Concurrent Thread-based Web Crawler Work Plan
CM0343

Michael Graham

Notes:

- Columns shaded in red are milestones
- Meetings with supervisor is every two weeks towards end of the working week

Date / Week	Objectives	Completed
15/10/2012 – 19/10/2012 (Week 1)	• Determine what is expected from project by having first initial meeting with Supervisor • Build initial plan for project	✔
19/10/2012	**MILESTONE – INITIAL PLAN SUBMISSION**	✔
22/10/2012 – 26/10/2012 (Week 2)	• Start initial research on existing web crawler techniques • Build notes on research in preparation of both interim report and start of implementation	
29/10/2012 – 9/11/2012 (Weeks 3 & 4)	• Continue research regarding web crawler algorithms • Finish research with a small overview of existing architecture and execution environments with current large-scale web-crawlers • Prepare and organize notes taken from research in preparation for starting Interim Report	
12/11/2012 – 23/11/2012 (Weeks 5 & 6)	• Plan Web Crawler architecture • Determine best scheduling technique, keeping technique relatively simple for preliminary implementation • Plan application structure ready for initial implementation	

Concurrent Thread-based Web Crawler Work Plan
CM0343

Michael Graham

26/11/2012 – 7/12/2012 (Weeks 7 & 8)	• Build simple implementation that meets the basic project requirement • Gather preliminary information from simple implementation such as execution times, testing and note any analysis for Interim report preparation	
10/12/2012 – 14/12/2012 (Week 9)	• Convert all notes from research into suitable section within Interim report • Portray initial simple implementation of web crawler analysis in an appropriate manner within the interim report • Build & complete interim report	
14/12/2012	**MILESTONE – ITERIM REPORT SUBMISSION**	
28/01/2013 – 01/02/2013 (Week 10)	• Determine areas within web crawler application that require improvement • Compile a list of areas of the crawler that will be improved, adjust architecture accordingly	
04/02/2013 – 15/02/2013 (Weeks 11 & 12)	• Make changes to initial implementation according • Make note of any changes	
18/02/2013 – 01/03/2013 (Weeks 13 & 14)	• Determine advanced features to be implemented into the crawler, such as different crawling algorithms, search constraints etc… • Justify choices of advanced features in preparation for Final report deliverable • Make changes to architecture plan etc accordingly	
04/03/2013 – 29/03/2013 (Weeks 15, 16 & 17)	• Implement advanced features determined in previous two weeks within the crawler	

Concurrent Thread-based Web Crawler Work Plan
CM0343

Michael Graham

	• Test new advanced features implemented to ensure correct execution	
01/04/2013 – 26/04/2013 (Weeks 18, 19, 20 & 21)	• Begin to prepare draft for final report submission • Gather runtime execution times ready for final report • Make comparisons with initial simple implementation against advanced implementations • Take note of any comparisons and discoveries	
29/04/2013 – 03/05/2013 (Week 22)	• Write up final version of final report, deriving from final report draft created in previous weeks • Prepare web crawler code for submission	
05/05/2013	**MILESTONE – FINAL REPORT SUBMISSION**	