

# バイナリー概念カバー率に基づく マルチクラス学習手法

1P-6

アルモアリム・フセイン、秋葉泰弘、山崎毅文、金田重郎  
NTT情報通信網研究所

## 1. はじめに

マルチクラス概念を事例から学習する手法として、バイナリー概念学習アルゴリズムを用いる手法が良く知られている。この従来手法では、ルールは各クラス毎に個別に学習され、テスト時にはこれらのルールが並列に利用される。この場合、テスト時に唯一のルールのみがマッチすれば問題はない。しかし、現実には、マッチするルールが皆無となったり、複数ルールが同時にマッチ(ルール競合)する問題点を有している。

上記問題点を回避するため、本稿では、バイナリー概念を学習する際に、事例のカバー率を保存しておき、テスト時にこれを利用してマルチクラスを判別する手法を提案する。更に、本手法を機械翻訳ルール学習に適用した結果を示す。

## 2. カバー率に基づくマルチクラス判別法

本提案の手法では、ルール学習時に事例のカバー率に関する統計情報を収集し、これを、テスト事例に対するマッチ時に利用する。以下、学習時とテスト時に分けて、処理の概要を示す。

### 2.1 学習時

本手法では、ルールCは、ある条件 $d_j$ のConjunction

$$C = d_1 \cap d_2 \cap \dots \cap d_j \cap \dots \cap d_n$$

で表現されるものとする。一般には、条件 $d_j$ の個数が多いほど、ルールCがカバーする領域が小さくなる。但し、本提案の手法では、ルールCを獲得する学習手法に制限は無く、どのようなバイナ

リー概念学習アルゴリズムを用いても良い。

ここで、条件 $d_j$ を追加した事により、ルールCのカバー範囲からはずれた負事例の、負事例全体に対する割合を $CN_j$ とする(本来、負事例は全てカバー範囲からはずれる事が望ましい)。また、条件 $d_j$ を追加した事によりカバーされなくなった正事例の正事例全体に対する割合を $CP_j$ とする(正事例は、本来、全てカバーされるべきである)。もし、ルールCが全く何の条件も持たなければ(この状態を $d_0$ で表わす)、全正事例はこの $d_0$ でカバーされ、全負事例が誤って $d_0$ でカバーされてしまう。本手法では、ルール学習時に、 $CP_j$ 、 $CN_j$ の値を記録しておかなければならない。

### 2.2 テスト時

本手法では、テスト事例に対して、 $d_1$ から始めて、 $d_2$ 、 $d_3$ というように、次々と条件を追加して照合を行う。追加して行くに従って、ルールCがカバーできる領域は小さくなり、マッチ困難となって行く。その結果、 $d_j$ でマッチしなくなったとする。この時、 $d_{r-1}$ までのルールの、正事例(学習事例)に対するカバー率 $w_p$ は、

$$w_p = 1 - (CP_1 + CP_2 + \dots + CP_{r-1})$$

であり、また、 $d_{r-1}$ までのルールが負事例に対するカバー率 $w_N$ は、

$$w_N = 1 - (CN_1 + CN_2 + \dots + CN_{r-1})$$

で与えられる。

これらから、テスト事例に対するルールの適合度 $\gamma$ を、例えば、

$$\gamma = w_p - w_N$$

A Multi-Class Learning Method Using Covering Ratio and Binary-Class Hypotheses.  
Hussein ALUMUALLIM\*, Yasuhiro AKIBA, Takefumi YAMAZAKI, and Shigeo KANEDA  
NTT Network Information Systems Laboratories,  
1-2356, Take, Yokosuka-shi, Kanagawa-ken, 238-03, JAPAN

\*)On leave from the Dept. of Information and Computer Science, King Fahd University of Petroleum & Minerals,  
Dhahran, Saudi Arabia

とする。この値が大きいほど、より良く「正事例をカバーし、負事例を排除した」ルールとマッチしている事になる。扱う問題が、マルチクラスであるから、ルール適合度 $\gamma$ を、各クラス毎に計算し、もっとも大きなルール適合度 $\gamma$ を持つルールをマッチ成功とする。

### 3. 応用例

本提案の手法を、著者らが研究を進めている翻訳事例（コーパス）からの翻訳ルールの学習[Alumallim93]に適用を試みた。翻訳ルールは、NTTが開発を進めている日英機械翻訳システムALT-J/Eが用いているルールであり、一つの日本語動詞に対して、複数の英語動詞を対応づける。翻訳ルールは、条件部に日本語ボタン、ルール実行部に英語ボタンを持つ。日本語ボタンは、一つの日本語動詞、格要素（「が格」、「を格」等）の主名詞が持つべき意味属性条件、および助詞表現から構成される。英語ボタンは、対応する英語動詞、及び、日本語文中の格要素との対応関係から構成される。「焼く」の翻訳ルールは、例えば、以下の様な形式を持つ。

```
IF                               THEN
J-Verb  ="焼く"                 Subj   = N1
N1(が格)=[人]                   E-Verb="bake"
N2(を格)=[パン] or [菓子]      Obj    = N2
```

```
IF                               THEN
J-Verb  ="焼く"                 Subj   = N1
N1(が格)=[人]                   E-Verb="roast"
N2(を格)=[肉]                   Obj    = N2
```

ここで、[パン]、[菓子]等は、意味属性であり、ALT-J/Eの場合には、約40万語の名詞が、約3000の意味属性に割り付けられている。現状のALT-J/Eでは、約1万5千個の翻訳ルールが利用されている。

現状の翻訳ルール学習研究では、翻訳ルールの左辺部を対象とし、学習アルゴリズムとして、Hausslerによる、Internal Disjunctive Conceptのためのバイナリー概念学習アルゴリズム[Haussler88]を利用している。これは、翻訳ルールの条件部の形式が、Internal Disjunctive Conceptとみなせるためである。但し、Hausslerアルゴリズムでは、ルールは、

常に正事例を満足する様に決定される。従って、ルール適合度 $\gamma = 1 - w_N$ であり、負事例のカバー率 $w_N$ が小さい程、すなわち、できるだけ、多くの負事例が正しく排除されたルールを優先する。

表1は、実際に6種類の和語動詞について、翻訳事例からの学習実験を行った結果である。学習アルゴリズムとしては、Hausslerアルゴリズムと、本提案のアルゴリズムを適用し、1事例のみをテスト事例として取り出す、cross-validationを行った。表1から分かるように、一部の動詞（応じる）で性能低下がみられるものの、平均で1%程度の性能向上をしめしている。

### 4. まとめ

学習時に事例のカバー率を保存しておく事により、バイナリー概念学習アルゴリズムを用いたマルチクラス概念学習アルゴリズムの性能を、向上させる手法を提案した。本手法によれば、従来手法におけるルール競合等の問題を解決できる。また、本手法を、翻訳事例からの翻訳ルール（英語動詞の決定ルール）学習に適用した結果、学習性能の向上を確認できた。本提案手法では、利用するバイナリー概念学習アルゴリズムに特段の制限はない。従って、本稿で示した、Hausslerによる概念学習アルゴリズム以外にも、種々の学習アルゴリズムに適用できると考えられる。

#### 【文献】

- [Alumiallim93]フセイン・アルモアリム他、「概念学習アルゴリズムの適用による意味属性体系を用いた翻訳ルールの学習」、93年度AI学会全大、15-3, 1993  
 [Haussler88]Haussler.D., 「Quantifying Inductive bias : AI Learning algorithms and Valiant's Learning Framework」, Artificial Intelligence, 26(2), 177-221, 1988

表1 翻訳ルール学習の結果

日本語動詞	誤り率 (%)	
	本手法	従来手法
使う	1.4	1.5
飲む	2	1.0
行う	6	6
応じる	1.3	1.0
焼く	7	7
解く	0	0