

コンピュータの国語力を大学入試問題で測る

Measuring Japanese Language Proficiency of Computer by Using University Entrance Examination

名古屋大学大学院工学研究科教授 **佐藤 理史**

PROFILE

京都大学大学院工学研究科電気工学第二専攻博士課程研究指導認定退学。博士（工学）。北陸先端科学技術大学院大学、京都大学を経て、2005年6月より名古屋大学大学院工学研究科教授

1 はじめに

日本において、大学入試問題は、学力（知力および知識力）を問う問題として定着している。この大学入試問題をコンピュータに解かせようという試みが、国立情報学研究所のグランドチャレンジ「ロボットは東大に入れるか（略称：東ロボ）」というプロジェクトとして2011年に開始された^[1]。このプロジェクトの中間目標は、2016年度までに大学入試センター試験で高得点を取ることであり、最終的な目標は、2021年度までに東京大学の入試を突破できるような人工知能システムを作ることである。

我々の研究室は、2013年5月からこのプロジェクトに参加し、大学入試センター試験の『国語』現代文の問題を解くシステムの開発に取り組んでいる。なぜ『国語』現代文なのか。それは、国語力の実現こそが、このプロジェクトの正否の鍵を握っていると考えるからである。国語力とは何か。それをどのような方法でコンピュー

タに身につけさせることができるか。我々がこのプロジェクトで考えたいことは、これらのことである。

2 センター試験『国語』現代文

当面のターゲットは、センター試験『国語』現代文である。センター試験『国語』の基本構成を表1に示す。この表に示すように、現代文に関する出題は、第1問「評論」と第2問「小説」であり、『国語』全体（200点満点）の半分を占める。それぞれの問題は、比較的長い本文（3000字～4000字程度）と、それに対する3種類の設問から構成されている。3種類の設問は、(1)漢字問題（評論）か語句問題（小説）、(2)傍線部問題、(3)総合問題である。これらのうち、傍線部問題が配点の3分の2を占める。傍線部問題の例を図1に示す。

典型的な傍線部問題は、傍線部の意味を問う問題（傍線部の内容を別な言葉で言い換える問題）と、傍線部の理由を問う問題である。さらに、小説の傍線部問題では、

表1 センター試験『国語』の基本構成

第1問	評論	本文 3000-4000字程度	問1	漢字の書き取り問題（計5問）	10点
			問2～問5	傍線部問題	32点
			問6	本文全体に関わる問題 （論の進め方、構成上の特徴）	8点
第2問	小説	本文 3000-4000字程度	問1	語句の意味を問う問題（計3問）	9点
			問2～問5	傍線部問題	31点
			問6	本文全体の趣旨、作者の意図、 表現上の特徴等に関する問題	10点
第3問	古文				50点
第4問	漢文				50点

第1問 次の文章を読んで、後の問い(問1～6)に答えよ。(配点50)

鐺というものを、ふとした機会から注意して見始めたのは、ここ数年来の事だから、未だ合点のいかぬ節もあり、鐺に関する本を読んでみても、人の話を聞いてみても、いろいろ説があり、不明な点が多いのだが。

鐺の歴史は、無論、刀剣とともに古いわけだが、普通、私達が鐺を見て、好き嫌いを言っているのは、室町時代以後の製作品である。何んと言っても、応仁の大乱というものは、史上の大事件なのであり、これを境として、A日本人の鐺というものの見方も考え方も、まるで違って了った。所謂鐺なるものは、この大乱の産物と言ってよいのである。(以下、省略)

問2 傍線部A「日本人の鐺というものの見方も考え方も、まるで違って了った」とあるが、それはどういうことか。その説明として最も適当なものを、次のうちから一つ選べ。解答番号は[6]。

- ① 鐺は応仁の大乱以前には富や権力を象徴する刀剣の拵の一部だったが、それ以後は命をかけた実戦のための有用性と、乱世においても自分を見失わずしたたかに生き抜くための精神性とが求められるようになったということ。
- ② 鐺は応仁の大乱以前には特権階級の富や権力を象徴する日用品としての美しさが重視されていたが、それ以後は身分を問わず使用されるようになり、平俗な装飾品としての手ごろさが求められるようになったということ。
- ③ 鐺は応仁の大乱以前には実際に使われる可能性の少ない刀剣の一部としてあったが、それ以後は刀剣が乱世を生き抜くために必要な武器となったことで、手軽で生産性の高い簡素な形が鐺に求められるようになったということ。
- ④ 鐺は応仁の大乱以前には権威と品格とを表現する装具であったが、それ以後、専門の鐺工の登場によって強度が向上してくると、乱世において生命の安全を保証してくれるかのような安心感が求められるようになったということ。
- ⑤ 鐺は応仁の大乱以前には刀剣の拵の一部に過ぎないと軽視されていたが、乱世においては武器全体の評価を決定づけるものとして注目され、戦いの場で士気を鼓舞するような丈夫で力強い作りが求められるようになったということ。

(2013年センター試験『国語』第1問より)

図1 傍線部問題の例

登場人物の心情を問う問題が出題される。このような問題を解く能力は、一般に「読解力」と呼ばれている。

このような読解問題を、はたしてコンピュータは解くことができるのであろうか。もし、そのような能力をコンピュータが持ち得ないのであれば、SFに出てくるような人工知能は、永遠に夢物語である。

3 『国語』現代文を解くために必要な能力

多くの受験参考書が、『国語』現代文を解くための技術を解説し、その能力を身につけるすべを指南している。しかしながら、私が見るところ、正解を導くのに必要な能力を体系立てて説明している参考書はほとんどないように思われる。

貴重な例外は、石原千明氏による一連の著作物である。最新の『打倒！センター試験の現代文』^[2]によれば、評論問題を解くための能力は次の3つであると述べられている。

- ① 書いてあることを過不足なくまとめること
＝情報整理能力
- ② 書いてあることを別の言葉で言い換えること
＝翻訳能力

③ 二項対立にそって論述されている道筋を整理すること＝二項対立参照能力

一方、小説問題を解くための能力は、①②の他に、次の2つが示されている。

- ③ 書かれていない「心情」を多くの人がそう思うだろうように想像すること＝小市民的感情力
- ④ 出題者と問題を共有すること
＝物語パターン化能力

これらのうち、①と②は、まさに一般知能(General Intelligence)と呼ぶにふさわしい能力である。多くの参考書は、それ以外の能力(評論の③と小説の③④)の必要性和その技術(悪く言えば、小手先のテクニック)について紙面を割いているが、その技術の適用には、①②の能力とそれを支える語彙力等の基本的な国語力が大前提となっている。それゆえ、受験参考書が推奨する技術の大半は、そのままコンピュータに実装することはできない。

石原氏は、小説の③④を主な理由に、センター試験即刻廃止を主張している。彼は、この本で、「(解き方を)無味乾燥に、つまり技術的にテクニカルにかいておこう(p5)」と述べているが、コンピュータに実装する立場から見ると、解き方はまったく機械的にはなっていない。能力①②を含む基本的な国語力がない限り、絵に描いた



餅なのである。つまり、センター試験『国語』は一般知能や基本的な国語力を測っており、大学入学試験としてそれなりに（あるいは、十分に）意味があるとみなすことができる。解けるかどうかはさておき、人工知能がチャレンジする問題としてふさわしい問題である。

これまでの言語処理の研究において、①や②の能力の実現を目指した研究がないわけではない。①に対してはテキストの自動要約、②に対しては含意認識あるいは言い換えの認識・生成がある。しかしながら、入試問題を解くために必要な能力は、現在の研究レベルを遥かに超えている。新しいブレークスルーが必要であることは、言うまでもない。

4 コンピュータの国語力の強化

人間の国語力の強化法は、私にとっては自明である。それは、文章をたくさん読み、たくさん書くことである。一方、どうすればコンピュータの国語力・読解力を強化することができるかは、よくわからない。単に、大量の文章をコンピュータに記憶させておくことで、能力が向上するとも思えない。この辺が他の教科（たとえば、日本史や世界史）と大きく異なる点である。

昨年度は、6か月という短期間で、とにかく動く解答器（ソルバー）を作らなければならなかったため、読解問題に対しては、非常識な方法を採用した。その方法とは、「本文（のある部分）に出現する文字を最も多く含む選択肢を選ぶ」という方法である。何人かの研究者からは、「イカサマのような方法」やら「まったく人を食った方法」などの非難（賞賛？）を受けたが、結果的には、この方法で、評論の傍線部問題の約半数を解くことができた^[3]。大手予備校の模試を利用した公開実験では成績は振るわなかったが^[4]、それでも英語よりは良い成績であった。

センター試験『国語』の平均点は100点換算で60点程度であり、直近の2年は50点前後である。つまり、平均的な受験者は、『国語』の問題を半分程度しか解けない。この事実をどう考えればよいのであろうか。すくなくとも、「平均的な受験者は、問題の本文に書かれている内容をすべて読み取ることができるわけではない」

と考えるべきだろう。言語処理研究者は、人間の言語能力を理想化して考えがちだが、平均的な人間の読解力は、それほど高くないのかもしれない。とするならば、「コンピュータはどのレベルの読解力を目指すべきか」について、真剣に考えてみる必要がある。

昨年度は、傍線部問題と平行して、漢字問題や語句問題を解くソルバーの開発も行なったが、これらの経験も、いくつかの知見をもたらした。たとえば、構文解析は、これらの問題を解くことにまったく寄与しない。これらの問題を解くために不可欠なことは、文の形式的構造を知ることではなく、文中に出現する語や句をグラウンディングすることである。グラウンディングとは、記号に意味を貼付けることであり、具体的には、たとえば、文中の語や句から国語辞書の項目へのリンクを確立すること（国語辞書へのグラウンディング）である。しかしながら、この処理は、それほど自明ではないし、その技術は確立されていない。

漢字問題は、カタカナに書き換えられた表記（「クウバクたる問題であろう」）から、国語辞書の項目（「空漠」）を同定する問題として捉えることができる。すなわち、正しくグラウンディングできれば、問題は解ける。語句問題は、フレーズの正しい言い換えを選択する問題（石原氏の分類の②のひとつ）である。「愛想を尽かしてた」というフレーズに対しては、「愛想」や「尽かす」に個別にグラウンディングするのではなく、「愛想を尽かす」という慣用句に正しくグラウンディングできれば、かなりの確率で正しい選択肢（「嫌になってとりあわないでいた」）を選ぶことができる。

現在の解析ツールは、このようなグラウンディングをサポートしているかといえ、否である。形態素解析において、形態素解析辞書の項目へのグラウンディングは行われているが、意味が記述されている辞書へのグラウンディングは視野に入っていない。つまり、構文解析等の処理で、文の形式的構造が得られたとしても、それを構成する要素（形態素や文節）は、品詞情報がついた記号に過ぎないのである。そのレベルでは、読解とよびべき処理が何もできないのは、明らかであろう。

常識的に考えれば、大学入試の『国語』現代文の問題は、コンピュータにとって、とてつもなく難しい。記述式の東大の二次試験に至っては、解けるはずがない、と

考えて当然である。しかし、本当にそうであろうか。もしそうだとすれば、コンピュータの言語処理は、いつまでたっても形式的な処理の範囲を超えることはないであろう。HAL9000 も鉄腕アトムも R2D2 もドラえもんも、すべて夢物語のままである。読解という方向を陽に目指すことにより、言語処理の新しい方向性とブレークスルーが見えてくることを期待したい。

参考文献

- [1] 新井紀子, 松崎拓也. ロボットは東大に入れるか? — 国立情報学研究所「人工頭脳」プロジェクト—. 人工知能学会誌, Vol.27, No.5, pp.463-469, 2012.
- [2] 石原千秋. 打倒! センター試験の現代文. ちくまブリマー新書 217, 筑摩書房, 2014.
- [3] 佐藤理史, 加納隼人, 西村翔平, 駒谷和範. 表層類似度に基づくセンター試験『国語』現代文傍線部ソルバー. 自然言語処理, Vol.21, No.3, pp.465-483, 2014.
- [4] 佐藤理史, 加納隼人, 西村翔平. 代ゼミ模試に挑戦 2013—『国語』現代文. 情報処理学会 第 215 回 自然言語処理研究会 (SIG-NL-215), 2014.