様式 C-19

科学研究費補助金研究成果報告書

平成２２年　５月１４日現在

研究種目：基盤研究（Ｂ）
研究期間：2007 ～ 2009
課題番号：19300042
研究課題名（和文）　テキストの自動評価システムの開発

研究課題名（英文）　Developing a system that automatically evaluates texts

研究代表者
奥村　学（OKUMURA　MANABU）
東京工業大学・精密工学研究所・教授
研究者番号：60214079

研究成果の概要（和文）：

1）文章としての質の評価では，文章中の文間のつながりの良し悪しを自動評価し，テキスト自体の文章としての良し悪しを計る統計的な手法を開発した.結束性の情報として利用可能なもののうち，接続詞，語彙的結束性の情報を，従来用いられているentity grid 手法に追加，拡張することで，従来よりも高精度にテキストの一貫性を判定できることを明らかにした．2）内容の情報量の評価では，これまでの研究成果により開発している，テキスト自動要約システムの出力したテキストの内容を自動的に評価する手法において，内容的な類似度尺度に，語彙的な言い換えの情報を導入することで，より高精度に類似性判定を行える枠組みを開発した.

研究成果の概要（英文）：

As the evaluation framework for text quality, we developed a statistical method that evaluates local coherence of a text. We made improvements to the entity grid local coherence model for Japanese text, and investigated the effectiveness of taking into account cohesive devices, such as conjunction, explicit reference relation, lexical cohesion, and refining syntactic roles for a topic marker in Japanese. As the evaluation framework for text content, we explored the use of paraphrases for the refinement of traditional automatic methods for summary evaluation.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
2007 年度	7,400,000	2,220,000	9,620,000
2008 年度	3,600,000	1,080,000	4,680,000
2009 年度	3,200,000	960,000	4,160,000
年度			
年度			
総　計	14,200,000	4,260,000	18,460,000

研究分野：総合領域
科研費の分科・細目：情報学・知能情報学
キーワード：情報の信頼性、テキスト自動要約、自動採点、テキストの質、内容の情報量

１．研究開始当初の背景

近年我々人間の周囲には，さまざまなメディアを通した情報が満ち溢れ，情報洪水という言葉が使われるようになってからかなりの歳月を経ている．WWW上も例外ではなく，新聞，雑誌の記事が電子化された形で提供されるようになり，スポーツ等のイベントは，テレビなど他のメディア同様インターネット上でも中継されるようになり，また，いわゆるマスメディアではない独自なメディアとして，掲示板(BBS)，chat，blog(Weblog)などのような「口コミ」としての一般大衆による情報発信も盛んになりつつある．

しかし，言うまでもなく，上述したような「口コミ」の情報発信を中心とする，個人が発信している大量の情報は玉石混交であり，それらの大量の情報を格付け，有用な情報のみをその中から選択する必要が出てきている．Web spam と言われる，ある種の大量の「ゴミ」情報が Web 上にはますます蔓延しつつあり，社会問題の 1 つともなり始めているが，これへの対処も緊急的な課題と言える．このように，インターネット上の大量の情報の中から，有用な情報のみを選別したり，ゴミと考えられるような大量の情報をフィルタリングしたりするためには，それらの情報を何らかの基準に基づき評価し，その評価に基づき，情報の格付けを行う必要がある．本研究課題では，情報の中で特にテキスト情報に着目し，システムにより自動評価を行う手法の開発を行う．

我々はこれまで様々な種類のテキスト自動要約システムの研究開発を進めてきている．また，我々，研究代表者と研究分担者のグループはこれまで，国立情報学研究所が主催するNTCIR Workshopのタスクとしてテキスト自動要約タスクを3回にわたり主催している．この過程で，テキスト自動要約システムの出力したテキストを自動的に評価する枠組みの構築に成功し，また，実際にその枠組みを用いて，日本語を対象とする多くのテキスト自動要約システムの出力するテキストを自動評価している．そして，人間が人手で評価した結果と非常に相関の高い評価を実現できている．この枠組みでは，テキストを1) 文章としての質，2) 内容の情報量，の大きく2つの評価基準で評価する．どちらの評価基準においても，模範的なテキスト(自動要約システムの出力するテキストを評価する場合，人間が作成した模範的な要約)や，あらかじめ人間がテキストを評価した結果を，ある程度の量用意しておき，自動評価システムは，評価するテキストと，模範的なテキストあるいは評価済みのテキストの内容的な類似度を計算し，その類似度に応じて，テキストの評価結果を計算する．

我々は，このテキストの自動評価システムをテキスト自動要約システムの改良のための枠組みとしてこれまで利用してきたが，システムの出力したテキストにとどまらず，広く人間が書いたテキスト一般について，システムが自動的に評価する枠組みとして利用できることに思い至った．そこで**本研究課題では，我々がテキスト自動要約システムの出力するテキストを自動評価する枠組みとして当初開発したものを援用，拡張することで，広く一**

般のテキストを自動評価する枠組みを開発することを目的とする．

一方，アメリカで開発が進められているE-ratorなどのように，近年様々な領域の教育分野においては，学習者の作文，小論文，レポート等を自動的に評点付けする手法，システムの開発が活発に行われ始めている．これらの手法，システムも，人間の書いたテキストを自動評価するものと言うことができ，本研究課題で開発する，言語処理技術を用いたテキストの自動評価手法が，これらの分野における自動評点付けにおいても利用可能であると考える．

２．研究の目的

本研究課題では，上で述べたように，1）文章としての質，2）内容の情報量という 2 つの評価基準に基づき，Web 上のテキストを含む，一般のテキストを自動的に評価し，格付けする手法を開発することを目的とする．後述するように，これまでの Web 上の情報評価に関する試みは，Web 上のリンク情報に頼る傾向が強く，書かれているテキスト自体を評価しようという視点に欠けていると言える．そこで我々は，情報評価のための強力なツールとなりうる，テキストを評価しうる言語処理技術の開発を進める．

具体的には，2 つの評価基準それぞれについて以下のような評価手法を開発する．

1）文章としての質の評価

Web 上のテキストは多様であり，また，書き手の多様性により，書かれるテキストの文体，用いられる用語等も様々であるが，概して，くずれた文章で書かれたテキストほど質が低いことが多いと考えられる．そこで，書かれたテキストの文章の質を評価することで，書かれているテキストの質を近似的に評価することを試みる．

書き手の書いたテキストの語彙，文法性，文章構造を解析することで，文章としての質を定量的に計る手法を開発する．研究成果欄で述べるように，我々はこれまでのテキスト自動要約システムに関する研究開発において，システムの出力したテキストの文章としての質を自動評価し，その評価に基づいて，システムの出力したテキストを自動的に推敲する枠組みを試作している．この枠組みを拡張，改良することで，テキストの文章としての質を評価するシステムを開発する．

2）内容の情報量の評価

テキストの情報内容を評価する上でもっとも重要と考えられるのは，その情報の内容の詳細さ，カバレッジ(情報のカバーする範囲の広さ)であろう．そこで，模範的なテキストが本来持っているべき内容の情報量を元に，それと比べてどの程度の情報量を持っているかを計ることで，テキストの内容の評価を行う手法を開発する．

このようなテキストの情報量の自動評価手法を用いる場合，模範的なテキストがあらかじめ評価の際必要になる．テキスト自動要約システムや機械翻訳システムなどの出力するテキストを評価する場合，模範となるテキストをあらかじめ人間が作成することで対処することができる．教育分野における作文，小論文，レポート等の自動採点の場合も，模範解答となるテキストを人間があらかじめ用意することで対処できる．

一方，Web上のテキストを自動評価する場合，すべてのテキストに対して逐一，模範となるテキストを人手で用意しておくことは現実的でない．そこで，テキストのトピックを自動推定し，トピックごとに，そのトピックのテキストが本来持っているべき情報量を推測し，内容的に理想的な仮想テキスト(正解テキスト)を自動的に作成する手法を開発し，この仮想の正解テキストとの比較により，Web上のテキストの情報量を自動評価する．

３．研究の方法

本研究課題では，

1) 文章としての質の評価，

2) 内容の情報量の評価，

それぞれを行う手法，システムの開発を並列して行う．以後，1)，2)の開発それぞれについて計画を述べる．

1) 文章としての質の評価手法の開発

書き手の書いたテキストの語彙，文法性，文章構造を解析することで，テキスト自体の文章としての質を定量的に計る手法を開発する．これまでの手法は語彙，文法性といった表層的な特徴のみを用いていたことから，本研究課題では，照応，省略解析，テキスト構造解析等の文脈解析技術を用いて，文章としてのつながりの良し悪しを計る尺度を導入する．

本研究課題では，照応，省略解析，テキスト構造解析技術を利用し，要約結果の推敲の場合と同様に，文章中の文間のつながりの良し悪しを自動評価し，テキスト自体の文章としての良し悪しを計る手法を開発する．

2) 内容の情報量の評価手法の開発

これまでの研究成果により，テキスト自動要約システムの出力したテキストの内容を自動的に評価する枠組みの構築に成功している．この枠組みでは，模範的なテキストあるいは，あらかじめ人間が評価したテキストと，内容的にどの程度類似しているかを元に，テキストの評価結果を計算する．人間が人手で評価した結果と非常に相関の高い評価を実現できている．

1年目(平成19年度)では，まずテキスト自動要約システムの出力するテキストを対象に開発した我々の手法が一般のテキストに適用可能であるかどうかの検討を行う．また，一般のテキストに適用可能にするための改良を行う．

一般のテキストの場合(小論文等，学習者があるテーマについて書いたテキストにおいても，Web上のテキストのように，書き手が自由に書いたテキストにおいても)，システムの出力したテキストを評価する場合と比較して，テキストの内容の自由度が大きいことが考えられる．このため，模範テキストあるいは，すでに人間が評価したテキストをあらかじめ一定量保持し，それらのテキストと，評価するテキストの類似度の計算を行い，類似したテキストの評価結果を元に，評価するテキストの評価結果を計算する，我々がこれまでに開発した手法を用いる場合，どの程度の量のテキストをあらかじめ保持している必要があるか，どのような類似度尺度を用いて類似するテキストを求めるのが適切であるか等，いくつか検討する必要がある点が存在する．

そこで，このような検討課題を考慮し，手法を拡張，改良することで，一般のテキストに適用可能な，内容の情報量を評価する手法を開発する．

４．研究成果

1) 文章としての質の評価では，文章中の文間のつながりの良し悪しを自動評価し，テキスト自体の文章としての良し悪しを計る統計的な手法を開発した．結束性の情報として利用可能なもののうち，接続詞，語彙的結束性の情報を，従来用いられているentity grid手法に追加，拡張することで，従来よりも高精度にテキストの一貫性を判定できることを明らかにした．また，この手法の拡張として，テキストの断片の文章としての良し悪しを計る統計的な手法を開発した．2) 内容の情報量の評価では，これまでの研究成果により，テキスト自動要約システムの出力したテキストの内容を自動的に評価する枠組みの構築に成功している．この枠組みでは，模範的なテキストあるいは，あらかじめ人間が評価したテ

キストと，内容的にどの程度類似しているかを元に，テキストの評価結果を計算する．人間が人手で評価した結果と非常に相関の高い評価を実現できている．この手法において内容的な類似度尺度に，語彙的な言い換えの情報を導入することで，より高精度に類似性判定を行える枠組みを開発した．

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　4 件）
①平原一帆,難波英嗣,竹澤寿幸,奥村学，言い換えを用いたテキストの自動評価，情報処理学会論文誌データベース，Vol.3, No.2,2010,査読有
②横野光，奥村学，テキスト結束性を考慮した entity grid に基づく局所的一貫性モデル，自然言語処理，Vol.17，No.1, pp161-182,2010,査読有
③Tsutomu Hirao，Manabu Okumura，Norihito Yasuda and Hideki Isozaki，Supervised automatic evaluation for summarization with voted regression model，Information Processing & Management,Vol. 43，1521-1535,2007,査読有
④平尾 努，奥村 学，福島孝博，難波英嗣，野畑周，磯崎 秀樹，抜粋による複数文書要約を評価するためのコーパスと評価指標，情報処理学会論文誌データベース,Vol.48, 60-68,2007,査読有

〔学会発表〕（計　7 件）
① Hikaru YOKONO，Manabu OKUMURA，Incorporating Cohesive Devices into Entity Grid Model in Evaluating Local Coherence of Japanese Text，CICLing2010, 2010.3.23，Iaşi，Romania
②平原一帆,難波英嗣,竹澤寿幸,奥村学，言い換えを用いたテキストの自動評価，言語処理学会第 16 回年次大会，2010.3.9,東京大学
③ Kazuho Hirahara，Hidetsugu Nanba，Toshiyuki Takezawa，Manabu Okumura，Automatic Evaluation of Texts by Using Paraphrases，the 4th Language & Technology Conference(LTC 2009)，2009.11.7，Poznań，Poland
④奥村学，テキストの自動評価システムの開発，電子情報通信学会思考と言語研究会，2009.6.18，機械振興会館
⑤平原一帆,難波英嗣,竹澤寿幸,奥村学，言い換えを用いたテキストの自動評価，情報処理学会自然言語処理研究会，2009.5.22,東京工業大学
⑥横野光，奥村学，テキスト結束性判定のための entity grid モデルの素性の検討，情報処理学会第 189 回自然言語処理研究会，2009 年 1 月 23 日，お茶の水女子大学
⑦平原一帆，難波英嗣，竹澤寿幸，奥村学，平尾努，原文からの抜粋度合を考慮した要約の自動評価法，言語処理学会第 14 回年次大会，2008.3.19,東京大学

６．研究組織
(1)研究代表者
奥村　学（OKUMURA　MANABU）
東京工業大学・精密工学研究所・教授
研究者番号：60214079

(2)研究分担者
高村　大也（TAKAMURA　HIROYA）
東京工業大学・精密工学研究所・助教
研究者番号：80361773

(3)連携研究者
平尾　努（HIRAO　TSUTOMU）
日本電信電話株式会社ＮＴＴコミュニケーション科学基礎研究所・知識処理研究グループ・研究員
研究者番号：40396148

難波　英嗣（NANBA　HIDETSUGU）
広島市立大学・情報科学研究科・講師
研究者番号：50345378