

Faculty of Science

Revised Syllabus for the

M.Sc. Bioinformatics Credit Systems

(From Academic year: 2006 - 2007)



University of Pune
Department of Bioinformatics
Pune 411 007

Present Course Structure of M. Sc. (Bioinformatics)

(T = Theory course, P = Practical course)

(Revised on May 18, 2006)

Semester – I

<i>Course No.</i>	<i>Course Name</i>	<i>Total Credits</i>
BIM 101 (T)*	Basic Biology	2
BIM 102 (T)*	Mathematics For Bioinformatics	2
BIM 103 (T+P)	Statistical Techniques	3
BIM 104 (T)	Biological Chemistry	3
BIM 105 (T)	Genetic Information Flow & Processing	3
BIM 106 (T+P)	Basic Concepts in Computing	3
BIM 107 (T+P)	Introduction to Database Systems	4
BIM 108 (T+P)	Biological Databases and Data Analysis	4
BIM 109 (P)	Programming in C	3
		25

* Students from Mathematics Stream will take BIM 101(T), while those from the Biology Stream will take BIM 102(T)

Semester – II

<i>Course No.</i>	<i>Course Name</i>	<i>Total Credits</i>
BIM 201 (T)	Cell Biology and Genetics	3
BIM 202 (T)	Structural Biology & Molecular Modelling	4
BIM 203 (T+P)	Chemoinformatics	3
BIM 204 (T+P)	Immunology for Bioinformatics	3
BIM 205 (T+P)	Programming in Object Oriented Languages	4
BIM 206 (P)	Programming in Perl	3
BIM 207 (P)	Protein Structure Prediction and Molecular Modeling	2
BIM 208 (P)	Laboratory Techniques in Modern Biology	2
		24

Semester – III

<i>Course No.</i>	<i>Course Name</i>	<i>Total Credits</i>
BIM 301 (T+P)	Taxonomy & Phylogeny	3
BIM 302 (T+P)	Genome to Drug and Vaccine	4
BIM 303 (T+P)	Comparative Genomics and Proteomics	4
BIM 304 (T)	Parasite Bioinformatics	2
BIM 305 (T)	Object Oriented and Relational Databases	2
BIM 306 (T+P)	Scientific visualization for data analysis	3
BIM 307 (T+P)	Tools & techniques for biological data mining	4
BIM 308 (P)	Computer Implementation of Data Structures, Data Integration	3
		25

Semester – IV

<i>Course No.</i>	<i>Course Name</i>	<i>Total Credits</i>
BIM 401 (T+P)	Advanced Techniques for Sequence and Structure Analysis	4
BIM 402 (T)	Metabolomes and Metabolic Pathway Engineering	2
BIM 403 (T)	Emerging Areas in Bioinformatics	2
BIM 404 (T)	Project work	16
Any one of the following		
BIM 405 (T)	Seminars on Applications of Bioinformatics in Agriculture	2
BIM 406 (T)	Seminars on Applications of Bioinformatics in Human Health	2
BIM 407 (T)	Seminars on Applications of Bioinformatics in Environment	2
BIM 408 (T)	Seminars on Applications of Bioinformatics in Biotechnology	2
BIM 409 (T)	Seminars on Applications of Bioinformatics Molecular Biology	2
BIM 410 (T)	Seminars on Applications of Bioinformatics in Neurobiology	2
BIM 411 (T)	Seminars on Applications of Bioinformatics Drug Designing	2
BIM 412 (T)	Seminars on Applications of Bioinformatics in Veterinary Sciences	2
		26

(Numbers in the brackets indicate total credits for respective courses and also indicates number of hours of classroom teaching for respective theory courses per week. Credits for practical courses indicate number of laboratory sessions per week (each session is of 3 hours).

Semester – I

BIM 101 (T): Basic Biology

(2 credits)

Objectives:

The course aims at exposing the students of the non-biology stream to the diversity of microbial, plant and animal life. At the conclusion of the course the student would have become familiar with outlines of the classification of the organisms, their structural organization and functional complexities.

Theory

Syllabus:

- Bacteria: Structure of bacterial cell; bacterial types; transformation, transfection, transduction and conjugation; nutrition; phylogeny (5)
- Viruses: Biology of viruses; bacteriophages, plant and animal viruses; replication of viral genome; HIV (3)
- Protists: protozoans, algae, slime and water molds (2)
- Fungal World: Feeding, reproduction, diversity and relationships (2)
- Plant Diversity: Broad classification and inter-relationships of non-vascular and vascular plants; tissue organization; reproductive patterns; transport mechanisms, growth, photosynthesis, hormones. (8)
- Animal Life: Major animal phyla, characteristics and interrelationships; tissues, organs and organ systems; principles of nutrition, digestion, thermoregulation, Osmoregulation and excretion, muscle contraction, neural reflexes, circulation, respiration and endocrines. (10)

References:

- The Science of Life by R. A. Wallace, G. P. Sanders & R. J. Ferl, Harper Collins Publishers, Biology, 1991.
- Big-time biology by Beardsley Tim, Scientific American, 1994.
- Biogenesis: Theories of life's origin by Lahav, N., New York. Oxford University Press, 1999.

BIM 102 (T) Mathematics for Bioinformatics

(2 credits)

Objectives:

- To upgrade the skills of the students (biology) in mathematics that are essential for learning Bioinformatics.

Theory

Syllabus:

- Trigonometry (2)
Trigonometric Functions, Series Expansion, Inverse, General Values, Graphs
- Calculus (3)
Limits, Continuity, Analysis, Differentiation (1D & Partial), Reimann Integration, Definite Integrals, Recursion Theorems.
- Ordinary & Partial Differential Equation (3)
1'st Order & 2'nd Order Ordinary Differential Equations. Self Adjoint Equations,

Special Functions, Nature of Partial Differential Equations, Method of Separation of Variables.

- 2Dimensional & 3Dimensional Geometry (2)
Cartesians and Polar Coordinates, Locus, Equation of a straight line, pair of straight lines, circle, ellipse, parabola, hyperbola. Asymptotic Analysis. Equation of sphere, cone. 3D Coordinate System., Spherical and Cylindrical Coordinates.
- Vector & Matrices (3)
Vector Algebra, Vector Calculus, Basic Computations, Matrices
- Introduction To Set Theory & Groups (2)
Introductory Notions, Composition, Boolean Logic Algebraic Structures, Groups, Vector Spaces
- Integral transform (2)
Fourier Series, Fourier Transform, Laplace Transform
- Series And Sequences (1)
Sequences, Infinite Series, Infinite Products
- Numerical Techniques (1)
Basic Formalism, Methods for Solving Equations, Finding Eigenvalues & Eigenvectors, Solving ODE & PDE, Differentiation and Integration
- Introduction to MATLAB (4)
Mathematical Operations, Numerical Programming, Illustration of all Mathematical Methods through MATLAB.
- Applications To Mathematical Biology (2)
Enzyme kinetics, Immunology, Population genetics, Tumor modeling, Applications of ordinary & partial differential equations to Biology.

MATLAB practicals : for 5h

References:

- Introduction to mathematical methods in bioinformatics by Isaev, Alexander Berlin Springer, 2004.
- Mathematics in chemistry by K. V. Raman & Pal, Sourav, New Delhi, Vikas publishing house Pvt. Ltd., 2005
- Calculations in molecular biology and biotechnology: a guide to mathematics in the laboratory by Stephenson, F.H. Amsterdam, Academic Press, 2003.
- Advanced mathematical methods for engineering & science students by Stephenson, G. & Radmore, P. M., 1990.
- Mathematics and computer science in medical imaging by Viergever, Max A.& Todd-Pokropek, Andrew.,1988.
- Mathematical Methods for Physicists by G. Arfken, Academic Press, New York,1970.
- Some Modern Mathematics for Physicists and Other Outsiders by P. Roman, Pergamon, New York, 1975, Vol. 2, p. 660.
- Mathematical Biology, by J. D. Murray Springer Verlag, 1989.
- Mathematical Models in Molecular and Cellular Biology by Segal, L., ed. 1980. Cambridge: Cambridge University Press.
- Numerical Methods by Balaguruswamy, TMH.

BIM 103 (T+P) Statistical Techniques

(2T+1P credits)

Objectives:

To upgrade the skills of the students in statistics that are essential for learning Bioinformatics

Theory

Syllabus:

- Introduction to principles of statistical sampling from a population (2)
- Probability Distributions Introductory Concepts, Conditional Probability, Bayesian Rules, Frequency Distributions and Statistical Measures, Random variable, Integration of random variables, Vector valued Random Variables (6)
- Correlation and regression analysis (4)
- Multivariate analysis (4)
- Hypothesis testing (2)
- Markov Models & Markov Chains (4)
- Cluster Analysis (3)
 - Nearest neighbour search
 - Search using stem numbers
 - Search using text signatures
 - Phylogenetic Analysis Tools: Maximum Likelihood, Parsimony methods, Distance methods, Model Comparison. (5)

Practicals

Syllabus:

- Using any of the software like SPSS, SAS etc. exercises will be done on (8)
- PAM, BLOSUM, Substitution Rates etc.
 - Analysis on clustering using Phylogenetic methods.
 - Use of correlation and regression analysis.

References:

- Probability statistics, and reliability for engineers by Boca Raton, Ayyub B. M. & McCuen, R H, CRC Press, 1997.
- Statistics: concepts and applications by Frank, Harry & Althoen, S. C., Cambridge University Press, 1995.
- Statistical methods in bioinformatics: an introduction by Ewens, W. J. & Grant, G. R., New York. Springer, 2001.
- Handbook of computational statistics: concepts and methods by Gentle, J.E., Hardle, W. & Mori, Y., Berlin, Springer-Verlag, 2004.
- Statistical design and analysis of industrial experiments by Ghosh, Subir, Ed., 1990.
- Scan Statistics by Glaz, J., Naus, J. & Wallenstein S, New York, Springer, 2001.
- Statistical design for research by Kish, L., Wiley series in probability and mathematical statistics, New York, John Wiley & Sons, 1987.
- Introduction to probability and statistics by Lipschutz, S. & Schiller, J. J., New York. McGraw-Hill, 1999.
- Schaum's outline of theory and problems of statistics by Spiegel, M. R. & Stephens, L. J., Ed. 3, New Delhi, Tata McGraw-Hill Publishing Co. Ltd., 2001.
- Mathematical and statistical methods for genetic analysis by Lange, K., 2nd Ed., New York. Springer-Verlag, 2002.

Objectives:

The objective of the course is to make the students understand the chemistry of different classes of biomolecules, their interactions in an aqueous environment, the structure-function relationships of macromolecules, the principles of enzyme catalysis and regulation, organization of metabolic systems and the pathways for the complete oxidation of glucose.

Theory**Syllabus:**

- Water as the universal biological solvent; concept of osmolarity. (1.5)
- Carbohydrates: monosaccharides, oligosaccharides, polysaccharides, proteoglycans and glycoproteins. (3)
- Lipids: fatty acids, acylglycerols; phospholipids, sphingolipids, cholesterol and membranes Isoprenoids, icosanoids and their biological importance. (4.5)
- Proteins: amino acids and peptides; primary, secondary, tertiary and quaternary structures; structure, function and evolutionary relationships; protein – protein interactions protein folding; allosteric proteins. (9)
- Nucleic acids: bases, nucleotides, RNA and DNA; different structural forms of DNA; denaturation, renaturation and hybridization of DNA; different types of RNA; Protein-nucleic acid interaction. (4.5)
- Enzymes: details of enzyme nomenclature and classification; units of enzyme activity; coenzymes and metal cofactors; temperature and pH effects; Michaelis-Menten kinetics; Inhibitors and activators; active site and catalytic mechanisms; covalent and non-covalent regulations; isoenzymes; osmolytes and intracellular modulation of enzymes. (13.5)
- Organization of metabolic systems: enzyme chains, multi-enzyme complexes and multifunctional enzymes; anaplerotic sequences and amphibolic pathways; pacemaker enzymes and feedback control of metabolic pathways; shuttle pathways; energy charge . (3)
- Oxidation of glucose in cells: high energy bond, glycolysis, citric acid cycle and oxidative phosphorylation. (6)

References:

- Fundamentals of Biochemistry (2nd edition) by D., Voet, Voet, J.G. & Pratt, C. W. John Wiley & Sons, 2006.
- Biochemistry (3rd edition) by Voet D. & Voet, J.G., John Wiley & Sons, 2004.
- Lehninger, Principles of Biochemistry (4th edition) by Nelson, D. L. & M. M. Cox, W. H. Freeman & Co., 2005.
- Principles of Biochemistry (4th edition) by Horton, H. R., Moran, L.A., Scrimgeour, K.G., Perry, M. D. & Rawn, J. D., Pearson-Prentice Hall, 2006.
- Biochemistry (5th edition) by Berg, J.M., Tymoczko, J.L. & Stryer, L., W.H. Freeman & Co., 2002.
- Biochemistry (3rd edition) by Mathews, C.K., Van Holde, K.E. & Ahern K.G., Pearson, Education, 2000.

Objectives:

The objective of the course is to make the student understand the current concepts in gene organization, transcription, translation and regulation of gene function as well as the biotechnological implications of recent developments in cloning and genome sequencing.

Theory**Syllabus:**

- Introduction: DNA as a genetic material – Experiments done to prove this (1)
- Nucleic acid structure: single stranded, double stranded; secondary structures in single stranded molecules; alternative double helical structures in double stranded DNA; Closed DNA as supercoiled molecule (2)
- Genome organization: Prokaryotic and eukaryotic genomes – C value paradox, repetitive and non-repetitive DNA., transposons and retroposons; Exons and introns – organization of interrupted genes, one DNA sequence may code for multiple proteins; Gene numbers – essential genes and total gene number, gene clusters, pseudogenes; Gene families – globin gene and immunoglobulin gene families; Organelle genome – mitochondrial and chloroplast. (6)
- Packaging of genome – Bacterial genome as nucleoid; Eukaryotic genome – nucleosomes, chromatin, solenoids, loops, domains, scaffolds, chromosomes. (3)
- Perpetuation of DNA: Prokaryotic DNA replication – DNA polymerases, origin of replication, initiation, elongation and termination of replication. Rolling circle model of replication; Eukaryotic DNA polymerases – multiple origins of replication, process of replication; Regulation of replication in both prokaryotes and eukaryotes. (5)
- DNA damage, repair and recombination: Different types of DNA damages; Variety of DNA repair systems in prokaryotes and eukaryotes – Base excision repair system, Nucleotide excision repair system, Mismatch repair system, Recombination repair system; Recombination – homologous and non-homologous recombination. (3)
- Gene Expression – Transcription and Translation: Transcription in prokaryotes – RNA polymerase, initiation, elongation and termination of transcription; regulation of transcription – operon concept, lactose and Tryptophan operons (4)
- Transcription in eukaryotes: Different RNA polymerases – requirement of promoters by these RNA polymerases, Initiation, elongation and termination by these polymerases; Processing of transcripts – 5' capping, 3' polyadenylation, splicing and editing; Regulation of transcription – Response elements, enhancers and silencers, HLH, Leucine zipper proteins., noncoding RNAs. (5)
- Translation in prokaryotes: protein synthesis machinery – mRNA, tRNA and rRNA molecules; initiation, elongation and termination of translation; Genetic code – interpreting genetic code; Accuracy of translation. (3)
- Eukaryotic translation: protein synthesis – initiation, elongation and termination; Post-translational modifications of proteins; protein degradation; Regulation of translation - mRNA stability, 5' and 3' UTRs, mRNA localization (5)
- Gene regulation by post-translational modifications of proteins (acetylation, methylation, ribosylation, phosphorylation etc.) and different intermediate RNAs (ribozymes, miRNAs, siRNAs etc.) (4)
- Genome sequencing, PCR, RFLP, Fingerprinting, RAPDs, Microarrays, ESTs (4)

References:

- Molecular Biology by D. Clark, Elsevier publishers, 2005
- Human Molecular Biology by R. J. Epstein, Cambridge University Press, 2003.
- Genes VIII by B. Lewin, Pearson-Prentice Hall, 2004.
- Genetics from Genes to Genomes by L. H. Hartwell, L. Hood, M. L. Goldberg, A.E. Reynolds, L. M. Silver & R. C. Veres, McGraw-Hill, 2004.
- Genetics: A Molecular Approach (2nd edition) by P.J. Russell, 2006.
- Essentials of molecular biology by D. Freifelder, New Delhi, Narosa Publishing House, 2001.

BIM 106 (T+P) Basic Concepts in Computing

(1T+2P credits)

Objectives:

- To introduce the students to fundamentals of computers, computing and software.

Theory

Syllabus:

- Overview and functions of a computer system, storage, devices, memory, etc (1)
 - Types of Processing: Batch, Real-Time, Online, Offline
 - Types of modern computers: The workstation, The Minicomputer, Mainframe Computers, Parallel Processing Computer, The Super Computer, etc
- Introduction to operating systems: Windows/Unix/Linux (2)
- The Internet and its Resources, World Wide Web (WWW): associated tools, services, resources and various terminologies (1)
- Searches on Medline, bibliographic databases, etc. (1)

Computer Networking; Network and Data security

(10)

- OSI Reference Model, TCP/IP, topologies and protocols, designing networks
- Networking gadgets (Router, Switch, etc); Data Communication (ISDN, VPN, DSL, cable modem, cellular modem, etc); Communication Links (Wire pairs, Coaxial cables, Fiber optics, Microwave, Satellite, etc)
- Network security fundamentals: types of attacks, firewall, packet filtering, classification of data security threats, protection mechanism (authentication, access control, access rules)
- Encryption/Decryptions techniques
- An overview of Computer viruses: How do they get transmitted? What are the dangers? General Precautions to be taken.
- Current & future technologies (Grid Computing, VPN, wireless, mobile computing, biometrics etc.)

Practicals

Syllabus:

- Tutorials (Windows, Linux, Internet, etc), applications and utilities of Windows, Browsers, various search engines and technologies, metadata search engines, using E-Mail/Web mail, ftp, searching on the WWW, word processing, creation of computer presentations with graphics (5)
- Basic Unix/Linux commands (3)

- Spreadsheet applications (worksheet basics, working with workbooks, working with formulae and functions, an introduction to the use of advanced spreadsheet concepts)
- Database management assignments (Designing and creating databases, sorting records, finding records, filtering records, etc) (5)
- Designing and creation of WebPages
- Introduction to Multi-Media tools & devices (2)

References:

- Introduction to Computers by A. Leon and M. Leon, Vikas Publishing House.
- Fundamentals of Computers by Rajaraman V., PHI.
- Computers Today by Sanders D. H., McGraw Hill.
- Computer Architecture and Organizations by J. P. Hayes, Mc Graw Hill.
- Modern Digital Electronics by R. P. Jain, Tata Mc Graw Hill.
- Computer Network by Andrew S. Tanenbaum, PHI.
- Inter Networking With TCP/IP: Principles, Protocol And Architecture by D.E. Comer Vol1, 2nd Edition Prentice Hall, 1991.

BIM 107 (T+P) Introduction to Database Systems

(2T+2P credits)

Objectives:

- To be able to understand the concepts of data, data models and relationships.
- To be aware of various data representation techniques and various types of databases.
- To appreciate and implement relational database design. The objective is to train the students to efficiently create the database systems with user-friendly front-ends for fast and efficient data retrieval and storage.
- To acquire the skills of using Oracle DBMS, SQL and basic skills in creating front end applications

Theory

Syllabus:

- Database designing, data capturing (4)
- Data Abstraction, Data Models, Instances & Schemes
- E-R Model
 - Entity and entity sets
 - Relations and relationship sets
 - E-R diagrams
 - Reducing E-R Diagrams to tables
- Basic concepts and applications of Network Data Model, Hierarchical Data Model, Multimedia Databases
- Indexing and Hashing (2)
- Basic concepts of (2)
 - ISAM
 - B+ Tree indexed files
 - B Tree indexed files
 - Static Hash functions
 - Dynamic Hash functions
- Textual Databases
- Introduction to Distributed Database Processing

- Data warehousing and Data mining
- Organizing biological species information using various database techniques. (8)
- Studying datasets in biodiversity informatics: Species 2000, Tree of life, ATCC, NBII, Species analyst collaboration, ICTV, Animal Virus Information System etc (as examples of biological databases)

ORACLE, SQL (14)

- Oracle Architecture
- Oracle objects - Tables, Views, Indexes, Sequences; Synonyms, Snapshots, Clusters
- Database - Table space, Data files, Blocks, Extents, Segments; Oracle Background Processes, control files; Oracle Memory Management; Rollback Segments; Redo logs/Archival; Security, Grants, Roles, Privileges
- Oracle Utilities & SQL *DBA - Oracle Server Manager; Export-Import/SQL Monitor Backup & Recovery (Archiving); Physical Storage & Logical Storage
- Oracle * Reports - Reports Features; Full Integration with Forms and Graphics; Data Model and layout editors
- Layout Objects - Frames, Repeating Frames, Fields, Boiler Plate, Anchor; Interface Components; Report Formats; Example Reports; Single Query, Multi Query, Matrix, Master-Detail etc.; User Defined Columns; PL/SQL Interface/ Triggers; Packaged Procedure; Calling Report from a Form
- Menu - Default Menus; Custom Menus; Menu Objects; Menu Module, Main Menu, Individual Menus, Sub Menus, Menu Items; Menu Editor, PL/SQL in Menu Modules, Menu Security
- Select Statements, Data Definition Statements, Data Manipulation Statements, Data Control Statements

Practicals

Syllabus:

Structured Query Language (6)

- Assignment based on “Data Definition Language”.
 - A set of SQL commands used to create table, modify table structure, drop table, rename table.
- Assignment based on “Data Manipulation Language”
 - A set of SQL commands used to change the data within the database.
 - It consists of inserting of records in the tables, updation of all or specific set of records in tables, viewing the attributes of table’s column.
- Assignment based on “Data Query Language”
 - It allows extracting the data out of the database. Selecting the data from table using
 - Arithmetic and logical operators.
 - Range searching and pattern matching.
 - Function, group function, scalar function.
- Assignment based on defining Constraints.
 - Types: I/O constraints like Primary Key, Foreign key, Null and Unique constraints. Business constraints like check constraints.
 - Levels: Table level constraints, column level constraints, creating and deletion of constraints using the Alter Table clause.
- Assignment based on using joins.
 - Joining multiple tables, joining a table to itself.

- Assignment based on using Indexes, Sequences
- Security Management using SQL
 - Granting rights on user objects such as Tables, Views, and Sequences.
 - Revoking rights on user objects such as Tables, Views, and Sequences.

PL/SQL (3)

- Writing PL/SQL blocks by using power of SQL with procedural statements
- Writing Database triggers.
- Writing PL/SQL blocks using Built-in and User-defined function.
- Writing PL/SQL blocks using procedure and package to organize PL/SQL code into logical groups for maintenance and implementation.
- Writing PL/SQL blocks using expressions with operators such as Arithmetic operators, Comparison operators, Logical operators, String operators.
- Handling with cursors in PL/SQL blocks
- Types of cursors: Implicit and Explicit cursors.

Introduction to Front-end application development (2)

- Introduction to Front-end application development, working with code and forms, variables, procedures and controlling program executor, standard controls, data access
- Establishing connection to various databases

Organizing biological species information: Design & normalization of data in database (4)

- Need of metadata standards & ontology

References:

- Database System Concepts by Hanery Korth and Abraham Silberschatz, Tata Mac-Graw Hill.
- An Introduction to Database Systems by C.J. Date, Addison-Wesley.
- Database system organization by J.M. Martin, Princeton-Hall.
- Introduction to Database systems by J.M. Martin; Princeton-Hall.
- ORACLE: Power Objects Handbook by Bruce Kolste, David Peterson.
- Oracle 8 SQL Programming and Tuning by P. Cassidy, 1998.
- SQL, PL/SQL: the programming language of oracle by I. Bayross, Ed. 2, New Delhi. BPB Publications, 2002.
- Oracle SQL & PL/SQL Handbook: a guide for data administrators, developers, and business analysis (With CD) by J. Palinski, Delhi, Pearson Education, 2003.

BIM 108 (T+P): Biological Databases and Data Analysis (2T+2P credits)

Objectives:

- Understand the nature of biological data and need for Biological databases
- Understand and explore the major biomolecular sequence databases (organization and contents) and their respective search engines and database searches
- Understand and appreciate the need and significance of sequence analysis and the bioinformatics approaches for the same
- Application of software analysis tools to sequence data

- To understand the application of methods for analysis of the biomolecular sequence data, retrieval and representation.

Theory

Syllabus:

- Nature of biological data (1)
- Introduction to Bioinformatics (1)
- Overview of available Bioinformatics resources on the web (1)
 - NCBI/EBI/EXPASY etc
- Biological Databases: Nucleic acid sequence databases (2)
 - GenBank/EMBL/DDBJ
- Biological Databases: Protein sequence databases (2)
 - PIR-PSD
 - SwissProt, UniProtKB
- Database search engines (2)
 - Entrez
 - SRS
- Overview/concepts in sequence analysis (2)
- Pairwise sequence alignment algorithms (3)
 - Needleman & Wunsch
 - Smith & waterman
- Scoring matrices for Nucleic acids and proteins (3)
 - MDM
 - BLOSUM
 - CSW
- Database Similarity Searches (3)
 - BLAST
 - FASTA
- Multiple sequence alignment (2)
 - PRAS
 - CLUSTALW
- Biological databases: Genome & genetic disorders (3)
 - Genome databases: Human, model organisms, microbes & viral
 - OMIM
- Biological databases: structural databases (3)
 - PDB
 - NDB
 - CCSD
- Derived databases (3)
 - Prosite
 - BLOCKS
 - Pfam/Prodom

Practicals

Syllabus:

- Exploring the integrated database system at NCBI server and querying the PUBMED and GenBank databases using the ENTREZ search engine (2)
 - Use of operators (AND, OR & NOT)
 - Use of limits

- Exporting GI list
- Batch retrieval
- Exploring the integrated database system at EBI server and searching the EMBL Nucleotide database using the SRS search engine (1)
- Exploring & querying SWISSPROT & UniProtKB (1)
- Exploring and querying the PIR database (1)
- Pair-wise global alignments of protein and DNA sequences using Needleman-Wunsch algorithm & interpretation of results to deduce homology between the sequences, use of scoring matrices (2)
- Pair-wise local alignments of protein and DNA sequences using Smith-Waterman algorithm and interpretation of results (1)
- Database (homology) searches using different versions of BLAST and interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences (2)
- Database (homology) searches using different versions of FASTA & interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences (1)
- Multiple sequence alignments of sets of sequences using web-based and stand-alone version of CLUSTAL. Interpretation of results to identify conserved and variable regions and correlate them with physico-chemical & structural properties (2)
- Exploring and using the derived databases: PROSITE, PRINTS, BLOCKS, Pfam and Prodom for pattern searching, domain searches etc. (2)
- Search & retrieval: genomic and OMIM data at NCBI server (2)
- Studying the format & content of structural databases & visualization of structures using Rasmol, Cn3D and other utilities (2)

References:

- Bioinformatics: A Practical Guide to the analysis of Genes and Proteins (2nd Ed.) by Baxevanis, A.D. & Ouellette, B., F. F., New York, John Wiley & Sons, Inc. Publications, 2002.
- Introduction to Bioinformatics by Attwood, T.K. & Parry-Smith, D.J., Delhi, Pearson Education (Singapore) Pte.Ltd., 2001.
- Bioinformatics: Sequence and Genome Analysis by Mount, David, New York, Cold Spring Harbor Laboratory Press, 2004.
- Current Protocols in Bioinformatics by Baxevanis, A.D., Davison, D.B., Page, R. D. M. & Petsko, G.A., New York, John Wiley & Sons Inc., 2004.

BIM 109 (P) Programming in C

(3P credits)

Objectives:

To be able to conceptualize and formulate the logic and flow for the implementation of a computational task and to code the same using the structured programming approach as provided by the 'C' programming language. Finally the implementation is to be made to various applications in the area of Bioinformatics.

Practicals

Syllabus:

- Concepts of flowcharting, algorithm development, pseudo codes etc. (1)
- Laboratory assignments based on the following topics in 'C' programming: (10)
Data types, operators and expressions, Hierarchy of operators, control statements including decision (if, if-else), loops (while, do-while, for), branching (switch, break, continue), functions, arrays (1D, 2D- all matrix operations including inverse of a matrix), strings, file handling, data structures etc.
- Writing C programs for Bioinformatics applications: (4)
 - Extract a protein or nucleic acid sequence from any of the databank files (GenBank entry, Swiss-Prot, EMBL entry etc.)
 - Interconverting the sequence from one databank format to the other eg. GenBank format to FASTA format, FASTA to PIR format etc.
 - Determining the base composition in a nucleic acid sequence and amino acid composition in a protein sequence.
 - Generating the complimentary sequence of a DNA sequence
 - Pattern search algorithms
 - Search for a specific oligonucleotide pattern (eg. GAACATCC) in a given DNA sequence.
 - Find the position where a specific sequence say "GGTCCCGAC" will hybridize a given DNA sequence.
 - Find the restriction enzyme cleavage sites eg. where PVUZ, ECORI etc. will cut the DNA.
 - Locate palindromic sequence stretches in a DNA sequence.
 - Count the number of Open Reading frames (ORF's) in a DNA sequence.
 - Calculate the codon usage in a nucleic acid sequence.
 - Translate a DNA sequence into protein sequence in the forward and reverse frames.
 - Implementation of the Needleman-Wunsch algorithm for pair wise alignment and testing alignment score with randomized pairs of sequences also.

References:

- The C programming language by Kerighan & Richie, PHI Publication.
- Schaum's outline of programming with C by Byron Gottorfried.
- Programming in ansi 'C' by E. Balaguruswamy, Tata McGraw Hill.
- Let Us C by Kanetkar, BPB Publications.
- Algorithms in bioinformatics by Guigo R. Ed. & Gusfield D., Ed.: Berlin. Springer-Verlag, 2002.

Semester – II

BIM 201 (T) Cell Biology and Genetics

(3 credits)

Objectives:

The objective is to expose the student to the principles of structure and function of cells, membranes and organelles, laws of inheritance and population genetics.

Theory

Syllabus:

A. Cell Biology

(1.5 credits)

- Prokaryotic and eukaryotic cells; membranes and cellular compartmentation; An overview of organelles, (mitochondria, chloroplasts, ER, Golgi, lysosomes and peroxisomes; nucleus and nucleolus) and organelle genetic systems. (3)
- Cellular membranes: Structure, transport, channels, carriers, receptors, endocytosis, membrane potentials. (5)
- Cell motility and shape: cytoskeletal elements, cilia and flagella; motor proteins. (3)
- Cell-cell interactions and signal transduction: Intercellular junctions, signaling by hormones and neurotransmitters; receptors, G-proteins, protein kinases and second messengers (5)
- Protein traffic in cells: Protein sorting and signal sequences; protein translocation in ER and vesicular transport to Golgi, lysosomes and plasma membrane; protein import into nuclei, mitochondria, chloroplasts and peroxisomes. (4)
- Cell cycle and its regulation; events during mitosis and meiosis (3)

B. Genetics

(1.5 credits)

- Science of genetics – objectives, terminologies, methods (1)
- Mendelian principles of inheritance, sex linked inheritance (2)
- Chromosomes, linkage, linkage maps and recombination (3)
- Mutations – molecular, gene/point and chromosomal (2)
- Genetics of viruses and bacteria (3)
- Phenotype and genotype relationships, role of environment, from gene to phenotype, gene interactions (3)
- Study of quantitative traits (3)
- Genetics of populations, genetics and evolution (3)
- Genetics and diseases, cancer (2)

References:

- Essential Cell Biology (2nd edition), by B. Alberts, D. Bary, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts & P. Walter, Garland Science, 2004.
- Molecular Biology of the Cell (4th edition) by B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts & P. Walter Garland Science, 2002.
- The Cell: A Molecular Approach by G. M. Cooper, A. S. M. Press, 2000.
- Molecular Cell Biology (5th edition) by H. Lodish, A. Berk, P. Matsudaira, C.A. Kaiser, M. Krieger, M. P. Scott, S. L. Zipursky & J. Darnell, W. H. Freeman & Company, 2004.
- Principles of Genetics (8th edition) by E. J. Gardner, M. J. Simmons & D. P. Snustad John Wiley & Sons, 2002.

- Concepts in Genetics (7th edition) by W. S. Klug & M. R. Cummings, Pearson Education, 2003.
- Genetics: A Conceptual Approach by B. A. Pierce, W. H. Freeman & Company, 2003.

BIM 202(T) Structural biology & Molecular Modelling (4 credits)

Objectives:

- Understanding the levels of structure organization of macromolecules and related methods of structure determination
- Knowledge of various methods of structure prediction
- Understanding the interaction between macromolecules
- To know the approaches for structure analysis
- To get the deep understanding of molecular modeling and its application

Theory

Syllabus:

- Macromolecular Structure (7)
 - Protein - Primary, Secondary, Supersecondary, Tertiary and Quaternary structure
 - Nucleic acid – DNA and RNA
 - Carbohydrates
 - 3D Viral structures
- Overview of macromolecular x-ray crystallography (5)
 - Methods to study 3D structure
 - Principles of crystallography
 - Co-ordinate systems
 - Fitting and refinement
 - Validation
 - Analysis of 3D structures
- Principles of protein folding and methods to study protein folding (6)
- Structure of Ribosome (2)
- Macromolecular interactions (4)
 - Protein – Protein
 - Protein – Nucleic acids
 - Protein - carbohydrates
- Mass spectrometry and computational approaches in structural biology (1)
- Overview of molecular modelling - Introduction and challenges (1)
- Molecular modelling methods (6)
 - Conformational searching, Potential energy maps,
 - Ramachandran maps, Ab-initio methods, Semi-empirical methods
 - Empirical methods
- Conformational analysis (3)
 - Introduction and Methods
 - Molecular fitting
- Energy Minimisation (5)
 - Non-derivative and derivative methods
 - Global optimisation (simulated annealing, Tabu search, genetic algorithms)
 - Applications of energy minimisation

- Molecular Mechanics (8)
 - Conformations: global vs. local
 - Force fields: expressions for stretch, bond, torsion, etc.
 - Description of various force fields: MM3, Dreiding, AMBER, CHARMM
 - Mechanics of Bio-macromolecules
- Molecular Dynamics (8)
 - Newton's equations for many particles
 - Verlet and related algorithms
 - Types of dynamics simulations: adiabatic, constant T, annealed, etc.
 - Conformational searching using MD and other methods
 - Free energy calculations
 - Dynamics of Bio-macromolecules
- Methods for 3D structure prediction (4)
 - Knowledge based & Fold recognition

References:

- Structural Bioinformatics - Methods of biochemical Analysis V. 44 by Philip E. Bourne (Editor), Helge Weissig (Editor) New Jersey. Wiley-Liss, 2003.
- Principles of protein X-ray Crystallography by Jan Drenth, Springer-Verlag, 1994.
- Introduction to Protein Structure by Branden, Carl & Tooze, John, Garland Publishing, 1991.
- Molecular Modeling: Principles and Applications by Andrew Leach, Prentice Hall, 2001.
- Computational methods for protein folding : advances in chemical physics vol. 120 by Friesner, R.A. Ed., Prigogine, L. Ed. & Rice, S.A. New York. John Wiley & Sons, Inc. publication, 2002.
- Dynamics of Proteins and Nucleic Acids by J.A. McCammon and S.C. Harvey Cambridge University Press, 1987.
- Protein Structure: A Practical approach by Creighton T. E., 1989.
- Protein Folding by Creighton T., 1992.
- Protein Structure Prediction: A practical approach by Sternberg M.J.E., 1996.
- Molecular Modeling: Basic Principles and application by Hans Dieter and Didier Rognan. Wiley VeH GmbH and Co. KGA, 2003.
- Prediction of protein structure and the principles of protein conformation by Fasman, G.D. New York. Plenum Press, 1989.
- Protein modules in cellular signaling edited by Heilmeyer, L. & Friedrich, P. Amsterdam . IOS Press, 2001.
- Metal sites in proteins and models by Hill, H.A.O., Sadler, P.J. & Thomson, A.J Berlin. Springer, 1999.

BIM 203 (T + P) Chemoinformatics

(2T+1P credits)

Objectives:

- To provide an introduction to the major aspects of chemoinformatics, with particular emphasis on applications in modern drug discovery.
- To provide hands-on experience in chemical enumeration, creation of databases and analysis of chemicals.
- To develop tools that aid in the design of drugs.

Theory

Syllabus:

- Role of Chemoinformatics in pharmaceutical/chemical research (2)
 - Integrated databases
 - HTS analysis
 - Ligand based design of compounds
 - Structure based design of compounds
- Structure representation systems, 2D and 3D structures (6)
 - General introduction to chemical structure-hybridization, tetrahedron geometry etc
 - The degeneracy of isomeric SMILES and introduction to unique SMILES. Reaction transformations notation like SMIRKS.
 - Introduction to graph theory, vertex partitioning algorithms- Morgan's and CANGEN algorithms and canonical labeling of the symmetrical vertex
 - Introduction to conformation generating methods. Various ring conformation (sugar) and ring closure problem. Method to identify SSR (smallest subset of ring)
 - Internal co-ordinates and introduction to calculation of Z matrix of simple small organic molecules.
- Chemical Databases – Design, Storage and Retrieval methods (2)
- Introduction to database filters, property based & (drug-like)-Lipinski Rule of Five
- Search techniques, similarity searches and clustering (4)
 - Introduction to molecular pattern finding language- SMARTS
 - Introduction to fingerprints (Daylight, MDL and BCI fingerprints-with special emphasis on the Daylight path based fingerprints algorithms). Introduction to dictionary (of fragments) based fingerprints methods. Strengths and weakness of the fingerprint methods. Introduction to structural keys and hashing and folding fingerprints.
 - Introduction to distance measurement methods from the bit-strings of fingerprints- Tanimoto index and Tversky Index
 - General introduction to clustering- K means and Hierarchical clustering of chemical database
 - Diversity analysis- BCUT descriptors
- Modeling of small molecules and methods for interaction mapping (2)
- Chemical properties 2D and 3D (4)
 - Introduction to adjacency, distance matrix and use of these matrices for calculating Weiner Index, Hosoya Index, Balban Index, Shultz Index, Randic Index (first and higher order).
 - Introduction electrotopological indexes- Kier-Hall Electronegativity, Intrinsic State, and calculation of E-state.
 - Introduction to shape indices- Kappa Shape index and calculation of molecular shape.
 - Physical interpretation of indices.
- Characterization of chemicals by Class & by Pharmacophore, application in HTS Analysis (4)
 - Introduction to pharmacophore
 - Identification of pharmacophore features
 - Building pharmacophore hypothesis
 - Searching databases using pharmacophores
- Quantitative Structure Activity Relationship & application to Hit to lead optimization

- Introduction to QSAR (4)
- Methods- Hansch analysis, Free-wilson QSAR, GFA.
- Correlation matrix, Statistical validation- cross validation techniques
- Practical Session: QSAR using some public domain tools
- Design & Analysis of combinatorial libraries (2)
 - Reagent and product base combinatorial library generation
 - Focus library and HTS library
- Chemoinformatics tools for drug discovery (2)
 - Integration of active drugs
 - Optimization techniques
 - Filtering chemicals
 - In silico ADMET; QSAR approach, Knowledge-based approach

Practicals

Objectives:

- To get familiarize with the chemical database and storage system, retrieval and analysis using physicochemical property and introduction to structure activity

Syllabus:

- Importance of storing chemical in the form of graph, linear notation (SMILES, WLN, ROSDAL-with special emphasis on SMILES and stereochemistry- both optical and geometrical isomerism), connection tables-sd and mol files. (1)
- Graph data structure and its implementation in the context of chemistry.
- Importance of 3D structure and methods available for 3D structure generation- CORINA and CONCORD (1)
- A brief introduction to database (ISIS Base) with special emphasis on the storage of chemical in the database format.
- Substructure searching and general property calculation-rotatable bonds, hydrogen bond donor, hydrogen bond acceptor, molecular weight, molecular refractivity, molecular volume, surface area and polar surface area. (3)
- Writing general SMARTS, Recursive SMARTS and Component level SMARTS and linear representation of chemical concepts like Pka, pH, HBA, HBD, Zwitterions, Functional Groups, Aromaticity (Huckel's Rule). (2)
- Using these properties to design combinatorial library (1)

References:

- Chemoinformatics by Johann Gasteiger and Thomas Engel, 2004.
- An introduction to Chemoinformatics by Andrew R. Leach and Valerie J. Gillet, Kluwer Academic Publisher, 2003.
- Handbook of Chemoinformatics. From Data to Knowledge by Johann Gasteiger.
- Chemometrics and Chemoinformatics by Barry K. Lavine, ACS Symposium series 894.
- Molecular modelling and prediction of bioactivity by Gundertofte, K. & Jorgensen, F.S. New York. Kluwer academic publishers, 2000.

Objectives:

- To provide the system-level understanding of the immune system.
- To take into account the growing relevance of immunology in the design and development of vaccines and immuno-diagnostic tools.

Theory**Syllabus:**

- Immune systems and systems biology (3)
 - Innate and adaptive immunity in vertebrates
 - Antigen processing and presentation
- Antibodies (2)
 - Immunoglobulins
 - Immunoglobulin classes and subclasses
 - CDR and LDR regions and sequence numbering
 - Immunogenetics
 - Hybridoma technology: applications and engineering
 - Humanization of antibodies by design
- Membrane receptors for antigen (5)
 - The B-cell surface receptor for antigen (BCR)
 - The T-cell surface receptor for antigen (TCR)
 - Antigen recognition diversity
 - The major histocompatibility complex (MHC)
- Contemporary challenges to the immune system (3)
 - Infectious diseases
 - Clustering of infectious disease organisms
 - Autoimmune diseases
- Epitopes (5)
 - Affinity Maturation
 - Recognition of Antigen by B cells
 - Neutralizing Antibody
 - Prediction of epitopes
- The primary interaction with antigen (4)
 - The nature of B-cell epitopes
 - Antigen and Antibody interaction
 - The specificity of antigen recognition
 - Binding of the peptides on the MHC
 - Superantigens
 - Tools and servers available
- Vaccine design (3)
 - Categories of vaccines
 - Polytope vaccines
 - Therapeutic vaccines
 - Evolution and escape due to variations
- HLA, immunogenomics and viral bioinformatics (3)
- Generating data for databases – the peptide repertoire of HLA molecules
- HLA nomenclature and IMGT/HLA sequence database
- Mathematical models of HIV and the immune system

- Reverse immunology and approaches in computer aided vaccine design
- Viral bioinformatics: computational views of host and pathogen
- MHC polymorphism (2)
 - Causes of MHC polymorphism
 - MHC supertypes

Practicals

Syllabus:

- Browsing and searching immunological databases (1)
 - Immunoglobulin: sequence and structure
 - Databases of epitopes
- Antibody (1)
 - Antibody numbering
 - Prediction of 3D structure using homology modeling
- Sequence analysis in immunology (1)
 - Alignments
 - Molecular evolution and phylogeny
 - Prediction of functional features of biological sequences
- Methods applied in immunological bioinformatics (2)
 - Sequence weighing methods
 - Pseudocount correction methods
 - Performance measures for prediction methods
- Prediction of cytotoxic and helper T cell epitopes (1)
 - MHC Class I epitopes
 - MHC Class II epitopes
- Web-based tools for vaccine design (2)
 - Databases of MHC ligands
 - Prediction servers
- Predicting immunogenicity (1)
 - Combination of MHC and Proteasome predictions
 - Combination of MHC, TAP, and Proteasome predictions

References:

- Immunological Bioinformatics by Ole Lund, Morten Nielsen, Claus Lundegaard, Can Kesmir, and Soren Brnak, The MIT press, ISBN 0-262-12280-4.
- Immunoinformatics: Bioinformatics Strategies for Better Understanding of Immune Function, ISBN 0-470-09075-8
- Essential immunology. Blackwell Science by Ivan M. Roitt and Peter J. Delves, ISBN 0-632-5902-8.

BIM 205 (T+P) Programming in Object Oriented Languages (2T+ 2P credits)

Objectives:

- To learn & implement the biological problems using object oriented languages

Theory

Syllabus:

- An introduction to JAVA programming (2)
- Object-oriented programming and Java
- Java Basics
- Working with objects (5)
- Arrays, Conditionals and Loops
- Creating Classes and Applications in Java
- More about methods (5)
- Java Applets Basics
- Graphics, Fonts and Color
- Simple Animation and Threads (5)
- Advanced Animation, Images and Sound
- Managing Simple Events and Interactivity
- Creating User Interfaces with AWT (2)
- Windows, Networking and other Tidbits
- Modifiers, Access Control and Class Design
- Packages and Interfaces (3)
- Exception
- Multithreading
- Streams and I/O (3)
- Using Native Methods and Libraries
- Under the Hood
- Java Programming Tools (5)
- Working with Data Structures and Java
- Image Filters

Practicals

Syllabus:

Programs related to biological data to be implemented using:

- Java Basics
- Working with objects (3)
- Arrays, Conditionals and Loops
- Creating Classes and Applications in Java
- More about methods (3)
- Java Applets Basics
- Graphics, Fonts and Color
- Simple Animation and Threads (3)
- Advanced Animation, Images and Sound
- Managing Simple Events and Interactivity
- Creating User Interfaces with AWT (1)
- Windows, Networking and other Tidbits

- Modifiers, Access Control and Class Design
- Packages and Interfaces (1)
- Exception
- Multithreading
- Streams and I/O (2)
- Using Native Methods and Libraries
- Java Programming Tools (2)
- Working with Data Structures and Java
- Image Filters

References:

- Java In a Nutshell by David Flanagan, Oreilly Publications.
- Java Examples in a Nutshell by David Flanagan, Oreilly Publications.
- Java 2: The Complete Reference by Patrick Naughton and Herbert Schildt, McGraw Hill.

BIM 206 (P) Programming in Perl

(3P credits)

Practicals

Objectives:

- To obtain in depth knowledge & implementation in Perl to solve biological problems.

Syllabus:

- Introduction: What is Perl? Why use Perl in Bioinformatics? History of Perl, Availability, Support, Basic Concepts (1)
- Scalar Data: What Is Scalar Data? Numbers, Strings, Scalar Operators, Scalar Variables, Scalar Operators and Functions (1)
- Arrays and List Data: What Is a List or Array? Literal Representation, Variables, Array Operators and Functions, Scalar and List Context (2)
- Control Structures: Statement Blocks (1)
- Hashes: What Is a Hash? Hash Variables, Literal Representation of a Hash, Hash Functions, Hash Slices (1)
- Basic I/O (1)
- Regular Expressions: Concepts About Regular Expressions, Simple Uses of Regular Expressions, Patterns, More on the Matching Operator, Substitutions, The split and join Functions (1)
- Subroutines: System and User Functions, The local Operator, Variable-length Parameter Lists, Notes on Lexical Variables (1)
- Miscellaneous Control Structures: (1)
- File handles and File Tests: What Is a File handle? Opening and Closing a File handle, Using Pathnames and Filenames, A Slight Diversion: die, Using File handles, The -x File Tests, The stat Function (1)
- Formats: What Is a Format? Defining a Format, Invoking a Format (1)
- Directory Access: Moving Around the Directory Tree, Globbing, Directory Handles, Opening and Closing a Directory Handle, Reading a Directory Handle (1)
- File and Directory Manipulation (1)

- Process Management: Using system and exec, Using Backquotes, (1)
- Other Data Transformation: Finding a Substring, Extracting and Replacing a Substring (1)
- Formatting Data: Sorting, Transliteration (1)
- System Information: Getting User and Machine Information, Packing and Unpacking Binary Data
Getting Network Information (1)
- Database Manipulation: DBM Databases and DBM Hashes, Opening and Closing DBM Hashes, Fixed-Length Random-Access Databases, Variable-Length (Text) Databases, Win32 Database Interfaces (1)
- CGI Programming: The CGI.pm Module, Your CGI Program in Context, Simplest CGI Program, Passing Parameters via CGI, Perl and the Web (1)
- Object oriented perl: Introduction to modules, Creating Objects (1)
- Bioperl: Introduction, Installation procedures, Architecture, Uses of bioperl (1)

A tutorial to be developed to include the numerical techniques/algorithms as part of the programs.

References:

- Beginning Perl for Bioinformatics by James Tisdall, O-Reilly.
- Developing Bioinformatics Computer Skills by Cynthia Gibas, Per Jambeck, O-Reilly
- Learning Perl by Randal L. Schwartz, Tom Phoenix, O-Reilly.
- Programming Perl by Larry Wall, Tom Christiansen, Jon Orwant, O-Reilly.
- Programming the Perl DBI by Alligator Descartes, Tim Bunce, O-Reilly.
- Advanced Perl Programming by Sriram Srinivasan, O-Reilly.

BIM 207 (P): Protein Structure Prediction and Molecular Modeling (2P credits)

Practicals

Syllabus:

- Structural data, databases and structure analysis (1)
Exploring the Database & searches on PDB and CSD, WHATIF
- Molecular visualization tools (2)
Visualization of tertiary structures, quaternary structures, architectures and topologies of proteins and DNA using molecular visualization softwares such as RasMol, Cn3D, SPDBV, Chime, Mol4D, etc.
- Structure prediction tools and homology modeling (7)
Prediction of secondary structures of proteins using at least 5 different methods with analysis and interpretation of the results. Comparison of the performance of the different methods for various classes of proteins.
The probable SW lists to be used are SOPM, Insight II, MMTK & Internet based modeling tools.
Prediction of tertiary structures of proteins using Homology Modeling approach: SWISSMODEL, SWISS-PDB Viewer
Prediction of tertiary structures of proteins using at least 3 methods for fold recognition along with analysis and interpretation of results
- Molecular dynamics simulation and docking (5)
Peptide building (PYMOL / DStools)

Minimization of a peptide using appropriate force field
Molecular dynamics analysis using any one of the SW (GROMACS, AMBER & CHARMM)

References:

- Protein structure prediction: methods and protocols by Webster, D. M., Ed. Totowa Humana Press, 2000.
- Modular protein domains by Gimona, G. Cesareni. & Yaffe, M. Sudol (EDS.), USA., Wiley-vch verlag gmbh & co.,3-527-30813-X , Aug. 2004.
- Molecular modeling: basic principles and applications by Holtje, H.D. & Folkers, G., Weinheim, VCH, 1997.
- Molecular Modeling: Basic Principles and application by Hans Dieter & Didier Rognan, Wiley VeH Gmbh and Co. KGA, 2003.

BIM 208(P): Laboratory Techniques in Modern Biology

(2P credits)

Practicals

Objectives:

- This course aims at giving hands-on experience to the students with some of the laboratory techniques necessary to understand genomics and proteomics.

Syllabus:

- A. General (2)
- Washing, Cleaning and sterilization of apparatus, chromic acid preparation.
 - Distillation of water
 - Media and buffer preparation
- B. Basics of Molecular Biology and Biophysics (8)
- Overnight culture inoculation, pouring of agar plates
 - Streaking and spreading of plates
 - Bacterial growth curve
 - Transformation by TSS method
 - Plasmid isolation
 - Gel analysis
 - DNA spectrum analysis
 - Restriction enzyme digestion
 - PH titration of glycine, histidine and glutamic acid
 - CD spectra
- C. Basics of Proteomics (6)
- Extraction and column (affinity, ion-exchange, gel filtration) purification of protein
 - Lowry/Bradford estimation of fractions
 - SDS-PAGE
 - PCR amplification of gene
 - PCR product analysis on agarose gel

References:

- Molecular cloning: A laboratory manual by Tom Maniatis.
- Protein purification applications: a practical approach by Harris, E. L. V., Ed. & Angal, S., Ed., 1989.

- Protein purification from molecular mechanisms to large scale processes by Ladisch, M.R. et.al. Ed., 1990.

Semester III

BIM 301 (T+P) Taxonomy & Phylogeny

(2T+1P credits)

Objectives:

- Understand the concepts in Systematics, Classical Taxonomy and Phylogeny in the viral, microbial, animal and plant universe: basis for classification, nature of characteristic properties used for the same
- Understand the concepts of molecular evolution and the nature of data for deriving molecular phylogeny
- Understand the statistical approaches and models that can be used for Phylogenetic analysis and tree reconstruction
- Understand the computational approaches for Phylogenetic analysis and their applications

Theory

Syllabus:

- Basic concepts in Systematics, Taxonomy and Phylogeny: Species concept; kingdom to species; the five kingdoms; classical, phenetic and cladistic approaches; taxonomic information on viruses, microbes, plants & animals (7)
- Concepts in Molecular Evolution (2)
- Nature of data used in Taxonomy and Phylogeny: Morphological and molecular character data (2)
- Phylogenetic trees and their comparison: Definition and description, various types of trees; Consensus (strict, semi-strict, Adams, majority rule, Nelson). Data partitioning and combination. Tree to tree distances, similarity (3)
- Phylogenetic analysis algorithms (6)
 - Maximum Parsimony
 - Distance-based: UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining
- Probabilistic models and associated algorithms (3)
 - Probabilistic models of evolution
 - Maximum likelihood algorithm
- Approaches for tree reconstruction (3)
 - Character optimization; delayed and accelerated transformation.
 - Reliability of trees. Bootstrap, jackknife, decay, randomization tests.
- Applications of phylogeny analyses (4)
 - Comparison of Phylogenetic Trees obtained using DNA seq. Vs. protein seq. Vs. Full genomes. Need for addition of other properties towards total phylogenetic analysis
 - Comparative methods for detection of species / organism relationships
 - Gene duplication, Horizontal transfer, Domain evolution
 - Study of co-evolution: Plant-insect interactions. Host-parasite interactions.
 - Viral evolution.

Practicals (at least 8)

Syllabus:

- Explore & study the taxonomic resources available for (2)
 - Animals, Plants, Microbes, Viruses etc
- Numerical taxonomy & Bacterial identification using matrices (1)
- Viral identification: ICTV key & probabilistic (1)
- Survey of software programs available for phylogenetic analysis (1)
 - Installation of at least 2 public domain packages for both Windows & Unix environment: Phylip, PAUP, MEGA
- MSA using ClustalW: writing files for phylogenetic analysis (1)
 - Informative and variable sites; Singleton sites etc.
 - Difference between dendrogram & phylogenetic tree
 - Converting sequence data into distance data
- Exploring Phylip package, its utilities and process flow (1)
- Preparing input files using sample datasets (1)
 - Numerical, binary & molecular data
 - Formats & format conversions
- Reconstruction of phylogenetic trees using molecular data (at least 2 datasets)
 - Distance-based methods: UPGMA, Neighbor-joining, Neighbor-relations & Transformed distance (2)
 - Maximum Parsimony (1)
 - Maximum likelihood (1)
- Using bootstrapping tool to generate multiple datasets from the original input data & generation of consensus tree (1)
- Plotting, visualizing & printing phylogenetic trees: TreeView and other tools (1)
 - Various rendering
 - Formatting & labeling
 - Interpretation of trees
- Comparison of trees drawn using (1)
 - RNA, Nucleotide & protein data
 - Gene trees & species tree
- Un-rooted & rooted tree
 - Rooting un-rooted tree using an out group (1)
- Reconstruction of phylogenetic trees using whole genome data of viruses (1)

References:

- Fundamentals of Molecular Evolution by D. Graur and W-H Li, 2nd Edition, Sinauer Associates.
- Molecular Evolution a Phylogenetic Approach by R. D. M. Page and E.C. Holmes, Blackwell Scientific, 1998.
- Protein Evolution by L. Patthy, Blackwell Scientific, 1999.
- Practical taxonomic computing by Pankhurst, R.J., 1991

BIM 302 (T+P) Genome to Drug and Vaccine

(2T + 2P credits)

Objectives:

- Appreciate and understand the changes in the approaches for computational analysis between the pre- and post-genomic era.

- Understand the role of Bioinformatics in the genome sequencing process and post genomic analyses for gene identification, full genome comparison, structural and functional elucidation of genomes, drug target identification etc.
- Appreciate the role of Bioinformatics in post-genomics technologies and areas such as DNA micro-array experiments, Proteomics, protein-protein interactions, pharmacogenetics, identification of disease genes, drug and vaccine design etc.

Theory

Syllabus:

- Overview of genome sequencing: strategies & approaches (2)
- Genome Assembly (2)
- Genome Databases and related data resources (EST, STS, GSS, HSS etc.) (4)
 - Nature and types of data
 - Organization of data in databases; OMIM
 - Genome Data Visualization (With emphasis on Human Genome)
- Tools for Genomic Data Mining (6)
 - Basic Aspects of Genome Annotation
 - Database Search Engines: Special tools for searching genomic data
 - Prediction of ORFs and Genes; Gene Modeling
 - Prediction of Signal sequences (Promoters, Primers, splice sites, UTRs etc.), Operons
- Identification of Disease Genes (8)

(in the context of Human Genetics and Genetics of Model Animals)

 - Identification of Drug Targets
 - Metabolic diseases and Pathogenic diseases
 - Gene Expression Analysis
 - Structural Genomics
 - Functional Genomics
- Pharmacokinetics (4)
 - Classification
 - Case study
 - Comparative screening
 - ADMET: Drug metabolism; Role of cytochromes P450; Elimination half-life; Toxicity screening
 - Pharmacogenetics (2)
 - The genetics of drug metabolism
 - The genetics of therapeutic targets
 - Interactions of small molecules and gene-based drug targets (1)
 - Proteome analysis and Prediction of epitopes on Genomic scale (2)

Practicals

Syllabus:

- Browsing & viewing genome data (1)
 - Ensembl@EBI
 - MapViewer@NCBI
- Viewing regions exhibiting Synteny
 - Genome assembly (1)
 - Genome Annotation (2)

- Using integrated genome annotation servers such as the server developed at IMTech, Chandigarh (<http://imtech.res.in/raghava/gp.html>) (2)
- Gene Prediction and Gene Modeling
 - Prediction of genes and gene structures (gene modeling) using online (web) servers of different methods tailored for prokaryotic and eukaryotic organisms such as GLIMMER, GeneMark, Grail, GENSCAN etc. Interpretation of results and comparison with known gene models (where available). Evaluation of accuracy of the methods. (2)
 - Prediction of promoters using methods such as Neural Network Promoter Prediction (NNPP) at Berkeley Drosophila Genome Project server, Genome inspector for combined analysis of multiple signals in genomes etc. Using Promoter databases. (1)
 - Prediction of alternate splice sites using methods such as Splice Site Prediction by Neural Network (at Berkeley Drosophila Genome Project server), GenScan, NetGene2 GeneSplicer etc. Prediction of PCR primers using Primer 3, ePCR etc. (1)
- Functional Genomics
 - Using primary databases (such as UniProt) and derived databases such as InterPro, PRINTS, BLOCKS, PRODOM, Pfam etc. along with advanced sequence analysis tools such as profiles searches, pattern searches for function annotation of genomic sequences. Validation and verification of results for known case studies (2)
 - Using sequence-based and structure-based Function Annotation Servers such as (2)
 - ProKnow (<http://www.doe-mbi.ucla.edu/Services/ProKnow/>),
 - Joined Assembly of Function Annotations (JAFA) at <http://jafa.burnham.org/learnMore.html> etc. which are integrated services for function annotation
 - ProFunc (<http://www.ebi.ac.uk/thornton-srv/databases/ProFunc>)
- Drug target and Vaccine target identification (2)

References:

- Guide to Human Genome Computing by Martin J. Bishop, Academic Press. ISBN 0-12-102051-7.
- From Genome to Therapy: Integrating new technologies with drug development by Novartis Foundation, John Wiley. ISBN 0-471-62744-5.
- Genome mapping and sequencing By Ian Dunham, Horizon, ISBN1-898486-50-6.
- The Genome by Ram S. Verma, VCH, ISBN 1-56081-043-2.
- Bioinformatics - from genomes to drugs (vol. 1), basic technologies (vol.1) by Lengauer, T., Germany, Wiley-VCH, 2002.
- Vaccines: New Approaches to Immunological problems by Ronald W. Ellis, BH. ISBN 0-7506-9265-0.
- Human, Vaccines and Vaccination by M. Mackett and J.D. Williamson, BIOS. Scientific, ISBN 1-872748-7-5.
- Principles of Genome Analysis And Genomics (3rd Ed.) by Primrose, S.B. & Twyman, R.M., UK. Blackwell Publishing Company, 2003.
- Bioinformatics approach Guide to the analysis of genes and proteins by Andreas Baxevanis and B.F. Francis Ouellette. John Wiley 2004.

- Pharmacogenomics: the search for individual therapies by Licinio, J. & Wong Ma-Li: Germany, Wiley-VCH, 2002.
- Vaccines: Prospects and Perspectives (Vol. I and II) by Harminder Singh and Rajesh Bhatia, ISBN 81-7367-000-5 and 81-7367-001-3.

BIM 303 (T+P) Comparative Genomics and Proteomics (2T+2P credits)

Objectives:

- Appreciate the importance of full genome comparisons
- Understand the various algorithms used for comparisons of full genome and gene order
- Use the methods based on various algorithms and interpret the results.
- Understand and appreciate the full genome comparative studies of viruses, microbes, pathogens and eukaryotes
- Understand and explore the comparative genomics databases
- Understand the concepts of SNPs and their significance
- Understand and appreciate the proteomics concepts and technology
- Understand the basis and nature of protein-protein interactions and related databases

Theory

Syllabus:

- Objective and Overview of Genome Comparisons (1)
- Genome Alignments (5)
 - BLAST2
 - MUMmer
 - PipMaker
 - VISTA
- Comparison of Gene Order (1)
 - GeneOrder
- Comparative Genomics (6)
 - Viruses
 - Microbes
 - Pathogens
 - Eukaryotes
- Comparative Genomics Databases (6)
 - COG
 - VirGen
 - CORG
 - HOBACGEN
 - Homophila
 - XREFdb
 - Gramene
- Single Nucleotide Polymorphism, dbSNP and other SNP-related databases (2)
- Overview of Proteomics (3)
 - Experimental Techniques
 - Bioinformatics Approaches
- Protein-Protein Interaction Networks, databases and software (6)
 - DIP (Database of Interacting Proteins)

- PPI Server
- BIND - Biomolecular Interaction Network Database
- PIM -Hybrigenics
- PathCalling Yeast Interaction Database
- MINT - a Molecular Interactions Database\
- GRID - The General Repository for Interaction Datasets
- InterPreTS - protein interaction prediction through tertiary structure

Practicals

Syllabus:

- Explore comparative genomics resources and NCBI and EBI (1)
- Comparison of full / partial genomic sequences using following methods to identify conserved genes and map/compare the annotations of the two sequences
 - BLAST2 (1)
 - MegaBLAST and Discontiguous MegaBLAST (1)
 - MUMmer (1)
 - PipMaker (2)
 - VISTA (1)
- Compare gene order of given genomic sequences using the GeneOrder tool (1)
- Explore and query the comparative genomics databases: COG, VirGen, CORG, HOBACGEN, Homophila, XREFdb, Gramene etc. (3)
- Explore and query SNP and SNP-related databases (1)
- Explore and query the protein-protein interaction databases: DIP, PPI Server, BIND, PIM, PathCalling, MINT, GRID, InterPreTS (3)

References:

- Proteomics of microorganisms by Hecker, M. & Mullner, S., Berlin, Springer-Verlag, 2003.
- Introduction to proteomics: tools for the new biology by Liebler, D.C. & Yates, J.R. III, New York. Humana Press, 2002.
- Proteomics: from protein sequence to function by Pennington, S. R. & Dunn, M. J.: New Delhi, Viva Books Private Ltd, 2002.
- Bioinformatics: sequence and genome analysis by David Mount, cold springer harbour press, 2004.
- Comparative genomics: empirical and analytical approaches to gene order dynamics, map alignment and the evolution of gene families by Sankoff, D. & Nadeau, J.H., Netherlands, Kluwer Academic Publishers, 2000.

BIM 304 (T) Parasite Bioinformatics

(2 credits)

Objectives:

- Understanding of life cycles and biology of select parasites as well as their interactions with the host and vector
- Understanding the role of Bioinformatics in combating parasitic diseases through parasite-specific databases and analysis of genomic, proteomic data
- Appreciating the need and approaches for novel drugs/vaccines for parasitic diseases in the context of multidrug resistance

Theory

Syllabus:

- Parasitic Diseases to be covered
 - Malaria
 - Leishmaniasis
 - Trypanosoma
 - Entamoeba histolytica
 - Toxoplasma
 - Filariasis
- Biology of Parasites
 - Life Cycle, Infectivity, Demographic distribution of strains (4)
- Parasite Genome and Proteome Databases (3)
- Vectors of parasites (3)
 - Biology of vectors
 - Genome & Proteome databases: Mosquito
- Application of Bioinformatics Data Mining tools for Identification of (6)
 - Parasite-specific genes/gene products (e.g. house-keeping genes, genes essential for survival)
 - Resistant Genes
 - Tools: Full Genome Comparison, Gene Prediction, Signal sequence prediction, Protein sequence comparison and analysis, Protein structure comparison and analysis
- Micro Array and Proteomics Data Analysis (2)
- Structural genomics of parasites (2)
- Host-Parasite and Host-Vector-Parasite Interactions (3)
 - Pathway databases
- Multi-Drug Resistance (3)
 - Mechanism of MDR: genomic, molecular, cellular
 - Identification of genes responsible for MDR
 - Approaches to novel drug discovery
- Challenges and opportunities in vaccine development (2)
- Plant Parasites and diseases (2)
 - Disease resistance genes of plants
 - Plant-pathogen interactions

References:

- Parasitism by Bush, A. O., Fernandez, J. C., Esch, G.W. & Seed, R. J., Cambridge. Cambridge University Press, 2001.
- Parasite genomics protocols by Melville, S.E., New Jersey. Humana Press, 2004.
- A modern textbook of Parasitology by Latey, A.N, Pune, Narendra prakashan, 1991.
- Modern parasite biology: cellular immunological and molecular aspects by Wyler, D.J. Ed., 1990.

BIM 305 (T) Object Oriented and Relational Databases

(2 credits)

Objectives:

- To learn the techniques in relational and object oriented databases.

Theory

Syllabus:

- Part 1: Basic Concepts (3)
 - Introduction
 - Database System Concepts and Architecture
 - Entity-Relationship Model
 - EER and Object Modeling
- Part 2: Relational Databases (3)
 - Relational Model, Algebra, Calculus
 - ER- and EER-to-Relational Mapping
 - Relational languages SQL and QBE
 - RDBMS Systems: SQL server and MS Access
- Part 3: Object-Oriented Database Systems (3)
 - Object-oriented concepts
 - Object Modeling
 - Object-Oriented Databases
 - Object Database Languages
 - Object Database Design
 - Object-Relational and Extended Relational Database Systems
- Part 4: Database Design (3)
 - Functional Dependencies
 - Normalization
 - Design Algorithms and Further Dependencies
- Part 5: System Implementation Techniques (3)
 - Query Processing and Optimization
 - Transaction Processing
 - Concurrency Control
 - Recovery
 - Security and Authorization

Selected Advanced Topics

- Distributed Databases and Client-Server Architecture (3)
 - Distributed Database Architecture
 - Data Fragmentation, Replication, and Allocation
 - Distributed Query Processing
 - Distributed Concurrency Control
 - Client-Server Architecture
- Active Databases (3)
 - Event Languages, Event Detection and Delivery
 - Rule Processing

- Applications of Active Databases
- Data Management in Mobile Computing
- Caching and Prefetching
- Data Replication
- Speculative Data Dissemination and Broadcast Disk
- Mobile and Distributed Query Processing
- Mobile Transactions
- Logic and Deductive Databases (3)
 - Prolog/Datalog
 - Rule Interpretation
 - Inference Mechanisms
 - Program Evaluation
 - Deductive Databases
- Multimedia Databases (2)
 - Multimedia Networking
 - Multimedia Storage and Indexing
 - Content-based Multimedia Information Retrieval
- Data Warehousing and Data Mining (2)
 - Data Warehousing
 - Data Mining
 - OLAP
- Database Systems and the World-Wide-Web (2)
 - Connecting Database to the Web
 - Web Search
 - XML and the New Generation Web

References:

- Database Management and Design by G.W. Hansen and J.V. Hansen; Prentice-Hall of India
- Database System Concepts by A. Silberschatz, H.F. Korth and S. Sudarshan; McGraw-Hill
- Database Systems: The Complete Book by Garcia-Molina, J. D. Ullman, and J. Widom; Prentice Hall.
- Fundamentals of Database Systems by Ramez Elmasri and Shamkant B. Navathe, Addison-Wesley.
- Database Management Systems by R. Ramakrishnan and J. Gehrke; McGraw-Hill
- Database Systems by T. Connolly and C. Begg.; Addison-Wesley.
- Database Management Systems by A.K. Majumdar and P. Bhattacharyya; Tata McGraw-Hill
- An Introduction to Database Systems by C.J. Date, Addison-Wesley.
- Fundamentals of Database Systems by R. Elmasri and S.B. Navathe; Addison-Wesley.
- Modern Database Management by R.F. McFadden and J.A. Hoffer, Benjamin-Cummins (Narosa).
- Data structures, algorithms, and object oriented programming by Heileman, G.L New Delhi, Tata McGraw-Hill Publishing Company Limited, 2002.

BIM 306 (T+P) Scientific visualization for data analysis (2T+1P credits)

Objectives:

- To introduce the basic concepts in computer graphics as applied to biomolecules.
- To be aware of the various algorithmic approaches implemented for graphical applications.
- To appreciate visualization packages used in Bioinformatics and to be able to develop better graphical and visualization aids for ones own requirement and application.

Theory

Prerequisites: Course on Knowledge on Mathematics for Bioinformatics or equivalence.

Syllabus:

- Introduction (1)
- Scientific & Engineering Opportunities (1)
- Visualization techniques, software standards and hardware (4)
- Interactive Graphics (2)
- Interaction devices & techniques (1)
- Frame buffer techniques (1)
- Geometric Transformations (4)
- 2D viewing (2)
- Viewing in three dimension (2)
- Surface and solid modeling (2)
- Rendering (2)
- Animation (2)
- User interface study (2)
- Image processing & 3D Image reconstruction (3)

Practicals

Prerequisites: Sound knowledge and practical experience with C / Java programming

Syllabus:

- Standard functions in the 'C' graphics module – Introduction to pixels, Drawing various geometric objects, filling polygons, images etc. (1)
- Implementation of Raster algorithms (1)
- Geometric transformations in 2D (1)
- Clipping algorithms (1)
- Geometric projections – parallel and perspective (1)
- Lighting and Shading – Implementing Phong illumination model, Gouraud and Phong shading, Animation (1)
- Scalable Vector Graphics (SVG) (2)
- OpenGL (1)

References:

- Computer Graphics - D. Hearn and M. Baker
- Fundamentals of Interactive Computer Graphics - J.D.Foley, A.Van Dam
- Procedural Elements for Computer Graphics - David F. Rogers.
- Mathematical Elements for Computer Graphics - David F. Rogers, J.A.Adams
- Graphics Gems - Andrew S. Glassner
- IEEE Computer Graphics & applications (monthly journal)

BIM 307 (T+P) Tools & Techniques for biological data mining (2T+2P credits.)

Objectives:

- To be able to understand and implement various advanced mathematical, physical and statistical techniques that are being applied for analyzing information biological data.

Theory

Syllabus:

- Quality of Biological Data & Data Accuracy (2)
- General issues regarding Biological Databases; Representation of errors due to (machines, 3D structural and sequence data of proteins and nucleic acid, Proteomics and Microarray data)
- Optimization Techniques (6)
 - Steepest Descent, Conjugate Gradient, Newton-Raphson
 - Simulated annealing in Biomolecular Structure Optimization
- Genetic Algorithms (3)
- *Ab initio* methods for structure prediction (4)
 - Lattice, SOM, etc.
 - Information theory, entropy and relative entropy
 - Stochastic Grammars & Linguistics
- Clustering & Classification Algorithms (4)
 - Hierarchical and non-hierarchical Clustering
 - K-Means clustering
 - Grid based clustering
 - Analysis of MD trajectories
 - Microarray and Protein Array data Analysis
- Dynamic Programming & application in (2)
 - Sequence Alignments
 - Structure Alignments
- Foundations for Machine learning Techniques: (8)
 - Hidden Markov Model
 - Neural Network
 - Bayesian modeling
 - The Cox-Jaynes Axiomes
- Support Vector machine & Ant colony optimization applied to (4)
 - Multiple Sequence Alignments
 - Biomolecular Structure Prediction
- Fuzzy logic system & application in (4)
 - Clustering and classifications
 - Microarray and Protein Array data Analysis

Practicals

Syllabus:

- Neural networks: (5)
 - Use of neural network tools like BrainBox, MATLAB etc.
 - Program in 'C' for implementing the perceptron learning rule and the delta learning rule (adaline network) for say logic gate patterns.
 - Implementation of the Backpropagation learning algorithm for the XOR gate and for an application such as protein secondary structure prediction, promoter prediction etc.
 - Estimator of transition probabilities for markov models based on various sample sizes.
 - Hidden Markov model implementation in C/ Java. (2)
 - Ant colony algorithm for the Travelling salesman problem (TSP) & Implementation of SVM (4)
- Optimization Algorithms: (5)
 - Programming in 'C' for implementation of Golden section algorithm, Steepest descent, Newton Raphson, Conjugate gradient etc for energy minimization applications.
 - Implementation of random walk and Monte Carlo algorithm.
 - Simulated Annealing algorithm for energy minimization.

References:

- Data Mining: Concepts and Techniques by Han and Kamber, Morgan Kaufmann.
- Machine Learning by Tom Mitchell, McGraw Hill.
- Data Mining: Practical Machine Learning Tools and Techniques by Witten and Frank, Elsevier.
- Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G. Cambridge Univ. Press, 1998.
- Optimization Theory and Application by Rao, S.S., 1984.
- Discrete optimization by Parker, R. G. & Rardin, R. L., 1988.
- Stochastic simulation by Repley, Brian D, Wiley series, 1987.
- Methods of microarray data analysis III by Johnson, K.F. & Lin, S.M. Boston. Kluwer academic publishers, 2003.
- Exploration and analysis of DNA microarray and protein array data by Amaratunga, D. & Cabrera, J. New Jersey. John Wiley & Sons Inc., 2004.
- Ant colony optimization by Dorigo, Marco & Stutzle, Thomas New Delhi, Prentice-Hall of India Pvt Ltd, 2004.
- Data mining: introductory and advanced topics by Dunham, M.H.: New Delhi, Pearson Education, 2003.
- An introduction to bioinformatics algorithms by Jones, Neil.C. & Pevzner, Pavel A. New Delhi, Anne Books, 2005.
- Fuzzy sets and fuzzy logic: theory and applications by Klir, G.J. & Yuan Bo, New Delhi. Printice Hall of India, 2002. 81-203-1136-1.
- Fuzzy and neuro-fuzzy systems in medicine by Teodorrescu, Horia-Nicolai, Kandel, Abraham. & Jain, Lakhmi, C New York, CRS Press, 1999.

BIM 308(P): Computer implementation of Data structures and Data integration (3P credits)

Objectives:

- To be able to understand the existing data structures in various databanks and databases in terms of the data model, relationships between entities etc., attempts being made to create relational databases, object oriented relational databases and the tools for data integration.
- The student should be able to use the knowledge gained in the various computational courses such as Database systems (Oracle, SQL, VB etc), Perl, CORBA etc for the development of value added databases.

Practicals

Syllabus:

- Drawing ER diagrams for the primary sequence databanks to convert them into relational models (Swiss-Prot, PIR, GenBank, EMBL, DDBJ) (2)
- Parse the records from various entries of the primary databank. Create tables with appropriate relationships using a RDBMS tool, populate tables and query them using SQL. (2)
- In-depth study of the PDB relational model as in the EBI Macromolecular structure database. Compare with the PDB flat file and study the various fields. (2)
- Study the secondary / derived databases and create suitable relational models for the same. (Prosite, ProDom, Profiles, PRINTS, Pfam, BLOCKS etc.) (2)
- Assignment on InterPro as a model for Data Integration. (1)
- PlasmoDB, ICTV DB, Species 2000, VirGen, MPE (Metabolic pathway engineering) etc. as other models for Data Integration. (2)
- Optimize /derive two of the above mentioned databases like GenBank & PDB and port it locally & make searchable with use friendly frontend. (4)

CORBA (7)

Provide a training for GRID computing concept

- Common Object Request Broker Architecture
Distributed computing, Introduction to CORBA, About Object management group, CORBA architecture, architectural similarities, CORBA versus Java RMI, CORBA services, CORBA facilities- (Vertical and Horizontal facilities) CORBA domains. IDL (Interface definition language): compiler, stubs, skeletons, Repositories, ORB (Object request broker); Naming service;
- Inter-ORB communication
Creating CORBA objects; IDL modules and interfaces, data members and methods; IDL and Java; Simple server class, helper class, holder class, client and server stubs; Initializing ORB, Registering with a naming service; Adding objects to a naming context; Finding remote objects; Initial ORB references; Getting objects from other remote objects, Stringified object references; Dynamic invocation interface.

References:

- Data Structures Using C by Tannenbaum, Y. Lanhgsam and A.J.Augenstein, Prentice Hall of India.
- Theory and Practice of Data structures by Seymour Lipschultz, Mc Graw Hill
- The Essential Distributed Objects Survival Guide by Orfali and Harkey, Wiley.

- Client/Server Programming with JAVA and CORBA by Orfali and Harkey, Wiley
- CORBA Fundamentals and Programming by Siegel, Wiley.
- Teach Yourself CORBA in 14 Days by Jeremy L. Rosenberger, SAMS.
- CORBA: A Guide to Common Object Request Broker Architecture by Ron Ben-Natan, McGraw-Hill.
- Understanding Corba by Randy Otte, Paul Patrick, Mark Roy, Prentice Hall.

Semester IV

BIM 401 (T+P) Advanced Techniques for Sequence and Structure Analysis (2T+2P credits)

Objectives:

- To be able to understand and implement various advanced level techniques for analyzing information at sequence and structure.

Theory

Syllabus:

- Advanced Techniques for Sequence Analysis (5)
 - Sequence Profiles: Derivation, Databases, Application
 - Gapped BLAST, PSI-BLAST, PHI-BLAST
- Advanced Techniques for Structure Analysis (3)
 - Molecular replacement method, Direct method & Fiber diffraction
 - Methods for Comparison of 3D structures
- Application of the optimization techniques in (3)
 - Sequence Alignments
 - Prediction of Protein Structure
 - Docking Simulations
- Advance techniques in Prediction of 3D Structure (4)
 - Hidden Markov Model
 - Neural networks
 - Rossetta Stone
- Molecular Dynamics Simulations & Monte Carlo Methods (6)
 - Electrostatics of biomolecules
 - Simulations of Bio-macromolecular Structures in Water & membrane
 - Free energy perturbation method
- Simulated Annealing (2)
 - Multiple Sequence Alignments
 - Simulations of Bio-macromolecular Structures
- Designing of molecules like drug, inhibitors using (4)
 - Structure based & ligand based docking methods
 - Different Scoring schemes

Practicals

Syllabus:

- Data mining: (5)
 - Determine the entropy in a DNA sequence, relative entropy in 2 sequences etc.
 - Implement the clustering algorithm for biological data like Phylogenetic Analysis
 - MD trajectories & Microarray and Protein Array data Analysis
 - Program the UPGMA and EM algorithm for Clustering.
 - Dynamic programming application to structural alignment.
- Molecular simulation in water and membrane (5)
- Characterization of active site, Substrate based design , docking & scoring of chemical data base (5)

References:

- Hidden markov models for Bioinformatics by Koski, T, Dordrecht Kluwer, Academic Publishers, 2001.
- Neural Networks: A Comprehensive Foundation (2nd Edition) by Simon Haykin.
- Genetic algorithms in search, optimization and machine learning by Goldberg, D.E. 2002 Delhi, Pearson Education Pte. Ltd.
- Molecular dynamics of lipid bilayers by Nagumo, M. & Bruce, P.G., 1996.
- Genetic algorithms in molecular modeling by Devillers, J., Ed, 1996.
- Molecular dynamics: applications in molecular biology by Goodfellow, Julia M., 1991.
- Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G.: Cambridge Univ. Press, Cambridge, 1998.
- Methods in Enzymology, (V. 183), Molecular evolution: computer analysis of protein & nucleic acid seqs. 1990.
- Sequence analysis in molecular biology: treasure trove or trivial pursuit by Heijne, Gunnar von, 1987.
- Simulation Modeling and Analysis by Averill M. Law and W. David Kelton, Tata MacGraw – Hill, 2000.
- Protein Bioinformatics: An Algorithmic Approach to Sequence and Structure Analysis by Ingvar Eidhammer, Inge Jonassen, William R. Taylor John Wiley, 2003.

BIM 402 (T) Metabolomes and Metabolic Pathway Engineering (2 credits)

Objectives:

- Learning the basic biochemical concepts of metabolic pathways
- Understanding the role of Bioinformatics in the study of metabolic pathways
- Learning the Bioinformatics-based approaches for predicting and engineering metabolic pathways

Theory

Syllabus:

- Enzyme Nomenclature and Classification (3)
- Major Metabolic Pathways (10)
 - Gluconeogenesis, Pentose phosphate pathway, Glycogen synthesis and degradation, Fatty acid oxidation and synthesis, Amino acid catabolism, Purine and pyrimidine nucleotide synthesis etc.
- Metabolic Pathways databases
 - KEGG (2)
 - EcoCyc and MetaCyc (2)
 - EMP, Malaria Parasite Metabolic Pathways, Boehringer Mannheim - Biochemical Pathways (1)
- Enzymes, Compounds and Reactions databases (2)
 - LIGAND - Biochemical Compounds and Reactions
 - ENZYME - Enzymes
 - BRENDA - Comprehensive Enzyme Information System
- Full Genome Annotation through knowledge of Metabolic Pathways (1)
- Organism Specific Metabolic Pathways (1)

- Comparison of Metabolic Pathways (1)
- Mathematical modeling of metabolic pathways (3)
- Engineering of Metabolic Pathways (2)
- Representation of Metabolic Pathways (1)
- Generation and Dynamic Representation of Metabolic Pathways (1)
- Deriving Common Principles from the Metabolic Pathways Knowledge
 - E.g. deriving sets of enzymes specific for various reactions (e.g. oxidation), alternative paths for synthesis of metabolites etc.

References:

- Fundamentals of Biochemistry (2nd edition) by D., Voet, Voet, J.G. & Pratt, C. W. John Wiley & Sons, 2006.
- Biochemistry (3rd edition) by Voet D. & Voet, J.G., John Wiley & Sons, 2004.
- Biochemistry (3rd edition) by D. Voet & J. G. Voet, John Wiley & Sons, 2004.
- Lehninger Principles of Biochemistry (4th edition) by D. L. Nelson & M. M. Cox, W. H. Freeman & Co, 2005.
- Principles of Biochemistry (4th edition) by H. R. Horton, L. A. Moran, K. G. Scrimgeour, M. D. Perry & J. D. Rawn, Pearson-Prentice Hall, 2006.
- Biochemistry (5th edition) by J. M. Berg, J. L. Tymoczko & L. Stryer, W.H. Freeman & Co., 2002.
- Biochemistry (3rd edition) by C. K. Mathews, K. E. Van Holde & K. G. Ahern, Pearson Education, 2000.
- The Enzyme Reference by D. L. Purich & R. D. Allison (2002) Academic Press. IUBMBM,1992.
- Enzyme Nomenclature, Academic Press, 2002.
- Metabolic Engineering by G.N. Stephanopoulos, A. A. Aristidou & J. Nielsen Academic Press, 2006.
- Gene regulation and metabolism: postgenomic computational approaches. By Collado-Vides, J. & Hofestadt, R. Cambridge, The MIT Press, 2002.
- Metabolic profiling: its role in biomaker discovery and gene function analysis by Harrigan, George G. & Goodacre, Royston, London, Kulwer Academic publishers, 2003.

BIM 403 (T) Emerging Areas in Bioinformatics.

(2 credits)

Objectives:

Recent developments in Life sciences in the context of Bioinformatics will be discussed. A series of lectures could be planned on any topics for introducing the subject.

Theory

Syllabus :

Series of talks on any 2-3 topics such as Biodiversity informatics, Neuro bioinformatics, Epigenetics, Advanced Glycobiology, Nano –biotechnology, System Biology, Agro bioinformatics etc.

PRACTICAL COURSES

<i>Course No.</i>	<i>Course Name</i>	<i>No. of credits</i>
BIM 404 (P)	Project	16

Any one of the followings:

<i>Course No.</i>	<i>Course Name</i>	<i>No. of credits</i>
BIM 405 (T+P)	Seminars on Applications of Bioinformatics in Agriculture	2
BIM 406 (T+P)	Seminars on Applications of Bioinformatics in Human Health	2
BIM 407 (T+P)	Seminars on Applications of Bioinformatics in Environment	2
BIM 408 (T+P)	Seminars on Applications of Bioinformatics in Biotechnology	2
BIM 409 (T+P)	Seminars on Applications of Bioinformatics in Molecular Biology	2
BIM 410 (T+P)	Seminars on Applications of Bioinformatics in Neurobiology	2
BIM 411 (T+P)	Seminars on Applications of Bioinformatics in Drug Designing	2
BIM 412 (T+P)	Seminars on Applications of Bioinformatics in Veterinary Sciences	2