Workshop on N Networks - Acad / IND. / NASA / Defense
Feb 11-13, '91 Auburn AL

NEURAL NETWORKS FOR ENHANCED HUMAN-COMPUTER INTERACTIONS

Alianna J. Maren, Ph.D.
The Tennessee Center for
Neural Engineering and Applications
The University of Tennessee, Knoxville, TN
and
The University of Tennessee Space Institute, Tullahoma, TN

Keywords: Human-Computer Interface, User Models, Adaptive Systems, Self-Organizing Networks, Adaptive Critic

ABSTRACT

One of the most significant and challenging technology areas for the decade leading into the year 2000 is that of Human-Computer Interfaces (HCIs). HCIs are crucially important because they directly mediate the ability of humans to extend their capabilities via interaction with intelligent systems. In order to reach their potential effectiveness, HCIs must be able to effectively model users, their communication with the users, and their task environment. They should be able to dynamically adapt their models to changes in conditions and their users. Neural networks are a powerful, yet still relatively unexploited, technology for improving the effectiveness of HCIs. While self-organizing networks and feedforward networks have much to offer in this area, the most likely avenue for extraordinary contribution to improved HCIs lies with neural networks which create models of dynamic processes, such as Adaptive Critic network systems.

1.0 THE CHALLENGE: IMPROVED HUMAN-COMPUTER INTERFACES

One of the most significant challenges which we will confront over the next decade, and into the next millenium, is our potential to extend our abilities via interaction with intelligent systems. This interaction will be mediated by the Human-Computer Interface (HCI). With improved HCIs, we will be better able to manage complex tasks and environments, ranging from integrating information in a C3I context to managing nuclear power plants. HCIs will be crucially important in control tasks where many sensors and high information loads are present, such as in piloting high performance aircraft. Further, much of our education may take place via interaction with a computer, in an advanced intelligent tutoring or coaching system environment. This educational impact will be felt not only in the K-12 realm, but will be important in training military and civilian personnel new skills.

The effectiveness of an HCI is very much influenced by its ability to form adaptive models - models of its user, the task environment, and the communication between the intelligent system and the user. In each of these areas for model-building, one of the most important factors is the ability of an HCI to dynamically adapt the model - to developmental changes in the user, to changes in the task environment, to differing emphases or to new information that contributes to the human-computer interaction.

While much work has been done in this area, there are few systems yet which allow for adaptive HCIs. This is in part due to the complexity of the task, and in part due to the fact that many researchers in this area are not yet familiar with the power that new technologies, such as neural networks, can bring to bear on such challenges as creating real-time adaptive situation or user-specific models. Often, this lack of knowledge is directly related to the personal cost of gaining sufficient mastery in a new technology area to be able to capably exploit its potential in a new applications domain.

The goal of this paper is to briefly identify some of the issues in developing HCIs, to point out where neural networks can be used in this area, and to describe specific types of networks in terms of their useful HCI applications.

2. ISSUES IN DEVELOPING ADVANCED HUMAN-COMPUTER INTERFACES

There are many issues which are important in developing better models for improved HCIs, such as: How can we create a useful basis for modeling basic differences in users, such as cognitive style? How can we dynamically adapt a model to take into account real-time variables (such as user alertness, difficulty of material to be learned, or "busyness" of a control environment) along with measures of student/user performance in that task or environment? How do we select, configure, monitor, and integrate different information-bearing "channels" between the human and computer to build up a model of the communication? The following subsections explore some of these issues.

2.1 Developing User Models

Several researchers have proposed different dimensions for making distinctions among learning and cognitive style among individuals. An example might be the work of Kolb [Kolb, 1971; Kolb & Fry, 1975], who has suggested some major distinctions in learning style, such as the distinction between conceptual and concrete learners. As another example, individuals vary in their ability to use spatial representations of information. Similarly, work by Snow [1989, 1989-90, 1986, 1983-5], Sternberg [1989, 1988], Hunt [1975] and others addresses this issue.

Despite interest and work along these lines, we are not certain what forms of individual differences may be significant in developing user models and in mediating human-computer interactions. Although there is quite a large literature discussing individual differences in learning styles, firm empirical support has not been established. Comparisons of studies such as the one by Sein and Bostrom [1989] with that by Vicentes and Willeges [1986] illustrate the difficulty of trying to design tools which may differentially aid different classes of users.

This does not discount the need for user adaptive interfaces. It just means that an interface should be able to rapidly build up a model of the learning style which a (new) user might employ. The system should be able to monitor how the student's range of expertise and forms of expression grow over time, and take advantage of new ways to interact with the student as the student develops greater competence and expressive ease.

There are several requirements involved in incorporating aspects of individual differences into HCI design. First, we need to identify which, of the many possible dimensions of cognitive variability, are those most important in creating an HCI system. We need to identify the extent of their potential significance as variables for consideration in designing user-adaptive systems. We need to identify the ways in which these dimensions are confounded with each other (as evidenced in the work of Sein and Bostrom [1989]) when designing a structured, goal-oriented interaction. We also need to identify how these dimensions may become confounded with the task domain, and how user characteristics may become labile as the student learns.

2.2 Developing Task Environment / Situation Models

One way in which an HCI might evolve is called "adaptive aiding," in which the intelligent system might off-load certain responsibilities from the human operator when the workload or other situational concerns dictate. In order to do this, the HCI must develop a model of the task environment or situation, and model the workload. It must also be able to identify how to differentially allocate different tasks to either the human or the computer. Many situational factors influence human performance, and the system needs to build a model of the situation which takes into account historical information, such as the immediately preceding workload and the length of time the user has been working. This amounts to a demand for complex spatio-temporal pattern recognition on the part of the HCI. Although previous approaches have used expert systems, mathematical models, and algorithms, their is room for a more robust, flexible approach.

2.3 Developing Models of the Human-Computer Communication

There are several issues here. First, what channels shall we use, monitor, and model? As the complexity of HCI systems grow, there are a greater variety of information-bearing channels from which to select. When we wish to model time-varying student/user characteristics such as attention, we need to infer this quality from both explicit and implicit information. We will need to corrolate information from multiple channels to create estimates. When working with multiple channels, and attempting to integrate information, more issues arise, such as identifying the extent to which the information passed by each channel is orthogonal.

The second major consideration deals with developing a real-time, dynamically-adjusted model of system user performance. This model will be task-dependent. To build such a dynamically-varying model, it will be necessary to infer information about the user from signals extracted from the man-machine interaction.

The information which can be extracted may be either or both signal or symbol, and may be explicit and/or implicit. For example, if a person accesses an editor within the C environment, one can infer different levels of expertise if this person uses the "ED" command than if he or she uses the "EMACS" command. This is an example of inference from a symbolic datum. From this, we can move to an increasingly comprehensive concept of the "channels" of information flowing between the human and the computer to include a model of the entire conversation or communication.

3.0 MEETING THE CHALLENGE - NEURAL NETWORKS AND RELATED EMERGING TECHNOLOGIES

To meet the challenge just described, we will have to draw on developments in multiple disciplines, including cognitive science, adaptive interface and interactive system design, and technologies for real-time pattern recognition, multisource information fusion, adaptative control, and human performance modeling.

The previous subsections addressed the issues of creating a (dynamically updatable) user model and selecting and configuring optimal communications channels between a person and a computer. Given that appropriate dimensions for creating a user model can be specified, and that appropriate communications channels can be selected and configured, there is still the issue of how models (of both the user and the communication) will be created, maintained, and updated, and how multiple types of information from the different channels will be extracted and corrolated. Neural networks, aided by existing rational inference methods (ranging from expert systems to fuzzy logic) may be useful in these tasks. Neural networks are useful for real-time pattern recognition, multisource information fusion, clustering and mapping data, adaptive control, and building up system models without the need for explicit parameter or model-equation specification.

It will be easiest to create ways in which a system can adapt to different users if we can characterize, a priori and on a broad scale, different classes of users. In designing system response characteristics, it will be useful to know the types and the characteristics of different classes of users. Two types of neural networks will be useful for this type of task. The Kohonen unsupervised Learning Vector Quantization (LVQ) network is a clustering network. It can adjust the vector element values of a set of quantizing vectors in order to create "exemplar" vectors which represent clusters in a set of data. The Kohonen Self-Organizing Topology-Preserving Map (SOTPM) is a more advanced and powerful version of the LVQ. It can create a topographic mapping of a set of vector data which creates a data clustering visible in a reduced dimensionality space from the original data. This will facilitate interpretation of the data clusters describing different user types. It provides a proportional representation of the data, which can be used to estimate the probability distribution of different data types [Gacem and Maren, in preparation; Kohonen, 1989]. Heger and Koen [1991] have taken a self-organizing neural network approach to creating an adaptive interface to a database. Their approach builds up user-specific activations of units of memory, leading to adapting the interface to individual users as well as creating a more robust approach to database query.

Although self-organizing networks can be modified to reflect changes in user characteristics, their primary use will be in representing the baseline and slowly-varying aspects of the user population. Different types of networks will be useful for recognizing patterns in real-time spatio-temporal data acquired from the user and for corrolating and fusing real-time data.

The data which will be acquired from the user may be either or both symbol or signal level data, and may be an explicit communication from the user, or it may be indirect. Neural networks such as the back-propagating Perceptron (and its many varients and specializations) will be best for extracting and recognizing patterns in real-time signal data (e.g. time between keystrokes). These networks excel at being able to work with partially incomplete and/or noisy data. In addition, certain neural networks (e.g. Adaptive Resonance Theory - ART - networks) can create a new pattern category when presented with a novel stimulus. Both network types can be used to associate multi-source data at different levels of abstraction, or to create a new, more abstract feature representing corrolated information.

The ability of neural networks (the back-propagating Perceptron in particular) to model different aspects of individual users will be an especially useful area to explore. For example, suppose that we wish to represent an abstract quality, such as "attention." Such user attributes as this can not be directly determined from a single data source, and are likely to be task-influenced. Further, each user will be characterized differently for any abstract quality (such as attention), depending on the user's familiarity with the task, response skill, and other factors which may confound any attempt to set parameters definining attention, alertness, or any other abstract characteristic that we wish to model. We can use neural networks to create models of specific users based on experience with each user. Recent work by Fix [1990] illustrates how a simple back-propagation network can be used to emulate human performance in a dynamic control task. Maren et al. [1990] describes other neural network emulations of human performance. The issue of creating a user model is, however, deeper than that of just modeling certain abstractions describing user characteristics. A rather highly evolved form of neural networks, actually a system of interacting networks, can be used to create more powerful system models.

One very powerful form of neural network system design for this purpose is the "adaptive critic" system [Werbos, 1990]. The adaptive critic system is composed of three types of networks - the Utility network, the Model network, and the Critic network. The Utility network expresses as utility function (which is maximized by the system) as a network. The Model network inputs current sensor readings and actions, and outputs simulated values for these readings at the next time period. The Critic network predicts what the future utility will be - summed up over future time - as a function of present actions. In the back-propagated Adaptive Critic, the strategic utility function J is back-propagated through the Model network, in order to pinpoint which actions need changing, and in which direction. This carries the potential for dynamically adaptating the Model network. In application to HCI systems, the Model network would generate predictions of the user actions (via the Action network). This would provide a basis for comparing the user's real actions with the predictions, and thus adjusting the Model. The Critic adapts the Model under the assumption that the user is attempting to optimize performance, which is not unreasonable as a first-order assumption in either a teaching or a system monitoring or control type of application.

4. SUMMARY

Although neural networks have not yet been exploited to meet the challenges of developing improved human-computer interfaces, their potential for use is very much in evidence. As knowledge grows of how neural networks can be used for pattern recognition, information fusion, self-organization of information, and adaptive model building, more HCI designs will undoubtedly incorporate neural network components.

ACKNOWLEDGEMENTS

Conversations with Drs. Alan Lesgold, Tony Norcio, John Hestenes, Joseph Young, and Paul Werbos have all been influential in bringing together the thoughts expressed in this paper. However, the author takes full responsibility for the expression of what is, essentially, a position paper and expression of personal views. While working on this paper, the author has been supported by research contracts from the U.S. Dept. of Energy (DoE-DEFG-88-ER-12824), the U.S. Navy (N66001-89-C-7063), and the U.S. Marines. She would like to express appreciation to Dr. Robert E. Uhrig, Mr. Robert M. Pap, and Ms. Mary Rich for their colleagial support.

BIBLIOGRAPHY

Fix, E. (1990). "Modeling human performance with neural networks," Proc. Third Int'l. Joint Conf. Neural Networks (San Diego CA; June 17-21, 1990), I-247 - I-252.

Gacem, A., & Maren, A.J. (Manuscript in preparation.) "The Kohonen Self-Organizing Networks: A comprehensive review of properties and varients." (Manuscript in preparation for the IEEE Trans. Neural Networks.)

Heger, A.S., & Koen, B.V. (1991). "KNOWBOT: An adaptive database interface," Nuclear Science and Engineering Journal, 107, #2. (To appear in February issue.)

Hunt, E. (1975-78). "Cognitive theory applied to individual differences," in W. K. Estes (Ed.), Handbook of Learning and Cognitive Processes, Hillsdale, NJ: Lawrence Erlbaum, 81-110.

Kohonen, T. (1989). Self-Organization and Associative Memory, New York & Berlin: Springer-Verlag.

Kolb, D.A. (1971). Individual Learning Styles and the Learning Process (Sloan School Working Paper 535-71). Cambridge, MA: MIT.

Kolb, D.A., & Fry, R. (1975). "Toward an applied theory of experienctial learning." In C.L. Cooper (Ed.), Theories of Group Processes, New York: Wiley, 33-57.

Maren, A.J., Harston, C.T., & Pap, R.M. (1990). Handbook of Neural Computing Applications, Academic Press, New York.

Norcio, A.F., & Stanley, J. (1989). "Adaptive human-computer interfaces: A literature survey and perspective," IEEE Trans. Systems, Man, & Cybernetics, 19 (March/April), 399-408.

Sein, M.K., & Bostrom, R.P. (1989). "Individual differences and conceptual models in training novice users," Human-Computer Interaction, 4, 197-229.

Snow, R.E. (1989-90). "Toward assessment of cognitive and conative structures," Educational Researcher.

Snow, R.E. (Expected 1989). "Cognitive-conative aptitude interactions in learning," in R. Kanfer, P.L. Ackerman, & R. Cudeck (Eds.), Abilities, Motivation, and Methodology, Hillsdale, NJ: Lawrence Erlbaum.

Snow, R.E. (1986). "Individual differences and the design of educational programs," American Psychologist (October, 1986), 1029-1039.

Snow, R.E. (1984). "Toward a theory of cognitive aptitude for learning from instruction," J. Educatioanl Psych., 76, 347-376.

Snow, R.E. (1983-5). "Aptitude-treatment interactions a a framework for research on individual differences in learning" in P.L. Ackerman, R.J. Sternberg, and R. Glaser (Eds.) Learning and Individual Differences, New York: W.H. Freeman, 13-59.

Sternberg, R.J. (1989). "Domain-generality versus domain-specificity: The life and impending death of a false dichotomy," Merill-Palmer Quarterly, 35 (January, 1989), 115-130.

Sternberg, R.J. (1988). "Mental self-government: A theory of intellectual styles and their development," Human Development, 31, 197-224.

Vicente, K.J., & Willeges, R.C. (1988). "Accomodating individual differences in searching a hierarchical file system," Int. J. Man-Machine Studies, 29, 647-668.

Werbos, P.J. (1990). "Neurocontrol and Related Techniques," in A.J. Maren, C.T. Harston, & R.M. Pap, Handbook of Neural Computing Applications, Academic Press, New York.