

InTREND：ユーザの探索行為と振り返り行為に着目したデータ分析支援システム

松下 光 範^{†1}

探索的データ分析は、データの明確な利用方針をあらかじめ持っていない分析者が、試行錯誤を繰り返しながら興味深い特徴や有意義な知見を見つけ出す際の分析手法の1つである。我々はこのような分析を行うユーザを支援するために、自然言語を用いて対話的にデータ分析を行うと同時に、その過程で表出されたオブジェクトやアノテーションを外在化し、振り返りの際の手がかりとすることで知見の整理や解釈に貢献することを目的としたシステム InTREND (Interactive Tool for Reflective Exploration through Natural Discourse) を提案する。このシステムを用いたユーザ観察を実施した結果、(1) 分析方針の違いが履歴のパターンやウィンドウの使用方に影響すること、(2) 被験者が振り返りの際の手がかりとして、付箋機能のような直接的な外在化だけでなくオブジェクトの位置移動やサイズ変更といった暗示的外在化を用いていること、(3) バックトラックを行う際、直接操作と自然言語の両方で操作できる冗長性が求められていることが明らかになった。

InTREND: A System for Supporting Data Analysis by Focusing on Exploration and Reflection Acts of Users

MITSUNORI MATSUSHITA^{†1}

Exploratory data analysis is a methodology that a user tries to make beneficial discoveries through viewing data from multiple perspectives and examining the data with a variety of emerging hypotheses by repeatedly interacting with the data. During the exploration, the user not only executes exploration but also reflects on his/her past exploration process in order to consolidate his/her findings or to organize the subsequent exploration strategy. To encourage the user's exploration and reflection activities, we have developed a system named InTREND. This paper describes a user study to investigate how a user conducts an exploratory data analysis with the system. Our user study reveals that (1) historical pattern of a participant's exploration varies depending on a style of the participant's analytic procedure, (2) participants often utilize allusive externalization (e.g., positions and sizes of objects modified by themselves) along

with direct externalization (e.g., annotations attached onto History Map), as clues for reflecting their exploration acts, and (3) redundant operability is desired for backtracking.

1. はじめに

我々は、蓄積された大量の情報の中から利用者（ユーザ）が自らのかかえる問題の解決やタスクの遂行に役立つ情報を分析する行為の支援技術について研究を行っている。特に、このようなユーザの分析行為の方法論の1つである探索的データ分析 (exploratory data analysis) に注目し、これを円滑に行うための支援システムの実現を目指している。

探索的データ分析とは、ユーザが解決すべき問題 (e.g. 「この地区に新しい店舗を開店すべきか?」) については漠然と認識しているものの、どのようなデータをどのように利用すればその問題解決に有益な知見を得られるか (e.g. 「同業他社の収益データを店舗新設の論拠に使う」) が判明しておらず、試行錯誤を繰り返してそれを見つけ出すという分析方法である⁵⁾。

この探索的データ分析を支援するシステムとして、我々は InTREND (an Interactive Tool for Reflective Exploration through Natural Discourse) の研究開発を進めている^{8),9)}。InTREND は、(1) 自然言語で表現されたユーザ質問を解釈する、(2) その質問と、それによって得られた検索結果に基づいてグラフを構成する、(3) そのグラフをアニメーションを用いて表示する、という枠組みに基づいてユーザの分析行為の支援を試みたシステムである。本研究では、ユーザの分析行為をより広範に支援することを目的としてこのシステムに拡張を施し、過去の探索過程を概観することで知見の整理を行ったり、ある時点から異なる観点で再度探索を行ったりできるようにした。

本稿では、拡張を施した InTREND の機能について説明するとともに、ユーザ観察を通じてそのデザインの検証と考察を行う。

2. 探索的データ分析におけるインタラクション

探索的データ分析は単なる one-shot の作業の集まりではなく、同じデータを様々な視点

^{†1} 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories, NTT Corporation

から多角的に眺める, 仮説を立てながらデータを検索するなどの試行錯誤を繰り返しつつ行う高度に知的な思考活動¹⁸⁾である. この過程においては, あらかじめ明確なゴールが定まっておらず, それまでの探索過程を俯瞰して再考したり, メタな視点から解釈したりする必要が生じることもある¹²⁾. したがって, 探索的データ分析を計算機システムで効果的に支援するためには, 探索行為 (exploration) だけでなく分析者 (ユーザ) が自らの探索の過程を顧みて知見を整理するという振り返り行為 (reflection) についても考慮しなくてはならない.

探索行為はユーザが試行錯誤しながら有益な知見を外在化することが主眼であるのに対して, 振り返り行為は外在化された知見を整理し解釈することが主眼であるため, その各々に適切なインタラクションは必然的に異なる. その一方で, これら2つの行為は互いに独立した行為ではなく深く関連した行為であり, ユーザは状況に応じてそれらの行為を使い分けながら分析を進めていくことが想定される. したがって, これら2つの行為をシームレスに行えるようにするには, 各々の行為に適したインタラクションをデザインするとともに, それを支援する機能の間でどのような情報やオブジェクトを共用すべきかについて検討しなくてはならない. 以下では, このような観点の下で各々の行為で想定されるインタラクションについて述べる.

2.1 探索行為におけるインタラクション

計算機を用いて探索的データ分析を行う場合, ユーザは, (1) はじめにデータから何を見出すかについての漠然とした考えに基づき, その考えの下で最初の質問を計算機に与える, (2) 計算機がその質問に基づいて検索・提示した結果を受け取る, (3) 計算機から提示された結果を見て, 必要なデータについての理解を深め, 新たな質問を計算機に与える, という行為を繰り返す¹⁶⁾. この行為の繰返しにより, ユーザは徐々に問題解決や意思決定にとって有用な情報を収集していく. 図1にユーザと計算機とのやりとりのモデルを示す. 我々はこの枠組みに基づくシステムとして, ユーザからの質問を受け付け, その質問に基づいてデータを検索し, その検索結果をグラフを用いてユーザに回答する可視化システムの実現を目指している.

本研究では, このシステムへの入力手段として自然言語を採用した. 従来の情報可視化システムでは, 画面上に表示されているオブジェクトを選択する直接操作 (direct manipulation) がよく用いられている^{6), 14)}. 直接操作は, 簡便な操作でシステムに要求を伝えることができるため, 多くのユーザにとって非常に敷居が低い方法であると考えられている. しかし, 探索的にデータを分析する場合, ユーザの関心は描かれているグラフをどのように操作し

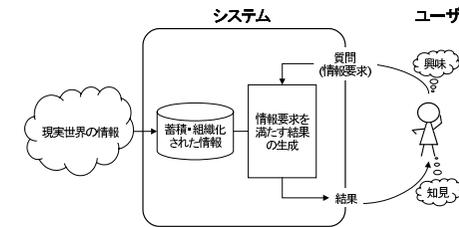


図1 探索過程におけるユーザと計算機とのやりとり

Fig. 1 Interaction model between a user and a system in exploration phase.

変化させるかではなく, どのデータをどのような視点から分析するかという点におかれているため, 直接操作による探索行為ではユーザ自身が情報要求をオブジェクトの操作に具体化しなくてはならない. これは, 本来ユーザが求めている「何を描画したいか (what to draw)」という要求を「どのように描画すべきか (how to draw)」という操作に変換する作業がユーザの負担として残されていることを意味する. 自然言語を用いた入力手法はユーザに柔軟な表現力を提供する手段であり³⁾, この問題の解決に寄与すると考えている.

また本研究では, 検索結果を出力する方法として統計グラフによる情報提示を採用した. 棒グラフや折線グラフなどの統計グラフは, 蓄積された莫大な観測データを分析して傾向や特徴を見出す際の有効な手段である¹⁹⁾. 提案システムはコンテキストや質問から推測されるユーザの関心を考慮し, それらを反映したグラフを描画することでユーザの理解を助ける⁸⁾.

図1に示したように, 探索の過程でユーザと計算機との間でやりとりされるのは, ユーザが表出 (articulate) した情報要求である質問文と, その回答として計算機が提示するグラフである. しかし, あとから探索行為を振り返る際に手がかりとなるのはそれだけではない. 探索の過程で, ユーザはシステムが提示したグラフを見ることで知見を得たり, 新たな興味を抱いたりする. このような, 探索の途中でユーザが得た知見や興味も振り返りの際に有効な手がかりになる. 我々はこの点を考慮して, ユーザが知見や興味を外在化 (externalize) できるように, グラフに対して付箋の形式でアノテーションを付与する機能を設けた.

2.2 振り返り行為におけるインタラクション

上述したように, 探索的データ分析は試行錯誤によって新たな知見や結果を得る手法であるため, その探索の過程は大局的には課題のゴールに向かって進んでいるが, 局所的に見れば必ずしも線条ではなく, 様々な分岐が存在する場合が一般的である. このようにして得ら

れた一連の探索結果から、意思決定に役立つ知識を取り出すためには、探索過程の枝葉や冗長さを削除したり、探索過程を精緻化するなどの整理が必要となる。そのため、それまでの探索過程で外在化されたオブジェクト（質問、グラフ、アノテーション）を手がかりとして自らの探索行為を振り返るというユーザ行為が想定される。

本研究では、これらの外在化されたオブジェクトを履歴として二次元空間に配置することで可視化し、ユーザがそれらのオブジェクトに対して直接操作でインタラクションできるようにすることで、ユーザの振り返り行為の支援を試みた。振り返りの際のインタラクション手段として自然言語ではなく直接操作を用いた理由は、振り返り行為の主眼がグラフや質問などの新たなオブジェクトの創出ではなく、探索の過程で創出されたそれらのオブジェクトの関係を整理し概観することであるため、直接操作のほうが直観性に優れると考えたためである。

Shneiderman の Visualization Seeking Mantra¹⁵⁾によれば、このような可視化表現を利用する際のユーザの行動は、“Overview first, zoom and filter, then details on demand”である。したがって、それまでの探索履歴を一覧し、それをズームイン/アウトしたり、興味に応じて不要なオブジェクトを削除/省略したりできる機能が必要になる。さらに、探索行為と振り返り行為は必ずしも1方向ではなく、ユーザが振り返りの途中で再度探索を行う場合も想定される。そのため、ユーザが指定した探索過程のある地点から異なる観点の下で探索を再開できる機能も必要である。

また、外在化された表現をユーザが振り返り整理する際には、探索する対象そのものではなく、探索過程の構造や関係に関するメタな視点からのコメント（メタコメント）を利用することが指摘されている¹⁷⁾。本研究でもこの点に鑑みてメタコメントを付与する機能を設ける。メタコメントは内容的には探索過程で利用するアノテーションとは異なるが行為自体は同一であり、それらのインタラクション手法を共通化するほうがユーザにとって理解しやすいものと考えられるので、本研究ではこれらを「付箋を貼る」というメタファを共通に用いて実現した。

3. 実 装

我々は、探索的データ分析を行うユーザを支援するシステムとして InTREND (an Interactive Tool for Reflective Exploration through Natural Discourse) を提案している⁸⁾。InTREND は、(1) 自然言語で表現されたユーザ質問を解釈する、(2) その質問と、それによって得られた検索結果を用いてグラフを構成する、(3) そのグラフをアニメーションを



図 2 Query Window の例
Fig. 2 Query Window example.

用いて表示する、という枠組みによりこのような循環的なやりとりを支援するシステムである。このとき、過去のやりとりをコンテキストとして保持することにより、ユーザと計算機とのやりとりの円滑化を図っている。我々はこの InTREND システムに対して、振り返りを容易に行えるように履歴を記録し再利用できる機能の拡張を施した。

InTREND のユーザインタフェースはユーザ要求を入力するウィンドウ (Query Window)、回答となるグラフを表示する出力ウィンドウ (Graph Viewer)、履歴を表示するウィンドウ (History Map) の 3 つのウィンドウから構成される。探索の場面では主に Query Window と Graph Viewer が用いられ、振り返りの場面では History Map と Graph Viewer が用いられることを想定してデザインを行った。

3.1 Query Window

Query Window はユーザが質問をテキスト形式で入力するためのウィンドウである (図 2)。InTREND は、ユーザが質問を表出しやすいように前の質問のコンテキストを利用した質問を許している。たとえば、ユーザが「2000 年から 2001 年までの近畿地方の降水量は？」と尋ね、システムがグラフで回答したとする。このグラフを眺めることにより、ユーザが近畿地方の県ごとの降水量に興味を持った場合、「県ごとに知りたい」と質問するほうが、「2000 年から 2001 年の近畿地方の県ごとの降水量は？」と質問するよりもユーザにとって自然であるためである。このような省略をとる断片的質問を解釈するために、InTREND はパターンマッチング処理によって自然言語で表現されたユーザ質問を形式表現に変換し、直前までのユーザ要求との構文的/意味的類似性を用いてその形式表現を書き換える方式²⁾を応用している。この処理の詳細は文献 7) を参照されたい。

3.2 Graph Viewer

Graph Viewer はユーザの質問に応じてグラフを提示するウィンドウである (図 3)。

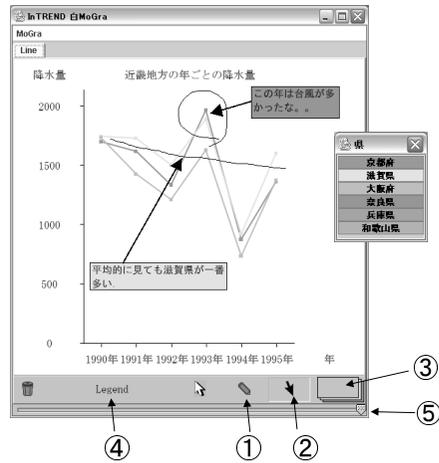


図 3 Graph Viewer の例
Fig. 3 Graph Viewer example.

InTREND システムは Query Window を介して入力されたユーザ質問を解釈して蓄積されたデータを検索・集約し、その質問の回答となるグラフを Graph Viewer に表示する。この際、前回のやりとりの結果として描画したグラフを考慮し、適切なアニメーションを生成してグラフを描き換える。

Graph Viewer では、ユーザが描かれているグラフに対してそれを眺めることで思いついた興味や知見などをアノテーションとして付与できるように自由曲線 (図 3 中の ①)、矢印 (②)、付箋 (③) の 3 つのツールを用意した。これらは Graph Viewer の右下に配置されている。付箋をグラフ上に貼り付ける場合は、グラフ上の任意の位置までドラッグして離せばよい。この付箋は 3 色 (緑、黄色、橙色) あり、クリックするたびにその色が変わる仕様になっている。

グラフの凡例 (legend) は Legend ボタン (④) をクリックすることにより現れる。凡例とグラフは連携しており、凡例のいずれかをクリックすると対応するグラフオブジェクトが強調表示されるようになっている。

Graph Viewer の下部にあるのはグラフ描き換えの際のアニメーションを操作するためのトラック (⑤) で、前に描かれていたグラフとの関係を再度確認したい場合に利用できる。

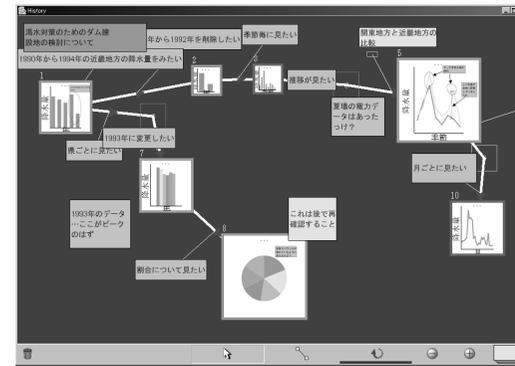


図 4 History Map の例
Fig. 4 History Map example.

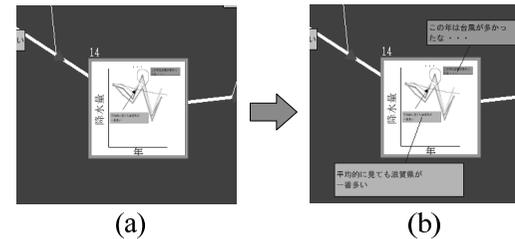
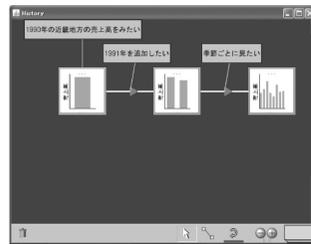


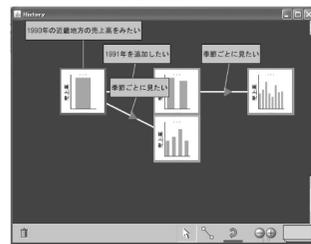
図 5 サムネイル化されたグラフ。(a) グラフタイトルと軸スケールが省略されている,(b) 付与されたアノテーションがクリックによって拡大されている
Fig. 5 A subnailed graph. (a) Graph titles and scales of axes are omitted. (b) Attached annotation appears outside of thumbnail by clicking on it.

3.3 History Map

History Map はユーザの探索履歴を表示するウィンドウである (図 4)。InTREND は Query Window からユーザが入力した自然言語質問文とシステムがその回答として Graph Viewer 上に表示したグラフを History Map 上に逐次サムネイル化して追加していく。たとえば、図 5 (a) は図 3 で Graph Viewer に表示されているグラフがサムネイル化されたものである。また、質問文は灰色の付箋としてオブジェクト化され、そのグラフの左側の位置に表示される。状況によっては History Map 上に多数のサムネイルが表示されることが想定されるので、視認性の向上のためにサムネイル化されたグラフにはグラフのタイトルや軸の目盛といった詳細な情報は省略されている。History Map と Graph Viewer は連動



(a)



(b)

図 6 生成された分岐の例

Fig. 6 An example of a branch generated.

しており、省略された情報を参照したい場合には History Map 上でサムネイルをクリックすることで元のグラフを Graph Viewer 上に表示することができる。この点に注目すると、History Map と Graph Viewer の関係は Overview+Detail¹¹⁾ と見なすことができる。

これらのサムネイル化されたオブジェクトは、探索の進行にともなって逐次 History Map 上に追加される。その過程で新たに創出された質問文やグラフなどのオブジェクトは、それまでアクティブになっていたサムネイルの右側にリンクをともなって追加されていく。二次元平面上の位置そのものに意味はなく、History Map 上に作成・配置されたサムネイルをユーザがポイントしドラッグすることで自由に場所を移動させることが可能である。さらに、サムネイルのサイズ変更やリンクの追加や削除、付箋を用いたメタコメントの付与、ズームイン・ズームアウトが可能である。

2.1 節で述べたように、ユーザが振り返りを行う際に利用可能な外在化表現には、これらのオブジェクトの他に探索の過程でユーザによってグラフ上に貼り付けられたアノテーションがある。これらは、サムネイルが表示された段階では図 5 (a) のようにグラフのどこに貼

られたかということだけ分かるようになっていたが、クリックすることで図 5 (b) のようにサムネイルの外側に配置され、アノテーションの内容が確認できるようになる。

また、過去の探索過程からあるグラフをクリックしてアクティブにし、そこから新たな質問を入力することで探索を行う場合、選択したグラフから自動的に分岐が生成される。このとき、すでに描かれている枝と重ならないように右斜め下に配置される仕組みになっている。図 6 は分岐が生成される様子を示している。図 6 (a) に表示されているグラフのうち、左のグラフをクリックしてアクティブにし、新たな質問文を入力すると図 6 (b) のように自動的に分岐が生成される。

4. ユーザ観察

探索的データ分析は個人の分析行為に関わる方法論であり、たとえ同じデータを分析対象にした場合であっても、すべての分析者が同じ結論にたどり着くことは期待できない。それゆえ、その行為を支援するシステムの評価においては、タスクの達成度や正答率といった指標を用いることができない。そこで本研究では、ユーザが InTREND を用いてどのように分析を行うかについてユーザ観察を行い、我々が意図した用いられ方との差異を調べることにした。

4.1 実験手続き

研究・開発に携わる 20 代～30 代の男性 5 人を被験者にしてユーザ観察を行った。実験に先だって、操作の習熟のために一定期間システムを利用してもらい、システムが解釈可能な自然言語表現と、システムが提供するすべての機能の操作方法について理解してもらった。実験では、まず課題の教示を行い、その課題に基づいた分析を InTREND を用いて行ってもらった。実験では各被験者は課題分析フェーズと説明生成フェーズを順に行い、最後に方針説明を行うよう促された。本実験で対象としたデータは、架空の企業の 1986 年 1 月から 1995 年 12 月までの県ごとの月別売上高である。

課題分析フェーズではデータの特徴を理解して知見を得るために、振り返り行為よりも探索行為が多く観察されると推測されるのに対し、説明生成フェーズでは他者に自分の見つけた知見を他者に論理立てて説明するために、探索の過程を概観したり整理したりする振り返り行為がより顕著に観察されると推測される。したがって、各々のフェーズでの振舞いの違いに着目することで、被験者の行う探索行為と振り返り行為のインタラクションの特徴が観察できると考えたため、このような手順で実験を行うこととした。

被験者に教示した課題スクリプトを以下に示す。

まず、今年は 1995 年だと思ってください。あなたは NTT 商事の経営企画担当者です。NTT 商事は各県に営業所があって、各営業所から毎月売上高の報告がされています。あなたはこの春より、中国地方を担当することになりました。あなたの目的は、過去 10 年間のデータを用いて中国地方の特徴を理解し、今後の経営方針を見出すことにあります。いまから InTREND を用いて分析を行ってください。その後、分析結果からどのような方針を立てたかについて説明してください。なお、開始から 10 分、15 分にベルが鳴りますが、考えがまとまるまで分析を継続していただいて結構です。また、分析終了後に報告を行っていただきますが、そのための整理の時間は別に設けます。

この教示の後、被験者は分析開始を促された。課題分析フェーズでは、10 分と 15 分でベルが鳴らされたが、被験者は自分の考えがまとまるまで分析を継続してよいこととした。これは、定まったゴールを持つような課題ではなく、各被験者が一定の結論に至ったと自覚することがゴールとなるような課題のためである。ただし、時間の制約をまったく与えずに実験を行った場合、課題終了までの時間が被験者によって大幅に異なり、実験結果の解釈に影響することが懸念される。そこで、ベルを鳴らすことで被験者に目安を与えることにした。ベルを鳴らすタイミングは、筆者が事前に行った予備調査の結果から決定した。また、この実験では、大量の情報を表示した際の視認性を確保するため 50 インチの PDP ディスプレイが用いられた。

課題分析フェーズが終了した後、被験者には報告する内容を整理するための時間が与えられた(説明生成フェーズ)。この説明生成フェーズには制限時間は設けなかった。この説明生成フェーズが終了した後、被験者に自らが策定した経営指針についての報告(方針説明)を行ってもらった。また、実験終了後に各被験者に聞き取り調査を行った。

4.2 実験結果

実験中のディスプレイの様子、および実験中の被験者の様子は各々 VTR で録画された。その書き起こし結果に基づいて、各被験者が課題分析フェーズと説明生成フェーズに費した時間、操作数を表 1 に示す。ここでは、マウス操作の場合はクリックないしドラッグを 1 操作、テキスト入力の場合は連続した文章入力を 1 操作と数えた。また、表 2 に、各フェーズで生成されたオブジェクトの数ならびに観測された振舞いの回数を示す。なお、表 1 および表 2 において、A 列は課題分析フェーズ、B 列は説明生成フェーズを示している。

表 2 のオブジェクトの項は各被験者が実験の際に外在化したオブジェクトの数を示したもので、「ユーザ質問」は被験者が Query Window に入力した質問の数(= History Map 上

表 1 課題分析フェーズおよび説明生成フェーズにおける被験者の操作時間と操作数(A 列は課題分析フェーズ、B 列は説明生成フェーズ)

Table 1 Each participant's operation in each phase (Column A: task analysis phase, column B: explanation generation phase).

	被験者 1		被験者 2		被験者 3		被験者 4		被験者 5	
	A	B	A	B	A	B	A	B	A	B
時間(秒)	1,570	277	1,204	471	1,519	332	1,867	248	2,162	350
総操作数 (内訳)	301	34	205	140	160	98	193	66	421	113
Input Window	40	0	19	0	47	0	45	3	82	2
Graph Viewer	96	0	98	17	47	1	67	20	162	12
凡例	115	0	38	3	42	0	15	2	13	1
History Map	50	34	50	120	24	97	66	41	161	98
その他	0	0	0	0	0	0	0	0	3	0

表 2 各フェーズで生成されたオブジェクトの数および観測された振舞いの回数(A 列は課題分析フェーズ、B 列は説明生成フェーズ)

Table 2 Generated objects in each phase (Column A: task analysis phase, column B: explanation generation phase).

	被験者 1		被験者 2		被験者 3		被験者 4		被験者 5	
	A	B	A	B	A	B	A	B	A	B
オブジェクト										
ユーザ質問	33	0	18	0	36	0	32	0	26	0
描画されたグラフ	29	0	15	0	33	0	31	0	19	0
G 付箋	2	0	13	0	0	0	0	0	9	0
G 自由線	0	0	10	0	0	0	0	6	0	0
H 付箋	5	2(+1)	9	9(+3)	0	0	0	0	0	0
H リンク追加	0	0	0	0	0	1	0	0	0	2
振舞い										
凡例強調表示	113	0	31	3	40	0	11	0	3	0
H オブジェクト移動	5	7	12	62	3	43	18	15	34	49
H オブジェクトサイズ変更	0	3	2	5	1	0	4	8	0	2
H オーバビュー操作	6	7	19	8	4	18	14	5	16	9
ウィンドウ切替え	32	1	9	12	13	4	30	14	153	13
ウィンドウ最大化	0	1	0	1	0	2	15	2	11	0

で生成された質問文オブジェクトの数)、「G 付箋」は Graph Viewer 上に付与された付箋数、「G 自由線」は Graph Viewer 上で付与された自由曲線ないし矢印の数、「H 付箋」は History Map 上に付与された付箋数、「H リンク追加」は History Map 上で追加されたリンク数を各々示している。ここで、表 2 の「H 付箋」の覽で括弧表記されている値は、Graph

Viewer 上で貼付された付箋が History Map 上に呼び出された回数を示している。

また、表 2 の振舞いの項は各被験者が実験の際に行った振舞いの回数で、「H オブジェクト移動」は History Map 上でサムネイルや付箋などを移動させた回数、「H オブジェクトサイズ変更」は History Map 上でサムネイルや付箋などのサイズを拡大/縮小させた回数、「H オーバビュー操作」は History Map 上でオーバビュー機能を利用した回数を各々示している。

なお、Graph Viewer のアニメーショントラッカ機能は課題分析フェーズ、説明生成フェーズのいずれにおいても利用されていなかった。また、本実験では被験者は発話について制約を受けなかったが、単独で行う分析作業であったため発話はほとんど観察されなかった。

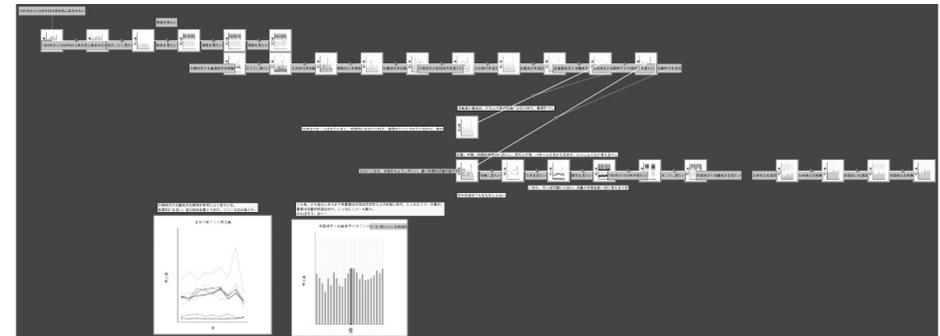
次に、各被験者の課題分析フェーズと説明生成フェーズの振舞いを以下に示す。紙面の都合上、すべての被験者の説明生成フェーズ終了後の History Map を掲載できないため、対照的な被験者 1 と被験者 4 の History Map を図 7 に示す。また、課題分析フェーズでの各被験者の History Map の利用状況を図 8 に示す。図 8 は各被験者の総操作数で正規化し並べたもので、a は History Map の操作全般が行われたとき、b はそのうちオーバビュー機能が使われたとき、c は History Map 上での付箋が貼付されたときを各々示している（被験者 3, 4, 5 は History Map 上での付箋の利用がなかったため c を省略している）。

被験者 1

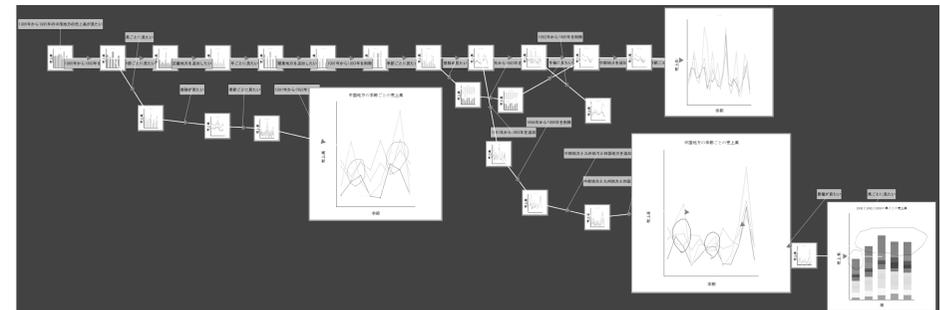
被験者 1 は、課題分析フェーズでは線条に探索を進める傾向が見受けられた（分岐は 2 カ所）。この被験者の場合、探索の途中で適宜 History Map を参照し、その上に付箋を残すことで、得られた知見や観点を外在化する様子が観察された。また、凡例を用いたグラフ要素の強調表示を多用し、ある要素とそれ以外の要素との関係を確認する行為が観察された。一方で、History Map を用いて前に描いたグラフに戻る行為はほとんど観察されなかった。説明生成フェーズでは、方針説明の際に用いるグラフを 2 つ選択し、そのグラフの前後のリンクを削除することで脱文脈化し拡大して並べた（図 7 (a)）。そして、説明内容の要点を付箋を用いて記述し、拡大した 2 つのグラフの上部に貼付していた。

被験者 2

被験者 2 は、課題分析フェーズでは線条に探索を進める傾向が見受けられた（分岐は 1 カ所）。この被験者の場合、探索の途中で疑問に感じたこと（e.g. 「岡山県が低いのは？」）や気がついたこと（e.g. 「割合だとあまり変わらない」）などを Graph Viewer 上に付箋として適宜残し、外在化する様子が観察された。History Map は探索の中盤から終盤にかけて用いられているが、中盤では、前に描画したグラフへの移動手段として用いられているのに



(a) 被験者 1 の History Map



(b) 被験者 4 の History Map

図 7 被験者 1 と被験者 4 の探索履歴（説明生成フェーズ終了後）

Fig. 7 Exporation histories obtained after explanation generation phase. (a) Participant 1, (b) Participant 4.

対して、終盤では付箋による要点の付与やオーバビュー機能による全体の把握のために用いられていた。説明生成フェーズでは、History Map 上で付箋を用いて説明内容のメモを記述していた。

被験者 3

被験者 3 は、課題分析フェーズでは線条に探索を進める傾向が見受けられた（分岐は 1 カ所）。この被験者の場合、付箋や自由線、サイズ変更、などの機能の利用はほとんど観察されなかった。History Map は、前に描画したグラフへの移動手段として探索の中盤で数回用いられていた。説明生成フェーズでは、分岐の一方をすべて消去し、残った一方の系列を

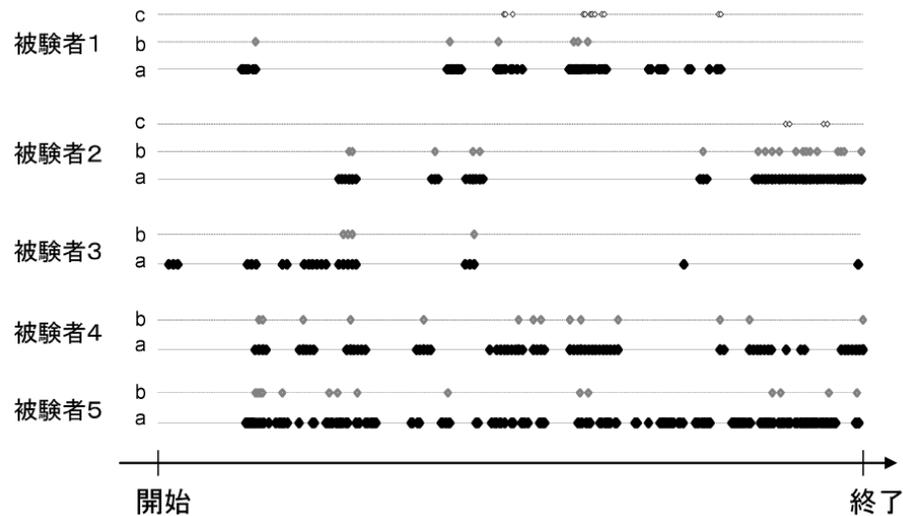


図8 課題分析フェーズにおける History Map の利用状況 (a: History map の操作, b: オーバビュー機能の利用, c: History Map 上で付箋の貼付)

Fig. 8 Utilization records of History Map in task analysis phase (a: History Map is manipulated, b: an overview function is utilized, c: an annotation is attached).

上下に振り分け, History Map 全体を使うようにノギリ歯状に配置していく様子が観察された。

被験者 4

被験者 4 は, 課題分析フェーズの前半では粒度変更 (地方ごと → 県ごと, 季節ごと → 年ごとなど) や要素の追加/削除などを繰り返しつつ, 様々な観点からのグラフ描画を試みていた (分岐は 6 か所)。一方, 課題分析フェーズの後半では, History Map のオーバビュー機能でそれまでの探索過程全体を概観し, 前に描画したグラフのうちいくつかを対象として, それらと新しく描いたグラフとを見比べる行為がしばしば観察された。説明生成フェーズでは, History Map 上でいくつかのグラフを選択し, サムネイルのサイズを拡大した。そして, Graph Viewer 上で自由線を付与し, それらのグラフにおける注目箇所を明示する行為が観察された (図 7 (b))。

被験者 5

被験者 5 は, 課題分析フェーズでは「1985 年から 1995 年までの中国地方の売上高」を

描画したグラフを起点として, 様々な観点の下でのグラフ描画を試みていた (分岐は 11 か所)。この被験者の場合, 課題分析フェーズの早い段階から History Map を多用し, 起点となったグラフへ移動し, 再度新しい探索を開始する行為が観察された。また, History Map のオーバビュー機能の使用やウィンドウの切替えが頻繁に観察された。説明生成フェーズでは, サムネイルの位置調整と不要なサムネイルの除去が観察された。

5. 議 論

5.1 課題分析フェーズにおける分析パターンの違いについて

各被験者の分析パターンを見ると, 被験者 1, 2, 3 は線条に探索を進める傾向があったのに対し, 被験者 4, 5 は様々な観点から眺める傾向があった。この違いは History Map の分岐数の違いに現れている。事後の聞き取り調査から, 被験者 1, 2, 3 は課題分析フェーズの早い段階で分析方針を決定し, それ以後はその方針に沿ったデータを探するという行為を行っていたことが分かった。たとえば被験者 1 は, 課題分析フェーズの早い段階で「中国地方と類似した地方を見つけてその連携を提案する」という方針を決定し, それ以後はその方針に合致するデータを見つけることに主眼をおいて分析を行っていた。一方, 被験者 4, 5 は, いろいろな可能性を確かめるために様々な観点からデータを視覚化し, ある程度それが蓄積された段階で分析方針を決定するという行為を行っていたことが分かった。

各被験者のウィンドウの用い方に注目すると, 被験者 1, 2, 3 が Graph Viewer と History Map を一覧できるようにウィンドウを配置していたのに対して, 被験者 4, 5 は Graph Viewer を最大化し, ウィンドウを切り替えながら探索を進めていた。被験者 4, 5 が Graph Viewer を最大化して利用していたのは, いろいろな観点から眺めている段階ではそれまでの探索過程 (履歴) を参照する必要がなく, 探索行為に必要なウィンドウを拡大してその視認性を向上させるほうがユーザの行為に即していたためだと解釈できる。これは上記の分析方針の違いにも符合する。

このように, 分析方針が異なると履歴のパターン (分岐数など) やウィンドウの使われ方に影響があることが観察された。このことから, 履歴パターンやウィンドウ切替えを手がかりとしてユーザのとった分析パターンや分析の進行状況 (e.g. 分析方針の決定がなされたか否か) などをシステムが理解し, それに基づいて支援方式を適応させる, といった機能の実現可能性が示唆される。

5.2 外在化の方法について

システム的设计段階では, 「被験者は考えの外在化に付箋を用いる」ことを想定していた。

実際、付箋を貼って考えの外在化を行う様子が被験者 1, 2 で観察された。しかし、被験者 3, 4, 5 は付箋をまったく使用しなかった。その代わりに、被験者 4 と 5 ではオブジェクトの位置を変えたりサイズを変更したりすることで、振り返りの手がかりにする行為が観察された。前者を直接的外在化 (direct externalization)、後者を暗示的外在化 (allusive externalization) と呼ぶことにする。暗示的外在化は内的認知資源である「記憶」に頼る外在化表現と考えられ、内的認知資源との相互作用によって振り返りの際の手がかりとして機能しうるものといえよう。

事後の聞き取り調査で直接的外在化を行わなかった 3 人の被験者にその理由を尋ねたところ、「付箋の付与が面倒だった」、「覚えているので必要性を感じなかった」という回答が得られた。しかし、被験者 4, 5 は振り返り行為の際、過去に描いたグラフを探すのに手間どり、あたかも操作に迷っているかのようにオーバビューやウィンドウ切替えを多用する様子がしばしば観察された。これは、暗示的外在化が探索行為の振り返りの手がかりとして十分に機能していない、もしくは不十分であることを示唆している。このような「操作の迷い」は直接的外在化を用いた被験者には観察されなかった。また、直接的外在化を利用した被験者 1, 2 では、説明生成フェーズでそれを参照しながら説明に用いるグラフを選択したり、方針説明の際に外在化された内容に言及したりする様子が観察された。これは、直接的外在化が探索行為の振り返りに寄与することを示唆している。

ただし、すべての場合に直接的外在化を利用できるとは限らない。たとえば被験者 4 は事後の聞き取り調査で「分析しているときは、これをポイントにしようってつもりはないんですが、ばあーっと見ていて、ポイントにするならそこかなあ、と思い出して戻っていく、とか、だーっと見ていった結果、なんかこうポイントとしてつかえそうかなあ、っていうんで、ヒストリを見ながら戻っていく、という感じ。後から、あそこが使えるかな、って」という意見を述べていた。この意見で示されているように、探索を進めている時点では考えが漠然としていて、あとでその重要性に気がつく場合には、そのグラフが描画された時点での直接的外在化を期待することは現実的ではないだろう。

このことから、(1) よりユーザに負荷の少ない、いいかえれば「面倒だ」と思わせない手段で直接的外在化を可能にすることと、(2) 内的認知資源を有効に活用できる暗示的外在化を探求することが、探索的データ分析支援システムに求められるデザイン課題として示唆される。

5.3 前に描いたグラフへのバックトラックについて

InTREND の大きな特徴は、システムとのインタラクションに自然言語を用いる点であ

る。ユーザは自らの興味や疑問に基づいて質問 (発話) をシステムに与えることで、グラフを描画する。システムの設計段階では、発話は探索を進める過程で利用され、振り返りのために前に描いたグラフへ戻る (バックトラックする) 場合は History Map が利用されると考えていた。しかし、実際の観察結果を見ると、一部の被験者で「言葉によるバックトラック」をする様子が観察された。たとえば、中国地方のグラフが描かれているときに、「四国地方を追加したい」と入力して中国地方と四国地方のグラフを描いたあと、再度「四国地方を削除したい」と入力して、直前に描いたグラフを再度描画するという行為である。事後の聞き取り調査で確認したところ、「ウィンドウを切り替えて戻るより、言葉として表現するほうが早い」、「同一のウィンドウで要素を追加したり削除したりして違いを見たかったのだが、ウィンドウを移ると画面上の位置が変わってしまい、思考が途切れてしまう」と考えていたためだということが分かった。また、「(過去に描いたグラフの識別性が低い) History Map でグラフを探している間に考えていることを忘れてしまう」という意見も得られた。これは、5.2 節で言及した「操作の迷い」の一因とも考えられる。これらに鑑みると、ユーザの思考を阻害しないためには、直前に描いたグラフだけでなく、いくつか前に描いたグラフに対しても自然言語でバックトラックすることを想定した自然言語解釈方式の拡充が必要であるといえる。

5.4 使用されなかった機能について

今回の実験では、「アニメーショントラッカ機能」がまったく使われていなかった。その理由としては、(1) 2 つのグラフの関係性を把握することが分析を進めるうえで必要ない、(2) トラッカの機能を把握していない、(3) トラッカの操作がユーザの分析行為に適していない、が考えられる。ユーザの課題分析フェーズでの振舞いを見ると、直前に描いたグラフを参照する場合、History Map で直前のグラフをクリックして Graph Viewer に表示させる、自然言語で質問を生成して直前のグラフと同じものを再度描画する、などの行為がしばしば観察された。加えて、事後の聞き取り調査でも、2 つのグラフを並べて比較したかった、という要望があった。また、事前の習熟訓練の段階でトラッカ機能の使い方を教示しており、事後の確認でもその操作方法是被験者に理解されていた。このことから、トラッカが用いられなかった主たる要因は (3) にあると思われる。事後の聞き取り調査で「現在描いているグラフと過去に描いたグラフとを対応づけて比較したい」という意見があったが、これに鑑みると、同一ウィンドウ内でアニメーションを用いて対応付けを把握させるのではなく、複数のウィンドウを並べて一覽させ、その描画要素の対応関係を提示する方法が、よりユーザの分析行為に即している可能性がある。

5.5 探索行為と振り返り行為でのインタラクションモダリティの違いについて

探索の履歴を空間的に配置することで探索過程を把握したり再利用したりするインタフェースは様々な提案されている。白井らの分類¹³⁾に基づけば、本稿で提案した履歴表示の方法は、空間マッピングの1つである軸マッピング方式を用いて履歴を表示する手法に属する視覚化方法である。同様の方式でユーザの探索行為に関する履歴を視覚化する先行研究としては、齋藤らの探索行動フィードバックシステム¹²⁾、AmantのAIDE/NIFT¹⁾、DerthickらのVisage⁴⁾などがあるが、本稿で提案するインタフェースは(1)探索行為と振り返り行為で用いるモダリティが異なっている点と、(2)振り返りの手がかりとなるアノテーションやメタコメントを付箋をメタファとして付与できる仕組みを持っている点で特徴的である。(1)のデザイン理由は、探索/振り返りの各々の行為に適切なインタラクションモダリティを利用できることが、効率的な分析の遂行に寄与すると期待したためである。課題分析フェーズでユーザがシステムに質問として与えた自然文は、History Map上では灰色の付箋としてオブジェクト化される。ユーザが振り返りを行う際、このオブジェクトの内容を確認し、その時点でどのような意図で探索を行っていたか確認する振舞いが一部のユーザで確認されており、自然文での入力が振り返りの際の役に立っていることが示唆される。

ただし、事後のインタビューで、「(質問は)その時点の文脈でしているの、前の質問も見ないとちゃんと分からなかった」という指摘があった。これは、現時点の実装が、省略をともなう断片的質問であってもそのままHistory Map上にオブジェクトとして表示する仕組みになっているためである。これを解決するには、システムが文脈を補完して解釈した結果を再度言語化してオブジェクトとして提示する方法が考えられるが、この場合、ユーザが実際に入力した質問文と表層的に一致しないため、振り返りの手がかりとして機能しない可能性があることと、文字数が多くなるために一瞥性の低下が懸念されることの2つの問題が生じる可能性がある。この点に関しては、今後どのような表示が適切であるか実験を通じて明らかにしてゆきたいと考えている。

また、5.3節で述べたように、前のグラフに戻る操作は、直接操作だけでなく言語表現による操作が適切である場合もあるため、いくつかの操作に関しては、自然言語による操作とマウスによる直接操作のどちらでも可能にする冗長性が求められていることが分かった。

これらの知見をふまえ、本システムのさらなる改良を進めてゆきたいと考えている。

6. おわりに

本稿では、探索的データ分析支援システムInTRENDを用いた被験者実験を実施し、分

析を遂行する被験者の振舞いを観察した。本システムは一般ユーザではなくデータの分析を業務の1つとしてかかえるユーザを対象としていることと、現状の自然言語対話処理技術は完全に自由な発話を解釈する能力の実現に至っていないためシステムの利用に十分な習熟を要することから、5人の被験者を対象として十分な習熟訓練を行ったうえでシステムの用い方を観察する方法を採用した。その結果、被験者は振り返りの際の手がかりとして、付箋機能のような直接的外在化と同様に位置の移動やサイズ変更といった暗示的外在化を用いている様子が観察された。また、分析方針の違いが履歴のパターンやウィンドウの使用方法に影響すること、一部の機能に関しては直接操作と自然言語の両方で操作できる冗長性が求められていることなどが明らかになった。

本研究の立場は、探索行為と振り返り行為の各々に適したインタラクションをユーザに提供し、それらを分析の進行状況に応じて使い分けることでより効果的な分析支援が可能になるというものである。この立場に立脚してデザインしたInTRENDシステムを用いて行ったユーザ観察の結果から、その有効性の一端を示すとともに改良すべき点も明らかになったと考えている。異なるインタラクションを用いた探索/振り返り行為のシームレスな遂行に関するデザイン指針は、探索的データ分析支援システムの枠を越え、発想支援システムや意思決定支援システムなど、様々なインタラクティブシステムのデザインにも関わる重要な論点であると考えられる。今後、さらなるシステム改良やユーザ観察を通じて議論を深化させてゆきたいと考えている。

今回の実験では売上高データのみを使用したが、実際の分析場面では、このようなデータを単独で用いるのではなく、経済指標や天候など他のデータを組み合わせる用いることが想定される。現在のシステムは複数の異なるデータを組み合わせる手段を提供していないため、今後はそのような利用を想定した機能の拡充を図る予定である。

インタラクションデザインの観点からすれば、人によって分析行為のパターンが異なるため、各々の分析行為パターンに応じたシステムデザインが施されるべきであるという指摘もある¹⁰⁾。しかし、分析行為のパターンは、たとえ同じ人物であっても、ドメインや対象などによっても異なる可能性がある。もちろん、機能を豊富に用意してあらゆる可能性に対処するという方法は、システムデザインとして適切ではなくユーザの混乱を招く結果にもなるが、同一人物であっても状況によって分析パターンが変化しうることを鑑みれば、ある程度の冗長性を持ったデザインにしておくことは必要であろう。そして、その範囲内で、ユーザの嗜好や当面のタスクによって機能を適宜選択し、円滑に分析が遂行できるというのが望ましいあり方なのではないかと考える。今後、この冗長性の幅をどの程度見込むことが汎用性

と簡便性の両立に寄与するかについての考察にも興味を持っている。

謝辞 本研究は文部科学省科学研究費(課題番号:17700168)の助成を受けています。ご支援を感謝します。本研究の遂行にあたり、有益な助言をいただきました名古屋大学齋藤洋典教授ならびに齋藤研究室の学生諸氏に感謝します。

参 考 文 献

- 1) Amant, R.S.: Navigation and Planning in a Mixed Initiative User Interface, *Proc. AAAI-97*, pp.64–69 (1997).
- 2) Carbonell, J.G. and Hayes, P.J.: Recovery Strategies for Parsing Extragrammatical Language, *American Journal of Computational Linguistics*, Vol.9, No.3-4, pp.123–146 (1983).
- 3) Cohen, P.R.: Natural Language Techniques for Multimodal Interaction, 電子情報通信学会論文誌 D-II, Vol.J77-D-II, No.8, pp.1403–1416 (1994).
- 4) Derthick, M. and Roth, S.F.: Enhancing Data Exploration with a Branching History of User Operations, *Knowledge Base Systems*, Vol.14, No.1-2, pp.65–74 (2001).
- 5) Hartwig, F. and Dearing, B.E.: *Exploratory Data Analysis*, SAGE Publications (1979).
- 6) Hutchins, E.L., Hollan, J.D. and Norman, D.A.: Direct Manipulation Interfaces, *Human-Computer Interaction*, Vol.1, No.4, pp.311–338 (1985).
- 7) 加藤恒昭, 松下光範: 探索的データ分析のためのマルチモーダル対話処理, 電子情報通信学会論文誌 D-II, Vol.J87-D-II, No.5, pp.1142–1152 (2004).
- 8) 松下光範, 加藤恒昭: コンテキスト保持による探索的データ分析支援の枠組, 知能と情報, Vol.18, No.2, pp.251–264 (2006).
- 9) Matsushita, M. and Shirai, Y.: Supporting Exploration and Reflection in Exploratory Data Analysis, *Proc. CDWS2005*, pp.3–8 (2005).
- 10) 中小路久美代: インタラクシオンデザイン, 情報処理学会研究報告, IPSJ-SIG-HI, Vol.2004, No.115, pp.1–3 (2004).
- 11) Plaisant, C., Milash, B., Rose, A., Widoff, S. and Shneiderman, B.: LifeLines: Visualizing Personal Histories, *Proc. CHI'96*, pp.221–227 (1996).
- 12) 齋藤ひとみ, 三輪和久: Web 情報検索におけるリフレクションの支援—探索行動フィードバックシステムの構築, 人工知能学会論文誌, Vol.19, No.4, pp.214–224 (2004).
- 13) 白井良成, 中小路久美代, 山田和明: インタラクシオンヒストリの要約と閲覧に関する一考察, 情報処理学会研究報告, IPSJ-SIG-HI, Vol.2004, No.14, pp.17–24 (2004).
- 14) Shneiderman, B.: Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces, *Proc. IUI'97*, pp.33–39 (1997).
- 15) Shneiderman, B.: *Designing the User Interface*, 3rd edition, Addison-Wesley (1998).

- 16) Williams, M.D., Tou, F.N., Fikes, R., Henderson, A. and Malone, T.: RABBIT: Cognitive Science in Interface Design, *Proc. CogSci'82*, pp.82–85 (1982).
- 17) 山本恭裕, 高田眞吾, 中小路久美代: Representational Talkback の増幅による「書いてまとめる」プロセスの支援へ向けて, 人工知能学会誌, Vol.14, No.1, pp.82–92 (1999).
- 18) Yamamoto, Y., Nakakoji, K. and Takada, S.: Hands-on Representations in a Two-Dimensional Space for Early Stages of Design, *Knowledge-Based Systems Journal*, Vol.13, No.6, pp.375–384 (2000).
- 19) Zelazny, G.: *Say It with Charts*, 3rd edition, McGraw-Hill (1996).

付 録

付録として、実験で得られた被験者 1 と被験者 4 が行った報告文を掲載する(口頭で行われた説明を書き起こしたものであり、フィラなどは除いてある)。

A.1 被験者 1 の行った方針説明

私がいいたいのは、全国的な視野の大きな視野で見なくてははいけません。そのときに各支店を見てみると、非常に傾向にばらつきがある。最初に結論をいいますと、中国支店と近畿支店は、ゆるやかな統合をして協調するとよいでしょう。その根拠となるのは、まず地理的に近いということが、統合してうまくやっていくポイントとなるだろう。けれども、売上高を見ていくと、中国地方と近畿地方は非常に似かよったパターンで推移しているのが分かります。非常に似たパターンでいってるので、ノウハウとかも似ているのだろう。それで、一緒にやっていったほうがいいんじゃないかな? という話です。じゃあ一緒に近い、九州地方や四国地方もあるだろうという話になりますが、九州地方は、中国・近畿に比べて非常に売上高が違うので、何をやっているのかわかりませんが、我々とはちょっとパターンが違うだろう、そこは分けて考えたほうがいいだろう。同様に四国地方に関しても、同じラインで若干、中国・近畿よりも落ちる。そうすると、当面のところそれも外して考えといたほうがいいのではないかな。こういった全体的な傾向から、まず中国地方と近畿地方が協同していくべきだ。そのときに、たとえば、リソースをお互いに融通しあうといった考え方があると思います。じゃあ、どういうふうのリソースの配分をするかということですが、こちらに、ここ 2 年の月ごとの売上高をプロットしてみました。オレンジの方が中国地方で、黄色の方が近畿地方です。そうすると、おもしろい傾向があって、寒い時期、冬に入ると中国地方の売上高が高くなる、逆に暖かくなってくると、割合として近畿地方の売上げが高くなっていく。そこでどういった戦略をとるかということですが、当面、とりあえずは、中国地

方で売上げが好調な冬期間には、近畿地方の方からリソースを中国地方の方に回す、で夏の時期には逆に売上げ好調な近畿地方へ中国地方からリソースを回す、と。こういう戦略で、なにも中国支店 1 店舗に投じるのではなくて、もうちょっと統合してやっていくと、全社的にもハッピーになるのではないかと私は考えております。

A.2 被験者 4 の行った方針説明

中国地方のこれまでの売上高を全体的に他の地方と比較しますと、全体的な年ごとの推移はそれほど変わらないのですが、季節ごとに見ますと、どうも中国地方、このオレンジ色の部分は、年にもよるのですが、夏から秋にかけてどうも他の地方と比べるとすこし推移にムラがある。他の地方では、このように夏から秋にかけてかなり伸びているのですが、中国地方は、たとえば、この部分この部分ですとか、夏から秋にかけて伸びている年と伸びていない年とムラがあるような気がいたします。では、それを詳細に分析してみますと、これは中国地方の季節ごとの売上高なんですけど、中国地方の夏から秋をこのように、去年、今年で見ると、まあ、このようになっていまして、全体的にあらかた、夏から秋にかけてこのようにある程度伸びてはいるんですが、たとえばこの岡山県ですとか、広島県ですとか、ここらへんはちょっと、夏・秋の売上高ののびが、他よりもどうも出足が鈍っているような気がいたします。実際に、これを全体的な県ごとの売上げで見ますと、たとえば岡山県とか、広島県は他の売上げと近いんですが、岡山県とか広島県はこのように、全体の県よ

り少し売上高が落ちています。ということは、今後の戦略としては、夏から秋にかけての売上高の伸びのカーブをできるだけ、他の地方と他の地域ですね、関東地方とかそういったところと同じように伸ばして行って、特に中国地方の中で伸びの率が悪い、岡山県ですとか、そこらへんに設備投資というか営業投資をしていくのがいいのではないかと、という考えに至った次第でございます。

(平成 19 年 10 月 4 日受付)

(平成 20 年 4 月 8 日採録)



松下 光範 (正会員)

1969 年生まれ。1993 年大阪大学工学部精密工学科卒業。1995 年大阪大学大学院基礎工学研究科物理系専攻制御工学分野博士前期課程修了。同年 4 月日本電信電話株式会社入社、現在に至る。自然言語理解、情報可視化、ヒューマンコンピュータインタラクションに関する研究に従事。博士(工学)。2002 年度情報処理学会論文賞、2005 年度電子情報通信学会ヒューマンコミュニケーション賞、2006 年度パーソナルリアリティ学会論文賞ほか各賞受賞。日本知能情報ファジィ学会、人工知能学会、日本パーソナルリアリティ学会、ACM 各会員。