4월 2주차 연구노트

RDF datasets 에서의 SPARQL 쿼리 최적화 관련 논문 조사.

관련논문 : [An analytical approach for query optimization based on hypergraph](http://ieeexplore.ieee.org/document/7207087/)

([Sangeeta Sen](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Sangeeta%20Sen.QT.&newsearch=true); [Anisha Agrawal](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Anisha%20Agrawal.QT.&newsearch=true); [Ankit Rathi](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Ankit%20Rathi.QT.&newsearch=true); [Animesh Dutta](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Animesh%20Dutta.QT.&newsearch=true); [Biswanath Dutta](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Biswanath%20Dutta.QT.&newsearch=true) [2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=7180538) Year: 2015)

Query processing

재배열된 SPARQL 쿼리를 사용하면서, 데이터사전으로부터 요구된 subjects와 objects를 아래의 단계를 거쳐 얻을 수 있다.

1 Loop over the match patterns(재배열된 쿼리안에서 패턴매칭을 반복)

패턴을 매칭은 3가지 경우의 수가 있다.

1.1 subject에 해당하는 문자열이 상수이고 object에 해당하는 문자열이 변수인 경우

1.2 subject, object 둘 다 상수인 경우

1.3 subject, object 둘 다 변수인 경우

1.1과 1.2의 경우에는 현재 트리플안의 변수들의 결과들이 담겨있는 precess(t, res, d, tempdict) 함수를 사용하고 1.3은 다시 3가지 서브 케이스로 가진다.

1.3.1 subject와 object가 패턴매칭 안에서 이미 처리되기 전에 마주칠 수 있음.

1.3.2 둘 중 하나가 패턴매칭 안에서 이미 처리되기 전에 마주칠 수 있음.

1.3.3 마주치지 않음.

1.3.1과 1.3.2의 경우에는 tempdict안에 변수들의 응답들을 저장하고 나서 FindVariables(p, tempdict, d, t)함수를 사용한다. 현재 패턴매칭 안에서 subject변수들의 답변에 대한 답을 찾고, 이전에 마주친 동일한 변수에 대한 답을 한 후, 그에 따라 p안에서 변수들의 답을 수정한다. d안에 모든 마지막 응답을 저장한다. 마지막 1.3.3의 경우, 현재 패턴매칭의 서술어에 해당하는 subject 와 object를 찾고 d안에 저장한다.

2. process(t, res, d, tempdict)

해당 함수는 subject or object에 해당하는 문자열이 변수인 경우를 처리하는 함수이다.

3. FindVariables(p, tempdict, d, t)

현재 패턴매칭에 영향을 받을 변수들을 찾고, 수정하는 함수이다.