

## 이산 확률 기법을 이용한 온톨로지 기반 교육 시스템

이윤수\*

# Ontology based Educational Systems using Discrete Probability Techniques

Yoon Soo, Lee \*

### 요 약

기존의 인터넷 시스템에서 사용자의 요구와 목적에 따른 프레젠테이션과 콘텐츠를 탐색하는 데 있어 풍부한 링크 구조 등으로 인해 일부 심각한 유용성 문제를 발생시키고 있다. 최근 많은 시스템에서는 온톨로지 기술을 이용해서 동적 적용이 가능한 다양한 시스템들을 연구하고 있다. 이 연구에서는 이산 확률 분포 함수와 사용자 프로파일 기반의 동적 적용 모델을 적용한 온톨로지 기반 교육시스템을 설계하였다. 이 시스템은 온톨로지를 이용해서 교육 콘텐츠의 재사용성을 향상시켰고, 이산 확률 분포 함수와 동적 적용 모델을 이용해서 학습자에게 동적 콘텐츠를 제공할 수 있도록 하였다. 이 모델은 응용 영역을 동적 적용 객체의 가중치 방향성 그래프로 표현하며, 사용자 행위를 이산 확률 함수를 동적으로 구축하는 접근 방식을 이용하여 모델링한다. 제안한 확률적 해석은 온톨로지 기반 환경에서 사용자의 탐색 행위를 추적하여 사용자 행위에 대한 잠재적 속성을 나타내는데 사용될 수 있다. 이러한 접근 방식은 사용자에게 가장 알맞은 프로파일을 동적으로 할당할 수 있다.

### Abstract

Critical practicality problems are cause to search the presentation and contents according to user request and purpose in previous internet system. Recently, there are a lot of researches about dynamic adaptable ontology based system. We designed ontology based educational system which uses discrete probability and user profile. This system provided advanced usability of contents by ontology and dynamic adaptive model based on discrete probability distribution function and user profile in ontology educational systems. This models represents application domain to weighted direction graph of dynamic adaptive objects and modeling user actions using dynamically approach method structured on discrete probability function. Proposed probability analysis can use that presenting potential attribute to user actions that are tracing search actions of user in ontology structure. This approach methods can allocate dynamically appropriate profiles to user.

▶ Keyword : 온톨로지 시스템(Ontology System), 이산 확률 분포 함수(Discrete Probability Distribution Function), 동적 적용 모델(Dynamic Adaptive Model), 교육 시스템(Educational System)

• 제1저자 : 이윤수

• 접수일 : 2007.1.15, 심사일 : 2007.1.24, 심사완료일 : 2007. 3.16.

\* 안산공과대학 컴퓨터정보과 부교수

※ 이 논문은 2005학년도 안산공과대학 학술연구비에 의하여 연구된 것임.

## 1. 서론

인터넷의 활용이 일반화 되고, 원격 교육에 대한 요구가 증대함에 따라 교육 분야에도 웹 환경으로의 변화가 급속히 이루어지고 있다[1]. 또한, 인터넷의 사용이 일반화됨에 따라 인터넷을 이용하는 통신 및 사이트의 수가 증가하고 있어, 여러 분야의 다양한 정보의 홍수 속에서 사용자가 원하는 정보를 찾는 것이 점차 어려워짐에 따라서, 사용자가 원하는 콘텐츠를 정확히 찾아주고 사용자의 특성에 따라서 적절한 자료를 제공해 주는 기술들이 필요하다[2,3,4,5].

일반적인 웹 시스템은 모든 사용자에게 페이지 내에 같은 링크를 제시한다. 탐색 문제(search problems)를 해결하기 위해, 시스템은 각 사용자별로 개인화된 링크나 사이트 맵이나 콘텐츠 목차 같은 탐색 도구들을 제공해야 한다. 따라서, 시스템은 사용자의 관심, 선호도, 전문성 등의 정보를 바탕으로 가장 알맞은 도구들을 제시해야 한다.

콘텐츠 개발자가 미처 예측하지 못한 탐색은 사용자에 대한 이해력 문제를 발생시키며, 개발자는 모든 페이지에 대해 사용자가 해당 페이지를 액세스할 때 사용자에게 대가정을 예측해 저장해야 한다. 그러나, 모든 사용자에게 가장 알맞은 링크 구조를 모두 예측하고 개발한다는 것은 어려운 일이다. 따라서, 페이지는 항상 같은 방식으로 제시되며, 이것은 사용자가 이미 방문한 페이지일 경우 중복된 정보를 포함하게 될 것이고, 사용자에게 대한 부족한 예측으로 사용자는 원하는 정보를 찾기 위해 많은 시간을 보내야 할 것이다.

이러한 문제를 해결하고 보다 정교한 서비스를 위해 사용자 모델링(user modeling)과 온톨로지를 이용한 개인화 기술에 대한 연구가 최근에 활발하게 진행되고 있다[6]. 개인화 시스템은 사용자 등록 정보의 이용뿐만 아니라 사용자 탐색 행위를 관찰해 자동적으로 사용자 프로파일을 변경하여 동적인 적응성을 갖고 있는 동적 적응 시스템(dynamic adaptive systems)이다. 동적 적응은 사용자 모델(user models)을 참조하여 영역 모델(domain models)로부터 적절한 콘텐츠와 탐색 지원을 제공하여 개인화 목적을 달성하기 위한 기술로서, 주로 규칙을 기반으로 한 생성 시스템과 확률(probability)과 베이저안 기법(bayesian techniques) 등의 통계적 기법을 사용한 평가에 대한 연구들이 진행되고 있다[7]. 또한, 온톨로지는 교육 시스템간의 정보 재사용과 공유를 가능하게 하며, 지능화되고 개인화된

교육을 가능하게 한다. 온톨로지가 가지고 있는 가능성은 e-Learning에 새로운 기회를 부여할 것이며, 학습자는 보다 쉽고 개인화된 방법으로 교육 시스템과 의사소통이 가능하다[8].

이 연구에서는 웹 미디어 시스템에서 학습 목표에 접근하기 위한 탐색 문제를 해결하기 위해 동적으로 개인화된 서비스를 할 수 있는 이산 확률 기법(discrete probability techniques)에 기반한 동적 적응 모델(dynamic adaptive models)을 이용한 온톨로지 기반 교육 시스템을 설계한다. 논문의 구성은 다음과 같다. 먼저, 2장에서는 이 연구와 관련된 연구들에 대해서 고찰하였고, 3장에서는 온톨로지 기반 교육 시스템을 설계하였다. 그리고 4장에서는 설계한 시스템을 평가하고, 5장에서는 결론과 향후 연구 과제를 알아 보았다.

## 2. 관련 연구

### 2.1 사용자 모델

사용자의 특성 파악이 정확하면 사용자가 원하는 정보를 보다 효과적으로 제공할 수 있을 것이다. 시스템에 저장된 사용자에게 대한 정보 형태를 사용자 모델(user models)이라고 하며, 사용자로부터 정보를 획득하고 추론하는 과정을 사용자 모델링(user modeling)이라고 한다.

일반적으로 사용자의 정보를 저장하고 유지하기 위하여 사용자 프로파일(user profiles)과 사용자 모델링을 이용한다. 사용자 프로파일은 사용자의 정보 요구와 선호를 표현한다. 사용자 모델링은 사용자에게 대한 정보를 자동화 시스템에서 어떻게 취득할 것이며, 이 정보의 성능향상을 위하여 어떻게 사용할 것인가에 대한 보다 넓은 개념이다.

사용자 프로파일은 사용자가 웹 사이트나 다른 시스템에 로그인 했을 때 사용자의 요구나 흥미와 같은 수집되는 정보로부터 얻어지는 사용자에게 대한 모든 정보를 의미한다. 개략적으로 검색 시스템이 사용자에게 의해 나타나는 목적을 자동적으로 수집하기 위하여 프로파일 기반으로 얻어지는 사용자 프로파일은 구조적이고, 정적인 표현이다[8].

사용자 모델링은 사용자의 흥미와 습관의 프로파일 생성을 포함하고, 사용자 모델링 시스템은 사용자 프로파일을 구성하는 방법에 따라 다르며, 동적으로 사용되어 진다. 사용자 모델은 사용자 프로파일과 다른 다음과 같은 부가적인 특성을 갖는다[9].

- 사용자 자료를 기반으로 사용자의 흥미와 선호를 배운다.
- 유사한 이웃과 동일한 사용자 하부 그룹(스테레오타입)에 대한 가정을 기반으로 각각 사용자의 흥미와 선호를 예측한다.
- 영역 지식을 사용하여 부가적인 흥미와 선호를 추론한다.
- 명시적으로 제공된 정보와 암시적으로 얻어진 가정을 저장하고 갱신하며, 삭제한다.
- 사용자 모델 내용의 일관성과 개인성을 보호한다.
- 사용자에 대한 현재의 자료로 권한이 부여된 응용을 제공한다.

사용자 프로파일은 사용자 모델링의 특수한 경우이며, 특정한 영역에서 사용자 모델을 만들기 위해 사용되고, 사용자의 목적과 독립적이며, 응용 분야에 초점이 맞춰져 있다. 반면에 사용자 모델링은 사용자에 초점이 맞춰져 있으며, 응용에 적용하기 위해 사용자의 목적을 사용한다.

일반적으로 사용자 모델은 성능 향상에 도움을 줄 수 있다. 보통 시스템에서 계산의 복잡성은 크며, 대부분의 경우 실시간 처리를 요구한다. 시스템이 사용자의 향후를 잘 알 수 있다면 전체의 자료로부터가 아닌 사용자의 흥미만을 고려해서 자료를 검색할 수 있다. 이것은 정확도와 효율성을 향상시킬 수 있다.

## 2.2 스테레오타입

사용자를 평가하기 위한 가장 단순한 방법은 스테레오타입을 기반으로 그들에 대한 예측을 수행하고, 사용자를 하나의 그룹으로 분류하는 것이다. 스테레오타입(stereotypes)은 분류에 속한 사람들의 특성을 가장하여 표준을 만든다. 컴퓨터 시스템에서 사용자의 모델에 대한 스테레오타입의 사용은 Rich의 Grundy 시스템에서 소개되었다(9).

사용자를 평가하기 위한 가장 간단한 방법은 사용자를 일정한 영역으로 분류하고, 그 영역에 고유한 스테레오타입 또는 속성에 따라 예측을 만드는 것이다. 따라서 스테레오타입은 그 영역에 속한 멤버에 관한 표준화된 가정을 포함한다. Grundy는 컴퓨터 시스템에서 스테레오타입을 이용해 사용자 모델을 유지한 예이다. 스테레오타입의 주요 요소는 다음과 같다.

스테레오타입은 스테레오타입을 적용할 사용자에게 전형적으로 참(true)인 정보를 포함하는 몸체와 사용자에게 스테레오타입을 적용하기 위한 활성화 조건(트리거)의 집합으로 구성된다. 스테레오타입 추론은 가용한 스테레오타입의 활성화 조건들을 규칙적으로 평가하는 방법이며, 조건이 어떤 사용자에게 만족된다면, 대응하는 스테레오타입의 컨텐

츠를 그 사용자의 사용자 모델에 통합한다. 스테레오타입은 사용자의 기초적인 자료로부터 사용자의 성향을 평가하여 사용자 모델을 초기화할 수 있다. 사용자 모델은 사용자의 행동양식을 모니터링하여 사용자의 프로파일에 반영한다. 이러한 양식들을 스테레오타입에 반영시키는 연구들이 다양하게 진행되고 있다.

스테레오타입의 프로파일은 사용자가 특정 스테레오타입에 속한 사용자를 나타내는 서술과 얼마나 관련되는가를 평가하기 위한 정보를 정의한다. 프로파일은 분류 데이터 집합을 포함하며, 각각에 대해 스테레오타입에 의해 표현된 사용자 클래스의 서술을 갖는 언어 변수값의 일치성을 기술한다. 어떤 데이터가 주어졌다면, 주어진 데이터가 언어 변수값을 얻기 위한 언어 변수값의 일치성을  $\{0, 1\}$ 안의 숫자이고, 특정 스테레오타입에 속한 사용자의 개별적 백분율에 대응한다. 예측 부분은 스테레오타입에 의해 기술된 사용자 클래스에 속한 전형적인 사용자의 특징과 선호도를 서술한다. 예측 값이 주어지고 사용자가 스테레오타입에 의해 기술된 클래스에 속한다는 것이 주어진다면, 언어 변수 값과 연결된 숫자는 예측값을 위한 언어 변수 값의 확률로 나타낸다.

이 연구에서는 이러한 스테레오타입 모델을 적용하여 웹 미디어 구조에서 사용자의 탐색 행위를 추적하여 얻을 수 있는 사용자 행위에 대한 잠재적 속성을 나타내는데 사용하는 확률적 해석을 제안하여 동적 적용 모델에 적용하였다.

## 2.3 온톨로지 기반 교육 시스템

현재 연구중인 많은 온톨로지 기반 교육시스템들은 현재 W3C의 웹 온톨로지 언어인 OWL(Ontology Web Language)에 기반하고 있다. 시맨틱 웹을 구성하는 핵심 기술요소로는 (1) 웹상의 자원의 성질과 다른 자원과의 관계를 표현하기 위한 방안으로서 RDF(Resource Description Framework)와 같은 자원 서술방안에 관한 기술과, (2) 웹상의 정보를 단순한 데이터 차원에서 처리하여 사람이 의미를 부여하는 현재의 상태에서, 정보 생성단계에서부터 의미를 부여하여 정보가 지식(Knowledge)으로서의 가치를 지닌 상태로 향상시킬 수 있는 지식의 체계적 표현방안인 온톨로지(ontology)에 관한 기술, 그리고 (3) 이들 자원에 관한 메타정보와 지식체제인 온톨로지를 이용하여 추론하고 사용자 대신에 유용한 일을 비교적 자율적으로 실행할 수 있는 컴퓨터 프로그램인 에이전트 기술로 분류할 수 있다(7). 이들 요소 기술들은 모두 기존의 unicode를 지원하는 XML(eXtensible Markup Language)을 기반으로 한다.

Devdezic은 그의 연구에서 보다 지능적이고 적응적인 교육을 지원하기 위해서는 온톨로지 지원이 필수적이라고 기술

하였다. 온톨로지 서술이 적용된 교육 콘텐츠의 생성은 온톨로지 기반 저작 도구의 지원에 의해서 가능하며, 이 저작 도구는 도메인 온톨로지 용어들과 학습 이론들 및 교수 설계 과정을 지원하도록 하였다(10). Mizoguchi와 Bourdeau는 온톨로지 지원 저작 도구 개발을 위한 로드맵을 제시하였다(11). 이러한 다양한 연구들은 온톨로지 기술이 지능형 교육 시스템에 매우 적합한 기술임을 보여주고 있다. 이 연구에서는 온톨로지를 이용해서 콘텐츠의 재사용성을 향상시키고, 학습자에게 동적 콘텐츠를 제공할 수 있도록 하였으며, 개인화된 교육을 지원하기 위해서 동적 적응 모델을 적용한 온톨로지 기반 교육시스템을 설계한다.

### 3. 동적 적응 모델을 이용한 온톨로지 기반 교육 시스템 설계

이 장에서는 이산 확률 기반 동적 적응 모델을 제안하고, 제안한 모델을 이용해서 온톨로지 기반 교육 시스템을 설계한다.

#### 3.1 시스템 설계

시스템은 그림 1. 과 같이 전체적으로 3개의 모델러와 2개의 데이터베이스로 구성된다. 영역 모델러(domain modeller)는 웹 미디어의 정적 특성들을 모델링하며, 사용자 모델러(user modeller)는 동적 특성들을 모델링하게 된다.

동적 적응 엔진은 정적 특성과 동적 특성들을 이용하여, 사용자에게 가장 알맞은 최적의 프로파일을 동적으로 할당하게 된다. 시스템은 윈도우 XP 서버 환경에서 웹 서버는 아파치 2.0, 데이터베이스는 MySQL 5.0, 웹 서버 스크립트 컨테이너는 JDK 1.5와 Tomcat 5.1, 문서의 동적 변환을 위한 COCOON 2.0을 이용하였다. 또한, 온톨로지 추론을 위해서 Jena 2.0을 이용하였다.

특정한 사용자를 위한 사용자 모델은 사용자 프로파일 에디터(user profile editor)를 통해서 제공되는 외부 정의에 기반한다. 사용자 온톨로지 데이터베이스에는 사용자와 사용자들간의 관계에 대한 정보를 포함한다.

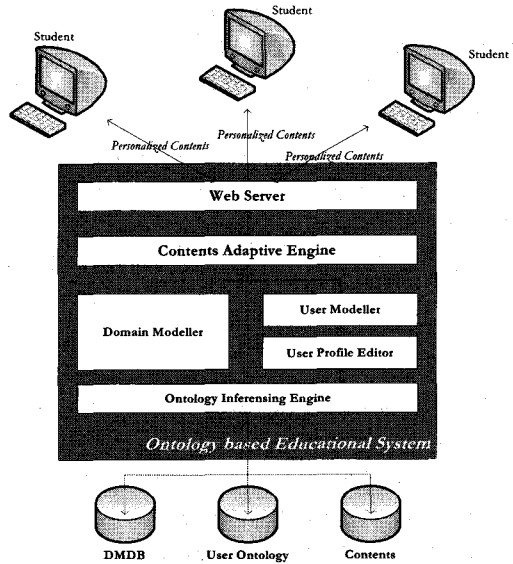


그림 1. 동적 적응 모델을 이용한 온톨로지 기반 교육 시스템 설계  
Fig. 1. Ontology based Educational System using Dynamic Adaptive Model

도메인 데이터베이스에는 도메인과 응용의 특성에 따라서 필요한 정보들과 그들간의 관계를 표현한다. 마지막으로 콘텐츠 데이터베이스에는 교육 콘텐츠에 대한 메타 데이터를 저장하고 있다. 이들 정보는 온톨로지 추론 엔진을 통해서 교육 시스템과 연결되며, 교육 시스템은 이들 정보를 이용해서 사용자에게 동적인 콘텐츠를 생성하고 이를 웹서버를 통해서 제공한다.

#### 3.2 확률적 해석기반의 동적 적응 모델

이 절에서는 웹 미디어 스키마 구조(webmedia schema structure)에서 아크의 가중치에 대한 확률적 해석(probabilistic analysis)을 제안한다. 이것은 웹 미디어 구조에서 사용자의 탐색 행위를 추적하여 사용자 행위에 대한 잠재적 속성을 바탕으로 사용자 분류를 하는데 사용된다. 즉, 사용자 모델은 초기에 스테레오타입으로 구축하고, 실행 시에 사용자 행위를 모니터링해서 사용자의 행위에 따라 사용자 모델의 프로파일을 동적으로 조정하는 것이다.

이 연구에서는 전체 응용 영역을 동적 적응 콘텐츠 객체(dynamic adaptive contents objects)의 가중치 방향성 그래프(weighted direction graph)로 생각한다. 이를 수식으로 나타내면, M개의 다른 프로파일을 갖는 응용 영역(application domain)과 동적 적응 콘텐츠 객체의 집합을 N이라 하면, 프로파일  $k=1, \dots, M$ 에 대해  $i \in \mathbb{N}$ 인  $i$ 라는 노

드에 j라는 도착 노드까지의 출력 링크를 집합 Lik라 하자.

가중치 방향성 그래프 G=(N,E)에서 N은 각 노드를 나타내고, E는 각 방향성 아크를 나타내며, 이를 수식으로 표현하면 다음과 같다.

$$G_k = (N_k, E_k)$$

$$E = L_{i1} \cup L_{i2} \cup \dots \cup L_{iM} = \bigcup_{i \in N} \bigcup_{k=1, \dots, M} L_{ik}$$

수식을 단순화하기 위해 가중치 방향성 그래프 G에서 각 프로파일에 대응하는 노드와 아크를 가중치 그래프 G<sub>k</sub> (k=1, ..., M)라 하고, 이 G<sub>k</sub>를 논리적 탐색 그래프(logical search graph)라 한다.

E<sub>k</sub>에서 (i,j) 아크의 가중치를 W<sub>k</sub>(i,j)라 하고 이것을 조건부 확률 P(j|k,i)로 정의하자. 즉, 프로파일 k에 속한 사용자가 i 노드에 도착해서 j 노드에 링크될 조건부 확률로 다음과 같이 정의된다.

$$W_k(i, j) : E_k \rightarrow [0, 1]$$

$$W_k(i, j) = P(j|k, i), (i, j) \in E_k, k = 1, \dots, M$$

그래프 G<sub>k</sub>에서 패스 S를 순서 있는 노드의 집합으로 다음과 같이 정의한다.

$$S = \{S_0, S_1, \dots, S_l | (S_j, S_{j+1}) \in E_k, j = 0, \dots, l-1\}$$

프로파일 k에 속한 사용자가 패스 S를 따를 확률은 다음과 같다.

$$\begin{aligned} P_s^k &= W_k(S_0, S_1) \times \dots \times W_k(S_{l-1}, S_l) \\ &= \prod_{j=0, \dots, l-1} W_k(S_j, S_{j+1}) \\ &= P_s^k(S_1|K, S_0) \times P_s^k(S_2|K, S_1) \times \dots \times P_s^k(S_l|K, S_{l-1}) \end{aligned}$$

여기서, S<sub>j</sub><sup>k</sup>는 프로파일 k에 속하는 아크를 통하는 노드 i와 j 사이의 일반적인 패스이다. P<sub>s</sub><sup>k</sup>는 패스 S에 속하는 아크에 확률을 계산하는 것으로, 주어진 프로파일 k에 대해 i에서 j의 노드 사이의 패스 중 가장 짧은 패스를 S<sub>ij</sub><sup>k</sup>라 하고 프로파일 k에 대해 i에서 j노드 사이에 패스 S의 확률 중 최대 확률을 P<sub>ij</sub><sup>k</sup>라 하자. 그러면, P<sub>ij</sub><sup>k</sup>는 다음과 같이 정의할 수 있다.

$$\widetilde{P}_{ij}^k = \max_{S_{ij}^k} (P_{S_{ij}^k}^k)$$

이 연구에서는 웹 미디어 구조의 본질적 속성과 3개의 이산 확률 분포 함수(PDF; discrete Probability Distribution Function)로 모델을 제안한다.

$$\mu(k) = \frac{\sum_{q=1}^M \left[ \sum_{(i,j) \in E_k} \widetilde{P}_{ij}^q \delta(k-q) \right]}{\sum_{q=1}^M \sum_{(i,j) \in E_k} \widetilde{P}_{ij}^q}$$

$$p(k) = \frac{\sum_{q=1}^M \left[ \sum_{(i,j) \in E_k} |S_{ij}^q| \delta(k-q) \right]}{\sum_{q=1}^M \sum_{(i,j) \in E_k} |S_{ij}^q|}$$

$$n(k) = \frac{\sum_{q=1}^M [|N_q| \delta(k-q)]}{\sum_{q=1}^M |N_q|} \dots\dots\dots (\text{식 1})$$

(식 1)에서 이들 값은 시간에 따라 변하게 된다. 웹 미디어 구조는 노드나 아크, 가중치 등이 추가되거나 제거될 수 있는 등 동적으로 수정될 수 있다. 이는 여러 사용자의 행위를 일부 자동화된 관찰에 근거하거나 저작자에 의해 응용 영역의 증가된 지식에 근거하여 수정된다. 프로파일의 본질적 관련성을 나타내는 가중치 중앙값(weighted medium value)은 다음과 같이 계산된다.

$$s(k) = \frac{\beta_0 \mu(k) + \beta_1 n(k) + \beta_2 p(k)}{\beta_0 + \beta_1 + \beta_2} \dots\dots\dots (\text{식 2})$$

여기서, μ(k)와 p(k)의 값은 노드가 없는 프로파일은 높은 확률을 갖는 패스를 가질 수 없기 때문에 서로 적절하게 조절되어야 한다. s(k)에서 각 항의 높은 값은 프로파일 k에 대한 높은 관련성을 나타내며, 따라서 β<sub>1</sub> > 0이다.

### 3.3 사용자의 동적 분류

이 연구에서 제안한 확률적 해석은 웹 미디어 구조에서 사용자의 탐색 행위를 추적하여 사용자 행위에 대한 잠재적

속성을 나타내는데 사용될 수 있다. 동적 적응 차원에서 사용자 행위에 관련된 속성은 사용자에게 관한 스테레오타입 모델에 연관해서 나타낼 수 있다.

제안된 시스템은 이산 확률 분포 함수  $A(k) (k=1, \dots, M)$  을 구축하는데, 각 그룹에 사용자의 "확률적 소속" 즉, 얼마나 많이 각 프로파일 사용자가 사용자에게 맞는지를 측정하는 것이다. 즉, 사용자 행위에 기반해서 시스템은 동적으로 사용자에게 가장 적합한 프로파일을 할당한다.

탐색은 시작 노드에 연결된 최종 페이지부터 시작한다. 만약 사용자가 이미 등록되어 있다면, 마지막  $A(k)$ 가 현재로 지정된다. 아니면, 사용자는 일반적 프로파일을 할당받거나 질문을 기반으로 해서 계산된 프로파일을 할당받을 수 있다.  $A(k)$ 의 초기값은  $A_0(k)$ 라고 한다.

노드  $R_{r-1}$ 를 방문한 사용자가 다음 링크를 요구했을 때, 시스템은 새로운 이산 확률 분포 함수  $A'(k)$ 를 사용자 행위 변수와 가중치 중앙값  $S(k)$ 를 기초로 계산하고, 사용자에게 할당될 새로운 프로파일을 결정한다. 사용자 행위는 사용자 행위 변수의 집합으로 저장된다.

- 현재의 프로파일  $K_c$
- 현재 이산 확률 분포 함수  $A(k) (k=1, \dots, M)$ 는 각 프로파일에 사용자의 소속 확률을 측정한다.
- 최근 탐색한 패스  $R = \{R_1, \dots, R_{r-1}, R_r\}$ 는 최근 방문한 노드를 포함한다.  $R_{r-1}$ 은 현재 노드이고  $R_r$ 은 다음 노드이다. 최근 아크  $(R_{r-1}, R_r, K_c)$ 는 사용자에게 의해서 선택된 출력 링크이다.
- 최근 노드에서 시간 소모는  $t(R_1), \dots, t(R_{r-1})$ 이다. 이것을 기초로 3개의 이산 확률 분포 함수를 정의한다.

$$c(k) = \frac{\sum_{i=1}^M [P_{R_i}^i \delta(k-i)]}{\sum_{i=1}^M P_{R_i}^i}$$

$$r(k) = \frac{\sum_{i=1}^M [\overline{P_{R_i, R_r}^i} \delta(k-i)]}{\sum_{i=1}^M \overline{P_{R_i, R_r}^i}}$$

$$t(k) = \frac{\sum_{i=1}^M [D^i[i] \delta(k-i)]}{\sum_{i=1}^M D^i[i]} \dots \dots \dots (식 3)$$

마지막으로 프로파일의 동적 관련성을 나타내는 가중치 중앙값은 다음 식으로 계산된다.

$$d(k) = \frac{\alpha_0 c(k) + \alpha_1 r(k) + \alpha_2 t(k)}{\alpha_0 + \alpha_1 + \alpha_2} \dots \dots \dots (식 4)$$

$d(k)$ 에서 각 항목의 높은 값은 프로파일  $k$ 에 대한 높은 관련성을 나타내며, 따라서  $\alpha_i > 0$ 이다.

사용자가 속한 프로파일을 계산하는 주요 아이디어는 사용자의 동적 행위를 결합하는데 있다. 사용자의 동적 행위는 주로 위상 기하학에 따른 웹 미디어 스키마  $s(k)$ 의 구조적 속성과 통합된  $d(k)$ 이다.

그림 2. 는 사용자 행동에 기반한 새로운 확률 분포 함수  $A'(k)$ 를 계산하는 알고리즘을 나타낸 것이다.

**INPUT**

이산 확률 분포 함수  $A(k), A_0(k), s(k)$   
 최근의 방문한 패스  $R = \{R_1, \dots, R_{r-1}, R_r\}$   
 최근에 방문한 노드에서 소비 시간,  
 $t(R_1), \dots, t(R_{r-1})$

**OUTPUT**

새로운 확률 분포 함수  $A'(k)$

**STEP**

새로운 이산 확률분포 함수  $d(k)$  계산  
 새로운 이산 확률분포 함수  $A'(k)$  계산

$$A'(k) = \frac{\gamma_0 A_0(k) + \gamma_1 A(k) + \gamma_2 d(k) + \Delta \gamma_3 s(k)}{\gamma_0 + \gamma_1 + \gamma_2 + \Delta \gamma_3}$$

만약,  $s(k)$ 가 변화하면  $\Delta = 1$ , 그렇지 않으면  $\Delta = 0$ 이다.

그림 2. 새로운 확률 분포 함수 계산 알고리즘  
 Fig 2. New Probability Distribution Function Calculation Algorithm

새로운  $A'(k)$ 는 4 가지 항목의 가중치 평균으로 계산되는데, 첫 번째 항목은 초기 사용자의 선택을 나타낸다. 두 번째 항목은 상호작용의 스토리이다. 세 번째 항목은 단일 사용자의 동적 행위를 나타내는 반면, 네 번째 항목은 웹 미디어의 구조적 속성을 나타낸다.

$A'(k)$ 에서 각 항목의 높은 값은 프로파일  $k$ 에 대한 높은 관련성을 나타낸다. 따라서  $\gamma_i > 0$ 이 된다. 새로운 프로파일은  $A'(k)$  분산에서 난수적 추출을 통해서 선택하거나 가장 높은  $A'(k)$ 값을 참고한다.

## 4. 평가

평가는 프로그래밍 영역에서 수행하였으며 초기에 구축된 스테레오타입의 사용자 모델을 실행 시에 사용자 탐색 행위를 추적하여 사용자의 행위에 따라 사용자에게 적합한 프로파일을 동적으로 적응시킴으로써, 학습 목표에 도달하는데 상당한 탐색 횟수를 줄일 수 있었다. 또한, 동적 적응 시스템이 사용자에게 효율성을 주고 있었으나 효율성을 저작 시 사용자의 스테레오타입 데이터와 추상 개념의 가중치 값의 정확성에 의존한다는 것을 확인하였다.

각 프로파일은 자바 프로그래밍에서 초보자, 중급자, 전문가로 분류하고 학습 목적별로 분류하였다. 하나의 노드는 동적 적응 콘텐츠 객체를 의미하고 하나의 동적 적응 콘텐츠 객체 즉, 하나의 페이지에는 여러 개의 링크가 프로파일 별로 다르게 구성되어 있다. 모든 프로파일 중 시작 노드 N1에서 목적 노드 N7로 갈 수 있는 일반적인 패스와 패스 S의 확률 값은 식 1)에 의하여 계산된 결과인 각 프로파일에 대한 하이퍼미디어의 구조적 속성값을 구한다. 현재 노드를 방문한 사용자가 다음 링크로 다음 노드가 올 경우 시스템은 새로운 이산 확률 분포 함수  $A'(k)$ 를 사용자 행위 변수와 가중치 평균  $s(k)$ 를 기초로 계산하고, 사용자에게 할당될 새로운 프로파일을 결정한다.

각 프로파일에 대해, 현재 이산 확률 분포 함수  $A(k)$  ( $k=1, \dots, M$ )는 각 프로파일에 사용자의 확률로 저작자가 초기에 할당된 값으로  $A(1) = 0.33$ ,  $A(2) = 0.33$ ,  $A(3) = 0.33$ 로 가정한다.

사용자가 속한 프로파일을 계산하는 주요 아이디어는 사용자의 동적 행위를 결합하는데 있다. 그림 2.의 알고리즘에 의해 사용자의 행동에 기반 한 새로운 확률 분포 함수  $A'(k)$ 를 계산하면,  $A'(1)=0.33$ ,  $A'(2)=0.31$ ,  $A'(3)=0.36$ 이 된다. 따라서, 사용자의 행동을 2가지로 분석할 수 있다.

첫 번째는 각각의 프로파일은  $A'(k)$ 의 값이 저작자가 설정한 초기값 0.33보다 차이가 10% 이내이므로 프로파일을 변경하지 않고,  $A'(k)$ 이  $A(k)$ 로 적용하여 다음 노드로 진행하고, 다음 프로파일은  $A'(k)$ 의 값이 0.33보다 차이가 10% 이상이므로, 다음 노드 진행 시 프로파일을 변경하고 사용자에게 적합한 화면을 제공하여 동적으로 적응시킨다. 그리고 다음 노드 진행 시에는  $A'(3)$ 은 초기값으로 설정한다.

계속해서 다음 노드를 방문한 사용자가 다음 링크를 선택

할 경우 그림 2.의 알고리즘에 의해 새로운 확률 분포 함수  $A'(k)$ 를 계산하면,  $A'(1)=0.34$ ,  $A'(2)=0.33$ ,  $A'(3)=0.32$ 이 된다. 따라서 다음 노드를 선택하였을 경우에는 각 프로파일별 확률 분포 함수가 0.03사이에서 있으므로 저작자가 설정한 초기값보다 차이가 10% 이내이므로 프로파일을 변경하지 않고, 다음 노드로 진행하게 된다.

## 5. 결론

이 연구에서는 웹 미디어 시스템에서 이산 확률 해석과 사용자 프로파일 기반의 동적 적응 모델을 제안하고, 제안한 모델을 이용해서 학습자에게 개인화된 교육을 지원하기 위한 온톨로지 기반 교육 시스템을 설계하였다. 확률적 해석 기법은 전체 응용 영역을 동적 적응 콘텐츠 객체의 가중치 방향성 그래프로 생각한 논리적 탐색 그래프를 통해, 웹 미디어가 갖고 있는 정적 특성을 이용한 이산 확률 분포 함수이다.

이러한 확률적 해석은 웹 미디어 구조에서 사용자의 탐색 행위를 추적하여 얻을 수 있는 사용자 행위에 대한 잠재적 속성을 나타내는 데, 동적 적응 차원에서 사용자 행위에 관련된 속성을 사용자에게 관한 스테레오타입 모델에 연관하여 나타내었다.

사용자 행위는 사용자 행위 변수의 집합으로 현재 프로파일, 현재 이산 확률 분포 함수, 최근 탐색한 패스, 최근 노드에서 시간 소모 등을 바탕으로 세 개의 이산 확률 분포 함수를 제안하였고, 이것을 바탕으로 사용자 분류를 하는 새로운 접근 방식을 제안하였다.

동적 적응 콘텐츠 객체는 사용자에게 제시될 페이지 개념의 단위로서, 동적 적응 프레젠테이션의 대상이 된다. 추상 개념과 응용 영역은 동적 적응 콘텐츠 객체를 방향성 그래프 구조로 나타내서 동적 탐색과 사용자 행위에 대한 평가를 할 수 있도록 하였다.

향후 연구로는 정확한 가중치 값을 저작자가 쉽게 결정할 수 있도록 하는 시뮬레이터에 대한 연구가 필요하며, 학습자에게 필요로 하는 보다 정확한 교육 자료를 제공하기 위해서 도메인별로 보다 세밀한 온톨로지 정의에 대한 연구가 필요하다.

## 참고문헌

- [1] Wattenberg, F. A National Digital Library for Science, Mathematics, Engineering, and Technical Education, D-Lib Magazine, Vol.5, No.10, October, 1999.
- [2] Brusilovsky, P., "Adaptive hypermedia," User Modeling and User Adapted Interaction 11(1/2), pp.87-110, 2001.
- [3] Eklund, J. and Brusilovsky, P., & Schwarz, E.(1997), "Adaptive Textbooks on the WWW," <http://ausweb.scu.edu.au/proceedings/eklund/paper.html>.
- [4] Adaptive Hypertext and Hypermedia Home Page, <http://www.wis.win.tue.nl/ah/>.
- [5] Emily Berk & Joseph Devlin, Hypertext/Hypermedia Handbook, McGraw-Hill, 1998.
- [6] Ardissono, L., and Goy, A., "Tailoring the Interaction With Users in Web Stores," in User Modeling and User-Adapted Interaction, Vol.10, No.4, Kluwer Academic Publishers, 2000.
- [7] Mario Cannataro and Andrea Pugliese, "XAHM: an XML-based Adaptive Hypermedia Model and its Implementation," Hypertext'01, ACM, 2001.
- [8] Dicheval, D., Sosnovsky, S., Gavrilova, T., Brusilovsky, P., "Ontological Web Portal for Educational Ontologies", International Workshop on Applications of Semantic Web technologies for E-Learning (SW-EL 05) July, 2005.
- [9] Kobsa, A., ed., "Special Issue on User Modeling Systems," User Modeling and User-Adapted Interaction, Vol. 4, No.2(3), 1995.
- [10] Devedzic, V. "Education and the Semantic Web," International Journal of Artificial Intelligence in Education 14, pp.39-65. IOS Press 2004.

- [11] Mizoguchi, R., & Bourdeau, J., "Using Ontological Engineering to Overcome Common AI-ED Problems," International Journal of Artificial Intelligence in Education, 11, pp.107-121, 2000.

## 저자소개



### 이윤수

1988년 2월 인하대학교 전자계산학과 졸업(학사)

1990년 2월 인하대학교 대학원 전자계산학과 졸업  
(이학석사)

2000년 8월 인하대학교 대학원 전자계산공학과 졸업  
(공학박사)

1995년 3월 ~ 1996년 2월 인하대학교 전임대우강사

1999년 1월 ~ 2006년 12월 안산공과대학 전자계산소장

1996년 3월 ~ 현재 안산공과대학 컴퓨터정보과 부교수

관심분야 : 컴포넌트 기반 소프트웨어 공학, 교육 시스템, E-Learning 및 디지털 콘텐츠, 멀티미디어 시스템