

多重参与を可視化するオンライン対話システムの開発

Development of an Online Dialogue System for Visualizing Multi-Dialogue Participation

市川 雅也[†], 竹内 勇剛[†]

Masanari Ichikawa & Yugo Takeuchi

[†] 静岡大学

Shizuoka University

ichikawa.masanari.18@shizuoka.ac.jp

概要

本研究では、参与構造の概念に基づいて複数の対話に同時参与可能なオンライン対話システムの開発を目的とする。具体的には、Goffman（1981）によって提案された参与構造をグラフ構造で表現し、参加者がこのグラフを操作することで相互に参与の程度を視覚的に表現したり、認識したりすることができるようにした。本稿では開発システムの設計および実装の概要を報告し、多重参与、対話場の可動性、および参与の可視化という三点から、新しいオンラインコミュニケーション環境の可能性を議論する。

キーワード：Computer Mediated Communication (CMC), 参与構造, エコーチェンバー

1. はじめに

一般に対面状況における教室やオフィスでは、授業や会議の進行のために公然と行われる対話と、その参加者同士で内々に行われる対話とを同時に存在させることが可能である（図 1(a)）。このように一つの空間に複数の対話が点在している対面状況は、参加者が複数の対話に掛け持ちして所属すること（多重参与）を許容している。参加者は公然と行われる対話を聞くだけでなく、自らが主体となって近傍の者と内々に対話することで情報交換をしたり、それらを傍観・傍聴したりすることで他者の意見や発想に触れることができ、しばしば個人／集団の相互理解や問題解決に役立てられている。

ところが多くのオンライン対話環境では、各対話はそれぞれ個別のチャンネルで管理されることが主流である。その構造から参加者は複数の対話に同時に参与することは許容されておらず、上述したような機会を得ることは極めて難しくなる（図 1(b)）。これによりオンライン対話環境では、多様な視点や異なる価値観に基づく、集団での問題解決や合意形成が困難と

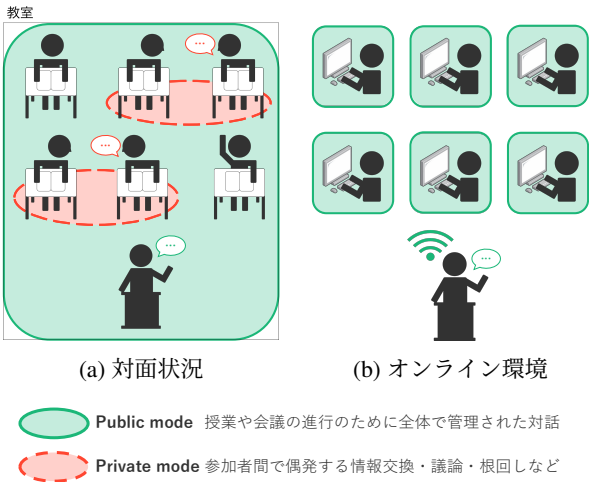


図 1: 対面／オンラインの対話場の構造の違い

なったり、集団内の議論や意思決定が、他者の視点に触れにくくなることで局所化（エコーチェンバー化）しやすくなったりするなど、知的な議論を実践する環境として課題がある。

本研究は多重参与を実現するオンライン対話システムの開発を目的とする。対話システムの設計に際し、枠組みとして参与構造を利用した。さらに、これを対話システムの設計に組み込むためにグラフ構造によって記述し、それを元に独自に対話システムを実装した。

2. 関連研究 – 参与構造

参与構造は Erving Goffman (1981) が指摘した対話の場における聞き手の役割の階層性に関する概念である（図 2）[1]。Goffman の参与構造によれば聞き手の役割は話し手になる権利がある「承認された参与者」と、そうでない「立ち聞き者」に大別される。承認された参与者とは一般的な意味での会話の参加者を指し、中でも発話をダイレクトに差し向けられている者を「受け手」、そうでない者を「傍参与者」と呼ぶ。立

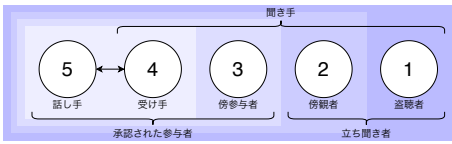


図 2: 参与構造 ([2] を元に筆者が加筆)

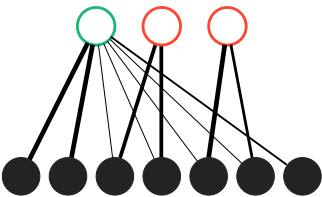


図 3: 参与構造グラフ

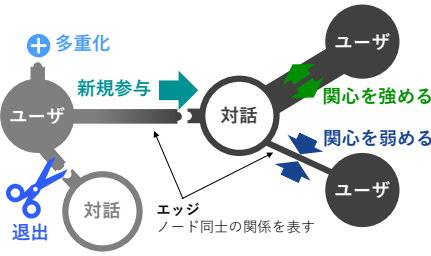


図 4: 参与構造グラフの操作

ち聞き者の内訳は「傍観者」と「盗聴者」であり、その名の通り対話場の外縁から遠巻きに関与する者である。このように参与構造は同じ聞き手の中にも役割の階層性があることを示している。

対面状況の対話において人はその身体配置 [3]、姿勢 [4]、物理的な距離 [5]、視線 [6, 7] などさまざまな身体動作を器用に調整することで、対話もとい参与構造を維持していることは広く理解されてきている。一般的なビデオ通話場面ではこうした非言語的なインタラクションが制約されており、本稿冒頭で指摘した問題の一因になっていると考えられている。オンライン対話の研究では 30 年以上も前から実世界／仮想世界においてアバターを利用した対話システムが多く開発されてきた (e.g., [8, 9, 10])。これらは対面状況を忠実に模倣するというアプローチにおいて、オンライン対話における参与構造の構築を試みる形になっていた。しかし現状、オンラインでの多人数対話の体験は対面状況のそれに匹敵しておらず、このことから、身体や視線など対面对話に適したインタフェースをオンライン対話にそのまま利用することの適切性について疑問が生じる。また、対面状況のインタラクションを模倣するアプローチは、対面状況特有の制約（身体的制約や地理的制約）を不要にオンライン環境に持ち込む可能性がある。

これまでにオンライン対話環境に適した形で参与構造を構築する試みはほとんど行われてこなかった。そこで本研究では、オンライン環境における参与構造の実現を目指す試みとして、グラフを利用した参与構造の記述方法を提案し、それに基づいてオンライン対話システムを設計・実装する。

3. 多重参与を実現する対話システムの設計

3.1 参与構造のグラフによる表現

グラフとは、ノードとエッジからなる離散的な要素間の関係を表す数学的な構造である。本研究で提案する参与構造グラフでは、参加者を表す「参与者ノード」(塗りつぶした丸)、対話の場を表す「対話ノード」(枠線の丸)、それらの関係を表す「参与エッジ」の 3 つの記号を定義する。すると図 3 に示すように、多重参与を含めた図 1(a) の様相を、重み (太さ) の異なる参与エッジによって記述できるようになる。さらに図 4 に示すようにエッジを繋いだり切ったり、その重みを増やしたり減らしたりすることで、グラフの操作によって参与者の各対話に対する状態を調整することができる。次節ではこれを実現するオンライン対話システム「Grape」(図 5) を紹介する。

ド」(塗りつぶした丸)、対話の場を表す「対話ノード」(枠線の丸)、それらの関係を表す「参与エッジ」の 3 つの記号を定義する。すると図 3 に示すように、多重参与を含めた図 1(a) の様相を、重み (太さ) の異なる参与エッジによって記述できるようになる。さらに図 4 に示すようにエッジを繋いだり切ったり、その重みを増やしたり減らしたりすることで、グラフの操作によって参与者の各対話に対する状態を調整することができる。次節ではこれを実現するオンライン対話システム「Grape」(図 5) を紹介する。

3.2 参与構造グラフに基づくオンライン対話環境の設計

Grape は、パソコン画面上の参与構造グラフを直接操作することで参与状態を表現できるオンライン対話システムである (図 5)。ユーザは画面中央に位置する対話ノードに対して、表 1 に示す操作を行うことで参与の状態をさまざまに調整することができる。以下にその主要な機能を紹介する。

参与する – 対話ノードをダブルクリックすると参与でき、対話を聴取できる。このとき、操作者の参与者ノードと対話ノードが半透明のエッジで結ばれ、全ユーザのグラフに同期する。これにより、新規参与者の存在が視覚的に認識可能となる。参与可能な対話場の数に制限はなく、複数の対話場に対して同様の操作を行うことで多重参与状態になる。なお、この時点では発言できず、「傍観者」としての状態にある。

参与の程度を調節する – 参与中の対話ノードを画面の左右中央に近づけるほど、その対話場から聞こえる音が大きくなる。このとき、音量の変化に伴ってエッジの太さも変化し、全ユーザのグラフが同期することで、他のユーザにも参与の強弱が視覚的に示される。これは同じ参与者の中でも参与意欲の高い／低い状態を表現している。

承認された参与者 (傍参与者) になる – 参与中の対話ノードを右クリックすると、マイクがアクティブになる。このとき、エッジのスクロール方向が逆転し、

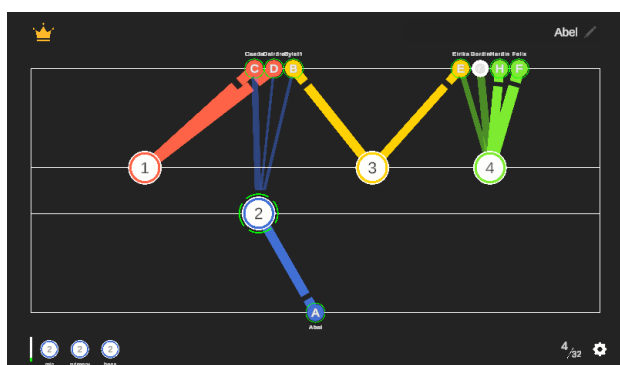


図 5: 参与構造グラフに基づいて設計されたオンライン対話システム

上段：他者の参与者ノード，下段：操作者の参与者ノード，中段下：自身が参与している対話ノード，中段上：それ以外の対話ノード。

操作者自身の参与者ノードの色が対話場のシンボルカラーに変化する。この「いつでも発言できる状態」は「承認された参与者」の状態を意味している。

発言する - 「承認された参与者」の状態で発話すると、音声に対話場の参与者に伝達される。発話中はエッジが不透明になり、話し手のノードに緑のリングが表示され、視覚的に強調される。この状態は「話し手」にあたり、最も際立って表現される。

3.3 多重参与状況での対話音声聴取を支援する音像定位

多重参与に関連して、人が複数の音から特定の音を選択的に聴取する上で、音の指向性が重要であることは広く知られている [11, 12, 13]。Grape では身体制御ではなく、対話音声自体を左右に定位させることで、多重参与状況での聞き分けを支援する仕組みを持っている。具体的には、図 5 の中段下の自身が参与している対話ノードをドラッグアンドドロップで左右に移動させると、音像が定位するようになっている。

4. 総合議論

本研究は、多重参与を実現するオンライン対話システムの設計と実装を目的としている。前章では開発した対話システム Grape について紹介した。本章では開発システムがもたらす特徴的なインタラクションやその効果について議論する。

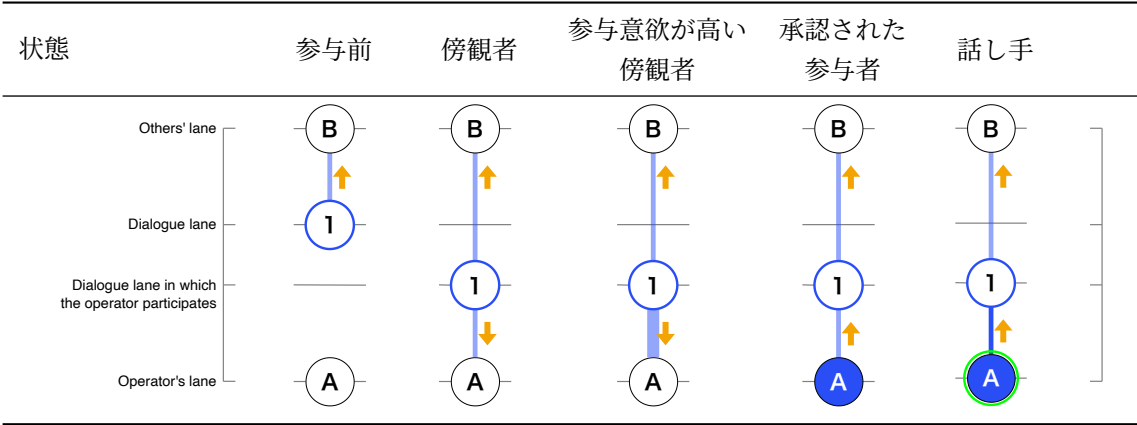
第一に本稿冒頭でも言及したオンライン対話環境における多重参与が可能になる。従来のビデオ会議システムでは、一人が発言している最中に別の参加者が発話すると、進行中の対話が妨げられてしまうことが普通であり、話し手でない者は聞き手にならざるを得なかった。しかし対面の会議であれば、特定の対話が

進行する背後で、参加者の一部が内々に議論や根回しを行っていることがあり、これが進行中の対話の円滑な進行や、有益な議論の布石になっていることが少なくない。複数の対話が並立可能な状況は、その対話場の数だけ話し手の数を増し、参加者がアウトプットする情報量が多くなることで、会議全体の了解性や創造性の向上する可能性がある。また特定の対話場に対する参与を強制させることで、意図的に対話を漏れ聞かせることができ、流動的な情報共有を促すと考えられる。これにより従来では困難であったオンライン対話環境での対話を通じた集団内／間の協調が促進され、学校の授業におけるグループワークや、オフィス内のオープンなコミュニケーションツール、エコーチェンバーを舵取りする仕組みとして役立てられることが期待される。

第二に身体や地理的な制約から解放されたコミュニケーションが実現する。対面状況では対話の場（対話を行なっている部屋、人が集まっている場所）は概ね不動であり、参与のためには人が動くことが主流である。このことは点在している対話場において、ある対話に近づくと、別の対話からは離れなければならないことを意味しており、多重参与における制約となっている。一方でオンライン環境では対話場が地理的に固定されていないため、人を不動にし、対話場を動かすというアプローチが可能になる。これにより離れた複数の対話に同時に参与できるという、いわば分身のような状態が生まれ、学術会議や合同企業説明会における情報の取得を大幅に効率化することが期待される。また上述したことは、任意のグラフを描くことであらゆる対話の場を自在にデザイン・アレンジできることを意味している。

第三に人の参与活動を可視化することができる。自分が参与していない対話場がどのような状況であるか認識することは、対面状況でも容易でなく、ビデオ通話では困難を極める。開発システムでは、何が話されているかはともかく、対話に対して誰が、どの程度参与しているかということは、各エッジの色の濃淡、明滅、線幅のリアルタイムでのフィードバックによって明らかとなる。このように参与に関わる非言語的な情報が補完されることで、テレワークやオンライン授業において損なわれがちな同僚（co-worker/co-student）の存在感が増すことで、オンライン環境特有の閉塞感を緩和され、通気性の高い対話環境を実現する。さらに可視化された参与活動は、同時にグラフという一般的なフォーマットにおいてデータ化し記録することができる。例えば図 6 は表 1 の 5 状態を左から順に実行

表 1: 参加者の状態に対するグラフの表示 †



† 黄色矢印は参加エッジのスクロールアニメーションの方向

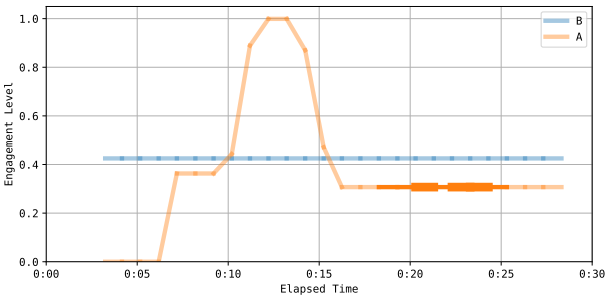


図 6: 参加状態の時系列データの例

表 1 の 5 状態を左から順に実行。Y 軸は参加の程度、色が濃い範囲はマイク起動中、線が太いところは発話中を表している。

した際の時系列データをプロットしたものである。このように、誰が、いつ、どの対話場で、どの程度の参加状態であったかの時系列データを収集・分析することで、特に学校のグループ活動における学び合いや協働、学習者間のインタラクションに関して、現実の状況を定量化したデータに基づいて議論することが可能になり、授業デザインや教育メソッド開発に寄与すると思われる。

以上のように本研究の開発システムは、オンライン対話環境における多重参加の実現を通して、情報共有の効率化や創造的な議論の促進に寄与する。さらに物理的な制約を伴わない対話場の構築や、参加活動の可視化など、新たなコミュニケーションのあり方を支える基盤となることが期待される。

謝辞

本研究は JST RISTEX の研究開発プロジェクト「ローカルエコーチェンバーをステアリングするトラスト調和メカニズムの認知的検討」(JPMJRS23L3)

による支援を受けたものである。

文献

[1] Goffman, Erving. (1981). Forms of talk. University of Pennsylvania Press.

[2] Clark, Herbert H. (1996). Using language. Cambridge university press.

[3] Kendon, Adam. (1990). Conducting interaction: Patterns of behavior in focused encounters. CUP Archive, 7.

[4] Schegloff, Emanuel A. (1998). Body torque. Social Research, 535–596.

[5] Hall, Edward Twitchell. (1966). The hidden dimension. Garden City, NY: Doubleday, 609.

[6] Duncan, Robert B. (1972). Characteristics of organizational environments and perceived environmental uncertainty. Administrative science quarterly, 313–327.

[7] Goodwin, C.. (1981). Conversational Organization. Interaction between speakers and hearers.

[8] Hirata, Keiji and Harada, Yasunori and Takada, Toshihiro and Aoyagi, Shigemi and Shirai, Yoshinari and Yamashita, Naomi and Yamato, Junji. (2006). The t-room: Toward the future phone. NTT Technical Review, 4(12), 26–33.

[9] Victor, Lyman and Porter, Lyman and Lawler, Edward. (2015). Expectancy theories. Organizational Behavior 1, 94–113.

[10] Zhang, Yizhong and Yang, Jiaolong and Liu, Zhen and Wang, Ruicheng and Chen, Guojun and Tong, Xin and Guo, Baining. (2022). Virtualcube: An immersive 3d video communication system. IEEE Transactions on Visualization and Computer Graphics, 28(5), 2146–2156.

[11] Bronkhorst, Adelbert W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. Acta acustica united with acustica, 86(1), 117–128.

[12] Bee, Mark A and Micheyl, Christophe. (2008). The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it?. Journal of comparative psychology, 122(3), 235.

[13] Hirsh, J. J.. (1950). The influence of spatial arrangement of sound sources on speech intelligibility. The Journal of the Acoustical Society of America, 22, 176–200.