修 士 論 文

収束的思考段階の構造を反映して
発想の支援を行うシステムの実現

指導教官　國藤進 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

野口裕史

1997 年 2 月 14 日

目 次

目 次

目 次

図一覧

図一覧

表一覧

第 1 章

序論

本論文では人間の思考をコンピュータで支援する研究について述べる. 人間の問題解決における発想プロセスではアイデアの生成と生成されたアイデア整理が行なわれるが, 実際には, このアイデアの生成と整理を一度だけでなく何度か繰り返し行うことにより発想の質が高められることが多い.

本研究ではこの点に着目し, 人間が一度アイデアを生成して整理した後に再びアイデアの生成に戻る過程をコンピュータにより支援する機構を提案し, 実験システムを構築した.

本章では, 研究の背景として人間の問題解決における発想のプロセスの特徴と発想支援システムについて述べ, 次に本研究の目的と研究の概要について述べる.

1.1 はじめに

人間の知的生産活動や創造的問題解決活動において, 発想のプロセスは最も上流の過程に位置する. これまでに情報処理技術, 人工知能技術などの発展や計算機の性能向上などにより, 人間の知的生産活動のいくつかの部分をコンピュータにより自動化する試みがなされてきたが, 発想プロセスは各個人が持つ経験や背景などが必要なため, 人間でなければできない活動として, コンピュータによる発想の自動化は夢の技術とされてきた.

しかし, コンピュータが急速に身近なものになるにつれ, さまざまな知的生産活動の現場からコンピュータ自体が発想をするのでなく人間の発想をコンピュータにより支援するという技術への期待が高まり, このうちのいくつかの技術は実現可能なものになってきている. 近年, 発想支援技術はコンピュータ科学, 人工知能の中の最も新しい研究分野の 1 つ

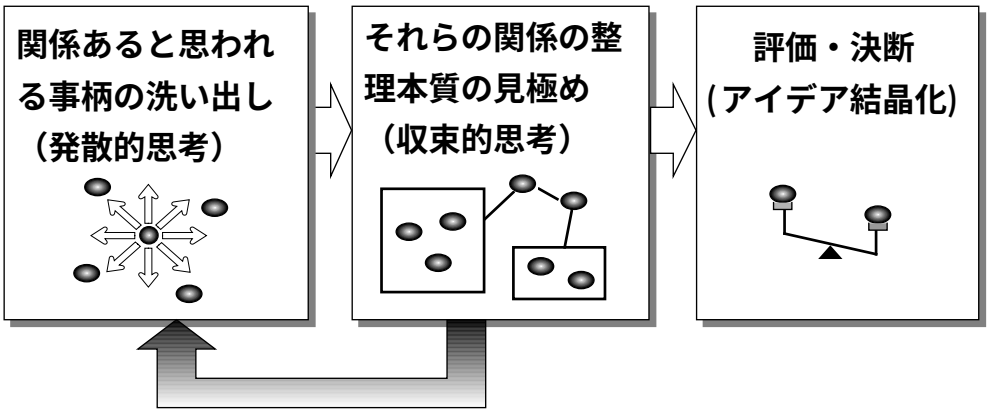

図 1.1: 発想のプロセス

として認知されつつある.

1.2 研究の背景

1.2.1 人間の問題解決における発想のプロセス

人間の問題解決における発想のプロセスは発散的思考, 収束的思考, アイデア結晶化の 3 段階に分けられる [5]. 発散的思考は, 頭の中でもやもやと問題に関連のあると思われるアイデアの断片の洗い出しを行う段階で, 情報収集活動なども含まれる. 次の収束的思考は, 発散的思考で出てきたアイデアの間に存在する関係をを整理して本質の追求をする段階である. 収束的思考を効率良く行うための方法論として KJ 法 [4] が有名である. 最後のアイデア結晶化は仮説評価・決断の段階で, 発散的思考, 収束的思考を通して得たアイデアから問題解決のための「ひらめき」が行なわれる.

この発想のプロセスを図にすると図 1.1のようにのようになるが, 実際の発想の現場においてはここで示すように発散的思考と収束的思考のそれぞれの段階を一度だけ行うのではなく繰り返し行うことによりアイデアの質が高められるということがよくある.

1.2.2 コンピュータによる発想支援

発散的思考, 収束的思考, アイデア結晶化からなる人間の発想のプロセスをを支援するコンピュータシステムは「発想支援システム」と呼ばれている. 発想支援システムは支援方法が主にアイデアを発散的に広げるか, 収束的に整理するかにより発散的思考支援システムと収束的思考支援システムに分けられる.

発散的思考支援システムとしては Keyword Associator[20] のようなキーワード提示機能を持つものが多い. Keyword Associator は連想辞書を持ち, キーワードやテキストから関連キーワードや関連テキストを提示することができる. 最近はこのようなアイデアの断片となる関連情報をただ単に与えるのではなく関係情報を 2 次元平面上へ配置してその配置の仕方により発散的思考を促進させるシステムも数多く研究されている [7][10].

それ以外にも発散的思考で重要な意味を持つ情報収集を考えると, 最近急速に広まってきた WWW の検索エンジンなども発散的思考を支援するための強力なツールとみなすことができる.

収束的思考支援システムについては日本では KJ 法の影響が強く D-ABDUCTOR[17], 群元 [19],ISOP(アイテック社) のように KJ 法図解を中心としたシステムが多い.

1.2.3 発想支援とデータベースからの知識発見

従来の発散的思考支援システムの多くはユーザに与えられたテーマに関連の強い情報を与えることによりユーザのアイデア生成の促進を試みてきた. これらはテキスト情報をはじめとする膨大な「情報資源」からキーワード間の関連の強さなどを計算して情報検索を行っていると言える. しかし発想に役立つ情報を得るための手段が関連度の高い情報の検索のみで良いのかという問題がある. そこで本研究では単に関連情報を得るだけでなく, ユーザの気付いていないような情報をより積極的に発見することを試みる. そのために最近盛んに研究されているデータベースからの知識発見の分野で利用されている技術を採り入れた.

データベースからの知識発見 (KDD：Knowledge Discovery in Database) は巨大データベースからの知識獲得を高速に行う技術で, 人工知能やデータ工学から派生した研究分野ある. データベースからの知識発見はデータマイニング (Data Mining:データ発掘) とも呼ばれる. 最近バーコード技術の進歩により商店の膨大な販売データを収集し記憶することが可能になったことにより, 膨大な情報資源の中に埋もれて隠れている情報の抽出をす

るこれらの技術はビジネスや知的生産活動の現場でも注目を集めている.

1.3 本研究の目的

発散的思考と収束的思考を繰り返すことによりアイデアが高められることがあるのは既に述べた通りである. そこで収束的思考段階にある人間が再び発散的思考段階に戻ってアイデアの生成を行う際に従来の発散的思考支援システムを利用することを考える. 従来の多くの発散的思考支援システムはあるテーマに関してユーザの発想の刺激となる関連情報を与えるものであったが, 発散的思考と収束的思考を繰り返し行うことを想定した場合でも単に関連情報をユーザに与える枠組だけで良いのかという疑問がある. つまり収束的思考から発散的思考に戻る際に, 従来の発散的思考支援システムをそのまま適用しただけではユーザの収束的思考段階での構造が無視されているので, そのままでは現実に即した支援を行っているとは言いがたい.

これまでの発想支援技術の研究では発散的思考から収束的思考へという枠組の支援機能は様々な形でなされてきたが収束的思考から発散的思考への支援機能は不十分であった. そこで本研究では収束的思考段階においてユーザの作成した図解に内在するユーザの気付かなかった情報をシステムが発見し, その図解を反映した方法で可視化することによりユーザの発散的思考の支援を行うシステムを提案し, 構築することを目的とする.

1.4 研究内容の概要

本研究では KJ 法の枠組を利用して発想の支援を行う. 具体的にはユーザの作成した KJ 法などの図解を元に, 前もって用意された関連のあるテキスト情報からユーザの気付いていない情報を発見し, 可視化することにより発想を促進させる. ユーザの気付いていない情報を見付けるために, 相関ルールの導出アルゴリズムというデータベースからの知識発見の技術をテキスト中のキーワードに対して適用し, 生成されたルールの中からユーザにとって役立つ情報のみを残すためのフィルタリング処理を開発した. さらに生成した情報をもとにユーザの発想の促進させるための可視化方法を開発した.

1.5 本論文の構成

本論文は本章も含め 6 章から構成される. 2 章では KJ 法の図解から発散的思考に役立つ関連情報を抽出して可視化するための機構について述べる. 3 章では 2 章で述べた機構を実装した実験システムについて述べる. ここでは各構成要素の特徴とユーザから見たシステムの利用の流れについて説明する. 4 章ではシステムの評価実験について述べる. 実験ではシステムの要素技術に関する評価とシステムによる利用者の思考レベルの評価について調べた. 5 章では関連研究との比較を行ない, 最後に本論文の結論と今後の課題について 6 章で述べる.

第 2 章

収束的思考段階の構造を反映した発散的思考支援機構

本研究では, 発散的思考と収束的思考を繰り返すことによりアイデアの質が向上することに着目し, 発想支援システムに収束的思考から発散的思考へ戻る際の支援をする機構を採り入れた. この方法は収束的思考段階において整理された思考を表す図解から情報を抽出し, そこから発散的思考の支援に役立つ関連情報の生成を行ない, 元の考えを考慮した方法で関連情報の提示することにより実現する.

本研究の目的は大きく分けて 2 つある. 1 つは, 従来の発散的思考支援システムのように単なる関連情報を生成するのではなく, ユーザの気付いていない情報をより積極的に発見すること. もう 1 つは, 発散的思考を支援する際に, ユーザの収束的思考段階の構造を反映させための可視化方法を行うことである.

本章でははじめに本研究において利用しているユーザの思考を図解にまとめるための方法論である KJ 法について述べた後, ここで述べた 2 つの目的を実現するための方法をそれぞれ述べる.

2.1 KJ 法

KJ 法 [4] はアイデアの断片をカードに書き出して図解にまとめることにより収束的思考を行うボトムアップ的な方法論である. KJ 法自体は問題提起, 現状把握から始まって結果の検証, 総括・味わいまでの問題解決のプロセス全体を対象にしているため, 情報収集の

ための取材活動なども含めた広範囲にわたる方法論であるが, 本研究ではこのうちの「狭義の KJ 法」と呼ばれる本質追求の過程のみ扱い, 以後単に KJ 法と呼ぶ. (狭義の)KJ 法の作業は以下のラベル作り, グループ編成, 図解化, 叙述化の各ステップからなる.

1. **ラベル作り**
   テーマに沿って素材となるデータを KJ ラベルと呼ばれるカードに書き写す.

2. **グループ編成**
   カードをランダムな順番に並べ (**ラベル広げ**), 並べられたカードを何度も繰り返し読む. これにより, お互い似ていると思うカードを近づけることによりグループ分けを行ない (**ラベル集め**), それぞれのグループに対してラベルが集まったゆえんを要約したものを別のカードに書き, そのラベルを一番上に乗せて輪ゴムで束ねる (**表札作り**). 全てのグループに表札をつけ終えたら, 束ねられた各グループを表札をラベルとした一枚のカードに見立て, 再びラベル広げに戻りこの作業を繰り返し, カードの束が数個になった時点で終了する.

3. **図解化**
   グループ編成を行なったカードの束を模造紙などの大きな紙の上で解き 2 次元上でバランス良く配置する (**空間配置**). その上で各グループを線で囲み (これを島と呼ぶ), 島間の関連づけを行うために関係線などをつける (**図解化**).

4. **叙述化**
   図解にして分かったことをストーリーにして文章にしたり口頭発表したりする.

以上のようなプロセスは一度だけ行うのでなく何度も繰り返すことにより問題をより深く掘りさげることができる. このように KJ 法を何ラウンドも繰り返す体系は川喜田により**累積 KJ 法**としてまとめられている.

2.2 関連情報を抽出する方法

本研究では, 発散的思考支援において収束的思考段階の構造を反映させるために, ユーザの作成した KJ 法の図解の関連情報を生成して可視化する. ここでは関連情報の生成方法について述べる.

本研究では関連情報の生成のために従来の発散的思考支援システムと同じように単語に着目している. 実際には,KJ 法の図解中のラベルに書かれている単語の中でキーワードになり得るものを抜きだし, 各キーワードに対してユーザの発想を促進させるようなキーワードの組を生成する. 従来のシステムはこの部分において類似度などを利用して与えられたキーワードに関係の深いキーワードを生成するものがほとんどであった. しかし, 本システムは従来のシステムのような関連の強いキーワードを生成するよりも一歩踏み込んで, キーワード間に潜んで隠れている関係をより積極的に発見することを試みるため, データベースからの知識発見の分野で使われている相関ルールの導出アルゴリズムを利用した. 以下の節では相関ルールの概要と本研究における関連情報としての相関ルールの特徴について述べ, 価値のある相関ルールのみを残すためのフィルタリングの処理について述べる.

2.2.1 相関ルール

相関ルール (Association Rule) は主にマーケティングにおける情報分析に利用されている. POS を用いることにより商店の膨大な販売情報を収集し, 記憶することが可能になった. 相関ルールはこれらの販売履歴データベースからから販売戦略に役立つ情報を生成するのに適している. 例えば図 2.1のようなデータベースから「この商店では, パンとバターを同時に購入する客は, ミルクも買うことが多い」というような情報を得ることができる.

相関ルールには支持度 (support) と確信度 (confidence) という値が付与されている.

支持度 相関ルールに含まれる全ての要素がデータベース中に出現する頻度を表す. 図 2.1の例の場合, パン, バター, ミルクが同時に出現する頻度が全体 (5 つ) のうちの 2 回あるため, 支持度は 40%となる.

確信度 データベースにおける相関ルールの正しさを表す. 図 2.1の例の場合, パン, バターが同時に出現するときにミルクも同時に出現するのが 3 回のうち 1 回なので, 確信度は 67%となる.

相関ルールの定義は以下のようになる.

定義 1 商品の集合を $\mathcal{I} = \{i_1, i_2, \ldots, i_m\}$, トランザクションの集合を $\mathcal{D}$ ($\mathcal{D}$の各要素 $\mathcal{T}$が $\mathcal{T} \subseteq \mathcal{I}$ を満たす) とすると, $X \subseteq \mathcal{I}, Y \subseteq \mathcal{I}, X \cap Y = \emptyset$ の時 $\mathcal{D}$の Xを含む要素の c%が

客	商品
1	{パン, バター, ミルク}
2	{ バター, 味噌 }
3	{ パン, 味噌 }
4	{パン, バター, 味噌, ミルク}
5	{パン, 米, バター}

図 2.1: 販売履歴情報の例

Yも含み, $\mathcal{D}$の要素の $s\%$が $X \cup Y$も含むならば, $X \Rightarrow Y$ を支持度 s, 確信度 c の相関ルールと呼ぶ.

支持度と確信度が与えられた値以上になる全ての相関ルールを生成すること相関ルールの導出という.

ここで, 生成したい相関ルールの支持度と確信度の最小値がそれぞれ $minsup$, $minconf$ であるとすると, 相関ルールの導出は以下のような 2 つのサブ問題に分割することができる.

1. 支持度が $minsup$ 以上の要素の組 (ラージアイテム集合) を全て発見する.

2. 全てのラージアイテム集合から確信度が $minconf$ 以上となる相関ルールを生成する.

サブ問題 1 を簡単に求めることを考えると, 表 2.1のように全ての要素の組合せのアイテム集合に関してデータベース中における出現回数を数えれば良い. しかしこのやり方だと, 要素数が m の時に組合せの数が 2^m となり, 通常は要素数 m の値が数千から数万程度あることを考えると現実的ではない.

そこで, この問題を高速に解くためのアルゴリズムがいくつか提案されており, 本研究ではそのうちの Apriori アルゴリズム [6] を利用している. Apriori アルゴリズムは高速化のために図 2.2のようにラージアイテム集合を要素数が少ない順に生成する. これは, ラージアイテム集合の全ての部分集合はラージアイテム集合であることを利用し, 要素数 i のラージアイテム集合から要素数 $i+1$ のラージアイテム集合になる候補を生成し, 生成した候補だけに対してデータベース中の出現回数を数えることにより不要な組合せのための処理をカットしている. Apriori アルゴリズムの詳細は付録 Aに載せる.

表 2.1: ラージアイテム集合を単純に求める方法

パン	バター	ミルク	味噌	米	出現回数
0	0	0	0	1	1
0	0	0	1	0	3
0	0	0	1	1	0
0	0	1	0	0	2
0	0	1	0	1	0
0	0	1	1	0	1
0	0	1	1	1	0
0	1	0	0	0	4
0	1	0	0	1	1
		⋮			⋮
		⋮			⋮

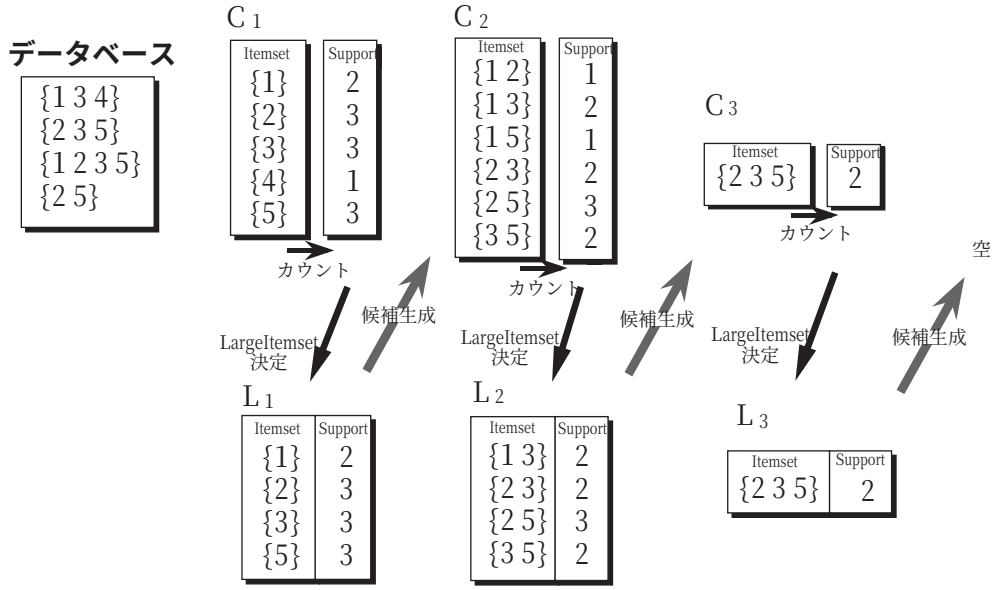

図 2.2: apriori アルゴリズムによるラージアイテム集合の生成

2.2.2 テーマ関連情報としての相関ルール

本研究ではユーザの作成した KJ 法図解の関連情報を作成するためにこれまで述べてきた相関ルールの導出アルゴリズムをテキスト中のキーワードに対して適用する方法を利用した. ここでは, テーマ関連情報としてのキーワードに関する相関ルールの特徴と相関ルールの導出アルゴリズムをテキスト中のキーワードに適用する方法について述べる.

テキスト中のキーワードに関する相関ルール

キーワードに関する相関ルールを生成すると, 電子ニュースなどの膨大な量のテキスト集合の中から

「キーワード $x_1, \ldots, x_i$ が同時に出現する文章には $y_1, \ldots, y_j$ も出現することが多い」

という情報を得ることができる.

キーワードについての相関ルールはテキスト中のキーワードの出現の仕方の統計情報 (キーワードの出現の偏り) をもとにして生成される. キーワードの出現の偏りは記事中の話題から作られるものが多いため, これらを何らかの概念を表したものであると解釈することができる. 本研究でテキスト中のキーワードに関する相関ルールを利用するのは, このような何らかの概念を内在するルールをユーザに提示することが発想の促進に役立つと考えられるからである.

キーワードファイルの自動構築

キーワードに関する相関ルールを生成するために, 電子ニュースなどのテキストに同時に出現するキーワードを記事ごとに並べたもの (キーワードファイル) を利用する. キーワードファイルは商店における販売履歴情報 (図 2.1) に対応づけることができる.

キーワードファイルを作成する手順を図 2.3に示す. はじめにシステムは各テキストからキーワードの切り出しを行う. このとき助詞, 助動詞, 接続詞, 記号文字などのキーワードに成り得ないものを無視し, それ以外の単語から不要語リスト法 [1] を用いて不要語を削除する. これにより残される単語を辞書順に並べたものがキーワードファイルとなる. ただし, テキスト中のキーワードの情報は, 相関ルールが主に利用されている販売履歴情報などと比べて個々の情報に含まれる各要素が強い関連を持つという特有の性質があるた

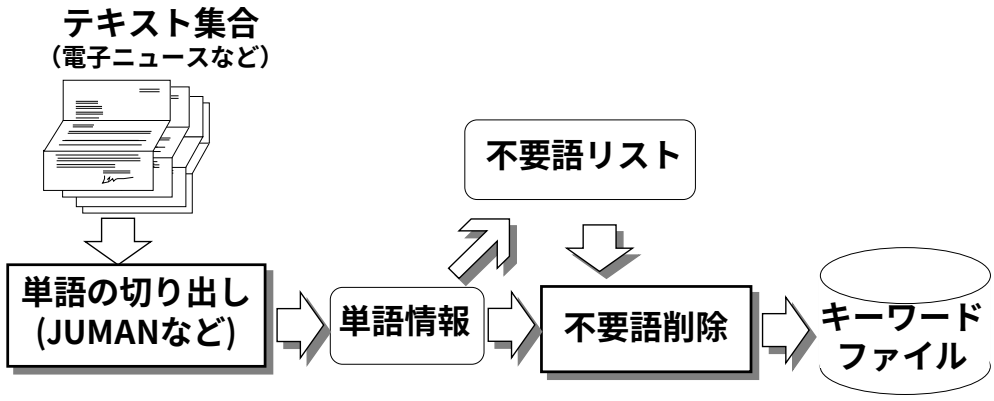

図 2.3: キーワードファイル作成の手順

め, キーワードファイルをそのまま用いて相関ルールを生成すると組合せ爆発の問題が生じることがある.

この問題は, キーワードファイルに含まれるラージアイテム集合 (キーワードの組に常に同時に出現するキーワードの組) があると生じる. シグネチャや定型文に含まれるキーワードがこれに該当する. 例えば, 本学のローカルニュースにおける就職の求人情報の定型文には「一連番号, 社名, 事業, 内容, 希望, 求人, 職種, 業務, 学校, 推薦, 期限, 特記事項, 連絡」などが含まれている. この常に同時に出現するキーワードの組 (ラージアイテム集合) のことを以下ではヒュージアイテム集合と呼ぶことにする. このようなサイズが大きなヒュージアイテム集合が複数存在するキーワードファイルについて考えると, 相関ルールの生成に必要なラージアイテム集合を生成する際の計算コストが膨大になり, 意味の無い不要なルールも数多く生成されてしまう. そこで, テキストからキーワードを抜き出す際に予め引用文やシグネチャのカットを行い, さらにキーワードファイルからヒュージアイテム集合を抽出し, これを他の 1 語に置き換えるという処理を施す. ここで, ヒュージアイテム集合を以下のように定義する.

定義 2 サイズ n のラージアイテム集合を $\mathcal{L} = I_1 I_2 \ldots I_n$ とすると, 全ての $I_i \in \mathcal{L}$ について, $\mathcal{L}$ の支持度と $(\mathcal{L} - \{I_i\})$ の支持度が等しい場合にラージアイテム集合 $\mathcal{L}$ をヒュージアイテム集合である.

ヒュージアイテム集合の抽出は以下のように行っている. はじめは, 指示度の最小値 (min-sup) を求めたい値より大きく設定してラージアイテム集合を生成し, その中からヒュージ

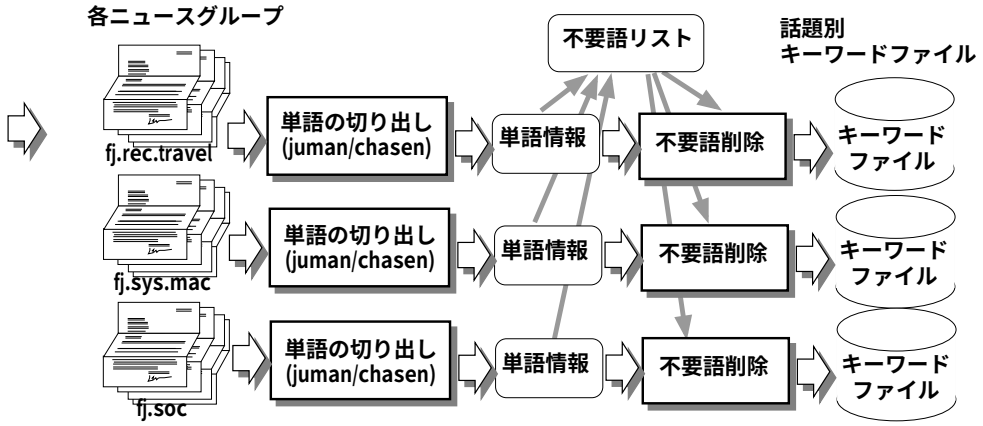

図 2.4: 話題別キーワードファイル作成の手順

アイテム集合を探す. ここで生成されたヒュージアイテム集合をキーワードファイルから抜きだして他の 1 語に置き換える. この処理を支持度を少しづつ少なくしながら繰り返し, 最終的に minsup が求めたい値になった時点で終了する. このようにヒュージアイテム集合の生成を一度に行なわず, 少しずつ行うことにより組合せ爆発を防いでいる.

話題別キーワードファイル

システムは図2.4のようにニュースグループごとに複数のキーワードファイルを作り, 話題別キーワードファイルとする. 話題別という名をつけたのは fj.rec.food ではたべものに関して, fj.soc は社会問題に関してというようにニュースグループによって扱う話題が異なるためである. 複数の話題別にキーワードファイルを用意することにより, ユーザは図解の話題に応じて一つ, または複数のキーワードファイルを選択することができる.

ただし, ユーザの扱いたい話題とニュースグループが 1 対 1 に対応するわけでない. ニュースグループに存在しないことがらに関するルールは生成されないので, 本方式には扱うことのできる話題が限定されるという欠点がある.

2.2.3 フィルタリング処理

抽出された相関ルールのフィルタリングは様々なレベル行われるが, これらは以下の 2 つの方法に分けることができる.

- ユーザの作成した図解と比べてそのルールがユーザにどの程度必要かというヒューリスティックスを用いた方法
- ルールの形・確信度・支持度をもとして冗長なルールを削除する方法

前者の方法については, 例えば「ユーザが既に理解していると判断できるルール (図解中にルールに含まれるキーワードを全て持つラベルが存在する場合など) は価値が低い」というようなヒューリスティックスを用いる.

後者の方法については, 以下のような基準でルールのフィルタリングを行っている.

1. $conf(A \Rightarrow R) < sup(R)$ ならば $A \Rightarrow R$を削除.
2. $conf(AL \Rightarrow R) \leq conf(L \Rightarrow R)$ ならば $AL \Rightarrow R$を削除.
3. $conf(L \Rightarrow AR) = conf(L \Rightarrow R)$ ならば $L \Rightarrow R$を削除.
4. $conf(L \Rightarrow R) \times conf(L \Rightarrow A) > conf(L \Rightarrow AR)$ ならば $L \Rightarrow AR$を削除.

(conf,sup はそれぞれ確信度, 支持度を意味する.)

1,2 は A の出現が確信度の上昇に貢献しないルールの削除, 3 はより制約の厳しいルール ($L \Rightarrow AR$) から導かれるルールの削除, 4 はより単純な 2 つのルールから導かれるルールの削除を行っている.

これらの式にあてはまるものは, 価値が無いと見なして削除され, 当てはまらなければ式の両辺の値が異なれば異なるほど片方のルールは他のルールからは得られない情報を与えるものとして価値があることになる. 逆に, 両辺の式の値があまり変わらなければ片方のルールは他のルールから予想できてしまうため価値があるとは言えない. そこで, 値がどの程度異なれば価値があるとするかを判断するために統計的検定を用いた基準 [15] も利用している (詳細は付録 Bに載せる).

関連情報の抽出とフィルタリングの処理をまとめると図 2.5の通りになる.

2.3　ユーザの視点の転換をはかるための可視化方法

ここでは, 生成された関連情報をユーザの作成した図解を反映して可視化する方法について述べる.

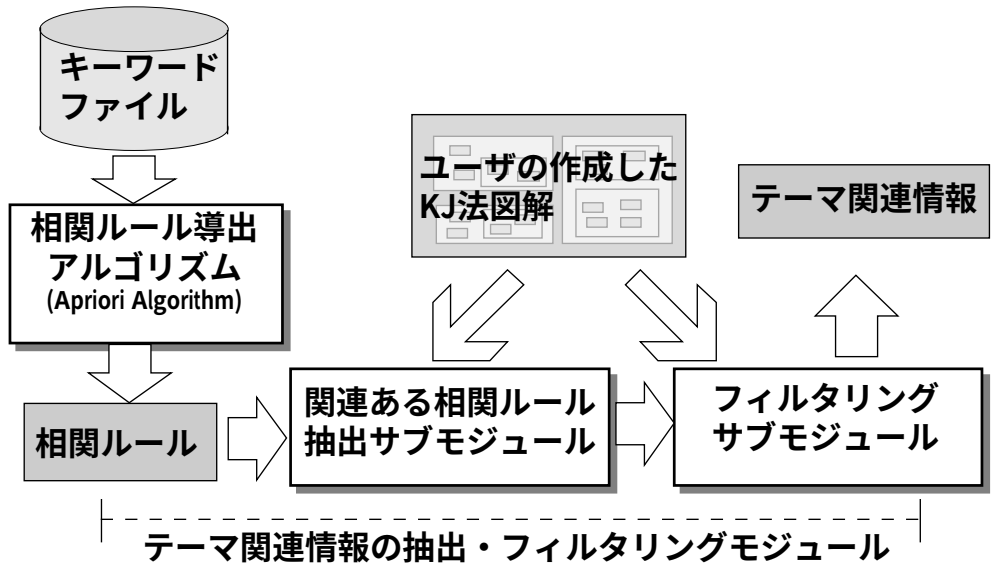

図 2.5: 関連情報を生成する方法

従来の発散的思考支援システムは, ユーザの設定したテーマからユーザの発散的思考を刺激するような関連情報を生成し, 提示するものが多かった. 最近はこれらに加えて生成された関連情報を直観的に分かりやすくするために 2 次元空間上に配置する研究も数多くされている. これらは, それぞれのキーワードやテキストについて類似度や双対尺度を求め, ばねモデルなどを利用して空間的配置をしている.

このような関連情報生成と可視化の方法は発散的思考の初めの段階においては効力を発揮するが, 本研究で対象とするように繰り返し思考活動を行うことを考えると得られた配置は収束的思考段階でのユーザの考えと比べて必ずしも直観に合っているとは言えない. これらのシステムはユーザ考えとは異なる視点からの配置を提供することにより発想の転換を促進させることを目的としているものであるが, 同じテーマを与えさえすれば収束的思考を行う前でも同じような配置を得ることが可能なため, そのシステムを繰り返し使っているユーザにとっては目新しい配置が得られない.

このような点から, 本研究ではシステムが生成した関連情報をユーザの収束的思考段階における構造を反映して可視化するために, 関連情報をユーザの作成した図解に埋め込む方法と関連情報をもとにユーザの作成した図解自体を変形させるという 2 つの方法を用いた.

2.3.1 関連情報の図解への埋め込み

既に述べたように, 従来の発散的思考支援システムによる連想キーワードのようなテーマ関連情報をユーザに提示するための方法のほとんどのものは, リスト形式を用いた方法と 2 次元空間上での配置による方法に分けることができる. それぞれ以下のような特徴がある.

リスト形式による提示の特徴

リスト形式ではヒューリスティックスを用いてソートをかけることによりテーマとの関連の強さや重要さのランキングを表現することができる. また, 関連情報が大量にあってもインターフェースを工夫することにより表示することが可能である.

2 次元空間上での配置の特徴

一度に表示することのできる関連情報の量は制限があるが, それぞれの関連情報に対して内容的にどのような関係があるかを示すことができ, 内容に関して空間上でおおまかな分類がなされるのでユーザの興味のある部分に注目することができる.

しかし, これらの方法によるランキングや配置などはシステムがあらかじめ持っている情報を元に実現しているためにユーザの思考を反映したものではない. 何度も繰り返し思考を行っているユーザにとっては, 生成される各関連情報の関係を示すのではなく, それぞれの関連情報がユーザの考えとどのような関係があるかを示したほうが良い. そこで, 本研究では図 2.6のように提示する関連情報をユーザの作成した図解の関係する部分に直接埋め込むことによりそれぞれの関連情報がユーザの作成した図解とどの部分に関係しているかを明らかにするという方法を採用した.

この方式を用いることにより, 2 次元空間配置の利点である内容により分けて表示することを可能とし, それぞれの埋め込む場所ごとにソートすることによりリスト形式での提示の利点である関連の強さや重要さによるランキングを表現できる.

この方法では, 図解全体に関係のある情報は外に近い島に, 図解中のある偏った部分に関係のある強い情報ほど内側の島に配置される.

各関連情報を図解のどの位置に配置するためのアルゴリズムを以下に示す.

ここで, 対象となる関連情報と図解中の各ラベルとの関係の強さは数値で表わされ, 図解中の各島 i と関連情報の関係の強さ a_i は, その島に含まれる全てのラベルの関連情報との関係の強さの総和であるとする. 関連情報と図解中の各ラベルとの関係の強さは, 各

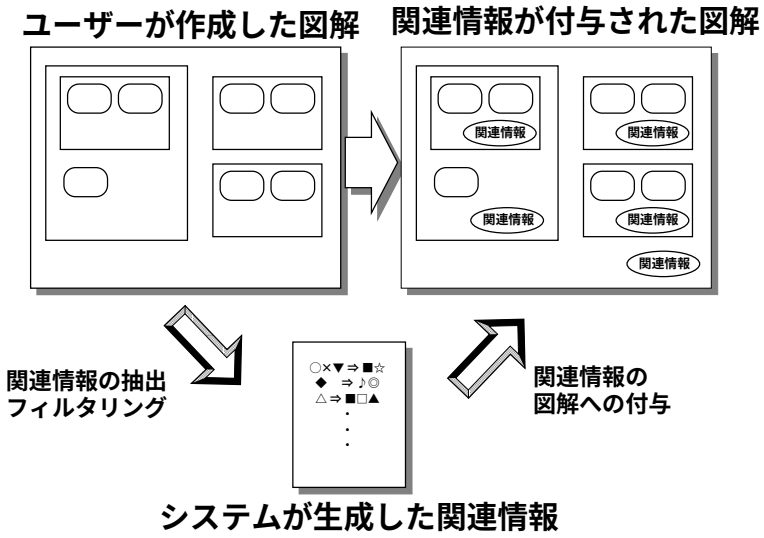

図 2.6: 関連情報の解釈支援

ラベルが関連情報の中のキーワードをどれだけ含むかによって自動的に計算される. このとき, 図解に頻繁に出現するキーワードほど影響が少なくなるように計算される. また, 各ラベルの関係の強さは図解全体の関係の強さが 1 になるように正規化される.

Step1: KJ 法の図解を木構造に変換し, その木のルートを x とする.

Step2: x の各ノード i の中から関係の強さ a_i が最大となるノード m を選ぶ.

Step3: もしも $a_m \geq S$ ならば x に対応する関連情報を配置する位置をノード x に対応する島に決定して終了. そうでなければ x に m を代入して Step2 に戻る.

S はどこまで厳密に行うかを決めるための閾値で,$0 < S \leq 1$ の値を取り得る. $S = 1$ ならば関連のある全てのラベルを包含する島に配置され, S が小さくなるほど関連情報の話題の中心に近い部分に配置されるが, 中心から離れたところに存在するラベルは無視される. S の値はユーザが指定することができる.

2.3.2 関連情報の重ね合わせによる図解の変形

発散的思考はアイデアの断片を生成するプロセスであるので, これまでに述べてきたような関連情報を提示するという方法で支援するのが一般的だった. しかし, 本研究の支援

の対象である一度収束的思考を行なったユーザは, 既に問題をある方向からの視点で見た構造を持っているため, システムがどのような関連情報を与えてもユーザにとってはその視点の方向に関係のある情報しか興味が無い状態に陥っている場合が多い. このような状態にある人間に対しては, 関連情報を与えるための支援をするより, むしろ視点を変えて, これまで見てきた情報を別の側面から見ることによって今まで見えていなかった新たなアイデアを引き出すことができると考えられる. そこで, ここではユーザの作成した図解を関連情報を元に変形することにより発想を促進させる方法について述べる.

ファジィ類似度関係を用いた KJ 法図解の重ね合わせ

文献 [2] において, 複数の KJ 法の図解をファジー類似度関係を利用してマージすることにより差異や共通部を可視化するアルゴリズムが提案されている. 本システムはこのアルゴリズムを用いて関連情報をもとに島を作りユーザの作成した図解とマージして図解を変形することによりユーザの視点の転換を支援する (図 2.7). ここでは KJ 法図解のマージアルゴリズムの概略について触れ, 詳細は付録 Cで述べる

ファジィ類似関係とは x と y の類似度 $\mu_R(x, y)$ が以下の条件を満たす関係である.

(a) **反射性**: $\mu_R(x, x) = 1$

(b) **対称性**: $\mu_R(x, y) = \mu_R(y, x)$

(c) **推移性**: $\mu_R(x, y) \leq \vee_y\{\mu_R(x, z) \wedge \mu_R(z, y)\}$

このような性質のある $\mu_R(x, y)$ を要素として行列に表したものをファジィ類似度行列という. ファジィ類似度行列は木構造との相互変換が可能である. そのため,KJ 法の図解の階層構造のみに着目して木構造とみなすと, 図解をファジィ類似度行列に変換することができる. KJ 法図解の階層構造 (木構造) をファジィ類似度行列にするために階層の深さに応じてレベルをつける. ここでは, 深さ i 番目のレベルを a_i とすると $a_i = \frac{i}{MC}$ (MC は最大の階層数) としている. このレベルを用いると, ファジィ類似度行列の各要素 $\mu_R(x, y)$ は, ノード x と y が共有する階層の深さを i としたときのレベル a_i となる.

KJ 法図解の重ね合わせるためには, このような方法で各図解をファジィ類似度行列に変換してから, 生成された各行列を合成する. 合成によって得られた行列を木構造に変換して図解に戻すことにより重ね合わせた図解が得られる. ファジィ類似度行列の合成は,

それぞれの行列の各要素の算術平均を行なった値を値の大きいものからあてはめ, 推移性の条件により自動的に決まる部分に値を埋めていくことにより作成する.

システムによる島の作成

本研究ではユーザの視点の転換を促進させるためにシステムが自動的に図解を作り, その図解とユーザの作成した図解をこれまでに述べた方法を用いて重ね合わせることによりユーザの図解を歪める. この処理のイメージを図 2.7に示す. システムはユーザの作成した KJ 法の図解から抽出したキーワードに対する関連情報 (実際には相関ルールの形式で表現される) を生成する. そして, 生成された各関連情報に対して自動的に島づくりを行ない, 関連情報の数だけ図解が作られる. (システムが生成した関連情報の数が多いときは, 各関連情報に含まれる単語の情報をもとにしたヒューリスティックスを用いてソートしたときの上位のものだけを扱っている.) 最後に, 得られたそれぞれの図解とユーザの作成した図解の重ね合わせをおこなうことにより新たな図解を得る.

島の作成方法は複数取り得る. 最も単純な方法は,KJ 法の各ラベルのうち, 関連情報と関係のあるラベルのみを一つのグループとして島を作る方法である. しかし, 関連情報と各ラベルとの関係の強さは, 各ラベルが持つ関連情報の中のキーワードと関係の深いキーワードから計算されるため, この方法による島作りは単に特定のキーワードを含むか含まないかという情報しか反映されず, あまり意味のある島作りを行うことができない. 従って, 実験システムでは以下のようなユーザによる島作りを反映した方法を用いた.

各関連情報に対して, 図解中のラベルを以下のように分ける.

(A1) 関連情報と関係のある (または関係の強い) ラベル

(A2) 関連情報と関係の無い (または関係の弱い) ラベル

先ほど述べた単純な島作りの方法はこの A1 にあてはまる全てのラベルをグループにまとめ,A2 にあてはまるラベルに関してはそのままにした図解である.

また,2.3.1節において関連情報を図解に埋め込む処理における情報を利用して以下のような別の分け方をする.

(B1) 関連情報を埋め込んだ島の中にあるラベル

(B2) 関連情報を埋め込んだ島の中に無いラベル

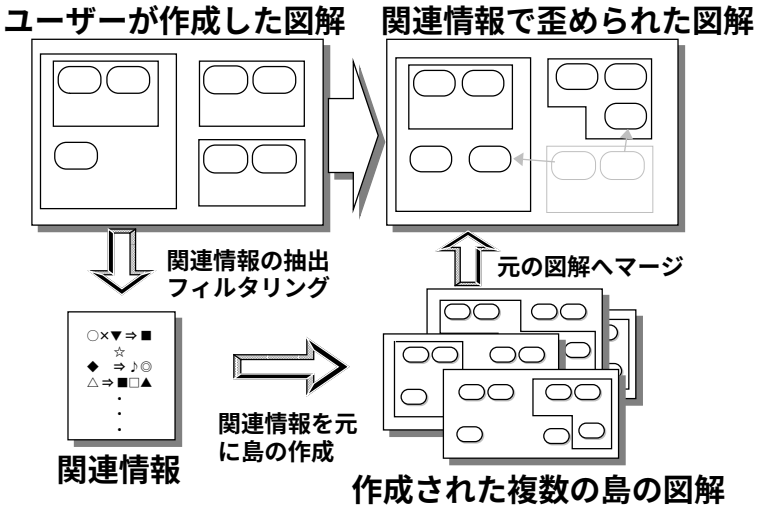

図 2.7: 関連情報をもとに図解を変形する

B1 に含まれる各ラベルはたとえキーワード情報からは関連情報と直接関係があるとされなくてもユーザが一つのグループにまとめたということから関連情報と何らかの関係があると考えることができる. そこで,A1 か B1 のどちらかに該当するラベルをグループにまとめることによりキーワードから関係あると判断できるラベルとユーザが関係あると判断したラベルの両方を採り入れる. このようにして, 各関連情報ごとに島を 1 つもつ図解を 1 つづつ作成し, これともとの図解をマージすることによりユーザの図解を変形させる. もとの図解とシステムが生成した島をマージする際, もとの図解の重みを大きくすることにより, 必要以上に図解が変形されることを防ぐことができる.

第 3 章

実験システムの構築

これまでに, ユーザの収束的思考を反映して発散的思考を支援する方法について述べてきた. これらの方法を用いて KJ 法をベースとした発想支援の実験システムを構築した [13][14]. 本システムはユーザが作成した KJ 法の図解から関連情報を生成し, それを提示することによりユーザの発散的思考を支援する.

本章では作成した実験システムの構成と各機能について述べた後にユーザからみたシステムの利用の手順について述べる.

3.1 システムの構成

システムの構成は図 3.1に示すようにキーワードファイルの自動構築モジュール, テーマ関連情報の抽出・フィルタリングモジュール, 生成情報の可視化モジュールからなる.

システムはあらかじめ電子ニュースのテキストから単語を抜きだし, キーワードファイルの自動構築を行う (**キーワードファイル自動構築モジュール**). ここで, さまざまな分野の話題に対応するためにキーワードファイルの作成をニュースグループごとに行ない, 話題別キーワードファイルとする. そして, それぞれの話題別キーワードファイルごとに相関ルールの生成を行ないファイルに格納しておく.

ユーザは使用するキーワードファイルを選択し, テーマに関して KJ 法の図解を作成する. KJ 法の図解作成には D-ABDUCTOR[17] を利用した. 図解作成時において, ユーザは D-ABDUCTOR が持つ直接操作, 図解の自動配置描画, フィッシュアイなどの図解作成支援機能 [8] を利用することが可能である.

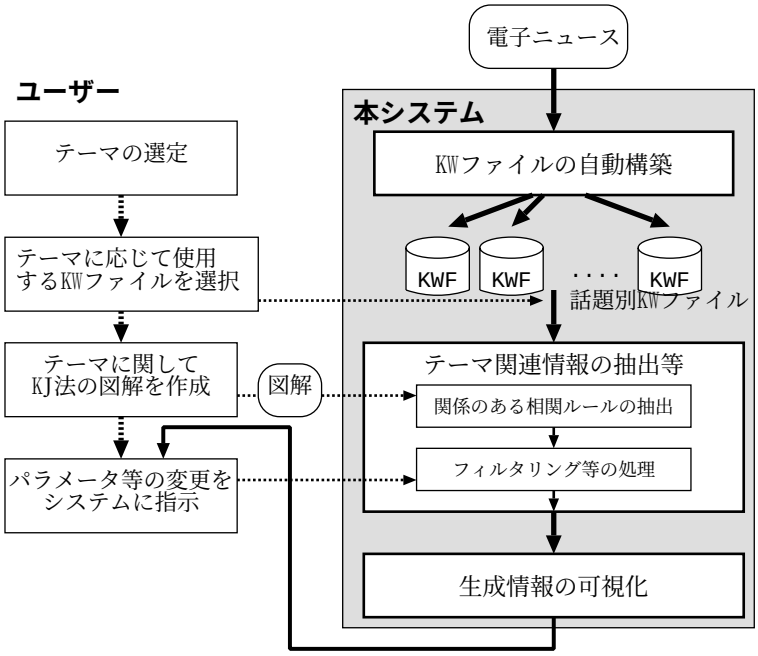

図 3.1: システム構成

システムはユーザが選択したキーワードファイルから作られた相関ルールのうち図解中のキーワードに関連するものをテーマ関連情報として生成する (**関連情報の抽出・フィルタリングモジュール**). これにより得られたテーマ関連情報をユーザの思考を反映した方法を用いて可視化する (**関連情報の可視化モジュール**).

3.2 システムの実装環境

これまでに述べたシステム構成による実験システムそ作成した.

開発は Sun Sparkstation 5 上で C++及び Perl を用いて開発した. ユーザインターフェイス部分は Motif を用いている. KJ 法支援ツールとして D-ABDUCTOR を用い, KJ 法図解からの情報抽出とキーワードファイル自動構築におけるキーワードの切り出しには日本語形態素解析システム茶筌 [16] を用いた. 茶筌は EDR 電子化辞書の日本語単語辞書 [12] を用いて辞書を強化したものを利用している. また, 日本語入力インターフェイスを mule 上で動作させるために一部 emacs-lisp も用いている.

図 3.2: 実験システムの画面例

3.3　システムの持つ発想支援機能

ここでは実装したシステムにユーザの発散的思考を支援するために組み込んだ機能について述べる.

本システムにおいて発散的思考を支援するための機能は以下の 2 つの観点をもとに分類することができる.

- アイデアの断片を生成するためにどのような情報を見せるか (与えられたテーマからどのような情報を生成するか)
- 情報をどのように可視化するか

以下の節において本システムで発想支援のために組み込んだ機能をここであげた 2 つの面 (何を, どのように表示するのか) に着目して紹介する.

3.3.1　システムが生成する情報

関連情報の生成に関しては, 以下のような機能を持たせた.

1. **関連情報 (相関ルール) 提示機能**
   テーマとなるキーワードに関連する相関ルールをユーザに提示する. ここで, 関連するキーワードとはテーマとなるキーワードを直接含む相関ルールだけでなく, 間接的に関連のある相関ルールも得ることができる.

2. **連想キーワード 提示機能**
   関連情報に含まれるキーワードを連想キーワードとしている. ここで, ユーザの指定したキーワード (X) と各連想キーワード (Y) の関連のしかたによって以下の 3 つに分けて提示することができる.

   (a) X から Y が連想される.(Y から X は連想されない.)

   (b) Y から X が連想される.(X から Y は連想されない.)

   (c) X から Y が連想され,Y から X も連想される.

3. **関連テキスト 検索機能**
   テーマとなるキーワードを電子ニュースの記事から検索して表示することができる.

図 3.3: テーマ関連情報の例

3.3.2 可視化に関する機能

生成された関連情報は図 3.3のような形式で得られるが, ユーザにとっては必ずしも理解しやすい出力であるとは言えない. これは, テーマに間接的に関連のあるルールは数多く存在するためシステムによるフィルタリングのみではユーザがその中から興味のあるものを見つけるのが困難となるからである. 本システムではユーザに分かりやすく提示するために以下のような可視化機能を実現した.

1. **リスト形式による表示機能**
   ルールの出力は, ユーザの興味に近いものが上位に来るように順番を並び替える. また, 一度に生成するルールはユーザが注目しているキーワード集合に直接的に関連するものに限定し, 生成されたルールの中でユーザが面白いと思うものを複数個選択することにより, ユーザが興味のある方向を直観をもとにたどっていくことができる.(図 3.4)

2. **相関ルールのグラフ表示機能**
   相関ルールは各キーワードにどのような関係があるのかを示しているが, リスト形

図 3.4: リスト形式のインターフェイス

式の情報だけでは複数のルールが複雑に関係するようになるとわかりにくい. そこで, グラフ形式を用いてキーワードの関係を表示させることができる. (図 3.5)

3. **関連情報の図解への埋め込みによる表示機能**
   関連情報をユーザの作成した図解に埋め込むことにより, それぞれの関連情報が図解のどの部分に関連があるかをあきらかにすることができる. 図 3.6の図解に関連情報を埋め込むと図 3.7の通りになる. (ここでは楕円に囲まれているものが関連情報である.)

4. **関連情報をもとにした図解の書き換え機能**
   生成した関連情報自体を提示するのではなく, ユーザの視点を転換させるために関連情報を用いてユーザの作成した図解を変形させる. 図 3.8は図 3.6を変型した例である.

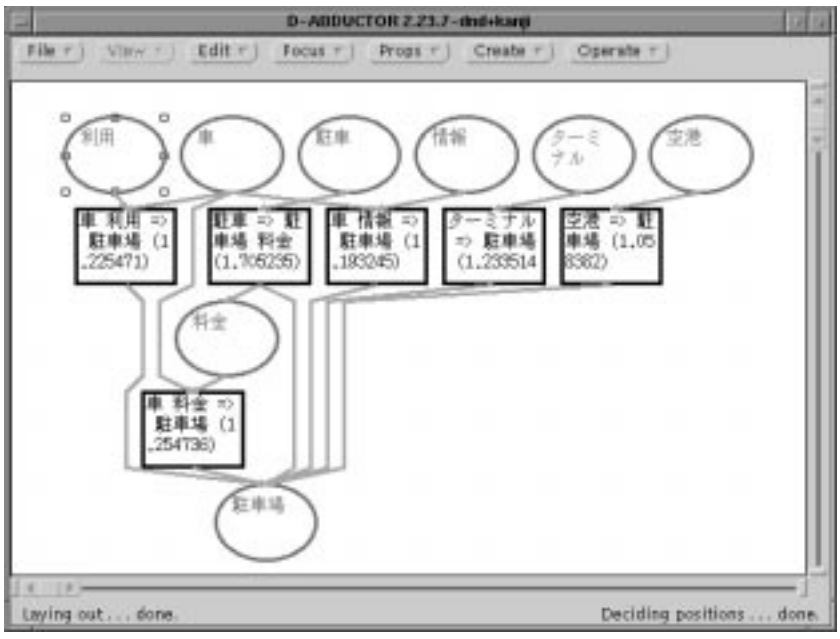

図 3.5: グラフ形式による表示

図 3.6: もとの図解

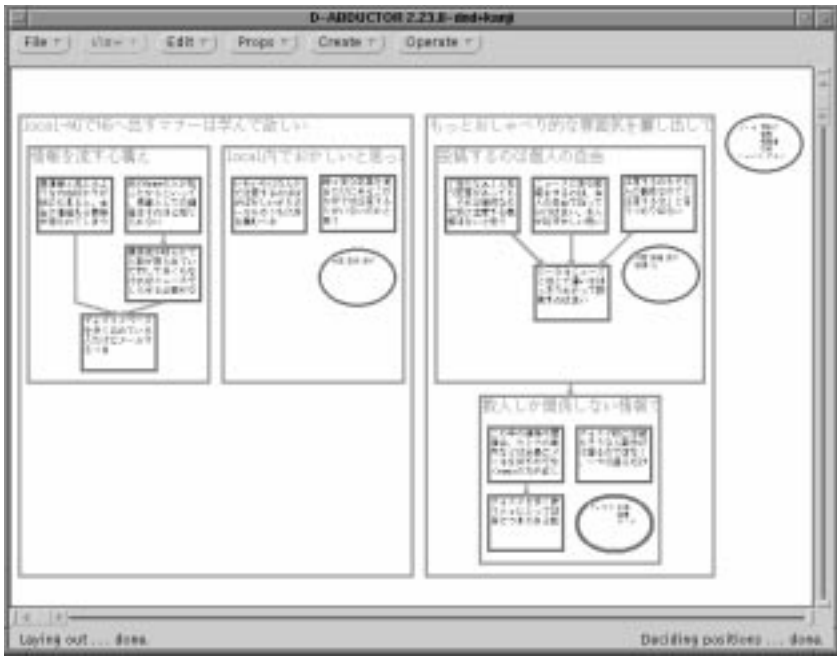

図 3.7: 関連情報を付与した図解

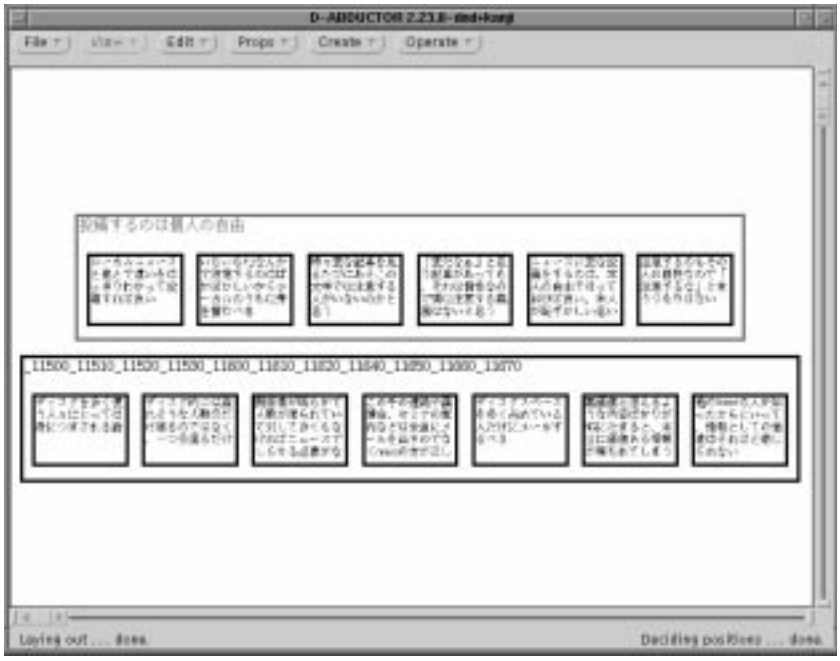

図 3.8: 変型された図解

3.4　システム実行のプロセス

これまでにシステムの持つ機能について述べてきた. ここではこれらの機能がどのように利用されるかについてユーザの立場から見たシステムの動きについて述べる.

本システムの利用の手順は, 前準備段階, 初期アイデアの入力段階, 図解作成段階, アイデアの追加段階, 関連情報生成段階, 図解の書き換え段階という 6 つのステップからなる. このうち, 実際にユーザが行うのは初期アイデア入力以降の段階である. はじめの前準備段階は管理者があらかじめ行っておく部分である. ユーザの作成する図解は図解マネージャー (図 3.9) により管理される. 図解マネージャはユーザの作成中の図解にユーザの作業のステップに応じて必要な操作を施す. ユーザは各段階ごとに作業が終了すると図解マネージャにマウスクリックにより知らせることにより次の作業に進むことができる. 以下では本システム利用における各ステップについて述べる.

1. **前準備段階**

   前準備段階として, キーワードファイルを作成し, 相関ルールの生成をおこないファイルに保存しておく. システムはニュースグループと minsup,minconf のパラメータ

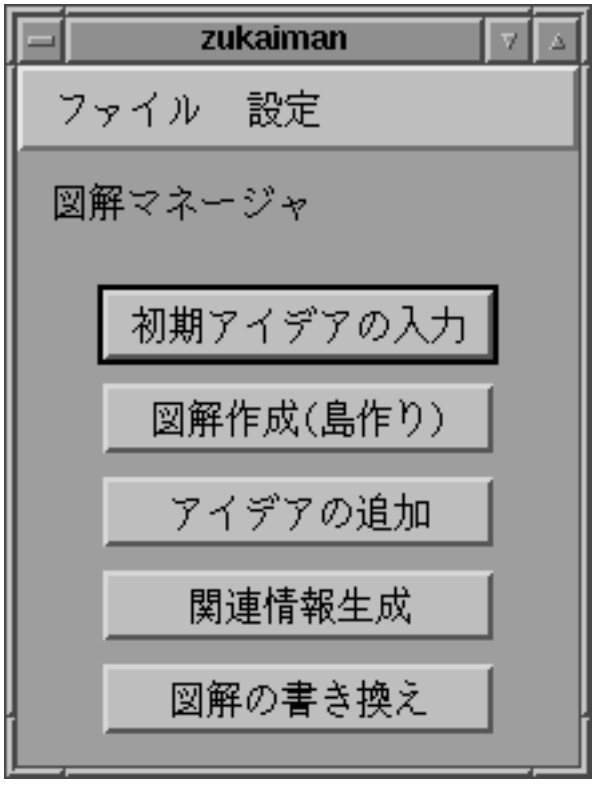
zukaiman
ファイル　設定
図解マネージャ
初期アイデアの入力
図解作成(島作り)
アイデアの追加
関連情報生成
図解の書き換え

図 3.9: 図解マネージャ

を指定すると自動的にキーワードファイルの構築を行ない, そこから指示度と確信度がそれぞれ minsup,minconf 以上の相関ルールを全て生成する.

既に述べた通りキーワードファイルはニュースグループごとに構築することにより複数の話題を反映させることができる. これによりユーザは必要に応じてキーワードファイルを使い分けることができる. (図 3.10)

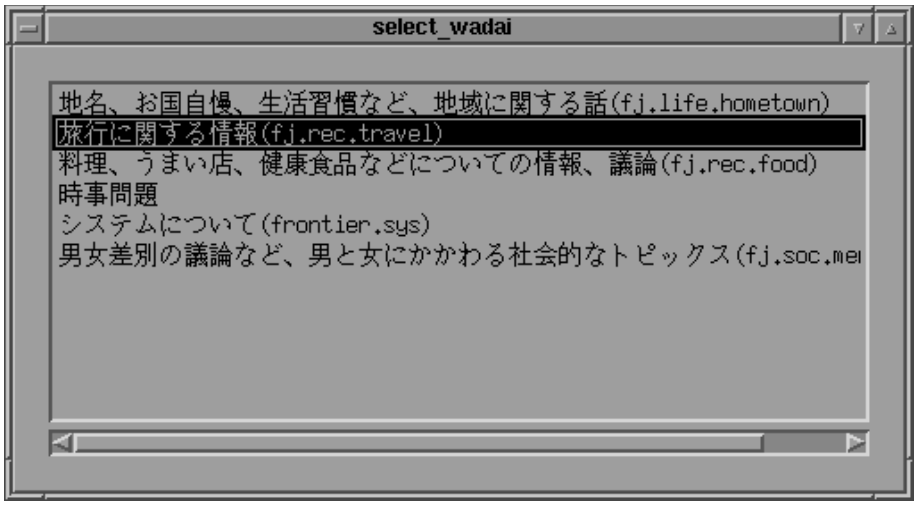

図 3.10: キーワードファイルの指定画面

2. **初期アイデアの入力段階**

   初期アイデアの入力段階は発散的思考に相当する. ここでは KJ 法のラベルとなるようなアイデアの断片知識を箇条書きに入力する.

3. **構造化段階**

   構造化段階では, ユーザが初期アイデアとして入力した断片知識をシステムがラベルとして D-ABDUCTOR に並べる. ユーザはこれをもとに図解を作成する.

4. **アイデア追加段階**

   図解化することにより新たに気づいたアイデアを追加する. この段階においてユーザはラベルの追加のみでなく, ラベルの削除や図解の構造の変更など自由に操作してよい.

図 3.11: アイデア追加画面

5. **関連情報の生成・閲覧段階**

   できあがった図解から関連情報を生成する. ユーザは生成された関連情報を前述のインターフェイス (図 3.4) を用いて閲覧することができる. そして, 関連情報のキーワードのうち興味深いものを複数個選択して, あらたななテーマとして関連情報を生成することができる. これによりユーザは直観に応じてキーワードをたどることが可能となる. また, 既に述べたように 3.3.2節に述べたリスト形式とグラフ形式で提示することができる.

   また, 関連情報の生成・閲覧段階では Association Rule マネージャーによる関連情報の提示と同時に,D-ABDUCTOR のユーザが作成した図解に関連情報が埋め込まれる. これによりユーザは, 生成された関連情報と自分が作成した図解にどのように関係があるのかを直観的にわかりやすくなる.

6. **図解変型段階**

   関連情報の生成・閲覧段階において一通りの関連情報を見てアイデアを広げた後に図形変形段階に入る. ここではシステムはユーザが作成した図解を変形することにより視点の転換をはかる.

   システムにより自動的に変形された図解はそのままでは意味のある図解とはいえない場合がある. そこでユーザはシステムによって変形された図解に手を加えることにより何らかの意味を持つ図解に書き換える.

   特にここで重要なのは, システムが自動的にグループ分けした島には名前が付けられていないためユーザがグループ分けされたラベルを見て名前付けを行うことである. この段階は新たな視点を発見することが目的であるため, この時点では必ずしも全体として意味の通った図解である必要はない.

このような一連の作業を行なった後に島を全て無くしてラベルだけ並べた状態から再度図解作りに戻る.

本システムではこの時点において新たに得られたアイデアや別の角度からの視点を反映した図解が作成されることを狙いとしている.

第 4 章

評価実験

これまでに発散的思考段階から収束的思考段階へ移るための支援機能の必要性と実現するための方法を提案し, 実際に開発した実験システムについて述べてきた. 本章ではこのシステムの有用性を調べることにより提案した方法が実際に機能するものであるかどうかを調べた評価実験について述べる.

一般に発想支援システムや思考支援システムの効果は明確な評価基準が無いために困難であるといわれている. 三末ら [18] は思考支援システムの評価のための以下の 3 つのアプローチを提案し,D-ABDUCTOR の評価を行なった.

1. **要素技術レベル：**システムを構成する各機能ごとにアルゴリズムの理論的な速さやプログラムの実行速度といった技術的な観点からの評価
2. **操作レベル：**思考作業 (例えば KJ 法) などに含まれるカードをさがす, 集める, 並べるといった「操作」をどう支援できるかの評価
3. **思考レベル：**思考作業に含まれる操作だけでなく「思考」をもどう支援できるかの評価

この中の操作レベルはシステム全体の操作性での評価であるが, 本システムを利用する際においてユーザが最も時間を費すのは D-ABDUCTOR を利用して KJ 法図解を作成する段階であり, 操作レベルの評価は D-ABDUCTOR の操作性に強く影響するため評価項目から外して要素技術レベルと思考レベルに関しての評価実験について述べる.

4.1　評価項目

本システムで行っている処理は大きく以下の 5 つに分けられる.

1. 電子ニュースからのキーワードファイルの自動構築
2. キーワードファイルから相関ルールの抽出
3. 冗長な相関ルールのフィルタリング
4. ユーザの図解の関連情報を抽出し, 提示
5. ユーザの図解の変形

このうち,1,2,3 はあらかじめ行っておく前処理で, 4,5 は実際にユーザの利用による処理である.

要素技術レベルの評価ではそれぞれの機能に関してプログラムの実行速度や生成されるデータ数などを比較する必要があるが, 現在の実験システムにおいてはユーザの待ち時間を少なくするためにリアルタイムに相関ルールを生成するのではなくあらかじめキーワードファイルから生成される全ての相関ルールを求めてファイルに保存しているため,4,5 に関しては要素技術の評価対象からはずしてこの部分の評価は思考レベルの評価に委ねることにした. 1 に関しては単なるキーワード抽出しか行っていないので, これも評価対象からはずした.

したがって, 要素技術レベルの評価は 2 のキーワードファイルからの相関ルールの抽出と 3 の冗長な相関ルールのフィルタリングを対象とした.

思考レベルの評価に関しては, 実際に被験者にシステムを利用してもらい, 被験者の作成した図解がシステム利用前と利用後とでどの程度変化しているかを調べることにより, 関連情報の抽出・提示や図解の変形機能が有効であるかを調べた.

4.2　関連情報生成機構の性能評価

ここではシステムの要素技術レベルの評価に関して述べる.

4.2.1 実験の目的

ここで評価するのはキーワードファイルからの相関ルールの生成と相関ルールのフィルタリングのそれぞれの処理を行うために, 提案した方式が効率良く機能しているかについてである.

キーワードファイルからの相関ルールの生成に関しては, キーワードファイルをそのまま用いると常に組となって出現するキーワードの組 (ヒュージアイテム集合) が組合せ爆発を起こして, 計算コストの増大と似たような意味のない膨大な相関ルールが生成されるという問題があった. そこで 2.2.2節においてキーワードファイル中のヒュージアイテム集合を他の一語に置き換えることによりこの問題を解決する方法を提案した. ここでは相関ルールの生成においてこの方法がどの程度機能しているかを調べた.

また, ルールのフィルタリング機構に関してはフィルタリング前後における相関ルールの量に関して調べた.

4.2.2 実験方法

実験 1 ヒュージアイテム集合の削除を行なった場合と行なわなかった場合について, 相関ルールの生成を行ない, 実行時間と生成されたルールの数を比較した.

実験にはニュースグループ fj.life.hometown から抽出したキーワードファイルを用いた.(記事数 1,757 件)

実験 2 フィルタリング機構の効果について調べるために複数のニュースグループのキーワードファイルから作成した相関ルールに対してフィルタリング処理を施し, 消去されたルール数を比べた. 冗長なルールのフィルタリング処理は統計的検定を用いた場合と用いない場合の両方について行なった.

実験 1 と実験 2 で評価する 2 つの処理はユーザがリアルタイムに行う処理ではないので瞬時に結果を得る必要は無い. そのためフィルタリング処理に関しては数秒から数分程度で処理が終わるため実行時間に関しては調査しなかった. しかし, 相関ルールの生成は条件によっては何時間たっても終了しなかったりメモリ不足などで結果が得られないことがあるので, 現実的な時間で答えが得られるかどうかも調査した.

4.2.3 実験結果

実験の結果は表 4.1の通りになった. ヒュージアイテム集合を除去したキーワードファイルを利用した場合とそのままで利用した場合とでは, 生成ルール数, 計算時間ともに明らかな差がある. ただし, ヒュージアイテム集合を除去したキーワードファイルから相関ルールを生成するためには相関ルールを除去するための処理も考慮する必要があるので, これも合わせた時間を括弧の中に示した.

キーワードファイルをそのまま利用した場合においては, 指示度を 0.6%にした時点で組合せ爆発が起こり, これ以上処理を続けることができなかった.

表 4.1: キーワードファイルから相関ルールの生成数と時間

指示度 (%)	生成ルール数 (個)		計算時間 (秒)	
	そのまま	除去	そのまま	除去 (Total)
1.0	277,298	134	254	46 (300)
0.9	367,029	134	324	52 (371)
0.8	368,234	165	370	65 (412)
0.7	967,612	320	978	101 (569)
0.6	生成不能	616	—	181 (895)

実験 2 の結果は表 4.2の通りになった. フィルタリングの方法は type1 と type2 の 2 通り用いた. type1 では統計的検定を用いずに, 完全に冗長であるルールのみの除去を行なった結果で type2 では統計的検定を用いて有意水準 2.5%で価値がないと判断したものを除去をした結果である.

4.2.4 評価・考察

実験 1 については, キーワードファイルからヒュージアイテム集合を除去しなかった場合には指示度を 0.6%に下げた時点で計算できない状態に陥ったため, 最終的に得られた相関ルールは実用レベルに達しなかったが, ヒュージアイテム集合を除去した場合に関してはこの後も計算を続けることができる. 実際に実験システムにおいてはこのニュースグループから指示度を 0.25%まで下げて得られた相関ルールを利用している.

表 4.2: フィルタリング前後の相関ルール数

ニュースグループ	フィルタリング前	フィルタリング後	
		type1	type2
fj.soc.copyright	227,453	111,474	53,689
fj.life.hometown	9,680	4,765	3,857
fj.soc.men-women	6,282	4,863	3,492

このように, キーワードファイルから相関ルールを求めようとすると, ヒュージアイテム集合の削除を行わなければ指示度を実用的なレベルまで下げることが出来ないことがわかった. また, ヒュージアイテム集合の除去は, 常に同時に出現するものを一つにまとめることから, 不要な組合せを生成しないのでフィルタリングの効果もある.

実験 2 については, ルールの数が多い場合と少ない場合とを比べてみると, フィルタリング前後において, 多少のばらつきはあるが半分程度まで冗長なルールが削除されることがわかった.

4.3 ユーザの思考の変化の評価

4.3.1 実験の目的

本システムは発散的思考と収束的思考を繰り返すことにより発想の質を高めるための支援を目指している. 本実験ではこの効果を調べるために一度ユーザが KJ 法の図解を作った後に本システムを利用し, 再度図解の作成をやり直してもらうことによりシステムの利用がユーザの思考にどのような変化を及ぼすかを調べた.

4.3.2 実験方法 (実験 3)

実験は被験者に課題を与えて本システムを利用した場合と他のツールを利用した場合とでユーザの発想の過程にどのような差があるかを調べるという形式で行なった. 実験に参加したのは國藤研究室の M1 の学生 4 人で, テーマとして「テーマ 1：本学の学食に関する問題について」と「テーマ 2：理想の結婚の条件」を設定した. 使用するキーワード

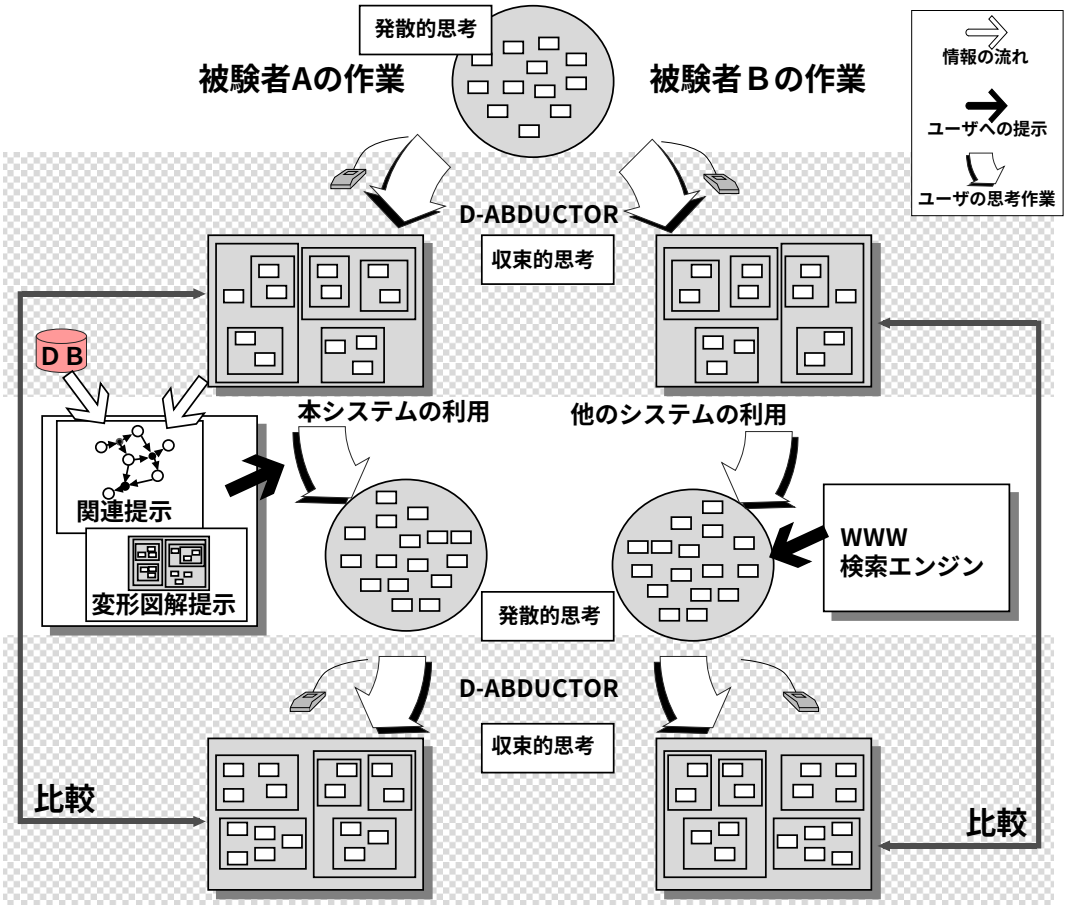

図 4.1: 実験における作業の流れ

ファイルは, それぞれ fj.rec.food, fj.soc.men-women を用いた.[1]

実験における作業の流れは図 4.1に示す通りである. 個人ごとで問題に関して観点が大きく異なると比較が困難になるので, 被験者を 2 人づつの 2 組のペア (図では被験者 A と被験者 B) に分け, それぞれのペアごとに与えられた問題に関して十分議論してもらい KJ 法のラベルとなるようなアイデアを生成してもらった. その後, それぞれ個人に分かれて, 相手と異なる意見に関しては自由に変更・削除を許して KJ 法の図解を作成してもらった.

被験者は全員 KJ 法の初心者であったので, 本来の KJ 法の作業より制約を緩め, 図解にまとめることにより新たに思いついたアイデアや不要な項目などがあった場合も変更, 削除, 追加を自由に行なえるようにした.

[1] ローカルな情報源を利用するのはシステムを利用しなかった場合と比較するのには適当でないと判断したため, 学内のニュースグループは使用しなかった.

被験者が満足のいく図解を作成し終えた後に各ペアのうち一人に本システムを利用してもらい, もう一人には関連情報を得るために WWW の検索エンジンを利用してもらった. (検索エンジンには普段使用している好きなものを利用することを許した. 被験者が実際に利用したものは mondou[11] と odin であった.)

本システムを利用した被験者には, システムによる関連情報提示機能や図形の変形提示機能などを自由に使ってもらった. ここで, 被験者にシステムが変形した図解を理解させるために変形によって新たに作られた島に名前付けを行わせた. この際, 変形された図形を自由に操作することを許した. また, 検索エンジンを利用した被験者には検索によって新たに得られたアイデアをメモしたり図解に追加するなど自由に作業をしてもらった. [2]

そして, 一通りシステムを利用した後に, はじめに作った図解から島を全てはずしてラベルをバラバラにならべた状態で再度アイデアの生成と図解を作り直してもらった.

以上の作業をテーマを変えて 2 セット行なった. 2 セット目に関しては 1 セット目で本システムを利用する人と検索エンジンを利用する人を入れ換えることにより,4 人の被験者全員が本システムと検索エンジンを 1 回づつ使ったことになり, 最終的にシステム使用前と使用後の計 16 枚の KJ 法図解が得られた.

4.3.3 システム使用前と使用後の図解の変化

システム使用前とシステム使用後とで被験者にどのような発想が生まれたかを調べる必要がある. ここではシステムの使用前と使用後で図解がどの程度変化しているかを測ることによりアイデアの断片だけでなく, ラベル間の関係に基づくアイデアについても評価した.

アイデアの断片の生成についてはシステム使用前後の図解におけるラベルの追加数をもとに評価した. ただし, 被験者は 2 回目の発散的思考においてラベルの追加だけでなくラベルの削除も多く行っているため, 単純な比較をすることはできない. また, 被験者のラベル間の関係に基づくアイデアがどのように生成されたかについては図解の構造がどの程度変化したかを調べた. しかし,2 つの KJ 法の図解がどの程度違うかを機械的に判断するのは困難である. そこで, 本実験では, 複数の人に図解を見せてアンケート形式で似てい

[2] 本システムを利用した場合と検索エンジンを利用した場合の両方の被験者に図解の自由な操作を許しているが, ここで作られた図解は被験者の作業過程を調べるために記録されたが, システム利用前と利用後との発想の変化を調べるための直接的な材料にはしない.

るかを答えてもらい, その平均をとることにより「類似度」を求めた. そして, 類似度の逆を「変形度」と定義して評価した. [3] ここで, アンケートの協力者には図解のラベルの追加や削除は考慮せずに図解全体の意味的な構造がどう変化しているかに着目するように求めた. アンケートは國藤研究室学生 10 人に参加してもらい, 各図解の組に関して以下の 5 段階の評価をしてもらった.

1. 全く似ていない (= 変形度 5)
2. あまり似ていない (= 変形度 4)
3. 少し似ている (= 変形度 3)
4. 良く似ている (= 変形度 2)
5. 全く同じ (=変形度 1)

被験者にははじめに作成した図解を参照することを許したが, その図解にはこだわらず現時点でもっとも良いと思う図解を描くように求めた. そのため図解の構造が変わったということは, これ以上アイデアが生まれない状態にまで思考が収束した時点からラベル間に存在する関係の新たな発見をもとに図解をより良く描き変えるための発想が生まれたことを意味する. 被験者がラベル間の関係の新たな発見をもとに図解の構造 (グループの分け方) を変えたことはラベルの見方に関して異なる観点を見出したと解釈することができる.

4.3.4 実験結果

実験結果を図 4.3に示す. 本システムを使用した場合と他のツールを使用した場合における図解の変形度の差はユーザごとにばらつきがあった. 被験者ごとにシステムを利用した時としなかった時のテーマが異なるため単純な比較はできないが, これをグラフにすると図 4.2の通りになる. またテーマごとの変形度の違いを示すと図 4.3の通りになる. いずれも本システムを利用した場合の方が変形度が上がっていることがわかる.

[3] 類似度 1 = 変形度 5, 類似度 2 = 変形度 4, ... , 類似度 5 = 変形度 1

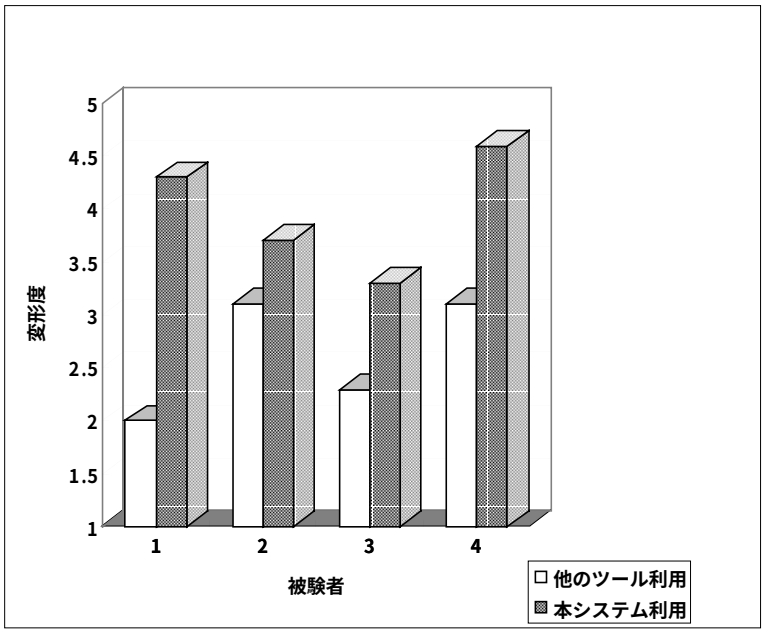

図 4.2: システム利用の効果 (ユーザごと)

1

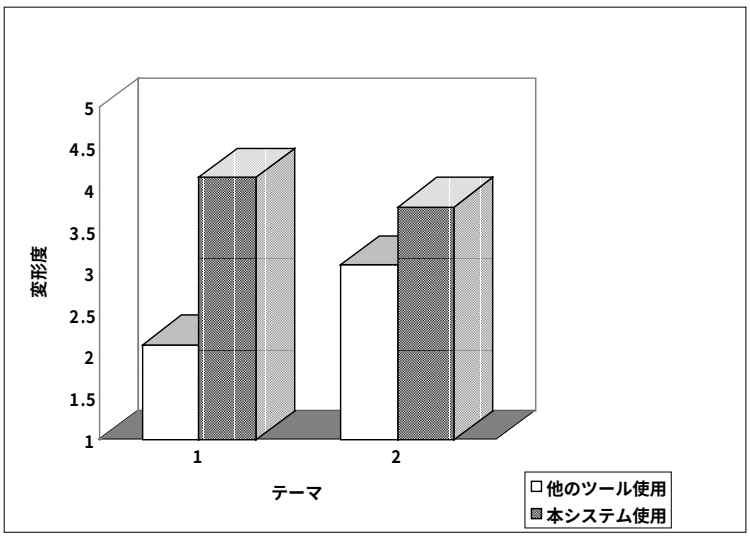

図 4.3: システム利用の効果 (テーマごと)

4.3.5 評価・考察

アイデア生成支援の効果

本実験では被験者に出来るだけ自然に発想させるため, 不要なラベルの除去や複数のラベルをひとつにまとめたり他の言葉に置き換えることを許したため, ラベル数の単純な比較によりアイデアの生成の効果を計ることはできないが, 本システムを利用した場合とWWW の検索エンジンを利用して関連情報を入手した場合とを比べると, 新たに生成されたアイデアの数はあまり差がなかった. この結果から, 本システムと検索エンジンがユーザに与える情報を比べると検索エンジンの方が圧倒的に大きな情報量を持っているはずであるが, それにもかかわらず同程度の効果があるということができる. これは, 被験者がはじめの発散的思考の段階でほとんどアイデアを出し切ってしまったことと, 一度収束的思考を行ったことによりその時点の図解にそのまま追加できるアイデアしか生成しなかったことが影響したと考えられる.

ラベル間の関係発見の効果

図 4.3の実験結果から, ユーザがはじめに作成した図解から最終的に完成した図解への変形度を本システムを利用した場合と利用しなかった場合で比べると本システムを利用した方があきらかに高くなっていることがわかる. これは, 本システムを利用したことによりユーザがラベル間に存在する関係を新たに発見し, それを反映して図解を書き直したことを意味している.

このことを詳しく見るために, 被験者の実験結果を例に挙げて説明する. 図 4.4は被験者 2 がテーマ 1 に関してはじめに作成した図解であるが, システムはこの図解を図 4.5のように自動的に変形した. 変形された図解を見た被験者は, この図に関して意味の無いと思われる島の削除やシステムにより作られた島への名前付けを行い, 図 4.6を作成した. 後のインタビューでこの被験者はこの図解に関してはじめの図解で気付かなかった点が反映され内容的にもこちらの図解の方が満足しているという意見を述べていた. その後, はじめの図解の島を全て無くしてバラバラにした状態で 2 度目のアイデア生成と図解作りをやり直してもらったところ, 前とは異なる図解 (図 4.7) が完成した. このように, 被験者 2 はシステムが作成した情報からラベル間の関係を発見し, それによって得たこれまで気付かなかった新たな問題の捉え方 (新たな観点) を考慮しながら再度図解作りを行った. 単に図解の構造が変わっただけでは必ずしも発想の質が高められたとは言えないが, この図解は

被験者自体が前より良いと思って作成した図解であるので本システムの利用により被験者の思考が一度懲り固まった状態から脱出し発想の質が高められたと考えることができる.

しかしシステムが作り出した図解に全ての被験者が満足したわけではない. 被験者 2 とは対象的に, 被験者 1 は実験後のインタビューにおいてシステムの作り出した図解が全く理解できなかったという感想を述べている. 被験者 1 が作成した図解は図 4.8で, システムはこの図解を図 4.9のように変形した. システムが作成した図解が被験者にとって満足いかなかった理由は, ユーザの作成した図解の各ラベルに書かれた文字数が全体的に少ないうえに辞書に存在しない単語が多く使われておりデータベースにマッチした単語が少なかったため, システムは意味の無い単語をもとに図解の変形を行なわなければならなかったのが原因であると考えられる.

それでも被験者 1 が最終的に作成した図解ははじめの図解と比べて大幅に変わった構造を持つ図解であった. これは, 実験において, システムが生成した図解に強制的に名前づけを行なわせたことにより, 何らかのひらめきがあったことと思われる.

被験者の作成した図解の構造が変化したことはシステムが収束的思考を支援したと解釈することもできるが, 被験者ごとの履歴を調べてみるとシステムが自動的に行ったラベルのグルーピングをそのまま利用しているわけではなく, むしろシステムが作成した図解を操作して島の名前付けなどを行わせたことにより今まで気付かなかったラベル間の関係を発見したということから発散的思考を支援したと考えることができる. しかし, 今回の実験においてはラベル数の大幅な増加には到らずに最終的な図解の構造にのみ反映された.

図解を作り直すことの効果

システムを利用しなかった場合 (検索エンジンを利用) の変形度は被験者によってばらつきがあり, 本システムを利用しなくても図解の構造を変えた被験者が見られた. 被験者は検索エンジンを利用する際に, 自分の作成した図解を自由に操作することが許されていたため, ユーザの図解作成の履歴を調べたところ, 検索エンジン利用時に図解の構造を大きく変えた被験者は一人もいなかった. 各被験者には図解を操作しながら検索エンジンの利用を十分に行ってもらい, 「これ以上変えるところは無い」と言うまで, アイデアが完全に収束した状態になった. それにもかかわらず, 完成した図解の島を全てはずして, ラベルをばらばらに並べた状態から新たにアイデア生成と図解の作り直しを行ったところ, 本システムを利用した場合程では無かったが, 図解の構造を大きく変えた被験者がいた. こ

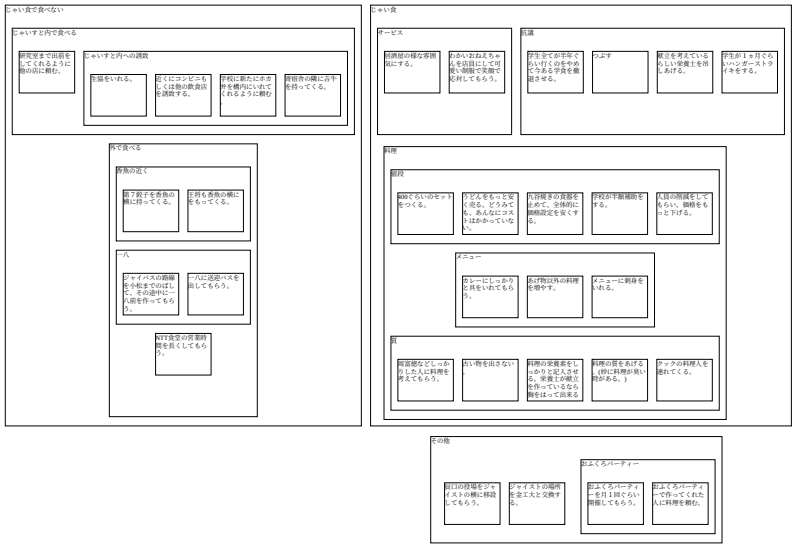

図 4.4: 被験者 2 がはじめに作成した図解

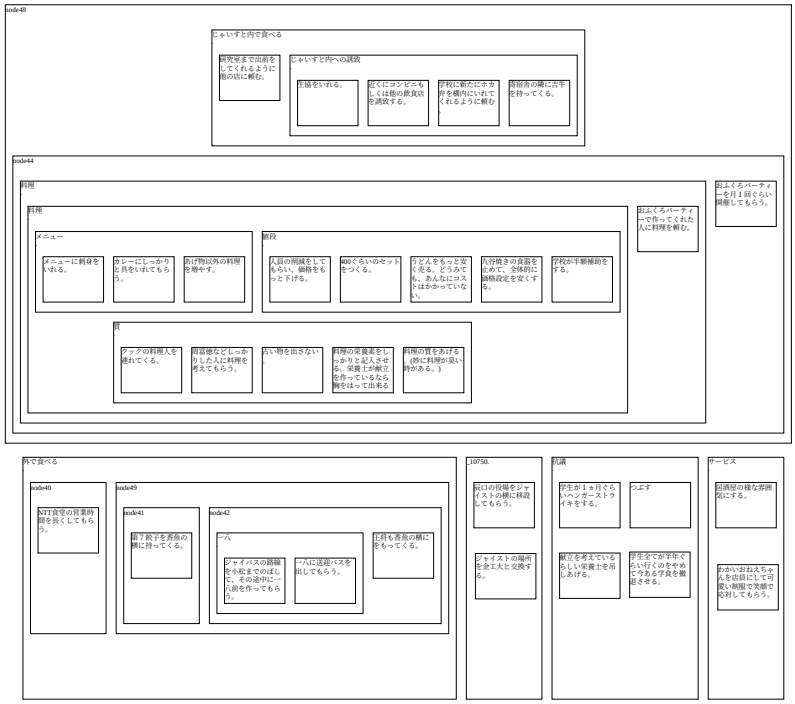

図 4.5: システムが被験者 2 の図解を変形した図解

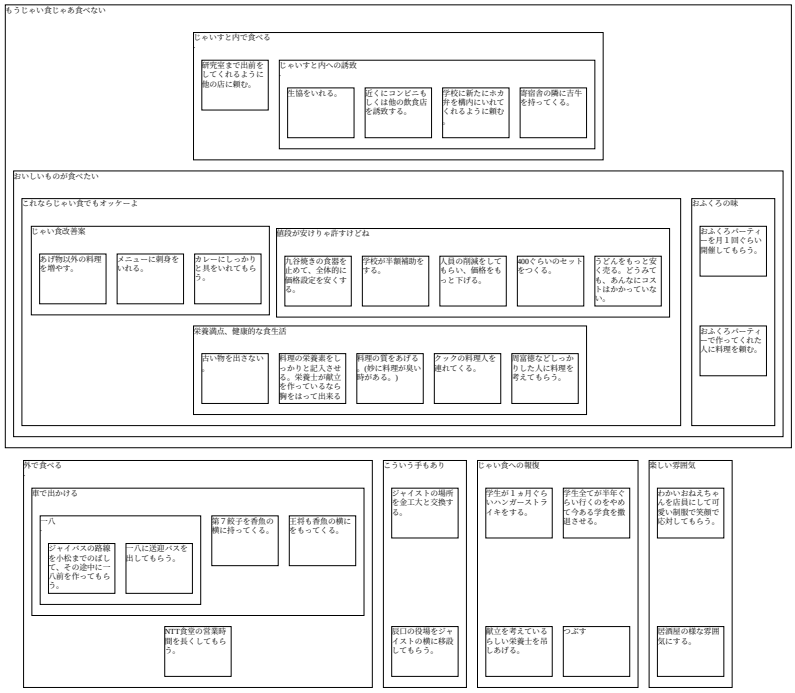

図 4.6: 被験者 2 がシステムに変形された図解に操作を加えた図解

のような結果を得たのは検索エンジンがラベル間の関係の発見を支援したのではなく, 一度ラベルをバラバラにしてまた並べ直す段階において新たな発想が生まれたものであると考えられる. これは, 被験者が D-ABDUCTOR と KJ 法の初心者であったために思考がふらついていたことが原因の一つに考えられるが, KJ 法を繰り返すことにより発想の質が高められることも作用したと考えられる.

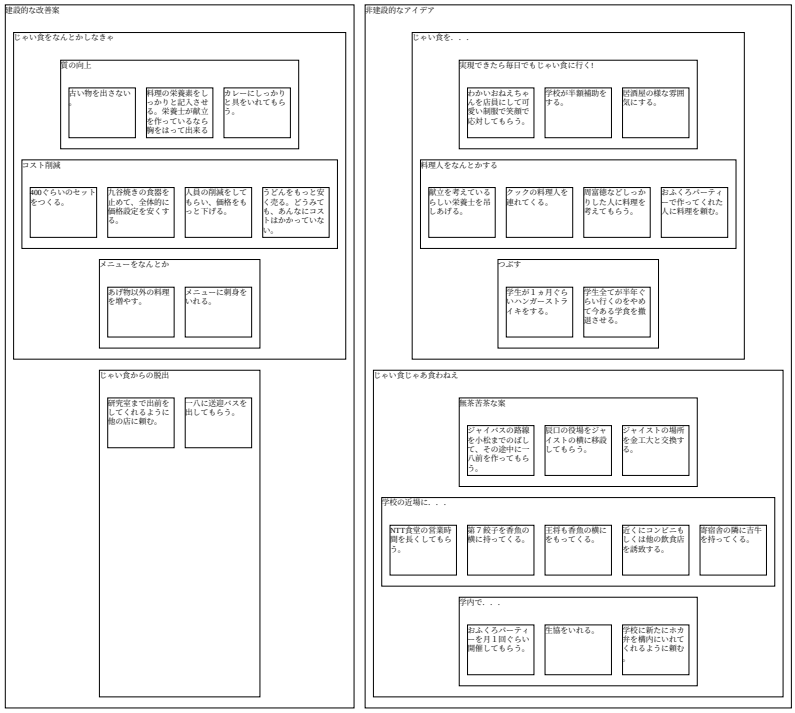

図 4.7: 被験者 2 が最終的に完成させた図解

図 4.8: 被験者 1 がはじめに作成した図解

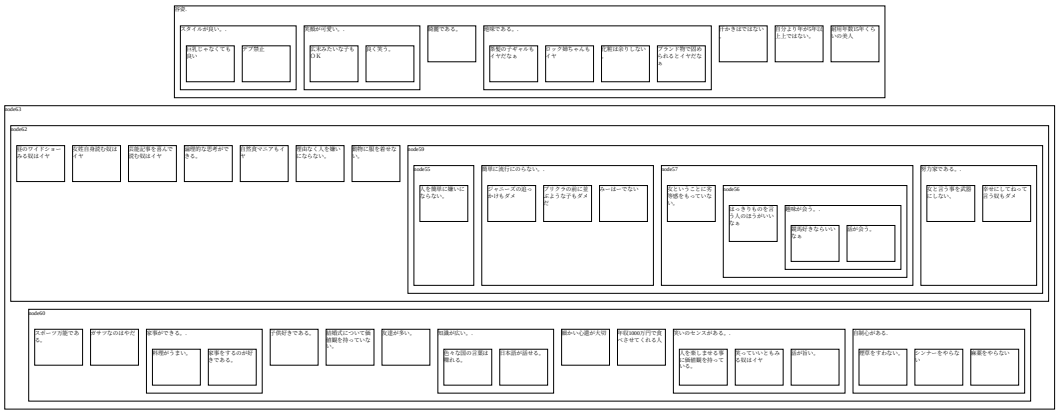

図 4.9: システムが被験者 1 の図解を変形した図解

表 4.3: 1 回目と 2 回目の図解の違い

被験者	テーマ	本システム	ラベル数の変化	図解の変形度平均 (分散)
被験者 1	1	使用しない	30 → 32	2.0 (0.20)
被験者 2	2	使用しない	37 → 38	3.1 (0.49)
被験者 3	1	使用しない	40 → 35	2.3 (0.21)
被験者 4	2	使用しない	32 → 20	3.1 (0.29)
被験者 1	2	使用した	46 → 44	4.3 (0.41)
被験者 2	1	使用した	33 → 31	3.7 (1.21)
被験者 3	2	使用した	19 → 21	3.3 (0.41)
被験者 4	1	使用した	53 → 31	4.6 (0.24)

第 5 章

関連研究との比較

本章では関連研究との比較を通じて本研究の特色をあきらかにする.

5.1 生成される関連情報に関する比較

本研究ではテキスト中のキーワードに関する相関ルールをもとに関連情報を生成する事により発散的思考支援を目指しているが, あるキーワードからそれに関連する情報を得ることを考えると WWW の検索エンジンなどの既存の情報ツールを利用することが考えられる. キーワードに関する相関ルールと検索エンジンのようなテキスト検索によって得られる情報は基本的に異なる性質を持っている. 検索エンジンではキーワードに関連するホームページが得られるため具体的な情報を得る事ができるが,「違法駐車を減らすためにはどうしたらよいか?」といったような議論に役立つ情報を得るのには適していない. これに対して, テキスト中のキーワードに関する相関ルールを利用すると, 電子ニュース中の個々の記事を得るのではなく, 記事全体を通して議論のながれや話題の傾向をもとにした全体的な情報を得ることができる.

議論の流れや話題の傾向を得るのか, 個々の情報を得るのかについては必要に応じて使い分けたり組み合わせたりすることが望ましい. 本システムのような枠組を用いたシステムと検索エンジンなどの検索ツールを組み合わせることにより関連するホームページの情報を KJ 法の図解に埋め込むといった拡張も考えることができるかと思われる.

5.2 関連情報の可視化に関する比較

本研究では生成された関連情報をユーザの図解に埋め込むことにより関連情報を提示する方法と関連情報自体を書き換える方法を用いて可視化を行った.

従来の発散的思考支援システムは, 似ていたり関連の強い情報が近くに来るように関連情報を 2 次元空間上に配置するという方法をとるものが多い. KJ 法も関連情報を空間上に情報を配置するという意味で, これらの方法との関係が深い.

発想支援システム HIPS[21] において D-ABDUCTOR と Keyword Associator を組み合わせることにより KJ 法の図解自動的を作成する試みがされている. このシステムではキーワード情報をもとに, ユーザの作成した各ラベルの関連度を求め, 関連の強いカードを自動的にグルーピングする. しかし,KJ 法の図解はラベル間の意味的な関係に深く関わっているため, キーワード情報のみを用いて意味のある図解を得るには限界がある. そのため, この方法は図解を作成するための初期配置を提供するに過ぎない. これは, ユーザが与えたテーマをシステムが持つ情報のみを利用して配置するという意味では, 包含グラフとして表現されるか, 2 次元空間上に配置するかの違いはあるが, 基本的には従来の方法と変わらない.

これに対して, 本研究で用いた方式は, 図 5.1に示すようにユーザが作成した図解からテーマとなるキーワードを抽出し, このテーマからシステムが情報を生成するところまでは従来の方法と同じであるが, ここで作られた情報をもとの図解に戻すことにより, ユーザの思考空間を反映させた情報を提供することを実現している. 本研究で提案したユーザの図解を変形させる方法は, システムが生成した島のみを用いて図解を作成すると単なるキーワードマッチングによるグルーピングになってしまうが, ユーザの作成した図解とを重ね合わせることが非常に有効なものであった.

5.3 KJ 法以外の思考技法を用いたシステムとの比較

本研究では KJ 法をユーザの思考の基板の技法として用いたが, KJ 法の発散的思考から収束的思考への処理は非常に協力であるが発散的思考と収束的思考を繰り返す処理は得意であるとはいえない. これに対して, この両者の繰り返しの処理に優れた発想法としては, マインドマッピング [3] がある. マインドマッピングは以下のように KJ 法と比べてややトップダウン的に思考を行う方法である.

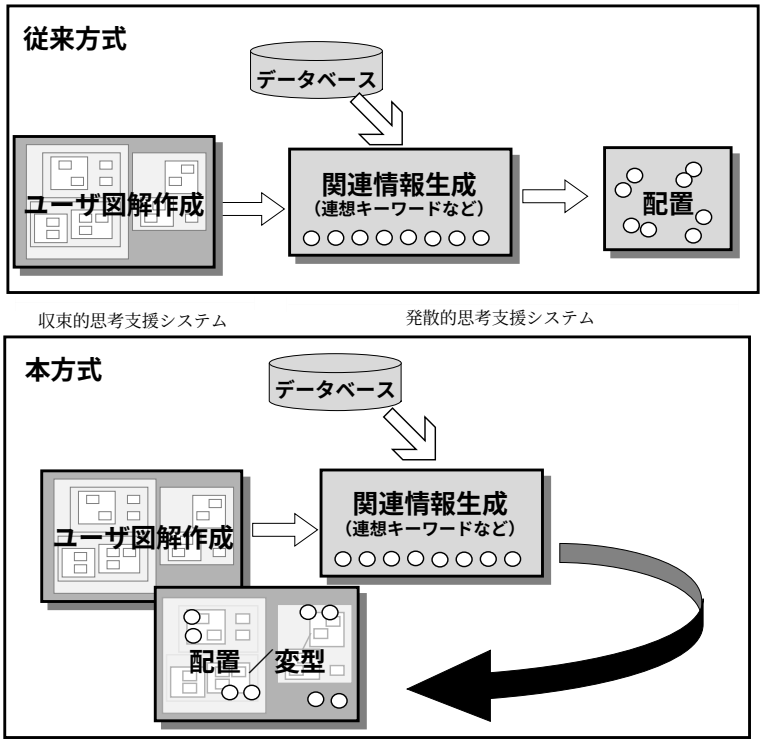

図 5.1: 本研究で用いた方式と従来の方式との比較

1. まず中央に話題のテーマとなるキーワード（メインテーマ）を書く.
2. そこからサブテーマとなるキーワードをメインテーマのまわりに書き, メインテーマから線を引いておく（線は関連をあらわす）
3. サブテーマのまわりに関連するキーワードを書いていき, 同様に線で結んでいく.
4. （必要ならば）キーワードを線で囲んでグループ分けする.

アイデアを生成した順に図解を書いていくため, 思考の過程をそのまま図解に書くことができるうえに,KJ 法と違って一枚の紙だけで作業ができるため手軽にできる点があげられる. これに対して,KJ 法は, 関連するアイデアを全て洗い出した後, 問題の分析や思考の整理という観点でみると生成されたアイデア全体のバランスを見ながら図解化を行うので, アイデアの生成順に依存する方法と比べて最終的に得られる図解の完成度が高い. これは, マインドマッピングの図解の作成は連想を用いて発散的思考と収束的思考の両方を

効率的に行っているのに対し, KJ 法の図解の作成は収束的思考に重点を置いているからであるといえる.

本研究で KJ 法を用いたのは, できるだけユーザが十分に問題を分析して得た問題認識の構造を反映した図解から情報を抽出するというアプローチをとったからであるが, 逆にマインドマッピングのような思考過程を反映した図解から情報を抽出するアプローチをとることも興味深いと考えられる.

コンピュータを利用してマインドマッピングの図解を作成するツールとしてはインスピレーション (図 5.2) などがあるが, 図解作成と関連情報の自動生成と組み合わせる研究はあまりされていない.

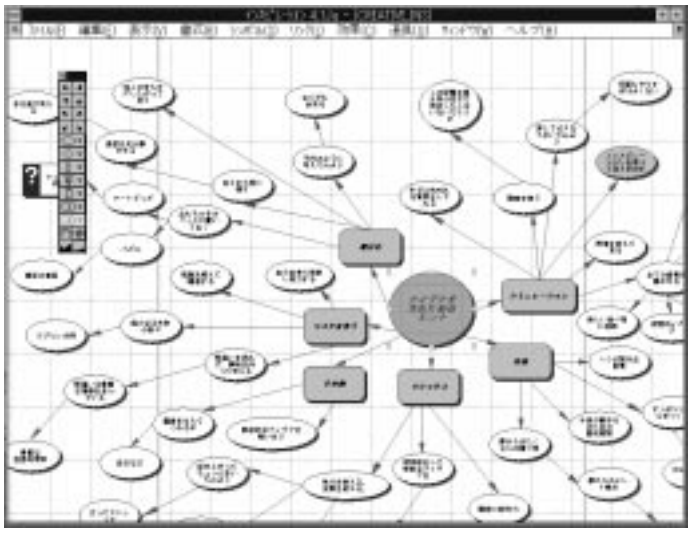

図 5.2: マインドマッピングの図解を作成するツール

本研究で行ったようなユーザの作成した図解とシステムの生成した情報を組み合わせることを考える。本研究においてユーザが図解を作成し, それに関連する情報をユーザの図解に反映して可視化するという枠組はマインドマッピングなどの KJ 法以外の発想技法に適用することも可能である. マインドマッピングは関係のあるキーワード同士を線で結ぶので関連のあるデータをデータベースから検索する際にその情報を利用することが可能である. ユーザの作成したマインドマッピングの図解と従来の研究におけるテーマの関連キーワードを 2 次元平面上に配置する方法を組み合わせることは興味深いと思われる.

第 6 章

結論

本研究では発散的思考と収束的思考を繰り返すことにより発想を高めていくというプロセスをコンピュータにより支援するシステム構築の目指し,KJ 法を基盤とした発想支援ツールの提案, 開発, 評価を行なった.

システムを構築において, 目的としたのは以下の点である.

- ユーザの作成した図解に対して, ただ単に関連のある情報を生成するのではなく, ユーザの気づいていない情報を積極的に発見する.
- 可視化において, ユーザの収束的思考を反映させる.

本章では, 本研究において達成したことがらをまとめ, さらには今後の研究の課題について述べる.

6.1 本研究においてあきらかになったこと

本研究ではユーザの作成した図解から関連情報を生成するために電子ニュースの記事に含まれるキーワードに関して相関ルールの導出アルゴリズムを適用した. そして, ユーザの収束的思考を反映して発散的思考を支援するために, ユーザの作成した KJ 法の図解への関連情報の埋め込みをおこない, さらにユーザの作成した図解と関連情報から作った島を重ね合わせることにより図解の構造を変形することにより新たな観点を発見するための支援機構を提案, 開発し, 開発したシステムの評価実験をおこなうことにより方式の有効性を確かめた.

評価実験は, 技術レベル思考レベルについての評価を行なった. 技術レベルの評価として電子ニュースの情報から関連情報が生成される過程を調べることにより, 提案した方法が実際に機能することをあきらかにした. また, 思考レベルとしては, 本システムの利用のユーザにより作成される図解による変化を調べることにより, 本システムが目的としているように本システムを利用する事により実際にユーザがこれまで気付かなかったラベル間の関係の発見をして新たな観点に着目することができることをあきらかにした.

6.2 展望

ここでは評価実験を通してあきらかになったシステムの問題点とともに, 本論文で提案したユーザが収束的思考段階で作成した図解の構造を関連情報を反映して発散的思考を支援するという新しい枠組の研究についての展望について述べる.

評価実験の結果システムによるユーザの図解変型が必ずしもユーザの満足する結果が得られない場合があることがわかった. これはユーザの作成した図解がシステムの用意した情報にマッチするものが少ないことが原因と考えられる. この問題に対処するためにキーワードファイルの構築や関連情報を生成する機構を改良する必要があると思われる.

また実験において被験者がもっとも苦労したのは KJ 法の図解を作成するのに手間がかかることであった. 一つの解決法としては本研究で扱おうとしている問題に特化したインターフェイスを持つ KJ 法図解作成ツールを作るという方法があげられる. しかしこの問題の本質はより深いところにあると思われる. それはこの種のシステムの最大の問題点はちょっとした発想をしようとしている人間がコンピュータに向かって KJ 法の図解を書かなくてはならないため, 気軽にシステムを利用しようという気にならないことである.

これは本システムを含むこれまで研究されてきた発想支援システムが普段コンピュータを使って行ういる業務の環境と独立していること自体に問題がある. これに対して Sugiyama ら [9] はザウルスのような携帯端末に入力されたメモや電子メールや WWW などからの日常の情報収集活動から得た情報を手軽に KJ 法のラベルに取り込むことができるシステムの開発を試みた. このように発想支援研究はこれまでに開発されてきた要素技術をいかにして従来のコンピュータ環境に融けこませるかを考える段階になってきたといえる. そこで, 今後, 本研究で提案している方法にこのような機能を持たせることが期待される.

謝辞

本研究を進めるにあたって, 多くの方にご支援をいただきました. この場を借りて感謝の気持ちを表したいと思います.

指導教官の國藤進教授には研究に関してさまざまな助言をしていただき, 自由な研究環境をはじめとする日頃の研究生活全般においても常に配慮していただき, 大変感謝しています.

実験システムの開発においては, (株) 富士通研究所の三末和男さんの開発された D-ABDUCTOR を使用させて頂きました. また, 図解の重ね合わせるという手法に関しては, 日本総合研究所の女部田武史さんに助言を頂き, 実験システムの初期バージョンにおいては D-MERGIN を利用させていただきました. 東京大学医科学研究所ヒトゲノム解析センタの佐藤賢二博士にはデータマイニングに関する助言をいただき, 本研究で用いた相関ルールのフィルタリング手法を紹介していただきました.

評価実験に参加して下さった小川剛志さん, 寺田賢二さん, 堀井宏祐さん, 森沢浩造さんには何時間にも渡る図解作成の実験に精力的に協力して頂きました. 知識工学講座の Thanaruk Theeramunkong 助手には論文執筆の際にご指導いただきました.

また, 國藤研究室の方々には日頃から研究に関して議論を重ねていただき, 研究活動以外にも私生活の面でも非常にお世話になりました. 心から感謝しております.

最後に私ごとで恐縮ですが, これまでの学生生活を金銭的・精神的に支えてくれた家族にも感謝の意を表させていただきます.

1997 年 2 月 14 日

野口 裕史

参考文献

[1] 伊藤 哲朗: 情報検索, 昭晃堂,(1986).

[2] 女部田武史: 複数の KJ 法図解の差異や共通部を可視化するグループ思考支援システムの研究, 北陸先端科学技術大学院大学 修士論文 (1996)

[3] 河合正義: 一枚の紙から発想する技法, 日本実業出版社, 1994

[4] 川喜田二郎: KJ 法, 中央公論社 (1986)

[5] 國藤 進: 発想支援システムの研究開発動向とその課題, 人工知能学会誌,Vol.8, No.5, pp.552-559 (1993).

[6] Rakesh Agrawal, Ramakrishnan Srikant: Fast Algorithms for Mining Association Rules, *Proc. of the 20th Int'l Conference on Very Large Databases*,pp.487-499 (1994).

[7] Sugimoto M., Hori K, Ohsuga S: A Document Rearieval System for Assisting Creative Research, ICDAR'95, Montreal,Canada,pp.167-170 (1995)

[8] 杉山公造: グラフ自動描画法とその応用- ビジュアルヒューマンインターフェース - , 計測自動制御学会 (1993).

[9] Kozo Sugiyama, Kazuo Misue, Isamu Watanabe, Kiyoshi Nitta, Yuji Takada: Emergent Media Environment, 人工知能学会研究会資料 SIG-J-9602-4, pp.19-24 (1996)

[10] 高杉耕一, 國藤進: ばねモデルを用いたアイデア触発システムの構築について, 人工知能学会研究会資料 SIG-J-9602-7, pp.34-39 (1996)

参考文献

[11] 西村英樹, 伊藤耕一郎, 河野浩之, 長谷川利治: 重み付き相関ルール導出アルゴリズムによる WWW データ資源の発見, 電子情報通信学会 第 7 回データ工学ワークショップ論文集 pp.79-84 (1996).

[12] 日本電子化辞書研究所: EDR 電子化辞書仕様説明書 (1995)
http://www.iijnet.or.jp/edr/TG.html

[13] 野口裕史, 國藤進: データベースからの知識発見法を用いた発想支援システムの提案, 平成 8 年度電気関係学会北陸支部連合大会予稿集 p.293 (1996)

[14] 野口裕史, 國藤進: データベースからの知識発見法を用いた発想支援システムの研究, 人工知能学会研究会資料 SIG-J-9602-14, pp.76-81 (1996)

[15] 福田剛志, 森下真一: 相関ルールの可視化について, 信学技報 DE95-6 pp.41-48(1995)

[16] 松本裕治, 今一修, 山下達雄, 北内啓, 今村友明: 日本語解析システム『茶筌』version 1.0b5 使用説明書 (1996)

[17] 三末和男, 杉山公造: 図形発想支援システム D-ABDUCTOR の開発について, 情報処理学会誌,Vol.35,No.9, pp.1739-1749 (1994).

[18] 三末和男, 杉山公造: 思考支援システムの評価法および D-ABDUCTOR の評価実験について, 計測自動制御学会 第 17 回システム工学部会研究会資料,pp.61-68 (1995)

[19] 由井薗隆也, 宗森純, 長澤庸二: 発想一貫支援グループウエア郡元, 計測自動制御学会第 17 回システム工学部会研究会資料, pp.101-108 (1995)

[20] 渡部勇: 発散的思考支援システム:Keyword Associator, 計測自動制御学会合同シンポジウム論文集,pp.411-418 (1991).

[21] 渡部勇 三末和男 新田清 杉山公造: ハイブリッド発想支援システム「HIPS」, 計測自動制御学会 第 17 回システム工学部会研究会資料, pp.101-108 (1995) pp.77-84 (1995)

付録 A

Apriori アルゴリズム

Apriori アルゴリズムは図 A.1のような手続きである。

1) $L_1 = \{$large 1-itemsets$\}$;
2) **for** (k = 2; $L_{k-1} \neq \emptyset$; k++) **do begin**
3) C_k = apriori-gen($L_k - 1$); // 新しい候補の生成
4) **forall** transactions $t \in \mathcal{D}$ **do begin**
5) C_t = subset(C_k,t) // t に含まれる候補
6) **forall** candidates $c \in C_t$ **do**
7) c.count++;
8) **end**
9) $L_k = \{\ c \in C_k | c.count \leq minsup\ \}$
10) **end**
11) Answer = $\bigcup_k L_k$;

図 A.1: Apriori アルゴリズム

候補の生成

apriori-gen 関数は引数として全ての large(k-1)-itemset の集合 L_{k-1} とる。これは、全ての k-itemset の集合の superset を返す。

始めに、join ステップで L_{k-1} の 2 つの要素を結合させる。

insert into C_k
select $p.item_1, ..., p.item_{k-1}, q.item_{k-1}$
from $L_{k-1}p, L_{k-1}q$
where $p.item_1 = q.item_1, ..., p.item_{k-2} = q.item_{k-2}, p.item_{k-1} < q.item_{k-1}$;

次に、削除ステップで c の (k-1)-subset のいくつかが L_{k-1} でないような全ての itemset $c \in C_k$ を削除する。

forall itemsets $c \in C_k$ **do**
　forall (k-1)-subsets s of c **do**
　　if $(s \notin L_{k-1})$ **then**
　　delete c from C_k

付録 B

相関ルールのフィルタリング

ここでは本研究で用いた相関ルールのフィルタリングの手法の詳細について述べる。相関ルールのフィルタリングのための方法は [15] において提案されている。本研究では基本的にこの方法を用いている。

以後の説明では x の指示度を s(x),x の確信度を c(x) とする。

B.1 判断基準 1

c(卵 $\Rightarrow$ ミネラルウォータ) $= 20\%$, s(ミネラルウォータ) $= 25\%$ とすると、卵を買う人は他の人よりミネラルウォータを買う割合が低いことになるため、卵 $\Rightarrow$ ミネラルウォータ というルールは価値が無い。

$$c(B \Rightarrow H) < s(H)$$

であれば、ルール $B \Rightarrow H$ は価値が無い。(**判断基準** 1)

B.2 判断基準 2

牛乳 $\Rightarrow$ ミネラルウォータ というルールがあった場合、s(牛乳) $= 16\%$, s(ミネラルウォータ) $= 25\%$, s({ 牛乳, ミネラルウォータ }) $= 4\%$ ならば、牛乳とミネラルウォータのそれぞれの支持度から { 牛乳, ミネラルウォータ } の支持度が予測できるためこのルールは価値が無い。

表 B.1: 分割表

条件	B	¬ B	計
H	$Ns(H,B)$	$N(s(H)-s(\{H,B\}))$	$Ns(H)$
¬ H	$N(s(B)-s(\{H,B\}))$	$N(1-s(B)-s(H)+s(\{H,B\}))$	$N(1-s(H))$
計	$Ns(B)$	$N(1-s(B))$	N

$B \Rightarrow H$ に対して、独立性の検定により、「B と H は関係が無い (独立である)」という仮説が棄却されればこのルールは価値があるとする。そうでなければ価値が無い。

この分割表から作られる

$$\begin{aligned} X &= \frac{(Ns(\{H,B\})-Ns(H)s(B))^2}{Ns(H)s(B)} \\ &+ \frac{(N(s(B)-s(\{H,B\}))-Ns(B)(1-s(H)))^2}{s(B)(1-s(H))/N} \\ &+ \frac{(N(s(H)-s(\{H,B\}))-Ns(H)(1-s(B)))^2}{s(H)(1-s(B))/N} \\ &+ \frac{(N(1-s(B)-s(H)+s(\{H,B\}))-N(1-s(B))(1-s(H)))^2}{N(1-s(B))(1-s(H))} \\ &= N\frac{(s(\{H,B\})-s(H)s(B))^2}{s(H)s(B)(1-s(H))(1-(B))} \end{aligned}$$

は B と H が独立ならば自由度 1 の χ^2 分布に従う。

ここで、有意水準 1%のとき $X > \chi_1^2(0.01) = 6.63$ であれば仮説は棄却される (このルールは価値があると言える)。(**判断基準** 2)

B.3 確信度の上昇

c(スポーツウェア $\Rightarrow$ スポーツシューズ) $= 30\%$, c(スポーツウェア卵 $\Rightarrow$ スポーツシューズ) $= 31\%$のとき、2 つめのルールの確信度は最初のルールとほとんど等しいので、2 つめのルールは価値が無い。

一般に、次のような 2 つのルールがあった時、

$$B \Rightarrow H \tag{B.1}$$

表 B.2: 観測度数と期待度数

	観測度数	期待値
HBC	$Ns(\{HBC\})$	$Ns(\{BC\})s(\{HB\})/s(B)$
$\neg HBC$	$N(s(\{BC\}) - s(\{HBC\}))$	$Ns(\{BC\})(1 - s(\{HB\})/s(B))$

$$BC \quad \Rightarrow \quad H \tag{B.2}$$

ルール (B.2) の確信度がルール (B.1) に比べて小さければアイテム C が確信度の上昇に全く貢献していない。従って、ルール (B.2) はルール (B.1) に比べて複雑になっただけで「価値が無い」と言える。(**判断基準** 3)

判断基準 3 において (B.2) の確信度が (B.1) より小さい場合が除去されるので、以下では (B.2) の確信度が (B.1) より大きい場合のみを考える。

(B.1) と (B.2) の確信度の違いが小さければルール (B.2) はルール (B.1) から容易に予想できる。

適合度の検定により、「ルール (B.1)、(B.2) の確信度に違いが無い」言い替えると「{BC} の上にもルール (B.1) が成立する」という仮説が棄却されればこのルールは価値があるとする。そうでなければ価値が無い。

表 B.2の期待度数と観測度数から作られる検定量

$$X = \Sigma \frac{(\text{観測度数} - \text{期待度数})^2}{\text{期待度数}}$$

は自由度 1 の χ^2 分布に従う。

ここで、有意水準 1%のとき $X > \chi_1^2(0.01) = 6.63$ であれば仮説は棄却される (このルールは価値があると言える)。そうでないものを価値が無いとする。(**判断基準** 4)

B.4 確信度

[15] では後件部が 1 つのアイテムのみのルールしか扱っていないため、判断基準 1〜4 までしか扱っていないが本研究では前件部も後件部も複数のアイテムからなるルールを扱うため判断基準 3,4 に類似した 2 つの判断基準 5,6 を追加した。

一般に、次のような 2 つのルールがあった時、

$$B \Rightarrow H \tag{B.3}$$

$$B \Rightarrow HC \tag{B.4}$$

ルール (B.1) の確信度がルール (B.2) に比べて違いがなければ後件部の C が無くなった分ルールが緩くなったにも関わらず確信度が全く上昇していないためルール (B.1) 「価値が無い」と言える。(**判断基準** 5)

ルール B.3の確信度がルール B.4の確信度より大きい場合は

$$B \Rightarrow C \tag{B.5}$$

についても考えなければならない。もし *ルール* B.3の確信度 × *ルール* B.5の確信度がルール B.4の確信度より小さければルール B.4はルール B.3とルール B.5から容易に予想できるためルール B.4は冗長であるとする。(**判断基準** 6)

付録 C

KJ 法図解の重ね合わせ

ここでは本研究で用いた KJ 法図解の重ね合わせの方法 [2] について述べる。本研究ではシステムの作成した図解とユーザの作成した図解を重ね合わせることによりユーザの作成した図解を変形させて発想の転換を行なうことを目的としてこの手法を用いているが、この手法は本来、複数のメンバが作成した図解の共通点や差異を求めることによりコミュニケーション活動に有益な情報を抽出することを目的に作られたものである。

KJ 法の図解の重ね合わせはファジィ類似度関係を用いて行なわれる。ファジィ類似度は木構造に対して相互変換可能なため、重ね合わせたい図解の階層構造を木構造に変換し、それをファジィ類似度行列に置き換え、ファジィ理論におけるファジィ類似度関係のもとで図解の比較・合成を行なう。

C.1 ファジィ類似関係

ファジィ類似関係とは以下の式で表せるファジィ関係である。

$$R = \int_{X \times Y} \mu_R(x, y)/(x, y)$$

$$\mu_R : X \times Y \to [0, 1]$$

$$(X \times Y = (x, y) | x \in X, y \in Y)$$

この式は類似度行列という形で表される。類似度行列は (a) 反射性、(b) 対称性、(c) 推移性の 3 つの性質を満たすものである。

(a) : $\mu_R(x,x) = 1$

(b) : $\mu_R(x,y) = \mu_R(y,x)$

(c) : $\mu_R(x,z) \leq \vee\mu_R(x,y) \wedge \mu_R(y,x)$
($\vee$, $\wedge$は max,min の意味)

この類似度行列は以下のようにして木構造と相互変換可能である。

- 類似度行列から木構造を作るには任意のレベル $\alpha \in [0,1]$ でαカットしたレベル行列を作り、同じ要素を持つ行同志をクラスタリングして木構造にする。
- 木構造から行列を作るには各分起点におけるレベル行列を作り、一つの類似度行列にまとめていく。

C.2 ファジィ類似度行列のマージ

それぞれの KJ 法図解を変換したファジィ類似度行列をマージしすることによりグループ全体のファジィ類似度行列を作成する。類似度行列をマージする方法にはさまざまな合成法を用いることができるが、各個人の意見を均等に反映するためには算術平均をマージアルゴリズムが適している。算術平均によるマージアルゴリズムは、基本的には行列の各要素について算術平均をとったものをグループの類似度行列とする。このままでは類似度行列の性質である (c) **推移性**を満たさなくなるので、算術平均を取った後に推移性を満たすために以下のような操作を行なう必要がある。

Step1: n 個の類似度行列から新しい行列 C を作成する。C の要素 c_{ij}は各類似度行列 A_k の対応する要素 $a_{k_{ij}}$ の平均で、以下の通りになる。

$$c_{ij} = \frac{\Sigma_{k=1}^{n} a_{k_ij}}{n}$$

Step2: C の全要素を降順に並べてリストを作る。

Step3: 新しい行列 D を作る。D の各要素を定めるために Step2 で作られたリストを順に取り出して x とし、Step3.1 から Step3.2 を繰り返す。

Step3.1: x と同じ値をもつ要素を行列 C から全て捜し出し、D の添字の要素に x を代入する。

Step3.2: Step3.1 で D に代入された全要素に関して、推移性をチェックして値の定まる部分については値を代入する。(d_{ij}に値をいれたときは、全ての k に対して $d_{i,k}$を調べ、$d_{i,k}$に値が定まっていたら d_{jk}の値を定める。

値の定め方は以下の通り。

d_{jk}の値も定まっている場合 推移性を満足している。

d_{jk}の値が未定で $d_{ij} = d_{ik}$の場合 $d_{jk} = d_{ij} = d_{ik}$

d_{jk}の値が未定で $d_{ij} \neq d_{ik}$の場合 $d_{ik} = min(d_{ij}, d_{ik})$

Step4: 終了。