

CSE 510

Web Data Engineering

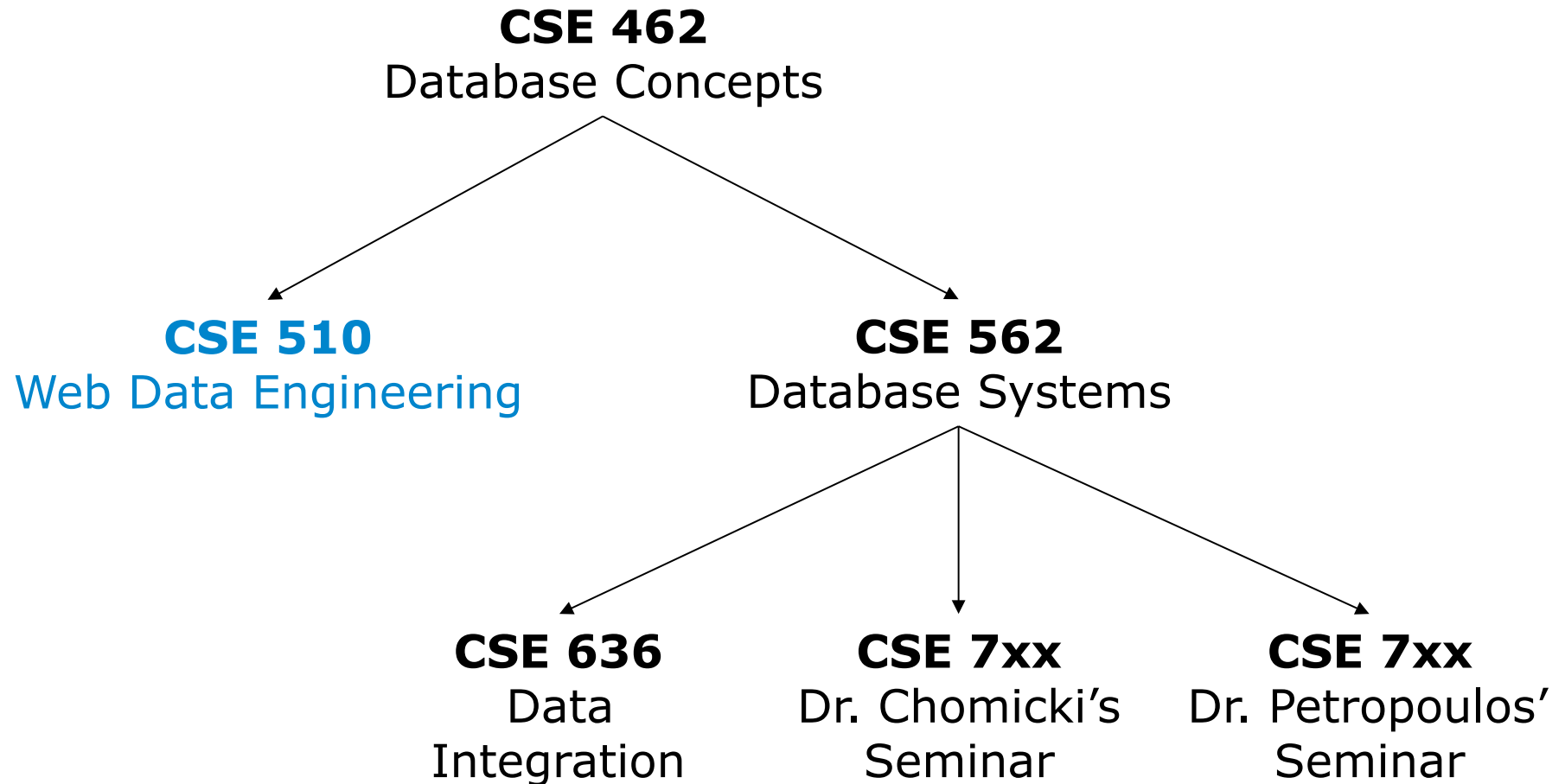
Introduction

cse@buffalo

Staff

- Instructor: Dr. Michalis Petropoulos
Office Hours: Mon & Wed @ 1-2pm
Location: 210 Bell Hall
- TA: Demian Lessa
Office Hours: Fri @ 1-3pm
Location: 329 Bell Hall
- Web Page
<http://www.cse.buffalo.edu/~mpetropo/CSE510-FA09/>
- Newsgroup
sunyab.cse.510

UB CSE Database Courses



Prerequisites

- CSE462 or equivalent undergrad database course
- Good knowledge of Java is needed for the project
- **Curiosity! You should ask a lot of questions!**

Relevant Material

Recommended Textbooks

- For introduction, servlets, JSPs and database/application server issues you may use **Tomcat Kick Start**, by Martin Bond and Debbie Law
- For Struts you may use **Struts in Action**, by Ted Husted, Cedric Dumoulin, George Fransiscus and David Winterfeldt
- A combination of class notes and online resources is probably better than textbooks since really nice online material on web programming can be found nowadays

Grade Computation

- Project: 70%
 - Teams of 2
 - Consists of three phases (at least)
- Final: 30% (in class)

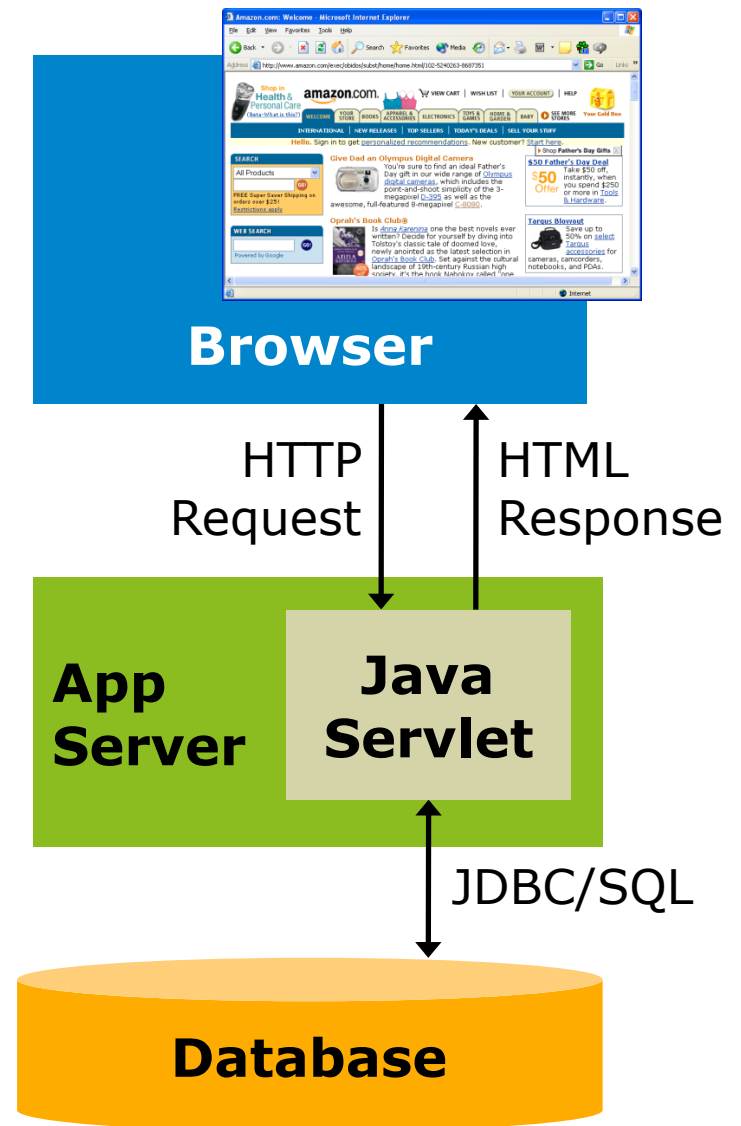
Content and Organization of the Course

Course Focus: Web Apps Providing Dynamic Content

- Web initially served static content
 - Pages are constructed in advance as html files and communicated with the http protocol
- Then most web sites served dynamic content
 - E-commerce, online banking, online auctions
 - Content typically comes from one or more databases

Course Focus: Web Apps Providing Dynamic Content

- We will learn how to build server-side applications that interact with their users and provide dynamic content
- Using the Java programming language and SQL-based databases
- Key ingredient: Application servers (Tomcat) that support Java-based server-side programs



Escalation of Java-Based Technologies for Server-Side Programming

- Discussion of network-level HTTP requests and responses
- **Java Servlets** are Java programs running inside the application server
- Servlet invoked using HTTP by client
 - App server provides HTTP request object that encodes the request information

Escalation of Java-Based Technologies for Server-Side Programming

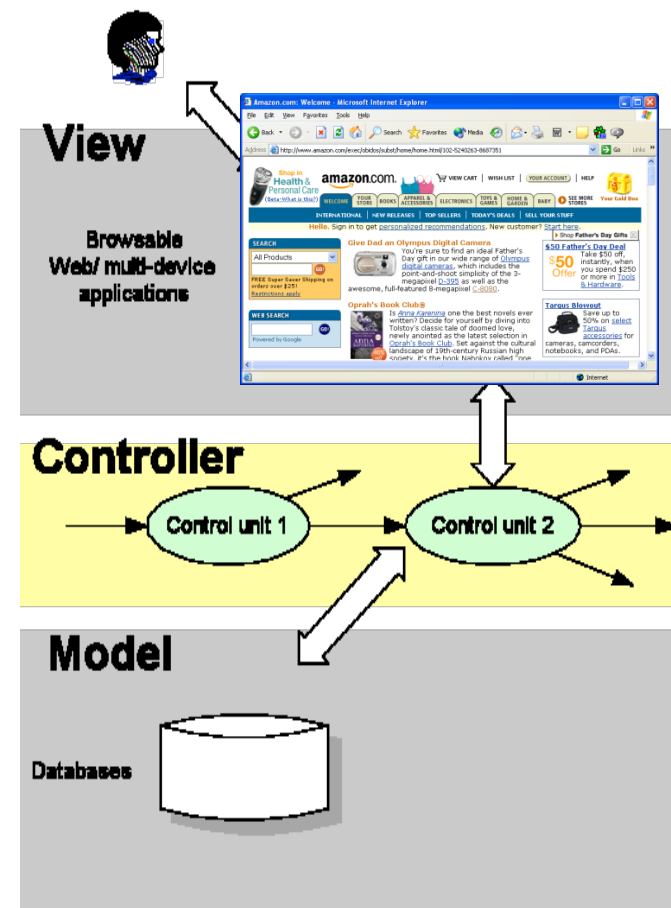
- Servlet typically (but not necessarily) returns HTML to client
 - Unfortunately, HTML response is created with many `println()` statements
 - Very hard to separate static HTML content from dynamic content and control flow
- Taught for educational purposes – nobody codes servlets directly

Next Technology: Java Server Pages (JSPs) & Java Beans

- HTML with embedded Java code
 - Easy to understand how the produced HTML looks
- Compiled into a Java Servlet
- Unfortunately, the business logic of the application (encoded in Java) is hard to understand and modify
- Java Beans provide a little remedy
 - Self-contained Java components (classes) with a bunch of restrictions

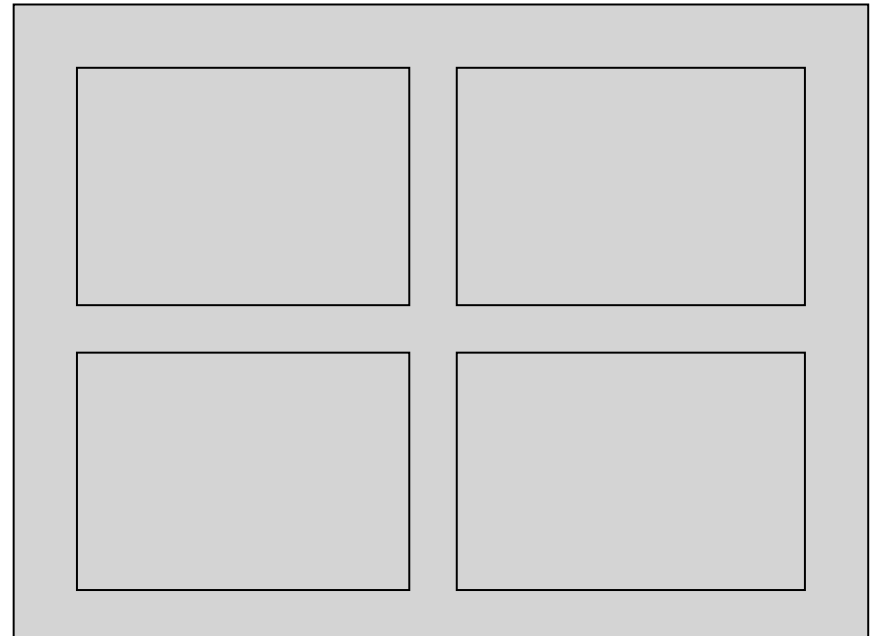
Next: Model-View-Controller (MVC) Programming, using Struts

- Emerging Development “Best Practice”
- **Model:** Access to Underlying Databases and Info Sources
- **Controller:** Control Flow of Web App
- **View:** Look-and-Feel



Next: AJAX and the Component-Based Page

- A new paradigm: Web applications providing the feel of desktop applications
- Essentially page consists of components
- Individually refresh themselves via XHR calls



Since Java, HTML and SQL Are Central To Examples & Project

- Database programming “fast track” course
 - Practical database design techniques
 - SQL programming
 - Use of JDBC in web applications
- Brief discussion of HTML of the examples

Project

- Think of the instructor and the TA as customers
- Go from our problem statement...
to a web app specification...
to a functional web application
- Project: Graduate Admissions Application
- Build using Struts framework

Many Dynamic Content Server-Side Technologies will NOT be Covered

- Common Gateway Interface (CGI)
 - Slow performance
 - No standard scripting language (Perl, PHP, ...)
- Microsoft's Active Server Pages (ASPs)
 - Very similar in principle to JSPs
 - Runs on Windows platforms only
- Other MVC frameworks
 - Spring, PureMVC
- AJAX architecture will be covered

Java Servlets vs. Java Applets

- Servlet runs on web application server
- Can access the (server-side) database and other resources
- Can only return data to browser
 - Interaction is based on user making HTTP requests and the servlet returning an HTML page
- Applet is downloaded on web client
- Accesses client-side resources
 - Due to security reasons resources are typically unavailable
- Better in some cases for interaction with user

Application Servers: The Essential Tool of Server-Side Programming

- Java servlet containers, responsible for:
 - facilitating HTTP communication
 - providing web application context
 - ...
- May also (but not necessarily) operate as web servers, that is, serve static pages
- Tomcat is an app server and the reference implementation of the Java Servlet and JSP specifications
 - Also serves static pages
 - The statement “Tomcat is a Web server” is not accurate