

---

# Preface

To be able to forecast future events, science wants to infer general laws and principles from particular instances. This process of inductive inference is the central theme in statistical modeling, pattern recognition, and the branch of computer science called “machine learning.” The minimum description length (MDL) principle is a powerful method of inductive inference. It states that the best explanation (i.e., model) given a limited set of observed data is the one that permits the greatest compression of the data. Put simply, the more we are able to compress the data, the more we learn about the regularities underlying the data.

The roots of MDL can be traced back to the notion of *Kolmogorov complexity*, introduced independently by R.J. Solomonoff, A.N. Kolmogorov, and G.J. Chaitin in the 1960s. These and other early developments are summarized at the end of Chapter 1 of this book, where a brief history of MDL is presented. The development of MDL proper started in 1978 with the publication of *Modeling by the Shortest Data Description* by J. Rissanen. Since then, significant strides have been made in both the mathematics and applications of MDL. The purpose of this book is to bring these advances in MDL together under one cover and in a form that could be easily digested by students in many sciences. Our intent was to make this edited volume a source book that would inform readers about state-of-the-art MDL and provide examples of how to apply MDL in a range of research settings.

The book is based on a workshop we organized at the annual Neural Information Processing Systems (NIPS) conference held in Whistler, Canada in December 2001. It consists of sixteen chapters organized into three parts. Part I includes six introductory chapters that present the theoretical foundations of the MDL principle, its various interpretations, and computational techniques. In particular, chapters 1 and 2 offer a self-contained tutorial on MDL in a technically rigorous yet readable manner. In Part II, recent theoretical advances in modern MDL are presented. Part III begins with a chapter by J. Comley and D. Dowe that describes minimum message length (MML), a “twin sister” of MDL, and highlights the similarities and differences between these two principles. This is followed by five chapters that showcase the application of MDL in diverse fields, from bioinformatics to machine learning and psychology.

We would like to thank our editor, Bob Prior, for the support and encouragement we received during the preparation of the book. We also thank Peter Bartlett, Alex Smola, Bernhard Schölkopf, and Dale Schuurmans for providing  $\LaTeX$ -macros to facilitate formatting and creation of the book. We also thank the authors for

contributing papers and the referees for reviewing the manuscripts. Finally, we thank our families for putting up with our foolishness.

Peter Grünwald, In Jae Myung, Mark A. Pitt  
Amsterdam and Columbus, Ohio, February 2004