

言語情報処理基盤の確立と 言語・知識コミュニケーションへの発展

黒橋禎夫

情報理工学系研究科電子情報学専攻

概要

言語の仕組み、それをを用いたコミュニケーションの仕組みを計算機が扱える正確さで解明するという理論的研究と、それによって情報検索、機械翻訳、マンマシンインタフェース等をより高度なものとし、人間の活動を支援するという工学的研究を有機的に結合させながら展開している。本稿では本年度の主たる研究成果を報告する。

1 はじめに

豊富な計算機資源に支えられ、大規模なテキスト、画像、動画コンテンツが蓄積されつつある。これらは、時空間をこえた情報・知識の伝達、創造を可能にする潜在能力があるが、現状では、これらを人間が有効に活用することや、人間同士のコミュニケーションの中に導入していくことはきわめて困難である。むしろ、情報洪水の中で、必要とする情報にたどり着くことすら難しいという状況である。

このような問題の根本解決のためには、コンテンツの中核である言語情報を計算機により深く理解させること、すなわち、より知的な言語情報処理基盤の確立が必要である。また、人と人、人と計算機のコミュニケーションにおいては、言語が基盤的役割をはたすものの、それを円滑で自然なものにするために多くの要素が関与しており、それらを人間の認知的な視点からモデル化する必要がある。

我々は、このような問題意識のもとに、言語情報処理の基盤技術の確立と、実践的アプリケーションへの適用、さらに、コミュニケーションのモデル化と、それに基づくコミュニケーション支援システムの構築を行っている。

2 超大規模コーパスからの常識の獲得

計算機による言語処理を高度化するためには、計算機に常識・世界知識を与える必要がある。我々はこれまでに、新聞記事 20 年分、約 2 千万文を構文解析し、そこから確からしい述語項関係

のみを取り出して、漸進的にクラスタリングを行うことにより、各用言の各用法の格フレーム（述語項構造のパターン）を自動構築する方法を開発してきた。

本年度は、この枠組みのさらなる改良を進めるとともに、クラスター計算機を用いて web から収集した超大規模コーパス、約 5 億文を構文解析し、格フレーム化した。新聞コーパスでは政治・経済にドメインが片寄っており、構築される格フレームにも片寄りがあったが、web コーパスによって広範囲の言語現象を収集することができ、常識といえるレベルの格フレームの獲得が可能となった。例えば「クロールで泳ぐ」「望遠鏡で見る」などのパターンも学習され、これによって「クロールで泳ぐ彼女を見た」と「望遠鏡で泳ぐ彼女を見た」（言語と常識の関係を示す古典的例文）の構文解析なども正確に行うことが可能となった。

ここで構築した格フレームは、あらゆる言語処理の基盤となるものであり、構文解析の高度化だけでなく、複合名詞句の解析、格・省略解析、言い換えパターンの自動学習などに利用している。

3 テキストの自動プレゼンテーション

人間は会話を通して知識を交換・創造している。一方、我々が知識を整理し蓄積するのはテキストであり、「会話」という形式との間にギャップがある。このギャップを解消するために、テキストを話し言葉調の自動ナレーションに変換し、同時に、テキストの文章構造を抽出して自動要約スライドを作成し、これによって、テキストの自動プレゼンテーションを実現した。

書き言葉話し言葉変換については、技術的には、文法規則に基づく「だ・である調」から「です・ます調」への変換、複合名詞の格解析に基づく展開、web からの書き言葉コーパス、話し言葉コーパスの自動学習と、それに基づく書き言葉特有語彙の抽出、国語辞典と大規模格フレームに基づく用言句の平易化などを統合して用いている。これにより、例えば次のような言い換え処理が実現さ

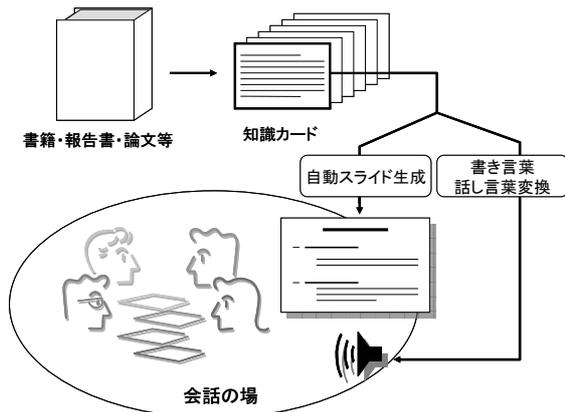


図1. テキストの自動プレゼンテーション

れている。

- (1) 液状化が生じず ⇒ 液状化しないで
- (2) 甚大な被害 ⇒ 大きな被害
- (3) 被害者の避難場所 ⇒ 被災した人の非難する場所
- (4) 災害発生後の疾病による死者 ⇒ 災害が発生した後の病気による死者
- (5) 精神科医による専門対応を行った ⇒ 精神科医による専門的な対応を行いました

自動要約スライド作成は、各文の主題を取り出し、さらに文間の主題連鎖、並列、対比、例示などの関係を解析して文章の談話構造を求め、これらの情報に基づいてタイトル、インデントなどを制御してスライドを生成する。

この技術を地震防災などの社会技術研究に適用し、人間の会話を通じた知識プロセスを支援する研究に発展させている。

4 言語情報と映像情報の統合的理解

計算機環境の進歩により大量の映像が蓄積されはじめたが、現在のところ映像のインデックス付けは手動であり、大量映像の有効利用は難しい状況にある。人にもものを教える映像、特に料理番組を対象として、言語情報（登場人物の発話やナレーション）と映像情報を統合的に解析し、どのような物体・動作が映っているかを検出し、これによって映像の自動インデックス付け、自動要約などを実現しようとしている。

言語情報は、現在のところ音声認識で獲得することは難しいことから、映像と同時に配信されているクローズドキャプションを用いている。クローズドキャプションでは、言い淀み、言い直しなどが若干修正されているものの、基本的には登場人物の発話がそのまま書きおこされている。

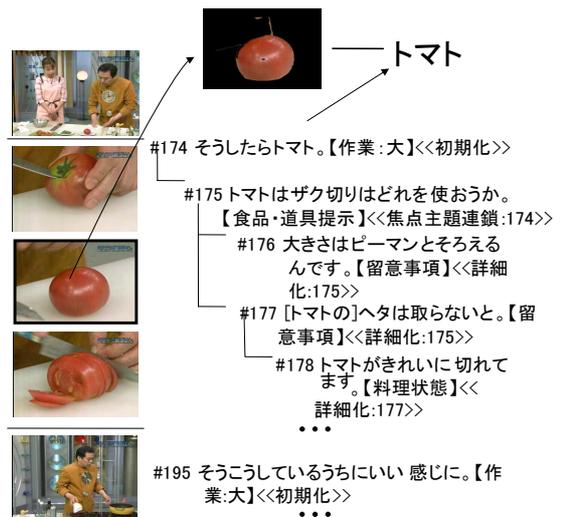


図2. 言語と映像からの物体モデルの学習

今年度は、言語と映像をリンクする取り掛かりとして、映像中の食材の認識を課題とした。映像中の物体の認識は、モデルの強い作り込みを行わなければならず、一般的にはきわめて難しい課題とされていた。一方、言語の解析からみると、話し言葉で多用される省略の復元が重要な課題であるが、現場の共有を前提とする発話中の省略を言語情報だけで解析することは困難であり、また不適切な問題設定であるともいえる。

このような問題を解決するために、映像から、物体を説明（導入）する発話と画像を抽出し、これによって物体の画像モデルと名前を同時に、教師無しで学習し、このモデルを用いて物体認識と省略解析を行う方法を考案した。物体の説明発話は、談話構造、発話タイプの解析結果等から絞りこみ、これと画像特徴量から抽出したアップ画像を大量に対応付けることにより、色、形状の学習を行った。今後は、この関係付けを、人の動作や、時々刻々変化する食材・料理状態の追跡などに発展させていく予定である。

5 確率的用例ベース翻訳の定式化と実現

利用可能な対訳コーパスの増大に伴い、用例ベース翻訳、統計ベース翻訳などのデータに基づく機械翻訳の研究が活発にすすめられている。

用例ベース翻訳の基本的なアイデアは、入力文の各部分に対して類似している用例を選択し、それらを組み合わせて翻訳を行うことである。このとき、入力文との一致が大きければ大きいほど良い用例である、すなわち、広い文脈を考慮するより正確な翻訳がえられると考えられる。ところが、これまでの用例翻訳では、この「良さ」を正確に

定式化しておらず、これが統計ベース翻訳に比べて不透明でアドホックであるという批判を受ける原因であり、また実際に翻訳システムを構築、最適化する際の障害であった。

この問題を解決するために、用例ベース翻訳の確率論的定式化を行った。すなわち、統計ベース翻訳のように語や句の一定の小さな単位だけを考えるのではなく、文全体までのあらゆるサイズを対象とし、それらが対訳コーパスの中でどのような確率で翻訳されるかを計算する。当然、大きなサイズの表現は安定した翻訳が与えられていて大きな翻訳確率を持つ。したがって、入力文に対して翻訳確率が高い用例への分割、翻訳を求めることで、自然と、用例のサイズを「良さ」とした用例選択が可能となる。

この方式を旅行対話翻訳、特許翻訳などで実験し、有効性を確認している。また、旅行対話翻訳では、国際的評価型ワークショップにおいて上位の成績をおさめており、特許翻訳においては、翻訳専門家が利用できる翻訳支援システムへの適用をすすめている。

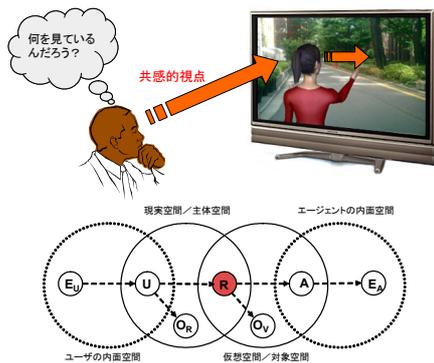


図3. 人間-コンピュータ間インタラクションにおけるユーザインボルブメント

6 人間-コンピュータ間インタラクションにおけるユーザインボルブメント

人間がコンピュータのディスプレイ内の仮想世界やロボットとのコミュニケーション、またコンピュータが媒介するコミュニティなどに自ら進んで関わったり、もしくは無意識的に没入していく認知的なあり方を「ユーザインボルブメント」として捉え、認知科学的な観点から理論化と検証を行った。特に、人間が他者の認知空間にアクセスする際、他者への「共感(empathy)」を通じて、相手の認知的視点を獲得し、その他者を認知主体として承認するという仮説を提出し、その「共感チャンネル」を人間-コンピュータ間インタラ

クション環境のデザインにおいてどのような形で実現するべきかを検討した(図3)。

7 振る舞いの冗長性を用いてジョイント・アテンションを確立する聞き手ロボット

カメラなどの人工物に対して説明を行う場面における人間の話し手の心理的負担の軽減を目的として、聞き手として自然な振る舞いをする聞き手ロボットの構築を行った。特に、この聞き手ロボットとのコミュニケーションを自然なものにするため、ジョイント・アテンションと呼ばれる現象に注目し、これを利用した。また、注目を表す振る舞いの意図性の認識・表現に、振る舞いの冗長性を利用することを提案した。

さらに、その枠組みをもとに、ジョイント・アテンション、振る舞いの冗長性の認識・表現、コミュニケーションのモードによる振る舞いの変化という3つの要素を提案した。その各要素をヒューマノイド・ロボットであるRobovieに実装し、家具の組み立ての説明におけるコミュニケーションを題材とした評価実験を行い、有効性を確認した(図4)。

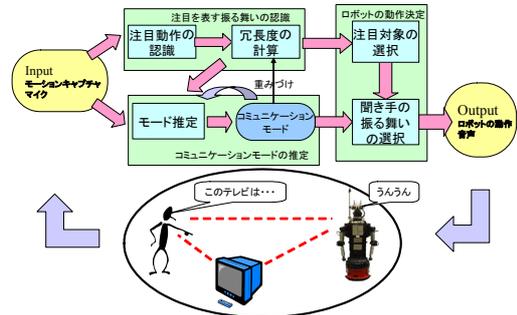


図4. 振る舞いの冗長性を用いてジョイント・アテンションを確立する聞き手ロボット

8 円滑なショット遷移を可能にする映像コンテンツ制作支援

視聴者の認知的な負担を軽減することで視聴者を引き込むような映像コンテンツを低コストで制作するための支援を目的とし、円滑なショット遷移を可能にするコンテンツ制作支援システムを構築した。

テレビ番組は、視聴者に見やすくわかりやすい映像コンテンツであるが、それは、テレビ番組など映像コンテンツを配信するプロのメディアに映像技法などのノウハウがあるからであると考えられる。そうしたノウハウを映像コンテンツ制作の過程に利用することを考える。そのために、

映像技法を検討し、視聴者の映像理解の観点から分類した。一方で、認知科学的な観点から、映像コンテンツに要求されるものについても検討を行い、その上で実際のテレビ番組を分析し、提案手法の有効性を実証した。最後に、仮説検証から得られる定型を映像コンテンツ制作支援システムに実装し評価実験を行った(図5)。

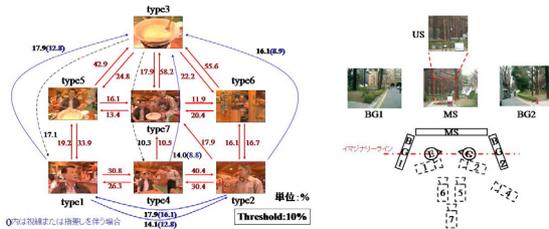


図5. 円滑なショット遷移を可能にする映像コンテンツ制作支援の概要

9 情報の空間化による知識編集環境

人間の知的活動を支えるシステムとその作業環境の双方を考慮した知識編集環境の方法論を提案した。特に、新たな情報を入手する<情報獲得>、新情報に適切な表示を与えて管理する<情報管理>、新情報と既存の情報を関連付ける<情報編集>という3つのフェーズが一つのサイクルとして知識形成プロセスとなっていることを主張し、その知識形成プロセスに沿って、どのような要件が知識編集システムやユーザサポートに求められるかを検討した。

具体的には、閉じた知識としてのマニュアル情報の空間整理実験と、人手によるカード配置実験の分析結果を元に、情報配置の傾向に合わせた情報配置支援、ユーザの主観評価による重要度の導入、時間的アクセスの支援、3次元ユニットによる情報表現支援など、それぞれのフェーズにおける知識編集支援の提案を行った(図6)。

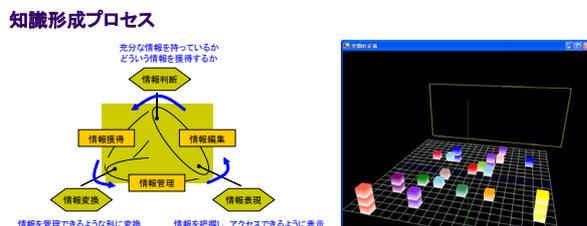


図6. 情報の空間化による知識編集環境

研究発表

- [1] Nobuhiro Kaji, Masashi Okamoto, Sadao Kurohashi: Paraphrasing Predicates from Written Language to Spoken Language using the Web, In Proceedings of HLT2003, pp.241-248, (2004.5).
- [2] Daisuke Kawahara, Ryohei Sasano, Sadao Kurohashi: Toward Text Understanding: Integrating Relevance-tagged Corpus and Automatically Constructed Case Frames, In Proceedings of The 4th LREC, pp.1833-1836, (2004.5).
- [3] 小笠原嘉靖, 田島敬士, 畠山誠, 西田豊明: 引き込み現象に基づく人間とロボットの暗黙情報のコミュニケーション, 人工知能学会, 第18回全国大会, 2B3-05, (2004.6).
- [4] 岡本和憲, 中野有紀子, 西田豊明: 台詞に基づく会話エージェントのジェスチャ自動生成, 人工知能学会, 第18回全国大会, 1E1-02, (2004.6).
- [5] 許幸源, 西田豊明: 相互適応的な知識外化手法を用いた知識編集支援, 人工知能学会, 第18回全国大会, 2E1-05, (2004.6).
- [6] 河原大輔, 黒橋禎夫: 自動構築した格フレーム辞書と先行詞の位置選好順序を用いた省略解析, 自然言語処理, Vol.11, No.3, pp.3-19, (2004.7).
- [7] Daisuke Kawahara and Sadao Kurohashi: Improving Japanese Zero Pronoun Resolution by Global Word Sense Disambiguation, In Proceedings of the 20th COLING, pp.343-349, (2004.8).
- [8] Ryohei Sasano, Daisuke Kawahara and Sadao Kurohashi: Automatic Construction of Nominal Case Frames and its Application to Indirect Anaphora Resolution, In Proceedings of the 20th COLING, pp.1201-1207, (2004.8).
- [9] Masashi Okamoto, Yukiko I. Nakano and Toyoaki Nishida: Toward Enhancing User Involvement via Empathy Channel in Human-Computer Interface Design, In Proceedings of International Workshop on Intelligent Media Technology for Communicative Intelligence (IMTCI 2004), pp.129-132, Warsaw, Poland, (2004.9).
- [10] Tomohide Shibata, Masato Tachiki, Daisuke Kawahara, Masashi Okamoto, Sadao Kurohashi, and Toyoaki Nishida: Structural Analysis of Instruction Utterances using Linguistic and Visual Information, In Proceedings of KES2004, pp.393-400, Wellington, New Zealand, (2004.9).
- [11] 鍛冶伸裕, 岡本雅史, 黒橋禎夫: WWWを用いた書き言葉特有語彙から話し言葉語彙への用言の言い換え, 自然言語処理, Vol.11, No.5, pp.19-38, (2004.10).
- [12] 黒橋禎夫, 大泉敏貴, 柴田知秀, 鍛冶伸裕, 河原大輔, 岡本雅史, 西田豊明: 会話型知識プロセスのための言語情報のメディア変換, 社会技術研究論文集, Vol.2, pp.173-180, (2004.11).
- [13] Masashi Okamoto, Yoshiyasu Ogasawara, Yukiko I. Nakano and Toyoaki Nishida: Enhancing User Involvement via Joint Attention with a Listener Robot, In Proceedings of Social Intelligence Design 2005 (SID 2005), (to appear), Stanford, USA, (2005.3).