

Thesis/Dissertation Sheet

Surname/Family Name : **Robinson**

Given Name/s : **Amos Stephen**

Abbreviation for degree as in the University calendar : **PhD**

Faculty : **Faculty of Engineering**

School : **School of Computer Science and Engineering**

Thesis Title : **The stuff that streams are made of**

Abstract 350 words maximum: (PLEASE TYPE)

To learn interesting things from large datasets, we generally want to perform lots of queries. If we compute each query separately, we may spend more time reading the data than we spend computing the answer. Instead of computing each query separately, we would like to amortise the cost of reading the data by performing multiple queries at the same time.

Two streaming models for executing multiple queries at a time are push streams and Kahn process networks.

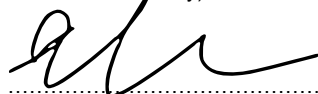
Push streams can execute multiple queries at a time, but these queries can be unwieldy to write as they must be constructed "back-to-front". We introduce a query language called Icicle, which allows programmers to write and reason about queries using a more familiar array-based semantics, while retaining the execution strategy of push streams. The type system of Icicle guarantees that well-typed query programs have the same semantics whether they are executed as array programs or as stream programs, and that all queries over the same input can be executed together.

However, push streams do not support computations with multiple inputs except for non-deterministically merging two streams. Kahn process networks support both multiple inputs and multiple queries, but require dynamic scheduling and inter-process communication, both of which introduce significant overhead. We introduce a method for taking multiple processes in a Kahn process network and fusing them together into a single process. The fused process communicates through local variables rather than costly communication channels. This fusion method generalises previous work on stream fusion and demonstrates the connection between fusion and synchronised product of processes, which is generally used as a proof technique rather than an optimisation.

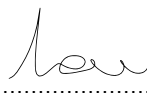
Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).



Signature



Witness Signature

2018/08/24

Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

**FOR OFFICE USE
ONLY**

Date of completion of requirements for
Award: