

会話量子を用いたミーティングのコンテンツ化支援

The Support of Creating Contents Using Conversation Quanta

齊藤 憲*¹
Ken Saito

久保田 秀和*¹
Hidekazu Kubota

角 康之*^{1*2}
Yasuyuki Sumi

西田 豊明*¹
Toyoaki Nishida

*¹ 京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University

*² ATRメディア情報科学研究所
ATR Media Information Science Laboratories

In this paper, we present a computational approach to understand and augment the conversational knowledge process. We introduce the concept of the conversation quantization, a technique of approximating a continuous flow of conversation by 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 the meeting videos by hand. As a result, we have obtained some profitable suggestions about conversation quantization and conversation quanta.

1. はじめに

会話は我々にとって最も一般的なコミュニケーション手段である。我々の知的活動の多くは会話を通じて行われるものであり、会話は知識共有・知識創造の場となっている。しかし、会話の場において生まれた知識は記録を残さない限り失われやすいものである。このためカジュアルな立ち話から生まれるような知識は再利用が困難である。また、議事録やメモのようなテキストベースの記録では会話の非言語的な側面を十分に残すことが出来ない。そこで実世界の会話を容易にコンテンツとして記録することが出来れば、知識の蓄積および再利用が加速されると考えられる。

西田[Nishida 05]は、実世界の会話をコンテンツ化し、流通可能とするシステムの枠組みとして会話量子化を提案している。会話量子化とは、連続した会話の流れから離散的な会話の粒を切り出して持続的に蓄積することにより、元とは異なる会話の場において元の会話のエッセンスを再利用可能とする枠組みである。この会話の粒のことを会話量子と呼ぶ。

本論文では、実世界の会話映像から会話量子を抽出し、別の場で再利用するまでのプロセスを手動でシミュレートすることにより、映像を用いた会話量子の持つ性質について議論する。会話量子の再利用例としては分身エージェント[Kubota 00]を考える。分身エージェントとは実在する人の代理として会話可能なエージェントのことを言う。以下、2章では会話量子化の概要について述べ、3章では2名によるミーティング映像を元にした会話量子獲得および分身エージェント構築について分析し、4章で映像を用いた会話量子の性質について議論する。5章はまとめである。

2. 会話量子化の概略

会話量子化とは、連続した会話の流れを会話量子と呼ばれる離散的な粒の流れとして近似することにより、会話を部品化する技術である。会話量子化の全体像を図1に示す。会話量子はある会話の場から抽出されたひとまとまりの会話のエッセンスである。会話量子は持続的に蓄積され、会話システムを構築する部品として再利用される。会話システムは会話量子を元とは異

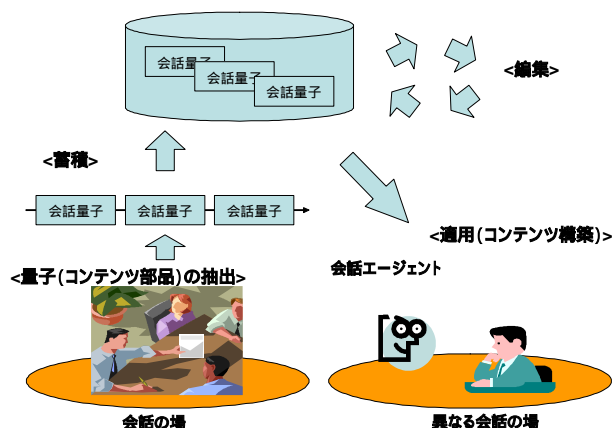


図1: 会話量子化の全体像

なる並びや組み合わせを用いて再構成することによって多様な会話を生成する。

会話量子化の実装は会話量子を構成するデータ構造に依存する。一般に、厳密な知識表現に従うデータを実世界から多数獲得することは困難である。会話量子化の狙いは実世界の会話を通じて生ずる大量の言語的・非言語的知識の流通支援であるため、会話量子を意味および談話構造の深い理解に基づく単位であると捉えることは目的にそぐわない。一方、会話量子が言語的な手掛かりを全く伴わない映像クリップであるとする、検索、編集、適用には多大なコストを要するものと考えられる。そこで、ここでは自然言語文によって注釈のつけられた映像クリップを用いて会話量子を表現するものとする。映像クリップは実世界会話における雑多な非言語情報をカプセル化可能であり、また今日の自然言語検索技術の発展を踏まえ、自然言語文は検索や編集の手掛かりを得る上では十分なものと考えられる。

3. 会話量子の分析

本論文では、実世界で生じた会話から抽出される会話量子の性質を分析するため、以下の手順で会話量子化をシミュレートした。

連絡先: 齊藤憲(Ken Saito), 京都大学大学院情報学研究科,
京都市左京区吉田本町工学部 10 号館 223 号室,
saitoh@ii.ist.i.kyoto-u.ac.jp

1. スライドを用いてミーティングをする
2. ビデオカメラを用いて会話をキャプチャする
3. 手作業でビデオ映像から会話量子を抽出する
4. 1~3を異なる場所, 時間で繰り返す
5. 獲得した会話量子を用いて, 新しい会話コンテンツを手作業で制作する

上記の手順に従い, まず, 筆者のうちの 2 名(研究室の修士 2 回生の学生(被験者 A), ポスドク研究員(被験者 B))によるミーティングを 3 回行い, ミーティング映像を 3 つ獲得した(計約 3 時間 30 分). 各々のミーティングは異なる場所, 異なる時間に行われ, PowerPoint スライドを利用し, 基本的にプレゼンテーション形式で進められた. その際, ノート PC を用いて各々の映像, 声を記録した. ミーティングの内容は 3 つ全てが会話量子の歴史, 問題点, 方法論, システムなどで, プレゼンテーションとディスカッションの割合が 1:1 ほどであった.

我々は会話量子の第 1 次近似モデルを図 2 のように考え, 次に示すポリシーに基づいて映像クリップを分割した.

1. スライドが変わった時点で分割
2. 会話スタイルが変わった時点分割

分割の第 1 のポイントは, スライドの変わり目, 第 2 のポイントは会話スタイルが変わった時点である. 手作業でビデオを分割していく試行の中で, 話者はスライドの内容とほぼ同等の内容を話していることと, 会話の始まりから, 該当スライドの終わりまでの映像クリップが会話スタイルの会話量子であるという示唆を得たためである. 会話量子の蓄積は基本的には重複を許さず, 発話者本人のアーカイブへ蓄積するものとするが, 会話スタイルの映像についてはやりとりをばらばらにするよりもまとめて扱う方が合理的であるため, やりとりを行ったすべての発話者のアーカイブへ重複して蓄積するものとした.

このモデルを用い, およそ 3 時間 30 分のミーティング映像から会話量子の抽出を試みたところ, 被験者 A の量子が 41 個, 被験者 B の量子が 66 個獲得できた. 表 1 にそれぞれの話者のアーカイブとして獲得された会話量子数を示す. “Single speech”は量子が話者を 1 名だけ持つ場合で, “Dialogue”は量子が話者を複数持つ場合である. 被験者 A の Dialogue スタイルの量子数は, 当然被験者 B の Dialogue スタイルの量子数と同数である.

これらの会話量子を元にして, 手作業で分身エージェントシステムをシミュレートした. 分身エージェントとは実在する人の代理として会話可能なエージェントのことを言い, いつでもどこでもその人物のように話すことが出来る. 分身エージェントは仮想的なキャラクターではなく, 実在の人物の代理として本人の記憶や経験について会話可能なキャラクターであり, 本人と時間的, 空間的, あるいは社会的な距離のある人物とを会話可能とすることにより, コミュニティにおける知識循環を促進する. 今回作成したシステムは, 被験者 A の会話量子を用いて, 被験者 A 本人の代わりとなって会話する分身エージェントである(図 3). まず, ユーザが被験者 A の顔が映し出されているシステムのスクリーンの前に来る. ここで, ユーザが何かを尋ねた時, システムは被験者 A の代理として話し始める(“Greeting”). このシステムは過去に獲得された被験者 A の会話量子の中から, ユーザの興味に関連した量子を組み合わせることでストーリーを生成する(“会話量子 1” “会話量子 2”). システムが話している途中で, ユーザは質問することが出来る(“Question”). ユーザが質問すると, システムは答えとなるであろう会話量子を探し, 提示

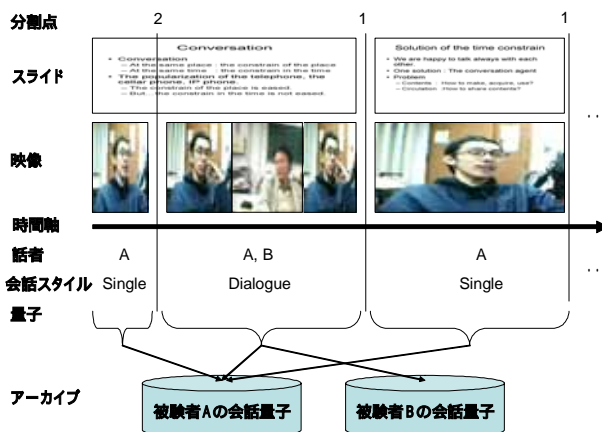


図 2: 会話量子抽出 第 1 次近似モデル



図 3: 会話量子を用いた分身エージェントのシミュレーション

表 1: Video A から獲得された会話量子数

被験者	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)

することによって質問に答える(“会話量子 3”). 回答が終了すると, 元の文脈に戻る(“会話量子 4”). ここで, ユーザは分身エージェントの母体となった人物と同僚関係であるため, コンテンツを理解可能であった.

分身エージェントのシミュレーションとその分析から会話量子についてのいくつかの有益な示唆を得た。1 つ目は文脈依存している会話量子であっても、会話量子が有している元の文脈をよく知っているユーザであれば理解可能であるということである。図 3 の“会話量子 1”、“会話量子 2”、“会話量子 4”はいずれも異なる時間、異なる場所で獲得したものである。すなわち、過去に異なったシチュエーションから得られた会話量子を用いて、新しい会話コンテンツ(分身エージェント)を作成できるということである。ユーザが求めている情報を有した会話量子を検索する時、ユーザの背景知識を考慮することは重要である。図 3 の会話は登場人物や背景に関する情報が足りないため、断片化されたままだが、ユーザがユーザとスクリーン上に映し出されている分身エージェントの母体となる実在の人間の同僚関係であるため理解できた。2 つ目は、会話スタイルの会話量子はテンポのよい Q&A、冗談などを含んでいるため、面白いということである。それらは会話のリズムや雰囲気、ダイナミクスといった会話の良い点を有している。さらに、分身エージェントの視点から見ると、個人的な経験談やノウハウを含んだ会話量子も再利用の際に面白いという示唆を得た。

4. 議論

我々は分身エージェントに焦点を当て、会話量子の性質を議論してきた。分析により、会話量子の性質についてのいくつかの有益な示唆を得た。

テンポのよい、スピーディーな Q&A、冗談、個人的な経験談やノウハウを含んでいるような会話量子は再利用する際に重要であり、面白い。文脈依存の残る会話量子から生成される分身エージェントは文脈を共有しているユーザにとっては簡単ではないが理解可能である。たとえば、あるコミュニティメンバーは同じコミュニティメンバーの会話ならば、その会話が断片的であったとしても理解できるだろう。これは、テキストベースの会話では平田[Hirata 00]が、音声ベースの会話では久保田[Kubota 00]がそれぞれ確認している。本論文では、映像断片には背景や服装などが変わるといった映像特有の空間的文脈の断絶があり、量子化は困難であると思われたが、上記のような結果を映像ベースの会話でも確認したといえる。背景や服装の変化はさほど気にならないが、図 3 の“会話量子 3”と“会話量子 4”のように同一話者の顔の向きが急に変わると、ユーザが混乱し、コンテンツを理解しにくくなるという知見を得た。

5. まとめ

本論文では、会話量子化のアプローチの全体像を提案すると共に、分身エージェントに焦点を当て、ミーティング映像から手作業で会話量子を抽出し、文脈依存の残る会話量子を用いて分身エージェントのためのコンテンツを作る事で、会話量子の性質について調査、分析した。その結果、文脈依存の残る会話量子から生成される分身エージェントは文脈を共有しているユーザにとって理解可能性は高い、といったいくつかの有益な知見を得ることができた。また、本論文の分析は会話参加者自身によってなされたため、会話量子の評価は主観的であった。そこで、会話量子化を客観視するために会話参加者以外による抽出・評価を現在行っている。

今後は、会話量子抽出の半自動化を目指し、会話量子の概念をより明確化していくとともに、コンテンツ化支援を含め、会話量子化のアプローチ全体の概念を詳細に考察し、実装していく。

参考文献

- [Nishida 05] Toyooki Nishida: Conversation Quantization for Conversational Knowledge Process, Special Invited Talk, S. Bhalla (Ed.): DNIS 2005, LNCS 3433, Springer, pp. 15-33, 2005 .
- [Kubota 00] Hidekazu Kubota, Toyooki Nishida, Tomoko Koda: Exchanging Tacit Community Knowledge by Talking-virtualized-egos. Fourth International Conference on AUTONOMOUS AGENTS (Agents 2000, Barcelona, Catalonia, Spain. June3 -June 7), pp.285-292, 2000 .
- [Hirata 00] Takashi Hirata, Hidekazu Kubota, and Toyooki Nishida: Talking virtualized egos for dynamic knowledge interaction. In Toyooki Nishida, editor, Dynamic Knowledge Interaction, chapter 6, pages 183-222. CRC press, 2000 .