Preface

I use the term *logical and relational learning* to refer to the subfield of artificial intelligence, machine learning and data mining that is concerned with learning in expressive logical or relational representations. It is the union of inductive logic programming, (statistical) relational learning and multi-relational data mining, which all have contributed techniques for learning from data in relational form. Even though some early contributions to logical and relational learning are about forty years old now, it was only with the advent of inductive logic programming in the early 1990s that the field became popular. Whereas initial work was often concerned with logical (or logic programming) issues, the focus has rapidly changed to the discovery of new and interpretable knowledge from structured data, often in the form of rules, and soon important successes in applications in domains such as bio- and chemo-informatics and computational linguistics were realized. Today, the challenges and opportunities of dealing with structured data and knowledge have been taken up by the artificial intelligence community at large and form the motivation for a lot of ongoing research. Indeed, graph, network and multi-relational data mining are now popular themes in data mining, and statistical relational learning is receiving a lot of attention in the machine learning and uncertainty in artificial intelligence communities. In addition, the range of tasks for which logical and relational techniques have been developed now covers almost all machine learning and data mining tasks. On the one hand these developments have resulted in a new role and novel views on logical and relational learning, but on the other hand have also made it increasingly difficult to obtain an overview of the field as a whole.

This book wants to address these needs by providing a new synthesis of logical and relational learning. It constitutes an attempt to summarize some of the key results about logical and relational learning, it covers a wide range of topics and techniques, and it describes them in a uniform and accessible manner. While the author has tried to select a representative set of topics and techniques from the field of logical and relational learning, he also realizes that he is probably biased by his own research interests and views on the

field. Furthermore, rather than providing detailed accounts of the many specific systems and techniques, the book focuses on the underlying principles, which should enable the reader to easily get access to and understand the relevant literature on logical and relational learning. Actually, at the end of each chapter, suggestions for further reading are provided.

The book is intended for graduate students and researchers in artificial intelligence and computer science, especially those in machine learning, data mining, uncertainty in artificial intelligence, and computational logic, with an interest in learning from relational data. The book is the first textbook on logical and relational learning and is suitable for use in graduate courses, though it can also be used for self-study and as a reference. It contains many different examples and exercises. Teaching material will become available from the author's website. The author would also appreciate receiving feedback, suggestions for improvement and needed corrections by email to luc.deraedt@cs.kuleuven.be.

The book starts with an introductory chapter clarifying the nature, motivations and history of logical and relational learning. Chapter 2 provides a gentle introduction to logic and logic programming, which will be used throughout the book as the representation language. Chapter 3 introduces the idea of learning as search and provides a detailed account of some fundamental machine learning algorithms that will play an important role in later chapters. In Chapter 4, a detailed study of a hierarchy of different representations that are used in machine learning and data mining is given, and two techniques (propositionalization and aggregation) for transforming expressive representations into simpler ones are introduced. Chapter 5 is concerned with the theoretical basis of the field. It studies the generality relation in logic, the relation between induction and deduction, and introduces the most important framework and operators for generality. In Chapter 6, a methodology for developing logical and relational learning systems is presented and illustrated using a number of well-known case studies that learn relational rules, decision trees and frequent queries. The methodology starts from existing learning approaches and upgrades them towards the use of rich representations. Whereas the first six chapters are concerned with the foundations of logical and relational learning, the chapters that follow introduce more advanced techniques. Chapter 7 focuses on learning the definition of multiple relations, that is, on learning theories. This chapter covers abductive reasoning, using integrity constraints, program synthesis, and the use of an oracle. Chapter 8 covers statistical relational learning, which combines probabilistic models with logical and relational learning. The chapter starts with a gentle introduction to graphical models before turning towards probabilistic logics. The use of kernels and distances for logical and relational learning is addressed in Chapter 9, and in Chapter 10 computational issues such as efficiency considerations and learnability results are discussed. Finally, Chapter 11 summarizes the most important lessons learned about logical and relational learning. The author suggests to read it early on, possibly even directly after Chapter 1.

An introductory course to logical and relational learning covers most of the materials in Chapters 1 to 4, Sects. 5.1 – 5.4, 5.9, and Chapters 6 and 11. The other chapters do not depend on one another, and, hence, further chapters can be selected according to the interests and the preferences of the reader. Given the interests in statistical relational learning, the author certainly recommends Chapter 8. Advanced sections and exercises are marked with a * or even with **. They are more challenging, but can be skipped without loss of continuity.

This book could not have been written without the help and encouragement of many persons. The author is indebted to a number of co-workers who contributed ideas, techniques, surveys and views that have found their way into this book, including: Maurice Bruynooghe for influencing the use of logic in this book and numerous suggestions for improvement; Hendrik Blockeel, Luc Dehaspe and Wim Van Laer for contributions to the upgrading methodology described in Chapter 6, Kristian Kersting for joint work on statistical relational learning presented in Chapter 8, and Jan Ramon for his work on distances in Chapter 9. This book has also taken inspiration from a number of joint overview papers and tutorials that the author delivered in collaboration with Hendrik Blockeel, Sašo Džeroski, Kristian Kersting, Nada Lavrač and Stephen Muggleton. The author would also like to thank the editor at Springer, Ronan Nugent, for his patience, help, and support during all phases of this book-writing project.

The author is grateful for the feedback and encouragement on the many earlier versions of this book he received. He would like to thank especially: the reading clubs at the University of Bristol (headed by Peter Flach, and involving Kerstin Eder, Robert Egginton, Steve Gregory, Susanne Hoche, Simon Price, Simon Rawles and Ksenia Shalonova) and at the Katholieke Universiteit Leuven (Hendrik Blockeel, Björn Bringmann, Maurice Bruynooghe, Fabrizio Costa, Tom Croonenborghs, Anton Dries, Kurt Driessens, Daan Fierens, Christophe Costa Florencio, Elisa Fromont, Robby Goetschalkx, Bernd Gutmann, Angelika Kimmig, Niels Landwehr, Wannes Meert, Siegfried Nijssen, Stefan Raeymaekers, Leander Schietgat, Jan Struyf, Ingo Thon, Anneleen van Assche, Joaquin Vanschoren, Celine Vens, and Albrecht Zimmermann), several colleagues who gave feedback or discussed ideas in this book (James Cussens, Paolo Frasconi, Thomas Gaertner, Tamas Horvath, Manfred Jaeger, Martijn van Otterlo), and the students in Freiburg who used several previous versions of this book.

Last but not least I would like to thank my wife, Lieve, and my children, Soetkin and Maarten, for their patience and love during the many years it took to write this book. I dedicate this book to them.

Destelbergen, May 2008 *Luc De Raedt*

http://www.springer.com/978-3-540-20040-6

Logical and Relational Learning
De Raedt, L.
2008, XV, 387 p., Hardcover
ISBN: 978-3-540-20040-6