Generating Nonverbal Indicators of Deception in Virtual Reality Training

Frederic (Rick) McKenzie
Electrical & Computer Engineering Dept.
Old Dominion University
231c Kaufman Hall
Norfolk, VA 23529
fmckenzi@ece.odu.edu

Mark Scerbo
Department of Psychology
Old Dominion University
250 Mills Godwin Bldg
Norfolk, VA 23529
mscerbo@odu.edu

Jean Catanzaro
Department of Psychology
Old Dominion University
250 Mills Godwin Bldg
Norfolk, VA 23529
jcatanza@odu.edu

ABSTRACT

Old Dominion University (ODU) has been performing research in the area of training using virtual environments. The research involves both computer controlled agents and human participants taking part in a peacekeeping scenario whereby various skills-based tasks are trained and evaluated in a virtual environment. The scenario used is a checkpoint operation in a typical third world urban area. The trainee is presented with innocuous encounters until a slightly noticeable but highly important change surfaces and the trainee must react in an appropriate fashion or risk injury to himself or his teammate. Although the tasks are mainly skill-based, many are closely related to a judgment that the trainee must make. In fact, judgment-based tasks are becoming prevalent and are also far more difficult to train and not well understood. Of interest is an understanding of these additional constraints encountered that illicit emotional response in judgment-based military scenarios. This paper describes ongoing research in creating affective component behaviors used to convey cues for anger, nervousness, and deception in Operations Other than War (OOTW) training.

Keywords

human behavior representation (HBR), virtual environments, training, emotion, body language, deception

1. 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 a peacekeeping scenario whereby various tasks are trained and evaluated in a virtual environment. The scenario used is a checkpoint operation in a typical third world urban area. The trainee is presented with innocuous encounters until a slightly noticeable but highly important change surfaces and he must react or risk injury to himself or others.

Specifically, the goal is to address both culturally independent and dependent cues of nonverbal communication and recreate them in training scenarios. The focus will be 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 may be the face. 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. There are also numerous vocal cues that convey information. These cues are fairly universal; thus, one does not need to understand a foreign language to interpret these cues. Higher fidelity behaviors are needed 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.

Component behaviors are needed for current human models to appropriately mimic nonverbal cues important to the judgment-based scenarios. When combined with the already available general kinesthetic motion behaviors in the human models used, these component behaviors will convey particular emotion in virtual agents thereby allowing a complete interpretation of the judgment-based vignettes. Figure 1 illustrates the component behavior architecture.

Journal of WSCG, Vol.11, No.1., ISSN 1213-6972
WSCG'2003, February 3-7, 2003, Plzen, Czech Republic.

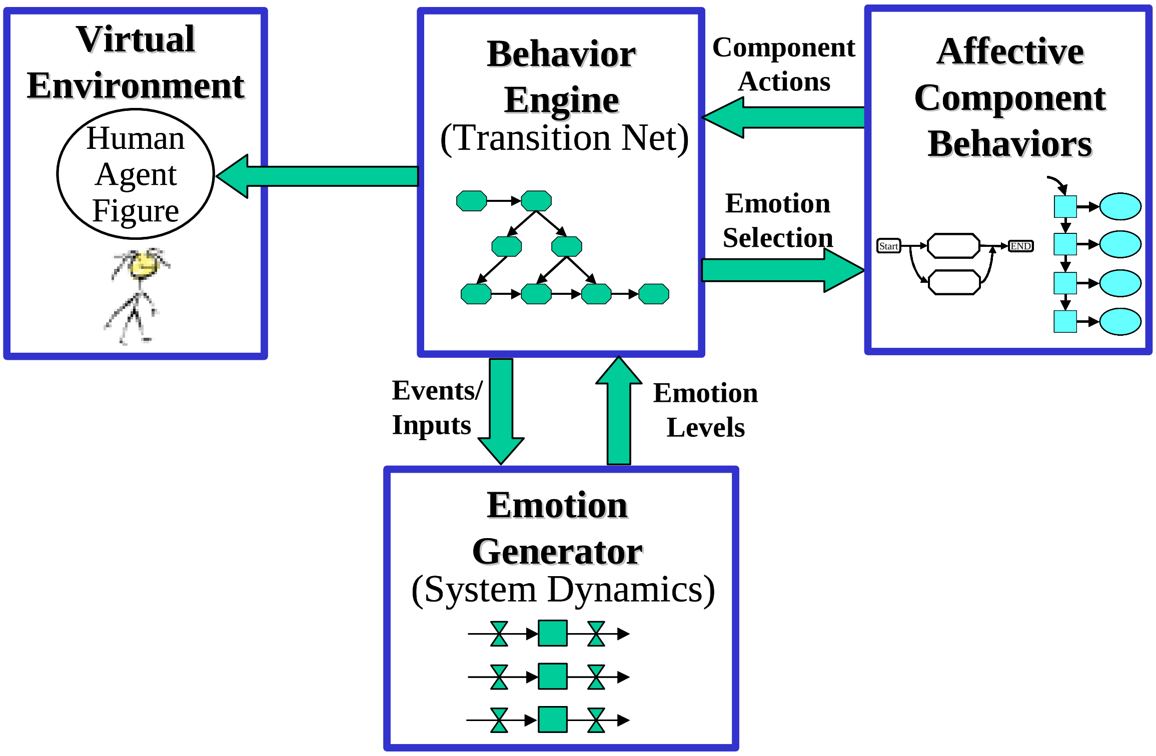

Figure 1: Component Behavior Architecture

The behavior engine drives the movements and interactions of the virtual human agents as well as other models in the virtual environment. These two components are an integral part of the Jack modeling methodology. Jack is a toolkit known for modeling high-resolution virtual humans. The scenarios in this research utilize the Jack toolkit and its flexible behavior modeling techniques to integrate affective component behaviors.

Such component body movement, body language, and interactions between the agent, its environment, and other agents, are critical to behavioral modeling. These behaviors are combined and imparted into the Jack agents in order to convey a particular emotion; thereby, making the scenarios more complex, realistic, and judgment oriented.

The behavior engine provides event and state information to the emotion generator which in turn determines the emotional reactions of a given agent. The associated emotion levels are returned to the behavior engine and a prevailing emotion is selected based on preset thresholds.

The prevailing emotion may also be associated with an indication of the intensity of that emotion. Upon selecting a particular emotion the behavior engine accesses the affective component behaviors module in order to activate the group of component actions needed to effectively convey the selected emotion at the indicated intensity. Also associated is the manner in which each action is to be activated. This paper discusses the need for and implementation of judgment-based scenarios and in particular distinguishing and conveying deceptive behavior in virtual reality human models.

2. Virtual Reality Training

Research has shown that humans are quite adept at identifying emotions in static line drawings [Weh00a] and remarkably proficient at gleaning critical information from even the most impoverished dynamic displays [Bar78a]. 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, hi-fidelity human agents have been utilized from the Jack project at the University of Pennsylvania. Concurrently, the research team has been developing an architecture that supports the incorporation of affective component behaviors into virtual environments.

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. Behaviors in Jack are supported through layers of interfaces with differing complexity. A network of these executable behaviors provides the activities and reactions that the agent will exhibit during part or possibly throughout an application's scenario. The network consists of basic transition nodes as well as nodes that can execute in parallel. Thus the behavioral network is called a Parallel Transition Network (PatNet) [Bad00a].

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 might decide to 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 exhibit nervousness or explicit cues of deception, in effect providing a training basis for those cues to the trainee.

The basic training objective is to monitor all ingress into a fictional town. A digital terrain database of the Quantico MOUT site called Combat Town (Figure 2) was enhanced with photographic textures of structures and features as well as additional buildings currently existing at the MOUT site to add realism to the geometry. Participants are briefed regarding their deployment, their duties and responsibilities, and the rules of engagement in effect. The participant's role is to act in the role of a guard and stop each vehicle as it approaches the checkpoint, check and verify the identities of all persons seeking access to a town, and clear and/or deny access to all vehicles that appear suspicious. Participants perform their duties in two-person teams comprised of the trainee and their Jack agent partner. Figure 3 shows an initial configuration of the scenario with a Jack driver navigating through concrete barriers in order to approach the checkpoint.

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.

Figure 2: Combat Town Model

Figure 3: Initial Scenario Configuration

3. THE SCENARIOS

In actuality, the process of manning a checkpoint can be a highly repetitive, mundane activity. Cars approach a gated checkpoint, a soldier stops the vehicle and asks the driver for his or her identification, verifies the diver's identity, and admits the driver. Nothing out of the ordinary transpires.

The general virtual scenario is designed to replicate that experience. The scenario begins with a car that approaches the checkpoint. The car comes to a halt. The trainee approaches the car and asks the driver, an avatar, for identification. The trainee's avatar partner provides cover for the trainee during the identity check. The driver produces an ID card and the trainee verifies that it is appropriate. The scenario ends when the trainee allows the driver passage to the town.

In actual checkpoint operations, it would be unusual to admit the same car and driver more than once a day. Therefore, a pool of neutral scenarios was generated that varies in vehicle type, vehicle color, driver's sex, skin color, hair color, and shirt color. In addition, the location where the driver's ID is kept also varies. It is important to remember that although the characteristics of the vehicle and driver can vary in each instance of the general scenario, the execution of the scenario remains the same. The trainee requests identification, verifies the information, and allows passage. Thus, the trainee can conduct numerous routine checks without encountering the same scenario more than once.

Critical scenarios were developed which address the specific training objectives. These are designed to exercise the trainee's skills and judgment with respect to their powers of observation, their ability to follow standard operating procedures, and their decision-making ability. In some scenarios, the trainee must respond in a specific manner while in other scenarios, the context was more ambiguous and the trainee must make a decision, act upon it, and subsequently be able to defend his/her actions.

Each critical scenario begins exactly the same way as the general neutral scenario, but then begins to deviate from that script. Thus, the trainee is presented with specific cues that require a different set of responses from those of the neutral scenarios. Some are quite obvious. For example, in one critical scenario when the trainee asks the driver for his/her identification, the driver responds in a language other than English. The trainee must then repeat the request using a translation card. In other critical scenarios, the cues are more subtle. For example, a driver may appear nervous. In this instance, the trainee must be able to distinguish nervous behavior from other potentially suspicious behaviors.

Powers of observation. These scenarios train the ability to detect suspicious cues. In some scenarios the trainee must be able to detect the presence of suspicious objects (e.g., a crow bar on the back seat of the car) and in other cases, the must check for the absence of objects (e.g., a license plate). Critical scenarios also require the trainee to detect suspicious behavior. Because this is the primary focus of the present paper, this class of scenarios will be treated in more detail below.

Standard operating procedures. Once a suspicious cue has been detected, the trainee must make decide whether to clear the vehicle and allow or deny passage. There are standard operating procedures for clearing a vehicle and some of the critical scenarios allow the trainee to perform those activities. Thus, in this study if the trainee decides to clear the vehicle, then he/she must ask the driver to step out of the vehicle, place the driver under his/her partner's cover, and inspect the inside the vehicle. The trainee must then ask the driver to empty his/her pockets. If no suspicious objects are found, the trainee then walks to the rear of the vehicle and asks the driver to open the trunk. The trainee must then inspect the contents of the trunk. If no suspicious objects are found, the driver is then told to close the trunk and to get back into the vehicle. If the trainee declares the vehicle and driver clear, he/she can then allow the driver to pass.

Judgment and decision making. The trainee's judgment and decision-making abilities are tested in critical scenarios in which the course of events unfolds less predictably. In some critical scenarios, information needed to make appropriate decisions may be invalid or not present at all. For instance, the driver may present inappropriate identification or background events may distract the trainee and prevent him/her from obtaining all of the necessary information. Other scenarios are designed to determine how the trainee will respond when standard operating procedures conflict with one's sense of ethics. For example, in one critical scenario an injured driver appears at the gate without proper identification. The driver pleads for access in order to seek medical attention. The trainee must weigh the requirement to follow proper procedure against the urgent needs of the driver and take appropriate action.

4. SUSPICIOUS BEHAVIOR

One of the most important skills for soldiers assigned to checkpoint duties is the ability to detect suspicious behavior. Most of the information conveyed 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. In terms of terrorism, the ability to interpret human nonverbal behavior is critical to survival. Accurately reading nonverbal indicators facial expression and body language of expressions is challenging, especially when a person is intentionally attempting to cover up their intent.

Research has shown that voice in and of itself contains useful predictive information independent of its semantic content. Research indicates that speech errors, speech rate, and verbal quantity are good indicators of deceptive activity. In general, it is found that persons who are nervous or attempting to be deceptive may manifest that nervousness with speaking faster and speaking more often to fill in silences that they are even more uncomfortable with than normal. On the other hand, Mehrabian (1971) has found that moderate discomfort such as in persons who were encouraged to be deceitful elicits more speech errors but with both a shorter duration of speech as well as a lower speech rate than those who were not instructed to be deceitful [Dru82a]. In one area, researchers have examined the 'ahs' in speech and have found that people are more prone to say 'ah' when in anxiety ridden situations than in non-anxiety prone situations. Such research has applications for military security in that the language barrier may bar certain guards from being able to detect anxiety utterances if they are not in the guard's native language.

Gratch and others [Gra01a, Vel97a] use facial features but do not focus much on the 3D models that exhibit the full body component behaviors that accompany various emotions while this research focuses on those component behaviors and addresses the issue of exhibiting deceptive behavior. Research in body language demonstrates nonverbal behavior is critical for detecting deceptive behavior. For example, Ford (1996) states that deceitful statements are often associated with a decrease in hand movements. In a scenario such as investigative decision-making or checkpoint training simulations, these factors could play an important role in generalizability of simulation training to real world application.

Nonverbal behavior is a significant clue for potential action. Therefore, it becomes necessary to be able to read even the most subtle of body movements. When conversing with another person, we gauge how our comments and conversation are received by reading each other's facial expression, and yet in studies of deceit, the human face is the one channel of communication that we are probably most familiar with and therefore, in deceitful behavior such as that made by a covert operation by a terrorist, the terrorist is likely to manage his facial expression very consciously.

Deceptive behavior refers to a behavior in which a person intentionally misleads another person to believe something that is not true of the real word. Research has indicated numerous indicators that are often associated with deceptive behavior. These nonverbal behaviors, in and of themselves do not prove deceit, but are useful when placed in situational context and also the context of other behaviors coinciding with the particular behavioral indicator.

In other words, identifying deceptive behavior is not tied to a single behavior, but rather it must be interpreted in terms of context [Ekm97a].

When a person is being deceptive, it is likely that the person is investing a lot more of their mental effort into what to say than in other situations and is less likely to use hand movements. Thus, decreased presence of hand movements is often associated with deception [For96a]. Ekman (1997) suggests that the hands will be used less to illustrate speech and furthermore, that voice intonation will flatten. Deception may involve other significant deviations from the person's normal behavior such as the appearance of "pauses, gaze aversion, speech disfluencies, and speech mannerisms," all or some of which "may all increase over what is usual for that person" [Ekm97a].

The indicators of deception are not black and white indicators, 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 might differ in occurrence depending on the complexity of the lie. "Liars made more speech hesitations and speech errors (compared to truth tellers) when the lie was cognitively difficult and made fewer speech hesitations (compared to truth tellers) when the lie was easy [Vri99a]. Table 1 shows various indicators of emotion that when put in context combines to exhibit a degree of emotion.

Depending on the complexity of the lie, the lie teller may require more practice or sophistication in order to carry out the deception successfully. A less sophisticated or practiced liar may be prone to demonstrate nervous responses such as fidgeting, gaze aversion, eye blinking, and sweating. In sum, deception is a complex behavior that can be behaviorally represented in numerous ways. Effective simulation of human deceptive behavior must regard this complexity by ensuring that the combination of deceptive cues is appropriate both in combination as well as ensuring that their intensity is properly matched with the environmental context and motives of the deceptive person.

5. GENERATING SUSPICIOUS BEHAVIOR

Nonverbal component behaviors that are combined with other agent actions to convey emotion are integrated using either parallel nodes or monitors. These are additional capabilities of Jack transition nets that respectively allow parallel and unsequenced actions to occur during the execution of a behavior. Figure 4 illustrates a parallel node.

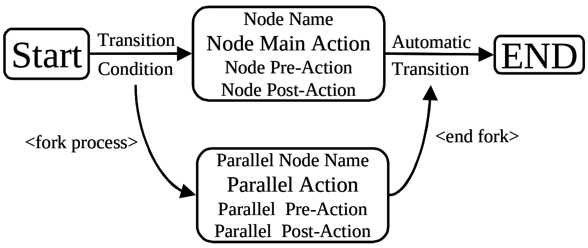

Figure 4: Transition Net Parallel Node

These may be used to combine a number of the nonverbal component behaviors together to achieve increased exhibition of a particular emotion.

Indicator	Neutral	Angry	Nervous	Deception
Trunk Swivels		Positive		
Fidgeting with Glasses			Positive	Positive
Speech Hesitations	Positive			
Rapid Eye Blinking				Positive
Looking Away While Speaking				Positive
Eyebrows Drawn Together		Positive		
Gaze Down				Positive
Facial Muscle Tension		Positive		
High Voice Pitch		Positive		
Stuttering			Positive	

Table 1: Non-Verbal Component Behaviors

Additionally, monitors may be used to trigger agent emotions when known events occur or given conditions are met. Monitors (Figure 5) are stored as a list of function pointer pairs with a trigger condition method and a corresponding behavior action method in each pair.

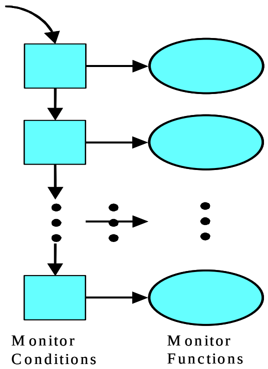

Figure 5: Transition Net Monitors

6. EMOTION GENERATOR

A generalized model of human emotion is indeed an intractable task yet computational models do exist [Vel98a, Ell92a]. So much depends upon the context, the human, and human experiences. It is important to provide a flexible methodology to encode limited, 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 which in turn may be translated to differential equations. 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.

Figure 6 shows a typical causal model that may be used to generate emotions as part of the emotion generator module. Many different models may be applicable depending upon the scenario used for training and the complexity of emotion to be conveyed. Models may be created a priori, stored, and invoked as needed. However, causal models must first be translated into their flow graph counterparts before they can be used. The figure shows causal influences that affect the increase or decrease of anger, nervousness, and deceptive behavior in the Jack agent driver that will interact with trainee checkpoint guards.

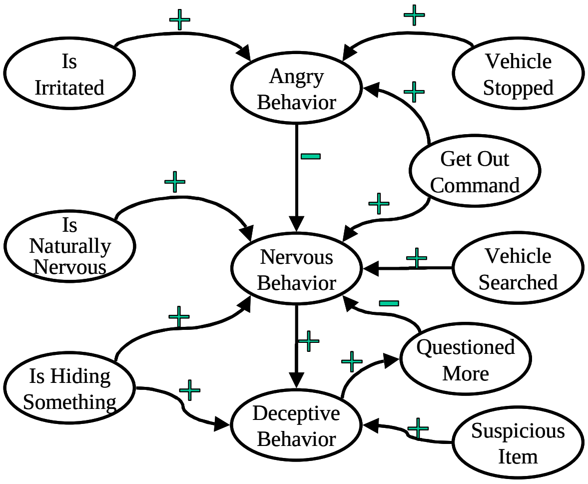

Figure 6: Causal Model

This simplified causal model is translated into the flow graph-based system dynamics model below. The triangles in the graph below denote constant values while the circles denote auxiliary variables that serve to combine various input values. The results show intuitive relationships as indicated in the causal model.

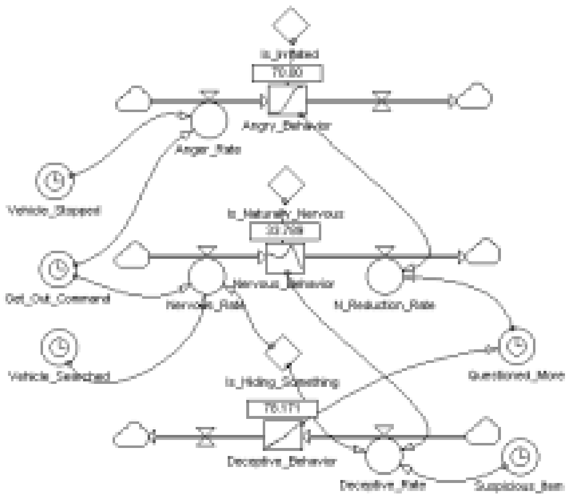

Figure 7: Systems Dynamics Model

Figure 8 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. In the figure, deceptive behavior exceeds both angry and nervous behavior at the time of more questioning. 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.

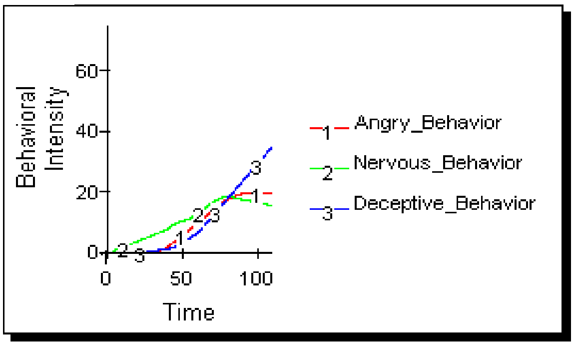

Figure 8: Emotion Level Output

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

$$\frac{dA}{dt} = k_1(a+b)$$

$$\frac{dN}{dt} = k_2(b+c+d) - k_3(eD+A)$$

$$\frac{dD}{dt} = k_4 d(f+N)$$

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

7. 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.

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.

8. ACKNOWLEDGMENTS

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.

9. REFERENCES

[Bad00a] 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.

[Bar78a] 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.

[Dru82a] Druckman, D.; Rozelle, R.M.; Baxter, J.C. (1982) Nonverbal Communication. Beverly Hills, CA: Sage Publications.

[Ekm97a] 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.

[Ell92a] Elliot, C. D. 1992. The affective reasoner: a process model of emotions in a multi-agent system. Ph. D. Thesis. Northwestern University.

[For96a] Ford, C.V. (1996). Lies! lies! lies!: The psychology of deceit. Washington, DC: American Psychiatric Press.

[Gra01a] 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. 2001.

[Meh71a] Mehrabian, A. (1971). Nonverbal betrayal of feeling. Journal of Experimental Research in Personality, 5, 127-138.

[Vel97a] Velasquez, J. 1997. Modeling emotions and other motivations in Synthetic agents. In Proceedings of American Association for Artificial Intelligence (AAAI) 1997.

[Vel98a] Velasquez, J. 1998. When robots weep: emotional memories and decision-making. In Proceedings of American Association for Artificial Intelligence (AAAI) 1998.

[Vri99a] Vrij, A., & Heaven, S. (1999). Vocal and Verbal Indicators of Deception as a Function of Lie Complexity. Psychology Crime & Law, 5(3), 203-215.

[Weh00a] 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.