

R2-723

奥村・高村研究室

ことば × 計算機

自然言語処理 / 計算言語学

研究テーマ

奥村・高村研究室では、ことばを計算機で処理する技術に関する研究と、その技術を用いた応用システムの開発を行なっています。

統計的手法や機械学習（深層学習）を用いた自然言語処理

Webや新聞等の大量の言語資源に対し、統計的手法や機械学習（深層学習）を用いることで、自然言語の抽出・分析・加工などを行う手法を開発しています。

自動要約

- 入力された文や文章を文脈を保ちながら短く表現
- 複数文書を単一の文書で表現

機械翻訳

- 入力された文や文章を別の言語に変換

実況の生成

- 入力された試合状況に沿う実況を生成

自動対話

- 入力された発話に対し、自然な返答を生成

語学学習者支援

- 誤りを含む文を自動で訂正

意見分析・テキストマイニング

- Web上のテキストから社会の動向・意見を分析
- 文章から書き手の意見や感情を推定

自然言語の解析

- 構文解析, 意味解析, 対話分析など

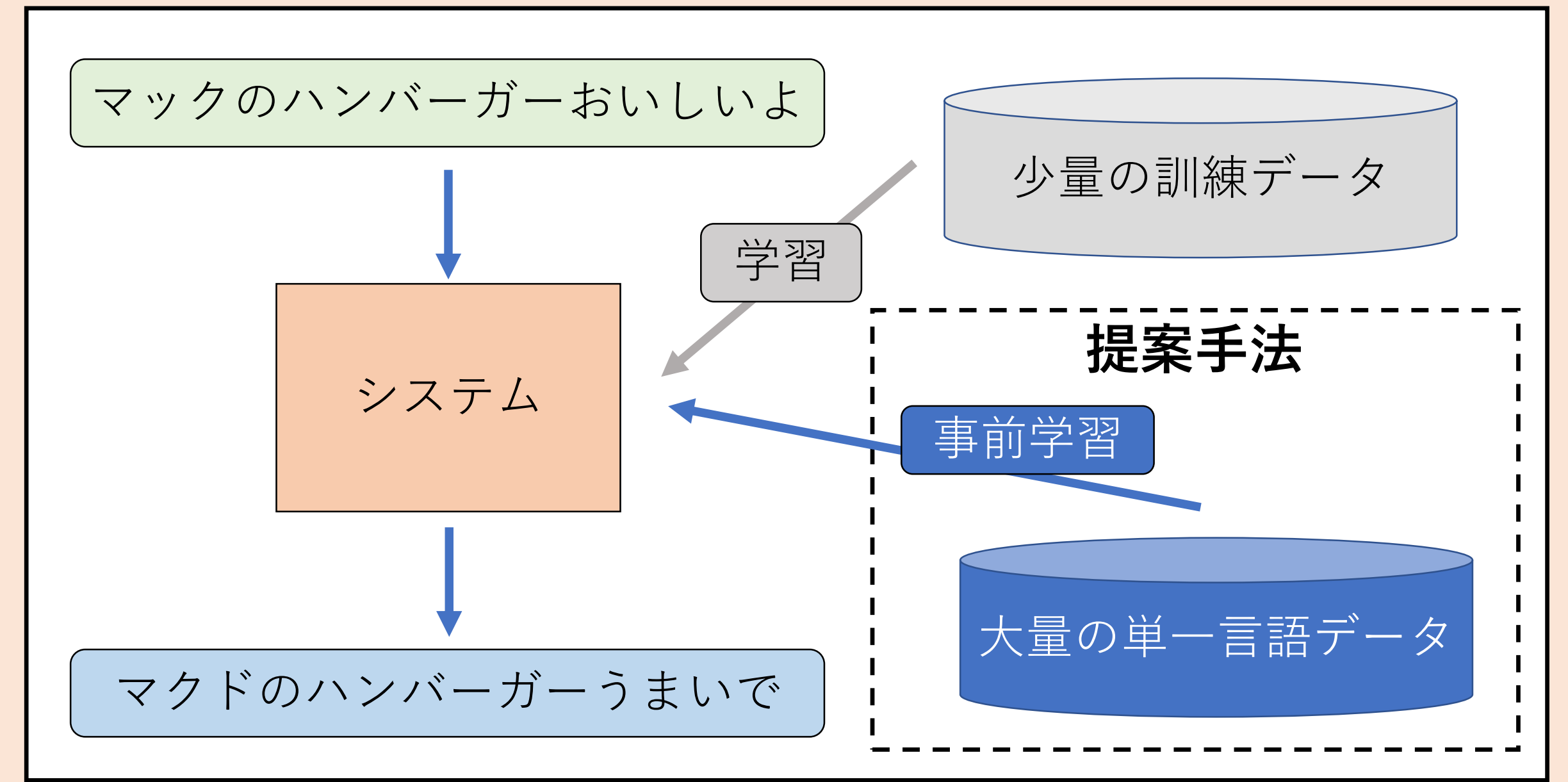
研究室の活動

月に一度の進捗発表会や毎週二回の論文輪読会があります。それらとは別に、機械学習とアルゴリズムの勉強会もあります。また、研究室メンバーと親睦を深めるために、年に一度の合宿があります。日常的なコアタイムはありません。

最近の研究成果

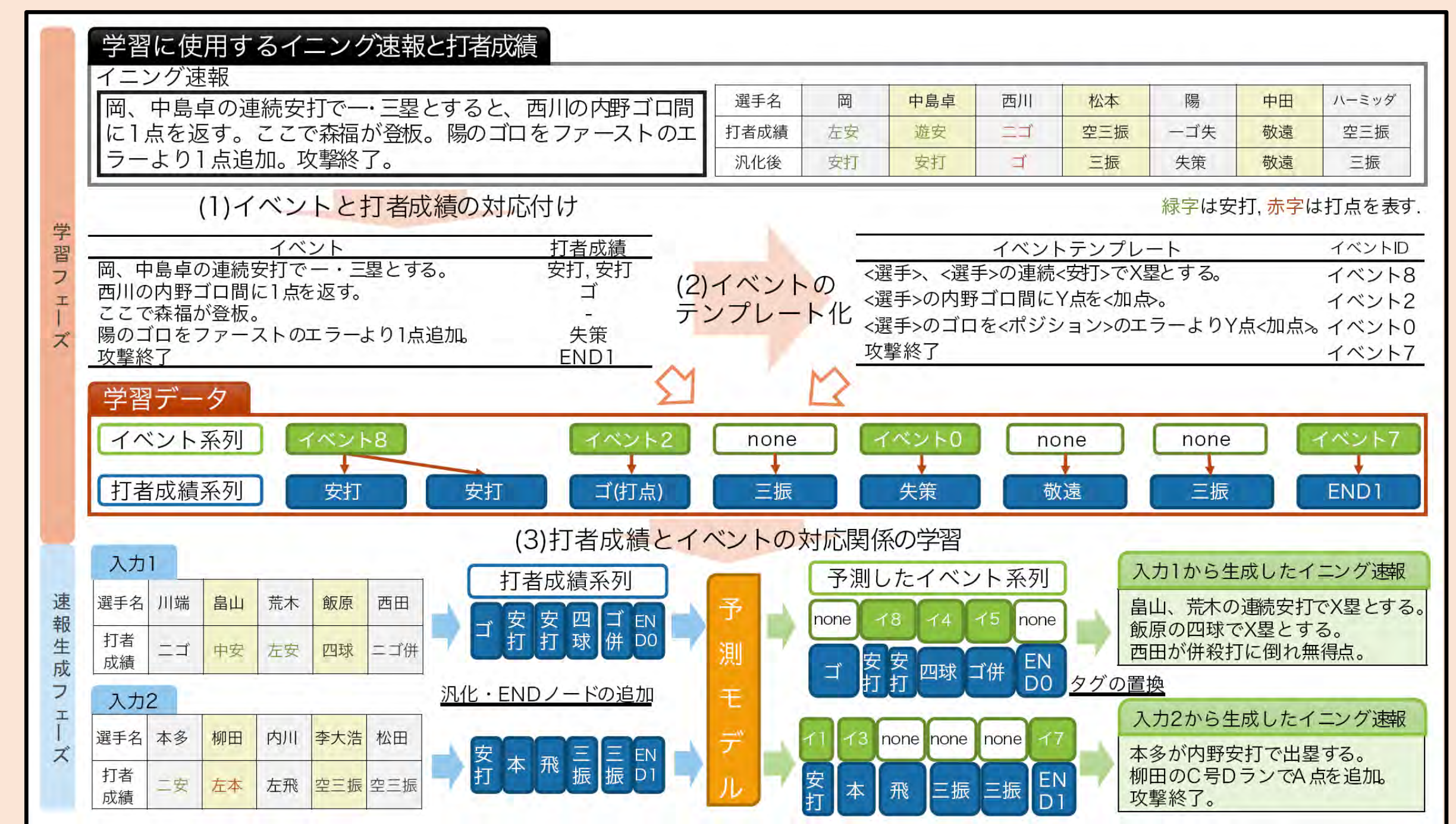
事前学習と汎化タグによる方言翻訳

標準語を方言に変換するシステムを、少量の訓練データから作成



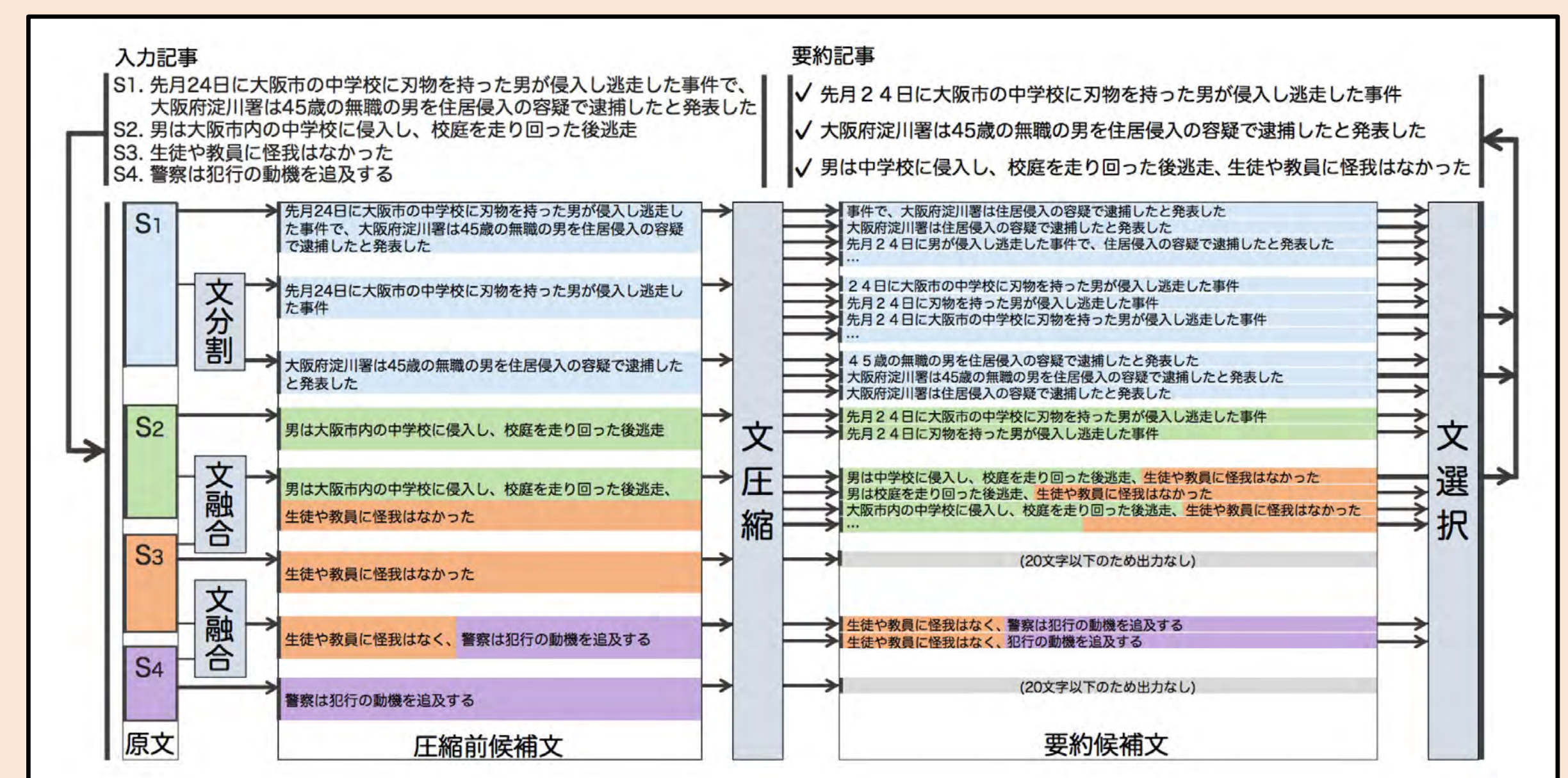
打者成績からのイニング速報の自動生成

打者成績から重要なイベントを予測し、テンプレートを用いてイニング速報を生成



要約長, 文長, 文数制約付きニュース記事要約

文融合, 文分割, 文圧縮を柔軟に用いて, 閲覧者にとって読みやすいニュース記事の要約文を生成



新カタカナ動詞の意味理解

Twitterなどのマイクロブログ中に含まれるカタカナ動詞の理解を, その単語の語源と言い換えを提示することで支援

