

Syllabus for P. G. Diploma in Sanskrit Computational Linguistics

Department of Vyakarana, Shastra Faculty, KSU

Semester – I

Paper – I	Natural Language Processing – I
Paper – II	Computer Programming – I
Paper – III	Vyakarana and Linguistics - I
Paper – IV	Introduction to Shabdabodha

Semester – II

Paper – V	Natural Language Processing – II
Paper – VI	Computer Programming – II
Paper – VII	Vyaakarana - II
Paper – VIII	Project

For each of the paper, we describe the objective, and the topics that are likely to be covered.

The syllabus, as well as the reading material and reference list is only indicative.

Detailed Syllabus

Paper -I Natural Language Processing – I

Objective: At the end of this course the students should be able to assess our traditional linguistic resources vis-a-vis the modern linguistic resources, also assess the relevance of fundamental principles and concepts in Indian traditional theories to the modern languages.

1. Introduction and brief History of NLP
2. MT in India and abroad
3. Linguistic issues in NLP
4. Morpheme, Word, Sentence, and Paragraphs
5. Morphological Analysis and finite State Transducers
6. Chunking
7. Parsing
8. Annotation of Sanskrit texts at various stages

Recommended Books and reading material:

- NLP: A Paninian Perspective by Akshar Bharati, Vineet Chaitanya, Sangal, Prentice Hall of India,
- 1995
- Speech and Language Processing By Daniel Jurafsky and James H Martin
- Annotation guidelines developed by Sanskrit Consortium
- Relevant research papers in the field of Machine Translation, Natural Language Processing,
- Computational Linguistics, Sanskrit Computational Linguistics, etc.
- A Key to Karaka

Paper-II Computer Programming -I

Objective: The goal of this course is to introduce the students to various Unix tools and scripting languages so that students can develop small interfaces on the top of existing tools, process corpus, do preliminary linguistic and statistical analysis of the corpus.

1. Introduction to Unix file system
2. Introduction to various Unix tools such as cut, paste, more, less, tr, diff, comm, locate, find
3. regular expressions - grep, sed, flex (lexical analyser)
4. Simple shell programmes command line arguments, loop, conditional statements
5. Introduction to HTML, and XML
6. Introduction to Apache, server programming
7. Philosophy behind GPL, Creative Commons and similar licences

Recommended Books:

- Unix Power Tools, by Jerry Peek, Shelley Powers, Tim O'Reilly, Mike Loukides
- Online tutorials for Apache, HTTP and Javascript

Paper-III Vyakarana -I

Objective: The aim of this course is to introduce the concepts of vyaakarana with reference to various issues in Natural Language Processing, and also to familiarise the students with the parallel Linguistic terminology and concepts so as to enable them to read and understand the latest research articles in the area of computational linguistics.

1. Phonology; Phonemics; Sandhi rules in A.s.taadhyaayii
2. Pada formation – subanta, tinganta, k.rt, taddhita; inflectional and derivational morphology, various approaches of morphological analysis
3. Syntactic Analysis, Kaaraka relations, theta roles
4. Akaanksha, yogytaa, sannidhi

Recommended Book:

- Siddhanta Kaumudi
- Ashtadhyayi
- Phonetics in Ancient India, W S Allen, 1971
- Sandhi, W. S. Allen
- Morphology
- Syntax

Paper -IV Introduction to Shabdabodha

Objective: This course aims at introducing the prominent concepts of Shabdabodha to the students.

1. शाब्दबोधः
2. प्रमाणम्, शब्दः
3. कारणानि - आकाङ्क्षा, योग्यता, सन्निधिः
4. शाब्दबोधोत्पत्तिक्रमः
5. पदज्ञानं तु करणं ...
6. वाक्यं वाक्यलक्षणम्
7. वाक्यार्थः
8. विशेष्यविशेषणभावः
9. पदम् - पदविभागः
10. वृत्तिः - शक्तिः, लक्षणा, व्यञ्जना
11. शक्तिग्रहोपायाः
12. शाब्दबोधे मतानि
13. मुख्यविशेष्यः कः?

14. क्रिया-भावना-प्रथमान्तार्थः
15. अन्विताभिधानम्
16. अभिहितान्वयः
17. संसर्गमर्यादावादः

Recommended Books:

- शाब्दतरङ्गिणी, सुब्रह्मण्यशास्त्री,- Prof. KTP edition - 2006.
- "The word and the world" - B.K.Matilal - 1992
- "Indian theories of Meaning" - Raja K. Kunjuni - 1963
- Philosophy of word and meaning, Gourinath Shastri - 1959
- "Sanskrit Philosophy of Language" JF Stall 1969
- "Logic, Language, Reality" - B.K. Matilal 1985

Paper -V Natural Language Processing – II

Objective: At the end of this course the students should be able to assess our traditional linguistic resources vis-a-vis the modern linguistic resources, also compare the relevance of fundamental principles and concepts in Indian traditional theories to the modern languages.

1. Corpus Linguistics
2. Corpus, collection, Digital Resources
3. Word Sense Disambiguation
 - Problems
 - Various approaches
4. Various Sanskrit Koshas, Amarakosha: Knowledge Structure
5. Electronic dictionaries and their linking
6. E-lexicons
7. WordNet, ConceptNet, PropNet, VerbNet
8. Lakshan Charts, Kaaraka Charts

Recommended Books and reading material:

- Speech and Language Processing By Daniel Jurafsky and James H Martin

- Amarakos ṣa: Sudhā Vyākhyāna
- Nirukta: durgā vyākhyā,
- Nirukta: laks ṣman ṣsarupa
- Lexicography: Rama-dhara SiMha
- Relevant research papers
- Online Lexical resources and their Documentation

Paper-VI Computer Programming -II

Objective: The basic aim of this course is to introduce basic concepts of programming and data structure to the students.

1. Introduction to Computer programming
2. Variables
3. Various Data structures scalar, array, hash, string, enumeration, set
4. String processing
5. Memory, pointers
6. Various constructs: Loop, conditional
7. Modularity, subroutines
8. Global Vs local variables
9. Parameter passing
10. Use of various libraries

Reference Books:

As decided by the instructor depending upon the language chosen.

Paper-VII Vyaakarana -II

Objective: The aim of this course is to introduce the concepts of vyaakarana with reference to various issues in Natural Language Processing, and also to familiarise the students with the parallel Linguistic terminology and concepts so as to enable them to read and understand the latest research articles in the area of computational linguistics.

1. Compounds – Analysis and generation
2. Derivation process in A.s.taadhyayii
3. Abhidhaa, lakshanaa, vyanjanaa
4. Meaning deciding linguistic factors

Reference Material:

- Ashtadhyayii
- Prathamaa Av.rtti of Yudhi.s.tiir miimaansaka
- Theories of Meaning: Kunjunni Raja

Paper-VIII Project

Objective: This course given Students and apparently to implement the thesis they studied and the will be a testing bed for thesis understanding. Students have to work on a problem selected on the guidance of hi/her teacher/ supervisor and submit a small dissertation of the end of the year in order to fulfill partial the requirement of the course.

Areas for Projects

1. Sanskrit Language Processing
2. Any language analysis based on Shastric approach
3. Machine Translation
4. Word-sense-disambiguation
5. Speech processing and so on.

Model Question Paper Pattern for all papers

I. Objective type Questions -	2X20 =	40
II. Short Notes	5X5 =	25 (with two extra Choices)
III. Small Essay	10X2 =	20 (with two extra Choices)
IV. Long Essay	15X1 =	15 (with two extra Choices)