

自然な発話に対する頑健なパーズングアーキテクチャ

3 B-8

小窪 浩明¹, 匂坂 芳典¹, 鈴木 紀子², 岡田 美智男²¹ATR 音声翻訳通信研究所, ²ATR 知能映像通信研究所

1. はじめに

コンピュータとの自然なインタラクションを実現する試みとして、音声による対話インタフェースの研究が進められている。音声対話とは本来、思いついたことをそのまま伝え、相手とのインタラクションを通して意図の伝達をはかる手段である。しかしながら、これまでの多くの対話システムでは、利用者はシステムの想定した対話モデルや文法に合わせ、あらかじめ伝えたい内容を整理してから話すことを強いられてきた。コンピュータに向かい、思いついたままを自由に話すことが出来るようになるまでは、音声による自然なインタフェースのメリットが十分に活かされることはない。

我々は、マルチエージェントの創発的な振る舞い、多重ゴールの扱い、動的な環境に対する親和性、頑健性などに着目し、マルチエージェントによって自然な発話の多様な挙動を捕えることを試みている。

本稿では、マルチエージェントによる動的な行動選択モデルに基づく、自然な発話に対する頑健なパーズングアーキテクチャを提案する。

2. 自然な発話に対する頑健なパーザ

自然な発話に対する頑健なパーザに求められる条件は、多様な言語現象に対応する能力を持つだけでなく、たとえシステムが解析を失敗した場合であっても、何らかの処理結果を返せる能力を備えていることである。システムがあらかじめ想定した範囲では扱えない文を不適格文と呼ぶ。自然な発話の中には、言い淀みや言い直し、省略など、不適格文に相当する文が多く存在している。

これまでに自然な発話に対する頑健な言語処理のアプローチとして、不適格文を処理するために緩和しなければならない制約をいかに発見し処理するかの研究が試みられてきた[1]。たとえば、全体の解析に失敗した場合に解析途中で得られた句や節の断片から意味の抽出を試みる部分解析法や、解析に失

敗した場合にのみ制約を緩和していく制約緩和法などがよく知られている。また、タスクを限定した手法としては、文中の断片をあらかじめ用意したテンプレートのスロットにマッチングさせる意味主導的なアプローチなどがとられてきた。これらのアプローチでは、制約を緩和することによって生じる解釈の曖昧さに対するヒューリスティックなルールやテンプレートなどを予め用意しておかなければならない。このように従来の自然言語解析のアプローチは、外からの手続きとして、事前に用意された戦略に基づき発話を解析する立場をとっていた。こうしたアプローチでは、多様に変化する自然な発話に対して、柔軟な制御をおこなうことは難しい。

音声による意思の伝達は、対話によって生じる互いのインタラクションの中から生まれてくるものである。我々は、音声言語の解釈は、外から手続き的にあたえられた戦略に基づくものではなく、対話の過程で作り出される環境によって、解釈そのものがボトムアップに組織化されるものという立場から捉える。このような視点に立ってパーザを考えたとき、パーズングにおける解釈の曖昧さは、発話の過程によってだんだんと明確化されていく相手の発話意図に対して、漸次的に解消されていく。また、その解釈の方略は、事前のプランに縛られず、むしろ対話の状況に応じて動的に立ち現われるものと捉えることができる。

3. 自律エージェントのパーザへの適用

近年、動的な環境に対する問題解決の新たなアプローチとして、環境に埋め込まれたプランニング(Situated Agent Planning)の研究がある。たとえば、Situated Agent の 1 つの実現例として、Maes [4]は ANA(Agent Network Architecture)を提案している。この ANA は、局所的な目的をもった自律エージェントが、お互いに協調競合し、外部の環境とのインタラクションの中で、全体としてある合目的な振る舞いをするネットワークアーキテクチャである。ANAにおいて、個々のエージェントは活性レベル(エネルギー)を持ち、他のエージェントや環境、ゴールとのインタラクションによってエネルギーを受けとるこ

とで、自らのエネルギーレベルを上げていき、活性化していく。エージェントは活性化することで、環境に対して影響を与え、ネットワーク全体の振舞いが決定される。

我々は、この自律エージェントモデル(ANA)における動的環境に対するアプローチを、「実世界」の一つである「自然な発話」に対するパーズングアーキテクチャに応用した。この概念図を図1に示す。自律エージェントに対する局所的な制約としてHPSG(Head-driven Phrase Structure Grammar)を採用した。HPSGにおいて主辞(head)は空でない下位範疇化素性を持っている。主辞は下位範疇化素性の値に合致する補足語と単一化することで、新たな部分解析木を生成することができる。この主辞の自律的な振る舞いをエージェントと定義すると、単一化による新たな部分木の生成はエージェントにとってのアクションとなる。また、単一化に必要な条件として、補足語を環境として考える。

パーザに単語列が入力すると、入力単語のもつ素性に応じて、環境とエージェントが次々に生成される。このとき、エージェントが単一化し、活性化するための条件は、下位範疇化素性の値に合致する補足語が環境に存在することである。このような補足語が環境に存在していれば、エージェントは環境の一つとして存在するその補足語からエネルギーを受けとることによって活性化し、アクションを起こすことができる。単一化によって生じた部分解析木もまたその素性に応じて、新たな環境やエージェントを生成する。下位範疇化素性の値に合致する補足語が環境に存在しないエージェントは、単一化によってこの補足語を生成する可能性のあるエージェントをサポートすることで、活性化に必要な条件を環境に揃える。このように、環境やエージェント間のイ

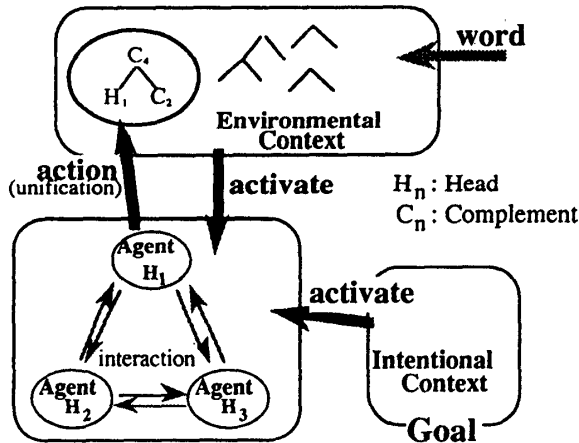


図1 自律エージェントを用いたパーズング

ンタラクションによって次々に単一化に必要な環境が満足され、パーズングが進んでいく。

このアーキテクチャは、環境とのインタラクションによって、次の動作が決定されるという特徴をもつ。我々は、このパーザを“Situated Parser”と呼んでいる。

図2にシステム構成を示す。音声認識部は、入力音声からワードグラフを生成する。“Situated Parser”はこのワードグラフを漸次的に入力し、動作する。ユニファイアは単一化文法に基づいて単一化操作をおこなう。ANAはゴールや環境とのインタラクションの重みによってシステムの変えることができる。ゴールからのインタラクションが大きい(goal-oriented)ほど、システムの動作はプラン依存性が強くなり、環境の変化に対して影響されにくいシステムとなる。逆に、環境とのインタラクションが大きい(situation-oriented)ほど、状況依存性が強く、環境の変化に柔軟に対応するシステムになる。

4. むすび

自然な発話に対する頑健なパーザとして、“Situated Parser”を提案した。今後は、多様な音声言語の挙動におけるパーザの有効性を明らかにしていきたい。

参考文献

- [1]河原, 松本: “音声言語処理における頑強性”, 情報処理 vol.36, No.11(95.11)
- [2]Henderson: “A Connectionist Parser for Structure Unification Grammar”, Proc. of ACL (1992)
- [3]岡田: 『口ごもるコンピュータ』, 共立出版(1995)
- [4]Maes: “Situated Agents Can Have Goals”, Robotics and Autonomous System 6 (1990)

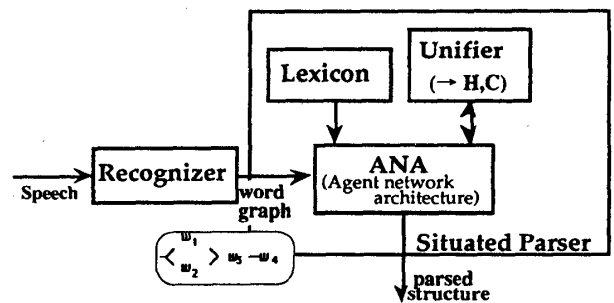


図2 システム構成