

不完全な検索クエリにおける 対話履歴からの検索ワード補完手法の検討

柴田 遼一[†] 吉野 哲平[‡] 松森 匠哉[‡] 福地 庸介[‡] 木本 充彦[†] 今井 倫太[†]

慶應義塾大学 理工学部[†] 慶應義塾大学 大学院開放環境科学専攻[‡]

1. 序論

本研究の目的は、ユーザがストレスを感じることなく利用できる検索システムの実現である。その中で我々は、ユーザが対話中に感じた疑問の解を、不完全文による入力での検索可能な Anytime クエリの実現を目指す。

従来研究では、単一のクエリに対してシステムがユーザに問い返しを繰り返すことでクエリを補完する研究がある[1]。固有表現や BERT を用いて新聞記事のような特定の文章を対象にしたクエリを拡張する研究[2][3]も存在する。

しかし、[1]では対話で明らかなキーワードもシステムが質問することによりユーザにストレスを与えてしまうことがある。さらに、進行する対話の中で生じる質問クエリに対して動的に補完するという問題は対象とされてこなかった。

本研究では、対話を逐次的に処理しながら不完全クエリに対して対話履歴中から適切に単語を補完する手法である SQuery を提案する。SQuery では、SCAIN[4]を用いることで、逐次的に進行する対話の中から補完するキーワードの候補群を抽出する。さらに本稿では、抽出された候補群からクエリに補うべき単語を選択する二種類の手法(E-QC, B-QC)を提案する。

2. システム構成

図1にSQueryのシステム図を示す。

2.1. SCAIN

時刻 t において対話文 c_t がSCAINへ入力されるとSCAINは時刻 t までにSCAINに入力されたすべての対話文 $c_{1:t} = (c_1, \dots, c_t)$ 中に含まれるすべての名詞に対して重要度を計算する。重要度が r 番目の名詞を n_{rt} 、その重要度を i_{rt} とし、それらの集合を $N_{1:t} = \{(n_{rt}, i_{rt})\}$ とする。Queryが入力されるとSCAINは、 $N_{1:t}$ を出力する。

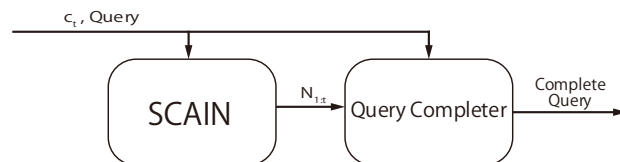


図1 システム構成図

2.2. Query Completer

Query Completerは、SCAINが出力する $N_{1:t}$ を入力として、検索に必要な単語がすべて補完された完全クエリを出力する。本稿ではQuery Completerの実装としてE-QCとB-QCの二種類を提案する。

2.2.1. E-QC

E-QCでは以下の4stepで完全クエリを生成する。

Step1: 固有表現の置換

まずQueryのうち名詞・動詞・形容詞・形容動詞の単語 q のみを抽出した集合 $Q(q \in Q)$ を定義する。次に n に対して固有表現抽出を行い、 n が固有表現の場合、要素に固有表現タグ e を追加した新しい集合 $E = \{(n_r, i_r, e_r)\}$ を作成する。

Step2: クエリと類似度の高い固有表現の判別

Step2では、wikipediaの単語を固有表現タグに置換した上で学習したword2vecモデルを用いて全ての q に対して類似度の高い固有表現をそれぞれ上位 k 番目まで計算する。この時の固有表現タグを e'_{qk} 、類似度を s_{qk} とする。

Step3: 単語を補完するかどうかの決定

s_{qk} の最大値が閾値 θ を下回った場合、必要な固有表現がないと判断して単語を補完せずにStep3で終了となり、検索クエリは Q となる。 θ は実験的に $\theta = k \cdot 0.15$ と設定した。

Step4: 補完語と完全クエリの決定

s_{qk} が最も大きい時の e'_{qk} に対して $e'_{qk} = e_r$ となるものがないか調べる。存在した場合は i_r の最も大きい n_r を補完語とする。存在しなかった場合は2番目に s_{qk} の大きいもので同様の作業を行う。これをすべての e'_{qk} に対して繰り返して $e'_{qk} = e_r$ となるものが見つからなかった場合は n_1 が補完語となる。これにより得られた補完語を Q に加えたものが完全クエリとなる。

2.2.2. B-QC

B-QCでは以下の3stepで完全クエリを生成する。

Step1: マスク単語予測用の文章生成

BERTで単語予測する箇所を[MASK]と表す。

An Examination of Search Word Completion Methods from Interaction History in Incomplete Search Queries

[†]Ryoichi Shibata, Mitsuhiro Kimoto, Michita Imai
Faculty of Science and Technology, Keio University

[‡]Tepei Yoshino, Shoya Matsumori, Yosuke Fukuchi
Graduate School of Science Technology, Keio University

Queryの文頭に補完語が発生すると仮定して、格助詞をjとしたとき, Bert での入力文 Q' を $Q' = c_{1:t} + [MASK] + j + Query$ とする。ここで“+”は文字及び文の連結を表すものとする。

Step2 : jごとの補完語予測

Q' を入力として[MASK]に入る単語を上位 100 位まで予測する。ここでnが上位 100 位までに含まれたとき, nの予測スコアが最大となるnを, 助詞jを用いて予測した補完語候補 $w(j)$ と表す。

Step3 : 補完語と完全クエリの決定

すべてのjについて Step1, 2 を繰り返し得られた $w(j)$ の集合のうち最も出現頻度の高い $w(j)$ を補完語としてQに加えたものが完全クエリとなる。

3. 実験

E-QC と B-QC の特性と性能を評価するためユーザ実験を行った。

3.1. データセット

評価のために対話とクエリがセットになった二種類のデータセットを作成した。以下の旅行対話とトピック対話の2種類である。また、実験参加者 40 人を対象にクエリごとに正しいと思う検索クエリを記述させた。参加者による平均補完語数が1以上となるクエリに対して、参加者の半数以上が一致して補完した単語を正解として定義した。全33クエリ中30クエリで正解となる補完語が得られた。

<旅行対話>

2人ずつ3ペアの計6人の実験参加者にオンラインチャット上で北海道旅行についての計画を立ててもらい、その際に疑問が生じた場合はロボットに質問をするよう指示した。

<トピック対話>

『日本語日常会話コーパス』モニター公開版 [5]からトピックの異なる10個の対話を選出した。

3.2. 実験方法

以下の3種類の実験を行った。

(実験1) E-QC と B-QC それぞれの場合で SQuery の出力とデータセットの正解を比較した。

(実験2) 実験参加者 100 人を対象にデータセットを参照しながら E-QC, B-QC の生成クエリについて7を最高とする7段階評価で評価させた。評価項目は以下の2種類とした。

- ①検索ワードは質問の趣旨と対応しているか
- ②検索に必要な単語が十分に含まれているか

4. 実験結果

(実験1) 各システムの正解数は表1のようになった。E-QC の性質上, 固有表現抽出されない語句を補完することができないため, 固有表現が正解

表1 各システムの生成した補完語の評価

システム	正解数(全30クエリ)		正解率 [%]
	旅行対話	トピック対話	
E-QC	8	2	33.3
B-QC	8	4	40.0

表2 各システムによる生成クエリの評価結果

システム	項目	評価	
		旅行対話	トピック対話
E-QC	①	4.21	3.51
	②	4.01	3.32
B-QC	①	4.86	4.05
	②	4.73	3.90

となる問題が少ないトピック対話では正解率が低くなっている。また, 正解が固有表現であるクエリに限定したところ E-QC の正解率は 47.6%となった。E-QC と B-QC を組み合わせることでより正解率の向上が期待できる。

(実験2) 結果は表2のようになった。二種類のデータセットともに B-QC の方が評価が高くなっている。また, E-QC と B-QC ともに評価項目①の方が高くなっている。これは, 同じ固有表現であるが単語が異なるもの(例: 札幌と北海道など)による補完語の違いが多かったため, 検索に必要な単語は十分ではないが, 趣旨にはおおよそ沿っていると判断されたためだと考えられる。

5. 結論

本稿では, 対話を逐次的に処理しながら不完全クエリに対して対話履歴中から適切に単語を補完する手法である SQuery を提案した。実験では, 人同士の対話の中で発生した不完全クエリのデータセットを取得して不完全クエリの補完を行いユーザによって評価させた。結果, E-QC と B-QC それぞれの特性が明らかになり, 両モデルが不完全クエリの補完に一定の有効性があることが示された。

参考文献

- [1] 油井 et al. (2012) “質問応答システムにおける曖昧な質問文の補完のための対話処理。”平成24年度 情報処理学会 関西支部 支部大会 講演論文集
- [2] Jeffrey et al. (2014) “Entity query feature expansion using knowledge base links.” SIGIR2014
- [3] Jiang et al. (2020) “Spoken Document Retrieval Leveraging Bert-Based Modeling and Query Reformulation.” ICASSP 2020
- [4] Takimoto, et al. (2020) “SLAM-Inspired Simultaneous Contextualization and Interpreting for Incremental Conversation Sentences.” arXiv preprint arXiv:2005.14662 (2020).
- [5] 小磯 et al. 「『日本語日常会話コーパス』モニター版の設計・評価・予備的分析」『国立国語研究所論集』18, pp. 17-33, 2020. 1.