

온톨로지 기반 학습 콘텐츠의 난이도 계산 방법

박재욱*, 박미화*, 이용규*

An Ontology-Based Method for Calculating the Difficulty of a Learning Content

Jae Wook Park*, Mee Hwa Park*, Yong Kyu Lee*

요약

이러닝 시스템에서 난이도를 이용한 학습추천 시스템 설계에 관한 연구가 활발히 진행 중이다. 난이도는 학습자의 수준에 맞는 후행학습을 추천하는데 매우 중요한 요소임에도 불구하고 현행 난이도 기반 학습 추천시스템은 각 학습 콘텐츠의 제작자가 주관적으로 정한 난이도를 적용함으로써 정확한 후행 학습 콘텐츠를 추천하기가 어렵다. 본 논문에서는 객관적인 난이도 지표를 제공하기 위하여 온톨로지에 기반한 새로운 학습콘텐츠 난이도 계산 방법을 제안한다. 기존 온톨로지나 지식맵을 이용한 난이도 계산 방법들은 선행학습과 후행학습 또는 주제간의 선후 관계를 표현하고 이를 이용하여 난이도를 계산하였으나, 이 방법들도 콘텐츠 작성자의 주관적인 판단에 의해 후행학습이 결정된다는 문제점이 있다. 본 논문에서는 이를 해결하기 위하여 콘텐츠를 구성하는 단어들의 상하위 관계 및 심화도를 나타내는 온톨로지를 이용하여 단어들 간 온톨로지의 경로상의 거리로 난이도를 계산한다. 이를 통하여 학습자에게 보다 객관적인 난이도 정보를 제공하고 학습자 수준에 가장 적합한 후행학습 콘텐츠를 추천할 수 있다.

▶ Keyword : 온톨로지, 학습 콘텐츠, 난이도

Abstract

Much research has been conducted on the e-learning systems for recommending a learning content to a student based on the difficulty of it. The difficulty is one of the most important factors for selecting a learning content. In the existing learning content recommendation systems, the difficulty of a learning content is determined by the creator. Therefore, it is not easy to apply a standard rule to the difficulty as it is determined by a subjective method. In this paper, we propose an ontology-based method for determining the difficulty of a learning content in order to provide an objective measurement. Previously, ontologies and knowledge maps have been used to recommend a learning content. However, their methods have the same problem because the difficulty is also determined by the creator. In this research, we use an ontology representing the IS-A relationships between words. The difficulty of a learning content is the sum of the weighted path lengths of the words in the

• 제1저자 : 박재욱 교신저자 : 이용규

• 투고일 : 2010. 12. 24, 심사일 : 2011. 01. 03, 게재확정일 : 2011. 01. 08.

* 동국대학교 컴퓨터공학과-서울(Dept. of Computer Science and Engineering, Dongguk University-Seoul)

※ 이 논문은 2009년도 동국대학교 연구년 지원에 의하여 이루어졌음

learning content. By using this kind of difficulty, we can provide an objective measurement and recommend the proper learning content most suitable for the student's current level.

▶ Keyword :Ontology, Learning Contents, Difficulty

I. 서론

이러닝 시스템이 빠른 속도로 확산되고 그 규모가 커짐에 따라 사용자가 방대한 콘텐츠 중에서 학습자 수준에 맞는 학습 콘텐츠를 스스로 선택하는 것이 쉽지 않다. 따라서, 이러닝 시스템에서 개인화를 위한 추천시스템 설계는 매우 중요한 요소 중 하나이며, 콘텐츠의 난이도는 콘텐츠를 추천하는데 매우 중요한 변수이다. 그러나, 기존 추천방법에서 콘텐츠의 난이도는 콘텐츠를 제작자의 주관적 판단에 의하여 정하고 있기 때문에 정확성이 떨어지고 객관적인 기준을 적용하기 어렵다.

이러한 단점을 해결하기 위해 지식맵을 이용한 방법 [9][10], 온톨로지(Ontology)를 이용하여 난이도를 계산하는 방법 [11] 등이 제안되었다. 기존연구들은 선행 및 후행 학습 또는 주제간의 선후관계로 관계로 표현하고 이를 이용하여 난이도를 계산하기 때문에 콘텐츠가 어느 주제에 속하는지에 대해서 주관적인 판단이 필요하다. 또한, 동일한 주제의 콘텐츠들의 난이도 값을 동일한 값으로 할당하는 단점이 있다.

본 논문에서는 이러한 단점을 해결하기 위하여 콘텐츠의 단어들로 구성된 온톨로지를 이용하여 난이도를 계산하는 방법을 제안한다. 제안한 방법을 통해 콘텐츠의 난이도를 자동적으로 계산할 수 있으며 난이도가 다르지만 동일한 주제에 속하는 콘텐츠의 난이도를 정확하게 계산함으로써, 객관적 기준에 따라 난이도 부여가 가능하다.

본 논문의 구성은 다음과 같다. 먼저, 2장에서 기존 연구에 대하여 기술하고 3장에서 온톨로지를 이용한 난이도 계산 방법에 대해 설명한다. 4장에서는 본 논문에서 제안한 방법의 성능을 분석하기 위한 실험 및 결과를 제시하고 5장에서 결론을 맺는다.

II. 관련 연구

1. 온톨로지

온톨로지에 대해서는 많은 정의들이 있으나 “개념의 형식적이고 명확한 명세”라는 정의가 가장 널리 알려져 있다 [13][15][16].

온톨로지는 개념, 개념간의 관계, 개념의 속성 등으로 구성된다. 이 중 관계는 개념들간의 상관관계로서 상하의 관계, 부분 관계 등으로 표현할 수 있다. 온톨로지는 시맨틱 웹을 구축하는데 있어서의 핵심기술이며 의료, 법률부분의 전문가 시스템, 검색서비스, 표준화, 상황인지 등 매우 다양한 분야에서 이용되고 있으며 학습 콘텐츠추천 분야에서도 이용되고 있다 [3][4][5][6][7][14][17].

2. 학습콘텐츠 추천

학습 콘텐츠 추천방법에는 콘텐츠에 대한 사용자의 평가에 의해 추천하는 협업 여과 추천방법과 인구통계학적 특성이나 프로필, 사례기반 추론 등을 이용하는 규칙기반 추천방법, 그리고 내용의 유사도를 계산하여 추천하는 내용기반 추천방법 등이 있다 [2][8]. 최근에는 특정한 방법만을 독립적으로 적용하지 않고 협업여과 추천방법 및 내용기반 추천방법 등을 함께 적용한 혼합형 추천방법을 많이 사용한다 [2][3][8].

내용기반 추천방법 중 온톨로지를 이용한 방법에는 학습 콘텐츠를 개체로 한 온톨로지를 구축하고 온톨로지와 1:1로 대응되는 역량(competency) 트리를 만들어 개체 즉, 학습콘텐츠 간의 역량 차이분석(competency gap analysis)을 이용하여 후행학습을 추천하는 방법 [14] 등이 있다. 또한, 사용자 프로파일을 온톨로지 구조로 만들고 이를 기반으로 추천엔진을 이용하여 추천 하는 방법 [3] 등이 제안되었다.

Mooney [12]는 협업 필터링을 이용하지 않고 문서분류를 위한 베이지안 기계학습 알고리즘을 적용하여 콘텐츠를 분류하고 이를 이용한 책 추천방법을 제안하였다.

3. 온톨로지를 이용한 콘텐츠 난이도 결정

온톨로지 또는 지식맵을 이용하여 난이도를 결정하는 방법에는 개념들과의 관계로 온톨로지 및 지식맵을 만들고 해당 콘텐츠가 몇 개의 개념과 관련이 있는지 판단하여 난이도를 결정하거나, 학습순서를 지식맵으로 표현하고 난이도를 계산하는 방법 등이 제안되었다 [9][10][11].

Kuo [9]는 지식맵에 기반하여 속성의 개수, 개념에서의 서브트리의 높이, 스키마에 저장되어 있는 학습순번, 문제 매트릭스에서 다루어지는 개념의 개수 등을 추출하고, 이를 이용하여 4가지 난이도를 계산하였다. [수식 1]은 이 중 학습순서와 개념의 깊이에 가중치를 부여하여 그 합을 난이도로 계산

한 식별난이도(identification difficulty)의 식이다.

$$\text{난이도} = w \times \text{learn_seq} + w \times \text{concept_depth} \dots [\text{수식1}]$$

.learn_seq : 학습순서 ÷ 학습순서의 최대값
 .concept depth : (루트의 높이-컨셉의 높이) ÷ 루트의 높이
 .w = 가중치

Gunel[10]은 지식맵을 이용하여 문제와 관련있는 개념의 개수를 측정하여 문제의 난이도를 계산하는 방법을 제안하였다.[수식2]는 Gunel의 난이도 계산식으로, 지식맵을 이용하여 해당 콘텐츠(문제)가 어느 개념에 속하는가에 대하여 개수를 세고 그것을 전체 개념의 개수로 나누어 1에서 10사이의 값을 갖도록 정규화하여 난이도를 계산한다.

$$\text{난이도} = 10 \times \frac{\sum_{i=1}^m N_{C_i}}{n} \dots \dots \dots [\text{수식2}]$$

.m : 콘텐츠(문제) 개수
 .n : 총컨셉 갯수
 .NC_i : 콘텐츠와 컨셉의 관련여부(있으면 1,없으면 0)

Khan[11]은 온톨로지를 이용하여 선행개념과 후행개념과의 관계를 결정해 놓고 각 개념에 가중치를 부여하고 해당 콘텐츠가 어떤 개념에 속하는지 결정하여 난이도를 계산하였다.

그러나, Kuo[6]의 지식맵을 이용한 방법과 Khan[8]의 온톨로지를 이용한 방법은 학습순서가 같거나 비슷한 내용의 콘텐츠들의 난이도를 같게 계산하는 단점이 있다. 또한, Gunel[7]의 난이도 계산 방법은 인위적으로 해당 콘텐츠가 어느 개념에 속하는지 판단해야 하므로 정확하지 못하다.

본 논문에서는 이러한 단점을 해결하기 위해 모든 콘텐츠에서 쓰이는 단어를 개체로 한 상하위 관계의 온톨로지를 구성한다. 그리고 각 개체의 심화도를 나타내는 속성을 추가시키고 단어별 심화도의 합으로 콘텐츠의 난이도를 계산한다.

III. 온톨로지를 이용한 난이도 계산 방법

1. 온톨로지 구성

온톨로지의 개체는 콘텐츠의 단어로 구성한다. 그리고 단어간의 관계는 상하위 관계만 고려하며, 단어의 심화도를 속성으로 부여한다.

그림 1은 콘텐츠의 단어들로 구성된 온톨로지의 예시이다. 그림 1에서 화살표는 단어간의 상하위 관계를 뜻하고 화살표 사이의 숫자는 단어간의 심화도를 표시한 것이다. 왼쪽의 숫

자는 단어의 레벨을 표현한 것이다.

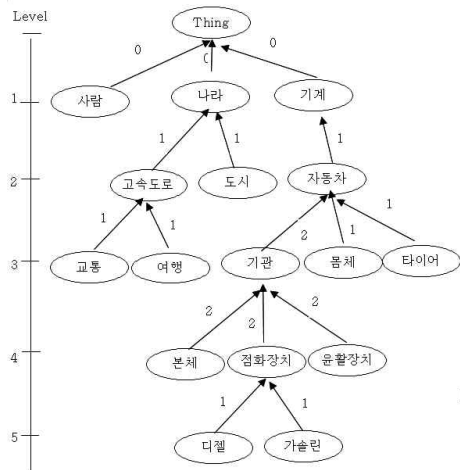


그림 1. 온톨로지 예시
 Fig. 1. An Ontology Example

2. 콘텐츠 난이도 계산 방법

본 논문에서는 온톨로지를 하나의 트리형태로 보며 난이도를 계산한다. 일반적인 온톨로지는 상위단어가 두 개 이상이 있을 수 있는 그래프 구조이므로 본 논문에서는 이를 트리구조로 변경한다. 각 콘텐츠의 난이도는 콘텐츠에서 쓰이는 단어들을 온톨로지에서 찾고 이들 단어들 간의 심화도를 거리로 표현하여 난이도를 계산한다.

온톨로지 루트(Thing) 바로 밑 하위 단어의 거리는 0으로 설정하고 단어들의 거리 합을 정규화하여 0과 1사이의 난이도를 가지며 1에 가까울수록 난이도가 높아지도록 [수식3]과 같이 표현한다.

$$\text{Diff}(c) = \frac{\sum_{i=1}^n d(t_i) \times f(t_i)}{(\sum_{i=1}^n f(t_i)) \times D} \dots \dots \dots [\text{수식3}]$$

. Diff(c): 콘텐츠 c의 난이도
 . n : 색인단어의 갯수
 . d(t_i): 루트에서부터 단어 t_i 까지의 거리의 합
 . f(t_i): 단어 t_i의 빈도수
 . D: 온톨로지의 단어의 거리 중 가장 큰 값

[수식3]에서 빈도수를 고려하지 않을 경우에는 빈도수(f(t))를 1로 계산한다. 단어간의 심화도가 모두 동일하다면 난이도는 [수식4]와 같이 같은 레벨에 있는 키워드의 개수의 합으로 계산한다. 빈도수를 고려하지 않을 경우에는 빈도수

(f(t))를 1로 계산한다.

$$Diff(c) = \frac{\sum_{i=1}^n L(t_i) \times f(t_i)}{\left(\sum_{i=1}^n f(t_i)\right) \times H} \dots\dots\dots [수식4]$$

- . Diff(c): 콘텐츠 c 의 난이도
- . n: 색인단어의 갯수
- . L(ti): 단어 ti의 레벨
- . f(ti): 단어 ti의 빈도수
- . H: 최대 레벨

3. 콘텐츠 난이도 계산사례

예를들어 그림 1의 온톨로지를 이용하고 ‘자동차’를 설명한 콘텐츠 A와 A보다 난이도가 높은 콘텐츠 B의 내용이 아래와 같다고 할 때 각각의 난이도를 계산하는 과정은 다음과 같다.

[콘텐츠 A]

원동기의 동력을 이용해 사람이나 화물을 운송하는 기계를 말한다.
 주로 휘발성 연료를 사용하는 내연기관으로 추진된다. 자동차는 20세기경부터 사용되기 시작해 현대의 어느 발명품보다도 사회를 변화시키는 데 큰 몫을 담당했다. 먼저 장거리 여행이 용이해졌으며, 도시와 멀리 떨어진 곳에서의 생활을 비롯해 원거리 통학이 가능해졌다. 증가하는 자동차의 교통량을 소화해내기 위해 모든 나라는 길을 확장하며, 고속도로를 만들고 있다.

[콘텐츠 B]

기관은 자동차를 움직이기 위한 동력원으로서 연료를 기관 내부에서 연소하는 내연기관인데, 본체 외에 연료장치, 흡·배기장치, 점화장치, 윤활장치, 냉각장치 등으로 구성되어 있다. 기관은 점화방식에 따라 가솔린 기관, 디젤 기관으로 나뉜다. 가솔린 기관은 기화기에서 만들어진 연료와 공기의 혼합 가스를 실린더 내부로 흡입해 피스톤으로 압축할 때 전기불꽃으로 점화·폭발시켜 동력을 얻는 기관이다.
 디젤 기관은 공기만 실린더 내부로 흡입해 압축한 후 고압으로 경우를 분사하여 자연점화로 폭발시켜 동력을 얻는 기관이다.

※ 출처 : 브리태니커 백과사전 (2008년판)

표 1은 콘텐츠 A의 단어와 그 빈도수, 레벨 등을 표현한 것이다. 빈도수를 고려하지 않았을 경우와 고려했을 경우의

콘텐츠 A의 난이도는 [수식3]에서 0.17과 0.20이 되며 [수식4]에서는 동일하게 0.38이 된다.

표 1. 콘텐츠 A의 빈도수와 거리 및 레벨
 Table 1. Frequency, Distance, and Level of Content A

단어	빈도	거리	빈도·거리	레벨	빈도·레벨
사람	1	0	1	1	1
기계	1	0	1	1	1
기관	1	3	3	3	3
자동차	2	1	2	2	4
도시	1	1	1	2	2
교통	1	2	2	3	3
나라	1	0	0	1	1
고속도로	1	1	1	2	2
합 계	9	8	11	15	17

표 2는 콘텐츠 B의 단어와 그 빈도수, 레벨등을 표로 표현한 것이다. 빈도수를 고려하지 않았을 경우와 고려했을 경우의 콘텐츠 B의 난이도는 [수식3]에서 0.74과 0.65이 되며 [수식4]에서는 0.77과 0.71이 된다. 위 결과로 콘텐츠 A보다 콘텐츠 B의 난이도가 더 높음을 알 수 있다.

표 2. 콘텐츠 B의 빈도수·거리 및 레벨
 Table 2. Frequency, Distance, and Level of Content B

단어	빈도	거리	빈도·거리	레벨	빈도·레벨
기관	10	3	30	3	30
자동차	1	1	1	2	2
본체	1	5	5	4	4
점화장치	1	5	5	4	4
윤활장치	1	5	5	4	4
가솔린	2	6	12	5	10
디젤	2	6	12	5	10
합 계	18	31	70	27	64

4. 그래프 구조의 온톨로지 난이도 계산

상위단어가 한 개 이상인 n개인 그래프 구조는 해당단어의 하위개념의 단어들을 n-1번 복사하여 상위단어의 서브트리로 만들어 트리구조로 변경한다. 단, 루트에서부터 단어까지 각 경로의 심화도의 합은 같아야 한다.

그림 2는 그림 1의 온톨로지를 변경하여 자동차의 상위개념의 단어를 기계와 나라로 변경한 그림이다. 그림 2에서 ‘자동차’라는 단어의 상위개념의 단어가 ‘기계’와 ‘나라’ 두 개 이므로 이 형태는 트리 형태가 아니라 그래프 형태이다. 이럴 경우 ‘자동차’의 하위개념들을 복사하여 상위단어인 ‘나라’와

‘기계’의 서브트리로 만들어 트리구조로 변경한다.

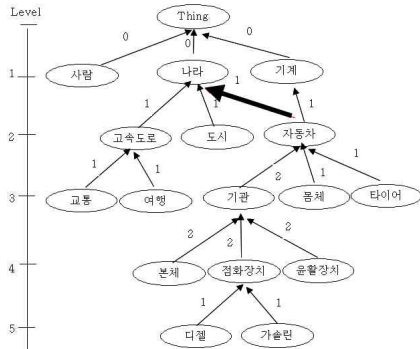


그림 2 그래프 형태의 온톨로지
Fig. 2 A Graph-Structured Ontology

이 때 루트(thing)에서부터 ‘자동차’라는 단어까지의 각 경로의 심화도의 합은 같아야 한다.

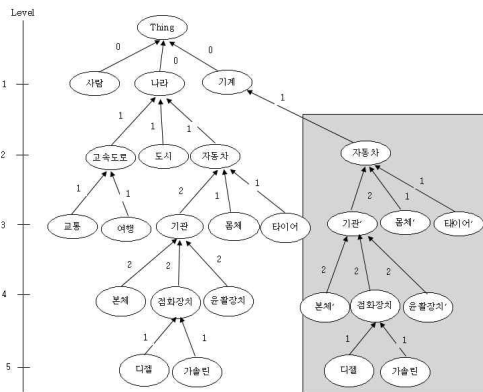


그림 3 그래프 형태를 트리형태로 변형한 온톨로지
Fig. 3. Transformation of a Graph-Structured Ontology into a Tree-Structured Ontology

그림 3은 그림 2를 트리구조로 변경한 것이다. 회색박스로 표시된 부분이 복사된 서브트리이다. 변형된 온톨로지를 이용하여 콘텐츠의 난이도를 계산할 때 복사된 단어의 거리와 레벨은 표 3과 같이 모두 동일하므로 난이도 계산 시 그래프 형태와 동일한 난이도를 계산할 수 있다.

표 3. 복사된 단어의 거리 및 레벨
Table 3. Distance and Level for Copied Words

단어	거리	레벨	단어	거리	레벨
자동차	1	2	본체	5	4
자동차	1	2	점화장치	5	4
기관	3	3	점화장치	5	4
기관	3	3	운할장치	5	4
몸체	2	3	운할장치	5	4
몸체	2	3	디젤	6	5
타이어	2	3	디젤	6	5
타이어	2	3	기술린	6	5
본체	5	4	기술린	6	5

IV. 시스템 구성 및 실험

1. 시스템 구성

학습자 수준에 맞는 후행학습을 추천하는 추천시스템의 구성은 그림 4와 같다. 그림 4에서 학습운영 응용프로그램은 로그인, 학습 및 테스트 기능을 제공하고 학습이 종료된 후 난이도 계산모듈에 의하여 난이도를 계산하고 후행학습을 추천한다. 온톨로지 저장소는 모든 콘텐츠의 색인단어가 온톨로지 형태로 저장되어 있고 콘텐츠 저장소는 모든 콘텐츠의 학습내용이 저장되어 있으며, 학습관리 저장소는 로그인정보, 학습관리 정보 등이 저장되어 있다.

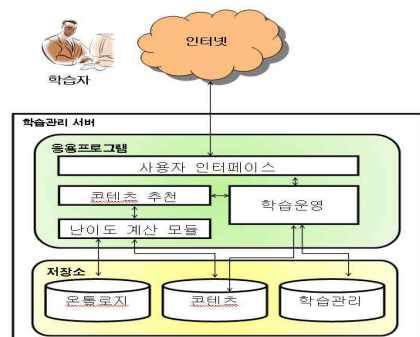


그림 4. 시스템 구성
Fig. 4. System Organization

2. 실험대상 및 방법

제안한 난이도 계산방법의 적정성 판단 및 기존방법과의 차이점을 비교하기 위하여 난이도 계산모듈에 관한 실험을 실시하였다.

실험은 YBM Si-sa에서 시행하는 JET(Junior English Test)[18]의 기출문제를 대상으로 하였다. JET는 초등학교 학생들의 영어능력평가를 위해 개발한 시험으로써 각 회차별로 초급,중급,고급의 3개 등급으로 나뉘어지며 듣기와 읽기영역으로 구성되어 있다.

실험은 4회분의 기출문제, 즉, 12개의 등급별 기출문제 중 읽기영역의 문제만 대상으로 하였다. 각 시험에 나오는 상하위 관계표현이 가능한 명사와 동사 614개 단어로 온톨로지를 만들고 단어별 빈도수를 체크하여 제안한 방법으로 난이도를 계산하였다. 그림 5는 온톨로지의 일부이다. 온톨로지에서 최대레벨은 8이었고, 각 단어간의 관계는 미 프린스턴 대학 워드넷 버전2.1[19]의 명사와 동사의 상하위 구조를 기반으로 구현하였다.

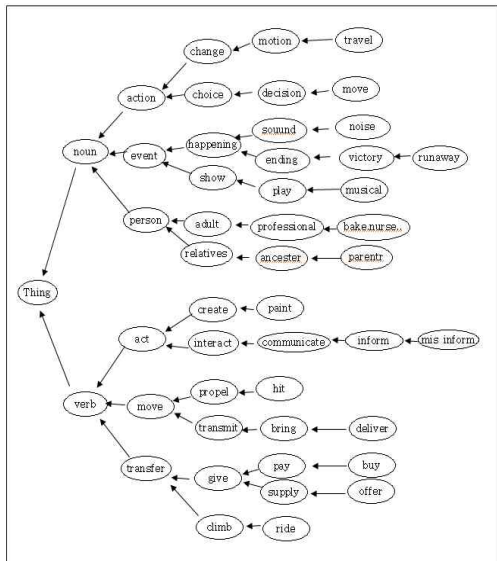


그림 5. JET 온톨로지(일부)
Fig. 5. A Part of the JET Ontology

기존방법과의 비교를 위해 학습순서를 초급->중급-> 고급 순으로 정하고 각 등급의 가중치를 임의로 지정한 지식맵을 이용하여 기존방법대로 난이도를 계산한 후 제안한 방법과 비교하였다.

3. 실험결과 및 분석

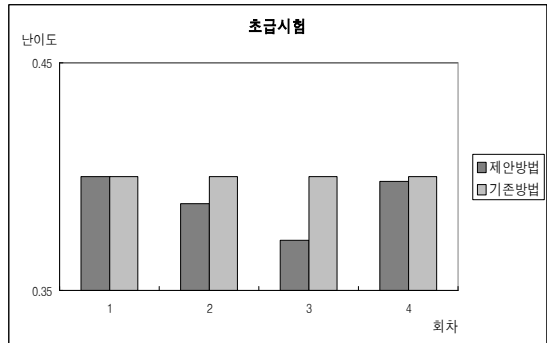
각 기출문제의 등급별 난이도 계산 실험결과는 그림 6과 같다. 그림 6의 (a)는 초급시험의 실험한 결과를 그래프로 비교한 것이다. 기존방법은 초급의 가중치를 0.4로 임의로 부여한 결과이고, 제안 방법은 각 회차별 빈도수의 합이 40, 39, 42, 43번이고, 심화도의 합은 128, 121, 125, 137이었고, 그 결

과 난이도는 각각 0.40, 0.388, 0.372, 0.398 이었다.

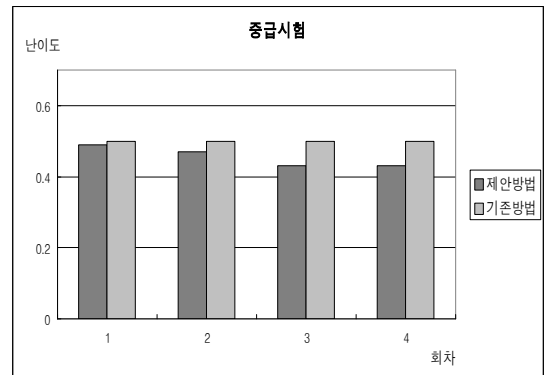
그림 6의 (b)는 중급시험의 실험한 결과를 그래프로 비교한 것이다. 기존방법은 중급의 가중치를 0.5로 임의로 부여한 결과이고, 제안한 방법은 각 회차별 빈도수의 합이 162, 179, 173, 159번이고, 심화도의 합은 644, 680, 596, 552 이었고, 그 결과 난이도는 각각 0.497, 0.475, 0.430, 0.434 이었다.

그림 6의 (c)는 고급시험의 실험한 결과를 그래프로 비교한 것이다. 기존방법은 고급의 가중치를 0.6으로 임의로 부여한 결과이고, 제안한 방법은 각 회차별 단어의 빈도수의 합이 259, 245, 227, 206번이고, 심화도의 합은 1,172, 1,029, 905, 857이었고, 그 결과 난이도는 각각 0.566, 0.525, 0.498, 0.520 이었다.

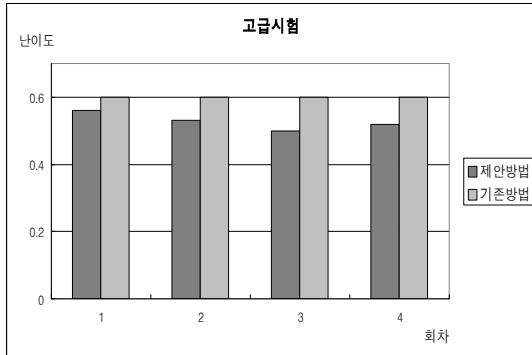
실험 결과, 제안한 방법은 각 회차별로 등급이 높을수록 난이도가 높았으며, 같은 등급이라도 회차별로 난이도의 차이를 명확히 계산할 수 있었다. 그러나 기존방법은 학습순서 및 가중치, 컨셉의 깊이가 같은 등급끼리 서로 같기 때문에 난이도를 모두 같게 계산하고 주관적인 판단이 필요하기 때문에 난이도를 명확히 계산할 수 없었다.



(a) 초급시험 비교



(b) 중급시험 비교



(c)고급시험 비교

그림 6. 난이도 계산 실험 결과

Fig. 6. The Result of the Experiments for Calculating Difficulties

표 4는 제안한 방법과 기존방법과의 차이점을 요약 비교한 것이다. 제안한 방법은 콘텐츠의 단어들로 난이도를 계산하기 때문에 기존방법처럼 지식맵 또는, 온톨로지를 이용하여 후행 학습을 사전에 결정할 필요가 없다. 또한, 기존방법은 동일한 주제 또는 등급의 난이도를 같게 계산하지만 제안한 방법은 이를 명확하게 구분하여 계산할 수 있다. 그리고, 기존방법은 콘텐츠의 주제의 개수, 학습순서 결정 등을 주관적인 판단에 의해 정하고 있는데, 동일한 대상을 놓고 여러명이 판단을 할 경우 서로 다른 결과가 나올 수 있기 때문에 주관적인 판단의 개입은 신뢰도와 타당성이 떨어지는 단점이 있다. 그러나 제안한 방법은 이러한 판단이 필요없기 때문에 객관적인 난이도 지표를 제공할 수 있는 장점이 있다.

표 4. 제안방법과 기존방법의 비교

Table 4. Comparison between the Proposed Method and the Previous Methods

구분	제안 방법	Kuo 방법	Guel 방법	Khan 방법
후행학습 사전 결정	필요없음	미리 결정	필요 없음	미리 결정
동일한 개념 (또는등급)의 콘텐츠 난이도	계산에 의해 객관적인 난이도 계산	같음	같음	같음
주관적 판단	없음	필요	필요	필요

V. 결 론

기존 학습콘텐츠 추천방법에서의 난이도는 후행학습 콘텐츠를 정하는데 있어 매우 중요한 변수이나 기존 온톨로지나 지식맵을 이용한 난이도 계산 방법들은 콘텐츠 제작자의 주관적 판단에 의하여 정하고 있어 신뢰도와 타당성이 떨어지고, 콘텐츠를 학습하는 대상의 수준에 따라 다르게 만든 동일한 주제의 콘텐츠의 난이도를 같게 계산하는 단점이 있었다.

본 논문에서는 이를 해결하기 위해 콘텐츠를 구성하는 단어들의 상하위 관계 및 심화도를 나타내는 온톨로지를 구성하고 단어들 간 온톨로지 경로상의 거리와 단어의 빈도수를 통하여 난이도를 계산하는 방법을 제안하였다. 제안한 방법에 의하여 콘텐츠 별로 0과 1사이의 난이도 값을 가지며 1에 가까울수록 난이도가 높은 콘텐츠라고 판단한다.

제안한 방법의 적정성과 기존 난이도 계산 방법들과 차이점을 비교하기 위하여 초등학교 영어시험인 JET시험의 등급별 기출문제를 대상으로 실험한 결과 기존방법에서의 난이도는 같은 등급의 문제지의 학습순서 및 가중치 등이 서로 같기 때문에 모두 동일한 값으로 계산되었다. 그러나, 제안한 방법은 문제 출제자의 주관적 판단없이 등급별 문제지의 난이도 및 같은 등급의 서로 다른 문제지의 난이도를 객관적 기준에 의해 명확히 계산함으로써 기존 단점을 해결하였다.

제안한 난이도 계산 방법은 문제지의 난이도 측정 뿐만 아니라 다양한 텍스트 기반의 콘텐츠 난이도 측정 및 비슷한 수준의 콘텐츠간의 변별력을 측정하는데 활용할 수 있다.

참고문헌

- [1] Hwa-Young Jeong, "A Design of The Tailored Learning Navigation based on The Learning Pattern of Learner," Journal of The Korea Society for Internet Information, Vol. 9, No. 6, pp. 109-115, Dec. 2008.
- [2] Yong Kim, Sung-Been Moon, "A Study on Hybrid Recommendation System Based on Usage frequency for Multimedia Contents," Journal of The Korea Society for Information management, Vol. 23, No. 3, pp. 91-125, Sept. 2006.
- [3] Je-Min Kim, Young-Tack Park, "Web Contents Recommendation based on Ontology," Proceedings

- of Korean Institute of Information Scientists and Engineers, Vol. 33, No. 2(B), pp. 294-299, Oct. 2006.
- [4] Sukyung Kim, Keehong Ahn, "Web Ontology Building Methodology for Semantic Web Application," Journal of The Korea Information Processing Society, Vol. 15-D, No. 1, pp. 47-60, Feb. 2008.
- [5] Kyungsil Moon, Suhyun Park, "Oriental Medical Ontology for Personalized Diagnostic Services," Journal of The Korea Society of Computer and Information, Vol. 15, No. 1, pp. 23-30, Jan. 2010.
- [6] Hyun-Chang Lee, "Building An Ontology Model for Warehouse Management using Sensor," Journal of The Korea Society of Computer and Information, Vol. 15, No. 1, pp. 201-206, Jan. 2010.
- [7] Yoon Soo Lee, "Ontology based Educational Systems using Discrete Probability Techniques," Journal of The Korea Society of Computer and Information, Vol. 12, No. 1, pp. 17-24, Mar 2007.
- [8] Burke R, "Hybrid Recommender System: Survey and Experiments," User Modeling and User Adapted Interaction, Vol. 12, No. 4, pp. 331-370, 2002.
- [9] Rita Kuo, Wei-Peng Lien, Maiga Chang and Jia-Sheng Heh, "Difficulty Analysis for Learners in Problem Solving Process based on the Knowledge Map," Proceedings of IEEE International Conference on Advanced Learning Technologies, pp.386-387, July 2003.
- [10] Korhan Gunel, Rifat Asliyan, "Determining Difficulty of Questions in Intelligent Tutoring Systems," Turkish Online Journal of Educational Technology Vol. 8, No. 3, pp. 14-21, July 2009.
- [11] Javed I. Khan, Manas Hardas, Yongbin Ma, " A study of Problem Difficulty Evaluation for Semantic Network Ontology Based Intelligent Courseware Sharing," Proceedings of IEEE International Conference on Web Intelligent, pp. 426-429, Sept. 2005.
- [12] Raymond J. Mooney, "Content-based Book Recommending Using Learning for Text Categorization," Proceedings of the fifth ACM conference on Digital libraries, pp.195-204, June 2000.
- [13] Gruber T., "A Translation Approach to Portable Ontologies," Knowledge Acquisition-Special issue: Current issues in knowledge modeling, Vol. 5, No. 2, pp. 199-220, June 1993.
- [14] Li-ping Shen, "Ontology-based Learning Content Recommendation," International Journal of Continuing Engineering Education and Life Long Learning, Vol. 15, No. 3, pp. 308-317, June 2005.
- [15] Devèzic V., "Understanding Ontological Engineering," Communications of the ACM, Vol. 45, No. 4, pp. 136-144, Apr. 2002.
- [16] Fensel D, Fvan Harmelen, LHrocks, DLMcGuinness, and P.F.Patel-Schneider, "OIL:An Ontology Infrastructure for the Semantic Web," IEEE Intelligent Systems, Vol. 16, No. 2, pp. 38-44, March 2001.
- [17] Binh An Truong, Young-Koo Lee, and Sung-Young Lee, "Modeling and Reasoning about Uncertainty in Context-Aware System," Proceedings of IEEE International Conference on e-Business Engineering, pp. 102-109, Oct. 2005.
- [18] JET, <http://exam.ybmsisa.com>
- [19] Wordnet, <http://wordnet.princeton.edu>

저 자 소 개



박 재 욱
 1999 : 서울산업대학교
 산업공학과 공학사
 2006 : 서울시립대학교 경영대학원
 이미지비즈니스학과 경영학석사
 2010 : 동국대학교
 컴퓨터공학과 박사수료
 현 재 : 새마을금고임직원상조복지회
 영업개발팀 과장
 관심분야 : 온톨로지, 데이터베이스,
 e-비즈니스
 E-mail : ssebbok@dongguk.edu



박 미 화
 1997 : 동국대학교
 컴퓨터공학과 공학사
 1999 : 동국대학교
 컴퓨터공학과 공학석사
 2008 : 동국대학교
 컴퓨터공학과 공학박사
 현 재 : 동국대학교
 컴퓨터공학과 전문연구원
 관심분야 : DB, 정보검색, 온톨로지
 E-mail : meehwap@dongguk.edu



이 웅 규
 1986 : 동국대학교
 전자계산학과 공학사
 1988 : 한국과학기술원
 전산학과 공학석사
 1996 : Syracuse University
 전산학 박사
 1978년~83년 : 행정직 국가공무원
 1988년~93년 : 국방정보체계연구소
 선임연구원
 1996년~97년 : 한국통신 선임연구원
 2002년~03년 : 콜로라도대학교
 컴퓨터학과 방문교수
 1997년~현재 : 동국대학교 컴퓨터
 공학과 교수
 관심분야 : 데이터베이스, 정보검색,
 웹사이언스, e-비즈니스 시스템
 E-mail : yklee@dongguk.edu