

ACKNOWLEDGMENTS

This dissertation marks the end of a long and eventful journey, which began in a rural area of North-Central Vietnam. There, my parents made tremendous sacrifices to ensure that I had a good education. For this and much more, I am forever in their debt.

After high school, I moved to Hungary, where I met Aiviet Nguyen. We became good friends, and he encouraged me to go to America for my future studies. At the time, there was no diplomatic relationship between the U.S. and Vietnam, the iron curtain in Europe had just barely fallen, and the idea sounded impossible. Nevertheless, I made a try. Fortunately, Dr. Judith Ladinsky at the University of Wisconsin-Madison introduced me to Peter Haddawy, who then took me under his wing. It was with Peter that I first learned how to do research and how to write. He was a fantastic advisor who was warm and caring, and has remained supportive to me until these days. Thank you, Peter, I am very grateful.

After a Masters degree with Peter, I moved to the Ph.D. program at the University of Washington, where I have been extremely lucky to work with several superb advisors: Steve Hanks, Alon Halevy, and Pedro Domingos. They always blow me away with their amazing sharpness, technical depth, knowledge, and communication skills. I learned a lot from them. I especially thank Steve for staying extremely supportive of my work, and for ensuring my continuous funding for the first five years – it helped enormously by allowing me to focus on my studies. I owe a special debt to my two main advisors, Alon and Pedro. Thank you for your guidance on this research topic, and on research and life in general. You are simply the best advisors, period. I thank Pedro for teaching me all that I know about machine learning, for being patient with my questions, and for the many lessons on writing. I thank Alon for guiding my entrance into the database world, for being warm and caring, for putting in encouraging words when I felt down, and for not raising the price you charged for each needless word I used in our papers.

I am grateful to Oren Etzioni, for working with Alon and me in the early state of this research, and for valuable comments later on. I am also extremely grateful to Phil Bernstein, for his feedback during the course of this research, for his support for my research career, and for being on my supervisory committee. I also thank Erhard Rahm for invaluable feedback on part of this research.

This research also benefited tremendously from the many friends at UW. Special thanks to Jayant Madhavan, for countless hours spent discussing schema matching and for your help on the GLUE project. Thank you, Geoff “Squid” “Ba Meo” Hulten and Matt “Octopus” “Ma Do” Richardson, you are ideal office mates. I’m glad I could partially repay your friendship by giving each of you *two* nicknames. Thank you, Oren Zamir, Fred Pighin, Vass Litvanov, Sujay Sparekh, Omid Madani, Adam Carlson, Matthai Phillipose, Markus Mock, and Dave Hsu, for your friendship and support over the years. I owe many thanks to Zack Ives and Rachel Pottinger, fellow “pioneer” database students at UW. I’m indebted to numerous members of the database and AI groups: Dan Weld, Dan Suciu, Corin Anderson, Tessa Lau, Steve Wolfman, Igor Tatarinov, Pradeep Shenoy, Todd Millstein, and many more, for your feedback and support.

As I write these lines, my wife and son have left for Vietnam several days ago. Their absence makes me realize more than ever how much they mean to me. They are the semantics behind all that I do. It is to them that I dedicate this dissertation.