

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 23 日現在

機関番号：20103

研究種目：挑戦的萌芽研究

研究期間：2014～2016

課題番号：26560124

研究課題名(和文)作文行動を分析できるマトリックス型文章モデルの研究

研究課題名(英文)A Matrix Model For Writing Activity Measurement And Analysis

研究代表者

大場 みち子 (Oba, Michiko)

公立はこだて未来大学・システム情報科学部・教授

研究者番号：30588223

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：作文(文章産出)に関するさまざまな研究や指導・学習支援の基盤とするために、ワープロなどによる作文過程を測定する技術を開発した。マトリックス型文章編集モデルと呼ぶモデルに基づいて、ITを使う作文行動をデータとして記録する。モデルを評価するために、それに基づいて編集操作を記録するTopic Writerを使って演習を行い、データを分析して評価した。前後して編集対象とした段落間の距離に基づく分析指標Editing Operation Indicator (EOI)を提案した。受講生同士のピアレビュー後に文章を修正する編集について、EOIが高い編集の方が、文章の論理的構成の評点に有意な改善がみられた。

研究成果の概要(英文)：To support various studies, teaching and learning on writing (text production), we developed technology to measure writing activities on word processors and other IT environment. In that technology, writing activities are recorded as data based on a model called a Matrix Model for Writing. We have developed "Topic Writer" which records editing operation based on the model. We analyzed the data and evaluated the model. We propose Editing Operation Indicator (EOI) which is calculated based on the distance between paragraphs which are edited in succession. There was a significant improvement in the score for the logical composition of the texts which are edited with high EOI after peer review between the students.

研究分野：教育工学

キーワード：作文 文章産出 測定 行動分析 モデル 共起

1. 研究開始当初の背景

近年、学習分析(Learning Analytics)研究の分析対象が、成績といった大きな単位からページめくり操作といった細かい単位へ移りつつある。この変化には、細粒度化と行動指向という2つの側面がある。

細粒度化とは、測定・分析・応用の対象がきめ細くなることを指す(図1)。学習分析は、より小さな粒度で行動を測定して、より小さな粒度で効果を上げる方向に向かっている。例えば、作文アプリケーションの操作の特徴と、書かれた文章の誤字脱字や論理的構成といった細かい評価との関係を分析して、作文の指導・学習にきめ細かく役立てようとしている。

| | | 分析・応用の対象の粒度 | | |
|-------|---|--------------|---------------------|---------------------|
| | | 大 | 中 | 小 |
| 測定の粒度 | 大 | 科目の成績 | ある科目の成績で、別の科目の成績を推定 | 講義への参画度(engagement) |
| | 小 | ページめくり編集対象の枠 | 成績を予測 | 予習・復習・講義への参画度 |
| | | | | 考え方の解き方 |
| | | | | 考え方の解き方の指導 |

本研究のターゲット

図1 測定と応用の粒度

行動指向とは、従来のAIや検索の分野では「書かれたテキスト」を分析したのに対して、「そのテキストを書く過程(process)・行動(behavior)・行為(activity)」に注目して測定・分析して活用することを指す(図2)。

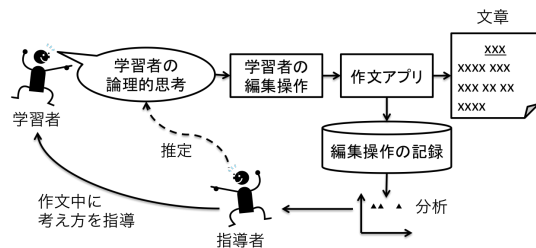


図2 行動指向

指導・学習環境にITが導入されることで、前記のような細粒度の測定データを、アンケート分析などと比べて低コスト・低バイアスで得られる可能性が高まる。また、個別に行われた研究で得られたデータを横断的に分析する、データサイエンスの可能性も広がる。

しかし、ただデータが大量にあるだけでは、期待した分析結果を得られるとは限らない。知的行動の細粒度のデータから何を見いだせるか知見が少なく、やってみなければ分からない。粒度の細かい学習分析研究の現状では、行動の傾向と結果の評価との関係に未知の部分が多い。作文であれば、行動つまり書き方と、書かれた文章の評価 -- 例えば論理的構成 -- との関係が、本研究の計画時点で

は不明であった。

また、学習環境がオンライン化することで、前記のようなデータを同時に大量に取得でき、またそれに基づいて大勢の学習者に同時に対応することができる。そのためには、測定・分析システムがコスト的にも性能的にもスケーラブルでなくてはならない。

2. 研究の目的

作文(文章産出、文章作成)に関するさまざまな研究・業務分析の基盤とするために、ワープロなどによる作文行動を測定する技術を開発する。具体的には、測定するデータを設計して評価する。

想定するデータの用途は次の通りである:

- ・ 作文の指導・学習支援
- ・ 指導・学習方法の評価・比較
- ・ ワードプロセッサやアイデアプロセッサなど作文ツールの評価・比較
- ・ 障がい者も含めて、言語横断的、国際的、PC/モバイルといった機器横断的、ライフステージ横断的な評価・比較
- ・ これらにおいて、企業内の業務やプライベートも含めて幅広い文章を対象とする

そこで、測定する技術の要件は次のとおりとなる:

- ・ 行動指向のモデル:
ITを使って文章を編集する行動のモデルである。これに対して、従来は、文章のモデルといえば、章立てや組版といった静的なものだった。
- ・ 指導・学習方法などから中立なモデル:
指導方法を比較するためには、特定の指導方法の考え方にのみ基づくのは好ましくない。
- ・ スケーラビリティ:
数人から数千、数万人規模の受講者を測定・分析できる低コストな技術であること。
- ・ アクセシビリティ、国際化、相互運用性の高い要素技術の採用
- ・ 機密保持・個人情報保護

3. 研究の方法

目的は測定データであるが、評価としてデータが役に立つ可能性を示す。このために、なんらかの分析手法を提案し、応用例を示す。実際に測定されるデータから見いだせる傾向などは、当初は不明である。小規模なデータに対して、傾向を可視化する手法を試し、記録するデータやアプリケーションの機能へフィードバックすることを繰り返し、スパイラルに進める。次の3.1から3.4を小規模に繰り返す(図3)。

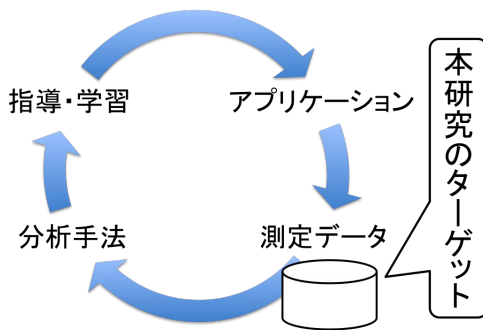


図 3 スパイラルなアプローチ

研究の計画、実験のデザイン、目論見、期待する成果などを、最初から公開し、研究会などで積極的に発表して議論する。

3.1 モデル設計

測定にあたっては、まず IT を使う作文行動(writing activity)をモデル化し、そのモデルに基づいて編集操作を記録して測定データとする。

3.2 リファレンス実装

モデルに従って編集操作を測定する作文アプリケーションを開発するか、あるいは既存のアプリケーションに測定機能をアドオンする。

アプリケーションのメモリ所要量などを測定することで、モデルが実装可能性を評価する。

3.3 分析手法の提案

測定データを分析する、なんらかの手法・指標の開発・提案する。書き手の意図や行動の意味を読み取れるような分析・可視化手法とする。

分析手法についても測定のモデル同様に、指導者などを想定したリファレンス実装を開発する。

3.4 応用と評価

テクニカルライティング演習などで3.2の作文アプリケーションを使って測定データを収集する実験をする。収集したデータを3.3の手法を使って分析・可視化する。

評価する。

4. 研究成果

4.1 モデル

開発したモデルをマトリックス型文章編集モデルと呼ぶ(図 4)。このモデルでは、段落といった文章の編集対象の部分と、いつ編

集した時刻と、編集操作の種類と、その部分に関連づけられた他の文章部分や(文章には含まれないが)執筆指針など(コンテキスト)と合わせて、機械的に記録する。コンテキストを軸として関連付けをマトリックス型に整理するのでマトリックス型と呼ぶ。典型的には業務帳票などのワークシートを想像すると近い。ワークシートの物理的なレイアウトを捨象して論理的な構造を記録する。

編集操作の種類としては、ツールは PC やスマホのアプリとして動作することから、アプリケーション一般が通常持っている追加・削除・コピーといった操作をモデルに採用することで、モデルの適用範囲を広めた。

編集対象の文章そのものを記録しないことで、機密保持や個人情報保護に寄与する。

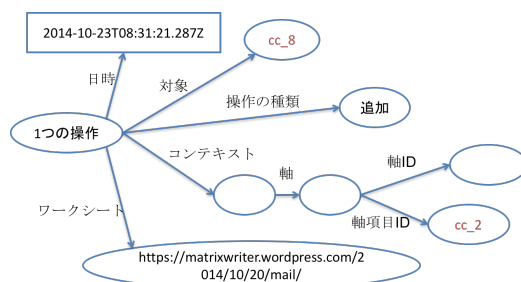


図 4 マトリックス型文章編集モデル

4.2 リファレンス実装

モデルに基づいて編集操作を記録する作文アプリケーション「Topic Writer」を新規に開発した。適用した演習でロジック・ツリーを教えるために、それに対応するアプリケーションを新規に開発し、既存のアプリケーションに測定機能をアドオンする方式は取らなかった。

「Topic Writer」はクラウド型の Web アプリケーションで、スケーラブルなシステムである。Web の技術を採用することで、アクセシビリティ・国際化・相互運用性を高めている。

評価として、開発したアプリケーションのクライアント側(Web ブラウザ上)のメモリ所要量を測定した。作文行動を測定する機能のメモリ所要量が許容範囲内であることを確認した。

また、情報系の学生でなくても、演習で導入して、その演習の時間内ですぐに使いこなすことができた。

4.3 分析手法

図 5 は、段落などが編集される時間的な前後関係に着目して、編集対象になった回数を集計した表である。n 回目(縦軸)に編集した記入欄の次(n+1 回目、横軸)にどの記入欄を編集したかの遷移数を行列で集計した。われわれは、これを編集操作の時間的共起行列と

呼ぶ。時間的に引き続いて編集対象となる枠同士は、書き手にとって関係が深い可能性が高いと考えられる。この原理は、テキスト・マイニングで行われる、文字列や単語の共起分析と同様である。

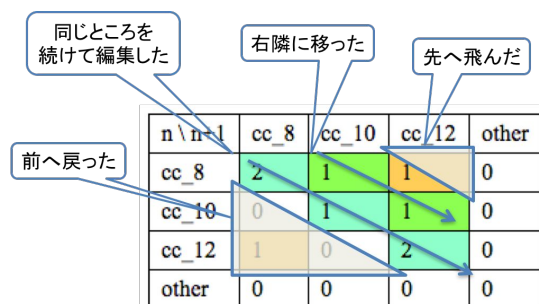


図 5 編集操作の時間的共起行列

図 5 で、cc_8 等は枠の ID である。対角線および 1 つ右のセルに集計された遷移は、単に左から右に書いたことを反映している可能性が高い。それ以外のセルに集計された遷移は、前後の枠間に特別な関係を書き手が感じている可能性が高い。

編集操作の時間的共起行列から、前後して編集対象となった段落間の遠さを積算した指標 Editing Operation Indicator (EOI) を提案した。ある段落を編集して、次に編集する段落が遠くにあるほど、EOI の値は大きくなる。

以上の可視化手法のリファレンス実装として「Writing Analytics」を開発した。「Writing Analytics」はクラウド型の Web アプリケーションである。Web の技術を採用することで、アクセシビリティ・国際化・相互運用性を高めている。

「Writing Analytics」は、講義・演習の講師が、特に指導を受けることなく、受講生の傾向を見るために使うことができた。

4.4 応用

大学・大学院でのライティング演習でリファレンス実装を使って受講生の作文を測定した。ロジック・ツリーを使う作文で「Topic Writer」を使い、測定データを「Writing Analysis」で分析・可視化した。

受講生同士のピアレビュー後に文章を修正する編集について、EOI が高い編集の方が、文章の論理的構成の評点に有意な改善がみられたので(図 6)、研究会で発表し議論した。

| | 全員 | 初期評価と EOI1 | 修正後評価と EOI2 | 改善度と EOI2 |
|----|-------------|------------|-------------|-----------|
| 構成 | ①定義 | 0.1 | 0.4 | 0.4 |
| | ②具体的な説明 | 0.0 | 0.2 | 0.1 |
| | ③メリットの記述 | 0.1 | -0.1 | -0.1 |
| 表現 | ①読み手に合わせた表現 | 0.0 | 0.0 | -0.3 |
| | ②一文一義 | 0.1 | 0.2 | 0.1 |
| | ③韻字脱字表現の統一 | -0.1 | 0.1 | 0.0 |

図 6 EOI と論理的構成の相関

4.5 まとめ

本方式では、データの記録・可視化はコンピュータによって機械的に行われ、大規模化でき、リアルタイムに分析・可視化できる。また、「段落間に『根拠』という関係がある」といった意味付けから中立である。

4.3 に提案した分析手法と 4.4 の応用によって、本研究で開発したマトリックス型文章編集モデルが、文章の論理的構成の指導や学習に役立つ可能性を示すことができ、本研究の目的を達成できた。

細粒度で取得した行動データを分析して傾向を可視化する研究は他にもあるが、行動の傾向とスキル評点との関係を示した研究は他にない。文章の書き方と文章の論理的構成の評点との関係を示すことで、取得したデータの有用性を示した本研究は、この点で特長的である。

4.6 今後

今後は、本方式による測定データを分析・可視化して、指導・学習に役立てる方法を明らかにしていく。このとき、他の分析手法や可視化手法を開発も試みる。

また、「Topic Writer」以外の作文アプリケーションも開発して、データを取得し、分析手法を開発して指導・学習に役立てる方法を検討する。すでに、本研究後に向けて、文でジグソーパズルを行うアプリケーション「ジグソー・テキスト」を提案し、それによる編集過程の測定・分析について考察して、研究会で発表し議論した。

5. 主な発表論文等

〔雑誌論文〕(計 1 件)

- (1) 大場みち子、ぺた語義：考える力を育てる、情報処理 58(4)、pp. 319-319, 2017, 査読無し

〔学会発表〕(計 22 件)

- (1) 山口琢、高橋慈子、小林龍生、大場みち子、作文行動を測定・分析するためのマトリックス型テキスト編集モデルの設計、研究報告コンピュータと教育(CE)、2017-03-11、津田塾大学小平キャンパ

- ス(東京都小平市)
- (2) 山口琢、高橋慈子、小林龍生、大場みち子、文章編集行動の測定・分析手法の考察：手順書とジグソー・テキスト、研究報告ドキュメントコミュニケーション (DC)、2017-03-10、東洋大学白山キャンパス(東京都文京区)
- (3) 高橋慈子、大場みち子、山口琢、小林龍生、テクニカルライティング教育におけるクラウドツール活用と相互レビューの効果分析、研究報告ドキュメントコミュニケーション (DC)、2017-03-10、東洋大学白山キャンパス(東京都文京区)
- (4) 大場みち子、山口琢、高橋慈子、小林龍生、文章作成における文章評価と編集操作との関係分析、研究報告コンピュータと教育 (CE)、2016-12-03、長崎県立大学シーボルト校(長崎県西彼杵郡長与町)
- (5) 山口琢、高橋慈子、小林龍生、大場みち子、文章を構成する過程を測定するジグソー・テキストの開発、研究報告コンピュータと教育 (CE)、2016-12-03、長崎県立大学シーボルト校(長崎県西彼杵郡長与町)
- (6) 山口琢、高橋慈子、小林龍生、大場みち子、作文行動の測定と分析:分析サーバの構築、情報教育シンポジウム 2016(SSS2016)、2016-08-23、グリーンピア大沼(北海道茅部郡森町)
- (7) 大場みち子、山口琢、高橋慈子、小林龍生、論理的文章作成における文章評価と編集操作との関係分析、情報教育シンポジウム 2016(SSS2016)、2016-08-23、グリーンピア大沼(北海道茅部郡森町)
- (8) 高橋慈子、山口琢、大場みち子、小林龍生、文書作成教育におけるトピックライティングツール活用と効果、情報処理学会 第 101 回ドキュメントコミュニケーション研究会、2016-03-24、東京工業大学大岡山キャンパス(東京都目黒区)
- (9) 山口琢、大場みち子、高橋慈子、小林龍生、作文行動の測定と分析: 大学生と社会人の比較例、情報処理学会 コンピュータと教育研究会 132 回研究発表会、2015-12-05、AOSSA(福井県福井市)
- (10) 山口琢、大場みち子、高橋慈子、小林龍生、オープンデータで実現する作文測定分析のシステム構成、情報処理学会 第 98 回ドキュメントコミュニケーション研究会、2015-07-14、公立はこだて未来大学(北海道函館市)
- (11) 小林龍生、鎌田博樹、山口琢、共観年代記に向けて、情報処理学会 第 98 回ドキュメントコミュニケーション研究会、2015-07-14 公立はこだて未来大学(北海道函館市)
- (12) 高橋慈子、山口琢、大場みち子、小林龍生、文章力向上教育におけるトピックライティングツールの活用、情報処理学会 第 98 回ドキュメントコミュニケーション研究会、2015-07-14、公立はこだて未来大学(北海道函館市)
- (13) 山口琢、大場みち子、高橋慈子、小林龍生、高橋修、編集操作の測定でアプローチする自然言語処理の提案、2015 年度人工知能学会全国大会、2015-05-31、公立はこだて未来大学(北海道函館市)
- (14) 山口琢、大場みち子、高橋修、ワークシート操作編集で追跡する組織の知識創造、2015 年度サービス学会 第 3 回 国内大会、2015-04-09、金沢歌劇座(石川県金沢市)
- (15) 山口琢、高橋慈子、小林龍生、大場みち子、高橋修、文章編集操作記録: 活用の展望、研究報告デジタルドキュメント (DD)、2015-03-30、東洋大学白山キャンパス(東京都文京区)
- (16) 山口琢、大場みち子、高橋慈子、小林龍生、高橋修、マトリックス型テキスト編集モデルによる作文行為の時間共起、情報処理学会第 77 回全国大会講演論文集、2015-1 pp.19-20、2015-03-19、京都大学吉田キャンパス(京都市左京区)
- (17) 山口琢、大場みち子、高橋慈子、小林龍生、高橋修、Linked Open Data で実現する作文活動リポジトリと分析、人工知能学会第 34 回セマンティックウェブとオントロジー研究会、2014-11-20、慶応義塾大学日吉キャンパス来往舎(神奈川県横浜市)
- (18) 山口琢、大場みち子、高橋修、相互運用可能な作文計測システムの設計、電気学会研究会資料. IS、第 60 回情報システム研究会、2014-09-25、琉球大学工学部(沖縄県中頭郡西原町)
- (19) 山口琢、大場みち子、高橋修、分析・モニタリングのために作文活動を測定するシステムの設計、日本ソフトウェア科学会第 31 回大会、2014-09-09、名古屋大学東山キャンパス(愛知県名古屋市)
- (20) 山口琢、大場みち子、高橋修、マトリックス型テキスト編集モデルによる編集操作ログの実際と分析方針、情報教育シンポジウム 2014(SSS2014)論文集、2014-08-25、オリビアン小豆島(香川県小豆郡土庄町)
- (21) 山口琢、大場みち子、高橋修、拡張現実 (AR) による 3D テキスト・エディタの提案 研究報告デジタルドキュメント (DD)、2014-07-24、東京工業大学 田町キャンパス キャンパスイノベーションセンター内 広島大学東京オフィス(東京都港区)
- (22) 山口琢、大場みち子、高橋修、マトリックス型文章編集モデルによる知的情報の形式化と活用状況追跡の提案、電気学会研究会資料. IS、第 58 回情報システム研究会 2014-05-16、電気学会(東京都

千代田区)

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

Topic Writer
<https://topic-writer.appspot.com/>

Writing Analysis
<http://writing-analytics.appspot.com/>

6. 研究組織

(1) 研究代表者

大場 みち子 (OBA Michiko)
公立はこだて未来大学・システム情報科学
部・教授
研究者番号：30588223

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：

(4) 研究協力者

()