

平成21年 4月 3日現在

研究種目：基盤研究(B)

研究期間：平成19年度 ～ 平成20年度

課題番号：19300292

研究課題名(和文) 日本語小論文の自動評価における総合的研究

研究課題名(英文) Synthetic research for automated Japanese essay scoring

研究代表者

石岡 恒憲 (ISHIOKA TSUNENORI)

独立行政法人 大学入試センター・准教授

研究者番号：80311166

研究成果の概要：

本研究者が試作した日本語小論文の自動採点システム Jess は、修辞、論理構成、内容の3つの観点から小論文を評価する、我が国で最初でかつ現時点で唯一の自動採点システムである。本研究では、さらにこれを推し進め、実用化に向けた改良を行った。具体的には以下の研究を実施した。(1)採点の論拠となる各メトリクス指標の影響評価、(2)接続表現に基づかない論理構成把握の方法の検討、(3)作文支援機能、(4)エージェントシステムについての研究、(5)新聞投稿文の調査研究。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	2,400,000	720,000	3,120,000
2008年度	1,600,000	480,000	2,080,000
年度			
年度			
年度			
総計	4,000,000	1,200,000	5,200,000

研究分野：総合領域

科研費の分科・細目：教育工学

キーワード：小論文・エッセイ、自動採点、作文評価、自然言語処理、統計的アプローチ

1. 研究開始当初の背景

小論文試験が Writing Ability を測定しているものだとすれば、その誤差要因には採点者による評定の違いを始め様々な要因が存在する。そのような誤差要因を排除するため、あるいは採点の手間を軽減する目的で、あるいは公平性の立場から、コンピュータによる自動採点が望まれている。また、ユーザに対して即座に診断情報を返すことができるので、作文指導に関する教育的効果が期待できる。

2. 研究の目的

本研究者が試作した日本語小論文の自動採点システム Jess は、修辞、論理構成、内容の3つの観点から小論文を評価する、我が国で最初でかつ現時点で唯一の自動採点システムである。本システムについては、2005年2月15日付けの朝日新聞夕刊1面トップに記事が大きく掲載されたのをきっかけとし、週刊アスキー、アサヒパソコン、ニッポン放送、韓国 KBS テレビなど多くのマスコミ各誌で紹介され、教育測定の専門家やテスト関係者だけでなく、大学、高校、予備校な

どの現場でも注目されている。

本研究では、さらにこれを推し進め、実用化に向けた改良を行う。

3. 研究の方法

(1) 採点の論拠となる各メトリクス指標の影響評価

初年度は毎日新聞の CD-ROM に納められている社説やコラム(「余録」)を模範エッセイとした場合に、文章がいかによく書かれているかの評価尺度である「修辞」と、文章の論理的構成尺度である「論理構成」の各評価尺度の特性値の分布がどのようなものであるのか、また代替特性となる評価尺度はどのようなもので、それらの定量的な関係はどうかを調査する。言い替えれば、模範エッセイの内容分析を行なう。これにより、よい小論文とは、どのような修辞上の特性と論理構成を持つのかを定量的に理解する。

第2年度は、初年度の研究結果である試作システムの定量的な評価を踏まえ、システムを改良する。改良したシステムを、現在運用している試作システムと置き換え、広く一般に公開する。一般公開用 Web サーバもこれにあわせてリプレースする。

(2) 接続表現に基づかない論理構成把握の方法の検討

我々のシステムは論理構成の把握に難があるという評価が一般にされており、本モジュールの改良の余地が大きいと予想される。前年度の研究成果を踏まえて、文末モダリティ処理の適用と星座グラフによる可視化を試みる。

(3) 作文支援機能

英文においては Criterion というシステムが既にその機能を実装しており、これを参考にして日本文においても実現できるよう現在でも調査を進める。その後、試作をし、これを改良システムに組み込む。必要に応じて、更なるアプローチについて検討を行う。これらの研究成果は、広く内外に紹介していく。

(4) エージェントシステムについての研究

エージェントは利用者や他のエージェントと知的に相互作用する自律的ソフトウェアであり、次世代の分散システムを構築する中核的技術である。本採点システムの利用において必須の技術であるため、そのための基盤研究を進める。

(5) 新聞投稿文の調査研究

エッセイの例として新聞投稿文を取り上げ、それらを特徴づける因子の抽出を試みる。他、関連研究を行なう。

4. 研究成果

(1) 採点の論拠となる各メトリクス指標の影響評価

よい小論文とは、どのような修辞上の特性

と論理構成を持つのかを定量的に理解するための計量分析をおこない、人工知能学会誌に解説論文を掲載するとともに、単行本「eラーニング」(培風館)に掲載した。また、短答式記述テストに対する自動評価システムへの応用を検討し、成果について入試研究ジャーナルほかに掲載した。

(2) 接続表現に基づかない論理構成把握の方法の検討

日本語小論文の論理構造をより正確に評価するために、従来、欧米語のみならず日本語においてもしばしば用いられてきた接続表現に加え、新たに前の段落を受ける指示代名詞や文末モダリティの利用を着想した。これは日本語では、接続表現が意識的に避けられるためへの配慮である。

この手法を毎日新聞の社説およびコラム(「余録」)に適用し、議論の構造/展開の仕方を統計学の分野で用いられている星座グラフで表現することを試みた。これにより、このグラフが全体の分布としての把握と、着目したデータの時系列な論理構造の変化の把握の両方に役立つことがわかった。成果については人工知能学会論文誌ほかに掲載した。

(3) エージェントシステムについての研究

エージェントは利用者や他のエージェントと知的に相互作用する自律的ソフトウェアである。これについての基礎研究を進め、20編以上もの学術論文として公開した。

(4) 新聞投稿文の調査研究

エッセイの例として新聞投稿文を取り上げ、新聞投稿文を特徴づける因子の抽出については言語処理学会 第14回年次大会で、新聞投稿文の接続関係の分析については第15回年次大会で発表した。

(5) システム公開

小論文の自動採点システムを以下のアドレスにて公開/運用している。

<http://coca.rd.dnc.ac.jp/jess/>

現在でも1日平均で500件以上のページビューがあり、広く利用されている。

(6) 特記事項

以下5節に示すように、非常に数多くの研究成果を挙げた。これら一連の成果が認められ、研究代表者である石岡が Marquis Who's Who in the World, 2008年版と2009年版に掲載された。2007年に韓国 KBS テレビの取材を受けた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計9件)

① Tsunenori Ishioka, Masayuki Kameda.

Automated Japanese Essay Scoring System based on Professional Writings, Joint Meeting of 4th World Conference of the IASC and 6th Conference of the Asian Regional Section of the IASC on Computational Statistics & Data Analysis, 679-688. (2008), 査読有

②Hideki Nagatsuka, Toshinari Kamakura, Tsunenori Ishioka. A Simple and Accurate Estimation of Ordered Means of Exponential Distributions Using Bayesian Approach, Joint Meeting of 4th World Conference of the IASC and 6th Conference of the Asian Regional Section of the IASC on Computational Statistics & Data Analysis, 1212-1220. (2008), 査読有

③石岡 恒憲, 日本語小論文の論理構成を把握する方法とその図式表現, 人工知能学会論文誌, 23 (5), 303-309. (2008), 査読有

④石岡 恒憲, 短答式記述テストにおける自動採点 —その採点ロジックと課題について—, 大学入試研究ジャーナル, 19, 57-63. (2008), 査読有

⑤石岡 恒憲, 小論文およびエッセイの自動評価採点における研究動向, 人工知能学会誌, 23 (1), 17-24. (2008), 査読有

⑥大石 哲也, 倉元 俊介, 峯 恒憲, 長谷川 隆三, 藤田 博, 越村 三幸: 関連単語抽出アルゴリズムを用いた Web 検索クエリの生成, 電子情報通信学会論文誌: データ工学特集号, Vol. J92-D, No. 3, 掲載予定 (2009), 査読有

⑦ Tetsuya Oishi, Shunsuke Kuramoto, Tsunenori Mine, Ryuzo Hasegawa, Hiroshi Fujita, and Miyuki Koshimura: A Method for Query Expansion Using the Related Word Extraction Algorithm, the 2008 International Workshop on Web Information Retrieval Support Systems, pp.41-44,(2008), 査読有

⑧Tetsuya Oishi, Yoshiaki Kambara, Tsunenori Mine, Ryuzo Hasegawa, Hiroshi Fujita, Miyuki Koshimura: Personalized Search using ODP-based User Profiles created from User Bookmark, The 10th Pacific Rim International Conference on Artificial Intelligence (PRICAI 2008),pp.839--848, (2008), 査読有

⑨Tetsuya Oishi, Shunsuke Kuramoto,

Hiroto Nagata, Tsunenori Mine, Ryuzo Hasegawa, Hiroshi Fujita, Miyuki Koshimura: User-Schedule-based Web Page Recommendation, the 2007 IEEE/WIC/ACM International Conference on Web Intelligence, pp.776-779, (2007), 査読有

[学会発表] (計 30 件)

①石岡 恒憲. テキストマイニング新展開 (特別講演), 新領域融合研究センタープロジェクトによる研究会「大規模データ・リンケージ, データマイニングと統計手法」. 統計数理研究所. 2008 年 12 月 11 日・12 日. <http://www.ism.ac.jp/~nobuo/sympo-0812.html> (2008), (特別講演のため) 査読無

②石岡 恒憲, 短答式記述テストにおける自動採点 —その採点ロジックと課題について—, 平成 20 年度全国大学入学選抜連絡協議会 予稿集, 51-56. (2008) 査読無

③石岡 恒憲. 日本語小論文の論理構成を把握する方法とその図式表現, 2007 年度統計関連学会連合大会 講演報告集, 132. (2007) 査読無

④石岡 恒憲. 日本語小論文の自動評価: 論理構成の把握とその図式表現, 合同エージェントワークショップ&シンポジウム 2007 (JAWS2007), マルチエージェントと協調計算研究会 (日本ソフトウェア科学会)・人工知能と知識処理研究会 (電子情報通信学会)・知能と複雑系研究会 (情報処理学会)・知識ベースシステム研究会 (人工知能学会)主催, 資料集 161-168. (2007), 査読有

⑤大石哲也, 峯恒憲, 長谷川隆三, 藤田博, 越村三幸: 関連単語抽出アルゴリズムを用いたクエリ拡張, 合同エージェントワークショップ&シンポジウム 2008 (JAWS2008), CD-ROM, (2008), 査読有

⑥山室元, 大石哲也, 峯恒憲, 長谷川隆三, 藤田博, 越村三幸: Web推薦システムの評価作業支援システムの提案, 合同エージェントワークショップ&シンポジウム 2008 (JAWS2008), CD-ROM, (2008), 査読有

⑦蔣偉, 永田廣人, 大石哲也, 長谷川隆三, 藤田博, 峯恒憲, 越村三幸: ODPを利用したユーザプロフィール作成と個別化検索システムの提案, 合同エージェントワークショップ&シンポジウム 2008 (JAWS2008), CD-ROM, (2008), 査読有

⑧神原義明, 大石哲也, 長谷川隆三, 藤田博, 峯恒憲, 越村三幸: ODPとブックマークを利用した個人化検索システム, 合同エージェントワークショップ&シンポジウム 2008(JAWS2008), CD-ROM, (2008), 査読有

⑨古後陽大, 峯恒憲, 雨宮聡史, 雨宮真人: ACP2P法におけるユーザフィードバックの利用とクエリ送信先決定法の提案, 合同エージェントワークショップ&シンポジウム 2008(JAWS2008), CD-ROM, (2008), 査読有

⑩堀憲太郎, 大石哲也, 峯恒憲, 長谷川隆三, 藤田博, 越村三幸: Wikipediaを利用したクエリ拡張によるWeb検索支援システム, 合同エージェントワークショップ&シンポジウム 2008(JAWS2008), CD-ROM, (2008), 査読有

⑪中尾好佑, 峯恒憲: ブラウジング支援に向けた情報推薦のための関連度計算方法の提案, 合同エージェントワークショップ&シンポジウム 2008(JAWS2008), CD-ROM, (2008), 査読有

⑫大石哲也, 倉元俊介, 堀憲太郎, 長谷川隆三, 藤田博, 峯恒憲, 越村三幸: 関連単語抽出アルゴリズムを用いたWeb検索クエリの生成, 情報処理学会研究報告, Vol. 2008-DBS-145・2008-FI-91, pp. 33-40, (2008), 査読無

⑬堀憲太郎, 大石哲也, 長谷川隆三, 藤田博, 峯恒憲, 越村三幸: Wikipediaへの関連単語抽出アルゴリズムの適用とその評価, 情報処理学会研究報告, Vol. 2008-DBS-145・2008-FI-91, pp. 81-88, (2008), 査読無

⑭神原義明, 大石哲也, 長谷川隆三, 藤田博, 峯恒憲, 越村三幸: ODPを利用したユーザプロフィールを用いた個人化検索システム, 情報処理学会研究報告, Vol. 2008-DBS-145・2008-FI-91, pp. 17-24, (2008), 査読無

⑮古後陽大, 峯恒憲, 雨宮聡史, 雨宮真人: ユーザフィードバックを利用したACP2P法とその評価, 電子情報通信学会技術研究報告, Vol. DE2008-26, PRMU2008-44, pp. 145-150, (2008), 査読無

⑯中井裕介, 古後陽大, 峯恒憲, 雨宮真人: ユーザのフィードバックを用いたACP2P法の改善, 情報処理学会九州支部火の国情報シンポジウム 2008(CD-ROM論文集), (2008), 査読無

⑰和田潤也, 大石哲也, 峯恒憲, 長谷川隆三,

藤田博, 越村三幸, 蔣偉: ユーザのスケジュールを用いたWeb ページ推薦, 情報処理学会第 70 回全国大会講演予稿集, CD-ROM, (2008), 査読無

⑱中尾好佑, 雨宮聡史, 峯恒憲, 雨宮真人: ブラウジング支援のための情報推薦手法の提案, 情報処理学会九州支部火の国情報シンポジウム 2008(CD-ROM論文集), (2008), 査読無

⑲種村新, 峯恒憲, 雨宮真人: Web際との構造とコンテンツを考慮した情報推薦手法, 情報処理学会九州支部火の国情報シンポジウム 2008(CD-ROM論文集), (2008), 査読無

⑳光行慶城, 高橋健一, 峯恒憲, 雨宮真人: 階層型コミュニティシステムにおけるセキュリティ機能の実装と評価, 情報処理学会九州支部火の国情報シンポジウム 2008(CD-ROM論文集), (2008), 査読無

21. 大石哲也, 峯恒憲, 長谷川隆三, 藤田博, 越村三幸: Web検索におけるスケジュール情報の利用, 第80回知識ベースシステム研究会 (SIG-KBS), Vol. A703, pp. 43-48, (2008), 査読無

22. 倉元俊介, 長谷川隆三, 藤田博, 越村三幸, 峯恒憲: 関連距離アルゴリズムを用いた関連語提案システム, 合同エージェントワークショップ&シンポジウム 2007(JAWS2007), (CD-ROM), (2007), 査読無

23. 大石哲也, 峯恒憲, 長谷川隆三, 藤田博, 越村三幸, 倉元俊介, 永田廣人: ユーザのスケジュールを考慮したWeb検索手法, 合同エージェントワークショップ&シンポジウム 2007(JAWS2007), (CD-ROM), (2007), 査読有

24. 種村新, 峯恒憲, 雨宮真人: Webサイトの階層構造を考慮した情報推薦手法, 合同エージェントワークショップ&シンポジウム 2007(JAWS2007), (CD-ROM), (2007) 査読有

25. 永田廣人, 長谷川隆三, 藤田博, 越村三幸, 峯恒憲: ODPを利用したユーザプロフィール作成と個別化検索システムの提案, 合同エージェントワークショップ&シンポジウム 2007(JAWS2007), (CD-ROM), (2007) 査読有

26. 種村新, 峯恒憲, 雨宮真人: Web際との構造とコンテンツを考慮した情報推薦手法, 情報処理学会九州支部火の国情報シンポジウム 2008(CD-ROM論文集), (2008), 査

読無

27. 中尾好佑, 雨宮聡史, 峯恒憲, 雨宮真人:
ブラウジング支援のための情報推薦手法の
提案, 情報処理学会九州支部火の国情報シ
ンポジウム 2008 (CD-ROM論文集), (2008),
査読無

28. 中井 裕介, 古後 陽大, 峯 恒憲, 雨宮 真
人: ユーザのフィードバックを用いたACP2P
法の改善, 情報処理学会九州支部火の国情
報シンポジウム 2008 (CD-ROM 論 文 集),
(2008), 査読無

29. 生田和重: 新聞投稿文の接続関係の分析,
言語処理学会, 第 15 回年次大会 (NLP2009),
(CD-ROM 論文集), (2009. 3), 査読無

30. 生田和重: 新聞投稿文を特徴づける因子
の抽出, 言語処理学会, 第 14 回年次大会
(NLP2008), (CD-ROM論文集), (2008), 査読
無

[図書] (計 1 件)

①石岡 恒憲(ほか(植野・永岡編): eテストイ
ング, 培風館, p.269, 2009.

[産業財産権]

○出願状況 (計 件)

○取得状況 (計 件)

[その他]

Web ページ

<http://coca.rd.dnc.ac.jp/jess/>

6. 研究組織

(1) 研究代表者

石岡 恒憲 (ISHIOKA TSUNENORI)
独立行政法人 大学入試センター・准教授
研究者番号: 80311166

(2) 研究分担者

峯 恒憲 (MINE TSUNENORI)
九州大学大学院・システム情報科学研究院・
准教授
研究者番号: 30243851

生田 和重 (IKUTA KAZUSHIGE)
徳島文理大学・文学部・教授
研究者番号: 60289311

(3) 連携研究者