File Ref.No.6912/GA - IV - J1/2013/CU

UNIVERSITY OF CALICUT

Abstract

Scheme and Syllabus of MPhil Computer Science Programme for the University Department- revised - implemented with effect from 2014 admissions -orders issued.

G & A - IV - J

U.O.No. 9878/2014/Admn Dated, Calicut University.P.O, 25.10.2014

Read:-1) Item No: 3of the Minutes of the meeting of the Board of Studies in Computer Science held on 6.8.14

2) Remarks of the Dean, Faculty of Science on 24.09.2014

3) Orders of the Vice-Chancellor in the file of even Number on 13.10.2014

ORDER

As per paper read as (1) above, the Board of Studies at its meeting, revised the syllabus of M Phil. Computer Science under University Teaching Department w.e.f 2014 admission onwards and has forwarded the Scheme and Syllabus to the University. As per paper read as (2) above, the Dean, Faculty of Science has recommended to approve item No 1 to 4 of the minutes of Board of studies and the syllabus of M Phil. Computer Science.

The Hon'ble Vice Chancellor, considering the exigency, exercising the powers of the Academic Council , has approved the item No 1 to 4 of the minutes of the meeting of the BOS in Computer Science, subject to ratification by the Academic Council, vide reference cited 5th.

Sanction has therefore been accorded to implement the scheme and syllabus of M Phil. Computer Science programme of University Teaching Departments w.e.f 2014 admission onwards.

Orders are issued accordingly. Scheme and Syllabus appended. (The syllabus is available in the website: universityofcalicut.info)

Muhammed S

Deputy Registrar

To

Controller of Examination

Exam Wing

Digital Wing

Forwarded / By Order

Section Officer

UNIVERSITY OF CALICUT

DEPARTMENT OF COMPUTER SCIENCE

Regulations, Scheme of Evaluation Course, Structure Syllabus for

MASTER OF PHILOSOPHY (M.Phil.)

in

COMPUTER SCIENCE

(Choice Based Credit Semester System – CCSS
For the students admitted from the Academic Year 2014-15 onwards)

Under the

FACULTY OF SCIENCE

August, 2014

© University of Calicut, 2014

UNIVERSITY OF CALICUT

Department of Computer Science

Regulations, Scheme of Evaluation Course, Structure Syllabus for

M.Phil. Computer Science

(with effect from 2014 Admission)

REGULATIONS

1. **Duration** of the course shall be 1 Year, divided into 2 semesters. The course work is based on credit system. Three papers, each carrying four credits (total 12 credits), shall be the course requirements in the first semester. In the second semester the student should carry out a research project under the supervision of a research guide. The project shall be on the area of specialization, which he/she takes up for study. The candidate need to submit a dissertation at the end of second semester and it carries 12 credits. In each paper there shall be an external and internal evaluation. The external evaluation shall be for 80% and internal for 20%.
2. **Selection and Eligibility** for Admission is based on the existing University rules (Candidates who have passed Masters Degree in Computer Science (M.Sc. Computer Science) or Masters Degree in Computer Application (MCA) or Masters Degree in Information Technology (M.Sc. IT) or Masters Degree in Computer Engineering (M.Tech. Computer Science/M.Tech. IT) recognized by the University of Calicut with not less than 55% marks is eligible).
3. **Evaluation** of all semester theory papers will be on the basis of existing University norms.
4. **Dissertation & Viva-voce**: The Dissertation should be carried out over the period of 16 weeks in the second semester in the Department/Institution. Every student should do the research work individually and no grouping is allowed. All the candidates are required to get the approval of their synopsis from the guide before commencement of the work. The research work will be reviewed periodically every month by the Department / Institution. At the end of the second semester the candidate shall submit the Dissertation (three bound copies and one soft copy) duly approved by the research Guide. If dissertation is found to be not up to the expected standard, the examiners can ask the candidate to modify and resubmit the project report after incorporating the suggestions of the examiners. Such reports shall be resubmitted within the stipulated period suggested by the examiner(s). After the evaluation of the dissertation, there will be a viva voce examination, jointly conducted by the examiner and the supervising teacher.

M.Phil. Computer Science

(with effect from 2014 Admission)

COURSE STRUCTURE AND SCHEME OF EVALUATION

Semester 1

Sl.No	Course Code	Course	Instructional Hrs/week		Duration of examination (Hrs)	Marks			Credits
			Lect.	Tutorial	Theory	External	Internal	Total	
1	MPCS1C01	Research Methodology	3	1	3	80	20	100	4
2	MPCS1C02	Advanced Trends in Computer Science	3	1	3	80	20	100	4
3	MPCS1E01 /02/..../08	Elective -1 (Specialization)	3	1	3	80	20	100	4
	Total		9	3	-	-	-	300	12

Elective 1 (Specialization)	Credit
MPCS1E01- Digital Speech Processing	4
MPCS1E02- Digital Image and Video Processing	4
MPCS1E03- Advanced Pattern Recognition	4
MPCS1E04- Neuro-Fuzzy and Soft Computing	4
MPCS1E05- Natural Language Processing	4
MPCS1E06- Remote Sensing and GIS	4
MPCS1E07- Web Data Mining	4
MPCS1E08- Grid and Cloud Computing	4

Semester 2

Sl.No	Course Code	Course	Marks			Credits
			External	Internal	Total	
1	MPCS2C03	Dissertation	120	30	150	12
		Viva-voce	40	10	50	
	Total				200	12

Grading :

The minimum required Cumulative Grade Point Average (CGPA) in each semester should be 5.5 out of a 10-point scale.

The grading shall be as follows :

Percentage of Marks	**Grade**	**Grade Points**
90 and above	A+	10
80-89	A	9
70-79	B+	8
60-69	B	7
50-59	C+	6
Below 50	D	Nil

Note : 0.5% Marks and above will be rounded to the next number.

UNIVERSITY OF CALICUT

Department of Computer Science

M.Phil. Computer Science

(with effect from 2014 Admission)

SYLLABUS

© University of Calicut, 2014.

MPCS1C01 – RESEARCH METHODOLOGY

Unit - 1

Research Methodology:

Introduction to Scientific Research, Meaning, Objectives and Significance of Research Motivation in Research, Types of research approaches, Quantitative research methods, Research methods versus methodology, Research process, Criteria of good research, Research problems, Necessity of defining the problem, Technique involved in defining the problem, Design and Development Research Methods, Meaning of research design, Need for research design, Features of a good design, Different research designs, Basic principles of experimental designs, Ethics in research, Building expertise in the areas of interest, generating the base content in the selected area, literature survey for research work, arriving at directions of research, Formulation of research title, development of criteria based research proposal.

Unit – II

Probability and Statistics:

Probability as a measure of uncertainty, probabilities for events, axioms, probability rules, Fail time data analysis, Hazard models, conditional probability, Bayes' rule, random variables, probability distributions, discrete and continuous distributions, univariate and multivariate distributions, joint, marginal, conditional distributions, expected values (mean, variance, covariance), sampling/simulation, study of a population or distribution, System reliability, Stochastic process, Software tools for Mathematical and statistical analysis, Scilab/SPSS.

Unit – III

Linear Algebra:

Introduction to vectors, vector and linear combinations, lengths and dot product, solving linear equations, vector space and subspace , orthogonality of the four subspaces, eigen values and eigen vectors, linear transformations.

Unit – IV

Scripting Languages:

Overview: The nature of scripting languages, scripting vs programming, Python Programming. Regular expressions, Network programming, Internet client programming, Multithreaded programming, GUI programming, Database programming, Web clients and servers, Web programming: CGI and WSGI , Web frameworks : Django, web services.

Unit – V

Technical writing using LaTeX:

Scientific Writing : Significance of report writing, Structure and Components of Research Report, Types of Report: research papers, thesis, Research Project Reports, Precautions for writing research reports, Pictures and Graphs, Citation Styles, Oral presentation, Exposure to LaTeX, Installation, MikTeX, TeXnicCenter, Creating reports and articles, Text environment, Math environment, Figures, Tables, BibTeX - reference manager, Camera Ready Preparation.

REFECENCES:

1. C.R.Kothari, *Research Methodology Methods & Techniques*, 2nd Edition, Wishwa Prakashan Publishers.
2. Misra R.P, *Research Methodology – A Hand Book*, Concept publishing Company, New Delhi 1988
3. Kai Lai Chung, *A Course in Probability Theory*, Third Edition, Academic Press.
4. Gilbert Strang, *Introduction to Linear Algebra*, 3rd edition, Wellesley-Cambridge Press and SIAM
5. David Barron, *The World of Scripting Languages*, Wiley Publications.
6. *Core Python application programming*, Third edition Wesley J Chun, PEARSON.
7. Leslie Lamport, *LaTeX: A Document Preparation System*, Second Edition.

MPCS1C01 – ADVANCED TRENDS IN COMPUTER SCIENCE

Unit – I

Design and Analysis of Algorithms:

Algorithm Analysis, Asymptotic Notations and Basic efficiency classes, Mathematical analysis of Recursive and Non-recursive algorithms, Master's theorem. Design- Divide and conquer, Greedy approach, Dynamic programming, Backtracking, Branch-and-Bound, Concepts of complexity classes P, NP, NP-hard, NP-complete.

Unit – II

Formal Languages and Finite Automata:

Chomsky hierarchy, Regular languages and Finite automata, Finite automata with epsilon moves, Regular expressions, Properties of regular languages, Applications. Context free languages and Push-down automata, Derivation trees, Simplification of context free Grammars, Chomsky normal form, Greiback normal form, Pumping lemma.

Unit – III

Distributed Computing and Web services:

Distributed Systems: Fully distributed processing systems, Networks and Interconnection structures, Design, Distributed databases, Challenge of distributed data, Loading factors, managing the distributed resources, division of responsibilities, distributed Computing introduction, Inter-process Communications, Distributed Computing Paradigms, The Socket API, The Client-Server Paradigm, Group Communication. Web Services Architecture – characteristics, core building blocks of web services, standards and technologies available for implementing web services, web services communication, basic steps of implementing web services, developing web services enabled applications.

Unit – IV

Web Data mining:

Introduction to web mining, Web content mining, Web structure mining, Web usage mining, Text mining, Text clustering, Temporal mining, Spatial mining, visual data mining, Knowledge mining, Various tools and techniques for implementation (weka, Rapidminer or Scilab).

Unit – V

Mobile Computing:

Wireless transmission, Frequencies for radio transmission, Signal propagation, Multiplexing, Cellular systems. Medium Access Control–SDMA, FDMA, TDMA, CDMA. Telecommunication systems– GSM-Mobile services, System architecture, Handover, Security, Satellite systems-Applications, Types, Handover. Broadcast systems - Digital Audio and Video Broadcasting. Wireless LAN-Infra red Vs radio transmission, Bluetooth. Mobile network layer- Mobile IP – IP packet delivery, Registration, Tunnelling and Encapsulation, Dynamic Host Configuration Protocol (DHCP). Transport Layer- Indirect TCP, Snooping TCP, Mobile TCP, and Transaction oriented TCP. WAP architecture.

REFERENCES:

1. Alfred V.Aho, John E.Hopcroft and Jeffrey D.Ullman, *Data structures and Algorithms*, Pearson Education Asia, 2002.
2. Horowitz E & Sahni S, *Fundamentals of data structures*, Computer Science press, 1978.
3. Cormen, Leiseson, RIvest and Stein, *Introduction to Algorithms*, 3rd Edn. PHI, 2010.
4. Richard Neapolitan and Kumarss Naimipour, Foundation of Algorithms, 4th Edn., Jones & Bartlett Publishers, 2011.
5. John E. Hopcroft, Jeffery D. Ullman, *Introduction to Automata Theory Language and Computation*, Addison-Wesley, 1999.
6. John A. Sharp, *An Introduction to Distributed and Parallel Processing*, Blackwell Scientific Publications, 1987.
7. Uyless D. Black, *Data Communications & Distributed Networks*, Prentice Hall, 1997.
8. Joel M. Crichllow, *Introduction to Distributed and Parallel Computing*, Prentice Hall, 1988.
9. S. Graham et., *Building Web Services with Java*, 2nd Edition, , Pearson Edn., 2008
10. Arun K Pujari, *Data Mining Techniques*, University press, Edition 2001.
11. T.Sushmita mitra, Tir ku Acharaya, *Data Mining Multimedia,Softcomputing & Bioinformatics*, Wiley Interscience publications , 2004.
12. C S R Prabhu, *Data Warehousing – concepts, techniques and applications*, 2nd Edition, Prentice Hall of India, 2002.
13. Schiller J., *Mobile Communications*, Addison Wesley, 2/e, Pearson Education, 2009.
14. William Stallings, *Wireless Communications & Networks,* Pearson Education, 2005.
15. C.Siva Ram Murthy, B.S. Manoj, *Ad Hoc Wireless Networks – Architectures and Protocols*, 2nd Edition, Pearson Education.
16. Ashok K Talukder, Roopa R Yavagal, *Mobile Computing*, Tata McGraw Hill, 2005.

MPCS1E01 – DIGITAL SPEECH PROCESSING

Unit – I

Introduction to speech processing, History, Speech production, Mechanism of speech production, Acoustic phonetics, Digital models for speech signals, Speech waveform representations, Sampling speech signals, Basics of quantization, Applications.

Unit – II

Short-time analysis of speech, Short-time energy and Zero crossing rate, Short-time auto correlation method, Short-time Fourier Transform, Speech spectrogram, Homomorphic speech analysis, Cepstrum, Complex Cepstrum, Short-time cepstrum, Computation of Cepstrum, Mel Frequency Cepstrum Co-efficients (MFCC), Linear predictive analysis.

Unit – III

Text to Speech Synthesis: Basic principles, Rule based speech synthesis, Corpus based Speech synthesis, Linguistic processing, Prosodic processing.

Unit – IV

Speech Recognition: Speech recognition architecture, Types and issues in speech recognition, - Speech databases, Performance evaluation, Feature extraction methods, Speech recognition methodologies, Acoustic-phonetic approach, Pattern recognition approach, Template based approach, Dynamic Time Warping, Hidden Markov Model (HMM), Vector Quantization, Support Vector Machine (SVM) - Artificial Neural Network (ANN) based approaches, Language Model, Trigram language model, Usage of HMM and CMU-SLM Toolkit.

Unit – V

Speaker Identification and Verification: Measuring speaker features, Statistical Vs Dynamic features, Cepstral analysis, Similarity Vs Distance measures, Constructing speaker models, Adaptation, Applications of speaker recognition, Text dependent/independent speaker recognition, Generative approaches: Rationale, Gaussian mixture model (GMM), Neural network approaches, Discriminative approaches: Support Vector Machine(SVM), Kernels.

REFERENCES:

1. L. R. Rabiner, R. W. Schaffer, *Digital Processing of Speech Signals*, Prentice Hall, 1978.
2. Jacob Benesty, M. Mohan Sondhi, Yiteng Huang, *Springer Handbook of Speech Processing*, Springer, 2007.
3. Douglas O'Shaughnessy, *Speech Communications: Human and Machine*, Wiley-IEEE Press, 1999.
4. L.R. Rabiner, B. H. Juang, *Fundamentals of Speech Recognition*, Prentice Hall, 1993.

MPCS1E02 – DIGITAL IMAGE AND VIDEO PROCESSING

Unit – I

Physiological characteristics of Human eye and image formation, fundamental steps in image processing - elements of digital image processing systems - digital image fundamentals, a simple image model – sampling and quantization - basic relationship between pixels - image geometry.

Unit – II

Image enhancement, Basic intensity Transformation functions, Histogram equalization, Spatial domain filtering, The continuous 2D Fourier transform and its properties ,The discrete 2D Fourier transform (DFT),Sampling and Quantisation, FFT, DCT, Filtering in Frequency Domain, Homomorphic Filtering.

Unit – III

Image restoration - model of Image degradation/restoration process, Noise models, Inverse filter, Weiner Filter. Colour models- CIE, RGB, CMYK, HSI and HSV. Morphological image processing- Erosion, Dilation, Opening and Closing. Image Segmentation - Fundamentals, Point, Line and Edge Detection, Thresholding, Region Based Segmentation. Feature Analysis and Extraction, Image Classification.

Unit – IV

Image Compression, Fundamentals, Models, Elements of Information Theory, Error Free Compression, Lossless compression - Huffman coding - Arithmetic coding - Bit plane coding, Run length coding, Lossy Compression, Multi Resolution Analysis, Image Pyramids, Multi resolution expansion, Wavelet transforms, Compression Standards, Applications of Image Processing.

Unit- V

Video Formation, Perception and Representation, Digital Video, Video Modelling, Fourieranalysis of Video Signals, Two-Dimensional Motion Estimation, Video Compression Standards- MPEG4 and MPEG7 and beyond.

REFERENCES:

1. R.C. Gonzalez and R.E. Woods, *Digital Image Processing – 3rd ed.*, Prentice Hall of India, New Delhi, 2008
2. A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall
3. Milan Sonka, V. Hlavac and R. Boyle, *Image Processing Analysis and Machine Vision*, Brooks/colic, Thompson Learning, 1999.
4. B. Chanda and D.D. Majumder, *Digital Image Processing and Analysis*, PHI
5. W.K. Pratt, *Digital Image Processing*, John Wiley, 2006
6. David Saloman, *Data Compression: The Complete Reference*, Springer
7. Tomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, Wiley InterscicencePublication, John Wiley & Sons, Inc.
8. Gilbert Held, *Data and Image Compression*, John Wiley & Sons Ltd.,
9. A.MuratTekalp, *Digital Video Processing*, Prentice Hall

MPCS1E03 – ADVANCED PATTERN RECOGNITION

Unit – I

Basics of pattern recognition, Introduction to statistical, syntactic and descriptive approaches, Bayesian decision theory, Classifiers, Discriminant functions, Decision surfaces, Normal density and discriminant functions, Discrete features.

Unit – II

Parameter estimation methods, Supervised learning, Maximum-Likelihood estimation, Gaussian mixture models, Expectation-maximization method, Bayesian estimation, Hidden Markov models for sequential pattern classification, Discrete hidden Markov models, Continous density hidden Markov models.

Unit – III

Dimension reduction methods, Fisher discriminant analysis, Principal component analysis, Non-parametric techniques for density estimation, Parzen-window method, K-Nearest Neighbour method.Linear discriminant function based classifiers, Perceptron, Support vector machines, Non-metric methods for pattern classification, Non-numeric data or nominal data, Decision trees.

Unit – IV

Unsupervised Classifications, Clustering for unsupervised learning and classification, Clustering concept, C-means algorithm, Hierarchical clustering procedures, Graph theoretic approach to pattern clustering, Validity of clustering solutions.

Unit – V

Recent Advances, Neural network structures for Pattern Recognition, Neural network based pattern associators, Unsupervised learning in neural Pattern Recognition, Self organizing networks, Fuzzy pattern classifiers, Pattern classification using Genetic Algorithms.

REFERENCES:

1. R.O.Duda, P.E.Hart and D.G.Stork, *Pattern Classification*, John Wiley, Second edition, 2006
2. Gonzalez R.C. & Thomson M.G., *Syntactic Pattern Recognition - An Introduction*, Addison Wesley.
3. Fu K.S., *Syntactic Pattern Recognition And Applications*, Prentice Hall, Eaglewood cliffs
4. Rajjan Shinghal, *Pattern Recognition: Techniques and Applications*, Oxford University Press, 2008.
5. Robert J. Schalkoff, *Pattern Recognition : Statistical Structural and Neural Approaches*, John Wiley & Sons Inc., New York, 1992.
6. Morton Nadier and Eric Smith P., *Pattern Recognition Engineering*, John Wiley & Sons, New York, 1993.
7. Theodoridis and K.Koutroumbas, *Pattern Recognition*, 4th Ed., Academic Press, 2009
8. C.M.Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006

MPCS1E04 – NEURO-FUZZY AND SOFT COMPUTING

Unit – I

Introduction, Soft Computing Constituents, Neuro-Fuzzy and Soft Computing Characteristics, Fuzzy sets, Membership functions, Basic operations, Measures and fuzziness, MF Formulation and Parameterization.

Unit – II

Fuzzy rules and fuzzy reasoning, Extension Principle and Fuzzy Relations, Fuzzy If-Then Rules, Fuzzy Reasoning, Fuzzy Inference Systems, Introduction, Mamdani Fuzzy Models, Sugeno Fuzzy Models, Tsukamoto Fuzzy Models, fuzzy models for pattern recognition.

Unit – III

Artificial Neural Network, Single layer perceptron, Multilayer Perceptron (MLP), Supervised Learning Neural Networks, MLP using Backpropagation of error, Radial Basis Function Networks, Unsupervised Learning and Other Neural Networks, Kohonen Network, Competitive Learning Networks, Kohonen Self-Organizing Networks, Learning Vector Quantization, Hebbian Learning, Principal Component Networks, Hopfield Network.

Unit – IV

Neuro-Fuzzy Computing, Various ways of integration, Mathematical foundation of fuzzy neural network, Incorporating fuzzyness in neural network framework, Neuro-Fuzzy Modeling, NFIS: Adaptive Neuro-Fuzzy Inference Systems, ANFIS Architecture, Hybrid Learning Algorithm, ANFIS Applications.

Unit – V

Advanced Topics: Evolutionary computation (EC), Introduction to Genetic Algorithm, Genetic Operators and Parameters, Genetic Algorithms in Problem Solving, Rough set, Characteristics, Knowledge encoding using Rough sets, Configuration of Rough-Fuzzy MLP for classification, Harmony search, Swarm intelligence.

REFERENCES:

1. J.S.R.Jang, C.T.Sun and E.Mizutani, *Neuro-Fuzzy and Soft Computing*, Pearson Education, 2004.
2. Sankar K. Pal and Sushmita Mitra, Neuro-fuzzy Pattern Recognition Methods in Soft Computing, John Wiley & Sons, 1999.
3. Li-Xin Wang, A course in Fuzzy Systems and Control, Prentice-Hall, 1997
4. M. Mitchell, *An Introduction to Genetic Algorithms*, Prentice-Hall, 1998.
5. D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.
6. S. V. Kartalopoulos, *Understanding Neural Networks and Fuzzy Logic: Basic Concepts and Applications*, IEEE Press - PHI, 2004.
7. S. Rajasekaran & G. A. Vijayalakshmi Pai, *Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis & Applications*, PHI, 2003.

MPCS1E05 - NATURAL LANGUAGE PROCESSING

Unit – I

What is Natural Language Processing, Introduction and Overview, Ambiguity and uncertainty in language, Regular Expressions, Chomsky hierarchy, Regular languages and their limitations, Finite-state automata, Practical regular expressions for finding and counting language phenomena, Morphology, Inflectional Morphology, Derivational Morphology, Finite-State Morphological Parsing, Combining an FST Lexicon and Rules, Porter Stemmer, Exploring a large corpus with regex tools.

Unit – II

Context Free Grammars, CFG definition, use and limitations, Chomsky Normal Form, Top-down parsing, bottom-up parsing, and the problems with each. The desirability of combining evidence, from both directions Non-probabilistic Parsing Efficient CFG parsing with CYK, Earley parser, Designing a little grammar and parsing with it on some test data.

Unit – III

Probabilistic language modeling and its applications, Markov models, N-grams, Estimating the probability of a word and smoothing, Generative models of language, Part of Speech Tagging and Hidden Markov Models, Viterbi Algorithm for Finding Most Likely HMM Path, Dynamic programming with Hidden Markov Models and its use for part-of-speech tagging.

Unit – IV

Probabilistic Context Free Grammars, Weighted context free grammars, Weighted CYK, Pruning and beam search, Parsing with PCFGs, Probabilistic version of CYK, Modern parsers, Maximum Entropy Classifiers, Maximum entropy principle and its relation to maximum likelihood, Maximum entropy classifiers and their application to document classification, sentence segmentation and other language processing tasks.

Unit – V

Maximum Entropy Markov Models & Conditional Random Fields, Part-of-speech tagging, noun-phrase segmentation and information extraction models that combine maximum entropy and finite-state machines, State-of-the-art models for NLP, Lexical Semantics, Mathematics of Multinomial and Dirichlet distributions, Dirichlet as a smoothing for multinomials, Information Extraction & Reference Resolution, Various methods, Machine learning methods for coreference.

REFERENCES:

1. D. Jurafsky and J. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, P Hall, 2008.
2. C. Manning and H. Schutze, *Foundations of Statistical Natural Language Processing*", MIT
3. James Allen, *Natural Language Understanding*, The Benajmins/Cummings Publishing Company
4. Cover, T. M. and J. A. Thomas, *Elements of Information Theory*. Wiley.
5. Charniak, E, *Statistical Language Learning*, The MIT Press.
6. James Allen. *Natural Language Understanding*, Addison Wesley, 1994.

MPCS1E06-REMOTE SENSING AND GIS

Unit- I

Concepts and Foundations of Remote Sensing: Definition, scope, types and History and recent developments in remote sensing, ideal and real remote sensing system, Comparison of conventional survey, aerial remote sensing and satellite remote sensing, advantage and limitation of satellite remote sensing, EMR and Remote Sensing: Energy sources, electro magnetic radiation, spectral regions, energy interaction in the atmosphere, atmosphere window, energy interaction with earth surface features, spectral reflectance patterns for different region of EMR, Actors affecting remote sensing signatures, Platforms, data capture types and systems, data recording methods.

Unit-II

GIS and spatial data, Definition, Issues of Spatial and non-spatial data collection, Representation and standardization, Components of GIS, Maps and spatial data, Computer assisted mapping and analysis, Thematic characteristics of spatial data, Geographic Grid, GPS coordinate system, Sources of spatial data; Census and survey data, Air photos, Satellite images, Field data, Data organization (location, attributes, consistency, scale), Data analysis operations in GIS, Data Interoperability, Data Classification, Terminologies in measurements of lengths, perimeter and area in GIS, Queries, Reclassification, buffering and neighborhood functions, Integrated data.

Unit- III

Data Models: Vector data model; Raster data model; Vector data – Objects and topology, Vector data input, editing, attribute, data input and management; Raster data – Types of raster data, Raster data structure, Data conversion; Integration of raster and vector data. Spatial Data Analysis: Vector data analysis – Buffering, Mapoverlay, Distance measurement, Map manipulation; Raster data analysis – Analysis environment, Local operations, Neighborhood operations, Zonal operations, Distance measure operations, Spatial auto correlation.

Unit- IV

Terrain Mapping and Spatial Interpolation: Terrain mapping and analysis, DEM, TIN, Operations in terrain mapping ; Spatial Interpolation, Control points, Global methods (Trend surface Analysis, Regression Models) , Local Methods (Theissen polygons, Density estimation, Inverse distance weighted interpolation, Thin-plate splines, krigging).

Unit- V

Land use/land cover: Corp assessment, disease detection, forestry: types – species identification and diseases detection. Soils: soil mapping – soil moisture – soil erosion – reservoir station – soil salinity – soil conservation. Water resources: surface water resources – water quality monitoring and mapping – water pollution, identification of ground water potential recharge areas – integrated watershed development.

REFERENCES:

1. Lilesand, TM John, *Remote sensing and Image interpretation*, Wiley.
2. Ian Heywood, Sarah Cornelius, Steve Carver, Srinivasa Raju, *An Introduction to Geographical Information Systems*, Pearson Education, 2007.
3. Bernhardensen Tor, *Geographic Information Systems: An Introduction*, John Wiley & So ns, Inc., 1999.
4. Campbell, James B., *Introduction to Remote Sensing* , (Second Edition) , Taylor & Francis, 1996.
5. Agarwal, C.S. and Garg P.K., *Textbook of Remote Sensing in Natural Resources Monitoring and Management*, Wheeler Publishing, New Delhi, 2000.
6. Avery T.E., *Interpretation of aerial Photographs*, Burgess Publishing Company, 1985
7. Bakker, Wim H., et al., *Principles of Remote Sensing – An Introductory Textbook*, ITC, 2001
8. Banerjee, R.K. and Banerjee, B., *Remote Sensing for Regional Development*, Concept Publishing Compan y, New Delhi, 2000.

MPCS1E07 – WEB DATA MINING

Unit – I

Introduction to internet and WWW, Data Mining Foundations, Association Rules and Sequential Patterns, Basic Concepts of Association Rules, Apriori Algorithm, Frequent Itemset Generation, Association Rule Generation, Data Formats for Association Rule Mining, Mining with multiple minimum supports, Extended Model, Mining Algorithm, Rule Generation, Mining Class Association Rules, Basic Concepts of Sequential Patterns, Mining Sequential Patterns on GSP, Mining Sequential Patterns on PrefixSpan, Generating Rules from Sequential Patterns.

Unit – II

Supervised Learning, Basic Concepts, Decision Tree Induction, Learning Algorithm, Impurity Function, Handling of Continuous Attributes, Classifier Evaluation, Rule Induction, Sequential Covering, Rule Learning, Classification Based on Associations, Naive Bayesian Classification , Naive Bayesian Text Classification - Probabilistic Framework, Naive Bayesian Model, Unsupervised Learning, Basic Concepts, K-means Clustering, K-means Algorithm, Representation of Clusters, Hierarchical Clustering, Single link method, Complete link Method, Average link method, Strength and Weakness.

Unit – III

Basic Concepts of Information Retrieval, IR Methods, Boolean Model, Vector Space Model and Statistical Language Model, Relevance Feedback, Evaluation Measures, Text and Web Page Preprocessing, Stopword Removal, Stemming, Web Page Preprocessing, Duplicate Detection, Inverted Index and Its Compression, Inverted Index, Search using Inverted Index, Index Construction, Index Compression, Latent Semantic Indexing, Singular Value Decomposition, Query and Retrieval, Web Search, Meta Search, Web Spamming.

Unit – IV

Link Analysis, Social Network Analysis, Co-Citation and Bibliographic Coupling, Page Rank Algorithm, HITS Algorithm, Community Discovery, Problem Definition, Bipartite Core Communities, Maximum Flow Communities, Email Communities, Web Crawling, A Basic Crawler Algorithm- Breadth First Crawlers, Preferential Crawlers, Implementation Issues – Fetching, Parsing, Stopword Removal, Link Extraction, Spider Traps, Page Repository, Universal Crawlers, Focused Crawlers, Topical Crawlers, Crawler Ethics and Conflicts.

Unit – V

Opinion Mining, Sentiment Classification, Classification based on Sentiment Phrases, Classification Using Text Classification Methods, Feature based Opinion Mining and Summarization, Problem Definition, Object feature extraction, Comparative Sentence and Relation Mining, Opinion Search and Opinion Spam. Web Usage Mining, Data Collection and Preprocessing, Sources and Types of Data, Key Elements of Web usage Data Preprocessing, Data Modeling for Web Usage Mining, Discovery and Analysis of Web usage Patterns, Session and Visitor Analysis, Cluster Analysis and Visitor Segmentation, Association and Correlation Analysis, Analysis of Sequential and Navigation Patterns.

REFERENCES:

1. Bing Liu, *Web Data Mining: Exploring Hyperlinks*, Contents, and Usage Data, Springer Publications
2. Jiawei Han, Micheline Kamber, *Data Mining: Concepts and Techniques*, Second Edition, Elsevier Publications.
3. Anthony Scime, *Web Mining : Applications and Techniques*
4. Kowalski, Gerald, Mark T Maybury: *Information Retrieval Systems: Theory and Implementation,* Kluwer Academic Press, 1997.
5. Frakes, W.B., Ricardo Baeza-Yates: *Information Retrieval Data Structures and Algorithms,* Prentice Hall, 1992.

MPCS1E08 – GRID AND CLOUD COMPUTING

Unit – I

Introduction to Grid Computing, Anatomy and Physiology of Grid, Early Grid Activities, Current Grid Activities, Grid Standards, Grid Challenges and Application area, Grid Computing Organization, roles.

Unit – II

Service Oriented Architecture, Web Service Architecture, Grid Architecture, Implementing Grid Architecture, Globus Toolkit, Services, Open Grid Services Architecture , Grid Scheduling and Resource Management, Framework, Grid Resource Management, Principles of Local Schedulers, Grid Scheduling with QoS, Data Management , Grid Security.

Unit – III

Cloud Computing , Overview, History, Key Terms and Definitions, Applications, Intranets and the Cloud, Cloud Today, Cloud Computing Services, On Demand Computing, Discovering Cloud Services, Software engineering fundamentals for Cloud Computing, Development Services and Tools.

Unit – IV

Cloud hardware and infrastructure, clients, security, network services, platforms, cloud storage, Key (Software) Technologies, Cloud software architecture issues, Classification of Cloud Implementations.

Unit – V

Operating System for the Cloud, Cloud Computing environments, Cloud Computing services available under various platforms, Application Patterns and Architecture, Comparison of Cloud Computing Platforms, Case Studies (Google AppEngine, Windows Azure Platform, VMware vCloud and Open Source Cloud Computing Platforms).

REFERENCES:

1. Joshy Joseph, Craig Fellenstein, *Grid Computing*, IBM Press, Pearson Education, 2004.
2. Ian Foster, Carl Kesselman (eds.), *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann Publishers, 2004.
3. Ahmar Abbas, *Grid Computing: A Practical Guide to Technology and Applications, Firewall Media*, 2009.
4. Anthony T.Velte, Toby J.Velte, Robert Elsenpeter, *Cloud Computing –A Practical Approach*, Tata McGraw Hill Education Pvt. Ltd, 2010.
5. Michael Miller, *Cloud Computing: Web based Applications that change the way you work and Collaborate Online*, Que Publishing, August 2008.
6. Haley Beard, *Cloud Computing Best Practices for Managing and Measuring Processes for on Demand Computing, Applications and Data Centers in the Cloud with SLAs*, Emereo Pvt. Ltd, July 2008.

General Pattern of Question Paper

Core and Elective courses in M.Phil. Computer Science Programme

(with effect from 2014 Admission)

Code: **Reg. No:**

Name :

First Semester M.Phil. Computer Science Degree Examination – 2014

Course Code : (eg: MPCS1C02) **Course : (Eg: Research Methodology)**

Time: 3 Hours **Total Marks: 80**

Answer five full questions; Each Question carries 16 marks.

Question Numbers 1 to 8 Total Marks = 5 x 16 Marks = 80 Marks

NOTE: Minimum one question from each of the five modules. Remaining three questions can be from any module. There should not be more than two questions from the same module.