

Integrating User Affective State Assessment in Enhancing HCI: Review and Proposition

Xiangyang Li*

Department of Industrial and Manufacturing Systems Engineering, University of Michigan – Dearborn, 4901 Evergreen Road, Michigan 48128, USA

Abstract: In the past decades, human computer interaction field began to witness studies exploring not only the rational characteristics of human users in making decisions but also their “extra-rational” aspects in interacting with environment and devices. The first task in exploiting one such facet, human affective states, is to accurately recognize and assess the internal state affective state. In this paper, we summarize the features, challenges and requirements of real-world HCI systems to user modeling and assistance. A survey of the state-of-the-art of affective state assessment focuses on computational assessment models underlying different studies. A user affective state assessment framework developed by the author based on dynamic Bayesian networks is also visited. The last part of the paper moves on to articulate what is perceived the future research direction contributing to the challenging research of affective state assessment in user modeling, which foresee integration of multi-disciplinary knowledge bases.

Keywords: Affective state assessment, user modeling, intelligent assistance, HCI.

I. INTRODUCTION

The field of human computer interaction (HCI) has moved from studies on friendly interfaces such as graphical user interfaces (GUIs), to those systems that actively and intelligently understand, explain, justify, and augment user actions, focusing on developing more powerful representations and inferential machinery [31, 35]. The rapid development of computing technology has cultivated emerging intelligent applications exemplified by pervasive computing, ubiquitous access, and ambient intelligence [1, 33]. Of them one important task is to design and deploy intelligent agents to provide personalized assistance to users in daily work and life, which are able of automatic feedback and interaction [20, 21]. Aided with new techniques acquiring physiological and behavioral information of users and the contextual information of their working environment, research on intelligent assistance systems has been extended into a broad spectrum of applications. Underlying these systems, an accurate user model plays the central role to capture the status and need of users and to estimate and reflect the impact of action/assistance, by constantly monitoring system states and ongoing activities. Based on often incomplete and uncertain sensory observations of multi-modalities, such intelligent assistance systems have to reason about the user’s internal state over time, under the requirement of high fidelity.

In this paper, we discuss the progress of intelligent user modeling and assistance systems considering human affective or emotional state. Being considered “extra-rational” aspect of human users interacting with the machine and other

surroundings, negative or positive affective states show impact on human cognitive and operating performance, recently raising interests among many researchers in HCI and cognitive science. As specified in affective computing [41], emotional machine intelligence comprises affective state recognition, understanding, and expression. As the foundation of affective computing applications, affective state assessment (ASA) uses observable evidence of human subjects, often from multiple modalities or channels, to classify and appraise the internal status into several affective state categories.

First of all we need to reason about the need for an integrated ASA model for efficient human machine interactions. In all existing studies random variables are used to provide situating of relevant factors to human user internal states, including goal, emotion, mentality, etc. For example, in order to differentiate varying affective states of the same user we often need contextual and profile variables in probabilistic models for more accurate assessment. Considering such a usually complex system, sometimes we do not need to or are unable to elucidate these factors in a system. An example in thermal theory is that we just focus on a set of equations involving a set of average system measurements of a chemical reaction process without describing the parameters of each participating molecule. But more and more, modeling at such single levels is not sufficient to achieve the required assessment for a HCI system personalized for individual human users or user groups. The evolution of this system is influenced by multi-fold information flows at different levels. Now we begin to see HCI applications that cannot rely on any more only a template of average human user behaviors. On the other hand, consideration and integration of heterogeneous knowledge bases at various levels and dimensions depends on the available techniques and capabilities. We have seen new computing technologies that offer the capabil-

*Address correspondence to this author at the Department of Industrial and Manufacturing Systems Engineering, University of Michigan – Dearborn, 4901 Evergreen Road, Michigan 48128, USA; Tel: 313-5836416; Email: xylum@umich.edu

ity we never had before, and psychological studies that provides techniques and findings essential for exploring rational or extra-rational aspects of human cognition activities. That is, those past random variables are not that random any more. It is the time for us to consider a more comprehensive “system of systems” solution to greatly push forward the study on pervasive, ubiquitous, customized, and intelligent access to and assistance from machines for human users. Thus the implementation can be positioned in an environment that is more practical than those controlled laboratory settings.

However, considering all the available does guarantee that we can always integrate them at varying details, in different dimensions, or scales successfully. The many challenges imposed as the obstacles as we will discuss provide the exact motivations for this article. Therefore, we will not focus on enumerating all relevant research sub-field considered in our integral framework. And we are doomed if we aim at generating a possible list of complete literature in each of those fields simply because of the enormous studies going on. Instead, we tackle the more essential inquires about *why* and *how* in integrating these research efforts. After the discussion of relevant studies that shed light on this topic, we will elaborate on our methodology of organically integrating heterogeneous models.

In this paper, we first present a survey of computational ASA models including a generic framework based on dynamic Bayesian network models that detects affective states based on observations, and profile and contextual information. The second part of the paper is devoted to the means to address the challenges of integrating heterogeneous models. We describe what we perceive the future directions of integrating computational models to improve the assessment performance. Whenever appropriate we also cover relevant issues in the broader themes of user modeling and assistance.

II. USER MODELING, ASSISTANCE AND AFFECTIVE STATE ASSESSMENT

Embodied by the explosive increase of computation capacity and new applications, computing system and human operating environment become extremely complicated and complex as never before. Ranging from commuting in a modern car providing remote assistance, to working on large office and management software, to steering a military vehicle or aircraft through highly dangerous battlefields, in order to aid human users to efficiently utilize such systems, intelligent agents coupled into these systems seem a necessity that automatically and intelligently provide feedback and help. These intelligent systems no longer carry a fixed working logic in response to static situations and limited configurations. Now they have to be able to dynamically catch and understand the change of human operators, with internal status as well as external behavior, and efficiently respond accordingly. A user model is always the core component of an intelligent assistance system, which captures, interprets and reacts to the goal, intention, need, and other internal states of the human user. This underlying user model fulfills these tasks through the exchange of knowledge with decision and actuating components and under the requirement of high

fidelity [19, 28, 54]. A variety of user modeling tools have been developed, falling into two major categories of paradigms: cognitive modeling from psychological and cognitive science, including GOMS, ACT-R, SOAR, EPIC [4, 6, 11] and statistical/probabilistic modeling based on various mathematical representations, such as rule-based systems, regression models, neural networks, and Bayesian networks [3, 52]. Many attempts have applied these models to human machine systems in computer operating, student tutoring, and vehicle driving [22, 25, 44, 51].

User modeling and intelligent assistance systems still face a set of great challenges in practical working settings because. First, user’s states and behaviors often show uncertainty that constantly changes over circumstance and time. Moreover, sensory observations of the user’s behaviors are often incomplete, uncertain, and from different modalities. Second, decisions about the user’s need and the assistance must be rendered appropriately and efficiently under various constraints. More important than ever, there are no “average” human subjects, calling for efficiently and incrementally adjusting the computational models to the specific user.

Furthermore, user modeling has traditionally focused on what is generally considered “rational” aspects of user behavior, typically the user’s knowledge and belief state. While useful, models focusing severely on these aspects of user state often miss critical components of user mental state and behavior: affective states. As “extra-rational” factors in human cognition and decision making, affective states, negative or positive, have been shown to strongly influence both reasoning and communication [23]. For example, certain anxiety or alertness level could improve the user’s preparation for reaction while too much anxiety could distract the same user from noticing important operational procedures. It is important to systematically integrate these factors into user modeling. Taken as another example, every year many people are injured in car accidents due to the fact that the drivers are in status including fatigue, nervousness, or confusion. If we could in a timely manner detect these dangerous states, understand the relevant causes and consequences, and engage in appropriate assistance and alerts accordingly, we may prevent many accidents from happening.

In traditional human factors and ergonomics research, the human operator is “fixed” once the initial configuration of the machine, only where there is consideration for the goodness of “average” operators, is done. However, human subjects may frequently enter abnormal or negative status, being inattentive, fatigued, or nervous. On the other hand, as in the vision-augmented driving assistance systems [49], the traditional “external” assistance system is based on the assumption of normal operators, fully capable of fulfilling tasks. Such an assistance system aims to modify working conditions or raise alarms based solely on context awareness, e.g. detection of obstacles or departure from the lane center. The assistance is passive and very limited. To provide appropriate assistance we need pay more attention to the “internal” situations because it is in these states that the human user’s operating performance deteriorates sharply, the operations

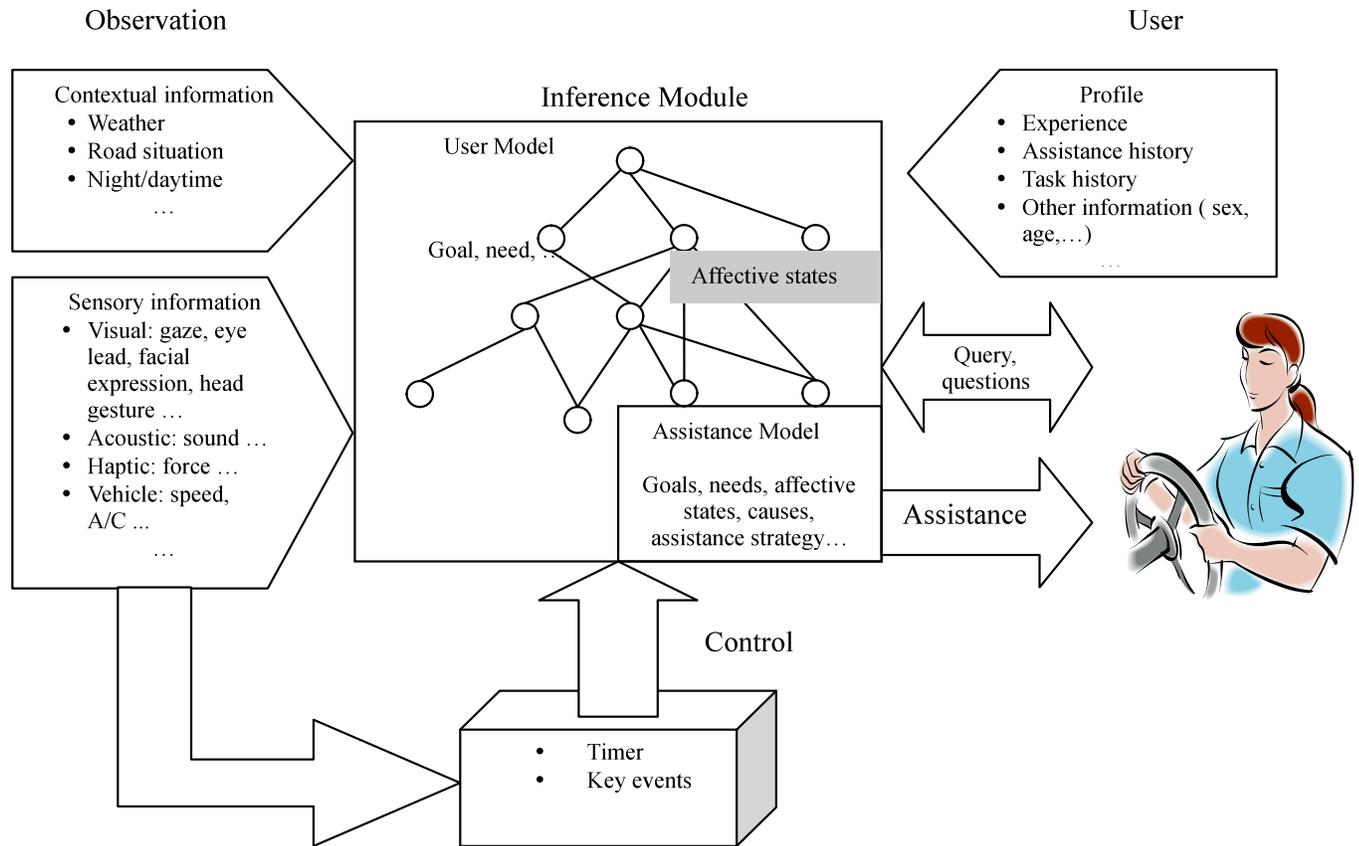


Fig. (1). An illustration of intelligent assistance, user modeling, and affective state assessment, where affective states and user goals, needs, etc. are assessed and fed into the assistance model.

tend to cause accidents, and thus the assistance is needed the most.

Affective state assessment is the core task of developing affection-aware user modeling systems, as shown in Fig. (1). ASA is also the foundation of emotional machine intelligence that involves the interpretation and exploitation of affective aspects in HCI systems. In the following sections, we present a survey of existing ASA models and then discuss a systematic methodology of affective state modeling and assessment.

III. AFFECTIVE STATE ASSESSMENTS MODELS

Affective state assessment is a pattern recognition or classification task. In this section, we first summarize several relevant issues in ASA including the classification of affective states and the type of measure that can be used as the predictors. Then the existing algorithms and models used for ASA are introduced and compared. Research efforts in ASA expand an extensive spectrum of psychology, physiology, linguistics, and computer science in the past several decades. The number of related systems and applications is too large to allow an exhaustive review. Also it could not be fully predicted at this time what else could contribute to the future research of affective state modeling and recognizing in user modeling. Therefore we limit this survey with a focus on current computational approaches in recognizing the affective state of a human subject, from observable information of the subject and/or surroundings.

A. Affective States and Useful Predictor Measures

Human beings have abundant emotions in terms of sadness, happiness, guilt, pride, shame, anxiety, fear, anger, disgust, etc. There are also mental states that are at lower consciousness level such as nervousness, anxiety, confusion, fatigue, etc. We use "affective state" to call all these internal states thereafter. Researchers in different fields choose different categorizations of these affective states that are closely related to the specific applications. Difficulty also exists in modeling if these affective categories are not independent because of the overlap among them. One helpful way of universally describing the emotion applies independent dimensions, such as valence and arousal [47]. Valence describes the "quality" of emotion, in terms of negative, neutral, and positive. Arousal describes the intensity or "energy degree," which may be "activated" or "not activated." Continuous values can be used for both dimensions. Classification assigns one categorical label for the affective status of a human subject at a specific time instant. However in practical HCI systems usually a quantitative value in terms of probability or confidence for multiple hypotheses is more useful. Other system components such as feedback and assistance activation can directly use these values in calculation and decision making.

Predictor measures used in ASA come from various informative modalities of a human subject, containing information implicitly or explicitly related to the internal affective

state. Information can be collected into discrete or continuous-valued variables with values evolving over time. A broad set of information falls into the group that a computing machine could use to assess the subject's affective state. To sympathize with the abundance and complexity of such measures, we could stretch our imagination to the context of active social setting such as a conference or party. The physiological measurements, facial and body expression, behavior, wording choice and sentence organization in dialogue could all be indications of a participant's affective state. Even the dressing of a guest could reflect the internal status prior to his/her arrival. Due to current limitation on the acquisition capability of instrument and the understanding of these measures in general application domains, most studies deal with only a small subset of these measures enumerated below.

The measures used in ASA can be classified into different categories according to different criteria, such as verbal and nonverbal, or intrusive and non-intrusive. The following categorization is based on the nature of information modality and the required type of acquisition instrument.

(1) *Self-report*: The subject could report his/her emotional state at specific point of time. Generally we consider self-report inapplicable in a practical user modeling and assistance system, where the self-report cannot be done in real-time, the interruption to the subject is at high risk or not tolerable, and the reliability of acquired assertions is greatly impacted by the human subject's personality and the reporting mechanism. However, self-report in terms of direct query to the subject may be of interests in some cases (e.g. in training of the detection models) if we have tremendous concern about the recognition accuracy and prefer a very conservative assistance strategy.

(2) *Physiological Measures*: These measures uses special acquisition instrument to get physiological measurements of human subjects. For example, skin conductance instrument measures the resistance of human body using two sensors, placed on two finger of one hand. This gives a measure of overall stress. They mainly include EEG (Electroencephalography) on brain, EMG (Electromyography) on muscle, SC (Skin Conductance) or GSR (Galvanic Skin Response) on skin resistance, GSA (General Somatic Activity) on human body, body temperature, and other types of biophysical feedback such as heart rate, respiration, and perspiration.

(3) *Physical Expression and Behavior*: Compared to the above physiological measures, these measures are non-intrusive but less reliable. These physical measures can be again classified into visual, acoustic, and behavioral groups.

- Visual modalities include eye movement, facial expression, head gesture, body gesture, etc. For example, the important coding systems in facial expression - the Facial Action Coding System (FACS) in Carnegie Mellon University, and the later Facial Animation Parameter Units (FAPUs) and the Facial Definition Parameter (FDP) set defined in the ISO MPEG-4 standard could be used to define facial shape and tex-

ture, as well as the animation of faces reproducing expressions, emotions and speech pronunciation.

- Acoustic modalities include the physical expressions in voice intonation, volume, pitch, and semantic content in speech such as wording, errors, and sentence organization.
- Behavioral modalities are highly domain specific. For example, they can be the mouse movement when the human subject is working on a computer, or the steering features when driving a car.

(4) *Social and Problem-solving Strategy*: Human beings' high level assessment of surroundings and decisions in choosing among alternative strategies to solve problems often reflect the change of internal state. Using social and problem-solving strategies in emotion recognition is still difficult and most research efforts have been in other fields such as management and psychology. Such modalities are highly domain dependent and obviously a little far away from current progress status of HCI study.

B. Computational Models

Affective state assessment labels the current user state with certain affective category based on the observed measurements coming from the informative channels as discussed in the above. Thus most computational models from statistics and probability, machine learning, and pattern recognition could be possible candidates for ASA. Existing assessment techniques include rule induction, fuzzy sets, case-based learning, linear regression, discriminant analysis, hidden Markov model, neural networks, and Bayesian models. Current ASA studies rarely use computational models from cognitive science.

(1) *Rule-based Systems*: As a common representation method of knowledge, rule based systems describe in condition-consequence pairs the relationship between predictor values and target classes. In fact, models built from different algorithms could be transformed into IF-THEN rules and match the patterns represented by these rules in assessment or classification stages. Moreover many emotion synthesis systems apply rule-based models to generate emotion expressions. However, because of the constraints in representation flexibility and modeling capability of the rule structure, researchers tend to ignore rule-based systems especially in feature-extraction and pattern recognition. Rule-based systems are not good at handling uncertainty and incompleteness in input measures. A research of interest is in facial expression recognition by Pantic *et al.* [37]. A knowledge base containing twenty rules was used to recognize the action units (AUs), defined in the Facial Action Coding System (FACS) (<http://www-2.cs.cmu.edu/afs/cs/project/face/www/facs.htm>). The input measures are the extracted mid-level feature parameters, e.g. distances between two points, which describe the state and motion of the feature points and shapes in the face profile contour. The action units, such as eyebrow raiser and lip corner depressor, relate to affective states closely and could be used in a higher-level model outputting affective state categories.

(2) *Fuzzy Sets*: Fuzzy sets incorporate probability by the degree of membership of an element in a class. It defines the input membership functions and fuzzy rules to process inputs into rule strengths, which represent the degree to agree on the consequence in each rule. These rule strengths are combined by output membership functions to get the output distribution. If necessary, a categorical class could be determined through defuzzification, e.g. by determining the center of this distribution gravity. In an effort to build an intelligent assistance system to pilots, Hudlicka and McNeese [24] used fuzzy rule knowledge base to assess the anxiety of a pilot from relevant static and dynamic data and observations that include task context, external events, personality, individual history, training, and physiological data. In this research, fuzzy rules are matched to generate numerical anxiety weight factors (AWFs) for different modalities expressed in the data. Then these AWFs are used to compute an overall anxiety level. This resulting anxiety value is mapped into three qualitative categories, i.e., high, medium, or low. Other works include emotion recognition from facial expression [50] and using facial and voice data [32]. Fuzzy sets theory is improved to handle uncertainty in input. It still has the similar problems as rule-based systems. Building up a complete and accurate rule base is an overwhelming task in a practical system.

(3) *Instance-based Learning*: Case/instance based learning is very straightforward in designing a classifier. Searching within a cases/instances pool for the most similar representatives to the new case does the classification. The closeness is determined by the distance between the new case from each representative, representing the dissimilarity between them determined by those predictor measurements. Scherer [45] designed an emotional analyst expert system, GENESE, based on the component process appraisal model [46]. The knowledge base consists of 14 vectors for 14 emotions, with quantified predictors for typical stimulus checks. In classification, subjects are asked 15 questions to determine the values for these checks. Then a Euclidean distance measure gives the distance of the new case to each emotion vector. The smaller this distance is, the more possible the current case belongs to the corresponding emotion. If the prediction is incorrect for current case, adjusting the corresponding emotion vector using the vector of this misclassified case does the training. Similarly, Petrushin [40] used k-nearest neighbors to predict emotions from a set of speech features. Instance/case based learning has strength in its natural way of designing classification algorithm and the decent performance in many applications. But the good performance of instance-based learning largely depends on carefully picking the predictor features and the representatives in the baseline model.

(4) *Regression*: Linear or nonlinear regressions in forms of logistic and probit regressions could be used in classification. These algorithms use regression functions to calculate the predicted target value for new cases, i.e. the affective category. In [34], authors described an emotion recognition and synthesis system using human speech input. In experiments, users repeat sentences in neutral and different emotional states. In training, the measured physical parameters

about the pitch contour and power envelop of these speeches are transformed into principal component forms. Then these parameters are used to estimate the coefficients in regression functions with these emotions as target variable, through multiple linear regressions. In recognition the input is the speech in the same words by the same person with unknown emotion or neutral state. The output from the regression functions indicates the corresponding emotion. Regression has been used as a common classification method. Like other augmented algorithms in the same family, regression suffers the distribution assumption, the computation cost for complex data, and the challenge from incomplete data.

(5) *Discriminant Analysis*: Discriminant analysis is a statistical classification algorithm used in multivariate analysis, which is based on comparing the Mahalanobi distance to different class representatives. A Mahalanobi distance is defined as the distance of the data point to the mean center of data points of the same class. Ark *et al.* [7] designed an emotion mouse to measure the user's affective state among happiness, surprise, anger, fear, sadness, and disgust. The four physiological measures are GSR, skin temperature, heart rate, and GSA. First the authors train a set of discriminant functions with these physiological measures as the predictors and the emotions as the classes. These functions are used to calculate the Mahalanobi distances to different emotion classes and accordingly to determine the membership of the current input. Again such algorithm is limited by its assumption of normal distribution. The reported prediction correct rate is only two thirds even in the case that uses the same training cases in testing data and no baseline normal cases in testing.

(6) *Neural Networks*: Neural networks (NNs) have been extensively used in pattern analysis. Using connection weights and processing functions, a collection of neurons could mimic complex input-output relationships between different node layers. A series of NN applications exist in facial expression recognition. Petrushin [39, 40] used acoustic features selected from human speech in call centers, including pitch, energy, speaking rate, formants, and also some descriptive statistics for them. The classifiers use a two-layer back-propagation neural network with 8, 10 or 14 input nodes, 10 or 20 nodes in the hidden sigmoid layer and 5 nodes in the output linear layer. This approach yields an average accuracy of about 65% with the following distribution of detection accuracies for emotional categories: 55-65% for normal, 60-70% for happiness, 60-80% for anger, 60-70% for sadness, and 25-50% for fear. In order to improve the performance, authors also used ensembles of NN classifiers applying voting strategy, and combinations of NN classifiers with each of them specially trained for only one emotion. In another research, Fellenz *et al.* [18] used a so called ASSESS system to process the acoustic features such as voice level, voice pitch, phrase, word, phoneme and feature boundaries, voice quality, temporal structure, and some facial features. Other efforts applied NN in facial expression analysis such as in [53]. Neural Networks have long achieved good performance in attacking many difficult problems. The disadvantages of them are mainly the required expertise in of the network structure and the training process, and the intense

computation. It is also hard to find clear explanations for the achieved models although studies already draw the mapping of various neural network structures to regression and clustering models.

(7) *Bayesian Models*: Bayesian approaches apply probabilistic theory and dependency into system modeling and learning. Given the evidence, Bayesian theorem calculates the posterior probability of a hypothesis using the prior probability of hypothesis and the conditional probability of the evidence on the hypothesis. In [42, 43], authors described a Bayesian classifier to predict the frustration level of users using the features of mean and variance of the sensory pressure on the mouse. The data distributions are modeled by a mixture of Gaussians. The Bayesian classifiers are augmented by switching among component classifiers according to different contexts. The training of the Bayesian classifiers is using an approximate algorithm, namely Expectation Propagation. The experimentation results show a little improvement compared with other global learning algorithms such as SVM, and are believed to be better than classical local learning algorithm such as k-nearest neighbors. Bayesian approaches provide a powerful modeling and prediction tool while normally the computation is intense. The Bayesian classifier in this case does not take the advantage of the local conditional independency among variables. In the much more simplified form, naïve Bayes classifier is used to predict emotions [48].

(8) *Bayesian Networks*: Bayesian networks use graphical models to summarize the prior knowledge of causal probability and conditional independency among the variables of a physical system. Bayesian inference based on Bayesian theorem is used to update the beliefs on the states of hidden and hypothetical variables given the observation of only partial evidence. In [8, 9], authors provided a Bayesian network model to assess the user's affective state in terms of the dimensions of valence and arousal, and the personality in terms of the dimensions of dominance and friendliness. The change of the states in these variables cause the change of observable evidence including facial and body movements, and acoustic and speech data of wording choice, speech characteristics, and input style characteristics. The network models the word selection more deeply by expanding it to an expression style including active, positive, terse, and strong expressions, and an interpretation layer of used paraphrases or concepts. This network could be implemented into a dynamic Bayesian network to capture the temporal emotion state structure. This research gives a good representation of a simple ASA model, based on which a more comprehensive Bayesian network model could be built to expand the representation structure both in depth and breadth.

Dynamic Bayesian networks (DBNs) add temporal causal links between hidden nodes in succeeding time slices. The hypothesis and hidden nodes in recent time slices could have impact on the corresponding nodes in current slice, thus simulating the decay or accumulation effects common in human emotional process. In [14, 15], authors designed a dynamic Bayesian network model for assessing students' emotion in educational games. Based on the OCC appraisal-

based emotion theory [36], this network models the emotions, as the appraisal results of how the current situation and action of help fits with the person's goal and preference. There are also body expressions and physiological sensors to provide additional evidence, such as the visual information, EMG, GSR, and heart rate measures. The emotion states in this study include joy, distress, pride, shame, admiration, and reproach. Each time the student performs an action, or the agent offers a help, a new time slice is added to the network. This research tries to combine an analytical emotion model into a synthetic assessment model using physiological measures. But the simple addition of two paradigms in a single Bayesian network is rigid in nature, and sometimes may be inappropriate, for example, if we consider the evidence can change the belief about the student action. The very fine grain size for describing the action and consequence is only appropriate for transient emotions when we consider the emotion of a user to some extent is stable. And the variant time interval between time slices may lead to requirement of representing time variant conditional probabilities. Furthermore in practice it is hard to know whether the action is satisfied or not.

(9) *Hidden Markov Model*: Hidden Markov Models (HMM) consider the situation where a Markov process is not directly observable. Instead, only emission signals from the underlying states are observed. A HMM is in fact equivalent to dynamic Bayesian network representation. Picard [41] conceptualized the use of HMM to model the emotion involvement process of users. The model has three emotional hidden states of interest, joy, and distress. The observation node could contain any sentic measures whose values change along with the underlying state, defined by the transmission probability. The transitional probability from one emotional state to another needs also to be defined. In [12, 13], the authors used a multiple HMM model to classify six emotions of happy, angry, sad, surprise, disgust, and fear. The input is the AUs defined in FACS for facial expression. At the lower level, six 5-state HMM models produce the state sequence from continuous AUs, one for each emotion. The high level HMM has seven states corresponding to the six emotions plus the neutral state, using the state sequence from the lower level as observation input. HMMs employ the Bayesian model in the basic form of two layer structure. They do not fully take into consideration the knowledge among variables and other variables influenced by or influencing emotion states. In Cohen's model, choosing the number of hidden states at the lower level HMM models is arbitrary, but important to the performance. The computational complexity will increase rapidly when more observation variables are combined into the model due to the full connection relationship between them and the hypothesis variables, because HMMs do not model the localization of dependency among variables.

C. Categorization of Affective State Assessment Models

Categorization of existing ASA models could take different merits. They could be intrusive or non-intrusive according to the awareness level of the human subject to the ASA task. Therefore the ASA models using physical expression

data such as those from visual, acoustic and behavioral modalities are non-intrusive at current technical level. The models could also be categorized according to the processing speed into real-time or post-processing applications. This depends on the specific information acquisition techniques used and the processing speed of assessment models. Here we categorize existing ASA models into two groups based on the usage and representation of prior knowledge. This categorization helps us understand the role and application of ASA models in HCI systems.

- The first group of ASA models uses the informative measures as predictors and applies classification algorithms without the prior and contextual knowledge among these variables and the affective states. Such approaches include regression, discriminant analysis, Neural Networks, Bayesian classifiers, and instance-based learning. Similar to them, decision trees, EM algorithm, and many other classification algorithms are potential candidates. These algorithms have very general and straightforward expressional and computation ability in the form of certain numerical functions. In assessment stage, these models could be easily converted into rule-based expert systems. The disadvantages include the lack of ability to handle uncertainty, complexity, and incompleteness involved in data sets. They could not take the advantage of special domain features. These features are critical to intelligent assistance systems where the uncertainty and complexity dominate and reshape the suitability of user models.
- The other group is represented by Bayesian networks and HMM models. They represent the prior domain knowledge and expertise into graphical networks. By incorporating the relationships among the subset of system variables, they also maintain the balance between the global and local representations. The localization of relevant variables and therefore the calculation resembles the intuition of human reasoning and logic processes. And the computational complexity can be reduced. On the other hand, most critiques come from the necessity of such domain knowledge for accurate models. In our view, however, such knowledge provides powerful capabilities in handling the complex situation in practical systems by the causal and uncertainty representations. Its representation mechanism is also close to the emotion and cognitive models from other expertise areas and thus it is easy to accommodate these models into a composite model as in the attempt to combine OCC emotion theory into Bayesian models by Conati [14].

IV. CASE STUDY: A GENERIC ASSESSMENT MODEL BASED ON DYNAMIC BAYESIAN NETWORK

Bayesian networks (BNs) are probabilistic graphical models representing dependencies as joint probabilities of random variables and independencies in their conditional relations within a network graph [38]. Those nodes characterize the hypothesis/goal variables, hidden state variables,

and evidence/observation variables in the physical system, while the arcs linking these nodes represent the causal dependency among these variables. Hypothesis nodes represent what we want to infer while the observation nodes represent sensory evidence. The intermediate hidden nodes are necessary to model the state generation process although in some cases they do not have explicit counterparts in the physical system. Static Bayesian Networks (SBNs) work with evidence and belief from a single time instant and thus are not particularly suited to modeling dynamic systems. Dynamic Bayesian networks (DBNs) overcome this limitation by connecting SBNs at different time slices. A Markov chain models the relationships between two neighboring time slices. The slice at the previous time provides prediction support for current slice through its temporal links, and it is used in conjunction with current evidence to infer the current hypothesis.

A. “Context-Affective State-Profile-Observation” Model

A generic assessment framework called “Context-Affective State-Profile-Observation” model to apply dynamic Bayesian networks in ASA is proposed in [30]. It is used to infer the user’s affective state from their observations (visual and other modalities if available). Shown in Fig. (2), this model captures the user’s profile, affective state, and the contextual information. Context component represents information about the specific environmental factors that can influence the user’s affective state, such as driving situations and user interface features. Profile component models user’s physical/personality character, ability and competitiveness in finishing the operations. This provides the adaptation capability of the model to individual users. Affective state component represents the user’s affective status. Typical affective states include but not limited to fatigue, confused, frustration, fear, sad, and anger. While we list negative affective states of safety interest, positive affective states may be useful in applications as in entertainment. Observation component includes sensory observations of different modalities describing user behaviors that are influenced by affective states. The figure shows some available visual modalities, while others such as verbal modalities could be used as well.

B. Advantages

This generic framework has the following advantages for user affective state assessment.

- Consideration of most relevant factors. In modeling the subject’s internal or affective state, this model incorporates the context, profile information into account. On the other hand, different from some other works previously discussed, we consider the stability of the subject’s affective state, and avoid the difficulty of depending on specific and transient task goals. We mainly rely on the power of external observations to recognize the human subject’s internal state, and let the profile and other context information help to improve the accuracy through online and offline training.
- Integration of more and more evidence. Applying Bayesian network model in recognizing affective

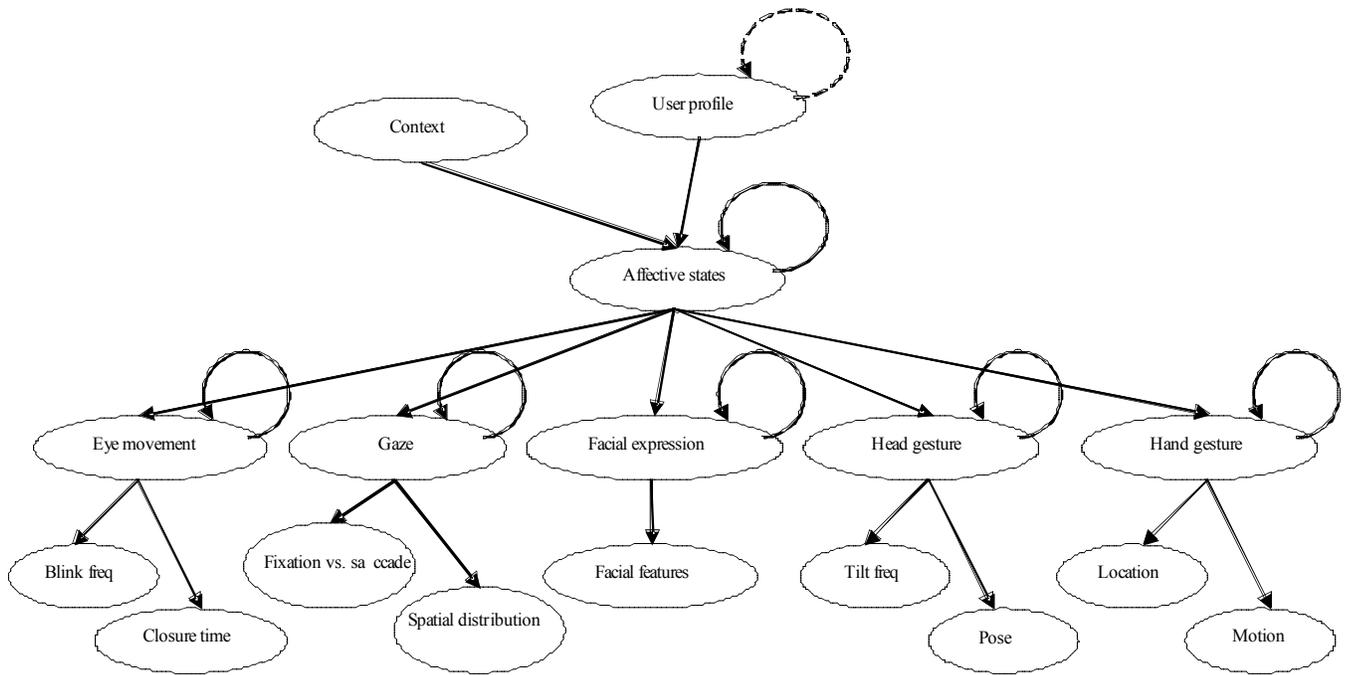


Fig. (2). The “Context-Profile-Affective State-Observation” model where self-pointing arrows indicate temporal links [30].

state has the advantage in handling the information uncertainty from multimodalities of the human subject. More new evidences could be integrated into this model once we identify their relationship to the affective state. It needs only a little effort to combine the new modules with the legacy system but will provide us with more accurate view. This could also include more context and profile aspects.

- Active strategy. Bayesian inference is normally considered as computationally intensive. In mission critical applications, computation and cost constraints associated with evidence acquisition and processing might prevent it from practice. Certain strategy could be used within this modeling framework to over time seek the most important evidence, in order to improve the efficiency of ASA tasks. Active fusion technologies selectively determine the best sensor subset based on the value of information and utility, which could distinguish the underlying affective state in a timely way while satisfying those constraints [26].

The proposed ASA model also provides a unified framework for intelligent assistance systems. Firstly this model does not only provide affective state detection but also certain diagnosis and synthesis functions. The beliefs about the hidden nodes, user profile and contextual nodes could be updated as well as the affective state nodes. These beliefs could provide partial insight about the causes for the affective problems. More nodes could also be added into the model to document relevant information such as the impact of the affective states on human performance and behavior. More important, assistance nodes and utility nodes can be integrated into the model to make it not only an ASA module but also the decision-making module to evaluate and choose

assistance when finding abnormal affective status. A simple assistance model is shown in Fig. (3).

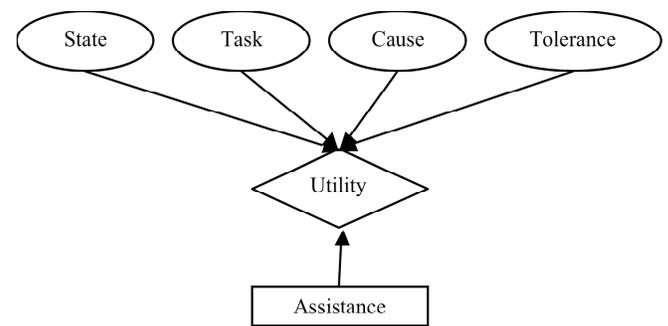


Fig. (3). A simple Bayesian assistance model taking in consideration the affective state, the current task, the cause explanation, and the tolerance of the human subject [30].

V. CHALLENGES AND FUTURE DIRECTIONS

A. Challenges and Problems

The state-of-the-art of HCI and user modeling is still embodied by the obvious gap between the abundance of computational models in research and the practicability of them in real-world task settings. For example, Microsoft carries out a lot of efforts into the user modeling research and applications based on Bayesian network models [22]. But we observe that most users soon turn off the automatic assistant that pops out in Microsoft Office, which runs a dynamic Bayesian network model to infer about the user’s need.

From the perspective of computational modeling, two major reasons relevant to affective state assessment are held in accounting for such disparity.

- Single user models are incapable of providing accurate enough and compatible results under the strict requirements of computing resources and speed. We make such a statement since we observe in our experiments that even with the carefully designed working procedures and paradigms, a single assessment model could not in some cases recognize the status of the subject very accurately and thus might fail to provide appropriate and urgent assistance. Moreover the validity of this ASA model cannot be verified and maintained by itself in practical applications. This is especially true when we consider the variability of individual personality, the configuration complexity for the large number of node states in Bayesian models, and the strict requirement on assistance systems for timely and appropriate decision making. This problem becomes more serious with the rapid increase of the complexity of computing systems in pervasive and distributed applications. Therefore an aggregate of these models is not only helpful but also necessary in addressing the challenges faced by intelligent systems. Further from the above assertion, the association and complementary alliance among the computational models have not been fully explored and exploited. In the past few years we begin to see the convergence of different expertise areas including AI, HCI, and cognitive science. The heterogeneous models capture different aspects of human beings and show different strengths for different tasks. For example, current statistical/probabilistic computational models are strong in hypothesis inference but lack cognitive fidelity and details; existing cognitive user modeling architectures provide plenty of situational information in simulating human cognitive process but assume normative performance and independent of user's emotional states. Systematically integrating these models of different types is the first critical step to gain insights and to stimulate further efforts into accurate and high-fidelity user models for practical intelligent systems.

B. Integrating Heterogeneous Models

We want to propose the study on organically integrating heterogeneous computational from various majors. As discussed before, cognitive models are another set of user modeling tools. However current ASA studies rarely use computational models from cognitive science, and thus do not exploit the potential of these models and the synergy with existing techniques to the full. So in this section we first give a brief introduction of cognitive models using ACT-R architecture as an example, and then give out the design framework of integrated ASA systems.

(1) *Cognitive Models*: Analytical cognitive models use production rule systems to simulate the memory and cognition management in human's cognition process by the rule retrieval and firing process. Of them, ACT-R has been used to model cognition processes during the interaction with task and artifact. The latest ACT-R 5.0 integrates perception and action with cognition and memory [5]. As shown in Fig. (4), ACT-R consists of several independent modules with associ-

ated buffers, and a pattern matcher, production selector, and production executor at its core. The visual and motor modules are responsible for interacting with the real world. The declarative memory module consists of facts while the procedural memory is made of condition/action productions. Buffers serve as the interface between the modules and the production system. The pattern matcher searches for the set of productions that match the current state of the buffers, selects one production, and executes it. In real world, the affective state of a human subject impacts the strategies in the above cognition processes and therefore these cognitive models can also function as affective state models.

In model tracing mode, the inference engine works to synchronize a computational cognitive model with the user's mind. Synchronizing model is accomplished by the cognitive state inference engine. The inference engine takes as input (a) the current state of the model and (b) the interpreted data stream. It attempts to find a series of actions that the model could take that would have produced that stream.

(2) *Integrated ASA Systems*: The following studies will greatly advance the whole area of affective state assessment and intelligent user assistance, emphasizing improving the performance of ASA. They include three coordination schemes of heterogeneous computational models from AI, HCI, and cognitive science, i.e. functionally coupled, independent, and hybrid schemes, according to the degree of functionality coupling.

- Independent. The most straightforward but also very efficient way to integrate multiple models is to combine the estimates of the same hypothesis from individual models. These models could be of the same type, e.g. two probabilistic models, or different types, e.g. a Bayesian model and a regression model. The individual models do not rely on each other in functioning. Since there is overlap among their inputs at each time instant, we can correlate the estimates from these models into a synergic score, which is better in accuracy than any estimate from these models. In the case of discrete estimates from individual models, a linear form of this fusion defines in fact a voting mechanism commonly used in meta-model management of multiple classifiers. A nonlinear fusion function could take many forms such as neural network models.
- Functionally coupled. Different models could fulfill different sub-functions during the course of user modeling. Data and beliefs are exchanged at various modeling element levels among these models, required for inference, induction, or usage of the hypothesis. An initial belief estimate of the hypothesis from one model could be input to another model. The second model may not be good at directly assessing the estimate from its input. But as a verifier it may use this estimate to adjust its modeling parameters and to evaluate such estimate in itself. During this process the conflict or mismatch could be detected and resolved, probably in a recursive mode, yielding a better estimate score. A design similar to that by Ji *et*

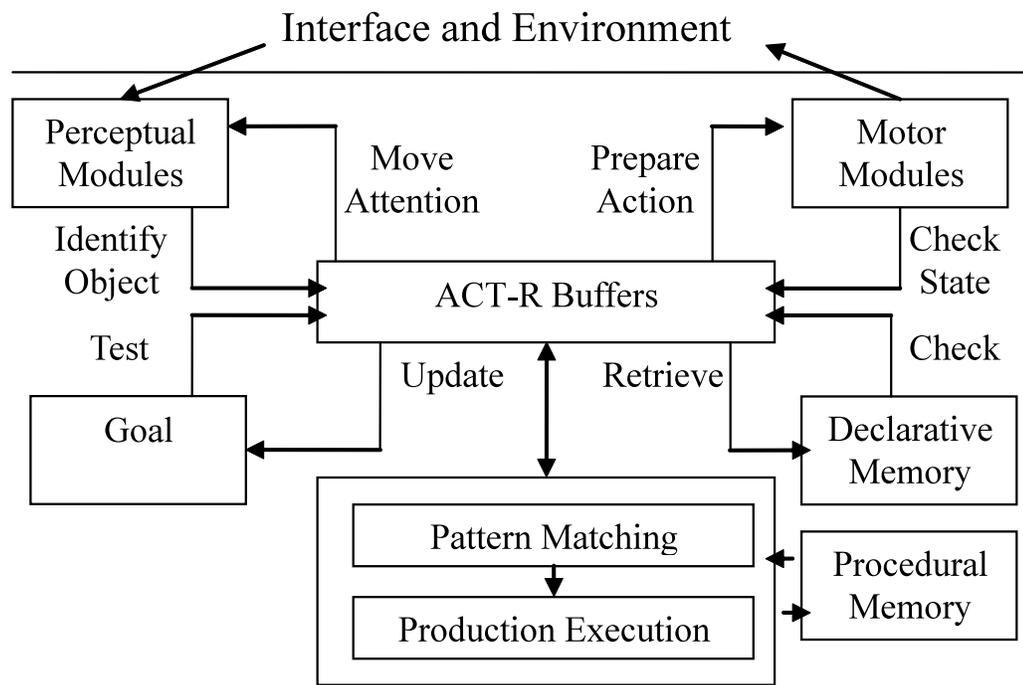


Fig. (4). The illustration of ACT-R/PM structure adapted from [10], where the information stored in the buffers associated with modules is used and changed by production rules.

al. [27] is given in Fig. (5), which uses a probabilistic model and a cognitive model. Integrating these two types of model in this way reflects the natural “guess-verification” strategy in human perceptual and cognitive activities. Take as an example the assistance to a driver from the view of an accompanying assistant. A human assistant constantly monitors the driver’s expression and behaviors such as facial and body expressions and driving operations, as well as the contextual information such as the time of day, the road and traffic situation. Combining the above evidence, the “guess” – the hypothesis about driver’s internal state of fatigue, distraction, or awareness of direction – is made. Then the assistant mimics the “suspicious” behaviors and intention of the driver through “common sense” or “sympathy” to confirm or reject such hypothesis. Surprisingly such a system works very well in most cases. Also in this process other side products such as the explanation and understanding of problem causes could be produced, which are essential in providing intelligent assistance.

- Hybrid. We can extend the integration into a hybrid scheme. More than one model fulfills the different sub-functions. For example, we use a set of models working independently to develop the initial estimate that is used in another set of models to verify and refine such estimate, until an accurate enough estimate is yielded.

C. An Integrated for Affective State Assessment and Intelligent Assistance

The overall structure of the proposed integrated fatigue modeling and performance prediction approach is shown in Fig. (6). Four organic modules fulfill the four core functions,

i.e., sleep regulation process modeling, fatigue detection, fatigue understanding and impact/performance assessment, and decision-making on countermeasures. These four functions apply sleep regulation process model, probabilistic model, cognitive/AI plan recognition model, and utility calculation, as underlying enablers. Other than the four modules, there is an information acquisition module in collecting and processing all necessary observation or information about the soldier. In our research, we will focus on the first three modules for sleep regulation, detection and understanding/performance assessment functions. The decision-making on countermeasures is out of the scope of this proposal, while our study will provide great insights and help for this function if applicable. An intelligent countermeasure decision-making module can be later developed and integrated into this comprehensive framework upon further grant support from appropriate funding agencies.

- The sleep regulation module models the fatigue accumulation and evolvment process in sleep regulation process. Variant biometmathematical models can be the underlying model for this module to describe the interaction of homeostatic, circadian, and ultradian variables. The output of this model emphasizes the qualitative description, in discrete values, of these and other relevant variables, and other associated soldier profile and working contextual knowledge that are used in the probabilistic fatigue detection module.
- The fatigue detection module uses a probabilistic dependency network model to model the dependency between the sleep regulation variables and profile/contextual variables, the fatigue hypothesis variable, and the observable evidence variables. This DBN model also incorporates temporal links. The evidence concentrates on physical and physiological

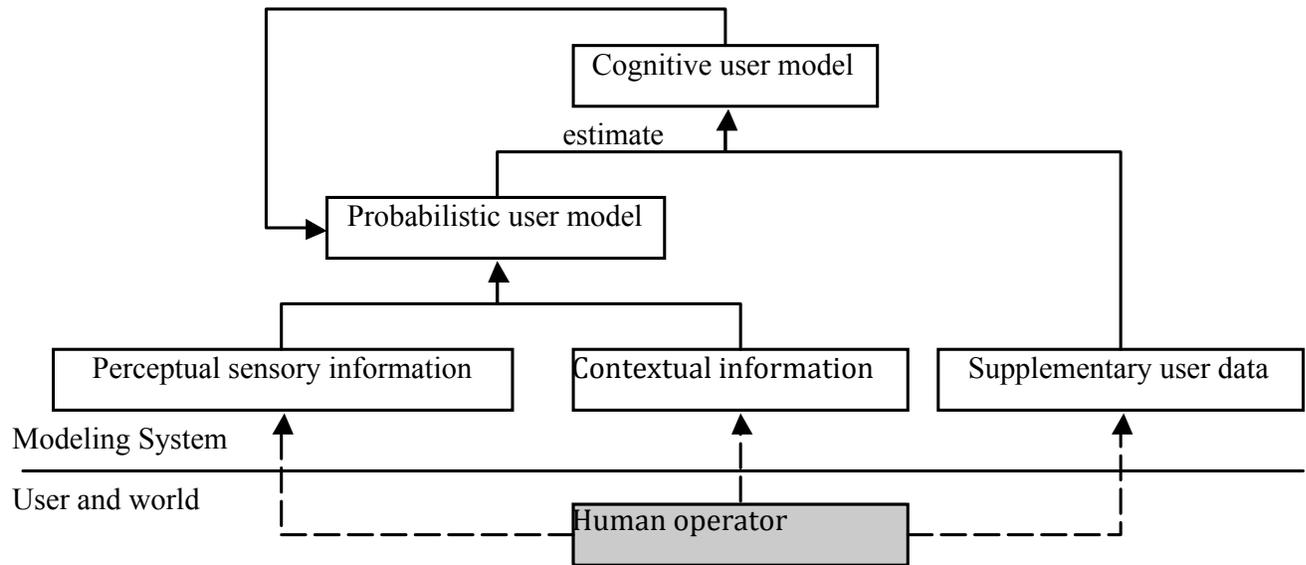


Fig. (5). Two user models that are functionally coupled, with one probabilistic model and one cognitive model.

measures of soldiers. Initial hypothesis belief is generated and updated over time with new evidence and feedback from the fatigue understanding module. The output from this module, the detected fatigue level, feeds the understanding and impact/performance assessment module.

- The fatigue understanding module applies a cognitive model built for specific tasks, such as a tank crew. This model reflects the influence of fatigue level to predict the interaction activities of soldiers with working artifacts, with the help of plan recognition model. The output of this module and the actual output of the soldier are compared. If there is no match achieved for these two, a conflict is reported. Based on this feedback, either a more accurate fatigue level is gained, or the validity of relevant models is checked to solve this conflict.
- The cognitive model and the plan recognition model in the impact/performance assessment module working together to assess the performance under the impact of fatigue presence. Quantitative metrics about the performance in the specific task can be defined and mapped from reaction times, errors, memory workload, choices, learning curves, and eye/hand movement in the cognitive model. Reliability measures in terms of confidence, probability and error ranges over time intervals can also be computed. This module and the above understanding module are necessary to provide domain specific knowledge in order to improve fatigue accuracy and interpreting performance. They are also important to making decisions on the type and timing of countermeasures, as discussed for the intelligent assistance system in [24].
- The countermeasure decision-making module determines the appropriate measures to counter the impact of fatigue. As we indicated before, currently this is out of the scope of this project although it can be an

organic component of this integrated framework. Based on information from previous modules, utility computation is a potential solution to support this function.

We can easily identify a generic set of key research issues associated with the above framework. The tasks in this project include: 1) customization to individual soldiers through the conflict resolution and model validation; 2) selecting or building appropriate cognitive model and plan recognition model for specific applications; 3) seeking feasible training, deploying, and managing strategy of the above core modules within this framework. Some details of these tasks will be given in the following sections.

D. Further Discussion

In future studies, the key task will become how to plan and manage the knowledge induction, exchange and fusion within and among these models. The strategy should be dynamic to accommodate the complexity of real-world. For this sake, a quality model need be developed, e.g., based on probability and utility theory to support the decision making in information fusion [16, 38]. Utility values represent the subjective estimate of the gain from the action/decision on fusion and information exchange, impacted by the uncertainty in the system represented by probabilities. With the definition of benefit and cost, we could choose the best decision by optimizing the expected utility. Such integration has the desired features and advantages summarized as follows.

- Integration of domain independent and dependent models. Some computational models such as Bayesian networks are domain independent and thus could be used in different applications. In the meanwhile, a domain specific model is necessary for explaining and understanding affective status in order to react to the subject's affective state correctly and accurately. This model has functions similar to the "Affective Understanding" described by Affective Computing

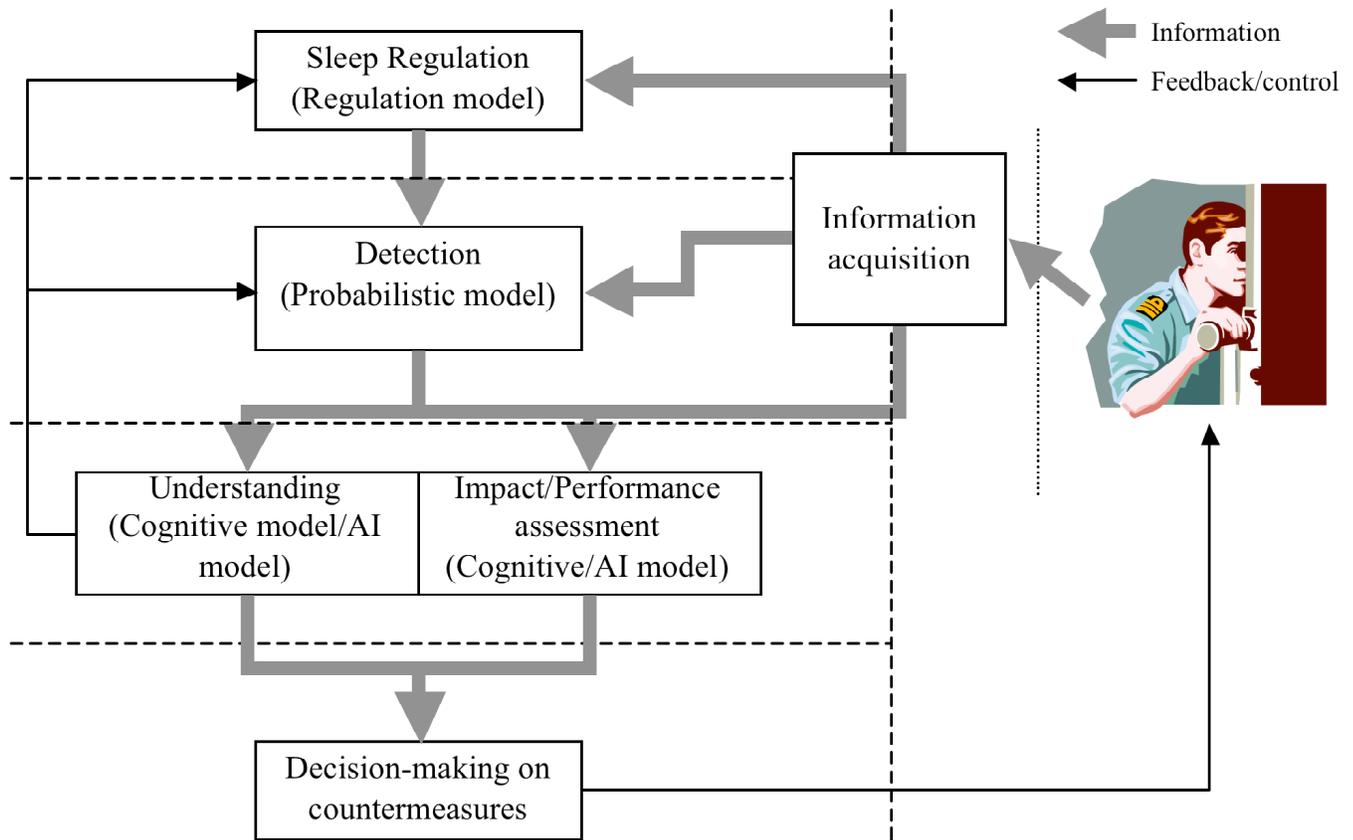


Fig. (6). The proposed integral framework for efficient affective state assessment.

Research Group [2] or the Belief Assessment component and the Impact Prediction component in ABAIS system by Hudlicka and McNeese [24]. Such a model is essential to provide task environment awareness for intelligent assistance. This domain dependent model captures the relevant application information, explains the causes of the problem, and predicts the impact of the problem in the task domain.

- Integration of analytical and synthetic models. Closely related to the above arguments, two weak points exist for a synthetic model such as stochastic models and must be addressed when we design and implement a practical HCI system. One problem is raised when we need the details in understanding the affect and related problem. Based on external observations only, the grain size of such models is normally not fine enough for this task. The other challenge is raised when we need to maintain the DBN model over the time and thus need to deal with the validity of model and parameters. We need a third party as an objective judge. An analytical user model would be very helpful to provide insight into the interactions among user, interface, and environment.

According to the type of input, ASA models can be domain specific, using knowledge of specific task and goal, or domain independent. In term of modeling mechanism, models can be analytical, which explicitly mimic the physical processes in the modeled system, or synthetic, which focus

on mapping input to output. The combination of statistical/probabilistic models, which are generally synthetic models and more domain independent, and cognitive models, which are analytical and domain specific, provides a very attractive solution.

Some relevant works towards the above research directions focus on cognitive models. Duric, *et al.* [17] present an initial design for intelligent HCI systems to integrate user behavioral and emotional information with a cognitive model implemented in ACT-R 5.0 architecture. The design proposed in [27] is a cognitive model able to reflect the influence of affective states, but preliminary in its consideration. Improving the accuracy of ASA is crucial to real-world systems and serves as the foundation of any practical deployment of such augmented cognitive systems.

VI. CONCLUSION

The research of affective computing and user modeling has been the thrust toward intelligent HCI systems with augmented cognition experience and performance of human users in dealing with the complexity of modern computing applications. Integrating and synergizing research efforts from various study fields would provide a promising playground for developing practical intelligent systems.

We have begun the study integrating our modeling framework with cognitive models in these directions. Our future work will be on the mechanisms to coordinate these heterogeneous models in understanding and explanation of

affective states, and in improving the user assistance efficiency.

REFERENCES

- [1] E. Aarts, R. Collier, E. van Loenen, and B. D. Ruyter, (Eds.) *Ambient Intelligence*, First European Symposium (EUSAI 2003), Veldhoven, Netherlands: Springer, 2003.
- [2] Affective Computing Research Group, "Affective Understanding: Modeling and responding to user affect", MIT Media Laboratory [online] Available: affect.media.mit.edu/AC_research/understanding.html [Accessed 2002].
- [3] P. E. Agre and S. J. Rosenschein, *Computational Theories of Interaction and Agency*, Cambridge, MA: MIT Press, 1996.
- [4] J. R. Anderson, *Rules of the Mind*, Hillsdale, NJ: Lawrence Erlbaum Associates, 1993.
- [5] J. R. Anderson, D. Bothell, M. D. Byrne, C. Lebiere, C., and Y. Qin, "An integrated theory of the mind", *Psychol. Rev.*, vol. 111, no. 4, pp. 1036-1060, 2004.
- [6] J. R. Anderson, A. T. Corbett, K. R. Koedinger, and R. Pelletier, Cognitive tutors: Lessons learned, *Learn. Sci.*, vol. 4, pp. 167-207, 1995.
- [7] W. Ark, D. C. Dryer, and D. J. Lu, "The emotion mouse", *Proceedings of HCI International '99*, Munich, Germany, 1999.
- [8] G. Ball and J. Breese, "Emotion and personality in a conversational agent", *Embodied Conversational Agents*, Cambridge, MA: MIT Press, 2000.
- [9] J. Breese and G. Ball, "Modeling Emotional State and Personality for Conversational Agents", Technical Report MSR-TR-98-41, Microsoft Corporation, 1998.
- [10] R. Budiu, "About ACT-R", ACT-R Web Site at Carnegie Melon University [online] Available <http://act-r.psy.cmu.edu/about/> [Accessed November 13, 2006].
- [11] S. Card, T. P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*, Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.
- [12] I. Cohen, A. Garg, and T. S. Huang, "Emotion recognition using multilevel-HMM", *NIPS Workshop on Affective Computing*, 2000.
- [13] I. Cohen, N. Sebe, A. Garg, and T. S. Huang, "Facial expression recognition from video sequences", *International conference on Multimedia and Expo (ICME '02)*, 2002.
- [14] C. Conati, "Probabilistic assessment of user's emotions in educational games", *App. Artif. Intell.*, vol. 16, no. 7-8, 2002.
- [15] C. Conati and X. Zhou, "Modeling students' emotions from cognitive appraisal in educational games", *Proceedings of 6th International Conference on Intelligent Tutoring Systems*, Biarritz, France, 2002.
- [16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.
- [17] Z. Duric, W. Gray, R. Heishman, F. Li, A. Rosenfield, M. J. Schoelles, C. Schunn, and H. Wechsler, "Integrating Perceptual and Cognitive Modeling for Adaptive and Intelligent Human-Computer Interaction", *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1272-1289, 2002.
- [18] W. A. Fellenz, J. G. Taylor, R. Cowie, E. Douglas-Cowie, F. Piat, S. Kollias, C. Orovas, and B. Apolloni, "On emotion recognition of faces and of speech using neural networks, fuzzy logic and the ASSESS system", *IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN)*, Como, Italy, July, 2000.
- [19] G. Fischer, "User modeling in Human-Computer Interaction", *User Model. User-Adapt. Interact.*, vol. 11, pp. 65-86, Kluwer, 2001.
- [20] M. A. Hearst, "Trends & controversies: Mixed-initiative interaction", *IEEE Intelligent Systems*, vol. 14, pp. 14-23, 1999.
- [21] E. Horvitz, "Uncertainty, action, and interaction: In pursuit of mixed-initiative computing", *IEEE Intelligent Systems*, Sept./Oct. Issue, IEEE Computer Society, 1999.
- [22] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, "The Lumiere project: Bayesian user modeling for inferring the goals and needs of software users", *Proceedings 14th Conference on Uncertainty in Artificial Intelligence*, pp. 256-265, San Francisco: Morgan Kaufmann, 1998.
- [23] E. Hudlicka, "To feel or not to feel: The role of affect in human-computer interaction", *Int. Hum. Comput. Stud.*, vol. 59, no. 1, pp. 1-32, 2003.
- [24] E. Hudlicka and M. D. McNeese, "Assessment of user affective and belief states for interface adaptation: Application to an air force pilot task", *User Model. User-Adapt. Interact.*, vol. 12, pp. 1-47, 2002.
- [25] A. Jameson, "Numerical Uncertainty Management in User and Student Modeling: An Overview of Systems and Issues", *User Model. User-Adapt. Interact.*, vol. 5, pp. 193-251, Kluwer, 1996.
- [26] F. V. Jensen, *Bayesian Networks and Decision Graphs*, New York: Springer, 2001.
- [27] Q. Ji, W. Gray, M. Guhu, and M. Schoelles, "Towards an integrated cognitive architecture for modeling and recognizing user affect", *AAAI Spring Symposium on Architectures for Modeling Emotion: Cross-disciplinary Foundations*, Stanford University, 2004.
- [28] A. Kobsa, "Generic user modeling systems", *User Model. User-Adapt. Interact.*, vol. 11, pp. 49-63, Kluwer, 2001.
- [29] X. Li and Q. Ji, "Active affective state detection and assistance with dynamic Bayesian networks", *3rd Workshop on Affective and Attitude User Modeling in conjunction with User Modeling 2003*, Pittsburgh, PA, USA, 2003.
- [30] X. Li and Q. Ji, "Active affective state detection and user assistance with dynamic Bayesian networks", *IEEE Trans. Syst. Man Cybernet.*, vol. 35, no. 1, pp. 93-105, 2005.
- [31] P. Maes and B. Schneiderman, "Direct manipulation vs interface agents: A debate", *Interactions*, vol. 4, no. 6, ACM Press, 1997.
- [32] D. W. Massaro, "Multimodal emotion perception: Analogous to speech processes", *Proceedings of the ISCA Workshop on Speech and Emotion*, pp. 114-121, Newcastle, Northern Ireland, 2000.
- [33] M. McCullough, *Digital Ground: Architecture, Pervasive Computing, and Environmental knowing*, Cambridge, Mass.: MIT Press, 2004.
- [34] T. Moriyama, H. Saito, and S. Ozawa, "Evaluation of the relation between emotional concepts and emotional parameters on speech", *IEICE J.*, vol. J82-D-II, no.10, pp.1710-1720, 1999.
- [35] B. A. Myers, "A brief history of Human Computer Interaction technology", *ACM Interactions*, vol. 5, no. 2, pp. 44-54, 1998.
- [36] A. Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotion*, Cambridge, UK: Cambridge University Press, 1988.
- [37] M. Pantic, I. Patras, and L. J. M. Rothkrantz, "Facial mimics recognition from face profile image sequences", Technical Report DKS-02-01, Data and Knowledge Systems group, Delft University of Technology, Netherlands, 2002.
- [38] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo, CA: Morgan Kaufmann, 1988.
- [39] V. A. Petrushin, "Emotion in speech: Recognition and application to call centers", *Artificial Neural Networks in Engineering (ANNIE '99)*, St. Louis, 1999.
- [40] V. A. Petrushin, "Emotion recognition agents in real world", *2000 AAAI Fall Symposium on Socially Intelligent Agents: Human in the Loop*, 2000.
- [41] R. Picard, *Affective Computing*, Cambridge, MA: MIT Press, 1997.
- [42] Y. Qi and R. Picard, "Context-sensitive Bayesian classifiers and application to mouse pressure pattern classification", *Int. Conf. On Pattern Recognition*, Quebec City, Canada, 2002.
- [43] Y. Qi, C. Reynolds, and R. Picard, "The Bayes point machine for computer-user frustration detection via PressureMouse", *Proceedings of the Workshop on Perceptive User Interfaces*, 2001.
- [44] D. D. Salvucci, E. R. Boer, and A. Liu, "Toward an integrated model of driver behavior in a cognitive architecture", *Transport. Res. Rec.*, vol. 1779, pp. 9-16, 2001.
- [45] K. R. Scherer, "Studying the emotion-antecedent appraisal process: An expert system approach", *Cogn. Emot.*, vol. 7, pp. 325-355, 1993.
- [46] K. R. Scherer, "On the nature and function of emotion: A component process approach", In K. R. Scherer & P. Ekman (Eds), *Approaches to Emotion*, pp. 293-317, Hillsdale, NJ: Erlbaum, 1984.
- [47] H. Schlosberg, "Three dimensions of emotions", *Psychol. Rev.*, vol. 61, no. 2, pp. 81-88, 1954.
- [48] N. Sebe, I. Cohen, A. Garg, and T. S. Huang, "Emotion recognition using a Cauchy naive Bayes classifier", *2002 International conference on Pattern Recognition (ICPR2002)*, 2002.
- [49] A. Shashua, Y. Gdalyahu, and G. Hayun, "Pedestrian Detection for Driving Assistance Systems: Single-frame Classification and System Level Performance", *Proc. of IEEE Intelligent Vehicle Symposium*, 2004.
- [50] N. Tsapatsoulis, K. Karpouzis, G. Stamou, F. Piat, and S. Kollias, "A Fuzzy System for Emotion Classification based on the MPEG-4 Facial Definition Parameter", *EUSIPCO 2000*.

- [51] O. Tsimhoni and Y. Liu, "Steering a Driving Simulator using the Queueing Network-Model Human Processor (QN-MHP)", *Proceedings of 2nd International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, 2003.
- [52] G. I. Webb, M. J. Pazzani, and D. Billsus, "Machine learning for user modeling", *User Modeling and User-Adapted Interaction*, vol. 11, pp. 19-29, 2001.
- [53] J. Zhao and G. Kearney, "Classifying facial movement by back-propagation neural networks with fuzzy inputs", *Proc. Intl. Conf. Neural Information Processing*, pp. 454-457, 1996.
- [54] I. Zukerman and D. Albrecht, "Predictive statistical models for user modeling" *User Modeling and User-Adapted Interaction*, vol. 11, no. 1-2, pp. 5-18, Kluwer, 2001.

Received: March 27, 2008

Revised: April 21, 2008

Accepted: May 15, 2008

© Xiangyang Li; Licensee *Bentham Open*.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.5/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.