

Models of Culture for Virtual Human Conversation

David Traum

Institute for Creative Technologies
University of Southern California
Marina del Rey, CA, USA traum@ict.usc.edu

Abstract. In this paper, we survey different types of Models of culture for virtual humans. Virtual humans are artificial agents that include both a visual human-like body and intelligent cognition driving action of the body. Culture covers a wide range of common knowledge of behavior and communication that can be used in a number of ways including interpreting the meaning of action, establishing identity, expressing meaning, and inference about the performer. We look at several examples of existing cultural models and point out remaining steps for a more full model of culture.

1 Introduction

Virtual Humans [1, 2] are artificial agents that include both a visual body with a human-like appearance and range of observable behaviors, and a cognitive component that can make decisions and control the behaviors to engage in human-like activities. Virtual humans often engage in conversation (embodied conversational agents [3]) and can play a number of roles, including providing a human-like interface to information services, acting as a roleplayer in a training system, acting as a tutor, or a “non-player character” in a game.

Virtual humans can have two main roles with respect to culture. First, they can be seen as artifacts that have been designed according to the guidelines of a specific culture, similar to the way the shape of electric outlets and the layout of keys on a keyboard vary from country to country. One finds it much easier to use an artifact that follows a familiar design than one which is equally efficient but follows a design from a different culture.

A more interesting role, though, is to portray the virtual human as belonging to a specific human culture. In some sense this is unavoidable in that every choice of appearance and behavior of the virtual human will indicate something about culture and roles within that culture. Having virtual humans with explicit models of specific cultures can offer a number of advantages, such as

- making it easier or more effective for people from that culture to interact with and understand the virtual human
- adding verisimilitude to an experience (game or training simulation) set in a particular locale, inhabited by people from that culture
- training people to recognize members of a specific culture
- training people how to behave with members of a culture and what results from different kinds of interaction

In this paper we examine several different approaches to modelling culture for virtual humans and the applications that they are appropriate for. In the next section, we discuss culture, and provide a working definition for study. In section 3, we examine some of the differences that arise in different cultures. In section 4, we look at different sorts of cultural models and examples of their application for virtual humans. Finally, we conclude in section 5 with some thoughts on the ideal cultural model and steps needed to achieve it.

2 Culture

There are literally hundreds of definitions that have been given for culture, already categorized in various ways by Kroeber and Kluckhohn [4]. For the purposes of this paper, and exploring different models of culture for virtual humans, we will use the definition in (1).

- (1) A culture is a set of rules that a group has common knowledge of and orientation towards.

These rules can be of several sorts, including

- (2)
 - a. Normative rules of how people in certain roles should behave
 - b. communicative rules of what behavior will signify which meanings in which contexts
 - c. inferential rules that specify what new knowledge can be derived from existing knowledge (including behaviors and communications)

Each of these rules can be used in several ways. For example, the normative rules can be used to recognize someone's observed behavior as conformant or discordant with those rules and thus decide whether they belong to the culture and what role they are adopting within that culture (e.g. as conforming member, nonconformist, leader, follower, ...). One could use the same rules to choose one's own behavior and project an identity as part of a culture or having a certain role. Likewise, one could use knowledge of a communicative rule in order to produce behavior such that it will convey a particular meaning in a given context. Or one could use the rule to infer a meaning from observed behavior in a particular context. Or one could infer from the apparent use of the behavior to convey the meaning that the context holds and/or that actor is a part of the culture.

The culture group can be of any size – it is often convenient to look first to national groups, because of obvious cultural markers such as distinct languages and legal force behind rules of behavior, however a cultural group can also be smaller or larger than this. At the extremes, universals among all humans would be indistinguishable from biological or physical universals, and at the other extreme, if only one individual is oriented to the rules, it may be said to be an individual trait. Anything in between this, however, may be part of the realm of culture. Any given individual will usually be part of many different cultures, including national and linguistic cultures, but also cultures based on occupation, local area, family, etc.

3 Aspects of Cultural Differences

Cultures can differ in many aspects of the rules. Some of these differences are:

- different appearance
- different behaviors
- different meanings
- different contexts
- different mappings from behavior and context to meaning
- different frequencies of behavior

Appearance is one of the most obvious markers of cultural difference. Cultures often have different conventions about clothing, hairstyle, jewelry, make-up, tattoos, etc., both in terms of what can be worn, and what the significance of wearing it is, e.g. group membership or status.

Cultures may also differ in the kind of behaviors that are performed. Some behaviors are performed in only some cultures but not others, while in other cases, the same behavior is performed but with different meanings. Cultures may also vary in terms of what kinds of meanings are important and expressed by behaviors. In still other cases, the same meaning is attributed to the behaviors, but the situations are different, so there is not the same opportunity to express the meaning. There are also cases in which the situations, behaviors, and meanings are the same but the frequency of occurrence are different across cultures. For example, as compared in [5], Canadian soldiers had a much higher incidence of acknowledgement acts (according to the HCRC coding scheme [6] than Scottish University students when performing the maptask. Likewise, English speakers had many more feedback moves (in the Verbmobil-2 coding scheme [7]) than German speakers, when performing the Verbmobil tasks.

In terms of conversational behaviors, there are a large number of behaviors that can vary across culture. Most obvious is verbal language - different languages and speaker groups within the languages can have noticeable differences at many linguistic levels, including the phones used to convey meaning, the way these sounds are carved up into phonemes, how meaning units such as words are put together, and syntactic and discourse levels. There are also a number of non-verbal behaviors that convey culture-specific meaning in different ways, including proxemics, gaze, facial expressions, body posture, hand gestures, and prosody and intonation. There are also patterns of interaction that vary from culture to culture, including turn-taking, greetings and closings, sequential interactions (such as how many turns it may take to propose and accept a course of actions), grounding behavior, attitudinal expressions such as boosting and downplaying when reporting significant events.

4 Types of Cultural Models

Because of so many differences and kinds of differences between cultures, there are thus many different ways of modelling culture-specific behavior for virtual humans. One key question about cultural models is whether they are intended as internal or external. An external model of culture is one that is meant to cause an observer of the agent to

see the agent as behaving according to that culture (and thus perhaps a member of the culture group with a specific role in that group). If a virtual human has only an external model built in to it, it may have no explicit model of the culture, how to interpret the behavior of others according to this culture, or how to draw culture-specific inferences. An internal model of culture would allow the agent to have a representation in some form of the rules of the culture and make decisions about both the meanings of the behaviors of others as well as its own behavior. The behavior that the agent produces might, however, not be recognizable as appropriate for that culture by a person familiar with the culture, if the model is partial in some ways, e.g. taking into account only some of the aspects of that culture. Both of these kinds of models can be further broken down into subtypes.

For external models, the simplest and arguably highest impact aspect is the physical appearance and set of behaviors that the virtual human can adopt. Here there is even more freedom for virtual humans than real people, since physical and biological characteristics can be abstracted away from toward cultural stereotypes. [8] shows differences in interpretation of static facial expressions of Avatars depending on the national origin of the coder. In the virtual humans in [9], the set of gestures used by the Spanish doctor and Iraqi elder were animated by observing actors from these cultures, and mapping is made from “universal” meanings to culture-specific behaviors.

Some models take into account not just the set of behaviors but the frequency, as well. E.g., [10] has models of inter-turn overlap and silence in which the relative frequency of behaviors depend on culture-specific parameters. These parameters can be set through observational studies of people in that culture, for example the corpora described in [11]. Some models also include different frequencies for different situational contexts. The model of proxemics in [10] has different culture-specific normal distances for different types of relationships, and the resulting behaviors of moving closer or further depend not just on these norms, but also other factors such as noise level and group cohesion, as described in [12]. The model of gaze in [10] also takes into account differences in conversational role, such as speaker, addressee, overhearer, other. The model of gaze in [13] relies on speech act, speaker type and function as well as culture group in order to set probabilities for gaze targets. In all of these cases, the main purpose of the model is to provide an impression to human viewers that the virtual human belongs to a particular culture group (and is behaving appropriately for one of that group). The model of [14] uses a bayesian planner and culture model to decide on behavior, and a culture-specific scheduler using frequency information about behaviors and cultures.

Internal models of culture for virtual humans are comparatively more rare. In one sense, every virtual human model that can engage in the full range of conversation (including producing and interpreting behaviors and making inferences and decisions) has some sort of culture model implicit in these processes. Few though explicitly separate the culture element from other aspects of behavior and meaning. There are some culture-training systems in which a user learns aspects of a culture by interacting with virtual humans e.g., [15–17]. In these systems, the choice of behaviors is important for the flow of the scenario: user choice of an inappropriate action can lead the characters to react badly and ultimately foil success in the simulated mission. One learns that certain actions, such as smalltalk before a business meeting, or removing sunglasses can lead

to greater success in such a situation. These models are generally limited to specific the utility of individual actions from a culture-specific point of view, and perhaps an aggregate utility computation based on the sum of actions. These models generally do not have a deeper notion of cultural constructs, however, such as roles within a culture, inferences about membership in one culture vs another, or culturally meaningful norms.

[18] explicitly models socio-cultural norms in a task model linking actions with effects in terms of not just physical states but mental models of culturally significant states. This allows an agent to compute tradeoffs between cultural and other goals in making decisions, as well as calculating the effect on a range of culturally meaningful states of the actions of others. Different cultures can be modelled by changing the states that are modelled, the links between actions, and the polarity and strengths of the effects of actions on states. This work has been used to allow different cultural variants in the framework in [17].

5 The way Forward

The ideal cultural model for virtual humans would include both internal and external models. A basic framework would be established to cover approximately the universal aspects of human behavior, with parameters for cultural specific constructs that can allow a range of different behaviors in different cultural situations. An agent should be able to learn these cultural constructs from observing and participating in interaction within that cultural context and should be able to decide on the appropriate cultural model to apply in a given situation. This learning may perhaps be short-cut by explicit programming or non-experiential forms of learning and inference.

In order to get to this state of affairs there are several different sorts of advances needed. First, we need more theoretical work on cultural constructs - what are the important concepts, how are they realized, how do they work with other parts of behavior and cognition? A lot of this is social science work, but there is also a role for computational modelling, so that they theories can be cast at a sufficiently concrete fashion for implementation. Second, we need a lot more empirical work on looking at the distribution of behaviors and cultures across culturally-relevant situations. We also need more modelling work so that agents can properly make use of observations of cultural phenomena and make full use of cultural rules, including using them for production, interpretation, and inference. Finally, we need evaluation work at all levels, verifying that the predictions that a model makes based on one set of data applies as the model says it should to a new situation.

This is an exciting time for this kind of research! Relatively new observational recording equipment and mathematical models can help bridge the gap between detailed micro-analysis of specific behavior in cultural situations and large statistical tendencies that are currently studied using indirect methods such as surveys of attitudes rather than observed behavior. Machine learning and formal modelling techniques can make use of large amounts of data to refine the models. Finally, and most importantly, virtual humans themselves can be a laboratory for studying the predictions of cultural theories and examining unforeseen consequences and gaps in the the theories that suggest areas for further study.

Acknowledgments

I would like to thank a number of people for interesting discussions on culture and virtual human models, including Jens Allwood, Elisabeth Andre, Justine Cassell, David Herrera, Dusan Jan, Stacy Marsella, Bilyana Martinovski, Yosuke Matsusaka, Yukiko Nakano, David Novick, Jennifer Overbeck, Matthias Rehm, Steve Solomon and Nigel Ward. This work was sponsored by ARO Grant number W911NF-08-1-0301 and the U.S. Army Research, Development, and Engineering Command (RDECOM), and the content does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

References

1. Rickel, J., Johnson, W.L.: Virtual humans for team training in virtual reality. In: Proceedings of the Ninth International Conference on Artificial Intelligence in Education, IOS Press (1999) 578–585
2. Gratch, J., Rickel, J., Andre, E., Cassell, J., Petajan, E., Badler, N.: Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems* (2002) 54–63
3. Cassell, J., Sullivan, J., Prevost, S., Churchill, E., eds.: *Embodied Conversational Agents*. MIT Press, Cambridge, MA (2000)
4. Kroeber, A.L., Kluckhohn, C.: *Culture: A Critical Review of Concepts and Definitions*. Random House (1952)
5. Traum, D.R.: 20 questions for dialogue act taxonomies. *Journal of Semantics* **17**(1) (2000) 7–30
6. Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., Anderson, A.: *HCRC dialogue structure coding manual*. Technical Report 82, HCRC (1996)
7. Alexandersson, J., Buschbeck-Wolf, B., Fujinami, T., Maier, E., Reithinger, N., Schmitz, B., Siegel, M.: *Dialogue acts in VERBMOBIL-2*. Verbmobil Report 204, DFKI, University of Saarbruecken (1997)
8. Koda, T.: Cross-cultural study of avatars' facial expressions and design considerations within asian countries. In Ishida, T., Fussell, S.R., Vossen, P.T.J.M., eds.: *IWIC*. Volume 4568 of *Lecture Notes in Computer Science*. Springer (2007) 207–220
9. Traum, D.R., Marsella, S., Gratch, J., Lee, J., Hartholt, A.: Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. [19] 117–130
10. Jan, D., Herrera, D., Martinovski, B., Novick, D.G., Traum, D.R.: A computational model of culture-specific conversational behavior. [20] 45–56
11. Rehm, M., Nakano, Y., Huang, H.H., Lipi, A.A., Yamaoka, Y., Grneberg, F.: Creating a standardized corpus of multimodal interactions for enculturating conversational interfaces. In: *Proceedings of the IUI-Workshop on Enculturating Interfaces (ECI)*. (2008)
12. Jan, D., Traum, D.: Dynamic movement and positioning of embodied agents in multiparty conversations. In: *In proceedings of AAMAS 2007: Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems*. (2007) 59–66
13. Iacobelli, F., Cassell, J.: Ethnic identity and engagement in embodied conversational agents. [20] 57–63
14. Rehm, M., Nakano, Y.I., André, E., Nishida, T.: Culture-specific first meeting encounters between virtual agents. [19] 223–236
15. Johnson, W.L., Beal, C.R., Fowles-Winkler, A., Lauper, U., Marsella, S., Narayanan, S., Papachristou, D., Vilhjálmsón, H.H.: Tactical language training system: An interim report. In Lester, J.C., Vicari, R.M., Paraguaçu, F., eds.: *Intelligent Tutoring Systems*. Volume 3220 of *Lecture Notes in Computer Science*. Springer (2004) 336–345

16. Deaton, J.E., Barba, C., Santarelli, T., Rosenzweig, L., Souders, V., McCollum, C., Seip, J., Knerr, B.W., Singer, M.J.: Virtual environment cultural training for operational readiness (vector). *Virtual Reality* **8**(3) (2005) 156–167
17. Jr, R.W.H., Belanich, J., Lane, H.C., Core, M., Dixon, M., Forbell, E., Kim, J., Hart, J.: Pedagogically structured game-based training: Development of the elect bilat simulation. In: 25th Army Science Conference. (2006)
18. Solomon, S., van Lent, M., Core, M., Carpenter, P., Rosenberg, M.: A Language for Modeling Cultural Norms, Biases and Stereotypes for Human Behavior Models. In: Proceedings of the 17th Conference on Behavior Representation in Modeling and Simulation (BRIMS 2008), Providence, RI (2008)
19. Prendinger, H., Lester, J.C., Ishizuka, M., eds.: Intelligent Virtual Agents, 8th International Conference, IVA 2008, Tokyo, Japan, September 1-3, 2008. Proceedings. In Prendinger, H., Lester, J.C., Ishizuka, M., eds.: IVA. Volume 5208 of Lecture Notes in Computer Science., Springer (2008)
20. Pelachaud, C., Martin, J.C., André, E., Chollet, G., Karpouzis, K., Pelé, D., eds.: Intelligent Virtual Agents, 7th International Conference, IVA 2007, Paris, France, September 17-19, 2007, Proceedings. In Pelachaud, C., Martin, J.C., André, E., Chollet, G., Karpouzis, K., Pelé, D., eds.: IVA. Volume 4722 of Lecture Notes in Computer Science., Springer (2007)