<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief Examiner</td>
<td>1</td>
</tr>
<tr>
<td>Lecturer(s) / Leader(s):</td>
<td>1</td>
</tr>
<tr>
<td>Caulfield</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>Unit synopsis</td>
<td>2</td>
</tr>
<tr>
<td>Learning outcomes</td>
<td>2</td>
</tr>
<tr>
<td>Contact hours</td>
<td>3</td>
</tr>
<tr>
<td>Workload</td>
<td>3</td>
</tr>
<tr>
<td>Unit relationships</td>
<td>3</td>
</tr>
<tr>
<td>Prerequisites</td>
<td>3</td>
</tr>
<tr>
<td>Teaching and learning method</td>
<td>4</td>
</tr>
<tr>
<td>Teaching approach</td>
<td>4</td>
</tr>
<tr>
<td>Timetable information</td>
<td>4</td>
</tr>
<tr>
<td>Tutorial allocation</td>
<td>4</td>
</tr>
<tr>
<td>Unit Schedule</td>
<td>4</td>
</tr>
<tr>
<td>Unit Resources</td>
<td>6</td>
</tr>
<tr>
<td>Prescribed text(s) and readings</td>
<td>6</td>
</tr>
<tr>
<td>Recommended text(s) and readings</td>
<td>6</td>
</tr>
<tr>
<td>Equipment and consumables required or provided</td>
<td>6</td>
</tr>
<tr>
<td>Study resources</td>
<td>6</td>
</tr>
<tr>
<td>Assessment</td>
<td>7</td>
</tr>
<tr>
<td>Overview</td>
<td>7</td>
</tr>
<tr>
<td>Faculty assessment policy</td>
<td>7</td>
</tr>
<tr>
<td>Assignment tasks</td>
<td>7</td>
</tr>
<tr>
<td>Examination</td>
<td>8</td>
</tr>
<tr>
<td>Due dates and extensions</td>
<td>8</td>
</tr>
<tr>
<td>Late assignment</td>
<td>9</td>
</tr>
<tr>
<td>Return dates</td>
<td>9</td>
</tr>
<tr>
<td>Feedback</td>
<td>9</td>
</tr>
<tr>
<td>Appendix</td>
<td>10</td>
</tr>
</tbody>
</table>
Introduction

Welcome to FIT5166 - Information retrieval systems. This unit is a postgraduate unit in the Faculty of Information Technology and a unit in the data management specialization of the Master of Information Technology and Master of Applied Information Technology. FIT5166 is running for the first time in 2010. As the unit is at level 5, students of FIT5166 will be assumed independent learners. Students will need to spend time outside of formal lectures and tutorials to do their own research and read widely in order to pass the unit.

In addition to the prerequisite knowledge listed in this unit guide (below), the entry requirements of the Master of Information Technology (a first degree in IT) or the first year of the MAIT will be assumed background of students. Students who do not have this background (for example, have not done more than basic programming) are advised to seek advice of the lecturer before enrolling in the unit. Students may also find the unit difficult if they are unfamiliar with basic linear algebra, vector space concepts and probability theory.

Please note the following important points:

- It is planned that 10 topics will be covered. Some topics may last more than one lecture week, and the number of weeks a topic will take won’t be determined before the semester starts. A constraint of the FIT unit guide editor is that this type of unit schedule is not supported. Thus, students must refer to the more detailed unit guide that will be provided in week 1 of the unit for the topic schedule.

- Lectures are timetabled for 2 hours, however the lectures will run between 1 and 2 hours, depending on the topic (with the rest of the lecture time generally then devoted to class discussion – it is expected that all students will participate in this discussion).

- Lecture slides will be provided after the lectures. Before the lectures, a list of topics to be discussed will be provided and students will benefit from participating in group discussion via Google Wave (details to be provided in class) before and after the lecture. During the lecture, a series of questions will be provided to students, which should be answered during note taking as the lecture proceeds.

- Tutorials will be theoretical in nature. Computers in the laboratory may be used for assignment coding when the tutorial discussion has concluded. The computer lab will be booked for 2 hours, with the tutorial portion normally going for 1.5 hours.

- Lectures will not be recorded live, however an enhanced podcast feed will be provided after the lecture.

Unit synopsis

This unit presents students with the theory and practice underlying computerised information retrieval. Topics covered include: history and context of information retrieval systems, retrieval models: Boolean, vector space and probabilistic, evaluation strategies and test collections, web search engines, very large databases, indexing, content-based multimedia retrieval, relevance feedback.

Learning outcomes

At the completion of this unit students will:

• appreciate the context and application of information retrieval systems;
• understand the different models of information retrieval and their comparative advantages and disadvantages;
understand how objective and subjective evaluation strategies are used with information retrieval systems;
• be conversant with the issues and challenges of managing very large databases of heterogeneous data for information retrieval;
• understand how web search engines and search algorithms are constructed, utilised and deployed;
• appreciate the different indexing strategies that can be used for the retrieval of information across different modalities;
• understand how information retrieval systems can be used for the retrieval of audio-visual information;
• appreciate the utility of and understand the theoretical and practical underpinnings of relevance feedback in information retrieval systems.

Contact hours

2 hrs lectures/wk, 2 hrs laboratories/wk

Workload

The university policy on workloads assumes that credit points are correlated with expected hours of workload by students (with 1 credit point being equal to approximately 2 hours/week of workload). While this would imply that a 6 point unit implies approximately 12 hours/week, the reality is that many students will need to expend much less or much more than this time to achieve a good grade in the unit. Ultimately the workload imposed on you as a student will depend on your background and your level of participation in discussions. In a postgraduate level 5 unit, you will be expected to be independently research topics when required, and thus workload will also depend on your research skills and time management.

Unit relationships

Prerequisites

Recommended knowledge: Basic familiarity with file organisation theory and UNIX.
Teaching and learning method

Teaching approach

Lectures, tutorials, online and in-class discussion. Google Wave will be piloted as a discussion medium in this unit (more detail will be provided in class).

Timetable information

For information on timetabling for on-campus classes please refer to MUTTS, http://mutts.monash.edu.au/MUTTS/

Tutorial allocation

On-campus students should register for tutorials/laboratories using the Allocate+ system: http://allocate.its.monash.edu.au/

Unit Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Date*</th>
<th>Topic</th>
<th>Key dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19/07/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>26/07/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>02/08/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>09/08/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>16/08/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>23/08/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>30/08/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>06/09/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>13/09/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>20/09/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mid semester break</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>04/10/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>11/10/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>18/10/10</td>
<td>REFER TO DETAILED UNIT SCHEDULE TO BE PROVIDED IN WEEK 1</td>
<td></td>
</tr>
</tbody>
</table>
*Please note that these dates may only apply to Australian campuses of Monash University. Off-shore students need to check the dates with their unit leader.
Unit Resources

Prescribed text(s) and readings

Readings will be provided during the semester. More will be said concerning this in week 1.

Recommended text(s) and readings

Equipment and consumables required or provided

Students studying off-campus are required to have the minimum system configuration specified by the Faculty as a condition of accepting admission, and regular Internet access. On-campus students, and those studying at supported study locations may use the facilities available in the computing labs. Information about computer use for students is available from the ITS Student Resource Guide in the Monash University Handbook. You will need to allocate time per week for use of a computer, including time for newsgroups/discussion groups.

Study resources

Study resources we will provide for your study are:

- Lectures and tutorials.
- Weekly lecture slides (provided after lectures). Note that the lecture slides will not be detailed, thus it is important that students take adequate notes in class.
- Pointers to reading.
- Enhanced podcast of lecture material.
- Assignment specifications (see assessment section). No sample assignment solutions will be provided.
- Tutorial questions and access to tutorial discussions. No tutorial answers will be provided other than solutions discussed during tutorial classes.
- A detailed unit guide including the full schedule of unit topics (in addition to this document).
- A Moodle website and/or ftp site through which resources can be downloaded.
- Other online discussion media (note Google Wave rather than Moodle will be used predominantly for online discussions in FIT5166).
Assessment

Overview

Examination (2 hours): 50%; In-semester assessment: 50%

Faculty assessment policy

To pass a unit which includes an examination as part of the assessment a student must obtain:

• 40% or more in the unit's examination, and
• 40% or more in the unit's total non-examination assessment, and
• an overall unit mark of 50% or more.

If a student does not achieve 40% or more in the unit examination or the unit non-examination total assessment, and the total mark for the unit is greater than 50% then a mark of no greater than 49-N will be recorded for the unit.

Assignment tasks

Assignment coversheets

Assignment coversheets are available via "Student Forms" on the Faculty website: http://www.infotech.monash.edu.au/resources/student/forms/
You MUST submit a completed coversheet with all assignments, ensuring that the plagiarism declaration section is signed.

Assignment submission and return procedures, and assessment criteria will be specified with each assignment.

Assignment submission and preparation requirements will be detailed in each assignment specification. Submission must be made by the due date otherwise penalties will be enforced. You must negotiate any extensions formally with your campus unit leader via the in-semester special consideration process: http://www.infotech.monash.edu.au/resources/student/equity/special-consideration.html.

• Assignment task 1

Title: IR Research Report

Description: Students must submit a 2,500 word research report on an approved topic concerning information retrieval. Subject to enrolment numbers, the research report will also be presented to other students and the lecturer as a component of its assessment. A list of possible topics will be provided during the second week of semester as will more detail on submission requirements.

Weighting: 25%

Criteria for assessment: Clarity of expression, grammar and spelling, appropriate structure, adequate coverage of research literature, quality of discussion and viability of conclusions drawn.

Due date: Week 11 (if presentations are required). Week 12 (if presentations are not required)
Remarks:
The due date will be set in week 2 of semester, depending on enrolment numbers in the unit.

• Assignment task 2

Title:
Retriever Implementation

Description:
Students must submit an implementation of a text retriever based on the vector space retrieval model. The retriever will implement at an absolute minimum, a tokenizing routine, tf/idf weighting of index vectors and the cosine similarity metric to retrieve from a supplied set of documents. Extra features (suggestions will be provided) will achieve extra marks. This retriever may be written in Java or C (or another language subject to lecturer approval). More detail on the submission requirements will be provided in week 2. This description should be considered subject to change and will be finalised in week 2 of semester. Students may be interviewed as part of their submission.

Weighting:
25%

Criteria for assessment:
Originality of submission and degree to which specified requirements are met by the program.

Due date:
Week 12

Examination

•

Weighting:
50%

Length:
2 hours

Type (open/closed book):
Closed book

Electronic devices allowed in the exam:
None

See Appendix for End of semester special consideration / deferred exams process.

Due dates and extensions

Please make every effort to submit work by the due dates. It is your responsibility to structure your study program around assignment deadlines, family, work and other commitments. Factors such as normal work pressures, vacations, etc. are not regarded as appropriate reasons for granting extensions. Students are advised to NOT assume that granting of an extension is a matter of course.

Students requesting an extension for any assessment during semester (eg. Assignments, tests or presentations) are required to submit a Special Consideration application form (in-semester exam/assessment task), along with original copies of supporting documentation, directly to their lecturer within two working days before the assessment submission deadline. Lecturers will provide specific outcomes directly to students via email within 2 working days. The lecturer reserves the right to refuse late applications.
A copy of the email or other written communication of an extension must be attached to the assignment submission.

Refer to the Faculty Special consideration webpage or further details and to access application forms: http://www.infotech.monash.edu.au/resources/student/equity/special-consideration.html

**Late assignment**

Late assignments will not be accepted.

**Return dates**

Students can expect assignments to be returned within two weeks of the submission date or after receipt, whichever is later.

**Feedback**

Types of feedback you can expect to receive in this unit are:

- Informal feedback on progress in labs/tutes
- Graded assignments with comments
- Interviews
Appendix

Please visit the following URL: http://www.infotech.monash.edu.au/units/appendix.html for further information about:

- Continuous improvement
- Unit evaluations
- Communication, participation and feedback
- Library access
- Monash University Studies Online (MUSO)
- Plagiarism, cheating and collusion
- Register of counselling about plagiarism
- Non-discriminatory language
- Students with disability
- End of semester special consideration / deferred exams