

## UNIVERSITY OF HYDERABAD

School/ Department/ Centres: Life Sciences/ Biotechnology

COURSE STRUCTURE FROM 2011-2012 BATCH ONWARDS

Course: M.Tech

Subject: Bioinformatics

Semester: First No of Courses: 9 Core No. of Prescribed Credits: 25

S.No	Course Code	Course Title	Credits	Core/ Elective	Theory/ Lab
1	BI701	Computational Techniques	3	Core	Theory
2	BI702	Proteomics	3	Core	Theory
3	BI703	Basic Mathematics & statistics	4	Core	Theory
4	BI704	Molecular Modeling and Chemoinformatics	4	Core	Theory
5	BI05	Bioinformatics	3	Core	Theory
6	BI706	Lab I-Computational Techniques	2	Core	Lab
7	BI707	Lab II- Molecular Modeling and Chemoinformatics	2	Core	Lab
8	BI708	Lab III- Proteomics	2	Core	Lab
9	BI709	Lab IV -Bioinformatics	2	Core	Lab
10	BT705	Molecular Biology*	0	Remedial	Theory
Total Semester I			25		

\*Remedial courses.

Semester: Second No of Courses: 10 core No. of Prescribed Credits: 23

S.No	Course Code	Course Title	Credits	Core/ Elective	Theory/ Lab
1	BI711	Genomics	3	Core	Theory
2	BI712	Mathematical Modeling of Biological Systems	3	Core	Theory
3	BI713	Machine Learning Methods and Biological Network	3	Core	Theory
4	BI714	Molecular Simulation and Drug Design	3	Core	Theory
5	BI715	Metagenomics	3	Core	Theory
6	BI716	Lab V-Genomics	2	Core	Lab
7	BI717	Lab VI- Mathematical Modeling of Biological Systems	2	Core	Lab
8	BI718	Lab VII- Machine Learning Methods and Biological Network	2	Core	Lab

**M. Tech Bioinformatics****Draft course from 2011-2012**

9	BI719	Lab VIII: Molecular Simulation and Drug Design	2	Core	Lab
10	BI720	Seminar/Journal Club	1	Core	Seminar
TOTAL Sem. II			24		

**Semester: Third****No of Courses: 2+2****No. of Prescribed Credits: 12**

S.No	Course Code	Course Title	Credits	Core/ Elective	Theory/ Lab
1	BI731	Elective-I	2	Elective	Theory
2	BI732	Elective-2	2		Theory
3	BI733	Project Proposal Presentation	2	Core	Project
4	BI734	Project Evaluation-1	6	Core	Project
TOTAL Sem. III			12		

**Electives being offered**

S.No	Course Code	Course Title	Credits	Core/ Elective	Theory/ Lab
1	BI752	Gene Expression	2	Elective	Theory
2	BI753	Java programming for Bioinformatics applications	2	Elective	Theory
3	BI754	Databases	2	Elective	Theory
4	BI755	Biological networks	2	Elective	Theory
5	BT	Bioentrepreneurship			

**Semester: Fourth****No of Courses: 2****No. of Prescribed Credits: 12**

S.No	Course Code	Course Title	Credits	Core/ Elective	Theory/ Lab
1	BI765	Project Evaluation-2	10	Core	Project
2	BT542	Project Seminar	2	Core	Project
TOTAL Sem. IV			12		

**Summary**

Semester	No. Courses	Credits
First	9	25
Second	9	23
Third	2	4
Fourth	0	0
Seminars	3	5
Project	2	16
Total	22	73

**University of Hyderabad**  
**M. Tech Bioinformatics (to be Revised from 2011-12)**

**DRAFT SYLLABUS FOR M. Tech. BIOINFORMATICS****Detailed Syllabus:**


---

<b>BI701: Computational Techniques</b>	<b>I Semester</b>	<b>3 Credits</b>
--	-------------------	------------------

---

- Computer fundamentals: Basic units of computer, Architecture
- Programming concepts: Reduction of a given problem to a computer program; flow chart concepts. Conversion of the flow chart to code
- Key concepts in programming: variables and constants; different kind or types of variables; storage needs; arrays; concept of address and value; keywords and reserved words; syntax.
- Programming tools: sequential execution; branching; conditional branching; different kinds of loops
- Machine dependent conditions: address; segment; offset; pointers; function calls; memory allocation; entry exit codes; large programs (overlays)
- Introduction to C: general features of a C program
- Arrays and pointers in C: detailed discussions
- Input and output: files: basic concepts of stream I/O
- Strings and string operations; string handling
- Scripting using Perl, Python
- Compilers; options and optimizations
- Loading; linking and debugging
- Good programming practices

---

<b>BI702: Proteomics</b>	<b>I Semester</b>	<b>3 Credits</b>
--------------------------	-------------------	------------------

---

Proteins to proteomes: Amino acids structure and general properties, types of structures of proteins, Protein classification, alpha-helix, beta-strands, loops and coils, Evolution of protein structures, protein stability and folding, Ramachandran plot, hydrophobicity and its applications.

Expression Profiling: Expression systems, Importance of transcription and transcriptomics, Microarray analysis; Two dimensional PAGE for proteome analysis

Protein interactions: Yeast two-hybrid system, protein-protein interaction, protein-DNA interaction, Protein chips technology, post-translational modifications

Introductory thermodynamics and kinetics: Handling and analyzing kinetic data and free energy: Graphical analysis of single and double substrate reactions, Enzyme inhibition and rate studies related to inhibition, allosteric enzymes, mutational analysis, subunit interactions and regulation

Sequence and structure determination: Mass Spectrometry, Protein digestion, Sample preparation, sample ionization, Mass analysis, types of mass spectrometers, Peptide fragmentation, approach to mass spectrometry, Tandem mass spectrometry and SALSA, Peptide mass fingerprinting and protein identification, Database Utilization, de novo peptide sequence information, amino acid sequence database searching, un-interpreted MS/MS data searching Crystal growth and X-ray Crystallography, NMR

Computational Proteomics: Methods of structure prediction (known folds and unknown folds), viewing protein structure, protein structure classification databases, Prediction of protein functions

---

<b>BI703: Basic Mathematics and Statistics</b>	<b>I Semester</b>	<b>4 Credits</b>
--	-------------------	------------------

---

### Mathematics

#### Unit -1: Calculus

- 1) Relations, functions, Examples, Graphs; Elementary functions: Polynomials, Periodic Functions- Trigonometric Functions; Exponential and Logarithmic Functions, Hyperbolic Functions; Linear, Convex and Concave Functions and Inverse functions.
- 2) Differentiation and Integration:
  - a) Limits, Derivative as Rate of Change; differentiation of Elementary Functions, Applications of derivative and Maxima, Minima problems.
  - b) Integration as antiderivative, integration of simple functions application of integration to areas etc.

#### Unit -2: Algebra and Geometry

- a) Algebraic Equations – Exact solutions in Lower degree and Numerical Solutions by Iteration Methods; Bisection, False Position, Newton-Raphson, Secant Method and Method of Successive Approximations, Prediction- Corrector Method.
- b) Systems of Linear Equations - Matrices, Operations on Matrices, Inverse of a Square Matrix, Determinant, Cramer's Method of Solutions; Eigenvector and Eigenvalues of as square matrix. Solutions by Iteration Methods: gauss Elimination Method and Gauss-Seidel Method.
- c)
  - i) Cartesian and Polar Coordinates – Lines and Conics in plane.
  - ii) Spherical and Cylindrical Coordinates in Space –Lines and Plans in Space; Curves and Surfaces in Space, Examples.
  - iii) Symbolic Logic and Boolean Algebras.

#### Unit -3: Differential Equations.

- i) Standard forms and Solutions; Systems of equations –Exact Solutions.
- ii) Numerical Solutions: Euler's Method, Taylor series Method, Runge-Kutta Methods with Examples.
- iii) Super position of Periodic Function, Fourier series and Transforms with Applications. And Lagrange Interpolation and Truncation of errors.

### **Statistics**

- Introduction to the concept of population, sample and random sampling, intuition behind laws of large numbers and advantage of sufficiently large sample

## M. Tech Bioinformatics

### Draft course from 2011-2012

- Preliminary handling of data: Graphical techniques, Measures of location and dispersion, coefficient of variation, quantiles and percentiles, outliers and robust measures, handling the grouped data
- Elementary Probability Theory: Basic notions of set theory Permutations and combinations, probability of an event and properties of probability Conditional probability, independent events, Bayes rule
- Random variables and probability distributions: Discrete and continuous random variables, probability density/mass functions and distribution functions, mean variance etc. Discrete distributions: Bernoulli, Binomial, Poisson. Continuous distributions: Normal distribution, sigma limits and probability coverage, importance of Normal distribution. (brief intro to central limit theorem). Exponential distribution
- Estimation: Relation between population and sample quantities, Methods of estimation: maximum likelihood, naive Bayes, Illustration using Normal sample. Introduction to interval estimation
- Concepts of hypothesis testing and p-value: Illustration using one and two sample tests and chi2 test for goodness of fit., Non parametric tests, Normal probability plot.
- Regression and Correlation: Simple linear regression, least squares fit, estimation and prediction.
- Analysis of Variance: One and two way analysis of variance.
- Introduction to designs of experiments, Multivariate statistical analysis.
- Introduction to random vectors and distance between vectors.
- Cluster analysis and pattern recognition: Dimension reduction through elementary techniques
- Introduction to Markov chains
- Linkage and association studies

#### Books:

1. E.Batschelet: Maths for Life Scientists, Narosa, 1975.
2. M.K.Jain and S.R.K. Iyenger: Numerical Methods for Scientists and Engineers.
3. **Rosner, B. (1982) Foundations of Biostatistics. Duxbury Press, Boston**
4. **Daniel, W. W. (1991) Biostatistics: A foundation for analysis in health sciences.) John Wiley, New York**
5. **Pollard, J. H. (1977) A handbook of Numerical and statistical techniques with examples mainly from the life sciences. Cambridge University, Cambridge**

Note: The emphasis will be on concepts and problems solving rather than rigor.

---

<b>BI704 Molecular Modeling and Chemoinformatics</b>	<b>Ist Semester</b>	<b>4 Credits</b>
--	---------------------	------------------

---

- Nomenclature, terms and atom-numbering used in molecular modeling
- Model-building and calculations: method of coordinate generation using vector algebra, fourth atom and tetrahedral fixing approaches, internal, fractional and orthogonal coordinates, Molecular geometry and conformation, bond length, angle and torsion angles, their calculations, improper and proper dihedral angles Steric criteria, Hard and soft spheres model, atomic, ionic, covalent and van der Waal radii
- Molecular representation and visualization: Different representations their use and limitations, color and labeling scheme, wire-Frame, stick, ball-and-stick models, ORTEP & polyhedral diagrams, Ribbon and cylindrical representations, molecular Surfaces - van der Waals (and dot) surfaces, Connolly and solvent accessible surfaces, calculation of molecular volumes and surfaces.
- Data representation and formats: Data formats for 2D, SMILES/InChI notations, matrix representation, connection tables, 3D data formats, PDB, CIF, RES, Mol2, SDF formats, Z-matrix, topology and parameter files, database redundancy
- Stereochemistry: Chemical diagrams and projections, rules for Fischer projections, absolute and relative configuration, Cahn-Ingold-Prelog (CIP) rules, R/S, E/Z and cys/trans nomenclature, CORN-rule, enantiomer and chirality, chirality of drugs and its consequences, diastereomers, meso compounds, stereogenic unit and stereogenic center, isomers and isomerism, Conformation analysis, puckering of alicyclic rings, group theory

## M. Tech Bioinformatics

### Draft course from 2011-2012

- Conformational and geometric parameters of protein and nucleic acid, hierarchy of protein structures, different types of helices, sheets and turns, dictionary of protein structures, , packing of helices and sheets, comparison of structures and mathematics of superimposition, RMSD calculations, Ramachandran map, different types of Ramachandran map an region in the map, use in conformation analysis, structure validation and consistency check, Side chain conformation, different rotamer libraries.
- Molecular graphics: basics of computer display systems, drawing shapes on screen, raster graphics, bitmap vs vector graphics, Color models – RGB/CMYK, mathematics of displaying still 3D molecules on to 2D screen, geometric, rigid-body and affine transformations, transformations involved in displaying objects – modeling, projection and view port transformations, mathematics of perspective projections, adding 3D effects to molecular models – shading, depth cuing, rendering, lighting etc., animating objects, stereoviewing, shutter glass and other technologies
- Techniques of three-dimensional structure determination and predictions: Small-and macromolecular crystallography, structure solution, phase problem and direct methods, model building and refinement methods, discussion on SHELX, temperature and occupancy factors, data and model quality –R-factor and resolution, Interpretation and use of X-ray crystallographic and NMR structures, protein structure predictions - homology modeling, fold recognition and *ab initio* structure prediction methods, structure consistency and validation tools; Procheck, WhatIF and verify3D.

#### Recommended texts

- (i) Chemoinformatics by J. Gasteiger and T. Engel, Wiley VCH
- (ii) Molecular Modeling and Simulation by Tamar Schlick, Springer
- (iii) Molecular Modelling: Principles and Applications by Andrew R. Leach, Pearson Education
- (iv) Stereochemistry Of Organic Compounds by Ernest L. Eliel, Wiley
- (v) Structural Bioinformatics by P. E. Bourne and H. Weissig, Wiley-Liss
- (vi) Crystal Structure Analysis: A Primer by J. P. Glusker & K. N. Trueblood, Oxford University Press
- (vii) Introduction to protein structure by Carl Branden & John Tooze, Garland Science
- (viii) Molecular modeling: basic principles and applications by H.-D. Höltje et al., Wiley-VCH.

---

**BI705 Bioinformatics**

**I Semester**

**3 Credits**

---

**Data sets:** Databases, DBMS, Biological databases- Primary and secondary.

**Pair-wise sequence comparison by DOTMatrix approach; Pair-wise sequence alignment by** Dynamic programming ; Global (Needleman and Wunsch) and Local alignments (Smith & Waterman); Measures of similarity (Alignment score, % sequence identity, % similarity, Statistical scores E, P and Z scores); Heuristic approaches for pair-wise sequence alignments; BLAST and FastA methods; Karlin-Altschul Statistics (Estreme Value Distribution; HSPs etc); Gap penalties (Linear and Affine gap models)

**Pair-wise substitution scoring matrices; PAM, BLOSSUM etc**

Multiple sequence alignment; comparison to pair-wise method; SP scoring; Muulti-dimensional Dynamic Programming and its limitation; Progressive sequence alignment approach

Identification of patterns from MSA; PROSITE; regular expressions

PSI-BLAST; PSSM; sensitivity and Specificity; Receiver operating characteristics curves

Sequence profile calculation from MSA; profile based MSA

Structures: Protein secondary structure prediction - DNA secondary structures

**Lab courses BI706, BI707, BI708, BI709 will be based on the theory.**

## Semester II

---

<b>BI711 Genomics</b>	<b>II Semester</b>	<b>3 Credits</b>
-----------------------	--------------------	------------------

---

Pre-requisite: Any Molecular Biology or Genetics Courses.

Introduction to Genomics: Structure and organization of prokaryotic and eukaryotic genomes, nuclear and organelle genome, mitochondrial and chloroplast Genomes, microbial genome, gene density, C-value paradox, intron-exon boundary, intergenic DNA, telomere and centromere, micro and mini repeats (satellites); selfish DNA; transposons

Genome mapping: Genetic markers, physical mapping (RFLP, AFLP, SNP etc.), BACs and YACs, genome sequencing, contigs, genome sequencing projects,

Functional genomics: concepts and application, genome mining, genome synteny, Gene annotation, finding genes and regulatory regions, Prediction of gene function, Annotations by comparison of sequence, phenotype, secondary structure and large scale mutagenesis.

DNA Microarray technology: Basic principles and design, cDNA and oligonucleotide arrays, Comparative transcriptomics, Differential gene expression, Genotyping/SNP detection, Computational analysis of microarray data, Clustering gene expression profiles, Expressed sequence tags.

Books recommended:

Principles of Gene Manipulation and Genomics Third Edition By Richard M. Twyman, Sandy Blackadder Primrose Blackwell Scientific Publications,

---

<b>BI712 Mathematical Modeling of Biological Systems</b>	<b>II Semester</b>	<b>3 Credits</b>
--	--------------------	------------------

---

Pre-requisite: Bioinformatics.

- Introduction to Systems Biology
- Types of data used in modeling
- Types of models
  - gene regulatory, metabolic, Signal pathway, disease, population
- Types of modeling frameworks
  - Deterministic vs Stochastic
  - Static vs. Dynamics
- Robustness and Stability of systems
- Methods used in systems biology
  - Ordinary differential equations (ODE)
    - Linear ODEs; Non-linear ODEs, steady states, phase plane analysis

## M. Tech Bioinformatics

### Draft course from 2011-2012

- Stability analysis - Linear systems; Non-linear systems; Phase plane analysis; Stable and unstable limit cycles; Oscillations; Bistability; Positive and negative feedback
- Parameter estimation and validation – Data sources, Regression techniques (maximum likelihood, least squares methods), optimization algorithms
- Michaelis-Menten and Hill functions
- Stochastic modeling and Simulation
  - Introduction; Chemical master equation; Gillespie algorithm; Stochastic ODEs
- Logical modeling
  - Basic introduction; Logic gates; Graph construction; Boolean networks
- Metabolic control analysis (MCA)
  - Basic introduction; Control and elasticity coefficients; Summation theorems, connectivity relations theorems
- Biochemical Systems Theory (BST)
  - Basic introduction; Power law representation; S-Systems and General Mass Action; Comparison of MCA and BST modeling approaches
- Flux Balance Analysis (FBA)
  - Basic introduction; Linear Programming; Constraints
- Neural Modeling
  - Introduction, Hodgkin-Huxley model, Markov models, Action potentials, Voltage-activated Ion channels, Nernst Equation, Electrical Properties of Neurons
- Tools and databases (1)
  - SBML; Modeling tools- Gepasi, Virtual cell, Cell Designer, GENESIS, MATLAB toolboxes; Model database- JWS etc.

---

**BI713: Machine Learning Methods and Biological Network****II Semester****3 Credits**

---

- Introduction to Machine learning approaches: Neural Networks: ; Hidden Markov Model; Support Vector Machines, Clustering algorithms of kNN, kMeans clustering etc. along with current topics on machine learning
- Introduction to Genetic Algorithms
  - Graph Theory; Cliques; Network motifs
- Protein structure prediction; protein secondary structure prediction; -Protein fold-recognition; 3D modeling; Model validation; VERIFY 3D and PROCHECK
- Structural classification of proteins: SCOP and CATH
- Protein biochips and micro arrays
- Metabolic pathway modeling
- Protein interaction maps
- Molecular Phylogenetics: Tree of Life; Molecular clock hypothesis; Rooted and unrooted trees; Nodes, branches, Topology of a tree, Methods: Maximum parsimony; Distance-based (UPGMA and NJ) and Maximum likelihood; Boot-strapping method.
  
- Gene identification; genome annotation; Gene Ontology



<b>B1509/724</b>	<b>Molecular Simulation and Drug Design</b>	<b>II Semester</b>	<b>4 Credits</b>
------------------	---	--------------------	------------------

- Molecular Mechanics: empirical potential functions; ball-and-spring model, harmonic approximation, bond length, angle, torsional, out-of-Plane and cross terms, popular, force field parameterization and various approaches for the problem, validation of force field, force fields and programs - MM, AMBER, CHARMM, OPLS, GROMOS, GROMACS, CVFF, Deriding and Universal force field, other methods of molecular energy calculations - ab initio, semi-empirical, density functional analysis
- Energy minimization: Optimization methods, gradient (derivative methods), Steepest descent, Conjugate gradient and Newton methods, constraint minimization, SHAKE and Tethering algorithms, Criteria for truncating minimization
- Molecular Dynamics: Steps in typical MD simulations, minimization, equilibration and data collection, velocity scaling, periodic boundary condition, Numerical integrators verlet algorithm, analysis of trajectories, water and membrane models for simulation, Monte carlo (MC) methods, conformation search procedure, protein folding problem, Anfinsen paradigm, folding pathway, Levinthal paradox and Folding Funnel
- Electrostatics in chemistry & biology: concept of electric field, multipole expansion of a charge distribution, permanent and induced dipole and their interactions, peptide dipole, partial charge, dielectric constant, continuum and discrete solvation model, intermolecular interactions, hydrogen bonds, weak intermolecular interactions, pi-pi interactions, geometry criteria
- Drug – discovery and development – target identification and validation, lead optimization, biochemical testing and clinical trials, structure based and ligand based approaches, molecular docking, scoring functions, introduction to Gold, autodock and Surflexdoc, virtual screening, structure similarity search, , concept of pharmacophore, *de novo* ligand design, Quantitative, structure an activity relationship, 2D/3D/4D/5D QSAR, Calculation of low energy pathways for reactions – structure refinement – computation of properties, Reactions of drugs and approaches for toxicity predictions
- Free energy methods, concept of free energy, perturbation method and thermodynamics integration, application of free energy methods for solvation and ligand binding affinity, QM/MM-FEP.
- Drug design case studies

### Recommended texts

- (ix) Computational Chemistry by Errol G. Lewars, Kluwer Academic Publishers
- (x) Molecular Mechanics across Chemistry by A. K. Rappé & C. J. Casewit, University Science Book
- (xi) Computational Chemistry by David C. Young, Wiley Inter Science
- (xii) Introduction to Computational Chemistry by Frank Jensen, John Wiley and Sons
- (xiii) Numerical Optimization by J. Nocedal and S. J. Wright, Springer
- (xiv) Computer simulation of liquids by M. P. Allen & D. J. Tildesley, Oxford Science Publications
- (xv) Encyclopedia of Computational Chemistry by Paul von Ragué Schleyer, John Wiley & Sons.
- (xvi) Chemoinformatics by J. Gasteiger & T. Engel, Wiley VCH
- (xvii) Drug Design: Structure- and Ligand-Based Approaches, Edited by Kenneth M. Merz, Jr., Cambridge University Press
- (xviii) Introduction to Electrodynamics by David J. Griffiths, Benjamin Cummings.

---

**BI715: Metagenomics**

II Semester

3 Credits

---

- Introduction to metagenomics
- Experimental methods used in metagenomics
- High throughput sequencing technologies
- Metatranscriptomics and Metaproteomics
- Bioprospecting approaches
- Computational challenges in metagenomics
- Estimation of taxonomic diversity
  - Marker gene based
  - Binning methods
- Gene prediction
- Functional characterization of metagenomic samples
- Community analysis
- Comparative metagenomics
- Metagenomics applications with case studies
  - Health, agriculture, environmental remediation, industrial biotechnology, bioprospecting

**Lab courses BI716, BI717, BI718, 719 will be based on the theory.**

**BI720 is seminar on the topic related to Bioinformatics**

Two electives from the following should be chosen. At least ten students should opt for an elective.

### Electives

---

<b>BI507 Principles of Gene Regulation</b>	<b>III Semester</b>	<b>2 Credits</b>
--	---------------------	------------------

---

Prokaryotic transcription, Prokaryotic gene regulation: lac Prokaryotic gene regulation:  $\lambda$ , Eukaryotic transcription: basal level & TFs activation, chromatin Negative reg. Epigenetic controls; Other aspects of gene regulation RNA processing, alternate splicing, Translational control, RNA stability RNA mediated gene regulation, Bioinformatics approaches to gene regulation.

---

<b>BI507 JAVA programming for bioinformatics applications</b>	<b>III Semester</b>	<b>2 Credits</b>
---	---------------------	------------------

---

Introduction to Java - Features of Java - Object oriented concepts - Data types - Variables - Arrays - Operators - Control statements

BASICSOFA JAVA - Classes - Objects - Constructors - Overloading method - Access control - Static and final methods - Inner Classes - Inheritance - Overriding methods - Using super abstract class-- String class- String objects - String buffer - Char Array

PACKAGES,INTERFACES&THREADS- Packages - Access protection - Importing packages - Interfaces - Exception handling - Throw and throws - Thread - Synchronization - Messaging - Runnable interfaces - Inter thread communication - Deadlock - Suspending, Resuming and stopping threads - Multithreading

APPLETS- Java Utilities -Applets - working with windows using AWT Classes - AWT Controls - Layout Managers and menus

INPUT/OUTPUT&NETWORKING - I/O streams - File streams -Networks basics - Socket programming - Proxy servers - TCP/IP sockets - Net address - URL - Datagrams

Bio Java- Announcing - Presentations - Publications - How Do I....? - Setup - Alphabets and Symbols Basic Sequence Manipulation - Translation - Proteomics - Sequence I/O - Annotations - Locations and Features - Counts and Distributions-Weight Matrices and Dynamic Programming- BioSQL and Sequence Databases- Genetic Algorithms

#### TEXT BOOKS

- Naughton and H.Schildt - "Java 2 - The complete reference" - Fourth edition.-2002.

#### REFERENCE BOOKS

- S.Horstmann, Gary Cornell - "Core Java 2 Volume I - Fundamentals" - Addison Wesley – 2001
- Arnold and J.Gosling - "The java programming language" - Second edition
- Art Gittleman – “Ultimate Java Programming” –Wiley Publications-2002

Link: <http://www.biojava.org/wiki/BioJava:CookBook>

---

BI507 Databases	III Semester	2 Credits
-----------------	--------------	-----------

---

- Database systems,                      • Entity relation model, • Relational model
- Relational database design, • Concurrency control
- Biological databases: GenBank-protein sequence database-EMBL Data Library- Brooke Heaven Database-Cambridge database
- Submission of database,            • Restructing of database
- Data acquisition, data distribution
- Compatibility of formats of data from different databases

---

BI507 Biological networks	III Semester	2 Credits
---------------------------	--------------	-----------

---

1. Network concepts and data- Basic graph theoretical concepts – Node degree, degree distribution and correlations –diameter-Clustering coefficient– Eccentricity- Motifs-survey of datasets (social, economic, transportation, computer, biological networks)-network types and representations (directed, weighted, bipartite). - types of network structure (regular, small world, scale-free, random)
2. Exploratory analysis - adjacency matrix; Laplacian matrix; distance matrix- small world property – scale free distribution – some examples- structural similarity; assortative mixing-visualizing networks- sub networks
3. Centrality measures- Degree; closeness; information; betweenness an. d Eigen vector centrality. - Other centrality measures - Its applications in natural sciences and social sciences
4. Algorithms for networks- Shortest paths algorithms - Eigen values and Eigen vectors. - Understanding community structures – some examples- Fractal dimension – some examples
5. Modeling network structure - Regular graphs-random graphs; networks with arbitrary degree distribution- small world networks; scalefree networks-network growth; densification; preferential attachment – network identification (which generative model fits better).
6. Biological Networks- protein contact networks- protein-protein interaction networks- metabolic networks- gene regulatory networks
7. Additional topics- Trees and optimization problems- Matchings - Connectivity and network flow - Colorings- Planarity – Edges and cycles