UNIVERSITY OF CALICUT

Department of Computer Science

Regulations, Scheme of Evaluation Course, Structure Syllabus for

**M.Phil. Computer Science**
(with effect from 2012 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 2012 Admission)

COURSE STRUCTURE AND SCHEME OF EVALUATION

Semester 1

<table>
<thead>
<tr>
<th>Sl.No</th>
<th>Course Code</th>
<th>Course</th>
<th>Instructional Hrs/week</th>
<th>Duration of examination (Hrs)</th>
<th>Marks</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lect.</td>
<td>Tutorial</td>
<td>Theory</td>
<td>External</td>
</tr>
<tr>
<td>1</td>
<td>MPCS1C01</td>
<td>Research Methodology</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>MPCS1C02</td>
<td>Advanced Trends in Computer Science</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>80</td>
</tr>
<tr>
<td>3</td>
<td>MPCS1E01 /02/..../08</td>
<td>Elective -1 (Specialization)</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>9</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Elective 1 (Specialization)  
- 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

<table>
<thead>
<tr>
<th>Sl.No</th>
<th>Course Code</th>
<th>Course</th>
<th>External</th>
<th>Internal</th>
<th>Total</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MPCS2C03</td>
<td>Dissertation</td>
<td>120</td>
<td>30</td>
<td>150</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Viva-voce</td>
<td>40</td>
<td>10</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>200</td>
<td></td>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>

**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:

<table>
<thead>
<tr>
<th>Percentage of Marks</th>
<th>Grade</th>
<th>Grade Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>90 and above</td>
<td>A+</td>
<td>10</td>
</tr>
<tr>
<td>80-89</td>
<td>A</td>
<td>9</td>
</tr>
<tr>
<td>70-79</td>
<td>B+</td>
<td>8</td>
</tr>
<tr>
<td>60-69</td>
<td>B</td>
<td>7</td>
</tr>
<tr>
<td>50-59</td>
<td>C+</td>
<td>6</td>
</tr>
<tr>
<td>Below 50</td>
<td>D</td>
<td>Nil</td>
</tr>
</tbody>
</table>

Note: 0.5% Marks and above will be rounded to the next number.
Department of Computer Science

M.Phil. Computer Science
(with effect from 2012 Admission)

SYLLABUS
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, PHP Basics- Features,
Embedding PHP Code in your Web pages, Outputting the data to the browser, Data types,
Variables, Constants, expressions, string interpolation, control structures. Function, Creating a
Function, Function Libraries, Arrays, strings and Regular Expressions, PERL- Names and
Values, Variables, Scalar, Expressions, Control Structures, arrays, list, hashes, strings, pattern
and regular expressions, text processing, subroutines, advance perl - finer points of looping,
pack and unpack, file system, eval, data structures, packages, modules, objects, interfacing to the
operating system.
Unit – V
Technical writing using LaTeX:

REFERENCES:
8. Larry Wall, T. Christiansen and J. Orwant, *Programming Perl*, O'Reilly, SPD.
Unit – I
Design and Analysis of Algorithms:
Elementary Data Structures, Algorithm Analysis, Asymptotic analysis (best, worst, average cases) of time and space, Design: Divide and conquer, Greedy approach, Dynamic programming, Tree and graph traversals, Multistage graphs, Optimal binary search trees, 0/1 knapsack Connected components, Spanning trees, Shortest paths, upper and lower bounds, Concepts of complexity classes P, NP, NP-hard, NP-complete.

Unit – II
Formal Languages and Finite Automata:
Regular languages and finite automata, Context free languages and Push-down automata, Derivation trees, Simplification of context free Grammars, Chomsky normal form, Greiback normal form, pumping lemma for context free languages, Finite state systems, definitions, Non deterministic finite automata, Finite automata with epsilon moves, Regular expressions, Applications.

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:

REFERENCES:
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

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:
**MPCS1E02 – DIGITAL IMAGE AND VIDEO PROCESSING**

**Unit – I**
The physics of light, Optics and image formation, Physiological characteristics of Human eye and image formation, Image acquisition, Digital image representation, Sampling, Quantization, Camera sensor technologies, Human color vision, Color models: CIE, RGB, CMYK, HSI, HSV, L*a*b*, fundamental steps in image processing, elements of digital image processing systems.

**Unit – II**
Point processing, contrast enhancement, 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, Image Deconvolution, Deconvolution of noiseless images, inverse filter, Weiner Filter.

**Unit – III**

**Unit – IV**

**Unit – V**
Video Formation, Perception and Representation, Digital Video, Video Modeling, Fourier analysis of Video Signals, Two-Dimensional Motion Estimation, Video Compression Standards- MPEG4 and MPEG7 and beyond.

**REFERENCES:**
8. Gilbert Held, *Data and Image Compression*, John Wiley & Sons Ltd.,
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

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

REFERENCES:
3. Fu K.S., Syntactic Pattern Recognition And Applications, Prentice Hall, Eaglewood cliffs
MPCS1E04 – NEURO-FUZZY AND SOFT COMPUTING

Unit – I

Unit – II

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

Unit – V

REFERENCES:
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


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

REFERENCES:


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 is 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
REFERENCES:

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

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:

2. Jiawei Han, Micheline Kamber, Data Mining: Concepts and Techniques, Second Edition, Elsevier Publications.
3. Anthony Scime, Web Mining : Applications and Techniques
MPCS1E08 – GRID AND CLOUD COMPUTING

Unit – I

Unit – II

Unit – III

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

REFERENCES:
General Pattern of Question Paper

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

(with effect from 2012 Admission)

Code: Reg. No:
Name:

First Semester M.Phil. Computer Science Degree Examination – 2012

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.