Computer Science Tripos – Part II – Progress Report

Exploring the structure of mathematical theories using graph databases

Dhruv C. Makwana, Trinity College

dcm41@cam.ac.uk

Project Supervisor: Dr. Timothy G. Griffin

Directors of Studies: Dr. Frank Stajano & Dr. Sean B. Holden

Project Overseers: Dr. David J. Greaves & Prof. John Daugman

Scheduling

The project is on schedule. Although the some of the details of the work to be done (especially regarding the script-to-CSV tool) have changed, most of it (compiled-to-CSV and library of queries) has been as predicted.

Firstly, work on the compiled-to-CSV tool is further along (with more detailed models, such as hierarchical modules and type-constructor relations.) than anticipated, reducing the need for the script-to-CSV tool. The current implementation suffers from some strange results (thanks to the representation of modules types/functors) but is overall, working as expected.

Work on the library of queries/tools is further along than expected thanks to pre-existing solutions (APOC: Awesome Procedures on Cypher) to some of the graph algorithm and query ideas considered. What remains to be done is either implementing a new metric or an existing one, aiming for better performance.

Lastly, work on the script-to-CSV tool has been temporarily suspended. This is because the model, as it stands, of Coq's proof-objects is sufficiently detailed to run some queries on and develop a library for. The only missing features are (a) inclusion of notation in the model and (b) appropriate de-duplications of module types (and dealing appropriately with opaque `:=` and transparent `<:` ascriptions).

Unexpected Difficulties

Understanding the Coq AST and API has been even more challenging than anticipated. The project is large and poorly documented, so a lot of time was spent trying to figure out how to find and chain appropriate AST transformations.

Once Coq was better understood, it became apparent that the full Coq AST was impractical to work with for this project and a easier alternative would be to process `glob` files, (global reference) which will provide the aformentioned missing features.

Accomplished Work

Accomplished work is detailed in Table 1.

~~Strikethroughs~~ represent work done. Green represents additions to the original schedule presented in the proposal and red shows work removed or rescheduled.

Exploring queries and starting the library was delayed (to January) as it became clear the compiled-to-CSV tool could provide more information than originally expected (and thus more time was reallocated to that).

CoqPL is a workshop (to which this project was submitted). It was removed from the schedule when the proposal submission of this project was rejected. The time for this was allocated to getting ahead on writing the final dissertation.

Table 1: Milestones: Updated (~~Done~~, Removed, Added)

Date	Milestone
21-10-2016	~~Complete Project Proposal~~
04-11-2016	~~Finish a prototype compiled-to-CSV tool.~~
	~~Get familiar with Neo4j Cypher.~~
	~~Understand how to use the Coq parser.~~
18-11-2016	~~Refine compiled-to-CSV tool: tests and documentation.~~
	Explore queries possible and start the library.
	~~Begin work on translating Coq constructs from proof-scripts.~~
02-12-2016	Finish a prototype script-to-CSV tool.
	~~Refactor compiled-to-CSV tool, add new features.~~
16-12-2016	Test and document script-to-CSV tool.
	~~Test and document refactored compiled-to-CSV tool.~~
30-12-2016	~~Begin work on integrating tools into one workflow.~~
13-01-2017	~~Stabilise and document whole project so far.~~
	Prepare presentation for CoqPL Conference.
	~~Write Introduction, Preparation chapters.~~
27-01-2017	Look at SSReflect and evaluate changes to be made.
	~~Explore tools and queries possible, start library.~~
10-02-2017	Incorporate changes from feedback/new features.
	Implement new graph query/algorithm.
24-02-2017	Test and document the new features.
10-03-2017	Write Introduction, Preparation and Implementation chapters.
	Write Implementation chapters.
	Implement script-to-CSV tool.
24-03-2017	Fix bugs/unexpected problems.
07-04-2017	Write Evaluation and Conclusion chapters.
21-04-2017	Fix bugs/unexpected problems.
05-05-2017	Complete Dissertation (references, bibliography, appendix, formatting).