

개인화 에이전트를 이용한 시멘틱 웹서비스 검색

하상범^o 박영택

숭실대학교 컴퓨터학과

terrie@ailab.ssu.ac.kr^o, park@comp.ssu.ac.kr

The Personalized Agent using the Semantic Webservice Discovery

Sang-Bum, Ha^o Young-Tack, Park

Dept. of Computing, SoongSil University

요 약

현재 소프트웨어들은 웹서비스라는 새로운 분야로 확대되어 발전되고 있다. 또한 기존의 웹이 가지는 단점을 보완하기 위해 등장한 시멘틱 웹을 차세대 웹이라 말하고 이러한 시멘틱 웹을 사용하기 위한 많은 연구가 진행되고 있다. 본 논문에서는 에이전트가 정보의 의미를 이해하고 자동화한 기법들을 사용할 수 있게 하는 시멘틱 웹에 대한 연구를 웹서비스(WebService)분야로 확대시키고, DAML-S 라는 독립된 연구분야로서 서비스내용에 대한 기술과 프로세스 컨트롤의 정의, 서비스를 접근하는 방법들을 에이전트가 수행함으로써 기존의 Syntactic 기반 웹 서비스 검색구조가 가지는 문제점을 개선하고 에이전트가 추론형 검색엔진을 통하여 보다 정확하고 적합한 웹 서비스 검색을 가능하게 한다. 또한 개인화 기법들을 활용하여 효율성이 한단계 높은 개인화 기능을 에이전트에게 부여하게 된다.

1. 서 론

웹서비스는 새로운 소프트웨어 분산객체기술의 패러다임으로 서비스 지향적인 컴퓨팅이라고 하는데 웹을 대상으로 하는 만큼 분산의 정도가 매우 크다. 현재 웹은 Syntactic 정보로 이루어져 있어서 사람은 그 의미를 이해 할수 있지만 기계 즉, 에이전트는 이해 할수 없었다. 이러한 관점에서 등장한 시멘틱 웹은 웹상의 정보에 Syntactic만이 아닌 잘 정의된 의미(Semantic)를 포함하고 있기 때문에 에이전트가 자동으로 문서로부터 의미를 추출해 낼 수 있게 되었다. 이러한 개념에서 웹서비스 또한 시멘틱 웹으로 확대되어 연구되고 있는데 기존의 웹이 아닌 시멘틱 웹을 바탕으로 구조화되어 더욱 에이전트의 시멘틱 검색, 태스크의 자동화 구조가 가능하게 되었다. 특히 웹의 사용자와 정보량이 무한적으로 증가하는 가운데 사용자가 필요한 정보를 얻기 위해서는 시멘틱 웹을 통한 에이전트 도입은 필수적이라고 말할수 있다. 또한 에이전트에 개인화 기법을 사용하여 한 단계 높은 정확성과 맞춤형성을 구현 할수 있게 되었다.

본 논문에서는 시멘틱 웹 기반에서의 웹서비스를 구현하고 이를 바탕으로 사용자에게 알맞은 웹서비스를 검색하게 해주는 에이전트 시스템에 대하여 제안한다.

2. 관련연구

2.1 UDDI 검색방식

현재의 웹서비스는 UDDI(Universal Description Language) 레지스트리에 등록이 되어야만 검색이 가능하다.

WSDL 서비스 인터페이스는 UDDI 레지스트리에 tModel 형태로 저장이되며 UDDI의 API의 find_tModel 메시지를 통해 WSDL 서비스 기술내용을 찾을 수 있지만 등록이 되지 않은 서비스는 찾을수 없으며 검색의 표현력 또한 부족하다.

2.2 DAML+OIL

DAML+OIL 은 시멘틱 웹을 구현하는데 있어서 가장 중요한 온톨로지를 구축하기 위해 웹 온톨로지 언어로서 DAML(DARPA

Agent Markup Language) 프로그램의 DAML-ONT와 주로 유럽에서 개발된 OIL(Ontology Inference Layer)의 장점만을 취하여 발전한 기술로서 RDF를 확장하여 디스크립션로직(Description Logic)을 기반으로 풍부한 표현력을 제공한다. Cardinality Restriction 과 InversOf, Transitivity 등의 확장된 표현력을 제공하며, OWL과 더불어 가장 많이 사용되고 있는 온톨로지 언어이다. DAML+OIL 웹 온톨로지 언어는 현재 W3C에 의해서 OWL 웹 온톨로지 언어로 계승 발전되고 있다.

2.3 DAML-S

DAML-S 기술은 DAML+OIL 기술을 기반으로 Web service를 제공하는 기술이다. DAML+OIL 기술을 기반으로 하는 확장된 레이어를 제공하여 Web Service의 모성(Unambiguous)를 제거하고, 상호운용성을 증가시킨다. DAML+OIL 기술을 기반으로 Service를 기술하기 위한 Upper 온톨로지를 정의하고, 이 Upper 온톨로지의 인스턴스를 만들어서 실제적인 서비스를 기술한다. DAML-S는 Upper 온톨로지는 Service, Service Profile, Service Model, Service Grounding의 4가지 온톨로지 로 구성되어 있으며, 서비스 접근해 대한 프로토콜을 정의하는 Service Grounding 온톨로지는 WSDL 문서와 연결된다.

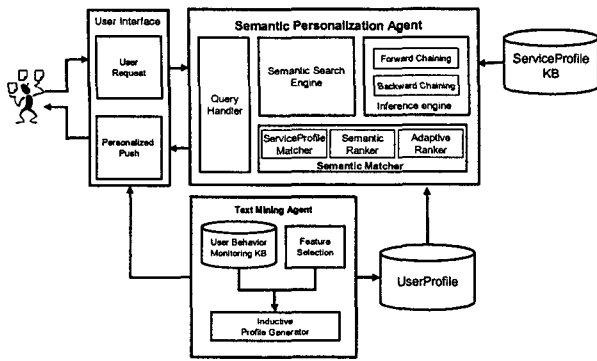
3. 시멘틱 웹서비스 검색을 위한 개인화 에이전트 시스템

본 장에서는 본 논문의 전개상 미리 언급되어야 할 전체 시스템 구조와 환경 및 배경에 대해서 설명한다.

시스템을 구성하는 모듈은 크게 세가지로 구분된다. 첫 번째는 DAML-S 서비스 온톨로지를 바탕으로 구현된 DAML-S 인스턴스와 두 번째는 사용자로부터 입력받은 질의와 변환된 DAML-S 인스턴스를 Rule을 바탕으로 추론을 하는 시멘틱 기반의 추론형 검색엔진 그리고 마지막으로 사용자 프로파일을 기반으로 하는 개인화 에이전트이다.

3.1 시스템 구조와 주요 컴포넌트

본 논문에서 제안하는 시스템의 구조는 다음과 같다.



[그림 1] 시스템 구조

주요 모듈인 시맨틱 개인화 에이전트 시스템은 세부적으로 두 단계의 계층으로 구성되어 웹서비스 검색을 수행하게 된다. 우선 첫 번째 레이어에서는 사용자의 요구사항을 입력받고 결과값을 보여주는 사용자 인터페이스와 DAML-S ServiceProfile 집단으로부터 웹서비스를 검색하는 추론형 검색엔진으로 구성되어 있고 두 번째 레이어에서는 사용자의 행동을 관찰하고 이것을 바탕으로 사용자 프로파일을 생성하는 개인화 에이전트로 구성되어 있다. 첫 번째 레이어의 웹서비스 검색결과가 충분하지 못하여 사용자에게 사용자에게 선택되지 않을 수 있기 때문에 사용자 프로파일을 바탕으로 하는 두 번째 레이어의 에이전트의 기능이 필요하게 된다.

3.2 DAML-S Service

본 논문에서는 DAML-S Upper 온톨로지를 참조하여 인스턴스를 생성하고 추론형 검색엔진에서 서비스를 찾는 연구를 수행하였다. 다음과 같이 ServiceProfile을 생성한다.

```
<profile:serviceName>PersonoalInfoAgent</profile:serviceName>
<profile:input>
  <profile:ParameterDescription df:ID="PersonallInfo_socialSecurityNumber_IN">
    <profile:parameterName>PersonallInfo_socialSecurityNumber_IN
  </profile:parameterName>
  <profile:restrictedTo rdf:resource="&xsd:integer" />
  <profile:refersTo rdf:resource="&my_process:#personallInfo_socialSecurityNumber_IN" />
  </profile:ParameterDescription>
</profile:input>
<profile:output>
  <profile:ParameterDescription rdf:ID="PersonallInfo_financialStatus_OUT">
    <profile:parameterName>PersonallInfo_financialStatus_OUT
  </profile:parameterName>
  <profile:restrictedTo rdf:resource="&xsd:string" />
  <profile:refersTo rdf:resource="&my_process:#personallInfo_financialStatus_OUT" />
  </profile:ParameterDescription>
</profile:output>
```

이러한 DAML+OIL 형태의 ServiceProfile은 RDF 문서의 서브셋으로서 Predicate, Subject, Object의 Triple로 재표현할 수 있다. Triple 형태로 재표현하는 이유는 추론형 검색엔진에서는 FOL(First Order Logic)형태의 미리 정의된 Axiom을 사용하여 추론을 수행하게 되므로 FOL 형태로 변환이 필요하다. 이와 같은 ServiceProfile은 다음과 같은 Triple 형태로 변환이 되어진다.

```
(has_process Profile_PersonalInfoAgent PersonalInfo_Process)
(inputProfile_PersonalInfoAgent PersonalInfo_socialSecurityNumber_IN)
(output Profile_PersonalInfoAgent PersonalInfo_financialStatus_OUT)
```

이렇게 변환된 Triple들은 추론형 검색엔진에서 주어진 규칙들을 바탕으로 입력된 질의들을 통해 새로운 정보를 추론하는데 사용하게 된다. 다음과 같이 질의값이 주어졌을 때 두가지 형태로 웹서비스를 검색하게 된다. 첫 번째는 서비스 이름을 바탕으로 검색을 하게 되고 두 번째는 입력값과 출력값을 대상으로 검색하게 된다.

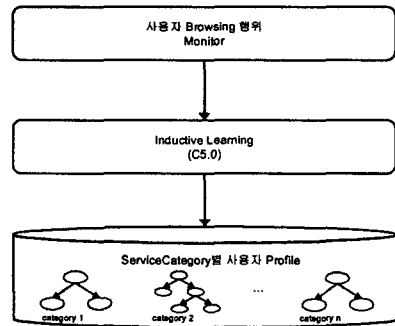
Query1	(serviceName Profile_PersonalInfoAgent ?K)
Query2	{(and(input ?K PersonalInfo_socialSecurityNumber_IN) (output ?K PersonalInfo_financialStatus_OUT))}

추론형 검색엔진의 결과는 다음과 같다.

$$K = |profile_personalinfoagent|$$

3.3 사용자 프로파일 구축

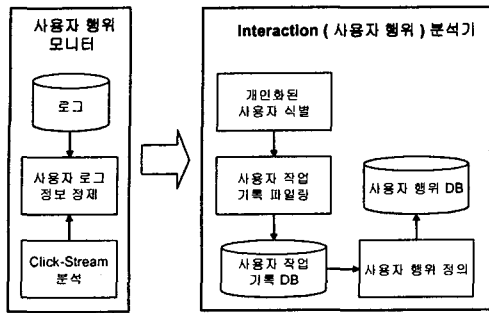
본 논문에서 제안하는 시스템을 사용하는 사용자의 개인 프로파일을 구축하는 작업은 에이전트의 성능을 향상시켜 각 개인에게 가장 적합한 정보를 제공할 수 있는 기능을 제공하게 된다. 사용자가 요구했던 사항들과 추론형 검색엔진의 검색결과와 그중에서 사용자가 선택했던 사항들에 대한 사용자 각각의 프로파일을 구축하여 검색결과에 반영한다면 보다 정확성이 뛰어난 개인화 에이전트를 구현할 수 있게 되는데 다음과 같은 세 단계의 작업을 요구하게 된다.



[그림2] 프로파일 생성 구조

3.3.1 사용자 Browsing 행위 모니터

Click-Stream 관점에서의 로그분석은 시스템 내에서 사용자들의 중요한 이벤트를 추출하여 중요 History로 기록한다. 로그문서는 사용자가 접속한 서버에서 행위한 모든 작업이 시간별로 저장되어 있는 문서이므로 로그를 통하여 얻어진 정보는 필요한 정보와 불필요한 정보로 구분되어 사용자가 행한 기록을 저장한다.

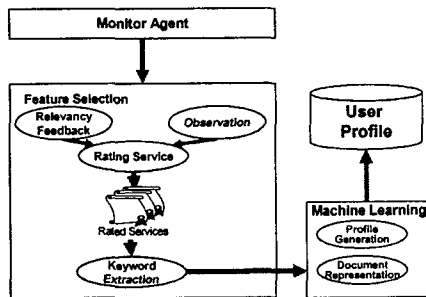


[그림3] 사용자행위 분석기

3.3.2 귀납적 기계학습

개인화 에이전트는 사용자 브라우징 행위 결과를 학습모델에 전달한다. 사용자에게 의해 선택된 서비스들은 양의 예제(positive example)로 구분되고 선택되는 않은 서비스들은 음의 예제(Negative Example)로 구별되어 귀납적 기계학습법에 의해 일반화(Generalization)과정을 수행하여 가설을 생성하고 이를 기반으로 새로운 예제가 입력되었을때 그 예제의 클래스를 예측할 수 있게 된다. 즉, 사용자가 원하는 웹서비스를 검색하고 추천형 검색엔진을 통해 나온 검색결과에서 사용자가 선택하는 웹서비스들을 통해 사용자의 선호도를 일반적인 규칙으로 생성하게 된다. 또한 DAML-S 에서의 ServiceProfile은 계층적 구조로 이루어져 있으므로 ServiceCategory별로 사용자의 선호도 파악이 쉽게 이루어 질 수 있다.

본 논문에서는 사용자 프로파일을 구축함에 있어서 먼저 사용자의 행위 정보에서 Feature Selection을 수행하여 중요 Feature들을 추출하고 이를 귀납적 기계학습 시스템으로 널리 이용되고 있는 C5.0 학습 시스템을 적용하여 Decision Tree로 구축한 후 이를 사용자 관심 정보 모델, 즉 사용자 프로파일의 형태로 생성하는 작업에 이용한다.

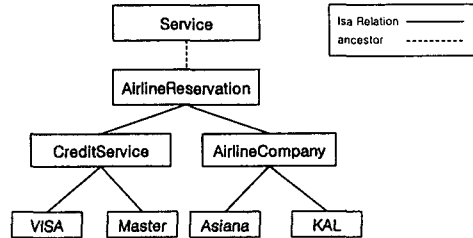


[그림4] 귀납적 기계학습을 통한 프로파일생성

3.3.3 온톨로지 계층하의 시멘틱 순위

DAML-S 의 ServiceProfile들은 모두 서비스 Upper 온톨로지 에서 정의되어진 계층 아래에서 선언되어 있다. 일반적으로 서비스들은 ProfileHierarchy 에 계층적으로 선언되어 있는데 이 점을 바탕으로 추천형 검색엔진의 결과를 시멘틱 매칭시켜 검색결과 순위를 조절 할 수 있게 된다. 예를 들어 사용자가 여행을 가기위해서 비행기표를 예약하는 웹서비스를 검색한다고 가정했을때 시멘틱 추천형 검색엔진의 검색된 결과가 Asiana 웹서비스와 KAL 웹서비스를 검색하였다면 사용자 프로파일을

통해 사용자가 선호하는 항공사의 순위를 높여서 검색결과를 출력 해줄 수가 있게 된다. 또한 비행기표를 예약하기위한 웹 서비스를 일반적인 검색방법을 통해서서는 찾을수 없는 비행기표를 예약하기위한 신용카드를 결제할 수 있는 웹서비스를 포함하여 검색결과를 출력하게된다. 이것은 서비스가 온톨로지의 계층적 구조를 바탕으로 정의되어 있기 때문에 에이전트를 통해서 구현이 가능한 것이다.



[그림5] 서비스 온톨로지 계층

4. 결론 및 향후 연구과제

본 논문에서는 시멘틱 웹을 기반으로 하는 DAML-S 의 추천형 검색엔진을 통한 검색방법과 개인화 에이전트를 통한 보다 정확하고 효율성이 높은 사용자에게 적합한 웹서비스 검색방법을 제안하였었다. DAML-S 온톨로지를 이용한 추천방법을 사용하였으며, 사용자 프로파일 사용하여 에이전트를 통한 맞춤형 서비스를 시도하였다.

향후 더 많은 DAML-S 인스턴스들의 확보를 통하여 본 논문에서 제시한 방법에 대한 추가적인 검증이 필요하겠고 서비스 구성(Service Composition) 분야에서 계획(Planning)을 사용하여 자동적인 실행이 가능한 에이전트로 발전시키는 연구가 필요하다.

5. 참고문헌

- [1] Deborah L.McGuinness, Richard Fikes, James Hendler, Lynn Andrea Stein, "DAML+OIL: An Ontology Language for the Semantic Web," IEEE Intelligent Systems, September, 2002.
- [2] Daniel J.Mandell, Sheila A.McIlraith, "A Bottom-up Approach to Automating Web Service Discovery, Customization, and Semantic Translation," In The Proceedings of the Twelfth International World Wide Web Conference Workshop on E-Services and the Semantic Web (ESSW '03). Budapest, 2003.
- [3] Tom M.Mitchell, "Machine Learning," McGraw-Hill, 1997
- [4] Mark Devaney, Ashwin Ram, "Efficient Feature Selection in Conceptual Clustering," Machine Learning : Proceeding of the Fourteenth International Conference, Nashville, 1997.
- [5] Kathleen Mckusick, Kevin Thompson, "COBEWEB/3:A Portable Implementation," NASA Ames Research Center, Technical Report FIA-90-6-18-2, 1990
- [6] Java Theorem Prover.
http://www.ksl.stanford.edu/software/JTP/
- [7] DAML Services Coalition. DAML-S version 0.9,
http://www.daml.org/services/
- [8] DAML, Darpa Agent Markup Language Program,
http://www.daml.org/