

代ゼミ模試に挑戦2013—『国語』現代文

佐藤 理史^{1,a)} 加納 隼人² 西村 翔平²

概要: 我々は、大学入試センター試験『国語』現代文の問題を解くシステムを実装し、代々木ゼミナールのセンター模試に挑戦した。本論文では、今回の挑戦に用いた解答器の技術的詳細と模試の結果について報告する。

キーワード: 大学入試センター試験, 国語現代文, 漢字問題, 語句問題, 傍線部問題

Yozemi-Challenge 2013—Contemporary Japanese Language

SATOSHI SATO^{1,a)} HAYATO KANOU² SHOHEI NISHIMURA²

Abstract: We have implemented systems that solves questions in contemporary Japanese language, National Center Test for University Admissions. This paper describes the solvers embedded in these systems and the result of Yozemi-Challenge 2013.

Keywords: National Center Test for University Admissions, contemporary Japanese language, Kanji questions, questions about words and phrases, questions about underlined passages

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計算機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年までに大学入試センター試験で、東京大学の二次試験に進めるような高得点を取ることである。

我々は、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。これまでに、いわゆる「傍線部問題」を解く方法を提案し、「評論」の「傍線部問題」が半数以上解けることを示した[2]。

本稿では、東ロボ・プロジェクトの一環として2013年

11月に実施したシステム評価(オープンテスト)のうち、『国語』現代文について報告する。このシステム評価は、作成した解答器(ソルバー)を用いて、実際に代々木ゼミナール(代ゼミ)の模試の問題を解くという試みであり、評価型ワークショップのフォーマルランと同じ方式で実施した。そこで使用したソルバーの技術的詳細と模試の結果を示すことが、本稿の主な目的である。

以下、本稿は、次のように構成されている。まず、2節で、今回の代ゼミ模試の結果を示す。3節から6節では、使用した4種類のソルバーの技術的詳細について述べる。最後の7節で、今回の挑戦についてまとめる。

2. 代ゼミ模試の結果

代ゼミのセンター模試『国語』の問題は、大学入試センター試験の『国語』の問題構成を踏襲しており、大問4題で構成されている。このうち、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(配点200点)の半分を占める。

我々は、以下のように、第1問と第2問の設問をそれぞれ

¹ 名古屋大学大学院工学研究科
Graduate School of Engineering, Nagoya University

² 名古屋大学工学部
School of Engineering, Nagoya University

a) ssato@nuee.nagoya-u.ac.jp

表 1 代ゼミセンター模試に対するソルバーによる実行結果と採点結果

大問	設問 ソルバー	2012 年度第 1 回模試							2013 年度第 1 回模試 (フォーマルラン)							
		番号	配点	0	1	2	3	4	番号	配点	0	1	2	3	4	
第 1 問 (評論)	問 1 (漢字) K0	A1	2	5						A1	2	2				
		A2	2	2						A2	2	3				
		A3	2	2						A3	2	3				
		A4	2	3						A4	2	5				
		A5	2	5						A5	2	4				
	(小計)	(10)	(10)						(小計)	(10)	(10)					
	問 2-5 (傍線部) E0-E4	A6	8	2	2	3	5	3		A6	8	5	5	5	5	5
		A7	8	3	3	3	3	1		A7	8	2	2	2	1	2
		A8	8	1	1	4	1	4		A8	8	4	4	4	4	4
		A9	8	4	4	4	4	4		A9	8	5	5	1	5	1
	(小計)	(32)	(16)	(16)	(24)	(16)	(16)		(小計)	(32)	(0)	(0)	(0)	(8)	(0)	
	問 6 (総合) F0	A10	4	2						A10	8	2				
		A11	4	2						(小計)	(8)	(0)				
(小計)		(8)	(4)						計	50	10	10	10	18	10	
第 2 問 (小説)	問 1 (語句) G0	A12	3	2					A11	3	3					
		A13	3	4					A12	3	4					
		A14	3	3					A13	3	4					
		(小計)	(9)	(9)					(小計)	(9)	(3)					
	問 2-5 (傍線部) N0-N3	A15	8	3	3	2	2		A14	8	2	5	2	2		
		A16	7	1	3	1	1		A15	7	2	2	4	2		
		A17	8	4	5	4	4		A16	8	4	4	4	1		
		A18	8	3	4	3	3		A17	8	3	3	5	3		
	(小計)	(33)	(8)	(7)	(16)	(16)		(小計)	(33)	(0)	(0)	(16)	(0)			
	問 6 (総合) F0	A19 A20	各 5	3 4					A18 A19	各 5	3 4					
		(小計)	(10)	(5)					(小計)	(10)	(5)					
	計	50	22	21	30	30		計	50	8	8	24	8			

数字はソルバーが出力した選択肢番号, □ は正解したことを表す。

れ 3 種類に分類し, そのそれぞれに対してソルバーを作成してフォーマルランを実行した。

第 1 問 「評論」

- 問 1 漢字問題 — ソルバー K0
- 問 2-5 評論傍線部問題 — ソルバー E0-E4
- 問 6 評論総合問題 — ソルバー F0

第 2 問 「小説」

- 問 1 語句問題 — ソルバー G0
- 問 2-5 小説傍線部問題 — ソルバー N0-N3
- 問 6 小説総合問題 — ソルバー F0

このうち, ソルバー F0 は, 総合問題に対して便宜的に解答を出力するためのもので, 選択肢の数と解答形式に応じて, あらかじめ決められた選択肢番号を出力する (たとえば, 5 択から 1 つ選ぶ場合は, 選択肢番号 2 を出力する)。

2013 年度第 1 回全国センター模試 (5 月実施) を対象としたフォーマルランの結果と, その直前に実施した 2012 年度第 1 回全国センター模試を対象とした実行結果を表 1 に示す。

2012 年度模試に対する結果 (図の左側) は, 最高点 68 点,

最低点 51 点である。我々は, ソルバー開発において, この模試を開発用問題群の一部として利用したが, 特にこの模試の点数が良くなるように調整したわけではない。そのため, 我々は, フォーマルランの直前, かなり楽観的な見通しを持つに至っていた。

しかしながら, フォーマルランの結果 (図の右側) は, 最高点 42 点, 最低点 18 点という惨澹たる結果であった。漢字問題は小問 5 つとも正解を導けたが, 半分は堅いと考えていた評論傍線部問題がほとんど解けず, 小問 3 つ中 2 つは堅いと考えていた語句問題でも, 1 問しか正解を導けなかった。我々の自己評価は「完敗」である。なお, 東ロボプロジェクトの公式発表は, プロジェクト事務局が最高点となるソルバーの組み合わせを採用したため, 『国語』現代文の成績は 42 点となっている。

3. 漢字問題ソルバー (K0)

3.1 漢字問題

漢字問題は, 第 1 問「評論」の問 1 として, 5 題出題される。漢字問題の例を図 1 に示す (オリジナルは縦書きで

表 2 設問側・選択肢文字列の階層構造

層	表層の特徴	言語的単位	設問側	選択肢 1	選択肢 4
K 層	傍線部	漢字 1 字	<u>タイ</u>	<u>タイ</u>	<u>タイ</u>
T 層	カタカナ文字列	語 (または語の一部)	<u>タイ</u> ザ	<u>タイ</u> マン	シン <u>タイ</u>
\hat{T} 層	前後の漢字列を含む	漢字複合語	—	—	出処シン <u>タイ</u>
X 層	直後のひらがな列を含む	—	<u>タイ</u> ザして...	<u>タイ</u> マンな	出処シン <u>タイ</u> を
S 層	文字列全体	フレーズまたは文	(下線部を含む文)	<u>タイ</u> マンなプレー を注意される	出処シン <u>タイ</u> を 明らかにする

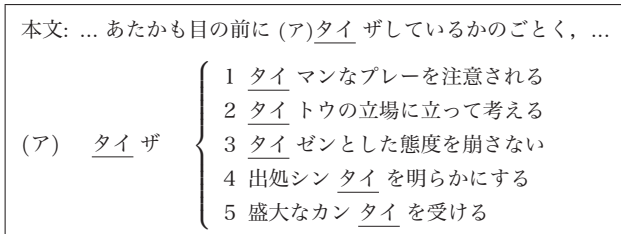


図 1 漢字問題の例 (2013 年度模試の第 1 問 問 1 (ア))

ある), この問題では, 本文中に出現する「タイザ」の「タイ」の漢字と同じ漢字を含むものを, 5 つの選択肢から選ぶことが求められる。このような問題を解くためには, 本文中に出現する傍線部 (これを, 便宜上, 設問側と呼ぶ) に対応する漢字と, それぞれの選択肢の傍線部に対応する漢字を推定し, 漢字が一致する選択肢を選ぶ必要がある。

3.2 基本戦略

漢字問題に対する我々の基本的なアイデアは, 設問側や選択肢の傍線部を含む文字列に対し, 表 2 に示すような階層構造を仮定する点にある。漢字問題を解くためには, K 層に対応する漢字 1 字を同定する必要があるが, それは, 実質上, T 層に対応する語 (または語の一部) の漢字表記を求めることに帰着される。

この漢字表記推定問題は, たとえば, 「タイマン (怠慢)」のように, T 層の漢字表記が一つしか存在しなければ, 容易に解ける。しかし, 「タイトウ (対等, 台頭, 帯刀, など)」のように, 漢字表記が複数存在する場合は, それらの中から正しいものを選ぶ必要が生じる。

この選択のために, 我々は以下の情報を得点化して利用する。

- (1) K 層に対応する漢字の難易度 (3.4.3 節の係数 α)
- (2) T 層に対応する漢字語 (漢字表記) の頻度 (3.4.2 節の頻度ディスカウント)
- (3) \hat{T} 層に対応する複合語の存在の有無 (3.4.2 節の複合語チェック)
- (4) T (または \hat{T} 層) に対応する漢字語の文法情報と, X 層との文法的整合性 (3.4.2 節の文法チェック表)
- (5) 候補の順位 (3.4.3 節の順位ディスカウント)

3.3 漢字問題の解法

漢字問題の解法の全体像を以下に示す。

- (1) 設問側の K 層に対応する漢字の候補 (複数) とそれらの候補の得点を求める。
- (2) 各選択肢に対して, 以下のことを行なう。
 - (a) その選択肢の K 層に対応する漢字の候補 (複数) とそれらの候補の得点を求める。
 - (b) 得られた漢字候補が設問側の漢字候補に含まれている場合, それを候補として残し, その得点を, 元の得点に設問側の得点を乗じたものに変更する。含まれていない場合は, その候補を破棄する。
 - (c) 残った候補の得点のうち, 最も高い得点を, その選択肢の得点とする。候補が残らなかった場合は, 0 点とする。
- (3) 最も得点が高い選択肢を出力する*1。

3.4 漢字候補の推定

漢字問題の解法の中核は, 上記の手続きの (1) および (2a) である。これらは, 基本的にほぼ同一の方法で解く。以下に, そのアルゴリズムを示す。

- (1) 入力として S 層が与えられる。
- (2) S 層に含まれる T 層で漢字語辞書を引き, 漢字語の候補を得る。同時に, それぞれの漢字語候補に対し得点を計算する。
- (3) 得られた漢字語候補リストから, K 層に対応する漢字候補リストを作成する。

3.4.1 漢字語辞書

漢字語辞書は, 短単位および長単位の語に対して, 以下の情報を定義したものである。

- (1) 語: 見出し語
- (2) 漢字部: 漢字で記述される部分
- (3) 漢字部読み: 漢字部の読み (カタカナ表記)
- (4) ひらがな部: 語幹の末尾のひらがなで記述される部分。空である場合もある。
- (5) 活用語尾: 活用する場合の活用語尾 (ひらがな表記)。活用しない語の場合は空。
- (6) 品詞・活用型: 品詞 (細分類) と活用型に関する情報。
- (7) 頻度順位: BCCWJ*2 を用いて計測した頻度順位

*1 最高得点をとる選択肢が 2 つ以上あった場合は, 選択肢番号が若いものを選ぶ。このタイブレイク法を, すべてのソルバーで採用している。

*2 http://www.ninjal.ac.jp/corpus_center/bccwj/

表 3 漢字語辞書のエントリーの例

語	漢字	漢字部読み	ひらがな	語尾	品詞・活用	順位
欺く	欺	アザム	-	く	動子カ	9875
朽ちる	朽	ク	ち	る	動母	25869
新体	新体	シнтаイ	-	-	名詞	ipa
真諦	真諦	シнтаイ	-	-	名詞	ipa
神体	神体	シнтаイ	-	-	名詞	42101
身体	身体	シнтаイ	-	-	名詞	824
進退	進退	シнтаイ	-	-	名サ	87433
出処進退	出処進退	シュツシヨ シнтаイ	-	-	名詞	four
対坐	対坐	タイザ	-	-	名サ	ipa
対座	対座	タイザ	-	-	名サ	37969
退座	退座	タイザ	-	-	名サ	ipa
怠慢	怠慢	タイマン	-	-	形ナ	24415

表 4 文法チェック表

品詞・活用	H の先頭が					
	空	「な」	「に」	助詞	整合	他
動詞	nil	nil	nil	nil	1.0	nil
形イ	nil	nil	nil	nil	1.0	nil
形ナ	1.0	1.0	1.0	1.0	1.0	0.4
形タル	0.4	0.6	0.6	0.4	1.0	nil
名サ	1.0	0.8	1.0	1.0	1.0	0.4
上記以外	1.0	0.8	1.0	1.0	-	0.4

「動詞」は「動母(母音動詞)」「動子 L(子音動詞 L 行)」「動ザ変」「動サ変」の総称、「整合」は H の先頭が活用語尾と整合すること表す。

(1-87,433 位)*3.

辞書のエントリーの例を、表 3 に示す。なお、語の表記が「漢字文字列(+ひらがな文字列)」という形式とはならない語は、辞書には含まれない。

漢字語辞書は、次の 4 つのソースの情報をマージして作成した。

- (1) JUMAN 辞書 (juman-7.0)
 - (2) IPA 辞書 (ipadic-2.7.0)
 - (3) BCCWJ から抽出した長単位語リスト [3] の一部
 - (4) クロスワードソルバー [4] の辞書データの一部
- 作成した漢字語辞書のエントリ数は、約 12 万件である*4。

3.4.2 漢字語候補とその得点

漢字語候補の生成と各候補の得点計算は、以下の手順で行なう。

- (1) T 層で漢字語辞書を引き、読みから漢字語候補を得る。
- (2) T 層の直後のひらがな列 H が、漢字語候補の文法的情報と整合するかどうかを調べる。漢字語候補の辞書エントリーに「ひらがな部」が存在する場合は、語幹の末尾にひらがなが列が必要なので、まず、その一致を確認し、一致しないものは候補外とする。次に、(語幹ひらがな部を削除した) H と「品詞・活用型」で文法チェック表(表 4)を引き、その結果に従う。この表に

*3 'juman' は 98,000 位, 'ipa' と 'four' は 99,000 位に換算する。

*4 本来ならば同一語とみなすべき重複エントリが存在する。

表 5 係数 α (漢字に対する重み)

カテゴリ	設問側	選択肢
1-2 年生配当の教育漢字	0.7	1.0
それ以外の常用漢字	1.0	1.0
常用漢字以外	0.1	0.1

において、「nil」は、その候補を候補外とすることを意味する。数字の場合は、候補に残すこと、および、その候補の基礎点を表す。

- (3) T 層の直前・直後に漢字文字列がある場合は、得られた各候補に漢字文字列を結合して、再度、漢字語辞書を引き、複合語が存在するかどうか調べる(このとき、上記の方法と同様に基礎点を定める)。これを複合語チェックと呼ぶ。

(a) 1 つ以上の漢字語候補に対して複合語が存在した場合、得られた複合語を漢字語候補とする。複合語が得られなかった(T 層のみに対応する)漢字語候補も候補として残すが、その基礎点は 0.6 倍する。

(b) 複合語がまったく得られなかった場合は、(T 層に対応する)漢字語候補をそのまま候補とし、基礎点もそのまま用いる。

- (4) それぞれの漢字語候補に対し、次式で得点を求める。

$$\text{得点}(w) = \text{基礎点}(w) - \frac{w \text{ の頻度順位}}{10^6} \quad (1)$$

これを頻度ディスカウントと呼ぶ。

以上の処理を完了すると、T 層(または \hat{T} 層)に対する(得点付き)漢字語候補リストが得られる。

3.4.3 漢字候補とその得点

次に、得られた漢字語候補リストから、K 層に対応する漢字候補リストを計算する。

- (1) それぞれの漢字語候補に対して、K 層の漢字を簡単なヒューリスティックにより推定した後、漢字毎にグルーピングする。
- (2) 各漢字の得点を、そのグループに属する漢字語候補の得点の最大値に、係数 α を掛けて求める。
- (3) 得られた漢字候補リストを得点の高い順にソートし、順位を利用した補正を行なう。

係数 α は、漢字に対する重みを表す。この係数は、常用漢字・教育漢字に基づき、表 5 のように定める。この表に示すように、設問側に対する得点計算では、1-2 年生配当の教育漢字*5に差を付けるが、各選択肢に対する得点計算では差を付けない。

順位を利用した得点補正には、次式を用いる。

$$\text{補正後得点}(k) = \text{得点}(k) \times \max \left(1 - \frac{r(k) - 1}{10}, 0 \right)$$

ここで、 $r(k)$ は、漢字候補 k の候補リストでの順位を表す。

*5 我々が調査したセンター試験の過去問題では、1-2 年生配当の教育漢字は出題されていない。模試では、出題されたことがある。

表 6 漢字問題ソルバーの性能

センター過去問 (2001-2013 年の奇数年)	68/ 70	97%
代々木ゼミナール問題集 [5]	33/ 35	94%
駿台予備校問題集 [6]	31/ 35	89%
合計	132/140	94%

この式は、1 位以外の候補のスコアを少し減少させることを意味する。これを順位ディスカウントと呼ぶ。

設問側の漢字候補の得点計算は上記のとおりであるが、選択肢の漢字候補の得点計算では、順位 $r(k)$ の計算法が異なる。具体的には、漢字候補 k の順位を、そのグループに属する漢字語候補の 基礎点 の最大値に基づいて定める。つまり、この順位決定では、頻度や漢字の難易度を考慮せず、文法チェック表と複合語チェックによって定まる基礎点だけで行なうということである。

3.5 ソルバーの性能と問題点

今回の挑戦の前に確認したソルバーの性能を表 6 に示す。この表に示すように、ソルバーの性能は 94% である。センター過去問よりも、予備校模試の方が、若干成績が悪い。

現在のソルバーは、基本的に、長単位語の文法的情報に基づいて、候補を定めている。そして、その範囲で順位が付かなかった場合は、語の頻度および漢字の難易度で(無理矢理)順位を定めている。X 層を越える手がかり、すなわち、隣接語とのコロケーションは用いていない。そのため、コロケーションを用いないと正しい漢字語を有効に絞り込めない問題は、正しく解くことができない場合が多い。しかし、そのような問題は比較的少数なので、今回は目をつぶることとした。

フォーマルランの結果は、表 1 に示した通り、小問 5 問全問正解であった。

4. 語句問題ソルバー (G0)

4.1 語句問題

語句問題は、第 2 問「小説」の問 1 として、3 題出題される。語句問題の例を図 2 に示す。この問題では、本文中に出現する語句「依然として固く」の本文中の意味として最も適切なものを、選択肢から選ぶことが求められている。

4.2 基本戦略

センター試験対策本 [7] によれば、センター試験の語句問題は、基本的に、辞書の定義に沿った選択肢を選ぶことが推奨されている。この指摘に基づき、我々は、次のような基本戦略を取る。

- (1) 設問傍線部の表現から、中心となる語・熟語・慣用句等 ($Q = \{q\}$) を抜き出す。
- (2) 抜き出した各 q で複数の国語辞書を引く。

本文: ... 峰の表情は (ア)依然として固く、どのような微妙な...

- | | | |
|-------------|---|------------------|
| (ア) 依然として固く | } | 1 これまでよりも緊張しており |
| | | 2 いつになく不機嫌そうで |
| | | 3 真面目な様子を崩すことなく |
| | | 4 いっそう頑固な様子になって |
| | | 5 あいかわらずこわばったままで |

図 2 語句問題の例 (2013 年度模試の第 2 問 問 1 (ア))

表 7 慣用句パターン

パターン	例
名詞+の+名詞	後の祭
名詞+の+形容詞	瘤の強い
名詞+を+動詞	不意をつく、体裁をなす、 目を瞞る、声を洩らす、目を疑う
名詞+(が は も)+動詞	はかがいく、憑物が落ちる

- (3) それぞれの選択肢において、得られた定義文のそれぞれとの照合スコアを計算する。それらのスコアの最大値と平均値を計算し、最大値に平均値の 1/10 を足したものを、その選択肢の得点とする。(得点は、1.0 を越えないものとする。)
- (4) 最も得点が高い選択肢を出力する。

4.3 ソルバー G0 の詳細

4.3.1 使用した辞書

以下の辞書を見出し語(表記および読み)で、語義毎の定義文を引けるような形式に変換し、使用した。

- (1) 岩波国語辞典第七版
- (2) 広辞苑第六版
- (3) クロスワードソルバー [4] の辞書データの一部

4.3.2 使用したツール

形態素解析器 mecab-0.994 を、unicdic-2.1.0 を独自に拡張した形態素解析辞書と組み合わせて使用した。この拡張では、各形態素に、派生語を同一視するための情報が追加されている。

4.3.3 語・熟語・慣用句候補の抽出

設問傍線部からの語・熟語・慣用句候補の抽出は、以下の手順で行ない、得られた q を順に集め、候補リスト Q を作成する。

- (1) 設問傍線部
設問傍線部全体を、そのまま q として採用する。
- (2) 形態素解析
設問傍線部を形態素解析し、形態素列を得る。
- (3) 長単位語の認定
簡単なヒューリスティックを用いて、形態素列を長単位語列に変換する。
- (4) 慣用句パターンによる抽出
表 7 に示す慣用句パターンに一致する部分列があれば、それらをすべて q として採用する。それらの順位

は、後方のものを優先する。

(5) 内容語の抽出

長単位の内容語*6をすべて q として採用する。それらの順位は、前方のものを優先する。ただし、その語がいずれかの選択肢に現れる場合は q として採用しない。

ここで、候補リストを優先順に並べておくのは、次の辞書検索ステップで、かならずしもすべての候補を使用するとは限らないからである。

上記の手順により、たとえば、下線部「後片付けのはかは行かず」からは、「後片付けのはかは行かず」という文字列以外に、〈はかは行かず〉、〈後片付けのはか〉、〈後片付け〉、〈はか〉、〈行かず〉が、 q として抽出される。これらは、長単位語の列、または、長単位語である。

4.3.4 辞書検索

先のステップで得られた Q に対し、以下の手順で国語辞書を引く、定義文の集合 P を得る。

- (1) Q の最初の要素を q とする。
- (2) それぞれの辞書に対し、以下を実行する。
 - (a) q が文字列の場合は、その文字列で辞書を引く。
 - (b) q が長単位語列、または、長単位語の場合は、後に述べる優先度順に表層文字列を生成し、辞書を引く。定義文が得られたら、それ以降の表層文字列は使用しない。
- (3) 得られた定義文の累計数がある閾値 (7 件) を越えたら、終了する。閾値未満の場合は、 Q の先頭要素を取り除き、この手順を最初から実行する。

長単位語の場合の表層文字列の生成は、次のような優先度で行なう。

- (1) 末尾の助動詞等の削除

できるだけ長い形態素列からの生成を優先する。すなわち、末尾の助動詞等は、一つずつ順に削除する。
- (2) 出現表記・語彙素表記・読み

同じ形態素列からの生成では、出現表記、語彙素表記、読みの順とする。

たとえば、長単位語〈耐えられなかった〉からは、「耐えられなかった」、「たえられなかった」、「耐えられない」、「たえられない」、「耐えられる」、「たえられる」、「耐える」、「たえる」という文字列を、順に生成する。なお、この例の場合、語彙素表記は出現表記と同一のため、省略した。

長単位語列の場合の表層文字列の生成では、末尾の長単位語に対して上記 (1) の正規化を適用するとともに、助詞の正規化 (「は」や「も」を「が」に置換する) も行なう。なお、長単位語列の場合は、読みしか生成しない。たとえば、長単位語列〈はかは行かず〉からは、「はかはいかず」、「はかがいかず」、「はかはいく」、「はかがいく」を、順に生成する。

*6 おおよそ、末尾の助詞を削除した文節に対応する。

表 8 語句問題ソルバーの性能

センター過去問 (2001–2013 年の奇数年)	29/42	70%
代々木ゼミナール問題集 [5]	19/21	90%
駿台予備校問題集 [6]	18/21	86%
合計	66/84	75%

表 9 正解できなかった語句問題

(ア) 依然として 固く		
3	0.341	真面目な 様子を崩すことなく
5	0.341	あいかわらずこぼったままで
(ウ) 毒々しい 笑いがふきあげた		
2	0.261	悪意 の感じられる笑みが浮んだ
5	0.268	腹黒さを隠しきれない笑い顔をしていた

□ が正解、もう一方がソルバーの出力

以上のような表層文字列の生成は、設問傍線部に現れる表現の形式・表記が、必ずしも辞書の見出し語の形式・表記と一致しないことに対処するためのものである。

4.3.5 照合スコア

辞書検索によって得られた定義文 p と選択肢 c の照合スコア $\text{score}(p, c)$ の計算には、ある集合 E の要素を単位とする、方向性を持ったオーバーラップ率 (付録 A の式 (A.1)) を用いる。

$$\text{score}(p, c) = \text{overlap_ratio}_D(E; p, c) \quad (2)$$

ここで、集合 E として、以下の 2 種類を使用する。

- (1) G : 内容語の集合。ただし、「動き (名詞)」と「動く」のような、同一語根を持つ派生語の一部を同一視する。
- (2) C : 文字集合 (句読点を含まず)

選択肢の得点は、基本的に、前者 (内容語の集合 G) を用いて計算するが、すべての選択肢の得点が 0 となった場合 (つまり、内容語が全く一致しなかった場合) には、後者 (文字集合 C) に基づく得点を採用する。

4.4 ソルバーの性能と問題点

今回の挑戦の前に確認したソルバーの性能を表 8 に示す。この表に示すように、ソルバーの性能は 75% である。センター過去問より、予備校模試の方が、少し成績が良い。

本ソルバーは、語句の意味を国語辞書に求めている。しかしながら、機械的な方法で「適切な項目を引く」ということは、それほど容易ではない。また、得られた定義文と選択肢との照合は、まったく単純な方法に依存している。それらの短所を、複数の国語辞書を用いることで (無理矢理) 補っている。さらに、本ソルバーでは、「本文中における意味」を全く考慮していない。つまり、複数の語義を持つ場合、最適な語義を選ぶことができない。

これらの問題は、ソルバー開発時に明らかであったが、代ゼミ模試の語句問題に対しては比較的成績が良かった (19/21) ので、問題の解決には取り組まなかった。

今回の模試で正解できなかった小問 2 問を表 9 に示す。

本文: A インターネットという技術が作り上げたひとつの功績は、空間と時間というものを無限に圧縮したことである。それによって、人間は遠方にいる友人とも、あたかも目の前に (ア)タイザしているかのごとく、時間差なしでコミュニケーションできるようになった。しかし、それなら無線でも、電話でも可能であったことである。それ以上に重要なことは、不特定多数、無限大の人々と、一瞬のうちにつながる事ができるということであり、このことが様々な新しいビジネスの可能性を切り拓いてきたわけである。...

問2 傍線部A「インターネットという技術が作り上げたひとつの功績は、空間と時間というものを無限に圧縮したことである」とあるが、どうしてこのようなことが可能になったのか。その説明として最も適当なものを、次の1~5のうちから一つ選べ。

- 1 インターネット環境において、テクノロジーは時間と空間という物理的な制約を乗り越え、人間という存在を否定するほどの無際限の発展を遂げつつあったから。
- 2 テクノロジーの進歩に伴うインターネット技術の飛躍的な向上によって、不特定多数の人々の間で瞬時に情報を共有することが可能になったから。
- 3 以前は実現不可能とされていた、時間と空間に縛られない自由な存在という人間の理想的なイメージが、テクノロジーの進化によって現実化されたから。
- 4 時間と空間の制約に縛られずに、ありとあらゆることを可能にする、人知を超えた能力を手に入れたという人間の欲望がテクノロジーの進展を促したから。
- 5 不特定多数の人々と瞬時につながることを可能にするテクノロジーの進歩によって、人間は自己の限界を超えたいという欲望を抱くようになったから。

図3 評論傍線部問題の例 (2013年模試の第1問問2)

それぞれの小問の下線部が、辞書引きに使用された部分であり、ゴシック体が辞書の定義文と一致した語である。この結果より、内容語の抽出法に問題があること(「まま」や「いた(いる)」が内容語と判断されている)、および、辞書定義文との間で適切な内容語(「あいかわらず」や「こわばった」)の一致がとれていないこと、など、設問傍線部と辞書定義文の照合の実装に大きな問題を抱えていることがわかる。

5. 評論傍線部問題ソルバー (E0-E4)

5.1 評論傍線部問題

第1問の評論では、何からの評論から抜き出された一部が「本文」として示される。評論傍線部問題は、本文中のある箇所(傍線部)を参照し、その箇所の内容や理由を問う問題である。評論傍線部問題の例を図3に示す。

我々は、この問題を解く一つの方法として、本文照合法を提案した[2]。フォーマルランに使用した5つの評論傍線部問題ソルバーのうち、E0, E1, E2の3つは、本文照合法をそのまま用いている。これに対して、E3とE4のソルバーでは、新たな機能を導入している。

5.2 本文照合法

傍線部問題を解く本文照合法は、

- 正解選択肢を選ぶ根拠は、本文中に存在する [8], [9],
 - 意味的に似ているテキストは、表層的にも似ていることが多い(根拠部分と最も似ている選択肢を選ぶ)、
- という考え方(仮説)に基づく方法で、次のような方法で出力する選択肢を決定する。

- (1) 入力: 本文, 設問, 選択肢集合が与えられる。
- (2) 照合領域の決定: 選択肢と照合する本文の一部を定める。

- (3) 選択肢の事前選抜: 考慮の対象外とする選択肢を除外する。
- (4) 照合: 考慮の対象とする選択肢のそれぞれを(2)で定めた照合領域と比較し、照合スコアを求める。
- (5) 出力: 照合スコアの最も高い選択肢を解答として出力する

この方法は、ステップ(2)-(4)をどのように行なうかによって、多くのバリエーションが存在する。実装したシステムでは、それらをパラメータによって制御する。フォーマルランに使用した3つのソルバーのパラメータは、次のとおりである。

E0 P-b-0=C1-z

E1 P-a-0=C1-z

E2 P-b-c=C2

これらのパラメータのうち、等号の前方の3つの記号が照合領域の決定法を表す。先頭から、

- (1) 使用する単位 (P: 段落, S: 文)
- (2) 照合領域の先頭位置 (傍線部を含む単位からどれだけ前を考慮の対象とするか)
- (3) 照合領域の末尾位置 (傍線部を含む単位からどれだけ後を考慮の対象とするか)

を表す。位置は、基本的に、当該設問の傍線部を含む単位を0としたときの相対位置(整数)で指定する。この他に、特別な位置として、a(本文冒頭)、b(1つ前の設問の傍線部を含む単位)、c(1つ後の設問の傍線部を含む単位)、e(本文末尾)が指定できる。たとえば、P-b-0は、「1つ前の設問の傍線部を含む段落から、当該設問の傍線部を含む段落まで」を照合領域とすることを意味する。

等号の直後の記号は、照合スコアの計算で使用する集合を表す。照合領域 t と選択肢 c の照合スコアの計算には、ある集合 E の要素を単位とする、方向性を持ったオーバー

表 10 評論傍線部問題のサブタイプ

サブタイプ	問われていること	タイプ判定の文字列正規表現	パラメータ
E	傍線部の言い換え・説明	/とあるが、(それは これは .*とは)?どういうことか./ /とあるが、(その それについての)説明として最も/ /とあるが、その.*を説明したものとして/ /とあるが、(それは これは)?(どういう どのような).*か. その説明として/ /とあるが、「.+」の説明として最も適当なものを/	P-b-0-x=C1
R	傍線部の理由説明	/とあるが、(それは (こう そう そのように)(い 言)えるのは)?なぜか./ /とあるが、なぜ(そう そのように このように 「.*」と) (言 い)える (言 い)う)のか./	P-b-0=C1
A	筆者の考え	/筆者/	P-b-c-j=C1
Unknown	上記以外		P-b-c-jw=C1

ラップ率 (式 (A.1)) を用いる.

$$\text{score}(t, c) = \text{overlap_ratio}_D(E; t, c) \quad (3)$$

集合 E として, 文字集合 C ($C1$), 文字 bigram 集合 C^2 ($C2$), 形態素出現形集合 W (W), 形態素原形集合 L (L) のいずれかを用いる. パラメータ $C1$ は, 文字オーバーラップ率, すなわち, 「選択肢 c に出現する文字のうち, 照合領域 t にも出現する文字の割合」を表す.

最後の記号 z は, 存在する場合, 選択肢の事前選抜を行なうことを表す. この事前選抜は, 「最も仲間外れ(他の選択肢と似ていない)の選択肢 1 つを除外する」. 具体的には, それぞれの選択肢 c_i に対し, 他の選択肢 c_j との間の方向性を持たない文字オーバーラップ率^{*7} $\text{overlap_ratio}_B(C; c_i, c_j)$ の平均値を求め, その値が最も低かった選択肢を除外する.

5.3 ソルバー E3

評論傍線部問題には, いくつかのサブタイプが存在する. ソルバー E3 では, 設問文に対する文字列照合により, このサブタイプを判定した後, 選択肢と照合する本文の領域を決定する.

ソルバー E3 で採用した 4 種類のサブタイプと, タイプ判定に使用する文字列正規表現, および, 本文照合法のパラメータを表 10 に示す. 以下では, 追加されたパラメータ x, j, w の機能について説明する.

5.3.1 パラメータ x (照合領域の修正)

パラメータ x は, タイプ E のために実装した新機能である. 傍線部の言い換えが問われるタイプ E の照合領域は, 傍線部周辺の比較的狭い領域でよいと思われる. そこで, デフォルトの照合領域 P-b-0 を, 次のような方法で変更する.

- (1) 5 つの選択肢を結合した文字列 c^* を作成する.
- (2) 当該傍線部を含む段落 p_0 と文字列 c^* の間で, 方向性を持たない文字オーバーラップ率 o_0 を計算する.

$$o_0 = \text{overlap_ratio}_B(C; c^*, p_0) \quad (4)$$

^{*7} 付録の式 (A.2) に定義を示す.

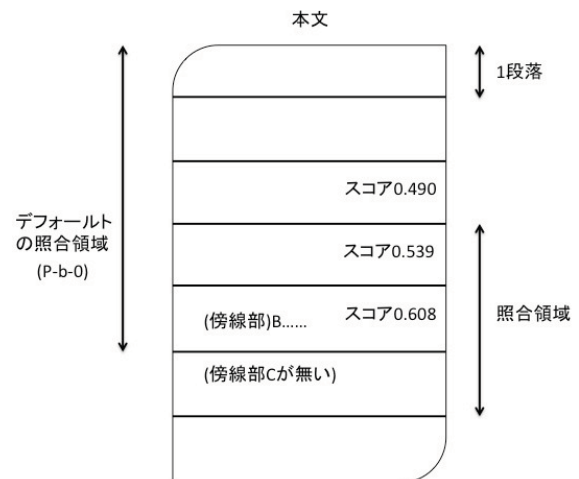


図 4 パラメータ x による照合領域の変更

- (3) 段落 p_0 から, 段落単位に本文を遡る. その過程において, それぞれの段落 p_{-i} に対し,

$$o_{-i} = \text{overlap_ratio}_B(C; c^*, p_{-i}) \quad (5)$$

を計算する.

$$o_{-i} < o_0 - 0.08 \quad (6)$$

が成立した場合, その段落 p_{-i} とそれより前の段落は, 照合領域から除外する.

- (4) 当該傍線部を含む段落 p_0 の直後の段落 p_1 に, 次の設問の傍線部が存在しない場合, 段落 p_1 を照合領域に含める.

照合領域の変更例を図 4 に示す.

5.3.2 パラメータ j と w (接続表現の利用)

評論文では, 論の展開に接続表現が重要な役割を果たす. そのため, 設問を解く際にも, 接続表現が重要な手がかりとなる [8], [10]. 我々は, 表 11 に示すような 6 種類の接続表現を設定し, これらの表現を, 照合領域の決定に利用することを考えた. (「並列・添加」は, 最終的には利用しなかった.)

パラメータ j は, 「逆接」と「転換」を利用する. 段落の先頭にこれらのタイプの接続表現が現れる場合, その直前

表 11 6 種類の接続表現

タイプ	例
並列・添加	同時に、さらに
因果関係 (理由)	なぜなら、というのも
因果関係 (結果)	それゆえ、したがって
逆接	しかし、ところが
換言	つまり、すなわち
転換	ところで、では

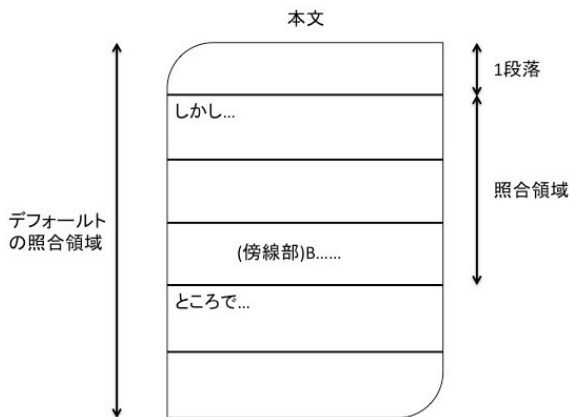


図 5 機能 j による照合領域の変更

の段落境界は、話の比較的大きな切れ目となる可能性が高い。パラメータ j は、デフォルトの照合領域 (P-b-c) から、この境界を排除する。すなわち、

- (1) 当該傍線部を含む段落より前に、「逆接」「転換」の接続表現で始まっている段落があれば、その段落より前は照合領域に含めない。
- (2) 当該傍線部を含む段落より後に、「逆接」「転換」の接続表現で始まっている段落があれば、その段落以降は照合領域に含めない。

パラメータ j による照合領域変更の例を図 5 に示す。

パラメータ w は、「換言」「因果関係 (理由)」「因果関係 (結果)」の 3 種類の接続表現を利用する。板野 [10] は、これらのタイプの接続表現で始まる文は、解答の根拠として重要であると述べている。それに基づき、パラメータ w は、これらのタイプの接続表現で始める文にボーナスを与える。すなわち、

- (1) 照合領域 t の中から、「換言」「因果関係 (理由)」「因果関係 (結果)」の接続表現で始まる文を全て抜き出す。これらをすべて結合した文字列を $imp(t)$ と書く。
- (2) 各選択肢の照合スコア (式 (3)) に、以下のボーナスを加算する。

$$bonus(E; t, c) = overlap_ratio_D(E; imp(t), c) \quad (7)$$

なお、集合 E は、照合スコアの計算で使用する集合 E を、そのまま用いる。

5.4 ソルバー E4

センター試験の第 1 問 (評論) では、ほとんどの場合、傍

表 12 評論傍線部問題ソルバーの性能

	問題数	正解数				
		E0	E1	E2	E3	E4
センター過去問	48	25	24	16	30	20
代々木ゼミナール問題集 [5]	28	10	12	13	12	12
合計	76	35	36	29	42	32

線部問題が 4 問出題される。それぞれの設問の解答の根拠部分が重ならなると仮定すれば、本文を 4 つのブロックに分割し、それぞれを各設問の照合領域として使用するというアイデアが生まれてくる。

ソルバー E4 は、このような考えに基づくソルバーである。各ブロックは、傍線部を 1 つだけ含み、ブロック境界の候補として、段落境界のみを考慮することにすれば、隣接する 2 つの傍線部の組のそれぞれに対し、その間に存在する段落境界のいずれかを、ブロック境界として採用すればよい。具体的には、次の方法でブロック境界を定める。

- (1) 候補となる段落境界において、直前段落を p_- 、直後段落を p_+ とし、 $overlap_ratio_B(K; p_-, p_+)$ を計算する。ここで、 K は、漢字カタカナ文字集合を表す*8。
- (2) この値が最小値をとる段落境界を、ブロック境界として採用する。

この方法は、表層的に最も異なる段落間の境界をブロック境界として採用することを意味する。

5.5 ソルバーの性能と問題点

今回の挑戦の前に確認した各ソルバーの性能を表 12 に示す。この表に示すように、センター過去問に対してはソルバー E3 が、代ゼミ模試に対してはソルバー E2 が最もよい成績を示している。全体としては、センター過去問より、代ゼミ模試の方が、成績が悪い。

実際には、ここに示した 5 つのソルバー以外にも、かなりの数のソルバーの性能を測定した。その結果、我々は、「センター過去問と代ゼミ模試では問題の傾向が異なる」という認識に至り、センター過去問を想定した本文照合法である E1、代ゼミ模試を想定した E2、センター過去問にチューニングした E3、本文を 4 分割するという新しいアイデアに基づく E4 を、フォーマルランに用いるという決定を行なった。(E0 はおまけである。)

フォーマルランの結果は、すでに表 1 に示したように、E3 が 1 問正解しただけという、惨澹たる結果であった。どうしてこのような結果になったかという、それは、「正解選択肢が本文記述と表層的に似ていなかったから」に尽きる。表 1 より、5 つのソルバーは、それぞれの設問で、最大でも 2 種類の選択肢しか選んでいない。これは、照合領域を変えても、表層類似度 (オーバーラップ率) が最大とな

*8 漢字とカタカナは、主に内容語を表すために使用されるので、単なる文字オーバーラップ率よりも、若干、内容的側面を重視した値となる。

【二】 次の文章は、小林勝の小説『軍用露語教程』の一節である。予科士官学校で軍隊式の生活をおくっている香取潔は、或る日の夕方、外出から学校へ戻る途中の坂道で、ロシア語教官の峰に出会った。峰は規律にうるさい教官であったが、潔はその峰からロシア語の発音をほめられたことがある。峰は教室では絶対にみせたことのない親し気な微笑を浮かべながら、「今度の外出の時はおれの家へ来んか。」と声をかけてきた。それに対して潔は考える余裕もなく、「ありがとうございます、教官殿。」と返事をした。その後、潔は学校へ戻って夜の自習時間を迎えた。以下はそれに続く部分である。これを読んで、後の問い（問1～6）に答えよ。

本文: ... 十二時少し前に、潔は峰の家を辞した。二人はぼつぼつと話を交したが、潔はろくに覚えてもいない。A 何故だか、自分が不当にも、手ひどく傷つけられた気がする。峰が憎かった、しかし、峰のどこを、何を憎めばよいのか。袴をはいていたことか、それならば、峰が言った事は区隊長が常々言っていることと変りはしない。疎開作業で、もうもうと白い埃のまじり道を、唇をぎゅつとひきしめて潔は歩いて行った。雑嚢の中には、行く時にははいていなかったものが、はいていることにふと潔は気付いた。お前はロシア語が出来るから、これでも読んでみる、と峰が言って貸してくれたかなり分厚い露語の原書だった。...

問2 傍線部A「何故だか、自分が不当にも、手ひどく傷つけられた気がする」とあるが、なぜそのような気がしたのか。その理由として最も適当なものを、次の1～5のうちから一つ選べ。

- 1 ロシア語の教官から自宅に招かれたのだから、ロシア語に関する何か面白い話でも聞かせてくれるものと期待してわざわざやって来たのに、峰はつまらない教訓的な話を聞かせるばかりなので、ひどく落胆させられたから。
- 2 峰はロシア語の発音をほめてくれた教官であり、自分が出掛けて行けば喜んでくれると思っていたのに、峰が学校にいる時と変わらぬ淡々とした態度をとり続けるので、自分の思い上がりに気付かされて、恥ずかしい思いをしたから。
- 3 峰の方から親しそうな態度で家に誘ったのだから暖かく迎えられるだろうと予想していたのに反して、峰が教官としての立場を崩さず、少しも打ちとけた様子を見せなかったことに納得がいかず、理不尽に感じられたから。
- 4 自分のロシア語の才能を見込んでくれたからこそ峰は自宅に招いてくれたはずであり、自分が尋ねて行けば大いに歓迎してくれるものと信じて疑わなかったが、説教ばかりを得々と聞かされ屈辱を味わわれたから。
- 5 峰がいつになくやさしい顔で家に来ないかと誘ってきたので、教官と親しくなれると思ってやって来たのに、峰が学校にいる時よりも厳しい表情で自分に対して威圧的な態度をとり続けるので、訳がわからず困惑させられたから。

図 6 小説傍線部問題の例 (2013 年模試の第 2 問 問 2)

る選択肢が、ほとんど変わらなかったことを意味する。

代ゼミの講師から伺った話によれば、問題を作成する際、

- 本文が難しい場合は選択肢をやさしく、本文がやさしい場合は選択肢を難しくすることによって、問題の難易度を調整する、
- 「選択肢を難しくする」ひとつの方法は、選択肢の表現に、本文とは異なる表現(言い換え)を用いる

とのことである。現在のソルバー群は、選択肢において、本文の根拠部分が全面的に言い換えられている場合、ほとんど無力である。

6. 小説傍線部問題ソルバー (N0-N3)

6.1 小説傍線部問題

第2問の小説では、何からの小説の一部が取り上げられる。評論との一つの違いは、問題の冒頭に、その後に示される「本文」の背景説明が示されるという点である。もう一つの違いは、小説傍線部問題では、傍線が付けられたある箇所を参照して、主に、その状況における登場人物の心情・人物像・行動の理由などが問われる点である。小説傍線部問題の例を図6に示す。

このような違いがあるにもかかわらず、小説傍線部問題は、形式的には、評論傍線部問題と同形式の問題とみなすことができるため、その解法として、先に述べた本文照合法を用いることが可能である。実際、今回のフォーメランに使用した4つの小説傍線部問題ソルバーのうち、N1を除く3つは、本文照合法に基づくソルバーである。使用

したパラメータは、次のとおりである。

N0 P-b-0=C1-z (ソルバー E0 と同一)

N2 P-a-0=C1

N3 P-b-0=C1

6.2 ソルバー N1

新たに開発したソルバー N1 は、小説傍線部問題を、(a) 文脈の論理関係を問う論理型と、(b) 登場人物の心理を問う心情型に分け [8]、それぞれに別の解法を適用する。具体的には、

- (1) 設問文に出現する表現に基づき、設問を論理型と心情型に分類する。分類に使用する表現を表 13 に示す。これらの表現が出現しなかった場合は、心情型に分類する。
- (2) 設問が論理型の場合は、本文照合法 S-a-0=W を適用して解く。
- (3) 設問が心情型の場合は、まず本文照合法 P-0-0=W を適用し、照合スコアを計算する。次に、後に述べる感情スコアを計算する。この2つのスコアの合計を選択肢の得点とし、最も得点の高い選択肢を解として出力する。

6.2.1 感情語辞書

感情スコアは、照合領域と選択肢に出現する感情語に基づいて計算する。スコア計算で考慮する感情語は、感情語辞書として定義する。

作成した感情語辞書の一部を、表 14 に示す。この辞書

表 13 小説傍線部問題の設問サブタイプ

サブタイプ	表現
心情型	なぜ, 心, 表現, 受け止め, 受けとめ
論理型	どのようになったこと, 違い, どのようなこと, ということ, 人物, どのような行為, どのような面

表 14 感情語辞書の一部

カテゴリ	例
喜び	嬉しい, 満足, 楽しむ
驚き	びっくり, 仰天, 動揺
好き	愛しい, 夢中, 慕う
昂り	興奮, 感動, 必死
怒り	腹立たしい, 苛立ち, 憤る
哀しみ	切ない, 残念, 嘆く
罪	後ろめたい, 謝罪, 悔やむ
恐怖	恐ろしい, 不安, 怯える
嫌悪	嫌い, 不満, 困る
恥	恥ずかしい, 屈辱, 照れる

は, センター過去問 13 回分と代ゼミ模試 1 回分から, 感情を表す語を人手で抽出し, それらを 10 種類の感情カテゴリに分類することによって作成した. 総語数は 467 語である.

その後, この感情語辞書を, 形態素解析用辞書の IPA 辞書 (ipadic-2.7.0) に統合した. この辞書を用いて形態素解析を行なうことにより, 同時に, 感情語を同定する.

6.2.2 感情スコア

照合領域 (P-0-0, すなわち当該傍線部を含む段落) と選択肢をそれぞれ形態素解析し, 感情語を抽出する. 照合領域 t と選択肢 c において, 同一の感情ラベルをもつ感情語が m ペア見つかった場合の感情スコアを, 以下のように定める.

$$e_score(t, c) = 0.15 m(t, c) \quad (8)$$

6.3 ソルバーの性能と問題点

今回の挑戦の前に確認した各ソルバーの性能を表 15 に示す. この表を表 12 と比較すると, 小説傍線部問題の正解率は, 評論傍線部問題と比べ, かなり低いことがわかる. ソルバー N1 は, センター過去問では, 他のソルバーに比べてかなり成績が良い. これは, 感情語辞書を, 主にセンター過去問に基づいて作成したことによると考えられる. 代ゼミ模試においても, 感情語辞書の効果は, 若干観察される.

フォーマルランの結果は, すでに表 1 に示したように, ソルバー N2 (本文照合法 P-a-0=C1) が 2 問正解したのみであり, 感情スコアを導入した N1 は, 1 問も正解できなかった. 感情語辞書の網羅性と感情カテゴリの弁別能力 (正解を導くのに十分な詳細度で感情を区分できるか) に問題があることは認識していたが, N1 の全問不正解は予想

表 15 小説傍線部問題ソルバーの性能

	問題数	正解数			
		N0	N1	N2	N3
センター過去問	50	8	24	10	8
代々木ゼミナール問題集 [5]	28	4	9	7	6
合計	78	12	31	17	14

外であった. 本文照合法の N2 と N3 は, フォーマルランの直前に採用を決めた, いわばバックアップ用のソルバーである. 平均的に 1 問正解するレベルであるが, 今回, N2 が 2 問正解したのは好運であった.

7. 今回の挑戦のまとめ

今回のセンター模試への挑戦では, 『国語 (現代文+古文*9)』以外に, 『世界史』, 『日本史』, 『政治・経済』, 『数学 I・数学 A』, 『数学 II・数学 B』, 『英語』, 『物理 I』を受験した. 偏差値では, 『国語 (現代文)』の成績 (偏差値 44.7) は, 『英語』 (偏差値 41.0) に次いで, 悪い成績であった.

図 7 に, 代々木ゼミナールによる『国語』の講評のうち, 現代文に関連する部分を示す. これは, 表面上は, ソルバーに対する講評であるが, 実質上は, 我々の取り組みに対する講評である. このような講評 (という名の励まし) をいただいたことに対し, この場を借りて, 感謝の意を表したい.

今年の挑戦を総括すると, それは, 「文章の『意味』に立ち入らず, どのくらい解けるか」ということに取り組んだということになる. 確かに, フォーマルランの成績は芳しくなかった. しかしながら, 現在のソルバー群の成績は,それほど悲観的なものではない. センター過去問 11 回分および代ゼミ模試 7 回分 [5] の成績を, 今回のフォーマルランと同様の方法で採点すると, 表 16 のようになる. センター過去問の平均点は 60 点, 代ゼミ模試 (7 回分+今回のフォーマルラン) の平均値は 53 点である.

この表からわかることの一つは, 今回のフォーマルランの問題 (FR) は, ソルバー群が非常に苦手とする問題だったということである. つまり, 「運が悪かった». しかし, 日本の大学入試は一発勝負であり, 「運が悪かった」では済まされない. 我々が今回の挑戦から学んだことの一つは, 「平均点を上げるよりも, 最低点を上げるの方が, より重要だ」ということである.

同時に, 表 16 は, いささか気になる傾向を示している. それは, 最近のセンター試験 (2011 年や 2013 年) の成績が悪いという事実である. 2013 年の本試験では, 第 1 問評論で小林秀雄の文章が出題され, それが『国語』の平均点を 10 点押し下げたとされている. それに合わせて, 代々木ゼミナールも模試の難易度を上げたのであれば, 今回の模試の成績は, 順当といえるのかも知れない.

*9 「古文」は, 国立情報学研究所の横野光氏が担当した.

【概評】

- ・ 第1問の現代文(評論)は18点(得点率36%)。漢字問題は満点ですが、読解問題が5問中1問しか解けなかったのが、得点が伸び悩んだ原因です。
- ・ 第2問の現代文(小説)は24点(得点率48%)。まずまずの出来です。ほぼ現役生の平均レベルに近い得点ですが、心情を問う問題の出来にムラがあるのが気になります。

【特徴的な設問について】

- ・ 第1問(評論)の漢字は全問正解です。今回の漢字問題はセンター本番よりも、やや難度が高く、正答率が5割を切っているもの(「享受-享楽」46.6%)もあり、訓読みから音読みへの変換を求めるもの(「従う」にもとづいて「従事」を選ぶ)もあったことを考えると、全問正解は立派です。
- ・ 第1問の間2では、「インターネットが空間と時間の圧縮を可能にした理由」が問われています。「人知をこえた能力を手に入れたという人間の欲望がテクノロジーを進展させた」という4が正解ですが、東ロボくんは「テクノロジーの進歩が人間に自己の限界を超えたいという欲望を抱かせた」という5を選んでいました。二つの要素はともに本文に書かれていますが、4は「人間の欲望→テクノロジーの進歩」であるのに対して、5は「テクノロジーの進歩→人間の欲望」となっています。東ロボくんはどちらがより根本的な原因であるかの判別ができなかったようです。
- ・ 第2問(小説)の語句の意味を問う問題(問1)では、教官の表情が「依然として固い」という表現の解釈が求められています。「表情が固い=こわばっている」とした5が正解ですが(正答率68.4%)、東ロボくんは「固い=真面目な様子」と取っています。傍線の直前にある「堅苦しさ」、直後の「どのような微妙な心の動きをも封じこめ」という表現を踏まえた文脈的な理解ができなかったようです。

【学習アドバイス】

- ・ 漢字力は高いレベルで安定しています。漢字問題は確実な得点源ですから、現在のレベルを維持しましょう。
- ・ 評論問題の課題は、論理的な読解力の養成です。本文にAとBという二つの要素が含まれている場合、AとBの関係(例:どちらが原因で、どちらが結果か)を正確に理解する必要があります。
- ・ 小説は、心情を問う問題が中心です。選択肢の中で本文に書かれていない内容、本文に反する表現を的確にチェックして、誤りの選択肢を落とす「消去法」の練習をしましょう。

図7 代ゼミによる講評(現代文に関連する部分のみ)

いずれにせよ、センター模試『国語』現代文では、

- 毎回、違う文章が本文として出題され、(漢字問題と語句問題以外は)異なる内容が問われる。つまり、問題の類型性は低い。
- 古い過去問は、あまり参考にならない可能性が高い。特に、学習指導要領が変更になれば、その可能性は高い。

このため、大量の実例から規則性を見つけて解くという機械学習のアプローチは、設問を解くというレベルでは、おそらくほとんど無力である^{*10}。利用できる実例は、せ

^{*10} 感情語を同定するなどの要素技術の実現には使用できる。

表16 センター過去問および代ゼミ模試の点数分布

試験	センター過去問			代ゼミ模試			
	評論	小説	計	試験	評論	小説	計
2001M	45	34	79				
2001S	37	31	68	Y1	38	30	68
2003S	37	31	68				
2005S	42	24	66				
2007M	42	24	66				
2009M	22	37	59				
2003M	30	28	58	Y7	26	32	58
				Y2	34	22	56
				Y4	26	30	56
				Y6	36	20	56
2011M	34	21	55				
2005M	26	27	53	Y5	30	20	50
2013M	26	21	47				
2013S	34	11	45				
				FR	18	24	42
				Y3	32	6	38
平均	34	26	60	平均	30	23	53

M:本試験, S:追試験, Y1-7:代ゼミ模試問題集 [5],

FR:フォーマルラン(代ゼミ模試2013年第1回)

いぜい数十であり、フォーマルランでは、たった数問で性能が評価される。このような厳しい問題設定に対する新しい戦術を考える必要がある。

謝辞 本研究では、国立情報学研究所のプロジェクト「ロボットは東大に入れるか」から、データの提供を受けて実施した。模試の採点と講評を担当した代々木ゼミナールに感謝する。本研究では、『現代日本語書き言葉均衡コーパス(BCCWJ)』(DVD版)を使用した。本研究の一部は、JSPS 科研費 24300052 の助成を受けて実施した。

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] 佐藤理史, 加納隼人, 西村翔平, 駒谷和範: センター試験『国語』現代文の傍線部問題を解くベースライン法, 情報処理学会研究報告 2013-NLP-212 No.5, 情報処理学会 (2013).
- [3] 佐藤理史: テキストの難易度と語の分布, 情報処理学会研究報告 2013-NLP-213 No.6, 情報処理学会 (2013).
- [4] 佐藤理史: 日本語クロスワードパズルを解く, 情報処理学会研究報告 2002-NLP-147-11, 情報処理学会 (2002).
- [5] 代々木ゼミナール: 国語 2014 年版(大学入試センター試験実戦問題集), 代々木ライブラリー (2013).
- [6] 全国入試模試センター: 大学入試センター試験 実戦問題集 国語 2014 (大学入試完全対策シリーズ), 駿台文庫 (2013).
- [7] 板野博行: 板野博行のターゲット現代文(改訂版) 2 センター突破編, 旺文社 (2010).
- [8] 船口 明: きめる! センター国語現代文, 学研教育出版 (1997).
- [9] 板野博行: ゴロゴ板野の現代文解法 565 パターン集 増補改訂版, アルス工房 (2009).

- [10] 板野博行：ゴロゴ板野のセンター現代文解法パターン集，
星雲社 (2010).

付 録

A.1 文字列のオーバーラップ率

2つの文字列 t_1, t_2 に対する，ある集合 E の要素を単位とするオーバーラップ率を，次のように定義する．

$$\text{overlap_ratio}_D(E; t_1, t_2) = \frac{\text{overlap}(E; t_1, t_2)}{\sum_{x \in E} f(x, t_2)} \quad (\text{A.1})$$

$$\text{overlap_ratio}_B(E; t_1, t_2) = \frac{2 \text{overlap}(E; t_1, t_2)}{\sum_{x \in E} (f(x, t_1) + f(x, t_2))} \quad (\text{A.2})$$

$$\text{overlap}(E; t_1, t_2) = \sum_{x \in E} \min(f(x, t_1), f(x, t_2)) \quad (\text{A.3})$$

ここで， $f(x, t)$ は，文字列 t における，ある集合 E の要素 x の出現回数を表す．

2つのオーバーラップ率のうち，文字列 t_2 の長さのみに正規化した式 (A.1) を，方向性を持つオーバーラップ率，2つの長さの和で正規化した式 (A.2) を，方向性を持たないオーバーラップ率と呼ぶ．

A.2 実装したシステム

本論文で述べたソルバーは，表 A.1 に示す独立した3つのシステムに組み込まれている．

表 A.1 フォーマルランで使した3システム

システム	内在するソルバー	実装者
システム S	K0, G0, E0, N0, F0	佐藤理史
システム K	E1, E2, E3, E4	加納隼人
システム N	N1, N2, N3	西村翔平