

対面同期型グループウェアの評価指標の提案

坂本 啓, 新井 克也, 桑名 栄二

NTT ソフトウェア研究所

近年グループウェアへの関心が高まっている。しかし実際に導入され成功したシステムは多くはない。この理由の 1 つとして、グループウェア導入の効果を明確に示せないということが挙げられる。

本稿では、対面同期型のグループウェアを対象とし、グループウェアが議論に与える影響に着目した。このとき議論の展開の変化をその効果と考え、Argument Structure に基づく評価指標を提案する。この指標の有効性を見るために少人数グループで行なわれる簡単なシステム設計会議を想定し、電子会議室を用いた場合と用いなかった場合について対照実験を行なった。今回用いた指標により、創造的議論に電子会議室が与える影響を明らかにすることができた。

Evaluation measures for face-to-face and synchronous groupware

Akira Sakamoto, Katsuya Arai, Eiji Kuwana

NTT Software Laboratories

NTT

3-9-11 Midori-cho Musashino-shi Tokyo 180 Japan

In recent year, groupware is being watched with keen interest. However, there are few successful systems that actually introduced. The various aspects on the reasons of groupware failures is indicated. One reason of them is that the effect of groupware is not clear.

In this paper, we developed argument structure-based measures for face-to-face groupware. In order to comfirm the abailability of this measures, we made comparison experiments in design meeting about simple system by small groups; (1) subjects used computer supported meeting environment (CSME) and (2) subjects did not use it. By using the measures, the effects of CSME for crative meetings was clear.

1 はじめに

近年グループウェアに関する関心が高まっている[3]。しかしその一方で実際に導入され成功したシステムはそれほど多くはない。この理由としては様々なものが考えられているが[6]、その1つとして、グループウェア導入によるメリットを明確に示すことが難しいという点があげられる。

従来グループウェアの効果は定性的な形で論じられていたが[10]、最近定量的な評価方法も提案され始めている[11]。定性的評価と定量的評価はお互いに補い合う関係にあり、グループウェアを評価していくためには、定量的に明らかになった特徴を定性的な評価結果で説明する、あるいは定性的な性質を定量的評価結果により裏付けするということが必要と思われる。また定量的評価が可能になると、グループウェア間で効果の比較を行なうことができ、グループウェアの効果を実際の利用を通じて実証的に示すことができる、という利点がある。

今回は特に対面同期型のグループウェアを対象とした。この研究の目的はグループウェアを定量的に評価するために、その効果を端的に示す評価指標を提案することであるが、一口にグループウェアといっても様々なものがある。その中で特に対面同期型を取り上げたのは、人間のコミュニケーション方法の最も基本的な形が対面同期的に行なわれる場合だと考えたためである。また今後、対面同期の場合を基準として、遠隔地間を結んだ協調活動の場合や、非同期の場合との差を調べることで、それぞれの場合を特徴付けていくことができると思われる。

対面型グループウェアについて提案する評価指標の有効性を調べるために、少人数グループで行なわれる簡単なシステムの設計会議を想定し、対照実験を行なうことと対面同期的なグループウェアが議論に対して与える影響を観察した結果について報告する。

2 設計プロセスと記述方法

2.1 方針

本稿では、創造性の要求される問題解決を少人数で構成されるグループにより協調的に行なう場合（例えば、設計会議でシステムの仕様を詰めていくような場合）について取り扱う。通常このような協調活動は、時間単位のショートレンジに一ヶ所に集まって行なわれ、議論が非常に重要な役割をはたしていると考えられる。したがってこのような協調活動を支援するグループウェアは、結局議論に対する支援を行なっていることになる。そこで本稿ではグループウェアを使った場合と使わなかった場合の議論の変化をグループウェアの効果と考えることにする。

以上のように考えるとこのようなグループウェアの効果を測るために

1. 議論を適切な形で記述する

2. 議論過程の変化が明確に見えるような特徴量を抽出する
3. グループウェアを使った場合と使わなかった場合で比較する

という手順を踏む必要がある。

そこで2.2節ではまず設計問題における議論がどう展開するか（議論過程）について考察し、2.3節で今回議論記述に用いた方法を紹介する。さらに、その記述方法から得られる特徴量について2.4節で述べる。

2.2 議論過程

少人数グループによりシステムの設計を行なう場合、議論は以下に示すような3つの段階を経て展開すると考えられる。

まず第一段階として、これから話し合う内容について話の糸口を見つけるために様々な項目を列挙していく、その項目に関係している情報が比較的無批判に提示される。ここで提示された情報から参加者間に今後の議論の前提事項が形成されていく。議論参加者全員に認られた前提事項を以下“共有知識”と呼ぶことにする。この共有知識から議論の焦点となる課題がこの段階で提示される。またその一方で、議論の方針を決める議論などメタな議論も行なわれる。第一段階とはこれから行なう議論の地盤作りの段階であると考えられる。

次に第二段階として、解決しなければならない課題について実際に決定していく段階がある。この段階で課題に対する様々な解決案が提示され、出された解決案に対して吟味が加えられる。議論の中で新たな問題点明らかになり、再びそれに対して解決案が提示される場合もある。第二段階は、第一段階で明らかになった課題が実際に解決され、参加者間で1つの合意が形成されていく段階と考えられる。

最後に第三段階として、第二段階で得られた結論を統合して全体的な構成を整える段階がある。それぞれの解決案は良いものでも統合すると全体として矛盾が生じたり冗長であったりする。この段階でそれらが検出され、各解決案が全体構成の中で再度吟味される。必要であれば第二段階に戻り、もう一度解決案を検討する。

このように議論過程は、始まりと終わりがそれぞれ一点に集まる木のような構造になっていると考えられる。ただしこの木は第一段階から第三段階へ整然と形成されていくのではない。時間経過に沿ってみると、第二段階から第一段階に戻ったり、第三段階から第二段階に戻ったりしながら全体として徐々に形成されていくと考えられる。

2.3 議論記述の方法

従来、社会心理学や社会学等コミュニケーションを主に取り扱う分野では議論を記述する試みは様々な形で行なわれている[2][5]。これらはそこで行なわれている

コミュニケーションの内容を詳細に分析する手法としては有効である。しかしこれらの手法は、議論をしている個人の方に興味の対象があり、ある発話が受け手に与える影響や1つの発言が議論全体の中で持つ意味などかなり詳細かつ意味に踏み込み過ぎる。またなかなか定量化しにくい。したがって今回の目的には使えない。

またソフトウェア設計プロセスを取り扱う分野では、設計時に行なわれた判断や理由を記録しておくために議論を明示的に構造化する手法 (Argument Structure : AS) が用いられている [1][4][8]。この手法には議論プロセス指向のものと議論構造指向のものの2つに大別される。プロセス指向の手法は途中経過を含めた議論中の発言を全て記述するため議論履歴を得ることができる。一方、構造指向の手法は議論の重要な部分のみを抽出し、論理的構造を得ることができる。ASモデルは単純なコーディングスキーマを用いて議論を記述していくので、インフォーマルな状況を量的に取り扱いやすい。そこで今回は議論全体の移動履歴を得る必要から、プロセス指向のASをベースとした記述方法を用いることにする。

具体的には、Olsonらが分析に使っていているカテゴリ [11] を参考に、表1. のような分類カテゴリを用いて議論を記述することにする。

表1: 分類カテゴリ

カテゴリ	内容
I	議論による意思決定が必要な課題・問題
A	あるIに対する選択肢の一つとなる解決案
C	意見・コメント
Cn	決定事項の確認
In	意思決定に際して必要な情報の提示
M	メタ議論
D	雑談
N	非発話的行為
T	議論をする際、使っているシステムに関する発言
O	上記のどのカテゴリにも入らない発言

このカテゴリを作成するに当たって以下の2点を考慮した。

(1) 情報提示を表す発話を記述するカテゴリの必要性：これは、議論においては情報提示による共有知識の増加は重要と考えられるためである。情報を議論の場に出すことにより、問題解決を行うとき暗黙の前提となっている事柄が明らかになり、共有知識が形成される。これが度重り、得られた共有知識は蓄積されて、問題解決を行なうための地盤が広がる。グループは蓄積された共有知識を用いて、問題点を明らかにし、また第二段階で話し合わなければならぬ課題を確定していく。共有知識が豊富になることにより、対象を多角的に見ることができ、問題解決に対するチェック項目がより多くなると考えられる。特に電子会議室といったシステムでは、議論過程全体をそのシステムを使って行なわれる情報提示に与える影響を積極

的に評価する必要があると思われる。

(2) 発話を伴わない図や文字・映像上の操作など非発話行為を記述するカテゴリの必要性：これは、非発話行為も発話と同様に記述することが重要であると考えられるためである。例えば、会議では説明の際、お互いの理解を円滑に進めるため、単なる発話だけでなく資料を配布したりホワイトボード上で図を描いたりする。特にコンピュータによる支援では視覚的効果の提供が多い。このようにグループで問題解決を行なう際、意味のある視覚的効果を発話と同様に記述していくことも必要であると考えられる。

分類・集計には発話の意味的な最小単位(チャンク)を用いる。発話及び発話と同程度の意味を持つ非発話行為をチャンクに分解し、各チャンクをカテゴリに分けて、各カテゴリ毎に集計する。さらに、得られた結果が比較可能な形になっている必要があるため、得られた個々のカテゴリに分類されたチャンク数の総チャンク数に対する割合を求めておく。

2.4 特徴量の抽出

前節で得られた結果から、グループウェアを使った場合と使わなかった場合の議論変化を比較し違いを見ていくために、適当な補正を行なう必要がある。これは個々の結果が完全に独立でなく、あるカテゴリの結果が別のカテゴリの結果に影響を与えている場合があるためである。例えば、解決案の数を表わすAをみるとことにより、ある問題に対してどれだけ多くの解決案が示されたか、どれだけ広く検討したかがわかるはずである。しかしAが同じ数であったとしても、その会議で意思決定が必要な問題Iが多い場合は検討した問題領域の広さは相対的に小さくなると考えられるし、反対にIが少ない場合は大きくなると考えられる。

ここでは以下に示すような6つの特徴量を抽出した。

- ・アイデア性 (A/I)：意思決定の必要な問題1つ当たりに提示された解決案の数。この値が大きいほうが課題に対して多くの解決策が提案されていると解釈される。
- ・確認性 (Cn/I)：Iに対する結論がグループとして採択されたことの確認がどれだけおこなっているか。この値が大きいほうが決定事項に対する各メンバの了解に誤解が生じにくくと解釈される。
- ・批判性 (C/A)：解決案に対して行われた批判やコメントの数。この値の大きいほうが各案に対する検討がよりよく行われていると解釈される。
- ・情報利用効率 (I/In)：提示された情報から引き出された意思決定の必要な課題数。この量が多いと提示された情報から効率的に課題が抽出されていることになる。なおこの値の逆数が、ある意思決定の必要な問題に対して提示された情報量である。
- ・メタ発話数 (M)：議論全体を通じてどれだけメタな議論を行っているかを表す量。この値が大きいほうが議論全体の見直し回数が多く、より視野の広い立場から議論

が行われていると解釈される。

- ・提示情報量 (In) : 議論全体を通じてどれだけ議論に関係する情報が提示されているかを表す量。この値が大きい方が提示された関係情報が多く、意思決定に際してベースとなる共有知識が多いと解釈される。

3 実験

3.1 実験目的

前節で示したカテゴリ分類基準の議論記述への有効性を検証し、得られる特徴量がグループウェアの評価に使えるか確認するために、電子会議環境 (Computer Supported Meeting Environment : CSME) [10][9][7] を用いた対照実験を行った。

3.2 実験方法

3.2.1 対象と実験環境

入社 1 年目の人たちを対象に、3 人のグループを 2 つ作った。一方は通常の会議室で、もう一方は CSME として COGENT[7] でそれぞれ実験を行なった。以下、前者を G_u 、後者を G_s と表記する。

COGENT 利用者は、一人に一台づつコンピュータを割り当てられ、自由に使うことができる。また各コンピュータはネットワークに接続されているので環境内から任意のコンピュータへアクセスすることができる。また映像レベルで情報を共有しているため、同じコンピュータの画面を参加者全員で見ることができる。大型のホワイトボードが設置されており、これを共同作業空間として使うこともできる。 G_s ではコンピュータの利用を前提としたので、事前にアンケート調査を行ない、各自が日頃使っている環境を用意した。

3.2.2 実験手順

実験は G_s, G_u 共に、約 1 時間の会議を 2 回、合計 2 時間行なった。初回の実験前に、実験趣旨の説明と課題の提示、 G_s の場合は更に COGENT の説明と試用を行なっている。

実験に当たって、両グループに同じ課題を与え議論方法に以下のようない制約を課した。 G_u は完全にフリーで特に制約を設けない方法で議論を行なった。 G_s は被験者の中から書記を定め、議論の進行状況をその場で記録した。この記録は COGENT の画像共有機能を用いて被験者全員が逐次参照できる。

3.2.3 課題

“Automatic Post Office (APO)” の設計は、設計プロセスを分析するような場合には良く使われる問題である。そこで APO の設計とそのおおまかな仕様の作成を課題とした。APO とは郵便局で提供される様々なサービスを 1 つの機械に集約し自動化したものである。

被験者は APO が提供するサービスやその形状などを自由に設定して良く、また設計仕様のフォーマットも自由である。ただし、最低限以下の項目を含んでいるように指示した。

- APO により提供されるサービス
- APO の完成イメージ (外観／システムの全体構成など)
- APO 導入によりどのような点よくなるのか

また各会議について議事録を作成してもらった。議事録作成に当たっては、 G_u の場合各会議の終了後に担当者が行ない、 G_s の場合は会議中に作成した記録を会議終了後若干修正して当てる。

3.3 結果

各実験の様子は VTR に記録した。分析に当たってこの VTR をテキストに落し、さらに発言内容に基づいてカテゴリ分類を行った。

G_u, G_s それぞれ先のカテゴリに分類した結果を、第一回目、第二回目および全体についてまとめ、表 2, 表 3, 表 4 に示す。N についてはサブカテゴリ毎に分配し、集計してある。N のみを集めた結果を表 4 に示す。

各カテゴリに分類されたチャック数の全チャック数に対する割合をパーセントで示し、括弧内に実際のチャック数も示してある。

総チャック数は G_s が 2454 に対して、 G_u が約半分の 1304 となっている。このように総チャック数に差が生じた原因としては、グループの特性として G_u の方が考え方やすく沈黙しがちなグループであったということ、 G_u では会話だけでなくキーボードによる入力操作や図形の編集操作など純粋に時間をとる操作が多かったこと、および手のあいている被験者が入力など操作をしている様子を見守る場面が多かったこと、などが考えられる。

表 2: 実験結果 G_u (単位:%)

カテゴリ	1 回目	2 回目	全体
I	7.1 (73)	4.5 (64)	5.6 (137)
A	13.6 (140)	8.9 (127)	10.9 (267)
C	15.6 (161)	8.2 (117)	11.3 (278)
Cn	4.7 (48)	1.3 (19)	2.7 (67)
In	26.2 (270)	41.8 (595)	35.2 (86)
M	6.9 (71)	3.4 (48)	4.8 (119)
D	26.1 (269)	31.8 (452)	29.4 (721)
N	0.0 -	0.0 -	0.0 -
T	0.0 -	0.0 -	0.0 -
合計	100 (1032)	100 (1422)	100 (2454)

() はチャック数

また分類結果を基に各特徴量について実際に得られた値を表 6, 表 5, 表 4 に示す。

表 3: 実験結果 G_s (単位:%)

カテゴリ	1回目	2回目	全体
I	6.5 (47)	4.4 (26)	5.5 (73)
Al	9.3 (67)	7.7 (45)	8.6 (112)
C	9.7 (70)	4.6 (27)	7.5 (97)
Cn	1.3 (52)	3.8 (22)	1.6 (74)
In	13.4 (96)	18.3 (107)	15.6 (203)
M	14.3 (103)	10.8 (63)	12.3 (166)
D	22.5 (162)	27.9 (163)	25.0 (325)
T	17.0 (122)	22.5 (132)	19.1 (254)
合計	100 (719)	100 (585)	100 (1304)

() はチャック数

表 4: 実験結果(非発話行為) G_s (単位:%)

カテゴリ	1回目	2回目	全体
N	7.5 (54)	2.1 (12)	4.9 (66)
-I	0.3 (2)	0.0 (0)	0.2 (2)
-A	0.0 (0)	0.0 (0)	0.0 (0)
-C	0.0 (0)	0.0 (0)	0.0 (0)
-Cn	6.1 (44)	1.5 (9)	4.1 (53)
-In	0.0 (0)	0.0 (0)	0.0 (0)
-M	0.6 (4)	0.3 (2)	0.3 (6)
-D	0.0 (0)	0.0 (0)	0.0 (0)
-T	0.5 (4)	0.2 (1)	0.4 (5)
-O	0.0 (0)	0.0 (0)	0.0 (0)

() はチャック数

4 結果の解釈および考察

4.1 結果の解釈

得られた特徴量に基づいて、各会議で行なわれた議論の分析を行なった。ここでは第一回目・第二回目を通して各特徴量を 2 つのグループについて比較した結果と、各グループ毎に第一回目と第二回目について各特徴量の時間的な推移を見た結果を示す。

4.1.1 議論過程の相違

各特徴量を G_u, G_s について比較してみると、特に確認性、情報利用効率、メタ発話数の 3 つの特徴量につ

表 5: 分類結果から得られる特徴量 G_u

特徴量	第一回目	第二回目	全体
アイデア性(A/I)	1.918	1.984	1.949
確認性(Cn/I)	0.658	0.297	0.489
批判性(C/A)	1.150	0.921	1.041
情報利用効率(I/In)	0.270	0.108	0.158
メタ発話数(M)	0.069	0.034	0.048
提示情報量(In)	0.262	0.418	0.352

表 6: 分類結果から得られる特徴量 G_s

特徴量	第一回目	第二回目	全体
アイデア性(A/I)	1.431	1.750	1.536
確認性(Cn/I)	1.123	0.864	1.018
批判性(C/A)	1.043	0.597	0.860
情報利用効率(I/In)	0.485	0.240	0.359
メタ発話数(M)	0.143	0.108	0.123
提示情報量(In)	0.134	0.183	0.156

いては G_s が大きくなっていることがわかる。

G_s には議論の模様をその場で記述していくように指示した。被験者たちは記述された過去の議論をその場で見ながら議論していくことで、まず骨格が定まりそこに肉付けしていくような形で議論が進展した。

議論の骨格を組み立てていく上で、何をどのように話し合っていくのかというメタな議論が行なわれる必要がある。この結果メタ発話数が大きくなっているのではないかと考えられる。

また議論の骨格がすでにわかつておりそれが目の前に見せられているので、今話し合っていることの位置付けがイメージし易くなっている。情報利用効率が大きくなっているのは、議論中に必要とされている情報が捉え易くなっているため、今行なわれている議論に関係がなく何に使われるかわからないような情報が出にくくなるためなのではないかと思われる。

さらにあやふやな結論や理解は文章化できない。確認度が大きくなっているのは、話し合いによって得られた結論を文章にする際、曖昧な部分を取り去る必要があったためと考えられる。

4.1.2 特徴量の時間的推移

両グループとも第一回目には機能仕様、つまり APO が提供するサービスの選定が行なわれた。また第二回目には詳細設計、つまり APO が提供する各種サービスの実現形態について検討が行なわれた。

双方とも程度の違いはあるものの、確認度、情報利用効率、批判性、メタ発話数は減少し、発散性、提示情報が増加している。これは機能設計および詳細設計の局面の違いを反映していると考えられる。

詳細設計を行なう際、確認度および批判性が小さくなるのは、話の流れの中で必然的に導かれてくるような案や結論が多くなってくるからであると考えられる。

また情報利用効率が減少するのは、具体的な機構の実現可能性を検討する際、既に実現されている似た機構を探してきて例示していくことが多くなるためであると考えられる。

さらにメタ発話数が小さくなるのは、詳細設計では、機能設計で決められた枠組に変更を加えずその中で具体化していくからであり、発散性が増加するのは話が詳細かつ具体的になっていくため、APO の機能を実現可能

性を検討していく際、取り得る選択肢が多岐に渡るためであると考えられる。

4.2 考察

分析結果を踏まえ、今回用いた特徴量がグループウェア評価の指標として有効かどうかについて考察する。

前節で示したように、今回用いた6つの特徴量により COGENT の利用による議論過程の変化あるいは議論段階の変遷による議論過程の変化を捉えることができた。

しかし現段階では、もともとのグループ特性の違いによる議論過程の相違とグループウェア導入により生じた議論過程の変化を明確に分離できていない。純粹にグループウェアの効果を見るためには、より多くの実験事例を集めることにより統計的な特徴をつかんでいく必要があると思われる。これは今後の課題である。

その他、今回用いた特徴量の問題点として以下のようないわゆるものが挙げられる。

- カテゴリ分類に評価者の主観が混入しやすい。
カテゴリ分類を行なう際、個々の判断は結局評価者の主観によって行なわれる。カテゴリ分類による評価結果に客觀性を持たせるためには、今回用いたカテゴリによる分類・評価結果が、評価者よりどの程度違いが生じるのか見積もっておく必要があると思われる。
- 並列的に行なわれるコミュニケーションの記述が難しい。
今回行なった実験では、 G_u については一度に1つの話題について議論が行なわれることが多かったが、 G_s については個々人がそれぞれ役割を持って独立に作業しながら議論が続いている場面が見られた。うまく役割分担が行なわれた場合には非常に密度の高いコミュニケーションが行なわれる。このような並列的コミュニケーションを記述するには今回用いた議論記述は不向きでありさらなる拡張を必要とする。

5 おわりに

本稿では、対面型会議を支援するグループウェアである CSME を用いた対照実験を行ない、結果の分析を通じてその効果をうまく表現できる指標について考察した。

その結果、議論の際 CSME を用いることにより、トップダウン的に議論が進んでいること、特に確認性、情報利用効率、メタ発話数の3つの特徴量が大きくなっていることがわかった。また機能設計と詳細設計の局面の相違により、確認度、情報利用効率、批判性、メタ発話数は減少し、発散性、提示情報は増加していることがわかった。

グループウェアの評価は様々な要因が絡み合うため一筋縄ではいかない。今回は対比による差分を見る上でグループウェアの効果を見ようとした。しかし実験サンプルが今のところ一組しかないので各グループの持っている特性とグループウェアによる効果が完全に分離されていないなどまだまだ未解決の問題が残されている。今後同様の実験を積み重ねていくことにより、今回得られた知見を検証し、前節で述べた問題点・課題に取り組んでいくつもりである。

謝辞

実験に協力して下さった皆様にこの場を借りてお礼を言わせて頂きたいと思います。また貴重な意見を頂いた島 健一主任研究員および浜田 洋 主幹研究員に心より感謝致します。さらに本稿をまとめるにあたり助力して下さった中村リーダ始め1Pの皆様に感謝致します。

参考文献

- [1] Carroll, C. et al., *Deliberated Evolution : Stalking the View Matcher in Design Space*, Human-Computer Interaction, Vol.6, No.3 & 4, pp.281-318, 1991.
- [2] Cappella, J.N., *Interpersonal communication : definitions and fundamental questions*, In Berger, C.R. and Chaffee, S.H. (Eds.), *Handbook of Communication Science*, pp.184-238, Newbury Park, CA : AGE Publications, Inc,
- [3] Ellis, C. et al., *Groupware : Some Issue and Experience*, CACM, Vol.34, No.1, pp.38-58, 1991.
- [4] Fisher, G., et al., "Making argumentation serve design", Human-Computer Interaction, Vol.6, No.3 & 4, pp.393-420, 1991.
- [5] Frey, L.R., *Interpreting Communication Research : A Case Study Approach*, Englewood Cliffs, N.J:Prentice-Hall, Inc, 1992.
- [6] Grudin, J., *Groupware and Social Dynamics: Eight Challenges for Developers*, CACM, Vol.37, No.1, pp.93-105, 1994.
- [7] 桑名栄二, et al., 協調作業支援システム (COGENT), NTT R & D, Vol.43, No.9, pp.963-971, 1994.
- [8] Maclean, A. et al., *Questions, options, and criteria: Elements of design space analysis*, Human-Computer Interaction, 1991.
- [9] Mantei, M., *Capturing the capture lab concepts : A case study in the design of computer supported meeting environments*, Proc. of CSCW '88, Portland, OR, 1988.
- [10] Nunamaker, J.F., et al., *Electronic Meeting Systems to Support Group Work*, CACM, Vol.34, No.7, pp.40-61, 1991.
- [11] Olson, J.S. et al., *How a Group-Editor Changes the Character of a Design Meeting as well as its Outcome*, Proc. of CSCW '92, Nov, 1992.