

1 はじめに

これまで、単語重要度、言語尤度、音声認識時における信頼度、および、原文の係り受け関係に基づき、単語抽出による音声自動要約手法を提案してきた [1]。日本語の音声を対象として音声自動要約を検討してきたが、提案手法はコーパスに基づく確率モデルを適用していることから、日本語に特化した手法ではなく、他言語への応用が可能であると考えられる。本稿では、英語のニュース音声の書き起こしを対象として、これまで提案してきた要約手法を拡張し、自動要約を行った結果を報告する。

2 音声自動要約手法

英語のテキストを対象とした文単位の要約の研究として、Knight らが確率付きの書き換え規則で表現された SCFG(Stochastic Context Free Grammar) に基づき、文圧縮 (sentence compression) による自動要約手法を提案している [2]。この手法は、原文と要約文が対になったコーパスを用いて、原文の構文木の書き換え規則が要約文で別の書き換え規則に変換、または削除される確率を学習し、自動要約に適用する。しかし、音声認識結果の自動要約を目的とした場合、必ずしも文法に当てはまらない音声認識結果を構文解析することは困難である。さらに、誤った構文解析結果に基づき音声認識結果を要約することは適切ではない。また、要約文と原文が対になっている大規模なコーパスを構築することも容易ではない。

これまで我々が提案している音声自動要約手法は、話題語を中心として言語的に尤もらしく単語を抽出するため、原文と要約文が対になったコーパスを必要とせず、既存のテキストコーパスを適用することが可能である。さらに、構文解析済みのコーパスから単語単位の係り受け SCFG(SDCFG:Stochastic Dependency Context Free Grammar) を学習することにより、確率的な構文情報を組み込むことが可能である。

2.1 要約スコアの定義

単語抽出により生成された要約文の適正を示すスコアとして要約スコアを定義する。要約スコアは、単語重要度スコア I と言語スコア L 、信頼度スコア C 、および、単語遷移スコア T_r に基づき、次式のように定義する。 N 個の単語からなる認識単語列 $W = w_1, w_2, \dots, w_N$ から要約文として M ($M < N$) 個の単語を抽出し接合した単語列 $V = v_1, v_2, \dots, v_M$ の要約スコアは次式によって示される。

$$S(V) = \sum_{m=1}^M \{L(v_m | \dots v_{m-1}) + \lambda_I I(v_m) + \lambda_C C(v_m) + \lambda_T Tr(v_{m-1}, v_m)\} \quad (1)$$

但し、 $\lambda_I, \lambda_C, \lambda_T$ は各スコアのバランスをとるための重み係数である。認識された単語列より抽出された部分単語列を $V = v_1, v_2, \dots, v_M$ ($M < N$) とするとき、要約処理は (1) 式で表される要約スコアを最大にする \hat{V} を求める問題となる。

単語重要度スコア $I(v_m)$ には、単語の出現頻度に基づく情報量を適用する。言語スコア $L(v_m | \dots v_{m-1})$ は、要約文内の単語連鎖の適正度を示すスコアである。本研究では、単語 trigram を用いる。信頼度スコア $C(v_m)$ は、認識結果に含まれる認識誤りを要約文に抽出しないよう、音響的、言語的に信頼度の低い単語に対しペナルティを与えるものである。本稿では、英語のニュース音声の書き起こしを対象としているため、信頼度スコアは適用しない。

単語間遷移スコア $Tr(v_{m-1}, v_m)$ は、要約文内の単語連鎖が原文において係り受け関係にあるか否かを示す単語間遷移確率の対数値で定義され、係り受け関係にない単語連鎖にペナルティを与えるものである。単語間の係り受け関係は、SDCFG の確率により推定する。

英語の SDCFG は、前から後ろへの係り受け関係も考慮した以下の規則からなる。

$$\begin{aligned} \alpha &\rightarrow \beta\alpha \\ \alpha &\rightarrow \alpha\beta \\ \alpha &\rightarrow w \end{aligned}$$

ここで、 α, β は任意の非終端記号、 w は終端記号 (単語) を表す。

これらの規則が適用される確率を基に、Inside-Outside 確率を計算し、二つの単語が係り受け関係にある確率を求める。図 1 に示すように、 w_m と w_n が係り受け関係にある確率は、非終端記号 S から文 $w_1 \dots w_L$ が生成されたときに、 $\alpha \rightarrow \beta\alpha$ が適用されて、 β から文節 $w_i \dots w_k$ が生成され w_m がその文節の主辞 (β に直接関連する単語) となり、かつ、 α から文節 $w_{k+1} \dots w_j$ が生成され w_n がその文節の主辞となる確率を、すべての α, β, i, j, k に関して総和をとることで計算できる ($\alpha \rightarrow \alpha\beta$ の場合も含む)。しかし、この係り受け確率をそのまま単語間遷移確率として用いると、要約時に w_m から $w_{k+1} \dots w_{n-1}$ (図の影の部分に対応する単語列) に遷移できず、そのような単語列が削除されてしまう。このため、 w_m と w_n の係り受け確率は、 w_m から w_{k+1}, \dots, w_{n-1} のそれぞれに遷移する確率も加算する。

2.2 動的計画法による音声要約

要約スコアが最大となる部分単語列を、自動要約文として動的計画法により決定する。さらに複数発話を要約する手法として、2 段 DP による要約手法を適用した。この手法は、各発話を可能性のある全ての要約率で要約後、全体として目的的要約率となる各文の要約文の組み合わせから、要約スコアを

* Automatic speech summarization of English news speech.

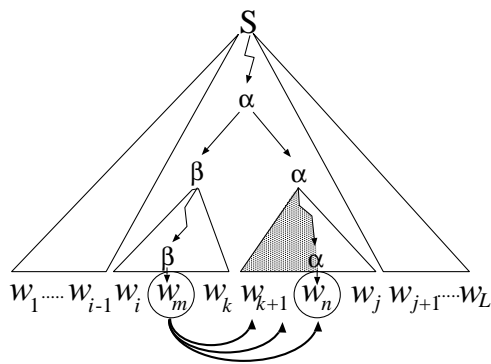


図 1. 英語の単語間遷移確率

最大とする組み合わせを決定する。

3 正解要約文単語ネットワークに基づく自動要約法の評価

被験者の作成した正解要約文をネットワークで表現し、その正解要約文ネットワーク上で自動要約文自身に最も近い単語列を正解として、要約正解精度により自動要約文を評価する [3]。

4 評価実験

4.1 実験条件

NIST 主催の Topic Detection and Tracking(TDT) のタスクから、CNN ニュース 5 記事の書き起こし文をテストセットとした。単語には品詞情報を付加して各モデルを学習し、評価時にも品詞情報を付加した。品詞付けには Brill tagger を用いた [4]。5 記事中の 50 発話文 (平均 30 形態素) について、50-70% と 20-40% の要約率で各文要約を行った。さらに、5 記事を記事単位で、60-70% と 20-30% の要約率で複数文要約を行った。

生成された自動要約文を、英語を母国語とする 17 人の被験者の作成した正解要約文の単語ネットワークに基づき評価を行った。

4.2 要約処理部の構成

単語重要度、要約用言語モデル、および単語間遷移確率に適用する SDCFG (非終端記号数 100) は、Penn Treebank コーパスに含まれる Wall Street Journal と BROWN コーパス 10681 文 (約 3.5M 単語) を用いて算出した [5]。

SDCFG は、Penn Treebank に付与された非終端記号を適用せず、括弧付きのコーパスとして非終端記号数のみを固定して学習した。ただし、非終端記号数は 100 である。

4.3 評価結果

図 2 に各文要約、図 3 に複数文要約の正解要約文単語ネットワークに基づく単語正解精度を示す。提案手法の有効性を検証するため、自動要約文と等しい要約率で単語をランダムに抽出した要約文 (RDM) に対して評価を行った。さらに、被験者各 17 人の被験者の正解要約文を他の 16 人の正解要約文で作成した要約文単語グラフに基づき評価した平均単語正解精度 (SUB) を示す。

全ての自動要約条件において、ランダムに単語を

抽出した場合と比較して、要約正解精度が有意に高くなることが示された。各文要約、複数文要約のどちらにおいても、要約率が低い条件では、単語重要度 I に比べ言語尤度を組み合わせた I_{LL} の精度が高い。さらに、単語間遷移スコアを組み合わせた I_{LLT} の要約精度が最も高いことが示されている。しかし、高要約率では言語スコアおよび単語間遷移スコアの効果はわずかである。また、被験者の作成した正解要約文の精度には至っていない。

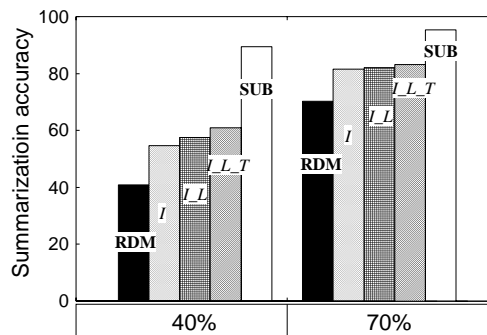


図 2. 40%,70%要約率における各文要約の要約正解精度

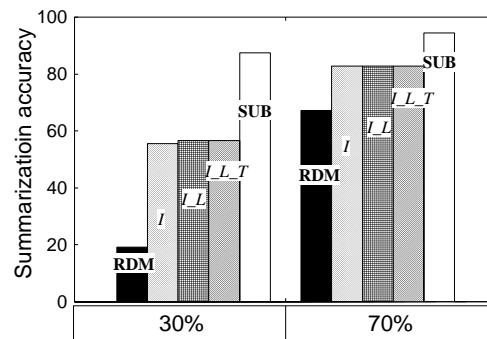


図 3. 30%,70%要約率における複数文要約の要約正解精度

5 まとめ

本稿では、英語のニュース音声の書き起こしを対象として、単語抽出により自動要約を試みた。日本語と同様、英語に対しても提案手法が有効であることを示した。今後、音声認識結果に対しても自動要約を行う予定である。

謝辞

正解要約文作成に協力してくださった Sheffield 大学の Yoshi Gotoh 氏に感謝致します。

参考文献

- [1] 堀, 古井, 信学技報, SP2000-95-116, pp.127-132(2000).
- [2] Knight and Marcu, Proc. National Conference on Artificial Intelligence (AAAI), 2000.
- [3] 堀, 古井, 音講論, Vol.1, pp.93-94(2001-3).
- [4] <http://www.cs.jhu.edu/~brill>
- [5] <http://www.cis.upenn.edu/~treebank>