

精神医学分野データベースを対象とした 抽象度依存計量空間と因果関係計量空間の実現方式

曽根 由希子[†] 吉田 尚史[†] 清木 康^{††}

本稿では精神医学分野データベースを対象とし、病名・症状の抽象度と因果関係性を反映した検索の実現方式を示す。本方式は事象の上位下位概念といった抽象度依存検索や原因結果といった因果関係検索などの、特定の方向を有したベクトル空間による検索を可能とする。本方式により、特定の方向性に関する計量を実現することにより、利用者の検索目的に応じた事象やその文書データを検索可能となる。本稿では、精神医学データベースを対象とし、本方式の実現可能性を検証する。

An implementation of a search method with calculation of abstractive and causal relationships for psychiatric databases

YUKIKO SONE,[†] NAOFUMI YOSHIDA[†] and YASUSHI KIYOKI^{††}

In this paper, we present a search method with calculation of abstractive and causal relationships for psychiatric databases. This method realizes a search space for events and documents with specific directions (e.g. searching causes and effects or abstracts and concretes of events) on vector spaces. Our method enables searchers to search events and related documents according to their purposes of the search with specific directions. We will clarify the feasibility of our method for psychiatric databases.

1. はじめに

現代社会では、日々の生活のストレスによって心の病に陥る人も少なくないが、それを軽視して、病を悪化させたり、自らの心的環境の不健康さに気づいたとしても周囲の目を気にするなどして精神科の受診やカウンセリングを拒む人は少なくない。

しかし、心の問題について、自ら知識を得ようと試みる場合、精神医学分野において検索者がこの分野のデータベース検索を行う際に、利用する一般的な検索語群と検索対象に含まれるメタデータとのマッチングが従来のパターンマッチングによる方法では困難である。

また、精神医学分野において、障害の分類などが複雑で、知識の少ない検索者には様々な抽象度や因果関係などの方向性を反映した検索が難しい。例えば、現在の症状の原因となる病名は、専門的な知識なくしては検索が困難である。また、精神医学分野において、障害の原因や障害同士の関係性など、まだ解明されていない事象は多く存在している。解明されていない事象は専門家には推測可能であっても、検索者には障害の分類や定義などの知識が不十分なため、解明されていない

事象の推測は大変困難であり、必要な情報を入手するのは難しい。

これまで多くのベクトル空間モデルでは特定の方向性に応じた検索（例：因果関係計量における原因検索など）が困難であった。文献¹⁾及び²⁾により、ベクトル空間モデルによる因果関係計量が実現されてきた。これにより、任意の事象群の直接的な因果関係の定義から、それらの事象群を対象として、与えられた文脈に応じた因果的関連性の相関量を伴ったランキングを獲得することが可能となった。

精神医学分野において、抽象度依存計量および因果関係計量の実現は、検索者に高度な情報獲得の機会を飛躍的に高めることが可能と考えられる。抽象度依存計量により症状や病名の抽象度を反映した検索が実現可能となる。因果関係計量により症状と病名の因果関係を反映した検索が実現可能となる。

本稿では、このような特定の方向性に応じた計量を可能とするベクトル空間の実現方式を示す。本方式によって、専門性の高い知識を持たない検索者が目的とする方向性（因果関係性、抽象度）に応じた情報獲得を行うことが可能になる。

2. 実現方式

本実現方式の基礎となる意味の数学モデルによる意味的連想検索方式は文献⁶⁾⁷⁾に示されている。また、こ

[†] 慶應義塾大学大学院政策・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

のモデルを応用した因果関係計量を実現する意味的連想検索方式は文献¹⁾²⁾に示されている。本稿で示す方式は、これらを基礎とし、抽象度に関する計量を加えることにより、抽象度の高さおよび因果関係性といった特定の方向性に応じた検索を実現するものである。

本方式は、次のステップ (Step-I ~ Step-IV) で実現される。

Step-I ベクトル空間マトリクスの生成 3種類の隣接行列 M_1, M_2, M_3 を生成する (図2, 事象間のデータ例は図1で示す。) これらのマトリクスは、意味の数学モデル⁶⁾⁷⁾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。 M_2 (図2 中心部) は方向1 (原因から

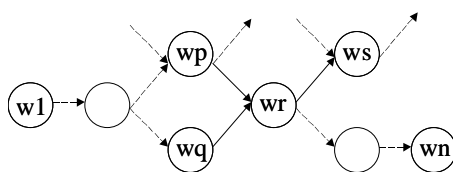


図1 事象群のデータの関連

		M1 特徴語					M2					M3							
		w1...	wp	wq	wr	ws...	wn	w1...	wp	wq	wr	ws...	wn	w1...	wp	wq	wr	ws...	wn
基本データ	w1	1...																	
	wp		1	0	1	0													
	wq		0	1	1	0													
	wr		1	1	1	1													
	ws		0	0	0	1	1...												

図2 ベクトル空間生成用行列 M_1, M_2, M_3

結果または上位から下位) に対応して、基本データに対応する行の特徴語に対応する列に、方向1に関連のある要素のみ1を有する行列である。ただし、自分自身に関しても1を有するように構成する。

M_3 (図2 右) は方向2 (結果から原因または下位から上位) に対応して、基本データに対応する行の特徴語に対応する列に、方向2に関連のある要素のみ1を有する行列である。ただし、自分自身に関しても1を有するように構成する。

M_1 (図2 左) は M_2 および M_3 を合成した行列であり、方向1と方向2の両方の性質を有する行列である。

Step-II 正規直交空間 MDS を生成意味的検索空間 MDS ⁶⁾⁷⁾ 生成のためのメタデータ行列としてマトリクス M_1 (もしくは M_2, M_3) を設定する。 M_1 の相関行列 $M_1^T M_1$ (もしくは $M_2^T M_2, M_3^T M_3$) 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。行列 X の転置行列を X^T 、行列 X の自己相関行列を得る関数を S_1 、自己相関行列 $X^T X$ を固有値分解し、正規直交空間 MDS を得る関数を S_2 とすると、本ステップは次のように表現できる。

$$S_1: X \quad X^T X$$

$$S_2: S_1(X) \quad MDS$$

$$(X = M_1, M_2, M_3)$$

MDS は、ある発生した事象について、その原因や結果

となる事象群を得るための状況を規定する文脈を与えることにより、文脈に応じた事象間の因果関係の計量を可能とする意味的検索空間である。

Step-III 検索語及び検索対象データベクトルの生成 VS-1 ~ VS-3 (図3) のマトリクスの3組を生成し、その中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。

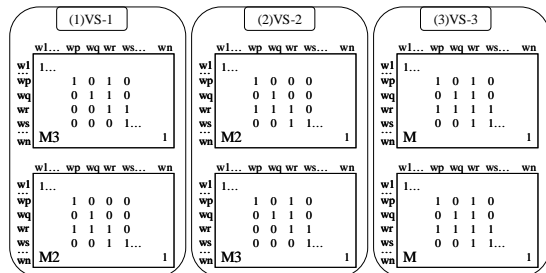


図3 ベクトル空間生成用行列の組 (上段:検索対象データのベクトル形成のための行列, 下段:問い合わせ語ベクトル形成のための行列)

(1) ある事象の方向1 (原因から結果または上位から下位) である事象群を検索する目的の場合

マトリクスの組 VS-1 (M_2, M_3) を検索語ベクトル形成のためのメタデータ行列, および検索対象データベクトル形成のためのメタデータ行列として設定する。VS-1 では、問い合わせ語ベクトル形成のための行列として M_2 を、検索対象データのベクトル形成のための行列として M_3 を用いることにより、方向1の検索が可能となる。

(2) ある事象の方向2 (結果から原因または下位から上位) である事象群を検索する目的の場合

(1)と同様に、マトリクスの組 VS-2 (M_3, M_2) を検索語ベクトル形成のためのメタデータ行列, および検索対象データベクトル形成のためのメタデータ行列として設定する。VS-2 では、問い合わせ語ベクトル形成のための行列として M_3 を、検索対象データのベクトル形成のための行列として M_2 を用いることにより、方向2の検索が可能となる。

(3) ある事象の方向1および方向2である事象群を検索する目的の場合

(1)と同様にマトリクスの組 VS-3 (M_1, M_1) を検索語ベクトル形成のためのメタデータ行列, および検索対象データベクトル形成のためのメタデータ行列として設定する。VS-3 では、問い合わせ語ベクトル形成のための行列, 検索対象データのベクトル形成のための行列として M_1 を用いることにより、方向1および方向2の両方向の検索が可能となる。

Step-IV 検索語, 及び検索対象データベクトルの MDS への写像

Step-III により、得られた検索対象データベクトルのた

めのマトリクス, および, 問い合わせ語形成のためのマトリクスを Step-II により生成した MDS に写像し, 意味的連想検索⁶⁾⁷⁾により与えられた検索語と検索対象との相関量を計量する。

これらのステップにより, 検索者の検索目的に応じた事象間の方向性を考慮した検索が, 文脈に応じて, 文脈解釈機能を伴った動的な情報検索を可能とする意味的検索空間上で実現される。

3. 適用と実現

3.1 精神医学分野を対象とした因果関係計量への適用と実現

事象間の因果関係を扱うベクトル空間生成方式の概要について述べる。本方式では, 検索対象領域における事象間の抽象度が計量可能なベクトル空間を生成する。本システムでは, システム生成の際, 精神医学分野において, 信頼性の高い障害分類体系として DSM-IV-TR⁴⁾を対象とした, 因果関係計量空間で対象としたのは, その中の解離性障害で, 基本データ, 特徴語ともに 52 事象を採用し, 関連付けを行った。図 4 はそのデータを「解離性障害」に特化して一部抜粋したものである。

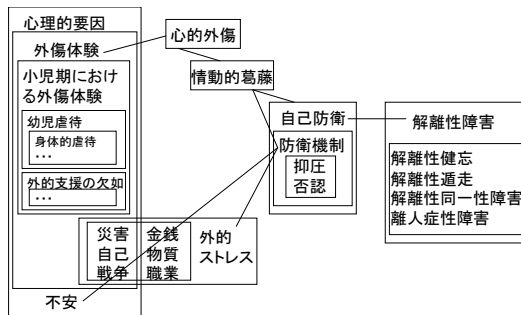


図 4 因果関係計量空間データ例 (データを一部抜粋し, 「解離性障害」の結果という方向性のみで特化し, 表示した。図中, 線で繋がれている事象が因果関係のあるもので, 線で繋がれている事象の左の方向にある事象が右の方向にある事象の原因となっている。)

3.1.1 マトリクスの生成

因果関係マトリクスとして 3 つの特徴行列 M, M_c, M_e を生成する。これらのマトリクスの生成ステップを, 次に示す。

Step-I 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 $w_1 \sim w_n$ を特徴語, 基本データとして, それぞれのマトリクスの横軸, 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 で特徴語群, 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づけの設定を行う。それぞれのマ

トリクスに対する特徴づけの設定方式を次に示す。

M : 基本データ中の事象語群について, 特徴語の中から, 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, それ以外の事象単語に 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。(表 1 に一部抜粋し, 掲載)

表 1 M (自分自身とそ直接の原因結果について関連付けを行ったマトリクス)

	情緒的 外傷体験	虐待	身体的 虐待	アルコール 乱用	自殺 企図	...	解離性 逃走	多重人格 障害
情緒的 外傷	1	0	0	0	0		1	1
虐待	0	1	0	0	0		0	0
身体的 虐待	0	0	1	0	0		0	0
アルコール 乱用	0	0	0	1	0		1	0
自殺 企図	0	0	0	0	1		1	0
...								
解離性 逃走	0	0	0	1	1		1	0
多重 人格障害	1	1	0	0	0		0	1

M_c : 基本データ中の事象単語について, 特徴語の中から, 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, それ以外の事象単語に 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。(表 2 に一部抜粋し, 掲載)

表 2 M_c (自分自身とそ直接の原因概念について関連付けを行ったマトリクス)

	情緒的 外傷体験	虐待	身体的 虐待	アルコール 乱用	自殺 企図	...	解離性 逃走	多重人格 障害
情緒的 外傷	1	0	0	0	0		0	0
虐待	0	1	0	0	0		0	0
身体的 虐待	0	0	1	0	0		0	0
アルコール 乱用	0	0	0	1	0		0	0
自殺 企図	0	0	0	0	1		0	0
...								
解離性 逃走	1	0	0	1	1		1	0
多重 人格障害	1	0	0	0	0		0	1

M_e : 基本データ中の事象単語について, 特徴語の中から, 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, それ以外の事象単語に 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_e を生成する。(表 3 に一部抜粋し, 掲載)

表 3 M_e (自分自身とそ直接の結果について関連付けを行ったマトリクス)

	情緒的 外傷体験	虐待	身体的 虐待	アルコール 乱用	自殺 企図	...	解離性 逃走	多重人格 障害
情緒的 外傷	1	0	0	0	0		1	1
虐待	0	1	0	0	0		0	0
身体的 虐待	0	0	1	0	0		0	0
アルコール 乱用	0	0	0	1	0		1	0
自殺 企図	0	0	0	0	1		1	0
...								
解離性 逃走	0	0	0	0	0		1	0
多重 人格障害	0	0	0	0	0		0	1

3.1.2 組の生成

3.1.1 節で生成した 3 つの抽象度依存マトリクス M, M_c, M_e を用いて、検索の方向性に応じた抽象度依存マトリクスの組 VS-1 ~ VS-3 の設定方法について述べる。

(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因果関係マトリクスとして M_e を、問い合わせ語ベクトル形成のための因果関係マトリクスとして M_c を設定する。この因果関係マトリクスの組を 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因果関係マトリクスとして M_c を、問い合わせ語ベクトル形成のための抽象度依存マトリクスとして M_e を設定する。この因果関係マトリクスの組を VS-2 とする。

(c) ある事象の原因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因果関係マトリクス、問い合わせ語ベクトル形成のための因果関係マトリクスともに M を設定する。この因果関係マトリクスの組を VS-3 とする。

このように設定した因果関係マトリクスの組 (VS-1 ~ VS-3) を内積計算やコサイン距離計算などのベクトル演算やベクトル空間モデルによる検索方式に適用することにより、検索者の検索の方向性に応じた因果関係による情報検索が可能となる。

3.2 精神医学分野を対象とした抽象度依存計量の適用と実現

事象間の抽象度を扱うベクトル空間生成方式の概要について述べる。本方式では、検索対象領域における事象間の抽象度が計量可能なベクトル空間を生成する。対象データ例は図 5 に一部抜粋して示した。実際に使用したデータは、基本データ、特徴語ともに 103 事象である。本システムでは、システム生成の際、精神医学分野において、信頼性の高い障害分類体系として DSM-IV-TR³⁾ を対象とした。抽象度依存計量空間で対象としたのはその中の全障害である。

3.2.1 マトリクスの生成

抽象度依存マトリクスとして、3 つの特徴行列 M_u, M_l, M_m を生成する。これらのマトリクスの生成ステップを、次に示す。

Step-I 特徴語群、基本データ群の設定 3 つのマトリクス M_u, M_l, M_m に対し、 n 個の事象語 $w_1 \sim w_n$ を特徴語、基本データとして、それぞれのマトリクスの横軸、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 で特徴語群、および基本データ群を設定した 3 つのマトリクス、 M_u, M_l, M_m

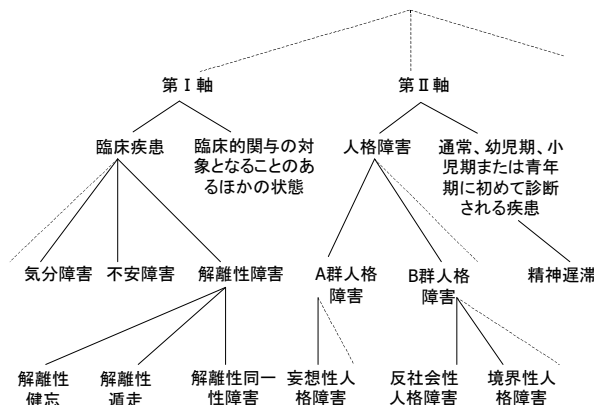


図 5 精神医学分野における病名の分類

についてそれぞれ特徴づけの設定を行う。それぞれのマトリクスに対する特徴づけの設定方式を次に示す。

M_m : 基本データ中の事象語群について、特徴語の中から、自分自身を表す事象単語、上位概念である事象単語、および下位概念である事象単語に 1 を設定し、それ以外の事象単語に 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_m を生成する。(表 4 に一部抜粋し、掲載)

表 4 M_m (自分自身とその直接の上位概念、下位概念について関連付けを行ったマトリクス)

	第1軸	臨床疾患	不安障害	外傷後ストレス障害	強迫性障害	...	解離性障害	解離性同一性障害
第1軸	1	1	0	0	0		0	0
臨床疾患	1	1	1	0	0		1	0
不安障害	0	1	1	1	1		0	0
外傷後ストレス障害	0	0	1	1	0		0	0
強迫性障害	0	0	1	0	1		0	0
...								
解離性障害	0	1	0	0	0		1	1
解離性同一性障害	0	0	0	0	0		1	1

M_u : 基本データ中の事象単語について、特徴語の中から、自分自身を表す事象単語、および上位概念である事象単語に 1 を設定し、それ以外の事象単語に 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_u を生成する。(表 5 に一部抜粋し、掲載)

M_l : 基本データ中の事象単語について、特徴語の中から、自分自身を表す事象単語、および下位概念である事象単語に 1 を設定し、それ以外の事象単語に 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_l を生成する。(表 6 に一部抜粋し、掲載)

3.2.2 組の生成

3.2.1 節で生成した 3 つの抽象度依存マトリクス M_m, M_u, M_l を用いて、検索の方向性に応じた抽象度依

表5 M_u (自分自身とその直接の上位概念について関連付けを行ったマトリクス)

	第1軸	臨床疾患	不安障害	外傷後ストレス障害	強迫性障害	...	解離性障害	解離性同一性障害
第1軸	1	0	0	0	0		0	0
臨床疾患	1	1	0	0	0		0	0
不安障害	0	1	1	0	0		0	0
外傷後ストレス障害	0	0	1	1	0		0	0
強迫性障害	0	0	1	0	1		0	0
...								
解離性障害	0	1	0	0	0		1	0
解離性同一性障害	0	0	0	0	0		1	1

表6 M_l (自分自身とその直接の下位概念について関連付けを行ったマトリクス)

	第1軸	臨床疾患	不安障害	外傷後ストレス障害	強迫性障害	...	解離性障害	解離性同一性障害
第1軸	1	1	0	0	0		0	0
臨床疾患	0	1	1	0	0		1	0
不安障害	0	0	1	1	1		0	0
外傷後ストレス障害	0	0	0	1	0		0	0
強迫性障害	0	0	0	0	1		0	0
...								
解離性障害	0	0	0	0	0		1	1
解離性同一性障害	0	0	0	0	0		0	1

存マトリクスの組 VS-1 ~ VS-3 の設定方法について述べる。

(a) ある事象の上位概念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抽象度依存マトリクスとして M_l を、問い合わせ語ベクトル形成のための抽象度依存マトリクスとして M_u を設定する。この抽象度依存マトリクスの組を VS-1 とする。

(b) ある事象の下位概念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抽象度依存マトリクスとして M_u を、問い合わせ語ベクトル形成のための抽象度依存マトリクスとして M_l を設定する。この抽象度依存マトリクスの組を VS-2 とする。

(c) ある事象の上位概念および下位概念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抽象度依存マトリクス、問い合わせ語ベクトル形成のための抽象度依存マトリクスともに M を設定する。この抽象度依存マトリクスの組を VS-3 とする。

以上設定した抽象度依存マトリクスの組 (VS-1 ~ VS-3) を内積計算やコサイン距離計算などのベクトル演算やベクトル空間モデルによる検索方式に適用することで、検索者の検索の方向性に応じた抽象度による情報検索が可能となる。

4. 実験

本方式の実現可能性を検証することを目的とし、各

実験において、それぞれ検索 a ~ j を設定し、本方式により実験を行った結果を専門家が設定した正解と比較することにより、実現可能性の検証を行う。

4.1 実験 1

4.1.1 目的と実験環境

本実験では 3.1 節で示した実現システムを用い、原因・結果の検索で、必要な結果が出ているかを検証する。

なお、対象とする空間は基本データ、特徴語ともに 51 語で 51 次元の空間となっている。

本実験で設定した検索 a, b は以下の通りである。

検索 a : 原因検索 : 解離性健忘の原因を検索する。

検索 b : 結果検索 : 情緒的外傷の結果を検索する。

4.1.2 結果と考察

検索 a, b の実験結果を表 7 に示す。

表 7 実験 1 検索結果 (下線部は正解データ)

検索 a: 解離性健忘の 原因の検索結果		相関量	検索 b: 心的外傷の 結果の検索結果		相関量
1	防衛機制	0.594826	1	離人症性障害	0.742250
2	外傷体験	0.592716	2	解離性障害	0.679682
3	心的葛藤	0.581153	3	自己防衛	0.673910
4	外的ストレス	0.546515	4	心的葛藤	0.562316
5	心的外傷	0.479928	5	防衛機制	0.559925
6	解離性健忘	0.436563	6	心的外傷	0.519371
7	不安	0.421955	7	解離性健忘	0.493735
8	うつ病	0.303318	8	不安	0.362852
9	内分泌障害	0.303318	9	外傷体験	0.337718
10	事故	0.303318	10	外的ストレス	0.294838

正解データについては、専門書に従って設定し、下線を付与した。検索 a に対しては、正解データに直接的な原因と間接的な原因が、検索 b に対しては、正解データに直接的な結果と間接的な結果が含まれる。

実験結果として、原因検索、結果検索共に良い結果が得られている。検索 a の原因検索では、解離性健忘の直接的な原因の「防衛機制」、「外傷体験」はもちろん、間接的にしか関連のない「心的葛藤」、「外的ストレス」などの事象も上位に検索できた。検索 b の結果検索でも、心的外傷の直接的な結果の「離人症性障害」、「心的葛藤」、「自己防衛」はもちろん、間接的にしか関連のない「解離性障害」、「防衛機制」などの事象も上位に検索された。

4.2 実験 2-1

本実験では 3.2 節で示した実現システムを用い、上位下位概念の検索で、必要な結果が出ているかを検証する。なお、対象とする空間は基本データ、特徴語ともに 103 語で 103 次元の空間となっている。

4.2.1 目的と実験環境

本実験で設定した検索 c,d は以下の通りである。

検索 c：上位概念検索：多重人格障害の上位概念を検索する。

検索 d：下位概念検索：解離性障害の下位概念を検索する。

4.2.2 結果と考察

検索 c,d の実験結果を表 8 に示す。

表 8 実験 2-1 検索結果 (下線部は正解データ)

検索a: 多重人格障害の 上位概念の検索結果		相関量	検索b: 解離性障害の 下位概念の検索結果		相関量
1	<u>解離性障害</u>	0.829917	1	<u>解離性健忘</u>	0.476209
2	解離性同一性障害	0.754050	2	解離性同一性障害	0.476209
3	多重人格障害	0.754050	3	解離性遁走	0.476209
4	解離性遁走	0.468455	4	心因性遁走	0.476209
5	心因性遁走	0.468455	5	多重人格障害	0.476209
6	解離性健忘	0.437349	6	心因性健忘	0.476209
7	心因性健忘	0.437349	7	離人症性障害	0.390771
8	臨床疾患	0.370639	8	解離性障害	0.227480
9	離人症性障害	0.347216	9	特定不能の身体...	0.150046
10	臨床的関与の対象...	0.267174	10	転換性障害	0.150046

正解データについては、専門書に従って設定し、下線を付与した。検索 c に対しては、正解データに直接的な上位概念と間接的な上位概念が、検索 d に対しては、正解データに直接的な下位概念と間接的な下位概念が含まれる。

検索 c の上位概念検索では、多重人格障害の上位概念「解離性障害」は 1 位に検索できたが、間接的に関連のある「臨床疾患」、「第 1 軸」などの事象が上位に検索できず、抽象度が多重人格障害と同じである「解離性遁走」、「解離性健忘」などがより上位に検索された。検索 d の下位概念検索では、解離性障害の下位概念である「解離性健忘」「解離性同一性障害」などが上位を占め、非常に良い結果が得られている。

4.3 実験 2-2

4.3.1 目的と実験環境

実験 2-1、検索 a でもそうであったように、自分自身と自分と直接関連のある事象のみに関連付けを行うと、間接的に関連のある事象の検索が難しい。そこで検索語ベクトルを拡張し、間接的に関連のある事象にも関連付けを行った。検索語ベクトルの生成については、各検索語の特徴づけにおいて自分自身と自分に直接関連のある事象に 1 を、間接的に関連のある事象には、それが間接的になるにつれて 1/2 倍し、関連付けを行った。上位概念の関連付けを例に出せば、ある事象 A には、自分自身と自分の直接の上位に 1 がつき、自分の直接の上位概念の更に上位概念に 0.5 を、更にその上位概念には 0.25 を関連付ける。本実験では以上のように生成された検索語ベクトルを用い(実験 1,3,4 でも検索語ベクトルは拡張したのものを使った。)、検索を行った。なお、対象とする空間は基本データ、特徴語ともに

103 語で 103 次元の空間となっている。

本実験で設定した検索 e,f は以下の通りである。

検索 e：上位概念検索：多重人格障害の上位概念を検索する。

検索 f：下位概念検索：解離性障害の下位概念を検索する。

4.3.2 結果と考察

検索 e,f の実験結果を表 9 に示す。

表 9 実験 2-2 検索結果 (下線部は正解データ)

検索b: 多重人格障害の 上位概念の検索結果		相関量	検索a: 解離性障害の 下位概念の検索結果		相関量
1	<u>第 1 軸</u>	0.741865	1	<u>解離性健忘</u>	0.476209
2	臨床疾患	0.490010	2	解離性同一性障害	0.476209
3	解離性障害	0.478842	3	解離性遁走	0.476209
4	多重人格障害	0.454470	4	心因性遁走	0.476209
5	解離性同一性障害	0.454470	5	多重人格障害	0.476209
6	通常、幼児期、小児...	0.426641	6	心因性健忘	0.476209
7	臨床的関与の対象...	0.413407	7	離人症性障害	0.390771
8	心因性遁走	0.270127	8	解離性障害	0.227480
9	解離性遁走	0.270127	9	特定不能の身体...	0.150046
10	身体表現性障害	0.265556	10	転換性障害	0.150046

正解データについては、専門書に従って設定し、下線を付与した。検索 e に対しては、正解データに直接的な上位概念と間接的な上位概念が、検索 f に対しては、正解データに直接的な下位概念と間接的な下位概念が含まれる。

検索 e の上位概念検索では、多重人格障害と直接関連のある「解離性障害」は 3 位検索結果が多少悪くなったが、その他に獲得したい「解離性障害」の上位概念である「臨床疾患」や「臨床疾患」の上位である「第 1 軸」が上位に検索され、有効な検索ができた。ここで、多重人格障害の直接の上位である「解離性障害」が実験 2-1 の結果と違い、1 位に検索されなかったのは、拡張の際の関連付けに原因があるため、間接的に関連のある事象をどのように検索語に関連付けるかが課題となっている。

検索 f の下位概念検索では実験 9 と同様良い検索結果が得られた。

4.4 実験 3

4.4.1 目的と実験環境

検索対象を文書に設定した場合、抽象度依存計量、因果関係計量での検索を行い、有効な文書が得られるかを検証する。文書データは精神医学分野におけるケースブック⁵⁾より抜粋した症例 8 件でそれぞれの文書より特徴語を抽出し、検索に使用する。特徴語群は表 10 の通りである。

本実験では 3.1 節で示した実現システムを用い、また、3.1 節で計量のために使用した 3 つの空間 M,Mc,Me と同じ基本データ、特徴語群を使用し、3 つの抽象度依存空間 Mm,Mu,Ml を新たに生成し、使用した。なお、対象とする空間は基本データ、特徴語ともに 51 語で 49 次

表 10 文書データ

	ドキュメントの主な特徴語群
doc1	嵐, 遭難, 心理的ストレス, 限局性健忘, 健忘期間中の困惑・失見当識, 解離性健忘
doc2	仕事のストレス, 失踪, 過去の記憶喪失, 不安, 解離性遁走
doc3	小児期の性的虐待, 身体的虐待, 子ども時代の記憶の空白, 自分の行動の記憶がない, 6つのパーソナリティ状態, 催眠にかかりやすい, 身体疾患の病歴, 解離性同一性障害
doc4	父の死, 抑うつ, 身体疾患, 二つの意識障害, 不安, 幻覚, 放心状態, 特定不能の解離性障害
doc5	幻覚, トランス状態, 民族精神病, 特定不能の解離性障害
doc6	事故, 生活のストレス, 情緒的ショック, 不安, 特定不能の解離性障害
doc7	結婚の失敗, 苦しい経済状態, 仕事上の問題, 行方不明, 健忘, 遁走状態
doc8	不安, 恐怖, 抑うつ, 父親の死, 身体的虐待, 健忘, 同一性の交代, 身体的虐待, 性的虐待, 感情的苦痛, 外傷後ストレス障害, 解離性同一性障害

元の空間となっている。

本実験で設定した検索 g,h は以下の通りである。

検索 g : 下位概念検索 : 外傷体験の下位概念の文書を検索する。

検索 h : 下位概念検索 : 解離性同一性障害の原因の文書を検索する。

4.4.2 結果と考察

検索 g,h の実験結果を表 11 に示す。

表 11 実験 3 検索結果 (下線部は正解データ)

外傷体験の下位概念のドキュメント	相関量	解離性同一性障害の原因のドキュメント	相関量
1 doc3	0.488085	1 doc3	0.615154
2 doc8	0.447914	2 doc4	0.205835
3 doc4	0.154988	3 doc8	0.202955
4 doc7	0.060029	4 doc6	0.031789
5 doc6	0.051428	5 doc1	0.027139
6 doc1	0.047489	6 doc7	0.022538
7 doc2	0.035930	7 doc2	0.017772
8 doc5	0.000000	8 doc5	0.008105

正解データについては、専門書に従って設定し、下線を付与した。検索 g に対しては、直接的、間接的な下位概念、検索 h に対しては、直接的、間接的な原因が含まれる文書を設定した。

実験の結果として、検索 g では外傷体験の直接的、間接的な下位概念を含んだ文書が上位に検索され、有効な結果が得られている。検索 h では解離性同一性障害の直接的、間接的な原因を含んだ文書も得られたが、特徴語群の解離性障害と関連付けた doc3,doc8 だけでなく、特定不能の解離性障害に分類されている doc4 も原因が関連しているため、上位に検索され、知識発見が広がったと言える。

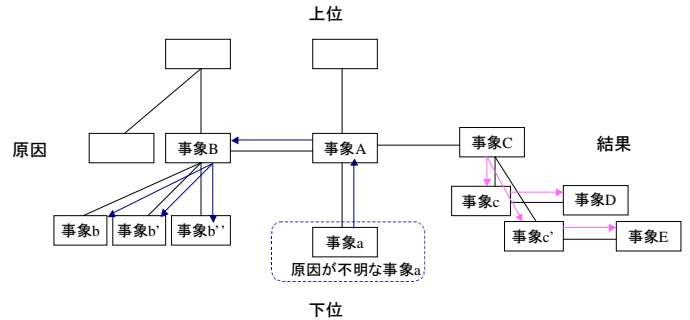
4.5 実験 4

4.5.1 目的と実験環境

本実験では実験 3 と同様に、3.1 節で示した実現システムを用い、また、3.1 節で計量のために使用した 3 つの空間 M,Mc,Me と同じ基本データ、特徴語群を使用し、3 つの抽象度依存空間 Mm,Mu,MI を新たに生成した。

本実験では原因・結果、上位下位の検索の連携で、有効な結果が得られるかを検証する。

検索の連携によって、原因が不明な事象の上位概念から原因を推測したり、情報が不十分な際に原因結果、上位下位を連携させることによって、ある事象に関連のある事象の可能性を示すことが出来る。実際の検索例を図 6 に示す。



検索例① 原因が不明な事象aの原因を上位概念より推測する。
 検索例② ある事象Cの具体例の結果として起こる障害。

図 6 上位下位概念検索、原因結果検索を連携した検索例

なお、対象とする空間は抽象度依存計量空間が、基本データ、特徴語ともに 51 語で 49 次元、因果関係計量空間が基本データ、特徴語ともに 51 語で 51 次元の空間となっている。

本実験で設定した検索 i,j は以下の通りである。

検索 i : 上位の原因の下位の検索 : 解離性同一性障害の上位の原因の具体的事象を検索する。

検索 j : 結果の上位の検索 : 外傷体験の具体のそれぞれの結果を検索する。

4.5.2 結果と考察

検索 i の実験結果を表 12 に示す。

表 12 検索 i の結果 (下線部は正解データ)

解離性同一性障害の上位概念の検索結果	相関量	解離性障害の原因として専門書で推測された事象
1 解離性障害	0.998534	小児期における外傷体験
2 解離性同一性障害	0.912813	てんかん
3 解離性遁走	0.436279	外的支援の欠如
4 解離性健忘	0.434547	身体的虐待
5 離人症性障害	0.434456	性的虐待
検索結果		近親相姦
解離性障害の原因		傷害
1 自己防衛	0.779070	死の目撃
2 心的外傷	0.532064	身近な親族などの死
3 解離性障害	0.526050	
4 外傷体験	0.356281	
5 防衛機制	0.309217	

正解データ

検索結果	相関量	検索結果	相関量	検索結果	相関量
自己防衛の下位概念		心的外傷の下位概念		外傷体験の下位概念	
1 抑圧	0.733212	心的外傷	1.000000	幼児虐待	0.524689
2 防衛機制	0.733212	アルコール乱用	0.000000	性的虐待	0.462315
3 否認	0.733212	境界性	0.000000	心理的虐待	0.438231
4 自己防衛	0.479731	シゾイド	0.000000	ネグレクト	0.438231
5 ネグレクト	0.000000	ネグレクト	0.000000	身体的虐待	0.438231

正解データについては、専門書に従って設定し、下線を付与した。検索 i の正解データには検索 a の原因として想定される事象が含まれる。

検索 i では、空間生成の際対象とした精神医学分野における知識体系⁴⁾に関して、原因不明とされている解離性障害の原因を上位概念より探る実験をした。実験結果は図 12 の通りである。本検索では解離性障害の上位概念として検索された結果の 1 位の事象を次の検索に使用し、その原因の上位 3 件についてさらに下位概念の検索を行った。また、同文献⁴⁾で専門家によって推測され、記載された原因を正解データとして採用し(図 12 上部右)、検証を行った。検索で得られた「外傷体験」の下位概念は正解データと合致しており、また「心的外傷」も、「自己防衛」とその下位も「外傷体験」の結果として関連付けられている事象で、十分に正解データとして推測される。つまり、実験の結果、上位概念の原因の下位概念の検索では、因果関係検索、抽象度関連検索を連携させることによって、原因の不明な解離性同一性障害の結果として推測される事象として有効な結果が得られた。本実験で示した方法では、実際的に次のように利用されることが有効であると考えられる。連携計量のプロセスであるそれぞれの上位概念検索、原因検索においては、病名、症状などの事象を対象として検索を行い、その後の下位検索において文書を対象とした本方式による検索を行った。

次に、検索 j の実験結果を表 13 に示す。

表 13 検索 j の結果(下線部は正解データ)

外傷体験の下位概念	相関量
1 幼児虐待	0.524689
2 性的虐待	0.462315
3 心理的虐待	0.438231
6 外的支援の…	0.347921
11 小児期にお…	0.252694
12 配偶者虐待	0.252614
20 戦争	0.047168

検索結果	相関量	検索結果	相関量	検索結果	相関量
外的支援の欠如の結果		小児期における外傷体験の結果		戦争の結果	
1 解離性同一性障害	0.997689	解離性同一性障害	0.997689	戦争	0.630675
2 外的支援の…	0.673674	小児期にお…	0.576688	解離性健忘	0.174106
3 小児期にお…	0.483827	てんかん	0.491172	心的葛藤	0.156328
4 てんかん	0.479962	外的支援の…	0.483827	離人症性障害	0.139424
5 脳腫瘍	0.189202	内分泌障害	0.157830	防衛機制	0.127388

正解データについては、専門書に従って設定し、下線を付与した。検索 j の正解データには検索 b の具体的事象の結果が含まれる。

検索 j では、外傷体験という検索語について、その具体的事象の結果を探った。外傷体験という事象は直接的、間接的に関連のある事象が多いため検索を連結する際の正解データは専門家などの手によって抽出され、次の検索にその結果を渡すことが出来る。(表 13 参照)。

5. ま と め

本稿では、精神医学分野データベースを対象とし、病名・症状の抽象度と因果関係性を反映した検索の実現方式について示した。本方式の特徴は、事象の上位下位概念や原因結果という特定の方向性を有した検索可能としたことである。これにより、検索者の検索目的を反映した事象やその文書データが検索可能となった。今後の課題として、本方式の適用領域についての考察、本方式の定量的・解析的評価、および、本方式によって得られる検索結果を検索者に分かりやすく表示する方式の実現が挙げられる。

謝辞 本論文の執筆にあたり数多くの貴重なご助言を頂いた、倉林修一氏、佐々木史織氏および鷹野孝典氏に感謝致します。

参 考 文 献

- 1) 図子泰三, 清木康, 鷹野孝典, 波内みさ, 但田育直: “事象データ間の因果関連性計量機能をともなったベクトル空間検索方式”, 情報処理学会論文誌: データベース, Vol.45, No. SIG 7 (TOD22), pp.124-136. (June, 2004).
- 2) 鷹野孝典, 図子泰三, 清木康: “事象間の因果関係を扱う動的な文脈解釈機能を有する意味的連想検索方式の実現”, 情報処理学会論文誌: データベース, Vol.46, No. SIG5 (TOD25), pp.40-55, (March, 2005).
- 3) 高橋三郎, 大野裕, 染矢俊幸訳: “DSM-IV-TR 精神疾患の分類と診断の手引き” 医学書院, 2002.
- 4) ベンジャミン J. サドック, パージニア A. サドック編, 井上令一, 四宮滋子監訳: “カプラン臨床精神医学テキスト 第 2 版 DSM-IV-TR 診断基準の臨床への展開” メディカル・サイエンス・インターナショナル, 2004.
- 5) 高橋三郎, 染矢俊幸訳: “DSM-IV-TR ケースブック” 医学書院, 2003.
- 6) Yasushi Kiyoki, Takashi Kitagawa, Takanari Hayama: “A Metadatabase System for Semantic Image Search by a Mathematical Model of Meaning”, ACM SIGMOD Record, Vol. 23, No. 3, pp. 34-41, September 1994.
- 7) 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.