# Finding the Real Patterns

Geoffrey Webb

Clayton School of Information Technology,
P.O. Box 75 Monash University,
Victoria 3800, Australia

**Abstract.** Pattern discovery is one of the fundamental tasks in data mining. Pattern discovery typically explores a massive space of potential patterns to identify those that satisfy some user-specified criteria. This process entails a huge risk (in many cases a near certainty) that many patterns will be false discoveries. These are patterns that satisfy the specified criteria with respect to the sample data but do not satisfy those criteria with respect to the population from which those data are drawn. This talk discusses the problem of false discoveries, and presents techniques for avoiding them.

## References

Webb, G.I.: Discovering significant rules. In Ungar, L., Craven, M., Gunopulos, D., Eliassi-Rad, T., eds.: Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD-2006, New York, The Association for Computing Machinery (2006) 434 – 443.

Webb, G.I.: Discovering significant patterns. Machine Learning (in-press).

## Biography

Geoff Webb holds a research chair in the Faculty of Information Technology at Monash University. Prior to Monash he held appointments at Griffith University and then Deakin University where he received a personal chair. His primary research areas are machine learning, data mining, and user modelling. He is widely known for his contribution to the debate about the application of Occam's razor in machine learning and for the development of numerous algorithms and techniques for machine learning, data mining and user modelling. His commercial data mining software, Magnum Opus, is marketed internationally by Rulequest Research. He is editor-in-chief of the highest impact data mining journal, Data Mining and Knowledge Discovery and a member of the editorial boards of Machine Learning, ACM Transactions on Knowledge Discovery in Data, and User Modeling and User-Adapted Interaction.