

## 不完全情報を含むフレームの階層化の一手法

5C-7

小泉和彦 大川剛直 馬場口登 手塚慶一

大阪大学工学部

## 1. まえがき

与えられた種々の情報を知識ベース化するためには、それらを体系的に整理することが要求される。それに対する1つの試みとして、フレーム形式で入力された観測情報から新たなクラスフレームを生成し、階層化を図る手法を提案した[1, 2]。その手法は、スロット値に着目して入力フレームを分類することにより階層を生成するものであるが、一般に与えられた情報は完全ではなく、欠落情報、不確定情報の存在が考えられる。フレーム形式の知識表現では、欠落情報、不確定情報は各々スロット値の未定義、スロット値の複数定義の形で出現する。本稿ではこのような不完全情報を含むフレームから階層を生成する手法について述べる。

## 2. 概念空間とフレームの抽象-具体関係

フレームは“属性-値”の形を成すスロットで構成されている。各スロットはそれぞれの属性、及び値によって1つの概念を表しており、フレームが表す概念はそれらのスロットが表す概念を同時に満たすものとして特徴づけることができる。そこでスロットの属性及び値によって決定される空間を考え、フレームの持つ概念をこの空間に対応させる。これをフレームFが表す概念空間CS(F)と呼ぶ。ここでフレームF1, F2にCS(F1) ⊂ CS(F2)なる包含関係が成立するならば、フレームF2はフレームF1より抽象的なフレームと考えることができる。

## 3. クラスフレームの生成法

クラス階層の生成は入力されたフレームを分類しながらクラスフレームを生成するプロセスと考えられる。ここでは分類対象となるフレーム集合から、より抽象的なクラスフレームを生成する手続きについて述べる[2]。

分類対象フレームと生成するクラスフレームの概念空間には包含関係が成立する。またフレームが表す概念はスロットの属性、及び値によって特徴付けられている。従ってクラスフレームは対象フレームが持つスロットから生成することができる。そこで分類対象フレームを構成するスロットの属性、値を参照し、以下に示す手続きに従って対象フレームを分類し、クラスフレームを生成する。

[クラスフレーム生成基本手続き]

(i) 分類基準となるフレームを生成する際に用いる属性Aを決定する。属性の決定方法は後述する。

(ii) Aのとりうる値Vi (i ≥ 0)を用いて、分類基準となるフレームF(A:Vi)を生成する。ここでF(A:Vi)は属性Aの値がViであることを記述した単一のスロットから構成されるフレームを表す。

(iii) 各F(A:Vi)に対し、分類対象フレーム集合の要素から

$$CS(F) \subset CS(F(A:Vi))$$

が成り立つフレームFを選択し、F(A:Vi)の下位フレームとして分類する。同時に、F(A:Vi)をこれらのフレーム集合のクラスフレームとして採用する。

以上の手続きは分類対象フレームの属性、値を用いて互いに素な概念を表すクラスフレームを生成できる場合にのみ適用可能であり、分類対象フレームが不完全な情報を含んでいる場合にはそのまま適用することができない。そこで次の2つの場合について手続きを拡張する。

## (1) 欠落情報の取扱い

入力として与えられる全てのフレームが同一の属性を用いて記述されているとは限らない。つまり特定のフレーム中の特定の属性について情報が欠落しており、値が得られない場合がある。ここではこのような属性を未定義属性と呼ぶ。

分類基準生成のために未定義属性が選択された場合、フレームの分類は、実際に値を持つ未定義属性以外の属性の値を比較し決定する。すなわち分類対象フレームが持つ未定義属性以外の全ての属性に対し、同じ値をより多く持つフレームが存在するクラスを分類対象フレームが属するクラスとする。そこで未定義属性を含むフレームからクラスフレームを生成するために、基本手続きを以下に示すように変更する。

(i) 分類対象フレームを未定義属性の有無に従って、次に示すフレームFi, Fj (i ≠ j) に分類する。

$$F_i \in FS(A:V_i), \quad F_j \in FS(A:nil)$$

ここで、FS(A:Vi)は、Viを値として持つ属性Aを含むフレーム集合、FS(A:nil)は未定義属性Aを含むフレーム集合を表す。

(ii) Fiを基本手続きに従って分類する。

(iii) Fjをもとに、FjにF(A:Vi)のスロットを追加したフレームFjiを生成する。これを分類対象フレーム集合に追加し、同時にFjを削除する。

(iv) Fjiを属性Aについて基本手続きに従って分類する。以下の分類ではFjiは他のフレームと同等に扱い分類する。

なお、クラス階層生成が全て終了した段階で、各jについてFjiから1つのフレームを選択し、それ以外は削除する。その際、各フレームの直接上位に位置するクラスフレームの概念空間を比較し、最も小さいものを選択する。

## (2) 不確定情報の取扱い

属性値が不確定で、1つのスロットが複数の値を持つ場合がある。この場合、同一属性を記述したスロットが表す概念の間に包含関係が成立する可能性がある。基本手続きは1つのスロットが1つの値を持つ場合、つまり同じ属性を記述したスロットが表す概念が互いに素となる場合にのみ適用可能である。そこで分類基準を生成するために不確定な値を持つ属性が選択された場合は、基本手続きを以下に示すように変更する。

(i) 分類対象フレームを参照し、

$$\bigcap_i CS(F(A:DV_i)) = \phi$$

を満足するフレームFiを求める。ここでDV\_iはフレームFiにおける属性Aの複数の値VijをORで結合したものである。

(ii) 分類対象フレーム中のFi以外のフレームから、

$CS(F(A:DVk)) \cap CS(F(A:DVl)) \neq \emptyset$   
 を満足するDVk(k≠l),及びDVkに対応するFkを求める。ここで単一のDVlに対してDVkがこの条件を満足するならば、DVkに対応するFkをFiと同クラスのフレームとして分類し、それらのクラスフレームとしてF(A:DVl)を採用する。また、複数のDVlに対してこの条件を満足するDVkに対応するFkには(ii')を適用する  
 (ii')各DVlについて、Fkが持つ属性Aの値のみをDVlに置換したフレームFkiを生成し、Fkiを各々Fiと同クラスとして分類する。以下(1)の(iv)以降と同様の方法で分類し、クラス階層生成終了後に単一のFkiのみを残し、それ以外を削除する。

#### 4. 分類属性の決定方法

クラスフレーム生成基本手続きの最初のステップで、分類基準を生成するための属性が決定すると、フレームは一意に分類される。従って生成されるクラス階層木は分類基準に用いる属性の使用順序に大きく依存する。本手法では、より好ましい階層生成結果として希薄度(sparseness) [3]の低い概念による分類を生成する。希薄度の低い概念とは以下の条件を満たすものである。

- ・より小さな概念空間に対応する概念
- ・より多くの入力フレームを下位フレームとして持つクラスフレームが表す概念

これらの条件を満たす分類を生成するために、属性選択の基準を以下のように定める[2]。

[選択基準]

- (1)より多くの分類対象フレームが持つ属性を優先する。
- (2)一つの属性に関して、その属性がとり得る値の種類がより少ない属性を優先する。
- (3)上位の分類に相関が強い分類基準を生成する属性を優先する。

#### 5. クラス階層生成アルゴリズム

クラス階層は以下の手続きで生成する。

(Step.1)属性の順序付け

分類対象となる全フレームを参照することにより、対象フレーム集合がスロットとして記述している全属性を用いて属性リストを生成する。各属性について、その属性に関して記述されたスロット数、そのスロットを持つフレーム名、そのスロットに記述された値の種類を求め、これに基づいて選択基準(1)、(2)の順に従い属性リスト中の属性を順序づける。

(Step.2)属性値に基づく分類

3. で述べた手続きに従いフレームを分類し、クラスフレームを生成する。同一基準で分類されたフレームを同クラスに属するフレームとしてまとめる。また、ここで分類に使用した属性を属性リストから削除する。

(Step.3)値の相関に従った分類

(Step.2)で生成した各クラスに属するフレーム集合について、属性リスト中の属性を順次用いてさらに下位へのクラス分けを行う。各属性を使用した場合について選択基準(3)を用いることにより、最適な属性を1つ選択する。基準(3)による評価が同一である分類を生成する属性が複数存在する場合は、選択基準(1)、(2)に従い、属性の優先度リストの上位にある属性を採用する。同時に、採用した属性を(Step.2)と同様に属性リストから削除する。

(Step.4)クラス階層木の拡張

(Step.3)で得られたクラス階層木の各ノードに位置するフレーム集合に対して(Step.3)の操作を繰り返し、属性リストが空になるかあるいは分類対象となるフレームが単数となった枝は階層木の拡張を停止する。

なお、図1に本アルゴリズムの適用例を示す。ここでは

入力フレームとして造園に用いられる植物に関する特徴を記述したものを与えている。

#### 6. まとめ

不完全なスロット値を含むフレームを取り扱うことのできるフレームのクラス階層生成法について述べた。本手法では、単一の属性に着目してフレームを分類し、クラスフレームを生成する手続きを基本とする。未定義属性を含むフレーム、及び複数值を持つ属性を含むフレームに対しては、着目点を他の複数の属性に移行し、クラスフレームの生成に利用し得る情報の拡張を図っていることが本手法の特徴である。

[参考文献]

- [1]小泉, 大川, 馬場口, 手塚: “フレームのクラス階層生成についての基礎的考察”, 信学会春季全大, D-389, pp. 109 (1989).
- [2]小泉, 大川, 馬場口, 手塚: “クラスフレーム生成によるフレームの階層化”, 情処知識工学と人工知能研究会資料, A165-7 (1989).
- [3]R.S.Michalski and R.E.Stepp: “Learning from Observation: Conceptual Clustering”, Machine Learning, Vol. 1, Morgan Kaufmann, pp. 331-364 (1983).

```
(PLANT1 (FRAME TYPE (INSTANCE))
  (REAF_STATE (EVER_GREEN))
  (HEIGHT (HIGH))
  (NATURAL_DISTRIBUTION (TYPE_E1))
  (USE (MAIN) (SHRUBBERY) (LINE))
  (LIGHT_REQUIREMENT (MEDIUM)) (MOISTURE (MOIST))
  (SOIL (MEDIUM)) (GROWING_SPEED (FAST))
  (ROOT_DEPTH (DEEP)) (TRANSPLANTATION (EASY))
  (SALT_RESISTANCE (MEDIUM))
  (AIR_POLLUTION_RESISTANCE (MEDIUM))
  (TREE_SHAPE (TYPE_C2)))
```

図1 (a) 入力フレームの一部

```
(CLASS1 (FRAME TYPE (CLASS))
  (CONSIST_OF (CLASS2) (PLANT3))
  (REAF_STATE (EVER_GREEN)) (HEIGHT (HIGH)))
(CLASS2 (FRAME TYPE (CLASS))
  (CONSIST_OF (CLASS3) (CLASS4))
  (MOISTURE (MOIST)))
(CLASS3 (FRAME TYPE (CLASS))
  (CONSIST_OF (PLANT2) (PLANT5))
  (PRUNING (MEDIUM)) (GROWING_SPEED (FAST))
  (TRANSPLANTATION (EASY))
  (NATURAL_DISTRIBUTION (TYPE_C2)) (SOIL (SANDY))
  (SALT_RESISTANCE (WEAK))
  (AIR_POLLUTION_RESISTANCE (WEAK))
  (USE (SHRUBBERY)))
(CLASS4 (FRAME TYPE (CLASS))
  (CONSIST_OF (CLASS5) (PLANT4))
  (PRUNING (STRONG)) (GROWING_SPEED (FAST)))
(CLASS5 (FRAME TYPE (CLASS))
  (CONSIST_OF (PLANT1) (PLANT6))
  (LIGHT_REQUIREMENT (MEDIUM)) (SOIL (MEDIUM))
  (ROOT_DEPTH (DEEP)))
```

図1 (b) 生成されたクラスフレーム

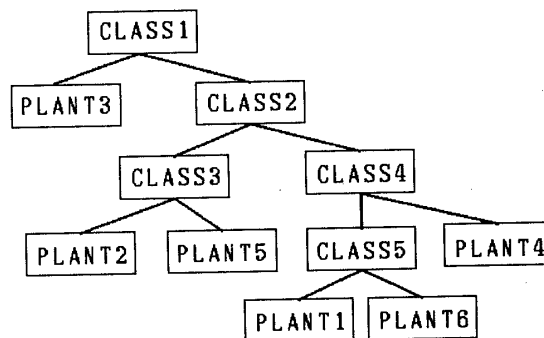


図1 (c) 生成されたクラス階層