

情報知識学会 2006 年度第 2 回研究会報告

日 時：9 月 23 日（土）14 時～16 時半
会 場：大阪市立浪速人権文化センター
テーマ：シソーラスとクラス階層と RDF/OWL
発表者：神崎正英氏（www.kanzaki.com）
共 催：日本図書館研究会整理技術研究グループ
出 席：24 名
内 容：

事前に連絡していた「主題表現としてのクラス階層とシソーラス」というテーマを上記のテーマに若干変更の上、以下の発表がなされた。

1. 「クラス階層」の考え方

・認識対象を、異なる性質に注目して分けたり、共有する性質に着目してグループ化したりして生成されるのが、一般的な「分類としてのクラス」である。また、オブジェクト指向言語に典型的な「型としてのクラス」（インスタンスを生成するための「型」）という視点もある。

・クラスの定義には、メンバー条件を記述する内包的定義と、メンバーを列挙する外延的定義がある。

・分類一般の階層は is-a、part-of、instance-of など多様な関係を含みうるが、「クラス階層」は is-a 関係（種類関係）のみで構築される必要がある。クラス階層は、(1)各クラスのインスタンス集合は上位クラスのインスタンス集合の部分集合となる、(2)上位クラスの性質は推移律に従って下位クラスに順次継承されていく、という性質を持つ。

2. RDF/OWL におけるクラス

・「主語、述語、目的語」のセットから構成される RDF 記述は、リソース間の関係を示すという側面と、主語にあたるリソースの属性を表現するという側面を持つ。OWL などオントロジーの視点からは後者が重視される。

・OWL におけるクラスは記述論理（DL）の Concept に相当し、名前付けのほか、インスタンス集合の論理演算やプロパティ制約によっても定義される。適切に構築されれば記述論理の推論が利用でき、包含関係の推論やオントロジーの整合性検証、個体の自動分類などの可能性が開ける。

・OWL のプロパティには数種類の型定義（推移型など）が可能で、またプロパティ間の関係も設定できる。これによってプロパティを用いた推論や他の語彙とのマッピングが可能である。

・プロパティを用いて、クラスの制約条件（とりうる値や数を）が設定できる。ただし、Web の開かれた世界を想定すると、クラスの全メンバーを拘束する強い制約の設定には十分な注意が必要である。

・記述論理に忠実な OWL DL では、クラスとプロパティを定義する「TBox」の世界と、個体（インスタンス）を記述する「ABox」の世界を峻別することが求められる。同じ URI をクラスとインスタンスの両方に使うことはできない。また、プロパティは（クラス間ではなく）個体間の 2 項関係を表現する。

3. 件名標目・シソーラスとクラス階層

・シソーラス型の階層構造を持つ件名標目表は、一見クラス階層に近いように見えるが、OWL DL を想定するといくつかの問題をはらんでいる。例えば、階層構造（BT, NT）中に固有名詞が現れることがあるが、固有名詞は個体であって本来クラスと同列には扱えない。また OWL DL では、subClassOf などの「メタ・プロパティ」に限りクラスを主語・目的語にできるが、通常のプロパティはインスタンス間の関係記述に用いられる。RT などのクラス間関係を直接扱うことはできず、クラスに対するプロパティ制約として記述しなくてはならない。

・したがって、ダブリンコアの「dc:subject」を考えたとき、このプロパティの目的語もクラスではなくインスタンスでなくてはならないが、主題を表す概念（クラス）に対するインスタンスとは何かという問題がある。

4. シソーラスの RDF 表現

・シソーラスの階層関係には属（種類）関係、階層的な部分全体関係、事例関係などが含まれ、推移律はなりたつものの必ずしも継承の関係ではない。さらに RT 関係等も考えると、すんなりと RDF/OWL に変換することはできない。

・既存シソーラスの資産を生かしながら RDF 表現する一つの戦略として、優先語（標目）をクラスではなく、Concept などのクラスのインスタンスとして位置づける方式が考えられる。インスタンス同士であれば様々な階層や関係もそのままマッピングできる。

・この方式でシソーラス等の概念スキームを記述する RDF モデルとして、SKOS（Simple Knowledge Organization System）が提唱され、W3C で草案段階にある。各概念をインスタンスとし、broader, narrower, related 等のプロパティを設定するもの

である。例えば、最近テキストデータでの研究用公開がはじまった NDLSH も、SKOS への変換は比較的容易にできる。

- ・辞書的なシソーラスである WordNet を RDF/OWL で表現しようとする試みも行われている。「語」を意味ごとに分割した WordSense、同一の意味を持つ WordSense をまとめた Synset、といった WordNet の独自の構造をそのまま RDF 表現するものである。

- ・RDF 表現を行う際には、各概念の URI をどう設定するのも、留意すべき問題である。ID 番号は明快だが使いにくく、標目を URI の一部に用いたほうがわかりやすくなる。ただし、件名標目表には指示や細目が組み込まれた独特の標目形も多く、問題が残る。

- ・クラス階層ならば RDF/OWL の組み込み伴意(推論)が使えるが、SKOS のようにインスタンスと位置づけた場合は別途クエリを工夫する必要がある。

当日の資料：<http://www.kanzaki.com/works/2006/pub/0923nal.html>