CENTRAL UNIVERSITY OF KERALA DEPARTMENT OF LINGUISTICS M.A. LINGUISTICS AND LANGUAGE TECHNOLOGY				

CORE COURSE				
COURSE CODE	COURSE TITLE	CONTACT HRS/WEEK		CREDITS
		LEC	LAB	TUT	
LLT5205	COMPUTATIONAL LINGUISTICS -I	2	2	1	4

Lec = Lecture, Tut = Tutorial, Lab = Practical

LLT 5205 COMPUTATIONAL LINGUISTICS I

Course Objectives:
1. To introduce the basic concepts in Computer and programming concept from Computational Linguistics perspective.
2. To understand the system of computer to understand language structure and the system analysis of language structures.
3. To develop the language data for the machine understanding and interpretation of language and their structures.

Learning Outcomes:
After successfully completing this course students will have

1. Understanding about the encoding language in computer.
2. Understanding about basic concept of Python and Perl Programming language.
3. Understanding about issues in developing computational grammar.
4. Able to prepare like POS, Treebank language data for purpose of language technology interface.
5. Developing Computational tools for analysis of natural language for the purpose of machine translation.

1. Introduction: Computational paradigm for Natural language Analysis and synthesis. Linguistics, Applied Linguistics and Computational Linguistics, Grammar formalisms and computational modelling for Natural Languages. Computational Techniques and Tools:

2. Computational Morphology: Morphological generators and analyzers for Indian Languages. Case Studies of morphological analyzers and testing Morphological Systems. Finite State machine for morphological computation. Taggers and shallow parsers.

3. Corpora Studies in Linguistics: Corpus as a basis for linguistic studies. Chomsky's criticism of early corpus linguistics. Modern corpus studies, corpus building. Applications of corpus. Digital corpora in computational studies. Characteristics of modern corpora and Applications of corpora.

4. Machine Translation: The earliest applications of computational Linguistics. Early models and latest developments. History and the Socio-political relevance of MT. Differences between MT, MAT and HT. The architectural issues. The feasibility and the relevance. The analytical pyramid in MT. Various MT experiments in India, Anusaraka, Anglabharti, ILILMT, EILMT etc.

5. Modeling: Students are expected to practice the existing models, and also work out certain models of Generators, Analyzers and taggers. Write short Algorithms for linguistic models of computational application. Use corpus for testing of models. Students need to submit a short write up for internal assignment and appear for a written test in the class room.

Suggested Readings:
*1. Cole, Ronald et al (eds.).1997. Survey of the State of the Art in Human language Technology. Cambridge: CUP.
2. Douglas Arnold, Lorna Balkan, Seity Meijer, R. Lee Humphreys & Loisa Sadler. 1994. Machine Translation : An Introductory Guide. New York: Blackwell Ltd.
*3. Jurafsky, D. and J. Martin. 2002. Speech and Language Processing. New Delhi: Pearson Education Inc.
4. King, Margaret. 1987. A Tutorial on Machine Translation. No. 53
5. Kucera, Henry and W. Nelson F.1967. Computational analysis of Present-day American English. Providence, RI: Brown University Press.
6. Litkowski C. Kenneth. 1992. A Primer on Computational Lexicography.
7. Miller, G. A., R. Beckwith, C. Fellbaum, D. Gross and K. A. Miller.1990. Introduction to Wordnet: An on-line Lexical Database. IJL, 3(4):235-44.
8. Raskin, Victor. Linguistics and Machine Translation. In Linguistics and Translation (ed.).
*9. Ruslan, Mitkov (ed.). 2002. The Oxford Handbook of Computational Linguistics. Oxford: OUP.
10. Uma Maheshwar Rao, G. and Amba Kulkarni. 2007. Natural Language and Computing. PGDCAIL, vol. 411. CDE, University of Hyderabad.