

拡張 SPARQL クエリのための 動的オントロジーマッチングとチューニング支援機構の試作

Toward a System of Dynamic Ontology Matchings and Tuning Supports for Extended SPARQL queries

足立拓也^{1*} 福田直樹²
Takuya Adachi¹ Naoki Fukuta²

¹ 静岡大学情報学部

¹ Faculty of Informatics, Shizuoka University

² 静岡大学大学院情報学領域

² Department of Informatics, Shizuoka University

Abstract: 本研究では、拡張 SPARQL クエリのための動的オントロジーマッチングとチューニング支援機構の実現のための試みとその課題について述べる。オントロジーマッチングを用いた SPARQL クエリ拡張である SPARQLoid クエリのコーディング時における不具合の原因調査およびその対処の困難さの改善のために、オントロジーマッチングおよびマッピング先の情報を必要に応じて提示するための機構と、動的オントロジーマッチングによる検索対象拡大・調整のための一時オントロジーマッチングに基づくクエリコードのチューニング支援機構の試作を行う。

1 はじめに

SPARQL は RDF 形式で記述する際に必要となる語彙の定義となるデータを検索するクエリ言語であり、RDF 形式で記述されたオントロジーを十分に理解しない状態では、効果的なクエリを記述することが容易ではない [Fujino 12a]。オントロジーマッピング [Atencia 12][Noy 09] を用いた SPARQL クエリを実行可能にするものに、SPARQLoid [Fujino 12a, Fujino 12b, Fujino 14] がある。SPARQLoid では、重み付きオントロジーマッピング [Atencia 12] を効果的に利用可能するために SPARQL クエリを拡張し既存の SPARQL エンドポイント上でも実行させることが可能な SPARQL クエリに変換する機構を備えるオントロジーマッピングを用いた拡張 SPARQL クエリを実行した結果が期待通りとはなかった場合使用したオントロジーマッピングやインスタンスの存在の多寡などといった様々な原因が考えられ、原因を特定することが難しい [藤野 13]。本研究では、オントロジーマッピングを用いた拡張 SPARQL クエリのコーディング時におけるオントロジーマッピングに起因した問題に対し、それを動的オントロジー

マッチングおよび記述パラメータチューニングの支援を行うことで、解決を試みる。

2 背景

本章では、既存の研究および技術である拡張 SPARQL クエリとその変換エンジンである SPARQLoid [Fujino 12a, Fujino 12b, Fujino 14] とその支援機構、および本研究で取り組む課題について述べる。

2.1 拡張 SPARQL クエリとその支援機構

オントロジーマッピングに基づいて SPARQL クエリを行う手法として SPARQLoid [Fujino 12a, Fujino 12b, Fujino 14] がある。SPARQLoid は重み付きオントロジーマッピング [Atencia 12] を利用するクエリを生成するために、既存の SPARQL の文法を拡張したクエリ言語、および、その変換エンジンのことを呼ぶ。拡張したクエリ言語の構文には、重み付きオントロジーマッピングに付与された確信度に基づくフィルタリングや並び替えが可能である。

拡張された SPARQL クエリである SPARQLoid クエリを記述することは、記述が SPARQL クエリより

*連絡先： 静岡大学
静岡県浜松市城北 3-5-1
E-mail: cs13007@s.inf.shizuoka.ac.jp

複雑になる部分があり、初めて SPARQL を知った人や SPARQL を利用する人にとって、それらの拡張構文の記述のみならず、適切な語彙の選択にも何らかの支援があることが望ましいと考えられる。これらの課題に対して、SPARQL クエリの語彙・エンドポイント検索支援の手法である SuPARQooL[Noguchi 13] と SuPARQooL を拡張した試作機構 [足立 16a] がある。試作機構において、幅広いエンドポイントに対して、語彙・エンドポイント推薦機構を行う際、処理に時間がかかってしまうことが課題であった。SPARQL Endpoint Status¹ から取得できるエンドポイントの特徴に基づく機械学習による手法 [Adachi 16b] 等を用いて処理時間の削減の実現が検討されている。また、ユーザがマッピングの詳細な情報を意識せずに、クエリを記述した場合にも、マッピングを効果的に利用可能なクエリへ拡張を行えるための機構 [足立 16b] が検討されており、使用するオントロジーマッピングに対して動的なオントロジーマッピングの生成を用いて、SPARQLoid クエリにおけるデバッグ支援を検討している [Adachi 16a]。

2.2 拡張 SPARQL クエリにおける課題

Listing 1 に、SPARQLoid クエリの例を示す。このクエリは、我々が現在準備を進めている木曽町オントロジーに基づく木曽町 LOD に対して、「yamada:子育てしやすい環境づくり」のインスタンスを取得するものである。このとき、SPARQLoid 内部では「yamada:子育てしやすい環境づくり」に対するマッピングの検索を行っている。「yamada:子育てしやすい環境づくり」に対するマッピングの検索結果を表 1 に示す。SPARQLoid では、拡張された SPARQL クエリを記述可能であり、これらの拡張されたクエリは新しい構文を含んでおり、オントロジーマッピングに付与された確信度を加味してクエリの動作を調整することができる一方で、対象となるオントロジーやマッピングの特性に関する知識が必要となる [足立 16a]。Listing 1 で示す SPARQLoid クエリは実行時に表 1 のマッピングを参照し、それらのマッピング先の要素を参照するように変換されて実行される。ところが、実際にこの SPARQLoid クエリをエンドポイントに対して実行した際、インスタンスを得られない。

先のクエリでインスタンスが得られない理由は、使用するオントロジーマッピングから参照されるクラスにはインスタンスが直接結びついておらず、そのオントロジーマッピングから参照されるクラスの下位クラスにインスタンスが結びついていた。エンドポイントが OWL-DL における推論等を扱えない設定になっている場合、このようなインスタンスの検索が成功しな

い場合がある。クエリ実行対象となるエンドポイントの設定を変更できない場合、オントロジーマッピングの情報および SPARQLoid クエリ中の記述のみからこれらの問題を解決可能となることが望ましく、そのための支援機構が必要となると考えられる。

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX yamada: <http://www.semanticweb.org/cs13098/ontologies/2016/4/kiso#>

SELECT DISTINCT ?s
WHERE {
  ?s rdf:type yamada:子育てしやすい環境づくり

  THRESHOLD { yamada:子育てしやすい環境づくり
    >= 0.48 }
  CRITERIA ?c { yamada:子育てしやすい環境づくり * 100 }
  RANKING ?score { ?c }
}

```

Listing 1: SPARQLoid クエリ例

表 1: 「yamada:子育てしやすい環境づくり」に対するマッピング例

MappedConcept	Confidence
adachi:住みやすい地域づくり	0.659
adachi:子育て環境整備施策	0.491
adachi:企業退職者の移住環境づくり施策	0.479
adachi:子育てカフェの整備・充実施策	0.429
adachi:結婚・出産・子育て戦略施策	0.324
adachi:妊娠・出産・子育て世代包括支援施策	0.289
adachi:子ども	0.050

3 本試作機構

本章では、マッピングデータおよびマッピング先の情報の提示するための機構と動的オントロジーマッチングの必要性、動的オントロジーマッチングに基づくチューニング支援について述べる。SPARQLoid クエリのコーディング時における不具合の原因調査のためにオントロジーマッピングおよびマッピング先の情報の提示を必要に応じて行い、不具合の原因の対処の困難さの解消のために、動的にオントロジーマッチングを用いて補助的な一時オントロジーマッピングを生成し、それを併用することを可能にする機構の実現に向け、複数のオントロジーマッピングを SPARQLoid クエリ内で効果的に扱うことを支援するための、クエリ・マッピングのチューニング支援について述べる。

3.1 マッピングデータおよびマッピング先の情報の提示

オントロジーマッピングマッピングを用いた拡張 SPARQLoid クエリでは、オントロジーマッピングを格納し

¹<http://sparql.es.ai.wu.ac.at>

表 2: 「yamada:子育てしやすい環境づくり」からマッピングされたクラスのインスタンスとサブクラス数

MappedConcept	Confidence	InstanceCount	SubClassOfCount
adachi:住みやすい地域づくり	0.659	0	1
adachi:子育て環境整備施策	0.491	0	2
adachi:企業退職者の移住環境づくり施策	0.479	3	0
adachi:子育てカフェの整備・充実施策	0.429	4	0
adachi:結婚・出産・子育て戦略施策	0.324	0	4
adachi:妊娠・出産・子育て世代包括支援施策	0.289	1	6
adachi:子ども	0.050	0	0

検索する記述を直接追加することにより、下位クラスに対してもクエリを行うことが可能である。オントロジーマッピングを用いた拡張 SPARQL クエリの記述では、たとえば検索に追加して含めたい関連語彙まで検索可能なようにクエリを書き換えるには、検索対象となるエンドポイント上で用いられるオントロジー中の目当てのクラス等の語彙との間でマッピングが存在しなければならず、それらが欠損している場合には、オントロジーマッピングを用いずに直接的に対象オントロジー上の語彙を参照してクエリを記述する以外に、解決の方法がない。また、そのような直接的な語彙の参照をクエリ中に行った場合には、拡張 SPARQL クエリにあるような、マッピングの確信度に基づく計算式を利用した並び替えやフィルタリング等と整合性の取れるクエリを作成することが難しく、拡張 SPARQL クエリから変換された複雑な SPARQL クエリを直接書き換えるといった対処が必要になる。本研究では、我々は、クエリの改変を直接的に行うことではなく、オントロジーマッピングを効果的に用いることに着目する。もしも、拡張 SPARQL クエリで用いられるオントロジーマッピングが必要な範囲の内容を十分に包含しない場合には、動的にオントロジーマッピングを用いて補助的な一時オントロジーマッピングを生成し、それを併用することを可能にする機構の実現を考える。この手法では、クエリを修正する必要がなく、オントロジーマッピングの内容を調整することで、直接マッピングされた概念等には結びついていないデータも、そのクエリの実行により取得することが可能になる。

今回の試作を行う機構では、動的オントロジーマッピングを行う設定となっている場合には動的に、補助的なオントロジーマッピングを生成し、それをクエリ中で使用できるようにする。

3.3 動的オントロジーマッピングに基づくチューニング支援

これまでの SPARQLoid のクエリ変換エンジンでは、動的に生成されたオントロジーマッピングに限らず、1

組のオントロジーの間に複数種のオントロジーマッピングが用意されている場合や、それらのオントロジーマッピングごとのカバー範囲が異なる場合などについては、その支援機構の検討が行われていなかった [Fujino 13]。本研究では、複数のオントロジーマッピングを SPARQLoid クエリ内で効果的に扱うことを支援するための、クエリ・マッピングのチューニング支援機構を用意している。動的オントロジーマッピングにより生成したオントロジーマッピングは、チューニング支援機構により、必要に応じてオントロジーマッピング格納エンドポイント上に読み込まれる。SPARQLoid クエリから SPARQL クエリの変換時において、オントロジーマッピングを格納したエンドポイントは、チューニング作業を効率的に行えるようにするために、チューニング支援機構に内包され、そこから直接制御される SPARQL エンドポイント上で動作し、チューニング作業中は、オントロジーマッピングの参照はこのエンドポイントを経由して行われる。本エンドポイントの実装には、fuseki²を用いている。

SPARQLoid クエリ実行時での動作のトレースを行いやすくするために、SPARQLoid クエリから SPARQL クエリに変換する過程において、必要に応じてチューニング用の付加情報取得のためのコードを付加できるようにした本試作機構を用いることによって、動的にオントロジーマッピング追加・改変しながら、そこで使用するオントロジーマッピングから参照されるクラス等にはインスタンスが直接結びついていない場合を検出したり、動的に生成する一時オントロジーマッピングに下位クラスや関連クラスを含めさせるようにマッピングを拡張させる機能を準備することで、OWL-DL における推論等を扱えない設定になっているエンドポイントに対しても、柔軟にインスタンスを検索するような SPARQLoid クエリを準備することが可能となる。

²https://jena.apache.org/documentation/serving_data/

4 おわりに

本研究では、拡張 SPARQL クエリのための動的オントロジーマッピングとチューニング支援機構の実現のための試みとその課題について述べた。オントロジーマッピングを用いた SPARQL クエリ拡張である SPARQLoid クエリのコーディング時における不具合の原因調査およびその対処の困難さの改善のためにオントロジーマッピングおよびマッピング先の情報を必要に応じて提示するための機構と、動的オントロジーマッピングによる検索対象拡大・調整のための一時オントロジーマッピングに基づくクエリコードのチューニング支援機構の試作について述べた。

本試作機構上では、ユーザがクエリの特定の部分に対応するインスタンスの有無などを参照したい場合において、それを提示可能なシステムの試作は行っているが、その効果的な方法には異なる複数の提示方法が考えられ、それぞれの優位性やその評価についての検討は、今後の課題である。

謝辞

本研究の一部は、JST CREST の支援を受けたものである。

参考文献

- [Adachi 16a] Adachi, T. and Fukuta, N.: Toward Better Debugging Support on Extended SPARQL queries with On-the-fly Ontology Mapping Generation, in *Proc. of The Eleventh International Workshop on Ontology Matching collocated with the 15th International Semantic Web Conference ISWC-2016 (OM2016)* (2016), (poster, to appear)
- [Adachi 16b] Adachi, T., Yamada, N., and Fukuta, N.: Toward Better Query Coding Support Utilizing Ontology Mappings, in *Proc. of 1st International Workshop on Platforms and Applications for Social problem Solving and Collective Reasoning (PASSCR2016)* (2016), (to appear)
- [Atencia 12] Atencia, M., Borgida, A., Euzenat, J., Ghidini, C., and Serafini, L.: A Formal Semantics for Weighted Ontology Mappings, in *Proc. of the 11th International Semantic Web Conference (ISWC2012)*, pp. 17–33 (2012)
- [Fujino 12a] Fujino, T. and Fukuta, N.: A SPARQL Query Rewriting Approach on Heterogeneous Ontologies with Mapping Reliability, in *Proc. of the IIAI-AAI International Conference on e-Services and Knowledge Management (IIAI-AAI-ESKM2012)*, pp. 230–235 (2012)
- [Fujino 12b] Fujino, T. and Fukuta, N.: SPARQLoid - a Querying System using Own Ontology and Ontology Mappings with Reliability, in *Proc. of the 11th International Semantic Web Conference (Posters & Demos) (ISWC2012)* (2012)
- [Fujino 13] Fujino, T. and Fukuta, N.: Utilizing Weighted Ontology Mappings on Federated SPARQL Querying, in *Proc. of the 3rd Joint International Semantic Technology Conference (JIST2013)* (2013)
- [Fujino 14] Fujino, T. and Fukuta, N.: Utilizing Weighted Ontology Mappings on Federated SPARQL Querying, in Kim, W., Ding, Y., and Kim, H.-G. eds., *Lecture Notes in Computer Science*, Vol. 8388, pp. 331–347, Springer-Verlag (2014)
- [Noguchi 13] Noguchi, H., Fujino, T., and Fukuta, N.: On Implementing a SPARQLoid Query Coding Support - Vocabulary Discovery for Queries with Weight Ontology Mappings, in *Proc. of International Workshop on Linked Data in Practice (LDPW 2013)*, pp. 33–38 (2013)
- [Noy 09] Noy, N. F.: Ontology Mapping, in Staab, S. and Studer, R. eds., *Handbook on Ontologies*, pp. 573–590, Springer-Verlag Berlin Heidelberg (2009)
- [足立 16a] 足立 拓也, 山田 直希, 野口 宙毅, 福田 直樹: オントロジーマッピングに基づく SPARQL クエリ記述支援システムの拡張機構の試作, 第 38 回セマンティックウェブとオントロジー研究会 (SIG-SWO38) (2016)
- [足立 16b] 足立 拓也, 福田 直樹: オントロジーマッピングを用いた SPARQL クエリにおけるマッピングに関する詳細知識を必要としないシステムの試作, 第 30 回人工知能学会全国大会 (JSAI2016), 1N4-OS-19a-5in2 (2016)
- [藤野 13] 藤野 敬久, 福田 直樹: マッピングの信頼度を利用した異種 LOD への横断的 SPARQL 検索におけるデバッグ支援手法の検討, 第 27 回人工知能学会全国大会 (JSAI2013), 1N4-OS-10b-2in (2013)