OMNISPECTIVE ANALYSIS AND REASONING

AN EPISTEMIC APPROACH TO SCIENTIFIC WORKFLOWS

SRINIVAS CHEMBOLI

Australian National University

A thesis submitted for the degree of
Doctor of Philosophy
of the
Australian National University

September 2012

version 12.09.04

DECLARATION

I declare that the work in this thesis is entirely my own and that to the best of my knowledge it does not contain any materials previously published or written by another person except where otherwise indicated.

Srinivas Chemboli
04 September 2012

[S]cience includes any approach that is open to reason, to rational discussion, investigation, skepticism, to critical thinking, to questioning. . . I wouldn't say you have to put on a white coat and go into a laboratory in order to [pursue science].

Dawkins [2012]

ACKNOWLEDGMENTS

The research presented in this dissertation could not have been done without the help of many colleagues and friends who supported me in many ways during all the years. Only the names of some of them can be listed here.

First and foremost, I would like to thank my supervisor, Clive Boughton, for his unstinting support and encouragement. Thank you, Clive, for skillfully ensuring the right mix of 'supervision' and 'independent latitude' in research. I have learnt much these past years, and that could not have happened without your guidance.

I have benefited greatly from the valuable feedback and suggestions of my advisors, Shayne Flint and Ramesh Sankaranarayana. Thank you Shayne, for the informed critique and many hours of brainstorming and discussion. I am grateful to Ramesh for steering me towards 'scientific workflows.'

I am also thankful to Elisa Baniassad, chair of my supervisory panel — although Elisa came on board at a later stage of my study, her advice and insights have been very helpful in shaping the research.

Over the past several years, I have had the great pleasure of interacting with an incredible group of people in the School of Computer Science.

In particular, I would like to thank Henry Gardner for taking an active interest in my research and offering valuable advice.

I am also thankful to Lynette Johns–Boast, Chris Johnson, Clem Baker–Finch and Alexei Khorev for giving me the opportunity to learn to teach. The skills acquired from this experience have been very useful in my research. I would like to thank Tom Worthington, Malcolm Newey and Jay Larson for numerous helpful and enlightening discussions about conducting and presenting research.

To Julie Arnold, Chelsea Holton, Fiona Quinlan, Bethany Flanders, Suzy Andrew, Debbie Pioch, Jill Mayo, Marie Katselas, Sue van Haeften, Jonathan Peters, Paul Melloy and Jadon Radcliffe: A Big Thank You for cheerfully handling all administrative niggles. Special thanks to Bob Edwards, Hugh Fisher, Steven Hanley and in particular to James Fellows, for technical support and wizardry.

During my PhD, I also had the opportunity to work as an educational technologist in the Division of Information at the ANU. I would like to thank Karen Visser, Jenny Edwards, Marina Lobastov, Peter Yates, Jaymie Parker, Sam Primrose, Grazia Scotellaro, Candida Spence, Hans Joerg–Kraus, Kim Blackmore and Lauren Thompson — it was a privilege to work with you.

In addition, I also participated in the Pinnacle Teaching Program, the Graduate Teaching Program and the Foundations of University Teaching and Learning course. I am thankful to Trevor Vickers and Beth Beckmann for their support in these programs.

It would be remiss of me not to express my thanks to Mark Drechsler for insights of Moodle and learning management systems.

I would also like to thank Deborah Veness for her comments and suggestions on the application of OAR to course design and for encouraging me to present the work at ascilite.

I am thankful to Tony Karrys, Gina Denman, Kaori Oikawa and Walter Sauer for their assistance and kindness.

This research has been supported by the Australian National University, and the Commonwealth of Australia, through the Cooperative Research Centre for Advanced Automotive Technology. I would like to thank Kate Neely for her support and assistance in coordinating with the AutoCRC.

I would like to thank the many people with whom I have shared my time at CECS: Agung Fatwanto, Normi Abu Bakar, Zoe Brain, Luke Nguyen–Hoan, Alvin Teh, Amir Hadad, Sukanya Manna, Josh Milthorpe, Jaison Mulerikkal, Derek Wang and Ian Wood. You have been excellent company. I must particularly thank my officemate and friend Ziyad Alshaikh. Thank you for the sparkling wit and interesting discussions — no research ideas were too outlandish for conversation!

Special thanks to:

> Aditi Barthwal and Kate Sullivan, for your friendship, support and encouragement.
>
> Josephine Wright, for your optimism and friendly advice.
>
> Ganesh Venkateswara, for taking time out and providing thoughtful insights and feedback.

Finally, I would like to thank my family for their continuous support and encouragement. Without you this work would never have been possible.

ACKNOWLEDGMENTS OF PERMISSIONS

I want to acknowledge the permissions of various copyright holders.

The Workflow Management Coalition kindly permitted me to use the following material:

> From D. Hollingsworth, "The Workflow Reference Model", Document Number TC00-1003, The Workflow Management Coalition, Hampshire, UK : 1995 http://www.wfmc.org/standards/docs/tc003v11.pdf,
>
> Figure 6 (Figure 2.1 in the thesis, redrawn for style and appearance)

Dr Shayne Flint of the Australian National University kindly permitted me to use the following material:

> From S. Flint, "Rethinking Systems Thinking", Proceedings of the 14th ANZSYS Australia New Zealand Systems Society Conference, Perth, Australia : 2008 http://www.anzsys.org/anzsys08/papers/Shayne%20Flint%20Rethinking%20Systems%20Thinking.pdf
>
> Figure 2 (Figure 3.1 in the thesis, redrawn for style and appearance)

The International Society for the Systems Sciences kindly permitted me to use the following material:

> From J.N. Warfield, "The Domain of Science Model: Evolution and Design", Proceedings of the Society for General Systems Research, Salinas, CA: Intersystems, 1986, H46-H59,
>
> Figure 1 (Figure 3.2 in the thesis, redrawn for style and appearance)

Lynette Johns-Boast of the Australian National University kindly permitted me to use the following material:

> "A Model of Curriculum Design", Personal communication : 2010,
>
> (Figure 5.1 in the thesis, redrawn for style and appearance)

ABSTRACT

This thesis presents the conceptualization, formulation, development and demonstration of Omnispective Analysis and Reasoning (OAR), an epistemic framework for managing intellectual concerns in scientific workflows.

Although scientific workflows are extensively used to support the management of experimental and computational research, intellectual concerns are not adequately handled in current practice owing to the focus on low–level implementation details, limited context support, issues in developing shared semantics across disciplines and lack of support for verification and validation of the underlying science of the workflow. The management of intellectual concerns in scientific workflows can be improved by developing a framework for providing a layer of abstraction to lift focus from low–level implementation details, adding context as a workflow parameter, introducing localized ontologies and abstracting and mapping intellectual concerns in the research–domain to workflow specification and execution semantics.

Following an examination of typical definitions of scientific workflow offered in literature, the Scientific Method is applied to develop an enhanced definition of a scientific workflow. This definition, which extends the scope of ordered analysis and investigation to a generic problem scenario, is utilized in the OAR framework. The design of OAR is modular like the Domain of Science Model (DoSM). The structure and working of OAR incorporate the evolving nature of science, hierarchy of conceptualization, omnispection, and the logical processes of analysis, reasoning and abstraction. These form the Foundation and Theory of OAR. Abstracting concerns in terms of unit knowledge entities (ukes) and groups of ukes (recipes), use of context to identify relation between recipes, the management of recipes in shelves, and the processes of concern refinement and context refinement constitute the Methodology.

A comprehensive and simple example of the application of OAR to the abstraction, analysis, formulation and orchestration of a scientific workflow at different levels of granularity is provided by applying it to the problem of origami paper folding. The use of OAR in capturing the rationale of design decisions and mapping them to desired outcomes is demonstrated by applying OAR for contextualizing course design. Another example illustrates the use of OAR in the analysis, understanding and management of complex systems. Localized ontologies enable the exposure of side–effects and emergent behavior in

large–scale systems due to the choice of any particular solution specification. These examples constitute a first step in building the Applications block of OAR. While OAR may be manually applied even to large–scale problems, it is expedient to avail of tool support. Soma — a simple and illustrative tool prototype is developed to indicate directions for a reference tool implementation.

The thesis concludes with a consideration of ideas for future work. The contribution in this thesis corresponds to an instance of the DoSM for scientific workflow management. The OAR framework has great potential for further development as a well–formed Science of Workflows.

PUBLICATIONS

Some ideas and figures have appeared previously in the following presentations and publications:

Chemboli, S. and Boughton, C. [2012a]. "Managing Large and Complex Systems with Omnispective Analysis and Reasoning." In: *Proceedings of SETE APCOSE 2012*. Brisbane, Australia. URL: http://hdl.handle.net/1885/9009

Chemboli, S. and Boughton, C. [2012b]. "Omnispective Analysis and Reasoning: A Framework for Managing Intellectual Concerns in Scientific Workflows." In: *Proceedings of the 5th India Software Engineering Conference*. Kanpur, India, pp. 143–146. DOI: 10.1145/2134254.2134279

Chemboli, S. and Boughton, C. [2011]. "Contextual Course Design with Omnispective Analysis and Reasoning." In: *Changing Demands, Changing Directions. Proceedings ascilite*. Ed. by Williams, G. et al. Hobart, pp. 210–219. URL: http://www.leishman-associates.com.au/ascilite2011/downloads/papers/Chemboli-full.pdf

Chemboli, S. [Oct. 2010a]. *Contextualizing learning outcomes and course design in Moodle*. Presented at Moodleposium AU 2010. Canberra, Australia. URL: http://hdl.handle.net/1885/9279

Chemboli, S., Kane, L., and Johns-Boast, L. [July 2010]. *Translating Learning Outcomes in Moodle*. Presented at Moodlemoot AU 2010. Melbourne, Victoria, Australia. URL: http://ubuntuone.com/4RBjlozHyEyITuy21aDCfe

Chemboli, S. [Feb. 2010b]. *Omnispective Analysis and Reasoning: An epistemic approach to scientific workflows*. Presented at the CECS Seminar Series, Australian National University. Canberra, Australia. URL: http://cecs.anu.edu.au/seminars/more/SID/2503

CONTENTS

LIST OF FIGURES

LIST OF TABLES