

文頭からの意味的区切りの抽出による意味表現の生成

4G-2

中川 真也 荒木 雅弘 堂下 修司

京都大学 工学部 情報工学教室

1 はじめに

従来の音声対話システムでは、発話文の区切りは明示されるという仮定をおき、発話文全体の音声信号が得られたものとして音声認識や意味解析を行なうものがほとんどである。しかし、実際の対話状況では、断片的に発話を続ける、発話を途中で止めて言い淀む、発話内容を考えて中断する、複数文をきちんと切らずに発話する、などの現象が見られ、発話文の区切りは必ずしも明確には現れない。また、実際の対話進行において、我々は発話途中の文に対しても、自然に相槌を打ったり、うなずいたり、相手が言い淀んでいれば適切な応答を返すなどして対話を進めており、発話文全体を理解の対象としているわけではない。

これらのことから、我々は発話の中に意味的なまとまりを見つけ、それを単位として段階的に理解を行なっていると考えられる。本稿では、このような人間の理解過程をモデル化するために、発話文の区切りを仮定した解析ではなく、発話の解析を発声単語順に順次進め、対話状況に見合う解析結果が得られた時点の意味的な区切りとして、意味表現を出力する発話解析システムについて述べる。

2 プラン認識を融合した発話解析

発話の意味的な区切りを認識するには、発話を文レベルではなく対話レベルで解析できなければならない。そのために、発話の背景にある話者のプランを認識し、現発話と話者のプランとを対応付けた解析が必要となる。

プラン認識に関して多くの研究がなされてきたが、その多くは観察された発話の意味や観察された行為を入力としている。しかし、ここでのプラン認識は発話の意味的な区切りを検出し、発話の意味を決定するためのものであり、プラン認識の入力として、発話の意味を仮定することはできない。ここでのプラン認識は、発話途中の文に対しても発話の断片的な情報を利用してプラン認識を進め、発話の部分と発話者のプランとの関係を抽出できる必要がある。

対話理解のためのプラン認識

Kautz は、イベントの抽象化とサブイベントへの分割概念に基づく階層性を表現したイベント階層を規定し、イベント階層上でプランを認識する手法を提案している [1]。これを対話理解に適用するなら、対話のある目的を達成するために協調的になされる一連の行為と捉え、発話によってなされる行為をイベントに対応させることになる。

しかし、発話とそれが表す行為は一対一に対応するとは限らず、いくつかの発話によって達成される行為や、一つの発話が複数の行為を表すと考えられる場合がある。また、断片的な発話も行為として捉えられる必要がある。

そこで、発話が表す行為を句レベルまで細分し、また、システムとユーザ双方の一連の行為からなる発話対も考える。我々が対象としている、個人のスケジュール管理というタスクに対しては図1のようなイベント階層を考える。

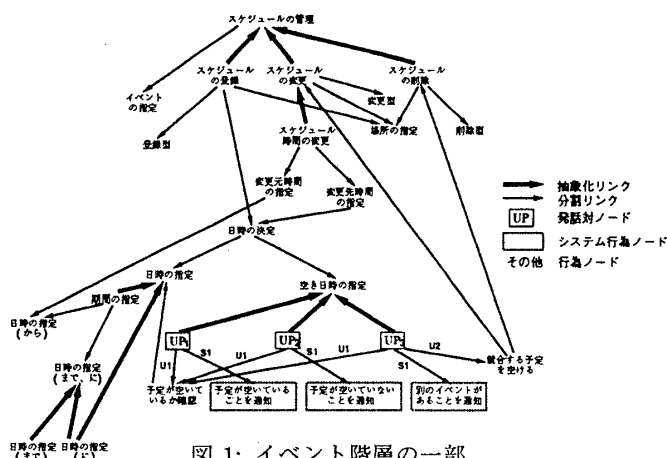


図1: イベント階層の一部

3 イベント階層上でのメッセージ伝達を利用した発話解析

Kautz[1] はイベント階層を一階述語論理で記述し、イベント階層上でのプラン認識手法としては、サーカムスクリプションにより各観察イベントを説明する被覆を求め、全ての観察イベントを最も簡単に説明するモデルである最小被覆をプランとして認識する手法を提示している。しかし、この方法ではかなりの計算量が必要となる。

これに対し、Lin[2] は同様な知識源に対し、メッセージ伝達の枠組でプランを認識する手法を提案しており、本研究では基本的に後者を採用している。

3.1 メッセージ伝達による発話解析

イベント階層上での発話解析は次の3ステップからなる。

Step1 音声認識部から Left-to-right で句単位に得られる解析結果をもとにメッセージ伝達を行ない、前発話までのプランと現発話とをうまく説明するプランを認識する。

Step2 認識されたプランから発話の各部分に対応する行為ノードを同定する。

Step3 行為ノードに定義された意味テンプレートに属性変数の値をバインドして意味表現を出力する。

処理の流れ、及び対話システム全体の構成を図2に示す。

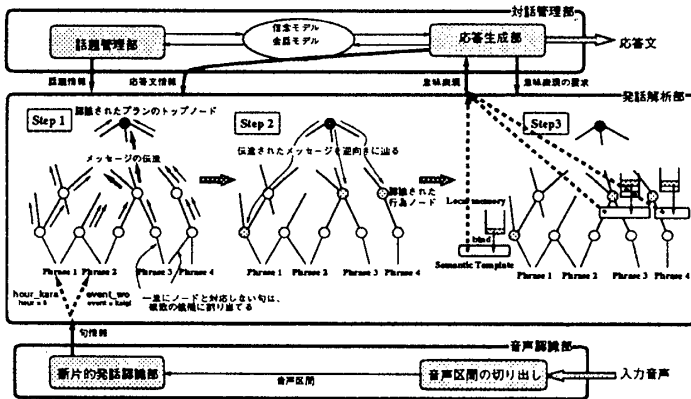


図2: 対話システムと発話解析の流れ

3.2 メッセージ伝達

各ノードには、ノードを規定する属性変数を割り当て、各ノードから伝達されるメッセージは、この属性変数と、どの句情報で起動されたかななどの情報を持つ。

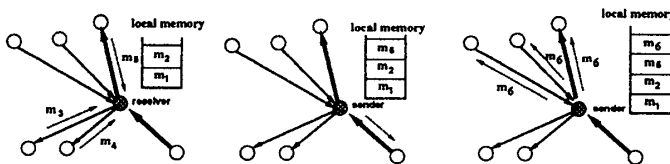
メッセージは図3に示す規則に従って伝達され、矛盾しないメッセージ同士を結合していくことで発話の部分に対するプランが認識されていく。また、行為ノードへの分割は、認識されたプランのトップノードからメッセージを逆向きに辿ることで行なう。

また、適当なプランが得られていない場合でも、長いポーズが検出され対話管理部から現時点までの意味表現が要求された場合には、前発話までで確定しているプランのトップノードから行為ノードへの分割が行なわれる。

(a) 初期メッセージ伝達 (b) 抽象化リンク方向でメッセージを受信した場合



(c) 分割リンク、抽象化リンクの逆向きでメッセージを受信した場合 (d) (c)でさらに、 m_1 と m_5 を結合し m_6 が生成可能な場合



(e) 発話対ノードにおけるメッセージ伝達

(e-1). システム-ユーザー型発話対 (e-2). ユーザー-システム型発話対

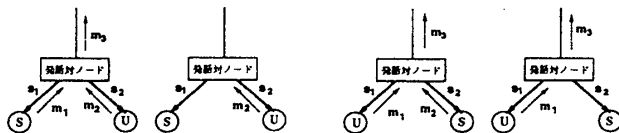


図3: メッセージ伝達の制御

3.3 行為ノードにおける制約の検査

メッセージ伝達の各過程において各行為ノードで満たすべき制約として以下を定義する。

- ノードに定義された属性変数間の関係に関する制約
- ノードに定義された属性変数が直属の分割行為の属性変数からどのように決定されるかに関する制約

あるノードに達したメッセージが、これらの制約を満たさなければ、そのメッセージの他への伝達は抑制される。このような制約をメッセージ伝達の各過程でチェックすることで、各ノード単位での独立な処理で全体としての一貫性が保たれる。また、前発話までに認識されたプランに対して、それに応じた属性変数が各ノードにバインドされているため、対話履歴に関する情報も制約として利用される。

4 解析例

例として、

U1: 明日の2時から4時まで2研で勉強会を行ないたい。

S1: 明日は3時から5時までミーティングがあります。

という対話を考える。U1は「スケジュールの登録」、S1は「別のイベントがあることを通知」と認識される。

この後に、ユーザが

- U2₁: では、ミーティングを3時から6時に、[して下さい。]
- U2₂: では、勉強会を6時から8時に、[して下さい。]
- U2₃: では、1時から3時に、[して下さい。]

という3通りの発話をする場合をみる。この場合、2つのイベント（ミーティング、勉強会）の時間に関する制約から、U2₁に対しては、「UP₃」のプランが認識されて、U2₁が「スケジュール時間の変更」に同定され、U2₂、U2₃に対しては、「スケジュールの登録」のプランが認識されて、「スケジュール（勉強会）の登録」に同定される。

5 おわりに

本稿では、発話途中の文に対して、メッセージ伝達の枠組でプラン認識を進め、前発話までのプランと関連した妥当なプランが認識された時点、意味的な区切りとして意味表現を出力する発話解析システムについて述べた。

今後は、メッセージ伝達のどの時点で意味的な区切りが抽出できたとするかの基準として、対話状況に応じて動的に変化する意味的な必須格を検討する。また、発話を文頭から解析する上で問題となる言い直しなどによる無用語の取り扱いについても検討する予定である。

参考文献

[1] H Kautz. A Circumscriptive Theory of Plan Recognition. In P.R. Cohen, J.L. Morgan, and M.E. Pollack, editors, Intentions in Communication. MIT Press, 1990

[2] Lin, D. and Goebel, R.: A message passing algorithm for plan recognition, in Proc. IJCAI-91, pp.280-285, Sydney, Australia (1991).