Toward Representing Agent Behaviors Modified by Personality and Emotion

Jan Allbeck and Norman Badler
Center for Human Modeling and Simulation
Computer and Information Science
University of Pennsylvania
200 S. 33rd St.
Philadelphia, PA 19104-6389
allbeck — badler @seas.upenn.edu

April 15, 2002

1 Introduction

In recent years, there has been a surge in research of embodied conversational characters. Across applications these characters vary widely. They differ in both appearance and ability. Their ability is not just what actions they can perform, but also how they can perform them. Many applications call for an embodied character with the ability to be an individual, a persona, and to interact with humans as such.

Many researchers have explored giving embodied characters a richer persona through personality and emotion models. To this point, each researcher has created their own representation of these models. Furthermore, they must each tie these models to animation and speech synthesis components. The creation of a standard representation for embodied character models would allow better cooperation between researchers and create a basis for designing standards for evaluation of the models.

Virtual humans can represent other people or function as autonomous helpers, teammates, or tutors enabling novel interactive educational and training applications. We should be able to interact and communicate with them through modalities we already use, such as language, facial expressions, and gesture. This paper discusses the representational basis for character believability, personality, and affect. We also describe, a Parameterized Action Representation (PAR) that allows an agent to act, plan, and reason about its actions or actions of others. Besides embodying the semantics of human action, the PAR is designed for building future behaviors into autonomous agents and controlling the animation parameters that portray personality, mood, and affect in an embodied agent.

2 Parameterized Action Representation

We have constructed a Parameterized Action Representation (PAR) and a system (PARSYS) which uses PAR as a knowledge base and intermediary between natural language and animation [2, 3, 5, 9]. The PAR parameterization was created out of information from computer graphics and animation, natural language processing, and movement observation science. Although the emphasis of our research has been on the representation and processing of actions, objects are also represented in our formalism.

As a representation for actions as instructions for an agent, the PAR has to specify (parameterize) the agent, any relevant objects, and information about paths, locations, manners, and purposes. Table 1 shows the highest level representation of actions and Table 2 for objects.

The applicability conditions of an action specify what needs to be true in the world in order to carry out an action. These can refer to agent capabilities, object configurations, and other unchangeable or uncontrol-

type parameterized action =

(name:	STRING;
participants:	agent-and-objects;
applicability conditions:	BOOLEAN-expression;
preparatory specification:	*sequence* conditions-and-actions;
termination conditions:	BOOLEAN-expression;
post assertion:	STATEMENT;
during conditions:	STATEMENT;
purpose:	purpose-specification;
subactions:	par-constraint-graph;
parent action:	parameterized action;
previous action:	parameterized action;
concurrent action:	parameterized action;
next action:	parameterized action;
start:	time-specification;
duration:	time-specification;
priority:	INTEGER;
data:	ANY-TYPE;
kinematics:	kinematics-specification;
dynamics:	dynamics-specification;
manner:	manner-specification;
adverbs:	*sequence* adverb-specification).
failure:	failure-data;

Table 1: High level Action PAR

lable aspects of the environment. The conditions in this boolean expression must be true to perform the action.

Preparatory specifications are a list of <CONDITION, action> statements. The conditions are evaluated first and have to be satisfied before the current action can proceed. If the conditions are not satisfied, then the corresponding action is executed–it may be a single action or a very complex combination of actions, but it has the same format as the other PARs. In general, actions can involve the full power of motion planning to determine, perhaps, that a handle has to be grasped before it can be turned. We presently specify the conditions to test for likely (but generalized) situations and execute appropriate intermediate actions. It would also be possible to add more general action planners, since the PAR represents goal states and supports a full graphical model of the current world state.

A PAR can describe either a primitive or a complex action. The subactions contain the details of executing the action after all the conditions have been satisfied. If it is a primitive action, the underlying motion generator for the action is directly invoked. A complex action can list a number of sub-actions that may need to be executed in sequence, parallel, or a combination of both. A complex action can be considered done if all of its sub-actions are done or if its explicit termination conditions are satisfied.

Termination Conditions are a list of conditions which when satisfied indicate the completion of the action. Post Assertions are a list of statements or assertions that are executed after the termination conditions of the action have been satisfied. These assertions update the database to record the changes in the environment. The changes may be due to direct or side effects of the action.

The object type is defined explicitly to represent a physical object and is stored hierarchically in a database, called the Actionary™. Each object in the environment is an instance of this type and is associated with a graphical model in a scene graph. An object type lists the actions that can be performed on it and what state changes they cause. Among other fields, a list of grasp sites and directions are defined with respect to the object. These fields help orient actions that involve objects, such as grasping, reaching, and locomotion.

The agent executes the action. The agents are treated as special objects, and their properties are stored in the hierarchical object database. Each agent is associated with an agent process, which controls its actions based on the personality and capabilities of the agent. Not only does an agent's personality affect his or her response to a situation, but it also affects the way these actions are performed. Two agents with different personalities would execute the same action in two different ways. For example, two agents could be waving at one another. A shy agent would wave his hand more slowly and with more hesitation than an extroverted agent would. This increases believability by preventing agents from reacting in the same manner in identical contexts and gives the impression that each agent has distinct emotions and personalities.

PARSYS uses this representation to animate embodied agents. Actions are represented as PARs and stored in their uninstantiated (UPAR) form. UPARs contain only default properties for the action. Instantiated PARs (IPARs) contain specific information about the agent, objects, and other properties. Similarly, objects are stored in the Actionary™ in their general form. Many of the details of the representation of an object

type object representation =	
(name:	STRING;
is agent:	BOOLEAN;
properties:	*sequence* property-specification;
status:	status-specification;
posture:	posture-specification;
location:	object representation;
contents:	*sequence* object representation;
capabilities:	*sequence* parameterized action;
relative directions:	*sequence* relative-direction-specification;
special directions:	*sequence* special-direction-specification;
sites:	*sequence* site-type-specification;
bounding volume:	bounding-volume-specification;
coordinate system:	site;
position:	vector;
velocity:	vector;
acceleration:	vector;
orientation:	vector;
data:	ANY-TYPE).

Table 2: High level Object PAR

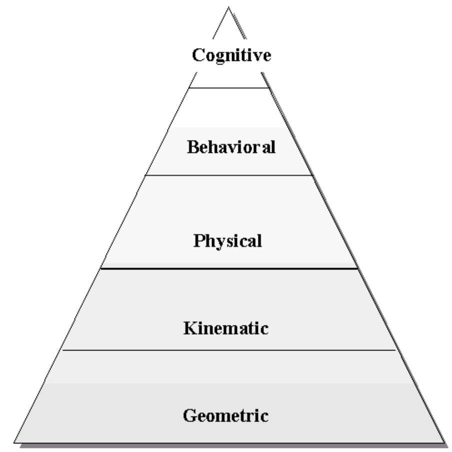

Figure 1: Funge's CG Modeling Hierarchy

can be filled in as the simulation begins (*e.g.* calculation of the bounding volume).

3 PAR for Agent Modeling

Given that PAR can be used to animate embodied agents, even from natural language instructions, can it be used to generate more "character" rich agents? In this section, we will show that PAR adequately represents the components necessary for modeling embodied agents and further that it is compatible with common methods for modeling emotion and personality in agents.

In [13], Funge et al. depicted a hierarchy of computer graphics modeling (see Figure 1). The bottom two layers were addressed early in computer graphics research with geometric models and inverse kinematics. Physical models generate realistic motion through dynamic simulation. Behavioral modeling involves characters that perceive environmental stimuli and react appropriately. Through cognitive modeling, *autonomous* characters can be given goals and react deliberatively as well as reactively.

PAR and PARSYS accommodate and enable each level in this hierarchy. While the actual geometry is assumed to have been created before the simulation begins, PAR does represent and PARSYS automatically recognizes some vital geometric constructs. Bounding volumes, for example, can be calculated as soon as the geometry is loaded into the system. Spatial properties, such as location and containment can also be recognized and stored. Updating and storing this information in a central location means that it does not have to be calculated by every object manipulator. Kinematics and dynamics are explicitly represented in PAR. Furthermore, PAR has been tied to a fast, analytic inverse kinematics program [16] that facilitates the generalization of actions such as *reaching*.

The behavioral component of embodied agency is at the foundation of PARSYS. The object hierarchy of the Actionary™ is updated to provide the necessary processes, agent processes and motion generators, with information on the current state of the environment. The embodied agents can be given goals directly in the form of a PAR or through natural language instructions. An agent tries to complete its goals by performing actions. Reactivity to the environment takes place in two forms. First, the agent processes and motion generators have quick access to the current state of the environment through the PAR allowing them to refine a motion or even terminate an action. Second, the PAR contains information about failure states and PARSYS has the ability to detect failures and notify the agent process with the information necessary to handle the failure. In PARSYS failures are anything that causes a motion generator to terminate before its termination conditions have been met. For example, a motion generator may check to ensure that the preparatory specifications of the action it is performing are maintained throughout. If the specifications are not maintained, a failure can be generated and returned to the agent process where a decision could be made to try to reestablish the specifications or abort the action.

The way in which an agent responds to changes in the environment, the way in which they pursue their goals, and even which goals are most important are aspects of cognitive modeling. The PARSYS contains mechanisms for planning and also filtering and prioritizing the actions that the planner can plan with, thereby individualizing the agent. During the planning process, the planner queries the Actionary™ for actions that match the conditions it is trying to meet. Before the satisfying actions are returned to the planner, an action filter removes any actions that the agents would not do in the situation and prioritizes the remaining actions. For example, *walking* might be prioritized over *running* or *skipping* in the satisfaction of a locomotion condition either because of the nature of the agent (businessman or child) or in sensitivity to motion goals or qualities (manner).

3.1 Personality and Emotions

The actions of the action filter may be dependent on any aspect of the agent, including its personality or current emotion level. Two popular models for personality and emotion are the OCEAN [17] and OCC [15] respectively.

Personality is a pattern of behavioral, temperamental, emotional, and mental traits that distinguish people from one another. Traits are basic tendencies that remain stable across the life span, but characteristic behavior can change through adaptive processes. The ways in which a person perceives, acts, and reacts is influenced by his or her personality. While there is no universally accepted theory, the Big Five or OCEAN model has gained some acceptance [17]. The "Big Five" represent a taxonomy of traits that some personality psychologists suggest capture the essence of individual differences in personality. The traits of the Big Five model are shown in Table 3.

Openness means a person is imaginative, independent-minded and has divergent thinking. Openness to experience describes the breadth, depth, originality, and complexity of an individual's mental and experiential life. Conscientiousness means a person is responsible, orderly, and dependable. Conscientiousness describes socially prescribed impulse control that facilitates task and goal-directed behavior, such as thinking before acting, delaying gratification, following norms and rules, and planning, organizing, and prioritizing tasks. Extroversion means that a person is talkative, social, and assertive. It implies an energetic approach to the social and material world and includes traits such as sociability, activity, assertiveness, and positive emotionality. Agreeableness means a person is good natured, co-operative, and trusting. Agreeableness contrasts a pro-social and communal orientation toward others with antagonism and includes traits such as altruism, tender-mindedness, trust, and modesty. Neuroticism means a person is anxious, prone to depression, and worries a lot. It contrasts emotional stability and even-temperedness with negative emotionality, such as feeling anxious, nervous, sad, and tense.

One of the most popular models for emotion is the OCC model, named after its authors [15]. In this model, emotions are generated through the agent's construal of and reaction to the consequence of events, actions of agents, and aspects of objects. Many researchers have based their work on this model [12, 8, 14].

Table 4 shows part of the PAR representation for agents. The parameters of the OCEAN model are represented as values along the scales of each of the characteristics. There is more information needed to implement the OCC model. First, the standards and values of the agent must be represented. These can be represented as statements that contain PAR actions. Essentially, each action can be associated with a number corresponding to the agent's thought of that action. Agents or classes of agents can also be associated with the actions to create more specific standards. Goals are actions with high priorities. Agents and objects can be tagged with information representing the agent's degree of cognitive unity and *liking* of the object.

	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 3: OCEAN Model of Personality

type parameterized agent =
(name: STRING;
personality: OCEAN-parameter-spec;
Openness INTEGER;
Conscientiousness INTEGER;
Extraversion INTEGER;
Agreeableness INTEGER;
Neuroticism INTEGER;
emotion: OCC-specification;
standards: *sequence* STATEMENT;
goals: *sequence* parameterized action;
appraisals: *sequence* cogn-unit-specification;
sequence appraisal-specification;

Table 4: Partial PAR Agent Representation

3.2 EMOTE for Displaying Affect

The implementation of personality or emotion for embodied characters must extend further than decision-making or action selection. The quality of movement in an action is also effected by personality and emotion. We have developed a parameterized system for creating more expressive gestures. The EMOTE system [18, 19, 4, 11] is based on movement observation science. Laban Movement Analysis (LMA) is a method for observing, describing, notating, and interpreting human movement. Two of LMA's components are *Effort* and *Shape*. Effort involves the dynamic qualities of movement. Shape describes the changing forms that the body makes in space. Effort comprises four motion factors: Space, Weight, Time, and Flow. Each motion factor is a continuum between two extremes: indulging in the quality or fighting against the quality. Table 5 describes the Effort qualities. Shape changes in movement can be described in terms of three dimensions: horizontal, vertical, and sagittal.

We have created many demonstrations of the EMOTE parameters. One such demonstration involved a virtual character hitting and touching a balloon (see Figure 2, http://hms.upenn.edu/software/EMOTE/balloon.html). Here the same basic animation data (from motion capture) for hitting was altered by the EMOTE system generating several different types of hitting and even touching.

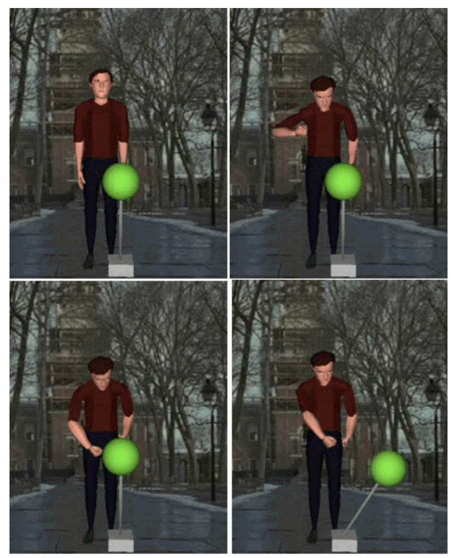

Figure 2: EMOTE alterations of hitting a balloon.

It is our goal to formally link these EMOTE parameters with OCEAN and OCC parameterizations. Table 6 shows an initial linking of EMOTE and OCEAN. This linkage is based on descriptions of LMA [7] and OCEAN [17] and is included only as an example of the type of mappings needed. We plan to verify or modify this linkage by showing agents exhibiting these qualities to naive observers and having them complete a questionnaire about the personality characteristics of the agent. We also plan to use a learning process to build the mapping between OCC and EMOTE. Automatically acquiring motion qualities from observation

Space:	attention to the surroundings
Indirect:	flexible, meandering, wandering, multi-focus
Examples:	waving away bugs, slashing through plant growth
Direct:	single focus, channeled, undeviating
Examples:	pointing to a particular spot, threading a needle
Weight:	sense of the impact of one's movement
Light:	buoyant, delicate, easily overcoming gravity, marked by decreasing pressure
Examples:	dabbing paint on a canvas, describing the movement of a feather
Strong:	powerful, having an impact, increasing pressure into the movement
Examples:	punching, pushing a heavy object, expressing a firmly held opinion
Time:	lack or sense of urgency
Sustained:	lingering, leisurely, indulging in time
Examples:	stretching to yawn, stroking a pet
Sudden:	hurried, urgent
Examples:	swatting a fly, grabbing a child from the path of danger
Flow:	attitude towards bodily tension and control
Free:	uncontrolled, abandoned, unable to stop in the course of the movement
Examples:	waving wildly, shaking off water
Bound:	controlled, restrained, able to stop
Examples:	moving in slow motion, tai chi, carefully carrying a cup of hot liquid

Horizontal	
Spreading:	affinity with Indirect
Enclosing:	affinity with Direct
Vertical	
Rising:	affinity with Light
Sinking:	affinity with Strong
Sagittal	
Advancing:	affinity with Sustained
Retreating:	affinity with Sudden

Table 5: Effort and Shape Elements

	Space	**Weight**	**Time**	**Flow**
Openness				
High	indirect	light	sustained	free
Low	direct	strong	sudden	bound
Conscientiousness				
High	direct	strong	sudden	bound
Low	indirect	light	sustained	free
Extraversion				
High	indirect	light	sustained	free
Low	direct	strong	sudden	bound
Agreeableness				
High	indirect	light	sustained	free
Low	direct	strong	sudden	bound
Neuroticism				
High	direct	strong	sudden	free
Low	indirect	light	sustained	bound

Table 6: Example EMOTE and OCEAN linkage

and validating them to make sure they are consistent with the LMA concepts and theories, are not only essential components to complete the EMOTE system in particular, but also can offer a powerful and valuable methodological tool for analyzing gestures and helping to create natural, personalized communicative agents. In [18] Zhao has developed a neural network based system to achieve this goal. The system inputs 3D motion capture and outputs a classification of EMOTE qualities that are detected in the input. The networks are trained with professional LMA notators to ensure valid analysis.

Future work in the EMOTE system and the motion quality recognizer will be to train the system to correlate captured motions with actor affect, behavior, mood, and intent. The critical problem in such training is setting up appropriate situations that truly elicit affective responses in individuals. We believe that the key ingredients to successful data generation are immersive experiences with both live and virtual agents. Engaging with either or both real and virtual agents in the same circumstances will be crucial to evaluating effectiveness and calibrating responses across the reality/virtual divide. Using the motion capture and post-session analysis, ground truth information can be supplied for training sets. The neural network models may then connect motion qualities to expressed affect and mood. Although the LMA community recognizes that such a mapping may exist it has not yet been possible to investigate it in a visually and computationally adequate environment.

3.3 Altering EMOTE Parameter Distributions

One problem that can result from parameterization is that rapid changes in the parameter values can cause inconsistent or unnatural looking movement. For example, an instantaneous change from a joint angle of 0 to 90 would appear quite unnatural. Treating the EMOTE parameters as a distribution and altering this distribution (scaling, shifting, amplifying, etc.) based on the personality and emotion parameters will lessen this inconsistency. A similar computational model has been used by Ball and Breese to model user mood based on user interface behaviors [6]. We start with neutral EMOTE parameters and alter them according to personality types. Distributions of EMOTE parameters for different personality traits will be created (probably through a learning process based on many observations). During simulation the agent begins interacting with its environment with actions modified by EMOTE values obtained from the parameter distributions. For example, an extremely extraverted personality may have a shift and amplification of the EMOTE value distributions to increase the likelihood of free, spreading gestures. If all channels of communication are affected by the EMOTE distributions for personality types, the agent's behavior will appear more consistent [1]. Currently, we have such parameterization for gestures and facial expressions. As emotional responses arise the EMOTE parameter distributions can be shifted or scaled to demonstrate the effect of emotions on movement behavior.

4 Conclusion

PAR was designed to be a flexible representation, meaning that many different types of information can be represented. Not all of the fields of the PARs need to be filled in for every action. When considering a representation for use with embodied conversational agents we should consider the trade-offs between parameterization specificity and program complexity. If you specify every joint angle for your character at every frame of the animation, your program needs only to display these angles on the figure. If you only specify that your agent needs to get some milk, then your program will need to figure out all the aspects of acquiring milk from high level planning to intricacies of movement. Our experience with the PAR and PARSYS leads us to conclude that they have the right balance of specificity and complexity.

That is not to say that there is not more work to be done. We would like to represent the PAR in XML format so that is more widely available to other researchers. Much work also needs to be done to establish the connection between EMOTE parameterization and models of personality and emotion. We are continuing to work on better planning and smarter motion generators for the PARSYS. Finally, although there is a natural language interface for the PARSYS, conversation and dialogue are not currently considered. A representation and system for modeling conversation and its timing, such as *BEAT* [10] would certainly enhance our system.

References

[1] J. Allbeck and N. Badler. Consistant communiction with control. In *Workshop on Multimodel Communication and Context in Embodied Agents at Autonomous Agents*, 2001.

[2] J. Allbeck, K. Kipper, C. Adams, W. Schuler, E. Zoubanova, N. Badler, M. Palmer, and A. Joshi. ACUMEN: Amplifying Control and Understanding of Multiple ENtities. In *Autonomous Agents and Multi-Agent Systems*, 2002.

[3] N. Badler, R. Bindiganavale, J. Allbeck, W. Schuler, L. Zhao, and M. Palmer. A parameterized action representation for virtual human agents. In *Embodied Conversational Agents*, pages 256–284. MIT Press, 2000.

[4] N. Badler, M. Costa, L. Zhao, and D. Chi. To gesture or not to gesture: What is the question? In *Proc. Computer Graphics International*, pages 3–9, Geneva, Switzerland, June 2000. IEEE Computer Society.

[5] N. Badler, M. Palmer, and R. Bindiganavale. Animation control for real-time virtual humans. *Comm. of the ACM*, 42(8):64–73, August 1999.

[6] G. Ball and J. Breese. Emotion and personality in a conversational agent. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill", editors, *Embodied Conversational Agents*, pages 189–219. MIT Press, Cambridge MA, 2000.

[7] I. Bartenieff and D. Lewis. *Body Movement: Coping with the environment*. Gordon and Breach, New York, 1980.

[8] J. Bates. The role of emotion in believable agents. *Comm. of the ACM*, 7(37):122–125, 1994.

[9] R. Bindiganavale, W. Schuler, J. Allbeck, N. Badler, A. Joshi, and M. Palmer. Dynamically altering agent behaviors using natural language instructions. In *Autonomous Agents*, pages 293–300, New York, June 2000. ACM Press.

[10] J. Cassell, H. Vilhjalmsson, and T. Bickmore. Beat: the behavior expression animation toolkit. In *Proc. ACM SIGGRAPH*, pages 477–486, 2001.

[11] D. Chi, M. Costa, L. Zhao, and N. Badler. The emote model for effort and shape. In *Proc. ACM SIGGRAPH*, pages 173–182, New Orleans, LA, 2000.

[12] M. El-Nasr, J. Yen, and T. Ioerger. FLAME – Fuzzy Logic Adaptive Model of Emotions. *Autonomous Agents and Multi-Agent Systems*, 3:219–257, 2000.

[13] J. Funge, X. Tu, and D. Terzopoulos. Cognitive modeling: Knowledge, reasoning, and planning for intelligent characters. In *SIGGRAPH '99*, pages 29–38, 1999.

[14] J. Gratch and S. Marsella. Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In *Proc. Autonomous Agents*, pages 278–285, Montreal, Quebec, 2001. ACM Press.

[15] A. Ortony, G. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

[16] D. Tolani and N. Badler. Real-time inverse kinematics for the human arm. *Presence*, 5(4):393–401, 1996.

[17] J. Wiggins. *The Five-Factor Model of Personality: Theoretical Perspectives*. The Guilford Press, New York, 1996.

[18] L. Zhao. *Synthesis and acquisition of Laban Movement Analysis qualitative parameters for communicative gestures*. PhD thesis, Computer and Information Science, Univ. of Pennsylvania, Philadelphia, PA, 2001.

[19] L. Zhao, M. Costa, and N. Badler. Interpreting movement manner. In *Proc. Computer Animation Conf.*, pages 112–120, Philadelphia, PA, May 2000. IEEE Computer Society.