



THE UNIVERSITY OF WINNIPEG

APPLIED COMPUTER SCIENCE

Course Number-ACS-7102/3-001

Course Name – Web and Document Databases

Instructor Information

Instructor: Dr. Yangjun Chen

Office: 3D27

E-mail: y.chen@uwinnipeg.ca

Office Hours: 15:00 - 16:00 pm Monday and Wednesday
10:00 – 17:00 Friday

Class Meeting Time: Monday and Wednesday 11:30 – 12:45pm **Room No:** 3D03

Instructor home page: <http://www.uwinnipeg.ca/~ychen2>

Important Dates

1. First Class Date: January 4, 2017
2. Midterm Exam/Tests/Quizzes (26%): Feb. 27, 2017
3. Final Exam (Comprehensive, 50%): The final examination may be replaced by a project, for which the students are required to implement some challenging algorithms for tree-pattern query evaluation and indexing mechanism (using a computer language) and make tests and comparison.
4. Final Withdrawal Date w/o academic penalty: March 01, 2017
(A minimum of 20% of the work on which the final grade is based will be evaluated and available to the student before the voluntary withdrawal date)
5. Reading Week: February 19th – 25th, 2017 (no classes)

Additional Course Related Information

1. When it is necessary to cancel a class due to exceptional circumstances, instructors will make every effort to inform you via uwinnipeg email, as well as the departmental assistant and Chair/Dean so that class cancellation forms can be posted outside classrooms.
2. Your uwinnipeg email address will normally be used for course related correspondence.
3. Please note that withdrawing before the VW date does not result in a fee refund.
4. Class make-up days are scheduled at the end of term for courses that conflict with holidays.

Course Objectives/Learning Outcomes

This course focuses on design and implementation of web and document databases, which can be defined to be a traditional database system plus the management of data distributed over Web servers. The course covers the theories of database design, different indexing techniques, semi-structured data management, eXtended Markup Language (XML), algorithms for evaluating tree-pattern queries, reachability queries, as well as some important graph algorithms related to web management. Through the study of this course, students will develop skills to solve hard problems in data organization.

Evaluation Criteria

6. Assignments (24%)
 - Number of Assignments: 3
 - Information about assignments
 - Two weeks after the assignments are issued.
 - Late work will receive a 25% penalty.
 - Work should be typed and handed in through e-mail.
 - Work should be prepared in English.
7. Midterm Exam/Tests/Quizzes (26%)
 - Feb. 28, 2017, 11:30 – 12:45
6. Final Exam (50%)
7. The final examination may be replaced by a project, for which the students are required to implement some challenging algorithms for tree-pattern query evaluation and indexing mechanism (using a computer language) and make tests and comparison.
9. Other course evaluation items (%) (*e.g., class participation, presentation*)

Final Letter Grade Assignment

Historically, numerical percentages have been converted to letter grades using the following scale. However, instructors can deviate from these values based on pedagogical nuances of a particular class, and final grades are subject to approval by the Department Review Committee.

A+	90+ - 100%	B	70 - 74%	F	below 50%
A	85 - 90%	C+	65 - 69%		
A-	80 - 84%	C	60 - 64%		
B+	75 - 79%	D	50 - 59%		

Test/Exam Requirements

- Photo ID is not required.
- The book is closed. Calculators/electronic translators can be used.

Students with documented disabilities, temporary or chronic medical conditions, requiring academic accommodations for tests/exams (e.g., private space) or during lectures/laboratories (e.g., note-takers) are encouraged to contact Accessibility Services (AS) at 786-9771 or accessibilityservices@uwinnipeg.ca to discuss appropriate options.

All information about a student's disability or medical condition remains confidential.
<http://www.uwinnipeg.ca/accessibility>.

Students facing a charge of academic or non-academic misconduct may choose to contact the University of Winnipeg Students' Association (UWSA) where a student advocate will be available to answer any questions about the process, help with building a case, and ensuring students have access to support. For more information or to schedule an appointment, visit our website at www.theuwsa.ca/academic-advocacy or call 204-786-9780.

We ask that you please be respectful of the needs of classmates and instructors/professors by avoiding the use of unnecessary scented products while attending lectures. Exposure to scented products can trigger serious health reactions in persons with asthma, allergies, migraines or chemical sensitivities. Please consider using unscented necessary products and avoiding unnecessary products that are scented (e.g. perfume).

Required Text Book(s)/Reading List

- The course uses a book:

No text book.

Reference books and articles:

- *Database Systems (the complete book)*, 2nd Ed. by Garcia-Molina, Ullman, Widom, Pearson Prentice Hall, 2009.
- *Introduction to Algorithms*, 2nd Ed. by Cormen, Leiserson, Rivest, & Stein (CLRS), McGraw Hill, 2002.
- Lecture slides online
- Chen, Yangjun: General Spanning Trees and Reachability Query Evaluation, in Proc. Canadian Conference on Computer Science and Software Engineering (C3S2E'09), Montreal, Canada, 2009, IEEE, pp. 243 – 252.
- Chen, Yangjun: General Spanning Trees and Core Labeling, in Proc. 4th Conf. on Software and Data Technology, 26 – 29 July, 2009, Sofia, Bulgaria.
- Chen, Yangjun: Unordered Tree Matching and Tree pattern Queries in XML databases, in Proc. 4th Int. Conf. on Software and Data Technology, 26 – 29 July, 2009, Sofia, Bulgaria.
- Chen, Yangjun: Bottom-up Evaluation of Twig Pattern Queries in XML Document Databases, in Proc. 20th Int. Conf. on Database and Expert Systems Applications, 31 August – 4 Sept., 2009, Linz, Austria.

Prerequisite Information (This information can be found in the UW Course Calendar)

- Consent of the Department Graduate Program Committee Chair or Instructor.

Misuse of Computer Facilities, Plagiarism, and Cheating

Academic dishonesty is a very serious offense and will be dealt with in accordance with the University's policies. Be sure that you have read and understood **Student Discipline #9** in the 2016-2017 UW Graduate Course Calendar.

Topics to be Covered (tentative)

- Database basics
 - Database system architecture
 - Enhanced entity-relationship diagram
 - Rules for mapping EERD to relational schema
 - Normalization theory
 - 1st, 2nd, 3rd, and 4th normal forms
 - Boyce codd normal form
- Indexes
 - B+-tree construction and maintenance
 - R-tree, Kd-tree, and Quad-tree construction
 - Bitmap
- Semi-structured data model
- Programming languages for XML
- Databases and Internet
- Evaluation of tree pattern queries
 - Tree searching and graph searching
 - Unordered tree matching
 - Ordered tree matching
 - XB-trees
- Reachability query evaluation
 - Network flow problem – node disjoint paths
 - Bipartite graphs – graph stratification
 - Graph decomposition
 - Transitive closure and reachability checking
- Finding most popular packages
 - Signature files
 - Signature trees, signature tree construction and search
 - SPD-trees, SPD-tree construction and search

Projects (tentative topics):

- Implementing an algorithm for constructing an XB-tree
- Implementation of an algorithm for decomposing a DAG into a minimized set of chains
- Implementing an algorithm for evaluating unordered tree pattern queries
- Implementing an algorithm for evaluating ordered tree pattern queries