Enhancing Computer Chat: Toward a Smooth User-Computer Interaction

Calkin A.S. Montero and Kenji Araki

Graduate School of Information Science and Technology, Hokkaidō University,
Kita 14-jo Nishi 9-chome, Kita-ku, Sapporo City, 060-0814 Japan
{calkin,araki}@media.eng.hokudai.ac.jp
http://sig.media.eng.hokudai.ac.jp

Abstract. Human-computer interaction (HCI) has fundamentally changed computing. The ubiquity of HCI can be seen in several kinds of application areas, such as text editing, hypertext, gesture recognition, and the like. Since HCI is concerned with the joint performance of tasks by humans and machine, human-computer conversational interaction plays a central role when trying to bring the computer behavior closer to the human conversational behavior. In this paper we introduce our idea of modeling computer *chat* based on the observed "critical" behavior of the structure of human chat.

1 Introduction

Theoretical and practical research on the structure of the human dialogue has been a topic of investigation for decades in Artificial Intelligence (AI). Since Alan Touring [1] presented his philosophical dissertation regarding the question "Can machines think?", conversational understanding has been considered the mean to determine whether the machine is intelligent. Turing proposed an "imitation game" in which an interrogator poses questions to an unseen entity (machine or person) then based on the responses the interrogator judges which is which. The main issue for Turing was that using language as humans do is sufficient, by itself, to determine intelligence.

Since, human-computer conversation (HCC) as part of human-computer interaction (HCI) has been growing until becoming nowadays one of the most important areas in AI, reaching a similar grade of development like that of a better-known area of language processing, i.e., machine translation (MT). One of the most famous examples of HCC systems is ELIZA [2] in which many of the issues raised by Turing became relevant. ELIZA is a pattern-matching natural language program that performs the role of a Rogerian therapist during its conversation with a user. The program worked by making use of a cascade of regular expressions substitutions that each matched certain part of the user's input and changed it into the system's reply.

A best performance was PARRY program [3]. PARRY attempted to simulate a paranoid patient, implementing a model behavior based on conceptualizations and beliefs - i.e., accept, reject, neutral: judgments of conceptualization. PARRY

R. Khosla et al. (Eds.): KES 2005, LNAI 3681, pp. 918–924, 2005.

introduced some important ideas for the development of conversational agents: topic and context understanding and conversational strategy or dialogue management.

In recent years a considerable amount of research has been carried out on conversational spoken systems. In this regard Jupiter [4], system that provides a telephone-based conversational interface for international weather information, stands as a good example. Other examples of conversational systems include airline travel information [5], speech-based restaurant guides [6], and telephone interfaces to emails or calendars [7]. The dialogue manager is the core of such systems, regulating the flow of the dialogue, deciding how the system's side of the conversation should proceed, what to ask or what assertions to make and when to ask or make them, being the dialogue constrained to certain specific domain.

Several approaches to modeling constrained dialogue have been proposed. Among those, dialogue grammar approach [8], based on the observation that there are regularities of sequence during the dialogue, as a question is follow by an answer, proposal by acceptances, and so forth. However, this approach assumes that only one state results from a transition, when actually utterances are multi-functional, i.e., an utterance can be both, a rejection and an assertion, which makes the approach unsatisfying. Plan-based approach [9], based on the observation that the user's utterances are not simple string of words but communicative actions or speech acts. This approach assumes that the user's speech acts are part of a plan and the goal is to uncover and respond appropriately to it [10]. However, one of the limitations of the plan-based approach is that the illocutionary act recognition, needed in order to infer the user's plan, has shown to be redundant [11]. Finally, joint action or collaborative approach, based on the assumption that both parties to a dialogue are responsible for sustaining it, existing a collaboration between the interlocutors in order to achieve mutual understanding [12]. However, all of those dialogue modeling approaches focused on creating dialogue agents in order to achieve specific goal, being their dialogues domain restricted, failing when it comes to hold a human-like dialogue.

Recently, the development of non-specific-goal[1] dialogue agents has exponentially increased. Those agents are called chat robots or chatbots. A chatbot is defined as a computer program that simulates human conversation through AI. In fact, ELIZA has been considered the first created chatbot. The dialogue modeling of a chatbot is based on certain words combinations that it finds in a phrase given by the user, and in the case of sophisticated chatbots, like ALICE [2] [13], AIML[3] is used to transcribe pattern-response relations. Chatbots are wide spread on the Web, being applied for electronic commerce customer service. An ELIZA-clone chatbot application to question answering has been proposed by C.S. Montero and K. Araki [14, 15]. However, it is highly improbable to create all the possible patterns-response that could appear during a dialogue, a shortcoming that makes the chatbot to easily lose the conversation flow.

[1] Open domain dialogue.
[2] Artificial LInguistic Computer Entity.
[3] Artificial Intelligence Markup Language, base on eXtended Markup Language, XML.

In this research a deep analysis of the human chat has been performed in order to apply its characteristics of criticality, described in Sect. 2, to modeling computer chat. Experiments show an improvement of the chatbot performance after providing its database with "critical categories," as it is described in Sect. 3. A conclusion with reflections on future directions of the research is given in Sect. 4.

2 Characteristics of Criticality in Human Chat

A human chat section has been considered *critically self-organized* [16]. Self organized criticality states that large interactive systems naturally evolve toward a critical state in which a minor event can lead to catastrophe. The critical state is established entirely due to the dynamical interactions among individual elements of the system, implying self-organization. Considering the system to be a human chat section, we argue that the chat is a dynamic system whose environment is formed by both interlocutors utterances. The chat evolves from an initial flat state to a state where a single utterance can completely change the direction of the conversation - catastrophe -, introducing a new topic, repeating the cycle again. This behavior has been called *Critically Self-Organized Chat* [16]. It could be argued that being aware of the "critical behavior" of the human chat allows to create more human-like dialogue flows when modeling computer chat, therefore its importance.

The critical behavior of the human chat can be observed in the following fragment of a transcription of a real chat between two friends[4]:

.
1A: Have you had your lunch
1B: No
2A: Uhm I haven't had my lunch also
2B: Uh busy uh
3A: Uhm busy playing with my nieces
3B: Both of them are in their own place eh at your home now
4A: Hah The younger one is always at my home mah
4B: Both of them – Orh
5A: My mum look after her
5B: Ah
6A: The baby is only two three months
6B: Um so weekends the parents come over lah
...
7A: You hear my little niece crying
7B: Heh uh Uhm so are you spring cleaning already
8A: Ah intend to do it today lor
8B: Orh
....
9A: Ah – Then weekdays definitely can not – Weekdays I'm so busy in the office you know
9B: Ah hah
10A: This whole this week itself ah
10B: Ah hah
.

In this chat, the utterances 1A to 2B form the flat initial state of the system: topic-introductory utterances. At this point there are not big changes in the chat, hence it could be said the system is in equilibrium. The next utter-

[4] University of South California, Dialogue Diversity Corpus, Dialogue 91, http://www-rcf.usc.edu/~billmann/diversity/DDivers-site.htm.

ance - 3A - initiates the chat toward a determined topic: A's nieces. This topic is remaining the main one, although small "avalanches" occur, it is to say, the focus of the topic is slightly changed, until the utterance 7B (the number of the utterances does not reflect the number of turns taken by the speakers) that crashes the main topic (catastrophe) giving birth to a complete new direction of the chat: laundry time. The utterance that leads to a catastrophe - changing the direction of the chat - has been called "critical utterance" [16]. It is worth noticing how the new topic rises from a flat equilibrium state to a state where a unique utterance - 9A - changes again the direction of the chat, starting a new cycle.

Taking in consideration the critical self organization of the human chat when modeling computer chat, we argue, a more human-like, smooth performance could be achieved. In Sect. 3 it is presented the basis of the application of criticality to modeling computer chat along with a design experiment.

3 Method: Modeling Computer Chat

In order to model computer chat, taking in consideration the properties of criticality of the human chat, it is needed to observe the behavior of the human chat in terms of the interrelationship between the speakers utterances. This interrelationship can be considered as the utterance co-occurrence during the chat. In order to achieve this goal, we applied a data mining tool called KeyGraph. The KeyGraph [17] identifies relationship between terms in a document, particularly focusing on co-occurrence relationships of both high-probability and low-probability events.

We analyzed the chat section[5] shown in Sect. 2. Each speaker's utterance was segmented into words, being those words filtered - eliminating common ones, i.e., I, you, is, and the like - and a sentence co-occurrence relationship was determined as:

$D=$

w_1:: S_1, S_2, S_4 ...

w_2:: S_9, S_{25} ...

w_3:: S_1, S_3, S_{10} ...

...

w_n:: S_{24}, S_{25}, ... S_m

where:

w_k $(k = 1, 2, 3, ..., n)$, represents a word in a sentence.

S_l $(l = 1, 2, 3, ..., m)$, represents a sentence (utterance transcription).

Feeding the KeyGraph with this preprocessed document (preprocessed chat section), the visual result showed clusters of interrelated sentences, where one critical utterance was leading to the other, and the links of the clusters were showing the shifting between topics during the chat.

[5] During the analysis the complete chat section was taken in consideration.

3.1 Design Experiment

In the experiment a native English speaker performed a chat with ALICE chatbot. The performed chat was analyzed, following the above method, in order to find co-occurrence between the user's utterance and the chatbot replies. The graphical view of this chat section could be seen in Fig. 1.

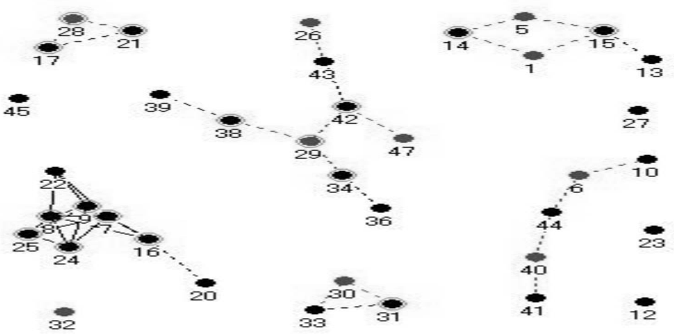

Fig. 1. Graphical View: Non Critical Chat.

In this figure, the clusters represent the relationship between the interlocutors utterances (user and computer) and the links between clusters represent the transition between one topic and the other. It can be observed that the main clusters are not interconnected, meaning the chatbot in many cases could not keep a natural flow of the chat, giving as a result vagueness during the dialogue. This chat section had 48 turns of the interlocutors (user - chatbot).

After analyzing this graph, in order to enhance the computer chat, we try to add criticality to the dialogue by making the chatbot to ask "intelligent questions" at certain points of the conversation, so as to make the shifts from one topic to the other, i.e., interconnecting the clusters shown in the graph, more naturally and in a human-like way. The chatbot database is made of "categories", each one including a pattern to match, i.e., user input, and a template, i.e., chatbot reply. Adding "critical categories" to the database the performance of the chatbot has shown improvement. For instance, if there is an utterance the chatbot does not know how to reply to, a general pattern for smoothly shifting the topic, by asking a question at this point, as to create a link in the graph, gives to the dialogue the desired criticality.

The same user was requested to performed a chat with the enhanced chatbot. The result of the analysis is shown in Fig. 2. This chat section was described by the user as "more interesting" than the previous one, having 82 turns of the interlocutors. In Fig. 2, the shown clusters are more interrelated than in Fig. 1 indicating a better flow of the chat section. For example, a fragment of the performed chat between the user and the chatbot is as follow:

Table 1. Improvement of the Chatbot

	No. of Turns	% Vague Replies	User's Opinion
Baseline Chatbot	48	14.6%	"loosy"
Enhanced Chatbot	82	3.7%	interesting to talk with

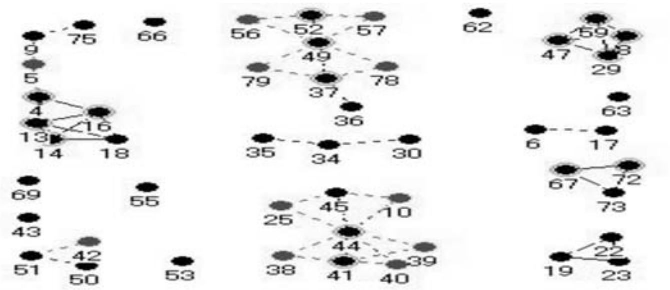

Fig. 2. Graphical View: Critical Chat.

Chatbot: What is like living in Sapporo?
User: sweet as
Chatbot: In the context of living in Sapporo I don't understand "what is like living in Sapporo?"

In this case there was not a pattern for matching the user's utterance "sweet as", being the rules for non-matching patterns of the chatbot a failure. After adding criticality to the chatbot, a reply for the expression is:

User: sweet as
Chatbot: "sweet as"! what is the weather like over there?
User: is a nice day so far

Table 1 shows a resume of the chatbot performance. The success of the method it is difficult to major mathematically due to its novelty, however, the satisfaction of the user while chatting with the enhanced chatbot indicates its betterment. Although the performance of the chatbot is not completely perfect, as it can be seen in Fig. 2 not all the clusters are linked to each other, a remarkable improvement in its performance can be observed, having less vague replies, showing the validity of the approach.

4 Conclusion

In this paper a novel method for enhancing computer chat has been proposed. Taking in consideration the observed critically self-organized behavior of the human chat when modeling computer chat has shown to be an useful approach. By making the chatbot to ask "intelligent question" at certain points of the chat, the criticality increases obtaining as a result a more natural flow of the dialogue, being the chatbot able to keep the chat section going longer time. Future works are oriented toward the application of the enhanced chatbot to Web-based question answering (QA) system in open domain.

References

1. Turing A.M.: Computing Machinery and Intelligence. Mind, Vol. 59, No. 236, pp. 433-460. (1950)
2. Weizenbaum J.: ELIZA-A Computer Program for the Study of Natural Language Communication Between Man and Machine. Communications of the ACM 9, No.1, pp. 36-45. (1966)
3. Colby K., Hilf F., Weber S.: Artificial Paranoia. Artificial Intelligence, Vol. 2, pp. 1-25. (1971)
4. Zue V., Seneff S., Glass J., Polifroni J., Pao C., Hazen T.J., and Hetherington L.: JUPITER: A Telephone-Based Conversational Interface for Weather Information. IEEE Transactions on Speech and Audio Processing, Vol. 8, No. 1, pp. 85-96. (2000)
5. Bratt H., Dowding J., Hunicke Smith J.: The SRI Telephone-based ATIS System, in Proceedings of the ARPA Spoken Language System Technology Workshop, pp. 22-25. (1995)
6. Lau R., Flammia G., Pao C., Zue V.: WebGALAXY: Beyond Point and Click - A Conversational Interface to a Browser, in Proceedings of the 6th International WWW Conference, pp. 119-128. (1997)
7. The University of Texas at Austin (2004) SmartVoice, http://www.utexas.edu/its/smartvoice/
8. Sacks H., Schegloff E., Jefferson G.: A Simplest Systematics for the Organization of Turn-taking in Conversation. In: Schenkein J, (editor) Studies in the Organization of Conversational Interaction. Academic Press, New York. (1978)
9. Searle R.J.: Speech Acts: An essay in the Philosophy of Language. Cambridge University Press, Cambridge. (1969)
10. Allen F.J., Perrault R.C.: Analyzing Intention in Dialogues. Artificial Intelligence, 15(3):143–178. (1980)
11. Cohen P.R., Levesque H.J.: Speech Acts and the Recognition of Share Plans, in Proceedings of the Third Biennial Conference, pp. 263-271. Canadian Society for Computational Studies of Intelligence. (1980)
12. Clark H.H., Wilkes-Gibbs D.: Referring as a Collaborative Process. Cognition, 22:1–39. (1986)
13. Wallace R.S.: A.L.I.C.E. Artificial Intelligence Foundation. http://www.alicebot.org
14. Montero C.A.S., Araki K.: Information Acquisition Using Chat Environment for Question Answering. KES 2004, Lecture Notes in Artificial Intelligence (LNAI) 3214, pp. 131-138, 2004. Springer-Verlag Berlin Heidelberg 2004.
15. Montero C.A.S., Araki K.: Improvement of a Web-Based Question Answering System Using WordNed-Upgraded Chat. Joint Convention Record, The Hokkaido Chapters of The Institutes of Electrical and Information Engineers, pp. 282-283, IEEE Sapporo Section, Japan. (2004)
16. Montero C.A.S., Araki K.: Discovering Critically Self-Organized Chat. The Fourth IEEE International Workshop on Soft Computing as Transdisciplinary Science and Technology. To appear. (2005)
17. Ohsawa Y.: *KeyGraph*: Visualized Structure Among Event Clusters. In: Ohsawa Y., and Mcburney P. (eds) Chance Discovery. Springer, Berlin Heidelberg New York.