

# 議事録コンテンツを用いた入力予測による議事録作成支援

井上 慧 松原 茂樹 長尾 確

名古屋大学 大学院情報科学研究科

{kinoue,matubara,nagao}@nagao.nuie.nagoya-u.ac.jp

## 1 はじめに

一般に、会議での議論内容は議事録として記録される。会議の参加者が過去の議事録を閲覧して議論内容を振り返ることは、今後の会議を円滑に行うための有効な手段である。

発言内容をテキストで記録する方法として、音声認識の利用が検討されているものの [1]、完全な自動化は難しい。一方、書記が会議中に発言内容を要約して文書化する、あるいは会議後に音声データを聴いて文書化するなどの方法では、入力にかかる書記の負担が大きい。また、入力予測機能を備えた日本語 IME (Input Method Editor) が開発されているものの [2][3]、議事録など、出現する語句や表現の分布が一般の文書とは異なる専門的な文書の作成において、効果的に機能するとは言い難い。

そこで、本稿では、現在の会議の情報と過去の議事録を用いて書記のテキスト入力を予測する議事録作成支援手法を提案する。提案手法では、書記のテキスト入力時に、現在の会議の情報と過去の議事録を利用して、次に入力される語句をリアルタイムに予測し、入力候補として提示する。書記は提示された入力候補から語句を選択することで、入力に要する時間的コストを削減できる。

## 2 関連研究

### 2.1 入力予測機能を備えた IME

入力予測機能を備えた日本語 IME がいくつか開発されている。

POBox[4] は、フルサイズキーボードが使用できない環境を想定して開発されたテキスト入力手法である。ユーザは、文字の綴り、発音、形状など、入力したい単語の検索キーを入力する。入力システムは検索キーとユーザ文書を利用して、辞書内の単語を動的に検索し、入力を予測し、予測結果を入力候補リストとして提示する。ユーザが候補リストから単語を選択すると、テキスト内に入力され、さらにその次の単語が予測される。

Mozc[2] は、マルチプラットフォームで動作するように設計されたクライアント・サーバ方式の日本語 IME であり、Google 日本語入力をオープンソース化したものである。Web から自動的にマイニングされた大規模な Web 辞書を内蔵しており、かな漢字変換・入力予測に利用している。

Social IME[3] は、Mozc と同様、インターネットを通してサーバ側でかな漢字変換・入力予測を行う日本語 IME である。Web 上の文章から抽出した大規模な統計量を利用して、かな漢字変換・入力予測を行う。また、ユーザが登録した単語をユーザ同士で共有することにより、専門用語や流行語なども変換できるようになる。

この他にも、Microsoft IME や ATOK など、一般に広く用いられている日本語 IME で入力予測機能が提供されている。これらの日本語 IME が備えている入力予測機能は、特定分野の文書の入力を想定して開発されたものではなく、一般的なテキスト入力を補助するものである。このため、議事録など、出現する語句や表現の分布が一般の文書とは異なる専門的な文書に対して、入力予測機能が効果的に機能するかは不明である。また、議事録のテキスト内に出現する語句や表現の分布は、会議の議題や書記によって大きく変化すると考えられるため、入力履歴やユーザ文書、Web から得た統計量などの情報だけでは、議事録の作成支援には不十分である。

一方、本研究では、現在の会議の情報と過去の議事録を利用することで、発言者や書記、会議の議題などの、現在進行中の会議の状況を考慮したテキスト入力予測を行う。

### 2.2 過去の文書データを利用した入力支援

特定の分野の文書を作成する状況を想定して、過去の文書データを利用した入力支援の研究が行われている。

海野ら [5] は、ビジネス・技術文書作成などの、既存の文書と用語や表現を統一させる必要のある文書の作成支援を目的とした入力予測手法を提案している。文字列検索結果の周辺文脈の圧縮手法 [6] を用いて、入力文字列から始まる頻出の表現を特定分野の文書集合か

ら動的に抽出し、入力候補としてユーザに提示することで、分野に依存した入力支援を行う。

本研究でも同様に、議事録という既存の文書を用いて入力候補を抽出・提示する。ただし、テキストそのものだけでなく、発言者や書記、会議の議題などの情報も予測に利用することで、より現在の状況に適合した入力候補の提示を行うことができる。

橋本ら [7] は、胃癌に関する病理診断報告書を対象とした入力支援システムを提案している。提案システムは、自由入力による報告書の作成を前提としており、Web ブラウザ上でのテキスト入力を支援する。人手で作成したオントロジーを利用して、入力補完候補を推定し、ユーザに提示する。

一方、本研究では、過去の議事録から機械的に生成した辞書および学習モデルを利用するため、人手による単語の辞書登録やオントロジー作成の必要はない。同一の属性をもつ議事録が一定数蓄積されていれば、入力候補の提示による議事録作成支援が可能である。

### 3 議事録コンテンツ

著者らの研究室では、ディスカッションマイニングと呼ばれる、テキスト・音声・映像を含む議事録コンテンツの作成、及び、その効果的な再利用を目的とした研究を行っている [8]。ディスカッションマイニングは研究室のゼミで実用化されており、大量の議事録コンテンツが作成・蓄積されている。2013年3月時点で489件の議事録コンテンツが存在し、総発言数は30,412、書記テキスト内の文の数は69,695である。

#### 3.1 議事録コンテンツの構成

議事録コンテンツは、以下の要素から構成される。

- 映像・音声
- 参加者情報
- 発表スライド
- 発言内容のテキスト（以下、書記テキスト）
- 発言へのマーキングやアノテーション

これらは、カメラ、マイクロフォン、リモコンデバイス、各種ツールを用いて記録されている。映像・音声・書記テキストは、発言の開始・終了時間に基づき発言ごとにセグメントされており、発言時に発表者によって表示されていた発表スライドのシートが関連付けられている。また、書記テキストは、発言ごとの属性に基づき構造化されている。

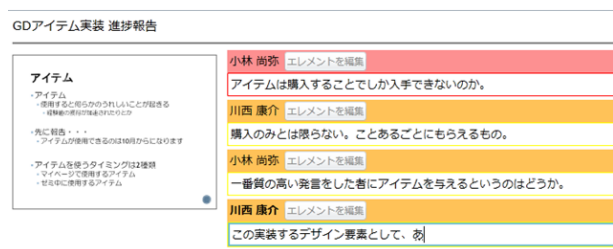


図 1: 書記ツール

#### 3.2 書記ツールによる書記テキスト作成

書記テキストは、書記に指名された参加者の一人が作成する。

書記は、会議中に、図 1 に示す Web ブラウザベースの専用ツール（以下、書記ツール）を用いて発言内容をリアルタイムに記録する。書記ツールは参加者全員が持つリモコンデバイスと連動しており、会議中、参加者の発言に関する情報がリモコンを経由して随時追加される。参加者が発言するたびに、書記は書記ツール内に自動的に追加されたフォームを選択し、発言内容を直接聴いて書記テキストを作成する。

また、会議終了後に、議事録閲覧用のブラウザを用いて、発言時の映像・音声を参考にして書記テキストを編集できる。

### 4 議事録を用いた入力予測

本稿では、前章で説明した議事録コンテンツを提案手法に適用する。すなわち、スライド、書記テキスト、発言者や書記、発表者などの議事録中の情報を用いて、現在の会議の状況に適合した入力予測を行う。

書記テキストの入力予測には、過去の議事録の書記テキストや現在の会議の発表スライドが有用であると考えられる。提案手法では、過去の書記テキスト、ならびに、発表スライド、それぞれに基づいて形態素列を予測した結果をまとめて入力候補とする。

#### 4.1 過去の議事録を用いた入力予測

過去の議事録の書記テキストから形態素列の入力を予測する。

研究室のゼミのような、同じ発表者が定期的に発表を行う種類の会議においては、発表者によって会議の議題が異なる。書記テキスト内に出現する語句や表現の分布は、会議の議題に大きく影響される。同様に、書記や発言者によっても、書記テキスト内に出現する語句や表現の分布は大きく変化すると考えられる。発言者が発表者であるときの発言は、議題の内容に特に深く関わっている。また、発言内で共起しやすい形態素の組み合わせがあると考えられるため、すでに現在の

フォームに記録されている書記テキスト中の形態素も予測に有効に働く。

会議の状況  $s$  のもとで、現在のフォームに形態素  $w_0$  が入力されたとき、次に形態素列  $w_1 w_2 \cdots w_n (= w_1^n)$  が入力される確率  $P(w_1^n | w_0, s)$  を次式で求める。

$$\begin{aligned}
 P(w_1^n | w_0, s) &= P(w_1 | w_0, s) \times P(w_2 | w_0^1, s) \times \cdots \times P(w_n | w_0^{n-1}, s) \\
 &\simeq P(w_1 | w_0, s) \times P(w_2 | w_1, s) \times \cdots \times P(w_n | w_{n-1}, s)
 \end{aligned}
 \tag{1}$$

$P(w_i | w_{i-1}, s)$  は、最大エントロピー法で求める。会議の状況として、発表者名、記録者名、発言者名、発言者と発表者が同一であるか否かを素性とする。また、形態素  $w_{i-1}$  に対しては、表層形、品詞、活用型、活用形を素性に用いる。

また、形態素列  $w_1^n$  の入力に必要なキーストローク数を  $key(w_1^n)$  とする。会議の状況  $s$  のもとで、現在のフォームに形態素  $w_0$  が入力されたとき、形態素列  $w_1^n$  の提示順を決めるスコアを次式で求める。

$$P(w_1^n | w_0, s) \times key(w_1^n) \tag{2}$$

ただし、書記が現在入力している文字列と前方一致しない形態素列のスコアを 0 とする。

すべての形態素列のスコアを求めることは効率的ではないため、本手法では、以下のアルゴリズムによって入力候補となる形態素列を選定する。

各形態素についてスコアを求め、上位  $N$  位以内の形態素列を入力候補として選定する。選定された入力候補  $w_1^n$  ごとに、入力候補の末尾に任意の形態素を加えた形態素列  $w_1^n w_{n+1}$  のスコアを求める。選定済みの入力候補  $w_1^n$  と、スコアを求めたすべての形態素列  $w_1^n w_{n+1}$  のうち、スコアが上位  $N$  位以内のものを新たに入力候補として選定する。上位  $N$  位の入力候補に変化が見られなくなるまで上記の手続きを繰り返す。なお、本研究では  $N = 10$  とした。

## 4.2 スライド内の文章を用いた入力予測

会議中の発言内容は、その大部分が現在の発表内容に対する質問や指摘、意見であるため、発言時に表示されているスライドのシートの内容に関連したものであることが多い。このことから、スライド内に含まれる形態素列は書記テキストの有力な入力候補になると考えられる。提案手法ではスライド内のテキストから入力候補となる形態素列を抽出する。

入力候補となる形態素列は、単体で意味をもつ必要がある。そのため、入力候補は、以下の条件を満たす形態素列とする。

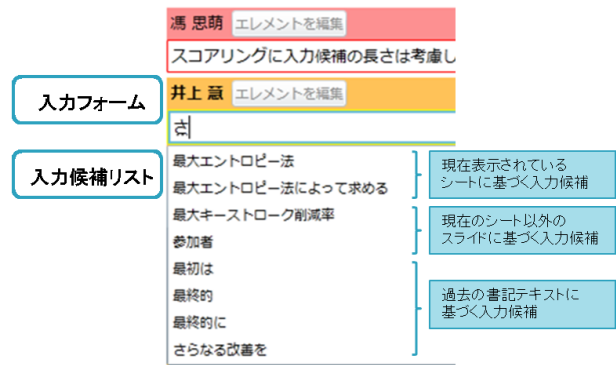


図 2: 書記ツールでの入力候補の提示の様子

- 一文中で接続する形態素列である
- 構文的にまとまっている

後者の条件を満たすために、スライド中の文を構文解析し、構文解析結果の任意の文節から葉ノードとなる文節に至るまでのパスで構成されている形態素列を抽出する。ただし、抽出した形態素列の末尾が助詞（接続助詞を除く）、または読点であった場合は、末尾の形態素を削除する。また、1 文節内で 2 形態素以上連続している名詞の列や未知語も抽出する。形態素列の読みは、形態素情報から機械的に生成する。

例えば、『過去の議事録を用いて予備実験を行う』という文からは、『過去の議事録』、『過去の議事録を用いて』、『過去の議事録を用いて予備実験を行う』、『予備実験を行う』、『予備実験』、『議事録』が抽出される。

抽出した形態素列のうち、書記が現在入力している文字列と前方一致するものを入力候補とする。

## 5 書記ツールでの入力候補の提示

議事録コンテンツを用いた入力予測手法を、書記ツールに実装した。

書記が発言内容を記録しているときに、書記ツールは現在の発言フォームに適合したテキスト入力を予測し、入力候補として提示する。図 2 に、実際の書記ツールでの入力候補の提示の様子を示す。

入力候補には、過去の書記テキストに基づくものと現在のスライドに基づくものがあるが、それらの入力候補の提示順は、以下のように決定する。

1. スライドに基づく入力候補のうち、現在表示されているシートから抽出したものを、辞書順に提示する。
2. スライドに基づく入力候補のうち、1. による入力候補に含まれない複合名詞と未知語を、1. による入力候補の下に、辞書順に提示する。
3. 過去の議事録に基づく入力候補を、1. と 2. による入力候補の下に、スコアの低い順に提示する。

## 6 評価実験

### 6.1 実験の目的と方法

議事録作成作業の効率化に対する提案手法の効果を確認するために、過去の議事録を用いて評価実験を行った。

実験では、過去の議事録の書記テキストを形態素解析し、各形態素の読みから議事録作成時の書記による入力を機械的に生成することにより、各入力での入力候補の提示を実現し、最適な選択をした場合にどの程度キーストローク数を削減できるかを調べた。

ここで、正解語句の入力に必要なキーストローク数を  $K_c$ 、書記による入力のキーストローク数を  $K_p$ 、正解語句の提示順位を  $R$  とするとき、キーストローク削減数  $G$  を以下のように定義した。

$$G = (K_c - K_p) - R \quad (3)$$

提示された入力候補の中から正解語句を選択するのに必要なキーストローク数が  $R$  であるため、 $(K_c - K_p)$  から  $R$  を引いている。また、10 位以内に正解語句の候補がなかった場合、あるいは上式で  $G$  が負になる場合、 $G = 0$  とする。

本研究室のゼミの議事録のうち、学習データとして議事録 464 件を、テストデータとして議事録 25 件を用いた。

### 6.2 実験の結果と考察

評価実験の結果を表 1 に示す。

最大キーストローク削減数とは、ユーザが最適な選択をした場合のキーストローク削減数の合計である。また、最大キーストローク削減率とは、書記テキスト全体のキーストローク数に対する最大キーストローク削減数の割合である。

書記テキスト全体にかかるキーストローク数の 16% 以上を削減できる可能性があり、議事録作成作業の効率化に対する本手法の効果を確認した。

また、過去の書記テキストに基づく入力候補とスライドに基づく入力候補が、それぞれどのように議事録作成作業の効率化に寄与しているのかを調査した。

実験において、入力候補を選択した方がキーストローク数を削減できるケースは 11,264 件あった。そのうち、過去の書記テキストに基づく入力候補を選択するケースは 10,433 件 (92.62%) あり、入力候補の選択による平均キーストローク削減数は 2.30 であった。一方、スライドに基づく入力候補を選択するケースは 831 件 (7.38%) あり、入力候補の選択による平均キーストローク削減数は 9.00 であった。過去の書記テキストに基づく入力候補は選択される頻度が高く、スライドに基づく入力候補は選択されたときのキーストローク削減数が大きいため、議事録作成作業の効率化に対してどちらの入力候補も有効であるとの知見が得られた。

表 1: 実験結果

書記テキスト全体のキーストローク数	194,023
最大キーストローク削減数の合計	31,436
最大キーストローク削減率 (%)	16.20

## 7 おわりに

本稿では、議事録コンテンツを用いて、会議の状況に応じて議事録作成者のテキスト入力を予測する議事録作成支援手法を提案した。また、実験によって、議事録作成作業の効率化に関する本手法の効果を確認した。

今後の課題として、文字入力時以外での入力候補の予測・提示の実現、及び、議論構造などの議事録コンテンツの構成要素の利用が挙げられる。

## 参考文献

- [1] 千代章, 上山学, 吉本大樹, 大塚隆宏, 議事録作成支援ソフトウェア VoiceGraphy, NEC 技報, Vol.63, No.1, pp.59-61, 2010.
- [2] 工藤拓, 小松弘幸, 花岡俊行, 向井淳, 田畑悠介, 統計的な漢字変換システム Mozoc, 言語処理学会第 17 回年次大会発表論文集, pp.948-951, 2011.
- [3] 奥野陽, 萩原将文, インターネットを用いた日本語入力システム, 情報処理学会研究報告, NL-190, pp.1-6, 2009.
- [4] Toshiyuki Masui, POBox: An Efficient Text Input Method for Handheld and Ubiquitous Computers, *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing(HUC-99)*, pp.289-300, 1999.
- [5] 海野裕也, 坪井祐太, 頻出文脈に基づく分野依存入力支援, 言語処理学会第 17 回年次大会発表論文集, pp.1107-1110, 2011.
- [6] 海野裕也, 坪井祐太, 文字列検索結果に対するコンパクトな文脈集合の高速抽出, NLP 若手の会第 5 回シンポジウム, 発表 7, 2010.
- [7] 橋本泰一, TAM Wailok, 鷹合基行, 荒牧英治, 宇於崎宏, 橋田浩一, 病理診断報告書作成のためのオンロジーを利用したテキスト入力支援, 言語処理学会第 17 回年次大会発表論文集, pp.940-943, 2011.
- [8] 土田貴裕, 太平茂輝, 長尾確, ゼミコンテンツの再利用に基づく研究活動支援, 情報処理学会論文誌, Vol.51, No.6, pp.1357-1370, 2010.