

|              |                                                                                 |
|--------------|---------------------------------------------------------------------------------|
| Title        | スプリングモデルを用いたアイデア触発のための思考支援システムの構築                                               |
| Author(s)    | 高杉, 耕一                                                                          |
| Citation     |                                                                                 |
| Issue Date   | 1997-03                                                                         |
| Type         | Thesis or Dissertation                                                          |
| Text version | author                                                                          |
| URL          | <a href="http://hdl.handle.net/10119/1066">http://hdl.handle.net/10119/1066</a> |
| Rights       |                                                                                 |
| Description  | Supervisor: 國藤 進, 情報科学研究科, 修士                                                   |

# 修 士 論 文

## スプリングモデルを用いたアイデア触発のための 思考支援システムの構築

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

高杉 耕一

1997 年 2 月 14 日提出

## 要 旨

本論文では、テキスト情報から抽出したデータを二次元空間に動的に配置することによりアイデアの生成過程を支援するシステムを提案する。

# 目次

|          |                             |           |
|----------|-----------------------------|-----------|
| <b>1</b> | <b>序論</b>                   | <b>1</b>  |
| 1.1      | 本研究の動機と目的                   | 1         |
| 1.2      | 従来の研究                       | 2         |
| 1.2.1    | 発想支援                        | 3         |
| 1.2.2    | 発散的思考支援                     | 3         |
| 1.3      | 本研究の位置づけ                    | 4         |
| 1.4      | 本論文の構成                      | 5         |
| <b>2</b> | <b>類似度の算出と空間配置への反映</b>      | <b>6</b>  |
| 2.1      | 類似度の算出                      | 6         |
| 2.1.1    | キーワード自動抽出                   | 6         |
| 2.1.1.1  | 英文字の処理の詳細                   | 7         |
| 2.1.1.2  | 日本語の処理の詳細                   | 8         |
| 2.1.2    | キーワードベクトルの扱い                | 9         |
| 2.2      | キーワードベクトルによる類似関係            | 10        |
| 2.3      | スプリングモデルによる空間配置             | 11        |
| 2.4      | 類似度のスプリングの理想距離への変換          | 13        |
| 2.4.1    | システムとのインタラクションをサポートするための仕組み | 14        |
| <b>3</b> | <b>実験システム DW の構築</b>        | <b>15</b> |
| 3.1      | システムの概要                     | 15        |
| 3.2      | システムの操作                     | 17        |
| 3.2.1    | キーワード入力ウインドウ                | 17        |
| 3.2.2    | 二次元空間表示ウインドウ                | 17        |

---

|         |                               |    |
|---------|-------------------------------|----|
| 3.3     | 実装例                           | 20 |
| 3.4     | 本システムの特徴                      | 20 |
| 3.4.1   | システムとのインタラクションのサポート           | 20 |
| 3.4.2   | ユーザの要求により同じテキスト集合に対して異った配置の提供 | 20 |
| 3.4.3   | スプリングモデルをつかったアニメーション的配置による刺激  | 23 |
| 3.4.4   | リアルタイムでの実行                    | 23 |
| 3.5     | システムの実行結果                     | 23 |
| 3.5.1   | 関連キーワード検索                     | 23 |
| 3.5.2   | 関連テキスト検索                      | 23 |
| 4       | 評価実験                          | 26 |
| 4.1     | 要素技術レベルでの評価                   | 27 |
| 4.1.1   | 要素技術レベルの評価に対する考察              | 29 |
| 4.2     | 被験者を用いた評価実験について               | 29 |
| 4.3     | 利用レベルでの評価                     | 30 |
| 4.3.1   | 利用レベルの評価に対する考察                | 30 |
| 4.4     | 思考レベルの評価                      | 33 |
| 4.4.1   | アンケート                         | 33 |
| 4.4.2   | 被験者のアイデア生成までの流れの観察            | 35 |
| 4.4.2.1 | Aさんの事例                        | 36 |
| 4.4.2.2 | Bさんの事例                        | 36 |
| 4.4.2.3 | Cさんの事例                        | 37 |
| 4.4.2.4 | Dさんの事例                        | 37 |
| 4.5     | 全体の評価と考察                      | 38 |
| 4.5.1   | 空間表示の効果                       | 38 |
| 4.5.2   | インタラクションの効果                   | 38 |
| 4.5.3   | テキスト集合の質の違いからくる影響             | 39 |
| 4.5.4   | アイデア生成への支援                    | 39 |
| 5       | 本システムの応用的利用                   | 40 |
| 5.1     | World Wide Web 上での利用          | 40 |
| 5.2     | 従来の発想支援ツールへの適用                | 41 |

---

|       |                                     |    |
|-------|-------------------------------------|----|
| 5.2.1 | ブレインストーミングでの利用 (発散的思考支援) . . . . .  | 41 |
| 5.2.2 | KJ エディタの前段階での利用 (収束的思考支援) . . . . . | 41 |
| 6     | 関連研究との比較                            | 43 |
| 6.1   | スプリングモデルを採用した空間表示ツールとの比較 . . . . .  | 43 |
| 6.2   | 統計的手法を採用した空間表示ツールとの比較 . . . . .     | 43 |
| 7     | 結論                                  | 45 |
| 7.1   | 本研究の成果 . . . . .                    | 45 |
| 7.2   | 今後の課題 . . . . .                     | 46 |
|       | 謝辞                                  | 47 |
|       | 参考文献                                | 48 |

# 第 1 章

## 序論

本論文では、人間がアイデアを生成する過程で、どのようなコンピュータ支援ができるかを考え、アイデア触発のための思考支援システムのありかたについて考察した。

実際に本研究ではアイデア生成をはじめとした知的活動を支援するため、ネットワーク上のテキスト情報から得たキーワード情報をもとにテキスト間、キーワード間の類似関係を二次元空間に表示するシステムが有効であると考えた。またさらに、このシステムではユーザの概念空間を刺激するため、システムとのインタラクションを従来の研究より強化し、ユーザのさまざまな視点を反映することを目指した。そして、そのために有効な方法を模索し、スプリングモデルを採用した。さらに、この手法を用い実験システムを構築し、評価実験を通して思考支援システムとしての効果を検証した。

まず本章では、本研究の目的を述べ、次に従来の思考支援システムについての研究を見渡すことにより、本研究の目的と位置づけを明確にする。

### 1.1 本研究の動機と目的

本研究の目的はアイデア触発のためのコンピュータを用いた思考支援のあり方を実際のシステムを検証することにより明らかにすることにある。

まず、アイデア触発のための思考支援のあり方を論じるまえに、アイデアとはどのようなことなのか考える必要がある。J.W.Young はその著書 [1] でアイデアについて二つの原理を主張している。

- アイデアは一つの新しい組合せである。
- 新しい組合せを作り出す才能は事物の関連性を見つけ出す才能によって高められる

また、アイデア生成までの全過程は以下のようなものであると主張している。

1. 資料集めの段階、当面の課題のための資料と頃から貯蓄した一般知識による資料を蓄える
2. 吟味の段階、上記の資料を吟味する
3. 意識の外で組合せの作業を行なう
4. アイデアの実際上の誕生。「分かった、見付けた」という段階
5. 現実の有用性に合致させるために最終的にアイデアを具体化し、展開する段階

本研究ではこれらの主張を前提として踏まえ、次のような動機からシステムを構築した。

人間がものを考えるとき一人で考えると行詰まってしまうことが多い。そこで、人は自分の環境を変えたり、他人の意見を聞いたりして、自分の認識や視点を変えることによって、関連する新たな情報を得る。しかし、いつも自分が求めている人が相談にのってくれるとは限らず、また誰がその情報をもっているか分からないので、このような方法で求める解が得られるとは限らないし、非常にコストがかかる。一方、メールやニュースといったインターネット上のテキストには電子化された人々の考えやノウハウといったものが大量に存在する。

そこで、これらの雑多な情報をシステムがうまく分類し、ユーザに提示することは、ユーザが大量の情報の中から自分の必要としている情報を効率よく引き出すことを促進するとおもに、今まで気付かなかった関係を把握できるようになると考えた。また分類の仕方を変化させたり分類の方法を変化させることにより、ユーザに新たな認識や視点をもたらしアイデアを触発するものとする。

つまり、ネットワーク上のテキストを用いるのはJ.W.Youngの言う資料集めの支援であり、これらから関連キーワード、関連テキストを算出し空間上に配置することは、新しい組合せの中から関連性を見出すことを支援するものである。

本研究ではこのシステムを検証することによりアイデア触発といった思考支援の可能性を明らかにする。

## 1.2 従来の研究

近年のコンピュータ技術の急速な発達により、コンピュータを単なる計算機としてだけでなく、さまざまな知的活動に用いられるようになった。

他方、インターネットに代表されるコンピュータネットワークの一般社会への広がりにもなつて、ネットワーク上に大量の情報が提示されるようになり、従来の文章や図を書くといったレベルの知的活動支援から、情報収集、加工といったレベルの支援へ広がってきている。

これらの大量の情報をすべて消費することは不可能であり、我々はうまく選択、利用することになる。しかし、蓄積されている情報についての知識が乏しい場合や目的が明確でない場合、このような選択し、利用は非常に困難である。

そこで、従来の発想支援システムをこのようなネットワーク上の情報に対して活用していく研究が注目されている。そこで、これらの研究へ至る過程を従来の発想支援システムから順に概観することにする。

### 1.2.1 発想支援

本節では、従来の発想支援に関する研究の概要を述べる。[2] では発想支援ツールを以下のように分類している。

1. 発散的思考支援
2. 収束的思考支援

発散的思考とはアイデアの断片そのものを作り出して行く過程である。この段階ではアイデアそのものの評価は行わず質より量が重視される。発散的思考支援の代表的な例はブレインストーミングである。これは集団での議論を通して刺激し合うことによりアイデアを生成するものである。

他方、収束的思考とは逆にアイデアの断片をまとめ上げていく過程である。日本においてはKJ法に関する研究が多い。これは、収集されたデータについて主題ごとにラベルを作り、これを何段階にもグループ化してことにより問題解決をはかっていく手法である。

しかし、収束と発散の分類は明確に区別できるものではない[3]。これは、収束的思考の結果である何らかのアイデアが次の発散的思考を促す原理や法則となりえるからである。また、多くの発想法において発散、収束は繰り返し行なわれている。

本研究ではアイデアの断片そのものの生成を支援して行こうという発散的思考支援を目指している。そこで発散的思考支援の枠組に関して次節で述べる。

### 1.2.2 発散的思考支援

全節に述べた発散的思考支援について L.F.Young は [4][5] において以下のように分類している。

- 秘書レベル (The Secretarial Level)

基本的には、計算機を動的な電子黒板として用いるものである。ユーザの思考がツールに表示され、システムはそれを操作するための機能をもつ、概念操作にまつわる雑用を計算機が肩代りするのである。ユーザは創造的な活動に専念できるので、結果として発散支援効果がえられる。

- 枠組-パラダイムレベル (The Framework-Paradigm Level)

ユーザが考えをまとめるに当たって適切な枠組を選んで提供するものである。枠組は、ユーザの思考を構造化するだけでなく、思考の流れをガイドする役割もはたす。このレベルではツールは思考の流れをガイドはするが、ユーザの考えている問題に対して直接解答を与えることはない。

- 生成レベル (The Generative Level) ユーザが入力した単語を統合したり関連づけることによって、新たなアイデアを生成して提供するものである。このレベルにおいても、ツールはアイデアを生成するための枠組を持っている。しかし、枠組-パラダイムレベルと異なり生成レベルのツールはユーザの考えている問題に対する解答となり得るアイデアを出力する。もちろん、それらを採用するかどうかはユーザの自由である。

この三つのレベルは、後のものほど支援の度合いや洗練度が高いが、高いレベルの支援が必ずしも効果的な支援でないと L.Y.Young は主張している。

文献 [3] ではさまざまな発想支援ツールをこれらの分類によって紹介している。ここで、折原はアイデア断片そのものを生成するよりもそれに通じる観点に重きをおいたツールが効果的であると指摘している。

### 1.3 本研究の位置づけ

本研究は L.F.Young のいう枠組-パラダイムレベルの思考支援である。秘書レベルでの思考支援ツールはワープロ、データベース、グループウェアを始め現在一般的に使われているものの多くはこのレベルであり、それなりに成功を納めている。しかし、これでは思考支援というにはレベルが低すぎてアイデアの断片生成を積極的に支援しているとは言えない。もう少しアイデアの断片生成という点に踏み込んで支援していく必要があるのではなかろうか。また、生成レベルでの思考支援ツールは無益なアイデアを大量にだし続ける可能性があり、探索空間の絞り込みなどテクニックが必要であるが、現在のところ、あまり有用な支援ができていないし、コンピュータがアイデアの生成そのものを行うことには疑問を持たざるおえない。

そこで、枠組 パラダイムレベルでの思考支援ツールが注目されており、従来からいくつかの思考支援ツールが提案されてきた。このレベルの思考支援ツールではなんらかの類似関係を示唆させるため、ネットワーク資源を利用したツールが多い。しかしながら、これらの研究における問題点として、関連発見のためのシステムとのインタラクションをサポートするしくみが弱いことがある。このインタラクションはユーザがシステムが与えたアイデアに通じる視点をより多様な形で獲得するためには必要不可欠のものである。

例えば、このレベルでの思考支援ツールとして、ネットニュースを元にした連想辞書や関連テキスト検索などが行われてきた。[6] しかし、連想辞書のようなものでは多くの場合関連がリストで表示され、個々の関連テキスト間の二次元の類似関係が分かりにくく、さらにでてきた関係に対してユーザが操作を行い、それをシステムにフィードバックする仕組みがまったくない。

また、論文などの文献を空間配置することにより情報検索を行うシステムも提案されてきた [2]。これらのシステムではユーザの視点によってマッピングしなおす仕組みもある。しかしながら、空間配置を行うのに双対尺度構成法など統計的手法が使われ、全体のテキストをある基準でマッピングすることには成功しているが、空間配置された個々に対して後からユーザの要求に追加、削除、移動などを行なうと配置の再構成により、空間配置が急激に変化してしまい思考が分断されてしまう。また、配置された個々の類似関係を直接二次元の距離空間に反映させていないため、人間の直観にあわない点もある。

そこで、本システムではシステムとのインタラクションをサポートする仕組みを強化するためにスプリングモデルを採用し類似関係を直接空間にマッピングした。これにより、アイデアの触発の鍵となる関連の発見の作業を従来の研究よりも積極的に支援できるものと考えた。

なお、関連研究との比較は6章において細かく検討することにする。

## 1.4 本論文の構成

本論文は本章も含め7章から構成される。第2章では、本研究で採用したキーワード、テキスト間の類似度を算出し、空間配置をする方法論について説明する。

第3章ではその方法論に基づき実装したプロトタイプシステムについて述べる。さらに、システムの特徴を詳しくみていく。

第4章では本システムの評価を行ない、本システムの効果について考察する。第5章では本システムの応用について検討する。

第6章では類似する研究との比較を行ない本研究の位置づけを明確にする。最後に本論文の結論と課題を第7章で述べる。

## 第 2 章

# 類似度の算出と空間配置への反映

本章では、本研究で用いた手法について述べる。本研究ではテキスト間、キーワード間の類似度を空間配置に反映させるため以下のようなステップを踏む。

1. キーワード自動抽出  
各テキストからキーワードとなる単語を抽出
2. キーワードベクトルの算出  
キーワードの頻度情報といくつかのヒューリスティックを基に各テキストそれぞれに対しキーワードベクトルを得る
3. 類似度の計算  
得られたキーワードベクトルを用いて各テキスト間、各キーワード間の類似度を算出する
4. 空間配置  
類似度を理想距離としスプリングモデルにより空間配置

これらの各ステップに対しその手法を述べ、最後にシステムとのインタラクションをサポートするための仕組みについて述べる。

### 2.1 類似度の算出

#### 2.1.1 キーワード自動抽出

ここでは各テキストから名詞を抽出する過程について述べる。類似度の算出を適切に行なうためには各テキストの特徴的な単語 (ここではキーワードと呼ぶ) を抽出することが重要である。か

ねてから筆者は情報共有化システム FISH に関連しキーワード自動抽出に関して検討を行った [7]。この FISH での数千に及ぶユーザによるキーワードづけや検索状況を解析することによりキーワードとなり得る語に関して検討をした。本システムは情報共有システムではないが、同様の処理が行なえると考えた<sup>1</sup>。

本研究ではテキストとして、電子ニュースの文章を対象とし図 2.1 に示すような算出方法を採用した。具体的な手順は以下に示す。

1. 不要な情報をあらかじめ削除する。  
(ヘッダ、括弧の中身、挨拶文)
2. テキストを日本語と英語に分ける。
3. 日本語は形態素解析にかけ名詞を抽出し、英語は汎用の不要語リストにかける。
4. 各単語に対し各テキストにおける出現頻度を各キーワードの重みとする。このとき、Subject と文頭の二行に出現したキーワードに対しては二倍の重みとする。

#### 2.1.1.1 英文字の処理の詳細

英語の不要語リストは [8] に掲載されているリストに必要と思われるのを加えた 471 語である。また、複数形の s や ing などは以下のヒューリスティックで原型にもどした。

- 複数形、三人称単数の s
  1. ss, ch, sh に続く es は削除
  2. ies で終わる単語を i を y に変えて削除 (例外 series 等)
  3. s, i, u, o 意外の文字に s が続いている場合 s を削除 (例外 news 等)
- ing 形
  1. nn, dd, tt, mm に続く ing は ing と一つ前の文字を削除
  2. z, at, od, ut, nc, iv, as, erg, ov, ud, rs, ed, os に ing が続く場合 ing を削除し e を付加
  3. 上記以外の ing は削除 (例外 ping 等)
- 過去、過去分詞形

---

<sup>1</sup>FISH では計算機環境の制約から形態素解析ではなく、字面解析となっている

1. ied で終わる場合 i を y で置き換えて ed を削除
2. z,at,od,ut,nc,iv,as,erg,ov,ud,rs,ed,os に ed が続く場合、最後の d を削除
3. 上記以外の ed は削除

#### 2.1.1.2 日本語の処理の詳細

日本語の部分は形態素解析で名詞を抽出した。ここでは日本語形態素解析として「茶筌」[9](以下 CHASEN とする) に EDR 電子化辞書等 [10] を付加したものを使用し、未定義語または名詞となる単語をキーワードとして採用した。

ただし、出てきた単語のなかで一文字のものやひらがなのみの単語はキーワードとして採用しなかった。

実際に CHASEN では以下のように文章から、名詞を抽出することができる。

入力文:

高杉は『スプリングモデルを用いたアイデア触発のための思考支援システムの構築』という題で修士論文を書いた。

出力:

|       |       |       |      |
|-------|-------|-------|------|
| 高杉    | たかすぎ  | 高杉    | 固有名詞 |
| スプリング | すぷりんぐ | スプリング | 普通名詞 |
| モデル   | もでる   | モデル   | 普通名詞 |
| アイデア  | あいであ  | アイデア  | 普通名詞 |
| 触発    | しょくはつ | 触発    | サ変名詞 |
| 思考    | しこう   | 思考    | サ変名詞 |
| 支援    | しえん   | 支援    | サ変名詞 |
| システム  | しすてむ  | システム  | 普通名詞 |
| 構築    | こうちく  | 構築    | サ変名詞 |
| 題     | だい    | 題     | サ変名詞 |
| 修士    | しゅうし  | 修士    | 普通名詞 |
| 論文    | ろんぶん  | 論文    | 普通名詞 |

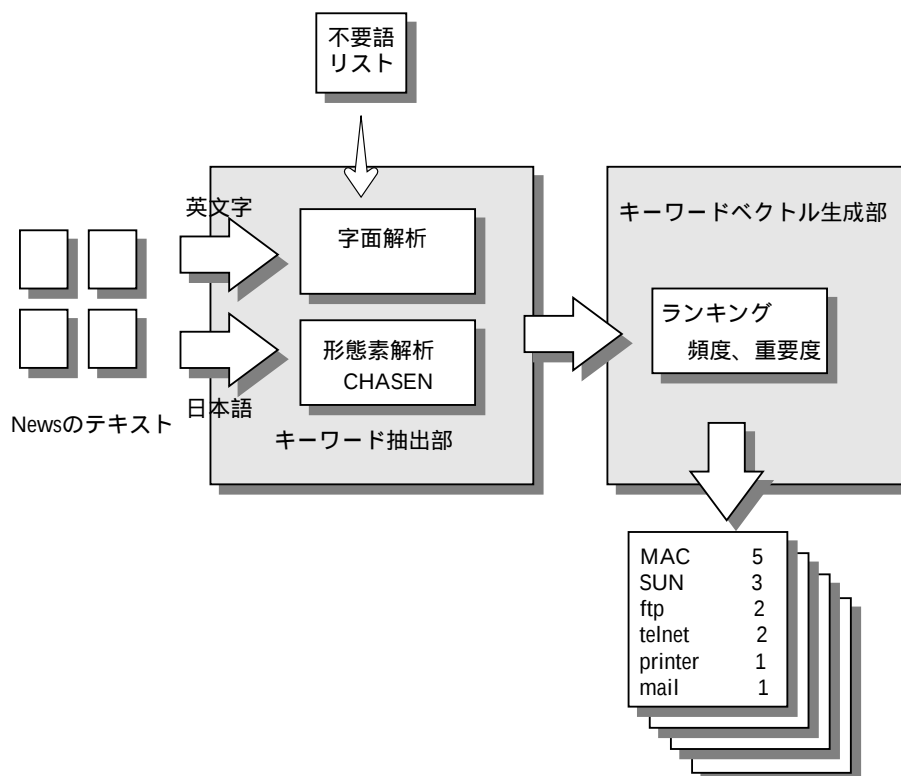


図 2.1: キーワード自動抽出→ キーワードベクトル算出

### 2.1.2 キーワードベクトルの扱い

キーワードベクトルの扱いについては従来から広く行われている手法をもちいた。前節で述べたように、各文章に対して名詞抽出ランキングを行った結果、出現したキーワードの総数を  $n$  とし、各キーワードに番号をふる。あるテキスト  $t$  の  $i$  番目のキーワードの重み (ランキングの結果) を  $W_i^t$  とするとこのテキスト  $t$  のキーワードベクトルは

$$\mathbf{a}_t = (W_1^t, W_2^t, \dots, W_i^t, \dots, W_n^t)$$

となる。

しかしながらこれでは、各キーワードの重みを平坦にみており各テキストに特徴的なキーワードの重みがうかびあがらない。そこでキーワードのテキスト郡全体の出現確率と各テキストにおける出現確率を考慮することによって重みをつけ直す。 $p_i^t$ はテキスト  $t$  におけるキーワード  $i$  の出現確率で  $p_i^{all}$ はテキスト集合全体におけるキーワード  $i$  の出現確率である。

$$p_i^t = \frac{W_i^t}{\sum_{\text{all } k} W_i^k}$$

$$p_i^{all} = \frac{\sum_{all\ k} W_i^k}{\sum_{all\ k} \sum_{all\ i} W_i^k}$$

重要度の計算はさまざまなものが存在するが、ここでは計算量を抑えるため Keyword Associator[6]と同様の方法を採用した。キーワード  $i$  の全体に対する出現確率とテキスト  $t$  の出現確率の比の対数を掛け合わせる。 $(p_i^{all}$  の  $p_i^t$  からのダイバージェンス)

$$W_i^{t'} = p_i^t \log \left( \frac{p_i^t}{p_i^{all}} \right)$$

これによって、全体における出現確率に対してテキスト  $t$  における出現確率が高いほど重みがつく。

$$\mathbf{a}_t' = (W_1^{t'}, W_2^{t'}, \dots, W_i^{t'}, \dots, W_n^{t'})$$

さらに、これを正規化するため、

$$\mathbf{v}_t = \frac{\mathbf{a}_t'}{|\mathbf{a}_t'|}$$

とする。(ただし、 $|\mathbf{a}_t'| = \sqrt{W_1^{t'^2} + \dots + W_n^{t'^2}}$  である。) 今後、この  $\mathbf{v}$  をテキスト  $t$  のキーワードベクトルとして扱う。

## 2.2 キーワードベクトルによる類似関係

キーワードベクトルから類似度を算出する方法はいくつかあるが計算方法が簡単であることからベクトルの内積で得ることにしてた<sup>2</sup>。本システムでは  $\mathbf{v}_{t_1}$ 、 $\mathbf{v}_{t_2}$  のキーワードベクトルをもつテキスト  $t_1, t_2$  間の類似度  $R_{12}$  を次のようにベクトルの内積で定義する。

$$R_{12} = \mathbf{v}_{t_1} \cdot \mathbf{v}_{t_2}$$

キーワード  $i, j$  間の類似度  $r_{ij}$  の定義は  $\mathbf{e}_i, \mathbf{e}_j$  をそれぞれ各単語  $i, j$  の単位ベクトル<sup>3</sup> とすると、

$$r_{ij} = \sum_t (\mathbf{e}_i \cdot \mathbf{v}_t) (\mathbf{v}_t \cdot \mathbf{e}_j)$$

となる。

<sup>2</sup>すでに正規化を行なっているのでコサイン関数を用いる方法と等価である

<sup>3</sup> $i$  または  $j$  番目の成分のみが 1 で他が 0 のベクトル

### 2.3 スプリングモデルによる空間配置

スプリングモデルはグラフの描画法として頂点を美しく配置する手法として定評がある。この手法を用いると、スプリングの自然長を頂点間の理想距離としておけば、スプリングの伸縮により頂点間の距離が理想距離に近づくように準最適化をして空間配置してくれる。今回はこのスプリングモデルを用いて類似関係を空間配置に反映させる。スプリングモデルはスプリングの張り方、伸縮したときの力のかかる度合などの違いによりいくつかの手法が提案されているが[11]、本システムではもっとも普通のスプリングに近いモデル[12]を採用した。ただし、本研究では類似度によって頂点間の距離を変化させたいので、理想距離の与え方が[12]とすこし異なる。頂点*i*と

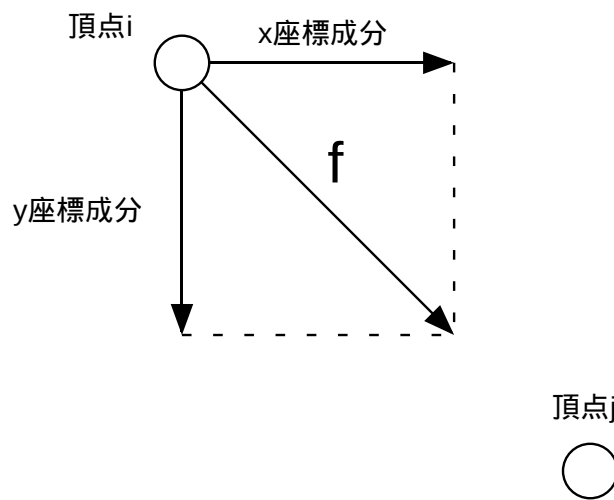


図 2.2: 頂点間のスプリング

頂点 *j* 間の理想距離 (スプリングの自然長) を  $l_{ij}$ 、実際の頂点間の距離を  $d_{ij}$ 、ばね定数を  $k_{ij}$  としたときスプリングの張られた頂点間の働く力は

$$f_{ij} = k_{ij}(d_{ij} - l_{ij})$$

となる。ただし、 $k_{ij}$  は任意の正数とする。 $k_{ij}$  の扱いはシステムとのインタラクションをサポートするための仕組みとして、この章の最後で述べるが、 $k_{ij}$  が大きい程そのばねの全体への影響が大きくなる。ここで  $f$  が正のとき引力となり負のとき斥力となる。このとき、全体のエネルギーの総計  $E$  は

$$E = \sum_i \sum_j \frac{1}{2} k_{ij} (d_{ij} - l_{ij})^2$$

となる。ある2つの頂点の座標を  $(x_i, y_i)$ 、 $(x_j, y_j)$  とすると、これらの間の距離  $d$  は

$$d = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

となる。よってこの間のエネルギー  $E_{ij}$  は

$$E_{ij} = \frac{1}{2} \left( \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} - l_{ij} \right)^2$$

である。これの  $x_i$  での偏微分を  $\frac{\partial E_{ij}}{\partial x_i}$  同様に  $y_i$  で偏微分したものを  $\frac{\partial E_{ij}}{\partial y_i}$  とするとそれぞれ、以下のようになる。

$$\begin{aligned} \frac{\partial E_{ij}}{\partial x_i} &= k \left( (x_i - x_j) - \frac{l_{ij}(x_i - x_j)}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}} \right) \\ \frac{\partial E_{ij}}{\partial y_i} &= k \left( (y_i - y_j) - \frac{l_{ij}(y_i - y_j)}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}} \right) \end{aligned}$$

これを用いて、スプリングの力のベクトル  $\vec{f}_{ij}$  は以下ようになる。

$$\vec{f}_{ij} = \left( \frac{\partial E_{ij}}{\partial x_i}, \frac{\partial E_{ij}}{\partial y_i} \right)$$

となる。したがって、各頂点の偏微分値はその頂点に働く力の各座標成分に相当する。そして、各頂点の座標で偏微分した値を移動量として定義すれば  $E$  を最小にすることができる。つまり  $c$  を任意の正定数として  $\delta_x = c \frac{\partial E_{ij}}{\partial x_i}$ 、 $\delta_y = c \frac{\partial E_{ij}}{\partial y_i}$  を移動量として定義する。 $c$  が大きければ速く収束に向かうが動きが粗くなり、 $c$  が小さければゆっくり滑らかに収束するが計算量がそれだけかかる。

実際のアルゴリズムは以下ようになる。

**描画アルゴリズム** (one step)

```

begin
  for  $i = 1$  to  $n$  do  $\rightarrow$  各頂点に対して ( $n$  は頂点の数)
    begin
      for 頂点  $i$  に設定されている各ばね do
        begin
           $\delta_x, \delta_y$  を計算する
           $\delta'_x \leftarrow \delta'_x + \delta_x$ ;
           $\delta'_y \leftarrow \delta'_y + \delta_y$ ;
        end
      end
       $x \leftarrow x + \delta'_x$  新しい頂点の座標に更新する
       $y \leftarrow y + \delta'_y$ 
    end
  end
  全ての頂点の重心を計算し、重心が画面の中心にくるように座標を平行移動
end

```

これを繰り返し各頂点を配置していくことにより、全体のエネルギー現象させていくことができる。これは頂点間の距離  $d_{ij}$  が理想距離  $l_{ij}$  に近付くような準最適解を求めることに相当する。

実際に画面に表示するには、このままでは、画面からはみ出たり、特定の方向に頂点がかたまってしまう。そこで、座標を更新するごとに頂点全体の重心をとり、それを表示画面の中央にくるように平行移動させる。それでも画面におさまらない頂点は画面の端に置く。

## 2.4 類似度のスプリングの理想距離への変換

テキスト間、キーワード間の類似度を距離空間に反映するため類似度  $R_{ij}$  から以下のようにばねの理想距離  $l_{ij}^1$  を算出する。

$$l_{ij}^1 = m \frac{1}{R_{ij}}$$

また、キーワード間の場合も同様に  $r_{ij}$  から次のように  $l_{ij}$  を与える。

$$l_{ij}^1 = m \frac{1}{r_{ij}}$$

このように類似度の逆数を理想距離とすることにより、類似度の高いテキストやキーワードは近くに、類似度の低いテキストやキーワードは遠くに配置するようにできる。 $m$  は任意の定数で空

間配置しようとする画面の大きさにより調整する。大きければ画面からはみ出るし、小さければ頂点が重なりあって見にくい。

また、必要に応じて対数をとったものも用意する。これは、頂点が重なりすぎるのをふせげる。

$$l_{ij}^2 = m \log \left( \frac{1}{R_{ij}} \right)$$

$$l_{ij}^2 = m \log \left( \frac{1}{r_{ij}} \right)$$

#### 2.4.1 システムとのインタラクションをサポートするための仕組み

ここでは、ユーザが空間配置に対して介入するための仕組みを述べる。システムが空間配置に用いたばねは仮想的なもので、スプリング定数  $k$  やスプリングの自然長  $d$  でさえ可変であることに注意されたい。

ラベルの移動では移動するラベル以外のラベルを再配置する必要がある。そのとき、ユーザはそのラベルに興味をもったということで、移動したラベルの影響を多少強くしながら移動させるのが望ましいであろう。そこでラベルに張られているスプリング定数  $k$  をその他のスプリングのスプリング定数の3~7倍に強くすることで容易にこれに対応できる。

ラベルの削除では逆にユーザにとって当面興味の少ないラベルであったとして、削除したラベルにはられているスプリング定数を  $k = 0$  にすることで削除したラベルの全体の空間配置に対する影響をなくすることができる。このようにすれば実際のスプリング自体は残っているため、スプリング定数をもとに元に戻すだけで、削除したラベルを復活させることも簡単にできる。

また、スプリングの自然長を変えることにより、複数の類似度を連続的に空間配置に表現できる。前節で述べたように同じ類似度から複数の理想距離(スプリングの自然長)を得ることができ、英語と日本語の両方のテキストが用意されているとき、個別に類似度を計算することにより、複数の理想距離が算出できる。これらの理想距離をユーザが選択していくことにより次々に切替えて行くことができる。

以上の操作を行っても新しい空間配置は古い空間配置から徐々に変化していくことに注意されたい<sup>4</sup>。

<sup>4</sup> 勿論ユーザが要求すればまったく異なる空間配置に飛ぶことも出来るが

## 第 3 章

# 実験システム DW の構築

第 2 章に述べた手法をもとに UNIX の X-Window 上に実験システム DW (Dancing Word) を実装した。最初にシステムの概観を紹介し、次にシステムのインターフェース部分の説明をする。最後にシステムの特徴についてまとめる。

### 3.1 システムの概要

システムの概要は図 3.1 に示す通りである。本システムはユーザがキーワードを入力することにより、そのキーワードに関連するテキストをラベルとしてスプリングモデルにより空間配置し、雑多な情報を見易くユーザに提示する。ユーザは発想または得たい情報のキーとなる単語を入力しシステムによって得られた空間配置から個々のテキスト間の距離を読みとる。その後、各テキストにアクセスしていき、得たい情報を効率よく得ることができる。また、ユーザはシステムによって得られた空間配置が気に入らなければシステムに何回でも空間配置をやり直すように要求できる。さらに、ユーザはシステムによる空間配置の途中でも、ラベルの移動、削除といったシステムへの介入が自由に行える。

本システムは関連キーワードに関しても同様の空間配置を提示する。これは、適切なキーワードが思い浮かばない場合の支援を目的としている。

また、検索速度を速めるため、キーワードベクトルの算出と類似度の計算はあるテキスト集合に対してあらかじめ行っておきテーブルとして保存しておく。

今回、テキストをネットニュースのテキスト及び修論の概要 (T<sub>E</sub>X 形式) とし、前章で述べた手法に基づき実装をした。文字処理は Perl、GUI は Tcl/Tk と Athena Widget、日本語形態素解析として CHASEN を用いた。

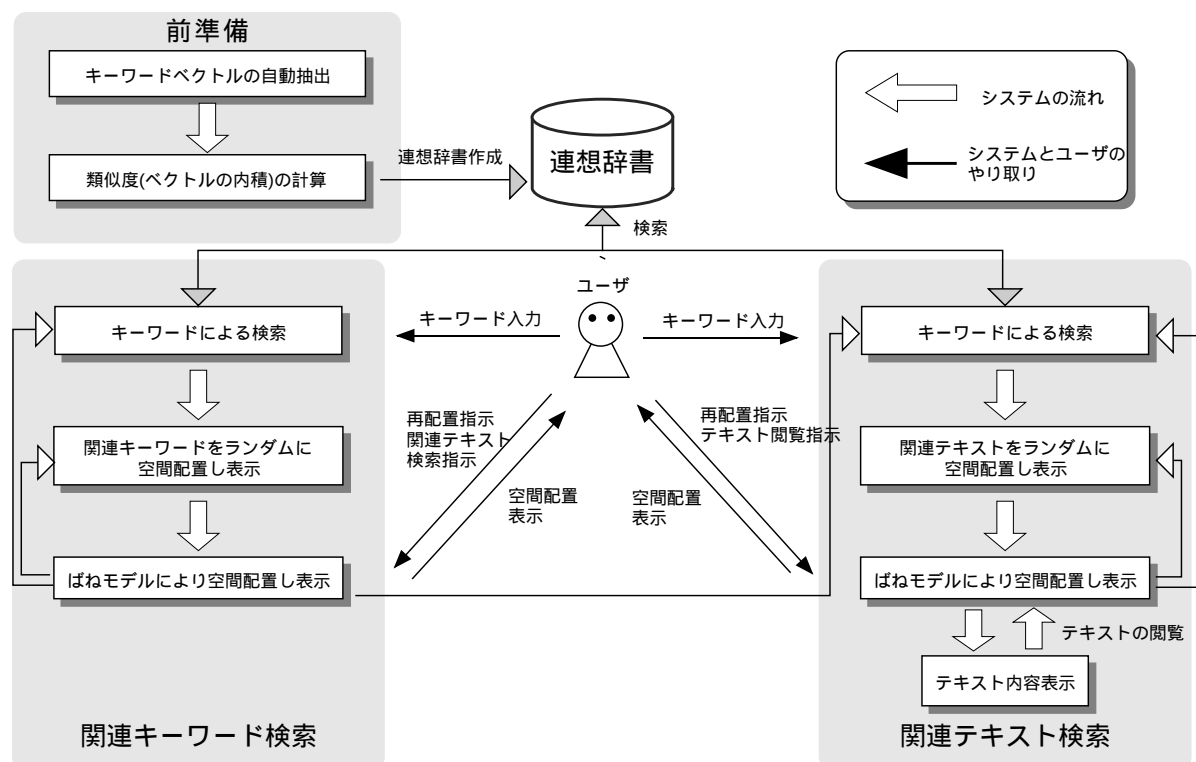


図 3.1: システムの概要

## 3.2 システムの操作

### 3.2.1 キーワード入力ウインドウ

ユーザがシステムを起動するとまず、キーワード入力ウインドウが開く 3.2。このウインドウでは対象とするテーマの選択とそのテーマに関するテキスト集合の中でキーワードによる関連テキスト、関連キーワード検索ができる。検索キーワードは二種類まで入力することができ、AND または OR 検索が可能である。検索結果はキーワード入力ウインドウの右下のリスト、または次節で述べる二次元空間表示ウインドウに表示することが出来る。



図 3.2: キーワード入力ウインドウ

### 3.2.2 二次元空間表示ウインドウ

これは前述のキーワード入力ウインドウで入力したキーワードに対する関連テキスト、関連キーワードを二次元表示するウインドウである。関連キーワードの二次元表示の例を図 3.3 に、関連テキストの二次元表示の例を図 3.4 に示す。二次元空間表示されたそれぞれのラベルをここではオブジェクトとよぶことにする。図からは読み取りにくい、各オブジェクトには色分けがしてあり、関連キーワードの場合は検索キーワード、関連テキストの場合は最も関連が深いテキストが赤で表示され、関連度が上位 30% のものは青、その他を濃緑に塗りわけてある。ユーザはウインドウの上部にあるボタンによってスプリング定数、表示スケール等のパラメータを調整するとともに、

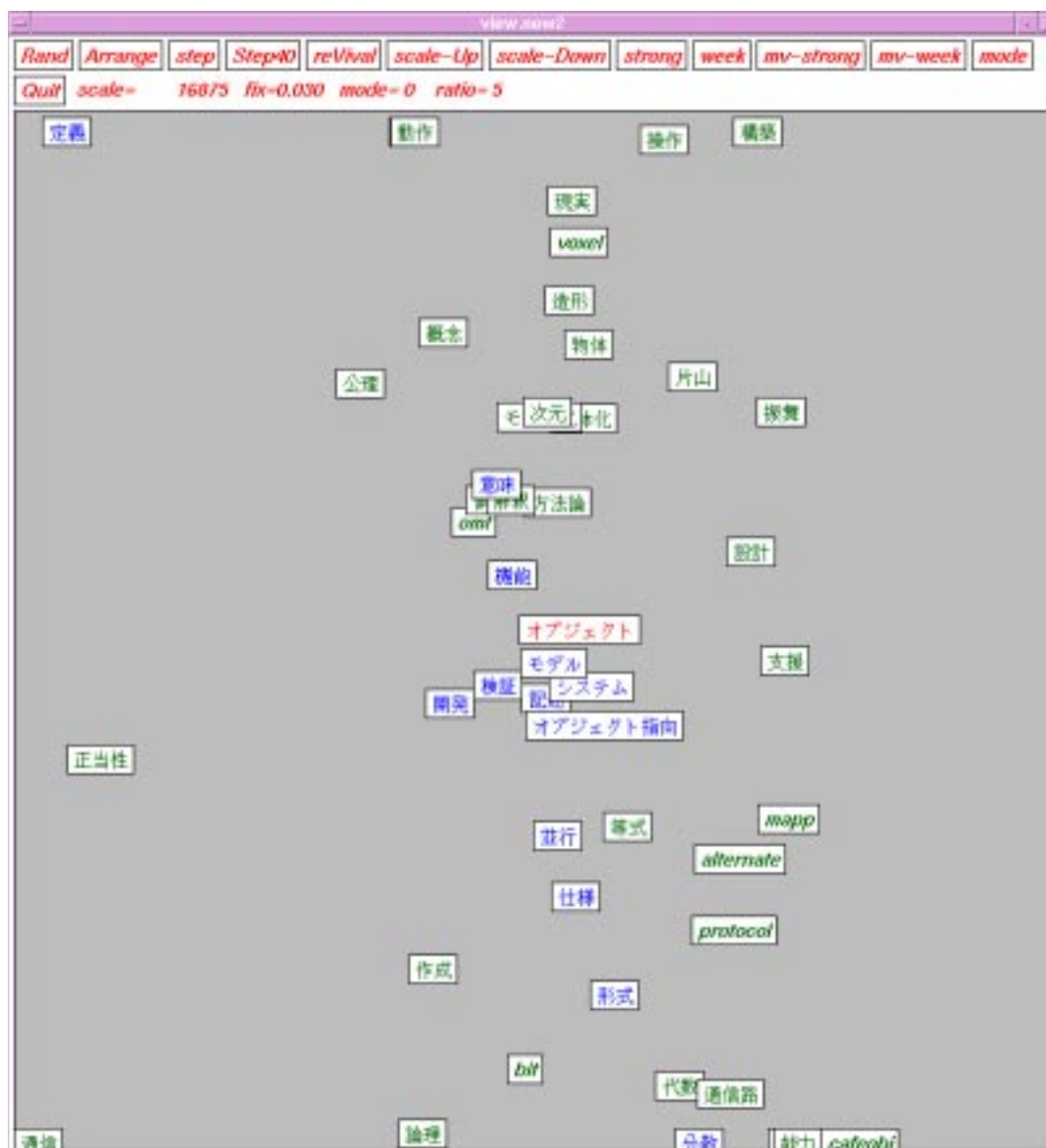


図 3.3: 二次元空間表示ウィンドウ (関連キーワード検索)

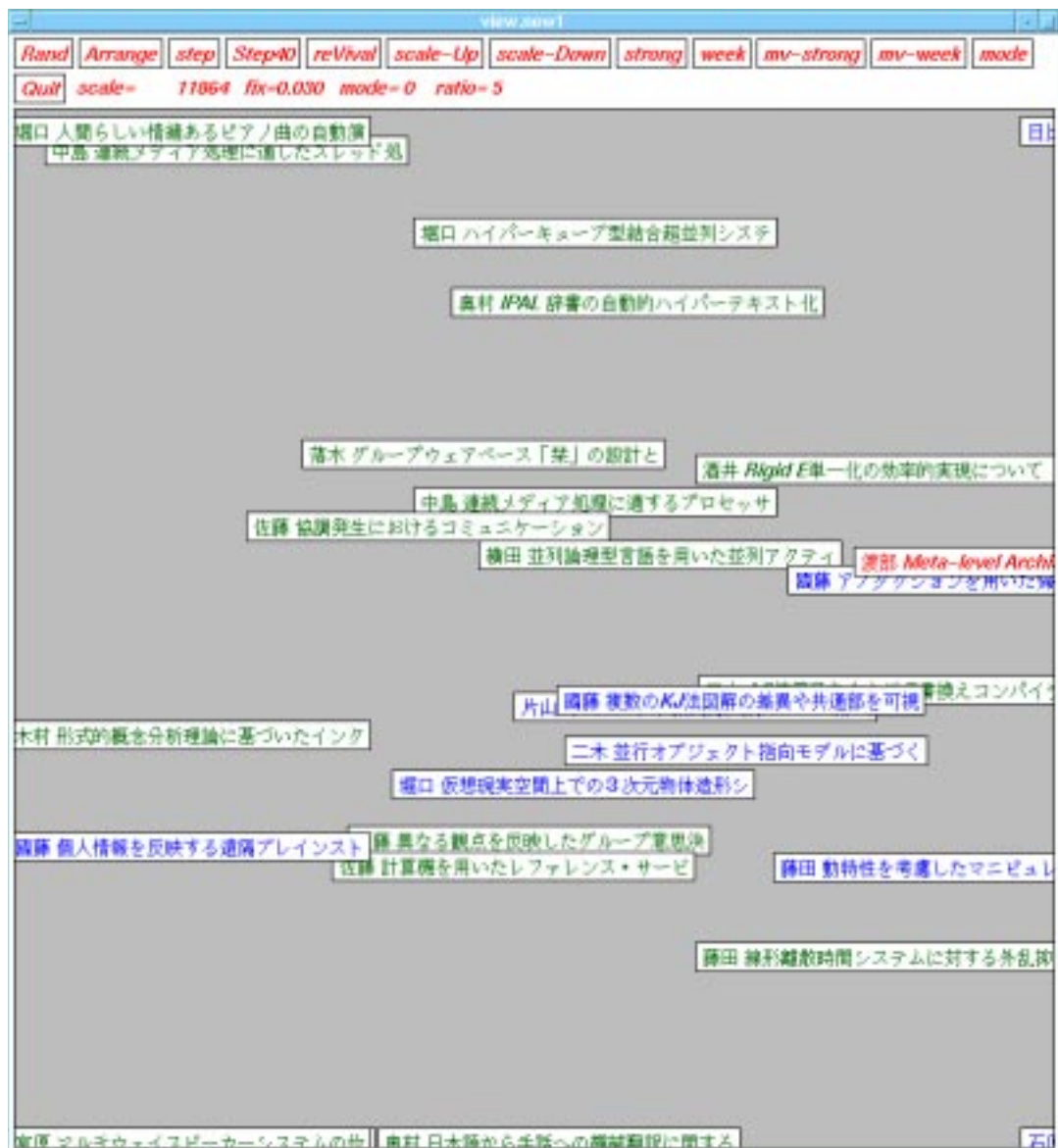


図 3.4: 二次元空間表示ウィンドウ (関連テキスト検索)

ランダムに配置したり、スプリングモデルによる配置を少しずつ行なったり収束段階まで行なったり出来る。さらに、ウインドウ内のオブジェクトをクリックすることにより、オブジェクトを削除したり、テキストオブジェクトの場合はテキストの内容を閲覧でき、キーワードオブジェクトの場合はそのキーワードに関連するテキストを表示できる。またオブジェクトをマウスでドラッグすることによりオブジェクトの移動ができる。この時、移動にともなって、スプリングモデルにより周辺のオブジェクトの再配置が自動的になされる。

### 3.3 実装例

実装例は図 3.5 である。左上のウインドウがキーワード入力ウインドウで、その下にあるのが空間配置ウインドウである。キーワード入力ウインドウからは複数の空間配置ウインドウを開くことができ、ユーザは並行して検索行動をとることが出来る。

また右下のウインドウは、テキストの閲覧ウインドウである。

### 3.4 本システムの特徴

#### 3.4.1 システムとのインタラクションのサポート

本システムではユーザに対話的なインターフェースを提供する。ユーザはシステムが空間配置を行っていく過程において、各ラベルを移動したり、興味のないラベルを削除するといった具合に空間配置に介入できる。これを繰り返し行うことにより、システムの一方向的な表示ではなく、ユーザ自身による思考錯誤を促進し、ユーザに多様な視点を示せる。より積極的な利用を促せると考えた。

また、今回スプリングモデルによる描画を行っているため、このような介入を行っても配置が徐々に新しい配置へ変化していくため、ユーザの思考がとぎれにくいと考えられる。

#### 3.4.2 ユーザの要求により同じテキスト集合に対して異った配置の提供

仮に適切な類似度が算出できたとしても、人間が把握できるのは 2 次元または 3 次元までで、多次元の類似関係は容易には把握できない。本システムでは多次元のことを二次元空間上に表示するため、ある程度のゆがみは必ず生じる。

そこでユーザの要求により複数の異った配置を提供する。これによりユーザは違った視点 (多次元から二次元への射影) を得ることができ、アイデアの触発に影響があるものと考えている。図

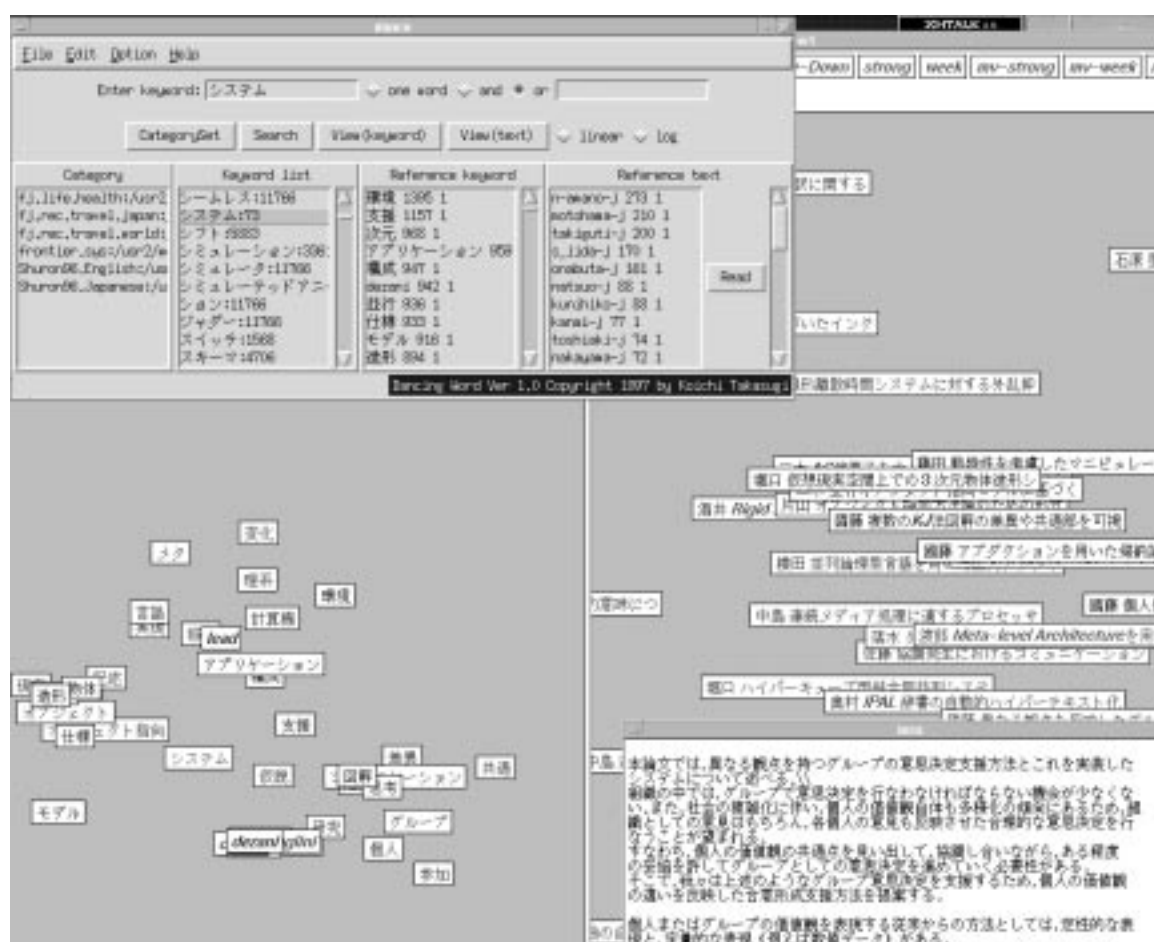


図 3.5: システムの実装例

3.9と図 3.8は同じ検索結果の空間配置であるが、異った初期配置から描画しているため、異なった視点を提供している例である。ここでの各ラベルは各記事の番号と Subject である。



図 3.6: 配置 1

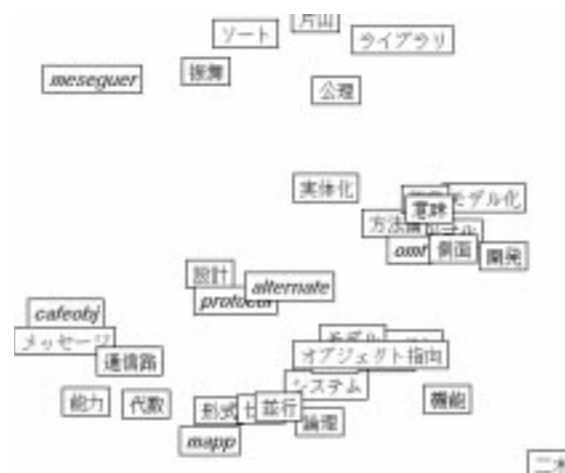


図 3.7: 配置 2

### 3.4.3 スプリングモデルをつかったアニメーション的配置による刺激

単に結果をユーザに与えるだけでなく、その過程をみせることがアイデアの触発に役立つと筆者は考えている。特に今回はスプリングモデルという親しみやすいモデルを用いているので、ユーザはオブジェクト間のひっぱりあいや、収束するときの速度を感じられるはずである。また、さまざまなユーザの空間配置への介入に対しても徐々に新しい空間配置へ移行できるので、ユーザの思考の思考を中断しないだけでなく移行するときの動きからも類似関係が読み取れるはずである。

### 3.4.4 リアルタイムでの実行

ここでの問題点は計算量である。今回はすべてのオブジェクト間に対してばねを設定したため、オブジェクト集合 (頂点集合) を  $V$  とすると  $O(|V|^2)$  の計算量となっている。今後頂点の数が増えた場合、計算量を減らすため、ある程度類似度の低い頂点間にはスプリングを設定しないなどの工夫が必要である。

## 3.5 システムの実行結果

計算機関係のローカルニュース (記事数 160) に対してシステムを適用してみた。

### 3.5.1 関連キーワード検索

http というキーワードで検索した結果が図 3.8 である。ホームページに関すること (ヤフー, サイト, ページ, http など) や、パケットとモニタリング service と port のように、近い意味のキーワードが固まっているのがわかる。

さらに、このニュースグループで良く発言する人や proxy などのサーバの管理にする話題が多そうなのも予想できた。

また、この分野に疎い人も関連するキーワードを広げるだけでなく、「ヤフーってデータベースなのかな? 」といった疑問も引き出せると推察される。

### 3.5.2 関連テキスト検索

また、図 3.9 は local というキーワードで関連テキスト検索した結果である。図からは分かりにくいですが、下の方には /usr/local の障害に関して、真中には /usr/local へのインストールに関して、左上の方には /usr/local の管理者会議に関すること、というように分類、表示されている。

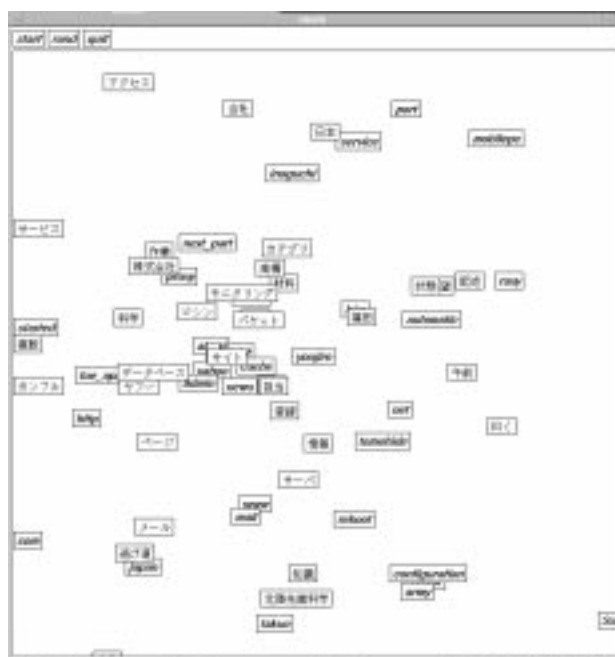


図 3.8: 関連キーワード検索例

`/usr/local` は本学のファイルサーバで `src-admin` といわれるグループによって管理されている。結局、各テキストの内容をいくつか閲覧すると、本学のファイルサーバの問題点は3つに分けられることがわかった。ファイルの上書きやファイルサーバのダウンなどの障害問題、どこへどのようなファイルを置くべきかというインストールに関する問題、どのような人が管理すべきかという問題といった具合である。特にインストールに関しては活発な議論が展開されていることも読みとれた。

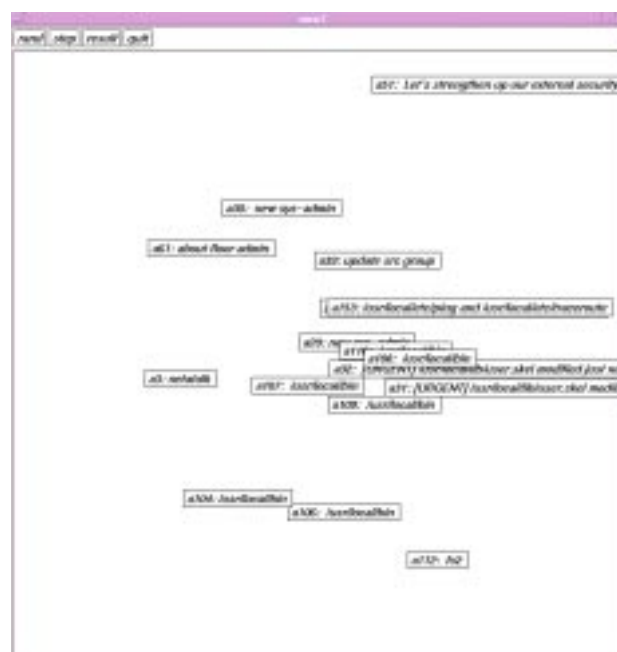


図 3.9: 関連テキスト検索例

## 第 4 章

# 評価実験

本章では、作成したシステムに対する評価について述べる。従来から、思考支援システムの評価については、その必要性がさげばれている [13]。それにも関わらず定量かつ客観的に計ることは困難であるとの指摘が多くなされている [14]。特に本システムのような発散的思考支援システムでは生成されたアイデアが適切かどうかを判断すべく客観的方法はなく、ユーザの満足度といった極めて主観的なものとなってしまう。また、生成されたアイデアがシステムの支援によってなされたか、ユーザ自身のひらめきによるものなのかを明確に区別することは困難であり、システムの有用性を立証することは難しい。

しかしながら、何らかの方法で評価を模索していくことが重要であり、十分ではないにせよ本システムを実際に被験者に使用してもらいそれを分析することで評価とした。

評価方法は文献 [15] に乗っ取り新たに機能レベルの評価を加えた下記の 3 点から行なった。

- 要素技術レベル: システムの性能に関する評価
- 利用レベル: システムに対しユーザがどのような機能をどのように使ったかを測定し評価
- 思考レベル: システムがユーザの思考活動をどのように支援できるかを評価

初めの 2 つのレベルの評価は定量的な評価が可能である。最後の思考レベルでの評価は定量的な評価は困難である。例えば単純にアイデアの数をかぞえたり、時間あたりのアイデア生成効率を計算してもシステムの有用性を示すことにはつながらない。何故ならばアイデアの生成能力は人によって異なるし、テーマによっても大きく変化するからである。さらにどのようにアイデア数をカウントするかも不明確である。よって、思考レベルの評価は定性的な評価とした。

## 4.1 要素技術レベルでの評価

ここでは本システムの性能を評価する。本システムの利用に大きく影響する。連想辞書のディスクに占める容量と検索時間を測定した。

まず、今回生成した連想辞書テーブルのファイル容量を表 4.1 に示す。ここで算出結果とあるのは検索キーワードが一つの場合検索を高速化するためにあらかじめ全てのキーワードに対して結果を算出したものである。これは、次に述べる検索の速度に貢献するだけなので、ディスク容量を減らしたければ、無理に作成する必要はない。テキスト集合の 1 から 4 まではネットニュース、5,6 は修士論文の概要である。

表 4.1: 連想辞書のテーブルの容量

| テキスト集合 | テキスト数 | テキスト間 | キーワード間 | キーワード → テキスト | 算出結果   |
|--------|-------|-------|--------|--------------|--------|
| 1      | 734   | 1051K | 1169K  | 93K          | 31891K |
| 2      | 759   | 1000K | 1742K  | 117K         | 46971K |
| 3      | 808   | 1129K | 3378K  | 128K         | 58251K |
| 4      | 722   | 2241K | 1215K  | 83K          | 25749K |
| 5      | 74    | 61K   | 1578K  | 65K          | 38442K |
| 6      | 74    | 62K   | 3105K  | 85K          | 60017K |

本システムはリアルタイムでの利用を前提としているが、連想辞書は前もって作成しておくのでこの部分の実行速度は無視する。<sup>1</sup>したがって、リアルタイムでの利用を行なう部分に関してだけ実行速度を考える。空間表示の部分の実行速度はオブジェクトの数が 50 から 80 ぐらいまでは十分に高速なので、システムの実行速度でボトルネックとなるのは空間表示を行なうために前述した連想辞書のテーブルを検索し新たなテーブルを作成する部分である。この実行速度を測定したのが図 4.1 である。図 4.1 の一番左側の項目は 1 キーワードで検索した結果をあらかじめファイルとしてもっていた場合の速度である。計算機として、SUN の Sparc Station5 を用いた。

<sup>1</sup>何時間というレベルの時間を要す

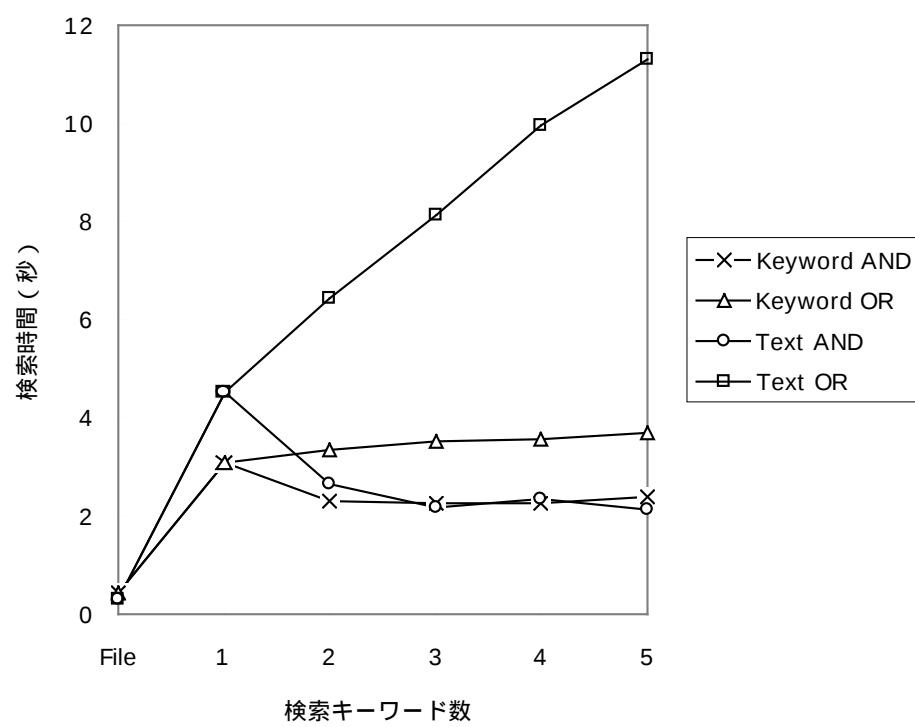


図 4.1: 検索速度

#### 4.1.1 要素技術レベルの評価に対する考察

一番検索要求が高いと思われる1キーワードによる検索に対してあらかじめ結果をファイルとして持っていたのはかなり効果があった。ただし、何千に及ぶキーワードに対してこれを用意するのは容量的にそれなりにディスクを浪費することになり、さらに作成するのに数時間もかかってしまう。また、テキストを新たに追加したときに再構成を余儀なくされる。

また、ネットニュースのテキストを用いたものと修士論文を用いたもので、テキスト数が少ないのに、キーワード間のテーブルが大きいのは個々のテキストの分量が大きいからである。

### 4.2 被験者を用いた評価実験について

機能レベル、思考レベルでの評価を行なうため被験者を用いた評価実験を行なった。ここでは被験者を用いた評価実験の概要について述べる。

被験者は研究室の学生4人を対象とした。各自に自由にテーマを設定してもらい、アイデアを生成してもらったのちアンケートに答えてもらうという形式で行った。各自興味のもてるようにいくつかテキスト集合を用意した。

- テーマ1:あたらしい健康法とは  
fj.life.health 記事数 734
- テーマ2:こんな旅行をしたい  
fj.travel.japan 記事数 759
- テーマ3:こんな旅行をしたい  
fj.travel.world 記事数 808
- テーマ4:JAIST のネットワークの改善点は？  
frontier.sys.admin(misc,queustion) 記事数 722
- テーマ5:この研究室でこんな研究がしたい  
1997 JAIST 修士論文日本語概要 文章数 74  
1997 JAIST 修士論文英語概要 文章数 74

時間は無制限にし、数日間思い付いたときにシステムを使ってもらった。これは、アイデアというものを短い一定時間の間に生ずるように強制することは序論で主張したアイデアの生成過程という観点と合い反するからである。

表 4.2: 実験のテーマと形態

| 実験番号 | 被験者  | テーマ | 形態      |
|------|------|-----|---------|
| A-1  | A さん | 2   | 空間表示    |
| A-1  |      | 5   | 空間表示    |
| B-1  | B さん | 4   | 空間表示    |
| B-2  |      | 2   | リスト表示のみ |
| C-1  | C さん | 2   | 空間表示    |
| C-2  |      | 3   | リスト表示のみ |
| C-3  |      | 5   | 空間表示    |
| D-1  | D さん | 2   | 空間表示    |
| D-2  |      | 5   | 空間表示    |

実際に各被験者がとりくんだ実験内容を表 4.2 に示す。

表 4.2 の形態の欄で、リスト表示のみというのはキーワード入力ウィンドウのみで実験を行なったものである。

### 4.3 利用レベルでの評価

実験においてはユーザがシステムに対してどのような操作を行なったか分かるように、利用履歴をとった。この利用履歴からいくつかの結果を得た。

まず、検索の仕方から見て行く、検索の種類を図 4.2 に示した。また、ユーザは空間配置を行なう過程を見るモードといきなり空間配置の結果を見せるモードを選択できる。ここで、どちらを選択したかを図に示した。

次にユーザがどれくらい空間配置に対して介入を行なっているかをオブジェクトの移動機能 (Move) と削除機能 (Delete) に関して利用状況を調べたのが図 4.4 である。

#### 4.3.1 利用レベルの評価に対する考察

検索の種類に関して、図 4.2 リスト形式における検索が多いのは操作履歴を細かく見るとリスト形式で検索してこれぞと思うものがあったら空間表示形式の検索に移行しているからであることがわかった。またキーワード空間表示検索よりテキストの空間表示検索が多いのはキーワードで

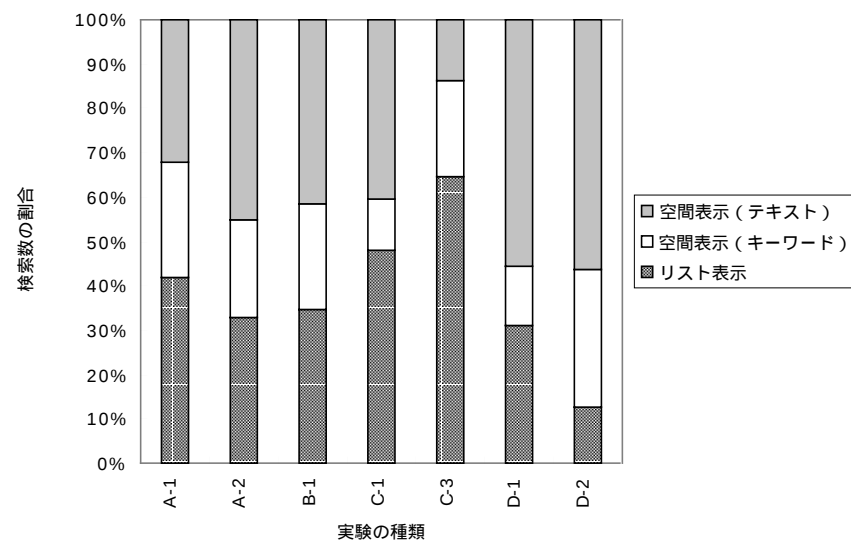


図 4.2: 検索の種類

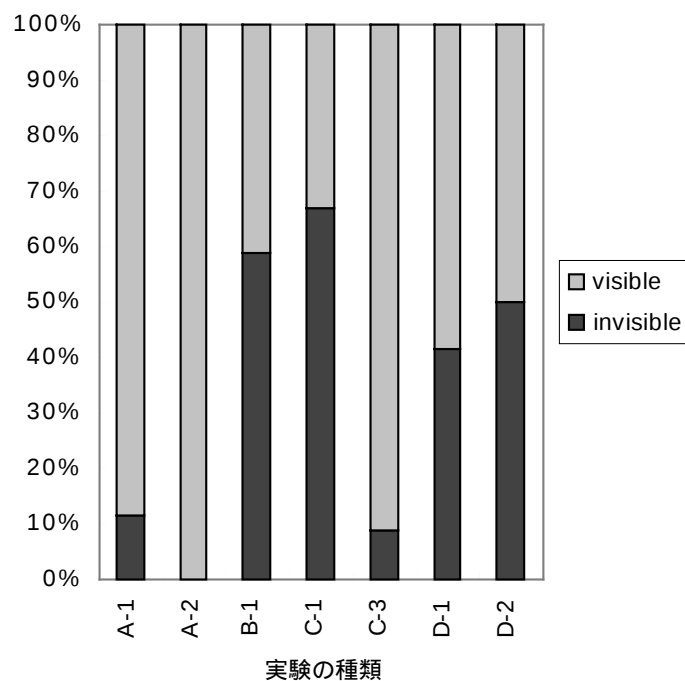


図 4.3: 表示過程における利用状況

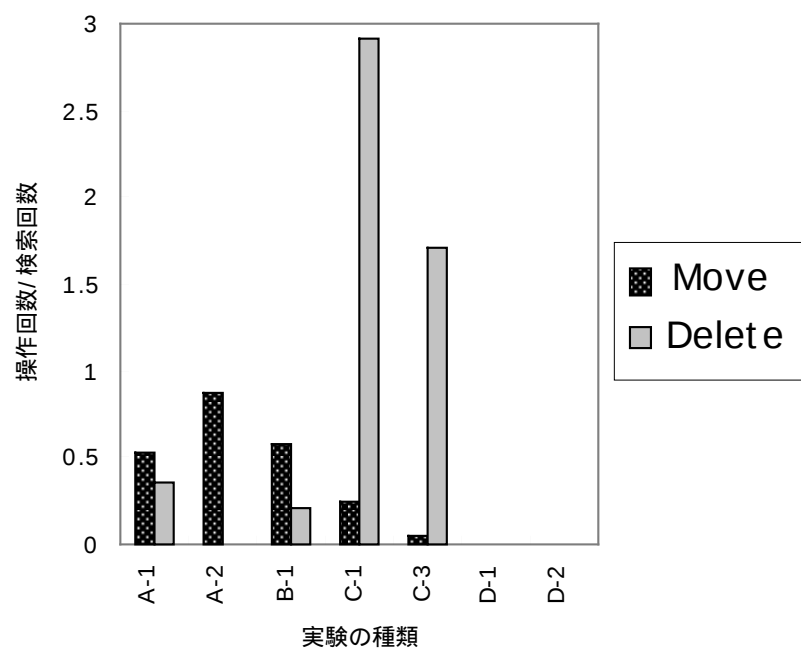


図 4.4: Move Delete 機能の利用状況

空間配置したのち配置されたキーワードそれぞれに対してテキストの空間表示検索を行なっているからであることがわかった。このことから、何回かリスト検索をして関連キーワード検索、何回か関連テキスト検索という流れがわかる。

空間配置をする際、ユーザがどのような表示の仕方を選択しているかについてを調べたのが図4.3である。この図から多くのユーザは空間配置を行なう過程を表示したがる傾向があることがわかった。いきなり結果を表示したほうが早いにもかかわらず、ユーザは空間配置をおこなう過程を見たがっていることがわかった。

システムとのインターアクション機能による効果は図から、個人差が大きいことがわかる。しかしながら、積極的に利用している被験者もあり、これらの機能が有効に働いていることがわかった。

## 4.4 思考レベルの評価

### 4.4.1 アンケート

アンケートは5段階で答えてもらった。

表 4.3: 実験の背景

|   | 質問項目                    | 最高点 | 最低点 | 平均点  |
|---|-------------------------|-----|-----|------|
| 1 | テーマの設定に関して満足していますか？     | 5   | 2   | 3.75 |
| 2 | テキストの内容は適切でしたか？(news)   | 4   | 3   | 3.5  |
| 2 | テキストの内容は適切でしたか？(thesis) | 5   | 3   | 4    |

表 4.4: キーワード抽出、類似度の生成等に関して

|   | 質問項目                | 最高点 | 最低点 | 平均点  |
|---|---------------------|-----|-----|------|
| 1 | キーワードの抽出は適切でしたか？    | 4   | 3   | 4    |
| 2 | 関連テキスト検索の結果は適切でしたか？ | 5   | 2   | 3.75 |
| 3 | 関連テキスト検索の結果は適切でしたか？ | 5   | 3   | 4    |

表 4.5: スプリングモデルによる表示に関して

|   | 質問項目                                   | 最高点 | 最低点 | 平均点  |
|---|----------------------------------------|-----|-----|------|
| 1 | 検索スピードは適切でしたか？                         | 5   | 4   | 4.25 |
| 2 | 表示スピードは適切でしたか？                         | 4   | 3   | 3.5  |
| 4 | オブジェクトの削除機能は役に立ちましたか？                  | 5   | 2   | 3.5  |
| 5 | オブジェクトの移動機能は役に立ちましたか？                  | 5   | 2   | 3.5  |
| 6 | オブジェクトの色分けは役に立ちましたか？                   | 4   | 3   | 3.3  |
| 3 | 表示はみやすかったと思いますか？                       | 4   | 5   | 3    |
| 6 | いきなり結果を表示されるより、配置される過程が見えるほうが良いと思いましたが | 5   | 3   | 4    |
| 7 | 関連キーワードの表示は気付きにくいような特徴を提示できたと思いますか？    | 5   | 4   | 4.25 |
| 8 | 関連テキストの表示は気付きにくいような特徴を提示できたと思いますか？     | 4   | 2   | 3    |
| 9 | 空間表示はリスト表示に比べて優れていると思いましたが？            | 5   | 4   | 4.25 |

### 空間表示とリスト表示の違い

- 空間表示の優れている点
  - キーワードで次々にたどって行きやすい
  - 全体を把握しやすい
  - ばねで表示するといきなり提示されるという不快感が和らぐ
- 空間表示の劣っている点
  - 対象となるテキストの数が少ないときは空間配置するより直接見に行ったほうがいい
  - 重なりが多いと見にくい

### 足りないと感じた機能

- 読んだ記事の履歴に応じて色をかえるなどして欲しい
- メールアドレスを取り除く機能が欲しい
- 関連テキストが完全に一致するキーワード同士をまとめて欲しい
- 表示したオブジェクトが画面の外にはみでないようにして欲しい
- パラメータの調節を自動的にして欲しい
- 記事の前後関係がわかりにくい

表 4.6: 総合評価

|   | 質問項目                      | 最高点 | 最低点 | 平均点 |
|---|---------------------------|-----|-----|-----|
| 1 | 全体的に操作は簡単でしたか？            | 5   | 4   | 4.5 |
| 2 | このシステムはアイデア生成に役立つと思いましたか？ | 4   | 4   | 4   |

#### 4.4.2 被験者のアイデア生成までの流れの観察

かならずしもアイデアと言えるものばかりではないが被験者がアイデアとして生成したものを列挙した。

## 4.4.2.1 Aさんの事例

- A-1
  - － 温泉 → 無料 → 露天風呂 → 秘湯 → 鯖江  
海辺  
鯖江の温泉にいきたい
  - － 割引 → 京都 → 変更  
乗車変更で旅行
  - － 弁当 → 発熱サリン
  - － 発熱弁当があるらしいので食べたい
- A-2
  - － 自律ロボットに期待することは並列作業、協調作業、自己組織化である
  - － 関係データベースの並列化には大別すると次の3つの方法がかんがえられる脱出した複数のトランザクションを並列に処理トランザクション間の一つのSQL命令を複数に  
関係代数演算に分割し演算用でパイプライン的に処理する
  - － 自律移動ロボットの視覚としてアクティブビジョンの基本的な考え方は「見やすいように見る」「必要なところだけ見る」「必要な程度に見る」「動くために見る」である

## 4.4.2.2 Bさんの事例

- B-1
  - － JAIST のネットワークにはいるのに fep がつかえる
  - － fep の運用にセキュリティ的な問題がありなんとかしたい
  - － ネットワークのことなら Y さんや K さんが詳しい
  - － Jaist の firewall についてもっと知りたい
  - － backdoor について調べて見たい
- B-2
  - － ハウステンボスに行くなら佐世保にもよろう
  - － アトラクションのノアの劇場がおもしろそう

- 土産はチーズケーキとクッキーがいい
- レンタカーもいいな
- 切符は自由周遊券が得だ

#### 4.4.2.3 Cさんの事例

- C-1
  - 国外に出るのは手続きが面倒だ
  - 冬にアラスカに行きたい、寒そうだけどオーロラがみれるし一日中夜というのを体験できる
  - 中東は紛争が多いので仕事でもいきたくないな
  - オーストラリアに行きたい、観光ツアーじゃなくて一般的な観光コースじゃないところに行きたい
  - ホノルルは日本人だらけでつままないとおもってたけど「気になる木」だけみたい
- C-2
  - スキーをする人って以外に多そう
  - スキーにさそわれるのは嫌だな
  - 車で旅行したらつかれるだろうな
- C-3
  - グループウェアの研究は面白そうだけど、人を使う場合かなり面倒そう
  - システム主体の研究は生理的にうけつけないそう
  - パターン系の研究は意味不明な式が並んでいてつかれそう

#### 4.4.2.4 Dさんの事例

- D-1
  - 忍者寺へ行きたい
  - 鰯組に飲みに行きたい
  - ローザンヌのケーキが食べたい

- D-2
  - 遺伝的アルゴリズムの研究においてアルゴリズムに頑健性をもたせる必要がある
  - 遺伝的アルゴリズムの研究において冗長性をもたせることも必要
  - 並列計算機を使ってプロセッサ間の協調発生もおもしろいかも
  - 並列計算機はプロセッサ郡を集団として考える。

## 4.5 全体の評価と考察

これらの結果を空間配置の効果、インタラクションの効果、テキスト集合の質の違いからくる影響の三つの観点から考察をし、最後に本システムの問題点を洗い出す。

### 4.5.1 空間表示の効果

Aさんの事例でAさんは「温泉」の関連キーワードから「海辺」「無料」というキーワードに注目していた。「温泉」というキーワードの関連キーワードからは必ずしも「海辺」「無料」といったキーワードは上位にランキングされるわけではない。しかしながら、これらのキーワードは空間配置をすると「温泉」の近辺に配置される。これは、リスト表示では温泉というキーワードとの類似関係しか見ていないので全体様子を大まかに見るができないからである。このような特徴がキーワードで次々にたどって行きやすいとか全体を把握しやすいといった評価につながった。

また、空間配置をする過程を見せることはシステムによる押しつけがましさを緩和することにより、ユーザの心理的な不快感を減らしているようである。

### 4.5.2 インタラクションの効果

空間配置への削除、移動などのシステムとのインターアクション機能による効果は機能レベルの評価でも思考レベルでの評価でも、個人差が大きいことがわかる。しかし、アンケートと図4.4の利用状況はほぼリンクしており、これらの機能にたいして否定的な評価を下した被験者は図4.4における被験者Dと一致し、これらの機能を全く使用していない。逆にいうと、他の被験者は積極的に利用してかつ役にたったという評価を下しているので、これらの機能が有効に機能していることがわかった。

#### 4.5.3 テキスト 集合の質の違いからくる影響

今回、ネットニュースのテキストと修士論文の概要という全く質の違うテキスト集合を用いた。全体的にネットニュース (得に fj) に関しては記事の内容的な質が悪いという指摘が多かった。またテキストにあまり適切なラベルがつけられず、Subject をラベルとしたために、いちいち内容を閲覧しないと、何が書いてあるか分からないという問題がおこった。

逆に修論概要ではタイトルと研究室名をラベルとし、テキストの内容もかなり洗練されているので、以上のような問題は緩和された。

アイデアの元資料となるテキストはなるべく洗練された、テキストを用いることが需要であろう。しかしながら、実際問題としては洗練されたテキストは往々としてフリーで公開されないという問題もある。

#### 4.5.4 アイデア生成への支援

最後に、本システムによってアイデア生成に役だったか考察する。

被験者の生成したアイデアからも分かるように、アイデアの質<sup>2</sup>は、全体的に見てその人のテーマに対するモチベーションの高さによるところが大きそうである。旅行に行きたくない人にどこに行きたいかといわれても、否定的なアイデアしかでてこない。逆にある程度モチベーションが高ければ、質のいいアイデアがでてきたようである。<sup>3</sup>

しかしながらそのテーマに関して関心がなかった場合でも、本システムからなんらかの関係を見出しそれをアイデアにつなげている。本システムの支援は枠組-パラダイムレベルによる支援であり、アイデアにつながる観点を示すことが目的なので、そういった意味では効果をあげている。以上のことがアンケートの総合評価でアイデア生成に役立つという評価につながったようである。

---

<sup>2</sup> 「アイデアの質ってどうやってはかるの」という大きな問題はあるがここでは筆者の主観である

<sup>3</sup> 中にはテキストにあった URL から WWW のページまでみにいった被験者もいた

## 第 5 章

# 本システムの応用的利用

本章では、本システムの応用的利用として WWW 上への実装と従来の発想支援システムへの適用の二つの点から見ていく。

### 5.1 World Wide Web 上での利用

3 章で作成したシステムは UNIX の X 上での利用であった。パフォーマンス的 (特に表示速度) にはこのような利用のほうが有効ではあるが、WWW 上に載せることにより、以下の利点が得られる。

- ソフトのインストールが不要
- プラットフォームが多様 (ウインドウズ、マック等)
- ユーザによるメンテナンスの必要がない

このような、利点からどこでもだれでもすぐに利用できる。とにかく、システムを使ってもらわなくては意味がないので、このような利用は大変効果的である。

二次元空間表示の部分は未完成ではあるが、WWW 上に CGI を用いて簡単な関連辞書検索を試作してみた。(図 5.1)

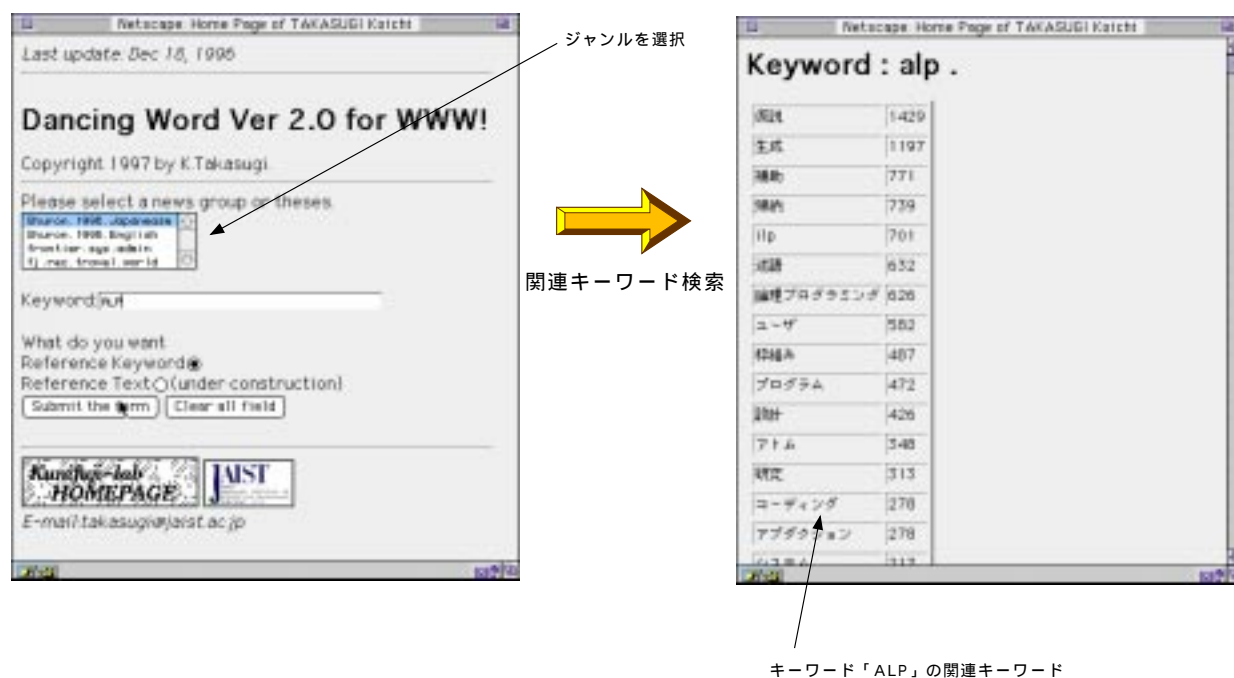


図 5.1: システムの概要

## 5.2 従来の発想支援ツールへの適用

### 5.2.1 ブレーンストーミングでの利用 (発散的思考支援)

ブレーンストーミングは集団で議論を通して刺激しあい、アイデアを出して行くものである。ブレーンストーミングではアイデアを出しやすくするための4原則(批判厳禁、自由奔放、質より量、結合発展)がある。このブレーンストーミングでだされたアイデアは本システムのようなもので次々と表示することができる。これによりいままでだされたアイデアを分類した形で表示するとともに、空白の部分に新たなアイデアを思い付くといった効果が期待される。また、これをそのまま次節で述べる収束的発想支援であるKJ法の段階に容易に移行できる。

### 5.2.2 KJエディタの前段階での利用 (収束的思考支援)

収束的発想支援としてKJ法がある。KJ法は収集されたデータについて主題ごとにラベルを作り、これを何段階にもグループ化することにより問題解決をはかっていく手法である。このKJ法をコンピュータ上に実現するKJエディタとして、D-ABDACTORが知られている[16]。

ユーザは個々のカードの中身を吟味しながら、グループ化していくわけだが、このとき、カー

ドを類似度別に空間配置してから行なえば、ユーザは似たキーワードをもつカードを容易に探し出せグループ化しやすくなるのではないか。現在 D-ABDATOR では、カードの初期配置は入力順であり、カードをユーザがマウスでドラッグしていくという面倒な操作をしいられている。また、いくつかの異なった配置からグループ化していけば、違った視点でカードを見ることができ、異なった図解を作り上げることができるだろう。

## 第 6 章

# 関連研究との比較

### 6.1 スプリングモデルを採用した空間表示ツールとの比較

文献 [17] ではスプリングモデルを用いた思考支援システムについて言及している。ここではユーザの概念空間を空間表示するのにスプリングモデルの各パラメータがユーザの概念のどのレベルにあるのかについて言及しているが、これらのパラメータを設定するのはシステムではなく、ユーザがみずから与えるものとなっている。これはユーザに過大な負担を強いる事になり、それにみあうだけの効果が期待できるとは考えにくい。というのは、ユーザがパラメータを与えることによりシステムに空間配置をさせるぐらいなら、ユーザが思った所に配置をしていけば良いからである。

文献 [18] ではスプリングモデルを用いてネットニュースの空間表示している。さらに、インタラクティブな操作環境を提供しており、配置されたものを移動するなど操作面でも本研究と似た点が多い。しかしながら、このシステムは分類表示を主眼においており、大量のニュースを一度に表示しようとしているため、類似度の計算を簡略化している。(類似度が高い、低いの二種類) ネットニュースのビューアとしてはそれでもよいが思考支援システムとしてはもっとこの類似関係をきちんと計算し反映する必要がある。なぜなら、アイデアの基本は類似関係の発見からくるものであると筆者は考えているからである。

### 6.2 統計的手法を採用した空間表示ツールとの比較

多次元尺度法や双対尺度構成法などの統計的手法を用いた思考支援のための空間表示ツールも多く研究されている [19] [20] [21] 。

これらのシステムでは空間配置された個々のオブジェクトに対して後から追加、削除などを行

なうと再計算がおこり、空間配置が大きく変化する。つまり、本システムのように徐々に新しい空間配置に変化するようなことはない。例え、新しい空間配置へ至る軌跡をシステムが補間することによって表示しても、あくまでも補間であり個々のオブジェクトを見失わないという効果はあるかもしれないが、その動きにはなんら意味はないであろう。また、ユーザが空間配置されたオブジェクトに対して移動させるといった操作はサポートしていない。また、類似度を直接空間配置に反映していないため人間の直観に訴えにくいのではないか<sup>1</sup>。

さらに、思考支援という観点からするとかならずしも数学的に最適な空間配置をする必要はないであろう。それよりも、本システムのように準最適解のようなものを数多く見せるほうが類似関係の発見や異なる視点の獲得につながると考える。そもそも、ここで議論している類似関係は主にキーワードベクトルによるのであり、これが絶対的な類似度であるとはいえない。

---

<sup>1</sup>もちろん、統計的手法が頭の中で思い描けるユーザは別である

## 第 7 章

# 結論

本章では本研究の成果をまとめるとともに今後の課題について述べる。

### 7.1 本研究の成果

本研究の成果をは以下である。

- 関連キーワード、関連テキスト検索のための連想辞書を生成した。
- スプリングモデルを用いてキーワード間、テキスト間の類似度を直接、二次元空間に反映させることができた。
- 連想辞書検索の高速化によりリアルタイムでの利用を実現した。
- スプリングモデルに対し独自の拡張をすることにより従来研究よりもシステムとのインタラクションを強化した思考支援システムを構築した。
- 評価実験を通して、ユーザによる本システムの利用状況を観察し思考支援としての本システムの役割を明らかにした。
- 評価実験を通して、本研究で提案したシステムとのインタラクション機能が有効に機能していることを確かめた。
- 評価実験を通して、本システムによりアイデアの断片生成の支援に効果があることを確認した。

## 7.2 今後の課題

統計的に処理をするほど十分な評価実験が行なえなかったので、評価実験を繰り返すとともに、問題点をシステムに反映させていく必要があるであろう。

また、5章で述べたように既存のさまざまな思考支援システムとの統合が考えられ、今後、模索していく必要がある。

# 謝 辞

本研究を進めるにあたり終始御指導頂きました國藤進教授に心から深謝致します。また、数多くの有益な助言を頂いた東条敏助教授、Thanaruk Theeramunkong 博士に厚く御礼申し上げます。

さらに、日頃御討論、御協力頂いた國藤研究室の方々に感謝の意を表します。特に同期生の皆さんには研究について多くの議論を行いかつ有益な助言を得ることができました。また、博士課程の皆さんには私の研究に関し鋭い指摘を多く頂きました。さらに、M1の皆さんには何時間にも及ぶ評価実験に協力してました。もちろん、私生活においても大変お世話になり、励みになったことはいうまでもありません。

また、本研究のキーワード抽出に関してはNTTの学外実習で行なわれた経験に基づいてています。実習を快く受け入れて下さったNTTマルチメディアネットワーク研究所の木下研作部長をはじめ、実習の際お世話になり、研究に関してさまざまな意見をいただきましたNTT研究開発推進本部の杉田恵三部長、NTTマルチメディアネットワーク研究所の爰川知宏氏に深く感謝致します。

そして、本システムはperl, CHASENをはじめとした様々なフリーソフトウェアに支えられています。これらのソフトウェアをフリーという形で提供して下さった開発者の皆様に感謝します。

最後に大学院生活をかげながら支えてくれた家族、友人に感謝したいと思います。

1997年2月14日

高杉 耕一

## 参考文献

- [1] James Webb Young. *A Technique for Producing Ideas*. NTC Business Books, 1975.
- [2] 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.
- [3] 折原良平. 発散的思考支援ツールの研究開発動向. 人工知能学会誌, Vol. 8, No. 5, pp. 160–567, 1993.
- [4] Young L.F. *Idea Processing Support*. Wm. C. Brown Publishers, 1988.
- [5] Young L.F. The metaphor machine a database method for creativity support. *Decision Support Systems*, Vol. 3, No. 4, pp. 309–317, 1987.
- [6] 渡部勇. 発散的思考支援システム「keyword associator」第二版. 第17回システム工学部会研究会「発想支援技術」資料, pp. 1–23. 計測自動処理学会, 1994.
- [7] 爰川和宏, 高杉耕一, 杉田恵三, 國藤進. 組織情報共有におけるキーワード抽出法の検討. オフィスシステム研究会資料, pp. 51–56. 電気情報通信学会, 1996.
- [8] W.B. and Baeza-Yates.R Frakes. *Information retrieval: data structure and algorithms*. Prentice Hall, 1992.
- [9] 松本裕治, 今一修, 山下達雄, 北内啓, 今村友明. 日本語形態素解析システム『茶釜』version 1.0b5, 奈良先端科学技術大学院大学松本研究室, 1996.
- [10] EDR 電子化辞書, 日本電子化辞書研究所.
- [11] 杉山公造. グラフの自動描画法とその応用. 計測自動制御学会, 1993.
- [12] Kamada T. and S. Kawai. Algorithm for drawing general undirected graphs. *Information Processing Letters*, Vol. 31, No. 1, pp. 7–15, 1989.

- [13] 堀浩一. 発想支援システムの効果を議論するための一仮説. 人工知能学会誌, Vol. 35, No. 10, pp. 1998–2008, 1994.
- [14] 三末和男. 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. 第 3 回『発想支援ツール』シンポジウム講演論文集. 計測自動処理学会, 1995.
- [15] 女部田武史, 國藤進. 複数の k<sub>j</sub> 法図解の差異や共通部を可視化するシステムの実現について. 第 18 回システム工学部会研究会「発想支援技術」資料, pp. 21–28, 1995.
- [16] 三末和男. 図的発想支援システム d-abductor の開発について. 情報処理学会学会誌, Vol. 35, No. 9, pp. 1739–1749, 1994.
- [17] 田中淳. 記号間の力学に基づく概念マップ生成システム SPRINGS. 情報処理学会誌, Vol. 33, No. 4, pp. 495–499, 1992.
- [18] 舘村純一. インタラクティブ視覚化による文献集合からの情報獲得支援. 日本ソフトウェア科学会第 13 回大会論文集, pp. 365–368. ソフトウェア科学会, 1996.
- [19] 角康之. テキストオブジェクトを空間配置することによる思考支援システム. Master's thesis, 東京大学, 1992.
- [20] 杉本雅則. 複数他者の視点を可視化するシステムとその知的活動支援への応用に間する研究. PhD thesis, 東京大学, 1994.
- [21] 杉本雅則. 主観的な概念の獲得、表現の支援と設計問題への応用に関する研究. Master's thesis, 東京大学, 1992.
- [22] 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. 第 7 回 AI シンポジウム予稿 (SIG-J), pp. 34–39. 人工知能学会, 1996.
- [23] 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. システム/情報合同シンポジウム'96 講演論文集, pp. 149–152. 計測自動処理学会, 1996.