

既存の雑多な情報を統合するための連想構造の提案

前田 晴美・梶谷 和人・西田 豊明

harumi-m@is.aist-nara.ac.jp

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

既存の雑多で構造の不均質な情報を統合する手法を提案する。基本となるアイデアとして、雑多な情報をゆるやかに関連づけるための連想構造とよぶデータ構造を用いる。連想構造は単純な表現であり、生データから容易に生成できることと、人間が直観的に理解しやすいことが特徴である。我々は、連想構造を用いて既存の情報源から情報を収集し、整理する過程を支援するシステム CM-2 を試作した。CM-2 では、(a) 既存の雑多で構造の不均質な情報源から情報を取り込み、CM-2 の連想構造を生成する情報キャプチャ機構、(b) 与えられた視点から情報の切り出しと構造化を行う知的情報統合機構、(c) キーワードによる連想的な検索を可能にする連想検索機構を実現した。CM-2 の有効性を実験によって確かめた。

A Proposal of Associative Structures for Reorganizing Diverse Information Obtained from Existing Information Sources

Harumi MAEDA, Kazuto KOUJITANI and Toyoaki NISHIDA
Graduate School of Information Science,
Nara Institute of Science and Technology
8916-5, Takayama, Ikoma, Nara 630-01 Japan

In this paper, we present a method to reorganize diverse information obtained from existing information sources. The basic idea is the use of a plain information representation called *associative structures* which enable to link information chunks using associative indexing. It is easy to generate associative structures from raw data and they are comprehensive to humans intuitively. We have developed a system called CM-2 which helps users to gather and reorganize information from existing information sources. We describe the system's three major facilities; (a) an *information capture facility* which allows users to gather information from heterogeneous information sources and generate CM-2 associative structures, (b) an *information integration facility* which helps users to reorganize information from the user's point of view, and an (c) *associative retrieval facility* which gives users access to information through associative indexing mechanisms using keywords. We verify our approach by analyzing results of experiments.

1 はじめに

研究活動に代表される創造的思考活動には、過去に作成した論文やOHPや研究メモ、研究動向を調査するための書籍・雑誌、広く知識を収集するための辞書、オンラインデータベースやWWW(World Wide Web)などの、雑多で構造の不均質な情報が必要である。文書として書き表されていない頭の中の断片的なアイデアや常識も重要である。このような情報を統合的に扱うためには、既存のデータモデルや知識表現などを用いて情報ベースを構築することは人間の負荷が高い。

本論文では既存の雑多で構造の不均質な情報を統合する手法を提案する。基本となるアイデアとして、雑多な情報をゆるやかに関連づけるための連想構造とよぶデータ構造を用いる。連想構造は単純な表現であり、生データから容易に生成できること、人間が直観的に理解しやすいことが特徴である。採用する方法はヒューリスティックなものであり、あらゆるケースに対応できるものではないが、予備実験の結果かなりの有効性が期待されたので、定式化、実現によって有効性を実験的に評価することとした。我々は、連想構造を用いて既存の情報源から情報を収集し、整理する過程を支援するシステムCM-2を試作した。CM-2では、

- 既存の雑多で構造の不均質な情報源から情報を取り込み、CM-2の連想構造を生成する情報キャプチャ機構
- 与えられた視点から情報の切り出しと構造化を行う知的情報統合機構
- キーワードによる連想的な検索を可能にする連想検索機構

を実現した。

本論文は以下のとおり構成される。2節では連想構造とCM-2の概要について述べる。3, 4, 5節ではCM-2の3つの手法について説明する。6節では実験結果を示し、有効性について議論する。7節では関連研究と比較する。

2 連想構造とCM-2

2.1 連想構造

連想構造はkeyと呼ぶいくつかのユニットとvalueと呼ぶいくつかのユニットの間に定義され、「keyが与えられるとvalueが想起される」というゆるやかな関連を表す。keyとvalueの間の連想関係は厳密に定義されるのではなく、多分に主観的なものであることを許している。これは既存の情報の雑多性、多様性に対応することを狙っている。

ユニットとは、情報を蓄積する情報ベースの基本構成要素である。ユニットはCM-2の外部のテキストファイルやイメージファイルを表す外部参照データユニットと、CM-2において情報を関連付けるための内部的な対象である概念ユニットに大別される。

本論文ではCM-2における連想構造を図1のように記述する。図1(a)はkey「奈良」が与えられるとvalue「奈良公園」「大仏」「鹿」が想起されることを示す。図1(b)は

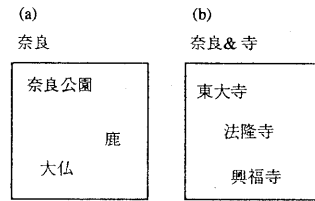


図 1: CM-2 における連想構造

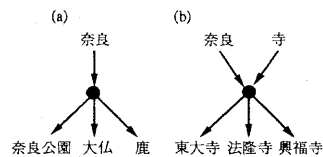


図 2: CM-2 における連想構造 (ドット表記)

key「奈良」と「寺」からvalue「東大寺」「法隆寺」「興福寺」が想起されることを示す。

図2は、図1の連想構造をドット表記と呼ぶ表現を用いて表したものである。ドット表記では、1つの連想構造を点を用いて表現し、keyとvalueの関係を示す矢印を用いて表現する。

連想構造には、特別な形態として、IS-A構造、辞書構造を定義できる。IS-A構造においては、クラス、サブクラスなどを区別することなく、上位概念をクラス、下位概念をインスタンスと呼ぶ。例えば、図3(a)のIS-A構造では、「東大寺」がインスタンスで「寺」がクラスである。IS-A構造は推論に使用される。辞書構造は、利用者の語彙を定義するもので、概念ユニット間の変換辞書として使用される。例えば、図3(b)の辞書構造がある場合、利用者が「奈良先端大」と入力すると「NAIST」というユニットに変換できる。

2.2 連想構造を用いた情報統合への枠組

図4に、連想構造を用いて既存の情報源の情報を統合する枠組の概要を示す。人間と機械の協調作業によって、(1) 既存の情報源から情報を抽出して連想構造を生成し、(2) 連想構造を再構成して、マップ、箇条書、表などの人間にとって理解しやすい形式で出力する。

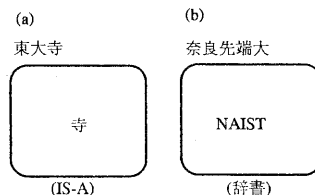


図 3: 連想構造の特別な形態

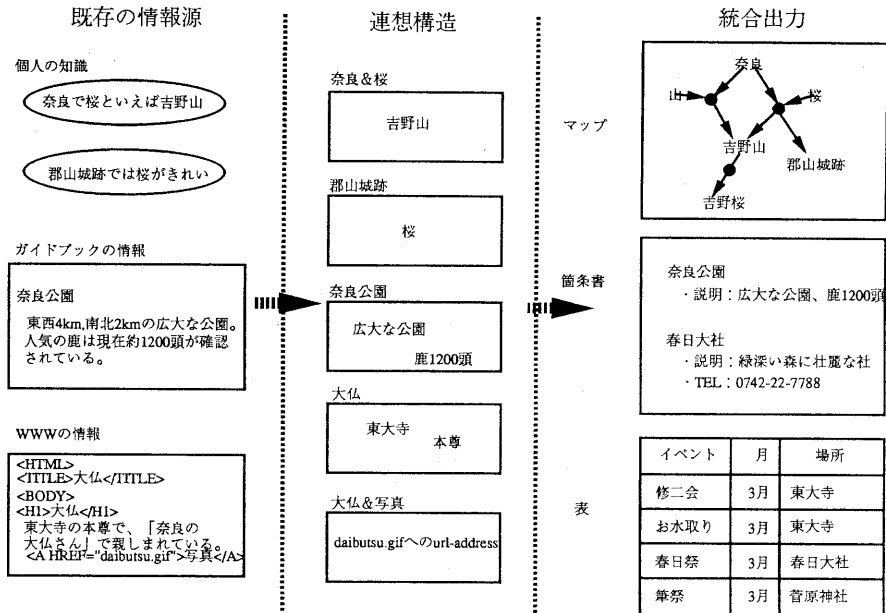


図 4: 連想構造を用いた情報統合の枠組の概要

2.3 CM-2 の概要

CM-2¹は、既存の雑多で構造の不均質な情報源から情報を収集・整理するための情報整理システムである。

CM-2の情報ベースは、内部知識ベース(概念ベース)とテキストデータやイメージデータなどがある外部データベースから構成される。利用者はワークスペースと呼ぶ空間を通してネットワークに接続された個人やグループの情報ベースを利用できる。

CM-2では、(a) 既存の雑多で構造の不均質な情報源から情報を取り込み、CM-2の連想構造を生成する情報キャプチャ機構、(b) 与えられた視点から情報の切り出しと構造化を行う知的情報統合機構、(c) キーワードによる連想的な検索を可能にする連想検索機構、の3つの機構を持つ。

2.4 CM-2 の適用例

2.4.1 例1: 人工知能情報ベース

例えば、自分の興味のある分野の研究者やプロジェクトに関連する情報を知りたいが適当なデータベースがないとする。このような時に、WWWのページを収集して必要な情報だけを得ることができると便利である。

しかし、WWW上のHTML文書は、情報の構造や表現方法が作者によって違うため、必要な情報を抽出することが難しい。例えば、ある人は、ページの最初に写真や名前を載せ、次に履歴や研究に関する情報を通常の文章として書いている。またある人は、名前、e-mailアドレス、肩書などを箇

¹“CM”は我々の長期的な研究目標である“Contextual Media(文脈メディア)”の略語である。超並列コンピュータとは無関係である。

条書で書いている。さらに、作者によって同じ概念に違う名前をつけたり、違う概念に同じ名前をつけることもある。

我々は、WWW上の人工知能研究者のホームページから、情報キャプチャ機構、知的情報統合機構を用いて人工知能情報ベースを作成した。図5に、人工知能研究者のホームページのURLを入力としてHTML文書を収集し、利用者の入力に応じて情報を整理した結果を示す。

図5(a)は、利用者がキーワード「reasoning」と、整理項目「研究者」「e-mail」「プロジェクト」「大学」を入力して、研究者毎に表形式に情報を整理してWWWブラウザに出力した結果を示す。

同様に図5(b)は、「プロジェクト」「研究者」「e-mail」「大学」を整理項目としてプロジェクト毎に箇条書形式で出力した結果を示す。

2.4.2 例2: 奈良観光情報ベース

我々は、奈良観光ガイドブックを参考にして、奈良観光情報ベースを手作業で構築した。構築にあたっては、概念の分類体系などの特別な基準を定めず、ガイドブックに出てきた文章や個人的な知識から連想構造を作成した。このようにして構築した構造化されていない情報ベースには、キーワードに基づく自然な問いにやわらかく答えることができるような連想検索機構が有効である。

図6(a)は、「大仏がある寺は？」という入力に対する活性伝播型連想検索(5節参照)の結果である。CM-2は、質問文から「大仏」「寺」の概念ユニットを抽出して、各ユニットから連想構造でつながれている「東大寺」をマップ形式で提示する。

図6(a)の状態で「東大寺」を選択する(近傍検索;5節参

(a) 研究者毎に表形式で出力

(b) プロジェクト毎に箇条書き形式で出力

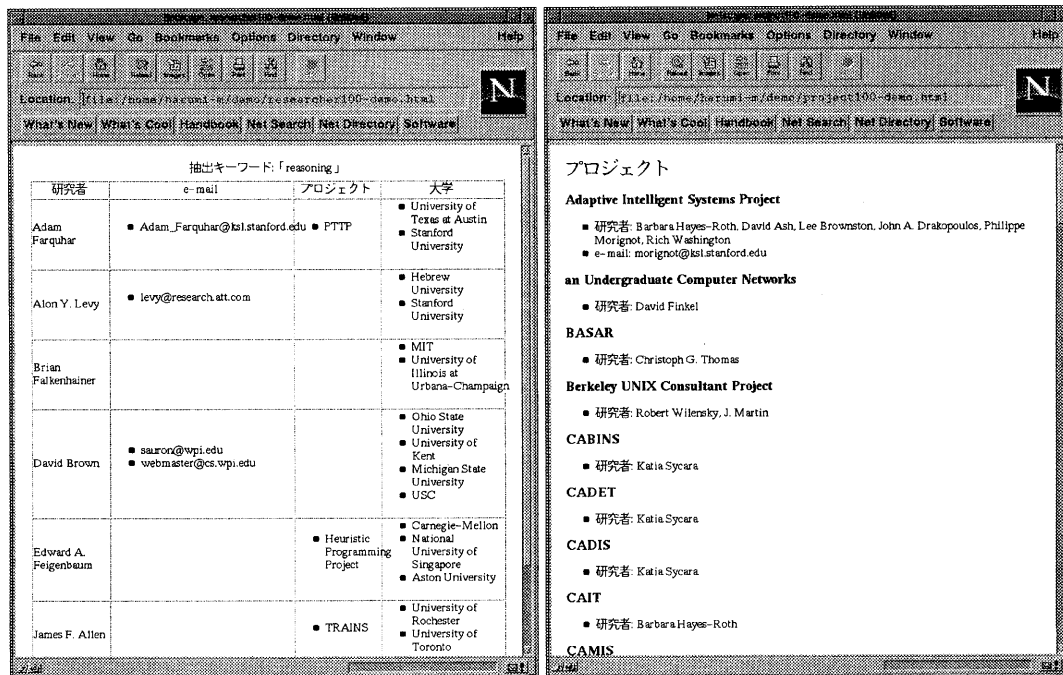


図 5: 人工知能情報ベース

(a) 活性伝播型連想検索

(b) 近傍検索

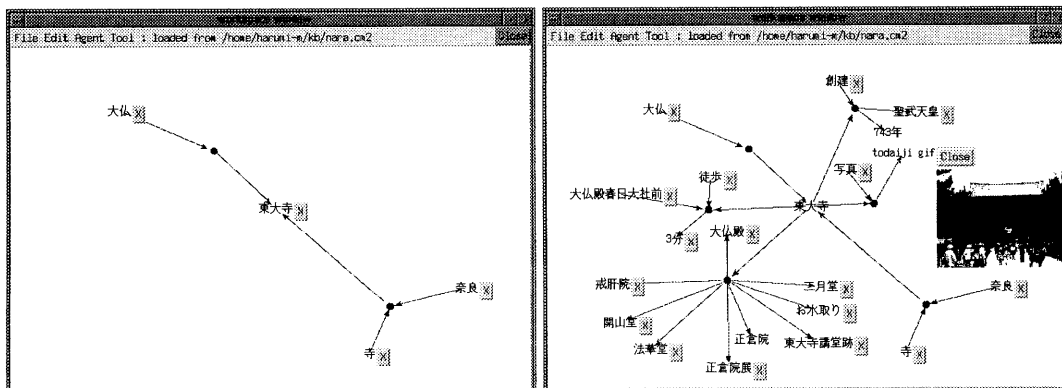


図 6: 奈良情報ベース

照) ことにより、「東大寺」から連想されるさらに詳しい情報を得ることができる(図 6(b) 参照)。例えば、「東大寺の創建が743年で聖武天皇によること」や「東大寺が大仏殿春日大社前より徒歩3分であること」などがわかる。

以下では、CM-2の3つの機構の実機能について述べる。

3 情報キャプチャ機構

情報キャプチャ機構は、既存の雑多で構造の不均質な情報源から情報を取り込み、CM-2の連想構造を生成する。

我々は、簡単なテキスト解析アルゴリズムとヒューリスティックスを用いて、WWW上のHTMLファイル、日本経済新聞全文記事データベース、Lispプログラム、UNIXファイルシステムなどを対象に、CM-2の連想構造を生成するプログラムを開発した。

CM-2 連想構造は、キーボードとマウスを用いて直接入力することもできる。

3.1 WWWからの情報キャプチャ機構

URLで指定されるHTML文書から情報を取り込み、CM-2の連想構造を生成する過程について述べる。アルゴリズムの概要は以下のとおり。

- **step 1** URL 解析による HTML 文書の取得
- **step 2** 形態素解析と HTML 構造解析を用いたユニットと連想構造の生成
 - **step 2.1** 形態素解析を用いたユニットの生成
 - **step 2.2** HTML 構造解析を用いた連想構造の生成
- **step 3** クラス判定ヒューリスティックスを用いた IS-A 構造の生成と連想構造の修正

step 2.1 においては、形態素解析(英語:Brill's Rule Based Tagger[1], 日本語:JUMAN)を用いて品詞情報を獲得し、名詞と、名詞が連続して出現する品詞群(間に前置詞や接続詞、冠詞を含んでもよい)を概念ユニットとして取り込む。

step 2.2 では、<h1> や <h2> などの見出しがある場合、また <dl>, , などのリスト表現を見出しとして用いている場合にタイトル-スコープを設定し、概念ユニットがどのタイトル-スコープに含まれているかを元に連想構造を生成する。

step 3 で用いるクラス判定ヒューリスティックスは、個別クラス判定ヒューリスティックから構成される。ユニットに対して、個別クラス判定ヒューリスティックを順番に適用し、最初の判定結果をそのユニットのクラスとする。個別クラス判定ヒューリスティックは、与えられたユニットの名前に含まれる文字列からクラスを推論する。

3.1.1 例

例えば、James Allen 氏のホームページの URL²が与えられた場合どのような処理が行われるのかを説明する(図7参照)。

²http://www.cs.rochester.edu/u/james/

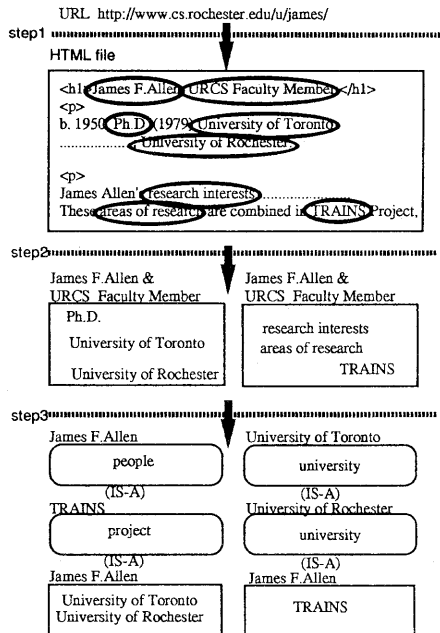


図 7: 情報キャプチャ機構のアルゴリズム動作例

URL 解析により HTML ファイル取得後、形態素解析を用いて、名詞句とそれに準じるものを概念ユニットとして抽出する。見出し表現 `<h1>...</h1>` に囲まれる範囲から生成されたユニット「James F.Allen」、 「URCS Faculty Member」を key とし、段落表現 `<p>...</p>` で囲まれる範囲から生成されたユニット「Ph.D.」、 「University of Toronto」 「University of Rochester」を value とする連想構造を生成する。

「James F.Allen」は「James」という英語名 of the 単語を含むため「people」であると推論し、IS-A 構造を生成する。「University of Rochester」「University of Toronto」は「university」という単語を含むため「university」であると推論し、IS-A 構造を生成する。

「James F.Allen」と「URCS Faculty Member」を key とする連想構造から、IS-A 構造によって連結されていない「URCS Faculty Member」と「Ph.D.」を削除して、連想構造を修正する。

4 知的情報統合機構

知的情報統合機構は、与えられた視点から情報の切り出しと構造化を行う。以下に基本的なアルゴリズムを示す。

- **step 1** ヒューリスティックスを用いたユニットと連想構造の統合
- **step 2** 利用者の項目入力による、経路探索を用いた連想構造の生成と結果の出力

表 1: 知的情報統合機構 (step 1) 主要ヒューリスティックス

同名概念統合	同名の概念ユニットを統合する
辞書参照概念統合	辞書構造を参照して概念ユニットを統合する
包含的連想生成	ある概念ユニットの名前が他の概念ユニットの名前に含まれている時、その概念ユニットを key として他の概念ユニットを value とする連想構造を生成する
文脈参照連想統合	ある連想構造と他の連想構造の key が同じ時、連想構造を統合する

step 1 では、別個に生成されたユニットと連想構造をヒューリスティックスを用いて統合する。利用者が選択可能なヒューリスティックスの中で、主要なものを表 1 に示す。

step 2 では、利用者の項目入力に従って、経路探索を用いて連想構造を生成し、結果を表や簡条書形式で出力する。概略は図 8 のとおりである。

基本的には、整理項目 1 で入力された概念を、利用者の視点として設定する。整理項目 2 以降の入力は、視点に関する属性を設定する。視点として設定されたユニットから IS-A 構造で連結されるユニット (インスタンス) に関して、属性情報を整理する。キーワードが入力されている場合には、上記のインスタンスの中から、キーワードの文字列が含まれているユニットのみを抽出し、それらに関する属性情報を整理する。

4.1 例

例えば、利用者が「reasoning」に関連のある研究者とプロジェクトについて知りたいとする。「reasoning」というキーワードと、整理項目 1 「研究者」、整理項目 2 「プロジェクト」を入力した場合の動作例を示す (図 9 参照)。

「reasoning」というキーワードを名前に含むユニット「plan reasoning」を検索する。「研究者」または「people」クラスのインスタンスであるユニットの中から、「plan reasoning」と連想構造でつながりがあるユニット「James F.Allen」を最初の key とする。項目 2 「プロジェクト」を 2 番目の key とする。「プロジェクト」または「project」クラスのインスタンスであるユニットの中から、「James F.Allen」と連想構造でつながりがあるユニット「TRAINS」を value として、連想構造を生成する。

5 連想検索機構

連想検索機構はキーワードによる連想的な検索を可能にするもので、近傍検索と活性伝播型連想検索がある。

CM-2 の基本的な情報アクセスは、指定した概念ユニットと連想構造で結ばれたユニットを検索する近傍検索である。近傍検索は連想構造の定義に基づく簡便な検索法であるが、選択された概念ユニットに多くの key や value が連結されている時に表示に時間がかかり、連想構造で直接関連付けられていない概念ユニット間の脈略を見つけにくい。

活性伝播型連想検索は、与えられたユニットを関連づける連想構造の連鎖を検索するもので、意味ネットワークにお

```

begin
整理項目 1 と同名の概念ユニット item1 から IS-A 構造で連結
されたユニット (インスタンス) を最初の key の候補集合
key1-candidates とする;
if item1 が辞書構造で連結されている
then
辞書構造で連結されたユニットから IS-A 構造で連結
されたユニット (インスタンス) を key1-candidates
に加える;
if キーワードが入力されている
then
キーワードの文字列を含む概念ユニットを探し連想をたどって
key1-candidates の中から最初の key の集合 key1s を確定
する
else key1s := key1-candidates;
for すべての key1 := key1s do
for すべての item := 2 番目以降の整理項目
do begin
key2 := item と同名の概念ユニット;
key2 から IS-A 構造で連結されたユニット (インスタンス) を
value の集合 value-candidates とする;
if key2 が辞書構造で連結されている
then
辞書構造で連結されたユニットから IS-A 構造で連結
されたユニット (インスタンス) を value-candidates
に加える;
key1 から連想をたどって value-candidates の中から value
の集合 values を確定する;
key1, key2 を key とし, values を value とする連想構造を
生成する
end;
結果をマップまたは簡条書または表形式にして HTML ファイルを
作成してブラウザに表示する
end.
    
```

図 8: 知的情報統合機構 (step 2) のアルゴリズム

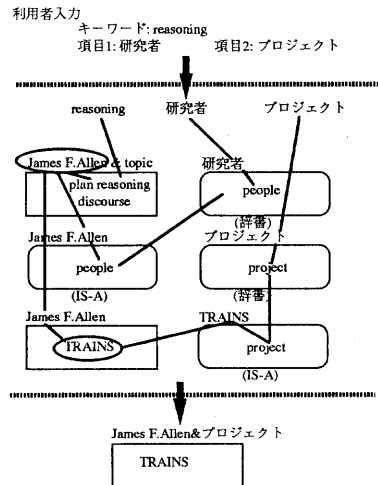


図 9: 知的情報統合機構 (step2) のアルゴリズム動作例

```

begin
n := 1;
質問文を解析して概念ユニット units を抽出する;
repeat
最初の unit から距離 n 以内にあるユニットを解集合 answers
に加える;
for unit2 := 2 番目以降の units do
answers := (answers と, unit2 毎に距離 n 以内のところに
あるユニットの集合の積集合);
n := n + 1
until (n < 3) and (answers = { });
answers, units, 及びこれらの間の経路をワークスペースに
表示する
end.

```

図 10: 活性伝播型連想検索のアルゴリズム

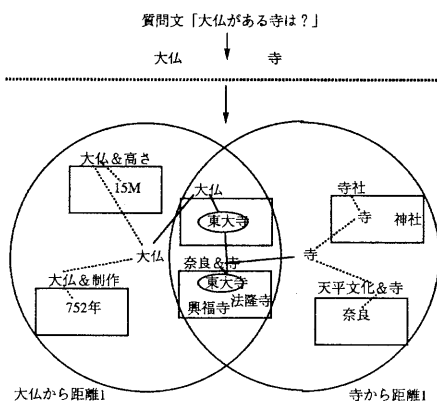


図 11: 活性伝播型連想検索のアルゴリズム動作例

ける活性伝播 [2] に基づいている。複数の概念を選択すると、概念の周囲に活性化の輪が広がり、輪が接触したところの概念を検索する。活性化の距離は key - value 間を 1 としている。図 10 にアルゴリズムの概要を示す。

5.1 例

「大仏のある寺は？」というような質問にどのように答えるかを示す (図 11 参照)。

質問文から概念ユニット「大仏」と「寺」を抽出する。距離 1 の場合の解の求め方は以下のとおりとなる。「大仏」から距離 1 にあるユニットの集合は { 「15M」, 「752年」, 「東大寺」, ... } であり、これらを answers とする。「寺」から距離 1 にあるユニットの集合は { 「寺社」, 「奈良」, 「東大寺」, 「法隆寺」, 「興福寺」, ... } である。この集合と answers の積集合は { 「東大寺」 } である。すなわち距離 1 の解は「東大寺」である。

表 2: 情報キャプチャ機構と知的情報統合機構の実験結果

		適合率	再現率
テスト1	people	93%	85%
	people以外のクラス	89%	83%
	総合	90%	83%
テスト2	project	78%	82%
	project以外のクラス	65%	69%
	総合	68%	73%

$$\text{適合率} = \frac{\text{正しく生成された連想構造のユニット数}}{\text{生成された連想構造のユニット数}} \times 100(\%)$$

$$\text{再現率} = \frac{\text{正しく生成された連想構造のユニット数}}{\text{生成されるべき連想構造のユニット数}} \times 100(\%)$$

6 実験と考察

6.1 実験 1: 人工知能情報ベース

まず、情報キャプチャ機構により、人工知能研究者のホームページを示す 100 個の URL を入力して HTML 文書を取得して、CM-2 連想構造を生成した。生成されたユニットは 764 個、連想構造は 625 個、IS-A 構造は 755 個である。クラス判定ヒューリスティックのために与えた個別知識は、people, project, e-mail, university, department, laboratory, topic など 7 つのクラスに関して 288 個である。

次に、知的情報統合機構により、people 毎に残りのクラスを整理するテスト 1 と、project 毎に残りのクラスを整理するテスト 2 を行った。情報キャプチャ機構と知的情報統合機構の統合実験結果を表 2 に示す。

適合率において people 毎で 90%、project 毎で 68%、また再現率において people 毎で 83%、project 毎で 73% の確度で情報整理が正しく行われることが確認できた。project 毎は、people 毎に比べると若干率が悪いが、これはもともとページが人間に関するページであり、project クラスのインスタンスと他のクラスのインスタンスの間に直接的な連想関係がないためである。

我々のアプローチでは、連想構造を用いることにより、特別な自然言語処理を行わなくても、HTML 文書からの重要なキーワード抽出と視点の変換がほぼ満足できる確度で行えると評価する。

6.2 実験 2: 奈良情報ベース

手作業で構築した奈良観光情報ベース (概念ユニット 1286 個、連想構造 862 個) に対して活性伝播型連想検索の実験を行った。事前の予備実験 (10 問) により、最短距離で正答 (抽出されるべき解がすべて得られ、かつ得られた解に誤りを含まない解答) が得られた場合距離をのばしても正解がふえないことと、距離 3 では正解だけでなく不要解が多量にまざることがわかったため、最初は距離 1 で探索を行い、正答が得られない場合のみ距離 2 探索を行った。表 3 に結果を示す。

50 問の質問のうち、距離 1 と距離 2 の総合評価において

表 3: 連想検索機構実験結果

情報ベース	距離	適合率	再現率
奈良観光情報ベース	距離1	64%	64%
	距離1+2	76%	92%

$$\text{適合率} = \frac{\text{正答(正解がえられ、誤りを含まない解答)数}}{\text{正解が情報ベースに含まれる質問数}} \times 100(\%)$$

$$\text{再現率} = \frac{\text{正解がえられた解答数}}{\text{正解が情報ベースに含まれる質問数}} \times 100(\%)$$

正答は 38/50(76%) あり、結果は良好であると評価する。また、解に誤りを含む場合でも、解にいたる経路が表示されるため利用者は表示された解が正しいかどうかを理解することができる。

7 関連研究

本研究は、知識メディア [3]・オントロジー・エージェントによる仲介を基本技術として知識共有と再利用の基盤の確立を目的とする知識コミュニティプロジェクト [4] の一つとして位置付けられる。本研究では、連想構造を基本構造とする知識メディアに焦点をあて、情報を収集、整理、蓄積するシステムを開発した。

本研究の関連研究として、インターネット上の異質な情報からの情報獲得の研究 [5],[6],[7],[8],[9] がある。また、WWW に関しては Yahoo をはじめとして、多くの検索エンジンが広く実用化されている。これらは主に情報収集の手法に焦点が当てられているが、本研究のように収集した情報を統合整理するところまでは研究が進められていない。

本研究は情報抽出の点から見ると、佐藤ら [10] の研究と関連する。この研究においては、あらかじめ抽出したい情報ターゲットを設定し、構造情報や必要な知識を与えて精度をあげるアプローチがとられている。本研究では、利用者がドメイン知識がない場合や最初からほしい情報がわからない場合も考慮して、不要キーワードを取り込む可能性があるがどのようなドメインでも汎用的に取り込み可能な段階と、ドメイン知識を与えてより有用な情報を取り込む段階から構成されるアプローチをとった。

本研究を、人間の創造的問題解決プロセスを支援するコンピュータシステムという広義の発想支援システムの枠組 [11] でとらえると、KJ エディタ [12] や CAT1 [13] などと類似性がある。これらは主にアイデアの発散・収束支援に焦点が当てられている。一方、我々は個人やグループの情報を収集、整理、統合して情報ベースを形成する手法に焦点を当てている。

8 おわりに

既存の雑多で構造の不均質な情報を統合する手法を提案した。基本となるアイデアとして、雑多な情報をゆるやかに関連づけるための連想構造と呼ぶデータ構造を用いた。連想構造は単純な表現であり、生データから容易に生成できること、人間が直観的に理解しやすいことが特徴である。我々は、連想構造を用いて既存の情報源から情報を収集し、

整理する過程を支援するシステム CM-2 を試作した。CM-2 では、(a) 既存の雑多で構造の不均質な情報源から情報を取り込み、CM-2 の連想構造を生成する情報キャプチャ機構、(b) 与えられた視点から情報の切り出しと構造化を行う知的情報統合機構、(c) キーワードによる連想的な検索を可能にする連想検索機構を実現した。CM-2 の有効性を実験によって確かめた。

連想構造を用いた我々のアプローチでは以下の長所がある。(1) 既存の情報源から情報を概念レベルで抽出して連想関係をつけることが容易である。(2) 連想構造は人間に直観的に理解しやすく、マップや箇条書や表などの形式に簡単に变换できる。(3) 連想構造を用いることにより、情報の構造や表現の違いを吸収し、情報内容の統合利用が行える。

今後は、情報源の種類と量をふやして、多様なドメインで実験して評価を行なう予定である。

参考文献

- [1] Brill, E.: Some Advance in Transformation-Based Part of Speech Tagging, *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, (1994).
- [2] Quillian, M. R.: *Semantic Memory. Semantic Information Processing* (Minsky, M.(ed.)), MIT Press, (1968).
- [3] Stefik, M.: The next knowledge medium, *AI Magazine* Vol. 7, No. 1, pp.34-46 (1986).
- [4] Nishida, T., Takeda, H.: Towards the knowledgeable community. *Proceedings of International Conference on Building and Sharing of Very Large-Scale Knowledge bases 93*, Japan Information Processing Development Center, pp.157-166 (1993).
- [5] Levy, A. Y., Sagiv, Y. and Srivasava, D.: Towards efficient information gathering agents, *Working Notes of the AAAI Spring Symposium on Software Agents*, pp.64-70 (1994).
- [6] Armstrong, R., Freitag, D., Joachims, T. and Mitchell, T.: A learning apprentice for the World Wide Web, *Working Notes of the AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments*, pp.6-12 (1995).
- [7] Balabanovi'c, M. and Shoham, Y.: Learning information retrieval agents: Experiments with automated web browsing, *Working Notes of the AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments*, pp.13-18 (1995).
- [8] Li, W.: Knowledge Gathering and Matching in Heterogeneous Databases, *Working Notes of the AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments*, pp.116-121 (1995).
- [9] Iwazume, M., Takeda, H., Nishida, T.: Ontology-based approach to information gathering and text categorization, *Proceedings of International Symposium on Digital Libraries*, pp.186-193 (1995).
- [10] 佐藤順, 佐藤理史, 篠田陽一: 電子ニュースのダイジェスト自動生成, 情報処理学会論文誌, Vol. 36, No. 10, pp.2371-2379 (1995).
- [11] 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.16-23 (1993).
- [12] 河合和久, 塩見彰隆, 竹田尚彦, 大岩元: 協調作業支援機能を持ったカード操作ツール KJ エディタの評価実験, 人工知能学会誌, Vol. 8, No. 5, pp. 583-592 (1993).
- [13] 角康之, 堀浩一, 大須賀節雄: 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, 人工知能学会誌, Vol. 9, No. 2, (1994).