Syllabus for M. Sc. (Scientific Computing)

LIST OF COURSES

Semester – I

SC-101 Programming Languages and principles I
SC-102 Software Engineering
SC-103 Foundation of Scientific Computing I
SC-104 Foundation of Scientific Computing II
SC-105 Computational Lab-I

Semester – II

SC-201 Programming Languages and Principles II
SC-202 Operating Systems
SC-203 Elective course
SC-204 Numerical Methods for Scientific Computing-I
SC-205 Computational Lab-II

Semester – III

SC – 301 Network Concepts
SC – 302 Parallel Processing and Grid Computing.
SC – 303 Elective course
SC – 304 Numerical Methods for Scientific Computing-II
SC – 305 Elective course

Semester – IV

SC-401 R & D/Industrial Project

Elective course will be offered from the following list.

Elective Courses

El-1 Application of Computer to Chemistry.

El-2 Scientific Visualization .
El-3 Statistical Computing.
El-4 Applications of Computer to Physics.
El-5 Biological Sequence Analysis
El -6 Modelling of Biological Systems
El -7 Artificial Intelligence
El -8 Quality Assurance and Software Testing
El -9 Soft Computing
El -10 Design concepts and Modeling
El -11 Business Analysis
El-12 Machine Learning
El-13 Data Mining

Detailed Syllabus:

SC – 101 Principles of Programming Languages

1. Introduction and Motivation **[8 hrs]**

Idea of analyzing an algorithm through examples, introduction to some notations, comparison of algorithms, notions of space and time efficiency and motivation for algorithm design methods, demonstration of algorithm analysis for some suitable example algorithm, say merge sort.

2. Algorithm Analysis Techniques **[12 hrs]**

(a) Asymptotic Analysis

Detailed coverage of asymptotic notations and analysis. Big Omicron, Big Theta, Big Omega, Small theta, Small omega. Comparison of the Insertion Sort and the Merge Sort Algorithms.

(b) Recurrence Analysis

Introduction to Recurrence equations and their solution techniques (Substitution Method, Recursion Tree Method, and the Master Method), Proof of the Master Method for solving Recurrences. Demonstration of the applicability of Master Theorem to a few algorithms and their analysis using Recurrence Equations. (Example algorithms: Binary Search, Powering a number, Strassen's Matrix Multiplication)

(c) Analysis of more Sorting algorithms: Quick Sort and Counting Sort

3. Algorithm Design Techniques **[12 hrs]**

(a) Types of Algorithms

(b) Dynamic Programming

Introduction and the method for constructing a DP solution, illustrative problems, e.g. assembly line scheduling problem using DP, or solution for the matrix chain multiplication.

(c) Greedy Algorithms

Greedy vs. DP, methodology, illustrative problems, e.g. the knapsack problem using a greedy technique, or activity selection Problem. Construction of Huffman Codes.

(d) Backtracking

Introduction to recursion, solving the 0-1 Knapsack problem using backtracking, pruning in backtracking and how it speeds up the solution for the 0-1 Knapsack problem.

(e) Branch and Bound

Description and comparison with backtracking, the FIFO B&B and the Max Profit B&B using the 0-1 Knapsack problem

4. Graph Theory **[12 hrs]**

(a) Breadth First and Depth First Search Algorithms

(b) Minimum Spanning Trees, Kruskal's Algorithm

(c) Minimum Spanning Trees, Prim's Algorithm

(d) Properties of Shortest paths.

(e) Dijkstra's Algorithm

(f) Bellman Ford Algorithm

5. NP-Completeness **[12 hrs]**

(a) Polynomial time

(b) Polynomial time verification (NP problems)

(c) Concept of NP-Hard with example (Halting problem)

(d) NP-Completeness and Reducibility (without proof)

(e) Some NP-Complete problems

(f) Overview of showing problems to be NP-Complete

Text Reference:

1. Introduction to Algorithms,
   T.H.Cormen, C.E.Leiserson, R.L.Rivest,
   Prentice Hall India, 2002.
2. The Art of Computer Programming, Vols. 1 and 3,
   D.E.Knuth,
   Addison Wesley, 1998.
3. Design and Analysis of Algorithms,
   A.V.Aho, J.E.Hopcroft, J.D.Ullman,
   Addison Wesley, 1976.
4. Fundamentals of Computer Algorithms,
   E.Horowitz, S.Sahni,
   Galgotia Publishers, 1984.
5. Data Structures and Algorithms, Vols.1 and 2,
   K.Melhorn,
   Springer Verlag, 1984.
6. The Analysis of Algorithms,
   P.W.Purdom, Jr. and C.A.Brown,
   Holt Rhinehart and Winston, 1985.

SC-102 SOFTWARE ENGINEERING

Part - I

1. Introduction to software engineering **[3 hrs]**

a. A generic view of process
b. Process models
c. An agile view of process

2. Software engineering practice

1. Introduction to software engineering **[3 hrs]**
   a. Need for a process
   b. Generic Process Model / Framework and umbrella activities
   c. Introduction to different process models: Waterfall, Incremental, Evolutionary, Component Based, The Unified Process, Agile
2. Software Development Lifecycle (SDLC) **[1 hr]**
   a. Phased approach
   b. The Waterfall process model
   c. Advantage and drawbacks
3. Requirements Analysis **[4 hrs]**
   a. Capturing Requirements
      i. Understanding different types of requirements
      ii. Requirement analysis steps
      iii. SRS or System Requirement Specification
   b. Analyzing Requirements
      i. Requirements modelling using use-case and use-case diagram, Activity diagram
      ii. Data Modeling: E-R Diagram , Sequence Diagram
      iii. Class Based modelling: Class Diagram, CRC
4. Design: **[7 hrs]**
   a. Architectural design
   b. Design Principles and fundamental design concepts
   c. Design Quality, Characteristics of a good design
   d. Component level Design: SOLID Principle, Cohesion and Coupling

   e. Introduction to Design Patterns

5. User Interface Design: **[1 hr]**
   a. Evolution of user interfaces
   b. Golden rules
   c. User interface analysis and design.
   d. User interface considerations
6. Construction / Development **[1 hrs]**
   a. Coding guidelines and standards
   b. Good coding practices
7. Testing or Quality Assurance **[4 hrs]**
   a. Software Testing Life Cycle (STLC): Test planning and Test case development
   b. Types of Testing: Static, Dynamic (Black box and White box)
   c. Black box Testing strategies
   d. Understanding Defect lifecycle, Severity and Priority
8. Software Project Management **[4 hrs]**
   a. Basic Principles, Project Planning Tasks
   b. Problem Decomposition, Task breakdown and Task network
   c. Estimation Techniques: Function Point Analysis,
9. Risk Management **[4 hrs]**
   a. Understanding Risk management: Proactive and Reactive strategies
   b. Risk Identification
   c. Risk exposure

Part - II

3. Review of Database Management Concepts **[6 hrs]**

   a. Types of database
   b. Normalization (1Nf, 2Nf, 3Nf, BCNF, 5NF),
   c. Data models and constraints,
   d. ER-model,
   e. Introduction to Scientific Database

4. Data Storage & Indexing techniques **[6 hrs]**

Architecture of DBMS, Storage of data on disk & files, File organization & type of file organization, Advanced storage devices – RAID, Type of Indexing

5. SQL query optimization **[8 hrs]**

Implementation and Evaluation of relational operations, Types of joins and join algorithms, Select of appropriate index, database workload, which index to create, guidelines for index selection, co-clustering, index on multiple attributes, Cost estimation and cost based optimization, Plan evaluation, Tuning conceptual schema

Text Reference:

1. Software Engineering a Practitioner’s Approach
   Roger S. Pressman
   McGraw – Hill
2. Software Engineering
   Richard Fairley
   Tata McGraw Hill
3. Software Engineering
   David Gustafson

4. Practical Guide in Structured System Design
   Meilier Page
5. Database Management System
   Raghuramkrishnan, Gehrke (3rd Edition)
   McGraw Hill
6. Database System Concepts(second edition)
   Silberschatz, Korth, Sudarshan
   McGraw Hill
7. Introduction to Database System
   C.J. Date
8. Fundamentals of Database systems (fourth edition)
   Elmaris, Nawathe
   Pearson Education
9. Distributed Database systems
   Tamer Ozsu
   Pearson Education

SC-103 : FOUNDATIONS OF SCIENTIFIC COMPUTING – I

1. Matrices **[10 hrs]**

Operations on Matrices, Inverse, System of equations, Eigenvalues and Eigenvectors

2. Vectors **[4 hrs]**

Operations on vectors

Scalar and vector product, Scalar Triple product.

3. Probability **[16 hrs]**

Outcomes, Events and Sample Space. Rules of probability. Combinatorics (Permutation and combination) Conditional probability and Independent events.

4. Statistics [**10 hrs]**

Simple Descriptive statistics Graphical statistics

5. Discrete Random Variable and Distributions **[16 hrs]**

Distribution of random variable, Expectation and Variance, Families of discrete distributions

6. Continuous Distributions **[4 hrs]**

Uniform and Normal Distributions, Central Limit Theorem.

Text Reference:

1. Linear Algebra (2nd edition)
   Hoffman and Kunz
   Prentice Hall International
2. Linear Algebra
   Peter D. Lax
3. Introduction To Probability And Statistics For Engineers And Scientists (4th Edition)
   Sheldon M. Ross
   Elsevier Academic Press
4. Probability and Statistics for Computer Scientists (2nd Edition)
   Michael Baron
   CRC Press

5. Advanced Engineering Mathematics (8^{th} Edition)
Erwin Kreyszig
John Willey and Sons

SC-104 : FOUNDATIONS OF SCIENTIFIC COMPUTING – II

1. Function **[2 hrs]**
Definition, examples, graphs of functions

2. Limit of a function **[4 hrs]**
Definition, right hand & left hand limits examples.

3. Continuity **[2 hrs]**
Definition, examples and properties of continuous functions ,Types of discontinuity.

4. Derivatives : **[6 hrs]**
Definition, geometrical interpretation, Derivatives of elementary functions by first principle. Problems on velocity and acceleration product, quotient and chain rule, implicit differentiation, derivative of inverse function.

5. Applications of derivatives **[10 hrs]**
Concavity and points of inflation, Maxima and minima of a function, Related rates, Roll’s theorem and Mean value Theorem, L′ Hospitals rule

6. Integration **[6 hrs]**
Introduction, elementary integration formulae, indefinite and definite integrals.

7. Integration Methods **[8 hrs]**
Substitution, Integration by parts, Integration of product and power’s of trigonometric functions, partial fractions.

8. Applications of integration **[10 hrs]**
Area under the curve, length of a curve, volumes and surface areas of solids of revolution.

9. Functions of two or more variables **[6 hrs]**
Definition, limits and continuity, partial derivatives, Directional derivatives, gradients and tangent planes, Second derivative, maxima, minima, saddle points

10. Differential Equations **[6 hrs]**
First order ODE, variables separable form, Solution of first order linear equation, Second and higher order equations, Solution of constant coefficient second order equation.

Text References:

1. Calculus and Analytical Geometry (9^{th} Edition)
Thomas and Finney
Pearson Education
2. Calculus (5^{th} Edition)
James Stewart
3. Advanced Engineering Mathematics (8^{th} Edition)
Erwin Kreyszig
John Willey and Sons
4. Linear Algebra (2^{nd} edition)
Hoffman and Kunz
Prentice Hall International
5. Linear Algebra
Peter D. Lax
6. Differentials Equations with applications and Historical notes.
Simmons G.F.

Interdisciplinary School of Scientific Computing

SC-105 : COMPUTATIONAL LABORATORY-I

A system development project should be taken and following should be done

(a) Analysis of the project

(b) Design of the project (using any design model)

(c) Coding (optional)

Experts from industry will guide these projects, which will be based on current technologies.

SC-201 : PROGRAMMING LANGUAGES AND PRINCIPLES-II

1. Object Oriented Programming Fundamentals [5 hrs]

(a) What is object?

(b) How to form an object?

(c) What is Object Oriented programming?

(d) Case study to identify objects - typically give a running system may be mechanical or any physically viewable system

2. Object Oriented Analysis and Design [6 hrs]

Object modeling, Inheritance, encapsulation, communication between objects

3. Working with classes [15 hrs]

Classes: Pointers and arrays, Dynamic memory, Expressions and statements, Various Types of Functions (Inline, Friend etc), Namespases and Exceptions, Concept of Classes, Types of Classes, Encapsulation, Conversions, type Promotion, Default Arguments And Type Casts.

4. Class Inheritance [10 hrs]

Operator Overloading, Inheritance, Virtual Functions, Templates, Exception Handling.

5. LISP [15 hrs]

(a) Introduction

Why LISP (is useful)?, Where LISP (is useful)?, Introduction to Symbol Manipulation, What is AI and where does LISP fit into it?

(b) The LISP Programming Language

Lists, Definitions, Predicates, Conditionals and Binding, Recursion and Iteration, Association Lists, Properties and Data Abstraction, Lambda Expressions, Printing and Reading, Macros, List Storage, Surgery and Reclamation, Arrays, OOP in LISP: CLOS (Awareness rather than formal teaching)

(c) Pattern Matching

Production Rules, Ordered PR, Limitations of Linear PR, Discrimination Nets, Pattern Matching in Production Rules, The MATCH Function, Application: Mathematical Formula Manipulation (Possible Assignment), Notion of Unification

(d) Knowledge Representation

Inclusion, Membership and 'ISA', Partial Orders and their Representation, An ISA Hierarchy in LISP, Inheritance (HAS Links), Propositional and Predicate Logic (A brief introduction), Constraints, Relational Databases, Problems of Knowledge Representation

(e) Searching

Elementary Search Techniques (DFS, BFS), Heuristic Search Methods (Best-First Search, Searching Graphs with Real Distances), Planning (A Robotics Example), Two Person, Zero-Sum Games (Tic-Tac-Toe/Checkers (Possible Assignment))

6. PROLOG **[6 hrs]**

(a) Syntax and Unification.
Prolog's slim syntax is described. Unification is described with examples to show how pattern matching is achieved.
(b) Lists, terms and arithmetic.
The Prolog syntax is used to create lists and terms, and to perform simple arithmetic.
(c) Graphs.
Classic graph algorithms are presented in the declarative Prolog style.
(d) Trees.
Classic tree algorithms are presented in the declarative Prolog style.
(e) Non-determinism.
The backtracking mechanism for finding alternative solutions is described and illustrated, with techniques to modify the default behavior.
(f) Negation as failure.
The reasons why Prolog saying ``No'' differs from ``false'' are discussed and illustrated.
(g) Difference structures.
Difference lists are described and students participate in a re-write of programs from classical lists to difference lists.
(h) Using difference structures.
Copying trees is used as an example to show the optimizations possible for large programs when the difference structure technique is used.

ASSIGNMENTS:

1. C++ Case studies should cover exception-handling, templates etc.
2. One project with sizable design should be implemented.
3. LISP Programming, Pattern Matching, Searching

Text Reference:

1. The C++ Programming Language
   B. Stroustroup
2. C++ Primer Plus
   S. Prata
   Galgotia Publications
3. LISP (3rd edition)
   Patrick Henry Winston and Berthold Klaus Paul Horn
   Pearson Education
4. The Elements of Artificial Intelligence: An Introduction using LISP
   Steven L. Tanimoto
5. Programming in Prolog for Artificial Intelligence (3rd edn)
   Bratko I
   Addison-Wesley, 2000
6. Programming in Prolog (4th edn)
   Clocksin W F & Mellish C S
   Springer, 1995

SC-202 : OPERATING SYSTEMS

1. Introduction to UNIX. **[10 hrs]**

Evolution of UNIX: Past, Present and future, Philosophy of UNIX: System’s Internal Structure, the process interface, OS features, OS systems Concepts: File Systems, Processes. Deviation from Unix: The GNU project.

2. Implementation of buffer cache. **[4 hrs]**

Structure and Philosophy of the cache implementation. Algorithms used by the buffer cache. Also, touch upon inode cache

3. File system. **[10 hrs]**

The file system switchers table, VFS architecture, File systems implementation on disk, File system handling kernel algorithms, Issues for file system handling, System calls for file system manipulations.

4. Process. **[7 hrs]**

State transitions, Process structure and layout, Multiprocessing details-Context, context switches, memory management concepts, System calls.

5. Process Scheduler. **[7 hrs]**

Class specific implementation, Priorities, system calls.

6. Memory Management Techniques. **[3 hrs]**

Swapping, Demand paging, Hybrid, Virtual Memory.

7. Time and Clock. **[3 hrs]**

8. I/O Subsystems. **[5 hrs]**

Concepts, data structures, device drivers, streams.

9. Interprocess Communication, and thread communication **[5 hrs]**

Threads- Thread creation. Inter thread control, Thread synchronization, Inter process communication.

ASSIGNMENTS:

1. Programming using Unix system calls
2. File system & directories, File operations.
3. Programming using Unix system calls
4. Process system, using commands like strace, top, lsof, ps.

Text Reference:

1. Design of the Unix System,
   M. J. Bach
   Prentice Hall
2. Operating System Concepts
   J. L. Peterson, A. Silberschatz, Galvin
   (Addision Wesley)
3. Modern Operating System
   Andrew Tanenbaum
   Pearson Education.
4. Advanced Concepts in Operating System
   Mukesh Singhal, Niranjan Shivrathri
5. Operating System Internals and Design Principles
   William Stallings

SC-204 : NUMERICAL METHODS FOR SCIENTIFIC COMPUTING-I

1. Number System and Errors **[4 hrs]**

Representation on integers and floating point numbers, Errors in computation, loss of significance.

2. Solutions of Equations in one variable **[6 hrs]**

Bisection Method, Newton Raphson Method, Fixed Point iteration, Error Analysis, Accelerating Convergence, Polynomial Evaluation – Horner's rule, Zeros of polynomials and Muller's Method

3. Systems of Linear Equations **[8 hrs]**

Gaussian Elimination, Triangular decomposition, Pivoting strategies, Error analysis and Operations count, Ill-conditioning and condition number of system, Evaluation of determinants

4. Eigenvalue Computations **[11 hrs]**

Diagonalization of system of ODE, Power Method, Gerschgorin theorem, Jacobi's Method, Given's and Householder's methods for Tridiagonalization, Method of sturm sequences for tridiagonal matrix, Lanczos Method, QR Factorization

5. Curve fitting and Approximation **[12 hrs]**

Lagrange's interpolation, Polynomial wiggle problem, Spline interpolation, Least Square Method – line and other curves, Orthogonal Polynomials, Tchebyshev interpolation, Fourier approximation and Fast Fourier, Transforms (FFT) algorithm.

6. Numerical Differentiation and Integration **[9 hrs]**

Numerical Differentiation – Richardson Extrapolation method, Numerical Integration – Newton Cotes quadrature for equidistant points, Gaussian Quadrature, Integration using Tchebyshev Polynomials

ASSIGNMENT:

1. Solutions of nonlinear equations
2. Fixed point iterations
3. Linear equation solvers
4. Eigenvalue computations
5. Curve fitting, interpolation
6. Numerical IntegrationText References:

1. Numerical Methods for Mathematics, Science and Engineering
   John H. Mathews.
2. Numerical Analysis (7th Edition)
   Richard and J. Douglas Faires
3. Numerical Analysis
   C. E. Froberg
4. Numerical Analysis – A practical Approach
   Maron M.J.
5. A First Course in Numerical Analysis
   Ralston and Rabinowitz.

SC-205 : COMPUTATIONAL LABORATORY-II

A system development project should be taken and following should be done

(a) Analysis of the project
(b) Design of the project (using any design model)
(c) Coding

Experts from industry will guide these projects, which will be based on current technologies.

SC-301 : NETWORK CONCEPTS

PART I:

1. Introduction to networks **[1 hr]**

OSI Layer model (Each layer with example)

2. ARP: Address Resolution Protocol **[2 hrs]**

Introduction, An Example, ARP Cache, ARP Packet Format, ARP Examples, Proxy ARP, Gratuitous ARP, ARP Command, Summary

3. RARP: Reverse Address Resolution Protocol [2 hrs]
Introduction, RARP Packet Format, RARP Examples, RARP Server design, Summary

4. ICMP: Internet Control Message Protocol [3 hrs]
Introduction, ICMP Message Types, ICMP Address Mask Request and Reply, ICMP Timestamp Request and Reply, ICMP Port Unreachable Error, Examples Ping, Traceroute

5. UDP: User Datagram Protocol [3 hrs]
Introduction, UDP Header, UDP Checksum, A Simple Example, IP Fragmentation, ICMP Unreachable Error (Fragmentation Required), Determining the Path MTU Using Traceroute, Path MTU Discovery with UDP, Interaction between UDP and ARP, Maximum UDP Datagram Size, ICMP Source Quench Error, UDP Server Design, Introduction to DNS, IGMP

6. Broadcasting and Multicasting [2 hrs]
Introduction, Broadcasting, Broadcasting Examples, Multicasting, Summary

7. TFTP: Trivial File Transfer Protocol [2 hrs]
Introduction, Protocol, An Example, Security, Summary

8. BOOTP: Bootstrap Protocol [2 hrs]
Introduction, BOOTP Packet Format, An Example, BOOTP Server Design, BOOTP through a Router, Vendor-Specific Information, Summary

9. TCP: Transmission Control Protocol [2 hrs]
Introduction, TCP Services, TCP Header, Summary

10. TCP Connection Establishment and Termination [2 hrs]
Introduction, Connection Establishment and Termination, Timeout of Connection Establishment, Maximum Segment Size, TCP Half-Close, TCP State Transition Diagram, Reset Segments, Simultaneous Open, Simultaneous Close, TCP Options, TCP Server Design, Summary

11. TCP Interactive Data Flow [2 hrs]
Introduction, Interactive Input, Delayed Acknowledgements, Nagle Algorithm, Windows Size Advertisements, Summary

12. TCP Bulk Data Flow [2 hrs]
Introduction, Normal Data Flow, Sliding Windows, Window Size, PUSH Flag, Slow Start, Bulk Data Throughput, Urgent Mode, Summary

13. TCP Timeout and Retransmission [3 hrs]
Introduction, Simple Timeout and Retransmission Example, Round-Trip Time Measurement, An RTT Example, Congestion Example, Congestion Avoidance Algorithm, Fast Retransmit and Fast Recovery Algorithm, Congestion Example (Continued), Per-Route Metrics, ICMP Errors, Repacketization, Summary

14. TCP Persist Timer [2 hrs]
Introduction, An Example, Silly Windows Syndrome, Summary

15. TCP Keepalive Timer [2 hrs]
Introduction, Description, Keepalive Examples, Summary

16. TCP Futures and Performance [3 hrs]
Introduction, Path MTU Discovery, Long Fat Pipes, Windows Scale Option, Timestamp Option, PAWS: Protection Against Wrapped Sequence Numbers, T/TCP: A TCP, Extension for Transactions, TCP Performance, Summary

17. Introduction to Network security [1 hr]
Firewall, VPN (IPSec), IDS/IPS

PART II:

1. Switching [2 hrs]

2. Network Architecture [2 hrs]
Layers, Address scheme

3. LAN technologies **[2 hrs]**
Ethernet , Token ring, FDDI

4. Transparent Bridges **[2 hrs]**
Principles of operations, Bridge Architecture, Bridge Address Table

5. Principles of LAN Switches **[3 hrs]**
Switched LAN concept , Cut-Through Vs Store-Forward operations, L3 switches, Switch Configuration(bounded/stackable chassis), Switch Application Environment

6. L3 Concepts **[2 hrs]**
CIDR, RIP V1, Intro to RIP V2, Intro to OSPF

7. Loop resolution **[2 hrs]**
STP, Intro to RSTP, Intro to MSTP

8. Advanced LAN switch concepts **[2 hrs]**
Full duplex operations, Flow control, Link-Aggregations, VLAN Application and concepts, VLAN IEEE standards

9. Switch Management **[2 hrs]**
SNMP concepts, RMON concepts

10. Advanced L3 concepts **[3 hrs]**

ASSIGNMENTS:
1. Open raw socket and read packets, display packet type. (Packet sniffer)
2. TFTP client/server using rfc 1350
3. Upgrade above UDP server to handle multiple client requests
4. Write file-server which process request from every client through same server port.
5. Write client/server using Open-SSL library
6. If time permitted extend the above library to write telnet client/server
7. Mini project: One of the above topics from chapter 17 can be taken up as project.
8. Implement BAT
9. Implement CAM (content addressable memory)
10. Study CLI command for some switch
11. Develop a prototype for the above
12. Implement STP algorithm in software
13. Involve relevant theory from Math (like Dijkstra, Shortest Path Algorithm, etc)
14. Mini Project (L2 Bridge) can be done using above assignment.

Text Reference:

1. Data and Computer Communication
   Williams Stallings
2. TCP / IP Illustrated Vol . I / II / III
   Richard Stevens
3. The Switch Book: The Complete Guide to LAN Switching Technology
   Rich Seifert
4. Computer Networks
   Andrew S. Tanenbaum
5. Data Communication and Computer Networks
   Prakash C. Gupta

SC-302 : PARALLEL PROCESSING AND GRID COMPUTING

1. Introduction **[3 hrs]**

Need for high-speed computing, need of parallel computers, Features of parallel computers, hardware requirements.

2. Solving Problem in parallel **[2 hrs]**

Temporal parallelism, data parallelism, comparison of temporal and data parallelism with specialized processors intertask dependency.

3. Structure of parallel computers **[5 hrs]**

Pipelined parallelism computers, array processors, a generalized structure of parallel computer, shared memory multiprocessors, message passing multicomputers multilink multidimensional computing system.

4. Programming parallel computers **[14 hrs]**

Programming message-passing multicomputers, programming shared memory parallel computer, programming vector computers.

5. Case Studies **[4 hrs]**

Matrix Multiplication, Graph theory, N-Body Simulation, Computer Vision and Image Processing applications

6. Grid Computing **[12 hrs]**

Introduction, grid architecture, resource sharing and allocation, job scheduling, grid security, Globus, Open grid services architecture (OGSA)

ASSIGNMENTS:

1. Matrix multiplication
2. Computing value of Pi
3. Solution to linear equation using Jacobi method
4. Finding patterns in a text file
5. Producer-consumer problem
6. All-pair shortest paths problem
7. N-body simulation

Text Reference:

1. Introduction to Parallel Computing, 2nd Edition
   Ananth Grama, Anshul Gupta, George Karypis, and Vipin Kumar
   Pearson Education Ltd., 2004
2. Parallel Computers-Architecture and Programming
   V. Rajaraman and C. Siva Ram Murthy
   Prentice-Hall (India), 2000
3. Designing and Building Parallel Programs
   Ian Foster
   Addison-Wesley Inc., 1995
4. Elements of Parallel Computing
   V. Rajaraman
   Prentice Hall, 1990
5. Introduction to Grid Computing
   Bart Jacob, Michael Brown, Kentaro Fukui, and Nihar Trivedi
   IBM Redbook, 2005

SC-304 : NUMERICAL METHODS FOR SCIENTIFIC COMPUTING-II.

1. Numerical Solutions of ODE **[18 hrs]**

Single step and Multistep methods Predictor-Corrector Methods. Boundary value problems stiff Equations.

2. Numerical Solutions of PDE **[12 hrs]**

Elliptic Equation, Parabolic Equation, Hyperbolic Equation

3. Numerical Optimization **[15 hrs]**

Linear Optimization: Simplex method, Transportation Problem.

Nonlinear Optimization: Golden ratio method, Nelder Mead method, Conjugate gradient method.

ASSIGNMENTS:
1. ODE solvers – Euler, Runge Kutta, Multistep methods
2. PDE solvers – Finite difference methods
3. Simple method and transportation
4. Conjugate gradient method, Golden search method

Text References:

1. Numerical Methods for Mathematics Science & Engineering. (Second Edition)
   John H. Mathews
   Prentice Hall of India.
2. Numerical Analysis (Seventh Edition)
   Burden and Faires
   Thomson Asia PTE. LTD
3. Numerical Analysis- A practical Approach
   Maron
   Mc Millan,1982
4. Numerical Analysis
   Forberg
   McGraw Hill,1979
5. A First Course in the Numerical Analysis of Differentials Equations
   Arieh Iserles
   The Press Syndicate of the University of Cambridge.
6. Numerical Solution of Partial Differentials Equations (Third Edition)
   G.D.Smith
   Oxford University Press
7. Numerical Methods for Scientists & Engineers
   H.M.Antia
   THM,1991

EL-1: APPLICATIONS OF COMPUTERS TO CHEMISTRY.

1. Computational Chemistry. [1 hrs]

Why learn Computational Chemistry? Applications areas: fundamental understanding. Predictions, design, structure of biomolecules, polymer design, catalyst and drug design.

2. Fundamentals of Chemistry. [2 hrs]

Concepts in chemistry like valency, hybridization, electronegativity, covalent bond, ionic bond, hydrogen bond, co-ordinate bond. Geometries of molecules like linear, angular, tetrahedral, etc.

3. Molecular Representations and Search [4 hrs]

Connectivity matrix, Adjacency matrix, SMILES notation, substructure search

4. Molecular Graphics and fitting. [4 hrs]

3-dimensional structures, steric pictures, CPK models, molecular dimensions and van der Waals volume.

5. Force Field (FF) Methods [4 hrs]

Molecular mechanics expressions for bond stretch, bond angle, torsion, improper torsion, Van der Waals, electrostatics and cross terms. Types of force fields, computational aspects in FF, parameterization in FF. Evaluation of number of energy terms for a given molecule

6. Classical energy minimization techniques [3 hrs]

Energy minimization by simplex, steepest descent, conjugate gradient and Newton-Raphson methods.

7. Conformational Analysis **[4 hrs]**

What is a conformation? Systematic, Monte Carlo and genetic algorithm based conformational analysis. Polling method of conformational analysis, simulated annealing method

8. Semi-empirical QM calculations. **[6 hrs]**

Cluster model calculations for the electronic structure of extended systems, Prospect and pitfalls in the usage.

9. Molecular Docking **[4 hrs]**

Concepts in docking. Parameterization in docking, Rigid docking, flexible docking, virtual screening, Scoring functions

10. Molecular Descriptors **[4 hrs]**

Molecular connectivity indices, topological indices, electro-topological indices, information theory indices, etc.

11. Quantitative Structure Activity Relationship **[6 hrs]**

Generation of training and test set methods, variable selection methods like stepwise forward, backward, etc. Regression methods like multiple regression, principal component regression.

12. Futuristic modeling techniques. **[2 hrs]**

Expert systems, Neural networks, Artificial Intelligence and virtual reality.

Text Reference:

1. An Introduction to Chemoinformatics,
A.R. Leach; V.J. Gillet,
Kluwer Academic Publishers, The Netherlands, 2003.
2. Introduction to Computational Chemistry
Frank Jensen
Chichester, Wiley, 2006.
3. Molecular Modeling: Principals and Applications
A.R. Leach
Pearson Education Limited, Essex, 2001
4. Essentials of Computational Chemistry: theories and models
Christopher J. Cramer
John Wiley, 2004
5. Optimization in Computational Chemistry and Molecular Biology: local and global Approaches
M. Panos, Christodoulos A. Pardalos
Floudas - Science – 2000
6. Chemoinformatics: A Textbook
J. (Johann) Gasteiger
Thomas Engel – 2003

EL-2: SCIENTIFIC VISUALIZATION

1. Introduction to Computer Graphics. **[5 hrs]**

Examples of Graphics applications. Key journals in Graphics input and output graphics devices, world coordinate systems viewports and world to viewport mapping.

2. Raster graphics technique **[5 hrs]**

Line drawing algorithms scan converting circles and ellipses, polygon filling with solid colors and filling patterns, half toning and dithering techniques.

3. Vectors and their use in graphics **[2 hrs]**

Operation with vectors, adding, scaling, subtracting, linear spaces, dot product, cross Product, Scalar triple product, application of dot product, cross product and scalar triple product, application of vectors to polygons.

4. Transformation of pictures **[3 hrs]**

2D Affine transformations, use of Homogenous coordinates, 3D affine transformations.

5. 3D viewing **[6 hrs]**

Synthetic camera approach, describing objects in view coordinates, perspective and parallel Projections, 3D clipping.

6. 3D graphics, Write frame models **[5 hrs]**

Marching cube algorithms for contour generation and surface generation from a given data over 2D/3D grid.

7. Hidden line and surface removal, backface culling **[7 hrs]**

8. Light and shading models, rendering polygonal masks Flat, Gouraud, phone shading. **[8 hrs]**

9. Ray Tracing **[11 hrs]**

Overview, intersecting ray with plane, square, cylinder, cone, cube and sphere, drawing shaded pictures of scenes, reflections and transparency.

ASSIGNMENTS:

1. 2D transformations
2. 3D transformations
3. Projections, Clipping, Shading
4. Contour, Marching cube
5. Creating a graphics package using above assignments.

Text References:

1. Computer Graphics: A programming approach,
   S. Harrington
   (McGraw Hill, 1986)
2. Computer Graphics – Principles and Practice
   Foley, Van Dam
3. Computer Graphics,
   F. S. Hill Jr.
   (Macmillan, 1990)
4. Procedural Elements for Computer Graphics.
   D.F. Rogers
   (McGraw Hill, 1995)
5. Principles of Interactive Computer Graphics
   William Newman, Robert Sproull

EL-3: STATISTICAL COMPUTING

1. Introduction **[3 hrs]**

Computation of averages and measures of central tendency, skewness and kurtosis. Preparation of frequency tables, computation of Pearson and rank correlation coefficients.

2. Statistical Distributions **[6 hrs]**

Evaluation of standard probability mass functions, cumulative density functions, and quantile functions.

3. Monte Carlo **[4 hrs]**
Random number generation including Markov Chain Monte Carlo (MCMC). Statistical simulation and Monte Carlo studies in statistics.
4. Numerical optimization and root finding methods, including the Newton-Raphson method and the EM algorithm. **[5 hrs]**
5. Resampling techniques, including the permutation test and the bootstrap. **[5 hrs]**
6. Classification using discriminant functions, rough sets, artificial neural networks, and decision trees. **[12 hrs]**
7. Clustering techniques **[10 hrs]**
8. Multiple regression Analysis **[6 hrs]**

Text References:

1. Simulation Modeling and analysis
   Averill M. Law and W. David Kelton
2. Applied Multivariate Statistical Analysis
   Richard A. Johnson and Dean W. Wichern
3. Data Mining: Concepts and Techniques
   Jiawei Han and Micheline Kamber
4. Data Mining : Multimedia, Soft Computing, Bioinformatics
   Sushmita Mitra and Tinku Acharya
5. Computational Statistics
   Geof H. Given and Jennifer A. Hoeting

EL4 : Applications of Computer to Physics

1. Introduction to Mathematical modeling

Any two of the following four topics :

2. Modeling Projectile Motion
Model trajectories using Euler's method for ODE's and Runge-Kutta Method.
Problems related with trajectory of cannon ball, motion of batted ball, effect of air resistance, drag coefficient on the trajectories.

3. Oscillatory Motion
Simple Harmonic Motion, Driven non-linear pendulum, chaos in non-linear pendulum

4. Solar System
Planetary Motion, inverse square law and stability of planetary orbits.
Three body problem and its effect on one planet

5. Waves
Waves on a string with free ends/ without free ends, frequency spectrum of waves on a string, vibrating strings with frictional loss

Any two of the following five topics :

6. Random Systems
Random number generation, distribution functions, Monte Carlo methods, sampling, integration, Metropolis algorithm

Random walks, self-avoiding walks, random walks and diffusion, cluster growth model, fractals, percolation Radioactive decay, discrete decay, continuous decay, decay simulation using Monte-Carlo

7. Molecular Dynamics
Dilute gas with L-J potential, Methods for many body systems, Verlet algorithm, structure of atomic clusters, elementary ideas of ensembles, constant pressure, constant temperature, simulated annealing, melting

8. Dynamical Systems
Linear, nonlinear models, Nonlinear growth - Logistic map, fixed points, period doublings, attractors, bifurcation diagram, Figenbaum constant,Henon Map, Lorenz Map, Chaos

9. Fractals
Self-similarity, fractal dimension, L-systems, Self affine fractals,diffusion limited aggregation, cellular automata and its applications.

10. Ising Model
Ising Chain -- analytic solution, numerical solution, approach to thermal equilibrium, beyond nearest neighbors.

Text References:

1. Computational Physics,
   Nicholas J. Giordano
   Prentice Hall (1997)
2. An Introduction to Computational Physics
   Tao Pang
   Cambridge University Press, (2006)
3. Projects in Scientific Computation
   Richard E. Crandall
   Springer (2000)
4. Introductory Statistics and Random Phenomena
   Manfred Denker, W. A. Woyczynski
   Birkhauser (2005)
5. Computational Physics
   Steven Koonin and Dawn Meredith
   Westview Press (1998)
6. Computational Physics
   J. M. Thissen
   Cambridge University Press (1999)
7. Computational Physics: Problem Solving with Computers
   Rubin. H. Landau, Manuel J. Paez, Cristian C Bordeianu
   Wiley-VCH (2007)

Assignments

1. Generate and plot trajectory of Cannon ball with and without air drag.
2. Effect of spin on trajectory of a thrown ball
3. Use Euler method for damped, nonlinear, driven pendulum. Plot trajectories.
4. Period-doubling and route to chaos in pendulum
5. Phase space plot of Lorenz model
6. Planetary trajectories with Euler-Cromer method.
7. Elliptical orbits and stability of orbits
8. Orbits of two planets under the influence of third planet
9. Signal from vibrating string.
10. Spectrum analysis of waves on string (fixed and free ends)
11. Random walk in one dimension
12. Two dimensional self-avoiding walk
13. Diffusion equation in two or three dimension
14. Diffusion limited aggregation in clusters with fractal dimensionality
15. Generate fractal curves recursively and obtain their dimensionality
16. Calculate M, E for the Ising model
17. Using Lennard-Jones potential, simulate a dilute gas, cluster of 10-20atoms.
18. Simulate cluster of atoms at a particular temperature. Plot the motion of a test particle, temperature of system as a function of time, pair separation and mean square displacement as a function of time.

EL-5. BIOLOGICAL SEQUENCE ANALYSIS

1. Analysis of DNA and Protein sequence-distribution, frequency statistics, pattern and motif searches-randomization etc.-sequence segmentation. **[10 hrs]**

2. Sequence alignment – scoring matrices- PAM and BLOSUM-Local and global alignment concepts- dynamic programming methodology- Needleman and Wunsch algorithm, Smith –Waterman algorithm –statistics of alignment score-Multiple sequence alignment –Progressive Alignment. Database searches for homologous sequences – Fasta and Blast versions. **[15 hrs]**

3. Fragment assembly, Genome sequence assembly – Gene finding methods:- content and signal methods- background of transform techniques – Fourier transformation and gene prediction – analysis and predictions of regulatory regions. **[10 hrs]**

4. Neural network concepts and secondary structure prediction Probabilistic models: Markov chain-random walk-Hidden Markov models. Gene identification and other applications. **[10 hrs]**

5. Evolutionary analysis: distances-clustering methods – rooted and unrooted tree representation – Bootstrapping strategies **[5 hrs]**

Text References:

1. Bioinformatics: a practical guide to the analysis of genes and proteins.
   A. Baxevanis and F.B.F. Ouellette(Eds.) John Wiley, New York (1998).
2. Introduction to computational biology: maps, sequences, and genomes.
   M.S. Waterman, Chapman and Hall, London (1995).
3. Proteome research: new frontiers in functional genomic
   M.R. Wilkins, K.L. Williams, R.D. Appel and D.H. Hochstrasser (Eds.), Springer, Berlin (1997).

Interdisciplinary School of Scientific Computing

EL-6. MODELLING OF BIOLOGICAL SYSTEM.

1. Concepts and principles of modeling. Limitations of models. [20 hrs]

Identifying the components of a process / system – variables, controlled variables, rate constants, relationships between variables. Writing a set of equations, describing a process or a system. Types of solutions – integration, equilibrium solutions, numerical solutions. Models involving space: spatial simulations.

2. Models of behavior. [15 hrs]

Foraging theory, Decision making – dynamic models, Game theory.

3. Modeling in Epidemiology and Public Health SIR models; Stochastic models and Spatial models. **[10 hrs]**

Text Reference:

1. Models in Biology.
   B. Brown and P. Rothary , John Wiley and Sons, New York.
2. Evolutionary Genetics
   J. M. Smith,. Oxford University Press, Oxford (1989).

EL - 7 Artificial Intelligence

1. Conventional AI - Reasoning and Belief Systems

a) Logical Inference **[3 hrs]**

Reasoning Patterns in Propositional Logic, Propositional inference, Predicate calculus, Predicate and arguments, ISA hierarchy, Frame notation, resolution, Natural deduction etc.

b) Reasoning under Uncertainty **[3 hrs]**

Belief and uncertainty handling mechanisms, certainty, possibility and probability, Dempster Schaeffer theory , fuzzy inference, structure knowledge representation, semantic net, Frames, Script, Conceptual dependency etc.

c) Goal Driven Intelligence(Planning, Search and Perception) **[6 hrs]**

i) Planning: Formulation of Planning Problem, decomposition, representation of states, goals and actions, action schema, partial order planning, planning graphs Block world, strips, Implementation using goal stack, Non linear Planning with goal stacks, Hierarchical planning, List Commitment strategy.

ii) Game Playing and Search: Heuristic search techniques. Best first search, mean and end analysis, A* and AO* Algorithm, Minimize search procedure, Alpha beta cutoffs, waiting for Quiescence, Secondary search, ***Perception* -** Action, Robot Architecture, Vision, Texture and images, representing and recognizing scenes, waltz algorithm, Constraint determination, Trihedral and non trihedral figures labeling

d)Expert systems **[3 hrs]**

Utilization and functionality, architecture of expert system, knowledge and rule bases, rule chaining strategies, conflict resolution, RETE algorithm, uncertainty handling in expert systems

2. Intelligent Agents and Computational Intelligence

a) Agent Oriented Programming and Intelligent agents **[3 hrs]**

Agent oriented programming as a paradigm, Agent orientation vs. object orientation, autonomous and intelligent agents, “Agency” and Intelligence, logical agents, multi agent systems, planning, search and cooperation using agents.

b) Evolutionary Algorithms **[3 hrs]**

Evolutionary paradigms, genetic algorithms and genetic programming, Ant colonies and optimization, evolutionary search strategies.

c) Agents , internet and Softbots **[3 hrs]**

Interface agents and reactive systems , Softbots and info agents, the three layer model , process automation and agents,

d) AI paradigms from biological, physical and social sciences **[3 hrs]**

Swarms and collective intelligence, programming with swarms, fault tolerant systems, spin glasses and neural networks, self organizing systems, cellular automata and amorphous computing.

3. Statistical Learning Theory

a) Learning Theory **[3 hrs]**

Formulation of learning as a statistical problem, estimation of probability measure, empirical and structural risk minimization, Linear methods, supervised and unsupervised learning, regularization and kernel methods, model selection, inference and averaging, boosting and additive methods

b) Applications and algorithms **[3 hrs]**

Perceptrons and Neural Networks, Support vector machines, Classification and regression trees, nearest neighbors and EM clustering, Kohonen maps.

c) Text Mining and Natural language processing **[2 hrs]**

Sentence, syntax and semantic analysis, document classification, sentiment perception, theme and association mining

4. Hybrid Systems **[6 hrs]**

Integration of data driven and concept driven methodologies, Neural Networks and Expert Systems hybrid, neural networks and game tree search hybrid, evolutionary systems and supervised learning hybrid, neuro-fuzzy systems, genetic programming for rule induction

Text References:

1. AI: a modern approach
   Russell and Norvig:
2. AI
   Winston
3. Mathematical Methods in Artificial Intelligence
   Bender
4. Reasoning about Intelligent Agents
   Woolbridge
5. Artificial Intelligence.
   Elaine Rich and Kerin Knight:
6. Artificial Neural Network
   Kishen Mehrotra, Sanjay Rawika, K Mohan;.

EL-8 Quality Assurance and Software Testing

1. Introduction to Software Quality Management Principles **[2 hrs]**

2. Software Quality Assurance & Quality Control **[3 hrs]**

What and how Software Quality, Quality Goals for Software, Process Quality Goal, Product Quality Goal, Quality policy and Quality Objectives: Linkage and control

3. Quality Models **[3 hrs]**

ISO 9001 – 2000 model, CMMI models, IS27000

4. Software Verification, Validation & Testing (VV & T) **[3 hrs]**

Understanding verification, validation & testing, Quality improvement through activities (like Reviews, Inspection, Walkthroughs, Testing), Process improvement through VV & T

5. Software Testing Principles and Concepts **[5 hrs]**

Software Testing Vocabulary, Testing and Quality, Who should Test, Independence in Testing, When Should Testing start?, Static versus Dynamic Testing, Testing and Debugging, The Cost of Quality, General Testing Process - Test Planning, Test Case Design, Test Case Execution, Test Analysis and Defect Reporting, Test Closure

6. Life Cycle Testing **[5 hrs]**

Various Software Development Models, Levels of Testing (Unit Testing, Integration Testing, System Testing, User Acceptance Testing), OO-oriented Testing, Model Based Testing, The “V” Model of Testing, Early Testing, Verification and Validation, Retesting and Regression Testing

7. Testing Techniques **[6 hrs]**

Static Testing Techniques (Reviews, Informal review, Walkthrough, Technical review, Inspection), Functional /Specification based Testing Techniques, Structural Testing Techniques, Experienced based techniques, Choosing test techniques, Test Oracle, Building Test Cases, Process for Building Test Cases, Test Case Execution, Recording Test Results, Problem Deviation, Problem Effect, Problem Cause, Use of Test Results

8. Test Reporting Process **[6 hrs]**

Prerequisites to Test Reporting, Define and Collect Test Status Data, Define Test Metrics used in Reporting, Define Effective Test Metrics, Test Tools used to Build Test Reports, Pareto charts, Cause and Effect Diagrams, Check Sheets, Histograms, Run Charts, Scatter Plot Diagrams, Regression Analysis, Multivariate Analysis, Control Charts, Test Tools used to Enhance Test Reporting, Benchmarking, Quality Function Deployment, Reporting Test Results, Current Status Test Reports, Final Test Reports, Guidelines for Report Writing

9. Test Management **[5 hrs]**

Testing in an Organization, Test management documentation, Test plan documentation, Test estimation and scheduling of Test Planning, Analyzing Testing metrics, Test progress monitoring and control, Testing and risk, Risk management, Software Configuration Management, Change Management

10. Test tools **[3 hrs]**

Types of Testing Tools and their use, Tool selection and implementation

Text References:

1. Software Testing Techniques: Finding the Defects that Matter (Programming Series)
   Michael Shannon, Geoffrey Miller, Richard, Jr. Prewitt,
2. Software Testing Fundamentals: Methods and Metrics
   M. Hutcheson
3. "Software Testing: A Craftsman's Approach, Second Edition,"
   Paul C Jorgensen,
   CRC Press, June 26, 2002.
4. "The Art of Software Testing," 2nd ed.,
   Glenford J. Myers,
   John Wiley & Sons, Inc.
5. "Lessons Learned in Software Testing: a Context-Driven Approach,"
   Cem Kaner, James Bach, and Bret Pettichord
   John Wiley & Sons, Inc.

EL-9 Soft Computing

1. Neural Networks **[15 hrs]**

a) Characterization
b) The brain, neural networks and computers
c) Neural networks and artificial intelligence
Background, Applications, Neural network software, Learning paradigms - Supervised learning, Unsupervised learning, Reinforcement learning, Learning algorithms
d) Neural networks and neuroscience
Types of models, Current research
e) History of the neural network analogy

2. Fuzzy Systems **[15 hrs]**

a) Antilock brakes
b) Fuzzy sets
Fuzzy control in detail, Building a fuzzy controller
c) History & applications
d) Logical interpretation of Fuzzy control

3. Evolutionary Computing **[15 hrs]**

a) Concept of Population genetics, probability, evolution principle
b) Genetic Algorithms or Evolutionary Strategies
i) Genetic Algorithms - General mechanism and terminologies, Selection, Crossover, Mutation
ii) Evolutionary Strategies - Two-membered Evolutionary strategy, Multi-member Evolutionary strategy, Recombination
c) Swarm Intelligence (one of the following three)

Ant colony optimization, Particle swarm optimization, Stochastic diffusion search Applications

ASSIGNMENT:

To implement at least two of these major computational methods.

Text Reference:

1. Neural Networks, A Classroom Approach
   Satish Kumar
   Tata McGraw-Hill Publishing Company Limited
2. Artificial Neural Networks
   Kishan Mehrotra, Chilkuri K. Mohan, Sanjay Ranka
   Penram International Publishing (India)
3. Neural Networks, A Comprehensive Foundation
   Simon Haykin
   Pearson Education
4. Genetic Algorithms, in Search, Optimization & Machine Learning
   David E. Goldberg
   Pearson Education
5. Artificial Intelligence and Intelligent Systems
   N P Padhy
   OXFORD University Press

EL 10- Design concepts and modeling

1. Introduction to design process. **[6 hrs]**

Building models suitable for the stages of a software development project. Introduction to UML.

2. Inception phase. **[8 hrs]**

Structured analysis, scenario structures.

3. Elaboration phase. **[10 hrs]**

Object modeling. Interfaces and abstraction. Information hiding.

4. Construction phase. **[10 hrs]**

Coupling and object interaction. Responsibilities, defensive programming and exceptions. Functional decomposition, module and code layout. Variable roles, object state, verification and assertions. Design patterns.

5. Transition phase. **[10 hrs]**

Inspections, walkthroughs, testing, debugging. Iterative development, prototyping and refactoring. Optimization.

Text Reference:

1. Code complete: a practical handbook of software construction.
   McConnell, S. (1993)
   Microsoft Press.
2. UML distilled. Addison-Wesley (2nd ed.).
   Fowler, M. (2000).

EL 11- Business Analysis

Introduction to DWH and OLAP **[8 hrs]**

A. Decision Support System: Introduction to Decision Support System (DSS), DSS Components, Decision Types; Data warehouse (DWH): Need, Definition, Advantages of DWH, OLTP Vs DWH, 3-tier Architecture, DWH Design Process, ETL Process, DWH Back-end Tools and Utilities, Metadata Repository, Models of DWH: Enterprise Warehouse, Data Mart, Virtual Warehouse, Comparison; OLAP: Data Cube and OLAP, Concept Hierarchies, OLAP Operations: Drill-Down, Roll-Up and Extreme Roll-Up, Slice-Dice and Pivot, OLAP Types, OLAP Query Processing, Computation of Data Cube.
B. ETL Tools, Commercial DWH Vendors/ Tools and their Comparison, Project Failure Reasons, Data Analytics, Business Intelligence, SAS Software.

Dimensional Modeling **[8 hrs]**

A. Dimensional Modeling: Dimensional Model Vs ER Model, DWH Schemas: Star, Snowflake, Fact Constellation, their Comparison, Techniques to Handle Changing Dimensions, Aggregation, Families of Fact Tables, Fact Less Fact Tables; Data Warehouse Indexing: Factors used to select an Indexing Technique, Properties of a Good Indexing Technique for DWH, Indexing Techniques: Projection Index, Bitmap Index (Pure and Encoded), Join Index and their Comparison.
B. Case Studies of Data Warehouse Applications in various Industry Segments.

Data Mining and Functionalities **[8 hrs]**

A. Introduction: Need of Data Mining, Knowledge Discovery in Database (KDD), Architecture of Data Mining System, Data Mining on Different kind of Data, Data Mining Functionalities; Data Preprocessing: Need, Cleaning, Integration, Transformation, Reduction, Discretization, Concept Hierarchy Generation; Cluster Analysis: Categories of Clustering methods, Partitioning methods: k-Means, k-Medoids ; Prediction: Numerical Prediction, Linear, Non-Linear Regression; Outlier Analysis: Applications, Techniques.

B. Data Mining Task Primitives, Query Language, System Classification, Data Mining Issues.

Classification **[9 hrs]**

A. Classification: Decision Tree Classifier, Rule Based Classification, Bayesian Classification, Neural Network Classification: Back Propagation Algorithm, Lazy Learner: kNN Classifier, Case-Based Reasoning, Other: Fuzzy Set Approach, Classifier Accuracy Measures, Techniques for Evaluating Classifier Accuracy; Frequent Itemset Mining: Interesting Item Set Mining: Market Basket Analysis, APriori Algorithm, Generating Association Rules, Types of Association Rules, Correlation Analysis.
B. Support Vector Machine, Associative Classification, other Classification Techniques: Genetic Algorithm, Rough Set, Constraints Based Association Mining.

Data Mining on different Databases **[7 hrs]**

A. Multimedia Data Mining, Web Mining, Text Mining, Spatial Data Mining, Mining on Social Networks, Multirelational Data Mining.
B. Data Mining Applications, Trends/ Challenges of Data Mining, Mining Sequence Patterns in Transactional Database, Graph Mining, Data Mining Tools- Dbminer/ WEKA/ Oracle DM Tools/ OLE DB/ Ida.

Text Reference:

1. The Data Warehouse Lifecycle Toolkit
   Kimball, Reeves, Ross, Thornthwaite
   John Wiley
2. Data Mining: Concepts and Techniques
   Jiawei Han and Micheline Kamber,
   Morgan Kaufman,

EL 12- Machine Learning

Unit-1: **[4 hrs]**

Meaning and need for Machine Learning (ML). Types of Learning – predictive or supervised
learning, descriptive or unsupervised learning, active and passive leranings, online and batch learning.
(most of the course is on supervised batch learning with passive learner.) Broad templates of ML applications – Learning associations, classification, regression, clustering and reinforcement learning. Relationship between ML and other data sciences.

Unit-2: **[4 hrs]**

A formal learning framework – Domain set, Label set, Training and test data, Learner's output, Data generation model, measures of success, Empirical risk and overfitting. Hypothesis classes.

Unit-3: **[15 hrs]**

Support Vector Machine (SVM) Learning – Definition, maximal-margin classifier, soft-margin classifier,. SVM kernels – Linear kernel SVM, Polynomial kernel SVM, Radial kernel SVM, Data preparation for SVM. SVM regression.

Unit-4: **[15 hrs]**

Artificial Neural Network (ANN) Learning – Perceptron learning, Learning boolean functions, XOR problem, Multilayer perceptron leraning, Back propogation algorithm and related concepts. Apllications to classification and non-linear regression.

Unit-5: **[4 hrs]**

Dimensionality reduction and application of rough set theory to classification and clustering.

Unit-6: **[4 hrs]**

Design and analysis of ML experiments, cross-validation and resampling methods.

Unit-7: **[6 hrs]**

Introduction to open source software such as Rosetta and Weka.

Text Reference:

1. Understanding Machine Language: From Theory to Algorithms
   Shai Shalev_Shwartz and Shai Ben-David (2014)
2. Introduction to Machine Language (Second Edition)
   Ethem Alpaydn (2010) MIT press
3. Course on Machine Learning
   Andrew Ng (Openclassroom.stanford.edu)

EL 13- Data Mining

Introduction: **[3 hrs]**

Need of Data Mining, Knowledge Discovery in Database (KDD), Architecture of Data Mining System, Data Mining on Different kind of Data, Data Mining Functionalities;

Data Preprocessing: **[2 hrs]**

Need, Cleaning, Integration, Transformation, Reduction, Discretization;

Classification: **[10 hrs]**

Decision Tree Classifier, Rule Based Classification, Bayesian Classification, Neural Network Classification Support Vector Machine: Back Propagation Algorithm, Lazy Learner: kNN Classifier, Classifier Accuracy Measures, Techniques for Evaluating Classifier Accuracy

Cluster Analysis: **[5 hrs]**

Categories of Clustering methods, Partitioning methods: k-Means, k-Medoids, Hierarchical Clustering, BIRCH, Chameleon

Outlier analysis: **[5 hrs]**

Proximity based approaches: Distance and Density based

Prediction: **[8 hrs]**

Numerical Prediction, Linear, Non-Linear Regression; Outlier Analysis: Applications, Techniques.

Frequent Itemset Mining: **[6 hrs]**

Interesting Item Set Mining: Market Basket Analysis, APriori Algorithm, Generating Association Rules, Types of Association Rules, Correlation Analysis.
Multimedia Data Mining, Web Mining, Text Mining, Multi-relational Data Mining
Mining graph and Social Networks, Mining data streams

Text Reference:

1. Data Mining: Concepts and Techniques (2nd Edition)
   Jiawei Han and Micheline Kamber
   Morgan Kaufman,