

会話音声認識における探索へのネットワークによる意味制約の利用

1E-7

額賀 信尾 荒木 雅弘 河原 達也 西田 豊明 堂下 修司

京都大学 工学部 情報工学教室

1 はじめに

従来の会話音声認識手法は、多レベルの知識を利用するために音素ラティスや単語ラティスのような中間表現を介してボトムアップに文仮説を生成するのが一般的であった。しかし、このような方法では、各処理レベル間の情報損失が大きいことが知られている。そこで、我々はHMMトレリス空間(時間×HMMの状態)上の探索によって、中間表現を経ない形での最適な文仮説の生成方法を提案している [1][2]。

音声言語の探索においては、その探索空間が膨大となるために、効率良く文仮説候補を生成できるアルゴリズムが求められる。探索における部分文仮説の評価値、すなわちその時点までの認識スコアは、音韻情報、語彙情報などにより算出される。しかし、スコア計算に用いる制約がゆるいために、結果として意味的に妥当でない候補が生成される可能性がある。この問題に対処するために、探索途上での部分文仮説候補の意味解析を行ない、早期に意味的に妥当でない候補を棄却し、探索空間を削減するという方法を考えている。

部分文仮説の意味解析においては、対象とする部分単語列の構文木を求めるといった解析的な手法は実現が困難であり、音声言語の意味処理に用いるには適切ではない。本研究では、ネットワーク構造により意味制約を実現し、ネットワーク上のメッセージ伝達による解析手法を提案する。また、計算機による個人スケジュールの管理というタスク [2] において、実際にネットワークを作成し解析を行なった。

2 ネットワークによる意味解析

2.1 基本構造

ネットワークは以下の基本的な構造を持つ。図1にネットワークの一部を示す。

ノード タスクに関する単語
 単語の上位概念
 単語の属性値

Incorporating Network-Based Semantic Constraints into Search Strategy in Continuous Speech Recognition
 Nobuo NUKAGA, Masahiro ARAKI, Tatsuya KAWAHARA, Toyoaki NISHIDA and Shuji DOSHITA
 Department of Information Science, Kyoto University

リンク 継承的な概念を示す isa リンク

あるクラスの要素を示す instanceof リンク

属性を示す attrib リンク

単語間の属性値の差異や語句間の意味的な不整合を検出するために、図1の黒丸で示されるノードを設ける。これを競合判定ノードと呼ぶ。例えば、イベント型での不整合を検出したい場合には、太線で囲まれるノードからリンクを張る。同様に、時間的制約に関しても制約判定ノードを構成する。以下の節で、本タスクにおける主要な意味制約と解析アルゴリズムについて説明する。

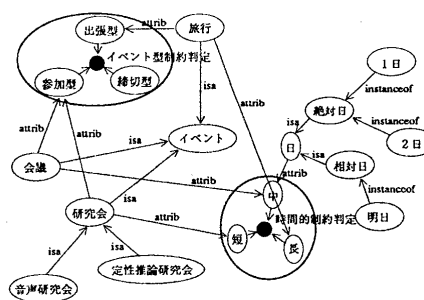


図1: 解析に用いるネットワークの例

2.2 意味制約の分類

タスクに関する意味制約、すなわち意味的な不整合の検出が可能な以下の二つの制約をここでは考える。

1. キーワードの属性に基づいた意味制約
2. タスクにおける意味制約

2.2.1 キーワードの属性に基づいた意味制約

これは、単語をある共通な属性(特徴軸と考えることもできる)に写像したものの一次元的な列をなす制約である。例えば、表1に示されるように、時間的な長さで単語を分類することにより、この種の制約が得られる。本研究で扱うタスクでは、単語もしくはその上位概念における時間的な制約、距離的な制約に基づくものを考えている。簡単のため、ここでは単語を近郊、

表 1: 距離に関する概念分類

近郊概念	中間概念	遠方概念
研究会 昼食会 報告会 打ち合せ 講義 発表会 談話会 勉強会 補講 面会 セミナー 予算会議 演習室 講義室 輪講 特別講演 会議室 堂下研	自然言語処理研究会 定性推論研究会 音声研究会 講演会 会議 分科会 参加 発表 京大 京都 阪大 大阪	信学会 人工知能学会 情報処理学会 出張 休暇 東大 名大 東京 バリ 名古屋 旅行

中間、遠方と三概念に分類しているが、本来はその特徴軸上にある分布を持つ連続関数であると考えられることができる。また、前後の文脈によってその分布を変化させることも可能である。

タスクの性質上、相違の大きい概念(先程の例では「近郊」と「遠方」)は、一文中に同時に存在しないと仮定できる。この意味制約をネットワークで実現し、各制約ノードに単語の属性に関する情報を伝達し、不整合が生じた部分文仮説を棄却する。

2.2.2 タスクにおける意味制約

構文制約を記述した文脈自由文法のみでは、構文的には正しいが意味的には誤っているような文を受理する可能性がある。逆に、係り受け規則や一文中に含まれる概念の個数などを文法に取り込もうとすれば、かえって文法自体が複雑になり、探索に組み込むには現実的でない。

このような問題に対処するために、文法では扱わないような経験的に妥当であると思われる意味制約を設定し、キーワードの属性に基づく意味制約と同様にネットワークで実現する。以下に典型的な制約に関しての例を示す。

- ・ イベント名詞は一文中に三個以上現れない
- ・ 休暇型イベントがあればその他の型のイベント名詞はない
- ・ 開始概念を示す句は高々一つ

2.3 制約に関するメッセージと競合判定

この節では、部分文仮説の意味的な不整合、すなわち競合を判定する機構の説明を行なう。まず、部分文仮説を構成する単語集合の要素から、競合判定ノードにその属性に関するメッセージを伝達する。各競合判定ノードは、蓄積されている情報をもとに不整合が生じていないかを、逐一判定する。その結果、意味的な整合性がないと判定されればその部分文仮説は棄却され、逆に全ての単語に対して不整合が検出できなかつたら、その仮説は意味的に妥当であると判定し棄却しない。競合判定ノードに伝達されるメッセージの例を示す。

- ・ 距離制約に関するメッセージ

`constraint_distance([near]).`

- ・ 開始概念に関するメッセージ

`constraint_timetypes([kaishi_hi]).`

`constraint_timetypes([shuuryou_ji]).`

- ・ イベント型に関するメッセージ

`constraint_eventtypes([sanka_gata]).`

実際に部分文仮説が棄却される例を示す。探索の進行中に「23日から25日まで一研に出張」という部分文を評価する場合を考える。このタスクでは、「一研」というのは「近郊」に含まれる概念であるので、「遠方概念」である「出張」と意味的に整合性がない。従って、この部分文仮説は意味的に妥当ではないとして、この時点で棄却される。

3 解析結果

上記の手法に基づいて、単語数219、上位概念および競合判定ノード数35のネットワークを作成し解析および評価を行なった。発声サンプルは、本タスクで発話され得るスケジュールの登録、変更、問い合わせに関する50の会話文で、成人男性話者8名によって発声された。表2に実験結果を示す。ここで、文理解率とは文の意味をなす自立語がすべて正しく認識された文の割合である。意味制約を利用しない場合と比べて、文認識率において8.5%、文理解率において6.8%の向上がみられた。

表 2: 文認識率

	文認識率		文理解率	
	第一	第十	第一	第十
意味制約なし	46.5%	76.5%	61.0%	85.3%
意味制約あり	55.0%	82.0%	67.8%	88.5%

第一：第一候補、第十：第十候補以内

参考文献

- [1] 河原達也. 探索アルゴリズム -A*探索を中心に - 信学技報 SP92-36, 1992.
- [2] 松本真治. 語彙・構文・意味制約を統合したA*探索による会話音声認識. 信学技報 SP91-93, 1991.