CMSC 724*Reading List

Sudarshan S. Chawathe†

Spring 2003

This document will change, so please check it often.

1 Schedule

This schedule is approximate and we may need to move topics around based on how fast we can cover material, feedback from the class, etc. It is almost guaranteed that we will not follow it exactly. *Do not plan major events (trips, weddings, proving P=NP, etc.) based on this schedule.* If you need to verify whether some date is OK, check with me.

<table>
<thead>
<tr>
<th>#</th>
<th>Date</th>
<th>Material due (numbers refer to the sections below)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>31 Jan</td>
<td>Introduction, 2.1</td>
</tr>
<tr>
<td>02</td>
<td>07 Feb</td>
<td>2.2</td>
</tr>
<tr>
<td>03</td>
<td>14 Feb</td>
<td>2.3</td>
</tr>
<tr>
<td>04</td>
<td>21 Feb</td>
<td>2.4</td>
</tr>
<tr>
<td>05</td>
<td>28 Feb</td>
<td>2.5</td>
</tr>
<tr>
<td>06</td>
<td>07 Mar</td>
<td>2.5 (continued) and review</td>
</tr>
<tr>
<td>07</td>
<td>14 Mar</td>
<td>5:00pm Take-home midterm assigned.</td>
</tr>
<tr>
<td></td>
<td>– 20 May</td>
<td>8:00am Deadline for submitting take-home midterm.</td>
</tr>
<tr>
<td>08</td>
<td>21 Mar</td>
<td>midterm discussion</td>
</tr>
<tr>
<td></td>
<td>– 28 Mar</td>
<td>Spring Break (no class)</td>
</tr>
<tr>
<td>09</td>
<td>04 Apr</td>
<td>2.6</td>
</tr>
<tr>
<td>10</td>
<td>11 Apr</td>
<td>2.9</td>
</tr>
<tr>
<td>11</td>
<td>18 Apr</td>
<td>Project Discussions</td>
</tr>
<tr>
<td>12</td>
<td>25 Apr</td>
<td>2.8</td>
</tr>
<tr>
<td>13</td>
<td>02 May</td>
<td>2.8 (continued)</td>
</tr>
<tr>
<td>14</td>
<td>09 May</td>
<td>Project Presentations</td>
</tr>
<tr>
<td></td>
<td>– 16 May</td>
<td>5:00pm Take-home final assigned.</td>
</tr>
<tr>
<td></td>
<td>– 20 May</td>
<td>8:00am Deadline for submitting take-home final.</td>
</tr>
</tbody>
</table>

*http://www.cs.umd.edu/class/spring2003/cmsc724/
†http://www.cs.umd.edu/users/chaw/
2 Topics

The following sections should be updated with details and summaries of the class discussions as we proceed. You should be able to find most of these papers very easily on the Web. (For more details, see Section 3.) You should read the material indicated below before the class meeting at which it is due so that you can actively participate in the discussion. You should read the papers critically, noting, for example, the advantages and limitations of the proposed methods. You should be prepared to both ask and answer questions intelligently. The class participation portion of your grade depends on such interactions. More importantly, if you do not do the readings before class, you will not benefit from the classroom discussions (which will assume you have read the material carefully).

2.1 Conjunctive and First-Order Queries
   - Chapters 4 and 5 of [AHV95].

2.2 Standard query processing techniques.
   - Chapters 15 and 16 of [GMUW02]
   - Graefe’s survey: [Gra93] Local copy
   - The classic System R paper [SAC+79]

2.3 XML Query Languages
   - Chapters 4-6 of [ABS99]
   - The XPath language specification [CD99]
   - The XQuery language specification [BCF+02]
   - Two papers on Lore and the Lorel language [MAG+97, AQM+96]
   - A paper on the UnQL language [BDHS96]

2.4 Datalog, Recursion, and Negation
   - Chapters 14, 15, and 16 of [AHV95]

2.5 Data storage and indexing
   - Chapters 12-14 of [GMUW02]
2.6 Describing Semistructured Data

- Chapter 7 of [ABS99]
- Representative Objects [NUWC97] Local copy
- Graph Schemas [BDFS96] Local copy
- DataGuides [GW97] Local copy
- Synopses
- DTD, RDF, XSchema, ...

2.7 Expressiveness and Complexity of Query Languages

- Chapters 16-18 of [AHV95]

2.8 Data Integration

- Chapter 20 of [GMUW02]
- Information Integration Using Logical Views [Ull97] Local copy
- Theory of Answering Queries using Views [Hal00] Local copy

2.9 Streaming Data

- Streaming XPath [PC03] Local copy

3 Resources

The ACM Digital Library\(^1\): Requires a subscription, but UMD has a site-wide subscription that gives access from all local machines.

The DBLP Bibliography Server\(^2\) has extremely good coverage of the Database and Logic Programming fields.

ACM SIGMOD\(^3\).

VLDB Foundation\(^4\).

SIGMOD Record\(^5\)

---

\(^1\)http://www.acm.org/dl/
\(^2\)http://www.purl.org/net/dblp
\(^3\)http://www.acm.org/sigmod/
\(^4\)http://www.vldb.org
\(^5\)http://www.acm.org/sigmod/record/
IEEE Data Engineering Bulletin\textsuperscript{6} 

Maryland Database Group\textsuperscript{7} with pointers to other relevant DB resources.

Modern Information Retrieval [BYRN99]. Use this book for an overview of Information Retrieval. The huge list of references is a big plus.

Readings in Database Systems [SH98]. This collection of papers is typically covered in CMSC 624 and similar courses. It includes many famous papers, such as "the System R paper," "the ARIES paper," and Gray et al.'s locking paper.

Principles of Distributed Database Systems [OV99]. Look here for distributed query optimization, distributed transaction processing, etc.

References


\textsuperscript{6}\url{http://www.research.microsoft.com/research/db/debull}

\textsuperscript{7}\url{http://www.cs.umd.edu/areas/db/}


