

認識誤りを含む音韻候補列を効率的に解析する方法

1E-2

小作浩美†

柳田益造†

西岡真吾†

山下洋一†

溝口理一郎†

† 郵政省通信総合研究所

‡ 大阪大学産業科学研究所

1 はじめに

近年、コンピュータの普及に伴い、音声によるインターフェースの要望が高まってきており、特に使いやすさの点から、連続音声理解システムの開発が強く望まれている。本稿では、著者らがこれまで行ってきた文節区切り発話の統合的言語処理の技術を踏まえて、現在開発を行っている認識誤りを含んだ文節区切りの無い音韻候補列を解析するシステムについて述べる。

2 従来のシステムについて

一般に音声認識では、ノイズ、個人差などに起因する誤認識を完全に避けることは出来ない。そのため、音韻認識の結果に多くの曖昧性が含まれ、それらを解消することが非常に重要な課題となっている。それに対し、これまで言語情報などの高次の知識を用いることで、これらの誤りを修正する方法が提案されてきている。このような技術は音声理解研究として位置付けられており、我々もその方向に沿って研究を続けてきている。本章では我々が今まで構築してきたシステムの概要とその問題点を述べる。

2.1 従来のシステムの概要

我々の構築してきたシステムは、文節毎に発声された音声を入力とし、局所的な係り受け処理と、文全体に渡る係り受け処理を独立に取り扱うことにより、一度に処理しなければならない係り受けの組合せ数を減少させている。これにより、今まで音声理解をする際に問題となっていた解析途中の組合せ爆発をある程度押えることが出来るようになっていた。また、システム全体は探索を管理する ATMS (Assumption-based Truth Maintenance System)[1] によって制御されるため、推論部と探索制御部との間のインタラクションが最適化され、探索木の枝刈りや、重複した推論の回避を効率的に行なうことが出来た [2]。

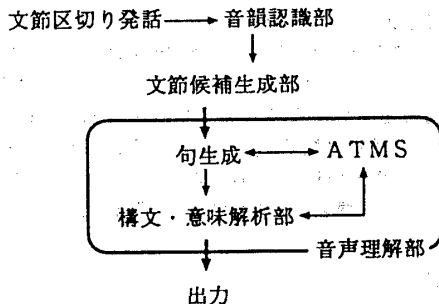


図 1: システムの概念図

音声理解部への入力には文節候補生成部によって生成された文節候補である。なお、形態素解析は文節候補生成部で行なう。音声理解部は、句生成、構文・意味解析部の二つのモジュールに分かれている(図 1)。句生成のモジュールでは文節候補生成部で得られた文節候補から小さな句を生成する。構文・意味解析部では小さな句から構文を予測し、その構文に基づいて文節を決定していく。

特に従来の方式 [3] では、文の構成単位である句を小さな句と大きな句の二種類に分類している。(小さな句とは連体形の動詞句を含まない名詞句、名詞句による動詞のスロットフィリングを含まない動詞句、連用修飾句などであり、それ以外を大きな句という。) 解析の際に、小さな句についてその句を生成する文節接続のみを考慮し、大きな句についてはその句を構成する小さい句同士の係り受けのみを考慮している。これにより、効率的な探索が実現されている。

2.2 システムの問題点

従来のシステムは、入力を文節区切り発話に限定していること、音韻認識の誤りに対する言語処理部からのフィードバックが間接的であることなど、連続音声入力に対する拡張に際して必要となる文節候補生成の柔軟な推論制御は出来ていなかった。そこで、このシステムを基礎として、連続発声の入力を受けられる連続音声解析システムの提案を行なう。

3 新しいシステムについて

連続発声入力の場合、文節区切りについてのボトムアップ的な情報がないため、文節区切りを何らかの方法で決定する必要があるが、これは、音韻候補列に認識誤りが含まれ、かつ、高次の知識が与えられていない場合非常に困難である。従って、適当な音韻候補を組み合わせで文節を仮定するとその候補の数は膨大なものとなり計算時間が増加するため、実用上問題を生じる。そこで、新しいシステムでは、文節生成とその候補数の制御を行なうために、文節候補生成部まで新しいシステムに含め、ATMSの管理下に置く。以下に、新しいシステムの概要と動作例を説明する。

3.1 新しいシステムの概要

本システムは、大きく分けて三つのモジュールより構成されている。それぞれのモジュールは、文節候補生成部、句生成部、構文・意味解析部となっている。さらに探索の管理に ATMS を用いている(図 2)。

本システムでは、(1) 文節候補及び文節境界を仮定とし、ATMS により探索の管理を行なう。これらの仮定は文節候補生成部が文節を生成した際に ATMS に通知される。この仮定により、(2) 時間的に重なりのある文節候補及び文節境界の組は矛盾として取り扱われる。

このような仮定及び矛盾を用いることで、本システムでは以下の二点を容易に実現することが可能である。

(A) 重複した推論の回避 一般に柔軟な探索戦略を用いたシステムでは、仮定の入れ替えが頻繁に起こる。その際、しばしば類似した仮定の下で推論が行なわれることがある。このような場合、本システムでは一旦行なった推論の結果は ATMS によりワーキングメモリ中に戻される

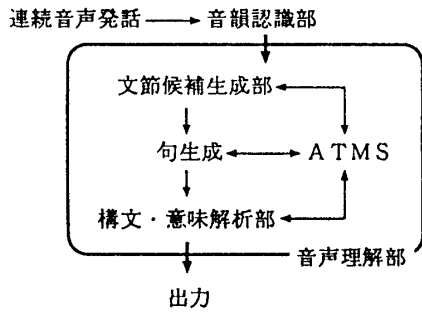


図 2: 新しいシステム概念図

ことが保証されている。従って、新たに増えた文節候補に関する推論のみを行なうことにより重複した推論を回避することが出来る。このように柔軟な推論の制御を行なう場合には、本システムは非常に有効であると言える。

(B) 矛盾した推論の回避、ワーキングメモリからの不要なデータの削除 これは (2) の矛盾を用いることにより ATMS が現在仮定されている文節境界や文節候補と時間的な重なりを持つ文節候補をワーキングメモリから取り除くことが可能となる。これにより、推論システムは時間的な重なりに関する矛盾を含まない推論が保障され、無駄な推論を自動的に回避出来ることになる。またワーキングメモリ中の不要な候補が取り除かれることにより、ワーキングメモリ中のデータの数も減少し、推論の高速化につながる。

ところで、文節候補生成部が文節境界、文節候補列のいずれを先に選択するかによって他方の選択に制限が加わる。このように、仮定の間に依存関係がある場合、ATMS では仮定の選択状況 (環境と呼ぶ) の妥当性は保障されない。この問題の解決には文節候補生成部が文節候補列を選択した時にその両端の文節境界も選択したこととし、一方、文節境界を選択した場合は単にその文節境界を環境に加えるようにすれば良い。以下に具体的な動作例を示す。

3.2 動作例

文節候補生成部 (以後文節部と呼ぶ) では、入力音韻候補列について韻律情報 [4] などの知識から得られる文節境界に関する情報や DP-matching の得点などを考慮することにより、文節の尤もらしい候補を逐次出力する。文節候補は、図 3 に示したように文節候補と文節境界と時間の三つの要素の組合せによって生成される。ここで、文節候補数の制限をするために、音韻候補に DP-matching による得点と単語出現頻度に基づいて計算した尤度に閾値を設けている。

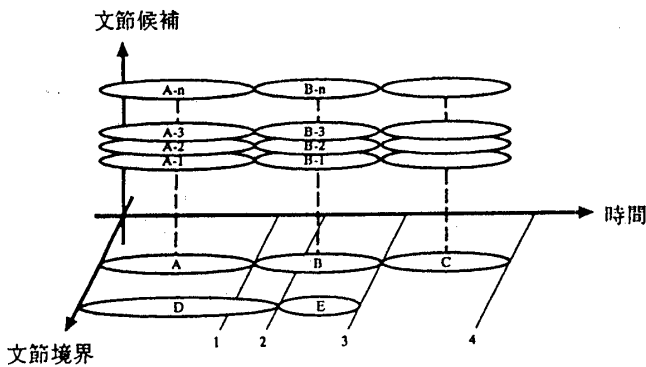


図 3: 解析される候補の例

句生成部においては、従来のシステム同様、小さい句の生成を行なう。構文・意味解析部においては、大きい句の推定を行ない、意味的なチェックを行なって矛盾のない解析結果を出力する。この間、無駄な処

理を制御するため、文節境界を仮定として、矛盾のない状態の探索空間を ATMS が管理する。

具体的な動作は下記の通りである。図 3 の B 列が文節候補列として尤もらしいと文節部で判断され、句生成部に送られたとする。その時、文節部においては、B 列に附属する情報を ATMS に通知する。それに伴い、B 列に時間的に重なる他の列 (D E) が矛盾知識により文節部のワーキングメモリから除かれる。これにより、文節部は B 列と矛盾なく組合せられる妥当な候補列 (A C) だけを取り扱うことができ、無駄な候補の生成することが自動的に抑制される。このようにして文節部において、時間的に矛盾なく組み合わせることができる文節候補列を順次句生成部に送っていく。

句生成部では、送られた一連の文節候補列を用いて小さな句を生成し、それを構文・意味解析部に送り、ATMS にその候補に関する情報を通知する。もし、文節候補列内で解を確定することが出来なかったなど不都合が生じた場合、その情報を文節部に伝える。

例えば、句生成部において、先ほどの B 列を用いた推論で高次の知識等を使用した結果、推論の中止が判断され、新たなる候補列が必要になったとする。句生成部は文節部に他の候補を要求する。それを受けて、文節部が B 列の代わりに D 列が妥当であると判断すれば、それを句生成部に送り、ATMS にそのことを通知する。それによって、ATMS は B 列に関するデータを文節部、句生成部のワーキングメモリから取り除き、D 列と矛盾がない過去に処理したデータを各々のワーキングメモリに入れる。

構文・意味解析部においても処理過程における同様な探索管理が行なわれる。

このように、連続音声に対しても、ATMS で文節境界を仮定として管理することにより無駄な処理を制御しながら、解析することが可能となっている。

4 まとめ

本稿では、文節候補生成部のモジュールを ATMS の管理下に置くことで、連続音声入力を処理出来るシステムについて提案した。また、ATMS を用いて探索管理をすることで無駄な処理を早い時期に回避するとともに、処理結果を再利用することが可能となり、効率的な推論が実現されている。

今後の課題として、本システムにおいて、設けている音韻候補の閾値を可変なものにすることがあげられる。つまり、閾値を厳しく設定すれば、正しい認識が出来ない可能性もあり、今後柔軟な対応を考える必要がある。また、高次の処理結果から得た知識を低レベルの処理にフィードバックさせ、音韻や単語レベルのスポットニングをさせる [5] などの制御を行なうことにより、統合的なシステムの構築も検討していく。さらに、システムを音韻認識部や、解析された入力からの文脈情報を抽出することや他の言語処理技術を組み込んで、ATMS により効率の良い音声理解システムの構築と必要な知識の整理を目指す。

参考文献

- [1] de Kleer, J.: "An Assumption-based Truth Maintenance System" *Artificial Intelligence*, 28, pp.127-162(1986).
- [2] 西岡他: "ATMS を用いた問題解決システムの構築技法" *信学技報*, AI90-72(1991).
- [3] 西岡他: "音声理解システムのための効率的な曖昧さ解消の枠組" *通信学会 2 種研*, SPREC-91-2(1992).
- [4] 小松他: "韻律情報を利用した構文推定およびワードスポットによる会話音声理解方式" *信学論*, Vol. J71-D No.7, pp.1218-1228(1988-7).
- [5] 伊藤他: "拡張 LR 構文解析法を用いた連続音声認識" *信学技報*, SP90-74:NLC90-46(1990).