M.TECH.
COMPUTER SCIENCE AND TECHNOLOGY WITH SPECIALIZATION IN BIO-INFORMATICS

WITH EFFECT FROM 2006-07 ADMITTED BATCH
(Tentative)

3rd Semester SYLLABI

CHAIRMAN
BOARD OF STUDIES
boschair.csse@aucevizag.ac.in

DEPARTMENT OF COMPUTER SCIENCE AND SYSTEMS ENGINEERING
COLLEGE OF ENGINEERING
ANDHRA UNIVERSITY
VISAKHAPATNAM-3
www.aucevizag.ac.in, www.andhrauniversity.info

M.TECH.
COMPUTER SCIENCE AND TECHNOLOGY WITH SPECIALIZATION IN BIO-INFORMATICS

COURSE STRUCTURE AND SCHEME OF EXAMINATION W.E.F 2006-07 ADMITTED BATCH

3rd SEMESTER

CODE	NAME OF THE SUBJECT	PERIODS/WEEK		EVALUATION MAXIMUM MARKS		
		Theory	Lab.	Internal	External	Total
MTCSTBI 3.1	Seminar on Advanced Topics	**3**	-	**100**	-	**100**
MTCSTBI 3.2	Software Engineering for Bio-Informatics	**3**	-	**50**	**100**	**150**
MTCSTBI 3.3	Elective 1.Algorithms on Strings, Trees and Sequences 2. Machine Learning for Bio-Informatics 3.Computer Vision	**3**	-	**50**	**100**	**150**
MTCSTBI 3.4	Bio-Informatics Lab	-	**3**	**50**	**100**	**150**
MTCSTBI 3.5	Web Programming Lab. (Common With M.TECH (CST))	-	**3**	**50**	**100**	**150**
			Total Marks	**300**	**400**	**700**

MTCSTBI 3.1 SEMINAR ON ADVANCED TOPICS

Instruction: 4 periods/week
Internal Assessment: 100 Marks

Purpose: To enable a student to be familiar with Communication skills

I. Student is expected to learn
   a. How to make a presentation
      i. Verbal
      ii. Non Verbal
      iii. LCD based Power Point

   b. How to write a report
      i. Abstract
      ii. Body
      iii. Conclusions
      iv. Executive Summary

   c. Group Discussion
      i. Share the work with a group
      ii. Modularization of the work
      iii. Shareware Development

   d. Communication
      i. Horizontal
      ii. Vertical

II. Students will be given a topic of importance and are expected

   a. To present the topic verbally in 30 minutes
   b. To present the topic as a report in 50 pages

MTCSTBI 3.2 **SOFTWARE ENGINEERING FOR BIOINFORMATICS**

Instruction: 3 periods/week **External Assessment:** 100Marks
Internal Assessment: 50 Marks **Time:** 3 Hours

1. Software Engineering and Bioinformatics
3. Project Definition.
4. Requirements Capture
5. Separating Function, Interface and Implementation.
6. Implementation Considerations.
7. Proof of Concept and Prototyping
8. Data in, Data out and Data Transformation.
9. Functional, then Optimized.
10. Coding Style.
11. Testing.
12. Rollout and Delivery.
13. Support and Feedback.
14. Planned and Unplanned Enhancements.

Text Book: **Paul Weston, Bio-Informatics Software Engineering, John Wiley & Sons,2004**

MTCSTBI 3.3: ELECTIVE - ALGORITHMS ON STRINGS, TREES AND SEQUENCES

Instruction: 3 periods/week
Internal Assessment: 50 Marks
External Assessment: 100Marks
Time: 3 Hours

PART I. EXACT STRING MATCHING:

1. The Fundamental String Problem:
2. Exact matching: fundamental preprocessing and first algorithms;
3. Exact matching: classical comparison-based methods;
4. Exact matching: a deeper look at classical methods;
5. Semi-numerical string matching;

PART II. SUFFIX TREES AND THEIR USES:

6. Introduction to suffix trees;
7. Linear time construction of suffix trees;
8. First applications of suffix trees;
9. Constant time lowest common ancestor retrieval;
10. More applications of suffix trees;

PART III. INEXACT MATCHING, SEQUENCE ALIGNMENT AND DYNAMIC PROGRAMMING:

11. The importance of (sub)sequence comparison in molecular biology;
12. Core string edits, alignments and dynamic programming;
13. Refining core string edits and alignments;
14. Extending the core problems;
15. Multiple string comparison: the Holy Grail;
16. Sequence database and their uses: the motherlode;

PART IV. CURRENTS, COUSINS AND CAMEOS:

17. Maps, mapping, sequencing and superstrings;
18. Strings and evolutionary trees;
19. Models of genome-level mutations.

Text Book: Dan Gusfield, Algorithms on strings, trees and sequences, Cambridge University Press, 1997

MTCSTBI 3.3: ELECTIVE - MACHINE LEARNING FOR BIOINFORMATICS

Instruction: 3 periods/week
Internal Assessment: 50 Marks
External Assessment: 100Marks
Time: 3 Hours

1. Introduction
Biological data in digital symbol sequences,Proteins & proteomics
Prediction of molecular function & structure

2. Machine-Learning Foundations: The Probabilistic Framework
Introduction: Bayesian modeling ,Bayesian inference & induction
Model structures: graphical models & other tricks

3. Probabilistic Modeling & Inference: Examples
The simplest sequence models ,Statistical mechanics

4. Machine Learning Algorithms
Dynamic programming ,Gradient descent,EM/GEM algorithms
Markov-chain Monte-Carlo methods,Simulated annealing
Evolutionary & genetic algorithms, learning algorithms: miscellaneous aspects

5. Neural Networks: The Theory
Universal approximation properties ,Priors & likelihoods
Learning algorithms: back propagation

6. Neural Networks: Applications
Sequence encoding & output interpretation,Sequence correlations & neural networks
Prediction of protein secondary structure ,Prediction of signal peptides & their cleavage sites
Applications for DNA & RNA nucleotide sequences,Prediction performance evaluation
Different performance measures

7. Hidden Markov Models: The Theory
Prior information & initialization,Likelihood & basic algorithms
Learning algorithms,Applications of HMMs: general aspects

8. Hidden Markov Models: Applications
Protein applications ,DNA & RNA applications
Advantages & limitations of HMMs

Text Book: Pierre Baldi and Søren Brunak, Bioinformatics: The Machine Learning Approach
MIT Press, February 1998

MTCSTBI 3.3: **ELECTIVE - COMPUTER VISION**

Instruction: **3 periods/week** **External Assessment:** **100Marks**
Internal Assessment: 50 Marks **Time:** **3 Hours**

IMAGE FORMATION and IMAGE MODELS:

Cameras
Geometric Camera Models
Color

ONE IMAGE and MULTIPLE IMAGES:

Linear Filters
Edge Detection
Texture
Geometry of Multiple Views
Segmentation by Clustering,

HIGH-LEVEL VISION:

Model-Based Vision
Finding templates using Classifiers

APPLICATIONS:

Finding In Digital Libraries
Image-based Rendering

Text Book:

David A.Forsyth and Jean Ponce, Computer Vision, A Modern Approach, PHI, 2003

MTCSTBI 3.4 BIOINFORMATICS LAB

Instruction:	**3 periods/week**		
Internal Assessment:	**50 Marks**	**External Assessment:**	**100**
		Time:	**3 H**

Students are expected to use 'Analysis Packages' available as 'Open Source' and develop Analysis' pertaining to Diabetes Mellitus.

MTCSTBI 3.5 WEB PROGRAMMING LAB
(COMMON WITH M.Tech(CST) with Specialization in Computer Engg.)

Instruction:	**3 periods/week**	**External Assessment:**	**100Marks**
Internal Assessment: 50 Marks		**Time:**	**3 Hours**

1. Design of the Web pages using various features of HTML and DHTML
2. Client server programming using servlets, ASP and JSP on the server side and java script on the client side
3. Web enabling of databases
4. Multimedia effects on web pages design using Flash.

Reference Books:

1. Internet and Web Technologies by Raj Kamal, Tata McGraw-Hill
2. Programming the World Wide Web by Robert W. Sebesta, Pearson Education