

なぜホワイトボードを使うと良い議論ができるのか？ホワイトボードを使用した共同作業における対話の流れの分析

藤枝佑喜 東中竜一郎

名古屋大学大学院情報学研究科

fujieda.yuki.x6@s.mail.nagoya-u.ac.jp

higashinaka@i.nagoya-u.ac.jp

概要

近年、創造的な共同作業を行う対話システムの研究が進んでいるが、その多くは言語中心のやり取りに留まり、言語以外のモダリティの考慮が限定的である。我々は先行研究において、ホワイトボードに着目し、ホワイトボードが活用可能な対話システムの実現のため、ホワイトボードを用いた対話データを収集した。しかし、ホワイトボードの使用が対話の流れに与える影響は明らかにされていない。そこで、本研究では発話の意味的類似度と描画位置の物理的距離を用いて対話の流れを定量化した。定量化およびその分析の結果、ホワイトボードは話題の一貫性維持に寄与し、多様な意見を整理する活用方法が対話の満足度向上に繋がることが分かった。

1 はじめに

近年、人工知能技術の発展に伴い、創造的な共同作業を行う対話システムの研究が盛んに行われている。[1, 2, 3] しかし、現在の対話システムの多くは、ユーザとの共同作業における相互理解の手段が言語中心となっており、言語以外のモダリティの考慮が限定的である [4, 5]。

一方、人間同士が創造的なアイデアを交換し、共同作業を行う際は、必ずしも言語のみを用いているわけではない。むしろ、状況や内容に応じて多様な表現手段を組み合わせながらコミュニケーションを行っている [6, 7]。例えば、ホワイトボードを用いた議論の場面では、参加者は文字による説明に加え、図形、矢印、記号、イラストなど自由な表現手段を活用している。対話システムにおいても、同様の多様で自由な表現手段を活用できるようになれば、人間との創造的な共同作業をより効果的かつ自然に行えると考えられる [8, 9]。

我々は、ホワイトボードを用いた対話に着目し、この領域における対話システムの実現を目指している。そして、これまでそのための基礎データとして、映像構成考案タスクを設定し、ホワイトボードを使用した共同作業の対話データを収集してきた。また、比較のため、同タスクでホワイトボードを使用しない条件でも対話データを収集した。さらに、ホワイトボードを使用した対話についての知見を得るため、発話と描画の時系列的な関係についてクラスタリングを行い、発話と描画のパターンが対話の満足度に与える影響の分析などを行った。

しかし、これらの分析では、ホワイトボードの使用が、対話の流れや構造そのものにどのような影響をもたらしているかまでは明らかにできていないという問題があった。そこで本研究では、対話の文脈的・意味的な構造を定量的に評価するため、対話データを一定の時間でセグメンテーションし、各セグメント間の発話類似度とホワイトボード上の描画距離を行列化する。この行列の特徴を調査することで、隣接する文脈の連続性だけでなく、離れた時点間の関係性も含めた話題の推移や一貫性といった対話全体の流れを分析する。

2 対話データの概要

我々は、先行研究において、ホワイトボードを活用した共同作業が可能な対話システムの実現に向け、映像構成考案タスクを設定し、対話データを収集した [10]。

映像構成考案タスクでは、「愛」や「自由」といった抽象的な単語をテーマに、対話を行いながらそのテーマを表現した短い映像の構成を考案する。実験は Zoom を用いたオンライン形式で実施し、ホワイトボードを使用可能な条件（以下、ホワイトボードあり対話）に加え、使用不可能な条件（以下、ホワ

表 1: 作業後アンケートの質問項目

Q1: 自分是对話の内容に満足した
Q2: 対話相手は対話の内容に満足した
Q3: 抽象的な単語について具体的なイメージを共有できた
Q4: 抽象的な単語が表す感覚を共有できた
Q5: 抽象的な単語が表す社会的な状況を共有できた
Q6: 抽象的な単語が表す心的状態や感情を共有できた
Q7: 具体的なイメージの共有にホワイトボードは役に立った

イトボードなし対話)の2条件を設定した。

本対話データには、クラウドワーカによって実施された計100対話(各条件50対話)が含まれている。一对話の制限時間は15分であった。収集データには、音声、映像、書き起こしに加え、作業後のアンケート(7項目の評価および考案した映像構成の100文字程度の記述)、考案した映像構成の記述に対する第三者評価が含まれる。作業後アンケートの質問項目は表1に示す。

3 発話間の類似度による対話の流れの分析

前述の対話データを用い、話題の推移や一貫性といった対話全体の流れを定量的に評価するために、対話のセグメンテーションおよびセグメント間の発話の類似度の測定を行い、類似度行列を作成した。そして、その結果をクラスタリングし、分析を行った。以下、その具体的な処理について述べる。

3.1 類似度行列の作成

以下の手順で対話のセグメンテーションおよびセグメント間の発話の意味的類似度の測定を行った。

1. 各対話データの開始から終了までの総時間を基準として、 k 個($k=5, 10, 15$)の等間隔のセグメントに分割する。
2. 各セグメントに含まれる発話を結合し、Sentence-BERT [11]を用いて埋め込みを得る。
3. 埋め込み間のコサイン類似度を算出し、 $k \times k$ の類似度行列を作成する。

さらに、対話の流れにおける局所的な話題の意味的連続性と対話全体での話題の拡散度合いを測定するため、隣接セグメント間の平均類似度および非隣接セグメント間の平均類似度を算出した。これにより、対話の主題が一貫しているか、あるいは話題が多岐にわたるかを定量的に捉えることができる。

結果を表2および表3に示す。全体的な傾向とし

表 2: 隣接セグメント間類似度の比較。(太字は各 k の値で類似度が高い方を示す。)

	ホワイトボードなし		ホワイトボードあり	
	平均	標準偏差	平均	標準偏差
$k=5$	0.47	0.093	0.51	0.090
$k=10$	0.49	0.064	0.53	0.071
$k=15$	0.50	0.060	0.54	0.058

表 3: 非隣接セグメント間類似度の比較。(太字は各 k の値で類似度が高い方を示す。)

	ホワイトボードなし		ホワイトボードあり	
	平均	標準偏差	平均	標準偏差
$k=5$	0.40	0.084	0.48	0.101
$k=10$	0.44	0.064	0.48	0.072
$k=15$	0.44	0.055	0.47	0.064

て、ホワイトボードあり対話の方がホワイトボードなし対話よりも高い類似度を示していることが確認できる。また、ホワイトボードあり対話は、 k に関わらず安定して高い類似度を維持した。これは、局所的な議論においても対話全体の主題との整合性が保たれていることを示唆している。

3.2 クラスタリング

各条件について作成した対話の 10×10 の類似度行列50個ごとに対してクラスタリングを行うことで、各対話条件での対話の流れのパターンの分析を行った。さらに、どのような対話の流れが話者の高い満足度につながるのかを明らかにするため、各クラスタに分類された対話における作業後アンケートの平均スコアを算出した。

具体的な手順として、まず各条件の対話における 10×10 の類似度行列を直列化し、100個の要素を含むベクトルとして扱い、Ward法による階層的クラスタリングを実施した [12]。クラスタの決定に際しては、分析に十分なサンプル数を確保するため、分割後の要素数が閾値(10件)を下回らないように制限を設けた。最終的に、得られた各クラスタと作業後アンケート結果を紐づけ、クラスタごとの対話の流れと対話の主観評価との関係性を分析した。

クラスタリング結果を図1および図2に示す。クラスタリングの結果、クラスタ数が3つずつとなった。また、各クラスタごとの対話の作業後アンケートの平均スコアを表4に示す。

各クラスタがどのような対話の流れになっていると考えられるかを説明する。ここでは、ホワイトボードなし対話のクラスタを $V_1 \sim V_3$ でホワイト

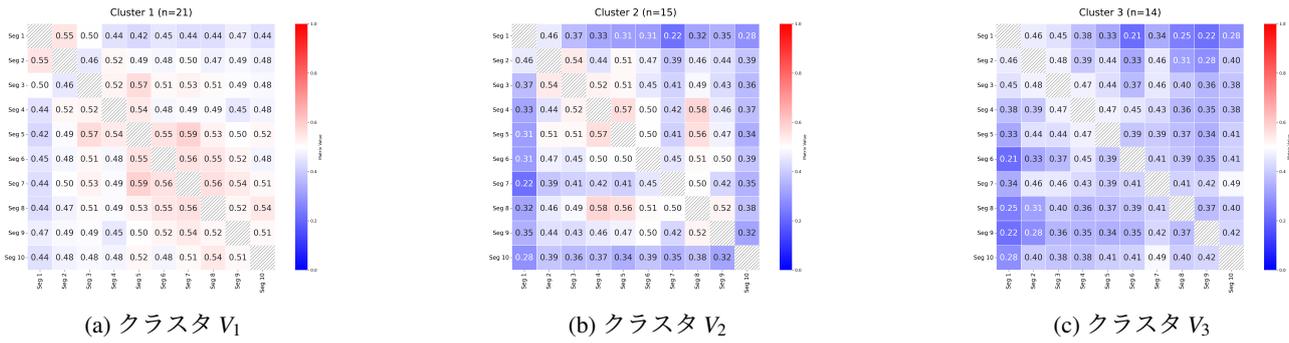


図 1: ホワイトボードなし対話のクラスタリング結果。(赤が濃いほど類似度が高く、青が濃いほど低い.)

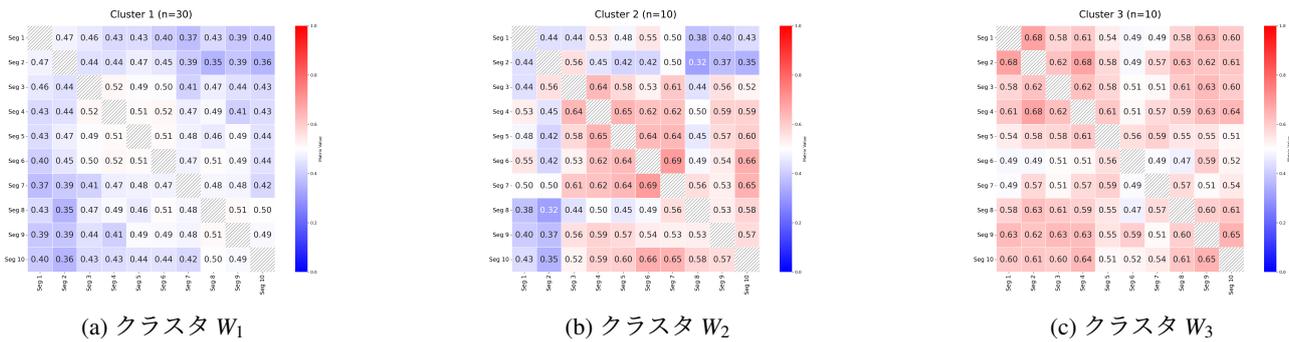


図 2: ホワイトボードあり対話のクラスタリング結果。(赤が濃いほど類似度が高く、青が濃いほど低い.)

ホワイトボードあり対話のクラスタを $W_1 \sim W_3$ で表す。

- V_1 ホワイトボードなし対話の中では、他のクラスタよりも類似度が低い箇所が少ないため、比較的高い類似度を示し、対話の一貫性が高い。
- V_2 隣接セグメント間の類似度が高く、直前の話題との意味的連続性を維持しつつ進行している。一方、非隣接セグメント間の類似度は低く、過去の話題への振り返りは少ないと考えられる。
- V_3 全体的に他のクラスタより類似度が低く、話題の推移が頻繁である。一つの主題に留まらず、多様な意見やアイデアが提示されている。
- W_1 全体的な類似度は低いものの、図 2(a) の 3 列目以降は隣接セグメント間の類似度が上昇する傾向にある。多様な意見を含みつつも、直前の話題との意味的連続性を保って対話が進行している。
- W_2 図 2(b) の 3 列目以降、隣接・非隣接を問わず高い類似度を示す。導入を経て本題に入った後は、一貫して特定の主題に集中し、議論を深めていると推測される。
- W_3 全体的に類似度が高いため、一貫した話題について対話していると考えられる。また、その上で、図 2(c) の右上部分である序盤と終盤のセグメント間に高い類似度が見られる。これは終盤

において初期の話題に回帰し、議論の総括やまとめを行っていることを示唆する。

表 4 は各クラスタに含まれる作業後アンケートの平均スコアである。ホワイトボードなし対話では、話題の意味的連続性を維持しつつ進行する V_2 が、Q1 (自身満足度) や Q2 (相手満足度) などでも最も高いスコアを示した。音声のみの対話においては、話題を頻繁に転換するよりも、着実に議論を積み上げることが満足度に寄与すると考えられる。

ホワイトボードあり対話では、 W_1 が Q1, Q2, および Q7 (ホワイトボード有用度) において最高値を示した。 W_1 は多様な意見を含みつつ話題の意味的連続性を維持するパターンであり、議論の整理にホワイトボードが有効に機能したと推測される。特定の主題を深掘りする W_2 は、Q3 (イメージ共有度) ~ Q6 (心的状態共有度) において最高値を示した。特定の主題を深掘りする際は、ホワイトボードは理解の共有に有効であると考えられる。その一方、Q7 においては最低値を示しており、これは深掘りの場面では議論の整理の必要性が相対的に低いためだと考えられる。 W_3 は全体的に類似度が高いものの、Q1 や Q2 において最低値を示した。これは、話題が展開せず停滞したことが、満足度の低下に繋がったことを示唆する。

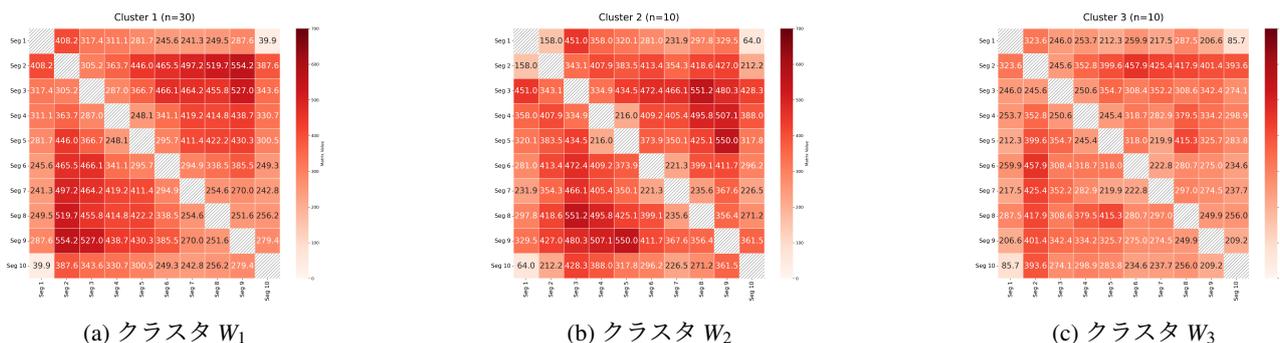


図 3: クラスタ $W_1 \sim W_3$ ごとのセグメント間のホワイトボードへの描画位置の重心の距離（色が濃いほど描画位置の重心の距離が離れている。）

表 4: クラスタごとのアンケート結果。（太字は各条件での最大値）

	Q1	Q2	Q3	Q4	Q5	Q6	Q7
V_1	6.33	6.40	6.57	6.62	6.40	6.48	N/A
V_2	6.67	6.67	6.83	6.73	6.67	6.80	N/A
V_3	6.50	6.39	6.79	6.64	6.61	6.61	N/A
W_1	5.87	5.90	6.33	6.10	6.20	6.18	6.10
W_2	5.83	5.78	6.50	6.39	6.44	6.61	5.56
W_3	5.72	5.67	6.00	6.11	5.83	6.00	6.00

以上より、満足度の向上には話題の意味的連続性の維持が重要であり、ホワイトボードは深掘り時の意見の共有に有効であるほか、多様な話題を整理する活用方法がその有用性を高めると示唆される。

4 ホワイトボードの描画位置による対話の流れの分析

言語的な対話の流れに加え、ホワイトボード上の描画活動が物理的にどのように推移したかを定量化し、対話の流れの分析を行った。

4.1 ホワイトボード描画距離行列の作成

動画データから描画活動の重心を抽出し、距離行列を作成する手順は以下の通りである。

1. ホワイトボードあり対話データを 10 個の等間隔のセグメントに分割する。
2. 各セグメント両端のフレーム差分を抽出し、その差分の画像の描画領域の重心を算出する。描画活動が検知されない場合は重心なしとする。
3. 各セグメントの重心間のユークリッド距離を計算し、 10×10 の距離行列を作成する。

この行列において、値が小さい場合は、ホワイトボード上の近い領域で書き込みながら議論している。値が大きい場合は、描画領域が大きく移動し、議論の対象が転換したと考えられる。

図 3 は 3 節において作成したクラスタ $W_1 \sim W_3$ ごとの行列の平均を示す。加えて各クラスタにおける描画活動の解釈は以下の通りである。

W_1 隣接セグメント間の距離は小さく、セグメント間の時間が離れるほど距離が増大する傾向にある。これは、近い場所で描画をしている、すなわち、書いている内容を広げながら議論していると考えられる。

W_2 他のクラスタと比較して全体的に距離の値が大きいいため、ホワイトボードを広範囲に使用して議論が行われていると考えられる。

W_3 他のクラスタと比較して全体的に距離の値が小さいため、ホワイトボード上の近い領域に連続して描画が行われており、一貫して意味的に近い話題について対話していると考えられる。

W_1 の満足度が高いことと併せて考えると、上記は、満足度の向上には話題の意味的連続性の維持が重要であるという 3 節での分析を裏付けるものであると考えられる。

5 おわりに

本研究は、ホワイトボードを用いた創造的な共同作業が可能な対話システムの実現に向け、ホワイトボードの使用が対話の流れにどのような影響を与えるかを分析した。対話の流れを定量的に評価し、発話の類似度および描画位置の分析から、ホワイトボードは話題の一貫性維持に寄与し、多様な意見を整理する活用法が対話の満足度向上に繋がることが分かった。特に、満足度の高い対話では、ホワイトボード全体を活用して多様な意見を整理しつつ、文脈を維持する傾向が見られた。今後は、本分析で得られた知見を活かし、対話の流れに応じて適切に描画や整理を行う対話システムの実現を目指す。

謝辞

本研究は、JST ムーンショット型研究開発事業、JPMJMS2011 の支援を受けたものです。

参考文献

- [1] Takuma Ichikawa and Ryuichiro Higashinaka. Modeling collaborative dialogue in Minecraft with action-utterance model. In **Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics: Student Research Workshop**, pp. 75–81, 2023.
- [2] Xulin Zhou, Takuma Ichikawa, and Ryuichiro Higashinaka. Collecting and analyzing dialogues in a tagline co-writing task. In **Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)**, 2024.
- [3] Yuki Hirota and Ryuichiro Higashinaka. Investigating feasibility of large language model agent collaboration in minecraft and comparison with human-human collaboration. In **Proceedings of International Joint Conference on Natural Language Processing Asia-Pacific Chapter of the Association for Computational Linguistics, 2025**, 2025.
- [4] Takuma Udagawa and Akiko Aizawa. A natural language corpus of common grounding under continuous and partially-observable context. In **Proceedings of the Thirty-Third AAIL Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAIL Symposium on Educational Advances in Artificial Intelligence**, 2019.
- [5] José Lopes, Nils Hemmingsson, and Oliver Åstrand. The spot the difference corpus: a multi-modal corpus of spontaneous task oriented spoken interactions. In Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, H el ene Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis, and Takenobu Tokunaga, editors, In **Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)**, 2018.
- [6] Barbara Tversky. Visualizing thought. **Topics in Cognitive Science**, Vol. 3, No. 3, pp. 499–535, 2011.
- [7] Almudena Fern andez-Fontecha, Kay O’ Halloran, Sabine Tan, Peter Wignell. A multimodal approach to visual thinking: The scientific sketchnote. **Visual Communication**, Vol. 18, pp. 5–29, 2019.
- [8] Sachit Menon, Richard Zemel, and Carl Vondrick. Whiteboard-of-thought: Thinking step-by-step across modalities. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, In **Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing**, pp. 20016–20031, 2024.
- [9] Apurva Gandhi, Ryan Serrao, Biyi Fang, Gilbert Antonius, Jenna Hong, Tra My Nguyen, Sheng Yi, Ehi Nosakhare, Irene Shaffer, and Soundararajan Srinivasan. SLATE: A sequence labeling approach for task extraction from free-form inked content. In **Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: Industry Track**, pp. 206–217, 2022.
- [10] 藤枝佑喜, 東中竜一郎. ホワイトボードを使用した共同作業タスクにおける対話の収集と分析. 人工知能学会全国大会論文集 第 39 回全国大会 (2025), pp. 3Win5–33, 2025.
- [11] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In **Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing**, 2019.
- [12] Joe H. Ward. Hierarchical grouping to optimize an objective function. **Journal of the American Statistical Association**, Vol. 58, pp. 236–244, 1963.