

通訳クラウドソーシングにおける不完全なタスク結果の統合手法

橋本 大空 †

白石 優旗 ‡

張 建偉 ‡

若月 大輔 ‡

関 洋平 †

森嶋 厚行 †

† 筑波大学

‡ 筑波技術大学

1 はじめに

近年、ネットワークを通じた問題解決の手法としてクラウドソーシングが注目を集めている。クラウドソーシングとは、問題解決のための作業（以下、タスク）を不特定多数の人々からなる作業員（以下、ワーカ）に依頼し、問題の解決を図る手法である。クラウドソーシングには様々な種類があり、その中にリアルタイムクラウドソーシングがある。リアルタイムクラウドソーシングとは、回答時間の制約や結果を出力する時間の制約が厳しいタスクを扱うクラウドソーシングであり、具体例として、講演の文字起こしを行うクラウドソーシング [1] が挙げられる。

クラウドソーシングの課題に、不特定多数の人々にタスクを依頼するために、タスク結果の品質が安定しないという点がある。品質の向上手法としてタスク結果を統合し、品質の高いタスク結果を選ぶことで、タスク結果全体の品質を向上させることができる。したがって、タスク結果の統合が重要になる。

統合の方法には様々なものがあり、選択式のタスクでは多数決、自由入力式のタスクでは複数ワーカのタスク結果を候補とした投票を行うといった手法が存在する。しかし、不完全なタスク結果が出現しやすいタスクでは、これらの統合手法では確からしい候補が少なくなってしまうため、統合が難しい。

ここでリアルタイムクラウドソーシングの一種である通訳クラウドソーシングを考える。通訳クラウドソーシングでは時間的な制約から、各ワーカのタスク結果である列挙された候補の中に確からしい候補の数が少なくなりやすい。この場合、投票を行っても確からしい候補の数が少ないために、品質の向上が期待できない。

そこで本稿では、ワーカによる不完全な通訳文を他のワーカの通訳文から補完することでより良い通訳文の候補を生成する手法を提案する。提案手法の概要を図 1 に示す。本手法は、「良い補完結果は、補完すべき文に存在する語の共起語を変えず元のタスク結果と似ている」という仮説を基に、日本語の通訳文が不完全だっ

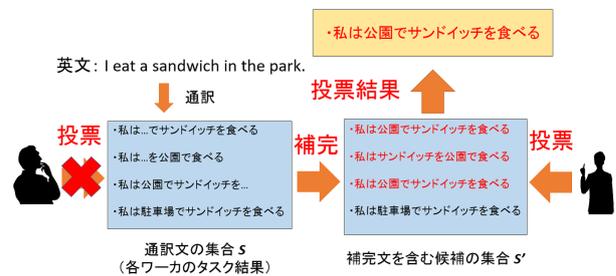


図 1: 提案手法の全体図

た場合、その文を補完した文を出力する手法である。関連研究. 論文 [1] は文字起こしにおける不完全なタスク結果の統合に関する研究である。また、論文 [2] はタスクを行う順序を変更して通訳クラウドソーシングの品質を向上を目指す研究である。我々の知る限り、本研究は通訳クラウドソーシングにおける不完全なタスク結果の統合手法を提案する初めての研究である。

2 本稿で扱う問題

不完全な通訳文. 本稿で扱う問題では完全なタスク結果が分かっていないため、タスク結果から情報が不足している箇所を自動で判別するのは難しい。本稿では、情報が不足している箇所に、明示的にワイルドカード*が配置されているものと仮定する。ワイルドカードは複数含まれていてもよい。例えば、「*は*でサンドイッチを食べる」は不完全な通訳文である。以降、不完全な通訳文のワイルドカードを他のワーカの通訳文中にある語で置換することを補完といい、補完によって生成された文を補完文という。また、補完語の候補を補完候補という。

問題の定義. 本論文で扱う問題の入力を次のようにモデル化する。

- あるタスクの結果として複数のワーカから得られた通訳文の集合 $S = \{s_1, s_2, \dots, s_n\}$
- 不完全な通訳文の集合 $R = \{s_{r1}, s_{r2}, \dots, s_{rm}\} \subseteq S$
- S 中に現れるワイルドカード以外の語の集合 $W = \{w_1, w_2, \dots, w_u\}$

上記の入力が与えられたとき、通訳文補完の問題とは、補完文の集合 $R' = \{r_S(s_{r1}), r_S(s_{r2}), \dots, r_S(s_{rm})\}$ を出力することである。ただし、補完文 $r_S(s_{ri})$ とは、 s_{ri} に現れる全てのワイルドカード*をいずれかの $w_i \in W$ で置

Integration method for incomplete task results of translation crowdsourcing

†Hirota Hashimoto ‡Yuhki Shiraishi ‡Zhang Jianwei

‡Daisuke Wakatsuki †Yohei Seki †Atsuyuki Morishima

†University of Tsukuba

‡Tsukuba University of Technology

き換えたものである。また、補完文を含む候補の集合を $S' = (S - R) \cup R'$ と表記する。

3 望ましい補完文の条件

通訳文補完を行うためには、望ましい補完文の条件を定める必要がある。その条件を説明するために、次の記法を用いる。

- 元の通訳文集合 S における共起語を表す関数を $C: W \rightarrow 2^W$ とする。例えば図1の S において $C(\text{食べる}) = \{\text{私}, \text{サンドイッチ}, \text{公園}, \text{駐車場}, \text{は}, \text{で}, \text{を}\}$ である。また、補完文を含む候補の集合 S' の共起語を表す関数を同様に $C': W \rightarrow 2^W$ とする。 $0 \leq \Omega(C, C') \leq 1$ をこれらの類似度とする
- 二つの文 s_i, s_j の類似度を $0 \leq \text{Sim}(s_i, s_j) \leq 1$ で表す

このとき、 R 中の不完全な通訳文 s_{ri} に対して、次の条件1, 2を満たす補完候補を補完した文が望ましい補完文 $r_S(s_{ri})$ である。

条件1: 補完する文中に既に存在する語の共起語をできるだけ変えない。すなわち、 $\neg \exists C'' : (\Omega(C, C'') > \Omega(C, C'))$ となるような共起語関数 C'' が存在しないことである。

条件2: 補完前のタスク結果の文とより類似している。すなわち、 $\text{SimSUM}(s) = \sum_{s_j \in S} \text{Sim}(s, s_j)$ とすると、 $\neg \exists s'_{ri} (\text{SimSUM}(s'_{ri}) > \text{SimSUM}(r_S(s_{ri})))$ (ただし、 s'_{ri} は s_{ri} の*を W 中の語で置き換えたもの) である。

条件1は、仮説中の「補完すべき文に存在する語の共起語を変えず」に対応し、条件2は「元のタスク結果と似ている」に対応している。

さらに、助詞、助動詞、および補完する文中に既に存在する語は共起語を変えないことが自明であるため、補完候補としない。

4 評価実験

評価実験では、提案手法を用いて補完した結果の通訳文が、元の不完全な通訳文と比較して品質が向上しているかを評価した。

使用データ。2種類のデータを使用した。

データセットA: 筑波技術大学で行われた手話の読み取り通訳タスクのデータの情報が不足している箇所に手作業で*を入れたもの。手話8文のそれぞれに対して1タスク(合計8タスク)を行い、1タスクあたり平均11.5文の通訳文が得られた。

データセットB: TOEICの参考書[3]に付属しているリスニング用CDの英文を通訳したタスクのデータ。英文18文のそれぞれに対して1タスク(合計15タスク)を行い、1タスクあたり平均4.0文の通訳文が得られた。

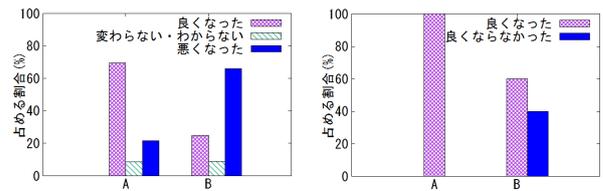


図2: 左: 補完した通訳文の評価 右: タスク単位の評価

利用する指標。提案手法では、共起語の類似度と文の類似度の2つの類似度を用いて、補完すべき語を選択している。本実験では、類似度算出指標として、共に Jaccard 係数 [4] を用いる。ここで共起語の類似度の算出には、共起語の集合を用い、文の類似度の算出には、文を形態素 bi-gram で分割したものの集合を用いる。評価結果。評価結果を図2に示す。左の補完した通訳文の評価では、データAでは元の不完全な通訳文と比較して品質が良くなった結果が多くなり、データBでは悪くなった結果が多くなった。

右のタスク単位での評価とは、各タスク中で元の不完全な通訳文と比較して品質が良くなった補完文が1文でも得られたタスクの割合である。データAでは全タスクで良くなった補完文が得られ、データBでも60%のタスクで良くなった補完文が得られた。また、両評価結果でデータAの方が良い結果が得られていることが分かった。これは、データBでは適切な補完語が他の文に存在する割合が低かったためである。

5 おわりに

本稿では通訳クラウドソーシングにおける不完全なタスク結果である不完全な通訳文を補完し、より良い通訳文を生成する手法を提案した。今後は、望ましい補完文の条件を再検討するなどの課題に取り組むことを予定している。

謝辞。本研究の一部は、JST CREST および JSPS 科研費 (#16K16460, #26870090, #15K01056, #16H02913, #25240012) の支援による。

参考文献

- [1] Iftexhar Naim, Daniel Gildea, Walter S Lasecki, and Jeffrey P Bigham. Text alignment for real-time crowd captioning. In NAACL-HLT, pp. 201–210, 2013.
- [2] Vamshi Ambati, Stephan Vogel, and Jaime Carbonell. Collaborative workflow for crowdsourcing translation. In Proc. of CSCW'12, pp. 1191–1194, 2012.
- [3] 菊間ひろみ. 新テスト対応!!これだけ!TOEIC テスト総合対策初めて 650 点. あさ出版, 2006.
- [4] Paul Jaccard. The distribution of flora in the alpine zone. New Phytologist, 11(2), pp. 37–50.