

議論マップを用いた議論におけるフィードバック機能の構築と検証

桐木平 龍之介[†] 嶋田 和孝[†]

[†]九州工業大学情報工学部 〒 820-8502 福岡県飯塚市川津 680-4

E-mail: [†]{r_kirikihira,shimada}@pluto.ai.kyutech.ac.jp

あらまし 近年, PBL(Problem-Based Learning) に代表されるような問題解決型の教育が注目されている. PBL において, 議論参加者は問題を解決し合意を得るために協力的に議論を行う. 本研究では, 議論マップシステムと呼ぶ, 複数人での議論における合意形成を支援するためのシステムを提案する. 議論マップは議論状態を可視化・整理するためのツールで, ノードとリンクから成る. ノードは選択肢と評価基準の2種類が存在し, ノード間のリンクは評価値を保持している. システムは, 個々の議論マップから得られる評価値を統合し, 参加者へフィードバックを行うことで議論の合意形成を支援する. 本論文では, 複数人による議論において, 議論マップシステムの有効性を検証する. キーワード 議論マップ, 合意形成支援システム, ファシリテーション, フィードバック

Feedback for supporting consensus-building on discussion map

Ryunosuke KIRIKIHIRA[†] and Kazutaka SHIMADA[†]

[†] Computer Science and Systems Engineering, Kyushu Institute of Technology

Kawazu 680-4, Izuka, Fukuoka, 820-8502 Japan

E-mail: [†]{r_kirikihira,shimada}@pluto.ai.kyutech.ac.jp

Abstract Problem-based learning (PBL) is one of the most important learning approaches in recent years. In PBL, participants in a group discuss collaboratively to solve a problem and to make a decision through a conversation. In this paper, we are developing Discussion Map (DM), which is a tool for supporting consensus-building on multi-party conversations. It consists of nodes and links between them and the nodes have two roles; alternatives and criteria in the discussion. Each criterion and each link between nodes on DM contain importance values. The DM system estimates the current consensus in the discussion by using these importance values. The estimated result leads to the improvement of discussions as a digital facilitator. We evaluate the effectiveness of the DM system with consensus-estimation experimentally.

Key words Discussion Map, Decision support system, Facilitation, Feedback

1. はじめに

近年, PBL に代表されるような問題解決型の教育の重要性が増している [1]. このような学習環境下では, グループでの議論が頻繁に行われ, 問題解決に関する多くの選択肢とそれらの評価基準が飛び交い, 最善の選択肢を決定するために奮闘する. しかしながら, 参加者だけではアイデアの幅が狭くなる場合や全員が満足のいく決定をできない場合が多く存在する. 円滑で効率的な議論を行うためには, 適切な議事進行などを行えるファシリテーターが必要であるが, PBL 環境下ですべてのグループにそのような人員を割り当てることは困難である.

本研究の目的は, 議論を円滑に進めるためのシステム構築の一環として, PC やタブレット端末を利用した議論の合意形成をデジタルファシリテーターとして支援するシステムを構築す

ることである. この目的のために, 我々は議論マップと呼ぶ議論を可視化・整理する Web ツールを作成している (図 1). 議論マップは, 話題やキーワードをノード, そのノード間の関係をリンクとするグラフ構造である. これまでに, その操作性や議論で合意される事項と参加者の評価傾向について調査を行った [2].

本論文では, これまでの調査を基に, デジタルファシリテーターとして実際に合意形成を支援するためのフィードバック機能について検証を行う. 本システムの目的は合意形成の支援であり, 一般的な議論という中盤から終盤にかけての, 議題に対する各選択肢を評価するための話題を膨らませ, 収束するまでの部分を支援対象としている. システムがデジタルファシリテーターとして機能するためには, まず議論参加者それぞれの評価と議論全体の評価を把握, 理解する必要がある. そこで本

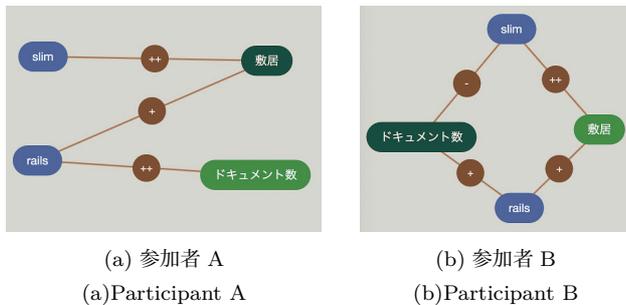


図 1 議論マップの例.

Fig. 1 An example of discussion maps.

システムでは、まず議論マップから得られる情報を基に、参加者全員の評価を統合することで議論全体の評価を推定する。そして、この統合された評価をいくつかのグラフで表示することで、潜在的な話題の発見を支援する。また、拮抗している選択肢の優劣をつけるため、議論すべき点を文章にて促す。これらのフィードバック機能 (DM アシスタント) を実装した本システムを用いて複数人による議論を行い、その有効性を検証する。

2. 関連研究

概念マップのようなグラフ構造を利用することは教育支援のための最も効果的なアプローチの一つである。Villalon と Calvo [3] は、概念マップを評価するための定義と枠組みを提案した。Yamasaki ら [4] は、概念マップに基づき、Kit-Build 方式を提案した。これらの研究の目的は、概念マップの自動抽出と、ゴールマップと学習者マップを比較することにあつた。それに対し、本研究の目的は、議論状態の管理と参加者の合意を評価することである。

タブレット端末を用いた議論支援として、Takagi と Shimada [5] のシステムがある。このシステムは参加者間の理解度の推定を目的としていた。桂ら [6] は、議題が論理的に正しいか否かの 2 値に収束させることを目的とするツールを提案した。本研究の目的は参加者の意見を集約して複数の候補から正解のない一つの解を導くところにある。この点が、これらの研究と異なる点である。

複数の評価基準から総合的な評価を得るためのツールとしては、Conati らの ValueCharts [7] や、Gratzl らの LineUp [8] がある。このツールは、選択肢に対して複数の評価基準を設け、その評価と重要度を基に、もっとも評価の高い選択肢を視覚的に見つけることが可能である。これらのツールはある程度評価基準が定まった場合に特に有効である。一方、本研究では選択肢や評価基準を動的に作成する。さらに、複数人の参加者における最善の選択肢、つまり合意を見つけることを可能とする。

3. 議論マップシステム

本研究で提案する議論マップシステムについて説明する。本システムでは、議論マップと呼ぶ、議論における話題やキーワードをノード、ノードの関係をリンクとするグラフ構造を構築することで議論の整理を行う Web ツールを用いる。議論マップ

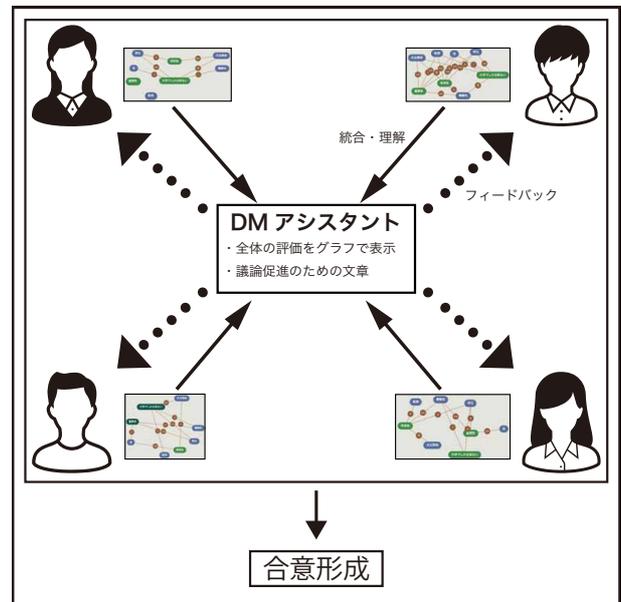


図 2 システムの流れ.

Fig. 2 The outline of our system.

から得られる情報を基に、参加者全員の評価を統合することで議論全体の評価を把握・理解する。システムは、参加者それぞれの評価と議論全体の評価を基に、議論を円滑に進めるためのフィードバックを返す。

図 2 に、本システムを用いた議論の流れを示す。合意形成を行うような議論において、参加者はそれぞれの考えを反映した議論マップを構築する。参加者は議論中、行き詰った時や収束に時間がかかっている時などに、DM アシスタントと呼ぶ、議論促進のためのフィードバック機能を使用して、円滑に議論を進める。

本節では、議論マップによる議論の整理と、議論全体の評価の把握について説明する。

3.1 議論マップ

議論マップの目的は、議論状態や参加者の意図及び理解状況を可視化することにある。前述の図 1 は、「Web サービスのフレームワーク比較」に関する議論マップを構築した例である。

議論マップはノードとリンクから構成される。ノードは、議題に対する選択肢と、それらを評価するための視点である評価基準の 2 種類ある。選択肢は青色のノードで示し、評価基準は緑色のノードで示す。図 1 では、「slim」と「rails」が選択肢ノードで、「数居」と「ドキュメント数」が評価基準ノードを意味している。図中で、評価基準ノードの色の濃淡が評価基準の重要度を示す。重要度は三段階で、色が濃いほど重要度が高く、淡いほど重要度が低いことを意味する。

選択肢ノードと評価基準ノードの間にはリンクを作成し、++ から -- まで 5 段階の評価値を持たせることができる。評価値は、リンク線上の茶色の円で示す。

議論マップは、ノードのみ参加者間で共有し、リンク関係や重要度、評価値は共有しない。これにより、各参加者が議論マップをどのように構築し、どのような評価をつけるかは自由

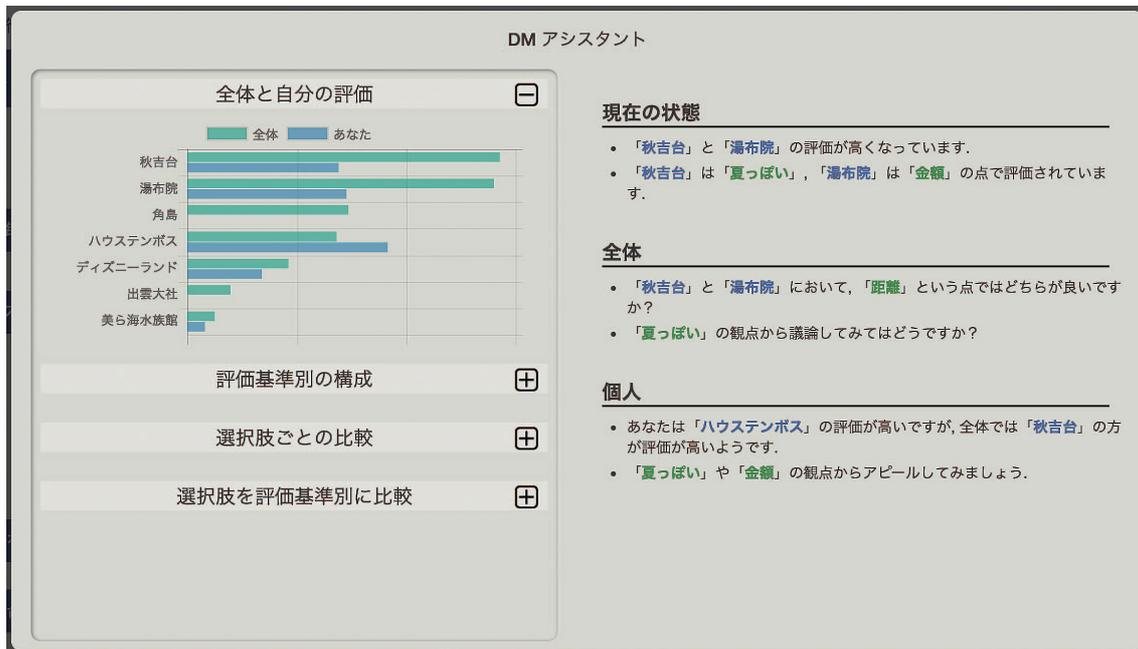


図 3 DM アシスタントの例.
Fig. 3 An example of the DM Assistant.

である。

図 1 を例に説明する。図 1 において、議論マップ上に存在するノード(「slim」, 「rails」, 「敷居」, 「ドキュメント数」)は参加者 A, 参加者 B とも同じものである。一方でその繋ぎ方と各ノードやリンクの重みは必ずしも同じではない。システムはこの差から各参加者の考え方の違いを推定できる。たとえば、選択肢「slim」については、両者とも評価基準「敷居」について高い評価(++)をつけている。しかし、参加者 B はこの「slim」について、評価基準「ドキュメント数」でも評価しており、その評価は - であり、高くない。加えて、参加者 B は「敷居」は重要ではない(色が薄い)と考えている一方、「ドキュメント数」は重要(色が濃い)だと考えていることが読み取れる。すなわち、両者とも「slim」には一定の評価を与えているが、参加者 A に比べると参加者 B の評価は必ずしも高くないということがこの議論マップの差からわかる。

3.2 評価値の統合

本システムでは、議論マップから得られる参加者全員の評価を統合することで議論全体としての評価を把握・理解する。本節では、参加者全員の各選択肢に対する評価を統合する手法について説明する。

単純な統合方法は、各選択肢の評価値を各評価基準に対する評価値の総和とすることである。本システムでは更に、時間的な特徴を考慮する。参加者は議論の序盤で、異なる意見を基に選択肢と評価基準を作り出す。そして中盤以降、重要な選択肢について絶えず議論しながら、重要でない選択肢を選別する。つまり、議論されている現在の話題は、それ以前の話題よりも重要であるといえる。そこで本研究では、この時間的な特徴を考慮するために Ebbinghaus の忘却関数 [9] を導入する^(注1)。

(注1)：予備実験により、忘却関数の有効性は確認されている。

Ebbinghaus は、最初の記憶量を 100 とした時の、 t 分後の記憶量 b を以下の式で表している。

$$b = \frac{100k}{((\log t)^c + k)} \quad (1)$$

c と k は定数で、Ebbinghaus はそれぞれ $c = 1.25$, $k = 1.84$ としている。本システムでは、(1) 式から算出できる記憶率 $sv = \frac{b}{100}$ を利用する。参加者 P_i の選択肢 A_j の評価値 $SvScore(P_i, A_j)$ を、

$$SvScore(P_i, A_j) = \sum_k^N w_{C_k} \times ev(P_i, A_j, C_k) \times sv_{A_j} \times sv_{C_k} \quad (2)$$

とする。 N は議論マップ中の評価基準ノードの数を表し、 C_k は評価基準、 w_{C_k} はその重要度を表す。 $ev(P_i, A_j, C_k)$ は A_j と C_k 間のリンクの値で、 sv_{A_j} , sv_{C_k} は、 A_j と C_k の記憶率を表す。経過時間 t は、それぞれ最後に操作(作成, 評価値・重要度の変更, ノードの移動)した時間からの経過時間とする。

最後に、選択肢 A_j の評価値 $Score(A_j)$ を、各参加者の評価値 $SvScore(P_i, A_j)$ の平均で表す。

$$Score(A_j) = \frac{\sum_i^N SvScore(P_i, A_j)}{N} \quad (3)$$

4. DM アシスタント

本節では、DM アシスタントと呼ぶ、3.2 節で統合した評価値 $Score(A_j)$ を基にするフィードバック機能について説明する。

本システムの目的は、議論の合意形成の支援であり、一般的な議論でいう中盤から終盤にかけての、各選択肢を評価するための話題を膨らませ、収束するまでの部分を支援対象としている。そこで、統合した評価値をいくつかのグラフで表示することで、潜在的な話題の発見を支援する。また、グラフの分析能

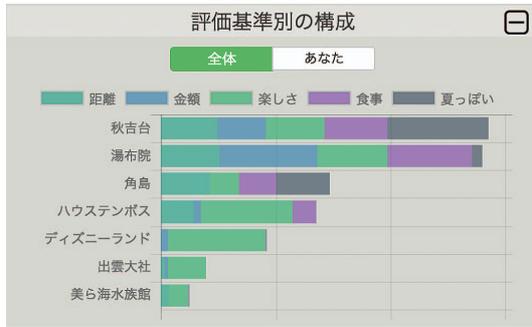


図4 「評価基準別の構成」グラフ。

Fig. 4 A graph of the distribution among the criteria.

力に左右されないように、いくつかの項目については文章で提示して、議論の促進を図る。

図3に、「次の旅行の目的地」に関する議論を行った時のDMアシスタント画面の例を示す。この画面は、統合した評価値のグラフ(左側)と、議論促進のための文章(右側)から成る。グラフは4種類あり、それぞれ異なる視点から評価値を閲覧、比較できる。各グラフは、グラフ名の右にある「+」、「-」ボタンで開閉できる。文章は、現在の議論状態、議論全体に対してのフィードバック、参加者個人に対してのフィードバックの3つの視点から生成する。

4.1 全体と自分の評価

参加者全員が満足する決定を行うためには、場全体の評価と自分の評価を把握し、その評価のギャップを埋める必要がある。そのためにまず、図3の左上に示すようなグラフを掲示する。このグラフは、全体の評価値と自分の評価値を比較するための棒グラフで、自分の評価が全体にどの程度共有されているかの把握に役立つ。図3の例では、全体では「秋吉台」の評価が高くなってきているのに対し、自分が高評価をつけている「ハウステンボス」は全体ではあまり評価されていないことが見てとれる。

4.2 評価基準別の構成

議論終盤において、いくつかの選択肢の評価が拮抗し、最終的な決定を決めかねるという状況がよく見られる。図4で示すグラフは、それぞれの選択肢の評価基準別の評価を比較するための棒グラフである。このグラフを見ることで、どの評価基準が重要か、ある評価基準についてどちらの選択肢が良いか、といったことを議論するきっかけになり、拮抗している選択肢の優劣をつけることができる。また、グラフ上部のボタンで自分の評価に切り替えることができ、自分の意見を整理することも可能である。図4の例では、評価が拮抗している「秋吉台」と「湯布院」の優劣をつけるために、評価が同じような「距離」の観点から見た時にどちらが良いか、といったように議論を膨らませることができる。

4.3 選択肢ごとの比較

図5に示すレーダーチャートは、4.2節の棒グラフと同様、選択肢の評価基準別の評価を比較するためのグラフである。このグラフは複数の選択肢に焦点を絞り、評価基準別の評価をより直感的に比較することができる。表示する選択肢は、グラフ

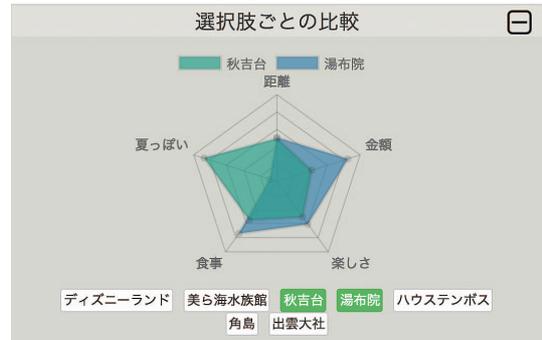


図5 「選択肢ごとの比較」グラフ。

Fig. 5 A graph of alternative comparison.

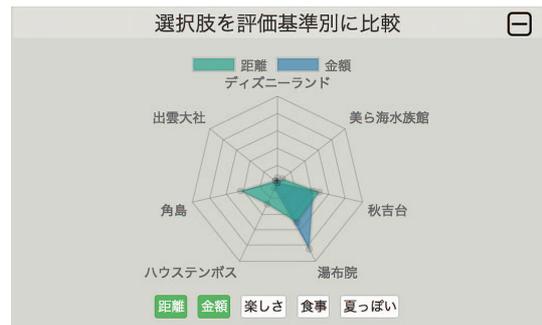


図6 「選択肢を評価基準別に比較」グラフ。

Fig. 6 A graph of criterion comparison.

下部のボタンで切り替えることが可能である。図5の例では、「秋吉台」が「夏っぽい」という点で評価が高いのに対し、「湯布院」は「金額」や「食事」の点から高評価であることが、直感的に理解できる。

4.4 選択肢を評価基準別に比較

ある評価基準について議論している際、各選択肢に対しての評価を把握したい状況は少なくないだろう。そこで、図6に示すような、各選択肢の評価基準別の評価を一見して比較するためのレーダーチャートを用意する。グラフ下部の評価基準ボタンで、表示する評価基準を切り替えることができ、評価基準同士の比較も可能としている。図6の例では、「距離」の点では「角島」が、「金額」の点では「湯布院」が高評価であることが把握できる。

4.5 現在の議論状態

議論の状態は、図3の「現在の状態」セクションのように、高評価の選択肢とその選択肢がどの評価基準で評価されているのかで表す。高評価の選択肢は、以下のルールにより選出する。

- 全体で最も評価の高い選択肢
- 選出された選択肢との差が5%以下^(注2)の選択肢

評価基準は、選出された各選択肢の評価の中で最も評価の高いものを選出する。

4.6 議論全体に対してのフィードバック

全体の評価値を基に、参加者全員に対して議論促進のためのフィードバックを行う。本システムの目的は合意形成の支援で

(注2) : 全ての選択肢の評価値の合計を100%とする。

ある。そこで、4.5節で選出する選択肢の優劣をつけるための情報を、図3の「全体」セクションのように掲示する。

評価基準は、選出する選択肢が評価を行っており、且つ最も評価の分散値が低いもの、つまり、優劣をつける余地があるものを選出する。また、全ての評価基準の中で、最も評価がつけられていないものも選出する。しかしながら、全選択肢の半分以上に評価が付けられていた場合は、選出しないようにする。

4.7 参加者それぞれに対するフィードバック

議論参加者それぞれの評価値を基に、高評価の選択肢、つまり、自分の主張を参加者全員に共有、理解させるためのフィードバックを行う。図3の「個人」セクションのように、全体の評価と自分の評価が異なった時、アピールしたほうがよい評価基準の掲示を行う。評価基準は、全体で最も評価の高い選択肢と自分の主張する選択肢を、全体の評価基準別評価と比較した時に、自分の選択肢の方が評価が低いものを選出する。

5. 実 験

提案する議論マップシステムを評価するために、議論実験を行う。本節ではその内容と結果について説明する。

5.1 実験設定

本実験は、4人1組で「日本の首都を移すならどこがいいか」というテーマのもと20分間議論を行い、合意事項を1つ決定する。参加者は議論中、自分の考えを反映した議論マップを構築し、適宜DMアシスタントによるフィードバックを得る^(注3)。

議論後にアンケートをとり、DMアシスタントの評価を行った。本研究の目的は合意形成の支援であるため、DMアシスタントによるフィードバック機能の満足度について評価、考察を行った。満足度は、DMアシスタントに関する次の9つの項目について5段階で評価した^(注4)。

- 各グラフ(4種類)
- 文章による議論促進(3種類)
- DMアシスタント全体のレイアウト
- DMアシスタント全体の操作性

5.2 実験結果と考察

本実験は、議論マップを用いた議論の経験があるグループ1組(議論1)、議論マップを初めて使うグループ1組(議論2)、合計2組で行った。表1に、各参加者の満足度の平均を示す。

まずグラフについて考察を行った。棒グラフで表現されるグラフ1及びグラフ2は高い評価を得られたが、レーダーチャートで表現されるグラフ3及びグラフ4は評価が下がっていた。この原因として考えられるのは、議論中に複雑な操作、分析はできない、という点である。グラフ3とグラフ4は、表示する項目をボタンで切り替えて比較することが可能であるが、どの項目を比較するか、どういった傾向があるか、といったことを議論をしながら考えるのは容易ではない。その反面、グラフ1とグラフ2は特に操作も必要なく、全体を直感的に把握できる

(注3)：議論序盤は情報が少ないため、DMアシスタントの機能は開始5分後から使用してもらうようにしている。

(注4)：5段階評価は、1を「とても不満」、5を「とても満足」としている。

表1 各議論の満足度

Table 1 The degree of satisfaction of each discussion.

項目	議論1	議論2	平均
グラフ1「全体と自分の評価」	4.25	4.50	4.38
グラフ2「評価基準別の構成」	4.25	4.50	4.38
グラフ3「選択肢ごとの比較」	3.67	3.50	3.58
グラフ4「選択肢を評価基準別に比較」	3.75	3.25	3.50
文章1「現在の状態」	4.25	4.75	4.50
文章2「全体」	2.75	3.50	3.13
文章3「個人」	3.00	3.25	3.13
DMアシスタント全体のレイアウト	4.25	4.25	4.25
DMアシスタント全体の操作性	4.50	3.50	4.00

ため、このような差が生まれたと考えられる。

続いて、文章による議論促進について考察を行った。「現在の状態」に関しては、高評価の選択肢と評価基準が一目でわかるため、高評価を得られた。しかし、「全体」と「個人」は評価が低くなっていた。これは、文章がほとんど表示されなかったために満足度が得られないことが原因であった。今回は、拮抗する選択肢や評価が異なる選択肢が存在するときにその優劣をつけるための情報を提示するようにしていた。そのため、拮抗する選択肢がない時、つまり議論がうまくいっている時は何も表示されない。今回行った議論はこのような状況だったため、文章がほとんど表示されず、評価が低くなった。一方で、自由記述のアンケートから、表示される内容そのものについては高い評価を得られていた。よって、本システムの目的としては良い結果を得られたといえる。

また、議論マップを用いた議論経験の有無についても大きな差はなく、レイアウトや操作性に問題がないことを確認できた。

6. おわりに

本研究では、円滑に議論を進めるためのシステム構築の一環として、タブレット端末を利用した議論の合意形成を支援するシステムの構築を目標としている。本論文では、実際に合意形成を支援するためのフィードバック機能の構築と評価を行った。結果として、その有効性を確認することができた。また、改善点も見つかри、今後のシステム改善の足がかりを得ることができた。今後は、DMアシスタント機能の改善を行い、議論支援の幅を広げていく予定である。

本研究の最終的な目的は、マルチモーダルの対話や議論を理解・管理し、効率的な議論を支援することである。今後は、議論マップシステムに加え、画像から得られる参加者の行動[10]や音声などから得られる対話の内容[11]を利用した議論の場の状態推定といった技術を用いて、円滑で効率的な議論支援システムの構築を目指す。

謝辞

本研究は科研費26730176の助成を受けたものです。

文 献

- [1] Cindy E. Hmelo-Silver: Problem-Based Learning: What and How Do Students Learn?, *Educational Psychology Review*, Vol.16, pp.235-266 (2004).

- [2] 桐木平 龍之介, 嶋田 和孝: タブレット端末を用いた議論の合意形成支援システムの開発, 人工知能学会, 先進的学習科学と工学研究会, Vol.75, pp.35-38 (2015)
- [3] Villalon J. Jorge, Calvo A. Rafael: Concept map mining: A definition and a framework for its evaluation. In *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008.*, pages 357-360, 2008.
- [4] Kazuya Yamasaki, Hiroyuki Fukuda, Tsukasa Hirashima, and Hideo Funaoi. Kit-build concept map and its preliminary evaluation. In *Proceedings of The 18th International Conference on Computers in Education, ICCE 2010*, 2010.
- [5] Hironobu Takagi, Kazutaka Shimada : Understanding Level Estimation Using Discussion Maps for Supporting Consensus-building, *Procedia Computer Science*, Vol.35, pp.786-793 (2014)
- [6] 桂 祐樹, 岡田 将吾, 新田 克己: 論証ダイアグラムを用いた動的議論支援ツールの提案, 人工知能学会全国大会論文集, Vol.29, pp.1-4 (2015)
- [7] Cristina Conati, Giuseppe Carenini, Enamul Hoque, Ben Steichen, Dereck Toker: Evaluating the Impact of User Characteristics and Different Layouts on an Interactive Visualization for Decision Making, *Computer Graphics Forum (Proc. EuroVis 2014)*, Vol.33, No.3, pp.371-380 (2014)
- [8] Samuel Gratzl, Alexander Lex, Nils Gehlenborg, Hanspeter Pfister and Marc Streit: LineUp: Visual Analysis of Multi-Attribute Rankings, *IEEE Transactions on Visualization and Computer Graphics*, Vol.19, No.12, pp.2277-2286 (2013)
- [9] Hermann Ebbinghaus: MEMORY: A Contribution to Experimental Psychology. *Tras. by H. A. Ruger & C. E. Bussenius*, New York: Dover Publications (1885)
- [10] 小松 和朗, 嶋田 和孝, 遠藤 勉: 話者の頭部及び姿勢変化に着目した複数人対話分析, 火の国情報シンポジウム 2013, B-5-1 (2013)
- [11] 嶋田 和孝, 楠本 章裕, 横山 貴彦, 遠藤 勉: 複数人談話における笑いの情報を考慮した盛り上がり判定, 電子情報通信学会, NLC研究会, Vol. 112, No. 110, NLC2012-7, pp.25-30 (2012)