Journal of Universal Computer Science, vol. 13, no. 2 (2007), 177-185
submitted: 31/7/06, accepted: 15/1/07, appeared: 28/2/07 © *J.UCS*

Analysis of Conversation Quanta for Conversational Knowledge Circulation

Ken Saito
(Dept. of Intelligence Science and Technology,
Graduate School of Informatics, Kyoto University, Japan
saitoh@ii.ist.i.kyoto-u.ac.jp)

Hidekazu Kubota
(Dept. of Intelligence Science and Technology,
Graduate School of Informatics, Kyoto University, Japan
kubota@ii.ist.i.kyoto-u.ac.jp)

Yasuyuki Sumi
(Dept. of Intelligence Science and Technology,
Graduate School of Informatics, Kyoto University, Japan
sumi@i.kyoto-u.ac.jp)

Toyoaki Nishida
(Dept. of Intelligence Science and Technology,
Graduate School of Informatics, Kyoto University, Japan
nishida@i.kyoto-u.ac.jp)

Abstract: In this paper, we present a computational approach to understanding and augmenting the conversational knowledge process. We introduce the concept of the conversation quantization, a technique of approximating a continuous flow of conversation by a series of conversation quanta that represent points of the discourse. To investigate what the nature of conversation quanta is, we attempt to extract conversation quanta from two types of the meeting videos by hand. As a result, we have obtained some profitable suggestions about conversation quanta.
Keywords: conversation, knowledge acquisition, knowledge circulation, conversation quantization
Categories: H.3.1, H.3.3, H.5.1

1 Introduction

The conversation is our primary method to communicate with each other. A lot of useful knowledge occurs in the conversation in the real world, however it almost fades away without the support of intelligent media technologies. In this paper, we present a computational approach to understand and augment the conversational knowledge process that is a collective activity for knowledge creation, management, and application where conversational communications are used as a primary means of interaction among participating agents. The key idea is conversation quantization, a

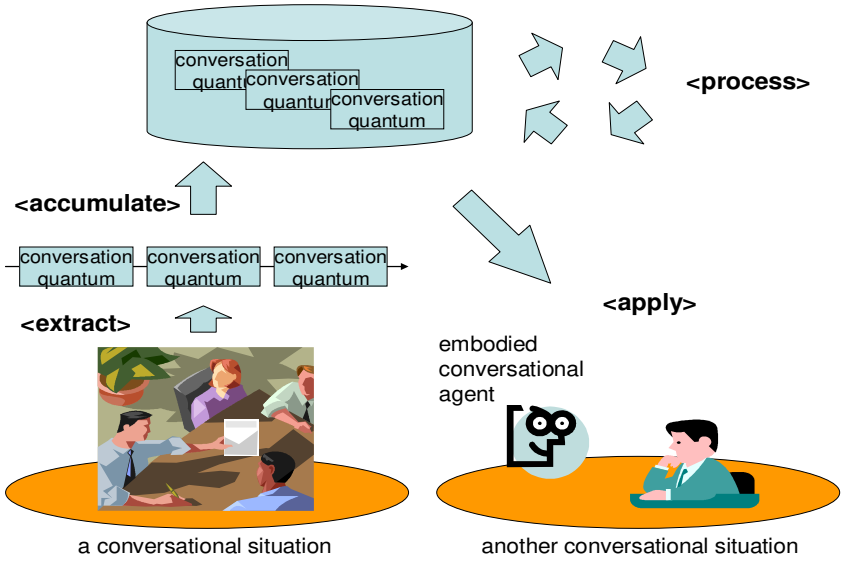

Figure 1: The concept of conversation quantization

technique of approximating a continuous flow of conversation by a series of conversation quanta that represent points of the discourse.

Previous works about conversation quantization surveyed by Nishida [Nishida, 05] include expansions of conversation quanta in knowledge circulation, embodied conversational agents, a virtual immersive environment, spatial content management and automated conversation capture. The former studies contribute a lot to knowledge circulation by using handcrafted conversation quanta. We are here concerned with the nature of conversation quanta from real world conversation on assumption that conversation quanta are utilized for a conversational agent that can talk on behalf of an actual person.

This paper describes the conceptual framework of conversation quantization and two analyses of conversation quanta. We have experimentally extracted conversation quanta by hand from practical situations to investigate their nature and reusability.

2 Conversation Quantization

Conversation quantization is a technique of articulating a continuous flow of conversation by a series of objects called conversation quanta each of which represents a point of the discourse. We define a conversation quantum to be an entity that contains a minimal amount of contextual information. In other words, each conversation quantum makes a minimal sense even though it may be presented in an inappropriate context. The granularity and size of conversation quanta essentially depend on the context and background knowledge of the observer.

Conceptually, the framework of conversation quantization consists of extraction, accumulation, processing, and application of conversation quanta (Figure 1). The extraction of conversation quanta results from identification and encoding of coherent segments of interactions in a conversational situation. The extracted conversation quanta are accumulated in a server, processed whenever necessary, and applied to other conversational situations. The processing of conversation quanta involves some interaction technologies that support the arrangement of the quanta. CSCW and computer visualization are examples of such technologies. The application of conversation quanta involves information retrieving, knowledge sharing and hands-on learning. In these situations, an embodied conversational agent is a good mediator because of its communicative abilities.

In this paper, we investigate the nature of the conversation quanta focusing on the virtualized-ego agent. The virtualized-ego is a conversable agent that functions as an alter-ego. A user can talk with a virtualized-ego of an actual person whenever and wherever he likes [Kubota, 00] [Kubota, 04]. A virtualized-ego is not a virtual character but a virtualized human that has autobiographical memory of an actual person. It is important that the aim of the virtualized-ego is the creation of a conversational agent that can talk about informal knowledge like personal experience. The virtualized-ego is expected to decrease the time-related, spatial and social constraint of the conversation.

3 Nature of Conversation Quanta

The implementation of conversation quantization depends on the data structure for representing conversation quanta. One could use plain video clips as representation, but the efficiency in retrieving and processing would be quite limited and a large cost would be require for retrieving, editing, and applying the conversation quanta. Alternatively, a deep semantic representation using logical formulas or case frames would not be ideal due to the expense and the limited capability of representing nonverbal information. A reasonable implementation appears to be to use annotated videos and images to represent a conversation quantum.

Firstly, we targeted a conversation using slides. We simulated conversation quantization by hand to investigate the nature of conversation quanta in a real situation. We gave shape to the concept of conversation quanta as follows:

1. Setting up a practical conversational situation
2. Capturing conversation by video camera
3. Extracting conversation quanta from the video stream by hand

We obtained the two types of video. One is the video in which two members of authors participated (Video A). The other is the video in which 4 people (including one of authors) participated (Video B). In the following, we will empirically analyze the nature of conversation quanta by extracting the conversation quanta from these videos, and creating new conversational contents using these conversation quanta.

3.1 Analysis 1 (Video A)

Video A consists of 3 meetings between two of the authors, a master course student (subject A) and a postdoctoral fellow (subject B). The subjects knew their conversations were later utilized as conversation quanta. Each meeting was held in a different place and at different times. Each of them talked using PowerPoint slides on mobile PC (with a web camera and a microphone) to capture his voices, faces and context. As a result of these meetings, we obtain three and a half hours video of subject A and subject B. In their conversation, presentational style and discussion style are half-and-half, and their topics were about conversation quantization – its history, problems, approaches, systems and so on.

Here, we suppose that we can create virtualized-ego by quantizing the video and mixing the quanta. To confirm this supposition, we made an experiment to extract conversation quanta arbitrarily from these videos and create a new presentation video in which the virtualized-ego of the participants talk about their study.

We have proposed the first approximate model for extracting conversation quanta. This first approximation aims to divide the conversation into units larger than sentences. We have already studied the virtualized-ego system by using a sentence as a conversational unit [Kubota, 00]. We extracted one important sentence from each post in a mailing list. We ensured that the virtualized-ego could generate a new conversation by using such sentences. We expect that we can extract conversational blocks that are more unique to a particular conversation by using a unit larger than a sentence; however, the larger unit would reduce the reusability of units because it depends more on the context.

The conversation is divided into relatively larger units to investigate the reasonable granularity of the conversation quanta in the first approximate model. First, the video is divided at the point of transition of the slides (Figure 2 “Division 1”) because we assume that the unit of topical contents in the conversation decides the rough unit of the conversation. The speeches seem to be almost coherent and united in a slide.

The second division point is the start of a dialogue (Figure 2 “Division 2”). We particularly intend to distinguish between a set of dialogues and a single speech unit. Our aim is for the conversation quantum to be a more holistic unit than a reductionistic unit like a morpheme, word, or sentence. A set of dialogues is expected to include a unique conversational scene, for example, a question and answer with a rich context, comic dialogue, or any other inter- and intra-personal synchronization of speeches and gestures.

We suppose that a video clip from the start of a dialogue to the end of a slide is a conversation quantum representing the dialogue style. A conversation quantum is stored into the archive of a speaker as a general rule. Only when a quantum is a dialogue style, it is stored into the archive of every speaker.

Using this model, we extracted conversation quanta from Video A. We have got 41 quanta for subject A and 66 quanta for subject B. Table 1 shows the number of conversation quanta in the archive of each subject. “Single speech” means the quanta include only the subject, and “Dialogue” means the quanta include two subjects. The number of dialogue quanta of subject A is same as that of subject B of course.

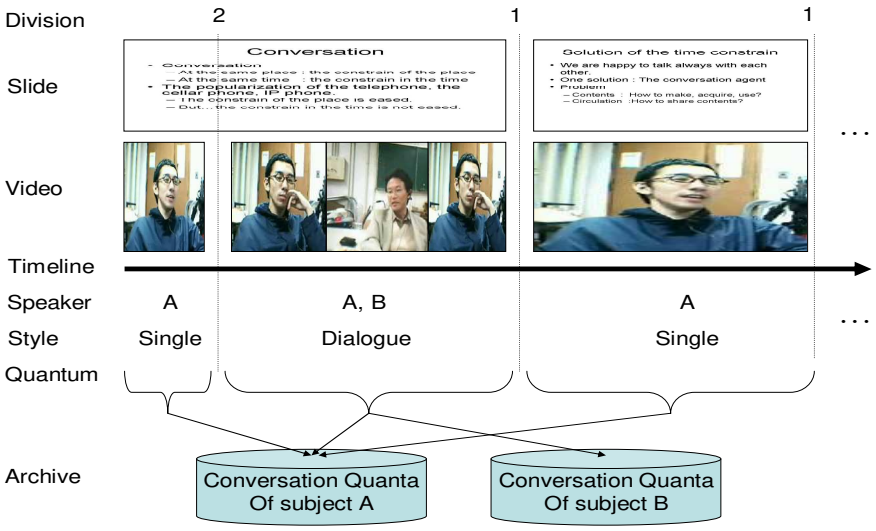

Figure 2: The first approximate model for extracting conversation quanta

Subject	Single speech (total time)	Dialogue (total time)
A	24 quanta(16 minutes)	17 quanta(21 minutes)
B	49 quanta(35 minutes)	17 quanta(21 minutes)

Table 1: Conversation quanta from three and a half hours videos

We have simulated a virtualized-ego system based on these conversation quanta. We arranged conversation quanta of subject A on the assumption that the system talks with a user on behalf of subject A. Figure 3 shows the overview of our simulation. Firstly, a user comes in front of a system screen where the face of subject A is displayed. Here, the system begins to talk on behalf of subject A when the user asks for his interest ("Greeting"). The system talks by arranging past conversation quanta that are related to the interest of the user ("Quantum 1" and "Quantum2"). While the system is talking, the user can ask any questions ("Question"). Then the system can answer the question by searching an answering conversation quantum ("Quantum3"), and keep on talking ("Quantum4").

We could not forecast whether the conversation quanta are reusable or not because the granularity of the conversation quanta adopted here is larger than that proposed by previous studies. We have obtained some profitable suggestions about conversation quanta from the simulation and analysis above. First, conversation quanta which depended on context are reusable in the situation where a user is familiar with the original situations of the conversation quanta. In Figure 3, Quantum1, Quantum2, Quantum4 were acquired in different rooms. Thus, we can make new conversational content from the past conversation quanta that were got in

different situations. When we searched conversation quanta that are suitable for a user, thinking about the background knowledge of the user was very important. The conversation in Figure 3 left fragment, however, it could be complemented by the user because he is a colleague of the speakers on the screen. Second, a dialogue style quantum which contains speedy Q&A pair, jokes is interesting. They have good points of conversation such as conversational rhythms, presence, and dynamics. In addition, on the viewpoint of virtualized-ego, a quantum which contains individual experiences and know-how is interesting too.

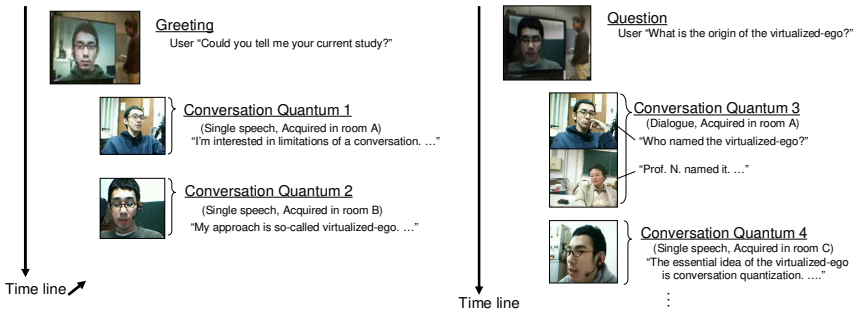

Figure 3: A simulation of a conversational video content using conversation quanta

3.2 Analysis 2 (Video B)

One of authors had the meeting in which four people participated (subject C, D, E, F). In this meeting, a doctor's course student of our laboratory (subject C) presented his studies with PowerPoint slides. This meeting was captured with a digital video, and then we have got one and a quarter hours video of the meeting. In their conversation, ratio of presentational style to discussion style is about 1:5. The subjects were not yet informed that their conversations were being utilized as conversation quanta during the meetings; however, they knew their conversations were being recorded as the minutes of the proceedings of a meeting.

In addition to the first approximate model, we adopted the third division point for dividing detailedly. The third division point (Figure 4 "Division 3") is the part in which no one speaks. In other words, this is silence from the end to the start of speech. We have proposed the second approximate model of extracting conversation quanta (Figure 4). Namely we divided videos by the following policy:

1. Divide at the point of the transition of slides (Figure 4 "Division 1")
2. Divide at the point of a start of dialogue (Figure 4 "Division 2")
3. Divide at the part in which no one speaks (Figure 4 "Division 3")

A conversation quantum is stored in the same way as in analysis 1. The quantum is stored into the archive of the speaker as a general rule. Only when the quantum is a discussion style, it is stored into the archive of every speaker.

On the third division point, we divided the part in which silence is more than 4 seconds. Table 2 shows the number of conversation quanta and quanta's average time.

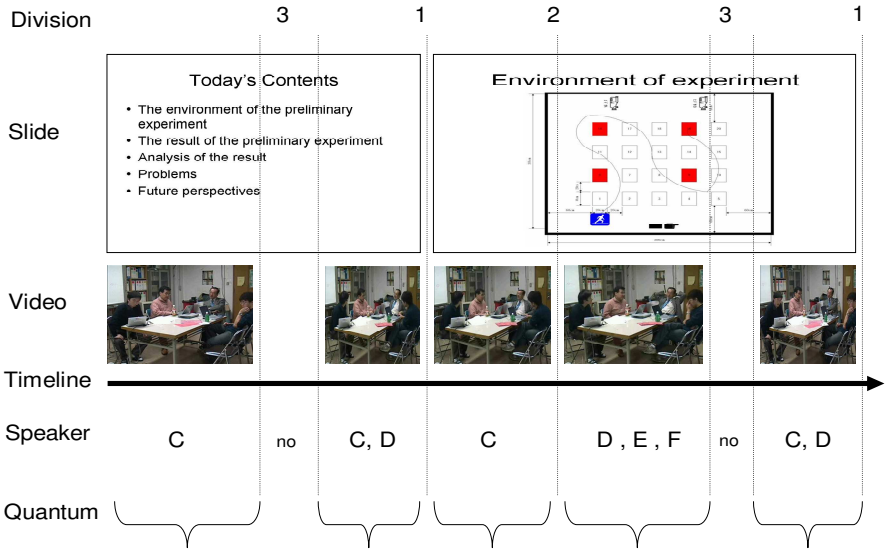

Figure 4: The second approximate model for extracting conversation quanta

Threshold of silence	Division	Silence (average time)	Quantum (average time)
More than 4 sec	111	47 (6.9 sec)	65(66 sec)

Table 2: Conversation quanta from Video B

As a result of the analysis, we found that the Video B includes two types of conversation quanta. One is the presentation in which subject C explained his slide. The other is the discussion by all subjects.

The presentation style quantum arises almost at the head of a slide at 71 %(11/14). The presentation style quantum does not very depend on the context in compare with the discussion style quantum. The presentation quantum gives the context to the discussion style quantum in right after because the presentation shows what to discuss. Some discussion style quanta require the presentation style quantum to be understood. Although it is difficult to understand the discussion style quantum when a user isn't given context, it is important and interesting. Because the discussion style quantum contains individual experience and know-how and joke.

4 Discussion

In the previous sections, we have discussed the nature of the conversation quanta on the viewpoint of virtualized-ego. As a result of two analyses, we obtain some suggestions about conversation quanta.

The conversation quanta which contain following contents, speedy Q&A pair, jokes, and individual experiences and know-how, are important and interesting in reusing conversation quanta. The virtualized-ego which is generated from conversation quanta that depend on context would be understandable for the user who shares the context. For example, a community member could easily complement fragmentized conversation of the members in the same community. This supposition was confirmed by Hirata [Hirata, 00] in only text conversation fragments. In this paper, we obtain the result above mentioned about video conversation fragments.

There are interesting works about conversation quantization. Conversation quanta can be extracted from the real world conversation by expanding the ubiquitous sensor room that is proposed by Sumi et al [Sumi, 04]. For spatio-temporal management of conversation quanta, Kubota is developing a system called the Sustainable Knowledge Globe [Kubota, 05]. Virtualized-ego [Kubota, 04] that can talk on behalf of an actual person (as mentioned in Section 3.1) is a good utilization of conversation quanta. Video and sound collage system of one's experience [Sumi, 04] would also be another good application.

The study of topic extraction from conversation has been growing. A deep understanding of discourse structure is indispensable to extract essence of conversation automatically. Shibata et al. [Shibata, 04] study the discourse analysis of the cooking program video by using linguistic and visual information. Our research object doesn't like the cooking program which is controlled by a video director but the casual conversation in any situations, so it is very difficult to understand the discourse structure automatically, especially the correspondence structure when there is the omission in the indication word in the video. We aim to make the conversation quantization feasible by supporting humans to understand topics in video in the loop of the conversation quantization (extraction, accumulate, apply, process) such as Q&A system of EgoChat [Kubota, 04].

The annotation of conversation quanta is the text that is mainly used for retrieving and combining quanta by using natural language processing. The annotation can be the sentences extracted from the slides, the text generated by speech recognition system, and the note annotated by people. We did not use the annotations in the two analyses because they were manually performed simulations. We plan to expand the way in which the virtualized-ego in the EgoChat system uses the annotation to generate conversation [Kubota, 04]. Here, the virtualized-ego answers the user's question by showing the slide with a synthesized voice that is retrieved by searching the text annotation. Future system will show the video, the slide, and the human voice instead of the slide.

There are many interesting work left for the future research. Among others, we need to build a more detailed and elegant theory of conversation quantization. A more sophisticated theory of conversation quanta will permit us to better design the representation and basic operation for conversation quanta. It may well enable us to predict the cost and effect of building a conversation system based on conversation quantization.

5 Conclusion

In this paper, we have presented a computational approach to understanding and augmenting the conversational knowledge process. We have introduced the notion of conversation quantization, a technique of approximating a continuous flow of conversation by a series of conversation quanta that represent points of the discourse. We obtained profitable suggestion about the nature of conversation quanta by extracting them from practical situations by hand.

References

[Hirata, 00] Hirata T., Kubota H., Nishida T.: Talking virtualized egos for dynamic knowledge interaction. In Nishida T., editor, Dynamic Knowledge Interaction, chapter 6, pages 183-222. CRC press, 2000.

[Kubota, 00] Kubota H., Nishida T., Koda T.: Exchanging Tacit Community Knowledge by Talking-virtualized-egos. In Proceedings of Fourth International Conference on AUTONOMOUS AGENTS (Agents 2000, Barcelona, Catalonia, Spain. June3 -June 7), pp.285-292, 2000.

[Kubota, 04] Kubota H., Hur J., Nishida T.: Agent-based Content Management System. In Proceedings of the 3rd Workshop on Social Intelligence Design (SID 2004), CTIT Proceedings, pp.77-84, 2004.

[Kubota, 05] Kubota H., Sumi Y., Nishida T. : Sustainable Knowledge Globe: A System for Supporting Content-oriented Conversation, in Proceedings of AISB 2005 Symposium Conversational Informatics for Supporting Social Intelligence & Interaction, pp.80-86, 2005.

[Nishida, 05] Nishida T.: Conversation Quantization for Conversational Knowledge Process, Special Invited Talk, S. Bhalla (Ed.): DNIS 2005, LNCS 3433, Springer, pp. 15-33, 2005.

[Shibata, 04] Shibata T., Tachiki M., Kawahara D., Okamoto M., Kurohashi S., Nishida T.: Structural Analysis of Instruction Utterances using Linguistic and Visual Information, In Proceedings of the 8th International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES2004), Wellington, New Zealand , pp. 393-400, 2004.

[Sumi, 04] Sumi Y., Mase K., Mueller C., Iwasawa S., Ito S., Takahashi M., Kumagai K., Otaka Y.: Collage of video and Sound for Raising the Awareness of Situated Conversations. In Proceedings of International Workshop on Intelligent Media Technology for Communicative Intelligence (IMTCI2004), pp.167-172, 2004.