

## テキストの構造化に基づく検索システム

岸 本 行 生<sup>†</sup> 須之内 美幸<sup>†</sup> 塚 田 康 博<sup>†</sup>  
千 葉 滋<sup>†</sup> 石 川 徹 也<sup>††</sup>

本論文では、テキストと質問文の内容を解析して知識データおよび検索意図データを抽出し、両構造化データの照合により記述文およびテキストを検索するシステムを提案する。本システムは、情報検索要求意図は質問文の用言に内在し、検索対象事項は5W1Hに相当する事項とその属性値を指示しているものとして、"情報検索要求意図"に対応し検索するために用言に対し時系列をもととする動詞ソーラスを用意し、属性値に対応し検索するために名詞ソーラスおよび属性辞書を用意し、テキスト内から知識データを抽出しデータベース化を図り、質問文を同様に解析し検索意図データを抽出し、両構造化データのパターン照合により記述文を検索し、当記述文記載のテキストを出力する。ニュース記事を対象にした評価実験の結果、96.2%の平均再現率が得られ、100%の平均適合率が得られた。また、従来のキーワード検索方式と比較実験した結果、本方式は検索精度の高いシステムを実現している。

### Information Retrieval System based on Text and Query Analysis

YUKIO KISHIMOTO,<sup>†</sup> MIYUKI SUNOUCHI,<sup>†</sup> YASUHIRO TSUKADA,<sup>†</sup>  
SHIGERU CHIBA<sup>†</sup> and TETSUYA ISHIKAWA<sup>††</sup>

In this paper, we propose a text retrieval system based on the semantic structure analysis of the contents of textual inquiries. In this system, the information retrieval intention is taken as being inherent in the terminology of the inquiry. The semantic item indicating a 5W1H correspondence and its related attribute value are captured. To enable retrieval according to the intention of the information retrieval request, we have prepared a verb thesaurus based on the time sequence logic of related vocabulary entries. For retrieval based on attribute values, we have prepared a noun thesaurus, as well as an attribute dictionary. Functionally, we extract knowledge data from text content and form a database. Similarly, we analyze interrogatives and extract retrieval intention data. Based on the results of a comparison of these two data structures, we retrieve a descriptive statement and output its text. In evaluation tests, in which the retrieval targets were news articles, a recall success rate of 96.2% was achieved with an accuracy rate of 100%. Based on these results, a retrieval system of high precision that compares favorably with the current keyword-search system has been achieved.

#### 1. はじめに

情報入手の対象の一つに、文献・文書等テキストが利用されている。テキストを検索するツールとして、従来から書誌データベースが実用に供されてきた。また、最近ではフル・テキスト・ファイルが利用され出している。いずれに対する現行の検索方式は、キー

ワード指示による検索方式であり、質問文の検索意図およびテキストの文意をもとに直接検索できない。このために、検索結果をもとに、利用者においてテキストを解読し情報を入力することが行われている。

本論文において、検索意図を反映したテキスト検索システムの提案を目的に、テキストの文意解析を行い、文意構造を抽出し、データベース化し、同じく質問文の文意解析を行い検索するシステム機能について報告する<sup>1)</sup>。

以下、2章で情報検索要求および従来のシステムについて考察を行い、3章でシステム機能を示し、4章で評価実験および考察を行い、5章で結論を示す。

<sup>†</sup> シャープ株式会社応用システム研究所  
Sharp Corporation Integrated Media Laboratories  
<sup>††</sup> 図書館情報大学図書館情報学部  
Faculty of Library and Information Science,  
University of Library and Information Science

## 2. 情報検索要求および従来のシステムについての考察

情報検索要求には、情報検索要求意図として、例えば「Aについて知りたい」という情報検索要求表現に対して、一つは検索対象事項「A」に“関連する情報”の検索結果を期待する場合と、もう一つは例えば「Aを購入したいので」といった“検索結果の利用目的”と「ただし、Aは1,000円以下」といった“検索条件”（検索対象事項の属性値）に見合う検索結果を期待する場合とがある<sup>2)</sup>。前者を関連情報検索要求と呼び、後者を知識情報検索要求と呼ぶ。すなわち、関連情報検索要求は指示検索語に対し照合するすべての検索結果を期待する検索要求である。このことに対して知識情報検索要求は検索意図および検索条件をもとに限定した検索結果を期待する検索要求である。

テキストを対象とする関連情報検索システムとして、検索システム内にソーラスを取り込み、検索指示語をもとにまずソーラス用語を検索し、関連するソーラス用語すべてをもとにテキスト内に出現する単語に対して検索を行うシステムが実用化されている<sup>3)</sup>。また、時間情報を考慮したソーラスによる検索方法も提案されている<sup>4)</sup>。知識情報検索システムについては、記事テキストを対象に5W1Hに相当する事項を検索するシステムが、絹川らによって実現している<sup>5)</sup>。しかし、5W1Hに相当する事項の属性値検索および質問文対応に至っていないことから、多様な情報検索に対応できていない、上記に示す情報検索要求意図を反映し知識データを直接検索するには、テキストの内容解析および質問文の内容解析を行い、属性値をも対象とする精緻な検索システムを実現する必要がある<sup>6)</sup>。このことに対して、質問文解析に基づくシステム化<sup>7),8)</sup>、属性値を対象とするシステム化の研究がなされてきている<sup>9)</sup>。また、テキストの見出し文を対象にキーワード検索結果を絞り込む試みがされている<sup>10),11)</sup>。しかし、“検索結果の利用目的”を対象とする、しかもトータルなシステム化の研究成果は未だない。

そこで、情報検索要求意図は、質問文の用言に内在し、検索対象事項は5W1Hに相当する事項とその属性値を指示しているものにとらえ、“検索結果の利用目的”に対応し検索するために時系列をもととする動詞ソーラスを用意し、属性値に対応し検索するために名詞ソーラスおよび属性辞書を用意し、テキスト

内から図1に示す知識データを抽出しデータベース化を図り、同じく質問文を同様に解析し、両構造化データのパターン照合によって記述文を検索し、当記述文記載のテキストを出力するシステム化を図った。

## 3. システム機能

### 3.1 質問文およびテキストからの意味構造化データの抽出

質問文およびテキスト内容を対象に文意解析を行い、図1のようなセマンティック・ネットワーク形式で表す。当データを意味構造化データと呼ぶ。ネットワークのノードには質問文およびテキスト内の語句が位置し、用言に対する5W1H相当語を格関係<sup>12)</sup>でリンクさせ、さらに用言について動詞ソーラス（時系列関係ソーラス）と5W1Hに相当する語句を名詞ソーラスおよび属性辞書とリンクし表す。

### 3.2 システム構成

検索システムは図2に示すように意味解析パーザ部、意味内容検索部および上記3種類の辞書から構成される。

意味解析パーザ部は、質問文およびテキストを解析し、それぞれを意味構造化データに展開する。意味内容検索部は、質問文の意味構造化データをもとにテキストの意味構造化データベースを検索し、適合する記述文および記述文記載のテキストを出力する。

以下、各々のモジュール機能と辞書のフォーマットについて示す。

#### 3.2.1 意味解析パーザ部

意味解析パーザ部では、下記(1)、(2)の処理により、質問文およびテキストを解析し、意味構造化データに展開する。

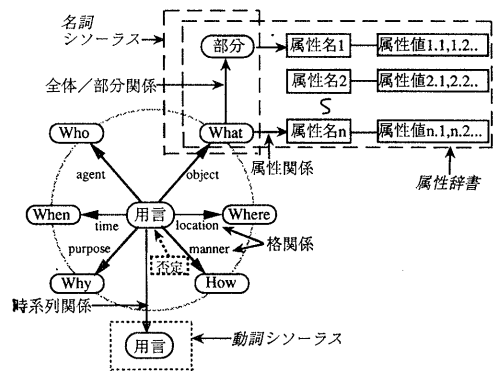


図1 意味構造化データ  
Fig. 1 Semantic structure data.

(1) 1文単位の意味構造化データ展開処理

1文単位の文末から昇順に単語辞書をもとに用言を探し、各用言に対して共起する5W1H相当語句を認識し、名詞ソーラスをもとに全体/部分関係を展開し、「部分」に相当する語句について属性辞書をもとに属性値を展開する。

例えば、「コンピュータを開発する」という文では、「開発する」に係る「コンピュータ」が単語辞書内の「を what/obj」に相当し object (対象) 関係でリンク付けられ、名詞ソーラスをもとに「コンピュータ」の「種類」が関係付けられ、その「種類」をもとに属性値が関係付けられる。なお、並列文等、テキスト・データおよび質問文に解析不可が生じては実用性がなくなることから、このことに対処するために、図6の画面にて、質問文を含め、テキスト・データ入力時にウィンドウを開き、人間介在によって正解を作り込む。

(2) 文間の構造化

文間の関係を名詞ソーラスおよび属性辞書を参照してリンク付けを行う。例えば、図3に示すように、第1の文に現れる名詞句「PS/2の新機種」と第2の文の名詞句「i486SX」の間を名詞ソーラスを参照して全体/部分関係のリンク付けを行う。

3.2.2 意味内容検索部

意味内容検索部は、下記(1)、(2)の処理により、質問文の意味構造化データをもとにテキストの意味構造化データベースに対し検索し、適合する記述文および記述文記載のテキストを出力する。数量表現に対しては、例えば、「5 kg」と「5,000 g」とを等しく扱

テキスト：「米IBM Corp.は、PS/2の新機種を開発した。i486SXを搭載した初めてのマシンである。」  
 テキストの意味構造化データ：

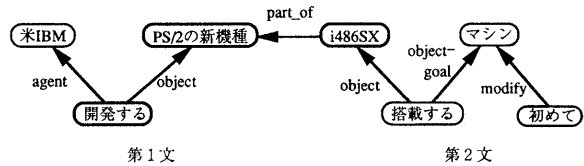


図3 テキスト文の意味構造化データ例(1)  
 Fig. 3 Example of a semantic structured data.

るように単位を揃える。

(1) 検索用意味構造化データの作成

質問文に「(コンピュータを) 購入したい」のような検索結果の利用目的が明示されている場合には、「購入する」という検索結果の利用目的を反映した検索用意味構造化データに変更する必要がある。そこで第1ステップでは、動詞ソーラスに記述されている時系列関係を参照して「購入する」という用言を「販売する」に置き換え検索用意味構造化データを作成する\*。ただし、「～について知りたい」等、単に知識を要求する質問文(例えば、「i486SXの開発について知りたい」)については、図4に示すように破線矢印で示した部分構造は削除する。第2ステップで、質問文中の語句の類義語も検索できるように名詞ソーラスの上位/下位関係を参照、下位に位置する語句を抽出しノードに追加する。

(2) 検索処理

質問文の意味構造化データの用言をもとに、テキストの意味構造化データベース内の用言を照合し、次に、用言が持つ格関係のラベルやそこにリンク付けられている語句を照合する。さらに、その語句に全体/部分関係のラベルや属性関係のラベルでリンク付けら

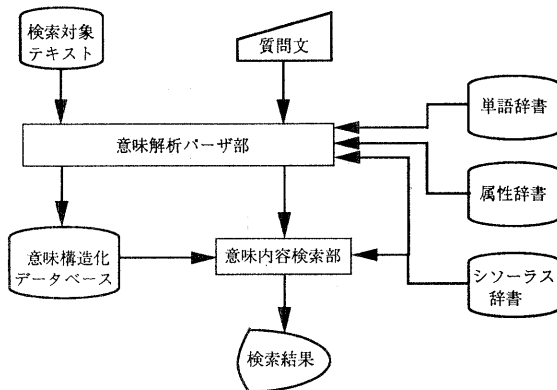


図2 システム構成図  
 Fig. 2 Structure of the system.

質問文：「i486SXの開発について知りたい。」  
 質問文の意味構造化データ：

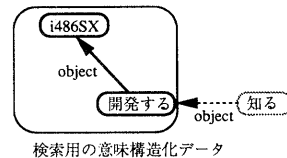


図4 検索用意味構造化データ例  
 Fig. 4 Example of a semantic structured data for retrieval.

\* 「購入する」ことを前提に情報検索を行っている以上、「販売されている」対象を検索結果として提示する必要がある。

れている語句を照合し、すべての語句とラベルとが一致する意味構造化データに対する記述文記載のテキストを検索結果として出力する。

例：図4に示す検索用意味構造化データと図5に示す意味構造化データでは、まず、用言「開発する」が一致し、object (対象) というラベルとそこにリンク付けられている「i486SX」という名詞が一致するため、「i486SXの開発について知りたい。」という質問文に対して「米 Intel Corp. は割込信号 SMI を備えた i486SX を開発中である。」という文を含むテキストが検索される。

なお、システムの実行画面を図6に示す。実行システムでは、質問文に対し該当する記述文を含むテキストの見出し文を出力表示し、ユーザが見出し文を指示し、そのもとで該当するテキストの内容を出力表示する形式をとっている。

3.2.3 システム用の辞書

以下に各辞書の構成式を示し、図7~図10にその例を示す。

(1) 単語辞書

(見出し語 (表記 読み)

(品詞 活用)

(シーソーラス ID 用言情報 \*1 (属性 ID) e) f)

\*1 = ((表層格 深層格 (意味カテゴリ: カテゴリ

ID) g) h)

ただしシーソーラス ID = 単語の意味を識別する ID

属性 ID = 属性辞書の見出しになる ID

e = 単語意味の数, f = 文法情報数

g = 意味カテゴリ数, h = 深層格数

(2) 名詞シーソーラス

(シーソーラス ID ((関係の種類 シソーラス ID) m)

n)

ただし関係の種類 = 全体/部分, 上位/下位, 前提

m = シソーラス ID の数, n = 関係の種類の数

(3) 動詞シーソーラス

(シーソーラス ID ((関係の種類 シソーラス ID) m)

n)

ただし関係の種類 = 上位/下位, 時系列関係

m = シソーラス ID の数, n = 関係の種類の数

テキスト: 「米Intel Corp.は割込信号SMIを備えたi486SXを開発中である。」

テキストの意味構造化データ:

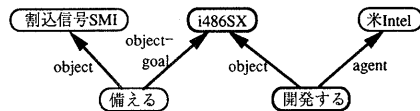


図5 テキスト文の意味構造化データ例(2)

Fig. 5 Example of a semantic structured data.

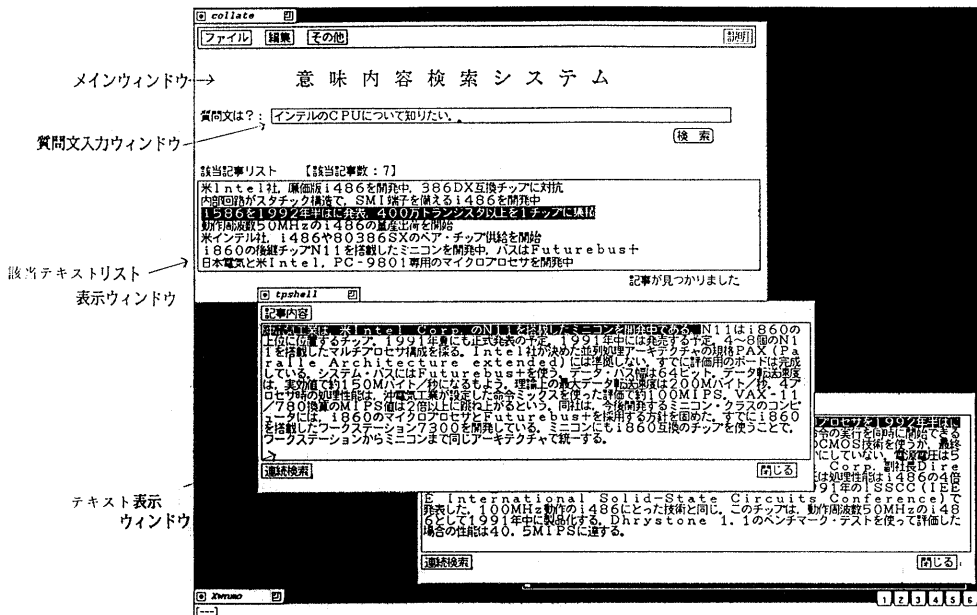


図6 システムの実行画面

Fig. 6 Example of a retrieval processing.

(4) 属性辞書

(属性 ID (属性名 (属性値)*i*)*j*)

ただし, *i*=属性値数, *j*=属性名数

4. 評価実験および考察

本論において提案した意味構造化法による検索方式が現行のキーワード検索方式と同等の検索を可能とし, しかも, 質問意図に適合するテキストのみを検索できるかを, 再現率と適合率<sup>13)</sup>をもとに確認する評価実験を行った。

4.1 実験評価用データ

『日経エレクトロニクス』の「技術速報」の約半年 (1991 年) 分の中でコンピュータのハードウェアに限定した 38 記事 (512 文) を評価実験対象テキストとした。

ただし, 並列構造を持つ文や重文などの複雑な構造を持つ文が含まれる文は, あらかじめ単文や複文などの単純な構造の文に人手で変換し利用した。また, 「i486SX (動作周波数は 20 MHz)」のように「( )括弧記号」を使って補足説明が加えられているような場合には, 括弧内の句を独立した文に人手で変換し, 対象データとした。

4.2 質問文の設定

評価実験対象テキストの内容に関し, 意味的な構造を区別した検索が行えるかどうかを調べるために質問文 80 件を用意した。例えば, 「i486SX の開発について知りたい」と「i486SX を搭載したコンピュータの開発について知りたい」の「開発」と「i486SX」のように, 同じ語句が質問文により異なった意味的な関係を取り得る文を含めている。また, 検索対象テキストと同様に, 解析可能な単文と複文を含む質問文を用意した。表 1 にその質問文例を示す。

4.3 辞書の構築

評価実験テキストおよび質問文を対象に人手により事前に分析し各々の辞書を構築した。この結果, 単語辞書の見出し語: 1800 語, 名詞ソーラス辞書の意味カテゴリ語: 75 カテゴリ, 属性辞書の属性名: 45 属性名の辞書を作成した。

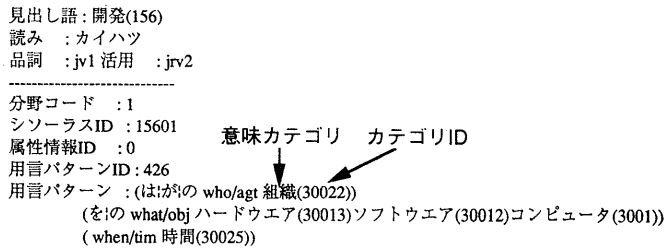


図 7 単語辞書例  
Fig. 7 Example of a word dictionary.

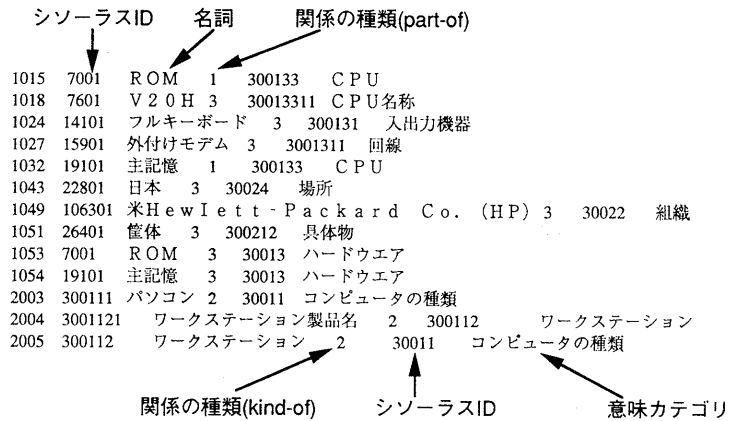


図 8 名詞ソーラス例  
Fig. 8 Example of a noun thesaurus.

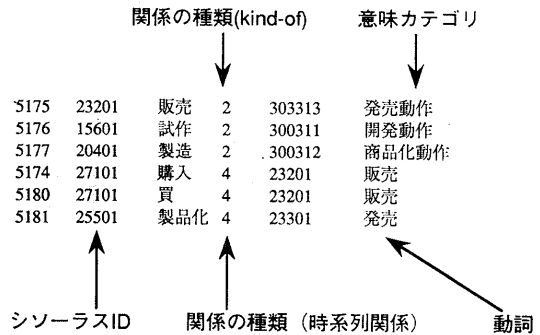


図 9 動詞ソーラス例  
Fig. 9 Example of verb thesaurus.

4.4 実験評価方法

意味構造化法の有効性を評価するためにキーワード検索方式と比較を行った (表 2 参照)。

図 11 にその結果を示す。

キーワード検索方式では, キーワードを減らす等, 検索論理式を操作することにより再現率を上げることが可能である。そこで, 基本となる検索論理式 (パ

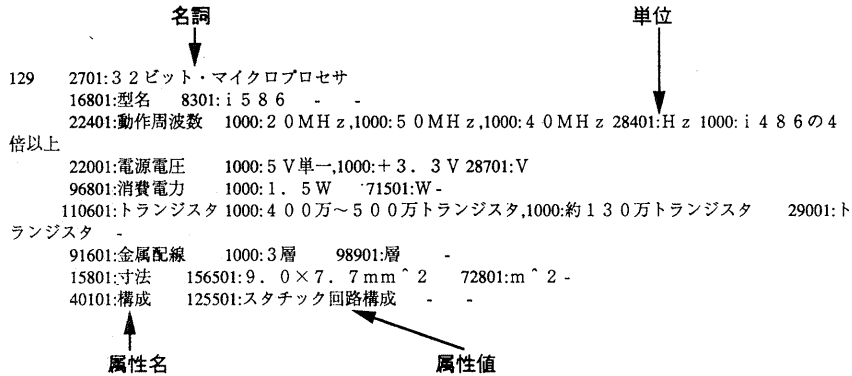


図 10 属性辞書例

Fig. 10 Example of an attribute dictionary.

ターン 1) をもとに、下記のように検索論理式を操作し、それぞれの検索結果を UNIX の egrep コマンドを用いて求めた。

- パターン 1 : 質問文から機能語 (助詞や助動詞) や活用語尾を除いて得られた語をキーワードとし、各々のキーワードに対して同義語を追加したすべてのキーワードの論理積
- パターン 2 : パターン 1 の検索論理式から用言や数字のキーワードを削除。
- パターン 3 : パターン 2 の検索論理式に類義語を追加。
- パターン 4 : パターン 2 の検索論理式から属性名のキーワードを削除。
- パターン 5 : パターン 4 の検索論理式に類義語を追加。
- パターン 6 : パターン 3 の検索論理式に具体的な製品名を追加。
- パターン 7 : パターン 5 の検索論理式に具体的な製品名を追加。
- パターン 8 : パターン 7 の検索論理式から単位記号を削除。

4.5 実験結果

(1) キーワード検索結果

基本となる検索論理式 (パターン

1) での平均再現率は 41.5% で、平均適合率は 82.1% であった。パターン 1 の検索論理式から用言や数字のキーワードを削除 (パターン 2) すると、平均再現率は 47.7%、平均適合率 74.3% になり、パターン 2 の検索論理式に類義語を付加 (パターン 3) すると、平均再現率は 62.2%、平均適合率は 67.1% になった。さらに、パターン 3 の検索論理式に具体的な製

表 1 質問文例  
Table 1 Example of queries.

質問類別	質問文例
関連情報検索	(a) i 486SX の開発について知りたい
	(b) 486SX を搭載したパソコンの開発は
	(c) FPU を内蔵するマイクロプロセサは
	(d) 486 を使ったコンピュータは
知識情報検索	条件検索
	検索結果の利用目的
	(e) 5kg より軽いパソコンについて知りたい
	(f) 60 万以下のパソコンは
	(g) 5000 ドル以下でコンピュータを購入したい
	(h) ノートパソコンを買いたい

表 2 評価観点および検索対象語  
Table 2 Viewpoint of evaluations.

	検索対象語	評価観点
パターン 1	機能語, 活用語尾を除く全語および同義語	全語照合方式評価
パターン 2	(パターン 1 に対し) 用言, 数字を除く全語	用言の必要性評価
パターン 3	(パターン 2 に対し) 類義語を付加	用言の必要性評価
パターン 4	(パターン 2 に対し) 属性名を除く全語	属性名の必要性評価
パターン 5	(パターン 4 に対し) 製品名を付加	属性名の必要性評価
パターン 6	(パターン 3 に対し) 製品名を付加	属性名の必要性評価
パターン 7	(パターン 5 に対し) 製品名を付加	属性名の必要性評価
パターン 8	(パターン 7 に対し) 単位記号を削除	単位変換の必要性評価

品名を追加 (パターン 6) すると平均再現率は 89.5%、平均適合率は 58.3% になった。最大限に検索論理式を操作 (パターン 8) すると、平均再現率は 100% になったが、そのときの平均適合率は 56.8% であった。このように検索論理式を操作すると、再現率は上がるものの、不必要なテキストが検索される件数が大きくなり、適合率が下がってしまう。

(2) 本システムによる検索結果

表 3 に示すように、意味構造化法による平均再現率は 96.2% で、平均適合率は 100% であった。

4.6 考 察

評価実験の結果、本システムは高い再現率と適合率をとともに実現できることがわかった。また、本方式は、意味的な構造の照合を行うことにより、質問文意に適合したテキストのみの検索を可能としている。例えば、「i486SX の開発について知りたい」という質問文に対して、「i486SX」が「開発」の object (対象) になる文を含むテキストのみを検索でき、「i486SX を使った製品の開発」に関するテキストは検索されない。これは、テキストおよび質問文の意味構造化データの照合において、図 3 と図 4 に示すように「i486SX」が「開発」という用言の object (対象) 関係になるか、object (対象) 関係になる語の part-of (部分) の関係になるかをセマンテック・ネットワーク上で区別しているからである。キーワード検索方式では「i486SX」と「開発」の間の object (対象) 関係を含んだ検索論理式を生成することはできないので、図 3 に示す「i486SX を搭載した PS/2 の新機種の開発」という文を含むテキストまで検索されてしまう。

実験の結果、以下の検索機能が実現されていることを確認した。

(1) 数値の範囲指定による検索

例えば、「5kg 以下のパソコン」のように属性値の

範囲を指定した質問文に対して属性値を持つ文を含むテキストのみの検索が可能である。これは、「パソコン」と「5kg」の間を属性関係でリンク付け、属性値である「5kg」を数値「5」と単位数詞「k (キロ)」と単位「g」に分割し、単位を揃えた「5,000g」という値を保持し、その数値との大小比較により実現している。このことに対してキーワード検索方式では、「5kg 以下」のような範囲を指定することはできない。

(2) 動詞の前提条件による検索

「購入したい」のような検索結果の利用目的を明示する検索意図が含まれている場合、「購入する」という事象の前提になる「販売する」などの前提条件を含むテキストの検索が可能である。キーワード検索シ

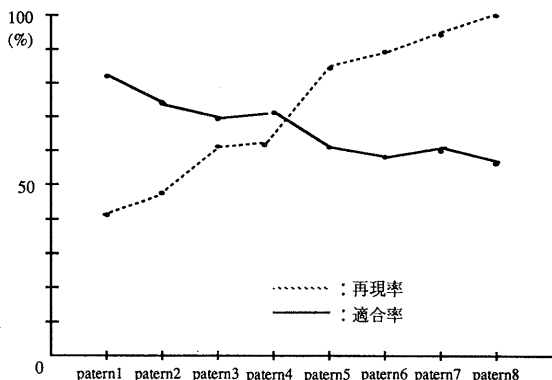


図 11 キーワード方式の検索結果  
Fig. 11 The result of retrievals by key-word.

表 3 検索実験結果表  
Table 3 The result of the experiment.

質問文例	意味構造化検索法		キーワード検索法							
			パターン1 (注1)		パターン2 (注2)		パターン4 (注3)		パターン8 (注4)	
	再現率 (%)	適合率 (%)	再現率 (%)	適合率 (%)	再現率 (%)	適合率 (%)	再現率 (%)	適合率 (%)	再現率 (%)	適合率 (%)
・5kg より軽いパソコンについて知りたい	100.0	100.0	0.0	0.0	80.0	100.0	80.0	100.0	100.0	33.3
・i486 の開発について知りたい	100.0	100.0	100.0	33.3	100.0	8.3	100.0	8.3	100.0	8.3
・486SX を搭載したパソコンの開発は	100.0	100.0	0.0	0.0	100.0	7.5	100.0	7.5	100.0	60.0
・5000 ドル以下でコンピュータを購入したい	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	13.6
・i486 と命令セット互換の MPU は	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0
Total	96.2	100.0	41.5	82.1	47.7	74.3	58.7	69.2	100.0	56.8

(注 1) 質問文中のすべての単語を組み合わせた検索式の場合

(注 2) 質問文中の単語から用言や数字を除いた語の検索式の場合

(注 3) 質問文中の単語から用言や記号や属性名など除いた語を組み合わせて作成した検索式の場合

(注 4) 質問文中の単語から用言や記号や属性名など除いた語に類義語を付与して作成した検索式の場合

テムでは、知識の利用目的を明示する動詞はキーワード検索式から省かれることが多い。また、仮に利用目的を明示する動詞を含むような場合でも、その動詞を含むテキストしか検索できない。

一方、今回の実験では検索結果が得られなかった質問文が3件あり、その理由は次の通りであった。

#### (a) 係受け構造の曖昧性処理問題

質問文中の連体修飾の用言の係り先が曖昧なために正しく解析されず、意味構造化データが照合しないものがあった。これは、用言に格関係でリンク付けられる語句の条件だけでは、複数の係り受け候補を絞り込むことができないことが原因と考えられる。これを解決するためには、用言と語句の格関係のラベル間に優先順位を設定したり、用言と語句の関連度を設定して、優先度や関連度の高い格関係構造を抽出する方法が考えられる。

#### (b) 程度性の処理問題

質問文の中に「互換」等の程度を表す語句を含むため、意味構造化データが照合しないものがあった。例えば、「IBM と互換性があるコンピュータについて知りたい」という質問文の検索用意味構造化データに対して、「互換性」という程度を表す形容詞の具体的な範囲を扱うことができなかったためである。このような程度性の表現を含む質問文に対しては、ファジィ関数等で程度を表現する必要があると考えられる。

#### (c) 省略語句の補完処理問題

本方式では、テキストの中に5W1Hに相当する語の省略があると、質問文の構造と完全照合が実現しない場合がある。5W1Hが欠落する場合は、文脈処理機能(照応処理)により補完する必要があるが、現行のシステムでは照応処理を行っていない。今回の検索実験では、省略が少ないテキストを評価実験対象テキストとして選択しており、照応処理を行わなくとも適合率が下がるという問題は生じていない。なお、照応処理は今後の課題として検討する。

### 5. おわりに

本論文では、テキストと質問文を解析して知識データおよび検索意図データを抽出し、両構造の照合により記述文およびテキストを検索するシステムを提案した。評価実験の結果、96.2%の平均再現率が得られ、100%の平均適合率が得られた。本実験システムでは、並列構造を含まない文の範囲の解析に留まっていたが、従来のキーワード検索方式に比べ検索精度の

高いシステムを実現している。

今後の課題としては、辞書作成方法が挙げられる。テキストおよび質問文の意味構造化は、辞書に依存する部分が多い。特に、時系列条件を維持する動詞シソーラスおよび属性辞書は、現在参考とするデータが提供されていないため、人手で作らざるをえず、シソーラスの整備を含めた辞書メンテナンスを支援するツールを研究開発することが必要となる。例えば、意味解析パーザ部を拡張して新規テキストを解析し、途中の意味構造化データをもとに人手を介し半自動的に語句を抽出し辞書をメンテナンスする方式が考えられる。

以上、今後は検索処理速度の問題を含め実用化に向け研究・開発を進めていく予定である。

**謝辞** 本研究の機会を与えてくださったシャープ(株)技術本部 藤本本部長、マルチメディア開発本部 新本本部長、応用システム研究所 中島所長に感謝いたします。

### 参考文献

- 1) 須之内美幸, 岸本行生, 塚田東博, 千葉 滋, 石川徹也: フルテキストの構造化に基づく検索システム, 情報処理学会データベースシステム研究会報告, 92-DBS-90, pp. 55-64 (1992).
- 2) 石川徹也: 知識データベースの構築と提供について—図書館の理念, 図書館情報学(情報学)の立場から—, 日本電子化辞書研究所ワークショップ論文集(TR-034), pp. 94-100 (1992).
- 3) 米田健二: メタモルフ(METAMORPH), テキストを理解する人工知能ソフトウェア, オンライン検索, Vol. 10, No. 3, pp. 117-126 (1989).
- 4) 黒岩真吾, ほか: テキストデータベースの知的検索手法の検討, 電子情報通信学会春季全国大会, pp. 6-7 (1990).
- 5) 絹川博之: 日本語情報検索システムにおけるキーワード自動抽出, 日立評論, Vol. 64, No. 5, pp. 25-38 (1989).
- 6) 拜原正人: 日本語文献データベースへの知的アクセス, 電子情報通信学会誌, Vol. 72, No. 7, pp. 797-806 (1989).
- 7) 森元 暉, ほか: 日本語質問文による情報検索エキスパートシステム, 情報処理学会データベースシステム研究会報告, 54-3, pp. 1-8 (1986).
- 8) Salton, G., et al.: A Vector Space Model for Automatic Indexing, *Comm. ACM*, Vol. 18, No. 11, pp. 613-620 (1975).
- 9) 松尾比呂志, ほか: 属性に基づくテキストベース検索方式, 情報処理学会論文誌, Vol. 32, No. 9, pp. 1172-1179 (1991).
- 10) 杉山健司, ほか: 自然言語理解に基づく情報検



索システム IRIS, 情報処理学会自然言語処理研究会報告, 58-8, pp. 1-8 (1986).

- 11) 秋山幸司, ほか: テキスト情報の知的検索における諸問題, 情報処理学会データベースシステム研究会報告, 64-3, pp. 1-8 (1988).
- 12) 日本電子化辞書研究所: 概念辞書 (第3版), p. 46 (1990).
- 13) 伊藤哲郎: 情報検索, p. 168, 昭晃堂 (1986).  
(平成5年7月20日受付)  
(平成6年1月13日採録)



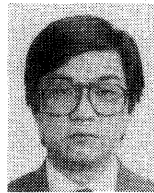
岸本 行生 (正会員)

1963年生. 1987年筑波大学第1学群人文学類卒業. 同年シャープ(株)入社. 1987年~1990年日本電子化辞書研究所に出向し, 概念辞書, 実証評価システムを担当. 現在, 応用システム研究所にて情報検索システム, 自然言語処理などの研究開発に従事.



須之内美幸 (正会員)

1964年生. 1987年筑波大学第1学群人文学類(応用言語学専攻)卒業. 同年, シャープ(株)入社. 1987~1990年, 日本電子化辞書研究所(EDR)に出向. 現在応用システム研究所にて意味理解を中心とした自然言語処理の研究開発に従事.



塚田 康博

1953年生. 1977年東京理科大学理学部卒業. 1982年シャープ(株)入社. 1986年~1989年日本電子化辞書研究所に出向. 現在応用システム研究所主任研究員. IEEE, ACM, AAAI 各会員.



千葉 滋

1946年生. 1969年東北大学工学部通信工学科卒業. 1974年同大学博士課程修了. 翌年通産省工業技術院電子技術総合研究所入所. 聴覚・音声情報処理の研究に従事. 1991年シャープ(株)入社. 現在同社応用システム研究所第一研究部長. 工学博士. 日本音響学会会員.



石川 徹也 (正会員)

昭和46年慶應義塾大学大学院修士課程(図書館・情報学専攻)修了. 富士写真フィルム(株)足柄研究所入社. 図書館短期大学, 昭和62年文部省在外研究員(UCLA)等を経て現在, 図書館情報大学教授. 工学博士. 情報管理システムの高度化の研究に従事. 人工知能学会, オフィスオートメーション学会, 日本経営工学会等会員.