Preface to UMUAI Special Issue on Machine Learning for User Modeling

It can be argued that every interactive software system utilizes a user model, albeit, in many cases, an implicit model of the user's objectives, and capabilities. Rather than such implicit models, research on user modeling has concentrated on explicit models that provide some form of assessment of specific attributes of the user. There are three main ways in which the content of such a model might be generated and maintained.

 – it might be specified by an external source, either through pre-session config-uration, or by externally specified update;
 – it might be specified by the user; or
 – it might be specified by the software, usually on the basis of observation of the user's performance.

The preferences settings to be found in many modern software packages serve to illustrate the first two methods. Such preferences are usually pre-set to reflect the manufacturer's assumptions about the software's community of users. The user may then modify them to suit his or her individual requirements. Much of the research in user modeling, however, has concentrated on the third method by which the software forms the user model.

Formation of a user model by observation of the user's actions usually involves a process of induction. The system infers a model of whatever aspects of the user are of interest – such as preferences, objectives, skills and aptitudes – from its observations of the user. Automated induction, such as this, has been extensively studied under the name 'machine learning'. This special issue of User Modeling and User-Adapted Interaction brings together a collection of papers presenting a wide variety of machine learning techniques and their use in a diverse range of user modeling applications.

Three of the papers use approaches in which the user model has a predefined structure. The task of the modeling system is to infer appropriate values for the various variables within the model.

For both Albrecht et al and Gymtrasiewicz et al these variables represent prob-abilities. Both use forms of Bayesian update to infer appropriate values for these variables. However, the types of models to which these Bayesian approaches are applied differ greatly. Albrecht et al use a Dynamic Belief Network, which supports the combination of evidence from multiple observed independent variables in order to assign appropriate levels probability to the potential values of the dependent

variables of interest. In their case the independent variables are a user's individual actions and the dependent variable is the ultimate objective of the sequence of actions being observed. Gymtrasiewicz et al develop models of multiple interacting agents. Each agent models each other agent's beliefs, desires, intentions and capabilities. These models are further complicated by the ability to include in agent A's model of agent B, a model of agent B's model of agent A. This may in turn, recursively, include agent B's model of agent A's model of agent B, and so on.

Balabanovic uses gradient descent parameter tuning to infer suitable values for the parameters to a model of a user's preference rankings in the context of the very topical subject of text recommendation on the world wide web. This paper examines the important issue of how a system that employs a user model to select appropriate system actions should manage the conflicting objectives of selecting actions that best satisfy the existing model, and selecting actions that best support refinement of the existing model.

The remaining papers explore machine learning techniques that infer both the appropriate structure and parameters for a model.

Sison et al use conceptual clustering to form bug descriptions when modeling student programming errors. From analysis of incorrect Prolog programs, their system generates a set of error classes, where each class represents a specific combination of underlying misconceptions and other knowledge errors.

Each class is represented by a characterization of discrepancies between the ideal solution to a task and the solution that will be generated in the presence of the error class.

Chiu and Webb use decision tree learning for modeling subtraction skills. They present and compare a series of techniques for increasing the numbers of predictions made by the FBM-C4.5 modeling system. The models of the initial system take the form of a set of decision trees, where each tree makes predictions about a specific aspect of a future actions from a description of the context in which the action will be performed. Most of the techniques examined vary in terms of how the disparate predictions are combined to create a specific single action prediction, although techniques are also considered for forming a single tree and for generating and combining predictions from multiple models each of the initial form.

The issue includes papers that variously present the subjects of the models formed as 'agents', 'users', and 'students'. It is worth clarifying why papers on agent and student modeling are included in an issue on machine learning for user modeling. Agents, users and students form part of a generalization hierarchy. Students are a type of user and users are a type of agent. Irrespective of the application domains utilized in the various papers, all of the techniques presented have the potential for broad applicability in a variety of user modeling contexts.

There is a long history of distinguished applications of machine learning for user modeling. The contributions to this issue show that this tradition has reached a stage of considerable maturity and sophistication.

Author's Vita

Dr G. Webb

Deakin University, School of Computing and Mathematics, Geelong, Vic., Australia, 3217. Home page: http://www.cm.deakin.edu.au/ webb

Dr Geoff Webb is a Reader in Computing at Deakin University where he has established and leads the Deakin University Knowledge Acquisition and Processing research group. He received his B.A. and Ph.D. degrees in Computer Science from La Trobe University. The author of more than sixty technical papers, he is active in several areas of artificial intelligence research, including machine learning, knowledge acquisition, agent modelling, and artificial intelligence in education.