

医師国家試験を自動解答するプログラムの構築

Construction of a question-answering program that automatically answers the medical licensing examination

伊藤 詩乃[†] 田中 佑岳[†] 佐藤 健吾[†] 洪 繁[†] 狩野 芳伸[‡] 榊原 康文[†]

Shino Ito Yugaku Tanaka Kengo Sato Shigeru Ko Yoshinobu Kano Yasubumi Sakakibara

1. はじめに

近年、電子カルテなどの診療データ、健康診断データ、服薬履歴データ（お薬手帳）、などに代表される医療健康データの電子化が著しい。とくに電子カルテは平成11年に厚生労働省が認可して以降、大規模病院を中心に利用率の高まりを見せている。平成25年には、全病院の31%、大規模病院の69.9%に導入されたという統計値もある。患者の問診記録、検査結果などを病院間で共有することで、患者の過去の疾病歴を調べることが容易となるだけでなく、睡眠・食生活などの患者の生活習慣の電子記録を合わせて考えることによって将来的に一人一人の患者に応じた医療を提供することが可能となる。また、今まで得ることができなかった様々な種類の情報を得ることによって、病気とある特徴量との間に相関を見つけることができれば、新たな医学的知見を得ることも可能となる。例えば、電子カルテから意思決定やデータマイニングのために情報を抽出するシステムの開発 [1]、日本語の医療文章で省略表現された単語を補完する手法の構築 [2]、ケースベース推論を用いた肝臓癌の再発予測 [3]、などの研究がすでに行われている。しかし、現状では、電子化された医療データに対して、それを利用するための統一された知的情報システムはいまだ模索段階であり、医師にとって使いやすく有用な医療情報システムの開発が急務となっている。

医療情報を用いた応用システムの一つとして、電子カルテに記述された内容から、患者の情報を適切に読み取り、そこから考えられる病名候補を予測するシステムの構築が考えられる。患者の状態を確認し、診断を与えるという作業は、専門性を必要とする難易度の高い経験の要る仕事である。したがって、プログラムによる完璧な診断を行うことは困難であると考えられるが、医師が最新の知識や情報にアクセスすることがより容易になり、見落としを防止できるなど医師の補助には有用であると考えられる。このような診断支援システムを実現する方法として、図1に示すように、ルールベース推論とケースベース（事例ベース）推論の二つの手法が主に考えられる。ルールベース推論は、医学的な専門知識を文献情報や専門家から取り出し、診断ルールとして明示的に実装し、そのルールを患者の症状に適用して病名を判定する方法である。一方、ケースベース推論は、過去に蓄積された電子カルテなどのデータベースから、患者の症状に類似の症例を検索して、その病名を出力する方法である。本論では、医学教科書 [4, 5] などから診断ルールを取り出して実装す

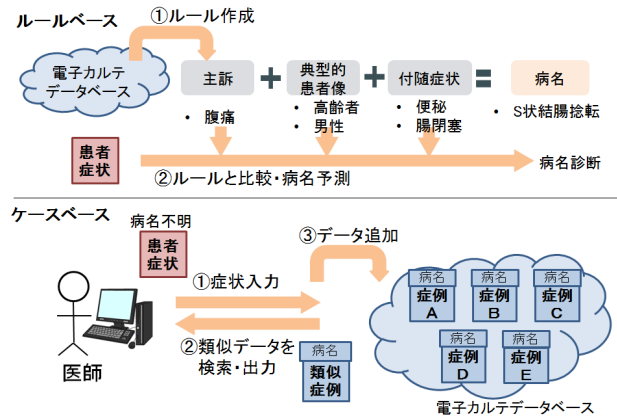


図1: 蓄積されたデータに基づく診断支援システム

ることにより、ルールベース推論を用いることとした。病名診断支援システムの構築を最終目標として、そのパイロット研究として、本研究では医師国家試験を自動解答するプログラムの構築を目的とする。医師国家試験問題は、多種の選択式の問題があるが、その中でも問題文として患者の情報や検査結果が与えられ、選択肢の中から適切な病名を解答する形式のものがあり、このタイプの問題を解答することを目指す。また、画像から読み取ることができる情報を抽出するという課題もあるが、これについては本論では取り扱わない。

医師国家試験を解答することによって、医療データの情報処理における課題とその対処法について考察し、最終的には患者情報を用いて病名診断支援を行うための基礎を築くことを目的とする。本論文では、実際に医師国家試験の一部の問題について解答を行い、正答できた問題、誤答となった問題について考察を行うことで、より精度のよいプログラムを構築するための改善点を明らかにする。

2. 背景

2.1. 医師国家試験

医師国家試験は、A問題からI問題まで9つのセッションに分かれ、全500問を解答するものである。医師国家試験問題の過去問題は、厚生労働省のホームページ [6] で公開されている。医師国家試験問題のうち、患者の病歴や症状などが問題文で与えられ、医学的な知識を問う設問（臨床実施問題と呼ばれる）は、医療の現場で用いられる病歴報告（第三者が患者の症例を把握することを目的として作成される文書）に近いものである [7]。

[†]慶應義塾大学, Keio University[‡]静岡大学, Shizuoka University

2.2. 試験問題に対する自動解答プログラム研究

類似の先行研究として、国立情報学研究所が中心となって行われている「ロボットは東大に入れるか」プロジェクト [8] がある。東ロボプロジェクトでは、大学入試センター試験の自動解答プログラムの構築を行っている [9]。医師国家試験においては、医療文書特有の処理が必要である点が大きく異なる。センター試験の各科目のうち、医師国家試験にもっとも性質の近い科目は社会科の問題であるが、仮に社会科の解答器をそのまま適用したとしても、医師国家試験では高得点が期待できない。たとえば、もっとも基本的な処理である形態素の分割に失敗すると、それをうけた後段のあらゆる処理が失敗してしまうが、病名や症状名などは医療文書特有であり既存のツールでは十分な性能で形態素分割ができない。また、医療においては年齢や血圧など、数値の抽出が必要であるうえ、その範囲が文脈上適切であるかを判断しなければならない。こういった医療文書特有の処理を行ってはいじめて、医師の補助を行えるような実用的なシステム構築に資する要素技術になる。

3. 方法

3.1. 問題解答プログラムの構成

問題解答プログラムの概要を以下に示す。

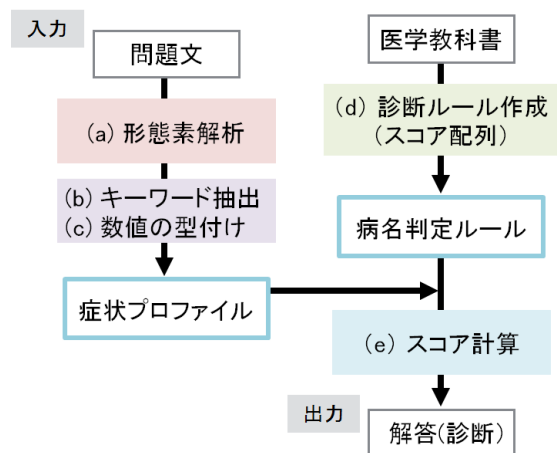


図 2: 問題解答プログラムの構成

図 2 に示した通り、問題解答プログラムは、(a) 問題文のテキストファイルを読み込み自然言語解析を行う関数、(b) その結果をもとに血圧や体温などの患者の属性を示す数値や腹痛などの症状を表す単語を抽出する関数、(c) 数値を属性分け (型付け) する関数、(d) 病名判定 (診断) ルール (スコア配列, 3.4 節) を構築する関数、(e) 患者の症状を表す単語・数値に病名判定ルールを適用することでスコアを計算して病名判定を行う関数に分かれている。以下の節では、その各関数について詳述する。

3.2. 形態素解析

解答対象となる問題文をテキスト形式で入力し、mecab[10] により形態素解析を行い、問題文を形態素に

分解した。医療分野の文書を扱うため、形態素解析を行う際にオプションとして、ライフサイエンス辞書、標準病名マスター、などをユーザー辞書として検討した。最終的に、ComeJisyoV5-1[11] の dic ファイルをユーザー辞書として設定した。これにより ComeJisyoV5 に登録された語彙と一致する単語が現れた場合、ComeJisyoV5 に記載された品詞と同一になるように優先的に解析が行われ、次に標準辞書である IPA 辞書が参照される。これにより、例えば「脳梗塞」のような医学専門用語が、「脳」と「梗塞」に分割されることなく、抽出されるようになった。

3.3. 診断キーワードの抽出

形態素解析を行った結果について、「名詞」である形態素については、これを症状プロファイルというリストに格納した。

また「数」である形態素について、その数値がどの「型」の特徴量 (年齢や血圧など) を表すのかを判別するために、あらかじめ判別ルールを決め、それに基づいて数値の属性解析を行った。判別ルールは以下に基づく。

- 特定の単位の形態素が直後に並ぶ場合、その直前の数の形態素は直後の形態素の属性をもつ。例えば“27/歳”という形態素の並びであれば、27 という数の形態素は年齢という属性をもつことがわかる。
- 「数」の形態素の直前の名詞の形態素を参照する。例えば、“血圧/ 140/90 /mmHg”という並びであれば、140, 90 といった数値は直前の名詞形態素を参照することで血圧についての数値であることがわかる。さらに、血圧のような属性の場合、高い数値を最高血圧、低い数値を最低血圧といったように判別をした。

3.4. 病名判定ルールの作成:スコア配列

病名判定を行うため、各病気について、その病気の症状、検査結果の数値などを表の形式でまとめたものをスコア配列と呼ぶ。医学教科書 [4, 5] などを用いて、各病気に対する症状・特徴などの単語と数値 (特徴量) をまとめてスコア配列を作成した。例えば、「医師国家試験のためのレビューブック 第 10 版」[12] には、頻度の高いまたは重要な病気に対して鑑別するための 1st impression が載せられている。このような病名判定ルールは、図 3(a) のように、主訴、典型的患者像、付随症状についてまとめることができる。このような形式で記載されている単語や数値を症状として、スコア配列に追加し、今回は単純に 1 というスコアをつけた。その病名では、特徴的ではない症状に関しては 0 のスコアをつけた。このようにして図 3(b) のようなスコア配列を作成した。

“338 病名”, “592 症状” について、スコア配列を構築した。

3.5. 病名判定ルールによるスコアの計算

作成したスコア配列の病名一つ一つに対して、患者の症状からスコアを計算し、最もスコアが高かったも

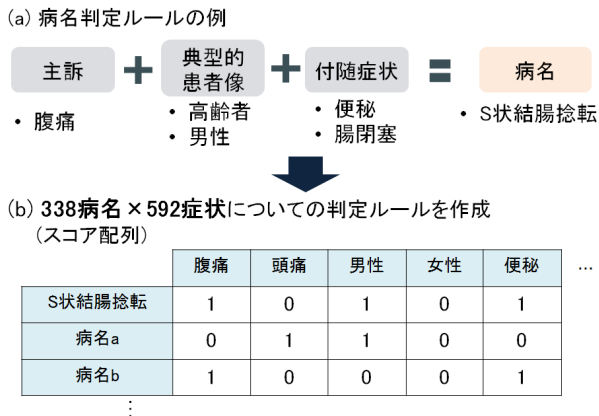


図3: 病名判定ルールに対するスコア配列の作成

のを解答として表示する。

3.5.1. 症状プロファイルとの一致

問題文の文章を形態素解析し作成した症状プロファイルの要素である症状を表す単語と、スコア配列に記述した症状の単語とが一致しているかを判定した。一致していた場合、解答候補である病名のその症状に付与されているスコアを加算した。

3.5.2. 数値の評価

問題文中に記述された数値について、その数値がどのような型を持つかを判定する。例として、問題文から「血圧:141/91 mmHg」であることを読み取ったとすると、最高血圧(収縮期血圧)が140以上150以下であり、最低血圧(拡張期血圧)が90以上99以下であるため、これはI度高血圧に分類される。このように数値からクラスに分類し、病名判定ルールのスコアを参照し、各病名のスコアに加算した。数値を評価し、クラス分類をする関数は下記のとおりである:

- ・年齢の数値を「年代」として分類し、スコアを計算する。
- ・身長、体重の数値からBMIを計算し、体型を判断する。
- ・体温の数値から、発熱しているか調べる。
- ・血圧の数値からクラス分類する。

たとえば、75歳という数値を文章中から読み取ったとすると、75歳という年齢そのものは病名判定ルールのスコア配列には載っていない。病名判定ルールでは、40代、20代のように、年代でだまかに区切りを入れている。75歳という年齢の情報から、75という数値が年代のどこに当てはまるかを判定し、それに従って“70代以上”のクラスのスコアを参照し足し合わせる(図4)。

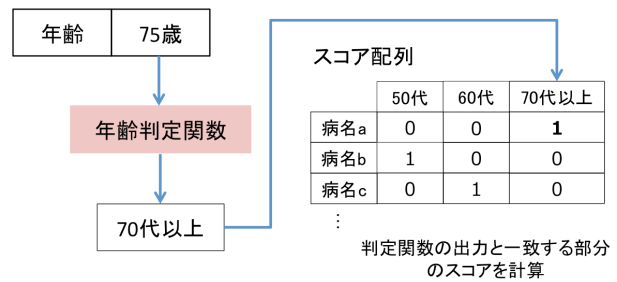


図4: 年齢の数値をクラス分類する年齢判定関数

3.5.3. 病名判定ルールの合計スコアの計算

解答となる5つの選択肢に対し、以上のような手順で各病名のスコアを計算し、スコアが最高となる病名を解答として選択した。複数の病気で同じスコアとなった場合は、そのどちらも表示させる。これらの選択肢について、プログラムが出力する解答と問題の正解を比較し、結果を評価した。

3.6. 解答対象

平成25年に行われた第107回・平成26年に行われた第108回の医師国家試験問題のうち、以下の条件に当てはまる問題を用いた。

- (1) 臨床実施問題のうち、解答として問われる知識が病名となるもの。
 - (2) 検査結果などに画像が使われていないもの。
- このような条件に当てはまる問題は全部で27問存在し、本研究ではこれらの問題に解答した。

4. 結果

第107回・第108回医師国家試験のうち、合計27題の問題について解答を行った。27題のうち26題は、5つの選択肢から1つを解答として選ぶ問題であり、残り1題に関しては2病名を解答すべき問題である。

解答結果を表1にまとめる。27題の問題のうち、正答のみを選ぶことができた問題が8題(うち1題は2病名解答問題)、解答として表示した2病名の中に正答が含まれていた問題が5病名あり、スコアがついたものの誤っていた問題が6題、スコアが付かず無解答の問題が8題あった。

表1: 第107回・108回医師国家試験問題解答結果

| | |
|-------------------------|-----|
| (a) 解答を1病名選択し、1病名正答した問題 | 8題 |
| (b) 解答を2病名選択し、2病名正答した問題 | 1題 |
| (c) 解答を2病名選択し、1病名正答した問題 | 5題 |
| (d) 誤った解答のみを選択した問題 | 6題 |
| (e) 解答を選択できなかった問題 | 7題 |
| 計 | 27題 |

問題の解答としてプログラムが複数解答した中に正答が含まれており、正答が1つである場合(表1(c))、

その問題の解答が複数解答のなかからランダムにひとつ出力されると考え、その割合を加算した。すなわち、表1(c)の5題に関しては、半分の確率で正答すると考えることとする。プログラムの正答率は下記のようになった：

$$\frac{(8+1)+5*\frac{1}{2}}{27} * 100 = 42.6\%$$

この問題をランダムに解答したときの正答率を示す。26題は5個の選択肢のうち1つが解答であるので、正答率は $\frac{1}{5} = 0.2$ (20%)、残りの1題は5題中2題選択であり、その正答率は10%となる。これらの重みを考慮し、ランダム正答率は

$$\frac{26*0.2+1*0.1}{27} * 100 = 19.6\%$$

と計算され、本プログラムはランダム正答率よりはよい結果を得られた。

一方で、実際の医師国家試験では、臨床問題の合格基準は第107回において71.5%となっており、これには届かない結果となった。

5. 考察

本節では、正答できた問題と正答できなかった問題について詳細を見ることにより、現状と今後の課題を述べる。とくに、正答できなかった問題の中で、診断ルールのスコアを改善すれば正答可能な問題と、症状の単語だけからでは正答が不可能で意味解析まで必要とされる問題を取り上げる。

5.1. 正答できた問題例

第108回問題 H34

問題

『42歳の女性。頭痛を主訴に来院した。現病歴：3年前から月1、2回の頭痛を自覚していた。頭痛は両側の後頭部を中心とした持続性の鈍痛で、肩こりを伴っている。夕方になると頭痛が悪化するが、生活に支障をきたすほどの痛みではない。家事などで体を動かしている時のほうが多少痛みは和らぐ感じがする。悪心はない。頭痛に大きな変化はないが、昨日知人が脳梗塞で入院したという話を聞き、心配になって受診した。既往歴：スギ花粉症。生活歴：事務職。喫煙歴はない。飲酒は機会飲酒。家族歴：父親が56歳時にくも膜下出血で死亡。母親が高血圧症で治療中。現症：意識は清明。身長158cm、体重52kg。体温36.3℃。脈拍76/分、整。血圧118/72mmHg。神経学的所見に異常を認めない。現時点で最も考えられるのはどれか。』

『選択肢』

- 片頭痛
- 群発頭痛
- 緊張型頭痛
- 三叉神経痛
- 側頭動脈炎

解答

この問題の正解と、問題解答プログラムの結果を表2と表3に示す。

表2: 第108回問題 H34の解答

| | |
|----------|-------|
| 正解 | 緊張型頭痛 |
| プログラムの解答 | 緊張型頭痛 |
| 最大スコア | 3 |
| 解答数 | 1 |

表3: 第108回問題 H34のスコア

| 選択肢 | スコア | 一致した症状を表す単語 |
|-------|-----|-------------|
| 片頭痛 | 2 | 頭痛, 女性 |
| 群発頭痛 | 1 | 頭痛 |
| 緊張型頭痛 | 3 | 頭痛, 40代, 女性 |
| 三叉神経痛 | 0 | - |
| 側頭動脈炎 | 2 | 頭痛, 女性 |

表3を参照すると、中年女性であり、頭痛が主訴という患者像から正答を選ぶことができていたとわかる。一方で、文章中の「持続性の鈍痛」、「体を動かしているときのほうが痛みは和らぐ」などは、緊張型頭痛の特徴的な所見である。これらの特徴は文章になっており、形態素解析の際に分割されてしまい、キーワードとして抽出することが困難であった。

第107回問題 B22

問題

『顔面神経麻痺、耳介周囲の皮疹、めまい及び難聴を特徴とする疾患はどれか。』

『選択肢』

- 耳下腺癌
- Bell 麻痺
- 小脳橋角部腫瘍
- 真珠腫性中耳炎
- Ramsay Hunt 症候群

解答

表4: 第107回問題 B22の解答

| | |
|----------|-----------------|
| 正解 | Ramsay Hunt 症候群 |
| プログラムの解答 | Ramsay Hunt 症候群 |
| 最大スコア | 4 |
| 解答数 | 1 |

この問題の実際の解答と、プログラムによる解答結果を表4と表5に示した。本問題は、顔面神経麻痺が現れる病気の中から、最も特徴に沿う問題を選択するものであるが、問題文章が短く、症状が単語として端的に示されているため、症状が正しく抽出できた。皮疹という症状は正答の Ramsay Hunt 症候群によく見られるものであり、これにより解答を一つに絞り込むことができた。

5.2. 正答できなかった問題例

第108回問題 A28

問題

『70歳の女性。左上腹部痛を主訴に来院した。昨夜、久しぶりに孫たちと遊んだり歌ったりして騒いだ。その3時間後から左上腹部に痛みを感じるようになった。診察室には前かがみの姿勢で入ってきた。食事摂取は良好であり、悪心や嘔吐はなく便秘も正常である。3年前

表5: 第107回問題 B22 のスコア

| 選択肢 | スコア | 症状を表す単語 |
|--------------------|-----|------------------------|
| 耳下腺癌 | 1 | 顔面神経麻痺 |
| Bell 麻痺 | 1 | 顔面神経麻痺 |
| 小脳橋角部腫瘍 | 3 | めまい, 難聴, 顔面神経麻痺 |
| 真珠腫性中耳炎 | 3 | めまい, 皮疹, 顔面神経麻痺 |
| Ramsay Hunt 症候群 | 4 | めまい, 難聴, 皮疹, 顔面神経麻痺 |

に脳梗塞を発症し, その後アスピリンを内服している。体温 36.5 °C. 脈拍 88/分, 整, 血圧 140/90 mmHg. 左上腹部に局限した圧痛を認めるが, 反跳痛はない。腹筋を緊張させると疼痛と圧痛とは増強する。腸雑音は正常である。』

『選択肢』

- 急性膵炎
- 腹壁血腫
- 腸腰筋膿瘍
- 虚血性大腸炎
- 穿孔性胃潰瘍

解答

表6: 第108回問題 A28 の解答

| 正解 | 腹壁血腫 |
|----------|----------------|
| プログラムの解答 | 急性膵炎 虚血性大腸炎 |
| 最大スコア | 2 |
| 解答数 | 2 |

表7: 第108回問題 A28 のスコア

| 選択肢 | スコア | 症状を表す単語 |
|--------|-----|---------|
| 急性膵炎 | 2 | 嘔吐, 悪心 |
| 腹壁血腫 | 0 | - |
| 腸腰筋膿瘍 | 0 | - |
| 虚血性大腸炎 | 2 | 女性, 70代 |
| 穿孔性胃潰瘍 | 1 | 嘔吐 |

この問題の実際の解答と, プログラムによる解答結果を表6と表7に示した。

この問題を解答する際には, Carnett 兆候(腹壁の緊張によって痛みが増強すること)が陽性であることを読み取り, ここから腹痛が腹壁由来であるということとを区別することが必要になる[13]。この問題では問題文最後の「腹筋を緊張させると疼痛と圧痛は増強する」という文章から Carnett 兆候を読み取る必要があったといえる。プログラムの解答は, 「急性膵炎」と「虚血性大腸炎」を選択した。「a. 急性膵炎」については, 表7から, スコアが加算された症状として「嘔吐」と「悪心」が挙げられている。問題文を読むと「悪心や嘔吐はない」と記述されているため, 本来否定的な表現であるが, 本プログラムは単語のみを抽出しているために否定表現を捉えることができていない。そのため誤ってスコアを上げる結果となってしまった。

これに対処するためには, Cabocha[14]などを用いて構文解析を行い, 否定表現がどこにかかっているか

という関係性を抽出していく必要がある。また, 「d. 虚血性大腸炎」については, 同様に表7を参照すると, スコアとして加算された症状は「女性」「70代以上」という文章中の患者像を抽出して加算しているが, 確かに虚血性大腸炎は60代以上の高齢者に多い病気であり, また女性に多いとする文献があるため[15], 誤ったスコア加算を行ってはいない。この問題に対しては, スコアについての工夫を行っていく必要がある。スコアについては次節でも議論を行う。

5.3. 正答できたが改良が必要な問題例

第107回問題 I65

問題

『58歳の女性。時々記憶がなくなることを主訴に夫に伴われて来院した。数年前から数秒間口をもぐもぐさせることがあり, 夫は気になっていたが本人は全く気付いていなかったという。昨日, 娘と買い物に出かけた際に, 娘が話しかけても数分間返事をしないことがあった。受診時の意識は清明, 身長 158 cm, 体重 52 kg, 血圧 130/76 mmHg. 神経学的診察で異常を認めない。自分では普通だと思うのですが, 夫と娘が私に物忘れがあると言うんですよという。受診日に行った頭部単純 MRI で異常所見を認めない。最も考えられるのはどれか。』

『選択肢』

- 不随意運動
- 逆行性健忘
- 解離性障害
- 一過性全健忘
- 複雑部分発作

この問題の正解と, プログラムによる解答結果を表8と表9に示した。

解答

表8: 第107回問題 I65 の解答

| 正解 | 複雑部分発作 |
|----------|--------|
| プログラムの解答 | 複雑部分発作 |
| 最大スコア | 2 |
| 解答数 | 1 |

表9: 第107回問題 I65 のスコア

| 選択肢 | スコア | 症状を表す単語 |
|--------|-----|-----------|
| 不随意運動 | 0 | - |
| 逆行性健忘 | 1 | 物忘れ |
| 解離性障害 | 1 | 物忘れ |
| 一過性全健忘 | 1 | 物忘れ |
| 複雑部分発作 | 2 | もぐもぐ, 物忘れ |

本問題は, 解答の選択肢中であれば表9のように正答を選択することができたが, 一方で全病名の中でスコアが最上位のものを選択しようとする, 同様にスコアが2となる病名が複数現れ, 解答を一つに絞り込むことができなかつた(表10)。

これらの病名の症状を表す単語を見ると, 正答以外の病名では「女性」「50代」など他の問題においても一般的に良く出てくる単語からスコアが加算されている

表 10: 全病名中でのスコアが最上位となった病名

| 選択肢 | スコア | 症状を表す単語 |
|---------|-----|-----------|
| 急性緑内障発作 | 2 | 50代, 女性 |
| 側頭動脈炎 | 2 | 女性, 50代 |
| 緊張型頭痛 | 2 | 50代, 女性 |
| 骨粗鬆症 | 2 | 50代, 女性 |
| 関節リウマチ | 2 | 50代, 女性 |
| 複雑部分発作 | 2 | もぐもぐ, 物忘れ |

ことがわかる。一方、正答である「複雑部分発作」については、「物忘れ」「口をもぐもぐさせる動作（[もぐもぐ]と表記）」など、他の病気においては当てはまるものが少ない特徴的な単語を取ってスコアが付いていることがわかる。病名を判定する際に特徴的な症状（他の病気に該当するものが少ない症状）のスコアを高く付けて、多くの病気で出てくる一般的な単語のスコアを下げることで、問題の状況とかけ離れた病名を表示することを避け、より正確に解答することができると思われる。

6. おわりに

医師国家試験問題について、問題文に患者情報が示され、最も考えられる病名を解答する形式である、臨床実施問題について解答するプログラムを構築した。各病名における特徴的な症状・状況を配列としてまとめ、338病名に対して592症状のスコア配列を作成した。また、問題文章を形態素解析し、文章中から症状を表す単語を抽出するとともに、検査結果等の数値について何の特徴を表しているかを判定することで患者情報を抽出した。これにより計算したスコアから第107回・第108回医師国家試験問題について解答を行った。問題解答プログラムが正答と誤答の複数解答を同時に選んだ場合、ランダムにどちらかを出力するとしたとき、正答率は42.6%であった。

また、正答・誤答の結果の考察から、かかり受け解析の必要性、否定表現の抽出、スコアの検討など、電子化された医療情報の利用技術を確認する上での課題点を見つけることができた。今後は上記課題点に対応した自然言語処理の手法を用いてより精度のよい情報抽出を行うとともに、さらなる病名・症状データベースの構築、機械学習手法を用いたスコアの自動調整を行っていく予定である。

謝辞

本研究は、文部科学省「革新的イノベーション創出プログラム (COI STREAM)」, 慶應義塾大学 COI トライアル拠点のプロジェクトの一環として行われた。また、科学研究費補助金新学術領域研究「ゲノム科学の総合的推進に向けた大規模ゲノム情報生産・高度情報解析支援」の支援を得ている。

参考文献

- [1] Meystre, S.M., et al.: Extracting information from textual documents in the electronic health record: a review of recent research, *Yearb Med Inform*, Vol.35, pp.128-44, 2008.
- [2] Shinohara, E., et al.: An easily implemented method for abbreviation expansion for the medical domain in Japanese text: A preliminary study, *Methods of Information in Medicine*, Vol.52, No.1, pp.51-61, 2013.
- [3] Ping, O., et al.: A multiple measurements case-based reasoning method for predicting recurrent status of liver cancer patients, *Computers in Industry*, Vol.69, 12-21, 2015.
- [4] 矢崎義雄編：内科学 第10版, 朝倉書店, 2013.
- [5] 加藤治文：標準外科学 第13版, 医学書院, 2013.
- [6] 厚生労働省：第108回医師国家試験の問題および正答について, (http://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryuu/iryuu/topics/tp140512-01.html).
- [7] 荒牧英治, 四方朱子, 島本裕美子, 久保圭, 宮部真衣, 大熊智子, 狩野芳伸, 森田瑞樹：日本語カルテをアノテートする, 言語処理学会 第20回年次大会 発表論文集, 2014.
- [8] 国立情報学研究所：ロボットは東大に入れるか, (<http://21robot.org/>).
- [9] 狩野芳伸：統合研究基盤：質問応答システムの互換コンポーネント化による再利用性向上と開発自動化支援, *人工知能学会誌*, Vol.27, No.5, pp.492-495, 2012.
- [10] 工藤拓, 山本薫, 松本裕治：Conditional random fields を用いた日本語形態素解析, *情報処理学会研究報告, 自然言語処理研究会報告 2004(47)*, pp.89-96, 2004.
- [11] 相良かおる：ComeJisyo の紹介と医療情報に含まれる誤字調査, *情報知識学会誌*, Vol.24, No.2, pp.204-209, 2014,
- [12] 国試対策問題編集委員会編：医師国家試験のためのレビューブック内科・外科 第10版, メディックメディア社, 2013.
- [13] 武田多一：Carnett's sign, *JIM (医学書院)*, Vol.19, No.3, pp.222-223, 2009.
- [14] 工藤拓, 松本裕治：チャンキングの段階適用による日本語係り受け解析, *情報処理学会論文誌*, Vol.43, No.6, pp.1834-1842, 2002,
- [15] 胡井智, 川本克久, 寺前直樹, 他：虚血性大腸炎 51 症例の臨床的検討, *日本消化器内視鏡学会雑誌*, Vol.32, No.12, pp.2839-2845, 1990.