - an e-version of the most referred Sanskrit thesaurus Amarakosha not only allows the search of synsets but also provides access to various other related words in the Amarakosha.
- Ashtadhyayi Simulator - shows the derivation of nominal forms of a noun by simulating the process Panini has employed in his Ashtadhyayi.
- Navya-Nyaya Diagramatic Representation - (Nyaayachitradiipikaa) is a tool to understand the Navya Nyaaya expressions. The tool segments a NN expression, semi-automatically parses the string and renders the expression using a graphical interface.
- Compound word generator - (Samastapadavyutpaadaka) generates a compound word from the praatipadikas following the Paninian process.
- Concordance of PāṇinianDhātuvṛttis - displays the dhaatus with pada and set/anit information with links to various vrttis.
- E-readers for saṃkṣeparāmāyaṇam, bhagavadgītā and śiśupālavadham - showing the pada pāṭha, pada paricaya, samasta pada paricaya and kārakaanvaya help in understanding the text in a better way. Also with Hindi and English translation are now available.
- Sanskrit-Hindi Machine Translation and Accessor - All these tools are joined together to get the translation from Sanskrit into Hindi. The output is Translation, with word by word meaning, and word analysis of each word.

  The link is - https://sanskrit.uohyd.ac.in

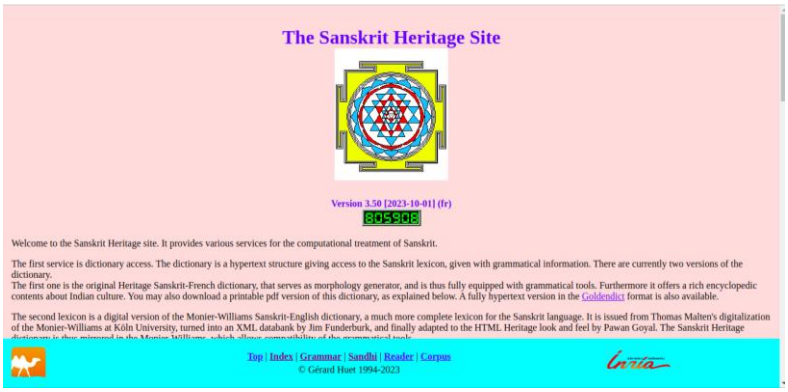
1.1.1.2 The Sanskrit Heritage Site

The Sanskrit Heritage Site

The Sanskrit Heritage site provides various services for the computational treatment of Sanskrit developed by Inria – a French institute for research in computer science and applied mathematics by Prof. Gérard Huet. The website looks like as the figure 2.

The tools available here is -

Figure 1: The Sanskrit Heritage Site



- Sanskrit Heritage dictionary - is a hypertext structure giving access to the Sanskrit lexicon, given with grammatical information.
- Sanskrit Reader - that parses Sanskrit text under various formats into Sanskrit banks of tagged hypertext. Various phonological and morphological tools are also provided.
- Morphological tools
- Grammar - gives you declined forms of nouns and conjugated forms of root verbs. It is the workhorse of morphological derivation.

The link is – https://sanskrit.inria.fr/

- Stemmer – provides the morphological tagging of a word.
- The Sanskrit Engine
- Sandhi - The Sandhi Engine takes two phoneme streams (input as transliterated strings) and gives as result their sandhi euphonic composition.
- Reader - The Sanskrit Reader Companion allows the analysis of Sanskrit sentences.

The link is - https://sanskrit.inria.fr/

1.1.1.3 Sambhasha (संभाषा)

Sambhasha is a Sanskrit computational toolkit

Sambhasha is a Sanskrit computational toolkit developed by the Computational Linguistics researchers at Karnataka Samskrit University, engaged in developing various computational tools for Sanskrit and allied languages under the Department of Vyakarana, KSU.

- Mahakosha - Kosha data navigation tool.
- Aayussamskritam - Courseware for Ayurvedic Medical Students.
- Tarkasaṅgrahaḥ - Ontological knowledge representation
- Vāṅmayī-e - Create text – explore the meaning.
- Sāṅgaṇikapāli - Pali language tools
- प्राकृत-स्त्रीप्रत्ययान्तरूप-निष्पादिका - Prakrit Verbforms
- Nāmadhāturūpikā - Denominative verbs
- Prācīnanyāyaḥ - Ontological knowledge representation
- शास्त्रीयस्फूर्तिस्पर्धा - Sanskrit Quiz App
- Bhaṭṭikāvya-nandinī - Explore Bhattikavya with Karaka dependency diagram
- Anvayacitraṇam - tool Draw relations
- Vaijayantīkośa Knowledge net – Synset and other semantic relations in the lexicon Vaijayantīkośa
- Śabdakautukam - Practice noun forms

The link is - https://sambhasha.ksu.ac.in/

1.1.1.4 TDIL

TDIL stands for Technology Development for Indian Languages

TDIL stands for Technology Development for Indian Languages. It's a program initiated by the Ministry of Electronics and Information Technology (MEITY) of the Government of India. The program's goal is to develop information-processing tools that facilitate human-machine interaction in Indian languages.

The TDIL program has established 12 Resource Centres for Indian Language. It has also created a repository of achievements in the form of tools, fonts, applications, and standards.

The TDIL program has created:

- Information-processing tools
- A repository of achievements
- Free Indian languages software and tools

The TDIL program is important because language technology, aided by artificial intelligence and machine learning, is going to be one of the most important disruptive technologies in the immediate future.

1.1.1.5 Other Tools

There are so many initiatives are happening in the present world using AI as base etc.

Sampark: Indian Language Machine Translation System

Centre for Development of Advanced Computing (C-DAC) have Multilingual Tools. Sampark: Indian Language Machine Translation System it is an Indian Language to Indian Language Machine Translation System. Also, C-DAC had developed so many other tools for language analysis.

Udaan is an initiative of IIT Bombay, Mumbai. The project aims to break the language barrier in education.

Anusaaraka is a Machine Translation tool

Language Technologies Research Centre, IIIT Hyderabad Anusaaraka is a Machine Translation tool that uses insights from traditional Indian knowledge of language along with modern technologies to make a machine translation tool. LTRC also runs other research projects like building corpora, building lexicons, and building speech processing tools.

Center for Indian Language Technology, IIT Bombay CFILT has undertaken the project of building WordNet—an international standard for online thesaurus and lexicon—or Indian languages. UNL and WordNet are both useful resources for machine translation projects. Obviously, IIT Bombay has many other projects in the field of NLP for Sanskrit and other Indian languages.

Computational Linguistics R&D at the School of Sanskrit and Indic Studies, JNU This R&D centre at JNU has produced multiple tools like the Multilingual Online Amarkosh, an online thesaurus for English, Sanskrit, Hindi, and many other Indian languages.

IIT Kanpur developed Anglabharti, a machine-aided translation

IIT Kanpur A machine-aided translation system for the specific domain of Public Health Campaigns called Anglabharti has been developed. The system analyses sentences and converts them into an almost disambiguated intermediate structure which can then potentially be used for translation of English into any Indian language. They are also involved in developing other projects, lexical resources to spell-checkers.

1.1.2 Sanskrit Literary Database

A literary database is an online collection of references of published works. These databases can include:

- Books
- Journal and newspaper articles
- Conference proceedings
- Reports
- Government and legal publications
- Patents

A literary database is an online colletion of references of published works

Literary databases for Sanskrit can focus on specific disciplines or be multidisciplinary. Some literary databases for Sanskrit include:

1.1.2.1 https://archive.org/ is a digital library which handles different file formats like pdf, video, audio, softwires etc. Sanskrit books are available in this service as enormous in number.

1.1.2.2 https://www.jstor.org/ is a digital repository of journals. Here we can find articles and research papers in the fields Humanities and Sanskrit studies.

1.1.2.3 GRETIL - Göttingen Register of Electronic Texts in Indian Languages - is having a collection of Sanskrit texts in editable format.

1.1.2.4 https://sanskrit.nic.in/ebooks.php is a collection of the books which are published or made as study material for Central Sanskrit University, New Delhi.

1.1.2.5 https://www.sanskritebooks.org/ is lists Sanskrit texts and can download it in PDF format.

1.1.2.6 'Samsaadhanii' – provides the digital corpora in Sanskrit

1.1.2.7 'The heritage site' is also providing a digital corpora in editable format.

1.1.2.8 https://sa.wikisource.org/ have a collection of Sanskrit texts in editable format.

1.1.2.9 https://www.ebharatisampat.in/ is having the collection of Sanskrit texts in editable format.

1.1.2.10 There are so many websites or portals or blogs which provides different kinds of Sanskrit E-texts in editable format or in the form of pdf available

Summarised Overview

Sanskrit language is highly productive and syntax is rule based. At this, understanding and studying any Sanskrit text will become easy with the help of different Sanskrit Tools at each stage of study. Some important tools are introduced at this point of view. Any literature study needs the availability of texts. At this modern age most of the texts are available in digital form. Sanskrit literary databases are introduced here in this thought. Important and available databases of Sanskrit is listed in this unit.

Assignments

1. Visit the following websites and list the facilities available in it.
   https://anuvadini.aicte-india.org/
   https://sambhasha.ksu.ac.in/
   https://sanskrit.inria.fr/
   https://sanskrit.uohyd.ac.in
2. Find out the morphological analysis of any 20 words in different tools and try to find the differences of style.
3. What are the uses of Sanskrit tools and categorize it.
4. List different Sanskrit Literary Database.

Suggested Readings

1. https://sanskrit.uohyd.ac.in/scl/scl_help.pdf
2. https://sanskrit.inria.fr/Symposium/DOC/Amba_intro.pdf

References

1. https://sanskrit.uohyd.ac.in/scl/#
2. https://archive.org/
3. https://anuvadini.aicte-india.org/
4. https://sambhasha.ksu.ac.in/
5. https://sanskrit.inria.fr/

Space for Learner Engagement for Objective Questions

Learners are encouraged to develop objective questions based on the content in the paragraph as a sign of their comprehension of the content. The Learners may reflect on the recap bullets and relate their understanding with the narrative in order to frame objective questions from the given text. The University expects that 1 - 2 questions are developed for each paragraph. The space given below can be used for listing the questions.

Sanskrit E-Resources and Online Resources

Learning Outcomes

- Knowing about the e-resources and online resources for Sanskrit
- Finding the pinpointed information using e-resources and online resources
- Analysing a problem using technology
- Enabling the critical thinking of the student

Background

In unit 1 we learned about Sanskrit websites and literary database. In this unit we go through E- resources and online resources in Sanskrit. E-resources are important for teaching, learning, research, and training. As well as online resources can be important for learning, communication, and developing information literacy. So knowing and using about these both are equally important in this digital era for an academic environment.

Keywords

E – resources, Online resources, E-journal, E-book

Discussion

Electronic resources developed for Sanskrit studies

Electronic resources, or e-resources, and Online resources refer to digital information and materials accessed and utilized through various devices, such as computers, smartphones, tablets etc. The materials in readable or editable format can call E-resources and the online tools, and other services etc. can call as online resource.

Most of the time the words e-resources and online resources are using as synonym. Both of these are created for educational purposes or any data collectional purposes. Here we are going to discuss the e- resources and the online resources which are developed for Sanskrit studies.

2.1 E-resources

E-resources are materials in digital format that are accessible electronically. They can be a collection of texts, videos, audios, images etc.

Examples of e-resources for Sanskrit can classify in to :

- Electronic journals (e-journals)
- Electronic books (e-books)
- Online databases in varied digital formats
- Online magazines
- Online newspapers
- Audio Video collections

2.1.1 Electronic journals (e-journals)

E-journals, or electronic journals, are periodicals that are published in electronic format. They are digital versions of scholarly publications that are available online.

E-journals are convenient and up-to-date research tools. They can be read 24 hours a day, regardless of the libraries' hours of operation.

E -journals

They also provide libraries with flexibility and convenience in terms of storage and accessibility. E-journals can be stored on servers or cloud storage and accessed by multiple users at the same time.

There are so many e-journals available for Sanskrit. Some of them are listed below.

2.1.1.1 Central Sanskrit University Journals

Central Sanskrit University (CSU) publishes e-journals from each campus. All of them are made available up-to-date in online mode. Here are the list

CSU publishes E-journals form each campus

- CSU Guruvayur campus (Kerala) publishes the e-journals for Sanskrit are - nibandhamālā, śikṣāvāhinī. The link is https://www.csu-guruvayoor.edu.in/journals.html
- CSU Gaṅgānāthajhā Campus, Prayāgarāja publishes the e-journals for Sanskrit are -
- Journal of Ganganath Jha Campus Link - https://www.csu-prayagraj.res.in/Journals.php
- uśatī

  Link-https://www.csu-prayagraj.res.in/Ushati%20volumes.php
- CSU K G Somaiyā Campus (Mumbai) publishes the e-journals for Sanskrit are vāgvai brahma, vidyāraśmi, śikṣāraśmiḥ, jyotiṣaraśmiḥ, kāvyalatikā, goṣṭhīgavyam

  Link - http://www.csu-mumbai.edu.in/publications.html
- CSU Bhopal Campus Sanskrit Journals Available - rāṣṭrī, vāgvilāsinī (vyākaraṇamīmāṃsā), sāhityamīmāṃsā, jyotiṣamīmāṃsā, jainadarśanamīmāṃsā, śikṣāvallarī

  Link -http://www.csubhopal.edu.in/publications.html
- CSU śrīraghunāthakīrti campus (Devaprayāga) Sanskrit Journals - śābdītripathagā, sāhitī, daivī, nyāyādiśāstrārtha-mañjarī, āmnāyahaimavatī, raghunāthakīrtipatākā, raghunāthavārtāvalī Link - http://www.csu-devprayag.edu.in/publication.php
- CSU śrīraṇavīra campus (Jammu) Sanskrit Journals Available - śrīvaiṣṇavī, śikṣāmṛtam Link - http://www.csu-jammu.edu.in/publications.php
- CSU Vedavyāsa Campus (Himacalapradesh) Sanskrit Journals Available - vedavipāśā, niḥśreyasī, khagolaḥ, padagaṅgā, sāhityavedikā

  Link - https://csu-balahar.edu.in/Journals.html

2.1.1.1 Other E-journals in Sanskrit

- Mahasvini – National Sanskrit University, Thirupati

  Link - https://nsktu.ac.in/index.php/mahashvini/

- Kalāvaibhava by Indirākalā-saṅgītaviśvavidyālaya

  Link - https://iksv.ac.in/pages/view/17

- Kiraṇāvalī - Sanskrit Research Foundation, Thiruvananthapuram

  Link - https://sites.google.com/view/kiranavalionline/home

- Pāṇinīyā - MaharṣiPāṇinīSaṃskṛtaEvaṃ Vaidika viśvavidyālaya

  Link https://www.mpsvv.ac.in/pANinIyA-traimAsikashodhapata/

- Prajñā – Kāśīhindūviśvavidyālayaḥ

  Link - https://internet.bhu.ac.in/research_pub/pragya.php

- Vedavidyā – Maharṣisāndīpanirāṣṭriyavedavidyāpratiṣṭhāna

  Link - https://msrvvp.ac.in/veda_vidya.html

- Vyāsaśrīḥ- MaharṣiVedavyāsa National Research Institute

  Link - http://www.vyasasri.com/issue.php

- Paachiprajnaa

  Link - https://sites.google.com/site/praachiprajnaa/home

- Jahnavi - Online Quaterly

  Link - www.jahnavisanskritejournal.in

- GurukulaKāngaḍī Sama viśvavidyālaya, Haridwar

- Vaidika vāgjyotiḥ

  Link - https://www.gkv.ac.in/vedic-vag-jyoti/

- Gurukulapatrikā

  Link - https://www.gurukulpatrika.in/downloads/

2.1.2 Electronic books (e-books)

E -book or electronic book is a book that is available in electronic form. E-books are designed to be read on devices like computers, tablets, or smartphones. They can contain text, images, or both.

E -books

E-books are usually downloadable as a PDF. They are not editable, which protects the author's work.

2.1.2.1 archive.org

This repository hosts directly downloadable PDFs as personal collections and others for example the copies from digital library of India is avalibale in https://archive.org/details/digitallibraryindia

The books are categorized in different sections/links, e.g. Langauge wise Sanskrit.

More than this scanned books as well as music, audiovisual and web contents. The books are available in different formats such as PDF, B and W PDF, text, djvu etc.

2.1.2.2 http://dli.sanskritdictionary.com

When the Digital Library of India discontinued then this site has complete mirror of it. It allows "deep OCR based searches" within documents activated by its own OCR engine.

http://dli.sanskritdictionary.com

2.1.2.3 The digital library at Osmania University, Hyderabad

At http://oudl.osmania.ac.in/ hosts thousands of directly downloadable PDF files of many of the books. The OUDL site appears to be available only during Indian office timings (IST 10 am to 6 pm).

2.1.2.4 The Universal Library at Pittsburgh, USA

At http://www.ulib.org/ has plenty of books scanned under Million Books project. Sanskrit books are also available here

2.1.2.5 Google Books

It havea list of Sanskrit books. All books are in PDF format.

2.1.2.6 Archaeological Survey of India

asi.nic.in has made available over two thousand scanned books in PDF format and is set to more than ten thousands. These are digitized at Indira Gandhi National Centre for the Arts (IGNCA).

2.1.2.7 E-books at National Sanskrit University, New Delhi

Sanskrit Texts are available here for download.

2.1.2.8 E-books at Library systems

Other than this the libraray systems makes it's own collection of e-books. In the library system where Sanskrit books are available are -

- The Library System, University of Delhi has Digital Library with many PDF files.
- West Bengal Public Library Network, has scanned books on various topics. Use dspace database.
- Dspace books at Sanskrit department at Pondicherry University.

2.1.3 Online magazines

Online magazines are magazines that are published on the internet. They are also known as ezines, e-zines, eMagazines, or webzines.

Online magazines

- **Saptavarṇā** is the World's first Sanskrit e-magazine for children. A fun, interactive magazine, it is a beautiful combination of technology and learnings of Sanskrit.

  Link - https://www.languagecurry.com/saptavarna.html
- Vākpatrikā(Fortnightly)
- Link - https://vaksanskrit.com/?page_id=2950

2.1.4 Online newspapers

Online newspapers

- Sudharma Daily

  Link - https://epapersudharmasanskritdaily.in/
- Samprativartha

Link - http://www.samprativartah.in/

- Sanskritvarta

  Link - https://sanskritvarta.in/

- **Telegram and Whatsapp Sanskrit News Channels**

  ●Ramdootah "संस्कृतसंवादः"
  Telegram: https://t.me/ramdootah

  - **Whatsapp New channel Sanskrit**

  ●Sajala Sandeshah (WhatsApp to subscribe)
  https://api.whatsapp.com/send?phone=+919540948998

  ●आह्निकवार्तापत्रम्(WhatsApp to subscribe)
  https://api.whatsapp.com/send?phone=+919415633297

2.1.5 Online databases in varied digital formats

Data in audio format and video format are also available in internet. They are also a kind of e-resource.

2.1.5.1 Audio collections

There are so many YouTube channels are available in this regard. The important links are given below.

Audio and Video collections

- All India Radio daily (twice) Sanskrit News audio, News on AIR (Audio + Text bulletin)

  Link - https://newsonair.gov.in/RNU-NSD-Audio-Archive-Search.aspx

- Saṃskṛtasāptāhikī - Weekly audio programme on News on AIR

  Link - https://newsonair.gov.in/Audio-Archive-Search.aspx

2.1.5.2 Video collections

- DD news (Daily vaarta video and weekly Vaartavali)

  Link - https://youtube.com/@ddnews

- Janam TV (Daily news video at around 5PM)

  Link - https://youtube.com/@JanamTVmedia

2.2 Online resources

Online resources

Online resources are web pages and documents in the form of online tools or other services that provide useful information. They can be educational, data-driven, or support software.

The online resources for Sanskrit can divide in to two according to it's process as Language Processing tools and Speech and text porcessing tools.

2.2.1 Language Processing tools

The language processing tools helps to process text in language aspect. In the unit 1 section 1.1.1 we came accross so many language computational tools. These all are language processing tools. There we focused on the website. Here we are describing each type of tools from different sources. These tools help to access the languge by our own.

2.2.1.1 Dictionaries

- Spokensanskrit: An online hypertext dictionary for Sanskrit-English and English-Sanskrit

  Link - https://www.learnsanskrit.cc/

- Sanskrit.Today : An online dictionary by Sankrit.Today 'Kosha' is a concodance of the dictionaries in Sanskrit like Vacaspatya, Sabdakalpadruma, Apte, Monier Williams, Spokensanskrit, etc. and the corpuses like Vedas, Puranas, Mahabharatha, Ramayana etc.

  Link - https://kosha.sanskrit.today/

- Cologne Digital Sanskrit Dictionaries

Online Sanskrit Dictionaries

The Cologne Digital Sanskrit Dictionaries is a collection of Sanskrit lexicons prepared since 1994 by the Institute of Indology and Tamil Studies, Cologne University.
The 38 dictionaries are organized primarily by the secondary language (English, German, etc.). Each dictionary has several types of display, as well as PDF scan and XML files for download. All dictionaries are also available for offline usage in android phones also.

Link - https://www.sanskrit-lexicon.uni-koeln.de/

- Sanskrit Dictionary

  Link - https://sanskritdictionary.com/

- Digital Dictionaries of South Asia

  Link - https://dsal.uchicago.edu/dictionaries/macdonell/

- Lexilogos Sanskrit Dictionary

  Link - https://www.lexilogos.com/english/sanskrit_dictionary.htm

2.2.1.3 Sanskrit E-readers

Sanskrit E - readers – machine-assisted analysis of sentences

- The Sanskrit Reader Companion by the Sanskrit heritage site. Sanskrit Reader allows machine-assisted analysis of Sanskrit sentences, that is segmentation (including sandhi viccheda), morphological tagging, and several parsers.

  Link - https://sanskrit.inria.fr/DICO/reader.fr.html

- **E-Readers by Samsaadhanii**

  Link - https://sanskrit.uohyd.ac.in/scl/

  The e-readers are developed by Department of Sanskrit Studies, University of Hyderabad. It captures the salient features of the book padapāṭha, padaparicaya, ākāṅkṣā, anvaya, the pictorial representation of the kāraka analysis, samāsa analysis and samāsa-vigra with the picture showing the composition.

- Saṅkṣēparāmāyaṇa
- Bhagavad Gita
- Śiśupālavadha
- Raghuvaṃśam

E-Readers are developed by Dept.of Sanskrit, Hyderabad University

- Bhaṭṭikāvya-nandinī is the e-reader of Bhaṭṭikāvya by Department of Vyakarana, Karnataka Sanskrit University, Bengaluru. It displays padapāṭha, padaparicaya, anvaya, the pictorial representation of the kāraka analysis, samāsa analysis and samāsa-vigra.

2.2.1.4 Transliteration Tools

Transliteration is the conversion of a text from one script or encoding to another.

- **Sanscript**

  Sanscripttranslitate from Indian languge scripts and other encoding schemes like IAST, SLP, Harvard-kyoto, ITRANS etc. to any of the same set of scipts.

  Link - https://www.learnsanskrit.org/tools/sanscript/

Translataration is the conversion of a text from one script to another

- **Aksharamukha : Script Converter**

  Aksharamukha aims to facilitate transliteration between various scripts. It also specifically provides lossless transliteration between the main Indian scripts including the scripts which are used in manuscripts like Brahmi, Sharada, Vatteluttu etc along with Sinhala. Aksharamukha as of now supports 120 scripts and 21 romanization methods.

- Link - https://aksharamukha.appspot.com/converter

- There are so many tools are available for transliteration according to the purpose. Like Samsaadhanii Transliteration tool helps to transliterate in between different textual encodings (WX encoding, SLP, Harvard-kyoto etc.)

2.2.1.5 Spell Checkers

There is no fully functional spellchecker for Sanskrit is available now. However, there are efforts are going on this way and some tools that can help with spelling:

Spell checker in Sambhashahttps://sambhasha.ksu.ac.in/spellchecker/

Spell Checkers

Sanskrit Spell Checker for Firefox is an add-on available for Firefox.

Sanskrit Spell Checker for Thunderbird is an add-on is available for Thunderbird.

2.2.1.6 Optical Character Recognition (OCR)

OCR is a technology that converts text in a digital image into a machine-readable text format. OCR is commonly used to recognize text in scanned documents and images.

OCR software can be used to:

- Convert a physical paper document or an image into an accessible electronic version with text
- Convert any kind of text or information stored in digital documents into machine-readable data
- Convert hard copies and paper documents into computer-readable file formats, suitable for further editing or data processing
- OCR converts the file formats JPEG, PNG, GIF, BMP, TIFF, and PDF in to computer-readable file format

OCR - Optical Character Recognition

Sanskrit OCR

The Sanskrit OCR is developed by Sanskritdictionary.com. In the web page we can paste or upload the image in any format as mentioned above and it will provide the out put in editable format.

Link - https://ocr.sanskritdictionary.com/

Devanagari OCR

Devanagari OCR converts image to text format.

Link - https://devnagri.com/ocr/

There are other OCRs for Sanskrit is avalibale like Nanonets, Iron OCR, Pramukh OCR for android phones etc.

2.2.2 Text and speech processing tools

Even though Sanskrit is one of the best languages for natural language processing (NLP) because of its highly organized structure and minimal ambiguity, Processing Sanskrit texts can be challenging because

Sanskrit is one of the best language for NLP

- High lexical productivity of words
- Free word order in poetry and sometimes in prose

- Euphonic (sandhi) assimilation of sounds at the word boundaries
- Phonemic orthography followed in writing

Different approches for text processing in Sanskrit and different tools for Sanskrit language are discussed here.

2.2.2.1 Machine Translation system

Machine translation (MT) systems uses to automatically translate text from one language to another. MT systems use machine-learning technologies or rule-based method to translate large amounts of text from and to any of their supported languages.

MT automatically translate text from one language to another

1. Anusaaraka

Anusaaraka is the MT system developed a consortium in different institutions. English to Hindi MT system and Sanskrit to Hindi MT system are pairs on Anusaarka.

English to Hindi MT system handles by IIIT Hyderabad and Sanskrit to Hindi MT system made by University of Hyderabad, Department of Sanskrit Studies. This helps the users to have access to the translated text in the source language(s) i.e. English or Sanskrit to Hindi.

2. Sampark: Indian Language MT System

Indian Language to Indian Language Machine Translation System funded by TDIL is an automated Translation among Indian Languages

Sampark: Indian Language MT System

Sampark is a multipart machine translation system developed with the combined efforts of 11 institutions in India under the umbrella of consortium project " Indian language to India Language Machine translation" (ILMT) funded by TDIL program of Ministry of Electronics and Information Technology (MeitY), Govt. of India.

Sampark uses Computational Paninian Grammar (CPG) approach for analyzing language and combines it with machine learning. Thus it uses both traditional rules-based and dictionary-based algorithms with statistical machine learning. Anusaaraka is a part of this project

3. Anuvaad

Anuvad is an AI based open source document translation platform

Anuvaad is an AI based open source Document Translation Platform to translate documents in Indic languages at scale. Anuvaad provides easy-to-edit capabilities on top the plug & play NMT models. Separate instances of Anuvaad are deployed to Diksha (NCERT), Supreme Court of India (SUVAS) and Supreme Court of Bangladesh (Amar Vasha).

4. Anuvādini

Anuvādini – Ministry of Education and All India Council for Tecnical Education

The ANUVADINI: Voice & Document AI Translation Tools consisting of a multitude of features and functionalities desires to close this gap arising due to language barriers. The tool has support for 22 regional Indian & foreign languages helping break language barriers & unifying India and the World under the principles of Ek Bharat Shrestha Bharat and One Earth, One Family, One Future!

Anuvadini -Voice and Document AI translation tools

- Text & Document Translation
  - Online document translation tool - Including Online Editor
  - Ananta – Infinite deep search generative AI engine with different file format like video, audio, pdf, pptx etc. image to generate Summary, Ask a question, Auto generate Question & Answer in Indian Languages
  - Chutki - Real-Time Doc Translation Service
  - Voice Based Multilingual Form - A web tool to create form to any Indian languages
  - Video Translation - A tool to make video translation to any Bharatiya languages
  - Virtual Keyboard - A web tool to type in any Indian languages
  - Dictation Tool - This tool is useful for making note taking easier

- Voice
  - Speech to Speech
  - Translation Services
  - Voice to Text, Text to Voice
  - Text to Text, Image to Text Translation
  - and many more like 3D Audio, Auto Panner, Bass Booster etc.
- File
  - PDF - Merge PDF, Split PDF, Compress PDF and more
- Education
  - Core Engineering Generative AI - Student friendly solution that seeks to disseminate information on core-Engineering streams in multiple Bharatiya Bhasha languages...
- Image
  - Image 23 - Translate photos into 30+ Languages In Seconds, whether they are posters, brochures, screenshots, advertisements etc.

  The link is - https://anuvadini.aicte-india.org/

Summarised Overview

In this chapter we came across different e-resources and online resources for Sanskrit in category basis. As e-resources described different e-journals, e-books, online databases, online magazines, online newspapers, audio and video collections. Online resources are divided in to two categories as language processing tools and speech and text processing tools. In this category we saw different approaches of MT System. Text or speech translation in different format. Usage of these technologies will help the Sanskrit language processing easy and time saving.

Assignments

1. List different e-resources and indicate how it help to Sanskrit.
2. List the Sanskrit journals and and magazines with its relevance.
3. Describe the facilities provided by Anuvadini

Suggested Readings

1. R. Raman Nair, L. Sulochana Devi, Sanskrit Informatics: Informatics for Sanskrit Studies and Research, Centre for Informatics and research and Development, 2011
2. https://sanskrit.inria.fr/Symposium/DOC/Amba_intro.pdf
3. https://egyankosh.ac.in/handle/123456789/85508
4. https://egyankosh.ac.in/handle/123456789/14110

References

1. https://sanskrit.uohyd.ac.in/scl/#
2. https://archive.org/
3. https://anuvadini.aicte-india.org/
4. https://sambhasha.ksu.ac.in/
5. https://sanskrit.inria.fr/

Space for Learner Engagement for Objective Questions

Learners are encouraged to develop objective questions based on the content in the paragraph as a sign of their comprehension of the content. The Learners may reflect on the recap bullets and relate their understanding with the narrative in order to frame objective questions from the given text. The University expects that 1 - 2 questions are developed for each paragraph. The space given below can be used for listing the questions.

Sanskrit Software

Learning Outcomes

- Introducing Sanskrit Software in various uses
- Supporting the Sanskrit learning application
- Simplifing operations and increasing the productivity

Background

In unit 1 and 2 we have gone through the online options for Sanskrit language processing. These all are working through internet connection only. Proceeding with an offline facility we need to use software only. Some of the language processing facilities need to be done through software only. This unit helps you to know about these kinds of facilities and the application software according to each specific goal.

Keywords

System software, Application software, Computer Application, Mobile Application

Discussion

As we covered so many e-resources for Sanskrit, online resources for Sanskrit, Sanskrit literary databases, Sanskrit online tools, etc. in this unit we are discussing the offline software which can be downloaded and used in different machines like laptops, desktops, tablets, and Mobile phones.

3.1 Software

Software is the bridge between the user and the computer's hardware

Software is used to increase productivity, and eliminate human errors. It also helps users complete specific tasks in a range of industries, including: Education, Business, Entertainment.

Software is the bridge between the user and the computer's hardware. It instructs the processor on what tasks to execute.

There are two basic categories of software:

System software: Provide access to the computer hardware and make system resources available. They are built-in softwares.

Application software: Designed to meet a specific goal.

The primary goal of application software is to simplify an operation and help users get their tasks done effortlessly.

Some benefits of software include:

- Managing the operations
- Increasing productivity, efficiency, and effectiveness
- Eliminating human errors
- Helping users complete specific tasks
- Simplifying operations
- Flexibility
- Scalability
- Better security

System software manages computer hardware

System software manages computer hardware and provides the basic services but the applications software helps to perform specific tasks.

Application software can divide into three:

1. General Application Software - for handling wide range of fundamental task

2. Business Application Software – to fulfil multiple functions and operations of businesses.

3. Customized Application Software – for specific business-based operations in the competitive world.

In this unit we discuss the general application software which is designed to possess special tasks related to Sanskrit language processing.

3.1.1 Sanskrit Software

Sanskrit Softwares- Computer and Mobile application

As the working platforms the Sanskrit softwares can categorize into two viz. Computer application and Mobile application.

3.1.1.1 Computer Applications

Computer application is a broad term that refers to any program, software, or set of computer programs designed to perform specific tasks for end-users. These applications are created to address various needs, ranging from productivity and entertainment to specialized tasks in different industries. Here we are looking at the software which is available for Sanskrit processes.

1. Input Method Editor

An Input Method Editor (IME) is a software component that enables a user to input text in a language that can't be represented easily on a standard QWERTY keyboard. This is typically due to the number of characters in the user's written language. Input method editors convert what you type into your desired script in real time. IME software is available for Sanskrit.

Baraha – An Input Method Editor for Indian Languages

Baraha - Baraha is an IME for Indian languages. Baraha was developed in 1998 by Sheshadrivasu Chandrasekharan with an intention to provide a free, user-friendly software to enable any non-computer-savvy person to use Kannada language on computer. Other Indian languages such as Hindi, Sanskrit, Tamil, Telugu, Malayalam, Gujarati, etc. were added in the later years. Baraha breaks the script barrier between different Indian languages. Indian scripts are mostly derived from the Brahmi script, which is also the underlying concept in Baraha. Baraha uses a common code to represent all the Indian languages which makes it possible to convert text from one script to another. Baraha breaks the keyboard barrier for Indian languages by providing a phonetic keyboard in which any Indian language word may be typed using the standard English keyboard.

Google IME - Google IME, also known as Google Input Tools, is a set of input method editors by Google for 22 languages. Sanskrit is one of it. Google IME is a virtual keyboard that allows users to type in their local language text directly in any application without the hassle of copying and pasting. It is available as a Chrome extension.

BhashaIME is an input method editor for Windows

BhashaIME - BhashaIME is an Input Method Editor for Windows. It helps in creation of Unicode text content in several Indian and Roman scripts. The supported scripts are Devanagari, Grantha, Kannada, Malayalam, Telugu, Tamil, IAST Roman and ISO15919 Roman. Vedic extensions are supported.

Lipika IME, ePandit IME etc. are the other IMEs for Devanagari script.

2. Intelligent Script Manager

Intelligent Script Manager (ISM) is from C-DAC GIST. ISM products are from GIST Research Labs. Using ISM is the easiest way to get started with Indian languages on your personal computer. This software consists of aesthetic Indian language fonts and tools that you require to start working with Indian Languages on computers.

ISM is developed to work with all Official Indian Languages

ISM is developed to work with all Official Indian languages. With the help of this tool users can type very easily official Indian languages. It supports UNICODE and Enhanced INSCRIPT keyboard layout. One can select any languages to start typing in that particular language on any editor on all available windows platforms (32 bit and 64 bit both.

ISM enables creation of Indian language content using Windows based applications. It can be used for typing in web-browsers to support applications such as e-mailing, chatting, searching, etc.

3.1.1.2 Mobile Application

Mobile apps are software applications that run on mobile devices like smartphones, tablets, and watches. They provide specific functions and services, and are designed to run on specific mobile operating systems. We can download it from the app providers like the play store in Android version, App

store in Apple mobile phones, Windows phone store for windows windows version.

Sanskrit language study, translation, dictionaries, competitive exam helpers, Message and poster makers, Sanskrit keyboard, Sanskrit library, textual study etc. are the different categories of apps developed for Sanskrit.

Sanskrit language study, translation etc. are the different categories of mobile apps

In Sanskrit language study category grammar helpers especially Sandhi, Dhaaturoopas, Sanskrit learners, School class learners etc. are the sub categories.

In the translation category we can find apps for translating Sanskrit from all Indian languages apps like Sanskrit to Hindi and vise-versa etc.

Dictionaries category apps are having so many apps particularly titled with Sanskrit dictionaries like Vacaspatya, Amarakosa etc. and language pairs Hindi – Sanskrit, Sanskrit – English etc. also

Competitive exam helpers category deals with UGC exams, UPSC exams, School class apps are also available.

Message and poster makers category apps allow us to make messages and posters in Sanskrit.

Textual study category apps deal with Sanskrit texts like Vaiyyakarana Siddhanta kaumudi, Ashtadhyayi etc.

Summarised Overview

In this unit we saw the two types of softwares System software and Application software. The application software can dvide into three viz. General Application Software, Business Application Software and Customized Application Software. The softwares which are using for Sanskrit comes under General Application software. These can divide again in to two viz. Computer Application and Mobile Application. Under computer application we saw some IMEs and ISM. Mobile applications available for Sanskrit are so many and according to the need of the user it differes. So here we discussed the types of mobile applications.

Assignments

1. List the computer applications for typing Sanskrit
2. List the mobile applications for Sanskrit dictionaries
3. Write the divisions of Software.

Suggested Readings

1. https://arshadrishti.org/sanskrit-tools/
2. https://egyankosh.ac.in/bitstream/123456789/10283/1/Unit-1.pdf

References

1. https://baraha.com/
2. https://www.google.com/intl/sa/inputtools/try/
3. https://sanskritdocuments.org/processing_tools/
4. https://egyankosh.ac.in/

Space for Learner Engagement for Objective Questions

Learners are encouraged to develop objective questions based on the content in the paragraph as a sign of their comprehension of the content. The Learners may reflect on the recap bullets and relate their understanding with the narrative in order to frame objective questions from the given text. The University expects that 1 - 2 questions are developed for each paragraph. The space given below can be used for listing the questions.

Sanskrit Fonts, Unicode Fonts, DTP in Sanskrit

Learning Outcomes

- Understanding different types of encodings for the Devanagari script
- Understanding the importance of Unicode fonts.
- Understanding different DTP operations

Background

The revolutionary use of computers in regular life made an important role of Desktop Publishing (DTP). Sanskrit texts were the first to describe many important scientific and mathematical concepts. Sanskrit fonts are important for preserving and promoting the ancient language. Making a brochure in Sanskrit to making a book needs different font and DTP setting in computers. Knowing about the use of different fonts and DTP setting are become must in the modern world.

Keywords

Font, ASCII, ISCII, Unicode, DTP in Sanskrit

Discussion

Sanskrit fonts

Letters of a script, with specific size and style can be called as font which is used to print or display the text character. Fonts can have different sizes, colors and designes.They can include:

• Symbols

- Numbers
- Alphabets (both lowercase and uppercase)
- Punctuation marks

Sanskrit fonts are used to type the text in Sanskrit

Sanskrit fonts are used to type the text in Sanskrit. Devanagari is the most widely used script to write Sanskrit. Here we are discussing about the devanagari script and it's character encodings

4.1 Character Encoding Systems

Encoding the characters helps to design interfaces that both humans and computers understand. It's used to represent text that's transmitted between different systems or devices. Different character encodings are discussed here.

4.1.1 ASCII

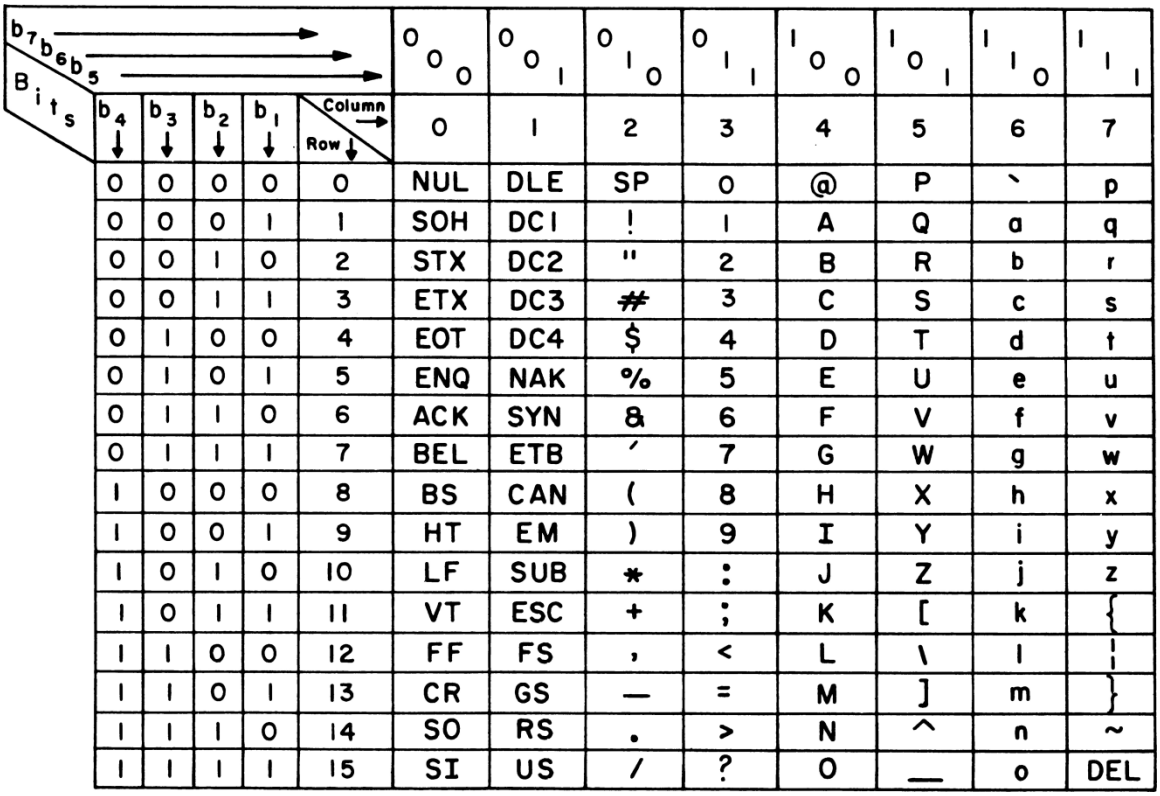
ASCII stands for American Standard Code for Information Interchange. It's a character encoding standard for electronic communication. The American Standards Association (ASA) introduced ASCII in 1963.

ASCII-American Standard Code for Information Interchange

ASCII assigns numeric values to letters, numbers, punctuation marks, and other characters.

For example, the ASCII code for uppercase A is 65, and the code for uppercase B is 66.

ASCII is an 8-bit code. It uses eight bits to represent a letter or punctuation mark. Eight bits are called a byte. ASCII code chart is given below.



Bits b_4	b_3	b_2	b_1	Row / Column	000	001	010	011	100	101	110	111
					0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	¦
1	1	0	1	13	CR	GS	—	=	M	]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
1	1	1	1	15	SI	US	/	?	O	_	o	DEL

When ASCII uses 1-byte character encoding it can only handle 256 characters that means after 128 characters it can accommodate 128 more. So it can hardly handle one or at most two languages and that shows its limitation to handle all world languages. So definitely at this juncture ASCII fails, especially to record Indian Languages.

4.1.2 ISCII

ISCII-Indian Standard Code for Information Interchange

ISCII (Indian Standard Code for Information Interchange) is an extended ASCII. ISCII also known as Indian Script Code for Information Interchange. The code represents various writing systems of India. It covers 10 Indic languages derived out of Bramhi script. ISCII does not encodes the scripts which are based on Persian script and the same encoded in the PASCII

ISCII caters to the following 10 Indian scripts – Devanagari, Gujarati, Punjabi, Bengali, Assamese, Oriya, Telugu, Tamil, Malayalam, Kannada. The ISCII code table is a superset of all the characters required for the above mentioned scripts. First version was released in 1983 and adopted by the Bureau of Indian Standards (BIS) in 1991 after revisions in 1986 and 1988.

ISCII follows a single encoding schema, which allows for easy transliteration between various writing systems.

For example, the ISCII code 0xB3 0xDB represents [ki]. It appears as:

കി in Malayalam

कि in Devanagari

கி in Tamil

ISCII is no doubt a solution for the Indian scenario but when it comes to global representation of characters both ASCII and ISCII fail to represent the characters at a stretch, as both use single byte code. So definitely at this juncture ISCII also fails and Unicode seems to be a promising solution which covers all the world languages scripts.

4.1.3 Unicode

Unicode is Universal code for representing world's scripts in computers

Unicode is Universal code for representing world's scripts in computers. A group of software companies like It is developed by Unicode consortium which is a group of software organizations like IBM, Microsoft, Xerox, Oracle, etc. They came out with a 16-bit code and called it Unicode, as it promises to cover all the world's scripts. The promising feature of Unicode is that it can represent 65,536 characters. The first version of Unicode came in 1991.

Unicode is efficient in character coding with no redundancies

Unicode can be represented using 7-bit, 8-bit, and 16-bit format standards. UTF-7 (Universal Character Set Transformation Format-7) is a format that breaks the Unicode code point into 7-bit values; these can be transferred through email (which uses the 7-bit ASCII encoding) and on the Internet. This format presents difficulties because some values are ambiguous, but it is still usable. The 8-bit UTF-8 (Universal Character Set Transformation Format-8) format breaks Unicode values into 8-bit sequences, which work well on the Web. UCS-2 (Universal Character Set -2) stores each character as a 16-bit value and each value corresponds directly to its value (code point) in the Unicode standard.

Unicode is efficient in character coding with no redundancies. It is a 16 bit code (allows for 65,536 characters) and has the following features:

- Plain text
- Dynamic composition
- Logical Order
- Unification
- Compatibility characters
- Equivalent sequence
- Convertibility.

Here are some fonts that are recommended for typing in Sanskrit:

- Arial Unicode MS
- CDAC-GISTSurekh
- CDAC-GISTYogesh
- JanaSanskrit
- Kalimati
- Mangal
- Sanskrit 2003
- SHREE-DV0276-OT

Arial Unicode MS

You can install Sanskrit fonts in windows operating system by:

1. Opening the Control Panel
2. Clicking on Region and Languages
3. Selecting Change Keyboards
4. Clicking on Add and scrolling down to Hindi (India) > Keyboards

We can install Sanskrit fonts in windows operating system

You can type in Sanskrit on a Linux machine by:

1. Open Settings

2. Clicking on Region and Languages

3. Add ‘Input Sources’ -> others -> Sanskrit

You can type in Sanskrit on a Mac by:

1. Clicking the Input Menu icon in the upper-right corner of the menu bar
2. Choosing "Devanāgarī-QWERTY" from the menu
3. Starting typing

You can type diacritical marks in Sanskrit by:

1. Setting up an alternative keyboard layout
2. Holding a modifier key to type letters with diacritical marks

For example, ***alt + a = ā***.

4.2 Sanskrit Fonts

Sanskrit fonts are used for typing Sanskrit texts using Devanagari script

Sanskrit Fonts are used for typing Sanskrit texts using Devanagari script. Unicode fonts are used to type Sanskrit now. Sanskrit Unicode fonts are designed to enable the representation of Sanskrit characters in a standardized and universally accepted way across different computer systems and platforms.

Sanskrit Unicode fonts are essential because:

- It provides a standardized encoding system for characters from all the world's major writing systems, including Sanskrit. This standardization ensures that text encoded in Unicode will be displayed consistently across various devices, applications, and platforms.
- It supports a vast range of languages, including Sanskrit, and it allows the mixing of characters from different scripts within the same document. This is particularly useful in multilingual contexts where Sanskrit might be used alongside other languages.
- Unicode fonts are continually updated to include new characters and scripts. Adopting Unicode for Sanskrit ensures that the language can benefit from ongoing developments in the standard, making it more adaptable to future technological advancements.

Unicode facilitates consistent rendering of Sanskrit texts

- Unicode facilitates consistent rendering of Sanskrit text across various operating systems (Windows, macOS, Linux) and devices (computers, tablets, smartphones). This cross-platform compatibility is essential for widespread usage.
- Unicode-encoded Sanskrit text allows for efficient search and information retrieval. Search engines and databases can process and index Sanskrit content consistently, making it easier for users to find relevant information.

The adoption of Sanskrit Unicode fonts promotes the digital representation,

The adoption of Sanskrit Unicode fonts promotes the digital representation, sharing, and preservation of Sanskrit content in a way that is both standardized and accessible on a global scale. It plays a crucial role in the integration of Sanskrit into the digital landscape.

4.2.1 Sanskrit Font Repositories

We can download and use Sanskrit Unicode fonts with its licensing terms and conditions from various font providers. Two of them are listed below. From these two websites we can download various kinds of Sanskrit Unicode fonts. There are so many websites which allow us to download the Sanskrit Fonts.

4.2.1.1 Google Fonts

Google fonts – a computer font and web font service by Google

Google Fonts (formerly known as Google Web Fonts) is a computer font and web font service owned by Google. This includes free and open source font families, an interactive web directory for browsing the library, and APIs for using the fonts via CSS and Android.

We can download the font family for Sanskrit free of cost and can install it in the system with very minimal steps.

Link :https://fonts.google.com/

4.2.1.2 Devanagarifonts.net

Devanagarifonts.net is a great resource for downloading Sanskrit fonts. We can download more than 200 various Devanagari fonts form here. The link is :https://devanagarifonts.net/

4.3 DTP in Sanskrit

DTP in Sanskrit refers to desktop publishing in Sanskrit. DTP creates a document on a computer using software.

DTP services use digital software to create and design documents. Partnering with a multilingual desktop publishing services provider can eliminate the need to invest in DTP equipment.

DTP in Sanskrit refers to desktop publishing in Sanskrit

DTP services can include:

- Sanskrit typing
- Document format conversion
- Content and copywriting
- Voiceover and subtitling in Sanskrit
- Editing and proofreading in Sanskrit
- Localization of software and documentation
- Brochures, flyers, visiting cards, and coupons

4.3.1 DTP Softwares

DTP software is used create and design documents for print and digital distribution. These tools provide advanced layout and formatting options, making them suitable for designing brochures, magazines, books, and other visually rich documents. All these documents are creating in Sanskrit language also.

Here are some popular DTP software options:

4.3.1.1 Adobe InDesign

Adobe Indesign

Adobe InDesign is one of the most widely used DTP software. It offers robust tools for layout design, typography, and image manipulation. It is part of the Adobe Creative Cloud suite.

4.3.1.2 QuarkXPress

QuarkXPress is another long-standing DTP software. It provides professional-level features for page layout and design. It's often used in publishing and graphic design industries.

4.3.1.3 Scribus

Scribus is an open-source desktop publishing software that is free to use. It's a powerful tool for creating layouts, and it supports various file formats.

4.3.1.4 Microsoft Publisher

Microsoft Publisher is a DTP software included in the Microsoft Office suite. It's user-friendly and suitable for smaller-scale desktop publishing projects.

4.3.1.5 CorelDRAW

While CorelDRAW is primarily known as a vector graphics editor, it also offers features for desktop publishing. It's suitable for creating a wide range of designs, including brochures and posters.

4.3.1.6 LaTeX

LaTeX is a free, open-source document preparation system that's used for typesetting technical and scientific documents. It was created in 1985 by Leslie Lamport, an American computer scientist. For some international publication of Sanskrit document demands this format.

LaTeX available for Mac OS, Windows and Linux

LaTeX is available for Mac OS, Windows, and Linux. It's distributed through CTAN servers or as part of TeX distributions from the TeX User Group (TUG) or third parties. There are so many LaTeX editors available. They are Dreamweaver, TeXmaker, Overleaf, LaTeX Base, TeXStudio, Authorea, Emacs. TeXworks is a LaTeX editor that uses a GUI. It has a built-in PDF viewer and auto-synchronization.

krit documents, Xelatex – the version of LaTex – is used normally.

Affinity Publisher, Scribus, PagePlus, etc are the other DTP softwares which are popular. Lucidpress, Canva etc. are some web-based design tools that simplify the creation of various visual content, including social media graphics, presentations, and posters.

When choosing DTP software, consider your specific needs, the complexity of your projects, and your budget. Each of these

tools has its strengths and may cater to different user preferences and project requirements.

Summarised Overview

In this unit, we saw the definition of font, different efforts of encodings to represent Indian Languages. Importance of Unicode fonts and Sanskrit typing. The websites from where the Unicode fonts for Sanskrit can be downloaded are also discussed in this unit. Importance of Desktop Publishing nowadays and different kinds of DTP softwares are also discussed here. Typing in Sanskrit is an essential to any Sanskrit student nowadays. Knowing about the fonts and typesetting for different kinds of documents with Sanskrit texts is also a must now. This unit helps get an awareness on this regard.

Assignments

1. List ten Sanskrit Unicode fonts.
2. Make a Sanskrit-typed document for at least five pages.
3. Make a poster on Sanskrit using any of the software.
4. Make a flier for the inauguration of a shop in Sanskrit using any of the DTP software.
5. Write the importance of using Sanskrit Unicode fonts for typing.

Suggested Readings

1. R. Raman Nair, L. Sulochana Devi, Sanskrit Informatics: Informatics for Sanskrit Studies and Research, Center for Informatics and research and Development, 2011
2. D Bott Woody Leonhard, Microsoft Office 2007, published by Pearson Education.
3. Comdex Desktop Publishing Course Kit by Vikas Gupta, Dreamtech Press, New Delhi
4. https://www.unicode.org/main.html

References

1. Sanjay Saxena, MS - Office 2000 for everyone, Vikas Publishing House.
2. Russell A.Stulz, MS – Office, BPB Publications.
3. M.C. Sharma, Desk Top Publishing on PC, DPB Publications
4. https://www.gammon.com.au/Unicode/

Space for Learner Engagement for Objective Questions

Learners are encouraged to develop objective questions based on the content in the paragraph as a sign of their comprehension of the content. The Learners may reflect on the recap bullets and relate their understanding with the narrative in order to frame objective questions from the given text. The University expects that 1 - 2 questions are developed for each paragraph. The space given below can be used for listing the questions.

Natural Language Processing

BLOCK - 2

Natural Language Processing

Learning Outcomes

- Understanding the importance of Natural Language Processing (NLP)
- Understanding the basic text processing tasks
- Understanding the possible NLP tasks for Sanskrit
- Understanding the need of NLP for Sanskrit literature.

Background

In block 1 we femiliarized the different online resources, software, tools to study language, etc. In block 2 we are going to experience the technical ways of processing of texts. In unit 1 we are going to learn about NLP, and it's importance in text processing especially for Sanskrit. Also we will learn the NLP tasks which we should do for the language processing area for Sanskrit language processing.

Keywords

NLP, Text Understanding, Information Retrieval, Automated Translation, Voice Assistant, Sentiment Analysis, Text Summarization, Speech recognition

Discussion

Natural Language Processing (NLP) involves the interaction between computers and human language. While most NLP research and applications focus on widely spoken languages

NLP involves the interaction between computers and human language

like English, there has been increasing interest in the application of NLP to Sanskrit and other classical languages.

NLP refers to the branch of computer science and linguistics or Computational linguistics. More specifically, NLP is the branch of artificial intelligence or AI. NLP enables machines to understand human language. NLP combines computer science, linguistics, and machine learning to study how computers and humans communicate in natural language.

1.1 Importance of NLP

NLP used in a wide variety of everyday products and services

NLP has existed for more than 50 years and has roots in the field of linguistics. It's used in a wide variety of everyday products and services. NLP's goal is to build systems that can automatically perform tasks like: Translation, Spell check, Topic classification, Summarization, Ticket classification, etc.

While NLP is focusing on the interaction between computers and human languages, there are several compelling reasons why NLP is important:

1.1.1 Communication with Computers

NLP enables humans to communicate with computers in natural language. This is especially important for those who may not have technical skills but still want to interact with technology in a more intuitive way.

1.1.2 Text Understanding

NLP allows computers to understand and interpret human generated text.

NLP allows computers to understand and interpret human-generated text. This is crucial for extracting information, sentiment analysis, and making sense of unstructured data, which constitutes a significant portion of the information available today.

1.1.3 Information Retrieval

NLP is vital for developing search engines and information retrieval systems that can understand user queries and retrieve relevant results. This improves the efficiency and accuracy of information retrieval.

NLP plays a key role in machine translation

1.1.4 Automated Translation

NLP plays a key role in machine translation, making it possible to automatically translate text from one language to another.

This is valuable for breaking down language barriers and facilitating global communication.

1.1.5 Voice Assistants

NLP is the technology behind voice assistants like Siri, Google Assistant, and Alexa. These systems understand spoken language, respond to queries, and perform tasks based on natural language commands.

1.1.6 Sentiment Analysis

NLP allows for sentiment analysis, which involves determining the emotional tone expressed in a piece of text. This is used in business to gauge customer satisfaction, public opinion analysis, and social media monitoring.

NLP is crucial for creating chatbots and virtual assistants

1.1.7 Chatbots and Virtual Assistants

NLP is crucial for creating chatbots and virtual assistants that can engage in conversations with users, answer questions, and provide assistance. This has applications in customer service, support, and various industries.

1.1.8 Text Summarization

NLP can be used for automatic text summarization, helping to distill large amounts of information into concise summaries. This is valuable for quickly understanding the main points of lengthy documents.

1.1.9 Autocorrect and Grammar Checking

NLP is used in autocorrect systems and grammar-checking tools to improve the accuracy of written text. This is essential for enhancing the quality of written communication.

1.1.10 Language Modelling and Generation

NLP models, such as language models and transformers, are capable of generating human-like text. This is utilized in content generation, creative writing assistance, and even in the development of conversational agents.

NLP can assist healthcare, Legal and Compliance

Also NLP can assist healthcare, Legal and Compliance, etc. In summary, NLP is essential for bridging the gap between human communication and computer understanding. It opens up new possibilities for human-computer interaction, information

processing, and automation in various fields, contributing to the development of more intelligent and user-friendly technologies.

1.2 NLP tasks and Sanskrit

Human language is filled with ambiguities that make it incredibly difficult to write software that accurately determines the intended meaning of text or voice data. Homonyms, idioms, metaphors, grammar and usage exceptions, variations in sentence structure—these just a few of the irregularities of human language that take humans years to learn, but that programmers must teach natural language-driven applications to recognize and understand accurately from the start, if those applications are going to be useful.

Natural language generation is sometimes described as the opposite of speech recognition or speech-to-text; it's the task of putting structured information into human language. Several NLP tasks break down human text and voice data in ways that help the computer make sense of what it's ingesting.

For Sanskrit language, these kind of works are needed to be adapt. In block 1 we saw the tools which are developed on this path. Here are some aspects of Natural Language Processing in the context of Sanskrit:

1.2.1 Speech recognition

Speech recognition, also called speech-to-text, is the task of reliably converting voice data into text data. Speech recognition is required for any application that follows voice commands or answers spoken questions. What makes speech recognition especially challenging is the way people talk—quickly, slurring words together, with varying emphasis and intonation, in different accents, and often using incorrect grammar.

Speech recognition

Speech recognition for Sanskrit involves converting spoken Sanskrit into written text. This can be applied to digitize oral recitations of Sanskrit texts or facilitate spoken interactions in Sanskrit.

1.2.2 Part-of-Speech Tagging

Assigning grammatical categories (parts of speech) to words is a crucial step in NLP. Part-of-speech tagging in Sanskrit involves determining whether a word is a noun or verb or

adjective, etc., which can be challenging due to the inflectional nature of Sanskrit.

Part of speech Tagging

Part of speech tagging, also called grammatical tagging, is the process of determining the part of speech of a particular word or piece of text based on its use and context. Part of speech identifies 'भवति' as a verb in 'अद्यकक्षाभवति।,' and as a noun in 'भवति, रमावा?'

1.2.3 Word sense disambiguation

Word sense disambiguation

Word sense disambiguation is the selection of the meaning of a word with multiple meanings through a process of semantic analysis that determine the word that makes the most sense in the given context. For example, word sense disambiguation helps distinguish the meaning of the indeclinable 'अपि' in 'रामःअपिश्यामःगच्छतः।' (and) vs. 'अपिभोजनंकरोति।' (Question word).

1.2.4 Tokenization and Morphological Analysis

Sanskrit is a highly inflected language with a complex system of word forms. Tokenization involves breaking down a sentence into individual words or tokens. Morphological analysis in Sanskrit NLP involves understanding the grammatical structure of words, including root forms, prefixes, and suffixes.

1.2.5 Named Entity Recognition (NER)

Named Entity Recognition

Identifying named entities such as names of people, places, and organizations is an essential NLP task. In Sanskrit, this involves recognizing names of characters, locations, and other entities in ancient texts. Named entity recognition, or NER, identifies words or phrases as useful entities. NER identifies 'विदर्भा' as a location or 'शङ्करः' as a man's name.

1.2.6 Co-reference resolution

Co-reference resolution is the task of identifying if and when two words refer to the same entity. The most common example is determining the person or object to which a certain pronoun refers (e.g., 'सा' = 'सीता'), but it can also involve identifying a metaphor or an idiom in the text (e.g., an instance in which 'देवानांप्रियः' isn't not in the meaning the who is favorite to diety but an ignorant or foolish person).

1.2.7 Syntax Parsing

Parsing the syntactic structure of Sanskrit sentences involves understanding the relationships between words, phrases, and clauses. Sanskrit sentences often have a free word order, making syntax parsing more challenging.

1.2.8 Sentiment Analysis

Sentiment analysis aims to determine the sentiment expressed in a piece of text. In the context of Sanskrit, this could be applied to classical texts to understand the emotional tone or sentiment conveyed in the writings.

Sentiment analysis aims to determine the sentiment expressed in a piece of text

Sentiment analysis attempts to extract subjective qualities - attitudes, emotions, sarcasm, confusion, suspicion - from text.

1.2.9 Machine Translation

Developing machine translation systems for Sanskrit involves translating Sanskrit text into other languages or vice versa. Translating Sanskrit, especially ancient texts, poses unique challenges due to archaic language forms and specialized vocabulary.

1.2.10 Text Summarization

Summarizing Sanskrit texts can be useful for extracting key information from lengthy manuscripts. Automatic text summarization involves identifying important sentences or passages to create concise summaries.

1.2.11 Language Modelling

Developing language models for Sanskrit involves creating statistical models that predict the likelihood of word sequences. These models are fundamental to many NLP tasks.

1.2.12 NLP for Sanskrit Teaching and Learning

NLP tools can be developed to aid in the teaching and learning of Sanskrit, providing interactive lessons, pronunciation guides, and automated assessments.

NLP tools can be developed to aid in the teaching and learning of Sanskrit

While NLP research in Sanskrit faces unique challenges due to the classical nature of the language, ongoing efforts are being made to develop NLP tools and resources specific to Sanskrit,

including annotated corpora, linguistic resources, and computational tools.

Summarised Overview

In lesson we saw what is NLP, and the importance of NLP in Sanskrit nowadas. Here we covered two major topics viz. Importance of NLP and the tasks in NLP. In the first portion we can see highly important ten reasons why NLP is important while Sanskrit as a natural language. In the second part the major tasks in NLP which are related to Sanskrit is discussed. Here almost twelve major fields which are more essential for Sanskrit is discussed. There are other fields also in NLP can concentrate for other disciplines of Sanskrit literature.

Assignments

1. Why NLP is needed today?
2. Write an essay, descibing the importance of NLP tasks in Sanskrit.
3. Entitle the NLP tasks and describe how it is carried out in Sanskrit.

Suggested Readings

1. Bharati, Akshar & Chaitanya, Vineet & Sangal, Rajeev (2002). Natural Language Processing: A Paninian Perspective.
2. https://www.ibm.com/topics/natural-language-processing
3. Kak, Subhash C.

References

1. Bharati, Akshar & Chaitanya, Vineet & Sangal, Rajeev & Gillon, Brendan. (2002). Natural Language Processing: A Paninian Perspective.
2. Basha, M. & Selvaraj, Vijayakumar & Jayashankari, J. & Alawadi, Ahmed & Durdona, Pulatova. (2023). Advancements in Natural Language Processing for Text Understanding. E3S Web of Conferences. 399. 10.1051/e3sconf/202339904031.
3. https://www.ibm.com/topics/natural-language-processing
4. Rick Briggs, Knowledge Representation in Sanskrit and Artificial Intelligence,DOI: http://dx.doi.org/10.1609/aimag.v6i1.466
5. Shashank Saxena and Raghav Agarwal, Sanskrit as a Programming Language and Natural Language Processing, Global Journal of Management and Business

Studies, ISSN 2248-9878Volume 3, Number 10 (2013), pp. 1135-1142

Space for Learner Engagement for Objective Questions

Learners are encouraged to develop objective questions based on the content in the paragraph as a sign of their comprehension of the content. The Learners may reflect on the recap bullets and relate their understanding with the narrative in order to frame objective questions from the given text. The University expects that 1 - 2 questions are developed for each paragraph. The space given below can be used for listing the questions.

Artificial Intelligence

Learning Outcomes

- Understanding the importance of Artificial Intelligence (AI)
- Understanding the application and uses of AI
- Understanding the application of AI in Sanskrit
- Understanding the need for AI for Sanskrit literature in modern era

Background

In the last chapter, we have gone through the features of NLP which is a branch of Artificial Intelligence (AI). Here we will look at what is AI and how Sanskrit can be useful in applying AI methods. Also in this unit, different tasks are carried out in the field of AI. Being AI is a larger sector than NLP the applications of AI with Sanskrit are discussed here with high priority. Finally, the open-source AI websites are also listed.

Keywords

Machine translation, Speech recognition, Text-to-Speech (TTS), Augmented Realms (AR)

Discussion

Artificial intelligence is the simulation of human intelligence by machines.

Artificial intelligence (AI) is the simulation of human intelligence by machines. It is a field of computer science that studies and develops intelligent machines. AI is also known as machine intelligence.

AI is not a single technology. It involves building machines that can: Learn, Reason, Act, Make decisions, and Carry out actions.

2.1 Applications and Uses of AI

ArtificialIntelligence (AI) has a wide range of applications across variousindustries, and its impact continues to grow. AI'sapplication can help in almost all the fields of human life. Forexample, In health care, it helps with Disease Diagnosis and DrugDiscovery. In Finance it helps with Fraud Detection and AlgorithmicTrading. In the field of Education, it helps to personalize learning and Automated Grading. In Customer Service, it can help as Chatbotsand Voice Assistants. Like this AI can help in any field of humanlife.

2.1.1 Artificial Intelligence in Language

In the language field, AI is used for different applications. Some applications of AI in languages include:

Artificial intelligence in Languages

- Expert systems
- Natural language processing
- Speech recognition
- Machine vision

Natural language processing (NLP) is a branch of artificial intelligence (AI) that enables computers to understand, generate, and manipulate human language.

Some applications of NLP include:

- Text generators: Compose coherent essays
- Chatbots: Fool people into thinking they're sentient
- Text-to-image programs: Produce photorealistic images of anything you can describe

2.2 AI and Sanskrit

In the quest to bridge ancient wisdom with cutting-edge technology, Artificial Intelligence finds its resonance in the realm of Sanskrit, unlocking new dimensions of understanding and preservation. Artificial Intelligence (AI) can be applied to

AI can be applied to Sanskrit

various domains, including natural language processing and understanding. Developing AI for Sanskrit, an ancient Indian language, involves addressing specific challenges related to the language's structure, script, and usage. Here are some ways AI can be applied to Sanskrit.

2.2.1 Natural Language Processing (NLP)

NLP techniques can be used to analyze and understand Sanskrit text. As we studied in detail in block 1, unit 1 NLP involves tasks such as tokenization, part-of-speech tagging, named entity recognition, and sentiment analysis. Adapting existing NLP models to handle Sanskrit's linguistic intricacies would be crucial.

Advancements in Natural Language Processing have empowered machines to unravel the intricacies of Sanskrit. From syntactic analysis to semantic comprehension, AI algorithms delve into the heart of Sanskrit texts, enabling a deeper understanding of classical literature and philosophical treatises.

2.2.1.1 Syntactic ability

Parsing the rich syntax of Sanskrit sentences, AI-driven tools dissect the grammatical structures, paving the way for a nuanced comprehension that goes beyond literal translations.

2.2.1.2 Semantic ability

Semantic ability

AI algorithms, finely tuned for Sanskrit semantics, disentangle layers of meaning embedded in philosophical texts, providing scholars with a deeper grasp of the profound concepts enshrined in Sanskrit literature.

Understanding the semantics of Sanskrit sentences is crucial for extracting meaningful information. AI algorithms can be employed to perform deep semantic analysis, enabling applications like question-answering systems and information retrieval.

2.2.2 Machine Translation

Machine Translation systems, trained on vast corpora of Sanskrit texts, facilitate the seamless exchange of ideas between ancient manuscripts and contemporary languages, fostering a global dialogue on India's intellectual heritage.

Machine translation for Sanskrit can help in translating Sanskrit texts

Building machine translation systems for Sanskrit can help in translating Sanskrit texts into other languages and vice versa. This can be particularly useful for scholars, researchers, and enthusiasts who may not be proficient in Sanskrit but are interested in its literature and philosophy.

2.2.3 Speech Recognition

The spoken word echoes through time with AI-powered speech recognition tailored for Sanskrit. Language learners, scholars, and enthusiasts now have a tool that not only understands spoken Sanskrit but aids in pronunciation and fluency, bringing the language to life in ways never imagined. Developing speech recognition systems for spoken Sanskrit can aid in language learning, enable voice-controlled applications, and enhance accessibility for those who prefer or require spoken interaction.

2.2.3.1 Conversational Proficiency

Beyond deciphering words, AI engages in meaningful conversations, adapting to the nuances of Sanskrit discourse and contributing to the preservation of oral traditions.

2.2.4 Text-to-Speech (TTS)

TTS systems can convert written Sanskrit text into spoken words

Sanskrit literature, once confined to written scripts, now resonates through digital voices. TTS systems, finely tuned for Sanskrit phonetics, transform written texts into audible expressions, fostering accessibility and a renewed appreciation for the rhythmic beauty of Sanskrit poetry.

Text-to-speech (TTS) systems can convert written Sanskrit text into spoken words. This can be beneficial for creating audio content, aiding in language learning, and making Sanskrit literature more accessible.

2.2.5 Educational Tools

AI and Educational Tools

AI takes the role of a knowledgeable guide in the pursuit of Sanskrit education. Intelligent tutoring systems, interactive exercises, and gamified applications make learning Sanskrit a dynamic and engaging experience, transcending the barriers of time and space.

AI-powered educational tools can be developed to assist in learning Sanskrit. This could include intelligent tutors,

interactive exercises, and gamified applications to make the learning process more engaging and effective.

2.2.6 Preservation of Texts

AI can play a role in the preservation and digitization of ancient Sanskrit texts

AI can play a role in the preservation and digitization of ancient Sanskrit texts. Optical Character Recognition (OCR) systems adapted for Sanskrit scripts can convert physical manuscripts into digital formats, making them more accessible for research and study and safeguarding them for future generations.

2.2.7 Cultural Heritage Applications

AI can be used to develop applications that promote and preserve Sanskrit literature, culture, and heritage. This could involve virtual museums, interactive exhibits, and augmented reality experiences.

2.2.7.1 Augmented Realms (AR)

Through AR-enhanced experiences, users embark on virtual journeys through historical settings, interacting with the luminaries of Sanskrit literature and experiencing the cultural tapestry firsthand.

As AI and Sanskrit weave a harmonious tapestry, the synergy between tradition and technology breathes new life into the linguistic and cultural heritage of India. The journey has just begun, and the collaboration between machines and the sacred language unfolds, promising a future where the echoes of the past resonate with the algorithms of the future.

2.3 AI and Working Process

Developing AI for Sanskrit involves linguistic features and challenges

It's important to note that developing AI for Sanskrit involves addressing the unique linguistic features and challenges of the language. Collaborations between linguists, Sanskrit scholars, and AI researchers would be beneficial in ensuring the accuracy and cultural sensitivity of these applications. Additionally, the availability of annotated datasets for Sanskrit and active participation from the Sanskrit community would contribute to the success of AI applications in this domain. How to start learning with AI, the companies that are working for AI research, and some AI chatbots that are available for free are listed here.

- Some companies that work with AI include IBM, Google, Amazon, People.ai, AlphaSense, NVIDIA, DataRobot and H2O.ai.
- To learn AI, you can:
  - Start with a foundation in computer science
  - Learn a programming language, like Python
  - Learn basic algorithms
  - Learn machine learning and data science principles
  - Apply your knowledge through AI projects

AI based online tools

- **AI-based online tools**
  - The AI-based tool ChatGPT by open AI is familiar to us. Google has also developed AI-enabled search nowadays. Here are some other AI tools -
    - https://copilot.microsoft.com/
    - https://www.copy.ai/
    - https://www.wordtune.com/
    - https://wordai.com/ - for rewriting
    - https://pdf.ai/ - allows the chat according to the PDF uploaded
    - https://aizzy.ai/

Summarised Overview

In this unit it is explored that what is AI?, the applications and uses of AI, AI for Sanskrit and some AI tools. As NLP is a branch of AI the tasks which are covered in unit 1 of block 2 comes under AI applications for Sanskrit. So here we can see some other applications which are really usefull and need to develop for Sanskrit Language. Also it is explained the advantage of doing these AI tasks for Sanskrit under every topic. Overall the possibilites of AI to Sanskrit is described here.

Assignments

1. Describe the applications of AI.
2. Explain the tasks for Sanskrit with AI
3. List five AI tools and their usage
4. How AI help Sanskrit? Describe.
5. Summarize a PDF text using https://pdf.ai/

Suggested Readings

1. Rick Briggs, Knowledge Representation in Sanskrit and Artificial Intelligence,DOI: http://dx.doi.org/10.1609/aimag.v6i1.466
2. https://www.ibm.com/topics/artificial-intelligence

References

1. Rick Briggs, Knowledge Representation in Sanskrit and Artificial Intelligence,DOI: http://dx.doi.org/10.1609/aimag.v6i1.466
2. https://copilot.microsoft.com/
3. https://www.copy.ai/
4. https://www.wordtune.com/
5. https://wordai.com/
6. https://pdf.ai/
7. https://aizzy.ai/
8. https://chat.openai.com/

Space for Learner Engagement for Objective Questions

Learners are encouraged to develop objective questions based on the content in the paragraph as a sign of their comprehension of the content. The Learners may reflect on the recap bullets and relate their understanding with the narrative in order to frame objective questions from the given text. The University expects that 1 - 2 questions are developed for each paragraph. The space given below can be used for listing the questions.

Major Task in NLP and The Future of NLP

Learning Outcomes

- Understanding the major tasks in NLP like machine translation (MT), Syntactic analysis, and Semantic analysis
- Understanding the different techniques of MT systems
- Understanding the different works to make Syntactic analysis and Semantic analysis
- Understanding the future works of NLP for Sanskrit

Background

In the covered units, we understood what is NLP and the different tasks in it too. This unit is going to discuss the major tasks in NLP such as machine translation, natural language generation, word and sentence segmentation, etc. which are the basic works for any NLP tasks. The future tasks on NLP for Sanskrit is also will be discussed. The improvement of the present systems also the need to create more systems for the needs of the hour.

Keywords

Machine Translation, Syntactic analysis, Semantic analysis, Natural language generation, Parsing, Word segmentation, Morphological segmentation

Discussion

As we experienced in unit 1 of this block Natural Language Processing (NLP) has plenty of real-world application. As we experienced NLP is an important tool for understanding and

processing the immense volume of unstructured data. NLP has made significant progress in recent years, and there are several major tasks within the field. Additionally, the future of NLP is likely to involve advancements in various areas.

3.1 NLP Tasks

NLP tasks- Machine translation, syntactic a and semantic analysis

NLP tasks can be divided into three major divisions according to the works happening in Sanskrit fields. The division is mainly on the basis of the techniques and methods used for NLP tasks. They are machine translation, syntactic analysis, and semantic analysis.

3.2 Machine Translation

Machinetranslation (MT) refers to the process of automatically translating text from one language to another using a computer application. The goal of machine translation is to produce output that is accurateand fluent in the target language, conveying the meaning of thesource language content.

MTsystem uses artificial intelligence to translate text without humaninvolvement to translate text or speech. It can also communicate thefull meaning of the original language text in the target language.

There are several approaches to machine translation, and advancements inthis field have been driven by developments in artificialintelligence and natural language processing.

3.2.1 Rule-Based Machine Translation (RBMT)

Anusarika is the rule based machine translation

This approach involves creating a set of linguistic rules and dictionaries to translate text. While it can be effective for specific language pairs and domains, it often requires extensive manual effort to develop and maintain the rules. Anusaraka is the rule-based machine translation approach that has to done for Sanskrit. The rules are made according to Ashtadhyayi. DESIKA by C-Dac is also a rule-based approach.

3.2.2 Statistical Machine Translation (SMT)

Google Translate for Sanskrit uses the SMT system to translate

SMT relies on statistical models that are trained on bilingual corpora to learn the probability of word sequences. It became popular in the mid-2000s and has been successful in many applications. However, it has limitations, especially when

dealing with languages with different word orders or idiomatic expressions. Google Translate for Sanskrit uses the SMT system to translate.

3.2.3 Neural Machine Translation (NMT)

Reva University, reported an NMT based system for translation

NMT is a more recent approach that uses deep neural networks, particularly recurrent neural networks (RNNs) or transformer models, to learn the mapping between languages. NMT has shown significant improvements over SMT, producing more fluent and contextually accurate translations. Reva University, Bengaluru reported in 2019 an NMT-based system for Sanskrit to English translation.

3.2.4 Other Methods

The other modern methods of Machine Translation approach are -

- Transfer Learning and Pre-training - is a methods of MT. Pre-training large language models on a broad range of language tasks has become a common practice. For example, models like GPT (Generative Pre-trained Transformer) are fine-tuned for specific tasks, including translation, leveraging the pre-existing knowledge.
- Transformer Architecture - allows the model to consider the entire context of a sentence, capturing long-range dependencies more effectively than traditional sequential models.

Major companies have developed machine translation systems

Major companies like Google, Microsoft, and Facebook, have developed and deployed machine translation systems. These systems are often integrated into various products and services, such as online translation platforms, language translation apps, and global communication tools.

It is important to note that while machine translation has made significant progress, it still faces challenges, particularly in accurately translating idiomatic expressions, preserving nuances, and handling low-resource languages. Sanskrit also facing such challenges. Researchers continue to work on improving the quality and capabilities of machine translation systems.

3.1.2 Syntactic analysis

Syntax – how grammar is generated and how sentence are formed

Syntax is the arrangement of words, phrases, and clauses in a sentence to make them meaningful. Syntax understands how grammar is generated and how sentences are formed.

The syntactic techniques in NLP are listed below.

3.1.2.1 Parsing

Parsing is the grammatical analysis of a sentence. ***Example***: A natural language processing algorithm is fed the sentence, "रमेशःअन्नंखादति।" Parsing involves breaking this sentence into parts of speech -- i.e., रमेशः = noun, अन्नं = noun, खादति = verb. This is useful for more complex downstream processing tasks.

3.1.2.2 Word segmentation.

This is the act of taking a string of text and deriving word forms from it. ***Example:*** A person scans a handwritten document into a computer. The algorithm would be able to analyze the page and recognize that the words are divided by white spaces (Space, tab, new line, etc.).

3.1.2.3 Sentence breaking

This places sentence boundaries in large texts. ***Example:*** A natural language processing algorithm is fed the text, "रमेशःअन्नंखादति।अहंपरिवेषयामि।" The algorithm can recognize the period that splits up the sentences using sentence breaking (full stop).

3.1.2.4 Morphological segmentation

Parsing, Sentence segmentation, Breaking, Morphological segmentation, stemming

This divides words into smaller parts called morphemes. ***Example:*** The word अत्युत्तमम् would be broken into [अति[उत्तम[अम्]]], where the algorithm recognizes "अति," "उत्तम," and "अम्" as morphemes. This is especially useful in machine translation and speech recognition.

3.1.2.5 Stemming

This divides words with inflection in them into root forms. ***Example:*** In the sentence, "रमेशःअन्नंखादति।," the algorithm would be able to recognize the root of the word "खादति" is "खाद्." This would be useful if a user was analyzing a text for all instances of the word खादति, as well as all of its conjugations.

The algorithm can see that they are essentially the same root even though the words are different.

3.1.3 Semantic analysis

Semantics evolves how meaning is constructed and communicated in spoken or written language. It also examines how meaning is interpreted by speakers and listeners. Semantics can address meaning at the word, phrase, sentence, and discourse levels of language. The most important methods in semantic analysis are Word sense disambiguation, Named entity recognition, and Natural language generation. The first two are covered in unit 1 of this block. The Natural Language Generation is explained here.

3.1.3.1 Natural language generation

NLG is the process of creating new text from a given input

Natural Language Generation (NLG) is the process of creating new text from a given input. This uses a database to determine the semantics behind words and generate new text. For example: An algorithm could automatically write a summary of findings from a business intelligence platform, mapping certain words and phrases to features of the data in the AI platform. Another example would be automatically generating news articles based on a certain body of text used for training.

3.2 Future NLP Tasks for Sanskrit

The Natural Language Processing for Sanskrit faces new challenges and opportunities. Sanskrit has a very rich linguistic structure and as well as it has certain characteristics that distinguish it from modern languages. For modelling the Sanskrit language for different NLP tasks need to handle the unique grammatical structure, noun and verb conjugations, syntactic features, and semantic features. Most of the major NLP tasks are discussed in earlier portions. The task that needs to be done for Sanskrit is listed here.

The first and foremost task in NLP needs to make for Sanskrit is a Machine translation system for Sanskrit to Indian languages and vice-versa and Sanskrit to English and other foreign languages. As we saw in 3.2 there are various approaches to develop a MT system for Sanskrit. Different approaches with different techniques also happened in the

Sanskrit field. But still, the researchers are improving the systems for betterment.

MEIT is now funding translation projects like Bhashini etc.

Ministry of Electronics and Information Technology (MEIT), is now funding translation projects like Bhashini, Sampark, Anuvadaksh etc. So a flowering future for Sanskrit is also foreseen.

The other major NLP task for Sanskrit is to make available a Spell-checker for Sanskrit. A text-to-speech system is also an important tool for Sanskrit.

There are OCR for Sanskrit available, but for doing OCR for old manuscripts are still not developed. In India, there are many manuscripts that are still unearthed. Even if it is available most users may not be able to read the old Devanagari script itself. Manuscripts are found in so many scripts. Making OCR for old scripts which are found in Manuscript is a valuable NLP task for the Sanskrit fraternity.

OHWR. NER for Sanskrit are the other important tasks for Sanskrit

Online handwriting recognition systems (OHWR), Voice search systems, etc. also need to be made for Sanskrit. Named Entity Recognition (NER) for Sanskrit, Text Summarization in Sanskrit, Sanskrit Speech Recognition, etc are the other important tasks for Sanskrit.

As we saw above most of the NLP works for Sanskrit have started and research is going on in these fields. Bringing it to a shape and in a usable format need to do more work on this field. Contributing to Natural Language Processing (NLP) tasks for Sanskrit can be a rewarding endeavor, and students can make valuable contributions to the field.

Summarised Overview

This unit discussed the major tasks in NLP and the tasks that need to be done for Sanskrit in detail. The major tasks are divided into three major divisions machine translation, syntactic analysis, and semantic analysis. Different approaches to MT are discussed. Different works on syntactic analysis with examples are also discussed. Semantic analysis and its types are also discussed in detail. Finally, the tasks need to be made for Sanskrit using NLP techniques.

Assignments

1. Differentiate, the different techniques of MT systems.
2. Describe the Syntactic analysis tasks and their use in Sanskrit.
3. Write semantical tasks in NLP for Sanskrit.
4. Write a short note on ‘how you can contribute to Sanskrit, by doing different NLP tasks

Suggested Readings

1. Mane, Deepak. (2013). Study of various approaches in Machine Translation for Sanskrit Language, International Journal of Advancements in Research & Technology, Volume 2, Issue4, April-2013
2. Koul, Nimrita& Manvi, Sunilkumar. (2019). A proposed model for neural machine translation of Sanskrit into English. International Journal of Information Technology.

References

1. Mane, Deepak. (2013). Study of various approaches in Machine Translation for Sanskrit Language, International Journal of Advancements in Research & Technology, Volume 2, Issue4, April-2013
2. Koul, Nimrita& Manvi, Sunilkumar. (2019). A proposed model for neural machine translation of Sanskrit into English. International Journal of Information Technology.
3. https://www.xenonstack.com/blog/evolution-of-nlp
4. https://subscription.packtpub.com/book/data/9781788478311/1
5. https://www.techtarget.com/searchenterpriseai/definition/natural-language-processing-NLP

Space for Learner Engagement for Objective Questions

Learners are encouraged to develop objective questions based on the content in the paragraph as a sign of their comprehension of the content. The Learners may reflect on the recap bullets and relate their understanding with the narrative in order to frame objective questions from the given text. The University expects that 1 - 2 questions are developed for each paragraph. The space given below can be used for listing the questions.

Computational Application of Sanskrit Grammar

Learning Outcomes

- Learning the structure of Sanskrit grammar
- Knowing the techniques used by Panini in Aṣṭādhyāyī
- Knowing the computational approaches of Panini

Background

In this unit we continue with a brief rendering of Sanskrit grammar cereated by Panini. The structure of Panini's Aṣṭādhyāyī and the techniques used by Panini to create this knowlege system. The computational accuracy of Panini is also indicated here.

Keywords

Aṣṭādhyāyī, Sūtra, Gaṇapāṭha, Dhātupāṭha, Anuvrtti, Kāraka

Discussion

Sanskrit grammar is highly systematic and structured

Sanskrit grammar is highly systematic and structured. It is codified in the ancient grammarian Panini's work "Aṣṭādhyāyī", composed around the 4th century BCE. The structure of Sanskrit grammar is characterized by its precision, logical organization, and comprehensive coverage of linguistic elements.

Panini's Aṣṭādhyāyī stands as a linguistic marvel, representing a cornerstone in the evolution of grammatical studies. This ancient Sanskrit text comprises eight meticulously crafted chapters, each housing aphorisms (सूत्रा's) that succinctly

encapsulate rules governing the language's phonetics, morphology, and syntax. Panini's approach, utilizing a generative grammar model, provides a systematic framework for understanding the intricacies of Sanskrit, with rules so comprehensive that they account for the entire language. The Aṣṭādhyāyī's enduring influence transcends its historical and cultural significance, extending into modern linguistic theory and even computational linguistics, where its rule-based structure has inspired researchers exploring applications in natural language processing. Panini's linguistic legacy endures through the Paninian tradition, with scholars and grammarians continuing to study and contribute to the understanding of Sanskrit grammar and its broader implications in the realm of language.

4.1 Aṣṭādhyāyī

Panini's Astadhyayi is organized into eight chapters

Panini's "Aṣṭādhyāyī" is organized into eight chapters, each referred to as an "Adhyāya." The structure of the "Aṣṭādhyāyī" is highly systematic, and each chapter is further divided into sections and rules.

4.1.1 Structure within Adhyāyas

Pāda

Each adhyāya is further divided into four quarters called "Pādas." These Pādas contain groups of Sutras that address specific aspects of grammar.

Sūtra

The core of the Astadhyayi is the sutra

The core of the "Aṣṭādhyāyī" is the sūtra, which is an aphorism or a rule. Panini's sūtras are concise and use a symbolic notation that requires interpretation and commentary for a complete understanding.

4.1.2 Five Parts of Panini's Vyākaraṇa

Panini's vyākaraṇa is divided into five parts or one main section and four supplementary sections, which provide additional information and classifications to aid in understanding the grammatical rules presented in the main text.

Sūtrapāṭha, dhātupāṭha, gaṇapāṭha, uṇādipāṭha, and liṅgānuśāsana are known as "pañcāṅgaṃvyākaraṇaṃ" (Five parts of paniniyan grammar)

Sūtrapāṭha

Sūtrapāṭha or śabdānuśāsana is again divided into two as - śivasūtras and aṣṭādhyāyīsūtrapāṭha.

Sutrapatha – Sivasutra and Astadhyayisutrapatha

The śivasūtras are a set of fourteen aphorisms that appear as an appendix to the "Aṣṭādhyāyī." These sutras, also called "māheśvarasūtras" are distinct from the main body of Panini's work. They provide a phonological and morphological code, offering a concise summary of certain aspects of Sanskrit phonetics and grammar.

The second part is Aṣṭādhyāyī – the eight chapters composed by Panini.

Gaṇapāṭha

Ganapatha

The gaṇapāṭha is an appendix found at the end of Panini's "Aṣṭādhyāyī ." It lists various groups of words based on specific criteria, such as the patterns of derivations or similarities in sound. In Panini's sūtra, he refers to the gaṇa names and needs to consult gaṇapāṭha to get the list of the words. The gaṇapāṭha serves as a helpful reference guide for understanding the application of grammatical rules in different contexts.

Dhātupāṭha

Dhatupatha is a list of roots organized systematically

The dhātupāṭha is a list of roots (dhātus) organized systematically. It provides a comprehensive catalog of verb roots, helping to understand the various forms and conjugations of verbs in Sanskrit. The dhātupāṭha is divided into ten classes according to the forms of the dhātus. The dhātupāṭha is an essential resource for students and scholars studying Sanskrit morphology.

Uṇādipāṭha

The uṇādipāṭha or uṇādisūtras are an additional set of aphorisms that appear as an appendix. They focus on the topic of nominal stems, particularly those derived from the roots mentioned in the dhātupāṭha. The uṇādisūtras contribute to the understanding of the formation and variations of nominal forms in Sanskrit.

Linganuśāsana

The "linganuśāsana" section in Panini's grammar provides rules for determining the gender of nouns based on certain

characteristics, such as their endings, meaning, or inherent gender. Panini's systematic approach to gender assignment is one of the aspects that makes his grammatical work highly detailed and structured.

These pañcāṅgas enhance the completeness of Panini's grammatical work by providing additional insights, explanations, and references. They are crucial for scholars and students seeking a more detailed understanding of the grammatical rules presented in the main body of the "Ashtadhyayi."

Linganusasana provides rules to determine the gender of Nouns

4.2 Panini's Approach to Sanskrit Grammar

Panini has written a grammar to Sanskrit and he recorded the concepts in it as a universal framework to write grammar for other languages.

Panini's approach to Sanskrit grammar is generative, meaning that his rules are formulated to generate all possible correct forms of a Sanskrit sentence. The generative power of grammar refers to its ability to produce, or generate, a vast and varied set of well-formed structures within a language. This systematic approach distinguishes the "Aṣṭādhyāyī" as a pioneering work in linguistics.

Astadhyayi as a pioneering work in linguistics

4.2.1 Features of the Paninian Grammar

Paniniyan grammar focuses on deriving grammatically correct sentences and phrases. It is based on the spoken language of Panini's time and includes rules on Vedic usage and regional variants. Panini uses many techniques to make the grammar a derivational one.

Pratyaharas

Using sivasutras Panini made pratyahara's which helped to achieve economy (**लाघव**) while creating sutras.

Pratyaharas helps to achieve economy while creating sutras

Meta-rules

Meta rules are the rules about rules. Panini named it as **परिभाषासूत्राणि**. These sutras help each sutra to avoid complications in action.

Technical Terms and Technical Language

Panini chooses technical terms that are not at all used anywhere to avoid confusion in the meaning such as **टि, भ, घि, घु** etc. Also,

a special language to facilitate the description. These also helped to achieve economy.

Anuvrtti

Anuvrtti is a concept used by Panini to achieve economy as well and it helps to understand a clear meaning of sutras.

The rules Pūrvatrāsiddhaṃ, paranityāntaraṅgāpavādānāmuttarottaraṃbalīyaḥ, vipratiṣedheparaṃkāryam etc are the best conflict resolution methods used by Panini.

Word creation method

Two categories of words – Noun and Verbs

Panini divides words into two categories nouns and verbs. Creating a word from a root word need adding suffixes, which have different meaning according to it. Also, a word contains so much technical information except its meaning or etymology. For example the word भवति = भू + अ + ति

1. The root भूhas the meaning of being

2. अindicates कर्तृकारकम्

3. तिcontains the informations -

A) It should have a कर्ता

B) The कर्ताwill be in प्रथमाविभक्तिएकवचन

C) It should have a कर्म

D) The कर्मwill be in द्वितीयाविभक्तिएकवचन

Kāraka Theory

Karaka

The Kāraka theory is a foundational aspect of Panini's linguistic framework, providing a systematic way to understand the relationships between different elements in a sentence. Karakas are grammatical cases or syntactic relations that define the roles of various nouns or noun phrases in a sentence. The kāraka theory of Panini is a universal theory that can apply to any language. The Kārakas are -

Kartā (Agent)

The Kartr kāraka represents the 'doer' or the agent of an action. It is associated with the noun or noun phrase that acts as the verb.

Karma (Object)

The Karma kāraka denotes the 'object' of the action. It is linked to the noun or noun phrase that undergoes the action of the verb.

Karaṇa (Instrument)

The Karaṇakāraka signifies the 'instrument' or means by which the action is performed. It is linked to the noun or noun phrase representing the instrument.

Sampradāna (Recipient)

The Sampradānakāraka indicates the 'recipient' or the one for whom something is given. It is associated with the noun or noun phrase receiving something.

Apadāna (Source of the motion)

The Apadana kāraka refers to the 'source' or 'origin' of an action or motion. It is linked to the noun or noun phrase indicating the starting point of an action.

Adhikaraṇa (Location)

The Adhikaraṇakāraka signifies the 'location' or 'setting' where an action takes place. It is associated with the noun or noun phrase representing the place.

4.3 Computational Process to Paninian Grammar

Computational Process to Paninian grammar

The Kāraka theory, as presented by Panini in the "Ashtadhyayi," provides a foundation for enriching existing grammatical structures through additional rules. To construct a new version of the grammar focusing on numerical or computing terms, one could introduce Kāraka rules and affixing markers that segregate these terms from ordinary Sanskrit grammar.

Expressing computations using Paninian grammar could involve transcribing a modern formalism into the grammar, leveraging its native techniques for categorical and rule definitions. It can use the grammar's metarules to formulate rules for numeric or computational tasks and symbolic manipulations of formal systems.

The parallel between a modern computing language or formalism and Paninian grammar can be understood in terms of four major functional components. These components include

the rewrite formalism for constructing Sanskrit expressions, metalinguistic paribhāṣa rules guiding operations, the finite inventory of phonemes, stems, and roots to which rules initially apply, and the verified sūtra codifying all rules in reduced form.

The sūtra formulation is likened to a computing language summarized in a terse programming manual intended for oral recitation, with linguists or grammarians as ancient programmers.

In conclusion, Panini's Kāraka theory, presented in the "Ashtadhyayi," provides a foundation for enriching grammatical structures, allowing the construction of a modified grammar focused on numerical or computing terms. This involves introducing Kāraka rules and affixing markers, showcasing parallels between modern computing languages and Paninian grammar in key components like rewrite formalism and metalinguistic rules.

Summarised Overview

This unit starts with an introduction to the structure and importance of Sanskrit grammar and continues with the structure of Panini's Ashtadhyayi. Describes the five parts of Paninian grammar and then moves to the techniques used by Panini. After detailing these techniques, it concludes with the computational processes of Sanskrit Grammar.

Assignments

1. Write an essay words on "the structure of Ashtadhyayi"
2. Describe the techniques used by Panini to achieve economy.
3. Describe how Paninian grammar is suitable for computational processing.

Suggested Readings

1. Subbanna, S., Varakhedi, S. (2008). Computational Structure of the Aṣṭādhyāyī and Conflict Resolution Techniques. In: Kulkarni, A., Huet, G. (eds) Sanskrit Computational Linguistics. ISCLS 2009. Lecture Notes in Computer Science, vol 5406. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-93885-9_5
2. Scharf, Peter. (2007). ModelingPaninian Grammar.

References

1. Subbanna, S., Varakhedi, S. (2008). Computational Structure of the Aṣṭādhyāyī and Conflict Resolution Techniques. In: Kulkarni, A., Huet, G. (eds) Sanskrit Computational Linguistics. ISCLS 2009. Lecture Notes in Computer Science, vol 5406. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-93885-9_5
2. Scharf, Peter. (2007). Modeling Paninian Grammar.
3. DESHPANDE, MADHAV M. "PĀṆINI AND THE DISTINCTIVE FEATURES." *Bulletin of the Deccan College Research Institute*, vol. 54/55, 1994, pp. 87–97. *JSTOR*, http://www.jstor.org/stable/42930460. Accessed 9 Dec. 2023.
4. https://bhavana.org.in/computing-processes-and-pan%CC%A3inis-as%CC%A3t%CC%A3adhyayi/

Space for Learner Engagement for Objective Questions

Learners are encouraged to develop objective questions based on the content in the paragraph as a sign of their comprehension of the content. The Learners may reflect on the recap bullets and relate their understanding with the narrative in order to frame objective questions from the given text. The University expects that 1 - 2 questions are developed for each paragraph. The space given below can be used for listing the questions.

Model Question Paper (SET – A)

SREENARAYANAGURU OPEN UNIVERSITY

Reg.NO..... **QP CODE**.....

Name.....

FIRST SEMESTER M.A. SANSKRIT LANGUAGE AND LITERATURE EXAMINATIONS
ABILITY ENHANCEMENT COMPULSORY COURSE - 1 M23SN01AC-
M23SN01AC - Computational Sanskrit
(CBCS- PG)
2023 Admission Onwards

Time: 3 Hours **Max Marks:70**

SECTION- A

I. Answer any five Questions in a Paragraph **(5 x 2=10)**

1. What is the structure of Sanskrit language?
2. Explain different kinds of tools that are available for Sanskrit
3. Write about Samasaadini
4. What is Sampark?
5. Give a brief note on electronic resources
6. What are E journals? Write the names of Sanskrit e journals
7. What is Artificial Intelligence?
8. Describe the facilities provided by Anuvadini

SECTION –B

II. Answer any six Questions in a page **(6 x 5=30)**

9. Give a brief description of Language processing tools
10. Explain Machine translation system
11. Explain Sanskrit E Resources
12. Write the importance of using Sanskrit Unicode fonts for typing
13. List five AI tools and their usage

14. Describe the applications of AI.
15. Write semantical tasks in NLP for Sanskrit.
16. Describe how Paninian grammar is suitable for computational processing.
17. Explain the features of Paninian grammar.
18. Write semantical tasks in NLP for Sanskrit.

SECTION – C

III. Write an essay on any two of the following (2x15=30)

19. Write an essay on "the structure of Ashtadhyayi"
20. Write an essay on the importance of NLP tasks in Sanskrit.
21. There are different applications for Sanskrit with AI.Explain
22. Differentiate, the different techniques of MT systems.

Model Question Paper (SET – B)

SREENARAYANAGURU OPEN UNIVERSITY

Reg.NO..... **QP CODE**.....

Name.....

FIRST SEMESTER M.A. SANSKRIT LANGUAGE AND LITERATURE EXAMINATIONS
ABILITY ENHANCEMENT COMPULSORY COURSE - 1 - M23SN01AC-
Computational Sanskrit
(CBCS- PG)
2023 Admission Onwards

Time: 3 Hours **Max Marks:70**

SECTION- A

I. Answer any five Questions in a Paragraph **(5 x 2=10)**

1. Write about Sambasha
2. What is E Book?
3. Write about the verbal system of Sanskrit.
4. What is Samasadhani?
5. Write about the different Sanskrit Heritage sites
6. List different Sanskrit Literary database
7. What is Artificial Intelligence?
8. List five AI tools and their usage.

SECTION –B

II.Answer any six Questions in a page **(6 x 5=30)**

9. Explain different types of encoding for Devanagari Script
10. Describe the applications of Artificial Intelligence
11. Write the relevance of Sanskrit E-Journals and Magazines
12. Describe different types of Machine Translation

13. Explain future NLP task for Sanskrit
14. Describe the Syntactic analysis tasks and their uses in Sanskrit.
15. What are the uses of Sanskrit tools and categorize it.
16. List and explain different Sanskrit Literary Database.
17. Describe the facilities provided by Anuvadini
18. List and describe the different computer applications which are used for typing Sanskrit

SECTION – C

lllWrite an essay on any two of the following (2x15=30)

19. Describe how Paninian grammar is suitable for computational processing.
20. Write an essay on the importance of NLP tasks in Sanskrit.
21. There are different applications for Sanskrit with AI. Explain
22. Write the importance of using Sanskrit Unicode fonts for typing

സർവ്വകലാശാലാഗീതം

വിദ്യയാൽ സ്വതന്ത്രരാകണം
വിശ്വപൗരരായി മാറണം
ഗ്രഹപ്രസാദമായ് വിളങ്ങണം
ഗുരുപ്രകാശമേ നയിക്കണേ

കൂരിരുട്ടിൽ നിന്നു ഞങ്ങളെ
സൂര്യവീഥിയിൽ തെളിക്കണം
സ്നേഹദീപ്തിയായ് വിളങ്ങണം
നീതിവൈജയന്തി പാറണം

ശാസ്ത്രവ്യാപ്തിയെന്നുമേകണം
ജാതിഭേദമാകെ മാറണം
ബോധരശ്മിയിൽ തിളങ്ങുവാൻ
ജ്ഞാനകേന്ദ്രമേ ജ്വലിക്കണേ

കുരീപ്പുഴ ശ്രീകുമാർ

SREENARAYANAGURU OPEN UNIVERSITY

Regional Centres

Kozhikode
Govt. Arts and Science College
Meenchantha, Kozhikode,
Kerala, Pin: 673002
Ph: 04952920228
email: rckdirector@sgou.ac.in

Thalassery
Govt. Brennen College
Dharmadam, Thalassery,
Kannur, Pin: 670106
Ph: 04902990494
email: rctdirector@sgou.ac.in

Tripunithura
Govt. College
Tripunithura, Ernakulam,
Kerala, Pin: 682301
Ph: 04842927436
email: rcedirector@sgou.ac.in

Pattambi
Sree Neelakanta Govt. Sanskrit College
Pattambi, Palakkad,
Kerala, Pin: 679303
Ph: 04662912009
email: rcpdirector@sgou.ac.in

Computational Sanskrit

COURSE CODE: M23SN01AC

Sreenarayanaguru Open University
Kollam, Kerala Pin- 691601, email: info@sgou.ac.in, www.sgou.ac.in Ph: +91 474 2966841

ISBN 978-81-966907-9-3