HUMAN PERFORMANCE IN AUTOMATED SYSTEMS: CURRENT RESEARCH AND TRENDS

Edited by

Mustapha Mouloua, Ph.D.
Cognitive Science Laboratory
The Catholic University of America
Washington, DC 20064

Raja Parasuraman, Ph.D.
Cognitive Science Laboratory
The Catholic University of America
Washington, DC 20064

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1994 Hillsdale, New Jersey Hove, UK

Recognizing User's Behavior Pattern through Associative Model

Qiyang Chen & A. F. Norcio

University of Maryland, Baltimore County

ABSTRACT

This paper presents a mechanism of modeling the user's characteristics in task performance. The associative network is utilized to store, maintain and reasoning task-related attributes of user. The causal relationships among the attributes are conditionally weighted. The modeling process extracts the assumptions and their weighted connections from the associative network to form a unique profile that fits a particular user or task. It is conducted by propagating the activation level throughout the network. We suggest that this approach can be expected to overcome some inherent problems of the conventional stereotyping approaches in terms of providing noise tolerance, pattern completion and learning capabilities. It can also avoid the complexity of truth maintenance in default reasoning which is required in previously known stereotyping approaches.

INTRODUCTION

User models have become important components of adaptive human-computer interfaces. It is recognized that a system will be able to exhibit flexible user-oriented behavior only if it has access to a model of user which consists assumptions about users' characteristics regarding certain task he or she is performing (Chen and Norcio 1993). These characteristics may vary depending on the types of tasks. Usually, they are related to a user's plans, goal, background knowledge as well as cognitive preferences. There are different taxonomies about these characteristics depending on the time period while they hold, the way by which their are elicited and represented as well as the degree by which they are specified.(Rich 1979).

It has been a common practice to use a set of predefined assumptions to initialize the system's beliefs about its users and the tasks they are performing. It is usually referred as stereotype approach (Clowes *et al.* 1985, Kobsa and Wahlster 1989, Rich 1979). The pre-defined stereotypical knowledge is organized into a generalization hierarchy in which the stereotypes inherit knowledge from their ancestors. The modeling process proceeds with a stereotype assignment in terms of default reasoning which allows the model to retain the stereotypical knowledge about a user in the absence of evidence to the contrary. This paper begins with a discussion on the problems of hierarchical stereotyping approaches, then presents an alternative approach which is based on the associative model.

STEREOTYPE HIERARCHY AND ITS LIMITATIONS

Generally, stereotype approach is established on a generalization hierarchy shown in Figure 1, in which each node in the hierarchy represents a stereotype regarding certain task. Each node can also come up with a set of attributes or assumptions that are particularly used to justify or describe the corresponding stereotype. A hierarchical user model S_h is a substructure of the generalization hierarchy S which can be represented as follows:

$S_h \quad S = \{ S_i \mid S_i : \text{a stereotype in the hierarchy} \}$

$S_i = \{ a_j^{(i)} \mid a_j^{(i)} : \text{an assumption in } S_i \}.$

Usually, the modeling process proceeds from the root of the hierarchy. If one or more assumptions are verified by certain facts or observation during a interaction, a stereotype that contains these assumptions is activated and then added into a current user model. As the interaction progresses, more and more stereotypes are activated and the user model become more informative. For example , in figure 1 if we define S1 as a computer user, then we may further define S2 as microcomputer user and S3 as mainframe user. Furthermore, we may specify S4 as a IBM PC user and S5 as Macintosh user respectively. If a user is modeled by systems as a PC user then his or her profile consists of all assumptions contained in S1, S2 and S4.

Although stereotype approach provides a simple way to initialize the modeling process and was successful in some applications, we belief that this approach limits the representation power of a user model in following aspects:

(i) Since the reasoning is conducted with extensive default assumptions that may conflict with the new evidence obtained as the interaction progresses, the revision of stereotypical knowledge is necessary to handle the inconsistencies. This in turn will cause the dependency-directed backtracking which is a non-monotonic process(Doyle 1979, de Kleer 1986). This process is called truth maintenance and has been a fertile field for AI research. Since the conventional truth maintenance approach examines one piece of evidence at a time, it is often ineffective or even impossible to detect noisy or inconsistent inputs that should be ignored (Pao 1989). This phenomenon, for instance, may be observed in a tutoring system where the student's input often carries some inconsistencies (typing mistake, grammar mistake or the way student raises the question, etc.) which may have nothing to do with understanding the assumption and therefore should be ignored in the modeling process.

In addition, the stereotyping approach must be able to examine the continuity and overall user behavior pattern in task performance. Otherwise, the modeling process may not reflect the real world situation and it is very possible that current effort of maintaining consistency may bring further conflicts in the subsequent interaction. Thus, model construction may fall into a dilemma where a non-monotonic process of conflict-resolution is frequently involved and even eventually no decision can be made after a period of interaction (Chen and Norcio 1992).

(ii) Hierarchical approach also provides a simple way to classify user into certain stereotypes. Since the user modeling must capture user's individual differences in order to tailor the system response to fit a particular user. Therefore, very often it requires this classification as personal as possible. But the hierarchical structure in the conventional stereotyping approaches may limit the degree of individualizing a user. Since the pre-defined attributes or assumptions are confined within each stereotype and can be only inherited by the descendant stereotypes, there is no effective way to update those attribute that are no longer significant in the context of task performance. Also, most stereotyping approaches solve conflicts by simply replacing the active stereotypes with their ancestors in the hierarchy (Rich 1979, Kobsa and Wahlster 1989). Thus some of the assumptions that are still consistent to current situation are lost. It is also possible that a user may fail to fit any set of stereotypes, so that the modeling process fails to associate any system decision to that user. In such situation, however, some of the assumptions distributed among the stereotypes might be still useful for characterizing that particular user. Hierarchical approaches lack the ability of reorganizing the assumptions distributed among the stereotypes (Chen and Norcio 1993).

(iii) Since the stereotype hierarchy is logically organized , it may not be appropriate for modeling user's short-term characteristics which change over time and temporally exhibit great varieties regarding the current task. For instance, it is very realistic that a user might demonstrate both expert and novice traits toward same task. In that case, the hierarchical classification of characteristics via few dimensions might not be able to provide sufficient information for the system to adapt the current interaction. As we discussed in *(i)*, the

hierarchical structure that only support the logical classification lacks the abilities of recognizing the *irrational* data about user's behavior . Whether such data need to be taken into account should depend on the examination of overall behavior pattern.

ASSOCIATIVE USER MODEL

We propose that associative network can be used as an efficient mechanism for stereotyping user and task. It can be expected to overcome limitations of hierarchical approaches discussed in above section. In associative network based user modeling process, the stereotypical knowledge is organized as a set of patterns. Since the system's beliefs about a user should be determined through the context of task performance, fragmented pieces of observation may not bring any meaningful implication. In addition, the observed user's behavior may be mixed with noise or inconsistencies. Therefore, all the aspects of the user's performance pattern have to be examined before any system decision can be made. In other words, stereotype assignment should be conducted in terms of pattern recognition, which requires that the modeling system has the capability of fault tolerance, graceful degradation, and signal enhancement. As we suggested in above section, the conventional stereotyping approaches, in which the inference proceeds a step at a time through sequential logic, may become seriously inadequate for processing pattern-formatted knowledge especially when there are incomplete, noisy or inconsistent observations involved (Kohonen 1988, Lippman 1987).

We have presented several operation primitives of artificial neural networks in user modeling (Chen and Norcio 1992). There are several paradigms of associative network where the associations between input and output patterns can be accurately captured despite incomplete or inconsistent inputs (lippman 1987)., In our approach associative network is implemented by a single set of interconnected units. Each unit represents a assumption which can be an assumption or an attribute that underlies the system's stereotypical knowledge about users and the task they are performing. Each unit serves as both an input unit and an output unit. Its feature of self-organizing is especially useful in recognizing user-task context.

In associative user model, the assumptions are not confined within any pre-defined stereotype. Actually there is no separately defined stereotype. All assumptions constitute a *universal stereotype* that covers all task-related characteristics of potential users. Given the representation of hierarchical stereotype (refer to previous section), an illustration on a associative user model S_A could be represented as follows :

An associative user model is a subset of universal stereotype **U**:

$\mathbf{S_A}$ U = , where N is the total number of stereotypes in S.

$\mathbf{S_A}$ = n is the total number of stereotypes involved in a user model, $n < N$.

A_i $\mathbf{S_i}$

It is obvious that the number of subset of U is much greater than the number of subset of S (refer to previous section), which implies three aspects:

(i) S_A could be any combination of the assumptions in **U**. Comparing with hierarchical approach, given certain number of assumptions, more states of user profiles can be established by associative modeling. Therefore the representation power of user models is enhanced. Ideally, 2^n different profiles could be reached .

(ii) The modeling process extracts some of assumptions to dynamically form a unique stereotype that fits a particular user and the task. Unlike the hierarchical stereotyping approaches that only associate the user with a single or a set of stereotypes, an associative user model includes various assumptions associated to a user. In other word, associative user

modeling proceeds at the level of an elementary assumption rather than the level of a stereotype. In this approach, all assumptions of the universal stereotype are considered to be relevant to each other in a spectrum which is valued from negative to positive (i. e. from contrary, via irrelevant, to consistent). Thus, it overcomes the limitation of the hierarchical stereotypes which is unable to extract assumptions from different stereotype structures to form a new profile about the current user and task. Therefore, it has better ability to personalize a user.

(iii) The associative user modeling can still make the same effect as hierarchical approaches does. If we let $A_i = S_i$ then S_A and S_h have the same assumtions. If the causal relationships are carefully distributed and weighted. The associative user modeling process can inherit all advantages of hierarchical approaches such as structured classification , attribute inheritance and automatic assumption assignment.

Figure 2 shows a structure of associative model. Hierarchical modeling activates the stereotypes whereas associative modeling activates the assumptions. Given the same amount of stereotypical knowledge (i.e. number of assumptions), it is obvious that the associative model can provide more individualized states than the hierarchical model. In the modeling process, the observations from user's input will activate some assumptions. The inference is conducted by propagating the activation level which functions as a mechanism of default reasoning. In the conventional stereotype approaches, if the inconsistency is detected, dependency-directed backtracking is evoked to solve the conflicts (i. e. truth maintenance process) . In our approach the symbolic reasoning in conducted over associative networks. The new input from user is combined with the previous input to form a current pattern of input to the modeling process. In other word, even though some of the input elements may conflict to each other, the network can still make correct classification because the neural networks always handles the input in the view of pattern recognition, and are robust to the inconsistent input. In addition, there is no truth maintenance procedure involved in this approach which greatly simplifies the modeling process.

Generally, two network status can be eventually created: the network reaches a single, stable state, or it reaches a mode in which it cycles through a constant series of states. The semantics of these status depends on the domain of applications. The set of those units reaching the highest activation level specifies a set of assumptions that are applicable to a user or a task. If the network cycles through a series of states, it may imply alternative sets of assumptions which are the likely results of a particular stimulus (McClelland and Rumelhart 1986).

The causal relationships between the assumptions must be defined to initialize an associative user model, . All assumptions are organized into associative network in which the relationships among the assumptions are weighted under certain conditions (see figure 2). Since it is easier to identify the contextual correlation between two assumptions than to define two complete stereotypes and their relationship, the initial knowledge acquisition tends to be easier in associative modeling. In addition, learning ability can be implemented in terms of automatically revising weights, adding or deleting a unit representing an assumption. One can expect that the self-organizing feature of the network enables the local change of the assumptions and the weights to make a consistent impact on model construction (Kohonen 1988).

The relationship between two assumptions could fall into three categories. Accordingly three types of the weight patterns are identified. If the two units tend to be on and off together, then it means that the corresponding assumptions are consistent regarding a user or task. Therefore the weight for their relationship can be a large positive value. If the two units come on and off independently, then the corresponding assumptions may be somehow irrelevant to each other and the weight of their connection should be almost zero. If the two assumptions contrary which is implied by a larger value difference of corresponding units, then the weight takes on a negative value. In order to implement such weight pattern in associative network, it might be necessary to conduct a Bayesian analysis on

$$p(x_i = a \;\; x_j = b \;) \; (a , b \; = \; 0{,}1 \;),$$

which is the probability that unit x_i is activated given unit x_j is activated and vise versa (McClelland and Rumelhart 1986). It is often difficult, however, to assign precise probabilistic values and maintain the formalism, we define value of p as the value of plausibility which is a fuzzy membership and is not required to satisfy the probability formalism.

RECOGNIZING USER'S BEHAVIOR

Associative approach provides a simple mechanism to model or predict the user's task performance in terms of two aspects. First, the state transition of the task performance can be transformed as a stored pattern in associate network in which each state is represented by a unit and the transitions are mapped to weighted connections. During an interaction, a user's operations will form an stimulus vector for further propagating activation throughout the network which can produce a relatively complete path of state transitions. This path can be used for predicting or navigating task performance. Secondly, this approach also provides a way of capturing user' mental model regarding certain tasks. Modeling process assumes a user's problem solving as a procedure of associative thinking. Once the input from a user triggers assumptions about what a user already knows, the correspondingly units will propagate their activation's through interconnected units until the network falls into an equilibrium state. The dynamics of activated units and the pattern of weights reflect the user's beliefs toward his/her task accomplishment. Therefore, this approach can establish a basis for comparing the user's real action with the predicated one and then revising the model dynamically. Such comparison is often necessary in the tutoring systems while a student model needs to be revised against the tutoring knowledge base.

CONCLUSION

An integrated framework of associative user modeling is under investigation (Chen and Norcio 1993). Two knowledge sources are necessary in order to tailor system response to an individual user: user profiles and the task profiles. Accordingly two associative memories are utilized: one that incorporates the assumptions about user's task-related characteristics and one that stores the assumptions underlying system actions or responses. From the view point of system adaptation, one network handles user modeling, while another handles task modeling. Therefore the issues of when, how and what to adapt to a user are all addressed by the modeling process (Norcio and Stanley 1989). The proposed framework is domain independent. However it is specially useful in the information retrieval systems or tutoring systems where the system response can be generated based on a set of activated units (e. g. the entities in a database or the assumptions in tutoring context).

REFERENCES

Chen, Q. and Norcio, A. F. (1993). An Associative Approach in Dynamic User Modeling , *Proceedings of 5th Intentional Conference of Human Computer Interaction,* pp 909-914

Chen, Q. and Norcio, A. F. (1992) Modeling Users with Neural Architectures. In*Proceedings of 1992 International Joint Conference on Neural Networks,* pp. I-547 - I-552.

Chen, Q. and Norcio, A. F. (1991). A Neural Network Approach to User Modeling, In *Proceedings of 1991 IEEE International Conference on Systems, Man, and Cybernetics,* pp.1429-1434

Clowes, I., Cole, I., Arshad, F., Hopkins, C., and Hockley, A. (1985). User Modeling Techniques for Interactive Systems," In *People and Computers: Designing the Interface,* Johnson P. and Cook, S. (Eds.), Cambridge University Press. pp. 35-45

Doyle, J. A Truth Maintenance System, *Artificial Intelligence Vol. 12,* pp 231-272 (1979).

de Kleer, J. (1986) An Assumption-Based TMS, *Artificial Intelligence Vol 28,* No.2 pp 127-162.

Kobsa, A and W. Wahlster (Eds). (1989)*User Models in Dialogue Systems,* Springer-Verlag, Berlin.

Kohonen, T. (1988). *Self-Organization and Associative Network,* 2nd ed., New York: Spring-Verlag,

Lippman, R. P. (1987). An introduction to computing neural nets. *IEEE ASSP Magazine, Vol 4,* pp 4-22

McClelland, J. L. and Rumelhart, D. E. (1986). *Parallel Distributed Processing Vol. 2,* Cambridge, MA: MIT Press

Norcio, A. F. and Stanley, J. (1989). Adaptive Human-computer Interfaces, A literature Survey and Perspective," *IEEE Transactions on System, Man and Cybernetics, Vol. 19,* No. 2, pp. 399-408.

Pao, Y. (1989). *Adaptive Pattern Recognition and Neural Networks,* Addision-Wesley Publishing Co.

Rich, E. (1979). User Modeling via Stereotypes," *Cognitive Science Vol 3,* pp 329-354.