

論文のチャート化による論文理解支援システムの提案

岡 孝明 武田 英明

奈良先端科学技術大学院大学 情報科学研究科

〒 630-01 奈良県生駒市高山町 8916-5

電話: 0743-72-5265 , fax: 0743-72-5269

e-mail: takaak-o@is.aist-nara.ac.jp

本稿では「論文を読んで、理解し、書く」行為の支援のため論文を要約してチャート化する手法を提案する。具体的には、「論文の内容を要約してチャート化すること」、「チャートを「イメージ・スキーマ」に基づいて再構成すること」で、支援を実現する。「イメージ・スキーマ」とは、Lakoff と Johnson によって提唱された概念で、人間の身体性に根差した、人間にとって普遍的な認知パターンのことである。本研究では、システム使用者がイメージ・スキーマを意識せずに書いたチャートを、イメージ・スキーマに基づいて再構成する。このことで、チャートを違った視点から見ることを可能にし、そのことによって、チャートのより直感的、本質的理解へとたどりつくことができる。

Keywords: 発想支援、論文理解、イメージ・スキーマ

A Proposal of a System for understanding technical papers by charting

*Takaaki OKA Hideaki TAKEDA

Graduate School of Information Science, Nara Institute of Science and Technology

8916-5 Takayama, Ikoma, Nara, 630-01 Japan

phone: 0743-72-5265 ext.5316, fax: 0743-72-5269

e-mail: takaak-o@is.aist-nara.ac.jp

In this paper, we analyze utilization of charts in human actions such as reading, understanding, and writing papers, propose a new method of understanding papers based on image schema by Johnson and Lakoff. Image schema suggested by Lakoff and Johnson is a set of cognitive pattern for human based on human embodiment. We conducted experiments on writing charts by five subjects, and proposal three different types of application shemata. Each application schema consists of a set of labels of arcs which is associated to elements of basic image schema. It makes people enable to see charts from another view point, and understanding charts more intuitively and essentially.

keyword: creative thinking support, understanding papers, Image-Schema

1 はじめに

人間の自由な発想、独創的な思考をいかに支援すべきかという問題、すなわち発想支援の問題は、最近のコンピュータ技術の進展から、注目を集めている課題といえる。発想支援には、さまざまな形態がすでに研究され、現在も研究されている（例えば [1]）が、本稿でとりあげるのは、「論文を読んで、理解し、書く」という行為の支援である。本研究では、特に、

- 論文の内容を要約してチャート化する。
- チャートを「イメージ・スキーマ」に基づいて再構成する。

ことによって、「読んで、理解し、書く」プロセスを支援する。イメージ・スキーマという概念は、言語学者の George Lakoff と 哲学者の Mark Johnson によって提案された [4]。Lakoff と Johnson は、人間の理解において、本質的で普遍的な、言語や文化の壁を越えた共通性があると指摘する。それは、我々人間の身体と、それにもとづく我々の経験である。そして、この身体性に根差した、人間にとって普遍的な認知パターンがイメージ・スキーマである。本研究では、システム使用者がイメージ・スキーマを意識せずに書いたチャートを、イメージ・スキーマに基づいて再構成する。このことで、チャートを違った視点から見ることを可能にし、そのことによって、チャートのより本質的理解へとたどりつくことができる。

2 「読んで、理解し、書く」プロセス

研究者は、常日頃、研究活動の際、論文を読み、理解し、それらをまとめて、整理し、独自の着想に至り、論文を書く。これらの行為そのものについては、[6] で詳細に述べられているので、ここでは概略を述べるにとどまる。

(1) 「読む」行為：文章を読む際に、文章にアンダーラインを引いたり、印をつけておく、などの行為がよく行われる。これを一般化すると、論文の重要な点を短いセンテンスで表現する、という行為になる。その短いセンテンスをメモとして取っておけば、その文章を「読んだ」ということが言える。すなわち、本論文での言葉の使い分けで言うと、「読む」行為は、「理解する」行為の前段階であり、情報の整理、圧縮行為である。

(2) 「理解する」行為：次に「理解する」行為について述べる。論文を理解するのに、論文の内容をまとめて、チャートを作るという行為が助けになると言われている [7]。

[7] においては、情報のチャート化が、知識を整理して、そして新たな文章を書く際に、重要な役割をなす、とある。ここでいうチャートとは、論文を「読んだ」ことによって出てくるメモを概念的に図解したものである。

チャートを作ることによって、現時点で自分が持っている知識の整理が可能となる。

[?] では、知識整理は、情報を再利用可能な形に体形立て有効に利用する場合には不可欠な活動、とある。さらに [?] では、システム設計をする場合に、あらかじめその領域に登場すると想定した概念実体・概念相互の関係からトップダウンに設計する方法と、対象とする領域に現れるデータ・情報を収集、分析してボトムアップに設計する方法の二つが挙げられている。本稿で取り上げる方法は、主にボトムアップ的手法に属する。トップダウン的設計方法は、対象領域、対象問題の性質があらかじめうまく整理・把握されている場合には効率的な方法である。しかし、システムの要求仕様決定のフェーズも含めて考えると、問題があらかじめうまくモデル化されていない場合も多く、本稿で取り上げる「論文を読んで、理解して、書く」行為の支援、という場合は、システムを底辺から作り上げる行為も含まれる。このようなことから、収集した情報を分析して、与えられた設計の枠組みを見直す、というボトムアップの方法が不可欠であると考えられる。

また [1] では、文章を作成、構成するという作業は、何を書きたいかの「理解」、その「表出」、その結果もしくは影響の「解釈」、という作業の繰り返し、とある。また、既存のワープロでは、その多くの支援が「表出」に向けられ、「理解」「解釈」の支援があまり行われていない、とある。チャートにする時（要約の特殊形）の「理解」を支援すれば、自分がよりその論文を深く理解することができ、[1] での言葉で言えば「解釈」までの行為にも影響を及ぼす、ということになる。

チャートと言っても、さまざまなものが考えられるが、本論文では、チャートは一種の有向グラフであるとする。ノードにあたる部分に、論文におけるキーワード、短いキーセンテンスが入る。アークにあたる部分には、ノード間の関係を示すラベルが添えられる。このようなチャートを作ることによって、論文の理解が深まるとするならば、チャートを作る行為を支援すれば、理解の支援が出来ることが予想される。

さらに理解を支援するために、本論文では、前述したイメージ・スキーマを導入する。イメージ・スキーマによって再構成されたチャートを「基本チャート」と呼び、元のチャートを「応用チャート」と呼ぶ。

応用チャートは、後述の実験でも明らかになったことだが、同じ論文を読んでも、作成者によって、全く違うものが出来あがる。これをイメージ・スキーマを用いて、基本チャートを作ることによって、人間の身体に基づいた普遍的なイメージで捕らえることが可能になる。このことによって、本質的な差異は残して、単なる用語の違い、などといった、些細な違いを切り落として、本質的には似ているものを似ていると認識することが出来るようになり、他人の作ったチャートを参考にする際に違いすぎることがないので、参考にしやすくなる。また、個人の単位でも、基

本チャートと応用チャートを比較することによって、これらチャートを生み出した元の論文のより深い理解が可能になると考える。

応用チャートから基本チャートへの変換は、論文の本質の抽出である。

(3)「書く」行為：[7]では、チャートは上手に作れば、一枚のチャートを説明することで、そのまま一篇の論説を書くことが出来る、とある。このような良いチャートを作ることの支援が、本研究での目的であり、それこそが、「読んで、理解し、書く」行為の支援である。良いチャートを書く、ということは、より深く「理解」する、ということである。すなわち「書く」行為は、より深く「理解」することによって、支援される、と考えられる。

[?]では、「人間が得意な部分は、人間に任せ、計算機が人間より優れた点があれば計算機に任せる。あくまで発想の主体は人間である。この立場に立つ利点は、計算機が発想とは何かを理解していなくても発想支援ができる点にある」とある。チャートを見て、イメージを膨らませて、文章を書く、という行為はコンピュータよりも人間の得意な分野である。このことから、「書く」行為を直接的にコンピュータが支援する方法を模索するよりは、チャートから文章を生み出す行為を助けることに専念すべきであり、そして、その方法は、より良いチャートを書くことを支援することに尽きる、と考える。

2.1 まとめ

以上より、チャートを作る行為を支援することによって、「読んで、理解し、書く」プロセスが総合的に支援される、と考えられる。

3 チャートによる論文の表現

3.1 人の手によるチャート作成実験

本論文におけるチャートとは、前述した通り、一種の有向グラフである。チャートは、論文のメモから作り出したノード群を、ラベルの付いた矢印によって、ノード間の関係を記述することで、作成される。ここで、どのようなシステムを用意すれば、チャートをより容易に作成できるかを調べるために、実験を行った。情報工学の初心者5名を被験者とする実験を行った。実験内容は、各被験者に10の情報工学の専門論文を与え、それぞれの要約チャートを書いてもらうというものである。著者が1つの例題を示した。そして、提出された5名、10枚ずつ、合計50枚のチャートを分析した。

結果を見ると、同じ論文でも他人と同じあるいは類似することはなかった。このことは一見チャート書きのプロセスは個人によって独特であるように見えるが、ノードの作

り方やアークの名前付け(ラベル)などの個々の部分では共通点が見られる。そこで、ここでは特にアークのラベルに注目して、その共通性を分析することにする。

3.2 応用スキーマの導入

ここでは、チャートに現れる要素間の関係をいくつかの応用スキーマに分けることを試みる。

技術系論文において本質的には言語に依存しない主張があると考えられる。しかし、文章によって表現するという制約から文章表現としての構成を持つ。すなわち、論文の中の構造は大きく分けて文章の主張にかかわらない構造と、文章の主張にかかわる構造が混在していると思われる。従って、チャートにおいてはこのような異なる構造が混ざって出現していることになる。ここでは、この視点から応用チャートを以下のように提案する。

- 文章の主張にかかわらず、文章の構造そのものを指し示す「文章論理」
これには、例えば A が B の必要条件であるという論理的つながりを示す「必要」や、B が A の内容を表すことを示す「説明」などが入る。この場合、A と B の関係は文章構成上の関係であり、主張そのものではない。
- 文章の主張のレベルで論理の展開を表す「因果関係」
これには、A が正しいと前提にした上で B を導く「そこで」(「A なので B である」)や、A が正しいにもかかわらず B が成立する、という「しかし」(「A だが B である」)などが含まれる。この場合、A と B がこのような関係にあることが論文の主張である。ただし、技術系論文特有ではなく、一般的に用いられるものである。
- 文章の主張にかかわり、かつ技術論文に特有の表現を集めた「技術論文系」
これには、技術論文に特有の表現、論理構成を表すための応用スキーマが入る。例えば、「解釈」(「A を解釈すると B となる」)「適用」(「A を B に適用する」)など。

チャートにおいて出現頻度の高いラベルはこれら3つのスキーマいずれかに分類される。頻度の高いラベルおよびその分類の一部は Table 1 を参照していただきたい。表 1 列目に書かれているのがスキーマの一番大きなスケールでの3分類、「因果関係」「文章論理」「技術論文系」である。順序が前後するが、表 3 列目に書かれているのがスキーマの一番細かいスケールでの分類、「分類名」である。これら分類名のいくつかは強い類似性を持ち、一つの群に属すると考えられるので、それを表したのが表 2 列目の「分類カテゴリ」である。例えば、分類名「変化」と「記述」は、一つの分類カテゴリ「状況変化」に属する。分類カテゴリ

応用スキーマ	分類カテゴリ	分類名	要素	基本スキーマ
因果関係	「因果」	「順列」	そこで、しかし、そのため	<経路>
		「因果」	さらに、原因、結果、重要	
文章論理		「説明」	説明、つまり、具体的には、内容	<経路>
		「相いれない」	相いれない	<連結>
		「必要」	必要	<経路>
		「不必要」	不必要	<経路>
技術論文系	「状況変化」	「変化」	解釈、改良、帰着、拡張、変化	<経路>
		「記述」	記述、表現、枠組み	
		「提案」	提案	<経路>
		「基づく」	基づく	<上/下>
		「関係」	関係、同値、適する	<連結>
		「特徴」	注目、特徴、ポイント、キーワード、問題点	<中心/周縁>
		「より困難」	より困難	<上/下>
		「列挙」	構成される、細分化、(暗黙)	<部分/全体>
	「主語A」	「自発動作」	増加、実現	オリジナル
		「意図的動作」	産み出す、支援、構築する、分析対象	
	「適用」	「適用」	導入、適用、広く用いられる、利用	オリジナル
		「本論文では」	実験結果、実験、本論文では	
		「対象」	対象	オリジナル
		「融合」	融合	変則<経路>

Table 1: 応用スキーマと基本スキーマの対応一覧

が空欄の分類名は、自分より1つ上位の分類が直接「因果関係」「文章論理」「技術論文系」になる。「要素」は、ラベルそのものである。これら(例えば「そこで」)がアークのラベルとなる。

3.3 応用スキーマの分析

現時点で提出した応用スキーマは今回の実験のみに基づくもので、完結したものではない。今後より広範な調査を行なって他の応用スキーマの可能性や応用スキーマ内の要素の検討をさらに行なっていく必要がある。

このような応用スキーマに相当するものは様々な分野で関係性の定義として表れているものを思われる。以下では、いくつかの論文におけるスキーマに対応する関係を取りあげて、本論文で提案した応用スキーマとの関係を議論する。

[2]においては、「FISMセッション」と呼ぶ発想支援プロセスが提案されている。このプロセスとは対象となる問題に対するアイデアを収集し、関係づけることである。そこでの関係とは擬順序を有する関係であり、因果関係をはじめとして10の関係が挙げられている。本稿の因果関

係における「因果」が[2]の因果関係と対応する。本稿で「そこで」や「しかし」を「順列」として分けたのは、[2]で言うところの因果関係とは厳密には違うと考えたからである。文章論理、技術論文系スキーマについては対応する関係が存在しなかった。

[8]においては、文章の要約のために、事象の枠組を記述する。そしてその事象の解析に必要なものとして基本要素、事象の役割を記述するものとして基本格を与えた。基本格としては、agent, object, 以下7つの格を与えている。これは[4]において与えられた、スキーマの一つAgentと類似している。本稿で、「主語A」として与えたスキーマは、これを論拠としている。

[5]では、自動抄録生成のための文書構造解析を行なっている。文書中の文章の構造解析のために修辞関係を解析しているが、修辞関係の例として、順接をはじめとする10の関係名を挙げている。ここで挙げられた関係は「そこで」「しかし」など本稿での応用スキーマの要素と重複するものがあり、関係の定義に類似性があると思われる。対応した要素は主に因果関係、文章論理であった。そこで[5]における関係の一部にTable 1では取り込んでいる。たとえば「換言」は、表現例として使われている「つまり」をそのまま用いている。その一方で「背景」は採り入

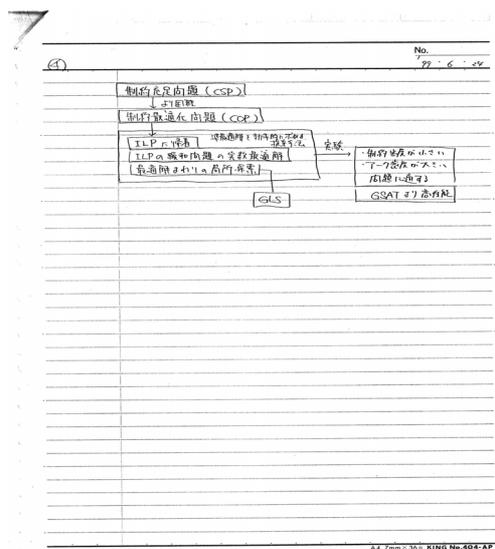


Figure 1: 被験者により提出された応用チャート

れず、別のスキーマを組み合わせて表現するようにしている。「対比」のように、出現頻度は低くても重要と思われるものもあり、今後本研究でのスキーマリストに入れることを検討している。

3.4 基本スキーマの導入

チャートのアークに貼りつけるラベルに付いては、前節で、頻度の高いものを選び出すことで、各被験者間の共通性を得た。ここでは、チャートのより深い理解のために、各応用スキーマが、どのような普遍的なイメージに裏打ちされているかを、対応づけることを提案する。この普遍的なイメージを持つスキーマを「基本スキーマ」と呼ぶ。書かれたチャートを、より人間に普遍的な構造をもって再構成することによって、視点を変えることができる。このことが、チャートのより深い理解と、チャートの不備の発見につながる。ここで人間に普遍的な構造と述べたが、Lakoff と Johnson の提唱する「認知意味論」の考え方に基づき、「イメージ・スキーマ」によるチャートの再構成をする。

このために、まず、応用スキーマの各要素を基本スキーマの各要素に結び付けることを試みる。Table 1の「要素」欄と「基本スキーマ」欄の対応に、その一部を示す。例えば、「ポイント」というラベルについて考えると、このラベルが示している事実は「AのポイントはBである」と

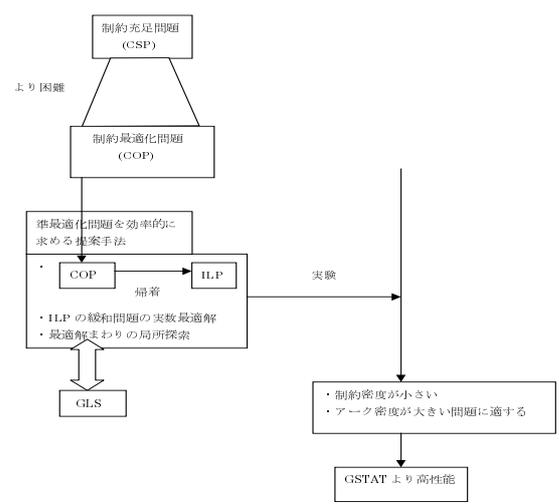


Figure 2: 再構成された基本チャート

いうことであるので、この事実は、Aというカテゴリーの示している領域の中に、Bという事実が含まれており、かつ重要である、ということイメージとして示唆するものであるから、「ポイント」に対応する基本スキーマは<中心/周縁>である。

本稿では、基本スキーマにオリジナルのスキーマを導入しているが、これは、現時点でのスキーマのリストが部分的である ([3]より) ことからである。[4]においても、数学の世界に認知意味論を導入する際に、オリジナルのスキーマを導入している。このことは、本論文のスキーマのリストのみならず、完璧なスキーマのリストが存在していないことを推察させる。

このようにして得られる対応を用いて、応用スキーマで表現されたチャートを基本スキーマのチャートに変換することができる。

3.5 再構成の例

この再構成は、具体的には Figure 2 のようなものが挙げられる。Figure 2 は、Figure 1 を前節の方法によって変換したものである。Figure 1 と Figure 2 を比較すると、応用チャートが一種の有向グラフであったのに対し、基本チャートはイメージ・スキーマが望むような図形へと変換されていることが分かる。ラベルに貼られた言葉による文章理解よりも、イメージ・スキーマに基づいた図形的イ

メージの方が、より人間にとって本質的な理解に近いことが期待される。この再構成によって、元になる論文の著者が、文章を使って表現していたイメージが、直接的に基本チャートによって示されるので、文体の違いや、文章の論理構造の順序の違いなどの非本質的な、著者と読者の差異が取り払われることが予想される。また、元論文に、論理的欠陥や、意味の定義の不明瞭さがある場合は、その点もこの基本チャートによって、明らかになる。また、論文に存在する、理解が難しい文章構造(4.2節)も、この再構成によって明らかになる。

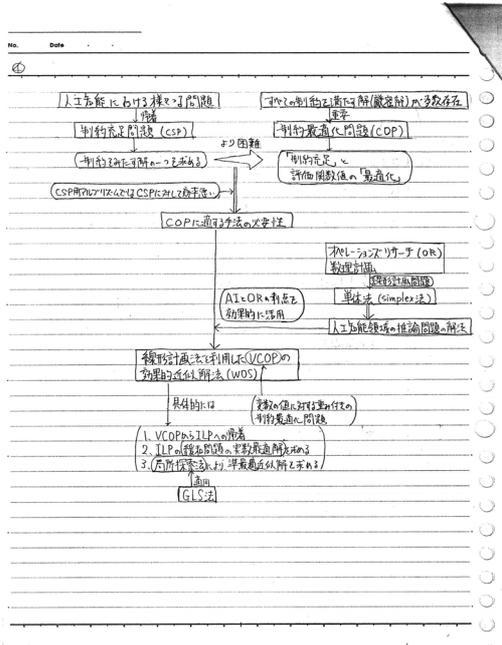


Figure 3: 被験者により提出された応用チャート 2

ここで、Figure1と同じ論文を違う被験者によってまとめてもらったFigure3と、Figure3を前節のFigure1とFigure2そして、Figure3とFigure4の関係を見てみると、それぞれ違う被験者によって提出されたチャートであるFigure1とFigure3は、同じ論文を表しているのに、かなり違った表現をとっていることが分かるが、それぞれをイメージ・スキーマに基づいた変換を施すと、論文中の重要部分(たとえばCOPはCSPよりも一般に困難であること)や、おおまかな構造が酷似してきていることが分かり、この変換が、被験者の違いを乗り越えたイメージを与えてくれることが期待できる。

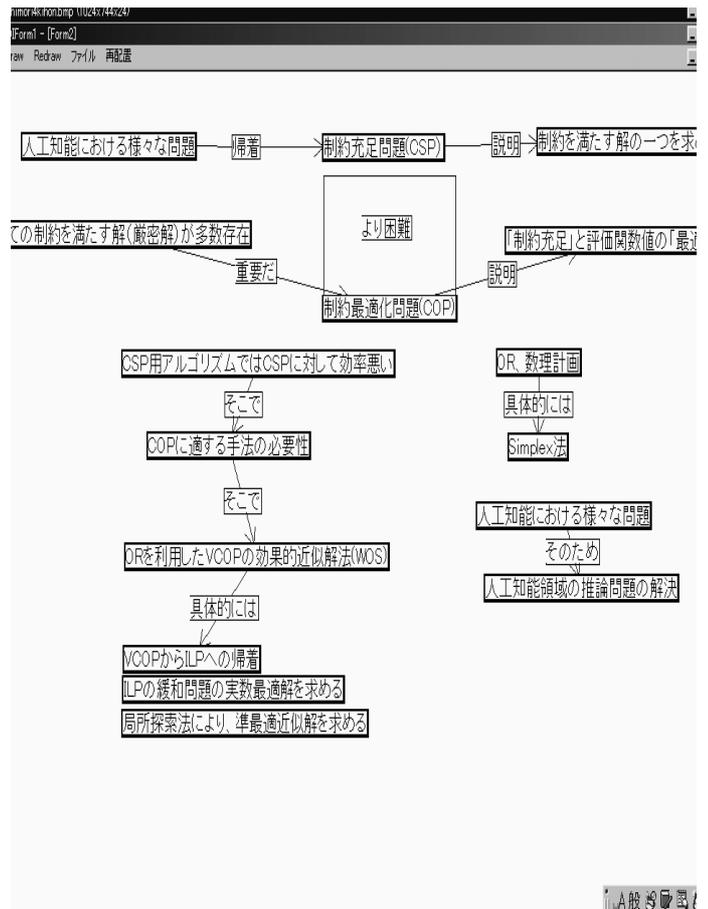


Figure 4: 再構成された基本チャート 2

4 チャート作成システム

4.1 システム化

以上の実験、考察、また[7]の内容をまとめ、イメージスキーマに基づくチャートを用いた論理解のプロセスを以下のようなシステムとする。

- ノード作り

まず、読んだ論文を要約すべくノードにメモをとる。ここでノードの数、大きさに1画面という制限が重要であるのには2つ理由がある。一つは、全ての情報を、一瞬のうちに視野に入れられる、ということ。もう一つは、制限があることで、情報の圧縮が行われ、より多くの情報を少ない記述で表すことが出来る、ということである。
- ノードをつなぐ

できたノードを、ラベル付き矢印で結んで、ノード間の関係を表す。この時、あらかじめ「技術論文スキーマ」「文章論理スキーマ」「因果関係スキーマ」の3つを用意しておいて、できるだけ、その中にあるものだけをラベルとして使って、チャートを作る。

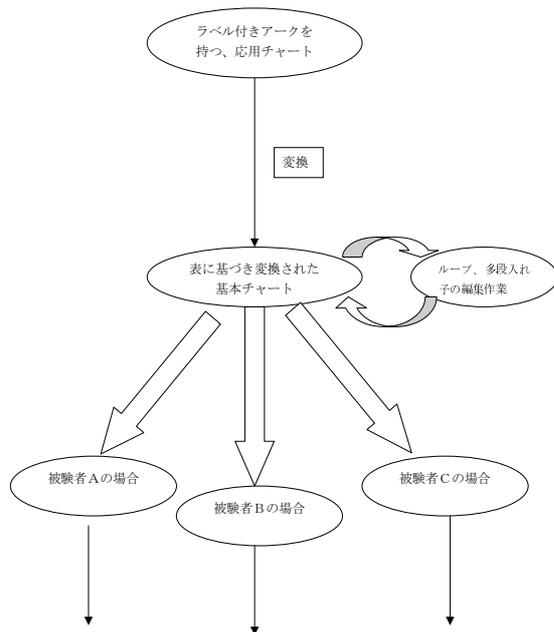


Figure 5: システムの概略図

どうしても、この3つに当てはまらないものを使いたいときは、個別に定義する。

- 整理する
ノード作りの際と同じように、1画面の制限を作ることで、全ての情報を一瞬のうちに視野に入れることが出来ること、なおかつ、情報の圧縮が出来るから、情報過多から逃れられる。
- チャートを違う視点から見なおす
「応用スキーマ」で関係を記述されたチャートを、「基本スキーマ」の立場から見なおすことによって、同じチャートを違う視点から見る事が出来、チャートの不備などを見つけ出すことに役に立ち、チャートの本質的理解にも役に立つ。

4.2 システムの実装

現在、システムのプロトタイプを、VisualBasic6.0 を用いて製作した。以下にこのプロトタイプ版の持つ機能について述べる。

(1) チャート作成機能：ノードを作り、動かし、繋げることが出来る。この時点で作り得るのは、ノードとラベル付きアークによる、単純な応用チャートである。この段階で、論文を読んで得た知識の外化をする。

(2) 基本チャートへの変換機能：手で作られた簡単な応用チャートを、応用スキーマと基本スキーマの対応表に基づき、新たな図形的な要素を持つ基本チャートに変換する。

Figure 6, Figure 7 に、それぞれ変換前の応用チャートと変換後の基本チャートを示す。このような応用チャートに対応する基本チャートを生成することで応用チャートに内在していた構造を明示化したり他の応用チャートと変換可能にすることができる。この例では、Figure 6 の「ポイント」が、Figure 7 では、＜中心/周縁＞のスキーマに基づいて、円の中心と周縁に位置されている。この変換によって、応用チャートで「ポイント」という関係で結ばれていた2つのノードが、基本チャートでは、身体性に基づいたイメージ・スキーマによる＜中心/周縁＞という関係によって結びつくことになり、より直感的に捉えやすいものとなり、また応用チャートでは明らかにならなかった本質的な関係、すなわち、「論文読解において、情報の要約が、中心的な役割を示す」ことの存在が図形的イメージによって明示化される。

(3) 構造に基づく推論機能：本システムでは、2つのチャートに対して、その構造を解析した結果に基づいた推論機能として、警告と、縮約の2つの機能を実装した。理解が難しいと考えられる構造としては、現時点では、ループ構造、入れ子構造を実装している。入れ子構造とは、例えば、応用チャートにおいては「... それで... それで... それで」の列や、あるいは、「Aにおける重要なことのポイントのキーワード」のような構造などのことである。この入れ子構造は、2つ目の例に挙げたもののように＜中心/周縁＞のスキーマである場合、直感的に分かりにくいので、基本チャートへの変換の際、修正するよう警告する。入れ子構造においても、＜経路＞のスキーマによるものは、最初と最後のノードをつないで、中途の部分を縮約することで、明確な理解が可能になる。

上の例は、応用チャートの段階では違うラベルに分類されているものが、基本チャートに変換されることによって、同じ基本スキーマ(例えば＜経路＞)に属していることが分かり、縮約が可能になるというようなケースもあることを示している。

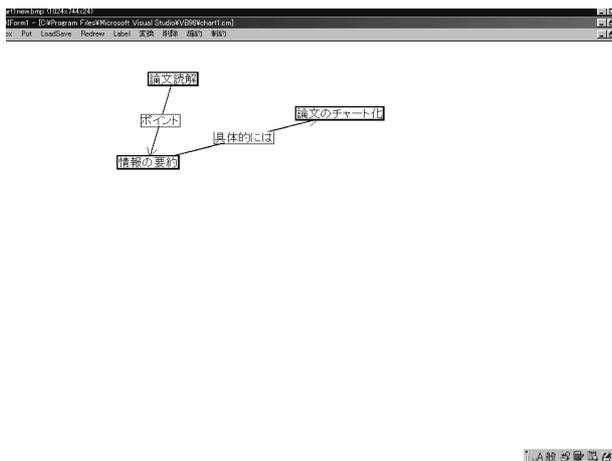


Figure 6: 変換前の応用チャート

5 今後の課題

応用チャートから基本チャートへの変換によって、より深いチャートの理解が得られるかを検証する。また、チャートを書くことによって、最終的に支援したい「論文を書く」ということの支援が可能なのかどうかを実験によって確かめる。

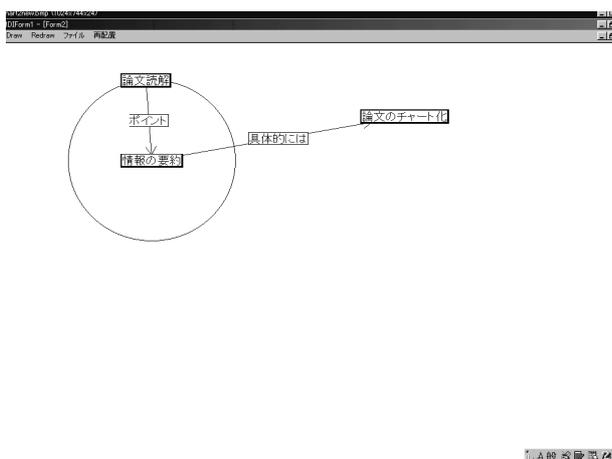


Figure 7: 変換後の基本チャート

参考文献

- [1] 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」プロセスの支援に向けて, 人工知能学会誌, Vol.14, No.1, pp82-92 (1999)
- [2] 遠藤聡志, 大内東: 統合型発想支援システム: FISM, 人工知能学会誌, Vol.8, No.5, pp611-618 (1993)

- [3] G.Lakoff, M.Johnson: レトリックと人生, 大修館書店 (1986)
- [4] G.Lakoff: 認知意味論 - 言語から見た人間の心, 紀伊国屋書店 (1993)
- [5] 住田一男, 知野哲朗, 小野顕司, 三池誠司: 文書構造解析に基づく自動抄録生成と検索提示
- [6] 岡孝明, 武田英明: 技術論文のチャート化による論理解の支援についての分析, 電子情報通信学会技術研究報告 AI99-61, pp27-34 (1999)
- [7] 立花隆: 「知」のソフトウェア, 講談社現代新書 (1984)
- [8] 稲垣博人: 事象解析による要約情報の抽出, 情報処理学会