Expressibility of Higher-Order Logics on Relational Databases: Proper Hierarchies

A dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Information Systems at Massey University, Wellington, New Zealand.

Flavio Antonio Ferrarotti

2008

Supervisor:
Associate Prof. Dr. José María Turull Torres

Co-Supervisor:
Prof. Dr. Klaus-Dieter Schewe

Internal Examiner:
Prof. Dr. Sven Hartmann

New Zealand Examiner:
Prof. Dr. Robert Goldblatt

Overseas Examiner:
Prof. Dr. Lauri Hella

Date of Examination:
12.06.2008

Dedicated to the memory of my grandfather, Reinaldo

Abstract

We investigate the expressive power of different fragments of higher-order logics over finite relational structures (or equivalently, relational databases) with special emphasis in higher-order logics of order greater than or equal three. Our main results concern the study of the effect on the expressive power of higher-order logics, of simultaneously bounding the arity of the higher-order variables and the alternation of quantifiers. Let $AA^i(r,m)$ be the class of $(i+1)$-th order logic formulae where all quantifiers are grouped together at the beginning of the formulae, forming m alternating blocks of consecutive existential and universal quantifiers, and such that the maximal-arity (a generalization of the concept of arity, not just the maximal of the arities of the quantified variables) of the higher-order variables is bounded by r. Note that, the order of the quantifiers in the prefix may be mixed. We show that, for every $i \geq 1$, the resulting AA^i hierarchy of formulae of $(i+1)$-th order logic is proper. This extends a result by Makowsky and Pnueli who proved that the same hierarchy in second-order logic is proper. In both cases the strategy used to prove the results consists in considering the set $AUTOSAT(F)$ of formulae in a given logic F which, represented as finite structures, satisfy themselves. We then use a similar strategy to prove that the classes of $\Sigma^i_m \cup \Pi^i_m$ formulae in which the higher-order variables of all orders up to $i+1$ have maximal-arity at most r, also induce a proper hierarchy in each higher-order logic of order $i \geq 3$. It is not known whether the correspondent hierarchy in second-order logic is proper. Using the concept of finite model truth definitions introduced by M. Mostowski, we give a sufficient condition for that to be the case. We also study the complexity of the set $AUTOSAT(F)$ and show that when F is one of the prenex fragments Σ^1_m of second-order logic, it follows that $AUTOSAT(F)$ becomes a complete problem for the corresponding prenex fragment Σ^2_m of third-order logic. Finally, aiming to provide the background for a future line of research in higher-order logics, we take a closer look to the restricted second-order logic SO^ω introduced by Dawar. We further investigate its connection with the concept of relational complexity studied by Abiteboul, Vardi and Vianu. Dawar showed that the existential fragment of SO^ω is equivalent to the nondeterministic inflationary fixed-point logic NFP. Since NFP captures *relational* NP, it follows that the existential fragment of SO^ω captures *relational* NP. We give a direct proof, in the style of the proof of Fagin's theorem, of this fact. We then define formally the concept of relational machine with *relational oracle* and prove

the exact correspondence between the prenex fragments of SO^ω and the levels of the *relational* polynomial-time hierarchy. This allows us to stablish a direct connection between the *relational* polynomial hierarchy and SO^ω without using the Abiteboul and Vianu normal form for relational machines.

Acknowledgements

It is very difficult to express how much I owe to José María Turull-Torres in just a few lines. I have been truly fortunate to have him as supervisor. He taught me not only most of what I know of finite model theory, but also to put first things first, to be precise, to call things by their name, to be rigorous and to be honest. Frequently, he would come over to my office and invite me for a coffee to not only discuss my work, but also to chat about life. I always found after these discussions with him the strength and motivation needed to continue with this project. He helped me immensely in every aspect of my work. More importantly, he was always very careful to give me just the right amount of guidance that would allow me to evolve as a researcher by finding my own path and research questions. Apart from being such an exceptional supervisor, he also gave me the invaluable gift of his friendship for which I am eternally grateful.

I am also deeply grateful to my co-supervisor, Klaus-Dieter Schewe. He made it possible for me to pursue my academic dream in New Zealand. He is a very generous man. He supported our work in every possible way, yet he never asked anything in return. He has been a guide to me from both an academic and a personal perspective.

I want to express my profound gratitude to Lauri Hella, Robert Goldblatt and Sven Hartmann, for serving as examiners, for a careful reading of the dissertation, and for their valuable comments.

I am very grateful to Leszek Kolodziejczyk for his important and detailed observations after a careful reading of a preliminary version of the work in Chapters 4 and 5. His remarks were very helpful. I have also learnt a lot from his excellent thesis. I thank Michal Krynicki for first drawing my attention to it.

I would also like to thank the anonymous referees for the Annals of Mathematics and Artificial Intelligence - their valuable observations have greatly helped me to shape the final version of fundamental parts of this dissertation.

I would like to thank all members of the Information Science Research Centre - it has been an extraordinary group to work with. A special thanks is due to the head of the department of Information Systems, Chris Freyberg, and to the secretaries, Merle Richardson and Julie Lyons. To Chris because of his constant support, and to Merle and July because they always had a solution for my problems.

My thanks also goes to the members of the doctoral subcommittee of the department of computer sciences at the University of Buenos Aires, especially to Alejandro Ríos. After careful evaluation of my application for admission as PhD student, they recommended me to take a paper on First Order Logic and Computability Theory taught by Roberto Cignoli, and a paper on Theory of Languages taught by Juan Miguel Santos. I benefited greatly from taking those papers and it was then that I developed my interest in the fascinating world of mathematical logic. I also thank Alejandro Vaisman who, without knowing me, accepted to act as my PhD advisor during the time I spent in the University of Buenos Aires.

My Friends in Argentina have always been with me through these years. They never complained for the long periods of silence on my part. Thanks are due to them for receiving me each time I went back and making me feel as if I never left. A very special thanks goes to Jorge Peri for putting me in contact with José María when I was starting my PhD. Jorge was the most extraordinary professor that I had during all my years as an undergraduate student and I will be always in debt to him for that. Also a special thanks goes to Stella Perillo and Fernando Lorge. When I decided to start my PhD, I had the good fortune of being working with them in my paid job. They strongly supported me in every imaginable way during that time. I also thank Gustavo Croch who has always been there when I needed a hand, and to Fernando Bordignon whom I admire for his enthusiastic and honest attitude toward research, which has been an important source of motivation for me.

I thank my dear friends, Peter and Mercedes Baxter and Raquel Direnzo, for all their support. They always made me and my partner feel at home in New Zealand and adopted us as part of their family here.

I thank my mother, Antonia, who encouraged me to pursue my dreams.

Last but not least, I am deeply in debt with my partner Lorena. I could not have done this work without her love, constant encouragement and support during all these years. She and our newborn son Luca are the two best things that ever happened to me.

Contents