

퍼지 온톨로지를 이용한 지능형 에이전트

박중민⁰ 양형정 양재동
전북대학교 케이테크 전북대학교
(jpark⁰, hjyang, jdyang)@ozzy.chonbuk.ac.kr

Intelligent Agent with Fuzzy Ontology

Jong-Min Park⁰ Hyung-Jeong Yang Jae-Dong Yang
Chonbuk University, K-Tech, Chonbuk University

요 약

현재 전자상거래 시스템에서 도메인에 대한 전문적 지식이 없는 사용자는 원하는 상품을 찾기 어렵다. 또한, 다양한 전자상거래 시스템간의 공통적인 상품 정보에 대한 표준이 부족하므로 원하는 상품을 찾기 위해선 많은 시간과 노력이 필요하다. 이를 위해 본 논문에서는 시멘틱 웹 기반의 온톨로지 언어를 확장하여 퍼지 상품 지식베이스를 구축하고 지능적 질의 처리가 가능한 지능형 에이전트를 제안한다. 상품 지식베이스를 퍼지 온톨로지를 통해 구축함으로써 도메인에 대한 전문적인 지식이 없는 사용자를 지원하고, 서로 다른 시스템간에 표준적인 상품 지식으로서 질의 처리에 사용될 수 있다.

1. 서론

인터넷상에서 전자상거래가 활성화됨에 따라 대상 품목도 종류가 다양해지고 이용자 층도 확대되어 가고 있다. 하지만, 여전히 전자상거래 시스템에서 도메인에 대한 전문적인 지식이 없는 일반 사용자들은 원하는 상품을 찾기 어렵다. 이러한 사용자를 지원하기 위해 지능형 소프트웨어인 에이전트[1]가 사용되지만 전자상거래 시스템에 따라 상이한 포맷으로 상품 정보를 기술하기 때문에 다른 시스템의 데이터를 접근해 정보를 공유하는데 한계를 가지고 있다. 이러한 문제를 해결하기 위한 기반 기술로 시멘틱 웹에 대한 연구가 활발히 진행되고 있다.

시멘틱 웹은 현재 웹을 확장한 버전으로 사람이나 컴퓨터가 데이터를 읽고 서로가 이해할 수 있도록 잘 정의된 의미와 구조를 가지는 웹이다. 이러한 정의를 만족하기 위해서는 해당 자원에 대한 표준화된 데이터 표현 방법과 표준화된 방법을 통해 기술된 자원의 내용을 서로간에 공통적으로 이해할 수 있도록 해주는 지식 표현 수단인 온톨로지를 필요로 한다. 이러한 온톨로지를 통해 자원내의 내용 뿐만 아니라 자원간의 관계를 통해 새로운 사실을 추론할 수 있는 기반이 된다[2].

전자상거래 시스템에서 도메인에 대한 온톨로지 구축시 전문적인 지식이 없는 일반 사용자들을 지원하기 위해서는 일상적으로 사용하는 애매 모호한 용어의 처리가 필수적으로 요구된다. 이러한 애매 모호한 용어를 온톨로지에 퍼지 이론을 응용하여 표현하고 지능형 에이전트에 의한 추론 과정을 거쳐 사용자에게 보다 의미적으로 정확한 결과를 제공할 수 있다. 이를 위해 본 논문에서는 시멘틱 웹 기반의 온톨로지 언어를 확장하여 상품에 대한 퍼지 기반 지식베이스를 구축하고 지능적 질의 처리가 가능한 지능형 에이전트를 제안한다.

퍼지 기술을 응용한 퍼지 상품 지식베이스를 이용함으로써 상품 선택 가이드에 전문가의 지식을 활용하여 사용자 중심의 지능적 상품 검색이 가능하다. 따라서, 상품 정보에 대한 전문적 지식이 없는 사용자를 지원하고, 사용자의 취향에 따른 동적인 상품 분류 계층을 제시할 수 있다.

본 논문의 구성은 2장에서 퍼지 온톨로지를 이용한 퍼지 상품 지식베이스 표현 방법에 대해 살펴보고, 3장에서는 구축된 퍼지 상품 지식베이스를 이용한 지능적인 질의 처리 방법에 대해 설명한다. 4장에서는 제약 사항 처리 방법에 대해 설명한다. 마지막으로 결론 및 향후 연구를 제시한다.

2. 퍼지 온톨로지를 이용한 퍼지 상품 지식베이스 표현

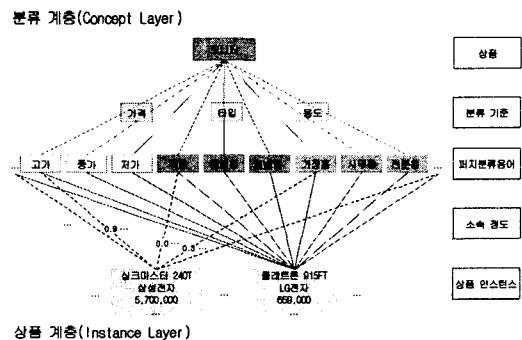


그림 1 퍼지 기반 상품 지식 구조

퍼지 상품 지식베이스는 전문가가 퍼지 분류 용어들 사이의 관계

를 일반화 관계로 분류하고 상품 인스턴스들과 퍼지 분류 용어들 사이의 관계는 클래스 관계로 분류한다. 그림 1은 퍼지 상품 지식베이스 구조를 도식화한 그림이다. 퍼지 분류 용어에는 분류 기준의 타입에 따라 퍼지 수, 퍼지 용어, 그리고 복합 퍼지 용어가 있으며 퍼지 분류 용어들의 멤버십 함수를 통해 상품 인스턴스가 퍼지 분류 용어에 소속될 퍼지 소속 정도 값을 계산하게 된다.

본 논문에서는 위에서 설명한 퍼지 분류 용어를 정의하기 위한 온톨로지로 현재 표준 온톨로지 언어로 대두되고 있는 DARPA 프로젝트의 일환인 DAML+OIL을 이용하여 정의한다[3]. 다음은 DAML+OIL에서 사용될 퍼지 소속 정도 값을 나타내기 위한 타입을 정의한 내용이다.

```

<!-- ***** User Defined Datatype ***** -->
<xsd:simpleType name="FuzzyValue">
  <xsd:restriction base="xsd:decimal">
    <xsd:minInclusive value="0.0"/>
    <xsd:maxInclusive value="1.0"/>
  </xsd:restriction>
</xsd:simpleType>
    
```

현재 DAML+OIL을 비롯한 온톨로지 언어들이 아직은 연산에 관한 정의가 되어 있지 않다. 따라서 본 논문에서 에이전트가 연산에 포함될 퍼지 멤버십 함수에 대한 정의를 갖고 퍼지 분류 용어 정의시 참조하는 방법으로 처리한다.

▪ 퍼지 수(FN : Fuzzy Number)

퍼지 수는 가격, 크기, 무게, 사이즈와 같이 분류 기준이 수치 값(Continuous Value)일 경우 사용한다. 예를 들어, 모니터의 가격에 따라 퍼지 수를 사용해 고가, 저가, 그리고 증가의 퍼지 분류 용어를 정의할 수 있다.

▪ 퍼지 용어(Fuzzy Linguistic Term)

퍼지 용어는 상품 회사나 상품 타입등과 같이 분류 기준이 문자열(Discrete Value)일 경우 사용한다. 예를 들어, 모니터 타입을 일반 모니터, 액정 모니터, 그리고 평면 모니터와 같이 분류할 수 있다. 여기서 액정 모니터는 $T_{액정} = \{0.7/LCD, 0.8/TFT\}$ 와 같이 퍼지 시소러스[4]나 전문가의 입력을 통해 정의할 수 있다.

▪ 복합 퍼지 용어(Composite Fuzzy Term)

복합 퍼지 용어는 기존에 정의된 퍼지 수, 퍼지 용어, 그리고 복합 퍼지 용어들을 사용하여 정의된다. 예를 들어, 모니터를 용도에 따라 분류할 경우 전문가용, 사무용, 그리고 가정용과 같이 분류할 수 있다. 여기서 전문가용 모니터는 “0.4/고해상도 AND 0.4/고색상 AND 0.2/(대형 OR 중형)” 와 같이 전문가의 입력을 통해 정의할 수 있다.

다음은 위의 세가지 퍼지 분류 용어를 DAML+OIL로 표현하기 위해 정의된 내용이다.

```

<!-- ***** Class ***** -->
<daml:Class rdf:ID="FuzzyTerm" />
<daml:Class rdf:ID="FuzzyInstance" />
<daml:Class rdf:ID="MembershipFunc" />
.....
<!-- ***** Enumeration ***** -->
<daml:Class rdf:ID="FuzzyTermType">
  <daml:oneOf rdf:parseType="daml:collection">
    <daml:Thing rdf:ID="FuzzyNumber" />
    <daml:Thing rdf:ID="FuzzyLinguisticTerm" />
    <daml:Thing rdf:ID="CompositeFuzzyTerm" />
  </daml:oneOf>
</daml:Class>
.....
<!-- ***** Datatype Property ***** -->
<daml:DatatypeProperty rdf:ID="Weight">
  <rdfs:domain rdf:resource="FuzzyInstance" />
  <rdfs:range rdf:resource="http://../FuzzyValue" />
</daml:DatatypeProperty>
.....
<!-- ***** Object Property ***** -->
<daml:ObjectProperty rdf:ID="Type-Of">
  <rdfs:domain rdf:resource="#FuzzyTerm" />
  <rdfs:range rdf:resource="#FuzzyTermType" />
</daml:ObjectProperty>
.....
    
```

다음은 위의 정의에 해당하는 퍼지 분류 용어와 퍼지 인스턴스에 해당되는 실제 인스턴스를 DAML+OIL로 예들들어 정의한 내용이다.

```

<!-- ***** Instance ***** -->
<FuzzyTerm rdf:ID="HighPrice">
  <Name> High Price </Name>
  <Type-Of rdf:resource="#FuzzyNumber" />
  <Apply-To rdf:resource="#Monitor" />
  <Concern-About> Price </Concern-About>
  <Membership_Func rdf:resource="#HighPrice_Func" />
</FuzzyTerm>

<FuzzyInstance rdf:ID="1">
  <Product-Of rdf:resource="#Monitor" />
  <FuzzyInstance-Of rdf:resource="#HighPrice" />
  <Weight> 0.8 </Weight>
</FuzzyInstance>
    
```

다음으로, 상품은 그 성격에 따라서 단일 상품(Primitive Product)과 복합 상품(Composite Product)으로 나눌 수 있다. 단일 상품은 세부 구성 상품이 없는 상품으로 퍼지 수, 퍼지 용어, 복합 퍼지 용어로 퍼지 상품 지식을 정의될 수 있으며, 복합 상품은 여러 개의 단일 상품으로 구성되는 상품으로 복합 퍼지 용어로만 퍼지 상품 지식을 정의할 수 있다.

3. 퍼지 상품 지식베이스를 이용한 지능적 질의 처리

사용자의 모호한 질의에 대해 퍼지 상품 지식베이스를 이용하여 지능적으로 상품을 검색한다. 사용자의 질의는 같은 퍼지 분류에 있는 용어들 간에는 논리합(∨)으로, 다른 분류에 있는 용어들 간에는 논리곱(∧)을 이용하여 부울 질의 형태로 변환되며, 변환된 질의는 지식 추론 엔진을 이용하여 분석하고 관련 상품 정보를 퍼지 정도에 따라 순위화 된 형태로 제시한다. 본 논문에서는 논리합과 논리곱을

위한 연산자로서 min, max 표준 퍼지 연산자를 사용하였다.

먼저, 가장 기본적인 질의로서 '가격이 보통인 모니터' 라는 질의는 ' $Q = \langle \text{모니터, 가격, 보통} \rangle$ ' 와 같이 작성된다. 이 질의는 모니터에 해당하는 상품 인스턴스들 $o_i \in \text{instance}(\text{모니터}), i=1,2,\dots, n$ 인 중에 $\mu_{\text{보통}}(o_i.\text{가격}) > \alpha$ 을 만족하는 상품 인스턴스를 결과로서 제시된다.

다음으로, 논리합(\vee)인 포함된 질의에 대해 살펴보면, '가격이 고가이거나 보통인 모니터' 라는 질의 Q 는,

$$Q_{m1} = \langle \text{모니터, 가격, 보통} \rangle$$

$$Q_{m2} = \langle \text{모니터, 가격, 고가} \rangle \text{ 라 할 때, } Q = Q_{m1} \vee Q_{m2} \text{ 이다.}$$

상품 인스턴스들 $o_i \in \text{instance}(\text{모니터}), i=1,2,\dots, n$ 중에서 $\max(\mu_{\text{보통}}(o_i.\text{가격}), \mu_{\text{고가}}(o_i.\text{가격})) > \alpha$ 을 만족하는 상품 인스턴스를 결과로서 제시해 준다.

마지막으로, 논리합(\vee)과 논리곱(\wedge)이 포함된 질의로 '가격이 고가이거나 보통인 LCD 모니터' 라는 질의 Q 는,

$$Q_{m1} = \langle \text{모니터, 가격, 보통} \rangle$$

$$Q_{m2} = \langle \text{모니터, 가격, 고가} \rangle$$

$$Q_{m3} = \langle \text{모니터, 타입, LCD} \rangle \text{ 라 할 때,}$$

$$Q = (Q_{m1} \vee Q_{m2}) \wedge Q_{m3} \text{ 이다.}$$

상품 인스턴스들 $o_i \in \text{instance}(\text{모니터}), i=1,2,\dots, n$ 중에서 $\min(\mu_{\text{LCD}}(o_i.\text{타입}), \max(\mu_{\text{보통}}(o_i.\text{가격}), \mu_{\text{고가}}(o_i.\text{가격}))) > \alpha$ 을 만족하는 상품 인스턴스를 결과로서 제시해 준다.

만약, 질의에 표현된 퍼지 용어가 퍼지 상품 지식베이스 내에 존재하지 않으면 기존의 퍼지 용어에 다음과 같은 퍼지 변경자(Fuzzy Modifier)를 적용하여 질의 처리에 이용한다.

$$\mu_{\text{low}}(x) = 1 - \mu_{\text{high}}(x)$$

$$\mu_{\text{very-high}}(x) = (\mu_{\text{high}}(x))^2$$

$$\mu_{\text{mid-low}}(x) = 1 - |\mu_{\text{mid}}(x) - \mu_{\text{low}}(x)|$$

마지막으로, 검색 대상이 복합 상품이고 질의에 표현된 퍼지 용어가 퍼지 상품 지식베이스 내에 존재하지 않으면 상품을 구성하는 각 단일 상품에 대해 질의에 나타난 퍼지 용어를 적용하여 검색한다.

4. 제약 사항 처리

온톨로지에 도메인과 범위 지정하는 슬롯을 두어서 제약사항으로 이용할 수 있다. 예를 들어, 복합 상품에 대한 질의시 복합 상품은 단일 상품들의 조합으로 구성되기 때문에 검색된 후보 상품들간의 호환성에 대한 문제가 있다. 이런 경우 단일 상품 지식들간의 정의된 제약 사항 관계를 이용하여 호환되지 않거나 적절하지 않은 단일 상품들 간의 결합을 방지하고 검색 속도 또한 향상시킬 수 있다

또한 개념의 속성과 도메인, 범위 등의 제약 사항을 이용하여 추론을 통한 필터링을 수행하여 온톨로지 구성시 인덱싱 공간을 줄여 줌

으로써 온톨로지 엔지니어의 전문성을 보완 할 수 있다. 예를 들어, "Pentium 4 CPU" 와 호환되는 "마더보드" 부품에 대한 정보를 온톨로지서 호환성 관계로 표현하고자 할 때 온톨로지 엔지니어에게 컴퓨터 부품에 대한 전문적인 지식이 요구된다. 이때 이에 해당되는 결과 부품의 상위 클래스 'Pentium 4 CPU 지원 마더보드' 가 'Pentium 4 CPU' 와 연관관계를 맺고 있다면 하위 부품들을 온톨로지 엔지니어에게 추천하여 온톨로지 엔지니어의 작업량을 현격히 줄일 수 있다.

5. 결론

본 논문에서는 시맨틱 웹 기반의 온톨로지 언어를 확장하여 상품에 대한 퍼지기반 지식베이스를 구축하고 지능적 질의 처리가 가능한 지능형 에이전트를 제안한다. 퍼지기반 지식베이스를 정의하는데 온톨로지를 이용함으로써 서로 다른 시스템간의 상품 지식 공유 문제를 해결하였으며, 퍼지 기술을 응용한 퍼지 상품 지식베이스를 전문가의 지식으로 이용함으로써 상품 선택 가이드에 사용자 중심의 지능적 상품 검색이 가능하다. 따라서, 상품정보에 대한 전문적 지식이 없는 사용자의 요구를 만족시키고, 사용자의 취향에 따른 동적인 상품 분류 계층을 제시할 수 있다. 또한, 상품에 대한 전문가 지식을 퍼지 용어들로 지식베이스에 저장함으로써 검색엔진의 기능 확장성을 용이하게 하였다. 즉, 지식베이스에 특정 도메인 지식을 퍼지 용어로 저장하기만 하면 에이전트는 원하는 지능을 부여 받을 수 있기 때문에, 원하는 지능이 필요할 때마다 응용 프로그램을 새로 구현할 필요가 없다.

향후 연구 과제로는, 응용분야에 적합한 퍼지 멤버십 함수의 정의와 멤버십 함수를 퍼지 온톨로지로 정의하기 위한 좀 더 정확한 방법론, 퍼지 분류 용어들의 좀 더 세밀한 추론 규칙과 제약 처리 규칙 정의가 필수적으로 요구된다.

참고문헌

- [1] Robert Guttman and et al, Agents as Mediators in Electronic Commerce, Matthias Klusch (Ed.), Intelligent Information Agents, pp.131-152, Springer-verlag, Berlin Heidelberg, 1999
- [2] Asuncion Gomez-Perez and et al, Ontology Languages for the Semantic Web, IEEE Intelligent System, pp.54-60, 2002
- [3] I. Horrocks and P. Patel-Schneider, The generation of DAML+OIL, In Proc. Of DL 2001, pp.30-35, 2001
- [4] 최재훈, 한종진, 박종진, 양재동, 구조적인 시소러스 구축을 지원하는 객체 기반 정보 검색 모델, 한국정보과학회논문지, Vol. 24, No. 11, pp. 1244-1256, 1997