Personality in Computer Characters

Daniel Rousseau

Knowledge Systems Laboratory
Gates Computer Science Bldg 2A
Stanford, CA 94305-9020, USA
Email: rousseau@hpp.stanford.edu
Phone: (415) 723-0948

Abstract

Personality characterizes an individual through a set of traits that influence his or her behavior. We propose a model of personality that can be used by intelligent, automated actors able to improvise their behavior and to interact with users in a multimedia environment. Users themselves become actors by exercising high-level control over their own intelligent agents. We propose different dimensions of personality that are based on the processes that intelligent agents usually perform. These dimensions are rich enough to allow the specification of an interesting number of characters able to improvise and react differently although they are put in the same context. We show the influence that the personality traits have on an actor's behavior, moods and relationships. An application of the Computer Virtual Theater, the Cybercafé, is used to test our model.

1. Introduction

Personality is an important domain of research in psychology. Unfortunately, there is no consensus on the definition of a personality or on its components. Personality theories differ in their degree of emphasis on the past and the present, the conscious and the unconscious, the directly observable and the relatively unobservable. Most of the definitions of a personality emphasize the unique or distinctive qualities of individuals. According to some psychologists, these qualities correspond to stable psychological traits that can be perceived through an individual's behavior and emotions. Such qualities motivate an individual to interact with others very often or rarely, to be autonomous or dependant, friendly or hostile, emotional or phlegmatic, etc.

Personality is also important in any domain fictive characters evolve in. An author has to determine the personality of a character she wants to create for a theater play, a movie or a narrative. An actor who personifies a character in a movie or a theater play must forget his own personality and act instead according to the personality of the character. An improvisation actor combines an author's and a simple actor's responsibilities, because he has to create a personality for the character he personifies and plays his role according to that personality.

In the Virtual Theater project (Hayes-Roth & van Gent 1996), we propose a model of personality that can be used by intelligent, automated actors able to interact with users in a multimedia environment in well-defined stories or in improvisational contexts. Users themselves become actors by exercising high-level control over their own intelligent agents. These agents improvise to meet the user's goals based not only upon their personality, but also upon their knowledge, their moods (including emotions) and their interpersonal relationships.

The remainder of this paper describes our model of personality and our experiments in more detail. In section 2, we present the dimensions of personality that we consider. We show the influence that the personality traits have on an actor's behavior in section 3. In section 4, we explain how personality can influence moods and relationships. In section 5, we present the implementation that we used to test our model: the Cybercafé. We mention important works related to personality in psychology and artificial intelligence in section 6 before evaluating our approach in section 7.

2. Dimensions of Personality

Works on personality traits are quite rare in artificial intelligence. A few researchers used a few traits in their work (Bates, Loyall & Reilly 1992; Hovy 1988; Carbonell 1980), but there is no model as complete and structured as the one that Ortony proposed for emotions (Ortony, Clore & Collins 1988). Our goal is to set a model that classifies personality traits in a structured way and that identifies their impact on a character's behavior, moods and relationships.

We base our classification on the different processes that our intelligent agents could perform in a conventional architecture: perceiving, reasoning, learning, deciding, acting, interacting, revealing, and feeling emotions. Although feeling emotions is not a real process but a capability, we include it in our model because emotions greatly influence a character's behavior. We consider each process at two levels: the natural inclination (tendency) that an agent has to perform the process, and the main aspect an agent focuses on while performing the process.

Process	*Inclination*		*Focus*	
	Level	*Value(s)*	*Aspect*	*Value(s)*
Perceiving	Low	Absentminded	Expectations	Imaginative
	High	Alert	Reality	Realistic
Reasoning	Low	Silly	Undesirable effects	Pessimistic
			Facts	Objective
	High	Rational	Desirable effects	Optimistic
Learning	Low	Incurious	What is learned only	Gullible
			Both what is learned and what is already known	Open-minded
	High	Curious		Selective
			What is already known only	Intolerant
Deciding	Low	Insecure	First reaction	Impulsive
	High	Self-confident	Good decision	Thoughtful
Acting	Low	Passive	Anything besides the task	Indifferent
	Intermediate	Delegating	Task	Diligent
		Active		
	High	Zealous	Result of the task	Perfectionist
Interacting	Low	Introverted	Addressee as a threat	Hostile
			Exchange of information	Neutral
	High	Extroverted	Addressee as a help	Friendly
Revealing	Low	Secretive	Lie	Dishonest
	High	Open	Truth	Honest
Feeling emotions	Low	Emotionless	Self	Selfish
	High	Sensitive	Others	Unselfish

Table 1: Dimensions of a personality

The focused aspects were chosen because of the nature of the processes and of the personality traits already proposed by psychologists (Cattell 1965; Allport 1966; Briggs-Meyers & Meyers 1981; Phares 1984). So, our model contains 16 dimensions, as we present in Table 1.

The table gives at least two examples of qualitative values for each dimension with the corresponding level of inclination or with the focused aspect. For instance, a character is absent-minded if he misses several events and details that he could perceive (low level of perception). He is alert if he is very inclined to perceive everything. Of course, an agent might have other values between these two extremes. He could perceive almost everything but miss a few details around him when he is very busy.

We can define very rich personalities with our model, because several combinations of traits are possible. For example, a courageous character is basically self-confident and active. A cowardly agent is insecure and passive. A courageous agent and a coward could both talk about their exploits if they are extroverted and open. The difference is that the first boasting agent would be honest and the latter would be dishonest.

3. Personality and Behavior

A personality trait has an impact on the process it is associated with. It could also influence other processes, since the processes are related to each other. We present in Table 2 the different types of behaviors we consider for each dimension in our model.

In any context, a character can react in different ways. An agent usually chooses actions that are consistent with his personality. For instance, a passive character tries to avoid performing a task himself if he can. He can ask someone else to perform the task or he can just do nothing.

The typical behavior that expresses a personality trait is specified in abstract rules that a character can follow when it is time to choose an action to perform in a given context. Such rules specify which types of behaviors an individual would likely perform according to his personality. For instance, a rule would specify that a passive agent prefers to do nothing rather than to act. As personality-related rules are abstract and do not contain specific actions, the same character can act differently in the same context, but

Processes	*Inclination*	*Focus*
Perceiving	Choice in perceiving or not something new	Perception or not of imaginary events and states
Reasoning	Search for inferring new knowledge or not from what is already known	Reasoning neutral or biased by positive or negative feelings
Learning	Search for learning or not	Acceptance or refusal of conflicting opinions or information
Deciding	Autonomy or dependency in decision and behavior	Immediate or delayed reaction
Acting	Preference to act, to delegate or to do nothing	Concern or not about performing a task well
Interacting	Choice in interacting or not	Choice in being friendly or not with others
Revealing	Choice in telling others about knowledge, feelings and intentions	Choice in telling the truth or not, in behaving honestly or not
Feeling emotions	Emotions affect an agent or not in his behavior	Importance given to self or others' situation and emotions

Table 2: Types of behaviors influenced by the personality traits

each action respects his personality and conforms to his style.

The resulting typical behavior of a character is obtained through the combination of the different behaviors described in the rules concerning his personality traits. As an example, we present the description of the typical behavior of four different waiters in a café when we consider five of their personality traits:

- *Waiter 1: realistic, insecure, introverted, passive, secretive*
  Such a waiter does and says as little as he can.
- *Waiter 2: imaginative, dominant, extroverted, active, open*
  This waiter takes initiative, comes to the customer without being asked for, talks much.
- *Waiter 3: imaginative, dominant, extroverted, passive, secretive*
  Such a waiter talks very often about others without revealing about himself; he does as little as he can.
- *Waiter 4: imaginative, insecure, introverted, active, secretive*
  This kind of waiter does his job well without saying much and without taking initiative.

What is very interesting is that the interaction between the waiter and the customer would be very different in each case. It is probable that a customer would be more satisfied with the second or the fourth waiter because he would likely get very good service. But, if the waiter was hostile, the customer's satisfaction would be pretty low. Moreover, a customer could show or hide his satisfaction or frustration, and the waiter could react in different ways to the customer's behavior. As a consequence, the interaction could get friendly, hostile, tense, relax, etc. We do not know in advance what will happen because each actor, autonomous or user-driven, must improvise according to his role, his personality, the directions he receives (see section 5) and the other actor's behavior.

4. Personality, Moods and Relationships

We assume that personality traits have a major influence on moods and on interpersonal relationships.

A mood is rather variable. It is an emotion, or a state of mind related to a physical state or to an interpersonal relationship. Emotions, such as happiness, sadness and anger, are mental states that are positively or negatively valued by an individual with respect to an object, an event or a situation depending on the pleasure or the displeasure felt by the individual. For instance, a character would likely get happy if he was complimented, but sad or angry if he was reprimanded. A character can also have physical moods: he can be tired or peppy, hungry or satiated, etc.

Interpersonal relationships are called in terms of the participants' social roles: employer-employee, mother-daughter, etc. They imply a certain kind of behavior that is expected for each agent who is part of the relationship (Cartwright 1974). They are also based upon the status of the participants (employer-employee, master-slave) and their level of attraction (lovers, enemies). Status (Johnstone 1992) and attraction (Ortony, Clore & Collins 1988) are two special variable moods. A relationship is considered from an agent's point of view that is not necessarily shared. For example, John believes that he loves Mary (lovers relationship), but Mary thinks that they are just friends. Both have attraction for each other, but

not with the same intensity. A relationship can be replaced by another under certain circumstances.

In a given context, we consider that agents with different personalities could experience different moods, or the same moods with various intensities. For instance, a self-confident character who is threatened would feel angry, while an insecure agent would be scared in the same situation. Sometimes, a relationship can change because of the characters' personality. For example, we could have a role reversal in the case where a master who is naturally insecure is challenged by a self-confident servant. The status of each character would change with time, and the insecure agent would serve the new master at the end.

On the other hand, agents having the same moods or involved in the same kind of relationship could react differently because of their personality. A courageous character who is scared nevertheless tries to struggle; the coward just tries to escape. An honest and unselfish character hardly cheats his friend, which is not the case for someone who is dishonest and selfish.

As for personality traits, abstract directions are used to describe preferable behaviors of an agent according to his moods and his relationships. For instance, an individual prefers to be honest than dishonest with his friend. If he is tired, he prefers to perform actions slowly or to do nothing. All the abstract directions, either related to the personality, the moods or the relationships, are considered by a character when it is time to choose an action to perform.

5. Application of the Model in the Cybercafé

We test our approach using an application of the Computer Virtual Theater (Hayes-Roth & van Gent 1996) which provides intelligent characters able to improvise their behavior. Those agents possess repertoires of physical and verbal actions that can be used as building blocks for constructing stories and plays. The users give the characters abstract directions down to which levels of abstraction they want. These directions are a kind of constraints that must be respected by characters as they are by an improvisation troupe in theater. Characters may be totally controlled by the system, or partially or totally controlled by users. There are two kinds of platforms for the Computer Virtual Theater: one where the characters are animated, and one where the environment as well as the interaction between the actors is described with text only as in a MUD application. A MUD (Multiple User Dimension, Multiple User Dungeon, or Multiple User Dialogue) is a computer program which users can log into and explore. Each user takes control of a computerized avatar through which he can walk around a fictive environment, chat with other characters, perform actions, etc.

We chose the Cybercafé as an application of the Computer Virtual Theater. The Cybercafé is based on Schank's well-known restaurant script (Schank & Childers 1984). In that application, a user can direct an avatar representing a customer and interact with an automated actor playing the role of a waiter. We present the architecture of the Cybercafé in Figure 1.

We distinguish the user who controls the avatar through a user interface and the author who defines the script of a session through an authoring tool. The script is the abstract specification of what should happen during a session. Such a definition can be specified in terms of high-level or specific actions, states, goals, events, personality, moods, relationships, and time. Actors, either autonomous or user-driven, are able to improvise their behavior at different degrees while respecting the constraints imposed by the script, the context and, in the case of the avatar, the user's selection of actions to perform.

At the beginning of a session, a user has also to setup the interaction through a setup interface that asks for information necessary to begin the session and to complete the current script. Personality traits, moods and relationships could be specified at this level. Afterward, the actors are ready to interact and improvise under the director's supervision. The director is a special agent that models the world where the actors evolve, communicates directions from the script and from the setup to the actors, perceives all the actions performed by actors, decides on the results of such actions in the world, and creates new events occurring independently from the actors.

The waiter can get personality traits at the beginning of a session in one of the four following ways:

- The values are directly specified by the user;
- They are selected by the user from a repertoire of personalities predefined by an author;
- They are specified by an author in a script;
- They are chosen by the system.

The first two options are used especially when the user is also the author of the application. Then, the user can observe the waiter's reactions and correct the set of abstract rules associated with a personality trait if the waiter has an inconsistent behavior with respect to his personality. The last two options offer more surprise for the user who does not want to know a priori which kind of waiter he will have to interact with.

We plan to test our approach using the aforementioned modes of setup in the Cybercafé. The first two modes have already been used to enable a user to test if an actor's behavior is pertinent for specific traits, and to correct the personality-related rules when problems are detected. The last two modes will be used to ask the user to identify the waiter's personality traits at the end of a session. This second type of experimentation will allow us to verify if the personalities we have modeled are believable.

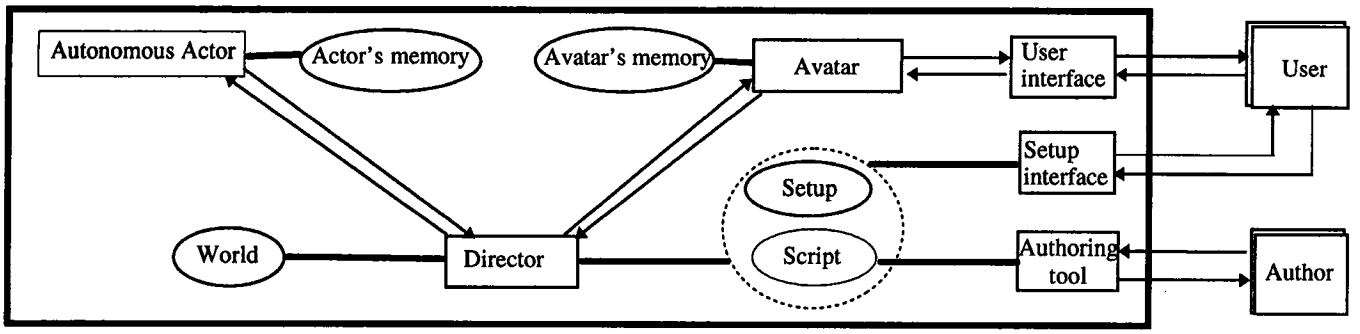

Figure 1: Architecture of the Cybercafé

Each actor has a repertoire of possible behaviors organized in a class hierarchy. Each action has a set of relevance conditions that can be satisfied by the occurrence of certain classes of events. The class hierarchy allows the actor to search efficiently for relevant behaviors. Each action has annotations specifying "typical" values on a set of personality or mood variables (e.g. extroverted, introverted, angry, happy, sad).

An autonomous actor iterates the following steps (Hayes-Roth, van Gent & Huber 1996): it incorporates perceptual information and directions into its memory; it instantiates the behaviors of its repertoire that are relevant in the current situation; it rates the behaviors according to its personality traits, its moods and some random variability; it performs the most rated behavior. The last two steps are different for an avatar. The relevant behaviors are presented to the user through the user interface. Then the user selects an action that is performed by the avatar afterward.

Moods can be changed through perception. A special rule is triggered when an actor having specific personality traits detects a certain event in a given context. For instance, a waiter who is self-confident can feel his status growing if the customer asks him for a suggestion.

The current version of the Cybercafé is not complete, but it is sufficient to test how personality traits can combine and influence an individual's behavior. We take into account only a few personality traits (inclination in deciding, interacting, learning, and focus in learning) and their influence on a character's moods. We do not consider the change of relationships between characters. We allow a user to specify the waiter's personality traits at the beginning of a session, because we want the user to be able to correct a character's behavior easily, the user being the author of the application. Only one avatar is available in the current version of the system. We expect to have several avatars able to interact together and with the waiter in future versions of the Cybercafé.

6. Related Works

Several psychologists were interested in personality (Cattell 1965; Allport 1966; Briggs-Meyers & Meyers 1981; Phares 1984). Unfortunately, there are many theories containing various personality traits and none of them is universally accepted. Nevertheless, works in psychology are good guidelines to establish a model of personality that can be used by computer actors able to improvise.

Few reseachers in artificial intelligence tried to use personality traits in a computer system. Unfortunately, most of the existing models are not very structured because they do not rely on a solid foundation. Let us mention some of them.

Loehlin (1968) implemented a limited model of personality: ALDOUS. In that model, Loehlin specifies how the system would react to a certain kind of object described according to its nature (human or not), its hair color, its sex and its age. Only three personality traits are taken into account for each type of object that the system may encounter: hostility, fear and attraction. These traits and the number of times the system already encountered a certain type of object influence its behavior.

Carbonell (1980) uses a set of personality traits derived from the goal trees in the POLITICS system. The traits proposed are based on deviations from the social normative goal tree characterized according to Schank and Abelson's taxonomy (Schank & Abelson 1977). For instance, an ambitious individual considers the acquisition goals concerning objects, status and knowledge as more important than the average character, but any preservation goals related to other people as less important. Although this work, which is the most complete that we know, relies on a well-known theory, it does not present personality traits in a structured manner.

In his natural language generation system PAULINE, Hovy (1990) uses stylistic rhetorical goals that express personality traits: timidity, open-mindedness, respect, etc.

But we do not really know where those rhetorical goals come from. The same remark applies to the behavioral features proposed in the OZ project (Bates, Loyall & Reilly. 1992) such as curious, content and aggressive, to modulate a character's behavior. Furthermore, the behavioral features can change with time, which is in contradiction with the personality definition.

Nass et al. (1995) wanted to verify if people would respond to machines endowed with personality-like characteristics as if they did have personality. Their studies were conclusive for the dominance and submission traits, but they did not test any other traits.

7. Conclusion

The model of personality that we propose provides a set of stable psychological traits that characterize a character through his or her behavior. Such traits are clearly distinguished from variable moods. They are related to different processes an agent, artificial or human, has to perform. Such relations provide a structure in our model that we use to create a rich variety of automated and user-driven actors able to improvise their behavior in a multimedia environment. We also consider the influence that the personality has on the behavior, the moods and the relationships.

We cannot claim that our model is complete. Personality is very complex and psychologists themselves do not agree about its components. Our goal is not to get a complete model of personality, but to propose a model providing a sufficiently rich structure based on the convention architecture of an intelligent agent, and available for computer actors. So, such actors can improvise their behavior in a consistent way that is driven by the personality and the moods of the characters they portray.

The model is still in development. Up to now, the results have been quite encouraging regarding the observation of personality traits through a character's behavior in the current version of the Cybercafé, although just part of the model was implemented. We will extend our tests to more complex and entertaining systems in the future.

Acknowledgments

This work was supported by NSERC and a gift from Intel. It has benefited from discussions with Barbara Hayes-Roth, Patrick Doyle and the other members of the Virtual Theater research group.

References

Allport, G. W. 1966. Traits Revisited. *American Psychologist* 21: 1-10.

Bates, J.; Loyall, A. B.; Reilly, W. S. 1992. *An Architecture for Action, Emotion, and Social Behavior.* Technical Report CMU-CS-92-144, School of Computer Science, Carnegie Mellon University.

Briggs-Myers, I.; Myers, P. 1980. *Gifts Differing.* Consulting Psychologists Press.

Carbonell, J. G. 1980. Towards a Process Model of Human Personality Traits. *Artificial Intelligence* 1-2: 49-74.

Cartwright, D. S. 1974. *Introduction to Personality.* Randy McNally College Publishing Company, Chicago, IL.

Cattell, R. B. 1965. *The Scientific Analysis of Personality.* Penguin.

Hayes-Roth, B.; van Gent, R. 1996. *Story-Making with Improvisational Puppets and Actors.* Technical Report KSL-96-05, Knowledge Systems Laboratory, Stanford University.

Hayes-Roth, B.; van Gent, R.; Huber, D. 1996. Acting in character. To appear in R. Trappl and P. Petta (Eds.), *Creating Personalities for Synthetic Actors.* Technical Report KSL-96-13, Knowledge Systems Laboratory, Stanford University.

Hovy, E. H. 1990. Pragmatics and Natural Language Generation. *Artificial Intelligence* 43: 153-197.

Johnstone, K. 1992. *Impro: Improvisation and the Theatre.* Routledge, New York.

Moulin, B.; Rousseau, D. 1995. An Approach for Modelling and Simulating Conversations. *Essays in Speech Act Theory*, D. Vanderveken and S. Kubo eds., John Benjamins Publishing Company. Also Technical report DIUL-RR-9402, Laval University.

Nass, C.; Moon, Y.; Fogg, B. J.; Reeves, B.; Dryer, D. C. 1995. Can Computer Personalities Be Human Personalities? *International Journal of Human-Computer Studies* 43: 223-239.

Ortony, A.; Clore, G.; Collins, A. 1988. *The Cognitive Structure of Emotions.* Cambridge University Press.

Phares, E. J. 1984. *Introduction to Personality.* Charles E. Merrill Publishing Company, Columbus, Ohio.

Schank, R. C.; Abelson, R. P. 1977. *Scripts, Goals, Plans and Understanding.* Lawrence Erlbaum, Hillside, NJ.

Schank, R. C.; Childers, P. G. 1984. *The Cognitive Computer on Language, Learning and Intelligence.* Addison-Wesley.