

# フレーム構造に基づく概念タグを利用した知識表現 複層意味フレーム分析からの知識構築を目指して

野澤元<sup>\*1</sup>

Nozawa, Hajime

<sup>\*1</sup> 独立行政法人情報通信研究機構

National Institute of Information and Communications Technology

This study attempts to represent human knowledge with semantic frame tags which are used in Multilayered Semantic Frame Analysis (Kuroda *et al.* 2005) for describing event meanings expressed in sentences.

## 1. はじめに

本研究は、複層意味フレーム分析(黒田他 2005)において、文の意味記述に用いられる意味フレームタグを利用して、世界に関する知識を表現する試みである。文の意味記述と知識表現が、概念フレームという共通の形式を持つことにより、より容易に意味タグ付きコーパスから知識を抽出することができるかと期待される。

## 2. 知識のデータベースとその構築

### 2.1 言語資料からの自動構築

世界に関する知識のデータベースを作成する場合、最も容易に利用できる情報源は、人類が歴史を通して蓄積してきた膨大な言語資料であろう。もし、テキストから自動的に知識を抽出できるのであれば、非常に低いコストで大規模な知識のデータベースを構築することができる。しかし、現状ではこのような試みにはいくつもの大きな問題がある。一つ目は、言語表現の意味を分析する手法であり、二つ目は、知識を表現するための汎用的な形式である。

### 2.2 意味分析

言語表現の意味を解析する技術はまだまだ発展途上である。これまで多くの場合、言語表現の統語構造に基づいて、語彙や形態素の意味を結合し、全体の意味を再構築する手法が試みられてきた。しかし、統語構造は伝達上の機能や制約のために、必ずしも意味の構造を正確に反映しておらず、そのような手法には限界があると考えられる。

これに対して黒田他(2005)は、Berkeley FrameNet (Baker *et al.* 1998)を参考にする形で、複層意味フレーム分析 (Multilayered Semantic Frame Analysis: 以下、MSFA と略記) という手法を提案した。MSFA では、文が表現する事態水準の意味を意味フレーム(Fillmore 1982)として記述し、また、文を構成する語彙や形態素の意味には、それぞれの意味フレームにおける役割を付与することで、統語構造に頼らずに、それらの間の関係を規定している。現在、最終的には分析の自動化を目指して、機械学習の基礎データとなる意味フレームタグ付きコーパスを構築中である。

### 2.3 知識表現の形式

筆者は知識ベースの分野の専門家ではないため、明確なことを述べることはできないが、知識表現の形式には XML を初めとして様々なものがあるが、それらの形式は

基本的には情報を分類することを主な目的として設計されているように思われる。しかし、これらの形式は、世の中で生じる事態や存在する物体の構造といった、世界に関する基本的な知識を記述するには適していないように思われる。また、言語資料から知識を抽出する場合、知識表現の形式は、意味分析の結果としての意味記述の形式と互換性があることが望ましい。もし、意味記述と知識表現の形式が共通であれば、個々の意味記述の間に関連性を与えるだけで、大規模な知識表現を得ることができるはずである。

## 3. 概念フレーム

### 3.1 概念フレームによる知識表現

本研究では、知識表現の形式として概念フレームを提案する。概念フレームとは、第一に現象の全体と部分の構成性を表現するフレーム構造を持ち、第二にフレームの範囲や構成として人間が状況を認識する際の時空間的範囲や際立ちを利用した知識表現の形式である。このような概念フレームの規定は、人間の知識を表現するには、人間の認識の構造と類似した性質を持つ形式が好ましいという思想に基づいている。また、MSFA における意味フレームタグは概念フレームの実装の一つと見なしうる。

### 3.2 フレーム構造

人間はある状況の認識において、状況を一つのまとまりのある現象として捉えるとともに、同時にその現象の構成要素を認識することができる。フレーム構造とは、このようなまとまりを持った一つの現象に対する部分の階層的構成性を表現する一般的構造である。ここでは、フレームを以下のように定義する。

- フレーム構造: 全体としてのある現象と、その部分である要素、そして、それらの要素間の関係を規定する一般的構造。

このようなフレームの定義は、基本的には Minsky (1975) によるフレームの定義と矛盾しないものであると考えられる。

### 3.3 認識の範囲と際立ち

概念フレームでは、フレームの範囲や構成に経験的に得られる人間の認識の構造を適用している。例えば、ピアノを演奏している状況にある認知主体が認識する場合、ピアニストとピアノと音楽を焦点化し、それらの要素を参与体

とする概念フレームを用いることがあるだろう。しかし、別の時には、認知主体は視点をより接近させて、指と鍵盤と音を参与体とする、より時空間的範囲の小さなフレームを用いて状況を認識するかもしれない。このように人間の認識は、恐らく何らかの適応的価値が存在すると思われる、特定の時空間的範囲や際立ちを持っている。知識の記述では、このような認識の構造が概念フレームの規定として利用される。

ピアノを演奏している状況の認識としての、上記のような二種類の概念フレームは、同一の状況に対する異なる解釈であり、基本的には異なる概念である。その意味では概念フレームを単位とする人間の認識は、非常に局所的である。しかし、それぞれの概念フレームの間には、様々な関連性が存在すると考えられる。このような概念フレーム間の関連性は、視点の転換や推論の指標となり、その結果、全体としての大域的な知識を実現していると考えられる。つまり、人間が持つ知識は、階層的に関係付けられた無数の概念フレームの体系だと言える。

### 3.4 概念フレームと意味フレームタグ

MSFA における意味フレームタグは、概念フレームの一つの実現形である。MSFA による意味フレームタグ付きコーパスの構築は、まだパイロット段階であるため、タグの明確な仕様は決まっていないが、基本的には、ある事態を表す意味フレームのフレーム名と、フレーム要素の役割名から構成される。本研究では概念フレームとして、意味フレームタグを再定式化し、一部拡張した次のような形式を提案する。

- (1) 事態を認識対象とする場合の基本形式  
(関係名 0 [役割名 1 <物体名 1>], [役割名 2 <物体名 2>])
- (2) 物体を認識対象とする場合の基本形式  
<物体名 0 (関係名 1 [役割名 1-1 <物体名 1-1>], [役割名 1-2 <物体名 1-2>])>

(1)は事態に関する知識を記述するための概念フレームであり、その形式は意味フレームタグと基本的には変わらない。関係名とはフレーム名と同じであり、事態の名称である。(2)は、物体に関する知識を記述するための概念フレームであり、認識された物体とその構成要素、そして、構成要素間の関係性を表現している。

### 4. 適用事例

人がたも網で魚を捕まえるという状況についての知識は、以下のような一連の概念フレームによって表現される。

- (3) 人がたも網で魚を捕らえる。  
(捕獲 [捕獲体 <人>], [被捕獲体 <魚>], [道具 <たも網>])

この状況に対する最も一般的な認識は、(3)のような概念フレームであろう。しかし、この概念フレームだけでは、この状況における人とたも網の関係や、たも網と魚の関係は詳細にはわからない。認知主体が(3)の概念を想起したときに、それらの関係を理解することができるためには、構成的階層性においてより下位の(4)や(5)のような概念フレームとの関連性が規定されている必要がある。

- (4) 人がたも網を持つ。  
(把持 [把持体 <人>], [被把持体 <たも網>])
- (5) たも網に魚が入る。  
(包囲 [包囲体 <たも網>], [被包囲体 <魚>])

(4)と(5)は、(3)と同じ状況に対する、異なる時空間的範囲と際立ちを持った、異なる解釈としての認識である。これらの概念フレーム間に関連性を与えることで、(3)は(4)と(5)の複合体と見なすことができる。

また、魚を捕まえる時のたも網の使用法を理解するためには、たも網の構造と機能についての知識が必要である。たも網のどの部分をどのように使うのかについて理解するためには、(4)と(5)は、それぞれ(6)と(7)のようなより詳細な概念フレームと関連付けられる必要がある。

- (6) 人が柄を持つ。  
(把持 [把持体 <人>], [被把持体 <柄>])
- (7) 網袋に魚が入る。  
(包囲 [包囲体 <網袋>], [被包囲体 <魚>])

ただし、(6)と(7)は柄や網袋がたも網の一部であることを表現していないため、このままではこれらは最終的な(3)の概念フレームとは無関係な概念フレームになってしまう。この問題を回避するには、(4)と(5)におけるたも網が、物体であるたも網の認識としての(8)のような概念フレームと関連付けられ、その上で(8)の柄や網袋が(6)と(7)と関連付けられる必要がある。

- (8) たも網  
<たも網 (付属 [基部 <棒>], [付属部 <網袋>]), (付属 [基部 <棒>], [付属部 <柄>])>

ちなみに、(8)はこの場合(9)と(10)の複合体として規定されている。

- (9) 棒に網袋が付いている。  
(付属 [基部 <棒>], [付属部 <網袋>])
- (10) 棒に柄が付いている。  
(付属 [基部 <棒>], [付属部 <柄>])

これらの概念フレームは、当該状況に関する知識のまだほんの一部を表現しているに過ぎないが、言語資料を参考として、存在する概念フレームと、概念フレーム間の関連性を記述することで、より人間の知識構造に近い知識表現を得ることができると思われる。

### 参考文献

- [1] Baker, Collin F. et al.: "The Berkeley FrameNet project", *Proceedings of the COLING-ACL*, 1998.
- [2] Fillmore, Charles J.: "Frame Semantics", *Linguistics in the Morning Calm*, Hanshin Publishing Co., 1982.
- [3] 黒田航・他: 複層意味フレーム分析の紹介, 人工知能学会 19 回大会論文集, 2005.
- [4] Minsky, Marvin: "A Framework for Representing Knowledge", *The Psychology of Computer Vision*, McGraw-Hill, 1975.