

<초청논문>

XRML 기반 비교쇼핑몰의 구조와 배송비 산정에 관한 실증분석 Architecture of XRML-based Comparison Shopping Mall and Its Performance on Delivery Cost Estimation

이재규*, 강주영**

한국과학기술원 테크노경영대학원
{jklee*, jy kang**}@kgsm.kaist.ac.kr

Abstract

인터넷 쇼핑이 성장하면서 비교쇼핑몰의 중요성도 그만큼 증가하고 있다. 그러나 현재의 비교쇼핑몰들은 웹으로부터 제품가격정보와 같은 단순한 XML 데이터만을 수집하여 비교할 뿐, 배송비와 같이 규칙을 기반으로 한 정확한 비교는 제공하지 못하고 있다. 각 쇼핑몰은 저마다 배송비에 관한 정책을 다양하게 제시하고 있으나 비교쇼핑몰에서는 이를 반영할 구조를 가지고 있지 못하기 때문이다.

따라서, 본 연구에서는 규칙 기반의 추론을 이용해 상품가격에 배송비를 포함하여 비교를 수행할 수 있는 비교쇼핑몰을 구현하고자 한다. 이를 위해 eXtensible Rule Markup Language (XRML)를 이용하여 각 쇼핑몰의 웹 페이지의 문장과 표로부터 규칙을 습득하는 방안을 제시하였다. 이 구조를 이용하면 웹페이지에서 완전히 가까운 규칙을 자동생성할 수 있을 뿐만 아니라, 각 사이트에서 변화가 발생하면 이를 반영하여 규칙을 일관성있게 수정하도록 지원할 수 있다.

본 연구에서는 인터넷 상의 대표적 서점인 Amazon.com, BarnesandNoble.com, Powells.com에 대해 XRML을 기반으로 설계한 비교쇼핑몰의 프로토타입 ConsiderD를 개발하였다. 이 과정에서 웹으로부터 규칙을 자동생성할 수 있는 잠재력을 검정하고, 배송비 효과의 중요성을 실험을 통해 예시하였다.

1. 서론

인터넷과 웹의 발전으로 인해 매년 전자상거래 시장은 폭발적으로 성장하고 있다. 이와 더불어 인터넷 쇼핑몰에서 획득 가능한 상품 종류의 수와 정보의 양도 엄청난 수로 증가하고 있다. 따라서 수많은 상품 정보 중 고객들이 원하는 제품을 검색하는 데 있어 시간과 비용을 줄이도록 도와주는 비교쇼핑몰의 역할이 점점 더 중요해지고 있다. Forrester (2001)의 조사에 따르면 2005년도에는 온라인 매출의 48%가 비교구매를 통해 형성될 것으로 예측되고 있다.

이러한 비교쇼핑몰들이 하는 일은 사용자가 선택한 물건들에 대하여 각각의 쇼핑몰들의 가격을 비교해 줌으로써, 사용자가 동일한 물건에 대해 최저의 가격으로 구매할 수 있도록 돕는 것이다. 그런데, 일반적으로 사용자가 지불해야 할 최종가격은 배송비용을 포함하고 있다. 대개의 비교쇼핑몰들은 동일한 고정가격으로 배송비용을 계산하여 비교하고 있지만, 쇼핑몰에 따라서 배송비용이 서로 다르며, 같은 쇼핑몰 내에서도 사용자가 선택한 물건의 조합, 배송 방법, 배송 지역에 따라 차이가 난다. 따라서, 비교쇼핑몰에서 보다 정확한 가격 비교기능을 제공하기 위해서는 쇼핑몰마다 각기 다르게 계산되는 배송비용을 포함하여야 한다.

그러나 현재의 비교쇼핑몰들은 웹으로부터 제품 정보나 가격 정보와 같은 단순한 데이터 형태만 수집하여 비교할 뿐, 배송정책과 같은 복잡한 형태의 규칙을 수집하여 규칙베

이스를 구축하고, 이러한 규칙베이스를 바탕으로 한 추론을 통해 최종가격을 계산하여 비교하지는 못하고 있다. 각 쇼핑물은 배송비에 대한 정책을 각기 다르게 웹 페이지에 설명하고 있지만, 이들 웹 페이지로부터 추론가능한 규칙을 습득하는 것은 쉬운 일이 아니다. 예를 들어 아마존의 경우, 배송비용을 설명하고 있는 웹 페이지가 무려 21개나 된다. 규칙의 습득 외에 또 어려운 점은 배송비용에 대한 규칙이 바뀌었을 때 이를 반영하기가 매우 어렵다는 점이다.

따라서, 본 연구에서는 확장형 규칙 표식 언어(eXtensible Rule Markup Language, 이하 XRML) [16] 구조를 이용하여 각 쇼핑물의 웹 페이지의 문장과 표로부터 규칙을 자동습득하는 방안을 제시한다. 이 구조를 이용하면 웹페이지에서 완전히 가까운 규칙을 자동생성할 수 있을 뿐만 아니라, 각 사이트에서 변화가 발생하면 이를 반영하여 규칙을 일관성있게 수정할 수 있도록 지원할 수 있다. 이렇게 수집된 규칙을 기반으로 한 추론을 통해 배송비를 계산하고 정확한 가격비교를 제공하기 위해서는 웹 기반 전문가시스템 구조가 요구된다. 이를 위해 본 연구에서는 대표적인 인터넷 서점인 Amazon.com, BarnesandNoble.com, Powells.com에 대해 XRML 기반 비교쇼핑물인 ConsiderD를 설계 및 구축한다. 마지막으로, 배송비의 효과성을 측정하기 위해 실험을 통해 일반 비교쇼핑물과의 배송비 차이에 대해 살펴본다.

본 연구의 구성은 다음과 같다. 먼저 2장에서 비교 쇼핑 및 지식 습득 관련 연구를 조사하고 3장에서 XRML을 활용한 지식 습득 방법론에 대해 살펴볼 것이다. 4장에서는 본 연구에서 제안한 방법론으로 구축한 ConsiderD라는 배송비를 포함한 가격 비교쇼핑물에 대해서 살펴보고 5장에서는 배송비의 효과에 대해 예시하고 끝을 맺는다.

2. 관련 연구

2.1 비교쇼핑 연구

Guttman (1998)이 제시한 소비자 구매의사 결정 모델을 살펴 보면, 인터넷에서의 구매의

사결정 단계 중 비교 쇼핑이 많은 비중을 차지하고 있음을 알 수 있다. 온라인 비교 쇼핑 기능은 여러 인터넷 쇼핑물로부터 사용자가 원하는 물건을 검색하여 사용자가 원하는 기준 (예를 들면 가격)으로 정렬해 줌으로써 사용자의 구매에 도움을 주는 기능이다. 이를 위해서는 웹 상의 상품 정보를 자동으로 추출하는 기능이 필요하다. 이러한 온라인 비교 쇼핑 연구 중 가격 비교 쇼핑 에이전트는 구매자가 원하는 상품을 최적의 가격으로 구매할 수 있도록 인터넷 쇼핑물들을 검색하여 가격 등의 정보를 찾아주어 비교구매를 할 수 있도록 도우미 역할을 대행해주는 기술이다 [1]. 널리 알려진 비교 쇼핑 에이전트로는 Bargain Finder와 Shopbot [9]이 있으며, 이러한 기술을 이용한 비교 쇼핑 사이트로는 mySimon.com, Compare.net 등이 있다.

그러나 이들 비교쇼핑 에이전트들은 단순히 물품 및 가격 정보만을 추출하여 내부 DB로 구축하여 제공해 줄 뿐 배송비와 같은 복잡한 형태의 지식을 추출하지 못하기 때문에 배송비를 포함한 정확한 가격비교를 제공해주지 못하고 있다.

2.2 지식습득 연구

웹 페이지로부터의 규칙 습득은 자연어로부터의 지식 획득과 많은 유사한 점을 갖고 있다. 기존의 지식 획득 방법론을 크게 지식분석도를 이용한 지식 획득 방법론, 지식획득을 위한 자연어 처리, 자동학습으로 나누어 조사하고 본 연구의 접근 방법과 비교하였다.

2.2.1 전문가 다이어그램 접근 방식

오랜 기간 동안 지식의 습득은 전문가 시스템의 구축과정에서 병목으로 자리해왔다. 특히 규칙 전문가와 도메인 전문가 사이의 의사소통이 원활하지 않을 경우에는 더욱 어려운 작업이었는데, 전문가 다이어그램 [17]은 도메인 전문가가 도메인 지식을 구조화된 형태로 표현하는 것을 돕기 위해 많이 사용되었다. 전문가 다이어그램은 추론과정을 시각적으로 표현함으로써 논리적 오류 없이 지식을 구체적으로 표현할 수 있도록 지원하였

다. 전문가 다이어그램 외에도 conceptual graph [3, 26], decision table [20, 24], influence diagram [6, 15, 19] 등이 지식 습득연구에서 이용되어 왔다. 이들 다이어그램은 전문가의 지식과 실제 추론에 사용되는 추론 규칙의 중간 매체로서 사용되었으며, 추론규칙으로 자동 변환될 수 있기 때문에 쉽게 규칙과의 연관관계를 보일 수 있다. 그러나 원본 문서와 규칙을 연결시키지는 못하기 때문에 둘 간에 일관성을 유지해주지 못하는 한계가 있다.

2.2.2. 온톨로지 획득을 위한 자연어 처리

Wetter and Nüse (1992)가 지적한 바와 같이 자연어는 그 중요성으로 인해 기계에 의해 완전하게 해석하는 것이 거의 불가능하다. 따라서 웹 페이지로부터 규칙을 습득하는데 있어서도 자연어 처리를 이용하는 것은 분명히 한계가 있다. 그럼에도 불구하고 최근 웹에서 특정 도메인에 대한 지식을 획득하기 위한 도구로 온톨로지가 활발하게 사용되고 있다 [10, 22, 27]. 도메인이 한정되어 있는 경우에, 온톨로지는 적절한 어휘를 선택하는데 도움을 줄 수 있다 [8, 13, 28]. 또한 지식이 서술되어 있는 문법적 패턴을 온톨로지 스키마로 정의하고 온톨로지의 형태로 지식을 자동 추출하는 연구에서 자연어 처리가 활발하게 사용되고 있다 [5, 18, 21, 22]. 그러나 자연어로부터 추출된 지식은 불완전하기 때문에, 지식관리자의 수정 및 보완이 필수적으로 요구되고 있다 [23, 29].

XRML 접근방식의 기본적인 아이디어는 자연어 처리가 아직은 완전하지 못하다는 가정 아래 지식관리자의 규칙습득과정을 보조하고자 하는 것이다. 그러나 자연어 처리의 기본적인 기술들이 웹 페이지의 문장으로부터 규칙을 식별하는 과정에서 규칙을 구성하는 단어나 동어를 검색하거나 규칙을 편집하는데 사용되었다.

2.2.3. 자동학습기법

텍스트로부터 지식을 습득할 때 패턴 분석, 연역적 학습, 인공지능망 그리고 통계적 모델과 같은 자동학습 기법이 많은 연구 및 도구에 활용되었다 [2, 4, 7, 14, 25]. 즉, 데이터의 구조

적인 집합이 일정한 틀로 표현되어 있을 때, 이러한 데이터의 틀을 이용하면 보다 일반화되고 추상화된 지식들을 추출하는 것이 가능하다. 그러나 대부분의 학습방법론은 추론 가능한 규칙보다는 간단한 개념이나 데이터에 가까운 지식 획득을 목표로 하고 있기 때문에 자연어로 된 텍스트와 표로부터 추론이 가능한 규칙을 습득하고자 하는 경우에는 이러한 방법론이 적절하지 않다. 따라서 자동학습을 이용한 지식의 추출과 XRML을 이용한 규칙의 추출은 그 추출 대상이 다르다고 할 수 있다.

이와 같이 기존의 지식 습득 방법론은 배송비와 같은 복잡하고 추론 가능한 형태의 규칙을 추출하는 방법을 제공해주고 있지 못하다. 또한, 추출된 규칙과 원문간의 일관성 유지를 지원하지 못한다. 이러한 한계점을 극복하기 위해 본 연구에서는 XRML을 활용한 지식습득 방법론을 채용함으로써 각 쇼핑몰로부터 배송비를 추출하여 정확한 가격비교를 제공해주는 가격비교 쇼핑몰을 구축하였다.

3. XRML을 활용한 지식습득 방법론

3.1 개요

본 연구에서는 자연어와 표를 중심으로 표현되어 있는 웹 페이지로부터 규칙을 습득하고 이를 관리하기 위해 XRML을 이용한 방법론을 제안하였다. XRML은 웹 페이지로부터 규칙과 규칙구성요소를 식별하기 위한 언어인 규칙 식별 표식 언어 (Rule Identification Markup Language, RIML)와 추론 가능한 규칙을 구조적으로 표현하기 위한 규칙 구조 표식 언어 (Rule Structure Markup Language, RSML)로 구성되어 있다.

XRML을 이용하여 지식을 습득하는 과정은 규칙에 대해 설명하고 있는 웹 페이지로부터 규칙구성요소를 추출하여 RIML문서를 작성하는 것으로 시작된다. 다음 과정에서 RIML의 형태로 식별된 규칙구성요소들은 RSML에서 정의하고 있는 추론 가능한 규칙

표현양식으로 자동 변환된다. 그러나 자동 생성된 규칙은 초안으로서 완전한 규칙은 아니다. 이러한 초안을 보완하여 완성된 RSML 문서는 일반적인 규칙기반시스템에서 사용되는 규칙으로 자동변환이 가능하다. <그림 1>은 이와 같은 규칙습득과정을 보여주고 있다.

RIML을 이용한 규칙식별을 통해 얻을 수 있는 혜택은 규칙식별의 결과로부터 규칙에 가까운 규칙 초안을 자동으로 생성할 수 있다는 점이다. 자동 생성된 규칙 초안은 규칙을 수정 및 보완하여 완전한 규칙으로 변환하여야 하는 지식관리자의 부담을 크게 줄일 수 있다. 또한 RIML과 RSML간의 연결성을 이용하여 웹 페이지에 변경이 발생했을 때 이를 쉽게 규칙베이스에 반영할 수 있도록 지식관리자에게 적절한 정보를 제공할 수 있다.

3.2 XRML을 이용한 규칙 습득 과정

XRML을 이용한 규칙습득과정은 구축하고자 하는 지식기반시스템의 규칙베이스를 설계하고 해당 웹 페이지를 선정하며, RIML을 이용하여 웹페이지로부터 규칙을 식별하는 과정과 식별된 규칙들을 RSML 문법에 맞도록 구조화하는 과정으로 이루어져 있다. 규칙의 구조화 단계는 다시 XRML 편집기에 의해 자동으로 규칙의 초안을 생성하는 과정과, 지식관리자가 수동으로 규칙을 완성하는 단계로 나누어진다. 각 과정에 대한 다음의 자세한 설명에서는 아마존으로부터 규칙을 습득하는 실제 예를 통해

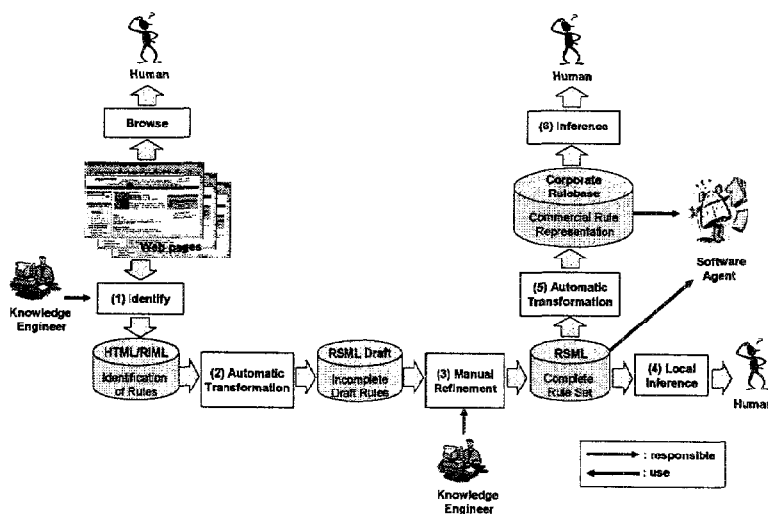
이해를 돕고자 한다. 습득된 규칙은 다음 장에서 보여줄 XRML 기반의 비교쇼핑몰인 ConsiderD를 구축하는데 이용된다.

3.2.1 규칙베이스의 설계 및 웹페이지의 선정

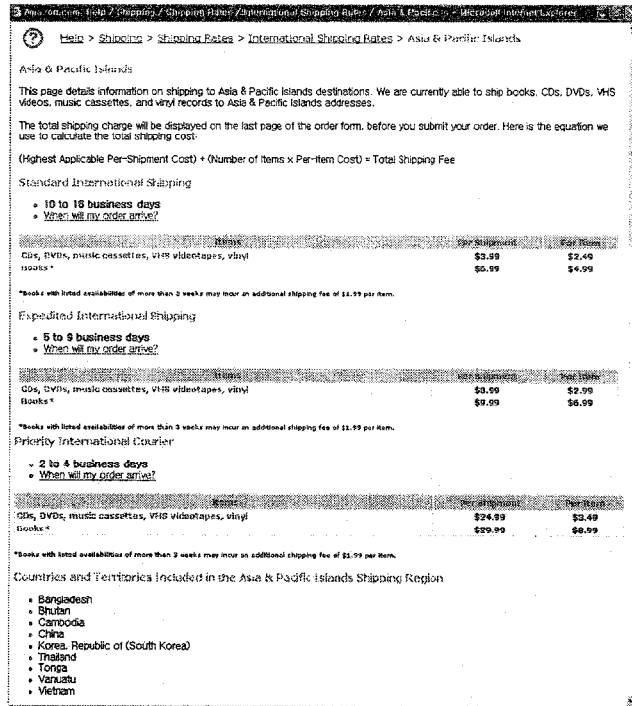
규칙을 습득하기에 앞서 전체적인 관점에서 구축하고자 하는 지식기반 시스템에 따라 규칙 베이스의 목표와 주제를 선정한다. 그런 다음 규칙 베이스를 구성하는 규칙의 그룹들을 설계한다. 규칙베이스의 큰 틀에 대한 설계가 끝나면 규칙베이스 설계에 맞추어 실제로 규칙을 습득할 웹페이지를 찾아 규칙베이스의 그룹과 웹페이지의 URL을 연결한다.

예를 들어 아마존 웹 페이지로부터 규칙을 습득하여 배송관련 전문가시스템을 구축하고자 할 때 가장 먼저 할 일은 배송과 교환 및 환불정책에 대해 설명하고 있는 웹 페이지를 아마존 웹사이트에서 찾는 것이다. 실제로 아마존에서 배달과 관련된 웹 페이지 21개를 찾을 수 있었다. 이들 웹 페이지는 배송이 가능한 지역, 배송 요금, 무료배송조건, 배송에 소요되는 시간, 그리고 교환 및 환불이 가능한 조건을 명시하고 있다. 여기 설명된 규칙은 *Shipping Rates*, *Modifying Orders*, *Shipping Guide*, 그리고 *Returns and Refunds*의 4개 범주로 분류될 수 있었다.

3.2.2 RIML을 이용한 규칙의 식별 단계



<그림 1> XRML을 이용한 규칙습득 과정



<그림 2> Amazon.com의 웹페이지

다음 단계에서는 선택된 웹 페이지들로부터 RIML 태그를 이용하여 규칙의 구성요소를 식별한다. 즉, 웹 페이지로부터 *RuleGroups*, *Rules*, *variables*, *values*, *IF-THEN* 관계 및 *AND*, *OR*와 같은 접속사를 식별하고, HTML 문서 위에 RIML 태그들을 추가하는 형태로 HTML/RIML 문서를 생성한다.

예를 들어, <그림 2>와 같이 아시아 지역으로의 배송요금을 표시하고 있는 웹페이지를 살펴보면 선택된 웹 페이지에 규칙을 설명하고 있는 자연어 문장들과 수식, 그리고 *Standard International Shipping*, *Expedited International Shipping*, *Priority International Shipping*으로 이루어진 세 개의 표가 있음을 볼 수 있다. 이러한 웹 페이지로부터 배송 요금이 배송 지역과 구매한 품목의 종류, 구매한 품목의 수량, 배송 방법에 따라 결정되는 것을 알 수 있다. <그림 3>은 선택된 웹 페이지에 RIML태그를 이용하여 규칙의 구성요소를 식별한 결과를 이탤릭체로 보여주고 있다.

이 예에서 웹 페이지 전체가 배송비용에 대해 설명하고 있기 때문에 전 범위를 *title = Shipping Rates*인 하나의 *RuleGroup*으로 식별하였다. 웹 페이지에서 처음으로 식별된 규칙은

<Rule rid=1>로 표기되었으며, 첫 변수는 <variable vid=1>able to ship</variable>로 식별되었다. 이 변수의 변수값인 True는 웹 페이지 상에 표시되지 않고 생략되었기 때문에 지식관리자에 의해 <value vid=1 name="True">로 식별되었다.

첫째 변수와 변수값은 규칙에서 *THEN* 절에 속하기 때문에 다시 <*THEN*>으로 식별되었다. 둘째 변수인 *Items*는 *books*, *CDs*, *DVDs*, *VHS videotapes*, *music cassettes*, *vinyl*의 여섯 개의 변수값에 의해 공유되고 있다. 공유된 변수인 <*Items*>가 *Rule 1* 과 *2*에서 각 변수값마다 반복적으로 사용되고 있음을 <그림 4>의 RSML 문서에서 볼 수 있다. RIML에서는 *Items*가 공유된 변수임을 나타내기 위해 여섯 개의 변수값에 *Items*의 *vid*를 동일하게 부여하였다.

이상에서 설명한 방식에 따라 아마존으로부터 4개의 *RuleGroup*, 120개의 *Rule*, 35개의 *RuleTable*, 313개의 *variable*, 808개의 *value*, 13개의 *operator*, 119개의 *IF* 절, 120개의 *THEN* 절, 107개의 접속사를 식별하였다. 그 결과, 총 1,635개의 규칙구성요소를 식별하였다 (표 1 참조).

3.2.3 RSML 규칙 초안의 자동생성 단계

앞 단계에서 생성된 RIML 문서의 웹 페이지로부터 식별된 규칙들은 RSML 문법의 규칙들로 자동 변환한다. 이렇게 자동으로 변환된 결과는 아직 불완전한 형태의 규칙 초안이 된다.

<그림 3>의 RIML문서는 자동생성을 통해 <그림 4>의 규칙 초안으로 변환된다. 아마존에서 식별된 RIML 문서를 이용해 총 2,520개의 규칙구성요소를 생성하였다. 표 1에서와 같이 규칙구성요소 중에서 Rule, RuleTable, IF, THEN, operator는 규칙식별과정에서의 식별된 수(IC)와 규칙 초안에서 자동 생성된 수(GC)가 동일하였다. 그러나 variable과 value는 각기 614개와 119개가 자동으로 추가되었다. 또한 접속사 중에서 AND는 152개가 추가되었다.

```
<HTML...><BODY>
<RIML version="2.0">
<RuleGroup title="Shipping Rates">
<P><H3> Help > Shipping > Shipping Rates > International Shipping
Rates > Asia & Pacific Islands </H3></P>
<URL_riml="http://xrml.kaist.ac.kr/amazon/Shipping_Rates.riml">
.....
<P>
<Rule rid=1>
This page details information on shipping to Asia & Pacific Islands destinations. <THEN>We
are currently <variable vid=1 name="ship">able to ship</variable> <value vid=1 name="True"/></THEN>
<P><OR><variable vid=2 name="Items"/><value vid=2 name="books"/><value vid=2 name="CDs"/><value
vid=2 name="DVDs"/><value vid=2 name="VHS videotapes"/><value vid=2 name="music
cassettes"/><value vid=2 name="vinyl"/></OR> records to <value vid=3 name="Asia & Pacific
Islands"/><value vid=3 name="Shipping Region"/>addresses</variable></IF>
</Rule></P>
.....
<RuleTable>
<P></IF rid="2, 3"><variable vid=4 name="Delivery Method"/>
<value vid=4 name="Priority International Courier"/></IF></P>
.....
<Table>
<TR></IF rid="2, 3"><td><variable vid=5 name="Items"/></td></IF>
<THEN rid="2, 3"><td><variable vid=6 name="Per Shipment"/></td>
<td><variable vid=7 name="Per Item"/></td>
</THEN></TR>
<Rule rid=2>
<P></IF><td><OR><value vid=8 name="CDs"/></value>
<value vid=8 name="DVDs"/></value>
<value vid=8 name="music cassettes"/></value>
<value vid=8 name="VHS videotapes"/></value>
<value vid=8 name="vinyl"/></value></OR></td></IF>
<THEN><td><value vid=9 name="$24.99"/></td>
<td><value vid=10 name="$3.49"/></td></THEN></TR>
</Rule>
<Rule rid=3>
<P></IF><td><value vid=11 name="Books"/></value>* </td></IF>
<THEN><td><value vid=12 name="$29.99"/></td>
<td><value vid=13 name="$8.99"/></td></THEN></TR>
</Rule>
</Table>
</RuleTable>
</P>
.....
<P>
<Rule rid=4><H3><font color="#cc6600">
<IF><variable vid=14 name="Countries"/></IF> and Territories Included in the
<IF rid="2, 3"><THEN><value vid=15 name="Asia & Pacific Islands"/>
<variable vid=16 name="Shipping Region"/></THEN></IF></font></H3>
</P><OR>
<UL>
<LI><value vid=17 name="Bangladesh"/></LI>
<LI><value vid=18 name="Bhutan"/></LI>
.....
<LI><value vid=21 name="Cambodia"/></LI>
<LI><value vid=22 name="China"/></LI>
.....
<LI><value vid=25 name="Korea, Republic of (South Korea)"/></LI>
.....
<LI><value vid=28 name="Vanuatu"/></LI>
<LI><value vid=29 name="Vietnam"/></LI>
</UL></OR></IF></Rule>
</P>
</URL></RuleGroup></RIML></BODY></HTML>
```

<그림 3> RIML이 추가된 HTML/RIML 파일

3.2.4 완전한 RSML 문서의 생성

<그림 4>과 같은 규칙 초안으로부터 문법

적으로뿐만 아니라 의미적으로 완성된 규칙을 만들어내기 위해서는 규칙들을 수정 및 보완하고 새로운 규칙을 추가하여야 한다. 즉, 규칙 식별과정에서 완전하게 식별되지 못한 규칙 구성요소들을 식별하고, 추론이 완결될 수 있도록 RIML로부터 생성된 규칙들과 추론의 결론을 연결하는 규칙을 추가한다.

표1의 rC 행은 규칙을 완성하기 위하여 지식관리에 의해 기존의 규칙에 추가된 규칙구성요소의 수를 보여주고 있다. 예를 들어, <그림 3>의 Rule 2와 3은 <Set_Shipping_Rates> Computed<Set_Shipping_Rates>가 각 규칙의 THEN 절에 추가되어 완성된다. 이와 같은 방식으로 81개의 변수와 변수값이 추가되었다. 그 외에 두 개의 산술연산자와 한 개의 IF 절, 5개의 접속사가 기존 규칙을 완성하기 위해 추가되었다.

```
<RSML version="2.0">
.....
<RuleGroup title="Shipping Rates">
<URL_riml="http://xrml.kaist.ac.kr/amazon/AsiaPacific.riml">
<Rule id=1>
<IF>
<AND>
<OR><Items>books</Items>
<Items>CDs</Items>
<Items>DVDs</Items>
<Items>VHS videotapes</Items>
<Items>music cassettes</Items>
<Items>vinyl</Items>
</OR>
<Shipping_Region>Asia & Pacific Islands</Shipping_Region>
</AND>
</IF>
<THEN><Able_To_Ship>True</Able_To_Ship></THEN>
</Rule>
<Rule id=2>
<IF>
<AND><Delivery_Method>Priority International Courier</Delivery_Method>
<OR><Items>CD</Items>
<Items>DVDs</Items>
<Items>music cassettes</Items>
<Items>VHS videotapes</Items>
<Items>vinyl</Items>
</OR>
<Shipping_Region>Asia & Pacific Islands</Shipping_Region>
</AND>
</IF>
<THEN><AND><Per_Shipment>$24.99</Per_Shipment>
<Per_Item>$3.49</Per_Item></AND>
</THEN>
</Rule>
<Rule id=3>
<IF>
<AND><Delivery_Method>Priority International Courier</Delivery_Method>
<Items>Books</Items>
<Shipping_Region>Asia & Pacific Islands</Shipping_Region>
</AND>
</IF>
<THEN><AND><Per_Shipment>$29.99</Per_Shipment>
<Per_Item>$8.99</Per_Item></AND>
</THEN>
</Rule>
<Rule id=4>
<IF>
<OR>
<Countries>Bangladesh</Countries>
<Countries>Bhutan</Countries>
.....
<Countries>Cambodia</Countries>
<Countries>China</Countries>
.....
<Countries>Korea, Republic of (South Korea)</Countries>
.....
<Countries>Vanuatu</Countries>
<Countries>Vietnam</Countries>
</OR>
</IF>
<THEN><Shipping_Region>Asia & Pacific Islands</Shipping_Region></THEN>
</Rule>
</URL></RuleGroup> ..... </RSML>
```

<그림 4> 그림 3으로부터 생성된 RSML 초안

<표 1> Amazon으로부터 생성된 규칙 구성요소 통계

Rule Component	Identified at RIML Stage (IC)	Automatically Generated from Shared Components (nC)	Automatically Generated from Default Operators (dC)	All Generated from RIML Statements (GC=IC+nC+dC)	Interactively Added for Rule Refinement (rC)	After Refinement (RC=GC+rC)	Effectiveness (%) = (GC/RC)*100	Added to Provide Linkages (lC)	Total in Complete Rules (TC=RC+lC)	Linkage Effort (%) = (lC/TC)*100	Overall Effectiveness (%) = (GC/TC)*100
RuleTable	35	0	0	35	0	35	100.0%	0	35	0.0%	100.0%
Rule	120	0	0	120	0	120	100.0%	4	124	3.2%	96.8%
Variable	313	614	0	927	0	927	100.0%	94	1,021	9.2%	90.8%
Value	808	119	0	927	0	927	100.0%	94	1,021	9.2%	90.8%
Operator	13	0	0	13	3	16	81.3%	1	17	5.9%	75.5%
IF	119	0	0	119	0	119	100.0%	5	124	4.0%	96.0%
THEN	120	0	0	120	0	120	100.0%	4	124	3.2%	96.8%
Connectives	107	4	148	259	5	264	98.1%	4	268	1.5%	96.6%
Total	1,635	737	148	2,520	8	2,528	99.2%	206	2,734	7.6%	92.2%

앞서 설명한 바와 같이 추론의 결론에 도달할 수 있도록 하는 규칙과, 규칙들을 서로 연결하기 위해 필요한 절은 웹 페이지에 명시되지 않기 때문에 RSML 단계에서 추가되어야 한다. 본 예에서는 4개의 규칙이 추가되었다. 새로 추가된 규칙에는 총 44개의 규칙구성요소가 관련되어 있으며 표1의 nC 행이 각 구성요소 별로 추가된 수를 보여주고 있다. 이 예에서는 완전한 추론을 위해 필요한 규칙구성요소의 92.2%가 규칙식별단계에서 식별되었다.

이상에서 제시된 규칙 습득 과정 중에서 지식관리자는 규칙베이스의 설계, 규칙의 식별, 규칙의 보완 과정을 담당하고 있다. 규칙 초안의 생성 과정이 자동으로 이루어지는 반면, 이 세 과정은 지식관리자가 XRML 편집기에 의해 도움을 받으며 진행하도록 되어있다.

4. XRML기반 비교쇼핑몰 구조 및 구현

본 장에서는 사용자가 선택한 배송방법에 따라 정확한 배송비용까지 포함하여 선택된 상품들의 최종가격을 비교해 주는 비교쇼핑몰인 ConsiderD의 구조 및 활용 방법에 대해 살펴보고자 한다. ConsiderD는 XRML 방법론을 활용하여 구축되었다.

4.1 개요

현재 인터넷 상의 대부분의 비교쇼핑몰들은 각 쇼핑몰의 배송비용에 대해 동일한 고정가격으로 계산해 제공하고 있으나, 쇼핑몰에 따라서 배송비용은 다양한 차이를 보인다. 같은 쇼핑몰 내에서도 사용자가 선택한 물건의 조합, 배송 방법, 배송 지역에 따라 다양한 배

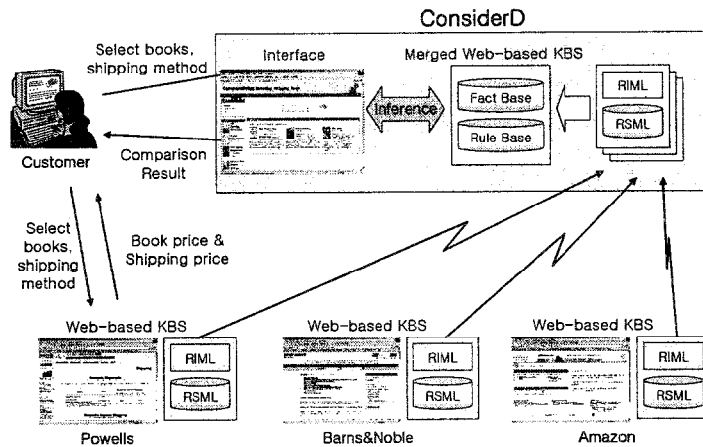
송비용이 있으며, 이러한 배송비용 계산 방법도 쇼핑몰마다 다양한 양상을 보이고 있다. 따라서 사용자의 입장에서 최저가라고 생각하고 직접 쇼핑몰에 가서 주문해보면 비교쇼핑몰에서 제안되었던 가격과 다른 최종가격이 계산되는 경우가 발생하며, 다른 쇼핑몰에서의 최종 가격이 더 낮은 경우도 발생하게 된다. 따라서, 비교쇼핑몰에서는 보다 정확한 최종 가격을 계산하기 위하여 쇼핑몰마다 각 쇼핑몰에서 제공하는 정확한 방법으로 배송비용을 계산하여야 한다.

쇼핑몰들은 배송비용을 계산하는 규칙들을 웹 페이지에서 설명하고 있지만 이 웹 페이지로부터 추론가능한 규칙을 습득하는 것은 쉬운 일이 아니다. 규칙의 습득 외에 또 어려운 점은 배송비용에 대한 규칙이 바뀌었을 때 이를 반영하기가 매우 힘들다.

따라서, 본 연구에서는 XRML을 활용하여 규칙을 습득하고 습득된 규칙을 기반으로 추론을 이용해 상품가격에 배송비를 포함하여 비교를 수행할 수 있는 ConsiderD 시스템을 구현하였다.

4.2 ConsiderD의 구조

ConsiderD의 목적은 온라인 서점들로부터 배송비용과 관련된 규칙들을 습득하여 사용자가 선택한 책과 선택한 배송방법에 대해 책가격과 배송비용을 함께 계산하여 비교하는데 있다. 이를 위하여 Amazon.com, BarnesandNoble.com, Powells.com, 세 개의 온라인 서점을 선택하고, 각 서점의 홈페이지를 검색하여 배송비용을 설명하고 있는 웹 페이지를 찾아내었다. 그리고 이 웹 페이지로부터 세 개의 서점에 대한 배송규칙을 앞서 설명



<그림 5> ConsiderD의 구조도

된 규칙습득 방법론에 따라 습득하였다. 이 과정에서 각 서점 별로 RIML과 RSML 문서가 생성되었다. 생성된 RSML문서는 추론이 가능한 UNIK-BWD 형태의 규칙으로 자동변환되었다.

습득된 규칙을 활용하여 비교쇼핑몰에서 정확한 가격 계산을 하기 위해서는 규칙베이스를 기반으로 추론할 수 있는 웹 기반의 추론시스템이 요구된다. ConsiderD는 이러한 웹 기반 추론시스템을 포함하고 있는 비교 쇼핑몰로서, 앞서 설명된 XRML을 이용한 지식의 공유 구조로부터 설계되었다. <그림 5>는 ConsiderD의 구조를 보여 주고 있다.

ConsiderD의 구조도를 보면, 세계의 인터넷 서점으로부터 규칙을 습득하여 RIML과 RSML이 생성되었으며, 이를 변환 및 통합하여 하나의 규칙베이스를 만들었다. 이 규칙베이스를 기반으로 추론이 진행되며, 추론엔진은 웹 인터페이스와 결합되어 웹 상에서의 추론이 가능하다. 사용자는 세 개의 서점에서 직접 책값과 배송비용을 알아볼 수 있지만, ConsiderD를 이용하면 세 서점에서의 책값과 배송비를 한 곳에서 함께 계산하고 비교하는 것이 가능하다.

4.3 ConsiderD의 활용시나리오

사용자가 ConsiderD를 이용하여 책값과 배송비용에 대한 비교를 하는 과정은 다음과 같다.

- ① 사용자가 ConsiderD에서 구매하고자 하는 책들을 선택하고 장바구니에 담는다.
- ② 사용자가 비교하기를 원하는 쇼핑몰들을 선택한다.
- ③ 각 쇼핑몰들에 대해서 원하는 배송방법을 선택한다.
- ④ 비교쇼핑을 실행하다.
- ⑤ ConsiderD는 정해진 쇼핑몰에 대해서 주어진 책에 대해 추론시스템을 이용하여 가격과 배송비용을 계산하고 결과를 사용자에게 출력한다.
- ⑥ 사용자는 필요한 경우, ConsiderD를 이용하여 계산에 직접 관련된 규칙에 대해 웹 페이지를 참조할 수 있다.

<그림 6>의 화면은 장바구니에서 사용자가 선택한 책들을 보여주고 있다. 이 화면에서는 일반 비교쇼핑몰과 같은 방법으로 배송비용을 보였다. 화면 오른쪽을 보면 각 서점에서의 책 가격과 평균 배송비용이 있는 것을 볼 수 있다. 일반 비교쇼핑몰의 경우, 선택된 책 전체에 대해 배송규칙에 따라 배송비용을 계산하여 보이지 않고 이와 같이 각각의 책 별로 평균 배송비용을 표시한다.

<그림 7>의 화면은 ConsiderD에서 사용자가 비교를 원하는 온라인 서점과 배송방법을 선택하는 화면이다. 서점 별로 정해진 배송방법의 수와 이름, 그리고 내용이 다르기 때문에 하나의 배송방법을 선택하는 것이 불가능하다. 따라서 각 서점 별로 배송방법을 선택하도록 하였다.

Compare Price including Shipping Cost - With List - Microsoft Internet Explorer

Shopping

CART

Qty	Description	Book Store	Price	Approx. Cost	Total Cost
1	What Should I Do With My Life Author: Po Bronson Publisher: Random House Publication Date: 2003-01-01 List Price: \$ 24.95	BarnesandNoble	\$ 14.97	\$ 3.5	\$18.47
		Amazon	\$ 14.97	\$ 3.99	\$18.96
		Powells	\$ 14.99	\$ 3.5	\$18.49
1	The Complete E-Commerce Book Design, Build & Maintain a Successful Web-based Business Author: Janice Reynolds, Roya Mofazali Publisher: CMP Books Publication Date: 2000-10-01 List Price: \$ 29.95	BarnesandNoble	\$ 23.96	\$ 3.99	\$27.95
		Amazon	\$ 20.97	\$ 3.99	\$24.96
		Powells	\$ 21	\$ 3.5	\$24.5
1	XML in a Nutshell, 2nd Edition Author: Eliotte Rusty Harold, W. Scott Means Publisher: O'Reilly & Associates Publication Date: 2002-06-15 List Price: \$ 39.95	BarnesandNoble	\$ 31.96	\$ 3.99	\$35.95
		Amazon	\$ 27.97	\$ 3.99	\$31.96
		Powells	\$ 27	\$ 3.5	\$30.5
1	The Lovely Bones: A Novel Author: Alice Sebold Publisher: Little Brown & Company Publication Date: 2002-07-01 List Price: \$ 21.95	BarnesandNoble	\$ 13.17	\$ 3.5	\$16.67
		Amazon	\$ 13.17	\$ 3.99	\$17.16
		Powells	\$ 14.98	\$ 3.5	\$18.48
1	Enterprise Knowledge Management: The Data Quality Approach Author: David Loshin Publisher: Morgan Kaufmann Publishers Publication Date: 2001-01-22 List Price: \$ 49.95	BarnesandNoble	\$ 35.96	\$ 3.99	\$39.95
		Amazon	\$ 49.95	\$ 3.99	\$53.94
		Powells	\$ 32.5	\$ 3.5	\$36

<그림 6> ConsiderD의 Shopping Cart 화면의 예

Compare Price including Shipping Cost - Select Shipping Method - Microsoft Internet Explorer

SHIPPING

COST

> These books are in your wishlist.

Book Store	Book Info	Total Book Price
Amazon	Enterprise Knowledge Management: The Data Quality Approach, \$ 49.95, Qty: 1	\$ 127.03
	The Complete E-Commerce Book Design, Build & Maintain a Successful Web-based Business, \$ 20.97, Qty: 1	
	The Lovely Bones: A Novel, \$ 13.17, Qty: 1	
	What Should I Do With My Life, \$ 14.97, Qty: 1	
BarnesandNoble	Enterprise Knowledge Management: The Data Quality Approach, \$ 39.96, Qty: 1	\$ 124.02
	The Complete E-Commerce Book Design, Build & Maintain a Successful Web-based Business, \$ 23.96, Qty: 1	
	The Lovely Bones: A Novel, \$ 13.17, Qty: 1	
Powells	Enterprise Knowledge Management: The Data Quality Approach, \$ 32.5, Qty: 1	\$ 120.43
	The Complete E-Commerce Book Design, Build & Maintain a Successful Web-based Business, \$ 21, Qty: 1	
	The Lovely Bones: A Novel, \$ 14.98, Qty: 1	

> Select a Shipping Country: Korea, Republic of

> Check bookstores which you want to compare shipping cost, and select a Shipping Method and Shipping Country

Select	Book Store	Shipping Method
<input checked="" type="checkbox"/>	Amazon	Priority International Courier (averages 2-4 days) <small>1. Free Shipping Guarantee: You must buy over \$25, ship everything to a single US address (including P.O. Boxes, APO/FPO, AIC, HI, PR, SVX, Guam, US Possessions, and all non-UPS addresses). You must also choose "Standard Ground Shipping" in 8 business days! See the shipping method drop-down menu above. 2. See Details about Shipping and Handling charges.</small>
<input checked="" type="checkbox"/>	BarnesandNoble	International Express (1 to 5 business days) <small>1. Free Shipping Guarantee: You must buy two or more qualifying items, ship everything to a single US address (including P.O. Boxes, APO/FPO, AK, HI, PR, SVX, Guam, US Possessions, and all non-UPS addresses). You must also choose "US Standard Ground (3 to 8 business days)" from the shipping method drop-down menu above. 2. See Details about Shipping and Handling charges.</small>
<input checked="" type="checkbox"/>	Powells	International Express (2-7 business days) <small>1. Free Shipping Guarantee: Place an order for \$15 dollar or more (not including shipping costs or any applicable taxes) at Powells.com, choose our Economy Mail (or USPS International Surface for packages going to Canada or Mexico), and we'll ship the order free of charge. 2. See Details about Shipping and Handling charges.</small>

<그림 7> ConsiderD의 배송방법선택 화면의 예

Compare Price including Shipping Cost - Shipping Result - Microsoft Internet Explorer

TOTAL COST ISBN: 0124568402 Search

* This is your total cost

BookStore	Book Info	Shipping Info	Total Cost
Amazon	Enterprise Knowledge Management: The Data Quality Approach, \$ 49.95, Qty: 1 The Complete E-Commerce Book Design, Build & Maintain a Successful Web-based Business, \$ 29.97, Qty: 1 The Lovely Bones: A Novel, \$ 13.17, Qty: 1 What Should I Do With My Life, \$ 14.97, Qty: 1 XML in a Nutshell, 2nd Edition, \$ 27.97, Qty: 1 Total Book Price: \$ 127.09	Shipping Method: Priority International Courier Trackable And Insured: Yes Time: 2 to 4 Business Days PerShipment: \$ 29.99 PerItem: \$ 8.99 Shipping Cost: \$ 74.94	\$ 201.97
BarnesandNoble	Enterprise Knowledge Management: The Data Quality Approach, \$ 39.96, Qty: 1 The Complete E-Commerce Book Design, Build & Maintain a Successful Web-based Business, \$ 29.96, Qty: 1 The Lovely Bones: A Novel, \$ 13.17, Qty: 1 What Should I Do With My Life, \$ 14.97, Qty: 1 XML in a Nutshell, 2nd Edition, \$ 31.96, Qty: 1 Total Book Price: \$ 124.02	Shipping Method: International Express Trackable And Insured: Yes Time: 1 to 5 business days PerShipment: \$ 30 PerItem: \$ 5.95 Shipping Cost: \$ 59.75	\$ 183.77
Powells	Enterprise Knowledge Management: The Data Quality Approach, \$ 32.5, Qty: 1 The Complete E-Commerce Book Design, Build & Maintain a Successful Web-based Business, \$ 21, Qty: 1 The Lovely Bones: A Novel, \$ 14.06, Qty: 1 What Should I Do With My Life, \$ 24.95, Qty: 1 XML in a Nutshell, 2nd Edition, \$ 27, Qty: 1 Total Book Price: \$ 120.43	Shipping Method: International Express Trackable And Insured: Yes Time: 2-7 business days PerShipment: \$ 35 PerItem: \$ 0 Shipping cost: \$ 75	\$ 195.43

<그림 8> 배송비가 포함된 비교 결과화면의 예

마지막으로 <그림 8>의 화면은 서점 별로 선택된 책들에 대하여 가격과 배송비용이 출력된 화면이다. 화면 오른쪽에 선택된 배송방법에 따른 배송추적가능 여부 (Trackable and Insured), 배송에 소요되는 시간(Time), 책 가격 합계(Total Book Price), 책 묶음에 대한 배송비용(Per Shipment), 책 권당 배송비용(Per Item), 그리고 최종 비용(Total Cost)이 출력된 것을 볼 수 있다.

5. XRML기반 비교쇼핑몰의 효과 분석

본 장에서는 본 연구에서 개발한 XRML기반의 비교쇼핑몰인 ConsiderD, 즉 배송비를 고려한 가격 비교쇼핑몰이 실제로 효과가 있는지를 실험을 통해 보여주고자 한다.

5.1 실험 설계

실험은 ConsiderD 비교쇼핑몰과 현재 실제로 온라인상에서 운영되고 있는 비교쇼핑몰을 선정하여 둘 간의 배송비용을 비교하는 방식으로 진행하였다. 배송비용 비교를 위한 비교쇼핑몰은 BestWebBuys.com (이하 BestWebBuys)을 선택하였다. BestWebBuys는 서적뿐만 아니라, 음반과 비디오 및 전자제품에 대해서도 비교쇼핑을 제공하고 있다. 대부분의 비교쇼핑몰들이

배송비용에 대해서는 아예 고려조차 없는 반면 BestWebBuys는 표준배송비용을 이용하여 비교해준다는 점에서 보다 나은 쇼핑몰이라 할 수 있다. 비교대상이 되는 온라인서점으로는 Amazon.com, BarnesandNoble.com, Powells.com를 선정하였다.

배송비용은 기본배송비용과 권당 추가되는 배송비용으로 이루어져 있다. 따라서 몇 권의 책을 주문하는가에 따라 배송비용이 달라진다. 그러나, 무료배송의 경우에는 주문한 책의 권수 외에 주문한 책의 총가격에 의해서도 결정된다. 따라서 배송비용을 계산하기 위해서는 주문한 책의 권수와 총가격을 모두 고려하여야 한다. 본 실험에서는 일반적으로 사람들이 주문하는 책의 권수 이내에서 실험하기 위해 책이 1권인 경우부터 10권인 경우까지 나누어 실험하였다. 매 권수마다 적절한 책 가격을 얻기 위하여 30회씩 주어진 권수만큼 무작위로 필요한 책을 데이터베이스로부터 추출하여 평균 책 가격을 구하였다. 이렇게 구해진 책 권수와 총 가격에 대해 BestWebBuys에서의 배송비용과 본 연구에서 구현한 ConsiderD에서의 배송비용을 각각 계산하였다. 세 개의 온라인 서점에 대한 비교를 해야 하므로 각 서점에 대해 모두 계산하였으며, 또한 서점 별로 다양한 배송방법을 제공하기 때문에 서점 내에서는 각 배송방법

별로 배송비용을 계산하였다.

<그림 9>는 책 권수를 5권으로 선택했을 때의 두 비교쇼핑몰간의 배송비용 비교결과이다. 화면에서는 Amazon.com의 결과를 가장 먼저 보여주고 있다. 표 상단에는 주어진 책 권수에 대한 평균 책 가격이 나타나 있고 그 밑의 표는 각 배송방법에 따른 배송비용을 서점의 배송방법 별로 ConsiderD에서 계산된 배송비용과 BestWebBuys에서 계산된 배송비용을 비교하여 보여주고 있다. 이 표를 보면 BestWebBuys에서 사용하는 표준배송비용과 실제 배송비용과는 상당한 차이가 있음을 알 수 있다. 표 아래의 MIN은 배송방법에 따른 다양한 배송비용과 표준배송비용 간의 차이 중에서 가장 작은 차이를 말하며 MAX는 가장 큰 차이를 말한다.

<그림 9> 두 비교쇼핑몰간의 배송비용 결과화면

5.2 효과분석 및 토의

<표 2>는 본 실험에서 제안한 비교쇼핑몰간의 배송비용 차이를 책 권수에 따라 최대차이 (MAX), 최소차이 (MIN), 평균값 (Average) 차이로 분류해 정리한 것이다. 평균값은 배송방법에 따른 정확한 배송비용을 계산해주는

ConsiderD와 BestWebBuys에서 사용하는 평균 배송비용 간의 차이를 평균한 값이다. 이 표를 보면 책의 권수가 많아질수록 실제 배송비용과 평균배송비용을 이용한 배송비용과의 차이가 커지는 것을 볼 수 있다. 가장 큰 차이가 나는 경우, \$79.99까지 차이가 나고 있으며, 책의 총 가격이 \$44.28인데 반해 배송비용의 차이는 \$39.01인 경우도 있었다.

<표 2> 비교쇼핑몰간의 배송비용 차이

# of Books	Amazon			BarnesandNoble			Powells		
	Min	Max	Average	Min	Max	Average	Min	Max	Average
1	\$1.49	\$34.99	\$11.32	\$0.49	\$32.45	\$7.20	\$1.51	\$39.01	\$8.95
2	\$0.51	\$39.99	\$12.16	\$0.90	\$34.90	\$8.03	\$0.02	\$43.02	\$9.70
3	\$0.99	\$44.99	\$13.32	\$0.54	\$37.35	\$8.76	\$1.03	\$47.03	\$11.23
4	\$0.01	\$49.99	\$14.72	\$0.95	\$39.80	\$9.16	\$0.96	\$51.04	\$12.97
5	\$1.51	\$54.99	\$16.66	\$0.56	\$42.25	\$11.67	\$0.05	\$55.05	\$14.94
6	\$3.01	\$59.99	\$18.60	\$2.07	\$44.70	\$13.35	\$0.94	\$59.06	\$17.12
7	\$4.51	\$64.99	\$20.55	\$3.58	\$47.15	\$15.03	\$1.93	\$63.07	\$19.31
8	\$3.99	\$69.99	\$22.49	\$5.09	\$49.60	\$16.71	\$2.92	\$67.08	\$21.50
9	\$2.99	\$74.99	\$24.43	\$5.05	\$52.05	\$18.39	\$3.91	\$71.09	\$23.69
10	\$1.99	\$79.99	\$26.38	\$4.50	\$54.50	\$20.07	\$4.10	\$75.10	\$25.88

BestWebBuys와 ConsiderD의 배송비용 차이가 통계적으로 유의한 지 살펴보기 위하여 t-test를 한 결과, 다음 표와 같은 결과를 얻었다.

<표 3> 서점별 배송비용 차이에 대한 t-test 결과

서점	ConsiderD와 BestWebBuys의 배송비 차이의 평균	t-statistic	p-value
Amazon	\$18.0623	9.7653	0.0000
BarnesandNoble	\$12.9359	10.4934	0.0000
Powells	\$16.4569	8.9028	0.0000

Amazon과 Powells는 9가지의 배송방법을 제공하고 있기 때문에 90개의 배송비용 차이를 얻을 수 있었고, BarnesandNoble은 11가지의 배송방법을 제공하기 때문에 110개의 값을 얻을 수 있었다. 따라서 t-test는 Amazon, BarnesandNoble, Powells에 대해 각각 90, 110, 90개의 표본 값을 사용해 시행하였다. 표에서 알 수 있듯이, 세 온라인 서점 모두 ConsiderD의 배송비용과 BestWebBuys의 배송비용 간에 뚜렷한 차이가 (p-value < 0.0001) 있는 것으로 나타났다. 즉, 규칙을 이용하여 정확한 배송비용을 계산하는 비교쇼핑몰은 현재까지의 비교쇼핑몰에 비해 확실히 차별화된 가격비교를 하고 있음을 알 수 있다.

5.3 실험의 한계

본 실험에서는 권수에 따라 동일한 표본수를 이용하였으며, 배송방법에 대해서도 표본수의 차이를 두지 않았다. 그러나, 보다 정확한 실험을 위해서는 실제 서점에서 구매되고 있는 권수별 분포와 배송방법별 분포를 이용하여 권수별, 배송방법별로 적절한 표본수를 구하는 것이 바람직할 것이다. 이렇게 할 경우, 사람들이 주로 구매하는 행태에 대해 보다 정확한 차이를 구할 수 있다. 다만 본 실험에서의 결과인 t 값이 매우 크기 때문에 분포를 반영한 실험에서도 유의한 차이가 있을 것으로 기대된다.

6. 결론

본 연구에서는 단순한 데이터에 기반한 가격비교뿐만 아니라 배송비와 같은 복잡한 형태의 규칙을 이용하여 정확한 가격비교를 제공할 수 있는 비교쇼핑몰인 ConsiderD를 설계 및 구현하였다. ConsiderD 비교쇼핑몰이 기존의 비교쇼핑몰보다 두드러진 점은 웹기반 추론엔진을 제공하여 가격비교 시 추론을 통해 정확한 비용을 계산해줄 수 있다는 점이다.

ConsiderD 비교쇼핑몰에서 이용되는 배송비 규칙은 XRML 기반 지식습득방법론에 의거하여 각 쇼핑몰의 웹페이지로부터 습득되었다. XRML 접근 방법을 활용함으로써 각 쇼핑몰의 웹페이지로부터 완전에 가까운 규칙을 자동 생성하고, 각 사이트에서 변화가 발생하면 이를 반영하여 규칙을 일관성있게 수정하도록 지원할 수 있었다.

본 연구에서는 실험을 통해 정확한 배송비를 포함한 가격비교가 의미가 있음을 보이고자 하였다. 이를 위해 실제 대표적인 가격비교 쇼핑몰인 BestWebBuys과 본 비교쇼핑몰간의 배송비 차이를 구하여 t -test를 통해 이 차이가 통계적으로 유의함을 보였다. 향후 좀 더 많은 비교쇼핑몰과의 비교를 통한 좀 더 의미있는 분석연구가 추가로 필요하다. 또한 B2C 쇼핑몰 외에 B2B 쇼핑몰과 같이 배송비의 단위가 큰 경우에는 보다 정확한 배송비의 계산이 더욱 큰 효과를 낼 수 있을 것으로 기대된다.

참 고 문 헌

- [1] 양재영, 김태형, 최중민, "MORPHEUS: 확장성이 있는 비교 쇼핑 에이전트", [정보과학회 논문지: 소프트웨어 및 응용], 제28권, 제2호(2001). pp.179-191
- [2] Alani, H., S. Kim, D. E. Millard, M.J. Weal, W. Hall, P.H. Lewis, and N.R. Shadbolt, "Automatic Ontology-Based Knowledge Extraction from Web Documents," *IEEE Intelligent Systems*, Vol.18, No.1(2003), pp.14-21.
- [3] Amati, G., and I. Ounis, "Conceptual Graphs and First Order Logic," *Computer Journal*, Vol.43, No.1 (2000), pp.1-12.
- [4] Apte, C., F. Damerau, and M.S. Weiss, "Automated Learning of Decision Rules for Text Categorization," *ACM Transactions on Information Systems*, Vol.12, No.3(1994), pp.233-251.
- [5] Babowal, D., and W. Joerg, "From Information to Knowledge: Introducing WebStract's Knowledge Engineering Approach," *Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering*, Edmonton, Alberta, May 1999, pp.1525 -1530.
- [6] Boose, J.H., J.M. Bradshaw, J.L. Koszar, and D.B. Shema, "Knowledge Acquisition Techniques for Group Decision Support," *Knowledge Acquisition*, Vol.5, No.4(1993), pp.405-448.
- [7] Craven, M., D. DiPasquo, D. Freitag, A. McCallum, T. Mitchell, K. Nigam, and S. Slattery, "Learning to Construct Knowledge Bases from the World Wide Web," *Artificial Intelligence*, Vol.118, No.1-2 (2000), pp.69-113.
- [8] Crow, L. and N. Shadbolt, "Extracting focused knowledge from the semantic web," *International Journal of Human-Computer Studies*, Vol.54(2001), pp.155-184.
- [9] Doorenbos, R.O. Etzioni, and D. Weld, "A Scalable Comparison-Shopping Agent for the World Wide Web," *Proceedings of the First International Conference on Autonomous Agents (Agents '97)*, Marina Del Rey, CA, 1997.

- [10] Forrester Research, "eCommerce Brokers Arrive", April 5, 2001.
- [11] Guarino, N., "Understanding, building and using ontologies," *International Journal of Human and Computer Studies*, Vol.46(1997), pp.293-310.
- [12] Guttman, R.H., P. Maes, and A.G. Moukas, "Agents as Mediators in Electronic Commerce", *Electronic Markets*, Vol.8, No.1, 1998.
- [13] Hulth, A., J. Karlgren, A. Jonsson, H. Boström, and L. Asker, "Automatic Keyword Extraction using Domain Knowledge," Proceedings of the Second Computational Linguistics and Intelligent Text Processing, Mexico City, Mexico, 2001, pp.472-482.
- [14] Kim, D., H. Jung, and G. Lee, "Unsupervised Learning of mDTD Extraction Patterns for Web Text Mining," *Information Processing & Management*, Vol.39, No.4(2003), pp.623-637.
- [15] Kim, J.D., and J.F. Courtney, "A survey of knowledge acquisition techniques and their relevance to managerial problem domains," *Decision Support Systems*, Vol.4, No.3(1988), pp.269-284.
- [16] Lee, J.K., and M. Sohn, "Extensible Rule Markup Language - toward Intelligent Web Platform," *Communications of the ACM*, Vol.46(May 2003), pp.59-64.
- [17] Lee, J.K., I.K. Lee, S.M. Ahn, and H.R. Choi, "Automatic Rule Generation by the Transformation of Expert's Diagram : LIFT," *International Journal of Man-Machine Studies*, Vol. 30(1990).
- [18] Maedche, A., and S. Stabb, "Mining Ontologies from Text," *Proceedings of the European Knowledge Acquisition Workshop, Lecture Notes in Artificial Intelligence*, Vol.1937, 2000.
- [19] McGovern, J., D. Samson, and A. Wirth, "Knowledge Acquisition for Intelligent Decision Systems," *Decision Support Systems*, Vol.7, No.3(Aug. 1991), pp.263-272.
- [20] Plant, R.T., "Techniques for Knowledge Acquisition from Text," *The Journal of Computer Information Systems*, Vol.35, No.1(1994), pp.64-70.
- [21] Ruiz-Sánchez, J.M., R. Valencia-García, J.T. Fernández-Breis, R. Martínez-Béjar, and P. Compton, "An Approach for Incremental Knowledge Acquisition from Text," *Expert System with Applications*, Vol.25, No.1(2003), pp.77-86.
- [22] Sánchez-Carreño, R., J.T. Fernández-Breis, R. Martínez-Béjar, and P. Cantos-Gómez, "An Ontology-Based Approach Knowledge Acquisition from Text," *Cuadernos de Filología Inglesa*, Vol.9, No.1 (2000), pp.191-212.
- [23] Schmidt, G. and T. Wetter, "Using natural language sources in model-based knowledge acquisition," *Data & Knowledge Engineering*, Vol.26(1998), pp.327-356.
- [24] Seagle, J.P., and P. Duchessi, "Acquiring expert rules with the aid of decision tables," *European Journal of Operational Research*, Vol.84, No.1(Jul. 1995), pp.150-162.
- [25] Soderland, S., "Learning Information Extraction Rules for Semi-structured and Free Text," *Machine Learning*, Vol.34, No.1(1999), pp.233-272.
- [26] Szpakowicz, S., "Semi-automatic Acquisition of Conceptual Structure from Technical Texts," *International Journal of Man-Machine Studies*, Vol.33, No.4(1990), pp.385-397.
- [27] van Heijst, G., A.T. Schreiber and B.J. Wielinga, "Using explicit ontologies in KBS development," *International Journal of Human-Computer Studies*, Vol.45(1997), pp.183-292.
- [28] Vargas-Vera, M., E. Motta, J. Domingue, S.B. Shum, and M. Lanzoni, "Knowledge Extraction by using an Ontology-based Annotation Tool," *Proceedings of the Knowledge Markup and Semantic Annotation Workshop*, Victoria B.C., Canada, 2001.
- [29] Wetter, T., and R. Nüse, "Use of Natural Language for Knowledge Acquisition: Strategies to Cope with Semantic and Programtic Variation," *IBM Journal of Research and Development*, Vol.36, No.3 (May 1992), pp.435-468.