

コールセンターの音声認識テキストを用いた テキスト要約技術の研究

Extractive Summarization of Call Center Conversation Using Voice-to-text Data

コールセンターでは、対応の品質向上などを目的として、音声認識システムにより蓄積された通話の音声認識テキストの活用が進められている。本研究では、音声認識テキストを要約して通話内容を容易に把握できるようにするテキスト要約技術を確立するため、単語の分散表現を用いた要約手法を通話の音声認識テキストに適用評価した。分散表現のコサイン類似度に基づく従来手法では質問内容に関する発話が要約に含まれない場合があるため、通話内の発話を時間に沿ったウィンドウで区切って要約を行う手法を考案した。従来手法と比較した結果、要約率約 20%のときに従来よりも約 10%高い再現率で質問内容に関する発話が要約に含ま

2.2 従来の抽出的要約手法

抽出的要約は、要約対象の文書 $D = \{s_1, \dots, s_N\}$ から文を抽出して要約を生成する技術であり、以下の式(1)の最適化問題として定式化できる³⁾。

$$\operatorname{argmax}_{S \subseteq D} F(S) \quad \text{subject to} \quad \sum_{s_i \in S} c_i \leq b \quad (1)$$

s_i は文書 D に含まれる文、 $S \subseteq D$ である S は文書 D の要約、 $F(S)$ は要約 S の良さを表すスコ

L_r が十分小さければ、各ウィンドウ W_k における主要な内容は一つの話題に限定され、要約 S_k にはその話題に関する発話が含まれる。特にそれが質問内容に関する話題のとき、要約 S_k には質問内容に関する発話が含まれる。その結果、通話全体の要約 S にも質問内容に関する発話が含まれるようになる。

なお、 T_r には正の値を設定することで s_1 や s_2 などの最初の方の発話が抽出されやすくなる効果がある。また、ステップ 8 で単語 w の分散表現 x_w に乗算している idf_w は要約対象の文書 D 全体における単語 w の IDF (Inverse Document Frequency) である。文書 D

