

大学入試問題と文章の理解

University Entrance Exams and Text Understanding



名古屋大学大学院工学研究科教授

佐藤 理史

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年より現職。

1 はじめに

2013年から2016年にかけて、「ロボットは東大に入れるか(東ロボ)」というプロジェクトに参加し、センター試験の国語の問題の自動解法に取り組んだ。その経緯と研究内容は[1][2]に記したが、「文章の読解とは結局のところ何なのか」という疑問が宿題として残った。自然言語処理の最終目標が「コンピュータに言葉を理解させること」にあるならば、それが何を意味し、どのようにすれば可能となるかは、まだよくわかっていないと言わざるを得ない。

我々が取り組んだのはセンター試験の最初の大問2問で、前者は評論、後者は小説を題材とした問題である(残りの2問は古文と漢文の問題)。いずれの大問も、その中核は読解問題とよばれる問題で、大問の冒頭に示された3000字から3500字程度の本文のある部分を参照し、その解釈を問う選択式問題(選択肢は5つ)である。評論の読解問題は、おおよそ50%の正解率に達したが、小説の読解問題は、ほとんど歯が立たなかった。

この差は、評論の読解問題では「本文に書かれていることが問われる」のに対し、小説の問題では「本文に書かれていないことが問われる」という違いによる。書かれていることについては、文章を理解せずとも問題を解く方法は色々あり、実際、機械的にかなり解ける。これに対して、文章に書かれていない(が、人間ならば常識的に想像できる)ことに答えるのは、容易ではない。

ここで、簡単な例を示そう。あなたは、次の文章を読

んで何を思うだろうか。

名古屋駅で15時すぎの「のぞみ」に乗った。東京に着いたのは真夜中近くだった。

ほとんどの人は、「何が起きたのだろう」、あるいは、「新幹線にどんなトラブルが起きたのだろう」と思うに違いない。しかし、それは文章のどこにも書かれてない。このように、我々は無意識のうちに、「書かれた内容」から「書かれていない内容」を推測しながら文章を読んでいる。

上記のようなプロセスを明らかにするために、どのような研究を行えばよいだろうか。そのような問題意識から取り組んでいるのが、本稿で紹介する2つの研究である。

2 熱化学方程式の文章題を解く

次のような問題を考えよう。

水素、炭素(黒鉛)、プロパン C_3H_8 の燃焼熱は、それぞれ 286kJ/mol 、 394kJ/mol 、 2219kJ/mol である。

- (A) プロパンの燃焼熱は?
(B) プロパンの生成熱を求めよ。

(A) は、与えられた問題文に書かれている内容を問う設問である。日本語の並列構造と「それぞれ」の文法

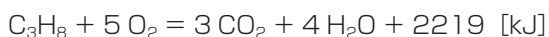
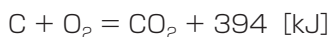
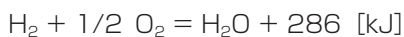
的機能を知っていれば、「2219kJ/mol」という答を導くのはたやすい。

一方、(B)の答「107kJ/mol」は問題文には存在しない。しかし、その答は、問題文に示された情報と化学の知識を組み合わせることによって導くことができる。実のところ、(A)は私が作問したが、(B)は高校の化学の教科書に掲載されている練習問題である。念のために付け加えると、(B)は知識(事実)を問う問題ではない(これらの値は近似値であり、真の値ではない)。この設問では、問題文で与えられた情報を使って答を導くことが求められている。

「燃焼熱」とは、物質 1 mol が完全燃焼するときの反応熱であり、その熱化学反応式は、次のような形式となる。

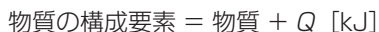


つまり、問題文は、以下の情報を読み手に提示している。

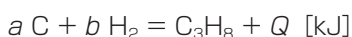


もちろん、問題文からこれらの情報(熱化学方程式)を引き出すためには、水素の元素記号が H で、標準状態では気体 H_2 であること、炭素の元素記号が C で、黒炭の標準状態が固体 C であること、そして、熱化学方程式の両辺の元素の数の比較により、係数が決定できること(連立方程式が解けること)が必要である。しかし、それらを知っているものにとっては、問題文が上記の3つの熱化学方程式を意味することは明快であり、これこそが、この問題を解くための問題文の解釈(理解)なのである。

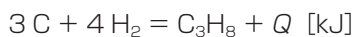
一方、「生成熱」とは、化合物 1 mol がその成分元素の単体から生成するときの反応熱であり、その一般式は、次のような形式である。



物質がプロパン C_3H_8 の場合、構成要素は炭素 C と水素 H_2 なので、次のようになる。



両辺の元素の数を比較して、係数の a と b を定めると、最終的には以下ようになる。



つまり、「プロパンの生成熱を求めよ」という問いは、「この式の Q を求めよ」と要求しているのである。この

Q の値は、この式と問題文から得られる3つの式を連立して解くことによって求められる^[3]。

さて、途中から言語(日本語)の話ではなく、化学の話になったように思うかもしれないが、先程の新幹線の例を思い出してほしい。新幹線の例の文章を読んで「何が起きたのだろう」と思うのは、「名古屋から東京まで『のぞみ』であれば2時間かからない」ということを知っているからである。「15時すぎに乗ったのに着いたのは真夜中近く」という文章から読み取れる情報と、この知識により、「何かトラブルが起きたのだろう」と推測したのである。

文章から読みとれる情報を、読み手が持っている知識と照らし合わせて「書かれてない情報を導き出す」。この点においては、新幹線の例も、熱化学計算の文章題も同じ構造をしている。そのプロセスは、新幹線の例ではいささか曖昧であるが、熱化学計算の問題では明快である。加えて、文章を正しく解釈できたかどうかを、答え合わせという形で確認できる。

東口ポにおいて、機械的にはほとんど解けなかった教科は理科である。それは、問題を解くために、専門知識を使って問題文を解釈し、内部表現(上記の例では熱化学方程式)を構築して、その表現上である種の推論を行う(上記の例では連立方程式を解く)必要があるからである。これを自在に行うことは、現在の技術ではまだまだ難しい。

3 確率・期待値の文章題を解く

東口ポ終了後も、継続して取り組んでいる研究に、数学の確率・期待値の自動解法の研究がある^{[4][5]}。具体的には、次のような問題である。

1 から 9 までの数字が書かれた、計 9 枚のカードがある。この中から、同時に 2 枚取り出す。

(1) 2 枚とも偶数である確率を求めよ。

取り出したカードの数字の大きい方を得点とする。

(2) 得点の期待値を求めよ。

数学は、東口ポでは最も注力され、比較的よく解けた科目であるが、確率・期待値の問題は、まったく解けなかったといってよい。その理由は、確率・期待値の問題



が、数学ソルバーの基本戦略（問題文を一階述語論理式に変換し、それを簡略化することによって答を導く）にフィットしないことにある。

確率・期待値の問題は、おおよそ5つの要素から構成される。

- (1) 仮想世界の記述：どんなオブジェクトがいくつあるか
- (2) 試行の記述：標本空間はどのように規定されるか
- (3) 確率の計算指示：ある事象が起きる確率を求めることを要求する
- (4) 属性の定義：特定の事象や（試行の）結果に対する属性を定義する
- (5) 期待値の計算指示：ある数値属性の期待値を求めることを要求する

確率・期待値の特徴は、「1 から 9 までの数字が書かれた、計 9 枚のカードがある」のような記述で、ある仮想世界が記述され、「この中から、同時に 2 枚取り出す」といった記述で、その仮想世界の中のオブジェクトの操作が記述される点にある。つまり、問題を解くためには、文章に従って仮想世界を構築し、それを操作できなければならない。さらに付け加えるならば、その操作で得られる特定の具体例だけでなく、その総体（可能なすべての具体例＝標本空間）が想像できなければならない。

実際に確率を求める際には、「2 枚とも偶数である（確率）」とはどういう事象（標本空間の部分集合）であり、かつ、その事象を作り出せなければならない。期待値の計算では、得点の定義を問題文から読み取り、それを計算できるようにしなければならない。

このように、確率・期待値の文章題を解くためには、文章解釈に関連する多くの問題を解く必要があるが、その中で我々が現在取り組んでいるのは、確率の計算指示の解釈である。ここでは、標本空間は与えられるものとし、計算指示中の事象を規定する条件記述「2 枚とも偶数である」を、標本空間（入力）から事象（出力）を求めるプログラムに変換する。事象が求まれば、確率は、それぞれの集合の要素の数を使って、次の式で計算できる。

$$\text{確率} = \text{事象の要素数} / \text{標本空間の要素数}$$

プログラムへの変換で重要な役割を果たすのは、標本空間の要素である「結果」の構造である。「この中から、

同時に 2 枚取り出す」という試行の記述から、結果は「2 枚のカード」（リストと呼ぶ）であり、その構成要素が「カード」（オブジェクトと呼ぶ）となる。一方、事象を規定する条件記述「2 枚とも偶数である」の「2 枚とも」はリストに対する制約であり、「偶数である」はオブジェクトに対する制約である。つまり、「あるカード（オブジェクト）の数字が偶数であるかどうかを調べる」プログラムと、「リストを構成するすべてのオブジェクトが、ある条件を満たす」プログラムを組み合わせれば、「2 枚とも偶数である」ことをチェックするプログラムを合成することができる。

このプログラム合成を実現するためには、おおよそ、以下のことを行えばよい^[5]。

- (1) 条件記述に用いられる基本要素（「N 枚とも」、「偶数である」）を整理し、それらに対応する基本プログラムを用意する。
- (2) 条件記述を基本要素に分解し、それぞれがどの階層（リスト、オブジェクト）に対する制約となっているかを、各種の手がかりを用いて決定する。
- (3) その決定に基づき、プログラムを合成する。

我々が作成したシステムの実行例を図 1 に示す。この例以外にも、「サイコロを 2 回振る。1 回目に 2 以下または 5 以上の目が出て、かつ 2 回目に奇数の目が出る確率を求めよ」といった記述でも、正しく解釈できる。

確率・期待値問題で重要なポイントの一つは、前の文の解釈が、次の文の解釈に必要となる点にある。先に述べた結果の構造は、試行の記述から読み取らなければならない（この部分は、まだ機械化できていない）。それが読み取れて、はじめて事象を規定する条件記述が正しく解釈できるようになる。文章の解釈には、読み手が持っている知識が関与するだけでなく、そこまでの文章から読み取った情報が関与する。確率・期待値問題は、そのような機能が実現できるかどうかの試金石でもある。

熱化学計算問題の内部表現が熱化学方程式だったのに対し、確率・期待値問題の内部表現は（実行可能な）プログラムである。言い換えるならば、文章で書かれた内容を手続きとして解釈し、それを実行するプログラムに翻訳することを行っている。確率の計算指示の解釈で行っていることは、広くは日本語で書かれた条件記述を計算可能にする技術に分類されよう。こうした技術が確立されれば、コンピューターは、単なる命令（「電話を

状況の設定 1から9までの数字が書かれた、計9枚のカードがある。

試行の設定 この中から、同時に **2枚** 取り出す。これを **1回** 繰り返す。ただし、取り出したカードを **元に戻す**。

枚数 2 枚

回数 1 回

取り出したカードを **元に戻す** **元に戻さない**

条件記述 2枚とも偶数である × 確率を求めよ。

計算 **クリア**

Step 1: ユニットの解釈

入力された条件記述をユニットに分解し、それぞれのユニットがどの階層で解釈できるのかを調べる。

#	ユニット	種類	階層
1	2枚とも	全部	List
2	偶数で	偶数?	Object
3	ある	肯定	Object

Step 2: 関数の定義

Step 1 で調べたユニットを合成して、関数を定義する。

#	関数名	処理内容	階層
1	偶数である	偶数?(属性値の取得(:数字))	Object
2	2枚とも偶数である	すべて(:式_偶数である)	List

Step 3: 確率の計算

Step 2 で定義した関数を用いて、確率を計算する。

$$\text{確率の値} = \frac{\text{事象}}{\text{標本空間}} = \frac{6}{36} = \frac{1}{6}$$

図1 条件記述を解釈してプログラムに変換し、確率の値を計算する

かけて)だけでなく、条件付き命令(「3時になったら、電話をかけて)を受け付けることができるようになる。

4 文章の理解とは

これら2つの研究で実感したことは、「文章の理解は一樣ではない」ということである。熱化学の問題文には熱化学の問題としての理解があり、確率・期待値の問題文には、確率・期待値問題としての理解がある。その本質は、分野固有の知識(専門知識)を利用した文の解釈、および、そこから得られる情報を専門知識と組み合わせることにより、明示的に書かれていない情報を導き出す

ことである。そして、問題文をそのように理解しなければ、これらの問題は決して解けない。

「いやいや、これは大学入試問題の問題文だからで、特殊な状況での文章の理解である」という反論があるかもしれない。確かに、使う知識は専門知識である。しかし、再度、新幹線の例を思い出してほしい。自分が持っている知識を使って書かれていない情報を引き出すことは、我々が文章を読むときに自然に行っていることであり、入試問題固有の現象ではない。

ここ数年、自然言語処理は、ディープニューラルネット一色であり、2000年頃には想像できないレベルで、色々なタスクを実現できるようになってきた^[6, 7, 8]。巨

大な言語モデルは、四則計算もあるレベルで実行できるという報告さえある^[9]。さらに、簡単な算数の文章題を解くニューラルネットモデルも提案されている^[10]。本稿で紹介した2つの問題も、ニューラルネットで解くことは原理的には可能であろうが、現時点ではまだまだ難しそうである。

一方、クラシックな記号処理で解くことができるのは、適切な内部表現が設定できる問題に限られる。熱化学計算問題が解けるのは、熱化学方程式という適切な内部表現が存在し、熱化学方程式に変換できれば、その後の処理を、方程式ソルバーに任せることができるからである。

それに対して、確率・期待値問題では、問題文をプログラムに変換する。実行可能なプログラムという内部表現は、汎用性という点では申し分ない。文章で書かれた内容を実行可能なプログラムに変換する技術は、文章の理解という難問に挑むためのひとつの方向性を示唆しているように思う。

参考文献

- [1] 佐藤理史. コンピュータが小説を書く日. 日本経済新聞出版社, 2016.
- [2] 新井紀子, 東中竜一郎, 人工知能プロジェクト「ロボットは東大に入れるか」. 東京大学出版会, 2018.
- [3] 経種直之, 佐藤理史, 小川浩平, 宮田玲. 熱化学計算問題の段階的解釈に基づく熱化学方程式の立式. 言語処理学会第27回年次大会, pp.674-678, 2021.
- [4] 神谷翼, 松崎拓也, 佐藤理史. 数学確率文章題の自動解答システムの開発. 言語処理学会第21回年次大会発表論文集, pp.365-368, 2015.
- [5] 岩間純輝, 佐藤理史, 小川浩平, 宮田玲. 数学・確率問題を対象とした条件記述の自動解釈. 言語処理学会第27回年次大会, pp.947-951, 2021.
- [6] A. Radford et al. Improving Language Understanding by Generative Pre-Training. 2018.
- [7] A. Radford. et al. Language Models are Unsupervised Multitask Learners. 2019.
- [8] J. Devlin. et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv: 1810.04805v2, 2019.
- [9] T. B. Brown et al. Language Models are Few-Shot Learners. arXiv:2005.14165v4, 2020.
- [10] J. Zhang et al. Graph-to-Tree Learning for Solving Math Word Problems. Proc. of ACL-2020, pp. 3928-3937, 2020.



5

産業日本語関連