

平成16年度 情報工学コース卒業研究報告要旨

吉川 研究室	氏 名	小野 貴博
卒業研究題目	会話マイニング —言語解析に基づく音声データからの情報抽出—	
<p>会話は最も重要な情報交換の手段の一つであり，多くの情報が会話を通して伝達される．会話を構成する発話は情報を話し手から聞き手に伝える媒体であり，そこには話し手が持つ知識が現れる．話し手から聞き手に伝わった情報は記憶として蓄えられるが，人間の記憶は時間の経過に従って曖昧になる．また，会話による情報交換は話し手と聞き手との間で行われるため，会話に参加していない第三者とは知識の共有が行われな い．話し手の知識を会話から抽出し，蓄積することができれば，過去の会話に出現した知識の検索や閲覧が可能となり，聞き手の記憶の補助や第三者との知識の共有が実現できると考えられる．</p> <p>本論文では，対話データから話し手の知識を抽出する手法を提案する．本手法では，各発話から抽出された情報を統合することにより，対話全体からの情報抽出を実現する．本研究で扱う対話には，情報を含む発話と含まない発話が存在する．これらを区別するために，発話に付与された発話意図タグを使用する．次に，係り受け構造に基づく規則を用いて，発話から情報を抽出する．係り受け構造を用いることによって，文節間の文法的なつながりに基づく情報抽出が可能であり，表層表現だけでは対応できない複雑な構造からの情報抽出が期待できる．最後に，各発話から抽出された情報を統合する．各発話はそれまでの文脈に依存すると考えられるため，統合は近隣の発話から抽出された情報間で，矛盾のない場合に限り実行するものとする．</p> <p>本手法を用いて対話データからの情報抽出実験を行った．対話データには，名古屋大学 CIAIR 車内音声対話コーパスのレストラン検索をタスクとする 200 対話 (2,096 発話) を用いた．対話の各発話には，発話意図タグがあらかじめ付与されている．係り受け解析には，話し言葉特有の言語現象に対応した解析ツール (当研究室で開発) を使用した．本実験では，これらのレストラン検索対話から話し手の知識として「店名」，「ジャンル」，「駐車場」，「予約」，「予算」に関する情報を抽出する．人手で正解を与えたところ，抽出すべき情報は 1223 個存在した．実験の結果，抽出性能は精度で 88.6% (714/806)，再現率で 58.4% (714/1223) であり，本手法による対話データからの情報抽出の実現可能性を確認した．</p>		