

UN SYSTEME D'APPRENTISSAGE SYMBOLIQUE POUR LE DECODAGE ACOUSTICO-PHONETIQUE

J. Guizol

G.I.A., Faculté de Luminy, 70 Route Léon Lachamp
13288 Marseille Cedex 09 FRANCE

I - INTRODUCTION

Le système que nous présentons constitue une phase d'acquisition automatique de connaissances symboliques en vue du décodage acoustico-phonétique de la parole. À partir d'un ensemble d'exemples constitués par un codage de portions de signal issues des réalisations de phrases types, l'apprentissage a pour tâche de fournir une caractérisation du concept induit par le choix de ces exemples.

Les règles de stratégie, de réécriture et de généralisation sont définies en PROLOG Colmerauer 83], de même que celles permettant de décrire les objets propres à l'application ou les contraintes de généralisation.

Les règles produites par la méthode comportent une information contextuelle et sont évaluées en fonction du nombre d'exemples qu'elles vérifient et de la précision de détermination des objets qu'elles utilisent [Michalski 80a].

II - CODAGE DU SIGNAL

Notre système opérant un apprentissage par acquisition de concept (événement acoustico-phonétique, phonème, trait acoustique, etc.), les exemples sont constitués d'une représentation symbolique d'une portion de signal. Celle-ci est obtenue grâce à un ensemble de prédicats évaluables permettant, pour chacun des paramètres du signal, de déterminer maxima, minima, moyennes, pentes afin de modéliser les évolutions temporelles sous la forme de collines, vallées et portions monotones reliées entre elles par des relations situationnelles (coïncidence, succession, chevauchement, etc.) [Meloni 86].

A l'issue de ce traitement, nous disposons donc d'une part, de formes élémentaires issues de l'analyse du signal, représentant l'évolution dans le temps des divers paramètres, et d'autre part, de la chaîne phonémique associée. A partir de ces données, le module d'apprentissage détermine les configurations des diverses formes caractéristiques du phonème, de la classe de phonème ou du trait acoustique que l'on désire étudier.

III - PRESENTATION DE LA METHODE

L'apprentissage inductif que nous réalisons sur les données définies précédemment s'effectue sur des exemples positifs. Si nous avons choisi dans un premier temps de n'opérer que sur de tels exemples, c'est parce que nous doutions de la pertinence et de la validité de contre-exemples dans le domaine étudié. Toutefois, on note des ambiguïtés sur les règles produites caractérisant des classes proches. Afin de les réduire, nous utiliserons pour préciser une classe donnée, des contre-exemples constitués de représentants des classes concurrentes.

III . 1 - Structures des Exemples

Conçu dans l'optique d'être indépendant de l'utilisation qui en sera faite, le système admet des exemples

dont la syntaxe est peu contraignante. Toutefois, dans un but d'efficacité, aucune interface n'a été prévue et ils doivent donc avoir une structure de termes PROLOG. Plus précisément, chacun d'eux est constitué d'une liste de n-uplets décrivant une conjonction de propriétés et de relations s'appliquant sur des objets.

Les propriétés sont des doublets formés d'une part d'un prédicat générique et d'une spécification, d'autre part d'un objet identifié par un terme. Les relations sont des triplets dont le premier élément est le symbole relationnel, le second un objet et le troisième une liste (de longueur quelconque) d'objets en relation avec le précédent.

Exemple :

```
<phoneme(ii),61>.<contexte-gauche(tt),56>.  
<contexte-droit(tt),76>.<colline-r0(niveau1),62>.  
<coincide,61,62>...
```

III . 2 - Organisation du Système

L'ensemble des règles constituant le système se sépare en deux parties totalement distinctes :

- une partie constituant le moteur d'induction proprement dit et dont les règles de production ou de stratégie présupposent uniquement les exemples mis sous la forme décrite précédemment ;
- une partie contenant la base de connaissance propre au domaine étudié (BCD) constituant un module facilement modifiable ou même interchangeable et qui permet de fournir des informations qui seront utilisées par le système (description arborescente des propriétés nécessaire à la généralisation par hiérarchie, modèle imposé à la forme généralisée, propriétés des relations, etc.).

III . 3 - Principe de Fonctionnement du Système

Le principe de fonctionnement s'inspire de la méthode proposée par R. S. Michalski [Michalski 80b]. La caractérisation d'un concept à partir des exemples s'opère pas-à-pas. A chaque étape, disposant d'une forme généralisée FG (issue de l'étape précédente) constituée d'une disjonction de règles, un nouvel exemple est proposé.

Dans un premier temps, le système s'assure que celui-ci n'est pas déjà inclus dans une des règles de FG. Ceci se fait très simplement sous PROLOG en transformant chaque terme de l'exemple en une clause unaire et en démontrant FG sur l'ensemble ainsi obtenu. Si l'exemple est vérifié, il est ignoré. Dans le cas contraire, le système va l'introduire dans chaque élément de la disjonction constituée par FG. En cas d'échec sur l'un d'entre eux, celui-ci demeure inchangé dans la nouvelle forme généralisée. En cas d'échec total, l'exemple est signalé à l'utilisateur.

Nous décrivons ci-dessous les étapes successives de l'introduction de l'exemple dans FG.

1) On procède tout d'abord à une factorisation de l'exemple en regroupant les arguments d'une propriété par conjonction interne :

$$P[A] \wedge P[B] \rightarrow P[A \wedge B]$$

2) On considère ensuite la disjonction de l'exemple factorisé avec chaque règle de FG transformée de la même manière. En utilisant la distributivité de \wedge par rapport à

∨ et la disjonction interne sur les arguments, on dégage les propriétés communes.

En supposant, par exemple que la règle considérée de FG est de la forme : $P/A \wedge R$ et l'exemple de la forme : $P/B \wedge E$, on aura la transformation suivante :

$$\{ P/A \wedge R \} \vee \{ P/B \wedge E \} \rightarrow P/A \vee B \wedge \{ R \vee E \}$$

A noter que A et/ou B peuvent être des conjonctions introduites par la factorisation préliminaire.

3) Des disjonctions d'arguments ainsi obtenues, on déduit des couples (formés d'un objet de FG et d'un objet de l'exemple considéré) qui sont ensuite évalués en fonction du nombre de propriétés puis de relations vérifiées. Cette valuation correspond à un degré de pertinence du couple considéré, nécessaire lorsque l'on ne dispose que d'exemples positifs. Seuls ceux dont la valuation est supérieure à une valeur fixée par l'utilisateur dans la BCD seront retenus.

4) Les couples restants peuvent être regroupés en classes d'équivalence, deux éléments d'une même classe vérifiant des propriétés et des relations déductibles les uns des autres ou compatibles, voire identiques. Pour cela, le système utilise les connaissances de la BCD indiquant les propriétés des relations (symétrie, transitivité, inclusion, etc.).

Exemple :

Si le couple $\langle A, B \rangle$ vérifie $P/A \wedge Q/B \wedge R(A, B)$

Si le couple $\langle C, D \rangle$ vérifie $P'/C \wedge Q'/D \wedge R(D, C)$

Si $P \Rightarrow P'$

Si R est symétrique

... alors, seul $\langle C, D \rangle$ sera retenu

puisque P' est plus générale que P.

5) Il peut être nécessaire selon l'application réalisée d'imposer des hypothèses sur le contenu de la forme généralisée. Ce modèle minimum sera décrit dans la BCD. Dans le cas où ce dernier est non vide, et après regroupement des couples liés par une relation, seuls ceux vérifiant les propriétés et/ou les relations contenues dans le modèle sont conservés.

6) A l'issue de cette série de filtres sur les couples, une forme généralisée est produite pour chaque regroupement obtenu. Elle est déterminée par la conjonction des propriétés et relations vérifiées par chacun des couples contenu dans le groupe.

A ce stade, chaque couple étant remplacé par une variable, si les propriétés de même prédicat générique ont une même spécification pour chaque élément du couple, le terme correspondant de la forme généralisée aura une structure identique. Dans le cas où les spécifications diffèrent, c'est leur disjonction qui spécifiera la propriété apparaissant dans la forme généralisée.

Ainsi, aucune information n'est perdue dans la généralisation, même si les propriétés diffèrent par leur spécification, chose que ne permettent pas certaines autres méthodes [Hayes-Roth 78, Guizol 85].

7) La généralisation par hiérarchie s'effectue en fin de traitement. On dispose pour cela dans la BCD d'autant de descriptions arborescentes des propriétés que de prédicats génériques, la spécificité des nœuds augmentant avec la profondeur. Les feuilles constituent en fait l'ensemble des valeurs possibles de la spécification d'une propriété dans les exemples de départ.

Par exemple, dans notre application, l'arbre décrivant les propriétés "phoneme", "contexte-gauche" ou "contexte-droit", est structuré selon la décomposition en traits acoustiques de Jakobson.

Cette généralisation va intervenir sur les propriétés dont la spécification est une disjonction. Après recherche du nœud de plus bas niveau, dont dépendent tous les éléments de la disjonction, elle s'opère de la façon suivante :

- Si ce nœud est la racine :

- si tous les identificateurs de feuilles sont présents dans la disjonction, la propriété, devenue non significative, est alors supprimée.
- dans le cas contraire la propriété demeure inchangée.

- Si ce nœud se situe en dessous de la racine, la disjonction est remplacée par l'identificateur affecté à ce nœud.

IV - CONCLUSION

Le système d'apprentissage que nous avons présenté constitue un outil très utile pour caractériser des concepts de façon automatique. En particulier, dans l'application que nous en faisons, il nous permet d'obtenir des règles décrivant des réalisations d'unités acoustiques ou phonétiques propres à un locuteur, nous dispensant ainsi de la laborieuse mise au point de règles ad-hoc.

Les temps de calcul sont assez conséquents, mais ce traitement devant être effectué une seule fois par locuteur, nous jugeons que cela ne constitue pas un réel problème et compense de toute manière le temps passé à déterminer les règles "à la main". D'autre part, le caractère systématique de la production des règles constitue un net progrès.

BIBLIOGRAPHIE

- Colmerauer A., Kanoui H., Van Caneghem M.
PROLOG : Bases théoriques et Développements actuels ; TSI, vol. 2, N° 4, pp 271-311, juin-juillet 1983
- Guizol J., Meloni H., Gispert J.
Inférence de règles d'adaptation au locuteur dans un système de R.A.P.C ; 14^{èmes} J.E.P., pp 315-318, Paris 10-13 juin 1985
- Hayes-Roth F., Mc Dermott J.
An Interference Matching Technique for Inducing Abstractions
Communications of the A.C.M, Vol. 21, N° 5, pp 401-410, 1978
- Meloni H., Bulot R.
Un système de traitement des connaissances pour le décodage acoustico-phonétique ; Symposium on Speech Recognition, Montréal 21-22 juillet 1986
- Michalski R. S.
Knowledge Acquisition Through Conceptual Clustering : A Theoretical Framework and an Algorithm for Partitioning Data into Conjunctive Concepts ; Policy Analysis and Information Systems, Vol. 4 N° 3, pp 219-244, 1980a
- Michalski R. S.
Inductive Learning as Rule-Guided Generalization and Conceptual Simplification of Symbolic Descriptions ; Workshop on Current Developments in Machine Learning, CMU, Pittsburgh, July 16-18, 1980b