

学習データを必要としない 日本語記述式問題の自動採点システム

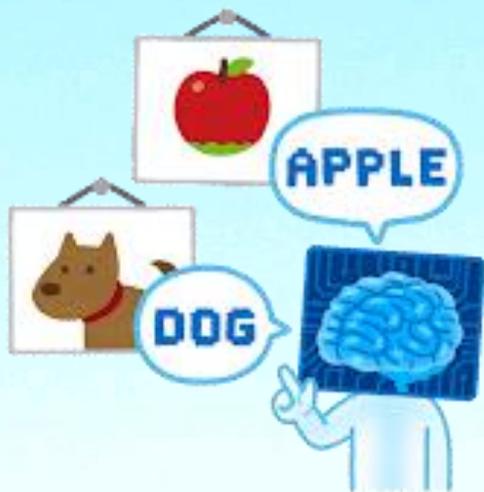
千葉大学 大学院工学研究院 情報科学コース
教授 須鎗弘樹

2022年2月15日

データの観点でみたAI

最近のAIを、**データの観点**で大きく分けると、

画像



言語



本発明

自然言語処理の応用

現場のニーズ＝人が書いた文章の評価

最もわかりやすい「人が書いた文章の評価」

記述式問題の答案の採点

実用化が想定される現場＝採点現場

- 学校
 - 模擬試験
 - 自宅でのオンライン学習
- (記述式問題は自己採点が難しい。)

従来技術の問題点(1/2)

AIでは、事前に、**大量の学習データが必要。**

(試験は未公開なので、学習データの事前準備不可。)

学習データが多いほど、精度は良い。

しかし、学習データの準備に、時間・人などのコスト大。

例：学習データとして、半分(実際よりも少なめ)を仮定。

1. 学校：1学年400名(多めを想定)

→ 全部、人が採点した方がトータルとしては早いかも。

2. 模擬試験：1万人(多めを想定)

→ 学習データの準備に非常に時間がかかる。

参考：大学入試共通テストの場合、毎年50万人以上の受験者

従来技術の問題点(2/2)

1. 学習データの**準備のためのコストが大きい。**
(時間と人的コスト)
 2. 学習データが**正しいとは限らない。**
(人の採点ミスが含まれることが多々ある。)
 3. 採点者が複数の場合、学習データに**バラツキを含む。**
(採点者が一人だと、準備に時間がかかる。)
 4. **学習データの選択**が自動採点の結果に影響あり。
(事前に点数の分布は不明。結局、全部採点の必要有。)
- など。

採点の現場では、学習データの準備は大きな問題

採点で考慮すべきこと

1. 情報を取り扱うときの3つの重要な視点

1. **効率性** できるだけコストを小さく **自動採点率↑**
2. **信頼性** できるだけ誤りなく **採点精度↑**
ただし、効率性と信頼性は、**トレードオフ**の関係
3. **安全性** 漏洩は許されない

2. 本発明(自動採点)で考慮したこと

信頼性を最優先 = 採点精度をできるだけ高く.
その範囲で、自動採点率をできるだけ高く.

本発明の利点

1. 学習データを**必要としない**.
⇒学習データの準備に関わる諸問題を回避できる.
(必要なもの: 答案, 模範解答, 採点基準)
2. **高い採点精度(最優先)と高い自動採点率**
(自動採点率: 自動採点できる答案の割合)
⇒確実に正誤判定できる答案を自動選択・自動採点
3. **一人目の採点者**として自動採点を設定できる.
4. **高速に実行できる**ため, 何度でも実行できる.
(それゆえ, 採点基準の見直しに対応可能)

本発明の自動採点システムの流れ

採点設定 (←採点基準, 模範解答)

キーワード, 文字数制限,
キーマス, NGワード,
NGエンドワード, 事前置き換えリスト

答案データ

自動採点

文字数チェック

事前置き換え

採点済み解答比較

キーワード, NGワード,
NGエンドワードチェック

キーマス比較

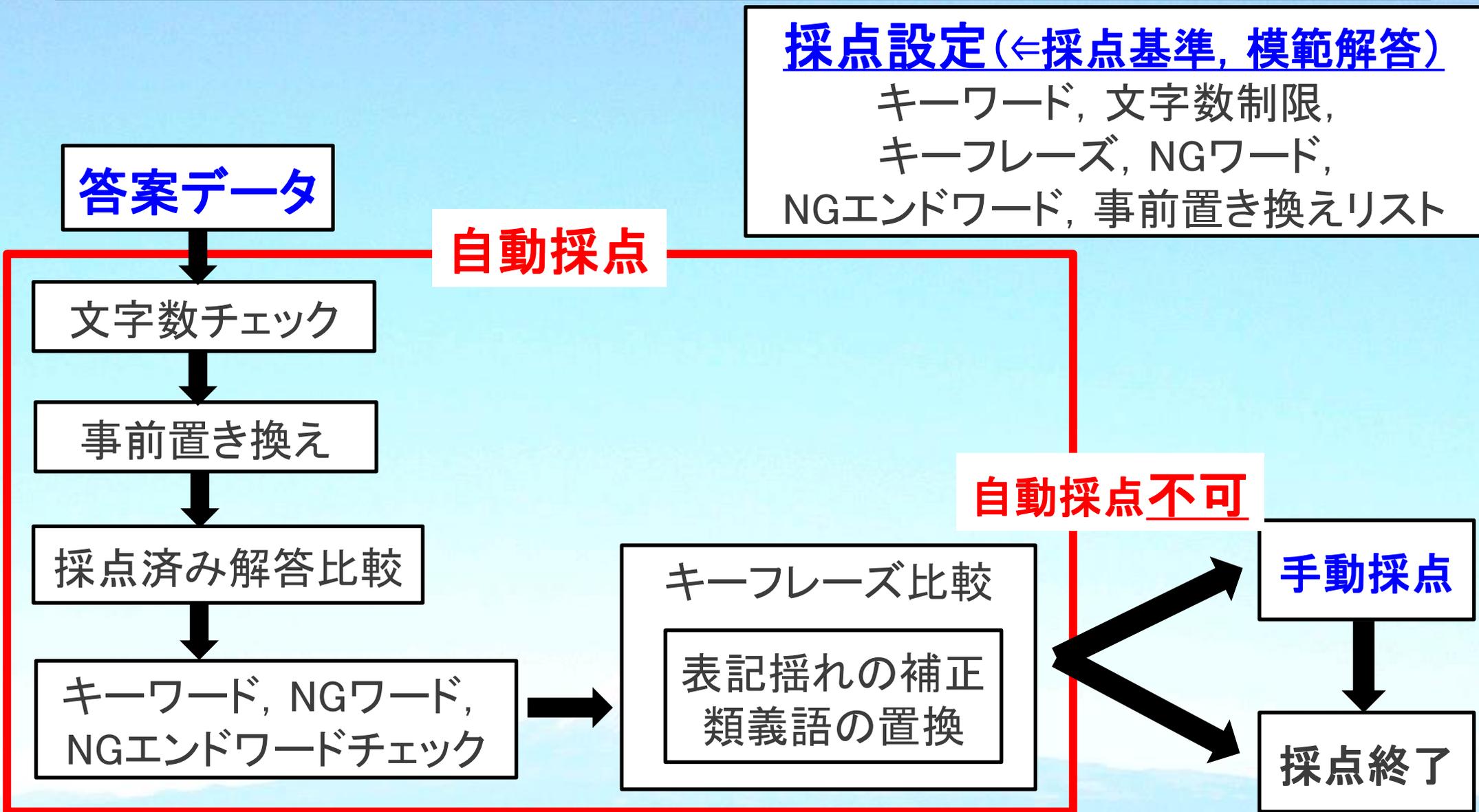
表記揺れの補正
類義語の置換

自動採点不可

手動採点

採点終了

自動採点可



具体的な適用例（中学理科）

中学生を対象に行われた模試の理科の答案データ
自由記述（字数制限なし） 約1,200人分

問題

イモリやカエルなどのなかまは、子と親とで呼吸のしかたが異なる。子と親の呼吸のしかたを、それぞれ簡潔に書きなさい。

模範解答例

子はえらで呼吸し、親は肺と皮ふで呼吸する。（21文字）

評価基準

自動採点率 (効率性の評価基準)

本発明のシステムで自動採点できた答案の割合

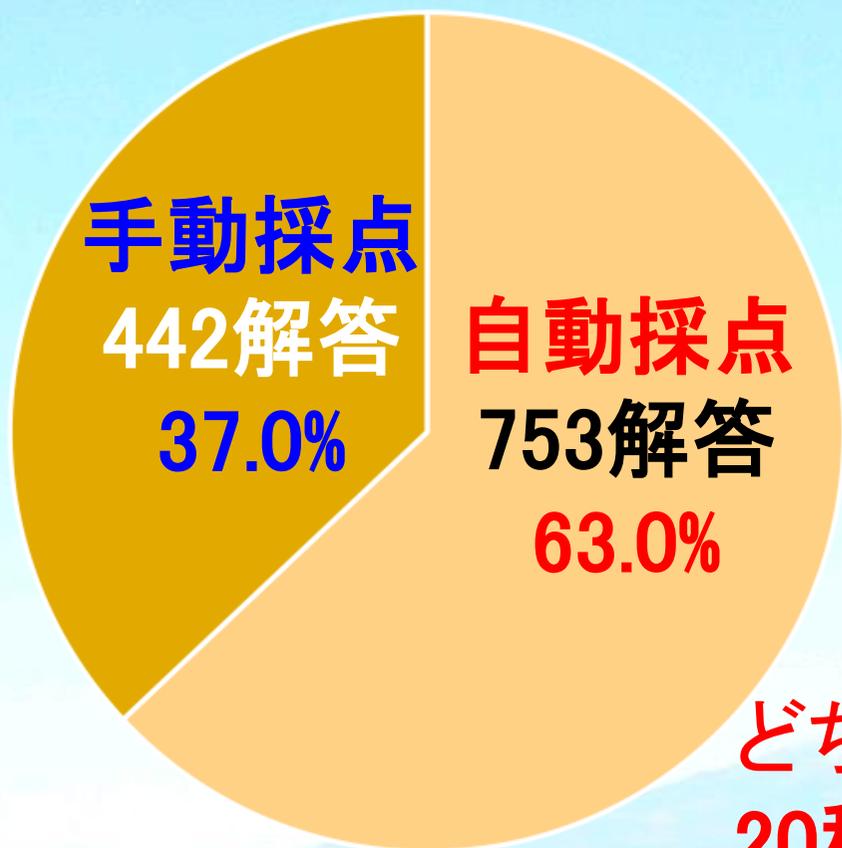
採点精度 (信頼性の評価基準, 最優先)

自動採点対象となった答案のうち, 正しい採点の割合

自動採点の結果（中学理科）

理科

採点精度 100%

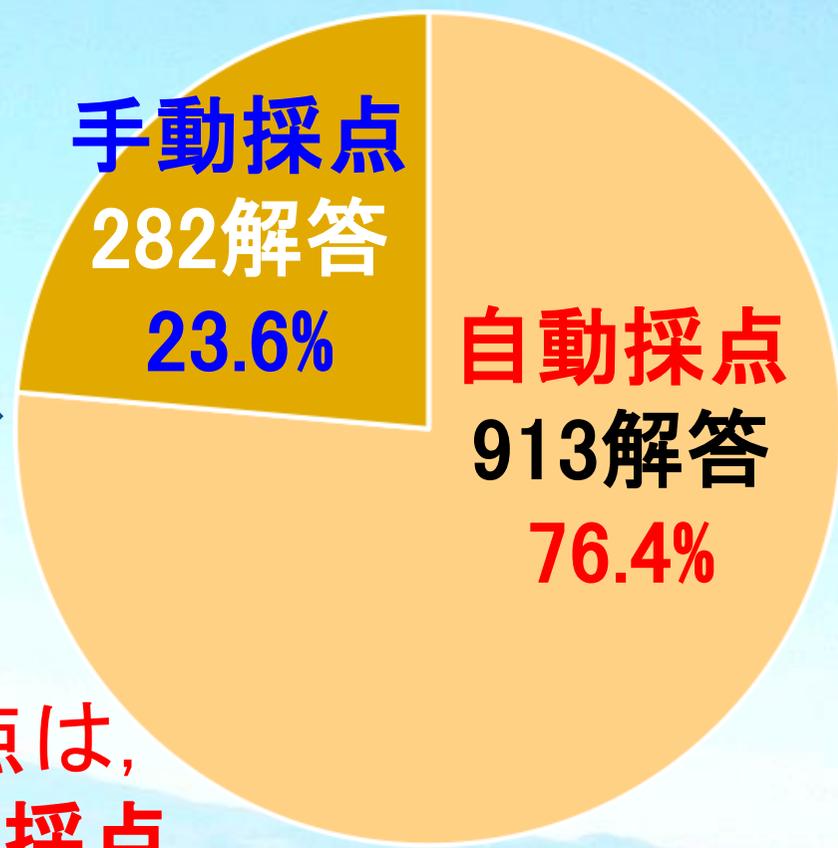


採点設定
の改善



理科

採点精度 100%



どちらも自動採点は、
20秒以下の高速採点

(Intel 第 6 世代 Core i5-6400, RAM:8GB, 2017年発売のPC)

ここからアルゴリズムの説明:

自動採点のための採点設定の準備

- キーワード 含まれていなければ不正解とする語
- NGワード 含まれていれば不正解とする語
- NGエンドワード 文末に含まれていれば不正解とする語
- 文字数制限 範囲内でなければ不正解とする
- 事前置換リスト 採点前に表記を置き換える語
- キーフレーズ 含まれていれば正解とする短文

事前設定ファイルの例

	A	B	C	D	E	F	G	H
1	文字数制限(最小)	文字数制限(最大)	事前置き換え(前)	事前置き換え(後)	キーワード	NGワード	NGエンドワード	キーフレーズ 1
2	5	50	えら	エラ	エラ	幼虫		子はエラで呼吸
3			鰓	エラ	肺	成虫		子はエラ呼吸
4			ひふ	皮膚		胚		小さい時はエラで呼吸
5			皮ふ	皮膚		肌		小さい時はエラ呼吸
6			皮フ	皮膚		肝		子がエラで呼吸

本発明の自動採点システム

採点設定 (←採点基準, 模範解答)

キーワード, 文字数制限,
キーマス, NGワード,
NGエンドワード, 事前置き換えリスト

答案データ

自動採点

文字数チェック

文字数制限内か?

事前置き換え

例: えら, 鰓 ⇒ エラで統一化

採点済み答案比較

採点済答案と同じなら,
同じ採点結果を返す.

キーワード, NGワード,
NGエンドワードチェック

キーマス比較

表記揺れの補正
類義語の置換

自動採点不可

手動採点

採点終了

自動採点可

キーフレーズ比較

答案の解答文とキーフレーズの比較

1. 形態素解析による分かち書き

例：子はえらで呼吸し，親は肺と皮ふで呼吸する。

⇒子/は/えら/で/呼吸/し/，

/親/は/肺/と/皮ふ/で/呼吸/する/。

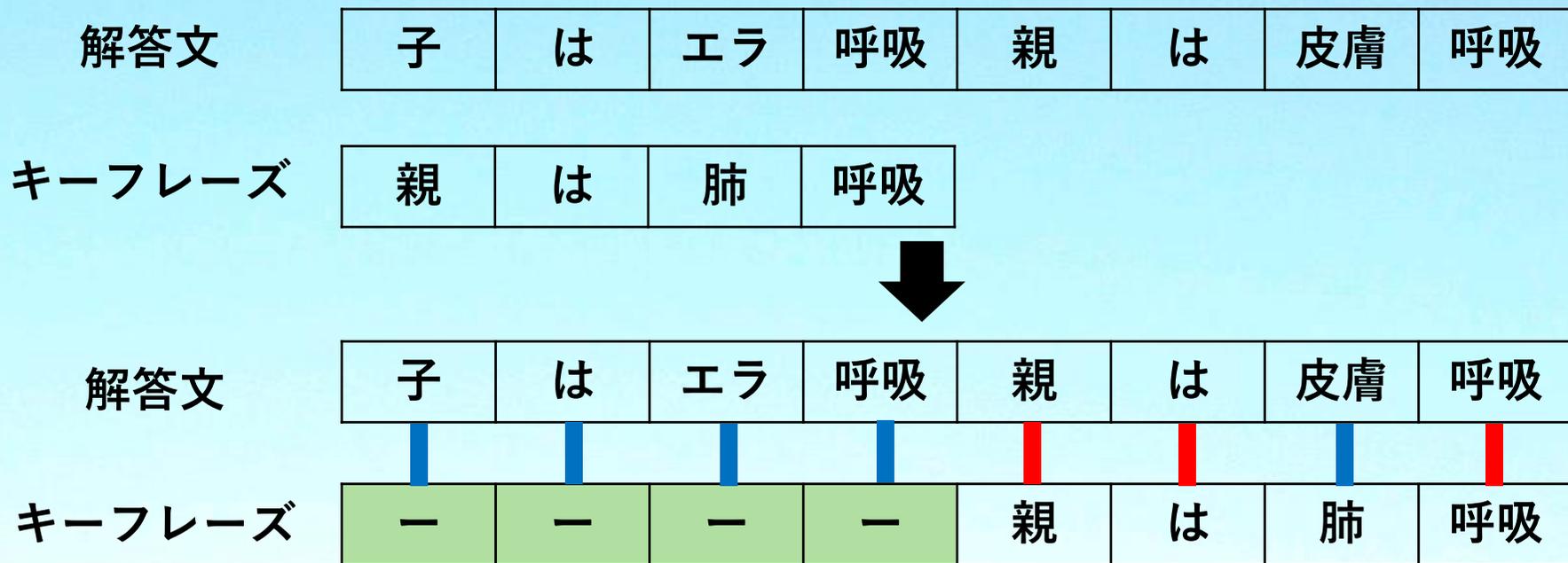
2. シーケンスアライメントを用いた比較

単語の並びを考慮して比較を行い，一致度を測る。

解答文とキーフレーズに シーケンスアライメントを施した結果

答案の解答文とキーフレーズの比較:

シーケンスアライメントの代表的手法 **Needleman-Wunsch Algorithm** を利用



単語に分かち書きされた配列に応用

→ キーフレーズの中から最も一致度の高いものを選択

他, 表記揺れの補正・類義語の置換も行う.

本自動採点システムの正誤判定

キーワードが

完全に含まれている ⇒ 正解と判定 (自動採点済みとして終了)

含まれていない ⇒ 自動採点不可として提示 ⇒ 手動採点

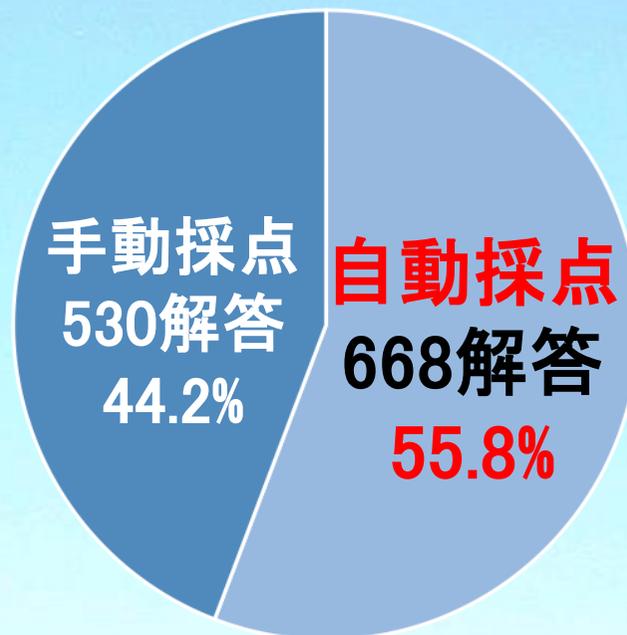
具体例

解答A	<u>子はエラ呼吸</u>	<u>親は肺呼吸</u>	⇒ 正解と判定
キーワード	子は エラ 呼吸	親は 肺 呼吸	(自動採点可)
解答B	<u>子は皮膚呼吸</u>	<u>親は肺呼吸</u>	⇒ 自動採点不可
キーワード	子は コ 呼吸	親は 肺 呼吸	⇒ 手動採点

自動採点の結果

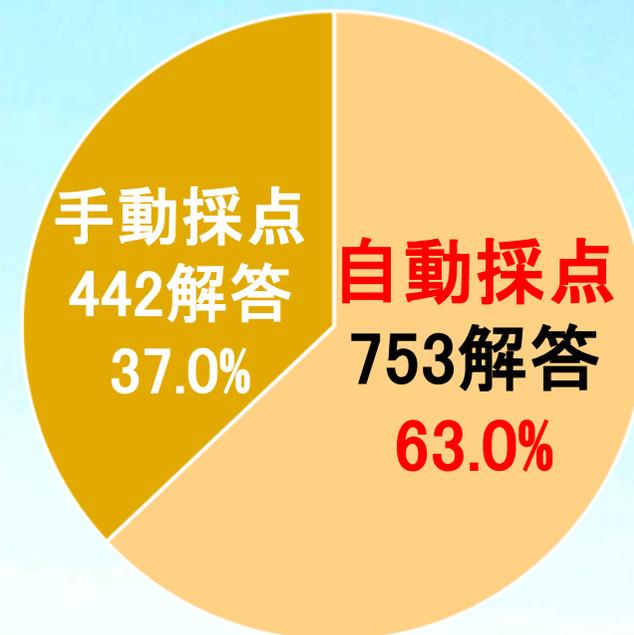
国語

採点精度 99.4%



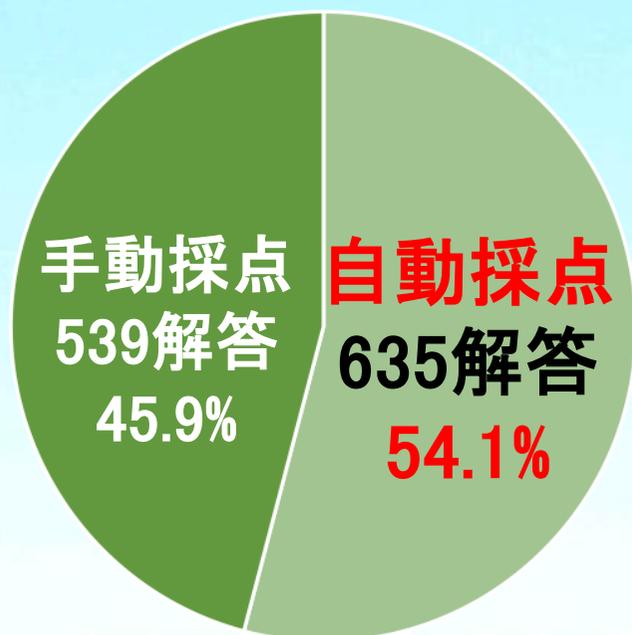
理科

採点精度 100%



社会

採点精度 98.6%



いずれも自動採点は、
20秒以下の高速採点

(Intel 第 6 世代 Core i5-6400, RAM:8GB, 2017年発売のPC)

採点設定の改善

採点結果付きテストデータがある場合(例:手動採点の採点結果)

1. 解答文をクラスタリング
2. 各クラスタをみて, 人手で採点条件を設定・改善



例:

子は肺呼吸で親も肺呼吸
親も子も皮膚呼吸をする

子は肺, 親も肺呼吸

親も子も肺とえらで呼吸

クラスタ 1

子は肺呼吸で親も肺呼吸
子は肺, 親も肺呼吸

→ 親も肺呼吸

クラスタ 2

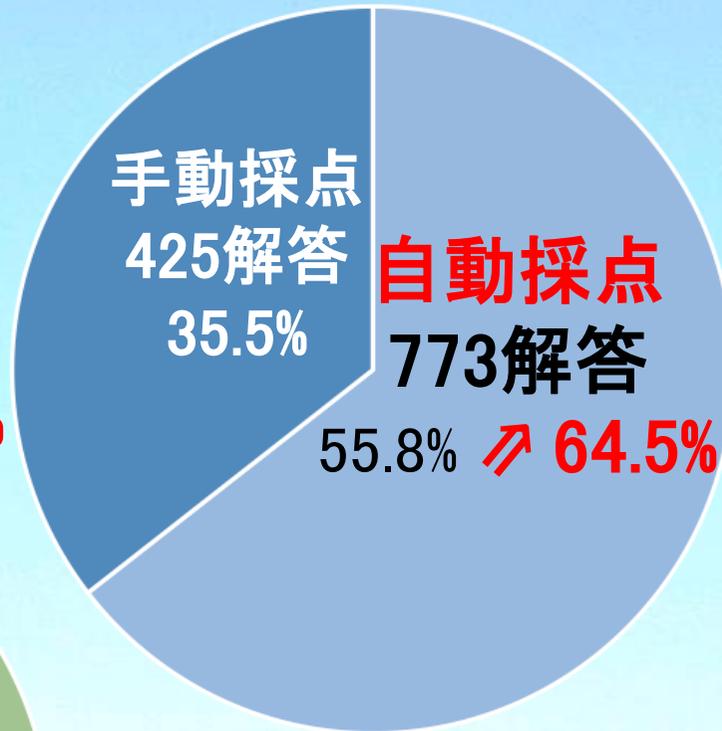
親も子も皮膚呼吸をする
親も子も肺とえらで呼吸

→ 親も子も

採点設定改善後の自動採点の結果

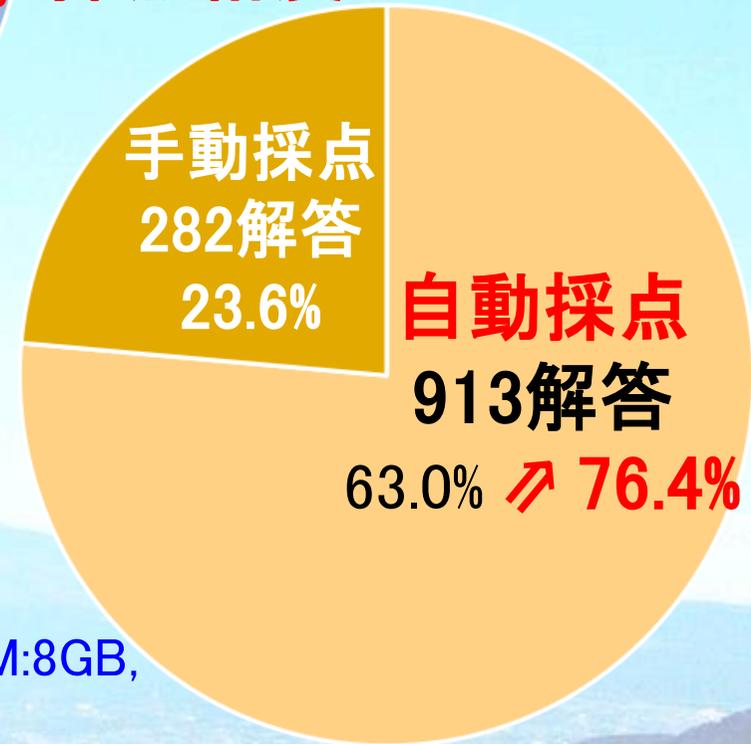
国語

採点精度 99.4% ↗ 99.9%



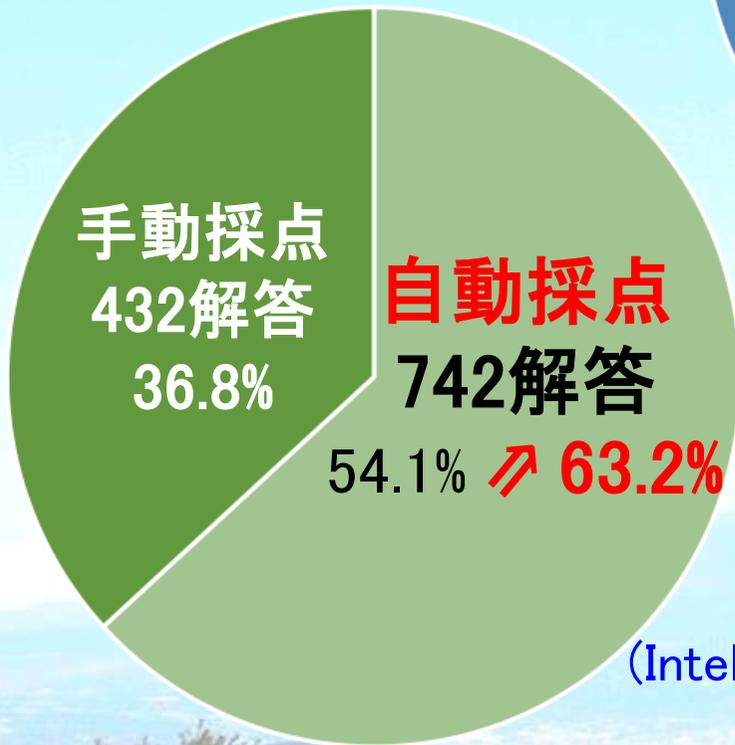
理科

採点精度 100% ⇒ 100%



社会

採点精度 98.6% ↗ 99.6%



いずれも自動採点は、
20秒以下の高速採点

(Intel 第6世代 Core i5-6400, RAM:8GB,
2017年発売のPC)

実用化に向けた課題

1. 手書き文字から、テキストデータへの変換
⇒既存のOCR技術で、対応可能.
2. 部分点への対応. ⇒現在, 研究中.
3. 大学入試以上の難易度の問題への対応.
⇒現在, 研究中.
⇒高校入試(含)までの問題では,
ほぼ実現可能. (国語, 社会, 理科)

企業への期待

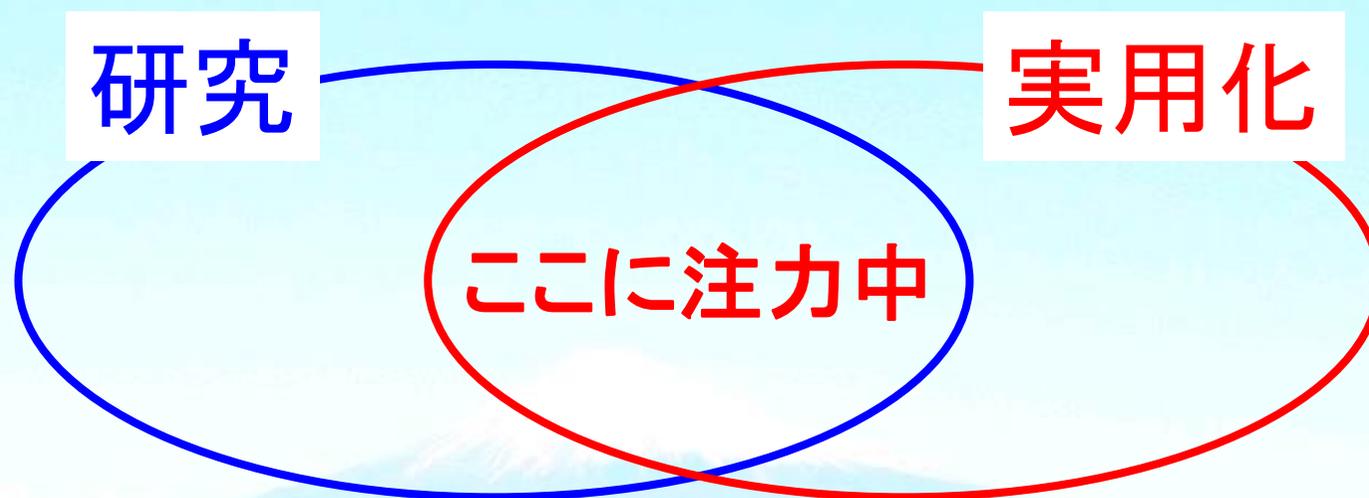
テキストデータの準備, 整備, 逐次更新

持続的な実用化にあたって, とても重要.

これさえできれば, 現場のニーズに応じた実用化
が可能な段階.

採点支援に関する産学連携の経歴

- 2018年-現在
模擬試験の大手運営会社と共同研究継続中
- 2019年-現在
大学入試センターと共同研究継続中



2022年以内に、新たに2件の採点支援の特許出願予定。

本技術に関する知的財産権

- 発明の名称 : 記述式試験採点プログラム
及び記述式試験採点方法
- 出願番号 : 特願2019-040212
- 出願人 : 千葉大学
- 発明者 : 須鎗弘樹, 森康久仁,
竹谷謙吾, 高井浩平

お問い合わせ先

千葉大学 学術研究・イノベーション推進機構
プロジェクト推進部門

TEL 043-290-3048

FAX 043-290-3519

e-mail ccrcu@faculty.chiba-u.jp