

複数ユーザ間の会話支援システムのための 文脈追跡手法の開発

白松 俊 (京都大学 情報学研究科 知能情報学専攻 奥乃研究室)

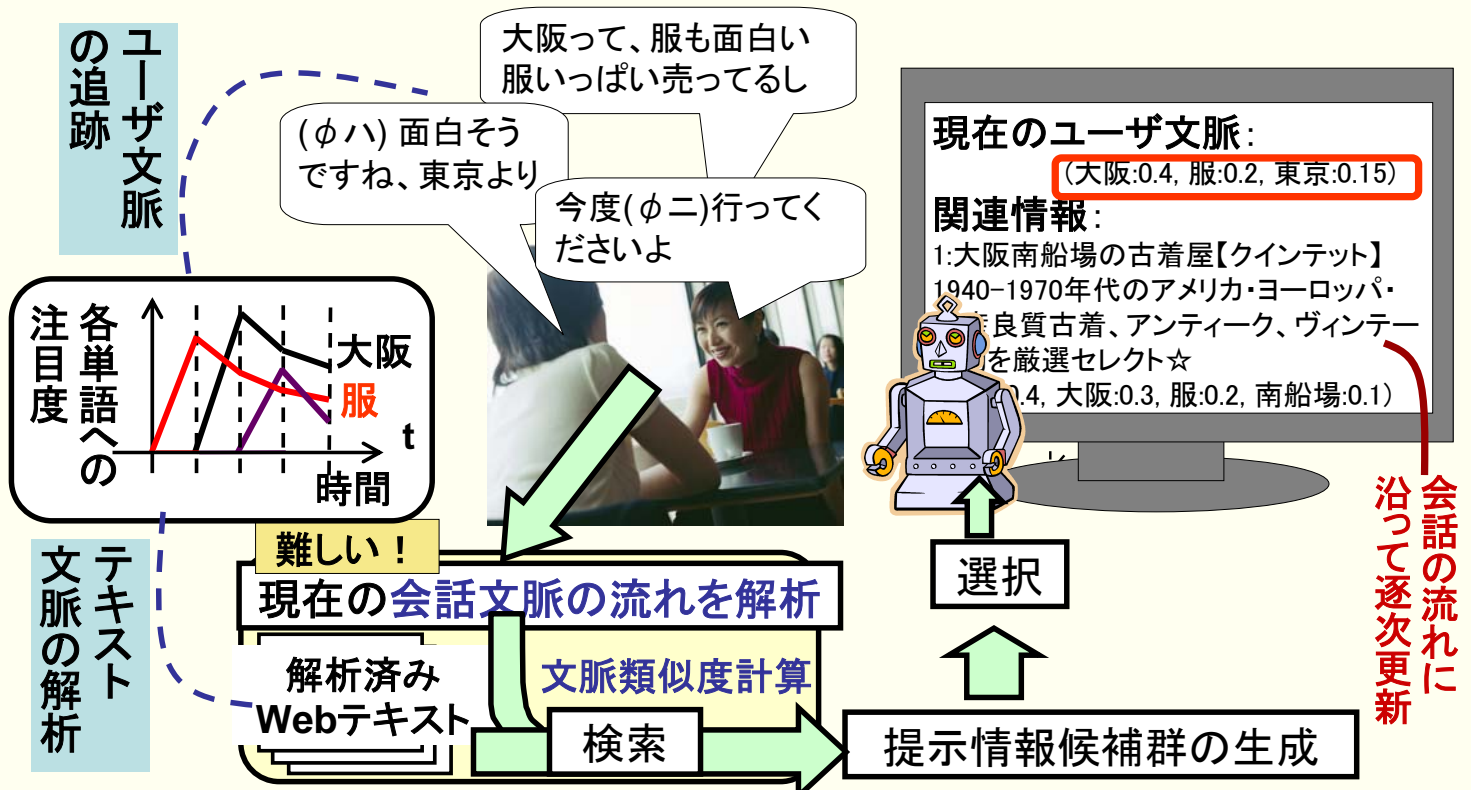
目標 会話支援のための話題提供サービス

会話の**文脈の流れ**に沿った話題を提示

↳ ユーザの**注目対象**の移り変わりを追跡


産業展開例

- 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
- 会議支援: 会議中に**関連する過去の議事録**を提示
- 文章執筆支援: 文章文脈に関連する情報を提示

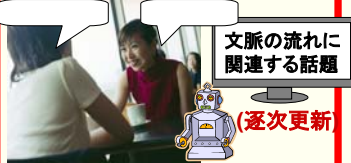


新しさ 目指す検索：
今までの検索とここが違う！

	従来の検索	話題提供に必要な検索
検索要求の時系列性	ユーザ文脈の流れ(注目対象の変遷)を追跡しない	ユーザ文脈の流れ(注目対象の変遷)を追跡
検索要求の指定方法	ユーザが明示的に検索語を指定	システムが自動的に検索語を推定(クエリーフリー)
検索対象	文書	文単位や段落(パッセージ)
利用形態	ユーザが検索インタフェースに向かって検索要求を入力	会話, チャット, 文章執筆時に, システムが「聞き耳を立てて」文脈を追跡し, 関連情報を検索



大阪 服 検索
検索結果



文脈の流れに関連する話題
(逐次更新)

技術的課題

文脈(注目対象の移り変わり)をどうやって追跡する？

単語への注目度の「ものさし」が必須！

アイデア 「注目されている単語ほど継続的に参照されやすい」

⇒ **参照確率** $\Pr(e | \text{pre}(U_i))$

文 U_i の先行文脈 $\text{pre}(U_i)$ が与えられた下で, 単語 e が次の文 U_{i+1} で参照される確率

	従来の「ものさし」			本研究
	語の頻度 (TF*IDF)	センタリング (Cf-ranking)	RAP (salience weighting)	参照確率
時系列的遷移	×	○	○	○
数値化	○	×	○	○
複数因子統合	×	△	○	○
統計的根拠	○	×	×	○

⇒ **文脈の類似度計算を可能に！**

成果

書き言葉の文脈(注目対象の推移)を文ごとに自動追跡するシステムを開発

⇒ 京大から**特許出願中**

プロジェクトの展開

話題提供サービス実現へ向けて

日本学術振興会の支援下で
プロジェクト推進予定

(平成19年4月～21年3月)

計画

- 話し言葉に特化した文脈追跡手法
- 話題提供に必要な検索の開発 (クエリーフリーな文脈連想パッセージ検索)
- ⇒ 2年後には実証実験を！