

時間的因果関係を扱う動的な文脈解釈を伴った 意味的連想検索方式の実現

鷹野 孝典[†] 関子 泰三[†] 清木 康^{††}
但田 育直^{†††} 波内 みさ^{†††}

本稿では、検索者の与える検索語を1つの事象としてとらえることが可能な領域を対象として、事象間の因果関係を扱う意味的連想検索方式を示す。本方式の実現により、検索者の与える事象に因果関係のある事象について記述された文書データを検索することが可能となる。本稿では、宇宙工学分野における不具合事象を対象とした意味的検索空間を生成し、擬似的な文書データを用いた実験により、提案方式の有効性を明らかにする。

A Semantic Associative Search Method with Temporal Cause-and-effect Relationship for Event Data Sets

KOSUKE TAKANO,[†] TAIZO ZUSHI,[†] YASUSHI KIYOKI,^{††}
IKUNAO TADA^{†††} and MISA NAMIUCHI^{†††}

In this paper we discuss a semantic associative search method with temporal cause-and-effect relationship between events. The method is used for the domain in which a query given by a user can be regarded as an event. The method enables to retrieve a document data set having the relationship with a query. We created a semantic search space for the space engineering field and clarify effectiveness of the method by an experiment using a dummy document data set.

1. はじめに

近年、様々な組織内において、大量の文書データが生成され、それらがデータベースに格納されている。また、それらの文書データを活用するために、組織ごとに検索エンジンが構築されている。データベースや情報検索の研究分野では、これらの文書データ群を対象とした検索方式として、ベクトル空間モデルによる検索方式が有効であると確認されている。

SMARTシステム⁴⁾やLatent Semantic Indexing(LSI)¹⁾等の従来のベクトル空間モデルでは、対象とするドキュメントデータ群の中に出現する各単語をベクトルデータとして表現し、単語間の意味的な同義性や類似性が計算可能な計量系を提供している。しかし、ある検索対象領域において、検索者が検索語として与える単語を事象(event)と捉えることが可

能な場合、事象間の類似性を計量するといった検索要求を満たした計量系だけでは十分ではなく、事象間の因果関係が計量可能な系が必要となる。

本稿では、事象間の因果関係が計量可能なベクトル空間生成方式を示し、意味の数学モデルによる意味的連想検索方式^{2),3)}に適用する際の実現方式について述べる。提案方式により、検索者の検索目的に応じて、「ある事象の原因である事象群」について検索する、「ある事象の結果である事象群」について検索するといった、事象間の因果関係を扱う動的な文脈解釈を伴った情報検索が可能となる。本稿では、宇宙工学分野における不具合事象を対象とした意味的検索空間を生成し、擬似的な文書データを用いた実験により、提案方式の有効性を明らかにする。

2. 時間的因果関係を扱うベクトル空間生成方式

提案方式は、検索対象領域における事象間の因果関係が計量可能なベクトル空間を生成する方式である。提案方式では、検索対象とする分野において起こり得る事象群を設定し、各事象間の因果関係に応じてベクトルデータを作成することにより、事象間の因果関係が計量可能な3つのベクトル空間マ

[†] 慶應義塾大学 政策・メディア研究科
Graduate School of Media and Governance, Keio University

^{††} 慶應義塾大学 環境情報学部
Faculty of Environmental Information, Keio University

^{†††} 宇宙開発事業団
National Space Development Agency of Japan

トリクスを生成する。さらに、生成した3つのベクトル空間マトリクスを組合せを、問合せ語、および検索対象メディアデータをベクトル空間上へマッピングするための特徴行列として設定を行なう。

提案方式により生成したマトリクスを、通常のベクトル間の内積計算・cos計算による検索方式や、意味の数学モデルによる意味的連想検索方式といったベクトル空間モデルによる検索方式に適用することで、検索者の検索目的に応じて、事象間の因果関係を扱う情報検索が可能となる。

2.1 因果関係を扱うベクトル空間マトリクスの作成

事象データ間の因果関係を計量するためのベクトル空間マトリクスとして、3つの特徴行列 M_c , M_e , M を生成する。これらのマトリクスの生成ステップを、以下に示す。以下の説明において、*feature*(特徴語)とは、ベクトル空間が扱う領域を表現(定義)するのに必要な単語の集合であり、基本データとは、ベクトル空間が扱う領域中の単語の集合である。

Step-I *feature*(特徴語)群, 基本データ群の設定

3つのマトリクス M , M_c , M_e に対し、 n 個の事象語群 $e_1 \sim e_n$ を *feature*, 基本データとして、それぞれマトリクスの横軸、縦軸に設定し、 $n \times n$ の正方行列を生成する。

Step-II 特徴付けの設定

Step-I で *feature* 群, および基本データ群を設定した3つのマトリクス M , M_c , M_e について、それぞれ特徴付けの設定を行なう。それぞれのマトリクスに対する特徴付けの設定方法を以下に示す。

M :基本データ中の事象単語について、*feature* の中から、自分自身を表す事象単語, 原因である事象単語, および結果である事象単語に1を設定し、それ以外的事象単語に0を設定する。この操作を基本データ中の全事象単語に適用し、 M を生成する。

M_c :基本データ中の事象単語について、*feature* の中から、自分自身を表す事象単語, および原因である事象単語に1を設定し、それ以外的事象単語に0を設定する。この操作を基本データ中の全事象単語に適用し、 M_c を生成する。

M_e :基本データ中の事象単語について、*feature* の中から、自分自身を表す事象単語, および結果である事象単語に1を設定し、それ以外的事象単語に0を設定する。この操作を本データ中の全事象単語に適用し、 M_e を生成する。

2.2 ベクトル空間マトリクスの組の設定

2.1節で生成した3つのベクトル空間マトリクス M , M_c , M_e を用いて、「ある事象の原因である事象群」について検索する、「ある事象の結果である事象群」について検索する等の、検索目的に応じたベクトル空間マトリクスの組 $VS-1 \sim VS-3$ (図2.2)の設定方法について述べる。

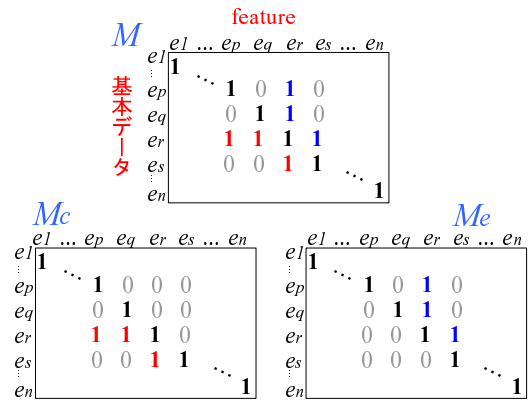


図1 ベクトル空間マトリクス M , M_e , M_c

- (a) ある事象の原因である事象群を検索する目的の場合
検索対象データベクトル形成のためのベクトル空間マトリクスとして M_e を、問合せ語ベクトル形成のためのベクトル空間マトリクスとして M_c を設定する。このベクトル空間マトリクスの組を $VS-1$ とする。
- (b) ある事象の結果である事象群を検索する目的の場合
検索対象データベクトル形成のためのベクトル空間マトリクスとして M_c を、問合せ語ベクトル形成のためのベクトル空間マトリクスとして M_e を設定する。このベクトル空間マトリクスの組を $VS-2$ とする。
- (c) ある事象の原因である事象群, および結果である事象群の両方について検索する目的の場合
検索対象データベクトル形成のためのベクトル空間マトリクス, 問合せ語ベクトル形成のためのベクトル空間マトリクスともに M を設定する。このベクトル空間マトリクスの組を $VS-3$ とする。

以上設定したベクトル空間マトリクスの組 ($VS-1 \sim VS-3$) を、意味の数学モデルによる意味的連想検索方式等のベクトル空間モデルによる検索方式に適用することで、検索者の検索目的に応じて事象間の因果関係を扱う情報検索が可能となる。

3. 意味の数学モデルによる意味的連想検索方式

意味の数学モデルによる意味的連想検索方式の概要について述べる。詳細は^{2),3)}に示されている。

3.1 意味の数学モデルによる意味的連想検索方式の概要

意味的連想検索方式は、メタデータ空間における文脈解釈、ベクトル計算により、利用者が指定した文脈に対して、意味的に近い情報を動的に検索することを可能とする検索方式である。ここで、メタデータ空間とは、分野別の専門知識を利用して、その分野の「意味」を形式的に計量することのできるベクトル空間である。意味的連想検索方式において、検索者は与える文脈を複数の単語を用いて表現する。検索者が与える単語の集合をコンテキストと呼ぶ。

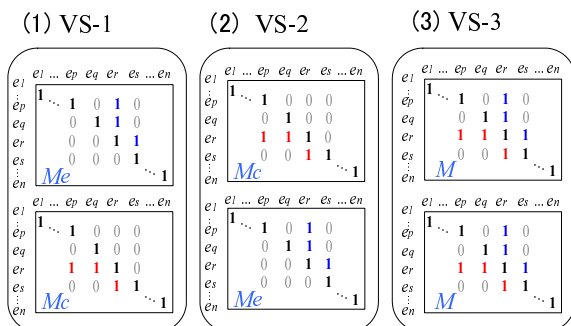


図 2 ベクトル空間マトリクスの組 (上段: 検索対象データベクトル形成のためのマトリクス, 下段: 問合せ語ベクトル形成のためのマトリクス)

3.2 メタデータ空間生成方式

以下に, メタデータ空間の生成プロセスを示す.

- (a) 対象とする分野を表現するために必要な特徴語 (以下, feature) 群を準備する. 対象分野の専門辞書等を用いて, 各見出し語を説明している説明文中の単語を抽出し, この集合を feature 群とする. これにより, その分野の意味を表現するのに必要な単語群が定義される.
- (b) 対象とする分野の基本的な用語である, 基本データ群を準備する (a) と同様に, 専門辞書を用いて, 見出し用語群を抽出し, この集合を基本データ群と定義する.
- (c) feature 群を用いて, 各基本データの特徴付けを行う. 同様の専門辞書を用いて基本データの説明文を調べ, 説明文をもとに, 関係のある feature には 1 を, 逆の意味で用いられている feature には -1 を, 関係のない feature には 0 を, それぞれ設定する. この方法で, すべての基本データに対して, feature による特徴付けを行う.
- (d) 以上の feature による基本データの特徴付けマトリクスから, 意味的連想検索のためのメタデータ空間を生成する.

以上のプロセスにより, 対象分野における意味の形式的な計算を可能とするメタデータ空間を生成する.

4. 提案方式の意味的連想検索への適用と実現方式

2章で示した提案方式により生成したベクトル空間マトリクスを, 意味の数学モデルによる意味的連想検索方式へ適用する際の実現方式について述べる.

Step-I ベクトル空間マトリクスの生成

2章で示した提案方式により, 3つのベクトル空間マトリクス M , M_c , M_e を生成する. これらのベクトル空間マトリクスは, 意味の数学モデルにおけるメタデータ行列として設定するために用いられる.

Step-II MDS の生成

意味空間生成のためのメタデータとして, マトリクス M を設定する. M の相関行列 $M^T M$ を固有値分解し, 固有ベクトルから構成される正規直交空間 MDS を生成する. ここで, MDS は, ある事象とその事象の原因である事象, およびある事象とその事象の結果である事象といった, 事象間の因果関係を計量可能な空間である.

Step-III 検索語, メディアデータのメタデータの MDS への写像

2章の提案方式で示した3つのベクトル空間マトリクスの組 ($VS-1$, $VS-2$, $VS-3$) を, 検索語ベクトル形成のためのメタデータ, およびメディアデータベクトル形成のためのメタデータとして設定する. これらのマトリクスの組から, 検索者の検索目的に応じて1つのマトリクスの組が選択される. 選択されたマトリクスの組を用いて, 検索者によって与えられた検索語, およびメディアデータのメタデータを MDS 上へ写像する.

- (1) ある事象の原因である事象群を検索する目的の場合
ベクトル空間マトリクスの組 $VS-1$ を, 検索語ベクトル形成のためのメタデータ, およびメディアデータベクトル形成のためのメタデータとして設定する. $VS-1$ を用いて, 検索者によって与えられた検索語, およびメディアデータのメタデータを MDS 上へ写像する. この場合の検索を, 以下「原因検索」と呼ぶ.
- (2) ある事象の結果である事象群を検索する目的の場合
ベクトル空間マトリクスの組 $VS-2$ を, 検索語ベクトル形成のためのメタデータ, およびメディアデータベクトル形成のためのメタデータとして設定する. $VS-2$ を用いて, 検索者によって与えられた検索語, およびメディアデータのメタデータを MDS 上へ写像する. この場合の検索を, 以下「結果検索」と呼ぶ.
- (3) ある事象の原因である事象群, および結果である事象群の両方について検索する目的の場合
ベクトル空間マトリクスの組 $VS-3$ を, 検索語ベクトル形成のためのメタデータ, およびメディアデータベクトル形成のためのメタデータとして設定する. $VS-3$ を用いて, 検索者によって与えられた検索語, およびメディアデータのメタデータを MDS 上へ写像する. この場合の検索を, 以下「原因・結果検索」と呼ぶ.

以上のステップにより, ある事象の原因について検索するための検索 (原因検索), ある事象の結果について検索するための検索 (結果検索), および事象の原因である事象群 / 結果で

表 1 各検索で設定されているメタデータ行列

	設定するメタデータ空間		
	MDS 生成用	検索語ベクトル 形成用	メディアデータ ベクトル形成用
原因検索	M	$VS - 1 (M_c, M_e)$	
結果検索	M	$VS - 2 (M_e, M_c)$	
原因・結果検索	M	$VS - 3 (M, M)$	

表 2 意味的検索空間 MDS の構成

	単語, 次元数
feature	365 語
基本データ	365 語
空間 MDS の次元	354 次元

ある事象群の両方について検索するための検索 (原因・結果検索) が意味空間上で実現される。各検索で設定されるメタデータ行列を表 1 に示す。

5. 実験

5.1 実験環境

4 章に示した実現方法により, 実験システムを構築した。意味の数学モデルへ適用するためのメタデータ行列作成のための単語群として, 宇宙開発事業団で提供されるドキュメントの不具合原因および不具合現象についての分類表の単語群 365 語を用いた。これらの単語群を基本データ, および *feature* としてそれぞれ設定し, 3 つのマトリクス M , M_c , M_e を生成した。それぞれのマトリクスにおいて, 各基本データの *feature* による特徴付けは, 宇宙工学の専門家が行った。 M を用いて生成した意味空間の構成を表 2 に示す。検索対象メディアデータとして, メタデータのみが設定された擬似的なドキュメントを 30 件設定した。各ドキュメントに設定したメタデータの内容を表 7 に示す。

5.2 実験 1

5.2.1 実験目的

提案方式により生成した意味的検索空間において実現した目的別検索, 原因検索 ($VS - 1$), 結果検索 ($VS - 2$) および原因・結果検索について, それぞれが検索目的に応じた検索に有効であることを検証する。

5.2.2 実験方法

事象単語を表す以下に示す context1-1, 1-2 を, 5.1 実験環境にて示した意味的連想検索システムへコンテキスト (検索語) として入力し, それぞれ原因検索, 結果検索および原因・結果検索を行なう。表 7 中のドキュメントの中から, context1-1, 1-2 について, それぞれ原因について述べたドキュメント 2 件, 結果について述べたドキュメント 2 件をあらかじめ設定しておく。これらのドキュメントについて, 指定したコンテキストに対する原因検索, 結果検索および原因・結果検索の検索

結果を比較する。

(context1-1) 色むら・光沢むら・色違い・しみ・汚れ・変色・斑点
(context1-2) 異常雑音

5.2.3 実験結果

5.2.2 実験方法で示した context1-1, 1-2 をコンテキスト (検索語) として入力し, 原因検索, 結果検索, 原因・結果検索を行なった実験結果を表 3, 4 に示す。表中の上段に示す 2 件のドキュメントは, 入力したコンテキストについて, それぞれ原因について述べているドキュメントである。また, 表中の下段に示す 2 件のドキュメントは, 入力したコンテキストについて, それぞれ結果について述べているドキュメントである。

5.2.4 実験考察

表 3 に示す実験結果は, context1-1 「色むら・光沢むら・色違い・しみ・汚れ・変色・斑点」に対する検索結果を示している。この検索結果において, 「原因検索」では, 「色むら・光沢むら・色違い・しみ・汚れ・変色・斑点」の原因についてのドキュメント doc09(1 位), doc08(5 位) を上位に検索し, 結果についてのドキュメント doc11(14 位), doc12(24 位) を下位に検索している。また, 「結果検索」では, 結果についてのドキュメント doc12(1 位), doc11(2 位) を上位に検索し, 原因についてのドキュメント doc08(25 位), doc09(20 位) を下位に検索している。さらに, 「原因・結果検索」では, 原因についてのドキュメント doc09(4 位), doc08(8 位), および結果についてのドキュメント doc11(3 位), doc12(5 位) それぞれを上位に検索している。ただし, 「原因・結果検索」では, 原因についてのドキュメント doc08, doc09 については「原因検索」での検索結果よりも順位が下位に検索されており, 結果についてのドキュメント doc11, doc12 については「結果検索」の検索結果よりも順位が下位に検索されていることが確認できる。これらの検索結果は, 提案方式による原因検索, 結果検索, および原因・結果検索が, それぞれの検索目的に応じて有効な検索が実現されていることを示している。

また, 表 4 に示す実験結果は, context1-2 「異常雑音」に対する検索結果を示している。この実験結果についても, context1-1 の示す実験結果についてと同様の考察ができる。

以上の実験結果は, 提案方式により生成した意味的検索空間が, 原因検索, 結果検索および原因・結果検索について, それぞれの検索目的に応じた検索に有効なことを示している。

5.3 実験 2

5.3.1 実験目的

提案方式により生成した意味的検索空間が, 複数の異なる側面の原因, 結果を持つ事象単語について, 同時に起こっていると考えられる事象単語を追加したコンテキスト (検索語) を与えることで, 文脈に応じた動的な事象間の因果関係の計量による情報検索が可能であることを検証する。

表 3 実験 1 の検索結果 (1)

context1-1: 色むら・光沢むら・色違い・しみ・汚れ・変色・斑点

document id	原因検索 VS-1		結果検索 VS-2		原因・結果検索 VS-3	
	順位	相関値	順位	相関値	順位	相関値
doc09	1	0.559260	25	0.181398	4	0.586840
doc08	5	0.422268	20	0.243327	8	0.527639
doc12	24	0.153958	1	0.600280	5	0.565532
doc11	14	0.224462	2	0.572128	3	0.664330

表 4 実験 1 の検索結果 (2)

context1-2: 異常雑音

document id	原因検索 VS-1		結果検索 VS-2		原因・結果検索 VS-3	
	順位	相関値	順位	相関値	順位	相関値
doc26	1	0.484424	20	0.178828	8	0.530468
doc25	3	0.421999	19	0.200069	13	0.455055
doc30	27	0.230328	1	0.514344	7	0.536498
doc29	12	0.293432	3	0.493342	4	0.574856

5.3.2 実験方法

ある事象単語に、同時に起こっていると考えられる 2 種類の事象単語を加えたコンテキスト (検索語) を、下記に示す context2-1, 2-2, および context3-1, 3-2 のように設定する。これらのコンテキストを、5.1 実験環境にて示した意味的連想検索システムへ入力して検索を行なう。context2-1, 2-2 についての検索は「原因検索」を行なう。context3-1, 3-2 についての検索は「結果検索」を行なう。表 7 中のドキュメントの中から、context2-1, 2-2 については、原因について述べたドキュメントをそれぞれ 1 件ずつ設定し (doc01, doc02), context3-1, 3-2 については、結果について述べたドキュメントをそれぞれ 1 件ずつ設定しておく (doc17, doc18)。これらのドキュメントについて、context2-1 と context2-2 の検索結果、および context3-1 と context3-2 の検索結果を比較する。

(context2-1) 消費電力異常 回転数規格外

(context2-2) 消費電力異常 炭化・熱分解

(context3-1) 亀裂 (割れ/クラック) 回転ムラ/トレンド変化

(context3-2) 亀裂 (割れ/クラック) 視野視線異常

5.3.3 実験結果

実験結果を表 5, 6 に示す。

5.3.4 実験考察

表 5 に示す結果は、context2-1 「消費電力異常 回転数規格外」、および context2-2 「消費電力異常 炭化・熱分解」をコンテキストとして、「原因検索」したときの検索結果を示している。doc01, doc02 はいずれも「消費電力異常」の原因についてのドキュメントである。doc01 のメタデータは、表 7 に示す通り「ガタ・緩み 異常振動 異常発振」が設定されており、

表 5 実験 2 の検索結果 (1)
原因検索 (VS-1)

document id	context2-1 消費電力異常 回転数規格外	
	順位	相関値
doc01	1	0.408987
doc02	4	0.325340

document id	context2-2 消費電力異常 炭化・熱分解	
	順位	相関値
doc02	1	0.439171
doc01	2	0.368898

表 6 実験 2 の検索結果 (2)
結果検索 (VS-2)

document id	context3-1 亀裂 (割れ/クラック) 回転ムラ/トレンド変化	
	順位	相関値
doc17	1	0.376697
doc18	8	0.277768

document id	context3-2 亀裂 (割れ/クラック) 視野視線異常	
	順位	相関値
doc18	1	0.475623
doc17	5	0.390772

doc01 は「消費電力異常」の振動異常に関する原因についてのドキュメントであることがわかる。また、doc02 のメタデータは「制御不十分 温度制御不良 温度規格外」が設定されており、doc02 は「消費電力異常」の温度異常に関する原因についてのドキュメントであることがわかる。表 5 に示す実験結果では、context2-1 「消費電力異常 回転数規格外」をコンテキストとして、「原因検索」を行なった場合には、doc01 が 1 位に検索されており、context2-2 「消費電力異常 炭化・熱分解」をコンテキストとして、「原因検索」を行なった場合には、doc02 が 1 位に検索されている。この検索結果より、回転に関する原因および温度に関する原因などの複数の異なる側面の原因を持つ事象単語「消費電力異常」について、同時に起こっていると考えられる事象単語「回転数規格外」または「炭化・熱分解」をコンテキストに追加することで、それぞれの文脈に応じた動的な事象間の因果関係の計量による検索が実現されていることが確認できる。

また、表 6 に示す結果は、context3-1 「亀裂 (割れ/クラック) 回転ムラ/トレンド変化」、および context3-2 「亀裂 (割れ/クラック) 視野視線異常」をコンテキストとして、「結果検索」をしたときの検索結果を示している。doc17, doc18 はいずれも「亀裂 (割れ/クラック)」の結果についてのドキュメントである。doc17 のメタデータは、表 7 に示す通り「異

表 7 ドキュメントのメタデータ

document id	metadata
doc01	ガタ・緩み 異常振動 異常発振
doc02	制御不十分 温度制御不良 温度規格外
doc03	異常雑音 回路設計不十分 ハンダ付不良
doc04	リニアリティ異常 過渡応答異常 光源 (測定系、校正系) のノイズ
doc05	発光スペクトル異常 帯域外特性異常 異常メッセージ出力
doc06	画像上のノイズ 画素欠陥 (白キズ等) 画素欠損
doc07	腐食・酸化・硫化・窒化・侵食・酸化膜の形成 炭化・熱分解 作業環境不良 (温湿度、清浄度、汚染ガス)
doc08	治工具不良 図面指示誤り、漏れ又は不明確 温度規格外
doc09	技量不足 不注意・ボカミス 発熱・オーバーヒート
doc10	物性 (屈折率、分散、泡、脈理、吸収率・輻射率) 異常 画像上のノイズ 透過率 (含反射率) 異常
doc11	コンタミネーション 素子感度 (含ブルーミング) 異常 コーティング劣化
doc12	材料特性異常 耐環境 (放射線、紫外線) 特性異常 周辺減光異常
doc13	強度 機構設計不十分 材料適用ミス
doc14	強度 剥がれ・剥離・浮き上り・ふくれ 孔内/外径不良
doc15	剥がれ・剥離・浮き上り・ふくれ 寸法不良 異部品装着
doc16	画素欠損 画素欠陥 (白キズ等) 画質異常
doc17	異常音 異常振動 画素間感度偏差異常
doc18	レベル特性異常 光源 (測定系、校正系) のノイズ 画素欠損
doc19	図面指示誤り、漏れ又は不明確 手順書指示誤り、漏れ又は不明確 作業指示不適
doc20	図面等の誤読 判断の誤り・勘違い 技量不足
doc21	腐食・酸化・硫化・窒化・侵食・酸化膜の形成 機械・設備不良 作業環境不良 (温湿度、清浄度、汚染ガス)
doc22	利得低下・利得増大 出力規格外れ 動作不安定
doc23	出力変動 波形不良・位相不良 画像上のノイズ
doc24	早期起動、早期停止 チャタリング システムダウン
doc25	回路設計不十分 実装不良 ハンダ付不良
doc26	接触不良 実装設計不十分 ガタ・緩み
doc27	実装不良 異常振動 図面等の誤読
doc28	出力異常 動作不安定 画像上のノイズ
doc29	動作せず、始動不能、停止不能 電圧の規格外れ・過渡電圧 画素間暗電流偏差異常
doc30	画質異常 利得低下・利得増大 光源 (測定系、校正系) のノイズ

常音 異常振動 画素間感度偏差異常」が設定されており、doc17 は「亀裂 (割れ/クラック)」の振動異常に関する結果についてのドキュメントであることがわかる。また、doc18 のメタデータは「レベル特性異常 光源 (測定系、校正系) のノイズ 画素欠損」が設定されており、doc18 は「亀裂 (割れ/クラック)」の光学異常に関する結果についてのドキュメントであることがわかる。表 6 に示す実験結果では、context3-1「亀裂 (割れ/クラック) 回転ムラ/トレンド変化」をコンテキストとして、結果検索を行なった場合には、doc17 が 1 位に検索されており、context3-2「亀裂 (割れ/クラック) 視野視線異常」を検索語として、結果検索を行なった場合には、doc18 が 1 位に検索されている。この検索結果より、振動異常に関する結果および光学異常に関する結果などの複数の異なる側面の結果を持つ事象単語「亀裂 (割れ/クラック)」について、同時に起こっていると考えられる事象単語「回転ムラ/トレンド変化」、「視野視線異常」をコンテキストに追加することで、それぞれの文脈に応じた動的な事象間の因果関係の計量による検索が実現されていることが確認できる。

以上の実験結果は、提案方式により生成した意味的検索空間が、複数の異なる側面の原因、結果を持つ事象単語について、同時に起こっていると考えられる事象単語をコンテキ

ストに追加することで、文脈に応じた動的な事象間の因果関係の計量による情報検索が可能であること示している。

6. 結 論

本稿では、検索対象領域で起こり得る事象間の因果関係を扱う意味的連想検索方式を示した。本方式は、事象間の因果関係をベクトル空間の形式で表現することが可能となることが特徴である。また、本方式では、3 つのベクトル空間マトリクスの組をそれぞれ用いることで、「原因検索」、「結果検索」、および「原因・結果検索」といった検索者の目的に応じた文書検索が可能となる。本稿では、宇宙工学分野における不具合事象を対象とした意味的検索空間を生成し、擬似的な文書データを用いた検索実験により、本方式の有効性を確認した。

今後は、大量の実事象データ・文書データを対象として、本方式の有効性評価実験、各ベクトルデータの重み付けの導入を行っていく予定である。

謝辞

本研究に関して、意味的検索空間作成および文書へのメタデー

タ付与について多くの貴重なご助言を頂いた, (株) 翔エンジニアリング仲春男氏, 宇宙開発事業団中塚大貴氏, 米山一彦氏に感謝申し上げます.

参 考 文 献

- 1) Deerwester, S. C., Dumais, S. T., Furnas, G. W., Landauer, T. K. and Harshman, R. A.: "Indexing by latent semantic analysis", Journal of the American Society for Information Science, Vol. 41, No. 6, pp. 391-407, 1991.
- 2) Kiyoki, Y., Kitagawa, T. and Hayama, T.: "A Metadatabase System for Semantic Image Search by a Mathematical Model of Meaning," Multimedia Data Management - using metadata to integrate and apply digital media -, McGrawHill, A. Sheth and W. Klas(editors), Chapter 7, 1998.
- 3) 清木康, 金子昌史, 北川高嗣: "意味の数学モデルによる画像データベース探索方式とその学習機構," 電子情報通信学会論文誌,D-II,Vol.J79-D-II,No. 4,pp. 509-519, 1996.
- 4) Salton, G., Wong, A. and Yang, C. S.: "A vector space model for automatic indexing", Communications of the ACM, Vol. 18, No. 11, pp.613-620, 1975.
- 5) 鷹野孝典, 清木康: "異分野データベース群を対象とした意味的検索空間統合プロセスの実現", DBSJ Letters, Vol.1, No.1, pp. 55-58, 2002.
- 6) 吉田 尚史, 関子 泰三, 清木 康, 北川 高嗣: "ドキュメントデータ群を対象とした文脈依存動的クラスタリングおよび意味的データマイニング方式," 情報処理学会論文誌: データベース, Vol. 41, No. SIG 1 (TOD5), pp.127-139, 2000.
- 7) 関子 泰三, 吉田 尚史, 清木 康: "ドキュメントデータ群を対象とした文脈依存動的クラスタリングの再帰的適用による意味的知識発見方式," 情報処理学会論文誌: データベース, Vol. 43, No. SIG 2(TOD13), pp.216-230, 2002.
- 8) 関子 泰三, 鷹野 孝典, 清木 康: "事象データ群の時間因果関係を扱う意味的連想検索方式," 情報処理学会研究報告 DBS-130, pp.71-77,2003.