

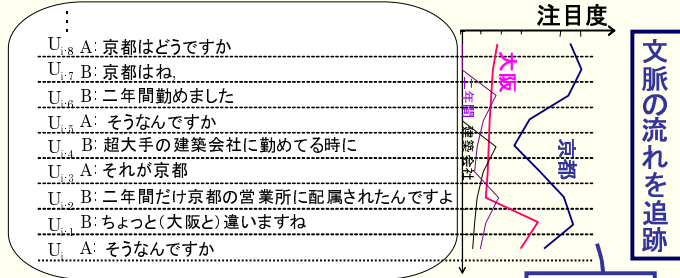
会話文脈の流れに応じた関連情報の自動提示システム

システムの目的 ユーザ間の**会話文脈の流れ**に沿った話題を自動提示

難しさ ユーザ達の**注目対象の移り変わり**を自動推定

産業展開例

- 会議支援: 会議中に**関連する過去の議事録**を提示
- コールセンター支援: 初心者オペレータに**関連ヒント**を提示
- 会話支援: 待ち時間に**会話のネタ振り**サービス
- 会話連動型広告: 文脈の流れにあわせて切り替わる**広告**
- 文章執筆支援: 文章文脈に**関連する情報**を提示



現在のユーザ文脈:
(京都:0.4, 大阪:0.3, 二年間:0.08)

関連情報:
...つまり極端な話、同じ関西人でも大阪人と京都人の会話では、「してんの?」&「したはりますのんか?」の違いがあるわけだから、どうしても京会話を使えば、かたつたるくなくなってしまうわけです...

(京都:0.4, 京会話:0.3, 大阪人:0.25, ...)

会話の流れに沿って
逐次的に**検索・更新**

プロトタイプシステムの開発

対象とした会話: 2者自由会話(書き起こしデータ)
提示情報の候補: Webからクロール(18,429ページ)

対象会話の例

B: 見て面白くても笑わないで我慢する
B: っていう人も多いらしいんですけど
B: 関西の人は受けてあげる
B: ちょっとでもおかしかったらわあわあ受けてあげてっていうお客さんのサービス精神があるらしくそれがね
B: だから演芸場とかじゃなくて普通の会話でも友達同士でもそうなんです
A: ちゃんと突っ込んでくれるとかそういうのは
B: そうそう突っ込んであげるとか突っ込んであげてっていうのは礼儀ですからね
A: 受けてあげる
A: おかしい
B: だってあたし東京に来たばかりの時寂しかったんですけど何か喋ってもみんな反応してくれない
B: 突っ込んで突っ込んでみないかね

今後の展開 **提示した関連情報の例**

- 会議支援への応用に向けて:
共同研究: 工学研究科 都市社会工学専攻 小林研究室
淀川コーパス(住民公聴会の議事録): 236ファイル, 44,703文
- コールセンター支援への応用に向けて:
協力: 株式会社アヴァンテ
まずはチャットベースのサポートサービスで開発・運用

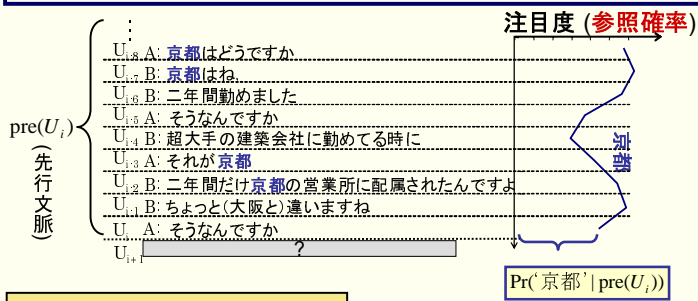
アイデア

文脈の流れを定量的に計算するために

単語へのユーザの**注目度**を自動推定
= その単語が**引き続いて言及されそう**な**割合**を予測
定義: 「**参照確率**」

$$\begin{aligned} (\text{発話 } U_i \text{ における実体 } e \text{ の注目度}) &= (\text{後続発話 } U_{i+1} \text{ で } e \text{ が参照される確率}) \\ &= \Pr(\exists w \xrightarrow{\text{ref}} e \text{ in } U_{i+1} \mid \text{pre}(U_i)) \\ &= \Pr(e \mid \text{pre}(U_i)) \end{aligned}$$

$\text{pre}(U_i)$: U_i までの先行文脈 $w \xrightarrow{\text{ref}} e$: 単語 w が実体 e を参照



計算の手がかり(入力)

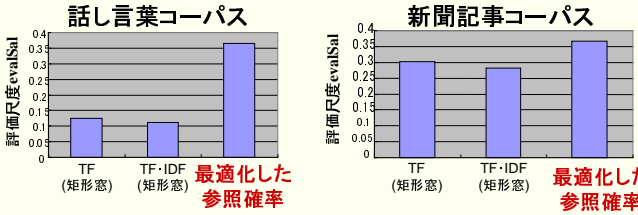
- ロジスティック回帰で以下の入力を統合して計算
- 先行文脈におけるその単語の**文法機能**(日本語なら助詞)
 - 品詞, **固有名詞の種類**(人名/地名/組織名/...)
 - 先行文脈でのその単語の**頻度**
 - 先行文脈でその単語が**出てきた箇所の近さ**

評価尺度(計算の最適化のため)

後続発話における言及の有無の予測能力で評価

$$\text{evalSal}(m) = \text{cor}([\text{sal}_m(e \mid \text{pre}(U_i))]_{(e,U_i)}, [\text{isRef}(e, U_{i+1})]_{(e,U_i)})$$

m : 計算方法 — 相関 — m による推定値 — 実際の言及の有無 (1 or 0)



特に話し言葉で評価尺度の向上幅が大きい
⇒ **特に話し言葉で有効な文脈計算手法を確立!**

文脈類似度の定義(検索のため)

各単語への注目度から成るベクトルの**コサイン類似度**
⇒ **会話文脈の流れに応じた関連情報の検索を可能に!**



白松 俊 <siramatu@kuis.kyoto-u.ac.jp>

京都大学 大学院情報学研究科
日本学術振興会 特別研究員 (DC2)