

連続音声認識エキスパートシステムSPREX II
における認識知識の自動生成

桜井 徹† 野村 康雄† 辻野 克彦‡ 千種 俊輔‡ 溝口 理一郎‡ 角所 収‡
† 関西大学工学部 ‡ 大阪大学産業科学研究所

エキスパートシステムを構築するためには、専門家から多くの知識を正確に聞き出す必要があるが、専門家にとって自らの知識を自発的に整理して述べることは、多くの場合困難であり、この点はエキスパートシステムを構築する上での知識獲得の問題としてしばしば議論されている。しかし、一方で専門家は具体的な例に対して自らの知識を使って見せることはできるのであるから、そのような例から帰納的に学習することにより、知識を獲得することができると考えられる。本稿では、エキスパートシステムとして実現されている連続音声認識システムSPREX IIの認識知識を学習するために必要とされる学習機構ARISの概要とこれを用いた学習に適した認識手法、及び学習された知識について述べる。

Inductive generation of knowledge for
a continuous speech recognition expert system : SPREX II

Toru SAKURAI†, Yasuo NOMURA†, Katuhiko TSUJINO‡, Shunsuke CHIGUSA‡,
Riichiro MIZOGUCHI‡, and Osamu KAKUSHO‡

† Faculty of Engineering, Kansai University

‡ Institute of Scientific and Industrial Reserch, Osaka University

Needless to say, an expert system requires a lot of heuristic knowledge of a human expert, but it is rather difficult for the expert to describe all of his knowledge of his own accord. On the other hand, he can use the knowledge for some examples, which suggests the possibility of acquiring knowledge by inductive learning. This paper describes an inductive generation of knowledge for a continuous speech recognition expert system : SPREX II and an inductive learning system : ARIS which is an extended version of a rule induction system and has many advantages that is necessary for SPREX.

1. まえがき

エキスパートシステムは、「専門家」と呼ばれる人間の知識を用いることにより、彼らとその分野で同程度の能力を発揮できるような計算機システムのことである。このようなシステムを構築するためには、専門家から多くの知識を正確に聞き出し、これを計算機で処理できるように表現しなければならない。ところが、この作業は容易ではなく、いわゆる Feigenbaum bottleneck と呼ばれる知識獲得に関する重要な問題とされている。すなわち、専門家は自分の持つ経験的知識を正確に自発的に述べるのが困難であり、概念的に表現することができても、それを計算機で処理できるような形式に具体的に表現できないことが多い。しかし、そのような場合においても、具体的な事例に対して用いられるべき知識を述べるができることが多く、この場合、そのようにして得られた多くの例題から帰納的に学習することによって知識を獲得することができる。また、その学習結果が専門家に対する知的刺激になって、新しい知識の獲得を支援することにもなる。筆者等は、音声理解システム SPURT-I (a SPEech

Understanding system with Rule-based and Topic-directed architecture)の音響処理部にあたる SPREX II (a SPEech Recognition EXpert system)の開発を行ってきた。このシステムは、上で述べたような知識獲得上の問題に対処するため、さまざまな知識ベース構築支援環境を備えており、本システムが採用した知識工学的実現方法が元来持つ透明性と自己説明性とあいまって、音声認識に対する知識工学的アプローチの有効性を確認することができた^{[1][2]}。しかし、システムの性能向上に伴い、より多くの話者に対応しようとすると、専門家は多くの音声試料を参照せねばならず、ルール作成が次第に困難になってきた。

本稿では、以上のような考えに基づいて開発された、大量データから認識ルールを帰納的に学習する適応的ルールインダクションシステム ARIS (Adaptive Rule Induction System)及び、その SPREX への適用方法について述べる。以下、2節では SPREX の認識処理について、3節では ARIS とルールインダクションシステムの概要について、4節、5節では ARIS を用いた SPREX の認識知識の自動生成について述べる。

2. SPREX の認識手法

音声の専門家は、図1に示したような音声のバ

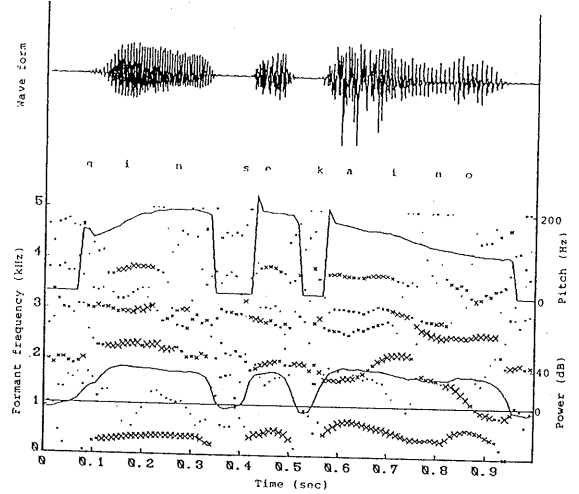


図1 連続音声とその特徴パラメータ

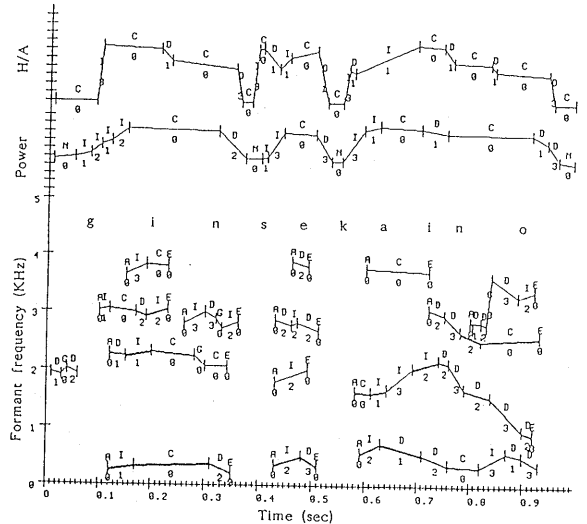


図2 図1の音声に関する記述の例

ワー、ピッチ、ホルマント周波数等の特徴パラメータの時間変化の様子を視察することにより、かなりの正確さでこれに含まれる音韻の位置とその種類を認識することができる。SPREXはこの専門家の知識を用いたエキスパートシステムである。SPREXの処理は、大きく分けて次の2つからなっている。まず、図1は図2で示したように状態記述される。即ち、専門家は各特徴パラメータの時間点ごとの値に注目しているのではなく、それらの変化の傾向、たとえば急激な増加、減少、不連続点等に注目して音韻を認識している。状態記述の処理は、このようなふるまいをシミュレートすることを目的としている。次に、これらの記号すなわち特徴を条件部に持ち、結論部に音韻カ

表1 SPREXの性能

		SPREX II	SPREX I
境界認識	挿入	2.6% 2.4%*	6.2% 3.0%*
	欠落	3.4% 2.7%*	3.5% 2.8%*
音韻認識**	母音	94%	92%
	子音群	85%	80%
	平均	90%	85%
試料音韻数		862	402
境界、音韻群共		87%	85%

*音韻認識時に誤りが回復されたものを除く

** SPREX Iでは音韻境界認識を含んだ性能
SPREX IIでは正解音韻境界を使用

テゴリー等を持つルールとして専門家の知識を表現し、それを図2に適用することにより認識を行わせることができる。この処理方法に基づき、人間の専門家が持つ知識をもとに構築された知識ベースから構成されるSPREXの性能を表1に示す。

3. 帰納的学習

帰納的学習を用いることにより、専門家から例題を通して知識を獲得することができることについては1節で述べた。本節では、SPREXに帰納的学習を適用する際に検討すべき問題点について述べる。SPREXの音声認識知識は、基本的には特徴パラメータの記号化された状態記述を用いて表現されるため、記号化された例題を扱うことのできる学習方式が必要とされる。そのためSPREXの学習機構であるARISでは、特徴ベクトルから、それらを分類するための決定木を帰納的に学習するルールインダクション（以下、RIと略す）と呼ばれる学習アルゴリズムを採用している。ところが、SPREXにおける学習対象が音声という物理量であること、専門家がしばしば不適切な「正解」を与えることにより例題に雑音が含まれ、RIだけでは対応できない幾つかの問題がある。本節では、RIの学習アルゴリズム及びRIを音声認識知識の自動生成に適用する際に生じてきた問題点について述べると共に、それに対処すべく拡張されたARISの機能について述べる。

3.1. 帰納的学習アルゴリズムRI

RIは、属性ベクトルとその属性ベクトルが属するクラスとで表現された例題から、その例題がどのクラスに属するかを分類する決定木を生成する帰納的学習アルゴリズムである。以下、この学習アルゴリズムの概要について述べる。まず、用

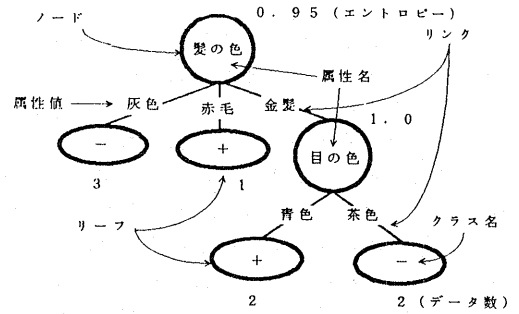


図3 図式表現された決定木

意された属性選択基準により、全例題を最も効率よく分類できる属性が1つ選択される。次にその属性の属性値により、例題をそれぞれの属性値に対応した部分集合に分割する。この作業をすべての部分集合が、単一クラスからなる部分集合になるまで再帰的に繰り返され、決定木が生成される。図3は、属性ベクトル（身長、髪の色、目の色）で表現された例を「+」「-」の2つのクラスに分ける決定木を示している。この例では、属性選択基準としてその属性を使って得られる情報量を用いている。

3.2. SPREXに適用する際の問題点と解決法

3.1で述べたように、RIは与えられた例題を効率よく、完全に分類する決定木を生成できるという特長をもつ反面、SPREXに適用しようとするとき、以下に示したような問題を生じる。

(1) 雑音問題

RIにより生成された決定木の性能が、与えられた例題の質に依存するものであることはいまでもない。しかし、音声という物理量を扱うため、特徴量そのものに物理的雑音の付いていることに加え、調音結合の影響により音韻境界の位置が確定的でなく、一つの例題に対して幾通りかの「正解」を与えることなどにより、例題の中に多くの雑音（曖昧性）が生じる。このような例題を用いて生成された決定木は、例題以外のデータに対して適応能力の低い特殊なものとなる傾向がある。

このような問題を解決するため、ARISでは、ヒューリスティクスに基づく一般化知識を用いて、生成された決定木に対して一般化を行う。一般化知識には、決定木生成の際に用いられた属性選択の様子やサンプル数等により一般化を行う汎用型一般化知識と、属性間の類似性や属性値の順序関係などのドメイン知識に基づくドメイン依存型一般化知識がある。

(2)属性選択基準

R Iは、もともと属性選択基準として情報量を用いているが、この属性選択基準は、同じ特徴量でも属性値の数の少ない属性と多い属性とがあれば、一般的に分離能力の高い属性値の数の多い属性を選択する傾向が高い。これは、例題に雑音がない場合、効率的であるが、属性値の多い属性ほど一度に細かい部分集合に分類するため、生成される決定木は雑音に対して弱いものになる。そこで、音声のように雑音が含まれる例題に対して適応能力を高めるため、A R I Sでは、情報量の代わりに情報伝送率を用いている。情報伝送率とは、各属性を使用することで得られる情報量をその属性の最大情報量で正規化したものであり、情報量を用いた場合と比べて、より大きな部分集合を生成する傾向を持つ。これにより、より多くの特徴を用いて大きく分類していくようになり、雑音に強い決定木を生成することができると考えられる。

(3)初期属性の与え方

例題は、それを特徴づける属性の集合で表現される。従って、その中に含まれる属性の概念レベルが高ければ高いほど、例題の質も高くなり、生成される決定木の性能も良くなる。しかし、概念レベルの高い属性を用意しようとする、ルール作成する際に生じる知識獲得の問題と同様の問題を生じると考えられる。一方で、概念レベルの低い属性は、雑音に弱く、問題点(1)に示した問題を生じる傾向が強い。

この両者を解決するために、A R I Sは、生成された決定木を評価することで、与えられた概念レベルの低い属性から、より概念の高い属性を生成し、それを新たな属性として付け加えることにより、より性能のよい決定木を生成していく機能を持っている。以下、A R I Sの属性生成について述べる。属性生成のトリガーは、「あるノードのリンクに結び付いているものがすべてリーフである場合、このノードに含まれる例題は、そこに至るまでの属性を用いても分離が困難であることを表している。しかし、それだけを分離する新しい属性があれば、識別が容易になるのではないか。」というヒューリスティックに基づくものであり、そのようなノードに含まれるクラスの例題のみを効果的に分類できる決定木を、再びA R I Sを用いて生成する。こうして、生成された部分決定木の判断結果を新しい属性とする。

(4)数値属性の取り扱い

音韻を認識するルールを学習する際、音韻時間長、パラメータ値などの数値特徴量を扱う必要がでてくる。しかし、このような数値特徴量をそのまま取り扱おうと、全く一般性のない条件を生成してしまうことは、数値の稠密性を考えると明らかである。

A R I Sでは、数値属性が使われる度に、分類しようとしている例題をその属性値を用いてクラスタリング及びラベル付けを行い、そのラベルを用いて数値属性を扱うことによりこの問題に対処している。

3.3. A R I Sの構成

A R I Sは図4に示したような6つのモジュールから構成されている。以下、各モジュールの機能について述べる。

(1)属性生成モジュール

与えられた例題と初期属性及び属性生成モジュールで生成された新しい属性を用いてR Iモジュールへ入力するための属性ベクトルを生成する。

(2)R Iモジュール

R Iを用いて決定木を生成すると同時に、新しく生成された属性の有効性の評価を行なう。

(3)クラスタリングモジュール

k-means法を用いて数値データを記号化する。R Iモジュールで数値属性が使用される度に本モジュールが呼び出され、適切なラベル付けが施される。k(クラスタ数)は、属性選択基準を最大にするように自動的に調整される。

(4)属性評価モジュール

A R I Sは、R Iと属性生成の過程を繰り返して新しい属性を発見する。しかし、生成される属性全てが使用されるわけではなく、処理が進むと使用される属性に変化がなくなってくる。本モジ

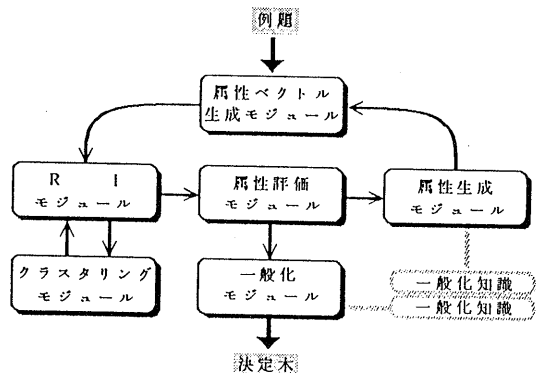


図4 A R I Sの構成

ジュールでは、生成された決定木を評価し、使用される属性に変化があれば処理を続行し、変化がなくなれば(5)で述べる一般化ジュールへ処理を渡す機能をもつ。

(5)一般化ジュール

R I モジュールで生成され、使用される属性の変化がなくなった決定木に対して一般化を行い、最終的な決定木を出力する。また、属性生成の際の部分決定木作成においても、本モジュールが使用される。

(6)属性生成モジュール

本モジュールは、R I で生成された決定木を評価することで新しい属性の生成を行なう。具体的には、R I モジュールと一般化モジュールを再帰的に用いることで3.3節で述べたような部分決定木からなる属性の生成を行っている。本モジュールにおける一般化の知識は、分離できない例題があるならば生成される属性に「分離しにくい」という属性値を持つように一般化する。これにより、この属性を生成する時に用いられなかったクラスのサンプルが細かく分割されてしまう傾向を、和らげることができる。

4. 子音認識知識の自動生成

本節では、ARISを用いて子音の分類を行う知識を学習する例について述べる。属性ベクトルとして、あらかじめ用意された正解音韻境界を基に、その区間中に含まれる特徴パラメータの遷移状態を、表2で示された属性、属性値により表現したものをを用いている。また、正解クラスとして、表3で示された音韻グループを用いている。

図5は、属性生成サイクルと認識率の関係を示したものである。①は、一般化された決定木を用いて学習サンプルを認識した場合、②は、一般化された決定木を用いてテストサンプルを認識した場合、③は、一般化する前の決定木を用いてテストサンプルを認識した場合の結果を示している。①、②、③共に属性生成されるごとに認識率が上がり、属性生成の効果が確認できた。また、②と

表2 子音認識知識の自動生成で使用された属性と属性値

パラメータ名	属性名	属性値
セグメント長	LEN	SS S H L LL
有音/無音	U/V-0 U/V-1 ... U/V-i U/V-LAST1 U/V-LAST0 U/V-DOMINANT	VOICE UNVOICE NIL
パワー	PWR-0 PWR-1 ... PWR-j PWR-LAST1 PWR-LAST0 PWR-DOMINANT	D3 D2 D1 C N I1 I2 I3 NIL
第一ホルマント周波数	F1-0 F1-1 ... F1-k F1-LAST1 F1-LAST0 F1-DOMINANT	D3 D2 D1 C N I1 I2 I3 G NON NIL
第二ホルマント周波数	F2-0 F2-1 ... F2-l F2-LAST1 F2-LAST0 F2-DOMINANT	D3 D2 D1 C N I1 I2 I3 G NON NIL
第三ホルマント周波数	F3-0 F3-1 ... F3-m F3-LAST1 F3-LAST0 F3-DOMINANT	D3 D2 D1 C N I1 I2 I3 G NON NIL
パワーの高域周波数成分比	I1/A-0 I1/A-1 ... I1/A-n I1/A-LAST1 I1/A-LAST0 I1/A-DOMINANT	D3 D2 D1 C I1 I2 I3 NIL

表3 音素のグループ化

音素グループ名	含まれる代表的音素例
VOWEL	/a/ /i/ /u/ /e/ /o/
PTK	/p/ /t/ /k/
BDGZ	/b/ /d/ /g/ /z/
R	/r/
SH	/s/ /c/ /v/ /v/
MN	/m/ /n/
W	/w/
Y	/j/

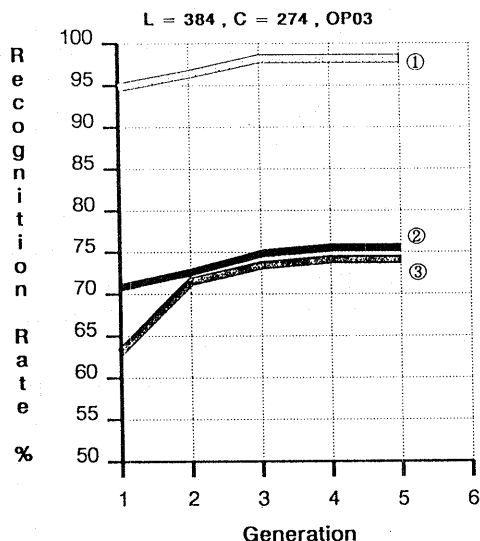


図5 属性の生成、追加による認識率の変化
③の認識率の関係より、一般化の効果が確認できた。

5. 音韻境界認識知識の自動生成

人間の専門家が開発した境界認識方式は音韻境界に見られる特徴的な変化の分布を総合的に考慮してパラメトリックな調整により境界を決定するものである。ところが、ARISは基本的に分類型の知識の学習を行うシステムであるため、このような知識の学習に適用するのは困難である。そこで、ARISにより音韻境界認識知識を学習するためには従来の認識方式を変えることが必要であると考えられる。本節では、人間が開発した音韻境界認識方式(以下、パラメトリックな音韻境界認識方式と呼ぶ。)及びそれをARISで学習する際の問題点について述べるとともに、これに基づき新しく開発されたARISによる学習に適した階層的音韻構造に基づく音韻境界認識方式について述べる。最後にその階層構造に基づくARISによる音韻境界認識知識の自動生成について述べる。

5.1. パラメトリックな音韻境界認識方式

パラメトリックな音韻境界認識方式は、音韻境界に見られる特徴パラメータの変化点の集中している部分を音韻境界と認識するエキスパートの振舞いをシミュレートしたもので、次のように処理が行われる。

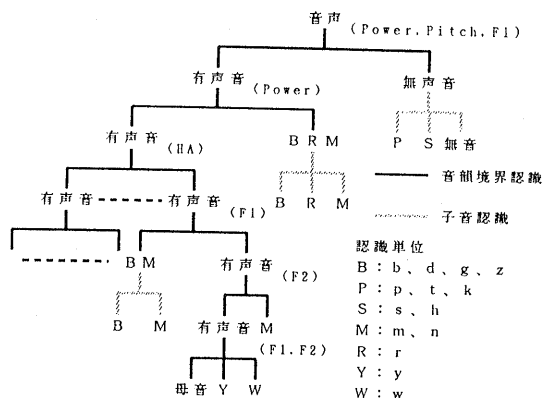
- (1) ルールを用いて音韻境界の候補と考えられる特徴パラメータの変化点にそのパラメータの重要性和変化の程度に応じた得点を与える。
- (2) 各特徴パラメータ間の音韻境界候補の時間的なずれを吸収するため、個々のルールの得点に適当な窓をかけたうえで、時間点毎に合計し得点分布を得る。
- (3) 得点分布は、境界らしさを表しているため、その中の極大点をルールを用いて選択し、音韻境界を決定する。

この認識方式は、各々の時間点における複数の特徴パラメータの変化を総合的に考慮できるという特徴を持っている反面、得点付けという手続き的な方法を用いているので、ARISのような分類型学習システムの適用が困難であるという問題を持つ。

5.2. 階層的音韻構造に基づく音韻境界認識方式

音韻境界で得点付けが必要であったのは、全音韻境界を一度に認識することを目指したため、全パラメータを総合的に判断する必要があったからである。ところが、知識を音韻の階層的関係に基づき構成し、その階層毎に境界を逐次認識していくようにすることで音韻認識知識の整理、作成が容易になり、得点付けの手続きを用いずに認識ルールだけで音韻境界を決定することが出来る。この方法は、ARISを用いた学習が容易であるだけでなく、以下のような特徴を持っている。

- (1) 音韻境界が決定された時点である程度音韻の候補が絞られているため、次に適用される子音認識ルールの記述を容易にできる。
- (2) 各階層に応じたドメイン知識（その階層の認識のために有効な特徴パラメータ等に関する知識）を用いることができるので、不安定な特徴パラメータの変動に影響されない雑音に強い認識が可能となる。
- (3) /y/, /w/といった特定の状況下で現れる半母音等に対して、音声に関する深い知識を基に境界認識知識が整理でき、これらの音韻に対する性能向上が期待できる。



() 内は、各階層において有効なパラメータ。
 HA: パワーの高域成分比。
 F1, F2: 第1, 第2ホルムント周波数。

図6 音韻の階層構造

5.3. 階層的音韻構造に基づく認識方式の処理手順

新音韻境界認識方式では、エキスパートが作成した音韻の階層構造に従って知識が整理されており、各階層には、その階層の音韻を認識する際に有効だと考えられる特徴パラメータを用いて構成された音韻境界決定ルールが用意されている。図6に、本方式で用いられている階層構造を示す。図中の()内は各階層において用いられている特徴パラメータを示す。

まず、全音韻区間は、パワー、ピッチ、第1ホルムント周波数等の特徴パラメータを用いて有声音、無声音区間に大きく分割される。たとえば、「ピッチがなくて、かつ第1ホルムント周波数がない場合は、無声音区間である。」というような知識が用意されている。階層構造に従って、このような知識を逐次用いることにより、区間を次第に細分割し、最終的に音韻単位に分割することができる。

5.4. 性能評価

表4は、人間により作成されたルールを用いて階層的音韻構造に基づく認識方式とパラメトリックな認識方式の性能評価を行なったものである。階層構造を導入することにより、実験試料が増したにもかかわらず、挿入誤り、欠落誤り共に減少し、新認識方式の有効性が確認できた。特に、語尾、無声音区間等に現れる不安定な特徴パラメータの変動に対して強く、挿入誤りを減少させるのに貢献している。

この実験により、新認識方式を用いることにより、得点付けの手続きを用いなくても、ルールだけで認識率の高い音韻境界認識が可能であること

表4 セグメンテーションの正当率

		境界数	挿入誤り	欠落誤り
得点付けによる音韻境界認識方式		5 2 5	6. 2%	3. 5%
階層的音韻境界認識方式	closed data	5 2 5	1. 7%	2. 9%
	open data	4 3 5	4. 1%	3. 0%
	全 data	9 6 0	2. 8%	2. 9%

が確認できた。また、このように階層化することで、各階層で必要とされる特徴パラメータ、特徴(属性)、例題を整理することができた。これにより、この認識方式に基づきARISを用いて学習する際、概念レベルの高い属性を比較的容易に用意することができ、雑音に強く性能のよい決定木を生成できると考えられる。

5.5. 新認識方式に基づく学習

パラメトリックな音韻境界認識は、境界そのものの特徴をとらえ認識しているのに対して、階層的音韻構造に基づく音韻境界認識は、境界を各階層区間の区切りとして認識する。すなわち、注目している音韻区間が、どの階層に属するかを判定し、そのような部分的な階層性のある程度、大局的に判断して境界を認識するものである。新認識方式に基づく学習というのは、上で述べた音韻区間を単位区間と考え、この単位区間を一つの例題として属性ベクトルを作成し、ARISを用いて、その単位区間を認識する決定木の自動生成を意味する。一方、実際に境界を認識する際は、このように生成された決定木で単位区間を認識し、それをマージし、境界を決定する。この際用いられるマージ用の知識は学習を必要としない単純なものが用意されている。

5.6. 単位区間認識用決定木の生成過程

属性ベクトルは、最初に設定された階層構造を基にして、各階層毎に用意された属性ベクトル生成知識により単位区間毎に生成される。この際、

表5 音韻区間認識知識の自動生成で使用された属性と属性値例

概念レベルの低い属性	属性名	属性値
ビッチ	UV-1 UV-2	VOICE UNVOICE EDGE
パワー	PWR-1 PWR-2 PWR-3 PWR-4	D3 D2 D1 C0 I1 I2 I3 NO EDGE
第1ホルマント周波数	F1-1 F1-2 F1-3 F1-4	NONE G14 G04 D3 D2 D1 C0 I1 I2 I3 EDGE
第2ホルマント周波数	F2-1 F2-2 F2-3 F2-4	NONE G14 G04 D3 D2 D1 C0 I1 I2 I3 EDGE
第3ホルマント周波数	F3-1 F3-2 F3-3 F3-4	NONE G14 G04 D3 D2 D1 C0 I1 I2 I3 EDGE
高域周波数成分比	HA-1 HA-2 HA-3 HA-4	D3 D2 D1 C0 I1 I2 I3 EDGE
概念レベルの高い属性	属性名	属性値
単位区間属性ベクトルの中に無声音区間と重なったものがあるか。	Cross-hierarchy-1	OK NO
現時点から無声音区間までの時間的距離が30msec以下であるか。	Time-diff-≤30-1 Time-diff-≤30-2	OK NO

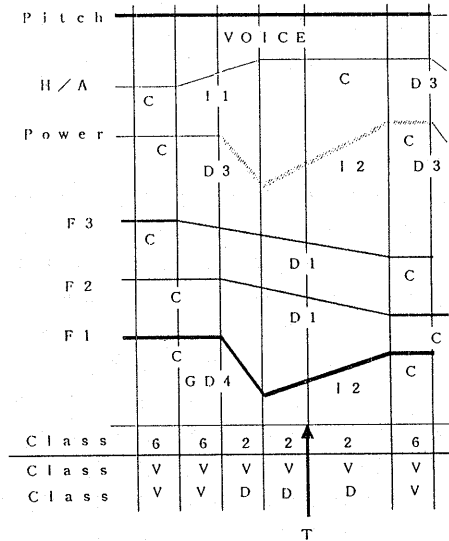


図7 特徴パラメータの概念図

同じ属性ベクトルを持った単位区間が続く場合、これらの単位区間を一つの単位区間として扱う。このように生成された属性ベクトルを学習サンプルとして、ARISを用いて音韻区間認識用決定木を生成する。表5は、音韻区間認識知識の決定木をARISを用いて生成する際に用いられる属性と属性値である。図7は、特徴パラメータの時間変化を概念的に表したものである。属性ベクトルは、各時点の前後それぞれ二つの状態、すなわち、計4つの区間からなるベクトルで表される。例えば、図7で時間点Tに対して、パワーに関する属性ベクトルは(D3 I2 I2 C)となる。また、階層的に属性ベクトルを生成していくので、上位階層の結果を利用した概念レベルの高い属性を利用することができる。例えば、有声音区間でパワーのDIP区間に注目して有声破裂音区間と有声音区間とを分類する場合、形状的にパワーのDIPは、無声破裂音の前後にある不安定なパラメータの動きと似ている。そこで、これらを区別するためには、その状態が無声音区間付近にあるかど

うかを確認する必要がある。このため、無声音区
間からの時間的距離を計算してこれを属性として
加える。

5.7. 認識結果

現在、各階層での音韻境界認識知識の自動生成
の実験を進めている。表6は、学習によって生成
された決定木により無声音区間を認識した結果で
ある。無声音区間の場合、特徴パラメータの有無
等で比較的簡単に分類ができるので、例題が少な
くとも比較的性能のいい決定木を生成することが
できた。しかし、パワーDIPにより音韻境界を
認識する場合、DIP自身の特徴はその形にある
ので、単位区間のようにDIPの部分的特徴しか
記述していない例題から、確実にDIP区間を認
識する決定木を生成することはむずかしいと考
えられる。また、パワーDIP区間が全有声音区
間と比べて少なく、学習で用いられる例題に片寄
りが生じている。このため、一般化の際、数の少
ないパワーDIP区間を表す例が雑音として扱われ、
過度に一般化される傾向が見られた。そのため、
学習サンプルを増やすか、一般化知識にドメイン
知識を用いて学習対象に応じた一般化知識を用
意するかで、性能向上が期待できると考えてい
る。

表6 決定木による認識率（無声音区間）

	境界数	挿入誤り	欠落誤り
決定木による無声音 境界の認識	90	5.6%	3.3%

6. 検討とまとめ

ARISを用いたSPREX IIにおける認識知
識の自動生成について述べた。音韻境界認識方
式に音韻の階層構造を導入することで境界認識の
性能向上、並びにARISへの適用が可能である
ことが確認できた。最後に今後の検討課題につ
いて述べる。

(1)階層構造と階層別例題の獲得

現在、音韻認識知識の階層構造は、エキス
パートの知識に基づいて構成されている。一方
でARISで生成された子音認識知識の決定木
は、音韻の階層構造に良く対応づけられること
が確認されており、これを用いた階層構造の自
動獲得が可能であると考えられる。また、この
決定木を生成する際に用いられた試料やその
識別条件等を参考にして、境界学習のための
例題や各階層で重要な特徴パラメータを獲
得することも可能であると考え

られる。

(2)音韻の階層構造に基づく連続音声認識エキス パートシステム

従来の認識方式は、特徴パラメータの変動から
境界らしさを検出し、これを基に音韻境界、及び
音韻認識を行っていくボトムアップな認識方式
であった。しかし、SPURT-Iの言語処理部に
あたるASP (Association-based Parser)^[3]と
の関連を深め、その情報を有効に利用するため
には、ある程度しぼり込まれた音韻の中からト
ップダウンに認識する方法が必要と考えられる。
そこで、階層的音韻構造に基づく音韻境界認
識で採用した方法を発展させて、音韻の階層
構造を用いてトップダウンに音韻を認識する
方式について検討を進めている。

[参考文献]

- [1] 溝口, 田中, 福田, 辻野, 角所: 連続音声
音声認識エキスパートシステム—SPREX—,
電子情報通信学会論文誌 (D) Vol J70-D, NO.6
PP.1189-1198 (1987).
- [2] 辻野, 千種, 桜井, 野村, 溝口, 角所: 連
続音声認識エキスパートシステム: SPREX II
—知識ベース構築支援環境と学習機能の統合—,
情報処理学会第35回大会, 11-1(1987).
- [3] 堀, 大隅, 尾崎, 溝口, 角所: 音声理解シ
ステムのための連想パーザASP—主題に基づ
く連想機構とその評価—, 情報処理学会第35回大会,
7t-1(1987).