

会話の均衡を保つ採点モデルの提案とシミュレーション

山腰 悠真^{1,a)} 安達 寛之^{1,b)} 明神 聖子^{1,c)} 島田 伸敬^{1,d)}

概要: 多人数による会話では、話しすぎてしまう人とあまり話せない人が得てして存在するものである。ひとりひとりが提供する発話量に偏りがあると、会話の方向が一意に定まってしまう、多人数で話している意味がなくなってしまう。本稿では、Adachi らの研究内容を簡単に説明し、そのシステムで採用されている採点モデルの問題と解決に向けた理論の提唱をした。そして実際に問題が改善されたかどうかをシミュレーションを通じて判断して、問題の解決を確認した。

Proposal and Simulation of Scoring Model to Keep Balance of Conversation

YUMA YAMAKOSHI^{1,a)} HIROYUKI ADACHI^{1,b)} SEIKO MYOJIN^{1,c)} NOBUTAKA SHIMADA^{1,d)}

Abstract: In conversation by many people, people often speak too much or too little. If the amount of utterance of people is unbalanced, there is a possibility that the direction of the conversation is decided uniquely, because people don't get various opinions. This paper describes the research content of Adachi et al. briefly, and proposes new models for solving the problem of the scoring models adopted in the Adachi's system. The evaluation results of our models in simulation show that the models have actually improved the problem.

1. 会話支援を行う背景

企業や教育現場などにおいて、ミーティングやブレインストーミングといった、多人数による会話を行う機会は少ない。複数人でユニークなアイデアを交わし合うことによって他に類を見ない新たな発見が生まれたり、会話を通して仲間との結束を固めたりと、多人数での会話はとても有益なものである。しかし、そのような満足のいく結果を生み出すには、みな平等に意見を交わすことが大切である。実際の多人数による会話では、発話量の多い人や少ない人が存在し、個人から得られる情報量には差異が生じてしまうことが多い。発話量の多い人は場を支配して、会話の方向性を一つに収束させる恐れがある。また、発話量

の少ない人は、場を与える影響が小さく、活発な会話が行われなくなる一因になりうる。会話参加者全員が一樣に発言出来てこそ、会話の質はより洗練されたものになると考える。そこで発話量の多い人には発話を控えてもらったり、発話量の少ない人には発話を推奨したりといった、会話支援機構による介入を行うことで少しでも会話の活性化に繋がるなら望ましいことであると考え、この研究を行う。

2. ScoringTalk の概要

理想的な多人数会話の実現を目指して、これまで多くの研究者らが前身となって会話支援研究を行ってきた [1],[2],[3]。今回はその中でも、Adachi らの ScoringTalk [4] の研究を受けて、その発話採点モデルの改善を目指す。

2.1 ScoringTalk のシステム詳細

ScoringTalk では、多人数での会話において、誰が・いつ・どこにいてどちらを向いて発話したかを、タブレット PC を用いて計測・可視化し、自分が話をした時間と相手に話しかけられた時間から会話に得点付けをして、ユーザ

¹ 立命館大学
1-1-1, Noji-higashi, Kusatsu, Shiga 525-8577, Japan
a) yamakoshi@i.ci.ritsumei.ac.jp
b) adachi@i.ci.ritsumei.ac.jp
c) seiko@i.ci.ritsumei.ac.jp
d) shimada@ci.ritsumei.ac.jp

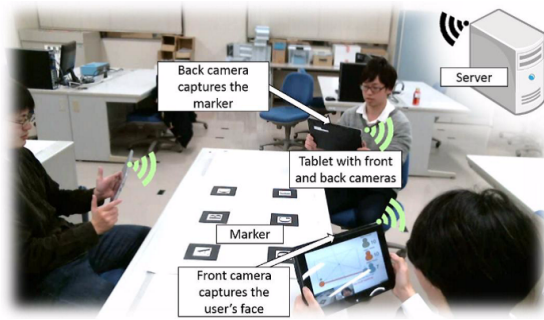


図 1 ScoringTalk のシステムの構成 [4]

Fig. 1 The structure of the system of ScoringTalk[4].

にフィードバックすることで発話量の調整を実現する。

誰が発話したかは、各ユーザが持つタブレット PC の ID から定まる。いつ発話したかは、顔計測で得られる開口度で判断する。どこにいてどちらを向いて発話したかは、ユーザ、タブレット、マーカーの位置関係から推定する。これらの情報から、誰が誰に話しているのか/誰を見ているのかを推定し可視化する。また、発話量の多い人には抑制、少ない人には推奨させるために、ユーザが話した時間と他のユーザに話しかけられた時間から、会話の得点を算出し、ユーザにフィードバックすることで、会話にゲーム性を持たせ、会話参加者から均等に意見を引き出せるようにする。

図 1 に ScoringTalk のシステムの構成を示す。これは、各ユーザがタブレット PC を持った状態で、マーカーが設置されたテーブルを囲んで会話を行う場面を表している。ユーザの発話を計測するため、タブレット PC のマイクでユーザの声の音量を計測し、前面カメラで顔検出、背面カメラでマーカー検出を行う。ユーザは前面カメラに自分の顔が映り、背面カメラに複数個あるマーカーのいずれかが映るように、タブレット PC を把持しながら会話する。このようにして計測した個々人の発話の情報は、無線 LAN で接続したサーバに送信される。サーバではそれらの情報を統合することで会話全体の特徴量として算出し、各ユーザにフィードバックする。ユーザは会話を行いながら、ときどき図 2 のような、タブレット PC の表示を確認することで、発話加減と得点を知ることができ、次の発話を調整することができる。

2.2 ScoringTalk の個人発話の採点モデル

ScoringTalk では会話量調整のために、発話の加減に基づいて得点を計算し、ユーザーにフィードバックして可視化するというゲーム的特性を会話に持たせる。その採点モデルは、発話量の少ないユーザにはより多く発話してもらえようように大きな得点を与え、発話量の多いユーザは会話の均衡を保つために、話しすぎると減点するといったもの

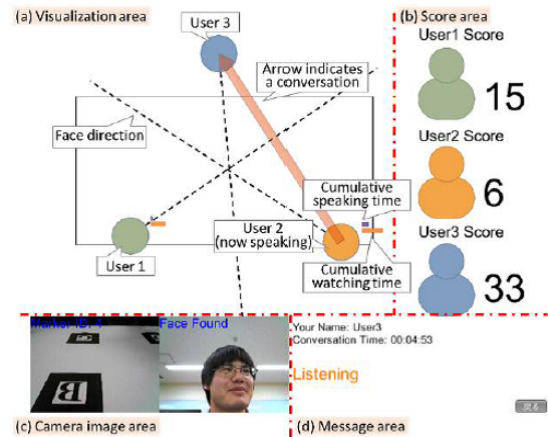


図 2 タブレット画面の例 [4]

Fig. 2 An example of tablet screen[4].

である。また、相手の話を聞くということも会話においては重要な要素であることを踏まえて、話しかけられた時にも得点が得られる。それぞれの採点方式の詳細は以下に示す通りのものである。^{*1}

• 自分が話した時の報酬

あるユーザ i が話したときの得点 SR_i は式 (1) のようにして算出する。

$$SR_i^{(t)} = \begin{cases} \alpha, & \text{if } t < T \\ \alpha * \left(1 - \frac{NS_i^{(t)}}{\sum_j^N S_j^{(t)}}\right), & \text{otherwise} \end{cases} \quad (1)$$

ここで、 α は話したときの得点比率のパラメータ、 S はユーザの累積発話時間、 t は計測開始時からの経過時間、 T は得点の計算式を切り替える時刻、 N は会話参加者の人数である。つまり、会話の最初の T 時間は発話した分だけ得点になるが、それ以降は、参加者全員の累積発話量の平均に対する自分の累積発話量の割合により得点が決定される。例えば、他の人と比べて発話量が少ない人は多くの得点を得るが、3 人で会話する場合に自分の発話量が全体の $1/3$ を超えている状態で発話すると、その分だけ減点される。これにより、各ユーザが自身の得点を最大化しようと振る舞うことで発話が加減され、参加者全員が均等に発言する方向に向かい、より多くの情報が相互にもたらされることが期待できる。

• 相手に話しかけられた時の報酬

相手に話しかけられたときには、一定の得点を得るものとする。すなわち、あるユーザ i が相手に話しかけられたときの得点 LR_i は、得点比率のパラメータを β として、式 (2) のようにして算出する。

$$LR_i = \beta \quad (2)$$

以上から、ある時刻 t におけるユーザ i の会話の得点

^{*1} ここでは [4] の採点モデルを厳密な式に直している

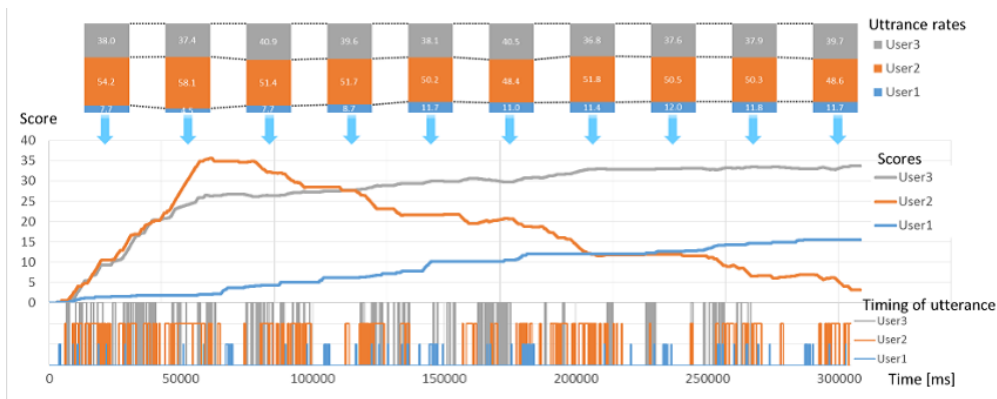


図 3 ScoringTalk の採点モデルを採用した時の各ユーザの得点の推移 [4]

Fig. 3 Transition of the scores of each user at the time of ScoringTalk[4].

$Score_i$ は、これら 2 つの得点の合計を時刻 0 から t まで足し合わせたものとする。

$$Score_i^{(t)} = \sum_{s=0}^t (SR_i^{(s)} + LR_i^{(s)}) \quad (3)$$

この得点をタブレット画面を通じてリアルタイムにユーザにフィードバックし、ユーザはそれを見ながら、各自の発話を調整する。

2.3 ScoringTalk の個人発話の採点モデルが抱える問題

式 (1) ~ (3) の採点モデルを用いて会話実験 [4] が行われたところ、ある問題点が浮き彫りとなった。図 3 は式 (1), (2) のパラメータを $T = 1[\text{min}]$, $\alpha = 0.05$, $\beta = 0.01$ とした時の、被験者 3 名による 5 分間の会話実験の計測結果である。最初の T 時間までにユーザ 2 の発話割合が全体の $1/3$ を大きく上回ったがために残りの時間ずっと減点を受けていることが見て取れる。この時ユーザ 2 がとるべきだった最善の行動は、最初の T 時間を超えた時点で、発話割合が均衡するまで一言も話さないことである。しかしそれは客観的に見て余りにも不自然かつ機械的な光景であり、そのような状況下で活発な情報交換が行われるとは考えにくい。また、減点を嫌うあまり、優れた意見やアイデアが思い浮かんだとしてもそれを話さずに棄却しなければならないといったことも想定され

3. 発話採点モデルの改善

ScoringTalk における採点モデルの問題は、発話割合の均衡に、強く重みを置きすぎたことによって生じたものである。会話の独占状態、話しすぎを判断するラインを別の基準で設ける必要がある。また発話量の少ない人は、自発的に話すことを不得手としている場合が多く、より多くの得点を与えるというだけでは発話を奨励する動機としては力不足である。それよりも周囲の人が折りをみて発言を促してあげた方がよいのではないかと考える。これらを踏まえ、問題解決に際して新たに「持ち時間」の概念を導入

することとする。これによって、話せる時間に上限を設けて、1 人が延々と話し続けるといった状況を回避することが出来る。さらに、自分より得点の低い相手に話しかけた場合、得点差に比例した大きさのボーナス補正が加算量にかかるという要素を新たに追加する。これは発話量の少ない人に積極的に話しかけることで、発話を促すというねらいがある。また、話しかけられた側には、相手のほうが得点が高い場合に限り、得点差に比例して回復速度が速くなるといった恩恵を与えることとする。そうすることによって、発話量が少ない人にもたくさん話してもらうための時間を設けることが出来る。持ち時間と加点行為の詳細は以下の通りである。

あるユーザ i の持ち時間の推移を表す式 A_i は式 (4) のようにして算出する。

$$A_i^{(t)} = \begin{cases} A_i^{(t)} - c_i^{(t)} & \text{if } A_i^{(t)} > 0 \\ A_i^{(t)} + nc_i^{(t)} & \text{if } A_i^{(t)} < \alpha \end{cases} \quad (4)$$

ここで、 α は任意に設定する持ち時間の上限のパラメータ、 t は計測開始時からの経過時間、 c は発話時間、 nc は非発話時間を表している。持ち時間は発話時間 c に伴って消費され、非発話時間 nc に伴って回復する。つまり話している時は持ち時間を消費し、話していない時は持ち時間が回復するということである。ただし持ち時間は 0 を下回って消費されることはなく、上限を超えて回復することもない。また、以下の式 (5) に表されるように、得点差によって持ち時間の回復速度が速くなる。なお、(5) における SR とはユーザの得点、 γ とは回復速度倍率のことである。

$$A_i^{(t)} = A_i^{(t)} + \frac{\gamma SR_j^{(t)}}{SR_i^{(t)} + SR_j^{(t)}} nc_i^{(t)} \text{ if } SR_i^{(t)} < SR_j^{(t)} \quad (5)$$

あるユーザ i の時刻 t における加算量を決める変数であるレベル $L_i^{(t)}$ は、式 (6) のようにして算出する。

$$L_i^{(t)} = \begin{cases} 0 & \text{if } S_i^{(t)} < 25 \\ \text{Ceil} \left(\log_2 \frac{S_i^{(t)} + 1}{25} \right) & \text{if } 25 \leq S_i^{(t)} < 200 \\ 4 & \text{if } S_i^{(t)} \geq 200 \end{cases} \quad (6)$$

また、あるユーザ i が話した時の得点 SR_i は式 (7) のようにして算出する。

$$SR_i^{(t)} = \begin{cases} 1 - \beta L_i^{(t)} & \text{if } A_i^{(t)} > 0 \\ -1 & \text{if } A_i^{(t)} = 0 \end{cases} \quad (7)$$

ここで、 β は話した時の得点比率のパラメータのことである。レベルは累積発話時間 S が一定の値に達するごとに上昇していき、レベルがあがるにつれてユーザが得られる得点の値は減少していく。それに対し、減少量はレベルに関わらず一律なので、レベルが高い人、すなわち累積発話量が多い人ほど相対的に減点行為の重みが大きくなる。よって、発話量が多い人は話しすぎて持ち時間を使い果たしてしまうことがないようにことさら気をつけなければならない。また、ある状況下で、加算される際に、加算量にプラス補正がかかる時の得点 SR_i は式 (8) のようにして算出する。

$$SR_i^{(t)} = \frac{\delta SR_i^{(t)}}{SR_i^{(t)} + SR_j^{(t)}} (1 - \beta L_i^{(t)}) \text{ if } SR_i^{(t)} > SR_j^{(t)} \quad (8)$$

ここで、 δ とは加算補正量の倍率を決めるパラメータのことである。自分より点数が低い相手に話しかけた時は、発話量の少ない相手に発話を促したとみなして加算量にボーナスがかかる。

以上の採点モデルの特性によって、各ユーザーが得点を最大化しようと振る舞えば、会話参加者全員の発言量が均等になった上でより多くの情報交換が行われることが期待される。

4. 発話採点モデルの評価

4.1 シミュレーション方法

今回作成した採点モデルが正しく機能するかどうかを確認するために、会話データにあててシミュレーションを行う。会話データは2種類のものを用意し、一方は3名の被験者による実際の5分間の会話を記録したものである。これを会話データ1とする。会話データ1の発話割合と、誰が、誰に、いつ話したかの発話時間の推移は図4に示す通りである。もう一方はプログラムによって模式的に3名の被験者による5分間の会話を作成したもので、これを会話データ2とする。会話データ2の発話割合と、誰が、誰に、いつ話したかの発話時間の推移は図5に示す通りである。また、旧採点方式における式(1), (2)のパラメータは $T = 1[\text{min}]$, $\alpha = 0.05$, $\beta = 0.01$ とし、新採点方式における式(4), (5), (7), (8)のパラメータは $\alpha = 30[\text{sec}]$, $\beta = 0.2$, $\gamma = 3$ (時間回復倍率を1~3倍に), $\delta = 2$ (加算補正倍率を1~2倍に)とした。また、簡単のために、式(6)の $L_i^{(t)}$ の変化を表1に記す。実験の手順を以下に示す。

(a) 旧採点モデルと、今回考案した新採点モデルとで採点結果を比較する

(b) 持ち時間の回復速度に関する式(5)の有無で採点結

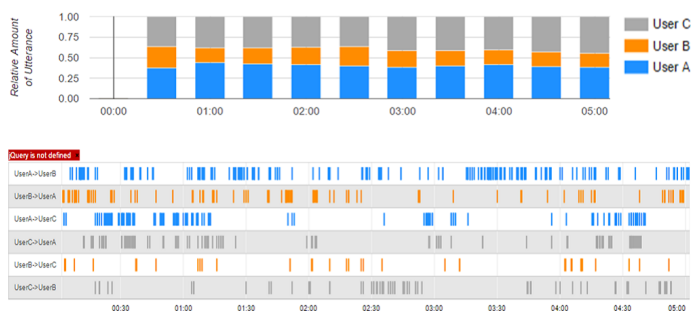


図4 会話データ1の詳細

Fig. 4 The detail of conversation data 1.

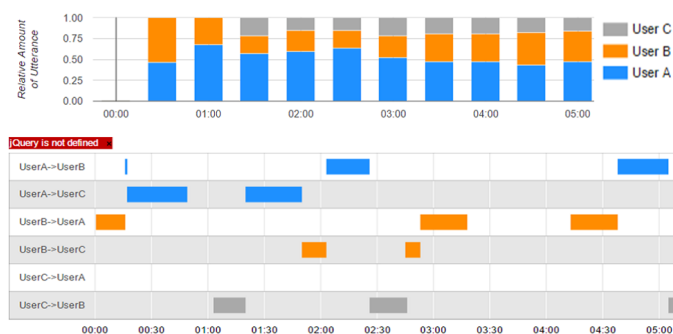


図5 会話データ2の詳細

Fig. 5 The detail of conversation data 2.

果を比較する(この時式(8)は適用しない)

(c) 加算量のボーナスに関する式(8)の有無で採点結果を比較する(この時式(5)は適用しない)

表1 レベル変化の推移

Table 1 Changes in the level.

$S[\text{sec}]$	0 ~ 24	25 ~ 49	50 ~ 99	100 ~ 199	200 ~
$L_i^{(t)}$	0	1	2	3	4

4.2 シミュレーション結果と考察

会話データ1に両採点モデルを適用した結果を図6と図7に、会話データ2に両採点モデルを適用した結果を図8と図9に示す。UserAに着目すると、旧採点モデルでは会話データ1, 2ともに序盤に話しすぎたことにより発話割合が均衡になるまで何を話しても減点を受けていることが分かるが、新採点モデルでは他のユーザの話の聞くといったインターバルを挟めば、後にまた自分が話せるだけの時間が必ず巡ってくるということが分かる。

会話データ1に時間回復式(5)を適用した結果を図10に、適用しなかった結果を図11に示す。会話データ2では両者ともに顕著な違いが見られなかったため、掲載を省略した。UserAに着目すると、減点を受け始めてから、時間回復式を適用したもののほうが適用しなかったものより得点が高くなることが分かる。これは、式(5)によって持ち

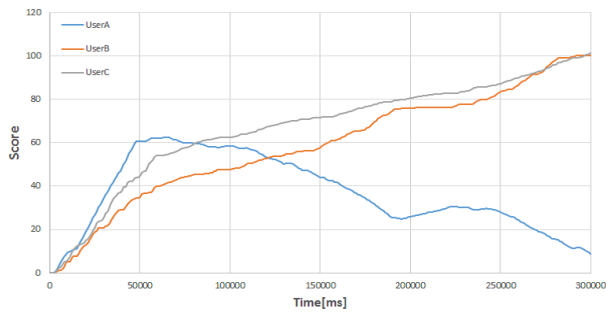


図 6 会話データ 1, 手順 (a) : 旧採点方式
Fig. 6 Conversation data 1, step (a)
 The conventional scoring models.



図 7 会話データ 1, 手順 (a) : 新採点方式
Fig. 7 Conversation data 1, step (a)
 The new scoring models.

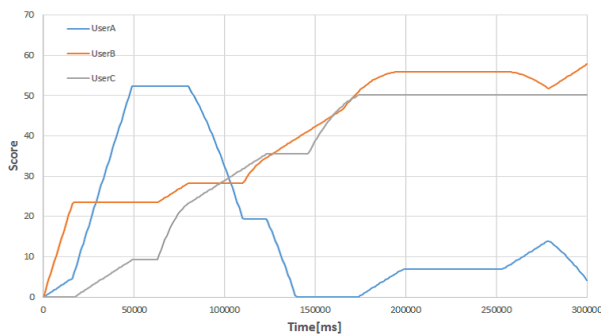


図 8 会話データ 2, 手順 (a) : 旧採点方式
Fig. 8 Conversation data 1, step (a)
 The conventional scoring models.

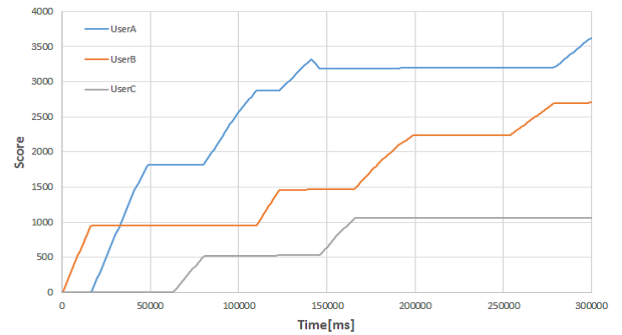


図 9 会話データ 2, 手順 (a) : 新採点方式
Fig. 9 Conversation data 1, step (a)
 The new scoring models.

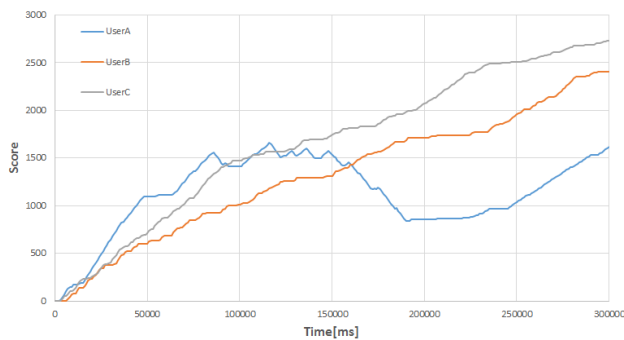


図 10 会話データ 1, 手順 (b) : 時間回復式あり
Fig. 10 Conversation data 1, step (b)
 With time recovery formula.

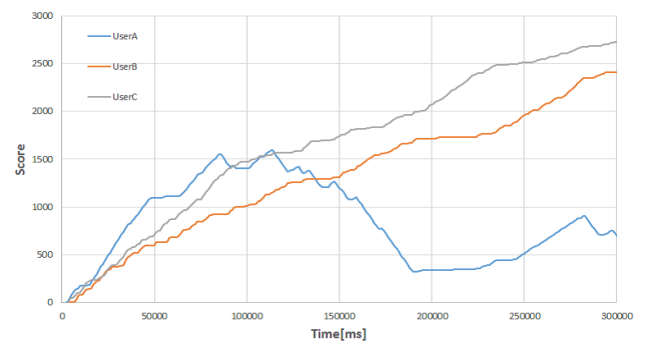


図 11 会話データ 1, 手順 (b) : 時間回復式なし
Fig. 11 Conversation data 1, step (b)
 Without time recovery formula.

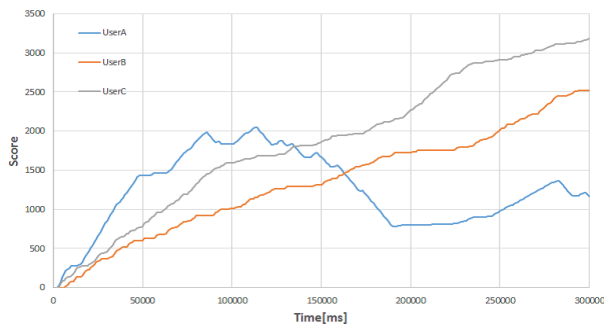


図 12 会話データ 1, 手順 (c) : 加分量補整式あり
Fig. 12 Conversation data 1, step (c)
With correction formula of additional score.



図 13 会話データ 1, 手順 (c) : 加分量補整式なし
Fig. 13 Conversation data 1, step (c)
Without correction formula of additional score.

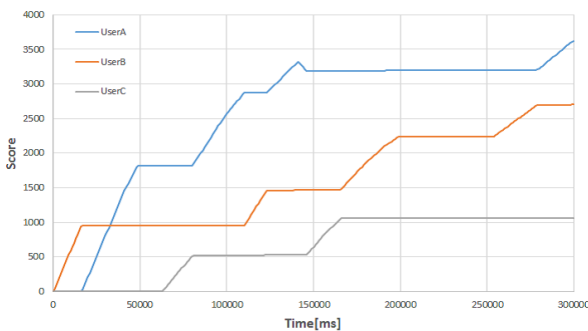


図 14 会話データ 2, 手順 (c) : 加分量補整式あり
Fig. 14 Conversation data 1, step (c)
With correction formula of additional score.

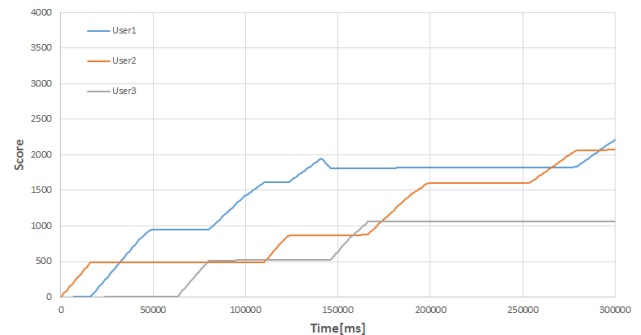


図 15 会話データ 2, 手順 (c) : 加分量補整式なし
Fig. 15 Conversation data 1, step (c)
Without correction formula of additional score.

時間の回復速度があがり、話せる時間が相対的に長くなったことを示している。よって時間回復式 (5) は正しく機能していると言える。しかし、UserA は余剰に発話することによって減点を受け、自らその得点を下げしており、持ち時間の回復速度をあげることでさらなる発話の奨励をうながすべきであるかどうかは疑問が残る。

会話データ 1 に加分量補整式 (8) を適用した結果を図 12 に、適用しなかった結果を図 13 に示す。会話データ 2 に加分量補整式 (8) を適用した結果を図 14 に、適用しなかった結果を図 15 に示す。会話データ 1 では、会話参加者全員が自分より点数が低い相手に話しかける機会があったので、図 12 の式 (8) を適用したもののほうが図 13 の適用しなかったものよりも全員得点量があがっているのが分かる。また、会話データ 2 では常に自分よりも点数の低い相手に話しかけていた UserA の得点が大きくあがっており、逆に常に点数が低かった UserC の得点量は適用時と非適用時で変わらないことが分かる。よって、加分量補整式 (8) も正しく機能していると言える。

以上の結果から、今回のシミュレーションでは考案した各採点モデルが正常に機能するということが分かった。ただし、あくまでシミュレーション上の動作として正常に機能するだけであり、この採点モデルが実際の多人数会話で

有意に機能するかどうかは分からない。時間回復式 (5) でも新たな疑問生じたため、今後さらに採点方式を煮詰めて実際の多人数による会話実験に臨む必要がある。

5. おわりに

本稿では、既存の研究システムの、その採点モデルにおける問題点を提示した上で、それを解消する方法を検討し、考案した新たな採点方式を用いてその効果を確認するべく、シミュレーションを行いその検証結果を述べた。その結果、考案した採点モデルはシミュレーションにおいて正しく機能することがわかったが、実際の多人数会話において有用性を発揮するかどうかはわからない。今後、この採点モデルを用いた会話実験を行って有意性を証明する必要があると感じた。

謝辞 本研究は、公益財団法人科学技術融合振興財団と、JSPS 科研費 15H0276 による助成を受けた。ここに厚くお礼申し上げる。

参考文献

- [1] 大島直樹, 岡澤航平, 本田裕昭, 岡田美智男. Tabletalkplus: 参加者の共同性や社会的なつながりを引きだすアーティファクトとその効果. ヒューマンインタフェース学会論文誌, Vol. 11, No. 1, pp. 105-114, 2009.

- [2] J. Terken and J. Sturm. Multimodal support for social dynamics in co-located meetings. *Personal and Ubiquitous Computing*, Vol. 14, No. 8, pp. 703-714, 2010.
- [3] G. Schiavo, A. Cappelletti, E. Mencarini, O. Stock, and M. Zancanaro. Overt or subtle? supporting group conversations with automatically targeted directives. In *Proc. of IUI '14*, pp. 225-234. ACM, 2014.
- [4] Hiroyuki Adachi, Seiko Myojin, Nobutaka Shimada, "ScoringTalk: A Tablet System Scoring and Visualizing Conversation for Balancing of Participation", *SIGGRAPH Asia 2015 Symposium On Mobile Graphics And Interactive Applications*, SA'15, pp. 9:1-9:5, ACM, 2015.