

번들상품추천시스템 개발을 위한 객체지향 사례베이스 설계와 유사도 측정에 관한 연구

정대울

경상대학교 경영학부 경영정보학전공 부교수
(dyeong@nongae.gsnu.ac.kr)

인터넷 쇼핑몰에서 사례기반추천기법을 통한 유사상품의 탐색과 사용자 요구에 적합한 상품추천을 위해서는 다양한 요구에 부응할 수 있는 사례베이스의 구축이 우선되어야 한다. 그리고 구축된 사례베이스로부터 유사한 사례를 검색하여 재사용하거나 필요시 수정하고, 그 결과를 다시 저장하는 기능이 요구된다. 사례기반 상품추천시스템 개발에 있어 가장 중요한 요소는 사례의 표현문제이다. 본 연구에서는 인터넷 수산물 쇼핑몰의 상품추천시스템에서 번들상품 구성문제(집안 이벤트 시 필요한 수산물의 집합)를 표현하는데 적합한 사례표현기법을 개발하며, 유사 사례를 추출하기 위한 유사도 척도의 개발에 연구의 첫 번째 주안점을 둔다. 본 논문에서는 번들상품추천을 위한 사례표현기법으로 객체모델링(OMT)기법을 사용하고 있다. 또한 다양한 사례 속성 유사도 측정방법을 적용하며, 유사도 측정에서 분류법(taxonomy)의 의미와 그 적용방법을 제시한다.

논문접수일 : 2002년 7월

게재확정일 : 2003년 4월

교신저자 : 정대울

1. 서론

최근, 전자 상거래 시장이 본격적으로 형성되고 있으며 이용 빈도 또한 높아지고 있는 추세이다. 하지만 이러한 추세에도 불구하고 현재의 많은 전자상거래시스템은 전통적인 상거래에 비해 제공되는 기능이 상대적으로 부족하여 사용자에게 많은 불편을 주고 있다. Vollrath 등(1998)의 연구에 의하면, 전자 상거래가 상당히 유용하지만 실제로 전자 상거래를 통한 비즈니스의 양이 많지 않은 이유로 ① 인터넷상에서 거래의 불완전한 보안 보증, ② 새로운 미디어

(internet)에 대한 불완전한 수용, ③ 경험과 예술적 감각에 의존한 온라인 카탈로그의 인터페이스 디자인, ④ 원하는 상품을 직접 찾기가 어려움 등을 들고 있다.

본 연구에서는 인터넷 쇼핑몰에서 Vollrath 등(1998)이 제시한 네 번째 문제를 해결하기 위한 상품추천시스템의 개발에 관심을 두고 있다. 즉, 인터넷 쇼핑몰에서 상품탐색에 대한 특별한 지식이 없는 고객이 원하는 상품을 손쉽게 찾을 수 있는 방법의 개발이 필요하다. 그런데 대부분의 인터넷 쇼핑몰에서 상품카탈로그는 데이터베이스에 담겨져 있으므로 원하는 상품을 찾기 위해

* 본 연구는 경상대학교 경영경제연구센터의 학술발표회 지원으로 이루어짐.

서는 상품 속성에 대한 정확한 정보를 사용자가 알고 있어야 한다. 왜냐하면, 일반적으로 사용되는 관계형 데이터베이스에 상품카탈로그를 저장할 경우 특정 질의에 완전히 일치하는 레코드(상품)만을 추출할 수 있기 때문이다. 이러한 문제점을 극복하기 위해서는 유사도(similarity)에 기반을 둔 탐색기법이 요구된다. 사례기반추론(Case-Based Reasoning: CBR)기법은 이러한 문제를 극복하는데 있어 매우 효과적이다(Volrath 등, 1999; Donner & Roth-Berghofer, 1999).

본 논문에서는 고객의 집안이벤트에서 필요한 여러 종류의 상품을 하나하나 상품 카탈로그로부터 찾아서 장바구니에 담는 것이 아니라, 서로 관련성 있는 상품들을 하나의 번들(bundle) 또는 패키지(package) 형태로 구입할 수 있도록 도와주는 번들상품추천 시스템(bundle products recommendation system)의 개발에 있어 CBR기법을 적용하고자 한다. 특정 문제 영역의 시스템 개발에 있어 CBR기법을 적용하기 위해서는 무엇보다도 효과적인 사례베이스의 설계와 추론을 위한 유사도 측정방법의 개발이 중요하다. 사례베이스 설계는 어떠한 사례표현법을 사용하느냐에 따라 그 구조가 결정되며, 추론을 위한 유사도 측정방법에 직접 영향을 미친다.

본 연구에서는 사례표현을 위한 문제영역으로 수산물 인터넷 쇼핑몰의 번들상품추천 분야를 선정하였다. 여기서 번들상품추천 문제는 고객의 특정한 집안이벤트(제사, 생일, 돌, 회갑, 집들이 등)에 필요한 수산물(예, 제수용 어물)을 고객의 요구사항에 가장 적합한 형태로 상품아이템을 구성하는 것이다. 따라서 번들상품추천 문제는 단일의 상품추천문제와 비교하여 복잡한 상품구성 문제를 해결해야 하며, 상품구성에 따라 다양한 번들이 존재할 수 있다. 이와 같이 다양한 변종

을 가진 복잡한 상품구성문제를 해결하는데 있어 상품 맞춤화(product customization)는 매우 중요하다. 사례기반추론기법은 유사한 특성을 지닌 상품의 탐색과 상품 맞춤화에 잘 적용될 수 있다(Schmitt & Bergmann, 1999a; Stahl & Bergmann, 2000).

수산물 인터넷 쇼핑몰에서 CBR기반의 번들상품추천 시스템을 개발하기 위해서는 무엇보다도 복잡한 사례 구조의 표현과 방대한 사례데이터의 처리를 요구한다. 이러한 특성을 갖는 사례베이스를 개발에 있어 객체지향 설계개념의 도입은 문제의 단순화와 개발의 효율성을 높일 수 있다. 특히, 사례베이스 설계를 위한 사례표현에 있어 객체지향 개념의 적용은 표현의 충분성과 의미성을 높일 수 있다. 객체지향 사례표현법은 공통된 사례의 특성에 따라 사례클래스를 형성할 수 있으며, 이들 클래스를 계층구조로 표현할 수 있다. 그리고 계층적으로 표현된 사례클래스들 간의 유사도 측정을 위한 일반적인 방법을 유도해 낼 수 있다(Bergmann & Stahl, 1998).

이에 본 연구에서는 사례베이스의 설계를 위한 기법으로 Rumbaugh 등(1991)의 객체모델링 기법(Object Modeling Technique: OMT)을 사용하고자 한다. OMT의 객체모델은 다양한 표기법을 제공함으로써 표현의 의미론적 충분성을 높여 줄 수 있으며, 관계형 데이터모델과 쉽게 결합할 수 있어 기존의 관계형 데이터베이스 기반의 인터넷 쇼핑몰 시스템에 CBR기법을 손쉽게 도입할 수 있는 개념적 틀을 제공해준다.

2. 상품추천시스템 관련연구

지금까지 상품추천을 위한 대부분의 시스템들

이 사용자의 요구사항과 사용자 프로파일을 기반으로 하고 있다. 상품추천을 위해 추천되는 기법으로는 ① 협업 필터링 기법(Konstan 등, 1997), ② 연관 규칙 탐사기법(Agrawal 등, 1993; Han & Fu, 1999), ③ 제약조건만족(CSP) 기법(Kim, S.H. 등, 2000), ④ 사례기반추론기법(Rahmer & Voss, 1996) 등을 들 수 있다.

현재, 개발된 대표적인 상품추천 시스템으로는 PersonaLogic(www.personalogic.com), Firefly(www.firefly.com), Techno Salesman(<http://www.technoagent.co.kr>) 등을 들 수 있다. PersonaLogic은 소비자의 구매 대상이 될 수 있는 다양한 제품군 중에서 소비자가 명시한 특정 조건(제품, 특성)을 충족시키는 제품만을 우선순위와 함께 제공해주며, 유저의 프로필에 의해 상품을 좁힐 수 있도록 하는 시스템이다. Firefly는 음반이나 서적과 같은 상품을 비교하여 소비자에게 추천해 주는 시스템으로 합작 필터링(collaborative filtering)기술을 이용하고 있다. Techno Salesman은 필터링(filtering)과 스코어링(scoring)을 통해 적절한 수의 상품을 추천하여 고객의 요구에 상응하는 제품이 없을 경우에는 이완(relaxation)을 통하여 고객의 요구에 가장 적절한 제품을 추천해 주는 기능을 제공하고 있다.

본 연구에서는 상품추천에 있어 사례기반추론 기법을 적용하고 있다. 사례기반추론은 일상 생활에서 인간이 생각하는 것처럼 과거의 경험적 사실들이 저장되어 있는 사례 베이스의 유사 사례(analogical case)를 기반으로 새로운 문제를 해결하는 지식 처리 방법이다. 즉, 사례 베이스로부터 해결할 문제에 적용할 최적의 유사 사례를 검색하는 일은 근사해의 산출과 밀접한 관련이 있다. 사례기반추론을 중심으로 한 시스템은 이

미 여러 응용 분야에서 성공적으로 운영되고 있으며, 기업경영의 여러 분야에서 다양하게 개발되고 있다(Watson 1997, Lenz 등 1998). 특히, 사례기반추론시스템은 지속적으로 발생하는 과거 경험들을 이용하여 유사한 문제를 해결하려는 분야에 적합하다. 사례기반추론기법을 이용한 상품추천시스템에 관한 연구로는 Vollrath 등(1998), Wilke 등(1998), Gardingen & Watson(1999), Schmitt & Bergmann(1999), Schmitt 등(2000), Cunningham 등(2001), 성백균 등(2000)을 들 수 있다.

본 연구와 관련하여 특히 관심을 끄는 것은 상품구성을 추천하는 지능형 에이전트에 관한 연구이다. 일반적으로 상품추천문제는 추천된 상품이 고정적인지 또는 변동가능한지에 따라 구분될 수 있다. 여기서 고정적이란 말은 추천된 상품에 대하여 고객이 수정할 수 없음을 말한다. 변동가능 추천상품은 상품 파라미터화(product parameterization)나 상품구성(product configuration)을 통하여 맞춤화 될 수 있다.

Wilke(1999)는 상품추천문제를 고객에 의해 맞춤화(customize) 될 수 있는 정도에 따라 ① 고정된 제품(fixed products), ② 파라미터화가능한 제품(parameterizeable products), ③ 구성가능 제품(configurable products), ④ 혁신적 제품설계(innovative product design), ⑤ 창의적 제품설계(creative products design)로 그 범주를 일련속선 상으로 분류하였다.

본 연구에서는 Wilke(1999)의 분류에 부가하여, 상품의 구성에 있어 제품의 특성과 구조에 의하여 그 구성이 사전에 결정될 수 있는지의 여부에 따라 폐쇄형 상품구성(closed products configuration)과 개방형 상품구성(open products configuration)으로 분류하고자 한다. 폐쇄형 상

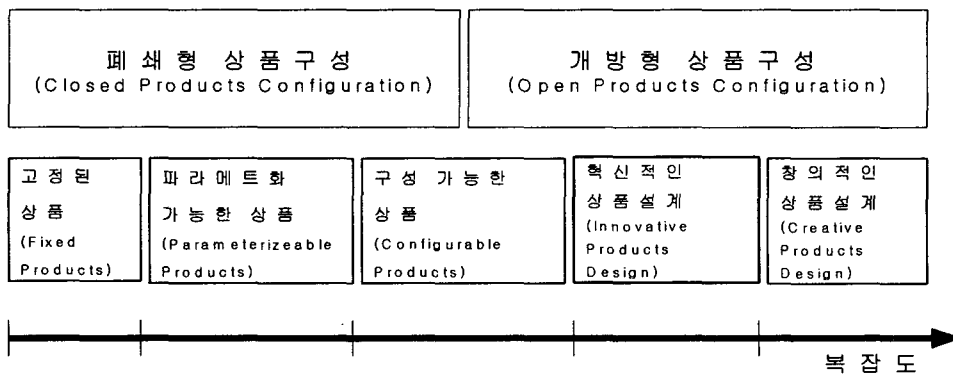
품구성의 예로서 PC의 구성을 들 수 있다. 폐쇄형 상품구성은 PC 조립을 위한 구성품 추천과 같이 그 상품의 구성을 위한 아이템과 그 구성을 위한 조합이 정형화된 경우의 문제이다. 반면 개방형 상품구성의 예로는 본 연구에서 제시하는 수산물 변들구성(상차림) 문제를 들 수 있다. 특정한 집안이벤트시 사용되는 수산물의 상품구성(예, 재사용 어물)은 정형성이 낮고 환경적 요소(구매예산, 참여인원의 수, 어물의 시세와 생산시기 등)에 의해서 그 구성이 가변적일 수 있다. 이러한 문제는 Wilke(1999)의 구성가능 상품의 일부와 혁신적 상품설계 및 창의적 상품설계의 문제에 해당한다. 이상의 상품구성문제를 그림으로 도식화하면 <그림 1>과 같다.

전자상거래에서 상품구성을 지원하는 시스템에 관한 연구로는 Rahmer & Voss(1996), Kim S. H. 등.(2000), Kim, Steven H. 등.(2000), Stahl 등(2000a, b), Cunningham 등(2001)의 연구이다. 이들의 연구를 요약하면 <표 1>과 같다.

3. 객체지향 사례표현과 추론

3.1 사례기반추론의 정의

사례기반 추론은 현재 해결해야 할 새로운 문제와 과거에 해결했던 문제(사례)를 비교하여 현재의 문제를 해결하는 방법이다. 사례기반 추론

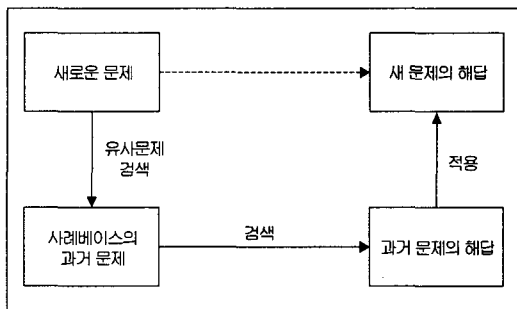


<그림 1> 상품추천문제의 유형

<표 1> 상품구성을 지원하는 에이전트에 대한 연구

연구자	연구내용
Rahmer & Voss(1996)	사례기반추론기법을 이용한 Telecooperation Systems의 구성
Kim S. H. 등(2000)	제약만족(CSP)기법을 이용한 컴퓨터 구성품의 추천
Kim, Steven H. 등(2000)	Multistrategy filtering 접근법을 이용한 개인화된 상품추천
Stahl 등(2000a, b),	순환적 사례기반 추론(Recursive CBR)기법을 통한 사례검색과 상품 맞춤화
Cunningham 등(2001)	사례기반 상품검색, 프로파일링과 협업적 추천, Operator에 기반한 상품맞춤화, 점진적 구성요소 대체

의 기본적인 가정은 사례와 해결해야 할 문제가 유사하며, 해결방법 또한 유사하다는 데 있다. CBR의 일반적인 과정을 도식화하면 <그림 2>와 같다.



<그림 2> 사례기반추론의 기본 모형

CBR은 기억장치에서 현재의 문제와 유사한 이미 해결된 문제를 찾고, 과거의 문제와 현재의 문제간의 차이를 분석하여 이전의 해결책을 현재의 문제에 맞게 조정하는 과정을 거친다. CBR의 학문적 배경은 크게 심리적인 것과 컴퓨터 공학적인 것으로 대별된다.

심리학적 모델로서의 CBR은 Schank(1982)에 의해 개발된 동적 메모리 이론(Dynamic Memory Theory)과 문제해결 및 학습에서의 회상(reminding)의 역할에 관한 연구를 기점으로 본격적으로 시작되었다. 동적 메모리 이론은 개별적인 경험이 어떻게 기억 속에 저장될 수 있는지, 그들이 어떻게 결합되고 추상화 될 수 있는지, 그리고 그들이 어떻게 검색되고 사용될 수 있는지를 기술하기 위한 골격을 제공하였다. 이러한 CBR의 기본원리는 사람이 복잡한 문제를 해결하기 위하여 과거에 문제를 해결했던 경험을 되살려서 비슷한 사례를 추출하여 주어진 문제를 풀고자 하는 유사추론, 혹은 경험적 추론을 하는

것에 기인한다. 여기에서 CBR과 유사추론은 때때로 동의어로 사용되어지기도 하지만 엄밀히 말해서 두 방법은 서로 다른 것이다. CBR은 한 적용분야(intra-domain)안에서 사용되는 방법이고, 유사추론은 서로 다른 적용분야(across-domain)에서 사용되는 방법이다(Aamodt & Plaza, 1994).

한편 컴퓨터 모델로서의 CBR은 규칙기반 시스템의 한계를 극복하기 위한 방안으로 연구되었다. 1970년대부터 개발되기 시작한 대부분의 전문가 시스템에서는 규칙을 지식의 기본 단위로 표현하고, 규칙을 통해 문제해결을 위한 추론을 수행하였다. 그러나 규칙기반시스템의 개발에 대한 경험이 증가할수록 지식 획득의 어려움, 과거 결과에 대한 기억 불가능, 불확실한 자료의 표현과 처리가 불가능, 지식의 수정 및 추가의 어려움 등 규칙기반 시스템의 단점을 발견하게 되었다. CBR은 인간전문가의 지식 및 추론과정에 대한 인지과학적 고찰을 토대로 제안되어 규칙기반 시스템의 문제를 해결하기에 적절하기 때문에 빠르게 발전하였다.

3.2 사례의 구조와 표현 형식

일반적으로 사례기반 시스템의 구축에서 가장 중요한 문제는 사례베이스의 구축을 위해 사례에 관한 정보를 표현하는 방법이며, 사례의 표현에서 고려해야 할 사항은 사례의 구조와 표현 형식이다. 즉, 사례를 어떻게 표현하여 좋은 사례베이스를 구축할 수 있는가 하는 것이다.

CBR에서 사례란 어떤 영역의 문제에서 특정한 상황을 표현한 것으로 신경회로망에서의 패턴에 해당되는 것이다. 그리고 사례 베이스는 이러한 사례들을 모아놓은 것으로 규칙기반 전문가 시스템에서 지식베이스에 해당하는 것을 말한다.

사례기반추론을 하기 위하여서는 먼저 사례에 대한 색인방법을 선택하여 사례기반추론의 구조를 결정해야 한다.

CBR은 일차적으로 저장된 사례에 기초해서 추론하므로 사례를 어떤 구조를 이용해서 저장하는가는 시스템의 전체적인 성능을 결정하는 데에 큰 영향을 끼친다. 또한 사례 베이스를 확장해나가는 과정에서 새로운 사례를 기존의 사례 베이스에 통합하는 것보다 사례의 저장 구조는 밀접한 관련이 있다. 사례의 저장 구조와 일반적인 문제 영역에서의 지식 모델의 통합 문제도 사례의 표현과 관련된 또 하나의 문제이다. <표 2>는 사례의 구조에 대한 내용이다.

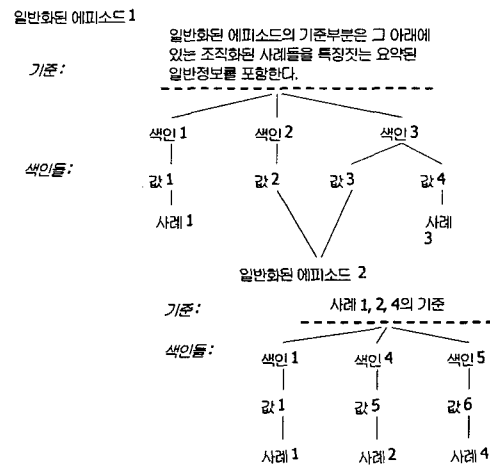
<표 2> 사례의 구조

구성 항목	내용
문제/상황 (Problem/Situation)	문제와 그 문제가 해결될 당시의 상황 묘사
해 (Solution)	문제가 발생했을 당시의 전문가의 반응 및 추론된 해결방법을 서술

사례의 표현 모델로서 많은 영향을 끼친 두 가지 연구로는 Schank(1982)의 동적 메모리 모델과 Kolodner(1988)의 범주-표본 모델을 들 수 있다.

동적 메모리 모델에서는 사례가 E-MOP (Episodic Memory Organization Packets)라는 요소들의 계층 구조의 형태로 저장된다. 이 모델의 기본 아이디어는 유사한 성질을 가지고 있는 구체적인 사례들을 일반화된 에피소드 (Generalized Episode: GE)라고 불리는 보다 일반적인 구조하에 정리한다는 것이다. 이 모델에서 사례의 전체적인 저장 구조는 식별 네트워크 (discrimination network)의 형태를 띠며, 새로운

사례가 저장될 때 이 사례의 특성 값이 기존의 사례의 특성 값과 일치하면 이들을 묶어 하나의 일반화된 에피소드를 형성하는 식으로 메모리 구조가 동적으로 변하게 된다. <그림 3>은 동적 메모리 모델에서 일반화된 에피소드의 한 예를 보여준다.

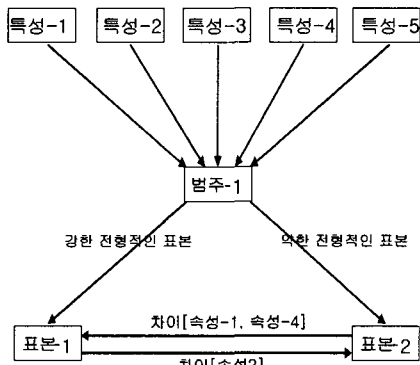


<그림 3> 동적 메모리 모델에서 사례와 일반화된 에피소드의 구조

범주-표본 모델에서는 사례를 표본(exemplar)이라고 부르는데 이 모델의 근거에 깔려있는 생각은 실세계의 개념들은 외연적으로 정의되어야 하고 일반화는 매우 조심스럽게 시도해야 하며 어떤 사례를 어떤 범주에 포함시킬 때 특성의 종류에 따라 중요도가 달라져야 한다는 것이다.

이 모델에서 사례 저장 구조는 범주, 사례와 색인 포인터의 네트워크 형태를 띤다. 색인 포인터는 문제의 특성 구분에서 범주나 사례로 연결되는 특성 고리(feature link), 범주에서 관련된 사례로 연결되는 표본 고리(exemplar link), 한 사례에서 약간의 특성 값만 차이를 보이는 이웃

한 사례를 연결하는 차이 고리(difference link)의 세 종류로 분류된다. <그림 4>는 범주-표본모델의 한 예이다.



<그림 4> 범주-표본 모델의 구조

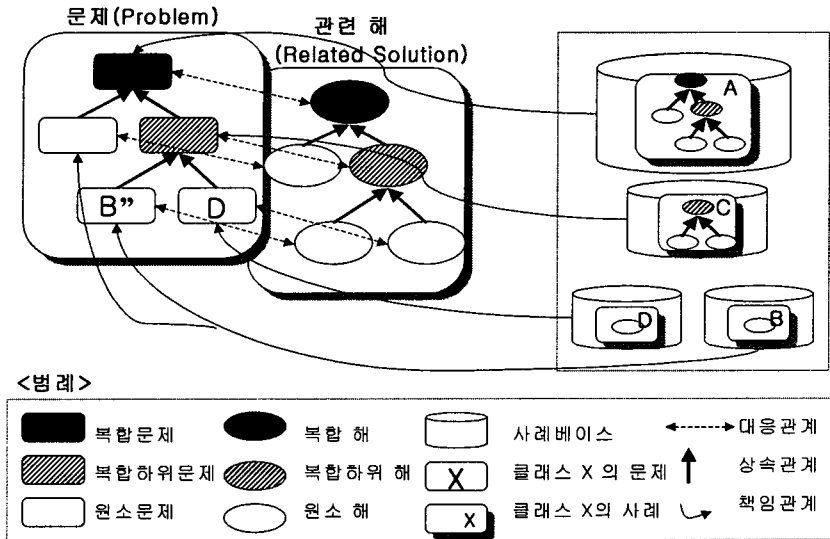
사례는 대개 특정한 결과를 이끌어내는 주요한 특성들(features)의 리스트(list)로서 표현된다. 사례를 표현하는 방법에는 ① 사례의 특성을 정의하는 속성의 단순한 집합으로 사례를 표현하는 특성표현(feature representation) 방법, ② 속성의 값을 포함하는 슬롯(slot)의 집합으로 구성된 프레임(frame)이 하나의 노드(node)가 되어 계층적 구조(hierarchical structure)로 사례를 표현하는 프레임 방법, 그리고 ③ 프레임의 확장된 개념으로 사례를 표현하는데 있어서 규칙(rule), 프레임, 추상화(abstraction), 상속(inheritance) 등과 같은 개념을 혼용 적용하여 사례를 표현하는 방법 등이 있다.

특히 사례표현에 있어 객체지향기법의 적용은 상이한 사례구조를 갖는 복잡한 문제영역에 적합하다(Bergmann & Stahl 1998). 객체지향 사례표현에서 사례는 객체들의 집합체로 표현되어지며, 각각의 객체들은 속성-값 쌍의 집합에 의해서 기

술되어진다. 객체의 구조는 각각의 속성에 대한 타입(가능한 값 또는 하위객체)을 지닌 속성(또는 슬롯)들의 집합으로 정의되는 객체클래스에 의해 기술되어진다. 객체클래스는 클래스 계층구조로 배열되어지고, 각 하위 클래스는 상위클래스로부터 속성과 오퍼레이션을 상속받는다. 이를 일반화(is-a)관계라 한다. 또한 객체클래스는 집합구성(part-of)관계를 가질 수 있다. 구성관계를 표현하는데 있어 관계속성(relational attribute)을 사용한다. 관계속성을 사용하여 복잡한 사례구조를 표현할 수 있다.

<그림 5>는 고도로 구조화된 복잡한 문제영역을 계층적 구조로 표현하고, 여러 개의 사례베이스가 어떻게 대응될 수 있는가는 보여주고 있다. <그림 5>에서 복합 문제(complex problem)는 다시 그 복합 하위문제(complex subproblem)로 나누어 질 수 있으며, 계층구조의 최하위(leaf node)에는 단위원소 문제(atomic problem)로 구성된다. 이러한 문제구조가 주어졌을 때, 해결해야 할 문제(CBR에서는 질의라고 함)와 대응되는 해(solution)는 구성관계계층(part-of hierarchies)으로 표현이 가능하다. <그림 5>의 계층적 구조 관계에서 최상층 노드(root node)는 복합문제 또는 해를 각각 나타낸다. 내부 노드(inner nodes)는 복합 하위문제 또는 복합 하위해(sub-solution)를 나타낸다. 이것은 다시 보다 더 작은 복합 하위문제 또는 해로 분해될 수 있다. 최하위 노드는 단위원소 문제와 관련 단위원소 해를 나타낸다.

따라서 주어진 복합 문제를 CBR을 통해서 해결하기 위해서는 각각에 대응되는 복합문제 또는 그 하위문제와 단위원소문제를 담은 사례베이스로부터 대응되는 사례를 검색(retrieval)하고, 이를 수정(adaptation)한다.



<그림 5> 객체지향 사례표현

3.3 사례의 유사도 측정과 맞춤화

3.3.1 사례 검색 방법과 유사도 측정

사례기반추론시스템은 사례베이스를 효율적으로 활용하여 관련 있는 사례들을 검색할 수 있는 기능들을 제공해야 한다. 따라서 규모가 큰 사례베이스로부터 정확한 사례가 검색될 수 있는 적절한 색인과 검색 절차가 필요하다. 숙련된 사용자들은 데이터베이스로부터 데이터를 검색할 때처럼 사례베이스로부터 특정의 사례를 검색하기 위하여 표준 질의어를 사용할 수 있다.

효율적인 사례의 검색을 위하여 사용되는 대표적인 색인 방법으로 다음과 같이 3가지를 들 수 있다.

- ① 가중합계 방법 : 이 방법은 입력 사례의 특성에 대한 가중합계에 기초하여 사례를 측정하는 방법이다. 주어진 유사도 측정방법에 따라 가장 근접한 사례를 매칭하게 된다. 이 방법

은 부동산 관련 데이터베이스로부터 유사한 문제들의 목록을 추출하는 것과 같이 사례의 추출요구가 특정 문제의 해결에 엄격하게 맞춰져 있지 않는 경우에 적합하다.

- ② 유도적 추출방법 : 이 방법은 채권 등급 평가 문제와 같이 추출목적이 명확하게 정의된 경우 가장 적합하다. 이 방법에서는 사례들이 데이터 자체로부터 파생되는 결과에 영향을 미치는 가장 중요한 특성에 기초하여 색인이 부여된다. 이 결과 생성되는 의사결정트리는 사례의 추출 시간을 상당히 단축시킨다.
- ③ 지식기반 추출방법 : 이 방법은 문제 해결에 관련성이 있는 사례의 추출을 위하여 이미 존재하는 영역지식을 적용한다. 이 방법은 규칙기반 전문가 시스템에서처럼 전문가가 사례를 분류하는데 사용되어지는 특성들을 결정하게 된다. 이 경우 시스템은 영역지식이 완전하지 않기 때문에 정확한 해결 방법을 찾기

위하여 다른 색인 방법을 활용하여 그 영역의 부분적인 모델을 자주 결합시키는 작업이 필요하다.

전통적 방식의 질의 처리는 명시된 조건과 정확히 일치하는 데이터만을 제공해 주는 엄밀 질의응답(exact query answering)기능을 따른 방식이라 할 수 있다. 따라서 엄밀 질의응답기능은 질의조건 q 를 완전히 충족시켜 주는 객체의 경우 1, 그렇지 않는 객체에 대해서는 0의 값을 발생시켜 1의 값을 갖는 객체들에 대한 관련 정보만을 제공해주는 특성 함수(characteristic function)로 볼 수 있다. 그러므로 전통적 방식인 엄밀 질의응답기능에 따라 시스템을 이용하려면 사용자는 문제영역과 관련 DB 스키마를 충분히 이해하고 있어야만 한다. 이는 검색 시스템의 이용률을 저하시키는 주요 원인 중 하나가 된다. 따라서 사용자의 부담을 줄여 주고 검색율을 높이기 위해서는 질의조건을 완화시켜 부분적으로 질의를 만족시키는 객체에 대한 정보도 제공하여야 한다. 이러한 질의응답방식을 근사 질의응답(approximate query answering)이라고 한다.

사례 베이스의 사례와 고객이 질의한 문제의 유사도를 평가하기 위해서 먼저 필요한 작업으로 유사도를 측정할 알고리즘을 작성하는 것이 필요하다. 유사도는 사례와 해결해야 할 질의의 비슷한 정도를 나타내는 것으로, 전체 유사도를 추론해 내기 위해 각 속성별 유사도 측정을 하여야 한다. 유사도를 표현하는 방식으로 0과 1사이의 숫자로 표현한다. 값 0이 뜻하는 것은 질의를 전혀 만족시키지 못함을 말하며, 1은 질의한 것에 일치한다는 의미이다(Vollrath 등, 1998). 그리고 0과 1사이의 값은 부분적으로 질의를 충족한다는 의미이다. 예를 들어, 질의의 내용 중 수산물을 사용하려는 이벤트가 제사일 경우 사례의 속성

중 이벤트가 '제사'인 사례의 유사도는 1이 된다. 하지만 실제 모든 속성 값을 숫자로 표현하는 데는 한계가 있다. 이런 경우에는 모든 가능한 속성값의 쌍에 대한 유사도를 정의한 유사도 테이블을 작성하여 사용할 수가 있다.

질의한 문제와 사례와의 유사도를 측정하는데 있어 크게 다음과 같은 두 가지 형태의 함수가 사용될 수 있다.

① 점 비교 함수(point matching function) : 주어진 두 값, 여기서는 해결해야 할 문제의 속성값과 과거 사례의 속성 값을 말하며, 이 두 값이 얼마나 유사한가를 차이 값의 범위에 따라 0과 1사이의 값으로 평가하는 함수이다.

② 구간 비교 함수(interval matching function) : 속성의 값이 일정한 구간을 가지는 값일 때 이용하는 것으로, 문제의 속성값과 사례의 속성값이 겹치는 부분(r)이 과거 사례의 속성값의 범위(R) 중 유사한 정도는 r/R 이 된다.

이와 같이 개별 속성에 대한 유사도 측정이 끝나게 되면 사례와 질의 전체에 대한 유사도를 추론할 수 있다.

전체 유사도를 추론하는 일반적인 방법으로 다음과 같은 예를 보자. 속성 q_1, q_2, \dots, q_n 으로 표현되는 질의와 c_1, c_2, \dots, c_n 으로 표현되는 사례 c 가 있다고 고려해 볼 때, q 와 c 사이의 유사도 σ 는 각 개별 속성의 유사도인 σ_i 로부터 계산된다. 다음 식은 각 항목간의 유사한 정도를 나타내는 유사성 매칭 함수이며 해결해야 할 문제의 속성과 과거 사례의 속성이 가지는 값에 따라 다른 두 가지 형태의 함수를 적용한다.

$$\sigma(q, c) = \sum_{i=1}^n w_i \sigma_i(q_i, c_i) \quad (\text{식 1})$$

여기에서 $\sum_{i=1}^n w_i = 1$, 모든 i 에 대해 $w_i \geq 0$

이 식은 사례 검색을 위해 최근이웃 추출방법을 사용한 것으로써 최근이웃 추출방법은 가장 가까운 거리에 위치하는 이전 사례를 발견하는 알고리즘으로 이미 여러 인공지능 분야에서 그 유용성이 검증되었으며, 기계 학습 분야에도 널리 이용되고 있다. 위의 식에서 w_i 는 가중치로써, 일반적으로 전문가에 의해 그 값이 주어지지만 개인적 선호도에 의해 조정될 수도 있다.

3.3.2 객체클래스 유사도 측정방법

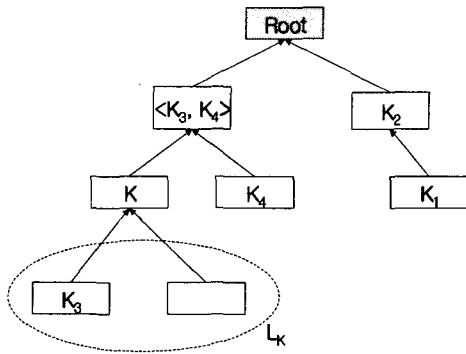
현실문제에 있어 사례의 구조가 단순히 평면적이지 않고 그 구조가 매우 복잡할 수 있다. 이 경우 속성들 간의 단순한 평면적인 가중합계만으로는 정확한 해를 찾지 못할 수 있다. 즉, 사례구조가 복합문제와 그 복합 하위문제로 구성될 경우 유사도 측정을 위한 체계적인 방법이 요구된다. 이 경우 유사도 측정은 문제영역의 복잡도에 따라 다양하게 나타날 수 있다. 이와 같이 복잡한 문제영역에서 가장 유망한 사례표현법으로 객체지향 사례표현법이 제시되고 있음을 앞에서 지적하였다.

객체지향 사례표현법을 이용하여 복잡한 복합문제가 계층적 구성관계(part-of 또는 is-a)로 표현될 경우 유사도 측정을 위한 일반적인 방법이 Bergmann & Stahl(1998)에 의해 제시되고 있다. 객체지향 사례표현에서 유사도 측정의 최종 목적은 두 객체간의 유사도를 결정하는 것이다. 두 객체 중 한 객체는 사례(또는 그 일부분)이고 또 다른 한 객체는 질의(또는 그 일부분)이다. 유사도는 객체 전체에 대한 전역적 유사도(global similarity)와 각 속성 또는 하위객체에 대한 지역적 유사도(local similarity)로 나눌 수 있다. 객체 유사도는 주로 상향식(bottom-up) 방식으로

순환적으로 결정된다. 즉, 각각의 단순 속성에 대하여 지역적 유사도 척도(local similarity measure)는 두 속성 값 간의 유사도를 결정하고, 각 관계 슬롯에 대하여 객체 유사도 척도(object similarity measure)는 관련된 두 하위객체를 순환적으로 비교한다. 그리고 지역적 유사도 척도와 객체 유사도 척도로부터 각각 계산된 유사도 값을 비교하고자 하는 객체간의 객체 유사도에 가중 합계되어 진다.

그런데 이러한 객체 유사도 척도는 다소 자의적일 수 있다. 그렇지만 클래스 계층구조로부터 유사도 측정을 위한 일반적인 지식을 찾을 수 있다. 즉, 클래스 계층구조는 일반적으로 분류법에 따르므로 객체의 유사도에 관한 지식을 내포하고 있다. Bergmann(1998)은 사례특성표현과 지역적 유사도 척도를 위한 분류법(taxonomies)의 사용에 관한 이론을 제시하고 있다. 이 이론을 확장한다면 계층적 구성관계를 갖는 사례표현의 유사도 측정이 가능하다(Bergmann & Stahl, 1998).

객체클래스가 계층적으로 조직화될 경우 클래스 유사도를 계산하는 방법을 간단히 예시하면 다음과 같다. <그림 6>에서 각 노드는 클래스를 나타내며, 화살표 마디는 클래스들 간의 구성관계, 즉 "is-a"관계 또는 "part-of" 관계를 나타낸다. 또한 K노드는 클래스 계층에서 내부노드(inner node)를 나타내며, L_K 는 K노드 아래로 구성되는 모든 최하위노드(leaf node)의 집합을 나타낸다. 그리고 $K_1 < K_2$ 는 K_1 이 K_2 의 계승노드(하위클래스)임을 의미한다. $\langle K_3, K_4 \rangle$ 는 K_3 와 K_4 의 가장 가까운 부모 노드(가장 가까운 공통 객체클래스)를 나타낸다. 즉, $\langle K_3, K_4 \rangle \geq K_3$ 이고, $\langle K_3, K_4 \rangle \geq K_4$ 이다.



<그림 6> 계층적 객체클래스 다이어그램

일반적으로 두 객체간의 유사도 계산은 클래스 내 유사도(intra-class similarity: SIM_{intra})와 클래스 간 유사도(inter-class similarity: SIM_{inter})로 나눌 수 있다.

(1) 클래스 내 유사도

클래스 내 유사도 측정을 위해서는 두 객체의 공통된 특성을 사용한다. 즉, 두 객체의 가장 공통된 속성을 중심으로 유사도를 계산한다. 공통성을 갖는 모든 속성에 대하여 지역적 유사도와 객체 유사도를 계산하여 총합(가중합계)한다.

$$SIM_{intra}(q,c) = \Phi(sim_{A_1}(q.A_1, c.A_1), \dots, sim_{A_n}(q.A_n, c.A_n)) \quad (식 2)$$

여기서, Φ 는 총합함수, $q.A_i$ 와 $c.A_i$ 는 각각 질의와 사례객체의 속성 A_i 의 값을 나타낸다. 그리고 sim_{A_i} 는 속성 A_i 의 지역 또는 객체 유사도이다.

(2) 클래스 간 유사도

두 객체간의 차이는 공유된 속성보다는 클래스 계층의 구조에 의해서 좌우된다. 즉, 클래스 간 유사도는 두 객체의 공통된 속성 값과는 독립

적이며, 계층구조 내에서 클래스의 위치에 의존한다. 따라서 클래스 간 유사도 계산을 위해서는 이 구조를 이용할 필요가 있다. 보통 클래스간 유사도 $SIM_{inter}(Q,C)$ 는 여러 객체 클래스에 걸쳐서 정의되어진다. 따라서 질의객체 q 와 사례객체 c 간의 최종 객체유사도 $sim(q,c)$ 는 클래스 간 객체 유사도와 클래스 내 객체 유사도의 곱으로 계산되어진다.

$$sim(q,c) = SIM_{intra}(q,c) \cdot SIM_{inter}(class(q),class(c)) \quad (식 3)$$

여기서 $class(q)$, $class(c)$ 는 각각 객체 q 와 c 의 객체클래스이다.

클래스 간 유사도는 클래스 계층구조 상의 위치와 질의된 값의 형태(구체적인 값, 아무 값, 알 수 없음)에 따라 달라진다. 클래스 계층구조 상에서 최하위 노드는 구체적인 인스턴스를 갖는 구체 객체(concrete objects)이다. 그러나 내부 노드들은 대부분 구체적인 인스턴스가 없는 추상객체(abstract objects)이다. 질의객체와 사례객체가 클래스 계층구조에 속하는 위치와 질의 형태에 따라 <표 3>과 같이 크게 7가지 유형의 클래스 간 유사도를 도출해 낼 수 있다.

<표 3> 클래스 계층구조에서 질의객체와 사례객체의 유형에 따른 유사도 분류표

질의 \ 사례	최하위 노드 구체객체	아무 값 추상객체	값 불확실 추상객체
최하위 노드 구체객체	유형 I	유형 II	유형 V
아무 값 추상객체	유형 III	유형 IV	
값 불확실 추상객체	유형 VI		유형 VII

[유형 I]의 유사도

<그림 6>에서 최하위 노드(구체 객체)들의 클래스 간 유사도에는 다음과 같은 일반적인 제약식이 존재한다.

IF <K, K₁> > <K, K₂> THEN

$$SIM_{inter}(K, K_1) \leq SIM_{inter}(K, K_2)$$

그리고 내부 노드 K_i의 유사도 값을 S_i ∈ [0..1]로 나타낸다면,

IF K₁ > K₂ THEN S₁ ≤ S₂. 가 성립한다.

S_i는 K_i를 부모로 하는 모든 최하위 노드 집합 L_{K_i} 내의 임의의 클래스 인스턴스의 클래스 간 유사도의 하한이 된다. 즉,

$$\forall X, Y \in L_{K_i} \quad SIM_{inter}(X, Y) \geq S_i \quad (\text{식 4})$$

따라서 이상의 의미론으로부터 [유형 I]의 클래스 간 유사도는 다음과 같이 계산되어진다.

$$SIM_{inter}(K_1, K_2) = \begin{cases} 1 & \text{if } K_1=K_2 \\ S_{<K_1, K_2>} & \text{otherwise} \end{cases} \quad (\text{식 5})$$

[유형 II]의 유사도

질의는 구체적인 객체를 포함하고 있으며, 사례는 추상객체(내부노드)를 포함하므로 속성에서 추상객체의 사용은 축약을 나타낸다. 질의의 구체적인 객체는 추상객체내의 하나의 객체인스턴스이다. 사례베이스 내에서 가장 유사한 사례를 찾는 것이 목적이므로, 질의와 추상객체를 포함하는 사례간의 클래스간 유사도는 질의와 구체객체 중의 하나 간에 가장 유사도와 높은 것과 같다.

$$SIM_{inter}(Q, C) = \max_{\{ C' \in L_C \}} (SIM_{inter}(Q, C')) = \begin{cases} 1 & \text{if } Q < C \\ S_{<Q, C>} & \text{otherwise} \end{cases} \quad (\text{식 6})$$

[유형 III]의 유사도

여기서 추상객체의 세부명세는 여러 질의에 대응되는 축약으로 볼 수 있다. 따라서 그 집합으로부터 가장 유사한 구체객체를 선택하면 된다.

$$SIM_{inter}(Q, C) = \max_{\{ Q' \in L_Q \}} (SIM_{inter}(Q', C)) = \begin{cases} 1 & \text{if } C < Q \\ S_{<Q, C>} & \text{otherwise} \end{cases} \quad (\text{식 7})$$

[유형 IV]의 유사도

이것은 [유형 II]와 [유형 III]의 결합형태이다. 따라서 추상객체에 의해 제시된 두 집합으로부터 두 구체객체간에 가장 높은 유사도를 갖는 것을 찾으면 된다.

$$SIM_{inter}(Q, C) = \max_{\{ Q' \in L_Q, C' \in L_C \}} (SIM_{inter}(Q', C')) = \begin{cases} 1 & \text{if } C < Q \text{ or } Q < C \\ S_{<Q, C>} & \text{otherwise} \end{cases} \quad (\text{식 8})$$

[유형 V]의 유사도

질의객체의 값은 구체적이거나 사례객체의 값이 불확실하므로 정보가 절대적으로 부족한 경우이다. 이 경우 가장 비관적인 접근법을 택한다면 유사도는 다음과 같다.

$$SIM_{inter}(Q, C) = \min(SIM_{inter}(Q, C') \mid C' \in Lc) = S_{\langle Q, C \rangle} \quad (\text{식 9})$$

[유형 VI]의 유사도

질의에 정보가 불확실하며, 사례에는 확실하므로 [유형 V]와 정반대의 상황이다. 이 경우 가장 비관적인 접근법을 택할 경우 유사도는 다음과 같다.

$$SIM_{inter}(Q, C) = \min(SIM_{inter}(Q', C) \mid Q' \in Lq) = S_{\langle Q, C \rangle} \quad (\text{식 10})$$

[유형 VI]의 유사도

질의와 사례 모두 불확실한 정보가 내재되어 있다. 이 경우 가장 낙관적인 접근법을 택한다면 유사도는 다음과 같다.

$$SIM_{inter}(Q, C) = \min(SIM_{inter}(Q', C') \mid C' \in Lc, C' \in Lc) = S_{\langle Q, C \rangle} \quad (\text{식 11})$$

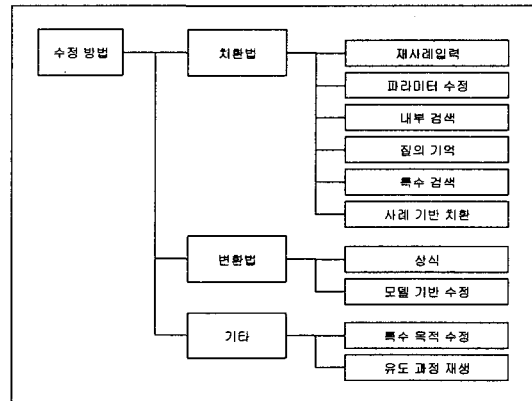
위 모든 경우에 있어 질의와 사례객체의 클래스 계층구조상의 위치를 알면, 두 객체의 가장 공통된 상위클래스의 유사도를 찾아 쉽게 클래스간 유사도를 계산할 수 있다.

3.3.3 사례의 수정(Adaptation)과 상품 맞춤화

유사도를 바탕으로 고객의 질의와 유사한 사례를 제시하였을 때, 고객이 이를 받아 들여 그대로 상품을 구입하는 경우는 드물 것이다. 여러 사례가 존재한다고는 하나 질의를 한 고객과 완전히 일치하는 사례는 없기 때문이다. 따라서 정도의 차이는 있지만 차이를 보정해 주기 위한 일

련의 과정을 거쳐야만 비로소 사례가 사용 가능한 경우도 있다.

사례기반 추론기법의 수정 방법은 <그림 7>과 같이 치환(substitution) 방법, 변환(transformation) 방법, 그리고 기타 특수한 방법으로 나눌 수 있다(Kolodner 1993).



<그림 7> 수정방법의 분류

- ① 치환방법 : 이 방법은 추출된 사례의 해의 일부분을 주어진 문제 상황에서 요구되는 새로운 값으로 바꾸는 방법이다.
- ② 변환방법 : 치환 방법은 추출된 사례의 해 중 부적절한 값에 치환될 항목이나 개념이 이미 존재하는 경우에 적당하다. 그러나 요구되어지는 항목이 존재하지 않는다면 사용될 수 없다. 이 경우 부적절한 해의 일부분을 현재의 문제 상황에 맞게 변환 시켜나아가야 한다.
- ③ 기타 방법 : 이 방법에 분류되는 수정 방법 중에는 특별한 목적을 위한 수정 방법 및 유도 유추에 의한 방법이 포함된다.

객체지향 사례표현에 의해 표현된 사례를 수정하는 방법으로는 수정규칙을 이용하는 방법(Bergmann 등, 1996), 수정오퍼레이터를 이용하

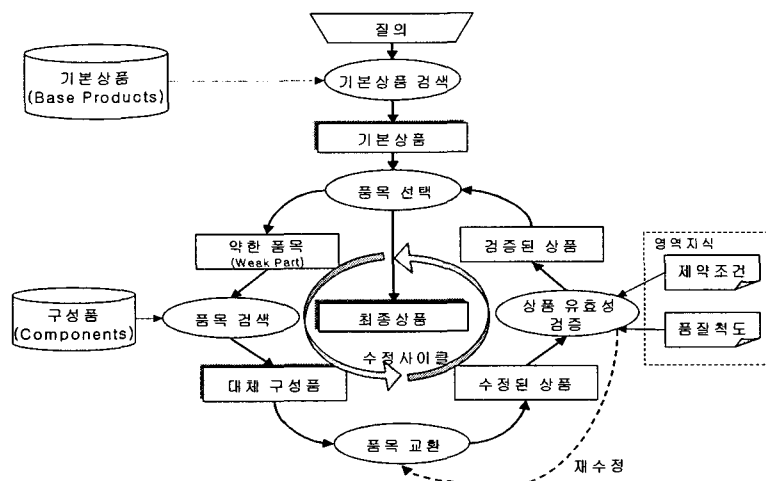
는 방법(Schmitt & Bergmann, 1999), 그리고 순환적 CBR을 적용하는 방법(Stahl & Bergmann, 2000) 등이 제시되고 있다. 이 중에서 수정규칙을 이용하는 방법과 수정오퍼레이터를 이용하는 방법은 변환방법에 해당한다.

변들상품과 같이 그 구성품을 가진 상품의 맞춤화(customization)를 위한 수정에 대한 일반적인 프로세스를 제시하면 <그림 8>과 같다. 변들상품의 수정과정은 크게 두 가지 부분, 즉 기본상품(base product) 검색 단계와 수정 사이클 단계로 나누어진다. 수산물 변들상품추천 시스템의 경우 기본상품은 표준상이 된다. 먼저, 첫 번째 기본상품 검색단계에서는 질의된 내용과 기본상품 사례베이스로부터 유사도에 근거하여 가장 알맞은 기본상품을 검색한다. 두 번째 수정 사이클 단계는 고객의 요구가 만족될 때까지 검색된 기본상품의 수정을 가하는 순환적 과정이다.

기본상품이 고객의 요구를 충족하지 못할 경우, 기본상품의 구성품 중에서 약한 품목(weak part)을 식별해 낸다. 부품질의와 관련된 질의가

있을 경우 유사도가 낮은 부품을 약한 품목이라 부른다. 일반적으로 약한 품목은 두 가지 경우, 즉 검색된 표준상품이 요구되는 부품을 포함하지 않는 경우와 부품이 고객의 기술적 요구에 미달하는 경우로 구분할 수 있다. 품목교환 프로세스는 약한 품목 중에서 교체할 품목(대체품목)을 구성품(components) 사례베이스로부터 검색하여 유사도가 높은 품목으로 교체하는 과정이다.

품목교체가 완료되면 수정된 상품(modified product)이 만들어지고, 이것이 유효한지를 체크하기 위해 상품 유효성검사(product validation) 과정을 거친다. 이 검사과정에서 여러 가지의 영역지식(domain knowledge)이 사용된다. 상품 유효성검사를 위한 대표적인 영역지식으로 상품의 기술적, 환경적 제약조건 대한 지식과 질의 내용과의 일관성 유지에 대한 지식 등이 있다. 제약조건의 예로서 PC의 경우 특정 메인보드에 장착될 수 있는 하드디스크와 같이 부품결합에 대한 제약사항이 있으며, 수산물의 경우 금어기, 계절 등에 대한 제약이 있을 수 있다. 만일 유효성 검



<그림 8> 변들상품의 맞춤화를 위한 수정 사이클

사를 통과하지 못할 경우 부품교환 프로세스로 되돌아가 다시 부품선택과정이 이루어진다. 유효성 검사과정을 통과된 상품은 고객에게 추천되어 지고, 이상의 과정은 고객이 만족할 때까지 반복되어진다.

려는 목적에 따라 다양한 수산물 구성을 사례베이스로부터 검색하여 고객이 가장 원하는 상품 구성을 추천해 주는 방식이다.

4. 번들상품추천 시스템의 사례베이스 설계

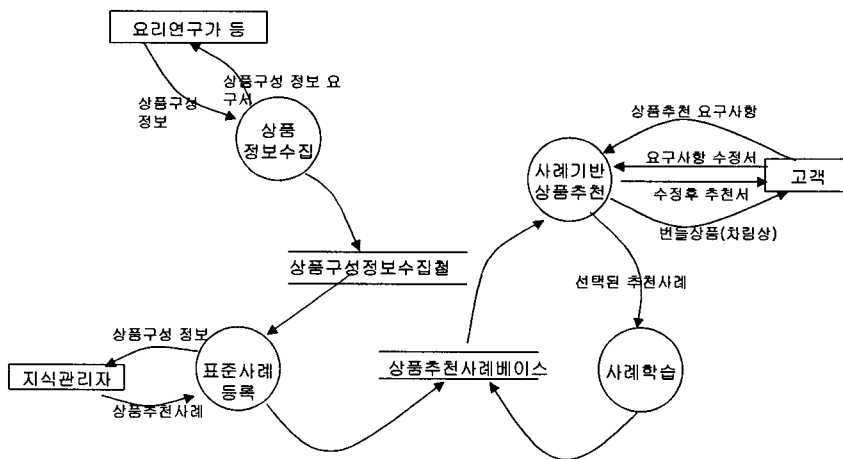
4.1 시스템 정의

본 논문에서 제안하는 사례기반 번들상품추천 시스템은 인터넷 쇼핑몰을 통하여 수산물을 구입하려는 고객에게 그가 원하는 상품을 차림상이라는 패키지형 상품으로 추천하는 시스템이다. 이 시스템의 자료흐름도를 제시하면 <그림 9>와 같다. 차림상은 고객이 수산물을 이용하려는 이벤트의 내용 또는 고객의 취향에 알맞은 단위상품들로 구성된 번들상품을 말한다. 즉, 이 시스템에서는 단일 상품 추천뿐 아니라 수산물을 사용하

본 연구에서 지향하는 바는 수산물 인터넷 쇼핑몰에서 고객이 자신의 각종 집안이벤트(제사, 생일, 돌, 회갑 등)시 구입해야할 수많은 수산물들을 큰 고민 없이 최적의 상품구성을 선택할 수 있도록 지원해주는 시스템을 만드는 것이다. 이를 위해 쇼핑몰 개발자와 운영자는 요리연구가 등으로부터 고객의 상품선택과 관련된 각종 정보들을 수집하여 다양한 상품의 구성을 준비한다. 한편 고객은 자신의 요구사항(이벤트시 참여인원, 가능한 예산액, 지역, 요리종류 등)을 입력하면, 쇼핑몰 시스템은 최적의 상품을 구성하여 고객에게 제시한다.

4.2 객체지향 사례표현

사례기반추론기법에서 문제해결은 현재 해결해야 할 문제와 유사한 문제를 찾아 그 문제의



<그림 9> 수산물 상품추천 시스템 자료흐름도

해결방법으로부터 해결해야 할 문제의 해결방법을 모색하는 것으로 과거 유사한 문제와 그 해결방법을 사례라고 한다. 사례기반 번들상품추천 시스템에서 추론과정은 번들구성 문제와 그 해결방법으로 구성된다. 먼저 번들구성 문제는 특정 고객이 자신의 이벤트에 사용될 수산물 품목들을 하나씩 선택하여 장바구니에 담는 수고를 하지 않고, 자신의 구미에 딱 맞는 장바구니를 번들형태로 구입하려는 상황이다. 그러한 상황에서 그 고객이 구입했던 상품들은 해결방법이 된다. 따라서 사례베이스에는 이러한 상황과 구입한 번들상품의 명세가 들어가게 된다.

사례 베이스에 저장될 정보에는 고객정보(직업, 나이, 지역, 연수입 등)와 이벤트에 대한 정보(이벤트의 종류, 인원수, 예산 등), 그리고 이벤트에 사용될 요리정보(요리종류, 요리재료, 요리방법), 그리고 번들상품에 대한 정보(주용도, 가격, 적정인분, 특징)와 그 구성품에 대한 정보(품명, 수량) 및 품목정보(단가, 산지, 가공방법, 생산방법, 생산계절, 포장단위)가 있다. 이들 정보는 숫자 또는 문자로 표시되며, 각 항목들은 유사도 측정에서 개별 속성으로 다루어진다.

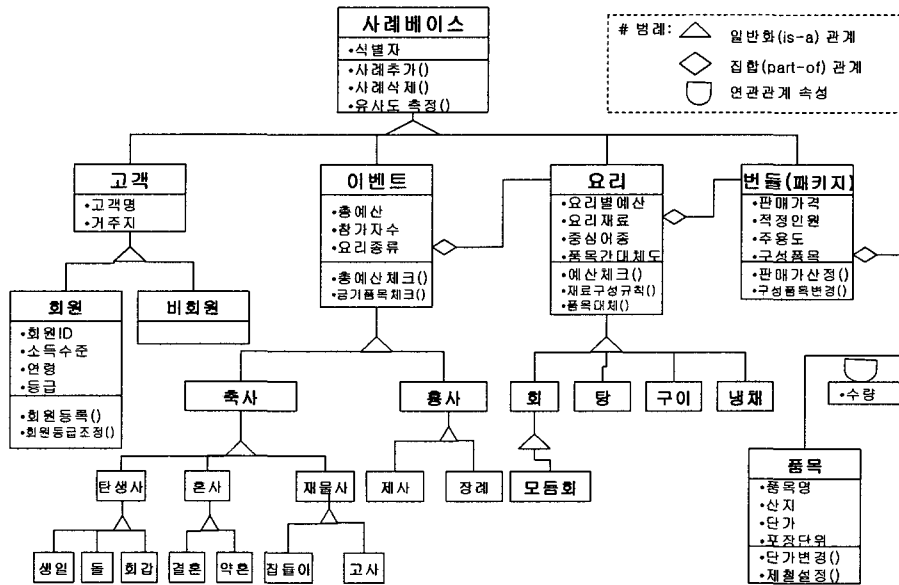
객체지향 사례표현의 관점에서 본다면, 위에서 나열된 각각의 정보들을 하나 또는 그 이상의 객체클래스로 볼 수 있다. 그리고 이들 객체클래스는 "is-a"관계 또는 "part-of"관계로 계층적으로 조직화할 수 있다.

수산물의 번들상품추천 시스템의 사례베이스를 설계하기 위하여 식별되어 질 수 있는 객체클래스와 이들 간의 일반화 및 집합관계를 OMT(Rumbaugh 등, 1991)의 객체모델로 표현하면 <그림 10>과 같다. 객체모델에서 객체클래스는 3단 박스로 표현되어지며, 제 1단에는 객체클래스 이름을 표시하며, 제 2단에는 객체클래스의

속성을 표시하고, 제 3단에는 객체클래스의 행위를 나타내는 오퍼레이션(operation)을 기술한다. 객체모델에서는 <그림 10>에서 제시된 일반화("is-a")관계와 집합("part-of")관계 외에도 객체들간의 관계에서 발생하는 다양한 요소들(multiplicity, ordering, link attribute, qualifier 등)을 표현할 수 있다. 그런데 객체지향 사례베이스 설계에서는 질의객체와 사례객체 간의 유사도 측정이 가장 중요한 요소이므로, 사례객체들간의 계층적 관계표현에만 중점을 둔다. 번들상품추천 시스템의 사례베이스 설계에 있어 가장 중심이 되는 객체클래스로는 <그림 10>에서는 보는 바와 같이 고객클래스, 이벤트클래스, 요리클래스, 번들(패키지)클래스, 품목클래스이다.

고객클래스는 고객의 유형에 따라 회원과 비회원으로 분류될 수 있다. 그리고 고객의 가장 중요한 속성으로는 고객의 거주지이다. 일반적으로 고객의 거주지에 따라 이벤트 시 다른 형태의 수산물을 선호하기 때문이다. 고객에 대한 또 다른 중요한 정보로는 고객의 소득수준이나 고객의 연령대 등으로 이들 속성에 따라 고객의 수산물 선호패턴이 달리 나타날 수 있다.

이벤트클래스는 그 유형에 따라 길·홍사로 나눌 수 있으며, 다시 하위클래스로 세분화할 수 있다. 이러한 분류는 통상의 집안이벤트 분류 체계에 따라 이루어질 수 있다. 사례기반 수산물 상품추천에 있어 이벤트의 분류는 매우 중요하다. 왜냐하면, 집안의 이벤트에 따라 요리의 종류와 사용가능한 수산물의 종류가 상당히 달라지기 때문이다. 즉, 집안이벤트에 따라 사용가능한 요리나 수산물의 종류에 대한 제약이 존재한다. 이러한 제약요소들은 각 이벤트마다 고유한 영역지식(domain knowledge)을 구성하는데 사용될 수 있다. 이벤트클래스 또는 그 하위클래스



<그림 10> 번들상품추천 시스템의 객체지향 사례 표현

의 영역지식에 대한 예로서 이벤트 시 금기해야 하는 금기품목체크를 들 수 있다. 이벤트 클래스의 또 하나 중요한 속성으로 이벤트에 참가할 인원수와 지출예산이다. 일반적으로 참가인원수와 지출예산 간에는 선형관계가 존재한다. 그러나 이벤트의 종류에 따라 그 기울기는 달리 나타날 수 있다.

각 이벤트별로 다양한 요리를 장만할 수 있다. 따라서 이벤트클래스의 요리종류 속성과 요리클래스 간에는 “part-of” 관계로 표현할 수 있다. 요리클래스는 요리의 종류에 따라 회, 탕, 구이, 냉채 등으로 분류될 수 있다. 회의 경우 다시 모듬회, 돛회, 광어회 등으로 분류될 수 있다. 그리고 이들회는 산지, 생산방법, 크기 등에 따라 구체적인 인스턴스(예, 남해자연산 광어회 大)를 가질 수 있다. 일반적으로 하나의 요리를 위해서는 여러 종류의 수산물들이 동시에 들어가며, 그 중심

어종이 있다. 예를 들어 모듬회의 경우 광어, 돛 등이 중심 어종이 되며, 이를 보조하는 여러 가지의 부산물(일명 찌개다시라고 함)이 들어간다. 모듬회에 사용되는 부산물로는 굴, 멍게, 다시마 등이 있다. 또한 중심어종은 요리에 따라 선호품목이 달라지며, 그 우선순위와 품목 간 대체도에 대한 지식을 별도로 만들 수 있다.

각 요리마다 사용될 수산물은 번들(패키지) 형태로 구성된다. 번들은 다시 여러 개의 구성품으로 구성된다. 번들의 판매가격은 각 구성품의 단가에 수량을 곱한 금액을 총합계하여 할인율을 적용한 가격이다. 번들의 중요한 오퍼레이션으로 번들구성품목 변경이 있다. 이것은 CBR을 통해서 제시된 번들이 고객이 원하는 것과 약간 다를 때, 고객맞춤(customization) 규칙, 즉 수정규칙을 담고 있다.

끝으로 고객의 예산과 관련하여 객체모델로부

터 아래와 같은 객체클래스 전역에 걸친 일반적인 지식을 찾을 수 있다.

$$\begin{aligned} \text{이벤트.총예산} &\leq \sum_{i=\text{요리종류}} \text{요리.요리별예산} \\ &\leq \text{번들.판매가격} \\ \text{요리.요리별예산} &= \text{이벤트.총예산} \\ &\quad * \text{요리 } i \text{의 예산비율} \\ \text{번들.판매가격} &= (\sum_{i=\text{번들구성품목}} (\text{품목.단가} \\ &\quad * \text{번들.구성품목.수량})) * \text{할인율} \end{aligned}$$

4.3 객체지향 사례표현과 관계형 데이터베이스의 결합

대부분의 인터넷 쇼핑몰들은 그 거래자료와 참조테이블을 관계형 데이터베이스에 저장하고 관리한다. 쇼핑몰에서 축적된 자료를 공유하기 위해서는 객체지향 사례표현 모델을 관계형 데이터 모델로 전환하여 구현할 필요가 있다. 객체지향의 관점에서 본다면 관계형 테이블을 하나의 객체클래스로 볼 수 있다. 따라서 <그림 10>의 번들상품추천을 위한 객체지향 사례표현을 관계형 데이터베이스에서도 손쉽게 구현할 수 있다.

<그림 10>의 객체모델을 관계형 데이터모델로 전환하는데 있어 아래와 같은 몇 가지 간단한 규칙을 적용할 수 있다.

첫째, 기본적으로 하나의 객체클래스 중 속성 부분이 하나의 관계형 테이블로 전환된다. 그리고 각 객체식별자(object identifier)가 기본으로 대응되는 관계형 테이블의 기본 키(primary key)가 된다. 그러나 관계유형에 따라 여러 개의 객체클래스가 하나의 테이블이 될 수도 있으며, 하나의 객체클래스가 두 개 이상의 테이블이 될 수도 있다.

둘째, “is-a”관계에서 발생하는 상속관계가 있는 경우 단순한 유형의 분류인 경우, 유형분류를

위한 필드와 계층관계를 표현하는 필드를 추가하여 모든 하위클래스를 하나의 관계형 테이블에 담을 수 있다. 계층관계를 표현하는 필드에는 부모 실체의 식별자를 저장한다. 그러나 계층관계에서 클래스간 유사도 값이 다를 경우에는 클래스간 유사도를 담은 참조테이블을 만든다. 이벤트나 요리 클래스의 경우 하위클래스를 묶은 하나의 테이블과 클래스 계층구조에서 각 클래스간 유사도를 담은 참조테이블을 만들 수 있다.

셋째, “is-a”관계에서 하위클래스가 추가적인 속성을 지니고, 서로 소(mutually exclusive)인 경우에는 각 하위클래스 수준에서 각각의 관계형 테이블을 만든다. 고객클래스의 경우 회원과 비회원 테이블을 독립적으로 만든다.

넷째, “part-of”관계에서 관계속성이 존재하지 않는 경우에는 구성품의 객체식별자를 구성품목 객체에 대응되는 관계형 테이블의 외래 키(foreign key)로 설정한다. 예를 들어 요리객체와 번들객체에 있어 요리재료를 구성하는 번들을 관계형 테이블로 표현하기 위해서는 번들테이블에 요리객체의 식별자를 외래 키로 설정한다.

다섯째, “part-of”관계에서 관계속성이 존재하지 하는 경우에는 관계를 표현하는 새로운 테이블을 생성하고, 두 객체클래스의 식별자를 기본 키로하는 복합 키(composite key)를 생성한다. 그리고 관계속성은 새로 생성된 테이블의 속성으로 된다. 예를 들어 번들클래스와 그것을 구성하는 품목클래스 간에는 수량이라는 관계속성이 존재한다. 이 경우에는 번들클래스의 식별자와 품목클래스의 식별자를 결합하여 복합 키를 만들고 수량을 속성으로 설정한다.

이상의 전환규칙을 중심으로 <그림 10>의 객체모델을 관계형 데이터모델로 전환하여 번들상품추천 시스템의 데이터베이스를 설계하면 <그

림 11>과 같다. <그림 10>의 사례표현에서 추가적으로 들어간 관계형 테이블로는 고객의 주문과 주문품목에 대한 실제 거래정보이다. 주문거래에 대한 데이터는 고객이 번들을 그대로 구입하였거나 또는 수정하여 구입한 경우 모두에 대한 정보를 저장하고 있다. 이 거래 데이터는 고객이 원래의 번들을 수정한 내역을 담고 있으므로 고객 수정사항을 분석하는데 사용될 수 있으며, 분석된 내용을 바탕으로 새로운 번들을 구성하는데 사용될 수 있다. 이 경우 수정된 번들 중에서 그 거래 빈도가 높은 또는 높을 것으로 예상되는 번들은 번들 사례베이스에 추가될 수 있다. 번들 사례베이스에 추가하는 방법은 기존 번들과의 유사도를 측정하여 색인생성규칙(indexing rule)을 이용하여 추가할 수 있다.

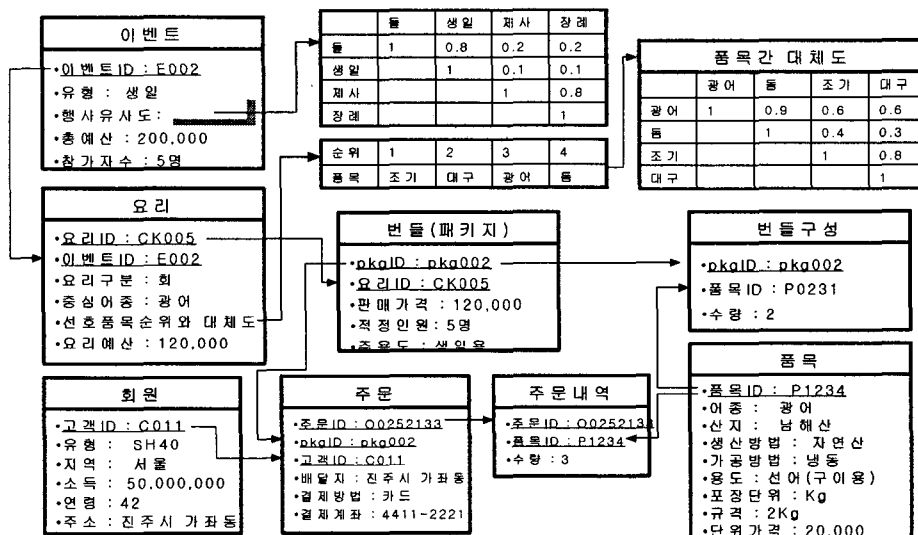
4.4 유사도 측정

객체지향 사례표현에서 질의객체와 사례객체

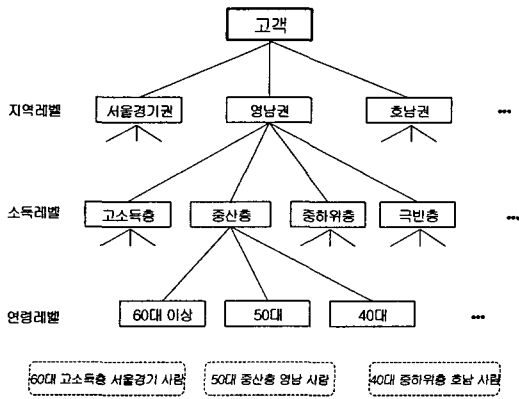
간의 유사도 측정은 전역적 유사도와 지역적 유사도로 나누어 진행된다. 먼저 지역적 유사도 측정을 위하여 고객객체, 이벤트객체, 요리객체와 그 하위객체에 대하여 각각의 지역적 유사도를 측정한다. 이들 객체클래스의 클래스 내 유사도와 클래스 간 유사도는 클래스 계층구조를 이용하여 제 3장 3절에 제시된 방법에 따라 객체클래스 간 유사도를 측정한다. 그리고 전역적 유사도는 고객, 이벤트, 요리클래스의 중요도를 가중하여 합계한다.

4.4.1 고객 유사도 측정

고객객체의 지역적 유사도를 측정하기 위하여 거주지, 소득, 연령속성을 범주형으로 전환하고, 이들 속성을 그 중요도에 따라 다시 계층형으로 전환한다. 이 때 지역을 가장 상위층에 두고, 그 다음에 소득, 그리고 연령순으로 배열한다. 이것을 그림으로 나타내면 <그림 12>와 같다.



<그림 11> 번들상품추천을 위한 데이터베이스의 구성



<그림 12> 고객유사도 측정을 위한 계층적 분류

4.4.2 이벤트 유사도 측정

이벤트정보는 사례구분에 있어 가장 중요한 영향을 미치는 정보이다. 이벤트정보의 유사도 측정은 기본적으로 <그림 10>의 이벤트유형별 계층구조에 따라 클래스 내 유사도와 클래스 간 유사도가 결정된다. 이벤트클래스와 그 하위클래스로 구성된 계층트리에서 내부 노드 K_i 의 유사도 값(S_i)을 결정하고 클래스 간 유사도를 측정하는 과정은 다음과 같다.

먼저, 이벤트별로 사용빈도가 높은 수산물의 선호 관계를 조사하여 <표 4>와 같이 작성한다. <표 4>을 바탕으로 각 이벤트별 수산물 소비 경향의 밀접도를 다음 <그림 13>과 같이 나타낼 수 있다. <표 4>은 실제 수산물 소비 경향을 가상으로 표현한 것이므로 실체는 이와 다를 수 있다. 물론 이러한 직관적인 방법보다는 <표 4>에서 이벤트별 수산물 선호도를 계수화하여 군집분석을 통하여 <그림 13>과 같이 그룹화 할 수 있다.

<표 4>과 <그림 13>를 볼 때, 생일과 회갑, 둘은 서로 비슷한 종류의 수산물로 구성되며, 제사와 장례식 또한 그러하다. 결혼식의 경우 쓰이는 수산물이 다른 이벤트와 차이가 있음을 알 수 있다. 이런 관계를 바탕으로 이벤트에 대한 유사도 테이블을 작성할 수 있다.

다음으로 <그림 13>을 바탕으로 이벤트별 유사도를 부여한다. 이벤트 계층구조에 따라 <표 5>와 같이 이벤트 계층별 유사도와 동일 객체클래스 내 유사도가 부여되었다고 한다면, 특정 질의에 대한 클래스 간 유사도 $SIM_{intra}(class(q), class(c))$ 값과 그 근거가 되는 3장 3절의 관련식을 예시하면 <표 6>과 같다.

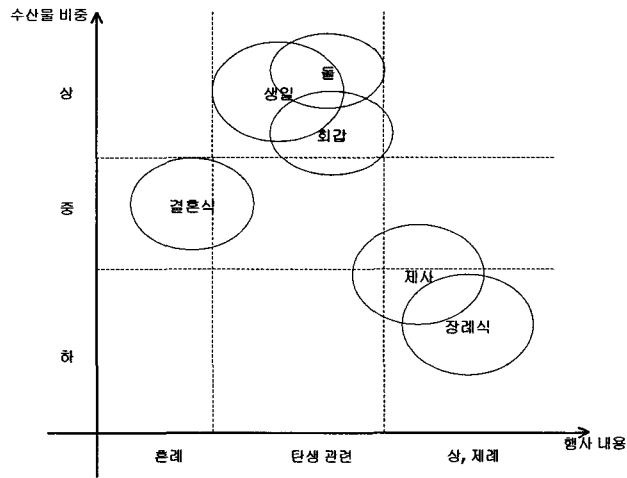
<표 4> 이벤트와 수산물 선호 관계

이벤트 종류	수산물 종류	돔	광어	참치	명게	대구	연어	고등어	도다리	오징어	문어	새우	게	조개	굴	김	미역	통조림류
		탄생축사	생일	○	○	○	△	△	×	△	○	△	×	○	△	×	○	○
	뚝	○	○	○	△	△	△	○	○	△	△	○	×	×	○	△	○	△
	회갑	△	△	○	×	○	×	△	×	○	△	△	○	○	△	×	×	○
혼례사	결혼식	○	○	△	○	△	×	△	○	△	×	○	△	×	○	△	○	×
홍사	제사	×	×	○	×	○	△	○	×	△	○	△	×	△	○	△	×	○
	장례	×	×	○	×	○	○	△	×	○	○	△	×	○	○	×	×	△

※ ○: 많이 쓰임, △:중간 ×:쓰이지 않음

<표 5> 이벤트 계층별 유사도 값(S)와 클래스 내 유사도

계층 0	계층 1	계층 2	계층 3	클래스 내 유사도
이벤트(0.2)	축사(0.6)	탄생사(0.9)	생일	생일->돌(0.9), 회갑(0.7)
			돌	돌->생일(0.8), 회갑(0.8)
			회갑	회갑->생일(0.6), 돌(0.9)
		혼사(0.8)	결혼식	결혼식->약혼식(0.8)
			약혼식	약혼식->결혼식(0.7)
		제물사(0.6)	집들이	집들이->고사(0.8)
			고사	고사-집들이(0.6)
	홍사(0.4)	제사	제사->장례식(0.8)	
		장례식	장례식->제사(0.6)	



<그림 13> 수산물을 기준으로 한 이벤트별 밀접도

<표 6> 질의에 대한 클래스 간 유사도 값 계산 예

질의 \ 사례	돌	결혼식	제사	탄생사	축사	탄생사 ?
돌	1 (식 5)	0.6 (식 5)	0.2 (식 5)	1 (식 6)	1 (식 6)	0.9 (식 9)
결혼식	0.6 (식 5)	1 (식 5)	0.2 (식 5)	1 (식 6)	1 (식 6)	0.6 (식 9)
탄생사	1 (식 7)	0.6 (식 7)	0.2 (식 7)	1 (식 7)	0.6 (식 7)	?
탄생사 ?	0.9 (식 10)	0.6 (식 10)	0.2 (식 10)	0.9 (식 10)	0.6 (식 10)	0.6 (식 10)

따라서 이벤트 클래스의 최종 지역적 유사도는 <표 5>의 클래스 내 유사도 $SIM_{intra}(q,c)$ 와 <표 6>의 클래스 간 유사도 $SIM_{intra}(class(q), class(c))$ 를 곱하여 계산한다.

4.4.3 전역적 유사도

요리 객체클래스도 이벤트 객체클래스와 동일한 방법으로 지역적 유사도를 계산할 수 있다. 이렇게 세 객체클래스의 지역적 유사도가 계산되어지면 특정 질의에 대한 전역적 유사도를 계산한다. 특정 질의에 대한 전역적 유사도 계산은 (식 1)과 같이 객체간의 가중치를 이용하여 결정한다.

$$\alpha(q,c) = \sum_{i=0}^n w_i \sigma_i(q_i, c_i) \quad (\text{식 1})$$

이 때 i 는 고객, 이벤트, 요리 객체이다.

최종 사례의 검색결과는 이들 세 가지 객체와 관련을 갖는 번들(패키지)이 전역적 유사도 값의 크기에 따라 내림차순으로 제시되어진다.

4.5 프로토타입

이상에서 설계된 사례베이스와 유사도 척도를 바탕으로 Inductive Solutions사의 CBR 도구인 Induce-It을 이용하여 프로토타입을 개발하였다. Induce-It은 MS-EXCEL을 기반으로 하고 있으며, 사례베이스의 작성과 추론을 위한 매크로들이 제공되고 있다. Induce-It은 객체의 지역적 유사도 측정을 위한 다양한 방법(선택형, 수치형, 순위형, 계층형)을 제공하고 있으며, 사용자가 직접 유사도 측정함수를 정의할 수 있다. 특히 계층형 객체를 <표 5>와 같이 테이블 형태로 저장

하고, 이를 참조하는 방법으로 손쉽게 유사도를 계산할 수 있게 한다. 또한 Induce-It은 MS-EXCEL을 기반으로 하므로 관계형 데이터 모델로 표현된 데이터베이스를 쉽게 수용할 수 있게 하며, VBA를 이용하여 손쉽게 사용자 인터페이스를 작성할 수 있으며, 객체클래스의 각종 메소드를 쉽게 구현할 수 있는 장점이 있다.

Induce-It을 이용하여 구현한 프로토타입 시스템의 질의입력을 위한 인터페이스는 <그림 14>와 같으며, 입력된 질의로부터 사례베이스에서 추론하는 부분은 <그림 15>와 같다. <그림 14>에서 입력된 질의 값은 <그림 15>의 질의영역(Reference 영역, 음영부분)으로 입력되며, 입력과 동시에 사례검색이 이루어진다. 사례검색을 위해 계산된 개별속성의 유사도는 Inductive DB 영역에 저장된다. 각 개별 객체와 속성에 대한 지역적 유사도는 가중합계를 통하여 전역적 유사도로 계산되어져 Scores 영역(<그림 15>의 왼쪽)에 저장되어지고, 각 개별 사례는 전역적 유사도 값에 따라 내림차순으로 정렬되어진다. 가중평가를 위한 각 속성별 가중치는 질의영역 바로

<그림 14> 질의 입력 인터페이스

위의 Weights 영역에 주어진다. 또한 이벤트정보(<그림 15>의 EventType 필드)나 주요리(<그림 15>의 MainDishes 필드)와 같이 클래스 계층 구조를 갖는 항목의 경우 클래스 내 유사도를 계산을 위한 데이터는 별도의 시트에 <표 5>와 같은 형태의 계층구조로 저장하며, 이를 참조하기 위하여 Maps 부분에 그 테이블의 범위를 지정한다. 그리고 그 객체 타입을 "H"로 지정한다. 또한

사례베이스에 사례가 많을 경우 유사도 점수(Score)의 하한을 적용하면, 그 하한값 이상의 점수를 갖는 사례들만을 보여줄 수 있다.

추론이 완료되면, 전역적 유사도 점수를 기준으로 그 순위가 높은 항목에 대하여 <그림 16>과 같이 추천된 번들(패키지)과 그 구성품, 그리고 품목 등에 대한 정보 및 가격을 추출하여 보여준다. 만일, 고객이 추천된 번들의 아이템과 수

10	Scores	PkgID	Event Type	MainDishes	FavoriteFish	NoOfPeople	AmtOfExpenditure	Region	Income	Age	SuccessRate
11	0.959	pkg003	회갑	자연산광어회	광어	10	280,000	경기	50,000,000	60	2
12	0.818	pkg001	생일	모듬회	광어	10	200,000	부산	30,000,000	35	2
13	0.690	pkg002	물	구이	조기	5	150,000	서울	30,000,000	35	5
14	0.618	pkg005	잔례	대구구이	대구	20	320,000	경남	40,000,000	50	5
15	0.574	pkg006	생일	굴비구이	조기	5	90,000	서울	50,000,000	50	3
16	0.525	pkg004	제사	대구구이	대구	8	150,000	경남	30,000,000	50	7

<그림 15> 사례기반추론부분

2	Rank	pkgID	sku	FisheryName	ComeFrom	KeepMethod	ProcMethod	UsedFor	Unit	Qty	UnitPrice	Amount	SubTotal
3	1	pkg003	FL001	광어	남해산	활어	포	회	2	2	50,000	100,000	260,000
4			SR003	석화	남해산	냉장	껍질제거	회	1	3	10,000	30,000	
5			SR005	새우	남해산	냉장	껍질제거	탕	0.5	3	10,000	30,000	
6			FR002	굴	남해산	냉장	껍질제거	탕	0.5	2	50,000	100,000	
7	2	pkg001	FL001	광어	남해산	활어	포	회	2	2	50,000	100,000	196,000
8			SR004	홍합	남해산	냉장	껍질제거	탕	0.5	4	3,000	12,000	
9			SR002	굴	남해산	냉장	껍질제거	탕	0.5	4	7,000	28,000	
10			SR003	석화	남해산	냉장	껍질제거	회	1	4	10,000	40,000	
11			AD001	김	남해산	상온	건조	반찬	0.2	2	8,000	16,000	
12	3	pkg002	FR005	조기	서해산	냉장	소금절임	구이	1.5	1	30,000	30,000	146,000
13			FF002	대구	동해산	냉동	미가금	구이	2	1	50,000	50,000	
14			FR010	명태	동해산	냉장	소금절임	탕	1	2	7,000	14,000	
15			SR006	바지락	남해산	냉장	껍질제거	탕	0.5	2	15,000	30,000	
16			AD002	미역	동해산	상온	건조	반찬	0.2	2	5,000	10,000	
17			SD001	홍합	남해산	냉장	건조	탕	0.5	3	4,000	12,000	

<그림 16> 추천된 번들(패키지)상품의 명세

량을 바꾸고자 한다면, <그림 8>과 같은 번들상품의 맞춤화 과정을 거쳐 새로운 사례로 사례베이스에 등록되어진다. 그러나 이 부분에 대해서는 현재 구현되어 있지 않다.

5. 결론

인터넷 쇼핑몰에서 상품을 판매하기 위해서는 상당한 노력이 필요하다. 특히 수산물과 같이 신선도나 상품의 질이 일정하지 않은 상품인 경우 더욱 더 인터넷을 통한 판매가 쉽지 않다. 더구나 각종 가족 이벤트에서 사용될 수산물을 제대로 알고 있지 않은 상황에서 기존의 구매 경험이나 시장 등에서 상인의 조언으로 수산물을 구입하는 사람인 경우 단순한 상품 사진만을 제공하는 인터넷 쇼핑몰에 대하여 신뢰가 가지 않을 것이다.

따라서 본 논문에서는 인터넷을 통한 수산물 구입을 원활하게 이루어지게 하기 위해 고객이 쓰고자 하는 목적에 알맞은 수산물 구성을 미리 알려주고 고객의 의사결정을 돕기 위해 사례기반추론을 이용한 번들상품추천 시스템을 제안하였다. 이 시스템은 차림상이라고 하는 팩키지형 상품을 고객이 선택할 경우 이것을 사례라는 형태로 사례 베이스에 등록하여 고객이 원하는 상품 구성을 찾아 제시하는 형식을 취한다. 사례기반추론기법은 규칙기반추론(rule-based reasoning)에 비해 지식의 전이와 설명이 용이하며, 항상 주어진 문제에 대한 근사해를 제공한다는 장점을 가지고 있다. 이로써 고객이 직접 상품을 고르고 비교해야 하는 불편함을 해소할 것으로 기대된다. 그리고 다양한 사례를 이용함으로써 다른 고객의 구매 정보를 손쉽게 얻어 상품 구매에 반영

할 수 있는 장점이 있다.

하지만 사례기반추론이 제대로 활용되기 위해서는 충분한 사례가 미리 사례 베이스에 들어 있어야 하며, 기계학습의 경우 상당한 시간이 필요하며, 수작업을 통한 사례의 등록 시에는 과도한 노력이 요구된다. 사례베이스 내에 방대한 량의 사례를 체계적으로 저장하고 관리하기 위해서는 사례의 표현기법이 중요하며, 이를 통해 표현된 사례베이스로부터 질의내용에 적합한 유사사례를 검색하기 위해서는 유사도 측도의 개발이 중요하다. 본 연구에서는 번들상품추천을 위한 사례표현기법으로 OMT에서 제시된 객체지향 표현방식을 사용하고 있다. 객체지향 사례표현방법은 "is-a" 또는 "part-of" 관계를 통하여 사례클래스를 계층적으로 분류하고 조직화함으로써 계층적 네트워크가 갖는 의미론(semantics)을 직접적으로 이용하여 클래스 간 유사도를 손쉽게 계산할 수 있는 매커니즘을 제공해준다. 또한 객체지향적으로 표현된 사례는 관계형 데이터베이스로 쉽게 전환할 수 있어 쇼핑몰과 같이 대규모의 관계형 데이터베이스를 이용하는 시스템과 쉽게 결합될 수 있는 장점이 있다. 더 나아가 객체지향 사례표현법을 이용할 경우 문제영역에 존재하는 여러 가지 제약사항과 영역지식을 표현할 수 있다. 또한 상위클래스에서 만들어진 영역지식은 클래스 계층관계를 통하여 하위 클래스에 상속하여 지식재사용의 가능성을 높일 수 있다.

본 연구에서는 번들상품추천 시스템의 사례베이스개발을 위한 사례의 표현에 있어 사례객체의 계층적 관계("is-a", "part-of") 표현에 중점을 두고 있다. 그런데 실제의 수산물 인터넷 쇼핑몰에서는 번들상품을 구성하는 데에는 다양한 제약요소들이 존재한다. 예를 들어 품목간의 궁합, 생산 및 출하시기 등과 같은 것이다. 그리고 번들상품

의 품목구성을 위한 어물전 상인들만이 갖고 있는 고유한 영역지식(예, 특정 어종 간의 대체가능성)이 있다. 이들 제약조건과 영역지식에 대한 표현은 제약조건들 간의 호환성 그래프(Compatibility Constraint Graph: CCG)와 규칙을 이용하여 표현하는 것이 효과적이다. 물론 본 연구에서 사용한 OMT의 객체모델에서도 객체들 간의 제약을 표현할 수 있는 표기법이 있다. 그러나 특정 영역에서 표현의 충분성을 높이기 위해 고안된 CCG가 더 효과적일 수 있다. CCG로 표현된 것을 추론하기 위한 엔진으로 Lee & Kwon(1995)의 충돌해결을 위한 제약규칙기반추론(CRSP)이 있다.

본 연구에서 설계된 사례베이스에 따라 번들상품을 추천할 경우, 하나의 질의에 대하여 동일한 유사도를 갖는 번들상품이 여러 개 추천될 수 있다. 이 경우 번들상품추천을 위한 우선순위가 제시되어야 한다. 일반적으로 구매횟수가 높은(성공율이 높은) 번들이 우선적으로 추천된다. 이 경우 또 다른 방법으로는 추천된 번들상품 중에서 가장 표준적인 상품구성을 먼저 추천할 수도 있다.

끝으로 추천된 번들상품에 대하여 고객이 그 구성을 수정하기를 원할 경우, 선택된 번들상품과 유사도가 높은 번들상품을 추천할 수 있다. 이 경우 현재 이용 가능한 번들상품들 간의 유사도를 측정해야 한다. 번들상품의 유사도 측정을 위해서는 번들상품 구성품목의 수와 종류, 구성품목들 간의 상관관계 또는 대체도 등 여러 가지 특성들이 고려되어야 한다. 따라서 사례의 질의에 대한 해(solution)로 제시되는 번들상품들 간의 유사도를 측정하는 방법에 대한 연구가 필요하다.

참고문헌

- 성백균, 김상희, 박덕원, "전자상거래를 위한 사례기반추론의 판매지원 에이전트," 한국정보처리학회 논문지, 제7권, 제5호, 2000, pp. 1649-1656.
- 정대율, "상차립중심의 지능형 수산물 인터넷 쇼핑몰 개발," 정보시스템연구, 제 10권, 제 2호, 2001. pp. 5-32.
- Aamodt, A., and E. Plaza, "Case-based Reasoning : Foundational Issues, Methodological Variations, and System Approaches," *Artificial Intelligence Communications*, Vol. 7, No. 1, 1994, pp. 39-59.
- Agrawal, R., T. Imielinski, and A. Swami, "Mining Association Rules between Sets of Items in Large Database," *Proceedings of the ACM SIGMOD Conference on Management of Data*, Washington, D.C., May 1993, pp. 207-216.
- Ansari, A., S. Essengaier, and R. Kohli, "Internet Recommendation Systems," *Journal of Marketing Research*, Vol. 37, August 2000, pp. 363-375.
- Bergmann, R., "On the Use of Taxonomies for Representing Case Features and Local Similarity Measures," In L. Gierl & M. Lenz (Hrsg) *Proceedings of the 6th German Workshop on Case-Based Reasoning (GWCBR'98)*. IMIB Series Vol. 7, Universitat Rostock, 1998, pp. 23-32.
- Bergmann, R., and A. Stahl, "Similarity Measure for Object-Oriented Case Representation," *Proceedings of European Workshop on Case-Based Reasoning, (EWCBR'98)*, Springer Verlag 1998.
- Bergmann, R., W. Wilke, I. Vollrath, S. Wess, "Integrating General Knowledge with Object-Oriented Case Representation and Reasoning," In: H. D. Burkhard, and M. Lenz (Hrsg.) *4th German Workshop: Case-Based*

- Reasoning - System Development and Evaluation*, Informatik-Berichte Nr. 55, Humboldt-Universität Berlin, 1996, pp. 120-127.
- Cunningham, P., R. Bergmann, S. Schmitt, R. Traphoner, B. Smyth, and P. MacAnUltaigh, "WEBSSELL: Intelligent Sales Assistants for the World Wide Web." *KI - Kunstliche Intelligenz*, 2001. pp. 28-32.
- Donner, M., and T. Roth-Berghofer, "Architectures for Integrating CBR-Systems with Databases for E-Commerce," *Proceedings of the German Workshop on Case-Based Reasoning(GWCBR '99)*, 1999.
- Gardingen, D., and I. Watson, "A Web Based CBR System for Heating Ventilation and Air Conditioning Systems Sales Support," *Knowledge-Based Systems*, Vol. 12, 1999, pp. 207-214.
- Han, J., and Y. Fu, "Mining Multi-Level Association Rules in Large Databases," *IEEE Transaction on Knowledge and Data Engineering*, Vol. 11, No. 5, 1999.
- Kim, S. H., J. E. Jung, and G. S. Jo, "Recommender System with Configuration in Electronic Commerce," *Proceeding of International Conference on Electronic Commerce 2000*, ICEC, pp.112-116.
- Kim, Steven H., S. W. Shin, and J. H. Kim, "Personalized Recommendations for Retailing in Internet Commerce: A Multistrategy Filtering Approach," *Proceeding of International Conference on Electronic Commerce 2000*, ICEC, pp. 103-111.
- Kolodner, J., *Case-based Reasoning*, Morgan Kaufmann, 1993.
- Kolodner, J., "Retrieving Events from Case Memory : A Parallel Implementation," *Proceedings from the Case-based Reasoning Workshop*, DARPA, Clearwater Beach, 1988, pp. 233-249.
- Konstan, J., B. Miller, D. Maltz, J. Herlocker, L. Gordon, and J. Riedl, "GroupLens: Applying Collaborative Filtering to Usenet News," *Communications of the ACM*, Vol. 40, No. 3, 1997, pp. 77-87.
- Lee, J. K., and S. B. Kwon, "ES*: An Expert Systems Development Planner Using a Constraint and Rule-based Approach," *Expert Systems with Applications*, Vol. 9, No. 2, 1995.
- Lenz, M., B. Bartsch-Spörl, H. D. Burkhard, and S. Wess(Eds.), *Case-Based Reasoning Technology - From Foundation to Applications*, Springer, Berlin, 1998.
- Osborne, H., and D. Bridge, "Similarity Metrics: A Formal Unification of Cardinal and Non-Cardinal Similarity Measures," In: David B. L., and E. Plaza, *Case-Based Reasoning Research and Development, Proceedings of the 2nd International Conference on Case-based Reasoning (ICCBR-97)*, 1997, pp. 235-244.
- Rahmer, J., and A. Voss, "Case-Based Reasoning in the Configuration of Telecooperation Systems," *AAAI 1996 Fall Symposium Workshop*, 1996.
- Resnick, P., and H.R. Varian, editors, *Special Issue on Recommender Systems, Communications of the ACM*, Vol. 40, No. 3, 1997.
- Rumbaugh, J., M. Blaha, W. Premerlani, R. Eddy, and W. Lorensen, *Object-Oriented Modeling and Design*, Prentice-Hall, Englewood Cliffs, NJ, 1991.
- Schank R., *Dynamic Memory: A Theory of Reminding and Learning Incomputers and Peoples*, Cambridge University Press. 1982.
- Slade, S., "Case-based Reasoning : A Research Paradigm," *AI Magazine*, Spring 1991, pp. 42-55.
- Schmitt, S. & R. Bergmann, "Applying Case-Based

- Reasoning Technology for Product Selection and Customization in Electronic Commerce Environments," *12th International Bled Electronic Commerce Conference*, 1999a.
- Schmitt, S., and R. Bergmann, "Product Customization in an Electronic Commerce Environment Using Adaptive Operators," *7th German Workshop on Case-Based Reasoning*, 1999b.
- Schmitt, S., R. Maximini, G. Landeck, and J. Hohwiller, "A Product Customization Module Based on Adaptation Operators for CBR Systems in E-Commerce Environments," *5th European Workshop on Case-Based Reasoning*, Springer Verlag, 2000.
- Stahl, A., and R. Bergmann, "Applying Recursive CBR for Customization of Structured Products in an Electronic Shop," *5th European Workshop on Case-Based Reasoning*, Springer Verlag, 2000a.
- Stahl, A., R. Bergmann, and S. Schmitt, "A Customization Approach for Structured Products in Electronic Shops," *13th International Bled Electronic Commerce Conference*, 2000b.
- Vollrath I., W. Wilke, and R. Bergmann "Intelligent Sales Support on the Web: Using Case-Based Reasoning Techniques for Online Catalogs and Product Databases," *IEEE Internet Computing*, Vol. 2, No. 4, 1998, pp. 47-54.
- Vollrath I., W. Wilke, and R. Bergmann "Intelligent Electronic Catalogs for Sales Support: Introducing Case-Based Reasoning Techniques to On-Line Product Selection Applications," In R. Roy, T. Furuhashi, P.K. Chawdhry (Eds.), *Advances in Soft Computing - Engineering Design and Manufacturing*, Springer, London, 1999.
- Watson, I., *Applying Case-Based Reasoning Techniques for Enterprise Systems*, Morgan Kaufmann Publisher, California, 1977.
- Wilke, W., "Knowledge Management of Intelligent Sales Support in Electronic Commerce," DISKI 213, Inflex Verlag, 1999.
- Wilke, W., M. Lenz, and W. Wess, "Intelligent Sales Support with CBR," In: M. Lenz, B. Bartsch-Sparl, H.D. Burkhard, and S. Wess(Eds.), *Case-Based Reasoning Technology - From Foundation to Applications*, Springer, Berlin, 1998. pp. 91-101.

Abstract

An Object-Oriented Case-Base Design and Similarity Measures for Bundle Products Recommendation Systems

Dae-Yul Jeong*

With the recent expansion of internet shopping mall, the importance of intelligent products recommendation agents has been increasing. For the products recommendation, This paper propose case-based reasoning approach, and developed a case-based bundle products recommendation system which can recommend a set of sea food used in family events. To apply CBR approach to the bundle products recommendation, it requires the following 4R steps : ① Retrieval, ② Reuse, ③ Revise, ④ Retain. To retrieve similar cases from the case-base efficiently, case representation scheme is most important. This paper used OMT(Object Modeling Technique) to represent bundle products recommendation cases, and developed a similarity measure method to search similar cases. To measure similarity, we used weight-sum approach basically. Especially This paper propose the meaning and uses of taxonomies for representing case features.

Key words : CBR(Case-Based Reasoning), Object-Oriented Case Representation, Similarity Measure, Bundle Productions Recommendation System.

* College of Business Administration, Gyeongsang National University