Nonverbal Indicators of Malicious Intent: Affective Components for Interrogative Virtual Reality Training

Frederic (Rick) McKenzie
Electrical & Computer Engineering
Old Dominion University
231c Kaufman Hall
Norfolk, VA 23529
Phone: (757) 683-5590
Fax: (757) 683-3220
Email: fmckenzi@ece.odu.edu

Mark Scerbo
Department of Psychology
Old Dominion University
250 Mills Godwin Bldg
Norfolk, VA 23529
Phone: (757) 683-4217
Fax: (757) 683-5087
Email: mscerbo@odu.edu

Jean Catanzaro
Department of Psychology
Old Dominion University
250 Mills Godwin Bldg
Norfolk, VA 23529
Phone: (757) 683-5000
Email: jcatanza@odu.edu

Mark Phillips
VMASC
Old Dominion University
7000 College Drive
Suffolk, VA 23435
Phone: (757) 686-6200
Email: mphillips@odu.edu

KEYWORDS: affective models, virtual environments, emotion, body language, deception

Summary
Models of affective behavior are critical for the development of training systems that are designed to exercise social interactions. Potential applications include various security-oriented operations such as police interrogation, airport security, border crossings, and military peacekeeping. Aside from speech, humans also communicate through vocalizations and inflections as well as with body language. Such nonverbal communication can convey affect such as anger or nervousness that is important in identifying deception. In this research, a trainee is asked to perform check point duty and question drivers of vehicles about their identity and reasons for entering a secured area. Most of the encounters are routine and innocuous, but occasionally a scenario unfolds that requires additional interrogation and rapid decision-making the part of the trainee. These special scenarios require the individual to draw upon his/her knowledge of social interactions in order to make the proper decisions and act appropriately. Virtual environments that address this form of training are few. Accordingly, the present paper describes an ongoing program of research designed to generate affective states for intelligent agents, create affective component behaviors to convey cues for anger, nervousness, and deception, and provide a complex interrogative training environment to exercise judgment-based decision-making.

1.0 Introduction
Old Dominion University (ODU) has been performing research in the area of training using virtual environments. The research involves computer controlled virtual humans and live human participants taking part in an interrogative scenario whereby various tasks are trained and evaluated in a virtual environment. The encounters include interchanges where affective states exhibited by the virtual humans are vital to the success of complex training tasks. These complex

training tasks require individuals to exercise judgment regarding social interactions and make quick decisions based upon those interactions. The scenario used is a checkpoint operation in a typical third world urban area. The trainee is presented with a series of innocuous routine encounters. Occasionally, a scenario unfolds that appears slightly different but incorporates one of several fundamental training objectives. The participant must react or risk injury to himself or others. Importance is placed on cues that are precursors to aggression and/or hostile activities.

There are numerous nonverbal cues that convey information. The most obvious source of information is the face (Ekman 1999). Beyond the face, body posture and movements can also convey information. Although individuals may learn to control their facial expressions, they rarely mask their body language. The focus of this paper is to describe the use of affective computing in the development of higher fidelity behaviors that include the aspect of emotion in order to create a more complex environment for the trainee -- an environment more conducive to the training of judgment-based decision-making in social interactions.

Research has shown that humans are quite adept at identifying emotions in static line drawings [Wehrle, Kaiser, Schmidt, & Scherer 2000] and remarkably proficient at gleaning critical information from even the most impoverished dynamic displays [Barclay, Cutting, & Kozlowski 1978]. Thus, even a low fidelity simulation can result in positive training benefits, provided that the critical cues are present and the key behaviors are exercised. A goal of this research is the integration of intelligent agents technologies with virtual environments. As a consequence, instead of concentrating of the fidelity of the graphical models, this research concentrates on the fidelity of the behavioral models. Thus, high-fidelity human agents have been utilized from the Jack project at the University of Pennsylvania. Using Jack as its base of human physical movement, the research team has been developing an architecture that supports the incorporation of affective component behaviors into virtual environments.

2.0 Intelligent Agents in Jack

Jack is a 3D modeling environment with support for high degree of freedom human models. The extent of motion of the human models is always within the physical constraints of selectable human body types. As a result, one is assured of gestures and positions that are within the realm of possibility given the particular human in a particular environment. Agent controls in Jack are supported through layers of interfaces with differing complexity. At the lowest level, one may manipulate the joint angles of a human graphical model. At the higher levels, one may develop behaviors such as cough or wink that are constructed from lower level physical movements. A network of these executable behaviors provides the activities and reactions that the agent will exhibit during part or an entire application's scenario. The network consists of action nodes and decision components. A decision transition is affected by environmental stimuli that would influence the validity of a condition for transition. Nodes can also execute in parallel. Thus, the behavioral network is called a Parallel Transition Network (PatNet) [Badler et al 2000]. Another layer of generalized capability called Parameterized Action Representation (PAR) is also available in Jack that supports more automatic behavioral animation activated through natural language interfaces [Badler et al 2000].

Decision points occur throughout the transition network. A trainee might decide to search a vehicle at the checkpoint by telling the driver to open the trunk or allow the driver to continue. It is at these decision points where more intricate behaviors may be used to illicit judgment-based decision-making on the part of trainees. The Jack driver agent can be made to assume an affective state such as nervousness or other explicit cues of deception, in effect, providing a training basis for those cues.

3.0 Cues for Deception

One of the most important skills for individuals assigned to checkpoint duty is the ability to detect suspicious behavior. Gratch and others [Gratch 2001, Velasquez 1997] use facial features and storylines to express emotion but do not focus on the highly interactive 3D human models needed to convey those component behaviors unique to deceptive behavior. Most of the information put forth by suspicious behavior is not communicated verbally. Instead, it is conveyed through facial expressions, body language, and non-speech characteristics such as vocal inflections, stammering, and rate of speech. The ability to cover one's actions with the intent of carrying out an unexpected attack relies, in part, on deception and the ability to mask nonverbal indicators such as anger and anxiety. Research in body language demonstrates nonverbal behavior is critical for detecting such affect. For example, Ford (1996) states that deceitful statements are often associated with a decrease in hand movements.

The indicators of deception are not wholly unambiguous. In fact, Vrij and Heaven (1999) note one particular finding in which vocal and verbal indicators such as hesitations, speech errors, repetitions of the wrong word, and word slips differ depending on the complexity of the lie. Specifically, they found that liars made more speech hesitations and errors as compared to truth tellers when the lie was cognitively difficult, but made fewer speech hesitations when the lie was easy. Depending on the complexity of the lie, an individual may require more skill to carry out the deception successfully. A less skilled liar may be more likely to demonstrate nervous responses such as fidgeting, gaze aversion, eye blinking, and sweating.

Identifying deception is not tied solely to behavior, but must be interpreted within the context in which it occurs [Ekman 1997]. In sum, deception is a complex behavior that can be represented in numerous ways. Effective simulation of human deceptive behavior must regard this complexity by ensuring that a combination of deceptive cues can convey affect that is appropriate and can ensure that their intensity is properly matched with the environmental context and motives of the deceptive person.

4.0 Generating Affect for Deception

So much depends upon the context, the human, and human experiences that a generalized model of emotion is indeed a difficult task but computational models do exist. For example, Velasquez (1998) uses excitatory and inhibitory inputs to nonlinear functions of emotions while Elliot (1992) takes a rule-based approach to reasoning about affect. It is important to provide a flexible methodology to encode relevant human personality types and experiences within the context of a given scenario. One method is to take a system dynamics approach to modeling such behavior.

System dynamics is a modeling methodology that utilizes causal models to generate flow graphs that in turn may be translated to differential equations (See Forrester 1975). First, positive and negative influences are labeled in the causal model. Nodes within the causal model are then attributed to variables that imply accumulation and rate. These nodes are mapped to flow graph equivalences such as valves for fluid flow rates and tanks for fluid accumulation levels. The transition to differential equations is governed by an algorithm which dictates that the change of a level over time is equal to the flow into the level minus the flow out. The flow in or out is a function of the input variables to a given node. For example, Figure 1 shows a flow graph with one level and two rates. The clouds indicate sources and sinks of flows. In following the algorithm, the change of Level_1 over time is equivalent to R1 – R2 where R1 is equal to Level_1 times a constant and R2 is equal to a constant.

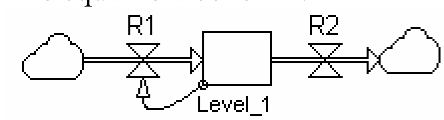

Figure 1: Flow Graph

This can be more aptly written in equation form as $\frac{dL1}{dt} = R1 - R2$, $R1 = k1 * L1$, $R2 = k2$ or

$\frac{dL1}{dt} = k1 * L1 - k2$, where $L1$ is Level_1.

Figure 2 shows a typical model that may be used to generate behavior. The figure shows causal influences that affect the increase or decrease of anger, nervousness, and deceptive behavior in a virtual agent. Models may be created a priori, stored, and invoked as needed. In this simplified flow graph-based system dynamics model, the triangles denote constant values while the circles denote auxiliary variables that serve to combine various input values. The model was executed using various input trajectories (circles with clocks) and initial conditions. The results show intuitive relationships as defined in the model.

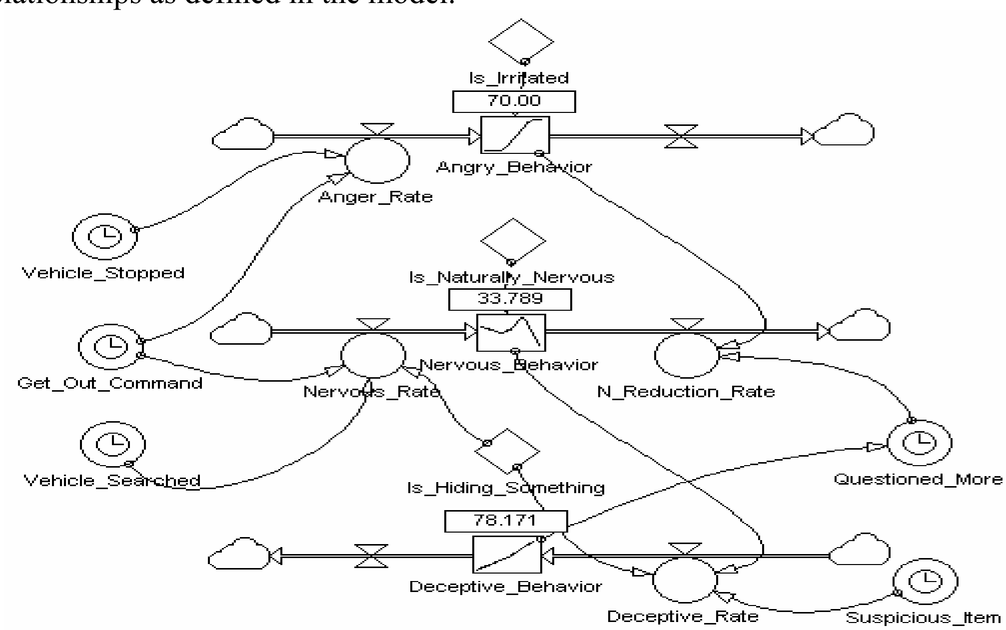

Figure 2: Systems Dynamics Model

Many different models may be applicable depending upon the scenario used for training and the complexity of emotion to be conveyed. Figure 3 shows the output of the model with initial values for anger and nervousness set to zero. The time scale indicated on the horizontal axis is in seconds. The driver is stopped and maybe told to get out of the car. The vehicle may be searched and a suspicious item may be found. All of these options are configurable in the scenario and the model. In the figure, deceptive behavior exceeds both angry and nervous behavior at the time of more questioning (about 80 seconds). Deceptive behavior continues to increase upon additional questioning and finding of a suspicious item. However, with increasingly more anger and nervousness levels, the exhibition of deceptive behavior may be masked resulting in less deceptive behavior.

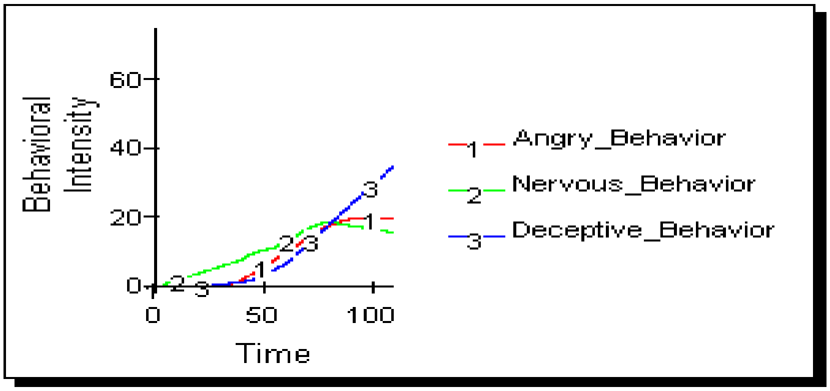

Figure 3: Emotion Level Output

The mathematical equations that are derived from the system dynamics model are shown below.

$$\frac{dA}{dt} = k_1(a+b) \qquad \frac{dN}{dt} = k_2(b+c+d) - k_3(eD+A) \qquad \frac{dD}{dt} = k_4 d(f+N)$$

A=Angry Behavior	N=Nervous Behavior	D=Deceptive Behavior
a=Vehicle Stopped	b=Get Out Command	c=Vehicle Searched
d=Is Hiding Something	e=Questioned More	f=Suspicious Item

At the moment, affect appraisal is accomplished by simply choosing the behavior that dominates in intensity by a given threshold. Once the type and intensity of behavior is determined, these are translated into movement elements such as fidgeting, gaze aversion, or eye blinking for the Jack agents in the VR environment by invoking the prescribed behaviors in sequence, repeatedly, or in parallel.

5.0 Ongoing Research

The training for this project takes place in a four wall immersive environment using CAVE technology. At present, the system incorporates speech recognition software and includes a focused natural language interface. The participants are armed with an inert replica of a handgun. Their movements within the environment are monitored by an Ascension Flock of Birds magnetic tracking system. This tracking information is provided back to the virtual agents. The technology allows for an extremely high level of interaction between trainee and the human models. These virtual agents answer questions, know where the trainees are in the environment, and reply while looking the trainees in their eyes.

The trainee approaches the car and asks the virtual driver for identification. The trainee's virtual partner provides cover for the trainee during the identity check. The driver produces an ID card and the trainee verifies that it is appropriate. A driver may appear nervous. At this point, the trainee must be able to distinguish nervous behavior from other potentially suspicious behaviors.

The component behaviors needed for exhibiting anger, nervousness, and deception are currently under development. It is intended that the system dynamics equations be encoded to directly influence the intensity and complexity of these behaviors.

6.0 Conclusions

Recent events have accentuated the need for more complex training involving the detection of individuals seeking to deceive. Deceptive behavior is difficult to discern and may be masked by common emotions. Research has shown that a number of distinct actions may contribute to the exhibition of a given emotion and that some behavior associated with deception is also shared with other emotions such as anger or nervousness. These actions or component behaviors serve as cues that help sensitize trainees to the nuances of deception. A flexible methodology of incorporating affective component behavior into agent models using system dynamics to drive the selection and intensity of these components will help to produce the complex scenarios needed to train and detect deception.

7.0 Acknowledgements

This research is made possible by Office of Naval Research Grant N00014-95-1-1044. The authors would like to thank R. Bowen Loftin for his insight and guidance.

8.0 References

Badler, N.; Bindiganavale, R.; Allbeck, J.; Schuler, W.; Zhao, L.; Palmer, M. *A parameterized action representation for virtual human agents.* In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill (eds.), Embodied Conversational Agents, MIT Press, 2000, pp. 256-284.

Barclay, C.D., Cutting, J.E., & Kozlowski, L.T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics, 23*, 145-152.

Ekman, P. (1997). Lying and deception. In Stein, N.L., Ornstein, P.A., Tversky, B., & Brainerd, C. *Chapter 14: Memory for Everyday and Emotional Events*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Ekman, P. (1999) Facial expressions. In T. Dalgleish and T. Power (Eds.), *The handbook of cognition and emotion* (pp. 301-320), Sussex, U.K.: John Wiley & Sons, Ltd.

Elliot, C. D. 1992. The affective reasoner: a process model of emotions in a multi-agent system. Ph. D. Thesis. Northwestern University.

Ford, C.V. (1996). Lies! lies! lies!: The psychology of deceit. Washington, DC: American Psychiatric Press.

Forrester, J. W. (1975). *Collected papers of Jay W. Forrester.* Cambridge, MA: Wright-Allen Press, Inc.

Gratch, J. & Marsella, S. 2001. Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In Proceedings of the Fifth International Conference on Autonomous Agents. Montreal, Canada. May 28 – June 01, 2001.

Velasquez, J. 1997. Modeling emotions and other motivations in Synthetic agents. In Proceedings of American Association for Artificial Intelligence (AAAI) 1997.

Velasquez, J. 1998. When robots weep: emotional memories and decision-making. In Proceedings of American Association for Artificial Intelligence (AAAI) 1998.

Vrij, A., & Heaven, S. (1999). Vocal and verbal indicators of deception as a function of lie complexity. *Psychology Crime & Law, 5*(3), 203-215.

Wehrle, T., Kaiser, S., Schmidt, S., Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality & Social Psychology, 78,* 105-119.