When a Machine Picks a Fight

notes on machinic male-dicta and synthetic hissy fits

Marc Böhlen
MediaRobotics Lab
Department of Media Study
University at Buffalo
marcbohlen@acm.org
www.realtechsupport.org

Abstract
In this paper I describe a rational and an experimental framework for aggressive synthetic agents.

Keywords
Aggressive synthetic agents, synthetic accents, synthetic foul language, transgressing human social and linguistic conventions in synthetic systems

ACM Classification Keywords
H. User/Machine Systems, H.1 Models and Principles, H.1.2 User/Machine Systems, Software Psychology. J. Computer Applications, J.5 Arts and Humanities, Linguistics.

Introduction
The history of HCI and social robotics is ripe with interaction scenarios based on benevolent and playful synthetic agents [2], [3] and robots [6]. Critical analysis of such assumptions has been previously voiced in Science Technology Studies and Cultural Theory [8], [9]. Within the HCI community, critical reflection on (embodied) synthetic agents and embodied conversational agents (ECAs) is more recent [1]. Interaction scenarios valid across cultural boundaries [5] and analysis of rude user reactions to ECAs have also been reported [1]. In the speech

processing community, synthetic accents are now under investigation [4].

In light of recent interest in negative emotions in computing, I describe here an agent scenario that transgresses accepted norms of polite behaviour.

Amy and Klara have similar interests. They both read Salon.com. But they do not get along. Not at all. Maybe Klara's thick German accent bothers Amy. And neither of them particularly likes the color pink. Unfortunately for Amy and Klara, they live on the same block and have pink houses! And when they become agitated they tend to fall into mutual accusations and rants. Yes, it can get rather nasty at times. Best then just to leave them be and to stay clear of the hissy fits.

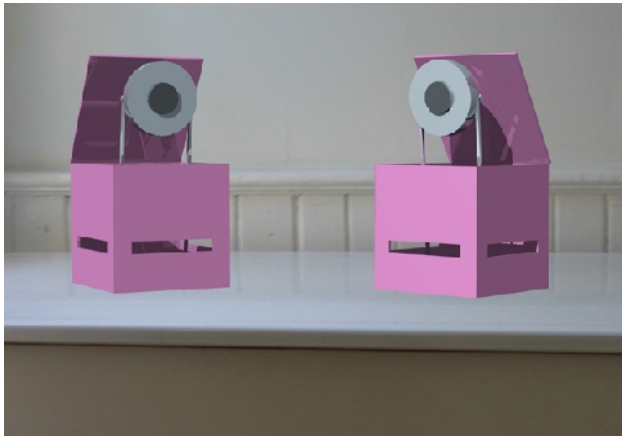

Figure 1: Amy and Klara (*work in progress*)

Against normalized interaction

Almost all HCI interaction schema filter conflict between humans and machines out of the exchange. This has resulted in a very one-sided, normalized interaction design strategy. By filtering conflict out of the concept of interaction we become unable to deal with it when it actually occurs. Furthermore, it makes sense to experiment with scenarios between synthetic agents and people that are not bound by conventions historically established between people alone. A machine-human future that does not actively seek alternate scenarios and is not willing to integrate impulsive -and other forms of irrational- behavior cannot, I believe, become a successful long-term and socially robust interaction paradigm.

Under such an assumption it is logical to add conflict, arguments and fights, including foul language, into the portfolio of interaction design schema. In this experiment, shallow human-like expression is contrasted with overtly un-human appearance. As some of Hollywood's most successful agent incarnations prove (*Hal, R2D2*), it does not take a human face to achieve short-term believable presence. By refusing physical anthropomorphism one can avoid the consequences of crossing the uncanny valley of imperfect mimesis.

Picking a fight

It is not particularly difficult to create aggressive actions in an agent; we humans deliver ample sample data. Amy and Klara are created almost identical to each other: They have the same architecture, they are fed by the same information sources (online life style magazines), they both are housed in pink boxes, and they both have a mechanism by which they make small talk and foul language. They share the results of their (statistical) evaluation of the online magazines with each other through text to speech and automated

speech recognition. However, the results from the speech recognizer as well as the physical transmission of utterances from speaker to microphone are error prone. Even the best speech recognizers offer often spotty recognition, particularly in noisy environments. Hence miscommunication is unavoidable. If several misunderstandings occur in a given time frame, aggression, for which the agents have a programmatic disposition, increases and foul language comes into play. The fact that one agent has a bad German accent only increases the potential for misunderstanding. Exposure to the color pink, to which they are negatively sensitized, compounds their respective aggression levels. This simple mix can lead to rather rough exchanges as the examples available online illustrate [10].

Guilty interaction

In this absurdist cabaret-like scenario people are seen by the agents as outsiders. The only kind of interaction that occurs between people and these two boxes is through verbal spillage. One can overhear the nasty exchanges between Klara and Amy and listen to the rants, much like one might listen to an argument amongst a couple at a nearby table in a restaurant. Curiosity, guilty voyeurism and the strange kind of satisfaction that can be obtained by listening to others wash their dirty laundry in public is the reward for those who participate. However, once both agents perceive the presence of real people through their built-in video cameras, they lower their voices, muffle their foul utterances or interrupt their nasty exchange and ask the gaffers to leave, temporarily altering the hierarchy between humans and synthetic agents. They then wait until they are alone again or slide back into their pink boxes where they eventually calm down (the aggression curves are modulated by a time-dependent decay function). Thereafter, they resume reading their magazine collections.

Foul language

Even intelligent beings are capable of dumb behavior. In foul language, people show some of their wittiest and stupidest traits at once. Foul language is a conduit into aspects of lived communication filtered from polite conventions. Foul language is the most obvious but least useful vocabulary expansion information centric agents, often specialized for commerce, might receive. However, this addition does allow one to reflect in new ways on how language relates to the world of synthetic beings. Many instances of foul language are derived from taboos in religion, sex and madness. Many taboos are directly related to the physical constraints of being human and have, as is the case in defecation, a close correlation between the degree of taboo in verbal usage and the degree of taboo in public exhibition [7]. Since machines lack our bodily functions, the corresponding taboos need not hold. Despite the logic there is likely to be little acceptance of machines cussing profusely in the presence of people. But will we map all our own taboos onto machines or might some taboos become acceptable? Might there be new curse words particular to the experience of being machine?

Conclusion

This work-in-progress is a small contribution to an interaction philosophy that includes irrational acts on the periphery of information exchange; acts for which there is no obvious need, but which can assist in imagining a less gentle and perhaps more realistic shared future between people and machines.

Technical notes

The agent programs are written in python and XML with the open source AIML environment. The sensitivity and agitation levels of the two boxes are set by evaluating texts from online life style magazines (*Cosmopolitan, Salon.com*). A machine vision module written in C with the open source OpenCV library checks for the presence of people and pink objects. Sound data captured by noise reducing microphones is piped to a speech recognition engine (FONIX). The agent programs running on each of the agent computers evaluate responses to the incoming sound and image data. When one box starts to speak, the other responds. If an instance of foul language is found in the utterance, it is countered with one of a similar flavor from a database of tagged curse words. The responses are synthesized with a proprietary speech synthesis engine (SVOX). The German accent is generated at run time by swapping select vowels and consonants between the SVOX language models for German and English and applying several ad hoc SAMPA alphabet based phonetic remappings for special cases. The resulting audio stream is then modified by a C program using an open source sound-processing library (SoX) before being sent to the audio output device.

Acknowledgements

This work is supported in part by SVOX, the ETH Switzerland and FONIX SPEECH.

Citations

[1] Angeli, D.A. and Carpenter, R., Stupid Computer! *Abuse and Social Identity. Abuse: The dark side of human-computer interaction*: pp. 19-25, 2005.

[2] Bartneck, C., Affective expressions of machines. *Conference on Human Factors in Computing Systems,* CHI '01, extended abstracts on Human factors in computing systems, pp.189 – 190, 2001.

[3] Bickmore, T. and Cassell, J., How about this weather? Social Dialogue with Embodied Conversational Agents, *Proceedings of the American Association for Artificial Intelligence* Fall Symposium on "Narrative Intelligence", pp. 4-8. November 3-5, Cape Cod, MA. 2000.

[4] Black, A. and Lenzo, L., Foreign Accents in Synthetic Speech: Development and Evaluation, *InterSpeech2005*, Portugal, 2005.

[5] de Rosis F., Pelachaud C. and Poggi I., Transcultural believability in Embodied Agents: a matter of consistent adaptation. In: S. Payr and R. Trappl (Eds): *Agent Culture: Designing Human-Agent Interaction in a Multicultural World*. Laurence Erlbaum Associates, 2003.

[6] Fong, T., Nourbakhsh, I., and Dautenhahn, K., A Survey of Socially Interactive Robots, *Tech. Rep. CMU-RI-TR-02-29*, Rob. Inst., CMU, 2002.

[7] Hughes, G., Swearing. A Social History of Foul Language, Penguin, London, 1991.

[8] Sengers, P., Symptom Management for Schizophrenic Agents, *AAAI-96*. Menlo Park, CA: AAAI Press. Vol 2., p. 1369, 1996.

[9] Zdenek, S., Artificial intelligence as a discursive practice: the case of embodied software agent systems. *AI & SOCIETY*, Springer-Verlag London, 2003.

[10] www.realtechsupport.org/new_works/male-dicta.html