



モデリング・解析基盤整備 「データマイニング」

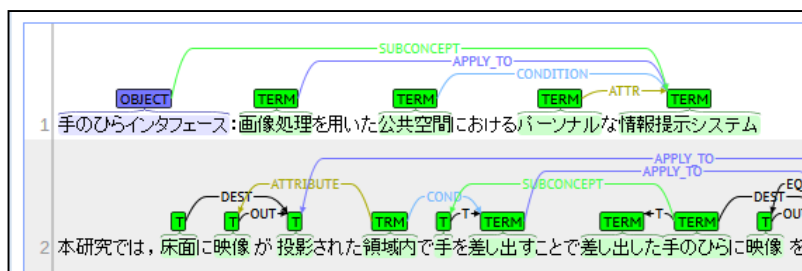
自然言語データ（新聞、論文、ウェブなど）の中に埋もれた「知識」を抽出する

「知識」とは？
→ **メタ知識構造**

因果関係：雨が降って → 地面が濡れている
目的：うどん屋に行つて → 昼ごはんを食べる
理由：電車が混んでいるので → 早起きする

メタ知識構造コーパスの開発

- コーパス：自然言語データに付加情報を人出で与えたデータ
- 機械学習のための学習データ・評価データとして必要
- 対象：学術論文のアブストラクト（日本語230件、英語300件）



メタ知識構造認識技術の開発、 かきこい検索システムへの応用

- 上記のデータと機械学習を利用し、メタ知識構造を自動認識
- 認識した構造を用いた論文検索
- 例：「機械学習を改良した研究」と「機械学習を利用した研究」を区別

波及効果

- 自然言語で書かれた「知識」が必要なあらゆる分野に応用可能
- 企業活動、ヘルスケア、研究開発、政策決定、etc.
- 情報分析が効率化され、人間はより創造的な活動にフォーカスできる
- 自然言語と知識の関係性は有史以来の難題であり、本研究はその一端を明らかにするもの

今後の課題

- 現状の解析精度では実用化は難しく、精度向上が必須
- 本研究の成果を有効活用できるインタフェースが必要

今後の発展

- 情報分析を超えて、新たな知識を創造する仕組みの探求