

**IJCNN
INTERNATIONAL
JOINT CONFERENCE
ON NEURAL NETWORKS**

**Baltimore, Maryland
June 7-11, 1992**

IEEE Catalog Number: 92CH3114-6
ISBN: Softbound Edition 0-7803-0559-0
 Casebound Edition 0-7803-0560-4
 Microfiche Edition 0-7803-0561-2
Library of Congress Catalog No. 91-59048

Modeling Users with Neural Architectures

Qiyang Chen and Anthony F. Norcio
Department of Information Systems
University of Maryland, Baltimore County
Baltimore, MD 21228

Abstract *This paper proposes and discusses a research framework on building user model system by utilizing artificial neural network (ANN) approaches. First some problems in user modeling are discussed which underlie the motivations of introducing ANN approaches. Secondly some considerations on ANN properties and their applications in task-related user modeling are presented. Finally an ANN-based, integrated user modeling system is proposed which incorporates conventional symbolic reasoning approaches in a multilevel processing environment.*

1 Introduction

User modeling, as a mechanism to capture the individual characteristics (goals, plans, beliefs, background knowledge, etc.) in a task environment, has been recognized as an important component in adaptive interface systems [14]. It helps systems tailoring their responses to individual users in task performance. Interfaces equipped with user models not only exhibit a wide range of cooperative functionalities, but also improve system efficiency and effectiveness.

Given a specific application domain, the sources of knowledge embedded in a user model may be of different importance and significance. Thus the current user modeling research exhibits great varieties. Several types of dimensional settings have been proposed to classify knowledge in user model [16][7][3]. Essentially these types are all consistent with each other in the sense that they attempt to characterize the knowledge about a user in terms of its persistency or temporal extent; the way of its elicitation and representation; and the degree of specification, as well as the relationship between these dimensions in model construction.

2 Some Problems in User Modeling

Although user modeling research has made great progress, some issues regarding both theoretical and practical aspects need to be explored further.

Feature Identification: First, in order to embody individual differences into user models, one needs to identify and characterize the significance of stable and unstable information about a user within the context of task. The hierarchical stereotyping approach [16] provides a simple way of identifying users by assigning them to one or more stereotypes along with some belief values. But it fails to update dynamically the attributes within a stereotype when they no longer correctly represent corresponding concepts or lose the significance as important characteristics affecting task performance. In addition, it provides no insight with respect to the context of task and underlying goals.

Default reasoning: Unlike expert systems, where the expertise is available toward a relatively narrow domain of discourse, a user modeling system usually lacks the support of expert knowledge[4]. In order to make a large number of inferences with a small number of observations, extensive default assumptions are established, which may not be true during the interaction. Therefore, the techniques for handling incomplete information and nonmonotonic reasoning are involved in modeling process. However, since the truth maintenance and the ad hoc approaches examine evidence one piece at a time, they may fail to prevent the model construction from the effects of noisy or inconsistent inputs that should be ignored. Thus the system decisions may not reflect the real change and the effort of maintaining consistency may bring some new conflicts. This greatly affects system stability, efficiency and effectiveness. The primary issue here is that information regarding a user's behavior should be viewed by pattern rather than fragmented pieces during the interaction and the historical data relevant

to interactive sessions should be utilized for modeling and prediction [6].

Managing uncertainties: By using conventional reasoning systems (e. g. production systems), incorrect conclusions may be drawn due to errors of input data which may stem from many sources of uncertainty [2]. It may be difficult to specify correctly the belief value of rules. The available expertise may be able to identify the correlation between observed behavior data and its implications or final hypothesis, but may be unclear with respect to the probabilistic assignment and some important intermediate concepts or conclusions that support inferences. In addition, even with a well understood situation, it is still possible to fail in maintenance of probabilistic formalisms (e. g. independence) over the whole production system.

Learning Learning in user modeling is another crucial issue for improving the adaptivity of a user model. It stems from two aspects: first, it is difficult to predefine an accurate set of knowledge about system's user community, and secondly the characteristics (familiarity with system, beliefs, goals and preferences, etc.) about both single or a class of user may change over time. Some techniques for handling uncertainties and domain dependent heuristics may be useful. The adaptation of user model usually involves the revisions such as modification of belief values, deletion or creation of evidences or rules. These revisions may in turn cause further data generalization and specification.

3 Implementations of User Modeling on Neural Networks

In artificial neural network (ANN) based user modeling the collection of a user's information is conceptually viewed as an individual, task-related "user image". Naturally, pattern recognition techniques can be used to process this image. In order to provide a complete and accurate understanding with respect to users and tasks, a user modeling system should have the ability of fault tolerance, graceful degradation and signal enhancement. Obviously, these are the properties of ANN techniques. In addition, with its strong capability of experience based learning, ANN can deduce the relationship between the user input and the output regarding the beliefs, goals and plans in the absence of the knowledge. The historical sessions are recorded as "log" and then sampled for initial training or retraining when necessary.

3.1 Some ANN Primitives

In user modeling, the following neural operation primitives (i.e. the generic functions represented by ANN) can be utilized to enhance the process:

Pattern classifier: It is able to judge whether a pattern of information about a user is a member of certain class. It is useful in stereotype assignment or classification of user's level of expertise with respect to certain task. Given stimuli input such as user's direct responses, inquiries and the ways he/she navigates a dialogue, etc., a preliminary pattern of a user's image is formed. The pattern classifier can produce a set of belief values corresponding to a set of stereotypes which a individual user is assigned. This assignment along with the distribution of believe values could be expected to be consistent due to ANN's properties. In contrast, rule based assignment is much more difficult to maintain the consistency and less effective. Feed-forward network trained by back propagation algorithm is an appropriate paradigm of pattern classifier.

Pattern associator: It captures the associations between input and output patterns despite incomplete or corrupted inputs. Also, an adaptive version can extract the ideal from a set of noisy exemplars which may occur in the user modeling process. Among various paradigms of associate memory, bidirectional associative memory (BAM) [11] is developed for learning pattern association. It accepts an input vector on one set of units and produces a related but different output vector on another set. Kohonen's LVQ framework for self-organizing auto-associate memory [10] is also very powerful in learning how to complete partial patterns of user input, based on the memory of pattern seen in the past. Both of the paradigms are especially useful in recognizing user-task context.

Considering a state transition diagram that predefines a user-task model, it can be easily transformed

into stored pattern in associative memory in which each state is represented by a neuron and the transition is mapped to weighted connections. During the real time interaction, a user's operations involving several nodes can be organized as an stimuli input vector. The associative memory can produce relatively complete path of state transitions as a prediction to navigate user toward task accomplishment. This provides a basis for comparing the user's real action with the predicated one and then modifying the models. This property can also be used in plan recognition while several traits are observed. For example in statistical database retrieval environment, if the path of several inquiries on some overlapped data fields is recognized, the assumptions about the presumable goals behind these operation can be made for preventing compromise individual data.

In addition, transforming a semantic network into associative memory provides a way to model what user knows. Considering that a user's problem solving could be viewed as a procedure of associative thinking, the representation power of a user's model could be enhanced from two aspects: first the adaptation can be conducted by adjusting the pattern of weights to reflect the real time change of user's mental model; and secondly the network is self-processing in the sense that it exhibits global system behaviors derived from concurrent local interaction of its numerous components[15]. In modeling process, a set of inputs from a user will trigger some assumptions about what the user already knew and then the corresponding nodes are fired. The spreading activation rule can be used to create an intersection search with certain conditions, while some form of activations propagates in a breath-first fashion until a set of nodes of intersection occurs between the expanding regions of action [1]. The set of those nodes reaching highest activation level specifies a possible range of concepts known to a user. The activation and propagation rules are usually ad hoc and the activated nodes compete each other in a multiple-winners-take-all fashion.

Regularity Detector: It classifies an input pattern based on pattern dissimilarities rather than predefined categories. Introducing this primitive in user modeling is very useful for two reasons: First the network is exposed to a constantly changing environment, it may never see the same training vector twice due to the nature of individualization in an interaction. Too often, learning a new pattern of user inputs impairs the previous training effect. Secondly, it is difficult to predetermine the type or number of user's characteristics as significant factors in model construction. The Adaptive Resonance Theory (ART) system [5] addresses the conflict in learning between adaptability and stability by temporarily shutting down learning when a large mismatch occurs between the input and projected output. It uses feedback from output to input to ensure a consistent categorization with no off-line training time required. This property can be used for creating a new stereotype and determining the significance of the attributes among all alternatives based on the regularities in the distribution of data types, meanwhile preventing the modification of patterns that have been learned previously.

3.2 Mapping Production System onto ANN : Capture the Uncertainties

The issues of mapping production system to an ANN, which is usually a feed-forward network trained by back propagation, has been proposed [7][12]. One of the motivations is to capture the uncertainties in knowledge acquisition. Since user modeling lacks the knowledge about relationship between the user's inputs and their implications on user's various characteristics and goals, ANN learning paradigms can be used to automatically obtain the belief values of these implications. Given the domain concepts and a set of exemplars from user's input, after training a set of certainty factors can be derived in terms of the combination of weights on relevant connections. The inference under the uncertainty becomes a process of combining the weights and propagating the activations throughout an ANN.

It is suggested that intermediate assertions in the inference be represented by the middle layers of a network. Thus an issue is raised of how to identify the semantics of each node in the hidden layers. The simple one-to-one mapping may lack the generality of use. The semantics of the nodes in hidden layer is not as clear as that of intermediate clauses in production system because of the highly correlated activities in network dynamics. It is possible to use object oriented mapping approach that encapsulates various network dynamics (i. e. directed connections, combining, activating and learning functions, etc.) as objects. However it still needs further justifications on the validity with respect to the complexity, convergency and semantic operations. One of the alternatives is, as proposed in [6], to subdivide an ANN into several submodules. Each module

accomplishes a subtask at an appropriate semantic level. Thus the concepts that summarize or categorize intermediate assertions are mapped to output layers of the submodules. This approach affords several advantages: Each module can be utilized separately for multiple purposes. Training process could be conducted for either single module or the whole ANN system, which makes system adaptation simple and flexible. It provides a clear insight into knowledge representation and inference, which enhances the capability of explanation.

3.3 An Integrated Framework of a User Modeling System

The integrated system framework proposed is shown in Figure 1. It functions as an adaptive interface in information retrieval system. The ANN techniques and production systems are incorporated in a multilevel processing environment. It is acknowledged that ANN cannot accomplish all tasks of user modeling. Symbolic reasoning system is still necessary to assist in preprocessing user-task images, analyzing and explaining ANN outputs. Meanwhile the declarative information in a user model needs to be stored in knowledge bases. The functions of each system component are briefly described below.

Dialogue Manager: As a front-end of the system, it plays major roles as follows:

(i) Identification of user input. It receives two types of user's input: user's direct input regarding a task (e. g. an inquiry on a database) or user's response to system questionnaire designed for modeling user or task (e.g. questions about user's general background on a presumed task). The task related input could be used to trigger the assumptions on task analysis and the input related to a user's characteristics are useful to build user profile. In most cases, users are allowed to control the modeling process in terms of participating or ignoring the system interview. A user can also only enter the most representative information in each category of questionnaire. Since ANN is robust in tolerating incomplete input, it can produce correct classification without complete input as long as the input characterizes the key information. This property can facilitate the stereotype based user modeling in which default reasoning is very much involved.

(ii) Mode analysis and information display. According to the modeling result, it designates a dialogue mode in which the provided information is individually tailored to a user.

(iii) It monitors ongoing sessions and coordinates other system components in terms of maintaining session log database and updating the assumptions stored in knowledge bases. Also, it delivers the original user input to the rest of system for processing.

Pre-processing: It is a production system that performs two functions. First, it transforms the input into appropriate numeric representation for each network paradigm, while a knowledge base containing the recommendations on transformation and normalization strategies is utilized. Secondly, it creates image vectors for ANN processing. Two types of images are necessary: user images that consist of various user's characteristics related to a task, and task images that consist of the traits of current dialogue paths in a session. At this stage, a limited default reasoning is required in assisting the extraction of features of both user's profiles and the tasks being performed. However, no further conclusion is made on how these features affect task performance and model construction. A rule's IF-part consists of the conditions satisfied by user input and its THEN-part is the action of triggering assumptions. Given the initial assumptions about a user and the tasks, some obviously inconsistent and noisy data could be filtered. Since user-system interaction is a gradually evolving procedure, the creation of both user and task images needs a period of dialogue. Preprocessing determines when the data is ready for ANN processing and what is needed from dialogue manager to form these images.

ANN processing: Once the image data is established, it is passed to several ANN modules for processing. As previously described, all network primitives could be implemented to serve different modeling purposes: The pattern classifier is used for stereotype assignment with the measure of beliefs. The pattern associator is used for task/goal detection and modeling what user knows within the context of task. The regularity detector is implemented to create new stereotypes or attributes and to determine the significance of attributes within a stereotype. For each primitive there are some alternative paradigms. The types of the ANN paradigms for the implementation of primitives depend on the application environment. In addition, it is suggested to introduce some paradigm as a functional module that synthesizes the outputs from other paradigms that process user/task images and produces an indication vector for system response in current dialogue. For example, in information

retrieval system the indication vector could be one that implies the depth and width of a database search.

Post-processing: The next level of processing is result analysis and explanation. Some numeric outputs of ANN need to be transformed into symbolic representations or linguistic variables. The knowledge bases that store various assumptions need to be updated according to the output of ANN processing. The knowledge of transformation and modification is represented by a set of production rules. The indication vector of system response should be also interpreted in terms of transforming numeric representations to semantic actions (e. g. a set of inquiry statements). After post-processing not only the models of user and task are established but also the actions that the system adopts are defined. This fulfills an adaptive activity.

Information Retrieval: This module is considered as a conventional retrieval system based on which user modeling system is established. From a user's perspective, the interaction is with this module, without realizing the existence of user modeling process. The output of post-processing forms a set of inquiries that evokes information retrieval activities. This set of inquiries may not be same as the original input from user but is expected to better reflect what a user really wants. Since this module treats the inquiries from post-processing just as they come from user's direct input, it could be independently developed. Thus it is possible that the user modeling component be appropriately developed so that it can be simply added to a previously developed information retrieval system.

Training Manager: Training is crucial in any ANN-based system. This module provides an interface to support off-line data acquisition and network parameterization. The initial training can start after training data is available, which is either sampled from log database or obtained from external data acquisition process. The training set must cover a sufficient variety so that the networks can make correlations for unfamiliar cases. This variety is embodied through the different ranges of characteristics of both users and the tasks they perform. Once the systematic observation shows that the ANN output (e. g. database search indication vector) does not serve the dialogue well, which is surveyed by dialogue manager in terms of targeting rate or some other metrics, the part of or whole ANN system may need to be retrained. Generally, retraining has no difference with initial training in process. The training process is mainly controlled by human expert since training management is still an art at present.

Knowledge Bases: There are several types of knowledge stored in system knowledge bases:

- . Declarative assumptions or stereotypes about users, tasks and their relationships (i. e. user-task metaphors).
- . Meta-knowledge of modeling, strategies for triggering the assumption about users, tasks and user-task metaphors.
- . Meta-knowledge of triggering system response such as the strategies of directing dialogue from one mode to another, according to the analysis of user-task metaphors.
- . Strategies of data transformation and normalization.

4 Conclusion

Feature identification, default reasoning and uncertainty management, as well as learning ability, are major aspects in determining the performance of user modeling system. The suggested ANN primitives have shown much promise for enhancement in these aspects. The proposed system framework provides a flexible platform for incorporating ANN approaches and conventional reasoning systems to user modeling, which constitutes a multilevel modeling process. Its implementation is now undergoing in a database retrieval system. Further research is aimed at verifying that this framework can scale to large, more complex information retrieval environment in which various ANN paradigms are investigated in terms of their efficacy in modeling users and context of the task.

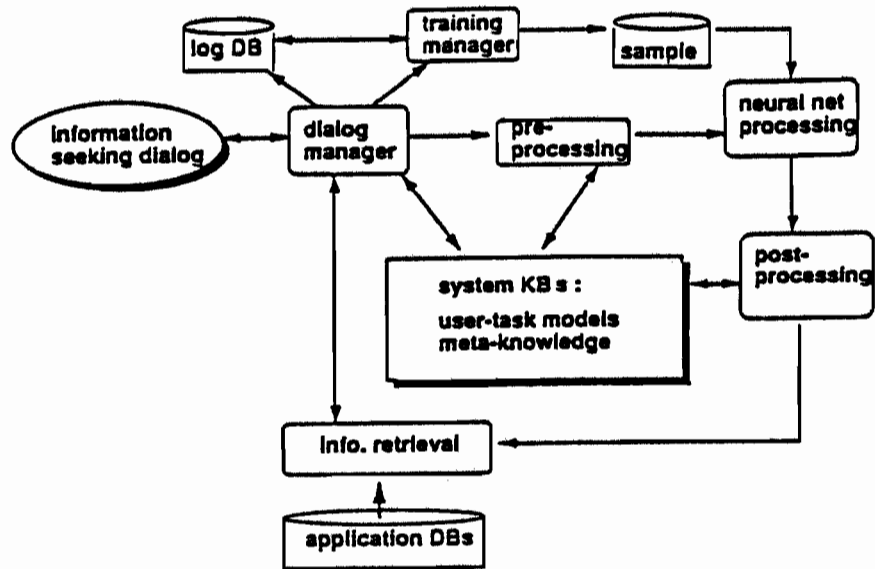


Figure 1: An ANN-based user modeling process

References

1. Anderson, J. R., "Retrieval of Information from long term Memory," *Science*, Vol. 220, pp.25-30, (1983).
2. Bonissone, P. and Tong, R. M., "Editorial: Reasoning with uncertainty in expert systems," *International Journal of Man-Machine Studies*, Vol. 30, 69-111 (1985)
3. Brajnik, G., Guida, G., and Tasso, C., "User modeling in expert man-machine interfaces : A case study in intelligent information retrieval," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 20, No. 1 pp 166-185 (1990)
4. Card, S. K., Morman, T. P. and Newell, A., *The psychology of human-computer interaction* (1983)
5. Carpenter, G. A. and Grossberg, S., "A massively parallel architecture for self-organizing neural pattern recognition machine," *Computer Vision, Graphics and Image Processing*, Vol. 37, pp. 54-115 (1987).
6. Chen, Q. and Norcio, A. F., "A neural network approach to User modeling," *Proceedings of 1991 IEEE International Conference on Systems, Man, and Cybernetics*, pp.1429-1434 (1991).
7. Finin, T. W., "GUMS - A general user modeling shell," in *User Models in Dialogue Systems*, ed. A. Kobsa and W. Wahlster, Springer-Verlag, Berlin (1989).
8. Fu, L., "Integration of neural heuristics into knowledge-based inference," *Connection Science*, Vol. 1, No.3, (1990)
9. Gallant, S. I., "Connectionist expert systems," *Communication of ACM*, Vol. 24, No. 2, pp 152-169 (1988)
10. Kohonen, T. *Self-Organization and Associative Memory*, 2nd ed., New York:Springer-Verlag, (1988)
11. Kosko, B., "Adaptive bidirectional associative memories," *Applied Optics*, Vol. 26, pp. 4947-60 (1987).
12. Kuncicky, D., "The transmission of knowledge between neural networks and expert systems," *Proc. of the WNN-AIND Workshop on Neural Networks* (1989).
13. Maren, A. J., "Neural network for enhanced Human-computer interaction," *IEEE control systems*, August, 1991
14. Norcio, A. F. and Stanley, J., "Adaptive human-computer interfaces, A literature survey and perspective," *IEEE Transactions on System, Man and Cybernetics*, Vol. 19, No. 2, pp. 399-408 (1989)
15. Reggia, J. A. and Sutton, G. G., "Self-processing networks and their biomedical implications," *Proceedings of IEEE*, Vol. 76, No. 6, pp. 680-692 (1988)
16. Rich, E. "Users are individuals: Individualizing user models," *International Journal of Man-Machine Studies* Vol. 18, No. 3 (1983)