

The Model to construct a Conceptual Structure

Tomoaki HIRANO †

Yuzuru FIJIWARA

## Abstract

Semantic processing is one of most important problems in various fields such as database, natural language processing, artificial intelligence etc, and many knowledge representation methods have been proposed to solve it.

In this paper, the focus is meaning of concepts especially description and structure. A model is proposed based on properties of concept and hierarchical relationship between concepts for representing meanings of concepts.

The model is suitable for dynamic properties of concepts, and is useful for construction of conceptual structure.

## 1 はじめに

人工知能の実現に当たって、知識表現の方法の確立は一つの重要なテーマである。これまでは、意味ネットワーク、フレームによる表現、意味素を基にしたオントロジー的なアプローチなどが提案されている。

知識は、概念間の関係により記述されるが、その意味とは単に記述ではなく、多くの他の事柄と関係づけられたその構造全体によって表現される。また、そこに得られた構造を用いることで、帰納推論、仮説推論、類推などの諸機能の実現が図れる。

知識表現を記述するのに、その表現の要素として概念が用いられる。意味処理を行なうには、概念自身の持つ意味を記述する必要がある。意味を記述された概念を用いて知識（事象）を記述し、始めて意味処理が出来るのである。

概念を表現するモデルとしては、Michell の Version 空間法 [Michell 77] において示されるモデルなどがある。例えば、このモデルの中では、概念は属性集合の要素の直積として定義されている。このような定義は簡潔で分かりやすい定義ではあるが、

- 属性集合を始めから与える必要がある。
- 概念の持つ属性数が、どの概念に対しても同じになる。
- 属性や、属性値で表現されているものは概念として扱うことができない。

などの問題点があり、特に属性集合を始めに決める必要があるのは、概念の生成や獲得に対して大きな制限を加えている。

本研究では、概念構造構築のため、各種の情報、特に辞書や用語集を解析し得られた情報に基づき、概念の持つ意味を記述するためのモデルを提案し、そのモデルを用いて概念構造がどのように形成されるかについて考察する。

\*筑波大学 電子情報工学系

†Institute of Information Science and Electronics, University of TSUKUBA

## 2 概念モデル

概念構造を構築するにあたって、概念とは何かを考える。まず始めに、概念を以下のように定義する。

**概念** 一つのものとして認知される対象。

抽象的な対象は、多くの概念間に共通する一般性がその対象として扱われる。また、事象は概念の意味的な組合せによって表現される。

概念は、対象単位の認知内容であるとの性質から、

- 共通の性質を持つ下位概念の集合としての表現
- 概念の持つ共通の性質による表現

の二通りが可能である。概念を表現するには、両方からのアプローチが必要である。

よって、概念は、その概念を特徴づける他の概念の集合としての記述と、概念間に与えられる階層関係の記述によって表現される。

概念の表現モデルを以下に示す。

### 概念空間

可算無限個の概念の集合  $X = \{x_1, x_2, x_3, \dots, x_n, \dots\}$  で表現する。

### 概念表現

概念空間上において、概念  $x_i$  は、その概念を性質を表す概念の集合  $x_i(x_1^i, x_2^i, \dots, x_n^i)$  で表現される。

### 階層関係

階層関係  $\geq$  を次のように定義する。二つの概念  $x_i(x_1^i, x_2^i, \dots, x_n^i)$  と  $x_j(x_1^j, x_2^j, \dots, x_m^j)$  の間の階層関係  $x_i \geq x_j$  が成り立つとは、すべての  $x_p^i$  に対して、 $x_p^i \geq x_q^j$  となる  $x_q^j$  が存在することとする。

概念間の階層関係は、「下位概念が、上位概念に比べ意味的な制約を明らかに多く受けている」ことが記述される。意味的な制約にはより特化された概念を性質として持つ場合と、上位概念にない性質を下位概念が持つ場合、の二通りが考えられる。概念のもつすべての性質に対してこれらの関係が成り立つ時に、階層関係が成り立つと言える。

上記の定義は、「下位概念が、上位概念の持つどの性質概念に対しても、その下位概念となるような性質概念を持っている時のみ、階層関係がある」と定義している。

現実世界においては、概念の説明として得られる記述は、その概念に最も近い概念を用いて、それとの差分記述で表現される。このモデルにおいて、それらの近い概念の階層関係は、下位概念の記述が、上位概念の持つ性質に他の性質が一つ加わった場合か、上位概念の持つ一つの性質概念が、それに最も近い下位概念に置き換わっている場合の二通りの記述によって行なわれる。

また、この階層関係は、反射律、推移律、反対称律を性質として持つ半順序関係であるとする。

**概念構造**

概念間に記述された階層関係を基に、概念空間全体に記述された構造。この構造を基に、概念の持つ意味を表現し、意味処理を行なう。

この概念モデルでは、概念は他の概念集合によって表現され、階層関係も性質概念に依存する。概念が概念によって記述されるはずであるが、他の概念を表現しているような基本的概念については、自分自身は他から表現されていない場合がある。このような、他の概念によって表現されない概念を概念プリミティブとし、また、そういった概念からなる階層を階層プリミティブとする。

概念構造は、多くの場合がプリミティブの組合せにより記述されるため、これらのプリミティブな部分での構造の記述の精度が、その構造全体の良否を決める一つの鍵となる。

なお、プリミティブな概念は絶対的な存在ではなく、対象と目的などにより変化する動的な存在である。その詳細は、別の機会に譲る。

また、概念の持つ他の性質として、例えば職業という概念は人間という概念以下の概念に対してのみ属性となり得るように、概念が他の概念の性質となる場合には、過度に一般化された概念に対しては性質として意味をなさなくなるという性質がある。

概念構造は、意味的に近い概念によって記述されたいくつもの概念構造を、概念の持つ性質の記述という形で組み合わせていくことにより形成される。概念構造がより良い概念構造となるには、部分的な概念構造の精度が全体としての精度に大きく依存してくる。

次に、概念モデルによる概念構造の簡単な例を図1によって示す。

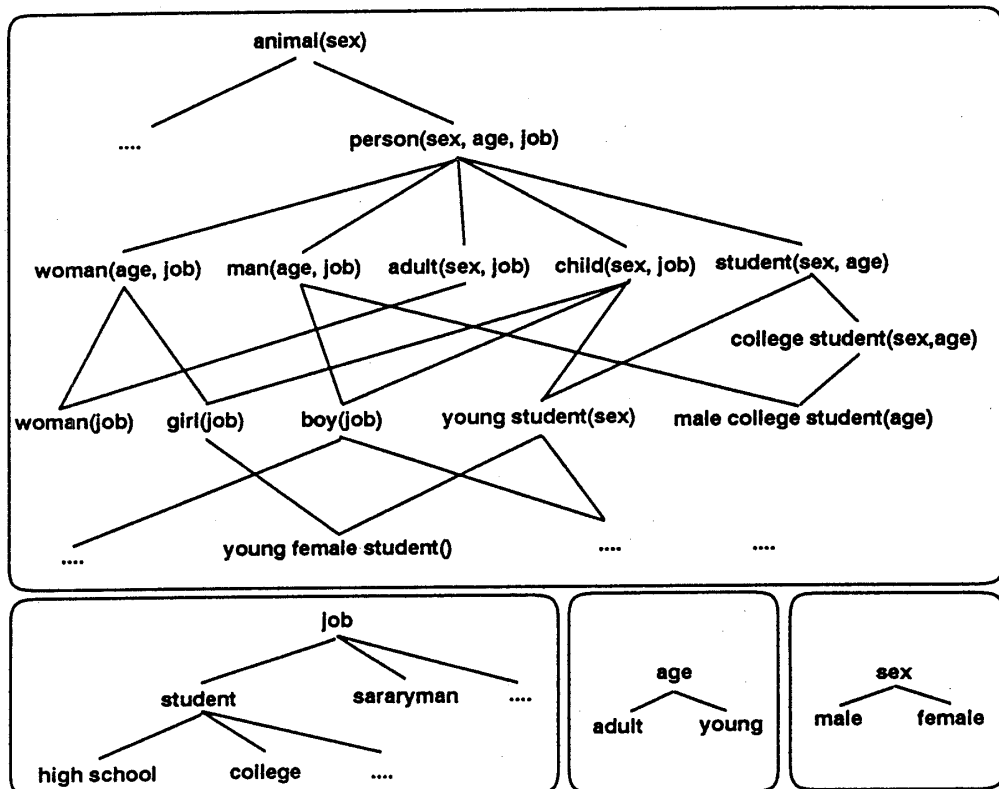


図 1: 概念構造の例

この例では、sex、age、job、の三つのプリミティブな概念構造を基に、それら三つの概念を属性として持ち得る概念 person と、その下位概念からなる概念構造の一部が示されている。対象として得られた概念は、それ自身に呼び名を持つものもあれば、性質の組合せにより表現されるものもある。

ここで、ラベルについて簡単にふれておく。概念とは、一つの対象として扱われるものであるから、「背が高くてカッコいい男の人」は一つ概念である。しかし、これを指し示す一つの名前は存在せず、複数の名前の組合せで一つ概念を表現されている。ラベルとは、概念に与えられた名前のことであるが、それはすべての概念に与えられるわけではない。

### 3 概念モデルの利用方法

以下の機能の実現に、概念モデルと概念構造が利用可能である。

#### 概念構造の自動構築問題

概念構造を自動的に構築する問題は重要な問題である。

概念モデルを用いて、概念構造を自動的に構築する問題がどのように表現されるのか、その枠組を与える。

システムに与えられるデータとしては、概念表現の断片であるとする。それらは、以下に示すようなデータになると考えられる。

1.  $x_i \geq x_j$
2.  $x_i(\dots x_n^i \dots)$
3.  $?( \dots x_n \dots x_m \dots)$

これらの断片的なデータを基に、概念構造を構築することを概念構造の自動構築問題とする。

ここで問題となるのは、与えられるこれらのデータは断片的であることから、そのまま概念構造の中に記述するわけにはいかないことである。

例えば、 $x_i \geq x_j$  が情報として得られたとしても、その間にはいくつもの概念が  $x_i \geq \dots \geq x_j$  のように隠されているかも知れない、また、 $x_i(\dots x_n^i \dots)$  のように概念  $x_i$  を  $x_n^i$  が説明しているように見えても、ある事例に限った話で、本来はより一般的な概念  $x_m^i (\geq x_n^i)$  が概念  $x_i$  を説明している、といったことも考えられる。

概念構造は概念の持つ一般的な性質と、それを基にした階層関係の記述であるので、まず、与えられたデータがその概念において必ず成立することなのかを判定する必要がある。もし、必ず言えるならば、その記述は概念構造の中に記述できるが、そうでなければ、得られた記述が概念構造中の何を表現しているかを推論しなければならない。

この問題は、機能推論を行ない概念の持つ性質を得ることであると言える。多くの高度な推論機能は、概念構造を基に行なわれることから、この問題は解決されるべき重要な問題である。

### 類似性の検出

類似性とは、複数の概念が共通に持っている性質であると言えるので、概念構造をたどり共通の上位概念の持つ性質を見ることで、概念間の類似の意味と程度を得ることができる。

**対象概念の生成** 意味処理の対象となる概念は、概念として一つの名前を持っていることは少なく、いくつかの名前を持つ概念の組合せによって表現され、対象となる概念の性質が明確化される。

概念構造内で、概念はそれが持つ性質によって記述されているため、いくつかの概念の組合せで表現された場合には、それらの性質の積によって対象概念の性質が表現できる。

### 概念の検索

概念を検索する時には、基本的にはその概念を指し示す言葉によって検索する。しかし、概念によっては特定の言葉で表されていないものがある。

こういった概念を検索する際には、記述された複数の性質に沿って概念構造を辿ることにより、期待する概念を検索することができる。

### 帰納推論

帰納推論とは、現実の多くの正しい事象が記述されている時に、この中に共通で、本質的な事象は何かを推論することである。

現実の事象は、本質的な事象の記述より、特化されたまたは一般化された概念を用いて記述されている。これらの記述の中から、どの概念を使って記述することが本質的な事象の記述なのかを見出す。

この時、多くの概念に共通する概念を見出すのに、概念構造が必要である。概念構造を用いることによって必要かつ十分な概念を見つけ出すことができ、もし、そういった概念がなければ、概念を作ることで概念構造自体がより有効なものになる。

## 4 考察

今回提案したモデルは、概念構造の構築を目的とし、その構造の持つ性質を表現するに耐えうるモデルとして構築した。

モデルに対する第一の要求は、概念構造の表現は階層関係を基にして表現すべきであることであり、そのためには概念間に階層関係を定義することで行なうことができる。

次の要求は、過去に提案されたモデルに存在する問題点への対応のためとも言えるが、概念として使えないような対象をなくすことである。過去のモデルにおいては、概念の表現が定義に合わないものや、属性値などとして与えられたものは、概念として扱われない場合があった。概念構造を構築する際には、考えられる概念のみを考えるのではなく、どのような概念でもそれを取り込んで同様に扱えるモデルである必要がある。

さらに、概念の持つ意味とは、その概念の持ついろいろな性質から記述できるので、概念の持つ性質に基づく概念の表現と概念間の階層関係を記述することで概念構造を表現した。

今回提案したモデルは、これらの要求を満たすものである。提案したモデルによって概念構造を表現することで、概念構造の持つ性質や、概念構造に対するいろいろな手続きの定義、概念構造の評価などを行なうことができるようになった。

本研究と関連したものに、概念全体に構造を持たせ、意味処理を行なえるようにするとのアプローチを行なった「EDR電子化辞書」がある。EDR電子化辞書には大規模であり多数の種類辞書が存在し、しかも計算機的な処理を行なうための表現がなされている。そのEDR電子化辞書のうちでも、特に概念辞書において、概念間に主要な20個程度の関係を表現し、それらが意味処理を行なうための下地として用意されている。

今回提案したモデルは、概念辞書の持つ階層構造のモデルに比べ、概念の意味がより表現されている。今回提案したモデルを用いて、EDR電子化辞書の概念辞書の持つ構造の評価を行なうこともできる。

また、EDR電子化辞書における概念間の関係は、その抽出に際し構造構築のための基準が明確でないこと、さらに多数の手入力されていることのために、データの欠落や偏り、不統一がある。そういった構造中に表現されていない概念の発見や矛盾の検出にもモデルは有用である。

## 5 結論と今後の課題

本研究では、概念構造を構築するため、概念の持つ意味の記述を、それが持つ性質の集合という観点から行ない、そこに表現される階層関係を示した。

構築されるべき概念構造の表現法、性質などが明らかになったので、今後の課題としては、概念構造を自動的に構築するアルゴリズムの確立、概念構造を用いた各種高度な推論機構の方法論の提示、さらには、概念の意味表現だけでなく、概念間の関係を基に事象を表現するモデルの確立などが挙げられるであろう。

## 参考文献

- [1] 宇陀 則彦、張 曉冬、藤原 謙  
“情報の自己組織化のための意味構造モデル”  
情報処理学会 第45回全国大会講演論文集 1992
- [2] N.Uda and W.G.lee and Y.Fujiwara.  
“Constraction of Semantic Structure in the Self-Organizing Information-Base Systems.” Journal of Japan Society of Information and KnowLedge. Vol.3, No1.1. 1994.
- [3] “EDR電子化辞書仕様説明書”  
日本電子化辞書研究所(1993)
- [4] Michell, T.M  
Version spaces : A Candidate Elimination Approach to Rule Learning  
Proc. 5th IJCAI
- [5] 志村 正道  
人工知能：基礎情報工学シリーズ17  
森北出版(1994)