

Perl Programming
Computer Information Systems Division
Wake Technical Community College

CSC 185 (2-3-3)	Chris James, Instructor
Course Syllabus Addendum	Fall Semester, 2003
Bioinformatics Certificate	Office: LeMay #21
Phone: 662-3416 (Secretary)	Office Hours: M,T,TH 5-6 PM
Email: ctjames@waketech.edu	Office Phone: 662-3656

Enrollment Information: This is a restricted enrollment course. The course is open to individuals possessing course work in the field of biology (baccalaureate degree recommended) or related sciences who wish to pursue the Bioinformatics certificate offered through Wake Technical Community College's High Performance Computing curriculum.

The Bioinformatics certificate is aimed at individuals who:

- are working in the field of biology and require bioinformatics skills.
- are enrolled in an undergraduate biology curriculum who wish to gain knowledge in the field of bioinformatics for later employment opportunities.
- wish to obtain academic preparation for the information technology portion of the Certified Bioinformatics Specialist certification offered by the National Bioinformatics Institute.

Other students may be admitted to the course by prior arrangement with the course instructor.

Email the course instructor (ctjames@waketech.edu) for additional information and registration procedures regarding this course or for more general information regarding the Bioinformatics certificate program.

Course Duration: This course meets for the entire 16-week semester.

Class Meeting Time: This is a hybrid web-based course combining distance education with on-campus class meetings. The course contact hours are equally divided between on campus meetings and online access. The online portion of the course is accessed through a web browser based interface known as Blackboard. A Blackboard tutorial will be given on the night of the first class meeting. The on-campus meetings will be on Thursday nights from 6:00 PM to 9:00 PM. The first on campus meeting will be **August 21, 2003**.

Course Catalog Description: This course introduces students to the Perl Programming language. Topics include programming techniques using CGI script, input/output operations, sequence, iteration, selection, arithmetic operations, subroutines, modules, integrating database, pattern matching and other related topics. Upon completion, students should be able to design, code, test, and debug Perl language programs.

Course Overview: Welcome to Perl Programming! This course is a comprehensive introduction to the Perl programming language and its application to the field of bioinformatics. Bioinformatics is an interdisciplinary field that uses computer algorithms to interpret biological data, such as DNA or protein sequences or structures. The course assumes a background knowledge in genetics and biochemistry, but assumes no prior programming experience. The course is taught under the UNIX (Linux) operating system.

Course Textbook: "Beginning Perl for Bioinformatics", by James Tisdall, O'Reilly Publishers, ISBN: 0596000804. From the publisher's website: "*This book shows biologists with little or no programming experience how to use Perl, the ideal language for biological data analysis. Each chapter focuses on solving a particular problem or class of problems, so you'll finish the book with a solid understanding of Perl basics, a collection of programs for such tasks as parsing BLAST and GenBank, and the skills to tackle more advanced bioinformatics programming.*"

Course Goals: The principal goals for the course are as follows:

1. Learn to write computer programs in the Perl programming language.
2. Manage the operation of an Apache web server using a Perl script.
3. Write, install, and execute simple CGI scripts on an Apache web server.
4. Understand how to use CPAN.
5. Locate reference materials and other resources related to Perl.
6. Use Perl to write computer programs to solve problems related to the life sciences.
7. Create a biological database accessed through a web interface.

Course Objectives: Upon successful completion of the course, students will be able to:

1. Learn the techniques and logic of transforming ideas into computer programs.
2. Use the principles of top-down design, stepwise refinement, pseudo-code, and structured methodology in developing efficient and functionally correct Perl algorithms.
3. Recognize and use the control structures, variable types, and statements used in defining the Perl programming language to solve the programming problem.
4. Develop appropriate documentation techniques so that programs are easy to read, easy to modify, and easy to check for correctness.

Grade Evaluation: Three tests (20 percent each), laboratory assignments (20 percent), plus a final project (20 percent). Ten-point grading scale.

Testing: All tests will be closed-book take home tests. The format of the tests will vary, but expect multiple choice, matching, fill in the blank type, code snippets, and short answer type questions.

Course Outline:

Part I. Perl Basics

The first part of the course introduces the fundamentals of the Perl programming language. Perl has become one of the most popular programming languages for biological data analysis because it makes the writing of biologically useful programs relatively straightforward. You will learn basic Perl program structure, syntax, data types, flow control and file handling. Lab assignments will focus on creating functional Perl programs of a bioinformatic nature.

Text chapters:

1. Biology and Computer Science
2. Getting Started with Perl
3. The Art of Programming
4. Sequences and Strings
5. Motifs and Loops
6. Subroutines and Bugs

Part II. Advanced Perl Features

This section of the course will examine the use of Perl for operating system interaction and the use of modules in Perl. Popular Perl modules such as BioPerl and DBI (Database Connectivity) will be described.

Text chapters: None. Lecture notes will be provided.

Part III. Biological Applications of Perl

This section will introduce some specific examples of Perl applied to bioinformatics as well as introduce some additional Perl programming concepts. Perl is widely used to automate repetitive tasks, work with DNA sequences and strings, find motifs in data, and parse annotations.

Text chapters:

7. Mutations and Randomization
8. The Genetic Code
9. Restriction Maps and Regular Expressions

Text chapters 10 (GenBank), 11 (Protein Data Bank), and 12 (BLAST) will be covered in HPC 270.

Part IV. Perl and the World Wide Web

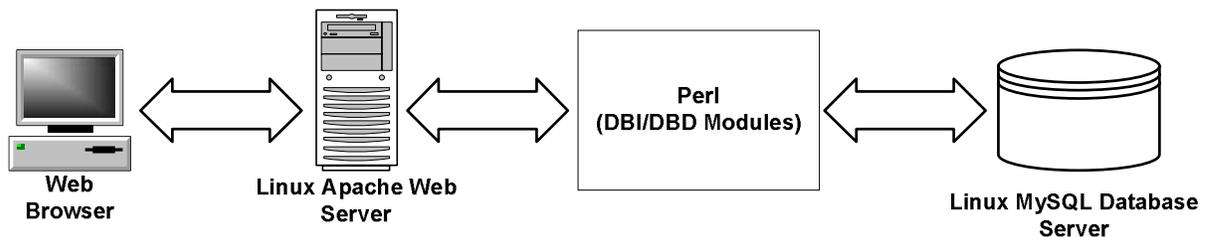
The fourth section of the course will focus on using Perl for the creation of dynamic websites on the World Wide Web using the Common Gateway Interface. CGI is an important World Wide Web technology that enables a client web browser to request data from a program executed on the Web server. CGI specifies a standard for passing data between the client and the program and is not a programming language in of itself. The lab assignments in this section will begin the creation of the capstone biological database project.

Text chapters: None. Lecture notes will be provided.

Part V. Web Database Integration with Perl

The final section of the course will involve the programming procedures required to tie a website to a backend database using Perl as a middleware component. The Perl DBI and DBD modules for interfacing with relational databases will be used. The capstone biological database project will be completed at the end of this section.

Text chapters: None. Lecture notes will be provided.



Instructional Methods: This is a hybrid web-based course combining distance education using Blackboard software with on-campus class meetings.

Course Webpage: All instructional materials for this course are available from the following location: <http://dist-ed.waketech.edu/>

Course Access: Students will be able to access the Blackboard interface to the course on the first day of the semester (**August, 19, 2003**).

Login ID: Your login name will be your first initial, middle initial and last name, plus the last two digits of your student ID. Your password will be your student ID. Additional information and trouble shooting your login ID is available here:

http://www.wake.tec.nc.us/dist_ed/internet/help.html

Getting Started: You should visit the course webpage after the first on-campus meeting.