Towards a Syllabus Repository for Computer Science Courses

Manas Tungare, *Xiaoyan Yu, William Cameron, GuoFang Teng,*
Manuel Pérez Quiñones, *Lillian Cassel, Weiguo Fan, Edward Fox*

Syllabi are Information-Rich

Title, Description, Instructor, Teaching Assistants, Prerequisites, Topics, Knowledge Units, Learning Objectives, Calendar, Readings, Books, Book Chapters, Articles, Papers, Instructor's Notes, Slides, Assignments

Closed and scattered

- Assortment of formats: HTML, PDF, closed
- Non-standard publishing mechanisms:
  - Instructor's website, CMS, catalogs
- Limited access outside the university
- So we get them by searching the Web
  - Enter the right keywords ⇒ Get syllabi

A Specialized Collection

or *"If Google is the answer, what was the question?"*

- Demarcating syllabi from other random documents on the Web
- Seeing interesting contexts and patterns for searches
  - *"plagiarism"*, *"evaluation"*, *"pair programming"*, professor's name
- Difficult to limit searches to syllabi otherwise

The Adoption Problem

- Chicken-and-egg situation
- Solution: crawl and parse existing syllabi
  - Develop a schema
  - Store in a repository
  - Develop tools & applications

⇒ Encourage creation of new syllabi conforming to schema

How we crawled

1. Searched for: "*computer science site:edu*"
   ~80 universities

2. Searched for: "*syllabus site:cs.vt.edu*", etc.
   ~8000 documents

3. Classified into:
   Full / Partial / Entry Page / Noise

Syllabus CS2704: Object-Oriented Software: Fall 2002

Course information

- Title: CS2704: Object-Oriented Software Design and Construction
- Index number: 91378
- Location: Pam 2030
- Time: MWF 1:25pm-2:15pm
- Final Exam: December 13th, 4:25pm-6:25pm
- Web page: http://perez.cs.vt.edu/cs2704/
- Class listserv: cs2704_91378@listserv.vt.edu

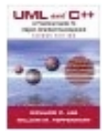
Instructor

Dr. Manuel A. Pérez-Quiñones
email: perez@cs.vt.edu
office: McBryde 621
phone: 231-2646
office hours: Wed 9:30-11:00, Fri 9:30-10:30, or by appointment

Graduate Teaching Assistant

Jesús Trespalacios
email: jtrespal@vt.edu
office: McB122
office hours: Mon 10:00-11:00 Tue 9:00-10:00 Wed 9:00-10:00, or by appointment

Textbook



UML and C++: A Practical Guide to Object-Oriented Development, by Richard C. Lee, William M. Tepfenhart, 2nd Edition. 2001, Prentice Hall, ISBN 0-13-029040-8. **Required**

The following are good reference books on UML, and C++ (in that order). They are not required for this course, just listed here as good supplements.

UML Distilled: Applying the standard object modeling language, by Martin Fowler with Kendall Scott. 1997, Addison-Wesley. ISBN 0-201-32563-2.

The C++ Programming Language by Bjarne Stroustrup, 3rd Edition. 1997, Addison-Wesley, ISBN0-201-88954-4.

Objective

Object-oriented programming concepts are studied and basic skills in software design are developed. Sound practices for design, construction, testing, and debugging of object-oriented software systems are emphasized. Object-oriented features of the C++ programming language are examined. The primary principles and language features studied are: objects, classes, inheritance, and polymorphism.

Prerequisites

CS 1704 or ECpE 2574. Computer Science majors and minors must have completed the prerequisite with a final grade of C or better. ECpE majors must have completed the prerequisite with a final grade of C- or better. We will grant no exceptions to these requirements.

Web page

http://perez.cs.vt.edu/cs2704/

Full Syllabus

Building High-Performance Servers

Home | Syllabus | Assignments | Readings

Course Description:

In depth study of modern server construction. Considers architectures for building high-performance, robust, scalable, and secure network servers. We will consider all aspects of "mission-critical" servers. Topics include multithreaded and asynchronous programming techniques, database access, performance profiling, embedded languages, and fault tolerance. Teams will build significant projects.

Schedule:

Mon, Wed, Fri: 10-11 in 345 (**note room change!**)

Instructor:

John Jannotti [web | email]
Office Hours: Mon & Thu: 11-12 in 379

Teaching Assistants:

Chris Erway [web | email] Office Hours: Thurs 12-2, 423
Ronald Tse [email] Office Hours: Wed 6-8 Fishbowl

Prerequisites:

CS32 and (CS168 or CS167)

Home | Syllabus | Assignments | Readings

Partial Syllabus

Syllabus

Syllabus.pdf

This page written an maintained by Carl Alphonce.

Entry Page

Leen-Kiat Soh

Teaching Vita

Department of Computer Science and Engineering
University of Nebraska
122E Avery Hall, Lincoln, NE 68588-0115

e-mail: lksoh@cse.unl.edu
tel: (402) 472-6738
fax: (402) 472-7767

List of Courses Taught

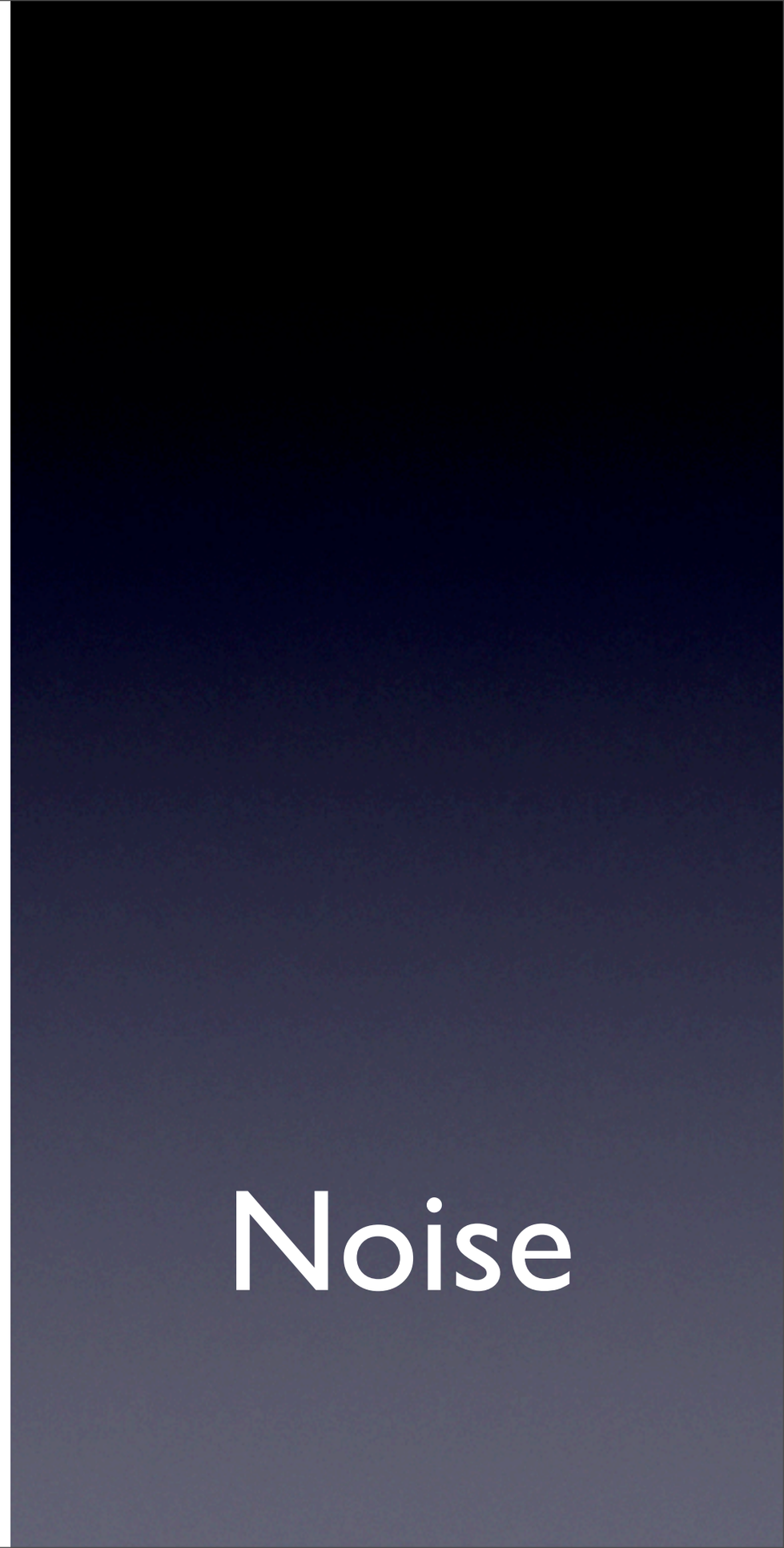
CSCE 410/810 Information Retrieval Systems (Fall 2001, Fall 2003)
CSCE 235 Introduction to Discrete Structures (Spring 2002, Spring 2003)
CSCE 496/896 Multiagent Systems (Fall 2002, Fall 2003)
CSCE 155 Introduction to Computer Science I (Fall 2004)

CSCE 996 Research other than Thesis: Decision and Utility Theories in Various Aspects of Multiagent Systems (Fall 2001) (Xin Li) (Completed)
CSCE 996 Research other than Thesis: I-CHOIR: Imagery Collaborative Hierarchical Ontology for Indexing and Retrieval (Spring 2004) (Chao Chen)
CSCE 996 Research other than Thesis: Building an Intelligent Online Survey Assistant (Spring 2004) (Jared Kite) (Completed)
CSCE 996 Research other than Thesis: GIS Applications in Drought Analysis (Fall 2003) (Songjie Wei) (Incomplete)
CSCE 891 Internship in Computer Practice (Fall 2003) (Ashok Thirunavukarras) (Completed)
CSCE 996 Research other than Thesis: Adaptive File Transfer (Spring 2004) (Eric Moss) (Completed)
CSCE 996 Research other than Thesis: Building an Intelligent Agent to Play Poker (Fall 2004) (Todd Blank)
CSCE 996 Research other than Thesis: Question Ranking, Classification, & Grouping for I-MINDS (Summer 2004) (Nobel Khandaker)

Curricular Development

Project Re-Inventing CS Curriculum Summer 2003 – Present
Joined the project in June 2003 to introduce laboratories to CS1/CS2, and help design and implement a placement examination, learning objects, and other relevant logistics for CS1/CS2. This is a very significant overhauling of introductory CS courses as we incorporate educational research, instructional design, traditional laboratories, and technology-based tools to help students learn as well as to investigate how they learn and how well they learn. Also assigned the mentorship of CS3 (CSCE 235), primary caretaker of the course for years to come.

CSCE 410/810 Introduction to Information Retrieval Fall 2001
Revamped all aspects of the class: course materials, subject topics, projects, presentations, reading lists, computer programming homework assignments. Reorganized the syllabus. Completely re-designed this class. Introduced several important, basic topics in Information Retrieval and designed four programming

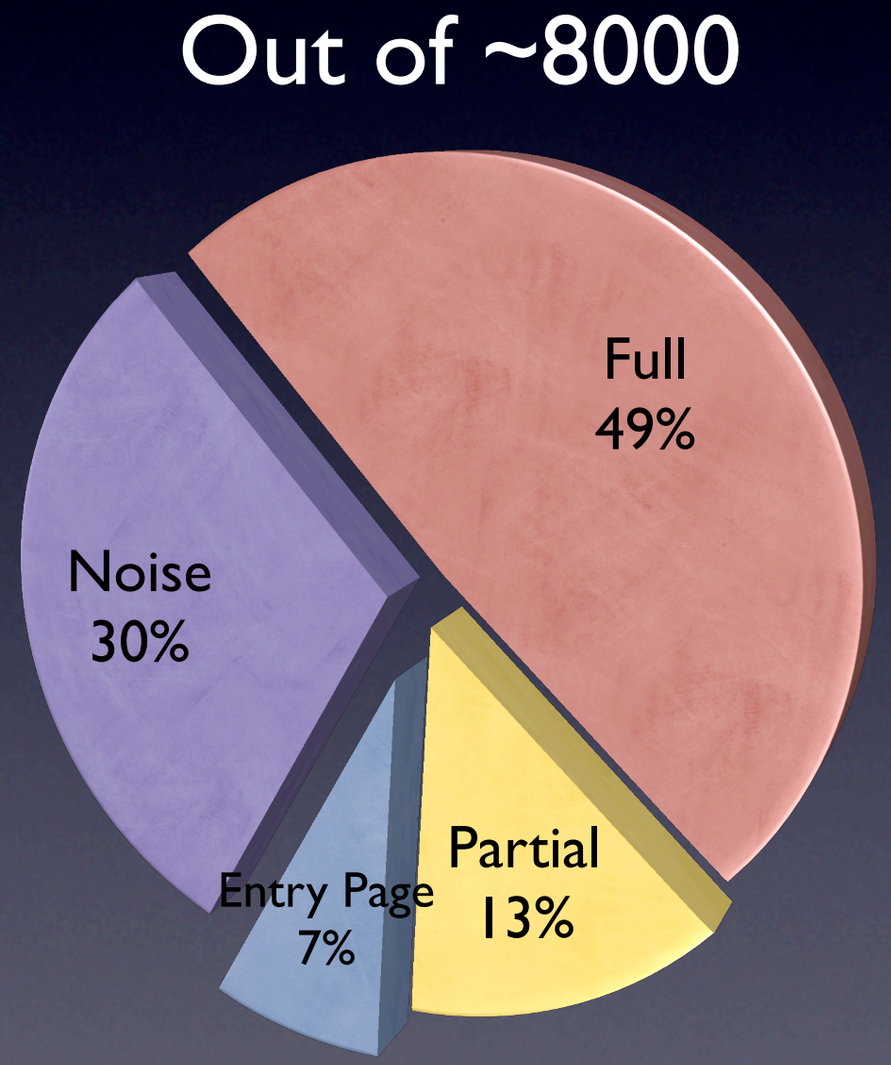


How we classified

- Training set: ~1000 documents
- Feature selection: 84 features considering the content and form of the syllabus genre
- Algorithm: Sequential Minimal Optimization with Polynomial kernel (SMO-P)
  - Found best by F-Measure test
- Automatically classified the rest ~7000

Results of Classification

	Content	Outgoing links
Full	✓	
Partial	✓	✓
Entry Page		✓
Noise		



The Syllabus Repository

- ~5600 documents
- Full cache available
- Transcoded PDF, PS & HTML to text
  - for searching
- Free-form search over all the content
- Search results show context plus other info

Other Services

- Tag using CC 2001 categories
- Browse syllabi by CC 2001 categories
  - Tag cloud
- Compare two syllabi
- Submit your syllabus URL
- View in standardized presentation format

http://doc.cs.vt.edu/syllabus/

Personal NSDL — Virginia Tech & Villanova University

- Home
- Syllabus Repository
  - Submit
  - Google Co-op
- CS Ed News
- News
- People
- Publications

Discrete Structures
Programming Fundamentals
Algorithms and Complexity
Architecture and Organization
Operating Systems
Net-Centric Computing
Programming Languages
Human-Computer Interaction
Graphics and Visual Computing
Intelligent Systems
Information Management

Syllabus Search

As part of our effort to personalize NSDL content and make it available as part of course websites, we have collected nearly 8000 syllabi available from the Web. This search engine allows you to search the content of these crawled syllabi.

Keyword: discrete structures [Search]

Show only phrase matches

Subscribe to these search results as Atom/RSS feed:
http://doc.cs.vt.edu/syllabus/feeds/?q=discrete+structures

Results **1** to **10** of **102**

1. **Borgida 205 Home page** - *Computer Science - Rutgers University*
   ## 198:205 **discrete structures** I -- SPRING 2002 ![][1] ### WHO, WHEN, WHERE * **Lectures:** TTh6 (Tuesday, Thursday 4:30pm) in [Engineering B120][2], Busch Campus * **Instructor:** Alex Bor
   Categories: Algorithms and Complexity, Discrete Structures, Intelligent Systems
   Original - Cached - Compare - DocID 8

2. **CSCE 310 Course Title Data Structures and Algorithms Credit Hours ...** -
   CSCE 235 Introduction to **discrete structures** Class Syllabus Spring 2002

Search Result Details

Course Title

Snippet

21. **Object-Oriented Software** -
Syllabus CS2704: ObjectOriented Software: Fall 2002 CS27041 Class Info2 Syllabus3 Calendar4 Assignments5 Course information Title: CS2704: ObjectOriented Software Design and Construction
Categories: Algorithms and Complexity, Information Management, Social and Professional Issues
Original - Cached - Extracted - Compare - DocID 6016

Links to various views

CC 2001 Categories

To compare two syllabi

http://doc.cs.vt.edu/syllabus/search/

Personal NSDL — Virginia Tech & Villanova University

- Home
- Syllabus Repository
- CS Ed News
- News
- People
- Publications
- Meetings

Syllabus Repository Home | Search | Browse | Submit

Syllabus Search

As part of our effort to personalize NSDL content and make it available as part of course websites, we have collected nearly 8000 syllabi available from the Web. This search engine allows you to search the content of these crawled syllabi.

Keyword: media computation Search

Show only phrase matches

Subscribe to these search results as Atom/RSS feed:
http://doc.cs.vt.edu/syllabus/feeds/?q=media+computation

Results **1** to **1** of **1**

1. **Suggested Syllabus for Media Comp in Java** -
   Media Comp and Pictures [Intro-MediaComp-Pictures-Mod5.ppt][40] (19 slides) * Topics: What is **media computation**. How does our color vision work? How do digital cameras and computer displays work? What is a pixel? How can you create and display a picture in Java? * In class look at the com
   Categories: Algorithms and Complexity, Information Management, Social and Professional Issues
   Original - Cached - Compare - DocID 578

http://doc.cs.vt.edu/syllabus/submit/

Personal NSDL — Virginia Tech & Villanova University

- Home
- Syllabus Repository
  - Submit
  - Google Co-op
- CS Ed News
- News
- People
- Publications

Discrete Structures
Programming Fundamentals
Algorithms and Complexity
Architecture and Organization
Operating Systems
Net-Centric Computing
Programming Languages
Human-Computer Interaction
Graphics and Visual Computing
Intelligent Systems
Information Management

Submit your Syllabus

Don't see your syllabus in the search results? Submit it to us for inclusion!

Syllabus URL:

Course Title:

Your email address (if you wish to be notified when your syllabus is added):

Enter the number from the captcha below:

Submit URL

http://doc.cs.vt.edu/syllabus/

Personal NSDL — Virginia Tech & Villanova University

- Home
- Syllabus Repository
  - Submit
  - Google Co-op
- CS Ed News
- News
- People
- Publications

Discrete Structures
Programming Fundamentals
Algorithms and Complexity
Architecture and Organization
Operating Systems
Net-Centric Computing
Programming Languages
Human-Computer Interaction
Graphics and Visual Computing
Intelligent Systems
Information Management
Social and Professional Issues
Software Engineering

Search the Syllabus Repos

Keyword:

Show only phrase matche

CC 2001 Categories

Syllabi in this collection have been categorized
categories. To view syllabi from a particular category
the sidebar.

Discrete Structures
Programming Fundamentals
Algorithms and Complexity
Architecture and Organization
Operating Systems
Net-Centric Computing
Programming Languages
Human-Computer Interaction
Graphics and Visual Computing
Intelligent Systems
Information Management
Social and Professional Issues
Software Engineering
Computational Science and Numerical Methods

Keyword: plagiarism Search

Show only phrase matches

Subscribe to these search results as Atom/RSS feed:
http://doc.cs.vt.edu/syllabus/feeds/?q=plagiarism

Results **1** to **10** of **462**

1. **CS 699 / IT 803 Syllabus, Spring 2005** - *Department of Computer Science - George Mason University*
   [GMU Honor Code][4] will be strictly enforced. We reserve the right to use [MOSS][5] to detect **plagiarism**. Violations of GMU Honor Code will result in an F. [1]: http://www.gmu.edu [2]: http://www.cs.gmu.edu [3]: http://cs.gmu.edu/~aydin [4]: http://www.gmu.edu/c
   Categories: Net-Centric Computing, Operating Systems, Social and Professional Issues
   Original - Cached - Compare - DocID 124

2. **cs306 - Syllabus** - *Department of Computer Science - George Mason University*
   individual efforts as well. **Failure to give proper credit by using quotes and cites constitutes **plagiarism****. [See here for an explanation of quotes and cites for this class][10]. If you have any questions about proper citation method, contact your instructor. **plagiarism** is governed by the GM
   Categories: Information Management, Social and Professional Issues, Software Engineering
   Original - Cached - Compare - DocID 127

3. **CS 571 OPERATING SYSTEMS - Spring 2006 - H. Aydin** - *Department of Computer Science - George Mason University*
   wed only for the group projects, within each group. We reserve the right to use [MOSS][5] to detect **plagiarism**. Violations of GMU Honor Code or a total score of 49 (or less) will result in an F. **Teaching Assistant:** Vinay Devadas (vdevadas@gmu.edu) **TA Office: **ST II, Rm. 268 **TA Of
   Categories: Architecture and Organization, Operating Systems, Software Engineering
   Original - Cached - Compare - DocID 132

4. **CS 773 Syllabus, Spring 2006** - *Department of Computer Science - George Mason University*
   d. [GMU Honor Code][4] will be strictly enforced. We reserve the right to use [MOSS][5] to detect **plagiarism**. Violations of GMU Honor Code will result in an F. ** Course Web Page: ** http://www.cs.gmu.edu/~aydin/cs773 [1]: http://www.gmu.edu [2]:

Search Results for *'plagiarism'*

http://doc.cs.vt.edu/syllabus/compare/

Personal NSDL — Virginia Tech & Villanova University

Discrete Structures

Programming Fundamentals

Algorithms and Complexity

Architecture and Organization

Operating Systems

Net-Centric Computing

Programming Languages

Human-Computer Interaction

Graphics and Visual Computing

Intelligent Systems

Information Management

Left: CS 415
Right: Computer Organization

CC 2001 Categories	CS 415	Computer Organization
Operating Systems	✔	✔
Programming Languages	✔	✘
Information Management	✔	✘
Architecture and Organization	✘	✔
Software Engineering	✘	✔

Content Syndication

- Search service provider to Google Coop for Syllabus Search
- To be made available to NSDL through Citidel
- RSS feeds for search & browsing

Summary

- Crawled for syllabi
- Classified into Full / Partial / Entry / Noise
- Created syllabus repository
- Built search
- Tagged collection ➠ browse, compare
- Syndicate content externally

Future work

- Better crawling
  - Less noise, focused crawler
  - Intensive: more syllabi per university
  - Extensive: more universities covered
- Services
  - Assisting instructors when creating new courses
  - Curriculum design & accreditation

Questions

More info and live service at:

http:// *doc.cs.vt.edu*