Social Interaction with a Conversational Agent: An Exploratory Study

Yun-Ke Chang, Nanyang Technological University, 31 Nanyang Link, Singapore 637718; E-mail: ykchang@ntu.edu.sg

Miguel A. Morales-Arroyo, Nanyang Technological University, 31 Nanyang Link, Singapore 637718; E-mail: miguel@yun-ke.com

Mark Chavez, Nanyang Technological University, 31 Nanyang Link, Level 3, Singapore 637718; E-mail: mchavez@ntu.edu.sg

ABSTRACT

Conversational agents that display many human qualities have become a valuable method business uses to communicate with online users to supply services or products, to help in online order process or to search the web. Gaming industry and education may benefit from this type of interface. In this type of interface, users could have different alternatives: text display, photo of a real person, or a cartoon drawing and others. This is an exploratory study that reviews five randomly chosen conversations that an animated chatbot has with web users. The character simulates human gestures, but they are stylized to reproduce animation standards. The goal of this exploratory study is to provide feedback that will help designers to improve the functionality of the conversational agent, identify user's needs, define future research, and learn from previous errors. The methodology used was qualitative content analysis. The results of the study suggest improving the flexibility of the conversational agent, and reducing repetitions in the interaction.

INTRODUCTION

For today's online business whether selling a service or a product, the main goal of web sites is to keep its users at the site as long as possible. As an interface, a conversational agent has to offer some features to maintain its audience interested. Concerns about agents' body design and personal sociability have become obvious. Users may favor an interface that suits their own personality. They also may be more predisposed in an animated exchange with an agent if the character's voice matches content tone with gesture that complements the context.

Animation synthesis procedure permits to create through numerous randomly interconnected cycles a dynamic web based agent. Nadia (http://CLONE3D. com), a conversational chatbot, was developed by the third author, and it is a human like agent able to perform dialogues with users by "comprehending", generating phonemes with automatic lip-sync, and expressing body language, including body movements, hand actions, and facial gestures. The lighting of the virtual agent is practically naturalistic and uses conventional illumination techniques (See Fig. 1).

The design of conversational agents has to face a set of challenges: promote trusting relationships with their audience (Cassell and Bickmore, 2003), body

Figure 1. Nadia - a virtual character

language matching (Cassell and Bickmore, 2000), ability to communicate in different languages, and adapt to different cultural contexts. An intelligent real-time 3D artificial agent unlocks additional opportunities for computer mediated communication. The facial expressions in the agent are critical in a dialog and could be used with hearing-impairment audience (Massaro, et al., 2001). The goal of this exploratory study is to provide feedback that will help designers to improve the functionality of the conversational agent, identify user's needs, define future research, and learn from previous errors.

DESIGN

Actions of the character are conceived to replicate human gestures, but behaviors are stylized to replicate animation standards. This design style builds a friendly web environment with web audience. The character employs AIML, Artificial Intelligence Markup Language, the Generation 4 Player, and Automated Personality Tool Set. Generation 4 is a player that simulates fluid full-body movement, mass and reactive actions, body tissue, gravity, and realistic physical responsiveness. The principles that guide the general design come from Lee and Lacey (2003).

The character is shaped with high production attributes in mind. Physical features containing an inner mouth with teeth and tongue, accurate fingers with fingernails and eyes are included in the model. The character also has over 50 facial points assembled into her expression database. These points are utilized to produce both a set of visemes, a basic unit of speech in the visual domain, to match phonemes for accurate lip-sync, other vocalizations, and present emotional changes in the facial expressions. Nadia's body features were developed for users' interest. The figure and components of the character's face and body are overstated to follow standards employed in caricature projects. For Nadia, the design target is to create an appealing, smart, young, female character, which is a standard method used in the development of female characters for broad appeal. To compliment the physical design, the personality is provided by ALICE bot, and its knowledge base was modified to answer with jokes by a process of trail and error.
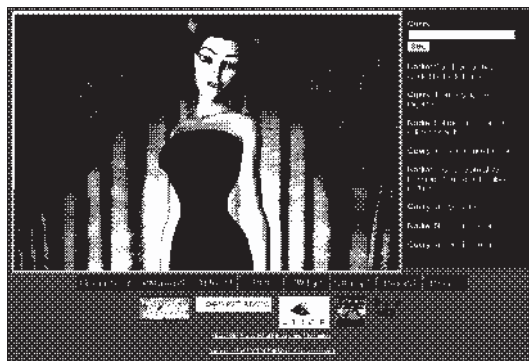
TECHNICAL ENVIRONMENT

Haptek supplies a 3D player that is mostly employed to produce the character's movement. This technology can connect the model to a skeleton and provide method for correcting volume problems that appear when the character's extremities are in action. However, it does not have a simple approach to improve the model and re-skin the geometry to the skeleton. It has a pipeline for importing motion captured from the current skeleton, but export tool does not export it. The bone framework is used to have an easy method to incorporate skinning techniques. The Haptek player's toolset is hampered by its incapacity to supply the correct adjustments for the characters body, its poses, and mass volumes.

Currently the ALICE, an open source chatbot technology used for the verbal/text based responses, is limited by the amount of content that the knowledge base has. The knowledge base can be quite large and needs to address the questions and issues the user would have to develop an interesting character. ALICE knowledge base was modified to provide jokes to user by trial and error.

PROJECT GOALS

The primary objective was to develop a project called Artificial Comedy. A number of other characters are to be developed with the goal of building a set of perform-

ers: actors, singers, comedians, and other entertaining characters. Each one would have their own environment and friends. These environments would become part of a collection of bizarre expressions and comical performances of representative characters and circumstances. Exploring the exchanges these avatars have with online users will help in fine tuning their design.

RESEARCH DESIGN
Visitors
Authors could not identify who specific visitors were or population demographics, but from the server statistics some information was gotten such as top referring sites, visitor's country, queries visitors utilized to search.

According to the server statistics visitors who have requested files had come from the following domains: .au (Australia) 0.88 %, .ca (Canada) 1.34 %, .com (Commercial) 30.46 %, .de (Germany) 2.48 %, .edu (USA Higher Education) 0.77 %, .fi (Finland) 0.77 %, .fr (France) 1.85 %, .it (Italy) 1.70 %, .net (Networks) 26.84 %, .nl (Netherlands) 6.62 %, .pl (Poland) 1.12 %, .uk (United Kingdom) 1.12 %, [unresolved numerical addresses] 15.20 %, Others 8.77 %.

The words people used in search engines to find the site in the last seven days: 3d 1.80 %, animated 0.20 %, artificial 3.51 %, character 0.60 %, characters 1.20 %, chavez 0.10 %, clone3d 0.20 %, download 0.30 %, girl 43.69 %, haptek 1.10 %, intelligence 0.30 %, mark 0.10 %, stories 0.30 %, time 0.20 %, virtual 40.58 %, virtualgirl 1.10 %, not listed: 1,124 search terms 4.71 %.

Visitors used the following queries in search engines to visit the site: artificial girl 5.68 %, virtual girl 82.10 %, virtual 3d girl 0.21 %, virtual girl 3d 0.21 %, 3d characters 0.42 %, artificial girl 2 0.42 %, clone3d 0.42 %, artificial girl download 0.42 %, haptek characters 1.05 %, 3d virtual girl 1.26 %, virtualgirl 2.10 %, not listed: 1,283 search terms 5.68 %.

METHODOLOGY
The dialogues between the chatbot and the online users were recorded in a log file. Five dialogs were chose randomly, some of them were more than two hours long, and all of them were more than one hour long. Chats took place from June 1st. to June 19 from a set of more than 170 dialogues. Qualitative content analysis was used to analyze the data, utilizing QSR N6, previously known as Nudist. The categories were created in an inductive process, common in content analysis methodologies. These categories were examined by three authors, each one initially working independently, and later coming together to get a consensus, but inter-code reliability was not estimated. These categories were developed from the conversational agent's perspective, for example, how to supply information and help its visitors, how well it "comprehended" visitors' contributions, how it managed imprecise input, et cetera.

RESULTS
Topics Addressed in Dialogs by Visitors
One user had an illogical conversation with correct grammar in general terms and no spelling mistakes. Other asked for obscene expressions and jokes, and another required affectionate conduct from the chatbot.

One user tested the factual knowledge the agent posses and focus his questions on literature. He asked about who Mark Twain was, who wrote Tom Sawyer and his nickname, Philip K. Dick, a science fiction writer, and other questions. The same user tested the agent's analytic aptitude in Mathematics: the ability to solve equations, the Riemann hypothesis, and the demonstration of last Fermat's theorem.

Three visitors were trying to persuade the chatbot to engage in some sort of virtual sex. Two of them were interested in movies; one of them in science-fiction pictures, like Star Trek, Star Wars, and Matrix, and the other one wanted to know where to download free sex movies and movies in general. Two users showed great interest in the conversational agent competence, and one of them was interested in its potential.

Description of Dialogs
A dialog that had sessions as long as 43 hours may not imply the dialog was that lengthy. Visitors in addition to chat with Nadia did other things. Conversation one lasted one hour 22 minutes in June 19, 2006.

Probably, visitor one knew what to expect from a chatbot and asked: "Nadia, can you tell me why Eliza Clone copies everything I say". He attempted to comprehend the potential this chatbot has, and asked if he gave his name, the chatbot could recall it. In the beginning, the agent did not follow properly the dialogue, even visitor one was making a significant effort to maintain the dialogue flowing and making sense of the conversation.

The agent check the initial word of the next expression: "nice talking with you" and answers, "thanks for the compliment". It creates the impression when a visitor provides congruent responses, the agent is more proactive, and looks like the agent is following the conversation.

Visitor one handled the agent courteously as he was interacting with another person; in addition, the interaction was very positive. This user was interested in the agent's level of intelligence and its learning capacity. He explained different issues, tried to reason, talked about movies, and asked the agent for help to find a game that was only sold in Japan. This user made a big attempt to maintain the agent in track with the conversation, but the lack of contextual knowledge and knowledge in general limited the conversation. He had some misspellings, and his behaviour was very polite.

Dialogue two took one hour 22 minutes in June 1, 2006. The conversation was trivial, but the interaction was intense. For example, Nadia said "Excuse me! You are a really tedious dope". Only a few interactions have continuity, and the dialogue was chopped and repetitive. There was no proper beginning and end, some incorrect sentences, misspellings, and insults from the users. Probably, visitor two was a chatbot. The dialogue was full of non-sense, which is someway appealing and amusing to observe how two chatbots may interact with each other.

Conversation three took twenty-one hours and 40 minutes in June 14 and 15, 2006. Initially, visitor three did not make sense. He used single word or short expression without giving context. The chatbot behaviour became reiterative, and this user criticized that behaviour by exclaiming "stop repeating".

Visitor three articulated sexual expressions, gave orders, conveyed obscenity, and wrote sexual abbreviations such as "omfg" that Nadia did not understand. This visitor became repetitive as he used the same expression several time and twice consecutively, "one point for you", and he requested affectionate conduct from the chatbot. When he did not get what he demanded, he insulted Nadia. This dialogue in general was not interesting because of the amount of single words written by the visitor, and his involvement was limited in most part of the conversation.

Dialog four lasted 43 hours and 42 minutes from June 1st to June 3rd, 2006. Visitor four had belligerent actions, attacking the agent for not being sharp enough and did not fulfill his request for virtual sex. His language was contradictory; he was curious in the bot's analytical capabilities, intelligence, and its potential as information agent.

Dialogue five took 38 hours and 30 minutes from June 11 to June 12, 2006. Visitor five was basically motivated in one topic, virtual sex, including coarse language. He used diverse tactics to accomplish his objective. Frequently, when he failed, he abused the chatbot. The tactics he used were the followings: asking to perform some actions in his subject of interest, coercion, adulation, argumentations, et cetera.

Categories Developed
The categories arose from the agent perspective, its specific behaviors and answers during its dialogs. The sample of users is not necessarily representative of the population. The creation of these categories follows an inductive process typical of content analysis. Each author did the analysis independently, met, and achieved consensus over the final categories. The categories developed were the followings:

Ambiguous Input - Answers to short expression without context;
Bot Talking - agent talking like machine;
Complements - Answering and providing complements;
Greeting & Bye - Welcoming and saying good bye;
Help - providing information and suggestions offered by the agent;
Insults - Answering insults;
Intelligence & Knowledge - Factual knowledge and pseudo - analytical abilities;
Jokes - Agent humor;
Leading Comments - Agent answers guide the user to some topics;

Memory - "Remembering" information provided by the user;
Miscommunication - failure to communicate clearly;
Personal Issues - Exchange of personal information;
Proactive Behavior - Learning possibilities;
Repetitive Behavior - Recurring answers; and
Technical Problems - Technical difficulties the user confront.

DISCUSSION

Some operational troubles in the agent are interconnected. When a user brings in ambiguous inputs in the conversation, the agent becomes "confused". It does not have explicit context to maintain an articulated dialog. A dialog needs at least two parties, but the user may not intend to participate in the social exchange. The chatbot can be proactive, offer help and suggestions, and make remarks that may direct the conversation into topics in which it has its main assets. For example, when a user protested about technical troubles, the chatbot offered some recommendations. When the user asked for information, the chatbot answered them to do any of the followings: to check the open directory, answers.com, or to send an email to. Although, this is an adequate alternative, due to the limited amount of programmed answers, sometimes the chatbot sounds reiterative.

Some users deplored chatbot's lack of functions and repetitive answers. These issues could be solved by providing the agent with a variety of answers, employing the contextual information provided by the user, and improving and handling agent memory. Lack of conversation flow and miscommunication were frequent practice, but not always due to agent problems. Sometimes, the users were distracted. Users became perplexed when the chatbot repeated phrases, providing wrong answers, or not following the dialogue sequence.

In general, the chatbot is polite, and it can compliment users. At the same time, it is able to receive compliments graciously, but sometimes it answers with similar phrases, which could make the communication tedious. It is capable to respond some offences, but not all of them. The chatbot present a dual behaviour, sometimes it expresses clearly that it a conversational agent to the user, but not always. One of the most critical difficulties that users described in the dialogues was the agent voice. At that time in some conversations, not only the voice did not appear natural, but also not as woman speech. This problem was solved, but it generated some confusion in the users.

CONCLUSION

From this research, the authors found out that users in these conversations did not employ the chatbot for humor purposes. Some of jokes the agent can tell may not be adequate for some users, and we will have to find what type of joke visitors would like to hear. The opening and the end of a dialog are important. In the beginning, the chatbot, besides providing its name to the visitor, may provide a presentation about what it is and is capable to do. In that sense, visitors will not have expectations that the conversational agent is not able to perform. Among the areas for improvement are having better welcoming segment, involving more diverse answers, and reducing the length of the goodbye segment when the visitor provide clues that he has to abandon the dialog. Although, miscommunication cannot be prevented, it may be decreased if the conversational agent becomes more proactive and has better logs. The latent technical troubles should be recognized, and the possible answers should be part of the assistance the chatbot provides to its visitors.

ACKNOWLEDGEMENTS

REFERENCES

Cassell, J., Bickmore, T. (2003). Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents, User Modeling and User - *Adapted Interaction*, 13(1) pp. 89-132.

Cassell, J., Bickmore, T. (2000). External manifestations of trustworthiness in the interface, *Communications of the ACM*, 43(12) pp. 50-57.

Massaro, D. W., Cohen, M. M., Beskow, J., Daniel, S. and Cole, R. A. (2001). Developing and Evaluating Conversational Agents, In Cassell, J., Sullivan, J., Prevost, S., Elizabeth Churchill, E. (Eds.) *Embodied Conversational Agents*, MIT Press, 287 – 318.

Lee, M. H., Lacey, N. J. 2003. The influence of epistemology on the design of artificial agents, *Minds and Machines*, 13 pp. 367-395.

Related Content

Fault Analysis Method of Active Distribution Network Under Cloud Edge Architecture

Bo Dong, Ting-jin Sha, Hou-ying Song, Hou-kai Zhaoand Jian Shang (2023). *International Journal of Information Technologies and Systems Approach (pp. 1-16).*

www.irma-international.org/article/fault-analysis-method-of-active-distribution-network-under-cloud-edge-architecture/321738

Reasoning on vague ontologies using rough set theory

(). *International Journal of Rough Sets and Data Analysis (pp. 0-0).*

www.irma-international.org/article//288522

Repurchase Prediction of Community Group Purchase Users Based on Stacking Integrated Learning

Xiaoli Xie, Haiyuan Chen, Jianjun Yuand Jiangtao Wang (2022). *International Journal of Information Technologies and Systems Approach (pp. 1-16).*

www.irma-international.org/article/repurchase-prediction-of-community-group-purchase-users-based-on-stacking-integrated-learning/313972

Consumer Adoption of PC-Based/Mobile-Based Electronic Word-of-Mouth

Akinori Onoand Mai Kikumori (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 6019-6030).*

www.irma-international.org/chapter/consumer-adoption-of-pc-basedmobile-based-electronic-word-of-mouth/184302

Power System Fault Diagnosis and Prediction System Based on Graph Neural Network

Jiao Hao, Zongbao Zhangand Yihan Ping (2024). *International Journal of Information Technologies and Systems Approach (pp. 1-14).*

www.irma-international.org/article/power-system-fault-diagnosis-and-prediction-system-based-on-graph-neural-network/336475