

지능형 통합에이전트를 이용한 검색시스템

A Search System Using The Intelligent Agent

박진희*, 허철회**, 정환목*-

*대구가톨릭대학교 컴퓨터정보통신공학부

**성덕대학 컴퓨터정보계열

*Jin-Hee Park, **Chul-hoi Her, *Hwan-Mook Chung

E-mail : aimajor@hanmail.net, hmchung@amare.ac.kr

요 약

전자상거래가 점차 활성화됨에 따라 다양한 형태의 쇼핑물들이 구축되고 있으나, 구매자가 상품을 구입하는데 있어 구매자 기호와 요구에 적합한 상품을 검색하기에는 미흡한 실정이다. 따라서, 본 논문에서는 CBR(Case Based Reasoning)과 RBR(Rule Based Reasoning)을 통합한 검색에이전트와 사용자 프로파일과 선호도를 관리하는 사용자 에이전트로 이루어진 멀티 에이전트를 이용하는 CARUBA 시스템을 설계하고, 검색에이전트가 사용자에이전트에서 보낸 정보를 이용하여 유사도를 산출하여 구매자의 요구에 적합한 상품을 신속하게 추천할 수 있는 방법을 제안한다.

Keywords : CBR, RBR, XML, Intelligent Agent

1. 서 론

인터넷사용자의 빠른 증가로 많은 정보가 창출되고, 관리되며 활용되고 있다. 이러한 많은 정보들 중에 사용자들에게 꼭 필요하고 유용한 정보를 검색하여 활용할 수 있는 많은 연구가 이루어지고 있다[2].

특히 월드와이드웹(WWW) 이용한 전자상거래(EC: Electronic Commerce)는 실세계의 상거래를 가상공간에서 이루어지게 함으로서 상품 홍보, 거래, 대금 지불, 상품 전달에 대한 새로운 정보와 기술의 발전을 가져오고 있다. 또한 기존의 상거래보다 시간, 매장, 거리등의 제약을 극복할 수 있도록 하였다[5].

전자상거래가 보다 효율적으로 활용되려면 사용자에게 이용하기 편리한 인터페이스를 제공하고, 상품검색이 쉽고 빠르며, 검색한 상품의 정보는 고객이 만족할 만한 정보를 가지고 있어야한다[4].

그러나 현재 전자상거래에서는 구매자가 상품을 구입하는데 있어, 구매자 기호와 요구에 적합한 상품을 검색하기에 미흡한 실정이다.

따라서 본 논문에서는 사례기반추론과 규칙기반추론을 통합한 검색에이전트와 사용자에이전트를 결합한 멀티 에이전트가 구매자의 기호와 요구에 가장 유사한 상품을 추천하는 검색 방법을 제안하고 유사도 산출 방법에 관하여 고찰한다.

2. 규칙기반추론과 사례기반추론

규칙기반추론은 가장 오래 사용되어 왔고 보편화된 지식습득 방법이다. 최근까지 대부분의 전문가 시스템에서는 규칙기반추론이나 사례기반추론을 따로 사용하였다.

규칙기반추론에서는 문제영역의 규칙을 인간 전문가로부터 모두 추출한 다음 관리자가 규칙을 정리하여 규칙베이스로 구현하고 이 규칙에 의해 추론함으로써 해를 얻는다. 규칙기반추론은 보통 70%이상이 'if-then' 형식을 취하고 있기 때문에 모듈성, 균일성, 자연성의 장점을 가진다[1]. 그러나 실제로 문제를 해결할 때 미리 모든 규칙을 구축할 수 없는 경우가 많으며, 문제와 규칙이 일치하지 않을 경우에는 문제를

해결하기 어렵다. 또한 규칙기반추론은 문제가 주어질 때마다 주어진 문제를 해결하기 위하여 관련된 규칙을 순서대로 처리하므로 규칙의 수가 증가할수록 성능이 저하된다.

사례기반 추론은 새로운 문제를 해결하기 위해 과거의 유사한 문제 해결 사례를 기반으로 해결 방법을 적절히 변형하여 유사한 사례로부터 결과를 찾아내는 효율적인 방법이다. 문제를 해결할 때 문제와 해를 기억하고 있기 때문에 경험으로부터 쉽게 해결할 수 있다.

그러나 탐색비용이 많이 들고 사례색인에 관한 문제점이 있다. 이러한 이유로 인해 서로의 단점을 보완하기 위하여 두 추론 방법을 결합하기 위한 많은 연구가 진행되어 왔다.

2.2 XML

XML은 데이터 자체를 구조적으로 표현할 수 있으며, 확장성이 뛰어나고, 문서에 대한 구조를 자체기술(self-describing)하는 형태로 되어 있다. 응용 프로그램간의 호환성이 뛰어나고, 에이전트와 클라이언트, 서버간에 통신을 용이하게 할 수 있으며 상호작용에 의한 처리 능력을 향상시킬 수 있다. 즉, XML태그가 정의하는 기능을 이용하여 웹브라우저에 표시되는 내용을 에이전트들이 쉽게 해석할 수 있게 함으로써 탐색이 용이하고 자동색인 기능을 지원함으로써 사례기반추론의 단점을 보완할 수 있다.

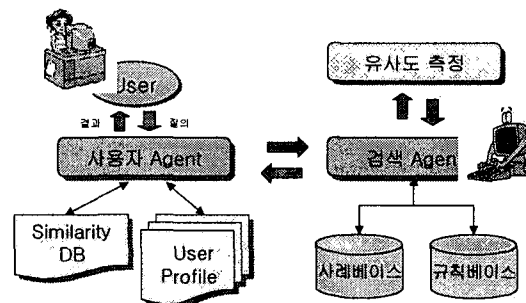
3. CARUBA 시스템의 구조

사례베이스추론(Case-Based reasoning)과 규칙베이스추론(Rule-Based reasoning)을 통합하여 추론하는 CARUBA 시스템을 <그림 1>과 같이 구성하였다.

대부분의 에이전트들은 환경변화에 대한 정보들을 관리할 필요가 있으나 사례, 특징, 성질 등의 자료를 브라우저에서 자동적으로 처리할 수 없다. 그러므로 사용자들이 쉽게 이용할 수 있는 사용자 인터페이스를 이용하여 정보를 받아들이고 보여주는 것이 필요하다. 이러한 기능의 인터넷 프로그램 도구로써 XML을 이용한다. XML은 대부분의 표준들을 포함하고 있어 인터넷 개발에 좋은 도구로써 사용될 수 있다.

사용자 에이전트는 사용자와 시스템과의 인터페이스로서 사용자의 질의를 받아 사용자의 요구사항을 분석하고, 사용자 프로파일과 유사도를 계산하기 위한 정보를 추출한다. 그리고 검

색 에이전트로부터 사용자가 요구하는 정보를 전달하거나 통합에이전트에서 찾은 사용자 원하는 정보를 사용자 인터페이스를 통해 전달한다.



<그림 1> CARUBA 시스템의 구조

검색에이전트는 사례베이스와 규칙베이스를 관리하고, 유사도를 계산한다. 사용자 에이전트에서 보낸 정보를 분석하고 분석된 정보의 유사도를 측정하여 사례베이스나 규칙베이스를 이용하기 위한 판단을 한다. 즉, 규칙베이스의 규칙에 의해 분석된 정보를 사례베이스가 이용하여 유사도에 가장 근접한 정보를 검색하여 그 처리된 결과를 사용자 에이전트로 전달한다.

4. 추론과정

4.1 추론

① 문제분석

사용자로부터 입력받은 정보를 적합한 규칙과 사례를 찾기 위해 주어진 문제를 분석한다. 사용자 에이전트에 의해 실행되며, 검색 에이전트에 정보가 전달된다.

② 규칙기반추론

사용자 에이전트에 의해 분석된 정보를 이용하여 규칙을 생성하고, 규칙 충돌의 해결 및 새로운 규칙을 규칙베이스에 저장하여 적절한 결론을 유도하는 역할을 한다.

③ 사례검색

사용자의 입력정보에서 추출한 문제분석 정보를 검색 에이전트가 사례베이스에 적합하다고 결정되면 적용된 규칙을 이용하여 규칙과 관련된 사례 중 가장 유사한 사례를 추출한다. 즉, 검색에이전트는 사례를 추론하기 위해 사용자 에이전트가 보낸 정보로 유사도 산출 알고리즘을 통해 관련 유사도를 산출하고 사례검색을 통해 추출된 사례에 적용하여 사례적용을 한다.

④ 사례적용

검색된 여러 개의 사례를 유사도 적용 기준에 가장 만족하는 하나의 사례를 선택한다. 또한 시스템에서 사용자들의 평가 결과를 반영하기 위해 성공 횟수를 사례에 첨가함으로써 제안된 횟수가 많은 사례를 선택할 수 있도록 한다.

⑤ 사례조정

규칙기반 추론에서 제안한 결론과 검출된 사례를 조정하여 사용자에게 해를 제안하는 역할을 수행하며, 새로운 사례를 사례베이스에 저장한다. 만약 사례베이스에서 사용자 질의와 유사한 사례를 검색하지 못한 경우 규칙베이스의 제안을 따른다.

4.2 유사도

Near Neighbor(NN)기술은 유사성검색을 위한 가장 오래되고 많이 이용되는 기술중의 하나이다. NN의 검색단계는 미리 정해진 유사도 기준에 따라 사례베이스에 저장된 사례들 중에 유사한 사례를 찾아 추천해준다. 유사도 측정과 가중치 조정은 사례를 측정하는데 중요한 요소이다.

$$SIM(q,c) = \sum_{i=1}^n w_i S_i(q_i,c_i) \dots (1)$$

SIM(q,c)는 질의(q)와 사례(c)사이의 전체유사도, S_i 는 속성 i에서 질의와 사례 사이의 부분 유사도이다. w_i 는 속성 i에 주어진 가중치이다. <식 1>은 부분 유사도 가중치의 합이다. 이것은 서로 다른 속성에서의 사용자 선호도를 보여준다[3]. 부분 유사성 함수는 추론사례의 특징 값들에 기초하여 적정 가중치를 재 정렬한다. 따라서 사용자 프로파일에 따라 관련성이 증가할 수도 있고 다른 속성에 의해 내려갈 수도 있다. 다시 말해서 부분 가중치는 선호도에 따라 변화될 수 있다. 다음 수식은 편차(D_i)를 나타낸 것이다.

$$D_i = \omega_i(1 - S_i(q_i,c_i)) \dots (2)$$

<식 2>에서 ω_i 는 속성 i의 상대적 중요도를 나타낸다. 사용자 속성에 의해 결정할 수 있으며 속성들을 평가할 때 이용된다.[3] 예를 들어 사용자들이 Style속성에 비중을 둔다면 Style의 가중치가 높아 질 것이다. <표 2>와 <표 3>은 사용자들이 선호하는 속성에 따른 가중치의 변화를 보여주고 있다.

<표 2> 가중치를 조절하기 전

Aattribute	Weight	Local Simmilarity	Deficiency
Number	0.20	0.90	0.02
Company	0.20	0.85	0.03
Style	0.20	0.80	0.04
Color	0.20	0.75	0.05
Price	0.20	1	0.0

<표 3> 가중치를 조절한 후

Aattribute	Weight	Local Simmilarity	Deficiency
Number	0.24	0.90	0.02
Company	0.19	0.85	0.03
Style	0.18	0.80	0.04
Color	0.17	0.75	0.05
Price	0.20	1	0.0

5. 실험 및 고찰

본 논문에서 제안한 CARUBA 시스템의 구조와 방법론을 평가하기 위해 소비자가 핸드폰을 구입할 때 추천을 해주는 에이전트를 웹에서 구현하였다. 사례베이스는 Number, Company, Style, Color, Price 등으로 구성되며, 103가지의 사례를 들었다. 5가지의 번호, 4가지의 brand, 4가지의 color, 2가지 style로 만들었다. 각 속성의 부분 유사도는 0과 1사이에서 정의되고 사용자 에이전트가 처리하게 된다. <표 5>는 사례베이스에 저장된 데이터의 일부분이다.

<표 4> 사례베이스에 저장된 데이터

Number	Price	Company	Style	Color	횟수
011	23만원	삼성	폴더	Red	150
011	23만원	SK	폴더	Red	130
019	15만원	한화	플립	Red	95
017	28만원	SK	플립	white	125
011	32만원	삼성	플립	Silver	82
018	40만원	SK	폴더	white	78

<표 5>는 사용자의 질의 정보의 예이다.

<표 5> 인덱스 예

Number	price	company	style	color
011	25만원	SK	폴더	red

<표 6>은 사용자의 질의에 대한 규칙베이스의 규칙을 정의하고, 사례베이스에 의해 추론하기

위한 구성 형식을 보여주는 예이다.

<표 6> 규칙베이스의 규칙

If Number(N)=011 and
Color=Red and Style=폴더
Then Price(P)= 23만원

<표 7>은 규칙베이스에서 정의한 정보를 이용하여 사례베이스에서 추론된 결과를 보여준다.

<표 7> 사례베이스 검색 예

	Number	price	company	style	color
A	011	23만원	SK	폴더	red
B	011	23만원	LG	폴더	red
C	017	25만원	삼성	폴더	white

5.1 유사도 계산

<표 7>에서와 같은 사례베이스 추론결과에 따라 다음과 같이 유사도를 계산한다.

· 번호의 유사도 : 가입자가 적은 번호순 019(1), 018(2), 016(3), 017(4), 011(5)로 한다.

$$\text{번호-가중치} * (1 - \frac{\text{고객의번호} - \text{사례베이스번호}}{5})$$

· 가격의 유사도

$$\text{가격} = \text{가중치} * (1 - \frac{\text{고객이원하는가격} - \text{사례베이스가격}}{5})$$

· 회사의 유사도 : 고객이 원하는 회사와 일치하면 가중치를 주고 일치하지 않을 경우 가중치를 달리 적용

· 스타일의 유사도 : 폴더(1), 플립(2), 모를 경우(3)

$$\text{스타일-가중치} * (1 - \frac{\text{고객이원하는스타일} - \text{사례베이스스타일}}{3})$$

· 컬러의 유사도 : 고객이 원하는 컬러와 일치하면 가중치를 주고 일치하지 않을 경우 가중치를 달리 적용한다.

<표 8>은 유사도 계산식에 의한 계산된 유사도를 나타낸다. 총 유사도가 높은 사례번호 1번

이 선택되어 사용자에게 보여진다.[3]

<표 8> 유사도 계산의 예

No.	Number	price	company	style	color	총유사도
1	0.24	0.12	0.19	0.18	0.17	0.9
2	0.24	0.12	0	0.18	0.17	0.71
3	0.192	0.24	0	0.18	0	0.612

6. 결론 및 향후과제

사례베이스와 규칙베이스를 통합한 검색 에이전트와 사용자들이 편리하게 이용할 수 있는 사용자 에이전트의 멀티 에이전트를 이용하여 시스템을 구성하였다. 검색 에이전트는 사용자 에이전트가 보낸 사용자 프로필 정보를 통해 유사도를 이용하여 결정을 쉽게 할 수 있으며 사용자가 원하는 가장 유사한 정보를 검색할 수 있다. 향후 연구 과제로는 본 시스템의 실용화를 위해서 복잡한 사용자 질의에도 응할 수 있는 다양하고 풍부한 사례베이스의 구축이 요망된다.

7. 참고문헌

[1] 이재필, 이상용, 김기태 “규칙기반 추론과 사제기반 추론의 결합” 한국정보과학회 가을 학술발표논문집 Vol.21, No. 2, 1994 pp.397

[2] 허철희, 조성진, 정환목, “사례기반추론 에이전트를 이용한 전자상거래” 한국전자거래 학회지(CALS/EC) 제5권 제2호 p.56~58

[3] Yanping Ma, Esmat Aimeur, “Intelligent Agent in Electronic Commerce-XMLFinder IEEE, 2001 p.275,276

[4] Katsumi Nihoi 외3인, “ExpertGuide for help Desk-An Intelligent Information Retrieval System for WWW Pages’, Proceedings of the Ninth International Workshop on Database and Expert Systems Applications, IEEE Computer Society, PP. 937-942, 1998

[5] T. Rolf, Wigand and I, Robert. Benjamin, “Electronic commerce “Effects on electronic Markets”, JCMC, Vo 1, 1999