

# セマンティックウェブ上の情報可視化フレームワーク

朴 斌 伊藤 正彦 田中 譲

北海道大学大学院情報科学研究科

コンピュータサイエンス専攻知識メディア研究室

## 1 はじめに

セマンティックウェブは、ウェブページにメタデータを導入することで、ウェブが「人間が理解可能なドキュメントの Web」から「機械が理解可能なデータの Web」へと変化をもたらした。近年、セマンティックウェブの発展により、メタデータの量が急激に増加しており、これらの膨大な量のメタデータをアクセスするには、SPARQL (SPARQL Protocol and RDF Query Language) が用いられることが多い。しかし、セマンティックウェブや SPARQL に関する深い知識を持ってないユーザが、複雑な検索要求文を書いて検索を行うことは困難である。

本論文では、SPARQL を用いて、セマンティックウェブ上の情報を三次元空間内で動的な操作により検索し、その検索の過程と結果を可視化する可視化フレームワークを提案する。本フレームワークでは、抽象化された RDF スキーマを作成し、それをスキーマレベルで可視化する、そして、その上で簡単な操作により、インスタンスの検索要求文を構築し、検索された結果をインスタンスレベルで可視化する。

## 2 可視化フレームワークの概要

ユーザにとって、複雑な SPARQL 検索要求文を記述し、目的インスタンスを検索し、さらに、そのインスタンスと関係あるほかのインスタンスを探し出すのは困難である。

この問題を解決するため、本フレームワークは RDF スキーマの可視化とインスタンス集合の可視化を組み合わせることでセマンティックウェブへのインタラクティブなアクセスを支援する。RDF スキーマの可視化部では入力されたキーワードに関する RDF スキーマの可視化をす

る。またインスタンス可視化部では RDF スキーマの上で、ノードを選択、条件を入力することでインスタンスの検索要求文を構築し、その検索結果のインスタンス集合を可視化する。

### 2.1 RDF スキーマの可視化

現在の Web には全てのコンセプトに関する RDF スキーマがあるとは限らない。本論文で提案するフレームワークではクラス中の一定量のインスタンスの全てのプロパティを抽出し、それに基づいて抽象化された RDF スキーマのグラフを作成し、さらに、動的な操作により RDF スキーマのグラフを展開していくことができる。

セマンティックウェブ上のリソースは全て URI で表記しているため、スタートポイントとして、クラスの名前を用いるのは困難である。そのため、本論文ではキーワードを入力することで、クラスのリストを求める。

[スタートポイントの検索要求文]

```
SELECT DISTINCT ?x
WHERE{?x rdfs:subClassOf ?object
      FILTER regex(str(?object),"キーワード入力")
}
```

そして、取得したクラスのリストからランダムに 10 個を選択し、各々のクラスのインスタンス全てのプロパティを抽出し、RDF スキーマを作成する。これによって得られるスキーマはクラスリストの一部から構築されたものであり、スキーマ全体を網羅していない。そのため、関心のあるノードがこの中に入らない場合には上の操作を繰り返すことで、足りないプロパティを追加していく方法を採用する。

### 2.2 インスタンスの可視化

本論文では、まず、RDF スキーマの可視化部分で作成された RDF スキーマのグラフを用いることでクラスとその間の関係を通じて関心を持つクラスを探す。そして、興

味のあるクラスを選択し、別のクラスに条件を課すことで、簡単にインスタンスの検索要求文を構築する。最後にインスタンスの検索結果を三次元空間で可視化する。

### 3 映画の検索例

実例として映画の例を示す。まずスタートポイントになるノードにキーワード“Movie”を入力し、検索を行うと、図1で示すように“Movie”を中心とする“Movie”のRDFスキーマが作成される。

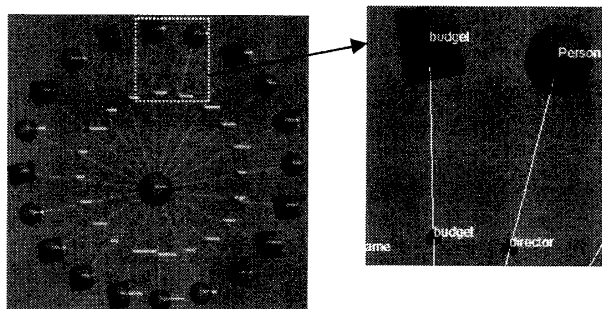


図1 映画のRDFスキーマ

しかし、ユーザにとって、これだけのスキーマ情報では足りない場合がある。例えば映画と関係ある主演、監督、国のスキーマも知りたいことがある。この場合は図1映画のスキーマ上でこれらに対応するノードをクリックすると、選択されたノードのスキーマが図2で示すように展開される。

作成されたRDFスキーマでノードを選択し、ノードに値の制約条件を入力することで、インスタンスの検索要求文を構築できる。例えば、2008年度に封切られたアメリ

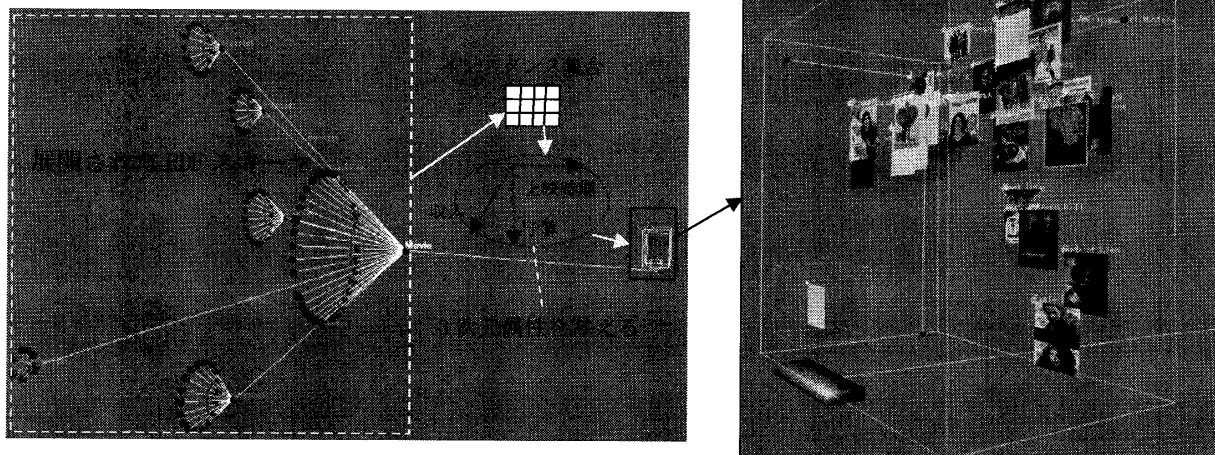


図2 展開されたRDFスキーマとインスタンス集合の可視化

カの全ての映画のタイトル、上映時間、予算、イメージ、収入を検索する場合はRDFスキーマ上で上映時期に2000を入力し、国の名前にアメリカを入力する。また映画のタイトル、イメージ、予算、収入を選択すると検索要求文が構築される。さらに、これを図2に示すように、検索結果に三次元の属性を与え、三次元空間で可視化することができる。

### 4 おわりに

本論文ではセマンティックウェブ上の情報の検索および可視化を行う新たなフレームワークを3次元知識メディア・アーキテクチャを基盤に用いて提案した。このフレームワークの特徴はRDFグラフの主語、述語、目的語の三つの要素を利用してユーザ自身による新たな検索要求文の構築を可視化結果の上で簡便に作成可能に点である。問題点としては膨大な量の情報により可視化が複雑になる可能性が考えられる。今後は、ユーザにとって分かりやすい可視化の形式について考察していきたい。

### 参考文献

- [1] 伊藤健太郎, 佐藤勇紀, 濱崎俊. ためしてわかるセマンティック Web ～次世代型データ活用術～. 株式会社技術評論社. 2007
- [2] SPARQL Query Language for RDF . <http://www.w3.org/TR/rdf-sparql-query/>