

# キーワード集合の位相構造とその可視化

## -ケアアセスメント文書への応用-

坂本 将樹

川島 健志

江藤 香

樺澤 康夫

日本工業大学

### 1. はじめに

介護保険の導入によりケアプラン策定が制度化されるなどケアの質の向上が不可欠となっている。我々はケアマネジメントの質の向上を支援するシステムの開発を目指している。ケアの質の向上にはケアニーズを的確に読み取るためのノウハウ情報の有無は重要な要因であり、ノウハウ情報の抽出・共有が必要となる。我々はノウハウ情報の抽出を支援する機能の一つは熟達者と初心者との観念の違いを感知させるための柔軟な観念の変更であると考えている。本論では概念ベースの階層上を移動させることにより、この機能を実現した。観念の変更ごとに読み取り文書の距離を計算し、その結果を可視化し、さらに読み取った文書の KOMI チャートも表示することにより観念の違いを気づかせることを試みた。

### 2. 概念ベースの構築

概念ベースとは概念を自身とその他の概念集合で表した知識ベースである。本研究ではシソーラスを用いて 6 階層の木構造のケアの分野に特化した概念ベースを構築した<sup>[1]</sup>。

#### 2.1 概念の抽出

概念は KOMI チャートのアセスメント項目とその説明文と KOMI チャートを読み取った事例の文書から抽出したキーワードからなる約 4000 個である。KOMI チャートとはケアアセスメントの結果を認識面と行動面の 2 つの円形チャートで表現するケアアセスメント方式である<sup>[2]</sup>。

#### 2.2 概念のコード化

抽出したキーワードに類語国語辞典<sup>[3]</sup>に基いて数字を割り振り、コード化を行った。病名など一部存在しないキーワードは独自にコード化を行った<sup>[4]</sup>。そして、レベル 1 は大分類、レベル 2 は中分類、レベル 3 は小分類のように階層を定義し、レベル 4、5、6 は必要に応じて細分化を行った。

### 3. 距離の定義と計算

キーワード間の距離はキーワード間の概念の近さにより測定される。ここではシソーラスの親子関係(階層)より兄弟関係(同階層の隣接するキーワード間)の方が概念的に近いとする。また、等しい兄弟関係でも、階層が深いものほど近いものとする。文書間の距離はキーワード集合間の距離により測定される。

#### 3.1 キーワード間の距離

キーワード間の距離はシソーラスの親子関係と兄弟関係、それぞれノードの枝に重み付け(親子関係のレベルを 2、兄弟関係のレベルを 1.5)とし、一つのキーワードから他のキーワードに達するのに経由したノードの枝の重みの合計を距離とする<sup>[5]</sup>。

### 3.2 観念の変更

観念の変更とは視点の拡大・絞り込み易さや関連度を見ることである。観念を変更する毎に距離を計算し、その結果を視覚化することでその差分を感知させる。これにより熟達者と初心者の観念の違いの抽出を支援する。本論では概念ベースの階層上を移動させることによりこの機能を実現した。

### 3.3 文書間の距離尺度

文書からキーワードを抽出し、キーワード集合を構成し以下では、キーワード集合を文書とみなす。観念の概念ベース上のレベルを  $L$  とする。キーワード  $a, b$  の間の観念  $L$  における距離を  $d(a, b | L)$  と表す。

キーワード集合を  $A, B$  とする。 $\forall a \in A$  と集合  $B$  の間の観念  $L$  における距離を

$$D_w(a, B | L) = \min_b \{d(a, b | L)\} \quad (1)$$

と定義する。キーワード集合  $A$  からキーワード集合  $B$  への観念  $L$  における距離  $D(A \rightarrow B | L)$  を

$$D(A \rightarrow B | L) = \frac{1}{|A|} \sum_{a \in A} D_w(a, B | L) \quad (2)$$

と定義する。 $|A|$  は集合  $A$  の要素数を表す。次にキーワード集合  $A, B$  の間の観念  $L$  における距離  $D(A, B | L)$  を

$$D(A, B | L) = D(B, A | L) = \max\{D(A \rightarrow B | L), D(B \rightarrow A | L)\} \quad (3)$$

により定義する。

### 4. 二次元可視化

キーワード集合(文書)間の距離が数値として求められても、多数の文書間の差を直観的に見ることはむずかしい。各文書をパーソナルコンピュータの画面上の点として配置することにより視覚化することを考える。このため我々は多次元尺度構成法のひとつであるクルスカルの方法<sup>[6]</sup>を用いて、平面(二次元空間)上に配置することを試みた。以下に示す図 1 は先に述べた平面上にキーワード集合を配置させたものである。配置された点には数字がそえてあり点をクリックすると数字の事例に対応した KOMI チャート、サマリーチャート、レーダーチャートを表示できるようにした。これにより他の事例との比較が容易に行えるようになる。

観念のレベルは最大 6 までであり (a) ではレベル 6 まで見たときの図であり、丸い印で示す通り 6 の点と 8 の点に近い関係にあることがわかる。また (b) はレベル 5 まで見たときの図であり (a) では近い関係にあった 8 と 12 が離れ 6 と 12 が近づいていることがわかる。

### 5. おわりに

本研究では、アセスメントの結果を読み取った文書の距離を概念ベースを用いて計算した。観念の変更は概念ベー

ス上のレベルを上下させることにより実現した。観点の違いを気づかせるために、観点の変更ごとに、その結果を二次元上に可視化することができた。さらに読み取った文書の KOMI チャートも表示した。

今後は現場のユーザーによる試用評価に基づきノウハウ抽出の機能の改善を図りたい。

#### 参考文献

[1] 江藤香, 樺澤康夫, 松居辰則, 岡本敏雄: KOMI チャートを用いたケアプランの策定過程におけるノウハウ情報-概念ベースにおける観点の変更と文書間の距離尺度の関

係について-, 信学技報, ET2003-112, 2004

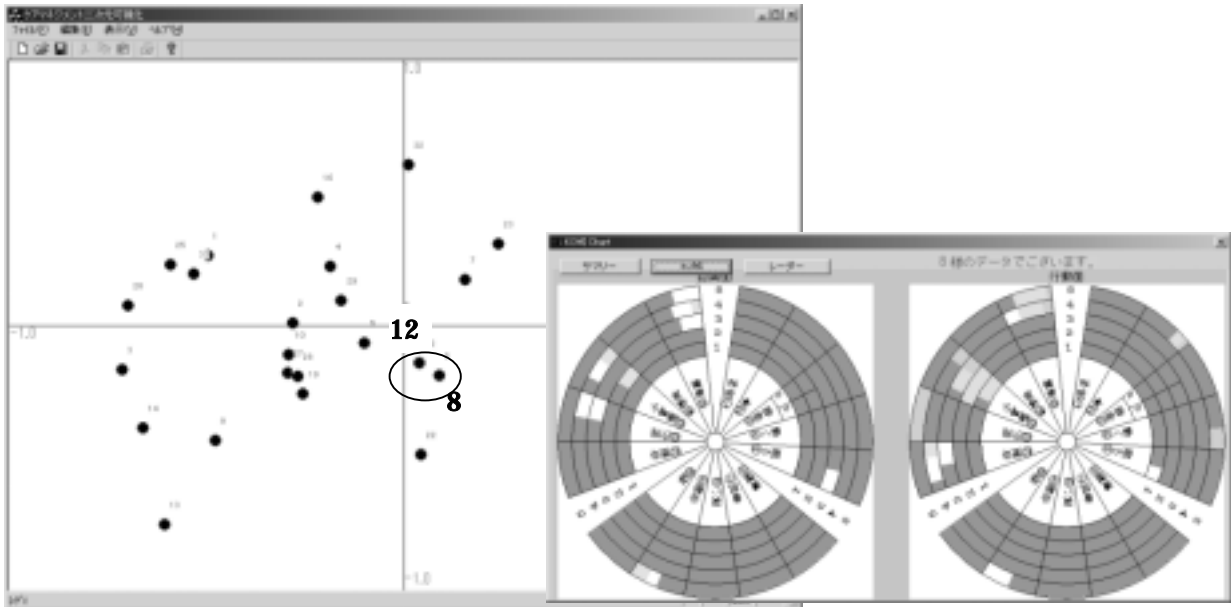
[2] 金井一薫編: KOMI チャートシステム・2000, 現代社, 1999

[3] 大野晋, 浜西正人: 類語国語辞典, 角川書店, 1985

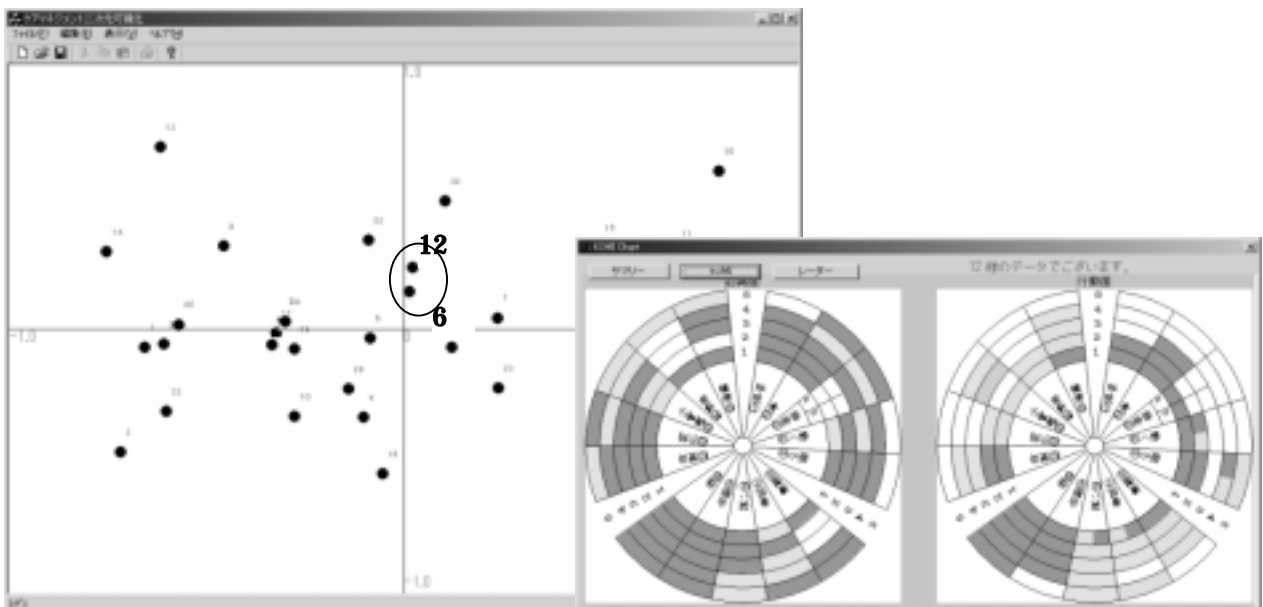
[4] 医学中央雑誌刊行会: 医学用語シソーラス, 2002

[5] 安部浩史, 岡村雄二, 江藤香, 樺澤康夫: 概念ベースを用いたケアアセスメント文書の分類, 情報処理学会第 66 回全国大会公演論文集 4ZB-2, 2004

[6] 岡太彬訓, 今泉忠: パソコン多次元尺度構成法, 共立出版, 1994



(a) レベル 6 までの二次元の可視化と KOMI チャートの例



(b) レベル 5 までの二次元の可視化と KOMI チャートの例

図 1、二次元可視化と KOMI チャートの例