

平成16年度 情報工学コース卒業研究報告要旨

石井(健)研究室	氏名	藤田幸久
卒業研究題目	テキスト対話戦略のための対話データ判別	
<p>近年、日本語による対話システムが注目を集めている。対話システムには、限られた話題の中である目標を達成するタスク達成型と、いわゆる「雑談」を行なう非タスク達成型がある。タスク達成型では自然な対話が実現され、天気予報や切符の予約などに応用されている。一方、非タスク達成型では自然な対話は実現されていない。その一因として、人間の発話の種類(挨拶、疑問、主張など)を考慮して発話内容を生成していないことが挙げられる。発話の種類を無視した対話では文脈がおかしくなり、対話として成り立たなくなる可能性が高い。</p> <p>本論文では非タスク達成型の日本語対話システムにおいて、一対の発話から人間の発話の種類を判別する手法を提案する。これによって、日本語対話システムにおける発話生成の手間を軽減し、システムによる発話の精度が高まると期待される。本手法は人間と対話システムが対一でテキストによる対話を行なっているものとし、対話中の発話を分類する NTOC-DAMSL(Non Task-Oriented Conversation Dialog Act Markup in Several Layers) タグを定義する。そして、対話例を学習し人間の発話を判別、タグを付与する。</p> <p>本研究ではまず最近傍法を用いてデータ判別を行なった。実験の結果、95.3%のデータが正しく判別されたが、一意に判別されたデータは55.9%であった。そこでSVMを用いてデータ判別を行なった。その結果、86.3%のデータが一意に正しく判別され、SVMによるデータ判別が有効であることが確認された。次にSVMを用いたデータ判別の精度を確認する実験を行なった。この実験では判別結果の曖昧さを排除するため、判別結果の信頼度が閾値以上のものを正解とした。その結果、信頼度が70%の時83.2%が一意に正しく判別され、精度が十分に高いことが確認された。</p>		