

日本語コーパスを使用した文章完成テストの解析

5 B-8

近藤 弓末 西野 順二 小高 知宏 小倉 久和 福井大学工学部

1 はじめに

精神科の専門医が、患者に対して行うテストに文章完成テスト (Sentence Complete Test 以下 SCT) がある。実施が簡便であり、わかりやすい内容であるため、様々な臨床状況に柔軟に利用され、例えば、言語の認識に重大な障害があると推測される分裂病患者にも適応可能である。しかし、各医師が評価するため、その結果は主観的かつ定性的である。そのため従来の評価結果では、疾患の進行の程度の客観的な評価や、治療効果の測定など、症状変化の指標に用いるには問題があった。

我々は、SCT を定量的に評価することを目的とし、SCT の解析方法を検討した。その方法として、テキスト処理による方法と概念距離を使用した方法の2つの方法を行った。

2 テキスト処理による評価

2.1 評価方法

本手法では、日本語コーパスと SCT の文章の文字の並びを比較し、一致する割合により評価を行う。本論文では、文字の並びを n 文字パターンといい、 n は文字数を表わす。日本語コーパスと評価すべき SCT の文章を n 文字パターンで表層的に区切り、その並びを比較して、両者が完全に一致する場合に得点 1 を加え、一致しない場合は得点 0 とする。これを全ての n 文字パターンで計算し、その合計を評価すべき文章のスコアとする。そのため、あらかじめコーパス中の n 文字パターンの出現頻度を調べて文字パターンファイルを作成した。この文字パターンファイルには、コーパス中の、文字パターンとその出現頻度を頻度が多い順に収められている。以降、このファイルを使用して n 文字パターンのマッチングを行う。

本論文においては $n=3$ とし、コーパス中に出現した文字パターンと SCT 中に出現した文字パターンの一致した割合によって文章を評価した結果をまとめる。 $n \neq 3$ の場合について行った結果は考察で簡単に述べる。マッチングのとり方については、以下の2つの方法をそれぞれ行った。

第1の方法では、SCT 中に出現した文字パターンがコーパス中に何回出現しているかを数え上げることで評価を行う。SCT1 問中の文字パターン数を a 、コーパス

中の文字パターンとマッチした回数を N とし、評価値 $H1=N/a$ とする。これを SCT60 問 (1 セット) について調べていく。この方法を評価方法 1 とする。

第2の方法は、コーパス中に出現した文字パターンが、それぞれ SCT 60 問 (1 セット) 中にどれくらい出現しているか、その回数で SCT の評価を行った。この方法ではコーパス中に出現した文字パターン全てに対してではなく、コーパス中の出現回数上位 200 個についてマッチング回数を調べた。SCT60 問中の文字パターン数を m 、コーパス中の文字パターンが SCT60 問中に出現した回数を M とし、評価値 $H2=M/m$ とする。この方法を評価方法 2 とする。

2.2 SCT の評価

今回の研究では、入院中の患者 21 人・通院中の患者 2 人、健常者 12 人の SCT について評価値を求め、比較検討した。ただし、健常者データとは 21 歳~24 歳の大学生のものである。また比較の対象とした病状診断は、臨床現場における担当医師の総合的な判断に基づくものである。

(1) 評価方法 1 の結果

被験者を入院中の患者、通院中の患者、健常者の3つのグループで比較を行った場合は、入院中の患者の評価値は低かったが、通院中の患者と健常者の評価値に差がみられなかった。1 人の患者において時間を追ってとった SCT の評価の場合は、患者 16 人中 8 人について、臨床現場で判断された病状の変化と SCT の評価値の変化が対応していた。

(2) 評価方法 2

(1) と同様、被験者を入院中の患者、通院中の患者、健常者の3つのグループで比較を行った。全体的に見ると、入院中の患者より通院中の患者、通院中の患者より健常者の評価値が高い傾向が見られる。しかし、中には、健常者であっても入院中の患者や通院中の患者の評価値よりも低いものもある。また、1 人の患者で時間をおいて病状と評価値の対応を比較した場合は、患者 16 人中 10 人が、病状の悪い方が評価値が低く、臨床現場で判断された病状と対応していた。全体的に見てみると、病状に変化があった場合は、評価値に大きな差がみられ、病状に変化がない場合には、評価値の変化が小さい。しかし、中には病状に変化がないのに、評価値に大きな差がみられたり、病状の悪い時の評価値が病状の良い時の評価値より高くなっているものもある。図 1 は、中でも病状の良し悪しによって評価値に差が見られたもので、左が病

状の悪い時、右が病状の良い時である。縦軸が各文字パターンとのマッチング回数、横軸がコーパスの中で出現回数の多かった文字パターンが出現回数の多い順に並んでおり、数字はその順位を表わしている。

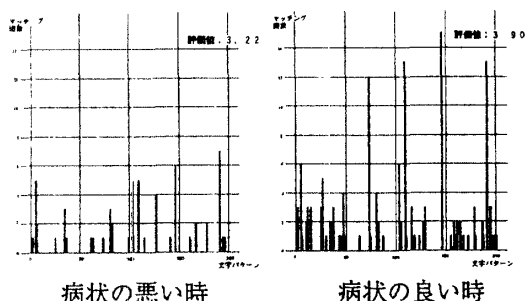


図 1: 評価結果

3 概念距離を用いた評価

3.1 概念距離

概念距離は、EDR 電子化辞書の概念体系辞書を使用して求めた [1]。単語とは、名詞・動詞・形容詞・形容動詞の自立語を指す。

単語と単語の間概念距離を調べることにより、単語と単語の意味的なつながりを見るのが可能である。SCT では文頭が与えられているために、被験者はその文頭に合うように後続の文章を続ける。そのため、その文頭に関係が深い単語がでてくる。例えば、「子供の頃、私は」という文頭に対しては、「遊ぶ」とか「やさしい」「おとなしい」等の状態を示す単語が現れやすい。そのため、概念間の距離は自然とよく似た値になると推測される。しかし、患者においてはその文頭に合わない文章を書くため、概念距離は概念距離が大きくなると思われる。これらのことより、概念距離を求めることにより、SCT を評価することが可能であると推測される。

3.2 評価方法

評価方法としては、SCT の各文章を形態素に切り分け、概念距離を測定し、その平均をとる。概念距離は以下の 3 つの組み合わせで測定を行った。

- 各問題の先頭の単語と以下に現れる単語との距離を測り、SCT60 問でその平均をとる。
- 隣合った単語同士の距離を測り、SCT60 問でその平均をとる。
- 各問題ごとの文章に現れる単語の全ての組み合わせで距離を測り、SCT60 問でその平均をとる。

例えば、「子供の頃、私は公園でよく野球をした。」という文章があったとする。第 1 の方法では、先頭の子供という単語と頃・私といった後続の単語との距離を求め、平均をとる。第 2 の方法では、子供と頃、頃と私といったように、隣合った単語同士の距離を求め、平均をとる。第 3 の方法では、子供・頃・私・公園といったようにに

章中に現れる全ての単語の全ての組み合わせで距離を測り、平均をとる。

3.3 結果

3.2 で述べた方法で求めた結果の中から、第 1 の方法で求めた結果の一部を表 1 に示す。表 1 で示されている被験者番号は被験者の SCT につけた番号で、同じ番号で (a)(b) と分かれているものは、同一の患者で時間を追って行われた SCT の結果である。また表の左側が患者、右側が健常者の結果である。

表 1: 概念距離を用いた評価値の例

被験者番号	評価値	被験者番号	評価値
1-1	0.0	2-1	16.0
1-2	20.7	2-2	18.4
1-3	16.3	2-3	14.8
1-4	25.1	2-4	16.4
1-5	15.5	2-5	20.5
1-6	18.6	2-6	18.2
1-7	19.1	2-7	16.9
1-8(a)	20.1	2-8	15.3
1-8(b)	16.3	2-9	17.4
1-9(a)	12.4	2-10	17.3
1-9(b)	19.3	2-11	19.1
1-10(a)	24.4	2-12	16.1
1-10(b)	22.3		

4 考察と今後の課題

テキスト処理を用いた評価において、 n 文字パターンのマッチングにより SCT を評価したが、1 人の患者で病状の善し悪しを比較した結果、患者 16 人中 10 人については臨床現場における医師の判断と一致していた。また、評価方法 2 においては、重症の患者の評価値がかなり低くなることもわかった。

今回は特に $n=3$ として行った実験結果について述べて来たが、 $n=2$ の場合や $n=4$ 、 $n=5$ の場合についても同様に評価して比較も行った。その結果をみると、 $n=2$ の場合には評価値と病状の対応をみることはできなかった。また、 $n=4$ 、 $n=5$ に増やして評価結果と病状との対応を見てみたが、 $n=3$ の結果と精度に変化はなかった。これらのことより、 $n=4,5$ より簡便な $n=3$ の出現頻度を用いて比較することが効率的であると考えられる。

一方概念間の距離を用いた評価では、あまり良い結果は得られなかった。概念距離が単語同士の意味のつながりを見ることは可能である。しかし、日本語の文章においては単語同士の意味的なつながりよりも、その間をつないでいる助詞が重要な役割を果たしているためであると推測される。

現在、より評価の精度を上げるために、形態素解析による新しい評価方法に取り組んでいる。

参考文献

- [1] 永松, 田中: 文章に対する重要度付与処理における視点情報の有効性の分析, 情報処理学会第 51 回全国大会, 7H-6, pp3-119 - 3-120(1995)