

일반화된 벡터 공간 모델을 적용한 주관식 문제 채점 보조 시스템

오정석, 추승우, 김유섭, 이재영
한림대학교 컴퓨터 공학과
e-mail : bil78@hallym.ac.kr

Subjective Tests Sub-System Applied with Generalized Vector Space Model

Jung-Seok Oh, Seung-Woo Chu, Yu-Seop Kim, Jae-Young Lee
Dept. of Computer Engineering, Hallym University

요 약

기존의 주관식 문제 채점 보조 시스템은 자연어 처리의 어려움으로 인해 채점의 자동화가 어려워 전자우편 등을 통하여 채점자에게 채점 의뢰를 하는 수준이었다. 본 논문에서는 이러한 문제점을 해결하기 위하여 문제 공간을 벡터 공간으로 정의하고 벡터를 구성하는 각 자질간의 상관관계를 고려한 방법을 적용하였다. 먼저 학습자가 답안을 작성할 때 동의어 사용을 한다는 가정하에 출제자가 여러 개의 모범 답안을 작성하고 이들 답안을 말뭉치에 첨가하여 구성된 다음 형태소 분석기를 통하여 색인을 추출한다. 그리고 학습자가 작성한 답안 역시 색인을 추출한 다음, 이들 색인들을 각 자질로 정의한 벡터를 구성한다. 이렇게 구성된 벡터들을 이용하여 답안들간 유사도 측정을 하고, 유사도 범위에 따라 답안을 자동으로 정답과 오답으로 분류하려는 시스템을 제안한다. 170 문항의 주관식 문제를 제안된 방법으로 실험하여, 기존 모델에 비해 성능과 신뢰성 향상을 이룰 수 있었다.

1. 서론

인터넷이 대중화가 됨에 따라 원격 교육은 인터넷을 통하여 실시간으로 시험을 치르는 영역으로까지 확대되고 있다[1]. 하지만 대부분의 학습자 수학 능력 평가는 작성된 답안을 자연어 처리를 통하여 분석하는데 많은 어려움이 있기 때문에 주관식 문제보다는 객관식 문제에 대한 평가만 이루어졌다[2]. 이러한 문제를 해결하기 위해 주관식 채점 보조 시스템이 제안되었다[4].

주관식 문제 채점 보조 시스템은 학습자가 시험에 응시하여 답안을 제출하면 출제자에게 채점 의뢰를 통지하여 채점하게 한 후, 채점 결과를 학습자에게 알려주는 것이다[4]. 하지만 이 시스템은 채점을 하기 위해 출제자에게 채점 의뢰를 통지해야 하고, 채점의 결과와 시간이 출제자에 의해 좌우 된다는 문제점이 있다[2, 3].

본 논문에서는 이러한 문제들을 해결하기 위하여

사람의 도움이 필요없이 자동으로 답안 채점이 가능한 시스템을 제안한다. 또한 자동 채점의 문제 공간을 벡터 공간으로 매핑하여 이들 벡터간의 문제로 단순화한 정보 검색 방법론을 사용하였다. 본 연구에서는 정보 검색 방법론으로 일반화된 벡터 공간 모델 (Generalized Vector Space Model: GVSM)을 사용하였다.

본 논문에서 제안된 시스템은 학습자의 답안과 학습자의 동의어 사용을 고려하여 출제자가 여러 개의 모범 답안을 작성하고, 이들 답안으로 구성된 말뭉치에서 형태소 분석기를 이용하여 각각 색인어를 추출한다. 이 추출된 말뭉치 색인을 출제자의 답안 용어들과 학습자의 답안 용어들간 상관관계를 측정하기 위한 벡터로 구성한다. 이렇게 구성된 학습자의 답안에서 추출된 색인 벡터와 출제자의 답안들에서 추출된 색인 벡터로 구분하여 이들 답안들간의 유사도 측정을 하고, 유사도 범위에 따라서 답안을 정답과 오답

으로 분류하려는 시스템을 제안한다. 그리고 답안들 간 유사도 측정 실험을 통해 정확도 향상을 확인한다.

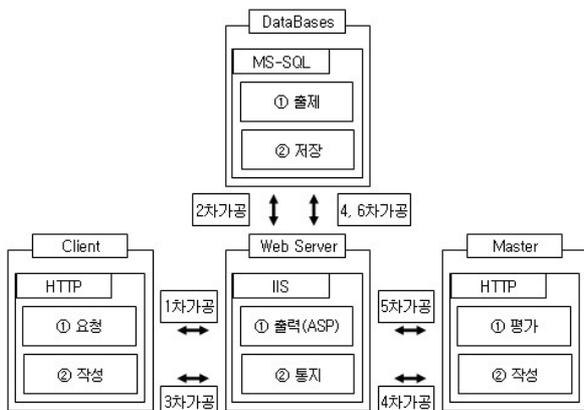
2 장에서는 기존의 주관식 문제 채점 보조 시스템의 문제점에 대해 설명한다. 3 장에서는 자동 채점을 위한 GVSM 에 대한 설명과 이를 적용시킨 시스템의 구성을 상세하게 설명한다. 4 장에서는 제안된 방법으로 실험한 결과를 살펴보고, 5 장에서는 결론을 밝힌다.

2. 주관식 채점 보조 시스템

주관식 문제 채점 보조 시스템은 학습자, 웹 서버와 데이터베이스 서버, 출제자의 역할이 분담되어 문제를 요청, 출제, 저장하고 채점한다.

다음의 [그림 1]은 기존의 주관식 문제 채점 보조 시스템의 개략적인 구성을 보여준다.

[그림 1]의 구성을 살펴보면 학습자가 웹 서버에게 서비스를 요청을 보내면, 데이터베이스 서버는 문제를 출제한 후 웹 서버에게 전달해 준다. 웹 서버는 이를 전송 받아 학습자에게 디스플레이 해준다. 학습자가 문제의 답안을 작성하여 웹 서버에게 전달하면, 웹 서버는 이를 데이터베이스 서버에 저장하는 동시에 전자우편이나 전화로 출제자에게 연락하여 채점 의뢰를 한다. 채점 의뢰를 받은 출제자는 웹 서버에 접속하여 학습자가 작성한 답안과 채점 기준을 가지고서 이를 평가한 후, 채점한 점수와 변경된 채점 기준, 틀린 문제의 해설 등을 웹 서버에게 전달한다. 그러면 웹 서버는 이를 데이터베이스 서버에 저장하고 학습자에게 디스플레이 해준다[4].



[그림 1] 주관식 문제 채점 보조 시스템의 개략적인 구성도

위의 주관식 문제 채점 보조 시스템을 이용하여 주관식 문제를 채점하면 학습자가 답안을 제출할 때마다 웹 서버는 출제자에게 채점 의뢰를 통지해야 한다. 또한 채점 의뢰를 받은 출제자는 답안을 채점할 때마다 채점 기준이 바뀔 수 있어 채점의 객관성을 유지하기 힘들다. 그리고 틀린 문제에 대한 해설도 직접 입력해야 하는 등 채점에 관한 출제자에 대한 의존성이 높아 출제자의 작업 능률에 따라 학습자가 채점 결과를 받아 볼 수 있는 시간이 일정하지 않게 된다. 이에 대한 문제점을 보완하기 위해 본 논문에서는 채점에 대한 출제자의 역할을 웹 서버가 담당하도록 하

는 자동 주관식 문제 채점 보조 시스템을 제안한다.

3. 일반화된 벡터 공간 모델을 적용한 주관식 문제 채점 보조 시스템

이 장에서는 문서 자동 분류 알고리즘 중 하나인 GVSM(Generalized Vector Space Model) 기법과 이를 적용한 시스템을 소개한다. 여기서 제안된 시스템은 답안 용어들 간의 상관관계를 고려한 유사도 측정을 통하여 답안들을 자동으로 정답과 오답으로 분류함으로써, 위에서 언급된 주관식 문제 채점 보조 시스템의 문제점을 해결할 수 있다.

3.1 일반화된 벡터 공간 모델

정보 검색을 위한 벡터공간 모델에서 두 문서 벡터 d_1 와 d_2 의 유사도는 일반적으로 다음과 같이 표현할 수 있다[6, 8].

$$sim(d_1, d_2) = (P^T d_1)^T \bullet (P^T d_2) = d_1^T P P^T d_2 \quad (1)$$

행렬 P 는 기본 벡터공간에서의 문서를 특정한 다른 자질공간으로 매핑시키기 위한 변환행렬로서 $P = I_M$ (M 은 문서 집합 내 어휘 개수이고 I_M 은 $M \times M$ 항등행렬) 일 때는 원래 용어 벡터 공간에서의 두 문서간 유사도를 의미한다[8]. 하지만, 이 경우에는 각 용어들에 의해 표현되는 축이 벡터공간상에서 직교를 이루게 되므로(orthogonal) 단어들의 상관관계를 고려할 수 없다는 점에서 문제점이 지적되곤 한다[6,8].

Wong[7]은 “generalized vector space model” 또는 GVSM 이라 불리는 벡터공간 모델의 대안을 제안하였다[8,10]. 이는 문서 집합 내에 나타나는 용어들 간의 공기정보를 이용하는 방식이다[8]. GVSM 는 문서들(서로 독립하여 나타나는 것을 가정 한 것)을 사용함으로써, 용어들을 벡터공간의 기초로 사용하는 것을 피하게 하였다[10].

언어 자료(corpus)는 용어-문서 행렬 A 로 나타내며, 행들은 문서들의 용어 가중치 벡터들이다[10]. GVSM 에서, 행렬 A 는 이차원 공간으로 보이며 용어들은 문서 가중치의 벡터들이다[10]. 행렬 A 를 구성하기 위해서는 우선, 문서들 내의 모든 용어들을 추출하여 아래와 같은 벡터를 구성한다[7].

$$[k_1, k_2, \dots, k_t], \quad t = 1, 2, \dots, n$$

문서들을 용어들의 출현 패턴에 따라 0 과 1 로 표시한 minterm $m_{2'}$ 을 구성하고, 각각의 minterm m_i 로 쌍방향 직교 벡터 \vec{m}_i 를 구성한다[7]. 여기서 m_1 는 용어들을 어느 것도 포함하지 않는 문서를 말하며, m_2 는 용어 k_1 만을 포함하고 있는 문서를 말한다[7].

$$m_1 = (0, 0, \dots, 0) \quad m_2 = (1, 0, \dots, 0) \quad \dots \quad m_{2'} = (1, 1, \dots, 1)$$

$$\vec{m}_1 = (1, 0, \dots, 0, 0) \quad \vec{m}_2 = (0, 1, \dots, 0, 0) \quad \vec{m}_{2'} = (0, 0, \dots, 0, 1)$$

용어들의 출현 패턴이 minterm m_r 과 일치하는 용

어 가중치 $w_{i,j}$ 을 합산한 상관계수 c_{ir} 는 식(2)처럼 계산된다[7].

$$c_{ir} = \sum_{d_j | g_i(\bar{d}_j) = g_i(m_r) \text{ for all } i} w_{i,j} \quad : w_{i,j} \rightarrow [k_i, d_j] \text{의 조합} \quad (2)$$

용어 k_i 가 포함되어 있는 모든 minterm m_r 의 벡터를 합산하여 정규화한다[7].

$$\bar{k}_i = \frac{\sum_{\forall r, g_i(m_r)=1} c_{ir} \bar{m}_r}{\sqrt{\sum_{\forall r, g_i(m_r)=1} c_{ir}^2}} \quad (3)$$

즉, \bar{k}_i 들로 행렬 A 가 구성된다. 또한, 행렬 A 는 질의 q 를 $q' = A^T q$ 처럼 이차원 공간으로 변환시킨다[10]. 문서도 마찬가지로 변환되며, 표준 코사인 측정을 사용하여 변환된 질의와 비교 한다[10].

$$sim(d, q) = \cos(A^T d, A^T q) = \frac{d^T A A^T q}{\|A^T d\| \|A^T q\|} \quad (4)$$

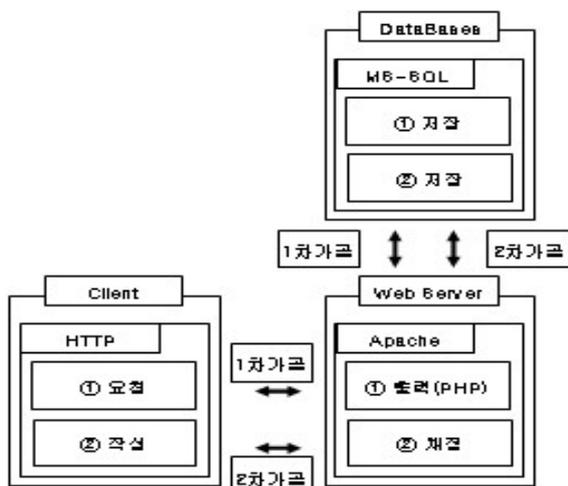
변환 $d' = A^T d$ 는, 문서들이 행렬 A 로 길이-정규화된다면, d 와 모든 문서들 A_i 사이의 유사점의 벡터가 산출된다[10].

$$A^T d = [A_1^T d, A_2^T d, \dots, A_n^T d] \quad (5)$$

본 논문에서는 위의 식들을 이용하여 행렬 A 을 만들고, 식(4)을 이용하여 학습자의 답안과 채점자의 답안의 유사도 측정을 하려고 한다.

3.2 시스템 구성

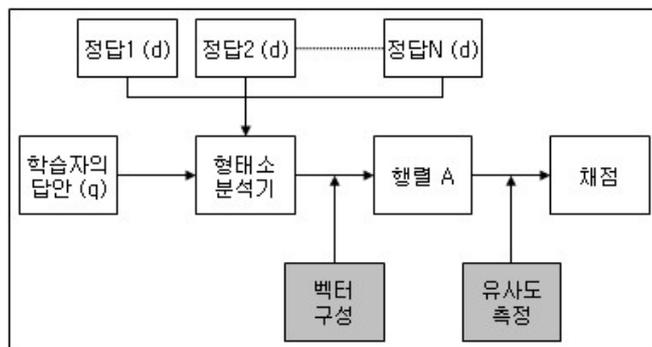
다음의 [그림 2]는 개선된 주관식 채점 보조 시스템의 개략적인 구성을 보여준다. [그림 2]의 구성을 살펴보면 기존의 주관식 채점 보조 시스템에서 출제자의 채점 역할이 웹 서버에게 이관되었다.



[그림 2] 개선된 주관식 채점 보조 시스템

다음의 [그림 3]은 웹 서버에서의 주관식 채점의 흐름도이다. [그림 3]을 살펴보면 채점을 하기 위해서는 용어들 간의 상관관계를 측정하기 위한 행렬 A 가

필요하다. 행렬 A 를 만들기 위해 출제자의 정답 문서와 말뭉치를 형태소 분석 프로그램인 ‘HAM’ [5]을 이용하여 용어들을 추출하고, 이를 용어 벡터로 만든다. 출제자의 각 정답 문서에서 추출한 용어들과 용어 벡터의 원소들을 비교하여 minterm m_i 를 구성하고, 이를 이용하여 쌍방향 직교 벡터 \bar{m}_i 도 구성한다. 구성된 각 벡터들과 출제자의 각 정답 문서로 용어 가중치 $w_{i,j}$ 를 계산하고, 식(2)에 대입하여 상관계수 c_{ir} 를 계산한다. 상관계수 c_{ir} 를 식(3)에 대입하여 문서내의 용어 가중치 \bar{k}_i 를 계산하고, 이를 행렬 A 로 만들어 각 벡터들과 함께 데이터베이스에 저장한다. 학습자가 답안을 제출하면 형태소 분석기를 이용하여 용어를 추출한다. 이 추출된 용어와 출제자의 정답 문서에서 추출된 용어들을 행렬 A 와 함께 식 (4)에 대입하여 유사도를 측정하도록 한다.



[그림 3] 주관식 채점의 흐름도

4. 실험 및 성능 평가

4.1 실험

본 실험을 위하여 형태소 분석기를 이용하여 신문기사들을 모아서 만든 말뭉치와 학습자의 동의어 사용을 고려한 출제자의 답안들을 첨가하여 247932 개의 용어들을 추출하였고, 170 문항의 ‘고사성어’ [11] 주관식 문제를 시스템에 적용하였다.

4.2 성능 평가

[표 1]은 5 명의 학생을 대상으로 출제자가 채점한 결과와 시스템이 채점한 결과를 비교한 것이다.

	학생 1	학생 2	학생 3	학생 4	학생 5
틀린 정답(개)	0	0	0	1	0
틀린 오답(개)	0	3	0	6	4
총합(개)	0	3	0	7	4
비율(%)	0	1.76	0	4.12	2.35

[표 1] 채점 결과 비교

다음의 [표 1]에서 “틀린 정답” 은 출제자가 오답으로 채점한 것을 시스템이 정답으로 채점한 것의 개수이다. “틀린 오답” 은 출제자가 정답으로 채점한 것을 시스템이 오답으로 채점한 것의 개수이다. [표 1]의 결과를 분석하면, 평균 오차는 1.65%이고 최고

오차는 4.12%이며 학생 1 과 학생 3 처럼 정확히 일치하는 채점 결과도 보여주고 있다.

위의 오차가 발생하게 된 원인을 살펴보면 다음과 같다.

- 경우 1 : 출제자가 채점시 오타나 띄어쓰기등 문법에 맞지 않는 경우를 정답으로 간주한 경우
- 경우 2 : 출제자 답안 작성시 학습자의 유의어 사용을 고려하지 않은 경우
- 경우 3 : 시스템이 채점시 핵심어는 일치하지 않지만 문맥이 일치하여 정답으로 간주한 경우
- 경우 4 : 시스템이 형태소 분석기에서 말뭉치로 용어를 추출시, 정답 문서에서 사용하는 용어를 추출하지 못하여 오답으로 간주한 경우

[표 2]는 위의 원인들을 표로 나타낸 것이다.

	경우 1	경우 2	경우 3	경우 4	총합(개)
학생 1(개)	0	0	0	0	0
학생 2(개)	1	1	0	1	3
학생 3(개)	0	0	0	0	0
학생 4(개)	4	1	1	1	7
학생 5(개)	3	1	0	0	4
총합(개)	8	3	1	2	14
비율(%)	0.57	0.21	0.07	0.14	1

[표 2] 오차 원인 비교

오차 원인들은 크게 출제자와 시스템에 의한 오류로 나눌 수 있다. [표 2]를 보면 출제자에 의한 오류는 0.78%이며 시스템에 의한 오류는 0.21%이다. 하지만 출제자에 의한 오류는 자동 채점을 통하여 출제자의 오타나 띄어쓰기로 인한 채점의 오류를 방지 할 수 있고, 출제자 개인의 답안뿐만 아니라 출제자가 답안으로 인정한 답안과 의미상으로 유사한 답안도 정답으로 인식하기 때문에 오히려 채점의 객관성을 유지할 수 있다는 점에서 매우 긍정적인 결과라고 볼 수 있다. 물론 보다 정확한 채점을 위해서는 시스템의 오류들 즉, 용어들의 누락을 방지할 수 있는 형태소 분석기와 핵심 키워드의 가중치 방식이 개발되어야 할 것으로 사료된다.

5. 결론

본 논문에서 제안한 시스템은 주관식 문제 채점 보조 시스템이 가지고 있는 문제점을 개선한 것이다. 벡터 모델을 활용함으로써 기존의 출제자가 담당하였던 복잡한 채점 시스템을 웹 서버로 이관하여 채점 시스템을 자동화시켜 학습자가 채점 결과를 바로 확인할 수 있게 되었다. 또한 출제자의 주관성이 개입될 소지가 있는 채점 기준을 입력된 답안들로만 제한하여 채점의 객관성을 확보할 수 있게 되었다.

앞으로의 과제로는 첫째 보다 정확한 채점을 위해 대규모의 말뭉치를 구축하여야 할 필요가 있으며 다양한 교과목에 관한 말뭉치를 구축해야 한다. 둘째, 문장에서 용어들이 누락되지 않게 추출되기 위해서는

보다 정교한 형태소 분석이 필요하다. 마지막으로 본 논문에서는 일반화된 벡터 공간 모델을 활용하였는데 이러한 모델을 확장하여 보다 강력한 의미 커널 모델을 활용한 시스템의 구축도 필요하다.

참고문헌

- [1] 김경아, “웹 기반 자동 문제 출제 시스템”, 동국대학교 교육대학원 석사학위논문, 2000
- [2] 정동경, “벡터유사도와 시소러스를 이용한 주관식 답안의 채점방법”, 동국대학교 교육대학원 석사학위논문, 2001
- [3] 박희정, 강원석, “유의어 사전을 이용한 주관식 문제 채점 시스템 설계 및 구현”, 컴퓨터교육학회 논문지, 한국컴퓨터교육학회, 2003
- [4] 방훈, 강태호, 김원진, 원대희, 이재영, “웹상에서 주관식 문제 채점 보조 시스템”, 춘계학술대회, 한국정보과학회, 2001.04
- [5] 강승식, 범용 형태소 분석기 “HAM Ver 6.0.0”, 국민대학교 자연언어 정보 검색 연구실, <http://nlp.kookmin.ac.kr>
- [6] Cristianini, N., Shawe-Taylor, J., and Lodhi, H., Latent semantic kernels, Journal of Intelligent Information System, vol.18, no.2/3, pp.127-152, 2002.
- [7] Wong, S. K. M., Ziarko, W., and Wong, P. C. N., Generalized vector space model in information retrieval, ACM SIGIR Conference on Research and Development in Information Retrieval, pp.18-25, 1985
- [8] 장정호, 김유섭, 장병탁, “헬름홀츠머신 학습 기반의 의미 커널을 이용한 문서 유사도 측정”, 추계학술대회, 한국정보과학회, 2003.04
- [9] 신동호, “Latent Semantic Analysis 를 이용한 내용기반 정보검색 시스템”, 서울대학교 대학원 협동과정 인지과학전공 석사 학위 논문, 2000.02
- [10] I. M. Soboroff and C. Nicholas. Collaborative Filtering and the Generalized Vector Space Model. In Proceedings of the 23rd Annual International Conference on Research and Development in Information Retrieval (SIGIR), 2000.
- [11] 김원중, “고사성어 백과사전”, 을유문화사, 2003.05