ANNA UNIVERSITY, CHENNAI
UNIVERSITY DEPARTMENTS
M.E. COMPUTER SCIENCE AND ENGINEERING
(SPECIALIZATION IN BIG DATA ANALYTICS)
REGULATIONS – 2015
CHOICE BASED CREDIT SYSTEM

PROGRAM EDUCATIONAL OBJECTIVES:

1. Prepare students to review and understand foundational Concepts in Computer Science and Engineering
2. Empower students to critically analyze current trends and learn future issues from a system perspective at multiple levels of detail and abstraction
3. Enable students to apply the interaction between theory and practice for problem solving based on case studies
4. Enable students to pursue lifelong multidisciplinary learning as professional engineers and scientists to effectively communicate technical information, function effectively on teams, and apply computer engineering solutions within a global, societal, and environmental context
5. Prepare students to critically analyze existing literature, identify the gaps in the existing literature, map the existing problems as Big Data and propose innovative and research oriented solutions.

PROGRAM OUTCOMES:
Students will be able to:

a. Apply mathematical foundations, algorithmic principles, and computer science theory in the modeling and design of computer based systems of varying complexity
b. Critically analyze a problem, identify, formulate and solve problems in the field of Computer Science and Engineering and Big Data Analytics considering current and future trends
c. Map Computer Science Engineering problems as Big Data problems and provide solutions for the same
d. Design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, ethical, health and safety, and sustainability in the field of computer engineering
e. Function effectively on teams to accomplish a common goal
f. Communicate effectively with a range of audiences and prepare technical documents and make effective oral presentations
g. Analyze the local and global impact of computing on individuals, organizations, and society by applying the concepts of Big Data Analytics of the domain
h. Recognize the need for and possess an ability to engage in lifelong learning continuing professional development
i. Use current techniques, skills, and tools necessary for computing practice and demonstrate advanced knowledge of Big Data concepts
j. Propose solutions for Big Data Analytics problems by handling the gaps in existing literature.

Mapping of Programme Educational Objectives with Programme Outcomes

A broad relation between the programme educational objective and the outcomes is given in the following table

PROGRAMME EDUCATIONAL OBJECTIVES	PROGRAMME OUTCOMES									
	POa	POb	POc	POd	POe	POf	POg	POh	POi	POj
1	√	√								
2	√	√	√	√			√	√	√	√
3	√	√	√	√		√	√	√	√	√
4			√	√	√	√		√	√	√
5		√			√	√	√	√		√

			POa	POb	POc	POd	POe	POf	POg	POh	POi	POj
YEAR 1	SEM 1	Advanced Mathematics for Computing	√	√								
		Advanced Data Structures and Algorithms	√	√		√					√	
		Advanced Software Engineering	√	√	√							
		Advances in Operating System	√	√	√	√						
		Multi-Core Computing	√	√		√						
		Advanced Data Structures and Algorithms Lab	√	√	√	√				√	√	
		Professional Practices					√	√	√			
	SEM 2	Big Data Mining and Analytics	√	√	√	√			√	√		√
		Machine Learning Techniques	√	√	√				√			√
		Cloud Computing Technologies	√	√	√				√			√
		Advanced Databases Management Systems	√	√	√				√			√
		Elective I										
		Elective II										
		Big Data Computing Laboratory	√	√	√	√				√	√	√
YEAR 2	SEM 3	Big Data Security	√	√	√					√		√
		Elective III Elective IV										
		Elective V										
		Project Work Phase I	√	√		√		√			√	√
	SEM 4	Project Work Phase II	√	√		√		√			√	√

ANNA UNIVERSITY, CHENNAI
UNIVERSITY DEPARTMENTS
M.E. COMPUTER SCIENCE AND ENGINEERING
(SPECIALIZATION IN BIG DATA ANALYTICS)
REGULATIONS – 2015
CHOICE BASED CREDIT SYSTEM
CURRICULA AND SYLLABI

SEMESTER - I

SL. NO	COURSE CODE	COURSE TITLE	CATEGORY	CONTACT PERIODS	L	T	P	C
THEORY								
1.	MA7153	Advanced Mathematics for Computing	FC	4	4	0	0	4
2.	CP7151	Advanced Data Structures and Algorithms	PC	3	3	0	0	3
3.	CP7152	Advanced Software Engineering	PC	3	3	0	0	3
4.	CP7153	Advances in Operating Systems	PC	3	3	0	0	3
5.	CP7154	Multi Core Architectures	PC	3	3	0	0	3
PRACTICALS								
6.	CP7161	Advanced Data Structures and Algorithms Lab	PC	4	0	0	4	2
7.	CP7162	Professional Practices	EEC	2	0	0	2	1
			TOTAL	**22**	**16**	**0**	**6**	**19**

II SEMESTER

SL. NO.	COURSE CODE	COURSE TITLE	CATEGORY	CONTACT PERIODS	L	T	P	C
THEORY								
1.	BD7201	Big Data Mining and Analytics	PC	3	3	0	0	3
2.	CP7253	Machine Learning Techniques	PC	5	3	0	2	4
3.	CP7251	Cloud Computing Technologies	PC	3	3	0	0	3
4.	SO7251	Advanced Databases Management Systems	PC	3	3	0	0	3
5.		Elective I	PE	3	3	0	0	3
6.		Elective II	PE	3	3	0	0	3
PRACTICALS								
7.	BD7211	Big Data Computing Lab	PC	4	0	0	4	2
			TOTAL	**24**	**18**	**0**	**8**	**21**

III SEMESTER

SL. NO	COURSE CODE	COURSE TITLE	CATEGORY	CONTACT PERIODS	L	T	P	C
THEORY								
1.	BD7301	Big Data Security	PC	3	3	0	0	3
2.		Elective III	PE	3	3	0	0	3
3.		Elective IV	PE	3	3	0	0	3
4.		Elective V	PE	3	3	0	0	3
PRACTICALS								
5.	BD7311	Project Work Phase I	EEC	12	0	0	12	6
			TOTAL	**24**	**12**	**0**	**12**	**18**

IV SEMESTER

SL. NO.	COURSE CODE	COURSE TITLE	CATEGORY	CONTACT PERIODS	L	T	P	C
PRACTICALS								
1.	BD7411	Project Work Phase II	EEC	24	0	0	24	12
			TOTAL	**24**	**0**	**0**	**24**	**12**

TOTAL NO. OF CREDITS: 70

FOUNDATION COURSES (FC)

SL. NO.	COURSE CODE	COURSE TITLE	CATEGORY	CONTACT PERIODS	L	T	P	C
1.		Advanced Mathematics for Computing	FC	4	4	0	0	4

PROFESSIONAL CORE (PC)

SL. NO.	COURSE CODE	COURSE TITLE	CATEGORY	CONTACT PERIODS	L	T	P	C
1.		Advanced Data Structures and Algorithms	PC	3	3	0	0	3
2.		Advanced Software Engineering	PC	3	3	0	0	3
3.		Advances in Operating System	PC	3	3	0	0	3
4.		Multi-Core Architectures	PC	3	3	0	0	3
5.		Advanced Data Structures and Algorithms Lab	PC	4	0	0	4	2
6.		Big Data Mining and Analytics	PC	3	3	0	0	3
7.		Machine Learning Techniques	PC	2	3	0	2	4
8.		Cloud Computing Technologies	PC	3	3	0	0	3
9.		Advanced Databases Management Systems	PC	3	3	0	0	3
10.		Big Data Computing Lab	PC	4	0	0	4	2
11.		Big Data Security	PC	3	3	0	0	3

PROFESSIONAL ELECTIVES (PE)

SL. NO	COURSE CODE	COURSE TITLE	CATEGORY	CONTACT PERIODS	L	T	P	C
1.	CP7085	Nano Computing	PE	3	3	0	0	3
2.	CP7074	Computational Game Theory	PE	3	3	0	0	3
3.	CP7075	Computational Geometry	PE	3	3	0	0	3
4.	CP7095	Virtualization Techniques and Applications	PE	3	3	0	0	3
5.	CP7081	Fault Tolerant Systems	PE	3	3	0	0	3
6.	CP7077	Database Administration and Tuning	PE	3	3	0	0	3

7.	CP7089	Real Time Systems Design	PE	3	3	0	0	3
8.	IF7071	Bio Informatics	PE	3	3	0	0	3
9.	CP7084	Models of Computations	PE	3	3	0	0	3
10.	CP7086	Network on Chip	PE	3	3	0	0	3
11.	CP7090	Secure Network System Design	PE	3	3	0	0	3
12.	CP7079	Domain Engineering	PE	3	3	0	0	3
13.	CP7071	ADHOC and Wireless Sensor Networks	PE	3	3	0	0	3
14.	CP7080	Ethical Hacking	PE	3	3	0	0	3
15.	CP7078	Digital Image Processing and Applications	PE	3	3	0	0	3
16.	CP7091	Service Oriented Architecture and Design	PE	3	3	0	0	3
17.	CP7252	Compiler Optimization Techniques	PE	3	3	0	0	3
18.	CP7155	Networking Technologies	PE	3	3	0	0	3
BIG DATA ANALYTICS ELECTIVES								
19.	CP7073	Cognitive Science	PE	3	3	0	0	3
20.	CP7082	Information Retrieval Techniques	PE	3	3	0	0	3
21.	CP7083	Internet of Things In the Cloud	PE	3	3	0	0	3
22.	CP7088	Parallel and Distributed Databases	PE	3	3	0	0	3
23.	CP7094	Statistical Natural Language Processing	PE	3	3	0	0	3
24.	CP7093	Soft Computing	PE	3	3	0	0	3
25.	BD7004	Data Visualization	PE	3	3	0	0	3
26.	BD7003	Data Intensive Computing	PE	3	3	0	0	3
27.	BD7005	R Programming	PE	3	3	0	0	3
28.	BD7007	Social Network Analysis for Big Data	PE	3	3	0	0	3
29.	CP7087	Parallel Algorithms	PE	3	3	0	0	3
30.	BD7006	Sentiment Analysis	PE	3	3	0	0	3
31.	BD7001	Big Data Acquisition	PE	3	3	0	0	3
32.	BD7002	Big Data Query Languages	PE	3	3	0	0	3
33.	BD7071	Text Mining	PE	3	3	0	0	3

A minimum of 3 of 5 electives have to necessarily be chosen from the list of “Big Data Analytics Electives”

EMPLOYABILITY ENHANCEMENT COURSES (EEC)

SL. NO	COURSE CODE	COURSE TITLE	CATEGORY	CONTACT PERIODS	L	T	P	C
1.		Professional Practices	EEC	2	0	0	2	1
2.		Project Work Phase I	EEC	12	0	0	12	6
3.		Project Work Phase II	EEC	24	0	0	24	12

MA7153 ADVANCED MATHEMATICS FOR COMPUTING L T P C
4 0 0 4

OBJECTIVES:

- To understand the basics of random variables and standard distributions
- To understand the arrival process and various queueing and server models
- To appreciate the use of simulation techniques
- To apply testing of hypothesis to infer outcome of experiments
- To apply mathematical linear programming techniques to solve constrained problems.

UNIT I RANDOM VARIABLES 12

Random variables – Bernoulli, Binomial, Geometric, Poisson, Uniform, Exponential, Erlang and Normal distributions – Function of a Random variable - Moments, Moment generating function.

UNIT II QUEUING MODELS 12

Poisson Process – Markovian Queues – Single and Multi-server Models – Little's formula – Machine Interference Model – Steady State analysis – Self Service Queue.

UNIT III SIMULATION 12

Discrete Event Simulation – Monte – Carlo Simulation – Stochastic Simulation – Applications to Queuing systems.

UNIT IV TESTING OF HYPOTHESIS 12

Sampling distributions – Estimation of parameters - Statistical hypothesis – Tests based on Normal, t, Chi-square and F distributions for mean, variance and proportion.

UNIT V LINEAR PROGRAMMING 12

Formulation – Graphical solution – Simplex method – Two phase method -Transportation and Assignment Problems.

TOTAL: 60 PERIODS

OUTCOMES:

Upon completion of the course, the student will be able to

- Identify the type of random variable and distribution for a given operational conditions/scene
- Study and Design appropriate queuing model for a given problem/system situation
- Simulate appropriate application/distribution problems
- Differentiate/infer the merit of sampling tests.
- Formulate and find optimal solution in the real life optimizing/allocation/assignment problems involving conditions and resource constraints.

REFERENCES:

1. Johnson, R.A. Miller and Freund's," Probability and Statistical for Engineers, Prentice Hall of India Pvt., Ltd., New Delhi, Seventh Edition, 2005.
2. Hamdy A. Taha, "Operations Research: An Introduction", Prentice Hall of India Pvt., Ltd. New Delhi, Eighth Edition, 2007.
3. Jay L. Devore," Probability and Statistics for Engineering and the Sciences", Cengage Learning, Seventh Edition, 2009.
4. Ross. S.M., "Probability Models for Computer Science", Academic Press, 2002.
5. Winston, W.L., "Operations Research", Thomson – Brooks/Cole, Fourth Edition, 2003.
6. Gross D. and Harris C.M., "Fundamentals of Queuing Theory", John Wiley and Sons, New York, 1998.
7. J.Medhi," Stochastic models of Queuing Theory", Academic Press, Elsevier, Amsterdam, 2003

CP7151	ADVANCED DATA STRUCTURES AND ALGORITHMS	L	T	P	C
		3	0	0	3

OBJECTIVES:

- To extend the students' knowledge of algorithms and data structures
- To enhance their expertise in algorithmic analysis and algorithm design techniques.
- To learn a variety of useful algorithms and techniques
- To extrapolate from them in order to apply those algorithms and techniques to solve problems

UNIT I FUNDAMENTALS **9**

Mathematical Proof Techniques: Induction, proof by contradiction, direct proofs – Asymptotic Notations – Properties of Big-oh Notation –Conditional Asymptotic Notation – Algorithm Analysis – Amortized Analysis – Introduction to NP-Completeness/NP-Hard – Recurrence Equations – Solving Recurrence Equations – Time-Space Tradeoff.

UNIT II HEAP STRUCTURES **9**

Min/Max heaps – Deaps – Leftist Heaps – Binomial Heaps – Fibonacci Heaps – Skew Heaps – Lazy-Binomial Heaps

UNIT III SEARCH STRUCTURES **9**

Binary Search Trees – AVL Trees – Red-Black trees – Multi-way Search Trees –B-Trees – Splay Trees – Tries.

UNIT IV GEOMETRIC ALGORITHMS **9**

Segment Trees – 1-Dimensional Range Searching – k-d Trees – Line Segment Intersection – Convex Hulls – Computing the Overlay of Two Subdivisions – Range Trees – Voronoi Diagram

UNIT V PARALLEL ALGORITHMS **9**

Flynn's Classifications – List Ranking – Prefix computation – Array Max – Sorting on EREW PRAM – Sorting on Mesh and Butterfly – Prefix sum on Mesh and Butterfly – Sum on mesh and butterfly – Matrix Multiplication – Data Distribution on EREW, Mesh and Butterfly

TOTAL : 45 PERIODS

OUTCOMES

Upon completion of this course, the student should be able to

- Have a basic ability to analyze algorithms and to determine algorithm correctness and time efficiency
- Master a variety of advanced data structures and their implementations and different algorithm design techniques in computational geometry and in parallel algorithms
- Apply and implement the learnt algorithm design techniques and data structures to solve problems

REFERENCES

1. Ellis Horowitz, Sartaj Sahni, Dinesh Mehta, "Fundamentals of Data Structures in C", Silicon Pr, 2007.
2. Gilles Brassard, Paul Bratley, "Algorithmics: Theory and Practice", Prentice Hall, 1988.
3. Mark de Berg, Otfried Cheong, Marc van Kreveld, Mark Overmars, "Computational Geometry Algorithms and Applications", Third Edition, 2008.
4. J.A. Storer, "An Introduction to Data Structures and Algorithms", Birkhäuser Boston, 2002.
5. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, "Introduction to Algorithms", MIT Press, 2009.

CP7152	ADVANCED SOFTWARE ENGINEERING	L	T	P	C
		3	0	0	3

OBJECTIVES

- To comprehend software development process and formal specifications
- To know advanced software development techniques and its application in real world context
- To understand how to manage complex projects
- To use advanced software testing techniques
- To understand process improvement and re engineering

UNIT I SOFTWARE ENGINEERING PROCESS AND FORMAL METHODS 9

Software Process models – Software Life Cycle – Development Activities – Managing Software Development – Unified Modeling Language – Requirement elicitation and specification – Understanding formal methods – motivation for formal methods – informal requirements to formal specifications – validating formal specifications – Overview of Z specification

UNIT II AGILE AND ASPECT ORIENTED SOFTWARE ENGINEERING 9

Agile Development: Agility – agile principles- Extreme Programming -– Agile process models – Agile modeling – Agile unified Process – tools set for agile process – Complex Projects: SCRUM – basics, SCRUM Process, Development using SCRUM – Aspect Oriented Software Development: Aspect-Orientation in the Software Lifecycle – Generic Aspect-Oriented Design with UML – Modeling for Aspect-Oriented Software Development-Developing Secure Applications Through Aspect-Oriented Programming.

UNIT III COMPONENT-BASED SOFTWARE ENGINEERING 9

Engineering of component-based systems, the CBSE process – Designing class based components – component design for WebApps – Component-based development – Component-level design patterns – Classifying and retrieving components, and economics of CBSE.

UNIT IV ADVANCED SOFTWARE TESTING TECHNIQUES 9

Software Review – Testing Strategies - Testing Conventional Applications – Testing Object-Oriented Applications – Testing Web Applications – Formal Modeling and verification – Metrics : Product, process, project, testing and quality metrics – Software Test Automation

UNIT V SOFTWARE PROCESS IMPROVEMENT AND REENGINEERING 9

SPI process – CMMI – SPI frameworks – SPI Trends – Emerging trends ion Software Engineering – identifying soft trends – Technology directions – Tool-related trends – Software Maintenance and Reengineering: software reengineering, reverse reengineering, restructuring, forward reengineering.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Analytically apply general principles of software development in the development of complex software and software- intensive systems
- Discuss methods and techniques for advanced software development and also to be able to use these in various development situations
- Apply testing techniques for object oriented software and web-based systems

REFERENCES:

1. Roger S. Pressman, “Software Engineering – A Practioner’s Approach”, MCGraw Hill, 7th edition, 2009.
2. Ian Sommerville, “Software Engineering”, Addison-Wesley 9th Edition, 2010
3. Bernd Bruegge, Allen H. Dutoit, "Object-Oriented Software Engineering", Prentice Hall, Third Edition, 2009.
4. Robert E. Filman, Tzilla Elrad, Siobhán Clarke, Mehmet Aksit, “Aspect-Oriented Software Development”, Addison-Wesley Professional, 2004.
5. Renu Rajni, Pradeep Oak, “Software Testing: Effective Methods, Tools and Techniques”, Tata McGraw Hill, 2004.
6. Jonathan Bowen, "Formal Specification and Documentation using Z - A Case Study Approach", Intl Thomson Computer Pr, 1996.
7. Antoni Diller, “Z: An Introduction to Formal Methods”, Wiley, 1994.
8. James Shore, Shane Warden “The Art of Agile Development - Pragmatic guide to agile software development”, O'Reilly Media, October 2007.
9. Ken Schwaber, “Agile Project Management with SCRUM”, Microsoft Press, 2004.

CP7153	ADVANCES IN OPERATING SYSTEMS	L	T	P	C
		3	0	0	3

OBJECTIVES

- To understand the concepts of distributed systems
- To get an insight into the various issues and solutions in distributed operating systems
- To learn about mobile and real-time operating systems
- To gain knowledge on the design concepts of mainframe operating systems

UNIT I BASICS OF OPERATING SYSTEMS **9**

Overview – Synchronization Mechanisms – Processes and Threads – Process Deadlocks – Issues in Distributed Operating Systems – Communication Primitives – Limitations of a Distributed System

UNIT II DISTRIBUTED OPERATING SYSTEMS **9**

Lamport’s Logical Clocks – Vector Clocks – Causal Ordering of Messages – Distributed Mutual Exclusion Algorithms – Centralized, Distributed and Hierarchical Deadlock Detection Algorithms – Agreement Protocols

UNIT III DISTRIBUTED RESOURCE MANAGEMENT **9**

Distributed File Systems – Design Issues – Google File System – Hadoop Distributed File System – Distributed Shared Memory – Algorithms for Implementing Distributed Shared Memory – Load Distributed Algorithms – Issues in Task Migration – Synchronous and Asynchronous Check Pointing and Recovery – Fault Tolerance – Two-Phase Commit Protocol – Nonblocking Commit Protocol

UNIT IV MOBILE AND REAL TIME OPERATING SYSTEMS **9**

Basic Model of Real Time Systems – Characteristics – Applications of Real Time Systems – Real Time Task Scheduling – Handling Resource Sharing. Mobile Operating Systems – Architecture – Layers – Microkernel Design – Kernel Extensions – Processes and Threads – Memory Management – File system – Android – iOS

UNIT V **MAINFRAME AND LINUX OPERATING SYSTEMS** **9**

Mainframe – z/OS – Overview of z/OS Facilities – Virtual Storage and other Mainframe Concepts – Workload Management – I/O and Data Management – Supervising the Execution of Work in the System – Cross-memory Services – Characteristics of z/OS. Linux – Design Principles – Kernel Modules – Process Management – Scheduling – Memory Management – I/O Management – File System – Interprocess Communication

TOTAL : 45 PERIODS

OUTCOMES

Upon completion of this course, the student should be able to

- Demonstrate the various protocols of distributed operating systems
- Identify the different features of mobile and real-time operating systems
- Discuss the various features of mainframe operating systems

REFERENCES:

1. Mukesh Singhal, Niranjan Shivaratri, “Advanced Concepts in Operating Systems – Distributed, Database and Multiprocessor Operating Systems”, Tata McGraw-Hill, 2001.
2. Rajib Mall, “Real-Time Systems: Theory and Practice”, Prentice Hall, 2006.
3. Neil Smyth, “iPhone iOS 4 Development Essentials – Xcode”, Payload Media, Fourth Edition, 2011.
4. Nikolay Elenkov, “Android Security Internals: An In-Depth Guide to Android’s Security Architecture”, No Starch Press, 2014.
5. Jonathan Levin, “Mac OS X and iOS Internals: To the Apple’s Core”, John Wiley & Sons, 2012.
6. Andrew S. Tanenbaum and Herbert Bos, “Modern Operating Systems”, Fourth Edition, Prentice Hall, 2014.
7. Mike Ebbers, John Kettner, Wayne O’Brien, Bill Ogden, “Introduction to the New Mainframe: z/OS Basics”, Third Edition, International Business Machines Corporation, 2011.
8. Abraham Silberschatz, Peter B. Galvin, Greg Gagne, "Operating System Concepts", Wiley, Eighth edition, 2008.

CP7154 **MULTI CORE ARCHITECTURES**

L	T	P	C
3	0	0	3

OBJECTIVES

- To introduce the students to the recent trends in the field of Computer Architecture and identify performance related parameters
- To understand the different multiprocessor issues
- To expose the different types of multicore architectures
- To understand the design of the memory hierarchy

UNIT I **FUNDAMENTALS OF COMPUTER DESIGN AND ILP** **9**

Fundamentals of Computer Design – Measuring and Reporting Performance – Instruction Level Parallelism and its Exploitation – Concepts and Challenges – Limitations of ILP – Multithreading – SMT and CMP Architectures – The Multicore era.

UNIT II **MEMORY HIERARCHY DESIGN** **9**

Introduction – Optimizations of Cache Performance – Memory Technology and Optimizations – Protection: Virtual Memory and Virtual Machines – Design of Memory Hierarchies – Case Studies.

UNIT III MULTIPROCESSOR ISSUES 9

Symmetric and Distributed Shared Memory Architectures – Cache Coherence Issues – Performance Issues – Synchronization Issues – Models of Memory Consistency – Interconnection Networks – Buses, Crossbar and Multi-stage Interconnection Networks.

UNIT IV MULTICORE ARCHITECTURES 9

Homogeneous and Heterogeneous Multi-core Architectures – Intel Multicore Architectures – SUN CMP architecture – IBM Cell Architecture. Introduction to Warehouse-scale computers, Cloud Computing – Architectures and Issues – Case Studies.

UNIT V VECTOR, SIMD AND GPU ARCHITECTURES 9

Vector Architecture – SIMD Extensions for Multimedia – Graphics Processing Units – Case Studies – GPGPU Computing – Detecting and Enhancing Loop Level Parallelism.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify the limitations of ILP and the need for multicore architectures
- Discuss the issues related to multiprocessing and suggest solutions
- Point out the salient features of different multicore architectures and how they exploit parallelism
- Critically analyze the different types of inter connection networks
- Design a memory hierarchy and optimize it

REFERENCES:

1. John L. Hennessey and David A. Patterson, “Computer Architecture – A Quantitative Approach”, Morgan Kaufmann / Elsevier, 5th edition, 2012.
2. Darryl Gove, “Multicore Application Programming: For Windows, Linux, and Oracle Solaris”, Pearson, 2011.
3. David B. Kirk, Wen-mei W. Hwu, “Programming Massively Parallel Processors”, Morgan Kauffman, 2010.
4. Wen– mei W. Hwu, “GPU Computing Gems”, Morgan Kaufmann / Elsevier, 2011

CP7161 ADVANCED DATA STRUCTURES AND ALGORITHMS LAB L T P C
0 0 4 2

OBJECTIVES:

- To understand heap and various tree structures like AVL, Red-black, B and Segment trees
- To understand the problems such as line segment intersection, convex shell and Voronoi diagram

1. Min/Max Heap
2. Leftist Heap
3. AVL Trees
4. Red-Black Trees
5. B-Trees
6. Segment Trees

7. Line segment intersection
8. Convex Hull
9. Voronoi Diagram

TOTAL : 60 PERIODS

OUTCOMES

Upon completion of this course, the student should be able to

- Implement heap and various tree structure like AVL, Red-black, B and Segment trees
- Solve the problems such as line segment intersection, convex shell and Voronoi diagram

CP7162 **PROFESSIONAL PRACTICES** **L T P C**
0 0 2 1

OBJECTIVES:

- To facilitate analysis, design and problem solving skills
- To have a thorough domain knowledge
- To understand the best Industry practices by reading case studies
- To kindle innovative and professional thinking
- To explore possible alternative solutions
- To estimate feasibility, cost, risk and ROI

Identify an application (may be of social relevance) – Understand customer requirements – analyze and understand customers and stakeholders – value additions – innovations and research component – preparing plan / SRS document indicating feasibility, cost, risk, ROI and related design – suggest implementation methodology – perform risk assessment and management

TOTAL : 30 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify and formulate the problem
- Describe the background of the problem
- Assess the needs of stakeholders
- Make estimates like cost, risk, ROI etc., to justify the business opportunity.
- Describe the industry standards and procedures
- Predict the business opportunity
- Suggest system implications

BD7201 **BIG DATA MINING AND ANALYTICS** **L T P C**
3 0 0 3

OBJECTIVES:

- To understand the computational approaches to Modeling, Feature Extraction
- To understand the need and application of Map Reduce
- To understand the various search algorithms applicable to Big Data
- To analyse and interpret streaming data
- To learn how to handle large data sets in main memory
- To learn the various clustering techniques applicable to Big Data

UNIT I DATA MININGAND LARGE SCALE FILES 9

Introduction to Statistical modeling – Machine Learning – Computational approaches to modeling – Summarization – Feature Extraction – Statistical Limits on Data Mining - Distributed File Systems – Map-reduce – Algorithms using Map Reduce – Efficiency of Cluster Computing Techniques.

UNIT II SIMILAR ITEMS 9

Nearest Neighbor Search – Shingling of Documents – Similarity preserving summaries – Locality sensitive hashing for documents – Distance Measures – Theory of Locality Sensitive Functions – LSH Families – Methods for High Degree of Similarities.

UNIT III MINING DATA STREAMS 9

Stream Data Model – Sampling Data in the Stream – Filtering Streams – Counting Distance Elements in a Stream – Estimating Moments – Counting Ones in Window – Decaying Windows

UNIT IV LINK ANALYSIS AND FREQUENT ITEMSETS 9

Page Rank –Efficient Computation - Topic Sensitive Page Rank – Link Spam – Market Basket Model – A-priori algorithm – Handling Larger Datasets in Main Memory – Limited Pass Algorithm – Counting Frequent Item sets.

UNIT V CLUSTERING 9

Introduction to Clustering Techniques – Hierarchical Clustering –Algorithms – K-Means – CURE – Clustering in Non – Euclidean Spaces – Streams and Parallelism – **Case Study:** Advertising on the Web – Recommendation Systems

TOTAL : 45 PERIODS

OUTCOMES

Upon completion of this course, the students will be able to

- Design algorithms by employing Map Reduce technique for solving Big Data problems
- Design algorithms for Big Data by deciding on the apt Features set
- Design algorithms for handling petabytes of datasets
- Design algorithms and propose solutions for Big Data by optimizing main memory consumption
- Design solutions for problems in Big Data by suggesting appropriate clustering techniques.

REFERENCES

1. Jure Leskovec, AnandRajaraman, Jeffrey David Ullman, “Mining of Massive Datasets”, Cambridge University Press, Second Edition, 2014.
2. Jiawei Han, MichelineKamber, Jian Pei, “Data Mining Concepts and Techniques”, Morgan Kaufman Publications, Third Edition, 2011.
3. Ian H.Witten, Eibe Frank “Data Mining – Practical Machine Learning Tools and Techniques”, Morgan Kaufman Publications, Third Edition, 2011.
4. David Hand, HeikkiMannila and Padhraic Smyth, “Principles of Data Mining”, MIT PRESS, 2001

CP7253 **L T P C**

MACHINE LEARNING TECHNIQUES 3 0 2 4

OBJECTIVES

- To understand the concepts of machine learning
- To appreciate supervised and unsupervised learning and their applications
- To understand the theoretical and practical aspects of Probabilistic Graphical Models
- To appreciate the concepts and algorithms of reinforcement learning
- To learn aspects of computational learning theory

UNIT I INTRODUCTION 8 + 6

Machine Learning - Machine Learning Foundations –Overview – Design of a Learning system - Types of machine learning –Applications Mathematical foundations of machine learning - random variables and probabilities - Probability Theory – Probability distributions -Decision Theory- Bayes Decision Theory - Information Theory

UNIT II SUPERVISED LEARNING 10 + 6

Linear Models for Regression - Linear Models for Classification – Naïve Bayes - Discriminant Functions -Probabilistic Generative Models -Probabilistic Discriminative Models - Bayesian Logistic Regression. Decision Trees - Classification Trees- egression Trees - Pruning. Neural Networks -Feed-forward Network Functions - Back- propagation. Support vector machines - Ensemble methods- Bagging- Boosting.

UNIT III UNSUPERVISED LEARNING 8 + 6

Clustering- K-means - EM Algorithm- Mixtures of Gaussians. The Curse of Dimensionality - Dimensionality Reduction - Factor analysis - Principal Component Analysis - Probabilistic PCA- Independent components analysis

UNIT IV PROBABILISTIC GRAPHICAL MODELS 10 + 6

Graphical Models - Undirected graphical models - Markov Random Fields - Directed Graphical Models -Bayesian Networks - Conditional independence properties - Inference – Learning- Generalization - Hidden Markov Models - Conditional random fields(CRFs)

UNIT V ADVANCED LEARNING 9 + 6

Sampling –Basic sampling methods – Monte Carlo. Reinforcement Learning- K-Armed Bandit- Elements - Model-Based Learning- Value Iteration- Policy Iteration. Temporal Difference Learning- Exploration Strategies- Deterministic and Non-deterministic Rewards and Actions Computational Learning Theory - Mistake bound analysis, sample complexity analysis, VC dimension. Occam learning, accuracy and confidence boosting

TOTAL : 45 + 30 = 75 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Design a neural network for an application of your choice
- Implement probabilistic discriminative and generative algorithms for an application of your choice and analyze the results
- Use a tool to implement typical clustering algorithms for different types of applications
- Design and implement an HMM for a sequence model type of application
- Identify applications suitable for different types of machine learning with suitable justification

REFERENCES:

1. Christopher Bishop, “Pattern Recognition and Machine Learning” Springer, 2007.
2. Kevin P. Murphy, “Machine Learning: A Probabilistic Perspective”, MIT Press, 2012.
3. Ethem Alpaydin, “Introduction to Machine Learning”, MIT Press, Third Edition, 2014.
4. Tom Mitchell, "Machine Learning", McGraw-Hill, 1997.
5. Trevor Hastie, Robert Tibshirani, Jerome Friedman, "The Elements of Statistical Learning", Springer, Second Edition, 2011.
6. Stephen Marsland, “Machine Learning - An Algorithmic Perspective”, Chapman and Hall/CRC Press, Second Edition, 2014.

CP7251 CLOUD COMPUTING TECHNOLOGIES L T P C
3 0 0 3

OBJECTIVES:

- To understand the concepts of cloud and utility computing
- To understand the various issues in cloud computing
- To familiarize themselves with the lead players in cloud
- To appreciate the emergence of cloud as the next generation computing paradigm
- To be able to set up a private cloud

UNIT I INTRODUCTION 9

Evolution of Cloud Computing -System Models for Distributed and Cloud Computing - NIST Cloud Computing Reference Architecture -IaaS - On-demand Provisioning - Elasticity in Cloud - Examples of IaaS Providers - PaaS - Examples of PaaS Providers - SaaS - Examples of SaaS Providers - Public , Private and Hybrid Clouds – Google App Engine, Amazon AWS - Cloud Software Environments -Eucalyptus, Open Nebula, Open Stack, Nimbus

UNIT II VIRTUALIZATION 9

Basics of Virtual Machines - Process Virtual Machines – System Virtual Machines – Emulation – Interpretation – Binary Translation - Taxonomy of Virtual Machines. Virtualization –Management Virtualization –– Hardware Maximization – Architectures – Virtualization Management – Storage Virtualization – Network Virtualization

UNIT III VIRTUALIZATION INFRASTRUCTURE 9

Comprehensive Analysis – Resource Pool – Testing Environment –Server Virtualization – Virtual Workloads – Provision Virtual Machines –Desktop Virtualization – Application Virtualization – Work with AppV – Mobile OS for smart phones – Mobile Platform Virtualization – Collaborative Applications for Mobile platforms

UNIT IV PROGRAMMING MODEL 9

Map Reduce Hadoop Distributed File Systems – Hadoop I/O – Developing Map Reduce Applications – Working of Map Reduce – Types and Formats – Setting up Hadoop Cluster

UNIT V CLOUD INFRASTRUCTURE AND SECURITY 9

Architectural Design of Compute and Storage Clouds - Inter Cloud Resource Management - Resource Provisioning and Platform Deployment - Global Exchange of Cloud Resources - Security Overview – Cloud Security Challenges – Software as a Service Security – Security Governance – Risk Management – Security Monitoring – Security Architecture Design – Data Security – Application Security – Virtual Machine Security.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Articulate the main concepts, key technologies, strengths and limitations of cloud computing
- Identify the architecture, infrastructure and delivery models of cloud computing
- Explain the core issues of cloud computing such as security, privacy and interoperability
- Choose the appropriate technologies, algorithms and approaches for the related issues

REFERENCES:

1. Kai Hwang, Geoffrey C Fox, Jack G Dongarra, "Distributed and Cloud Computing, From Parallel Processing to the Internet of Things", Morgan Kaufmann Publishers, 2012.
2. John W.Rittinghouse and James F.Ransome, "Cloud Computing: Implementation, Management, and Security", CRC Press, 2010.
3. Toby Velte, Anthony Velte, Robert Elsenpeter, "Cloud Computing, A Practical Approach", McGraw-Hill Osborne Media, 2009.
4. Jim Smith, Ravi Nair, "Virtual Machines: Versatile Platforms for Systems and Processes", Elsevier/Morgan Kaufmann, 2005.
5. Danielle Ruest, Nelson Ruest, “Virtualization: A Beginner’s Guide”, McGraw-Hill Osborne Media, 2009.
6. Tom White, "Hadoop: The Definitive Guide", Yahoo Press, 2012.
7. Rajkumar Buyya, Christian Vecchiola, and Thamarai Selvi, “Mastering Cloud Computing”, Tata McGraw Hill, 2013.

SO7251	ADVANCED DATABASE MANAGEMENT SYSTEMS	L	T	P	C
		3	0	0	3

OBJECTIVES:

- To understand the underlying principles of Relational Database Management System.
- To understand and implement the advanced features of DBMS.
- To develop database models using distributed databases.
- To implement and maintain an efficient database system using emerging trends.

UNIT I RELATIONAL MODEL 9

Data Model – Types of Data Models: – Entity Relationship Model – Relational Data Model – Mapping Entity Relationship Model to Relational Model – Structured Query Language – Database Normalization – Transaction Management.

UNIT II PARALLEL AND DISTRIBUTED DATABASES 9

Centralized and Client-Server Architectures – Parallel Systems – Distributed Systems – Parallel Databases – I/O Parallelism – Inter- and Intra-Query Parallelism – Inter- and Intra-operation Parallelism – Distributed Database Concepts: – Distributed Data Storage – Distributed Transactions – Commit Protocols – Concurrency Control – Distributed Query Processing.

UNIT III XML DATABASES 9

XML Databases: XML Data Model – DTD – XML Schema – XML Querying – Web Databases – Open Database Connectivity.

UNIT IV MULTIMEDIA DATABASES 9

Multidimensional Data Structures – Image Databases – Text / Document Databases – Video Databases – Audio Databases – Multimedia Database Design.

UNIT V CURRENT ISSUES 9

Active Databases – Deductive Databases – Data Warehousing – Data Mining – Database Tuning – Database Security

TOTAL: 45 PERIODS

OUTCOMES:

On successful completion of this course, the student will be able to:

- Design and implement relational databases, distributed databases, XML databases and multimedia databases.
- Implement the concept of database connectivity with the applications.

REFERENCES

1. R. Elmasri, S.B. Navathe, “Fundamentals of Database Systems”, Addison-Wesley, 2011.
2. Thomas Cannolly and Carolyn Begg, “Database Systems, A Practical Approach to Design, Implementation and Management”, Third Edition, Pearson Education, 2007.
3. Henry F Korth, Abraham Silberschatz, S. Sudharshan, “Database System Concepts”, Fifth Edition, McGraw Hill, 2006.
4. C.J.Date, A.Kannan and S.Swamynathan,”An Introduction to Database Systems”, Eighth Edition, Pearson Education, 2006.
5. V.S.Subramanian, “Principles of Multimedia Database Systems”, Harcourt India Pvt. Ltd., 2001.

BD7211 BIG DATA COMPUTING LAB L T P C
0 0 4 2

OBJECTIVES

- To understand setting up of Hadoop Cluster
- To solve problems using Map Reduce Technique
- To solve Big Data problems

LIST OF EXERCISES

1. Set up a pseudo-distributed, single-node Hadoop cluster backed by the Hadoop Distributed File System, running on Ubuntu Linux. After successful installation on one node, configuration of a multi-node Hadoop cluster(one master and multiple slaves).
2. MapReduce application for word counting on Hadoop cluster
3. Unstructured data into NoSQL data and do all operations such as NoSQL query with API.
4. K-means clustering using map reduce
5. Page Rank Computation
6. Mahout machine learning library to facilitate the knowledge build up in big data analysis.
7. Application of Recommendation Systems using Hadoop/mahout libraries

TOTAL: 60 PERIODS

OUTCOMES

Upon of completion of this course, students will be able to

- Set up multi-node Hadoop Clusters
- Apply Map Reduce algorithms for various algorithms
- Design new algorithms that uses Map Reduce to apply on Unstructured and structured data

BD7301 BIG DATA SECURITY L T P C
3 0 0 3

OBJECTIVES:

- To understand the mathematical foundations of security principles
- To appreciate the different aspects of encryption techniques
- To understand the role played by authentication in security
- To understand the security concerns of big-data.

UNIT I SYMMETRIC TECHNIQUES 9

Probability and Information Theory - Algebraic foundations – Number theory - Substitution Ciphers – Transposition Ciphers – Classical Ciphers – DES – AES – Confidentiality Modes of Operation

UNIT II ASYMMETRIC TECHNIQUES 9

Diffie-Hellman Key Exchange protocol – Discrete logarithm problem – RSA cryptosystems & cryptanalysis – ElGamal cryptosystem – Elliptic curve architecture and cryptography - Data Integrity techniques.

UNIT III AUTHENTICATION 9

Authentication requirements – Authentication functions – Message authentication codes – Hash functions – Security of hash functions and MACS – MD5 Message Digest algorithm – Secure hash algorithm.

UNIT IV SECURITY ANALYTICS I 9

Introduction to Security Analytics – Techniques in Analytics – Analysis in everyday life – Challenges in Intrusion and Incident Identification – Analysis of Log file – Simulation and Security Process.

UNIT V SECURITY ANALYTICS II 9

Access Analytics – Security Analysis with Text Mining – Security Intelligence – Security Breaches

TOTAL: 45 PERIODS

OUTCOMES

Upon Completion of the course, the students will be able to

- Design algorithms in a secure manner for Big data applications
- Use available security practices in big-data analytics.
-

REFERENCES

1. William Stallings, “Crpyptography and Network security: Principles and Practices”, Pearson/PHI, 5th Edition, 2010.
2. Behrouz A. Forouzan, “Cryptography and Network Security”, Tata McGraw Hill Education, 2nd Edition, 2010.
3. Douglas R. Stinson ,“Cryptography Theory and Practice ”, Chapman & Hall/CRC, 3rd Edition, 2006.
4. Mark Talabis, Robert McPherson, I Miyamoto and Jason Martin, “Information Security Analytics: Finding Security Insights, Patterns, and Anomalies in Big Data”,Syngress Media, U.S., 2014

CP7085 NANO COMPUTING

L	T	P	C
3	0	0	3

OBJECTIVES

- To understand the basics of quantum computing, membrane computing, molecular computing, DNA computing and nano computing.
- To understand the models and the theory involved in the biologically inspired computing techniques.
- To explore the applications of these computing models.

UNIT I INTRODUCTION 9

Nanoscale Science and Technology- Implications for Physics, Chemistry, Biology and Engineering-Classifications of nanostructured materials- nano particles- quantum dots, nanowires-ultra-thinfilms-multilayered materials. Length Scales involved and effect on properties: Mechanical, Electronic, Optical, Magnetic and Thermal properties. Introduction to properties and motivation for study (qualitative only).

UNIT II DNA COMPUTING 9

Structure of DNA – Operation on DNA molecules – Adleman's experiments – Other DNA solutions to NP problems – Two dimensional generalization – Computing by carving – Sticker systems – Extended H systems – Controlled H systems – distributed H systems

UNIT III MEMBRANE COMPUTING 9

P systems with labelled membranes – examples – Power of P systems – decidability results – Rewriting P systems – P systems with polarized membranes – Normal forms – P systems on Asymmetric graphs – P systems with active membranes – Splicing P systems – Variants, Problems, Conjectures

UNIT IV QUANTUM COMPUTING 9

Reversible computation – Copy computers – Quantum world – Bits and Qubits – Quantum calculus – Qubit evolution – Measurements – Zeno machines – Randomness – EPR conundrum and Bell's theorem – Quantum logic – Quantum computers – Quantum algorithms – Quantum Complexity – Quantum Cryptography

UNIT V NANO AND MOLECULAR COMPUTING 9

Defect tolerant nano computing – error detection – Non-traditional computing models – Reliability trade off for nano architecture – Molecular recognition – storage and processing of molecular information

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Comprehend the different computing paradigms
- Write Grammar rules for the different models of computing
- Design applications to incorporate one or more computing models
- Solve problems and prove the application of the computing models

REFERENCES:

1. CrisCalude Gheorghe Paun, "Computing with Cells and Atoms: An Introduction to Quantum, DNA and Membrane Computing", CRC Press, 2000.
2. Sandeep kumar Shukla, R Iris Bahar, "Nano, Quantum and Molecular Computing: Implications to High Level Design and Validation", Kluwer Academic Publishers, 2013.
3. Tanya Sienko, Andrew Adamatzky, Michael Conrad, Nicholas G. Rambidi, "Molecular Computing", MIT Press, 2005.
4. Kamala Krithivasan and Rama R, "Introduction to Formal languages, automata theory and computation", Pearson Education India, 2009.

CP7074	COMPUTATIONAL GAME THEORY	L	T	P	C
		3	0	0	3

OBJECTIVES

- To introduce the student to the notion of a game, its solutions concepts, and other basic notions and tools of game theory, and the main applications for which they are appropriate, including electronic trading markets
- To formalize the notion of strategic thinking and rational choice by using the tools of game theory, and to provide insights into using game theory in modeling applications
- To draw the connections between game theory, computer science, and economics, especially emphasizing the computational issues
- To introduce contemporary topics in the intersection of game theory, computer science, and economics

UNIT I INTRODUCTION 9

Introduction – Making rational choices: basics of Games – strategy - preferences – payoffs – Mathematical basics - Game theory – Rational Choice - Basic solution concepts-non-cooperative versus cooperative games - Basic computational issues - finding equilibria and learning in games- Typical application areas for game theory (e.g. Google's sponsored search, eBay auctions, electricity trading markets).

UNIT II GAMES WITH PERFECT INFORMATION 9

Games with Perfect Information - Strategic games - prisoner's dilemma, matching pennies- Nash equilibria- theory and illustrations - Cournot's and Bertrand's models of oligopoly- auctions- mixed strategy equilibrium- zero-sum games- Extensive Games with Perfect Information- repeated games (prisoner's dilemma)- subgame perfect Nash equilibrium; computational issues.

UNIT III GAMES WITH IMPERFECT INFORMATION 9

Games with Imperfect Information - Bayesian Games – Motivational Examples – General Definitions –Information aspects – Illustrations - Extensive Games with Imperfect -Information - Strategies- Nash Equilibrium – Beliefs and sequential equilibrium – Illustrations - Repeated Games – The Prisoner's Dilemma – Bargaining

UNIT IV NON-COOPERATIVE GAME THEORY 9

Non-cooperative Game Theory - Self-interested agents- Games in normal form - Analyzing games: from optimality to equilibrium - Computing Solution Concepts of Normal-Form Games – Computing Nash equilibriaof two-player, zero-sum games -Computing Nash equilibria of two-player, general-sum games - Identifying dominated strategies

UNIT V MECHANISM DESIGN 9

Aggregating Preferences-Social Choice – Formal Model- Voting - Existence of social functions - Ranking systems - Protocols for Strategic Agents: Mechanism Design - Mechanism design with unrestricted preferences- Efficient mechanisms - Vickrey and VCG mechanisms (shortest paths) - Combinatorial auctions - profit maximization Computational applications of mechanism design - applications in Computer Science - Google's sponsored search - eBay auctions

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Discuss the notion of a strategic game and equilibria, and identify the characteristics of main applications of these concepts.
- Do a literature survey on applications of Game Theory in Computer Science and Engineering.
- Discuss the use of Nash Equilibrium for other problems.

- Identify key strategic aspects and based on these be able to connect them to appropriate game theoretic concepts given a real world situation.
- Identify some applications that need aspects of Bayesian Games
- Implement a typical Virtual Business scenario using Game theory

REFERENCES:

1. M. J. Osborne, “An Introduction to Game Theory”, Oxford University Press, 2003.
2. N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, “Algorithmic Game Theory”, Cambridge University Press, 2007.
3. M. J. Osborne and A. Rubinstein, “A Course in Game Theory”, MIT Press, 1994.
4. A.Dixit and S. Skeath, “Games of Strategy”, W W Norton & Co Inc, 3rd Edition 2009.
5. YoavShoham, Kevin Leyton-Brown, “Multi agent Systems: Algorithmic, Game- Theoretic, and Logical Foundations”, Cambridge University Press, 2008.
6. Zhu Han, Dusit Niyato, Walid Saad, Tamer Basar and Are Hjorungnes, “Game Theory in Wireless and Communication Networks”, Cambridge University Press, 2012.

CP7075 COMPUTATIONAL GEOMETRY

L	T	P	C
3	0	0	3

OBJECTIVES

- To understand geometric problems.
- To learn the algorithmic solutions for geometric problems.
- To map problems in various application domains to a geometric problem.
- To learn to solve problems in various application domains.

UNIT I INTRODUCTION 9

Introduction – Application Domains – Line Segment Intersection – Intersection of Convex Polygons – Polygon Triangulation

UNIT II GEOMETRIC SEARCHING 9

Geometric Searching – Range Searching – K d-Trees – Range trees – Point-Location Problems

UNIT III CONVEX HULL PROBLEM 9

Convex hull Problem – Preliminaries – Convex hull Algorithms in the Plane – Graham’s scan - Jarvis’s March – Quick Hull – Divide-and-conquer – Dynamic Convex Hull Maintenance – Delaunay Triangulation

UNIT IV PROXIMITY PROBLEMS 9

Proximity Problems – Fundamental Algorithms (Closest Pair – All Nearest Neighbours – Euclidean Minimum Spanning Tree – Nearest Neighbour Search) – Lower bounds – Closest Pair Problem : A Divide-and-Conquer Approach

UNIT V VORONOI DIAGRAM 9

Voronoi Diagram – Proximity Problems Solved by the Voronoi Diagram – Planar Applications

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify problems that can be mapped to geometric problems
- Transform problems in different applications to geometric problems
- Use the algorithms learnt for solving the transformed problems

REFERENCES

1. Franco P. Preparata, Michael I. Shamos, “Computational Geometry: An Introduction”, Springer, 1993.
2. Mark de Berg, Otfried Cheong, Marc van Kreveld, Mark Overmars, “Computational Geometry : Algorithms and Applications”, Springer, 3rd Edition, 2008.
3. Satyan L. Devadoss and Joseph O'Rourke, “Discrete and Computational Geometry”, Princeton University Press, 2011.
4. Herbert Edelsbrunner, “Algorithms in Combinatorial Geometry, EATCS Monographs in Computer Science”, Springer Verlag, 2011.

CP7095	VIRTUALIZATION TECHNIQUES AND APPLICATIONS	L	T	P	C
		3	0	0	3

OBJECTIVES

- To understand the concepts of virtualization and virtual machines
- To understand the implementation of process and system virtual machines
- To explore the aspects of high level language virtual machines
- To gain expertise in server, network and storage virtualization.
- To understand and deploy practical virtualization solutions and enterprise solutions

UNIT I OVERVIEW OF VIRTUALIZATION 9

System architectures - Virtual Machine basics - Process vs System Virtual Machines - Taxonomy. Emulation: Basic Interpretation - Threaded Interpretation - Precoded and Direct Threaded Interpretation - Binary Translation. System Virtual Machines - Key concepts - Resource utilization basics

UNIT II PROCESS VIRTUAL MACHINES 9

Implementation – Compatibility – Levels – Framework – State Mapping – Register – Memory Address Space – Memory Architecture Emulation – Memory Protection – Instruction Emulation – Performance Tradeoff - Staged Emulation – Exception Emulation – Exception Detection – Interrupt Handling – Operating Systems Emulation – Same OS Emulation – Different OS Emulation – System Environment

UNIT III HIGH LEVEL LANGUAGE VIRTUAL MACHINES AND SERVER VIRTUALIZATION 9

HLL virtual machines: Pascal P-Code – Object Oriented HLLVMs - Java VM architecture - Java Native Interface - Common Language Infrastructure. Server virtualization: Partitioning techniques - virtual hardware - uses of virtual servers - server virtualization platforms

UNIT IV NETWORK AND STORAGE VIRTUALIZATION 9

Design of Scalable Enterprise Networks – Layer2 Virtualization – VLAN - VFI - Layer 3 Virtualization – VRF - Virtual Firewall Contexts - Network Device Virtualization - Data- Path Virtualization - Routing Protocols. Hardware Devices – SAN backup and recovery techniques – RAID – Classical Storage Model – SNIA Shared Storage Model – Virtual Storage: File System Level and Block Level

UNIT V APPLYING VIRTUALIZATION 9

Practical Virtualization Solutions: Comparison of Virtualization Technologies: Guest OS/ Host OS – Hypervisor – Emulation – Kernel Level – Shared Kernel, Enterprise Solutions: VMWare Server – VMWare ESXi – Citrix Xen Server – Microsoft Virtual PC – Microsoft Hyper-V – Virtual Box, Server Virtualization: Configuring Servers with Virtualization – Adjusting and Tuning Virtual servers – VM Backup – VM Migration, Desktop Virtualization: Terminal services – Hosted Desktop – Web-based Solutions – Localized Virtual Desktops, Network and Storage Virtualization: Virtual Private Networks – Virtual LAN – SAN and VSAN – NAS

TOTAL : 45 PERIODS

OUTCOMES

Upon completion of this course, the student should be able to

- Deploy legacy OS on virtual machines.
- Analyze the intricacies of server, storage and network virtualizations
- Design and develop applications on virtual machine platforms

REFERENCES

1. James E. Smith, Ravi Nair, “Virtual Machines: Versatile Platforms for Systems and Processes”, Elsevier/Morgan Kaufmann, 2005.
2. David Marshall, Wade A. Reynolds, “Advanced Server Virtualization: VMware and Microsoft Platform in the Virtual Data Center”, Auerbach Publications, 2006.
3. Kumar Reddy, Victor Moreno, “Network virtualization”, Cisco Press, July, 2006.
4. Chris Wolf, Erick M. Halter, “Virtualization: From the Desktop to the Enterprise”, APress 2005.
5. Kenneth Hess , Amy Newman, “Practical Virtualization Solutions: Virtualization from the Trenches”, Prentice Hall, 2010

CP7081 FAULT TOLERANT SYSTEMS

L	T	P	C
3	0	0	3

OBJECTIVES

- To provide and appreciate a comprehensive view of fault tolerant systems
- To expose the students to the methods of hardware fault tolerance
- To understand the different ways of providing information redundancy and the ways of providing software fault tolerance.
- To expose the students to concept of check pointing and their role in providing fault tolerance.
- To understand how to handle security attacks.

UNIT I INTRODUCTION 9

Fault Classification, Types of Redundancy, Basic Measures of Fault Tolerance, Hardware Fault Tolerance, The Rate of Hardware Failures, Failure Rate, Reliability, and Mean Time to Failure, Canonical and Resilient Structures, Other Reliability Evaluation Techniques, Processor level Techniques

UNIT II INFORMATION REDUNDANCY 9

Information Redundancy, Coding, Resilient Disk Systems, Data Replication, Voting: Hierarchical Organization, Primary-Backup Approach, Algorithm-Based Fault Tolerance, Fault-Tolerant Networks: Measures of Resilience, Common Network Topologies and Their Resilience, Fault-Tolerant Routing

UNIT III SOFTWARE FAULT TOLERANCE 9

Acceptance Tests, Single-Version Fault Tolerance, N-Version Programming, Recovery Block Approach, Preconditions, Post conditions, and Assertions, Exception-Handling, Software Reliability Models, Fault-Tolerant Remote Procedure Calls

UNIT IV CHECKPOINTING 9

Introduction, Checkpoint Level, Optimal Checkpointing - An Analytical Model, Cache-Aided Rollback Error Recovery, Checkpointing in Distributed Systems, Checkpointing in Shared-Memory Systems, Checkpointing in Real-Time Systems, Case Studies: NonStop Systems, Stratus Systems, Cassini Command and Data Subsystem, IBM G5, IBM Sysplex, Itanium

UNIT V FAULT DETECTION IN CRYPTOGRAPHIC SYSTEMS 9

Security Attacks Through Fault Injection – Fault Attacks on Symmetric Key Ciphers – Fault Attacks on Public (Asymmetric) Key Ciphers – Counter Measures – Spatial and Temporal Duplication – Error Detecting Codes

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Define the traditional measures of fault tolerance
- Point out the processor level fault tolerance techniques
- Critically analyze the different types of RAID levels
- Discuss techniques like recovery blocks and N-version programming
- Identify techniques for check pointing in distributed and shared memory systems.
- Provide techniques to detect injected faults in ciphers.

REFERENCES:

1. Israel Koren, Mani Krishna, “Fault Tolerant Systems”, Morgan Kaufmann, 2010
2. Parag K. Lala “Fault Tolerant and Fault Testable Hardware Design”, Prentice-Hall International, 1984.
3. LL Pullam, “Software Fault Tolerance Techniques and Implementation”, Artech House Computer Security Series, 2002.
4. Martin L Shooman, “Reliability of Computer Systems and Networks: Fault Tolerance, Analysis and Design”, Willey, 2002.

CP7077 DATABASE ADMINISTRATION AND TUNING

L	T	P	C
3	0	0	3

OBJECTIVES

- To design and implement relational database solutions for general applications
- To develop database scripts for data manipulation and database administration
- To understand and perform common database administration tasks such as database monitoring, performance tuning, data transfer, and security
- To balance the different types of competing resources in the database environment so that the most important applications have priority access to the resources

UNIT I INTRODUCTION TO DATABASE ADMINISTRATION 9

Database Administration - DBA Tasks - DBMS Release Migration - Types of DBAs - Creating the Database Environment – Defining the organizations DBMS strategy - Installing the DBMS - Upgrading DBMS Versions and Releases

UNIT II DATABASE SECURITY, BACKUP AND RECOVERY 9

Database Users - Granting and Revoking Authority - Authorization Roles and Groups - Using Views for Security - Using Stored Procedures for Security – Auditing - External Security - Backups - Recovery - Determining Recovery Options - Types of Recovery – DBA Tools – DBA Rules of Thumb

UNIT III FUNDAMENTALS OF TUNING 9

Review of Relational Databases – Relational Algebra – Locking and Concurrency Control – Correctness Consideration – Lock Tuning – Logging and the Recovery Subsystem – Principles of Recovery – Tuning the Recovery Subsystem – Operating Systems Considerations – Hardware Tuning

UNIT IV INDEX TUNING AND QUERY OPTIMIZATION 9

Types of Queries – Data Structures – B+ Tree - Hash Structures – Bit Map Indexes – Clustering Indexes – Non Clustering Indexes – Composite Indexes – Hot Tables – Comparison of Indexing and Hashing Techniques. Optimization Techniques - Tuning Relational Systems - Parameter Cache - Query Tuning – Triggers – Client Server Mechanisms – Objects, Application Tools and Performance –Tuning the Application Interface – Bulk Loading Data – Accessing Multiple Databases

UNIT V TROUBLESHOOTING 9

Query Plan Explainers – Performance Monitors – Event Monitors – Finding “Suspicious” Queries – Analyzing a Query’s Access Plan – Profiling a Query Execution – DBMS Subsystems

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Apply advanced features of databases in design, administration, and applications
- Provide techniques to improve the performance of a database
- Optimize the use of existing resources within the database environment

REFERENCES:

1. Craig S. Mullins, “Database Administration: The Complete Guide to Practices and Procedures”, Addison-Wesley Professional, 2012.
2. Dennis Shasha and Philippe Bonnet, “Database Tuning, Principles, Experiments and Troubleshooting Techniques”, Elsevier Reprint, 2005.
3. Avi Silberschatz, Henry F. Korth, S. Sudarshan, "Database System Concepts", Sixth Edition, McGraw-Hill, 2010.
4. Thomas Connoly and CarlolynBegg, “Database Systems, A Practical Approach to Design, Implementation and Management”, Fifth Edition, Pearson Education, 2009.

CP7089 REAL TIME SYSTEMS DESIGN

L	T	P	C
3	0	0	3

OBJECTIVES:

- To learn real time operating system concepts and the associated issues & techniques.
- To understand design and synchronization problems in Real Time System.
- To understand the evaluation techniques present in Real Time System.

UNIT I REAL TIME SPECIFICATION AND DESIGN TECHNIQUES 9

Introduction– Structure of a Real Time System –Task classes – Performance Measures for Real Time Systems – Estimating Program Run Times – Issues in Real Time Computing – Task Assignment and Scheduling – Classical uniprocessor scheduling algorithms –Fault Tolerant Scheduling.

UNIT II SOFTWARE REQUIREMENTS ENGINEERING 9

Requirements engineering process – types of requirements – requirements specification for real time systems – Formal methods in software specification – structured Analysis and Design – object oriented analysis and design and unified modelling language – organizing the requirements document – organizing and writing documents – requirements validation and revision.

UNIT III INTERTASK COMMUNICATION AND MEMORY MANAGEMENT 9

Buffering data – Time relative Buffering- Ring Buffers – Mailboxes – Queues – Critical regions – Semaphores – other Synchronization mechanisms – deadlock – priority inversion – process stack management – run time ring buffer – maximum stack size – multiple stack arrangement – memory management in task control block - swapping – overlays – Block page management – replacement algorithms – memory locking – working sets – real time garbage collection – contiguous file systems.

UNIT IV REAL TIME DATABASES 9

Real time Databases – Basic Definition, Real time Vs General Purpose Databases, Main Memory Databases, Transaction priorities, Transaction Aborts, Concurrency control issues, Disk Scheduling Algorithms, Two – phase Approach to improve Predictability – Maintaining Serialization Consistency – Databases for Hard Real Time Systems

UNIT V PROGRAMMING LANGUAGES 9

Assembly language – procedural languages – OO languages – Brief survey of languages – Faults, Failures and bugs – Fault Tolerance – Software integration – refactoring Real time code.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Apply principles of real time systems design.
- Make use of architectures and behavior of real time operating systems and database in real time applications.

REFERENCES

1. C.M. Krishna, Kang G. Shin, “Real-Time Systems”, McGraw-Hill International Editions, 1997.
2. Philip.A.Laplante, “Real Time System Design and Analysis”, Prentice Hall of India, 3rd Edition, 2004.
3. Rajib Mall, “Real-time systems: theory and practice”, Pearson Education, 2009.
4. Stuart Bennett, “Real Time Computer Control-An Introduction”, Prentice Hall of India, 1998.
5. R.J.A Buhur, D.L Bailey, “An Introduction to Real-Time Systems”, Prentice Hall International, 1999.
6. Allen Burns, Andy Wellings, “Real Time Systems and Programming Languages”, Pearson Education, 2003.

IF7071	BIOINFORMATICS	L T P C
		3 0 0 3

OBJECTIVES:

- To learn bio-informatics algorithms

UNIT I **9**

What is Bio-Informatics – Overview- Major databases in Bio Informatics- Molecular biology – Central DogmaData retrieval tools – gene structure - Prokaryotic and Eukaryotic Genomes – Sequence Assembly – Gene mapping – Physical maps – cloning — the genetic material — chemical bonds – molecular biology tools – genomic information content.

UNIT II **9**

DNA: working with single DNA sequence : removing vector sequences- verifying restrictio n maps – PCR design – GC content – counting words – internal repeats – protein coding regions – ORFing – Genomescan Protein: predicting properties – primary structure analysis – transmembrane segments – PROSITE patterns – interpreting scanprosite results- finding domains – CD server results – pfscan results. – Alignment of Pair of Sequences

UNIT III **9**

Phylogenetics – phylogenetic trees –Parsimony – ancestral sequences – strategies for faster searches – consensus trees – tree confidence – comparison of phylogenetic methods – molecular phylogenie. Dot matrix – using scoring matrices – PAM matrices – BLOSUM. - Working with FASTA – Algorithm – output – E-values – Histogram. Working with BLAST – algorithm – output – services – gapped BLAST- PSIBLAST – comparison of FASTA and BLAST. - Multiple sequence alignment - Criteria for Multiple sequence alignment – applications – choosing the right sequences; FASTA, ClustalW, TCoffee methods

UNIT IV **9**

interpreting multiple sequence alignment – getting in right format – converting formats – using Jalview – preparing for publication. - Protein Classification & Structure Prediction - Structure of amino acids – primary structure – secondary structure – folds and motifs – alpha and beta helix – structure based protein classification – protein structure Data bases – folding problem – PROPSEARCH – primary structure analysis and prediction – secondary structure analysis and prediction – motifs – profiles – patterns and fingerprints

UNIT V **9**

Drug Discovery – components – process – Perspectives – Numeric considerations – Algorithms – Hardware – Issues – Protein structure – AbInitio Methods – Heuristic methods – Systems Biology – Tools – Collaboration and Communications – standards - Issues – Security – Intellectual property

TOTAL:45 PERIODS

OUTCOMES:

Upon the completion of this course the student should be able

- To design and implement bio-informatics algorithms

REFERENCES

1. Arthur M. Lesk, “Introduction to Bioinformatics”, Second Edition, Oxford University Press, 2005.
2. T. K. Attwood, D. J. Parry-Smith, and S. Phukan, “Introduction to Bioinformatics”, Pearson Education, 1999.
3. Vittal R. Srinivas, “Bioinformatics – A Modern Approach”, Prentice-Hall of India Pvt. Ltd., 2005
4. S.C Rostogi , Mendiratta, P.Rasogi, “ BioInformatics: methods and applications”,second edition, PHI 2006.
5. Jean Mickel Clavere & Cadrienotredom “Bio Informatics– A beginners guide” Wiley DreamTech, 2003.
6. T.K. Attwood and D.J Perry Smith, “ Introduction to Bio Informatics”, Pearson Education, 1st Edition, 2001.

CP7084	MODELS OF COMPUTATIONS	L	T	P	C
		3	0	0	3

OBJECTIVES

- To understand computation and computability concepts.
- To study different approaches to facilitate computing
- To learn the abstractions of computation and their implementations

UNIT I TURING MACHINE MODEL **9**

Turing Machine Logic, Proof, Computability

UNIT II QUANTUM COMPUTATION **9**

Quantum Computing History, Postulates of Quantum Theory, Dirac Notation, the Quantum Circuit Model, Simple Quantum Protocols: Teleportation, Superdense Coding, Foundation Algorithms

UNIT III NATURE INSPIRED COMPUTING **9**

Nature-Inspired Computing Optimization and Decision Support Techniques, Evolutionary Algorithms, Swarm Intelligence, Benchmarks and Testing

UNIT IV SOCIAL COMPUTING **9**

Social Computing Online communities, Online discussions, Twitter, Social Networking Systems, Web 2.0, social media, Crowdsourcing, Facebook, blogs, wikis, social recommendations, Collective intelligence

UNIT V EVOLUTIONARY COMPUTING **9**

Evolutionary Computing Introduction to Genetic Algorithms, Genetic Operators and Parameters, Genetic Algorithms in Problem Solving, Theoretical Foundations of Genetic Algorithms, Implementation Issues

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify the terminology of the theory of computing
- Predict the major results in computability and complexity theory.
- Prepare the major models of computations

REFERENCES:

1. Danah Boyd, “It's Complicated: The Social Lives of Networked Teens”, Yale University Press, 2015
2. John E. Savage, “Models Of Computation - Exploring the Power of Computing”, Addison-Wesley, 2008
3. Margaret M. Fleck, “Building Blocks for Theoretical Computer Science”, University of Illinois, Urbana-Champaign, 2013.
4. Michael A. Nielsen & Isaac L. Chuang, “Quantum Computation and Quantum Information”, Cambridge University Press, 2010
5. M. Mitchell, “An Introduction to Genetic Algorithms”, Prentice-Hall, 1996.
6. G.Rozenberg, T.Back, J.Kok, Editors, “Handbook of Natural Computing”, Springer Verlag, 2012.

CP7086	NETWORK ON CHIP	L	T	P	C
		3	0	0	3

OBJECTIVES

- To understand the various classes of Interconnection networks
- To learn about different routing techniques for on-chip network
- To know the importance of flow control in on-chip network.
- To learn the building blocks of routers
- To provide an overview of the current state-of-the-art research

UNIT I ICN ARCHITECTURES **9**

Introduction - Classification of ICNs - Topologies - Direct networks - Indirect networks- Performance analysis.

UNIT II SWITCHING TECHNIQUES **9**

Basic switching techniques - Virtual channels - Hybrid switching techniques Optimizing switching techniques - Comparison of switching techniques - Deadlock, livelock and Starvation

UNIT III ROUTING ALGORITHMS **9**

Taxonomy of routing algorithms - Deterministic routing algorithms - Partially adaptive algorithms - Fully adaptive algorithms - Routing in MINs - Routing in switch-based networks with irregular topologies - Resource allocation policies- Flow control.

UNIT IV NETWORK-ON-CHIP **9**

NoC Architectures - Router architecture - Area, energy and reliability constraints - NoC design lternatives - Quality-of Service (QoS) issues in NoC architectures

UNIT V EMERGING TRENDS **9**

Fault-tolerance issues - Emerging on-chip interconnection technologies- 3D NoC- Simulation

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify the major components required to design an on-chip network
- Compare different switching techniques
- Evaluate the performance and the cost of the given on-chip network
- Demonstrate deadlock-free and live lock free routing protocols
- Simulate and assess the performance of a given on-chip network

REFERENCES:

1. Jose Duato, Sudhakar Yalamanchili, Lionel Ni, "Interconnection Networks: An Engineering Approach", Morgan Kaufmann, 2002.
2. William James Dally, Brian Towles, "Principles and Practices of Interconnection Networks", Morgan Kaufmann, 2004.
3. Giovanni De Micheli, Luca Benini, "Networks on Chips: Technology and Tools", Morgan Kaufmann, 2006.
4. Natalie D. Enright Jerger, Li-Shiuan Peh, “On-Chip Networks (Synthesis Lectures on Computer Architecture)”, Morgan and Claypool, 2008.
5. Fayez Gebali, Haytham Elmiligi, Mohamed Wathed El-Kharashi, “Networks-on-Chips: Theory and Practice”, CRC Press, 2009.

CP7090 **SECURE NETWORK SYSTEM DESIGN** **L T P C**
3 0 0 3

OBJECTIVES

- To understand best security practices and how to take advantage of the networking gear that is already available
- To learn design considerations for device hardening, Layer 2 and Layer 3 security issues, denial of service, IPSec VPNs, and network identity
- To understand security design considerations for common applications such as DNS, mail and web
- To identify the key security roles and placement issues for network security elements such as firewalls, intrusion detection systems, VPN gateways, content filtering, as well as for traditional network infrastructure devices such as routers and switches
- To understand the various testing and optimizations strategies to select the technologies and devices for secure network design

UNIT I NETWORK SECURITY FOUNDATIONS 9

Secure network design through modeling and simulation, A fundamental framework for network security, need for user level security on demand, Network Security Axioms, security policies and operations life cycle, security networking threats, network security technologies, general and identity design considerations, network security platform options and best deployment practices, secure network management and network security management

UNIT II IDENTIFYING SYSTEM DESIGNER'S NEEDS AND GOALS 9

Evolution of network security and lessons learned from history, Analyzing top-down network design methodologies, technical goals and tradeoffs – scalability, reliability, availability, Network performance, security, Characterizing the existing internetwork, characterizing network traffic, developing network security strategies

UNIT III PHYSICAL SECURITY ISSUES AND LAYER 2 SECURITY 9

Control physical access to facilities, Control physical access to data centers, Separate identity mechanisms for insecure locations, Prevent password-recovery mechanisms in insecure locations, awareness about cable plant issues, electromagnetic radiation and physical PC security threats, L2 control protocols, MAC flooding considerations, attack mitigations, VLAN hopping attacks, ARP, DHCP, PVLAN security considerations, L2 best practice policies

UNIT IV IP ADDRESSING AND ROUTING DESIGN CONSIDERATIONS 9

Route summarizations, ingress and egress filtering, Non routable networks, ICMP traffic management, Routing protocol security, Routing protocol authentication, transport protocol management policies, Network DoS/flooding attacks

UNIT V TESTING AND OPTIMIZING SYSTEM DESIGN 9

Selecting technologies and devices for network design, testing network design – using industry tests, building a prototype network system, writing and implementing test plan, tools for testing, optimizing network design – network performance to meet quality of service (QoS), Modeling, simulation and behavior analysis of security attacks, future issues in information system security

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Follow the best practices to understand the basic needs to design secure network
- Satisfy the need for user and physical level security on demand for various types of

network attacks
- Use best practice policies for different network layer protocols
- Understand the network analysis, simulation, testing and optimizing of security attacks to provide Quality of Service

REFERENCES:

1. Sumit Ghosh, “Principles of Secure Network System Design”, Springer, 2012.
2. Sean Convery, “Network Security Architectures”, Pearson Education, 2011.
3. Priscilla Oppenheimer, “Top-Down network Design”, Cisco press, Third edition, 2012.
4. Larry L. Peterson, Bruce S. Davie, “Computer Networks: A Systems Approach”, Morgan Kauffmann Publishers Inc., Fifth Revised Edition, 2011.
5. William Stallings, “Cryptography and Network Security Principles and Practices”, Pearson / PHI, Fourth Edition, 2006.
6. Wade Trappe, Lawrence C Washington, “Introduction to Cryptography with Coding Theory”, Second Edition, Pearson, 2007.

CP7079	DOMAIN ENGINEERING	L	T	P	C
		3	0	0	3

OBJECTIVES

- To maximize the satisfaction of the requirements of its stakeholders
- To include the core set of concepts and terminology understood by practitioners in a given area
- To include the knowledge of how to build software systems (or parts of software systems) in a given area
- To evolve consensus amongst the stakeholders

UNIT I DOMAIN ANALYSIS AND SCOPING 9

Domain Analysis. Domain Design and Domain Implementation. Application Engineering. Product-Line Practices. Key Domain Engineering Concepts. Domain. Domain Scope and Scoping. Relationships between Domains. Features and Feature Models. Method Tailoring and Specialization. Survey of Domain Analysis and Domain Engineering Methods

UNIT II DOMAIN ENGINEERING METHODS 9

Feature-Oriented Domain Analysis (FODA). FODA Process. Organization Domain Modeling (ODM). The ODM Process. Draco. Capture. Domain Analysis and Reuse Environment (DARE). Domain-Specific Software Architecture (DSSA) Approach, Algebraic Approach. Other Approaches. Domain Engineering and Related Approaches

UNIT III FEATURE MODELING 9

Feature Modeling. Feature Models. Feature Diagrams. Mandatory Features. Optional Features. Alternative Features. Or-Features. Normalized Feature Diagrams. Expressing Commonality in Feature Diagrams. Expressing Variability in Feature Diagrams. Other Infon-nation Associated with Feature Diagrams in a Feature Model. Assigning Priorities to Variable Features. Availability Sites, Binding Sites, and Binding Modes Sites. Relationship between Optimizations and Availability Sites, Binding Sites, and Binding Modes.

UNIT IV GENERATIVE PROGRAMMING 9

Generative Domain Models. Main Development Steps in Generative Programming. Adapting Domain Engineering for Generative Programming. Domain-Specific Languages.

DEMRAL: Example of a Domain Engineering Method for Generative Programming. Outline of DEMRAL. Domain Analysis. Domain Definition. Domain Modeling. Identification of Key Concepts. Feature Modeling. Feature Starter Set for ADTs. Feature Starter Set for Algorithms. Domain Design. Scope Domain Model for Implementation. Identify Packages

UNIT V LANGUAGES AND TOOLS 9

Hume, DSL Paradigm, Stratego/XT, Run-time Code Generation in C++

TOTAL : 45 PERIODS

OUTCOMES

Upon completion of this course, the student should be able to

- Collect, organize, and store past experience in building systems or parts of systems in a particular domain in the form of reusable assets
- Reuse assets when building new systems
- Develop an architecture for family of systems in a domain and provide production plan

REFERENCES:

1. Krysztof Czarnecki, Ulrich Eisenecker, “Generative programming: Methods, Tools and Applications”, Addition Wesley, 2000.
2. Christian Lengauer, Don Batory, Charles Consel, Martin Odersky, “Domain- Specific Program Generation”, Springer, 2003.

CP7071 ADHOC AND WIRELESS SENSOR NETWORKS

L	T	P	C
3	0	0	3

OBJECTIVES:

- To learn about the issues in the design of wireless ad hoc networks
- To understand the working of protocols in different layers of mobile ad hoc and sensor networks
- To expose the students to different aspects in sensor networks
- To understand various security issues in ad hoc and sensor networks and solutions to the issues

UNIT I MAC & ROUTING IN AD HOC NETWORKS 9

Introduction – Issues and challenges in ad hoc networks – MAC Layer Protocols for wireless ad hoc networks – Contention-Based MAC protocols – MAC Protocols Using Directional Antennas – Multiple-Channel MAC Protocols – Power-Aware MAC Protocols – Routing in Ad hoc Networks – Design Issues – Proactive, Reactive and Hybrid Routing Protocols

UNIT II TRANSPORT & QOS IN AD HOC NETWORKS 9

TCP’s challenges and Design Issues in Ad Hoc Networks – Transport protocols for ad hoc networks – Issues and Challenges in providing QoS – MAC Layer QoS solutions – Network Layer QoS solutions – QoS Model

UNIT III MAC & ROUTING IN WIRELESS SENSOR NETWORKS 9

Introduction – Applications – Challenges – Sensor network architecture – MAC Protocols for wireless sensor networks – Low duty cycle protocols and wakeup concepts – Contention-Based protocols – Schedule-Based protocols –Zig bee – Topology Control – Routing Protocols

UNIT IV TRANSPORT & QOS IN WIRELESS SENSOR NETWORKS 9

Data-Centric and Contention-Based Networking – Transport Layer and QoS in Wireless Sensor Networks – Congestion Control – In-network processing – Operating systems for wireless sensor networks – Examples.

UNIT V SECURITY IN AD HOC AND SENSOR NETWORKS 9

Security Attacks – Key Distribution and Management – Intrusion Detection – Software based Anti-tamper techniques – Water marking techniques – Defense against routing attacks - Secure Ad hoc routing protocols – Broadcast authentication WSN protocols – TESLA – Biba – Sensor Network Security Protocols – SPINS

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify different issues in wireless ad hoc and sensor networks
- Analyze protocols developed for ad hoc and sensor networks
- Identify different issues in wireless ad hoc and sensor networks
- Identify and critique security issues in ad hoc and sensor networks

REFERENCES:

1. Subir Kumar Sarkar, T G Basavaraju, C Puttamadappa, “Ad-Hoc Mobile Wireless Networks”, Auerbach Publications, 2007.
2. Holger Karl, Andreas Willig, “Protocols and Architectures for Wireless Sensor Networks”, Wiley India Private Limited, 2011.
3. Erdal Çayirci ,Chunming Rong, “Security in Wireless Ad Hoc and Sensor Networks”, John Wiley and Sons, 2009.
4. C. Siva Ram Murthy and B.S. Manoj, “Ad Hoc Wireless Networks – Architectures and Protocols”, Pearson Education, 2004.
5. Carlos De Morais Cordeiro, Dharma Prakash Agrawal, “Ad Hoc and Sensor Networks: Theory and Applications”, World Scientific Publishing, Second Edition, 2011.
6. Waltenegus Dargie, Christian Poellabauer, “Fundamentals of Wireless Sensor Networks Theory and Practice”, Wiley India Private Limited, 2014.
7. Adrian Perrig, J.D. Tygar, “Secure Broadcast Communication: In Wired and Wireless Networks”, Kluwer Academic Publishers, Springer, 2002.

CP7080 ETHICAL HACKING

L	T	P	C
3	0	0	3

OBJECTIVES

- To learn about the importance of information security
- To learn different scanning and enumeration methodologies and tools
- To understand various hacking techniques and attacks
- To be exposed to programming languages for security professionals
- To get familiarized with the different phases in penetration testing

UNIT I INTRODUCTION TO HACKING 9

Introduction to Hacking – Importance of Security – Elements of Security – Phases of an Attack – Types of Hacker Attacks – Hacktivism – Vulnerability Research – Introduction to Footprinting – Information Gathering Methodology – Footprinting Tools – WHOIS Tools – DNS Information Tools – Locating the Network Range – Meta Search Engines.

UNIT II SCANNING AND ENUMERATION 9

Introduction to Scanning – Objectives – Scanning Methodology – Tools – Introduction to Enumeration – Enumeration Techniques – Enumeration Procedure – Tools.

UNIT III SYSTEM HACKING 9

Introduction – Cracking Passwords – Password Cracking Websites – Password Guessing – Password Cracking Tools – Password Cracking Counter measures – Escalating Privileges – Executing Applications – Keyloggers and Spyware.

UNIT IV PROGRAMMING FOR SECURITY PROFESSIONALS 9

Programming Fundamentals – C language – HTML – Perl – Windows OS Vulnerabilities – Tools for Identifying Vulnerabilities – Countermeasures – Linux OS Vulnerabilities – Tools for Identifying Vulnerabilities – Countermeasures.

UNIT V PENETRATION TESTING 9

Introduction – Security Assessments – Types of Penetration Testing- Phases of Penetration Testing – Tools – Choosing Different Types of Pen-Test Tools – Penetration Testing Tools

TOTAL : 45 PERIODS

OUTCOMES

Upon completion of this course, the student should be able to

- Defend hacking attacks and protect data assets
- Defend a computer against a variety of security attacks using various tools
- Practice and use safe techniques on the World Wide Web

REFERENCES:

1. Ec-Council, “Ethical Hacking and Countermeasures: Attack Phases”, Delmar Cengage Learning, 2009.
2. Michael T. Simpson, Kent Backman, James E. Corley, “Hands-On Ethical Hacking and Network Defense”, Cengage Learning, 2012.
3. Patrick Engebretson, “The Basics of Hacking and Penetration Testing – Ethical Hacking and Penetration Testing Made Easy”, Syngress Media, Second Revised Edition, 2013.
4. Jon Erickson, “Hacking: The Art of Exploitation”, No Starch Press, Second Edition, 2008.

CP7078 DIGITAL IMAGE PROCESSING AND APPLICATIONS

L	T	P	C
3	0	0	3

OBJECTIVES

- To understand the basic concepts of digital image processing and various image transforms
- To familiarize the student with the image processing facilities in Matlab
- To expose the student to a broad range of image processing techniques and their applications, and to provide the student with practical experience using them
- To appreciate the use of current technologies those are specific to image processing systems
- To expose the students to real-world applications of Image Processing

UNIT I FUNDAMENTALS OF IMAGE PROCESSING 9

Introduction – Steps in Digital Image Processing – Image sampling and Quantization – Basic relationships between pixels – Color Fundamentals – File Formats – Image Transforms: DFT, DCT, Haar, SVD and KL- Introduction to Matlab Toolbox

UNIT II IMAGE ENHANCEMENT AND RESTORATION 9

Image Enhancement in the Spatial Domain: Basic Gray Level Transformations, Histogram Processing, Enhancement Using Arithmetic/Logic Operations, Spatial Filtering , Fuzzy sets for spatial filters – Image Enhancement in the Frequency Domain: Frequency Domain Filters - Image Restoration: Model of Image Degradation/Restoration Process, Noise Models, Linear and non linear image restoration techniques, Blind Deconvolution

UNIT III MULTI RESOLUTION ANALYSIS AND IMAGE COMPRESSION 9

Multi Resolution Analysis: Image Pyramids – Multi resolution expansion – Fast Wavelet Transforms, Lifting scheme. Image Compression: Fundamentals – Models – Elements of Information Theory – Error Free Compression – Lossy Compression-wavelet based image compression techniques – Compression standards-JPEG/MPEG, Video compression

UNIT IV IMAGE SEGMENTATION AND DESCRIPTION 9

Image Segmentation: Detection of Discontinuities, Edge Linking and Boundary Detection, Thresholding, Region Based Segmentation, Basic Morphological Algorithms, Morphological Water Sheds - Description: Boundary Descriptors, Regional Descriptors

UNIT V CURRENT TRENDS AND APPLICATIONS OF IMAGE PROCESSING 9

Applications: Image Classification, Object Recognition, Image Fusion, Steganography – Current Trends: Color Image Processing, Wavelets in Image Processing

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Have a clear impression of the breadth and practical scope of digital image processing and have arrived at a level of understanding that is the foundation for most of the work currently underway in this field
- Implement basic image processing algorithms using MATLAB tools
- Explore advanced topics of Digital Image Processing, Ability to Apply and develop new techniques in the areas of image enhancement-restoration segmentation-compression-wavelet processing and image morphology
- Make a positive professional contribution in the field of Digital Image Processing

REFERENCES:

1. Rafael C.Gonzalez, Richard E.Woods, “Digital Image Processing”, Pearson Education, Third Edition, 2013.
2. S. Sridhar, “Digital Image Processing”, Oxford University Press, 2011.
3. Milan Sonka, Vaclav Hlavac, Roger Boyle, “Image Processing, Analysis and Machine Vision”, Brooks Cole, Third Edition, 2014.
4. Anil K.Jain, “Fundamentals of Digital Image Processing”, Prentice-Hall India, 1994.
5. Madhuri A. Joshi, ‘Digital Image Processing: An Algorithmic Approach”, Prentice-Hall India, 2006.
6. Rafael C.Gonzalez, Richard E.Woods, Steven L. Eddins, “Digital Image Processing Using MATLAB”, McGraw Hill Education, Second Edition, 2010.

CP7091 SERVICE ORIENTED ARCHITECTUREAND DESIGN L T P C
3 0 0 3

OBJECTIVES

- To understand the SOA architecture
- To understand the service oriented analysis and design
- To understand the development of deployment of web services
- To understand the security issues of SOA

UNIT I SOA FUNDAMENTALS 9

Principles of Service Orientation - Client-Server Architecture - Distributed Internet Architecture - SOA Characteristics - Anatomy of SOA - Components - Interaction - Technical and Business Benefits - Multi-channel access - Business Process Management

UNIT II SOA AND WEB SERVICES 9

Web Service Platform - Web Service Description - Service Contracts - Service Level Data Model - Service Discovery - Service Level Security - Service Level Interaction Patterns: SOAP basics - Messaging with SOAP - Message Exchange Patterns - Web WSDL basics, Writing a Java Web Service, writing a Java Web Service Client ,Describing Web Services: WSDL, Representing Data Types - XML Schema, Communicating Object Data, SOAP Related Technologies

UNIT III SERVICE ORIENTED ANALYSIS AND DESIGN 9

Design principles - Business Centric SOA - Deriving Business services - Service Modeling - Coordination - Atomic Transaction - Business activities - Web Service Orchestration Business Process Execution Language (BPEL) - Choreography - Metadata Management- Entity centric business service design - Application Service design - Task centric business service design

UNIT IV WEB SERVICES DEVELOPMENT AND DEPLOYMENT 9

XML and Web Services - WSDL basics - SOA support in J2EE - Java API for XML-based Web Services (JAX-WS) - Java Architecture for XML Binding (JAXB) - Java API for XML Registries (JAXR) - Web Services Interoperability Technologies - SOA support in .NET - Common Language Runtime - ASP.NET - Web forms - ASP.NET Web Services - Web Services Enhancements

UNIT V SOA APPLICATIONS AND SECURITY 9

Security Overview: e-commerce based security (public key cryptography) – Public key encryption – Security issues in XML document – SOAP security issue – XML Security framework: XML Digital Signature (Enveloped, enveloping and detached) – Signature validation - XML Encryption – Types – Canonicalization - XML Key management.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Develop and deploy simple and composite web services with SOA design principles considering the security issues
- Understand and describe the standards and technologies of modern web service implementations
- Efficiently use leading development tools to create and consume web services
- Implement a service oriented application

REFERENCES:

1. Eric Newcomer, Greg Lomow, “Understanding SOA with Web Services”, Pearson Education, 2004.
2. Thomas Erl, “Service Oriented Architecture: Concepts, Technology, and Design”, Pearson

Education, 2006.
3. Shankar Kambhampaly, “Service Oriented Architecture for Enterprise Applications”, Wiley India Pvt Ltd, 2008.
4. Mark O’ Neill, “Web Services Security”, Tata McGraw-Hill Edition, 2003.
5. Frank Cohen, “Fast SOA”, Morgan Kaufmann, 2010.
6. Sandeep Chatterjee, James Webber, “Developing Enterprise Web Services”, Pearson Education, 2003.

CP7252 COMPILER OPTIMIZATION TECHNIQUES L T P C
3 0 0 3

OBJECTIVES

- To understand different forms of intermediate languages and analyzing programs
- To understand optimizations techniques for single program blocks
- To apply optimizations on procedures and low level code
- To explore and enhance inter procedural optimizations
- To enhance resource utilization

UNIT I INTERMEDIATE REPRESENTATION OF PROGRAMS AND ANALYSIS 9

Structure of an Optimizing Compiler – Compiler Construction tools - LIR, MIR, HIR – DAG – Syntax Tree – Postfix – Control Flow Analysis – Iterative Data Flow Analysis – Static Single Assignment – Basic Block Dependence DAGs – Alias Analysis.

UNIT II LOCAL AND LOOP OPTIMIZATIONS 9

Early Optimizations: Constant-Expression Evaluation - Scalar Replacement of Aggregates - Algebraic Simplifications and Re-association - Value Numbering - Copy Propagation - Sparse Conditional Constant Propagation. Redundancy Elimination: Common - Subexpression Elimination - Loop-Invariant Code Motion - Partial-Redundancy Elimination - Redundancy Elimination and Reassociation - Code Hoisting. Loop Optimizations: Induction Variable Optimizations - Unnecessary Bounds Checking Elimination.

UNIT III PROCEDURE OPTIMIZATION AND SCHEDULING 9

Procedure Optimizations: Tail-Call Optimization and Tail-Recursion Elimination - Procedure Integration - In-Line Expansion - Leaf-Routine Optimization and Shrink Wrapping. Code Scheduling: Instruction Scheduling - Speculative Loads and Boosting - Speculative Scheduling - Software Pipelining - Trace Scheduling - Percolation Scheduling. Control-Flow and Low-Level Optimizations : Unreachable-Code Elimination - Straightening - If Simplifications - Loop Simplifications -Loop Inversion – Un-switching - Branch Optimizations - Tail Merging or Cross Jumping - Conditional Moves - Dead-Code Elimination - Branch Prediction - Machine Idioms and Instruction Combining.

UNIT IV INTER PROCEDURAL OPTIMIZATION 9

Symbol table – Runtime Support - Interprocedural Analysis and Optimization: Interprocedural Control-Flow Analysis - The Call Graph - Interprocedural Data-Flow Analysis - Interprocedural Constant Propagation - Interprocedural Alias Analysis - Interprocedural Optimizations - Interprocedural Register Allocation - Aggregation of Global References.

UNIT V OPTIMIZING FOR MEMORY 9

Register Allocation: Register Allocation and Assignment - Local Methods - Graph Coloring – Priority Based Graph Coloring - Other Approaches to Register Allocation. Optimization for the Memory Hierarchy: Impact of Data and Instruction Caches - Instruction-Cache Optimization - Scalar Replacement of Array Elements - Data-Cache Optimization - Scalar vs. Memory-Oriented Optimizations.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify the different optimization techniques that are possible for a sequence of code
- Design performance enhancing optimization techniques
- Manage procedures with optimal overheads
- Ensure better utilization of resources

REFERENCES:

1. Steven Muchnick, “Advanced Compiler Design and Implementation”, Morgan Kaufman Publishers, 1997.
2. Alfred V. Aho, Ravi Sethi, Jeffrey D. Ullman, "Compilers: Principles, Techniques, and Tools", Addison Wesley, Second Edition, 2007.
3. Andrew W. Appel, Jens Palsberg, "Modern Compiler Implementation in Java", Cambridge University Press, Second Edition, 2002.
4. Keith Cooper, Linda Torczon, "Engineering a Compiler", Morgan Kaufmann, Second Edition, 2011.
5. Randy Allen and Ken Kennedy, “Optimizing Compilers for Modern Architectures: A Dependence based Approach”, Morgan Kaufman, 2001.

CP7155 NETWORKING TECHNOLOGIES L T P C
3 0 0 3

OBJECTIVES

- To learn about integrated and differentiated services architectures
- To understand the working of wireless network protocols
- To study the evolution made in cellular networks
- To get familiarized with next generation networks

UNIT I NETWORK ARCHITECTURE AND QoS 9

Overview of TCP/IP Network Architecture – Integrated Services Architecture – Approach – Components – Services – Queuing Discipline – FQ – PS – BRFQ – GPS – WFQ – Random Early Detection – Differentiated Services.

UNIT II WIRELESS NETWORKS 9

IEEE802.16 and WiMAX – Security – Advanced 802.16 Functionalities – Mobile WiMAX - 802.16e – Network Infrastructure – WLAN – Configuration – Management Operation – Security – IEEE 802.11e and WMM – QoS – Comparison of WLAN and UMTS – Bluetooth – Protocol Stack – Security – Profiles

UNIT III CELLULAR NETWORKS 9

GSM – Mobility Management and call control – GPRS – Network Elements – Radio Resource Management – Mobility Management and Session Management – Small Screen Web Browsing over GPRS and EDGE – MMS over GPRS – UMTS – Channel Structure on the Air Interface – UTRAN –Core and Radio Network Mobility Management – UMTS Security

UNIT IV 4G NETWORKS 9

LTE – Network Architecture and Interfaces – FDD Air Interface and Radio Networks –Scheduling – Mobility Management and Power Optimization – LTE Security Architecture – Interconnection with UMTS and GSM – LTE Advanced (3GPPP Release 10) - 4G Networks and Composite Radio Environment – Protocol Boosters – Hybrid 4G Wireless Networks Protocols – Green Wireless Networks – Physical Layer and Multiple Access – Channel Modelling for 4G – Introduction to 5G

UNIT V SOFTWARE DEFINED NETWORKS 9

Introduction – Centralized and Distributed Control and Data Planes – Open Flow – SDN Controllers – General Concepts – VLANs – NVGRE – Open Flow – Network Overlays – Types – Virtualization – Data Plane – I/O – Design of SDN Framework

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify the different features of integrated and differentiated services
- Demonstrate various protocols of wireless and cellular networks
- Discuss the features of 4G and 5G networks

REFERENCES:

1. William Stallings, “High Speed Networks and Internets: Performance and Quality of Service”, Prentice Hall, Second Edition, 2002.
2. Martin Sauter, "From GSM to LTE, An Introduction to Mobile Networks and Mobile Broadband", Wiley, 2014.
3. Savo G Glisic, “Advanced Wireless Networks – 4G Technologies”, John Wiley & Sons, 2007.
4. Jonathan Rodriguez, “Fundamentals of 5G Mobile Networks”, Wiley, 2015.
5. Martin Sauter, “Beyond 3G - Bringing Networks, Terminals and the Web Together: LTE, WiMAX, IMS, 4G Devices and the Mobile Web 2.0”, Wiley, 2009.
6. Naveen Chilamkurti, Sherali Zeadally, Hakima Chaouchi, “Next-Generation Wireless Technologies”, Springer, 2013.
7. Erik Dahlman, Stefan Parkvall, Johan Skold, “4G: LTE/LTE-Advanced for Mobile Broadband”, Academic Press, 2013.
8. Thomas D.Nadeau and Ken Gray, “SDN – Software Defined Networks”, O’Reilly Publishers, 2013.

CP70783 COGNITIVE SCIENCE

L	T	P	C
3	0	0	3

OBJECTIVES:

- To learn the basics of Cognitive Science with focus on acquisition, representation, and use of knowledge by individual minds, brains, and machines
- To study the mind and intelligence, embracing psychology, artificial intelligence, neuroscience and linguistics
- To understand the role of neuro science in the cognitive field

UNIT I INTRODUCTION TO COGNITIVE SCIENCE 9

The Cognitive view –Some Fundamental Concepts – Computers in Cognitive Science – Applied Cognitive Science – The Interdisciplinary Nature of Cognitive Science – Artificial Intelligence: Knowledge representation -The Nature of Artificial Intelligence - Knowledge Representation – Artificial Intelligence: Search, Control, and Learning

UNIT II COGNITIVE PSYCHOLOGY 9

Cognitive Psychology – The Architecture of the Mind - The Nature of Cognitive Psychology- A Global View of The Cognitive Architecture- Propositional Representation- Schematic Representation- Cognitive Processes, Working Memory, and Attention- The Acquisition of Skill- The Connectionist Approach to Cognitive Architecture

UNITIII COGNITIVE NEUROSCIENCE 9

Brain and Cognition Introduction to the Study of the Nervous System – Neural Representation – Neuropsychology- Computational Neuroscience - The Organization of the mind - Organization of Cognitive systems - Strategies for Brain mapping – A Case study: Exploring mindreading

UNIT IV LANGUAGE ACQUISITION, SEMANTICS AND PROCESSING MODELS 9

Milestones in Acquisition – Theoretical Perspectives- Semantics and Cognitive Science – Meaning and Entailment – Reference – Sense – Cognitive and Computational Models of Semantic Processing – Information Processing Models of the Mind- Physical symbol systems and language of thought- Applying the Symbolic Paradigm- Neural networks and distributed information processing- Neural network models of Cognitive Processes

UNITV HIGHER-LEVEL COGNITION 9

Reasoning – Decision Making – Computer Science and AI: Foundations & Robotics – New Horizons - Dynamical systems and situated cognition- Challenges – Emotions and Consciousness – Physical and Social Environments – Applications

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Explain and analyze the major concepts, philosophical and theoretical perspectives, empirical findings, and historical trends in cognitive science, related to cultural diversity and living in a global community.
- Use cognitive science knowledge base to create their own methods for answering novel questions of either a theoretical or applied nature, and to critically evaluate the work of others in the same domain
- Be proficient with basic cognitive science research methods, including both theory-driven and applied research design, data collection, data analysis, and data interpretation

REFERENCES

1. Neil Stillings, Steven E. Weisler, Christopher H. Chase and Mark H. Feinstein, “Cognitive Science: An Introduction”, Second Edition, MIT press ,1995
2. José Luis Bermúdez, “Cognitive Science: An Introduction to the Science of the Mind”, Cambridge University Press, New York, 2014
3. Robert L. Solso, Otto H. MacLin and M. Kimberly MacLin, “Cognitive Psychology, Pearson Education, 2007.
4. J. Friedenberg and G. Silverman, “Cognitive Science: An Introduction to the Study of Mind”, 2011
5. Steven Pinker, “How the mind works”, W. W. Norton & Company; Reissue edition, 2009
6. Carolyn Panzer Sobel and Paul Li, “Cognitive Science: An Interdisciplinary Approach”, 2013
7. Paul Thagard, “Mind: Introduction to Cognitive Science”, 2nd Edition, MIT Press, 2005

CP7082 INFORMATION RETRIEVAL TECHNIQUES L T P C
3 0 0 3

OBJECTIVES:

- To understand the basics of information retrieval with pertinence to modeling, query operations and indexing
- To get an understanding of machine learning techniques for text classification and clustering.
- To understand the various applications of information retrieval giving emphasis to multimedia IR, web search
- To understand the concepts of digital libraries

UNIT I INTRODUCTION: MOTIVATION 9

Basic Concepts – Practical Issues - Retrieval Process – Architecture - Boolean Retrieval – Retrieval Evaluation – Open Source IR Systems–History of Web Search – Web Characteristics– The impact of the web on IR ––IR Versus Web Search–Components of a Search engine

UNIT II MODELING 9

Taxonomy and Characterization of IR Models – Boolean Model – Vector Model - Term Weighting – Scoring and Ranking –Language Models – Set Theoretic Models - Probabilistic Models – Algebraic Models – Structured Text Retrieval Models – Models for Browsing

UNIT III INDEXING 9

Static and Dynamic Inverted Indices – Index Construction and Index Compression. Searching - Sequential Searching and Pattern Matching. Query Operations -Query Languages – Query Processing - Relevance Feedback and Query Expansion - Automatic Local and Global Analysis – Measuring Effectiveness and Efficiency

UNIT IV CLASSIFICATION AND CLUSTERING 9

Text Classification and Naïve Bayes – Vector Space Classification – Support vector machines and Machine learning on documents. Flat Clustering – Hierarchical Clustering –Matrix decompositions and latent semantic indexing – Fusion and Meta learning

UNIT V SEARCHING THE WEB 9

Searching the Web –Structure of the Web –IR and web search – Static and Dynamic Ranking – Web Crawling and Indexing – Link Analysis - XML Retrieval Multimedia IR: Models and Languages – Indexing and Searching Parallel and Distributed IR – Digital Libraries

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Build an Information Retrieval system using the available tools
- Identify and design the various components of an Information Retrieval system
- Apply machine learning techniques to text classification and clustering which is used for efficient Information Retrieval
- Design an efficient search engine and analyze the Web content structure

REFERENCES:

1. Ricardo Baeza – Yates, Berthier Ribeiro – Neto, “Modern Information Retrieval: The concepts and Technology behind Search” (ACM Press Books), Second Edition, 2011.
2. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schutze, “Introduction to Information Retrieval”, Cambridge University Press, First South Asian Edition, 2008.

3. Stefan Buttcher, Charles L. A. Clarke, Gordon V. Cormack, “Information Retrieval Implementing and Evaluating Search Engines”, The MIT Press, Cambridge, Massachusetts London, England, 2010.

CP7083 INTERNET OF THINGS IN THE CLOUD L T P C
3 0 0 3

OBJECTIVES:

- To understand the basics of Internet of things and protocols
- To get an idea of some of the application areas where Internet of Things can be applied
- To understand the middleware for Internet of Things
- To understand the concepts of Web of Things
- To understand the concepts of Cloud of Things with emphasis on Mobile cloud computing

UNIT I INTRODUCTION 10

Definitions and Functional Requirements –Motivation – Architecture - Web 3.0 View of IoT– Ubiquitous IoT Applications – Four Pillars of IoT – DNA of IoT - The Toolkit Approach for End-user Participation in the Internet of Things. Middleware for IoT: Overview – Communication middleware for IoT –IoT Information Security

UNIT II IOT PROTOCOLS 8

Protocol Standardization for IoT – Efforts – M2M and WSN Protocols – SCADA and RFID Protocols – Issues with IoT Standardization – Unified Data Standards – Protocols – IEEE 802.15.4–BACNet Protocol– Modbus – KNX – Zigbee– Network layer – APS layer – Security

UNIT III WEB OF THINGS 10

Web of Things versus Internet of Things – Two Pillars of the Web – Architecture Standardization for WoT– Platform Middleware for WoT – Unified Multitier WoT Architecture – WoT Portals and Business Intelligence. Cloud of Things: Grid/SOA and Cloud Computing – Cloud Middleware – Cloud Standards – Cloud Providers and Systems – Mobile Cloud Computing – The Cloud of Things Architecture.

UNIT IV INTEGRATED 9

Integrated Billing Solutions in the Internet of Things Business Models for the Internet of Things - Network Dynamics: Population Models – Information Cascades - Network Effects – Network Dynamics: Structural Models - Cascading Behavior in Networks - The Small-World Phenomenon

UNIT V APPLICATIONS 8

The Role of the Internet of Things for Increased Autonomy and Agility in Collaborative Production Environments - Resource Management in the Internet of Things: Clustering, Synchronisation and Software Agents. Applications - Smart Grid – Electrical Vehicle Charging

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify and design the new models for market strategic interaction
- Design business intelligence and information security for WoB
- Analyze various protocols for IoT
- Design a middleware for IoT
- Analyze and design different models for network dynamics

REFERENCES

1. Honbo Zhou, “The Internet of Things in the Cloud: A Middleware Perspective”, CRC Press, 2012.
2. Dieter Uckelmann, Mark Harrison, Michahelles, Florian (Eds), “Architecting the Internet of Things”, Springer, 2011.
3. David Easley and Jon Kleinberg, “Networks, Crowds, and Markets: Reasoning About a Highly Connected World”, Cambridge University Press, 2010.
4. Olivier Hersent, Omar Elloumi and David Boswarthick, “The Internet of Things: Applications to the Smart Grid and Building”, Wiley, 2012.
5. Olivier Hersent, David Boswarthick, Omar Elloumi , “The Internet of Things – Key applications and Protocols”, Wiley, 2012.

CP7088 PARALLEL AND DISTRIBUTED DATABASES L T P C
3 0 0 3

OBJECTIVES

- To realize the need of parallel processing
- To cater to applications that require a system capable of sustaining trillions of operations per second on very large data sets
- To understand the need of data integration over data centralization

UNIT I INTRODUCTION TO PARALLEL DATABASES 9

Need of Parallelism - Forms of parallelism – architecture – Analytical models. Basic Query Parallelism – Parallel Search- Parallel sort and Group By- Parallel Join

UNIT II ADVANCED QUERY PROCESSING IN PARALLEL DATABASES 9

Parallel indexing. Parallel Universal Qualification – Collection Join Queries. Parallel Query Scheduling – Optimization, Applications

UNIT III INTRODUCTION TO DISTRIBUTED DATABASES 9

Overview - Promises of DDB –Design Issues – DDB Design – DDB Integration – Data and Access Control

UNIT IV QUERY PROCESSING IN DISTRIBUTED DATABASES 9

Overview- of Query Processing – Query Decomposition and Data Localization – Optimization of Distributed Queries, Multi-database Query Processing

UNIT V TRANSACTION MANAGEMENT AND OTHER ADVANCED SYSTEMS 9

Introduction – Concurrency Control - Distributed DBMS Reliability – Data Replication – DDB Applications, Distributed Object Database Management – Peer -to-Peer Data Management – Web Data Management – Streaming Data and Cloud Computing

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Get good knowledge on the need, issues, design and application of both parallel and distributed databases
- Know how to write optimal queries to cater to applications that need these forms of databases
- Fragment, replicate and localize their data as well as their queries to get their work done faster
- Get idea on other similar trends of optimal data processing

REFERENCES:

1. David Taniar, Clement H.C.Leung, Wenny Rahayu, Sushant Goel , “High Performance Parallel Database Processing and Grid Databases” (Wiley Series in Parallel and Distributed Computing), Wiley-Blackwell, 2008.
2. M. Tamer Ozsu and Patrick Valduriez, “Principles of Distributed Database Systems”, Springer Science + Business Media , Third Edition, 2011.

CP7094 STATISTICAL NATURAL LANGUAGE PROCESSING L T P C
3 0 0 3

OBJECTIVES:

- To understand the representation and processing of Morphology and Part-of Speech Taggers
- To appreciate various techniques used for speech synthesis and recognition
- To understand different aspects of natural language syntax and the various methods used for processing syntax and disambiguating word senses
- To appreciate the various representations of semantics and discourse
- To know about various applications of natural language processing

UNIT I MORPHOLOGY AND PART-OF SPEECH PROCESSING 9

Introduction –Regular Expressions and Automata- Non-Deterministic FSAs. Tranducers – English Morphology - Finite-State Morphological Parsing - Porter Stemmer - Tokenization- Detection and Correction of Spelling Errors. N-grams – Perplexity - Smoothing - Interpolation - Backoff . Part-of- Speech Tagging – English Word Classes - Tagsets - Rule-Based - HMM - Transformation-Based Tagging - Evaluation and Error Analysis. Hidden Markov and Maximum Entropy Models

UNIT II SPEECH PROCESSING 9

Phonetics – Articulatory Phonetics - Phonological Categories - Acoustic Phonetics and Signals - Speech Synthesis – Text Normalization – Phonetic and Acoustic Analysis - Diphone Waveform synthesis – Evaluation- Automatic Speech Recognition –Architecture - Hidden Markov Model to Speech - MFCC vectors - Acoustic Likelihood Computation - Evaluation. Triphones – Discriminative Training - Modeling Variation. Computational Phonology-Finite-State Phonology – Computational Optimality Theory - Syllabification - Learning Phonology and Morphology

UNIT III SYNTAX ANALYSIS 9

Formal Grammars of English – Constituency - Context-Free Grammars –Grammar Rules – Treebanks - Finite-State and Context-Free Grammars - Dependency Grammars. Syntactic Parsing – Parsing as Search - Ambiguity - Dynamic Programming Parsing Methods –CKY-Earley and Chart Parsing- Partial Parsing-Evaluation. Statistical Parsing – Probabilistic Context-Free Grammars – Probabilistic CKY Parsing of PCFGs –Probabilistic Lexicalized CFGs –Collins Parser. Language and Complexity -The Chomsky Hierarchy -The Pumping Lemma

UNIT IV SEMANTIC AND PRAGMATIC INTERPRETATION 9

Representation of Meaning – Desirable Properties - Computational Semantics -Word Senses - Relations Between Senses – WorldNet - Event Participants- Proposition Bank -Frame Net -– Metaphor. Computational Lexical Semantics – Word Sense Disambiguation- Supervised Word Sense Disambiguation - Dictionary and Thesaurus Methods- Word Similarity - Minimally Supervised WSD - Hyponymy and Other Word Relations - Semantic Role Labeling - Unsupervised Sense Disambiguation. Computational Discourse - Discourse Segmentation - Unsupervised Discourse - Segmentation - Text Coherence - Reference Resolution –Phenomena – Features and algorithms - Pronominal Anaphora Resolution

UNIT V APPLICATIONS 9

Information Extraction – Named Entity Recognition - Relation Detection and Classification – Temporal and Event Processing - Template-Filling - Biomedical Information Extraction. Question Answering and Summarization -Information Retrieval -Factoid Question Answering - Summarization - Single and Multi-Document Summarization - Focused Summarization - Evaluation. Dialog and Conversational Agents – Properties of Human Conversations - Basic Dialogue Systems - VoiceXML - Information- State and Dialogue Acts - Markov Decision Process Architecture. Machine Translation –Issues in Machine Translation - Classical MT and the Vauquois Triangle -Statistical MT - Phrase-Based Translation Model - Alignment in MT –IBM Models –Evaluation

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Identify the different linguistic components of given sentences
- Design a morphological analyser for a language of your choice using finite state automata concepts
- Implement the Earley algorithm for a language of your choice by providing suitable grammar and words
- Use a machine learning algorithm for word sense disambiguation
- Build a tagger to semantically tag words using WordNet
- Design a business application that uses different aspects of language processing

REFERENCES:

1. Jurafsky and Martin, “Speech and Language Processing”, Pearson Prentice Hall, Second Edition, 2008.
2. Christopher D. Manning and Hinrich Schütze, “Foundations of Statistical Natural Language Processing”, MIT Press, 1999.
3. Stevan Bird, “Natural Language Processing with Python”, Shroff, 2009.
4. James Allen, “Natural Language Understanding”, Addison Wesley, Second Edition, 2007.
5. Nitin Indurkhya, Fred J. Damerau, “Handbook of Natural Language Processing”, (Chapman & Hall/CRC Machine Learning & Pattern Recognition), Second Edition, 2010.
6. Alexander Clark, Chris Fox, Shalom Lappin, “The Handbook of Computational Linguistics and Natural Language Processing”, Wiley-Blackwell, 2012.

CP7093 SOFT COMPUTING L T P C
3 0 0 3

OBJECTIVES:

- To learn the key aspects of Soft computing and Neural networks
- To study the fuzzy logic components
- To gain insight onto neuro fuzzy modeling and control
- To know about the components and building block hypothesis of genetic algorithm
- To gain knowledge in machine learning through Support Vector Machines

UNIT I INTRODUCTION TO SOFT COMPUTING 9

Evolution of Computing - Soft Computing Constituents – From Conventional AI to Computational Intelligence - Machine Learning Basics

UNIT II GENETIC ALGORITHMS 9

Introduction to Genetic Algorithms (GA) – Applications of GA - Building block hypothesis-Representation – Fitness Measures – Genetic Operators-. GA based Machine Learning.

UNIT III NEURAL NETWORKS 9

Machine Learning using Neural Network, Adaptive Networks – Feed Forward Networks – Supervised Learning Neural Networks – Radial Basis Function Networks - Reinforcement Learning – Unsupervised Learning Neural Networks – Adaptive Resonance Architectures – Advances in Neural Networks.

UNIT IV FUZZY LOGIC 9

Fuzzy Sets – Operations on Fuzzy Sets – Fuzzy Relations – Membership Functions-Fuzzy Rules and Fuzzy Reasoning – Fuzzy Inference Systems – Fuzzy Expert Systems – Fuzzy Decision Making.

UNIT V NEURO-FUZZY MODELING 9

Adaptive Neuro-Fuzzy Inference Systems – Coactive Neuro-Fuzzy Modeling – Classification and Regression Trees – Data Clustering Algorithms – Rule base Structure Identification – Neuro-Fuzzy Control – Case Studies.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of this course, the student should be able to

- Discuss on machine learning through neural networks
- Apply knowledge in developing a Fuzzy expert system
- Model Neuro Fuzzy system for clustering and classification
- Discover knowledge to develop Genetic Algorithm and Support vector machine based machine learning system

REFERENCES:

1. Jyh-Shing Roger Jang, Chuen-Tsai Sun, Eiji Mizutani, “Neuro-Fuzzy and Soft Computing”, Prentice-Hall of India, 2002.
2. Kwang H.Lee, “First course on Fuzzy Theory and Applications”, Springer, 2005.
3. George J. Klir and Bo Yuan, “Fuzzy Sets and Fuzzy Logic-Theory and Applications”, Prentice Hall, 1996.
4. James A. Freeman and David M. Skapura, “Neural Networks Algorithms, Applications, and Programming Techniques”, Addison Wesley, 2003.
5. David E.Goldberg, “Genetic Algorithms in Search, Optimization and Machine Learning”, Addison Wesley, 1989.
6. Mitchell Melanie, “An Introduction to Genetic Algorithm”, MIT Press, 1996.
7. S.N.Sivanandam, S.N.Deepa, “Introduction to Genetic Algorithms”, Springer, 2008 edition.

BD7004 DATA VISUALIZATION L T P C

3 0 0 3

OBJECTIVES:

- To understand how accurately represent voluminous complex data set in web and from other data sources
- To understand the methodologies used to visualize large data sets
- To understand the process involved in data visualization and security aspects involved in data visualization

UNIT I INTRODUCTION 9

Context of data visualization – Definition, Methodology, Visualization design objectives. Key Factors – Purpose, visualization function and tone, visualization design options – Data representation, Data Presentation, Seven stages of data visualization, widgets, data visualization tools.

UNIT II VISUALIZING DATA METHODS 9

Mapping - Time series - Connections and correlations - Scatterplot maps - Trees, Hierarchies and Recursion - Networks and Graphs, Info graphics

UNIT III VISUALIZING DATA PROCESS 9

Acquiring data, - Where to Find Data, Tools for Acquiring Data from the Internet, Locating Files for Use with Processing, Loading Text Data, Dealing with Files and Folders, Listing Files in a Folder, Asynchronous Image Downloads, Advanced Web Techniques, Using a Database, Dealing with a Large Number of Files. Parsing data - Levels of Effort, Tools for Gathering Clues, Text Is Best, Text Markup Languages, Regular Expressions (regexps), Grammars and BNF Notation, Compressed Data, Vectors and Geometry, Binary Data Formats, Advanced Detective Work.

UNIT IV INTERACTIVE DATA VISUALIZATION 9

Drawing with data – Scales – Axes – Updates, Transition and Motion – Interactivity - Layouts – Geomapping – Exporting, Framework – T3, .js, tablo.

UNIT V SECURITY DATA VISUALIZATION 9

Port scan visualization - Vulnerability assessment and exploitation - Firewall log visualization - Intrusion detection log visualization -Attacking and defending visualization systems - Creating security visualization system.

TOTAL : 45 PERIODS

OUTCOMES:

Upon completion of the course, the students will be able to

- Design and use various methodologies present in data visualization
- Discuss the process involved and security issues present in data visualization

REFERENCES:

1. Scott Murray, “Interactive data visualization for the web”, O’Reilly Media, Inc., 2013.
2. Ben Fry, “Visualizing Data”, O’Reilly Media, Inc., 2007.
3. Greg Conti, “Security Data Visualization: Graphical Techniques for Network Analysis”, No Starch Press Inc, 2007.

BD7003 DATA INTENSIVE COMPUTING L T P C
3 0 0 3

OBJECTIVES

- To understand the basics of the various database systems including databases for Big data
- To learn about the architecture of data intensive computing
- To learn about parallel processing for data intensive computing
- To learn about the applications that involve Data intensive computing

UNIT I INTRODUCTION 9

Introduction to Distributed systems – Databases Vs. File Systems - Distributed file systems (HDFS) – Distributed Machine-Learning System - Data Parallelism – Characteristics -Hadoop – Execution Engines -Map Reduce- Distributed Storage System for Structured Data – No SQL databases -Casandra, MongoDB-Developing a Distributed Application

UNIT II ARCHITECTURES AND SYSTEMS 9

High performance Network Architectures for Data intensive Computing – Architecting Data Intensive Software systems – ECL/HPCC: A Unified approach to Big Data – Scalable storage for Data Intensive Computing - Computation and Storage of scientific data sets in cloud- Stream Data Model - Architecture for Data Stream Management-Stream Queries –Sampling Data in a Stream-Filtering Streams

UNIT III TECHNOLOGIES AND TECHNIQUES 9

Load balancing techniques for Data Intensive computing – Resource Management for Data Intensive Clouds – SALT - Parallel Processing, Multiprocessors and Virtualization in Data-Intensive Computing - Challenges in Data Intensive Analysis and Visualization - Large-Scale Data Analytics Using Ensemble Clustering - Ensemble Feature Ranking Methods for Data Intensive Computing Application - Record Linkage Methodology and Applications - Semantic Wrapper

UNIT IV SECURITY 9

Security in Data Intensive Computing Systems - Data Security and Privacy in Data-Intensive Supercomputing Clusters - Information Security in Large Scale Distributed Systems -Privacy and Security Requirements of Data Intensive Applications in Clouds

UNIT V APPLICATIONS AND FUTURE TRENDS 9

Cloud and grid computing for data intensive applications **-**Scientific applications - Bioinformatics - Large science discoveries - Climate change - Environment - Energy - Commercial applications - Future trends in data intensive computing

OUTCOMES

Upon completion of this course, the students will be be able to

- Design applications that involve data intensive computing
- Decide on the appropriate techniques of Map Reduce, Mongo DB, for the different applications
- Decide on the various security techniques that are necessary for data intensive applications

REFERENCES:

1. Tom White , “Hadoop: The Definitive Guide”,. O'Reilly Media. October 2010.
2. Hector Garcia-Molina, Jeffrey D. Ullman, and Jennifer Widom., “Database Systems: The Complete Book, Pearson, 2013
3. Handbook of Data Intensive Computing, byFurht, Borko, Escalante, Armando, Springer 2011.

BD7005 R PROGRAMMING L T P C
3 0 0 3

OBJECTIVES

- To understand the basics in R programming in terms of constructs, control statements, string functions
- To learn to apply R programming for Text processing
- To understand the use of R Big Data analytics
- To able to appreciate and apply the R programming from a statistical perspective

UNIT I INTRODUCTION 9

Introducing to R – R Data Structures – Help functions in R – Vectors – Scalars – Declarations – recycling – Common Vector operations – Using all and any – Vectorized operations – NA and NULL values – Filtering – Vectorised if-then else – Vector Equality – Vector Element names

UNIT II MATRICES, ARRAYS AND LISTS 9

Creating matrices – Matrix operations – Applying Functions to Matrix Rows and Columns – Adding and deleting rows and columns – Vector/Matrix Distinction – Avoiding Dimension Reduction – Higher Dimensional arrays – lists – Creating lists – General list operations – Accessing list components and values – applying functions to lists – recursive lists

UNIT III DATA FRAMES 9

Creating Data Frames – Matrix-like operations in frames – Merging Data Frames – Applying functions to Data frames – Factors and Tables – factors and levels – Common functions used with factors – Working with tables - Other factors and table related functions - Control statements – Arithmetic and Boolean operators and values – Default values for arguments - Returning Boolean values – functions are objects – Environment and Scope issues – Writing Upstairs - Recursion – Replacement functions – Tools for composing function code – Math and Simulations in R

UNIT IV OOP 9

S3 Classes – S4 Classes – Managing your objects – Input/Output – accessing keyboard and monitor – reading and writing files – accessing the internet – String Manipulation – Graphics – Creating Graphs – Customizing Graphs – Saving graphs to files – Creating three-dimensional plots

UNIT V INTERFACING 9

Interfacing R to other languages – Parallel R – Basic Statistics – Linear Model – Generalized Linear models – Non-linear models – Time Series and Auto-correlation - Clustering

OUTCOMES:

Upon completion of this course, the students will be able to

- Create artful graphs to visualize complex data sets and functions
- Write more efficient code using parallel R and vectorization
- Interface R with C/C++ and Python for increased speed or functionality
- Find new packages for text analysis, image manipulation, and perform statistical analysis of the same

REFERENCES

1. Norman Matloff , “The Art of R Programming: A Tour of Statistical Software Design”, No Starch Press, 2011.
2. Jared P. Lander, “R for Everyone: Advanced Analytics and Graphics”, Addison-Wesley Data & Analytics Series, 2013.
3. Mark Gardener, “ Beginning R – The Statistical Programming Language”, Wiley, 2013
4. Robert Knell, “ Introductory R: A Beginner's Guide to Data Visualisation, Statistical Analysis and Programming in R”, Amazon Digital South Asia Services Inc, 2013.

BD7007 SOCIAL NETWORK ANALYSIS FOR BIG DATA L T P C

3 0 0 3

OBJECTIVES:

- To understand the graph essentials in the context of Social Network Analysis
- To understand the classification of social network behavior
- To identify social network communities and community structures
- To understand the recommendation systems
- To study few social networking sites

UNIT I INTRODUCTION 9

Introduction of Social Network Analysis –Graph Essentials –Graph Basics-Graph Representation-Types of Graphs-Connectivity in Graphs-Special Graphs-Graph Algorithms-Network Measures-Network Models: Properties of Real-World Networks-Random Graphs- Small-World Model-Preferential Attachment Model.

UNIT II BIG DATA ANALYSIS AND PREDICTIVE ANALYTICS 9

Evolution of analytic scalability – parallel processing systems – map reduce – enterprise analytic sand box – analytic data sets – Analytic methods – analytic tools- Predictive Analytics – Supervised – Unsupervised learning – Neural networks – Kohonen models – Normal – Deviations from normal patterns – Normal behaviours – Expert options – Variable entry – Mining Frequent itemsets – Market based model – Apriori Algorithm – Handling large data sets in Main memory – Limited Pass algorithm – Counting frequent itemsets in a stream – Clustering Techniques – Hierarchical – K- Means – Clustering high dimensional data Visualizations – Visual data analysis techniques, interaction techniques

UNIT III COMMUNITIES AND INTERACTIONS 9

Community Analysis-Community Detection-Community Evolution-Community Evaluation-Information Diffusion in Social Media-Herd Behavior-Information Cascades-Diffusion of Innovations-Epidemics.

UNIT IV RECOMMENDATION IN SOCIAL MEDIA AND BEHAVIOR ANALYTICS 9

Challenges-Classical Recommendation Algorithms-Recommendation Using Social Context-Evaluating Recommendations-Behavior Analytics: Individual Behavior- Collective Behavior.

UNIT V SOCIAL MEDIA INSITES 9

Introduction: Hacking on Twitter Data-Twitter: Friends, Followers, and Set wise Operations-Analyzing Tweets-Visualizing tons of tweets-LinkedIn: Motivation for clustering –Blogs: Natural Language Processing –Face book: Visualizing face book data.

OUTCOMES:

Upon completion of this course, the students will be able to

- Design algorithms for processing social networking data by using graph based algorithms
- Design recommendation systems for the various data available in social networking sites
- Design algorithms for identifying the underlying communities and their structure in social networking sites
- Apply recommendation algorithms for arena that has a Big Data component.

REFERENCES

1. Social Media Mining: An Introduction, R. Zafarani, M. Abbasi, and H. Liu, Cambridge University Press, 2014.
2. Matthew A. Russell, Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkeeIn, and Other Social Media Sites, O'Reilly Media, 2011.
3. Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", Wiley and SAS Business Series, 2012.
4. Colleen Mccue, "Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis", Elsevier, 2007

CP7087 **PARALLEL ALGORITHMS** **L T P C**
3 0 0 3

OBJECTIVES:
- To learn parallel algorithms development techniques for shared memory and DCM models
- To study the main classes of fundamental parallel algorithms
- To study the complexity and correctness models for parallel algorithms.

UNIT I INTRODUCTION **9**
Introduction to Parallel Algorithms – Models of computation – Selection – Mergin on EREW and CREW – Median of two sorted sequence – Fast Merging on EREW – Analyzing Parallel Algorithms

UNIT II SORTING & SEARCHING **9**
Sorting Networks – Sorting on a Linear Array – Sorting on CRCW, CREW, EREW – Searching a sorted sequence – Searching a random sequence – Bitonic Sort

UNIT III ALGEBRAIC PROBLEMS **9**
Permutations and Combinations – Matrix Transpositions – Matrix by Matrix multiplications – Matrix by vector multiplication.

UNIT IV GRAPH & GEOMETRY **9**
Connectivity Matrix – Connected Components – All Pair Shortest Paths – Minimum Spanning Trees – Point Inclusion – Intersection, Proximity and Construction Problems.

UNIT V OPTIMIZATION & BIT COMPUTATIONS **9**
Prefix Sums – Job Sequencing – Knapsack - Adding two integers – Adding n integers – Multiplying two integers – Selection.

TOTAL : 45 PERIODS

OUTCOMES:
Upon completion of this course, the student should be able to
- Familiarize with design of parallel algorithms in various models of parallel computation
- Familiarize with the efficient parallel algorithms related to many areas of computer science: expression computation, sorting, graph-theoretic problems, computational geometry, etc
- Familiarize with the basic issues of implementing parallel algorithms

REFERENCES:

1. Selim G. Akl, “The Design and Analysis of Parallel Algorithms”, Prentice Hall, New Jercy, 1989.
2. Michael J. Quinn, “Parallel Computing: Theory & Practice”, Tata McGraw Hill Edition, 2003.
3. Joseph JaJa, “Introduction to Parallel Algorithms”, Addison-Wesley, 1992.

BD7006 **SENTI MENT ANALYSIS** **L T P C**
3 0 0 3

OBJECTIVES
- To give an overview on the need for sentiment analysis
- To explore the various methodologies necessary to perform sentiment classification
- To learn about opinion summarization
- To learn the various tools used for sentiment analysis

UNIT I INTRODUCTION 9

Need for Sentiment Analysis – Problem of Sentiment Analysis - Subjectivity – Stance – Words to Discourse – Pragmatics – Natural Language Processing issues – Opinion Definition – Sentiment analysis Tasks – Opinion Summarization – Types of opinion – Subjectivity and emotion – Author and Reader Standpoint

UNIT II DOCUMENT SENTIMENT CLASSIFICATION 9

Sentiment classification using Supervised learning – unsupervised learning – rating prediction – cross-domain sentiment classification – cross-language sentiment classification – Sentence subjectivity and classification – subjectivity classification – sentence sentiment classification – conditional sentences - sarcastic sentences – cross-language subjectivity and sentiment classification – Discourse information for sentiment classification

UNIT III ASPECT BASED SENTIMENT ANALYSIS 9

Aspect sentiment classification – rules of opinions and compositional semantics – aspect extraction – identifying resource usage aspect – simultaneous opinion lexicon expansion and aspect extraction – Grouping aspects into categories – entity, opinion hold and timing extraction – coreference resolution and word sense disambiguation – aspect and entity extraction - sentiment lexicon generation – corpus based approach – dictionary based approach – desirable and undesirable facts

UNIT IV OPINION SUMMARIZATION 9

Aspect based opinion summarization – improvements to aspect-based opinion summarization – contrastive view summarization – traditional summarization – Analysis of comparative opinions – identifying comparative sentences – identifying preferred entities – opinion search and retrieval – opinion spam detection – types of spam detection - supervised and un-supervised approach – group spam detection

UNIT V TOOLS FOR SENTIMENT ANALYSIS 9

Detecting fake or deceptive opinions - Quality of Review – Quality as regression model – other methods – Case study – sentiment analysis applications – tools for sentiment analysis – Semantria – Meltwater – Google Analytics – Face book Insights – Tweetstats.

OUTCOMES:

Upon completion of this course, the students will be able to

- Apply the various algorithms to perform opinion mining and classification
- Identify the sentiment of any document, web-page or social networking site.
- Compare and contrast the various tools necessary for performing sentiment analysis
- Use the apt tools to perform sentiment analysis for any given application

REFERENCES:

1. Bing Liu, “Sentiment Analysis and Opinion Mining”, Morgan and Claypool publishers, 2012.
2. Bing Liu, “ Sentiment Analysis – Mining opinion, Sentiments and Emotions”, Cambridge University Press, 2015
3. Bo Pang and Lillian Lee, “ Opinion Mining and Sentiment Analysis”, Now Publishers Inc, 2008
4. Roy De Groot, “Data mining for Tweet sentiment classification – Twitter sentiment analysis”, LAP Lambert Academic Publishing, 2012

BD7001	BIG DATA ACQUISITION	L	T	P	C
		3	0	0	3

OBJECTIVES:

- To understand the complexity and volume of Big Data and their challenges
- To analyse the various methods of data collection
- To comprehend the necessity for pre-processing Big Data and their issues
- To understand predictive analytics and descriptive analytics

UNIT I INTRODUCTION TO BIG DATA ACQUISITION 6

Big data framework - Fundamental concepts of Big Data management and analytics - Current challenges and trends in Big Data Acquisition.

UNIT II DATA COLLECTION AND TRANSMISSION 9

Big data collection- Strategies- Types of Data Sources- Structured Vs Unstructured data- ELT vs ETL - storage infrastructure requirements -Collection methods-Log files- Sensors- Methods for acquiring network data (Libcap-based and zero-copy packet capture technology) -Specialized network monitoring softwares (Wireshark, Smartsniff and Winnetcap)- Mobile equipments- Transmission methods- Issues.

UNIT III DATA PRE-PROCESSING 9

Data pre-processing overview-Sampling- Missing Values -Outlier Detection and Treatment - Standardizing Data- Categorization - Weights of Evidence Coding -Variable Selection and Segmentation.

UNIT IV DATA ANALYTICS 12

Predictive Analytics (Regression, Decision Tree, Neural Networks) - Descriptive Analytics (Association Rules, Sequence Rules), Survival Analysis (Survival Analysis Measurements, Kaplan Meir Analysis, Parametric Survival Analysis) - Social Network Analytics (Social Network Learning-Relational Neighbor Classification).

UNIT V BIG DATA PRIVACY AND APPLICATIONS 9

Data Masking – Privately Identified Information (PII) -Privacy preservation in Big Data- Popular Big Data Techniques and tools- Map Reduce paradigm and the Hadoop system- Applications- Social Media Analytics- Recommender Systems- Fraud Detection.

OUTCOMES:

Upon Completion of the course, the students will be able to

- Identify the various sources of Big Data
- Design new algorithms for collecting Big Data from various sources
- Design algorithms for pre-processing Big Data other than the traditional approaches
- Design methodologies to extract data from structured and un-structured data for analytics

REFERENCES:

1. Bart Baesens," Analytics in a Big Data World: The Essential Guide to Data Science and its Applications"**,** John Wiley & Sons, 2014
2. Min Chen, Shiwen Mao, Yin Zhang, Victor CM Leung ,Big Data: Related Technologies, Challenges and Future Prospects, Springer, 2014.
3. Michael Minelli, Michele Chambers, Ambiga Dhiraj , “Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends”, John Wiley & Sons, 2013
4. Raj, Pethuru, “Handbook of Research on Cloud Infrastructures for Big Data Analytics”, IGI Global.

BD7002 **BIG DATA QUERY LANGUAGES** **L T P C**
3 0 0 3

OBJECTIVES:

- To familiarize the Data Systems of various organizations
- To perform data analysis using HADOOP and RHADOOP
- To understand the PIG Data model
- To learn the HIVE QL way of querying Big Data.

UNIT I INTRODUCTION TO BIG DATA 9

Introduction to BigData Platform – Challenges of Conventional Systems – Web Data – Evolution Of Analytic Scalability – Analytic Processes and Tools – Analysis vs Reporting – Modern Data Analytic Tools – Statistical Concepts: Sampling Distributions – Re-Sampling – Statistical Inference – Prediction Error.

UNIT II ANALYTICS FOR BIG DATA 9

IBM PureData Systems – Netezza's Design Principles – The Netezza Appliance – Extending the Netezza Analytics – Real-Time Analytical Processing – InfoSphere Streams Basics – InfoSphere Streams Working – enterprise class – industry use cases – Indexing Data from Multiple Sources – Creating Information Dashboards

UNIT III DATA ANALYSIS USING R AND HADOOP 9

Features of R language – HADOOP Features – HDFS and MapReduce architecture – R and Hadoop Integrated Programming Environment (RHIPE) Introduction – Architecture of RHIPE – RHIPE function reference – RHADOO Introduction – Architecture of RHADOOP – RHADOOP function reference, SQL on HADOOP.

UNIT IV PROGRAMMING PIG 9

Introduction – installation and execution – PIG Data Model – PIG Latin – Input, Output- Relational Operators – User Defined Functions – Join Implementations – Integrating Pig with Legacy Code and Map Reduce –Developing and Testing Pig Latin Scripts – Embedding Pig Latin in Python – Evaluation Function in Java- Load Functions – Store Functions

UNIT V HIVE QL 9

Introduction – Data Types and File Formats – Databases in Hive – HiveQL: Data Definition – Data Manipulation – Queries – Views – Indexes – Schema Design

OUTCOMES:

Upon Completion of the course, the students will be able to

- Design applications and apply HADOOP and RHADOOP
- Identify and apply PIG data model for applications
- Design Big data applications schema and use HIVE QL

REFERENCES:

1. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, 2007.
2. Tom White " Hadoop: The Definitive Guide" Third Edition, O'reilly Media, 2011
3. Zikopoulos, P., Parasuraman, K., Deutsch, T., Giles, J., & Corrigan, D.v *Harness the Power of Big Data The IBM Big Data Platform*. McGraw Hill Professional, 2012
4. Prajapati, V. *Big data analytics with R and Hadoop*. Packt Publishing Ltd, 2013
5. Gates, A. *Programming Pig*. " O'Reilly Media, Inc.", 2011.
6. Capriolo, E., Wampler, D., & Rutherglen, J., *Programming hive*. " O'Reilly Media, Inc.", 2012.

BD7071 **TEXT MINING** **L T P C**
3 0 0 3

OBJECTIVES:

- To understand the basic issues and types of text mining
- To appreciate the different aspects of text categorization and clustering
- To understand the role played by text mining in Information retrieval and extraction
- To appreciate the use of probabilistic models for text mining
- To appreciate the current trends in text mining

UNIT I INTRODUCTION **8**

Overview of text mining- Definition- General Architecture– Algorithms– Core Operations – Pre-processing– Types of Problems- basics of document classification- information retrieval- clustering and organizing documents- information extraction- prediction and evaluation-Textual information to numerical vectors -Collecting documents- document standardization- tokenization- lemmatization- vector generation for prediction- sentence boundary determination -evaluation performance

UNIT II TEXT CATEGORIZATION AND CLUSTERING **10**

Text Categorization – Definition – Document Representation –Feature Selection - Decision Tree Classifiers - Rule-based Classifiers - Probabilistic and Naive Bayes Classifiers - Linear Classifiers- Classification of Linked and Web Data - Meta-Algorithms– Clustering –Definition- Vector Space Models - Distance-based Algorithms- Word and Phrase-based Clustering -Semi-Supervised Clustering - Transfer Learning

UNIT III TEXT MINING FOR INFORMATION RETRIEVAL AND INFORMATION EXTRACTION **10**

Information retrieval and text mining- keyword search- nearest-neighbor methods- similarity- web-based document search- matching- inverted lists- evaluation. Information extraction- Architecture - Co-reference - Named Entity and Relation Extraction- Template filling and database construction – Applications. Inductive -Unsupervised Algorithms for Information Extraction. Text Summarization Techniques - Topic Representation - Influence of Context - Indicator Representations - Pattern Extraction - Apriori Algorithm – FP Tree algorithm

UNIT IV PROBABILISTIC MODELS **9**

Probabilistic Models for Text Mining -Mixture Models - Stochastic Processes in Bayesian Nonparametric Models - Graphical Models - Relationship Between Clustering, Dimension Reduction and Topic Modeling - Latent Semantic Indexing - Probabilistic Latent Semantic Indexing -Latent Dirichlet Allocation- Interpretation and Evaluation - Probabilistic Document Clustering and Topic Models - Probabilistic Models for Information Extraction - Hidden Markov Models - Stochastic Context-Free Grammars - Maximal Entropy Modeling - Maximal Entropy Markov Models -Conditional Random Fields

UNIT V RECENT TRENDS **8**

Visualization Approaches - Architectural Considerations - Visualization Techniques in Link Analysis - Example- Mining Text Streams - Text Mining in Multimedia - Text Analytics in Social Media - Opinion Mining and Sentiment Analysis - Document Sentiment Classification - Opinion Lexicon Expansion - Aspect-Based Sentiment Analysis - Opinion Spam Detection – Text Mining Applications and Case studies

TOTAL: 45 PERIODS

OUTCOMES:

Upon Completion of the course, the students will be able to

- Identify the different features that can be mined from text and web documents
- Use available open source classification and clustering tools on some standard text data sets
- Modify existing classification/clustering algorithms in terms of functionality or features used
- Design a system that uses text mining to improve the functions of an existing open source search engine
- Implement a text mining system that can be used for an application of your choice

REFERENCES:

1. Sholom Weiss, Nitin Indurkhya, Tong Zhang, Fred Damerau "The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data", Springer, paperback 2010
2. Ronen Feldman, James Sanger -" The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data"-Cambridge University press, 2006.
3. Charu C. Aggarwal ,ChengXiang Zhai**,**Mining Text Data, Springer; 2012