

Consistent Communication with Control

Jan M. Allbeck and Norman I. Badler
Center for Human Modeling and Simulation
University of Pennsylvania
200 S. 33rd St., Philadelphia, PA 19104-6389
allbeck@graphics.cis.upenn.edu

ABSTRACT

We are seeking to outline a framework to create embodied agents with consistency both in terms of human actions and communications in general and individual humans in particular. Our goal is to drive this consistent behavior from internal or cognitive models of the agents. First, we describe channels of non-verbal communication and related research in embodied agents. We then describe cognitive processes that can be used to coordinate these channels of communication and create consistent behavior.

1. INTRODUCTION

Consciously or not, people spend their entire lives observing other people. We have unconscious and cultural norms of human behavior and are more likely to notice the unexpected rather than the natural or expected. In the *embodied agent* research community, agent behaviors created with an attempt to conform to our nominal expectations are termed *believable*. Unfortunately, this term is awkward to define. *Believable* means “to accept as real.” But *real* is itself a loaded term, as there are numerous aspects of real people that embodied agents do not or cannot portray. Usually, it is the character’s *actions* and *communications* that ought to appear similar enough to those of real people that we accept the animation as having believable thoughts or emotions. If these conditions are satisfied we are capable of ignoring significant non-human variants in form, appearance, or structure. One only need look at the wide range of animated cartoon characters that communicate their presumed thoughts and feelings to see that reality in expression is the stronger determiner of believability [29].

If actions and communications are the triggers for our understanding of animated characters, then these must be manifest on the character in human-like ways. For example, mechanical speech may destroy the believability of an otherwise accurately rendered character (unless we need to believe that it is a robot!); an awkward mechanical walk can distract us from seeing well-executed and subtle facial expres-

sions. Herein lies the first major goal of this study. What we are seeking is an animated embodied agent with *consistency* both in terms of our expectations of human actions and communications in general and our expectations of individual humans in particular. The agent’s behavior must be consistent from moment to moment and from situation to situation. There should be no wild mood swings or complete loss of focus. Departures from consistency might be interpreted as dramatic effects or, more likely, as internal conflicts within the agent’s own cognitive state¹. Normally, we should expect the cognitive state of the agent to be consistent with every level of its behavior: the expression on its face, the affect of its movements, the actions it performs, and the goals which it pursues. Also, cognitive state (and thus actions) must be consistent with the context or situation in which the agent finds itself.

Inconsistencies at any level can cause mixed messages and miscommunication. As Burgoon et al [6] indicates, “When enacting multichannel nonverbal presentations, common sense says that one should coordinate the channels to produce a consistent message.” Sometimes mixed messages are deliberate as in the case of jokes and sarcasm. Other times, mixed messages are indicative of internal confusion [6]. There may, of course, be times when internal confusion is our intent, but it seems more prudent to first model unconfused internal states and demonstrate consistent communication. If the theory works, we should be able to portray inconsistent behaviors and have observers infer conflicted or unbelievable states within the agent.

Our second major goal is to drive this consistent behavior from internal or cognitive models of the agent. It is somewhat surprising that agent modelers have been so heavily influenced by Ekman and Friesen [11] that many concentrate only on the basic facial expressions of happy, sad, anger, fear, and disgust and do not include internal reflective states such as determination, confusion, vacillation, and anxiety. Similarly, human gesture performance appears to reflect internal agent state in subtle but observable ways [2]. Only by representing the agent’s internal cognitive state and thus the information, beliefs, desires, and intentions that motivate it, can we achieve consistent externalized actions.

In this paper, we will concentrate on nonverbal communi-

¹In this paper we abuse the term “cognitive.” By cognitive state or process, we also include other internal, mental processes such as psychological, social, and cultural processes.

cation and the cognitive states or parameters that effect it. We will propose a parameterized agent model that creates consistent behaviors and allows controllability at different levels. We will describe the importance and interaction of the components and their manifestation in the channels of nonverbal communication.

2. MANIFESTATIONS OF NONVERBAL COMMUNICATION

Though verbal communication is the standard channel of communication used by people, nonverbal communication also contains valuable information. In fact, the information contained in nonverbal communication may be more valuable in some situations. In *Snow Crash* [28] Neal Stephenson describes an international business meeting taking place in a virtual world:

They come here [The Black Sun] to talk turkey with suits from around the world, and they consider it just as good as a face-to-face. They more or less ignore what is being said—a lot gets lost in the translation, after all. They pay attention to the facial expressions and body language of the people they are talking to. And that's how they know what's going on inside a person's head—by condensing fact from the vapor of nuance.

The internal or cognitive state of a person can manifest itself in all of the channels of nonverbal communication. According to [19] the channels of nonverbal communication are:

- facial expressions (smiles, nods)
- gestures (especially hand and arm movements)
- body movements
- posture
- visual orientation (especially eye contact)
- physical contacts (handshakes, patting)
- spatial behavior (proximity, distance, positions)
- appearance (including clothes)
- non-verbal vocalizations

We will briefly describe each of these channels and some of the research that has been done in the embodied agents community.

2.1 Facial Expressions

Facial expressions are known to express emotion [11], but facial expression can indicate what a person is thinking as well as feeling. The face reflects interpersonal attitudes, provides nonverbal feedback on the comments of others, opens and closes channels of communication, complements or qualifies verbal responses, and replaces speech [17].

Both Brand [5] and Poggi and Pelachaud [24] have researched ways of generating facial expressions for speech. Brand generates facial animation from information in an audio track. Poggi and Pelachaud concentrate on the visual display of intentions through facial animation based on semantic data. They model performatives, which are the type of action a sentence performs, such as requesting or informing. They also discuss how the degree of certainty, the power relationship, the type of social encounter, and the affective state effect the facial animation.

Cassell et al [7] present a system which automatically generates and animates conversations between multiple agents. A dialogue planner creates the conversation and generates and synchronizes appropriate facial expressions, intonation, eye gaze, head motion, and arm gestures.

2.2 Gestures and Body Movements

Gestures are voluntary or involuntary movements that are intended to communicate. They may involve any part of the body. They are used to emphasize, clarify, or amplify a verbal message. They can also regulate or control a human interaction, or display affect [19].

By contrast body movements are not intended to convey information. Body movements include, walking, reaching, turning, bending, etc. The manner in which these actions are done can help convey the cognitive state of the performer. People can walk in dramatically different ways: fast, slow, straight, swerved, proudly, sadly, joyfully, etc. [25].

Gestural communication and body language have been studied by several groups e.g., Morawetz and Calvert [21], Kurlander et al [18]. The EMOTE tool of Chi et al [9] controls the expressive shape and effort characteristics of gestures. Amaya et al [1] studied the expression of emotion on the body.

2.3 Postures

Posture is an indicator of the degree of involvement, the degree of status relative to the other participants, or the degree of liking for the other interactants. A forward leaning posture, for example, can indicate higher involvement, more liking, and lower status in situations where the participants do not know each other very well. Posture is also a key indicator of the intensity of some emotional states. A drooping posture is associated with sadness and a rigid, tense posture is associated with anger. The extent to which the communicators reflect each other's posture may also be an indication of rapport or an attempt to build rapport [17].

Becheiraz and Thalmann [3] present a model of nonverbal communication where agents react to one another in a virtual environment based on their postures. Relationships between the agents evolve based on the perceptions of postures.

2.4 Visual Orientation

What a person pays attention to and how much attention they pay is another channel of communication. A person's gaze and even the dilation and constriction of their pupils can be an indicator of interest, attention, or involvement [17].

Johnson and Rickel [15] present an animated pedagogical agent, which uses both gestures and attention to aid in the instruction of manual tasks. Cassell et. al [7] created an interface for chat room avatars that allows the user to give conversational cues through attention control. If a user sees an agent that he is talking to begin to look away from his avatar more and more, then this is probably an indication that the agent no longer wants to participate in the conversation.

2.5 Physical Contacts and Spatial Behavior

Physical contacts may be self-focused or other-focused. Self-focused touching may reflect a person's cognitive state or a habit and include nervous mannerisms. There are many kinds of other-focused touching, including, irritating, condescending, comforting, and electric. The meaning of a touch behavior is often derived more from its context and manner than from its configuration.

Spatial behavior refers to social and personal space. Spatial behavior can vary based on many aspects of individuals including, age, gender, status, roles, culture, personality and context. Studies show that conversational distance is related to general comfort level [6].

Physical contacts and spatial behaviors are types of behaviors that animation artists do well, but embodied agents researchers do not focus on. In order to create consistent communication these two channels of communication will have to be coordinated with the other channels.

2.6 Appearance

Among other things, appearance can provide information about, behavior, values and attitudes, and occupation. An immaculate appearance can indicate that a person pays attention to details. Wearing hiking boots can indicate that the person likes the outdoors. An old-fashioned appearance sometimes indicates old-fashioned values, and excessive jewelry can indicate materialism. Wearing attire that is functional and protective can indicate a blue collar job, whereas white collar workers wear more formal clothing [19].

There are many companies and research laboratories, including Blaxxun Interactive [4] and MIRALab [20], working on modeling virtual human bodies, skin, hair, and clothing.

2.7 Nonverbal vocalizations

Nonverbal vocalizations are vocal sounds other than words. This includes tone of voice which is known to convey emotional information [17]. For example, depressed people speak in a low, slow voice, with falling pitch.

The Sims [27] is a good example of the use of nonverbal vocalizations in embodied agents. In this game, the characters live their daily lives including, participating in polite conversations and angry discussions, but the characters have no discernible spoken language. The game's characters communicate through gestures, thought-bubbles, and nonverbal vocalizations. In the game, it is easy to distinguish a polite conversation from an heated argument by the volume and frequency of the nonverbal vocalizations.

3. AGENT COGNITIVE MODEL

In order to create an agent whose cognitive state is reflected in these channels of communication and therefore create consistent communication and behavior, we need to examine what cognitive processes effect the channels of communication. Books concerning nonverbal communication [17, 19, 6] often talk about the effects of the following cognitive processes on communication: age, status, gender, culture, role, context, emotion, mood, and personality.

3.1 Age and Status

In any interpersonal situation, one person's status is always at least a little above or below the other person's [16], and age is often a component of status. Age and status are reflected in many different communication channels. In order to present consistent agent behavior, these channels should all indicate the same age and status cognitive states.

For example, gestures change and become more subtle with age [19], and people of higher status seem to gesture less frequently [19]. Interpersonal distance also changes with age. Distance seems to increase with age, but is always closer with peers than with those that are younger or older. Older people are more likely to touch younger people than vice versa [6], which is probably a factor of both status and age. People of more dominance are more likely to engage in unwavering, direct looks. People tend to lower their eyes to show deference to authority figures, and submission is often marked by raised eyebrows, which connote deference [6]. Proper posture signals dominance. High status people are more confident and therefore comfortable in their space. Nonverbal vocalizations are also effected by status. A short "er" at the beginning of a sentence is weak, but a long "er" is strong [17].

The agents research community has, to some extent, modeled status. Hayes-Roth et al [12] have explored the use of status with embodied agents in the form of a master-slave relationship. They illustrate how the postures and actions of the characters change as the servant becomes the dominant character. Poggi and Pelachaud [8] model status through facial expressions called performatives, which are facial expressions that accompany and add interpersonal relationship information to speech. Musse and Thalmann [22] included dominance in their crowd simulations.

3.2 Gender

Physical appearance is an obvious channel to communicate gender, but gender should also be consistent with the other channels of communications. For instance, pairs of women tend to engage in more eye contact than pairs of men [6]. Burgoon et al [6] also discusses many other gender differences that effect the channels of nonverbal communication, including: postures in which males tend have more dominant, less affiliative, and less intimate postures than woman, and spatial behavior in which in small groups and interpersonal interactions, women require less personal space than men.

Though both men and women have been modeled in virtual environments, we currently know of no implementation that models gender as a component of the cognitive state of the agents.

3.3 Culture and Role

It is said that cultural information is a minimum prerequisite for human interaction—in the absence of such information communication becomes a trial and error process [17]. Cultural differences in communication can be extensive and do not only include the language spoken. First, different cultures have different distances for interacting. In some cultures standing close and directly in front of a person while speaking is considered either an intimate or a hostile act. In other cultures, not standing close and directly facing a person would be considered rude. There are also different touching behaviors, gestures, and eye gaze patterns [17].

It is also well known that there are some similarities across cultures. Studies have shown that the six basic facial expressions can be distinguished across cultures [11]. Also, some behaviors have cross-cultural similarities, e.g. coyness, flirting, embarrassment, open-handed greetings, and a lowered posture for showing submission [17].

While culture is a very important component of human behavior and communication, it has been neglected as a focus for the embodied agents research community, perhaps due to its complexity.

Every character in a virtual environment should have a role that it is playing, whether it is a professor of astrophysics, a tour guide, or just a man walking down the street. Roles involve expectations, both from the individual playing the role and from those interacting with the individual playing the role. In order for a character in a virtual environment to be consistent, it must meet the expectations of the role it is playing.

Roles are learned, generalized guidelines for behavior. Among other things, a role can stem from an individual's occupation, kinship, age, sex, prestige, wealth, or associational grouping. In a situation, one participant normally establishes his or her role and the other participant(s) must either go along or counter with a different role definition. There must be an agreement on the roles in order to effectively interact. Otherwise, communication will break down [6].

Roles influences many of the channels of nonverbal communication. Take for example the roles of doctor and mechanic. We have certain expectations about these roles. The appearance of a doctor is expected be clean and neat, while a mechanic may be very messy. We would also expect the interpersonal distance with a doctor to be smaller and the physical contacts more frequent (when comforting as well as examining). Confusion and alarm might result from a mechanic standing too close or touching too often (even if try to comfort someone after showing them the bill).

Isbister and Hayes-Roth [13] have explored roles in relation to intelligent interface agents. They found that making the role of an interface agent clear helps to constrain the actions users will take in their corresponding roles.

3.4 Context

People all perceive situations differently, and form different mental representations of the environment, people, and actions of a situation. This implies that their behavior is

predicated on their knowledge and understanding of the situation. An embodied agent's behavior should be consistent with the current context (or its perception of it). We would not expect the same behavior in an opera-house as a football stadium.

The problem is that context is a difficult thing to represent. Not only must we take into account all of the people and objects in the environment, and the embodied agent's feelings about them, and all of the action taking place in the environment, and the feelings about the actions, and feelings about past events, and the overall feeling of the environment, but we must then decide what the significance of all of these factors are.

Although context is an important feature for agents in virtual environments, it has not been heavily researched by the community. It requires attention, synthetic vision, a representation of the situation, and a way to determine what is important in the situation based on the agent's current cognitive state. Once the environment has been perceived and the situation represented, the context can be used to create behavior which is contextually consistent.

3.5 Emotion, Mood, and Personality

Emotions and mood effect many of the channels of nonverbal communication. The effect of emotions on facial expressions is well-known and well-studied [11], but other channels are effected as well. Lewis [19] indicates that tense moods cause postures that are rigid and upright, or slightly leaning forward. Extreme inhibition tends to cause withdrawal movements and general motor unrest. When depressed, movements are slower, fewer, and hesitating. By contrast, elation causes fast, expansive, emphatic, spontaneous movements.

The embodied agents research community has studied emotion and mood more than any of the other cognitive processes [10, 8]. Personality is a pattern of behavioral, temperamental, emotional, and mental traits for an individual. There is still a lot of controversy in personality research over how many personality traits there are, but the OCEAN model by is popular [31]. See Table 1

Like the other cognitive processes described, the modeling of personality may lead to more consistent communication, and because personality is a pattern of behavior (longer temporal extent) it should lead to more consistent behavior from situation to situation. This may aid in observers of the character developing a sense of *knowing* the character. It may become an individual instead of just another computer character.

In spatial relations, introverts generally prefer greater interpersonal distances. Aggressive and violence-proned (not agreeable) individuals tend to need even greater interpersonal distances in order to feel comfortable. Introverts also tend to resist visual interaction. People who are more neurotic and introverted have more restrained and rigid behavior, and display more uncoordinated, random movements [6].

Though often personality traits are confused with emotions in embodied agents research, there has been research done in embodied agents with personality [30].

	High Score Traits	Low Score Traits
Openness	Creative, Curious, Complex	Conventional, Narrow interests, Uncreative
Conscientiousness	Reliable, Well-organized, Self-disciplined, Careful	Disorganized, Undependable, Negligent
Extraversion	Sociable, Friendly, Fun-loving, Talkative	Introverted, Reserved, Inhibited, Quiet
Agreeableness	Good natured, Sympathetic, Forgiving, Courteous	Critical, Rude, Harsh, Callous
Neuroticism	Nervous, High-strung, Insecure, Worrying	Calm, Relaxed, Secure, Hardy

Table 1: OCEAN Model of Personality

3.6 Interaction of Cognitive Processes

These cognitive processes can influence and even conflict with one another. An extremely introverted person, for example, is unlikely to express anger in the same way as an extroverted person. An agreeable person is less likely to feel anger or to feel it as intensely as a disagreeable person. Perhaps personality also influences the types of roles a person performs. Would we want an unconscientious, neurotic person as a doctor? An introverted person who is forced into a public role would feel uncomfortable. Only by representing the agent's internal cognitive state, can we hope to depict such interactions and contradictions that result in anxiety, vacillation, or confusion.

3.7 Individuals

What is important to people, what they value, and what they desire are important aspects of their individuality. At any moment a person's actions are motivated by their goals and the interactions and conflicts of their goals. In order to achieve consistent external actions for embodied agents, we also need to model their goals and the processes involved in planning for goals and resolving conflicts between goals. AI research has studied many aspects of planning and conflict resolution in planning [26], but what is more important for consistent communication is the manifestation of these processes in the channels of nonverbal communication. Imagine a young child whose mother asks if he pulled up all of her newly-planted flowers. The child values being honest with his mother, but he also values the dessert which will be taken away as punishment. He will express confusion and anxiety as he decides what to do. The manifestations of his cognitive processes will communicate valuable information to his mother.

A person's goals and their other cognitive processes are related. Age influences an individual's perceptions, actions, decisions. Dominant individuals tend to claim scarce and desirable resources. In our society, males are traditionally thought of as more task oriented, while females are considered more interpersonal oriented. Culture helps in determining the importance and immediacy of the activities of life. Roles can be defined by what goals are valued while the person is performing the roles, and personality can be defined by what goals are valued and how those goals are achieved through time.

4. CONCLUSION

Studies have shown that consistency influences people's reactions to embodied agents. Both [14, 23] showed that people prefer consistent characters and that consistent characters had greater influence over people's behavior. We are

seeking an animated embodied agent with consistency both in terms of our expectations of human actions and communications in general and our expectations of individual humans in particular. We believe that modeling the cognitive processes of embodied agents is a step in this direction, and will facilitate the communication of internal reflective states such as determination, confusion, and anxiety.

We have discussed the type of cognitive processes an embodied agent should have in order to create consistent communication. We must also address how to create and control these cognitive processes. Ideally our model will provide varying levels of control. There are times when a virtual environment create wants to specify every detail of the characters behavior, and there are other times when he or she wants autonomous characters. We envision a system where a user sets only the parameters that they are interested in, and the system sets the rest. For example, if a user only desires to create a character who is close-minded, unconscientious, extroverted, disagreeable, and neurotic, then the system would set the other parameters based on these personality traits. This character, for example, might tend toward anger. If the roles in the system included nun and boxer, boxer would probably be chosen. The user could always go back and fix settings that were undesirable.

This paper has focused on nonverbal communication, but these channels of communication would have to be coordinated with verbal communication including vocabulary, tone of voice, and intonation.

5. REFERENCES

- [1] K. Amaya, A. Bruderlin, and T. Calvert. Emotion from motion. In *Graphics Interface '96*, pages 222–229, Toronto, Ontario, 1996. Canadian Human-Computer Communications Society.
- [2] N. Badler, M. Costa, L. Zhao, and D. Chi. To gesture or not to gesture: What is the question? In *Proceeding of Computer Graphics International*, pages 3–9, Geneva, Switzerland, 2000. IEEE.
- [3] P. Becheiraz and D. Thalmann. A model of nonverbal communication and interpersonal relationship between virtual actors. In *Computer Animation '96*, pages 58–67. IEEE Press, 1996.
- [4] Blaxxun interactive. <http://www.blaxxun.com/>.
- [5] M. Brand. Voice puppetry. In *SIGGRAPH '99*, pages 21–28, 1999.

- [6] J.K. Burgoon, D.B. Buller, and W.G. Woodall. *Nonverbal Communication, the unspoken dialogue*. Harper and Row, New York, 1989.
- [7] J. Cassell, T. Bickmore, L. Campbell, H. Vilhjalmsson, and H. Yan. *Embodied Conversational Agents*, chapter Human Conversation as a system Framework: Designing Embodied Conversational Agents, pages 29–63. The MIT Press, Cambridge, MA, 2000.
- [8] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors. *Embodied Conversational Agents*. The MIT Press, Cambridge, MA, 2000.
- [9] D. Chi, M. Costa, L. Zhao, and N. Badler. The emote model for effort and shape. In *SIGGRAPH '00*, 2000.
- [10] EBAA Workshop of Autonomous Agents'99. <http://www.ai.mit.edu/people/jvelas/ebaa.html>, May 1999.
- [11] P. Ekman and W.V. Friesen. *Manual for the Facial Action Coding System*. Consulting Psychologists Press, Palo Alto, 1977.
- [12] B. Hayes-Roth, R. Van Gent, and D. Huber. Acting in character. Technical Report KSL-96-13, Knowledge Systems Laboratory, Stanford University, Stanford, CA, 1996.
- [13] K. Isbister and B. Hayes-Roth. Social implications of using synthetic characters: An examination of a role-specific intelligent agent. Technical Report KSL-98-01, Knowledge Systems Laboratory, Stanford, Stanford, California, 1998.
- [14] K. Isbister and C. Nass. Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies*, 53:251–267, 2000.
- [15] W. Lewis Johnson and Jeff Rickel. Steve: An animated pedagogical agent for procedural training in virtual environments. *ACM SIGART Bulletin*, 8(1-4):16–21, 1997.
- [16] Keith Johnstone. *Status: Impro: Improvisation and the Theatre*. Theatre Arts Books, New York, NY, 1979.
- [17] M.L. Knapp and J.A. Hall. *Nonverbal Communication in Human Interaction*. Harcourt Brace Jovanovich College Publisher, Fort Worth, TX, 1992.
- [18] D. Kurlander, T. Skelly, and D. Salesin. Comic chat. In *SIGGRAPH '96*, pages 225–236, 1996.
- [19] H. Lewis. *Body Language, a guide for professionals*. Response Books, New Delhi, India, 1998.
- [20] Miralab. <http://miralabwww.unige.ch/>.
- [21] C. L. Morawetz and T. W. Calvert. Goal-directed human animation of multiple movements. In *Graphics Interface '90*, pages 60–67, 1990.
- [22] S. R. Musse and D. Thalmann. A model of human crowd behavior: Group inter-relationship and collision detection analysis. In *Proc. Workshop of Computer Animation and Simulation of Eurographics '97*, Budapest, Hungary, 1997.
- [23] C. Nass and L. Gong. Maximized modality or constrained consistency? In *AVSP'99*, 1999.
- [24] I. Poggi and C. Pelachaud. *Embodied Conversational Agents*, chapter Performative facial expressions in animated faces. The MIT Press, Cambridge, MA, 2000.
- [25] C. Rose, M. F. Cohen, and B. Bodenheimer. Verbs and adverbs: multidimensional motion interpolation using radial basis functions. *IEEE Computer Graphics and Applications*, 18(5):32–40, 1998.
- [26] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Englewood Cliffs, NJ, 1995.
- [27] The Sims. <http://www.thesims.com/>.
- [28] Neal Stephenson. *Snow Crash*. Bantam Books, New York, 1992.
- [29] F. Thomas and O. Johnson. *The Illusion of Life*. Abbeville Press, New York, 1981.
- [30] R. Trappl and P. Petta, editors. *Creating personalities for synthetic actors : Towards autonomous personality agents*. Springer Verlag, Berlin, Germany, 1997.
- [31] J.S. Wiggins. *The Five-Factor Model of Personality: Theoretical Perspectives*. The Guilford Press, New York, 1996.