

QA 시스템에서 질의 패턴을 이용한 질의 확장 기법

(A Query Expansion Technique using
Query Patterns in QA systems)

김혜정*, 부기동**
(Hea-Jung Kim, Ki-Dong Bu)

요약 QA(질의응답) 시스템은 질의에서 요구하는 정답 유형 및 질의에 사용된 용어를 적용하여 보다 정확한 답을 추출하고자 한다. 그러나 질의에 사용된 용어들이 문서에 그대로 사용되지 않고 같은 의미의 다른 어휘로 출현하기도 하며, 혹은 다른 문법적 정보를 가진 카테고리로 등장하여 정답 추출에 어려움이 따른다. 따라서 본 논문은 질의에서 사용된 의미적으로 더 가까운 단어들로 구성되는 심층적 질의 카테고리의 질의 패턴을 이용한 질의 확장 방법론을 제안한다. 제안한 방법은 질의 유형에 따른 개념 리스트를 우선 구축하고, 학습 알고리즘에 의해 각 질의 카테고리에 대한 개념 리스트를 구축한다. 실험의 결과로서 제안한 방법의 성능이 향상되었음을 입증하였다.

핵심주제어 : QA 시스템, 질의 패턴, 질의 확장

Abstract When confronted with a query, question answering systems endeavor to extract the most exact answers possible by determining the answer type that fits with the key terms used in the query. However, the efficacy of such systems is limited by the fact that the terms used in a query may be in a syntactic form different to that of the same words in a document. In this paper, we present an efficient semantic query expansion methodology based on query patterns in a question category concept list comprised of terms that are semantically close to terms used in a query. The proposed system first constructs a concept list for each question type and then builds the concept list for each question category using a learning algorithm. The results of the present experiments suggest the promise of the proposed method.

Key Words : Question Answering System, Query Pattern, query expansion

1. 서론

자연어는 사용자가 필요한 정보를 가장 잘 표현할 수 있는 방법이다. 현대의 자연어 처리 시스템인 정보 검색 시스템에서는 주어진 사용자의 질

의에 대해 가장 관련이 있을 정답이 포함된 문서들이 추출되어지지만, 사용자 의도를 잘 파악하고 좀 더 주어진 질의에 명확한 대답을 줄 수 있는 시스템의 필요성이 대두되었다. 따라서 이러한 요구를 만족시키기 위하여 질의와 밀접한 연관성을 갖는 단어, 문단, 절을 추출하고 순위화한 후 가장 연관성이 있는 정답을 추출하는 QA(Question Answering) 시스템에 관한 연구가 활발히 진행되

* 계명대학교 교양학부 초빙전임강사

** 경일대학교 컴퓨터공학부 교수

었다[1]. 그러나 질의와 정답에 사용된 어휘들이 문서의 정답 문장에 그대로 사용되지 않고 같은 의미의 다른 어휘로 출현하기도 하며, 혹은 다른 문법적 정보를 가짐으로써 정확한 정답 추출에 어려움이 따른다.

아래 예제 질의와 정답 문장을 살펴보자.

■ 질의문장 : Who is the inventor a paper?

■ 정답문장 : A devised paper from China.

예제 질의를 QA 시스템으로 처리를 할 경우 먼저 사용자 질문을 분석하여 질의를 의미 분류 체계인 카테고리별로 분류하여야 한다. 기존의 QA 연구들은 질의 확장을 위하여 주로 시소러스의 동의어 및 상하위어를 사용하는 방법론을 제안하였다[2]. 그러나 이러한 기존 연구들은 질의어의 명사 "inventor"를 동사 "devise"로 문법 카테고리를 넘어 확장하는데 있어 어려움이 있을 뿐만 아니라, 상하위어로의 확장에서 확장 범위에 대한 경계가 불확실하며, 질의어에 동사 "invent"가 사용되었을 경우, 시소러스의 사용자가 같은 개념을 표현하기 위해 주로 사용하는 어휘인 동사 "make"로의 확장은 거의 불가능하다.

만약 질의에서 요구하는 정답유형과 개념적으로 유사한 문장을 대상으로 정답추출을 수행할 수 있다면 보다 정확한 정답을 추출 할 수 있을 것이다. 기존의 QA 시스템도 이러한 가정 하에 질의에서 요구하는 정보 유형, 혹은 질의 유형을 의미 체계별로 나누어 정답 추출을 수행하고 있으나 기존의 QA 시스템에서 정의하여 사용하고 있는 질의 유형은 "person, Location, Organization" 등 상대적으로 상위의 개념으로 체계화 되어있다. 질의 유형이 상위 개념일수록 해당 질의 유형을 표현하기 위해 사용되는 어휘의 범위도 넓게 분포된다. 그러나 질의 유형이 하위의 개념 즉, 보다 깊게 표현될수록 해당 질의 유형을 표현하기 위해 사용되는 어휘 집합 또한 좁게 분포될 수 있다. 하나의 깊은 질의 유형을 표현하기 위해 동원되는 어휘 집합은 해당 개념을 표현하는 다양한 어휘들로 각각 상위어, 동의어, 하위어 관계를 이룰 수 있으며 그러한 의미 관계를 제외한 다른 연관성을 가지는 어휘일 수도 있다.

본 논문에서는 이렇게 같은 질의 개념을 표현하

기 위해 동원되는 어휘 집합을 질의 확장을 위해 응용하고자 하며, 질의 패턴을 학습하여 얻어진 심층적 질의 카테고리의 개념 커버리지에 기반한 질의 확장 기법을 제안하고자 한다.

2. 관련 연구

QA 시스템에서는 정답 추출과 성능향상을 위해 정확한 정답 유형의 분류와 불일치되는 단어 문제를 해결하기 위한 질의 확장이 필요하다. 또한, 질의 확장을 위해서는 시소러스가 많이 이용되는데 Voorhees[3]는 워드넷(WordNet) [4]을 사용하여 질의 내의 모든 어휘들에 대해 동의어, 반의어, 상위어 등을 확장해서 비교적 질의의 길이가 짧은 경우에 대한 성능 향상을 보였다. Moldovan과 Mihalcea[5]는 워드넷 용어 풀이에 포함된 정보가 문맥 해석에 도움이 된다는 것에 착안하여 워드넷에서 원 단어 정의문에 기술된 단어들, 상위어, 하위어, 유사어 등을 가중치를 주어 연결하여 관련된 단어들의 사슬을 형성하였다. 이러한 단어 사슬을 이용하여 토픽이 관련된 문헌을 찾을 수 있음을 보였다. Mandala et al.[6]는 여러 개의 이질적(heterogeneous) 시소러스를 사용하여 용어들의 가중치를 결합 평균값을 계산함으로써 가장 높은 확률을 가진 용어에 대해 확장을 하였다.

Prager, et al[7][8]는 질의 확장을 위해 워드넷(WordNet) 시소러스를 사용하였는데 태깅된 코퍼스에서 주어진 질의 어휘의 태깅된 유사집합을 찾을 때 까지 의미간의 상하위 관계를 이용하여 질의를 확장하였으며 정답 추출을 위해 미리 정답 후보들을 찾아 색인하는 방법(Predictive Annotation)을 적용하였으나, 간혹 어떤 용어에 대해서는 워드넷의 상하위 관계를 이용하여도 상위어를 찾을 수 없을 경우 다른 외부적인 시소러스의 사용이 필요하다고 하였다. 그러나 이러한 기존 연구들은 문법 카테고리를 넘어선 질의 확장에는 어려움이 있을 뿐만 아니라, 시소러스의 동의어나 상하위어 정보에서는 찾을 수 없지만 사용자가 같은 개념을 표현하기 위해 주로 사용하는 어휘로의 질의 확장은 거의 불가능하다. 따라서 본 논문은 질의의 패턴에 검색하고자 하는 정보의 유형을 표

현하기 위해 동원되는 어휘 집합이 있다고 간주하고, 이러한 정보 유형을 표현하는 어휘 집합을 질의 확장에 적용하여 정답 추출의 효율을 높이고자 한다.

3. 질의 패턴을 이용한 질의 확장 기법의 제안

3.1 심층 질의 카테고리 학습

질의 패턴의 학습과 개념 분류를 위해서는 베이시안(Bayesian) 분류 알고리즘을 사용한다. 먼저 학습할 질의에 대한 패턴을 구한 다음, 패턴 내 모든 단어들에 대해 각각의 단어가 카테고리에 속할 확률을 구한 후 학습한다. 학습 방법은 다음과 같다: 어휘 w 는 주제를 분류하고자 하는 패턴에서의 키워드 단어이고 주변 각 단어를 v_j , 어휘 w 가 가질 수 있는 모든 질의 카테고리를 S_k 라고 하자. 이때, 단어 w 가 가지는 모든 질의 카테고리 S_k 에 대하여 v_j 가 나타날 확률($P(v_j | S_k)$)은 어휘 w 가 나타났을 때 v_j 와 S_k 가 함께 발생한 빈도를 v_j 가 발생한 빈도로 나눈 것이다. 또한, 어떤 어휘 w 가 주제 카테고리 S_k 가 될 확률 ($P(S_k)$)은 주제 카테고리 S_k 가 발생한 빈도를 어휘 w 가 발생한 빈도로 나눈 값이다.

$$P(v_j | S_k) = \frac{C(v_j, S_k)}{C(v_j)}, \quad (1)$$

$$P(S_k) = \frac{C(S_k)}{C(w)}$$

또한, 새로운 질의가 주어졌을 경우 학습된 지식을 기반으로 질의의 카테고리 S^* 를 분류하는 방법은 다음과 같다.

$$\text{Decide } S^* \text{ if } S^* = \text{argmax}_{S_k} [\log P(S_k) + \sum_{v_j \in C} \log P(v_j | S_k)] \quad (2)$$

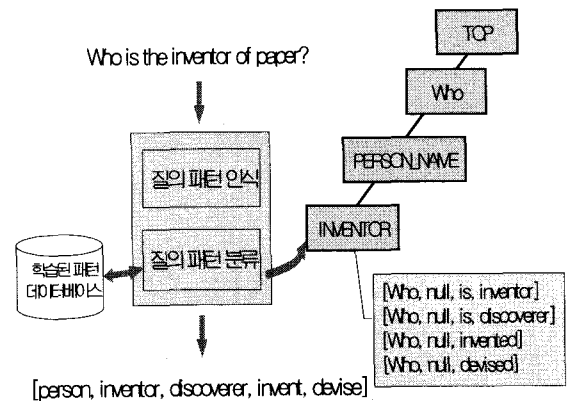
즉, 새로운 질의가 주어졌을 경우 해당 질의 패턴이 어떤 카테고리인지의 확률 S^* 은 학습된 지식(모든 단어에 대한 $P(v_j | S_k)$) 및 $P(S_k)$ 을 적용하여 어떤 단어에 대해 특정 카테고리일 확률

($\log P(S_k)$)과 카테고리 S_k 일때 패턴 내 주변단어 v_j 가 나타날 확률($\log P(v_j | S_k)$)을 더한 값 가운데 가장 큰 값을 해당 단어에 대한 카테고리로 정한다.

3.2 의미적 질의 확장

제안된 시스템은 먼저 입력 질의의 패턴을 추출하고, 학습 지식을 기반으로 추출 패턴의 질의 카테고리를 분류한다. 분류된 질의 카테고리로부터 해당 개념을 공유하는 확장 어휘들을 획득한다. 질의 확장 과정은 <그림 1>에서 보여주는 바와 같다.

예를 들어, 질의 "Who is an inventor of a paper?"와 같은 질의가 입력되면 질문의 개념을 파악하기 위해 먼저 질의 패턴을 추출한다. 그런 후 학습된 패턴 지식을 기초로 추출된 질의 패턴의 개념 카테고리를 분류한다. 입력 질의의 개념은 심층 질의 카테고리 중 "inventor"에 해당하고, "inventor" 카테고리로부터 해당 개념을 공유하는 확장 어휘들을 사용하여 질의를 확장한다. 따라서 입력 질의는 "inventor, discoverer, invent, devise"등으로 확장될 수 있다.



<그림 1> 개념별 질의 패턴 할당 및 확장

3.3 질의 패턴에 대한 학습 정확률

질의 패턴에 대한 학습을 위해 본 논문에서는 TREC-9[9]의 201~893번 질문 692개 중에서 who 질의 117개를 사용하여 질의의 개념을 표현 할 수

있는 세분화된 심층 질의 카테고리 질의를 나누고 각 카테고리의 개념을 커버리지 할 수 있도록 동원되는 단어들의 집합인 카테고리별 개념 커버리지 리스트를 구축하였다. 구성된 개념 커버리지 리스트는 나이브 베이시안 분류 알고리즘을 이용하여 각 질의 카테고리별로 학습하였다. 베이시안 접근 방식을 적용하는 방법은 기본적으로 주어진 문맥 상황에서 중의성을 가진 단어의 각각의 의미의 확률을 계산하여 그 중 가장 확률이 높은 의미를 선택하는 것이다.

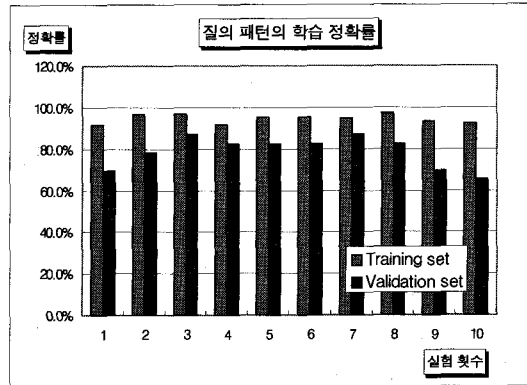
정확한 학습 성능을 검증하기 위해서는 패턴에 대해 전체를 학습한 경우와 10-fold 교차 검증(cross validation)을 수행하여 학습한 경우에 대해 정확률을 기준으로 비교하였다. 전체를 학습한 패턴에 대한 정확도는 94.4%, 학습되지 않은 패턴에 대한 정확도 즉, 주어진 패턴의 90%를 학습하고, 10%는 성능 검증을 위한 validation set으로 테스트하는 방법을 10번 반복하여 평균값을 계산한 결과의 정확도는 78.7%를 얻을 수 있었다. 이때 각 패턴에 대한 정확도는 <표 1>과 같고, 이를 그래프로 표현하면 <그림 2>와 같다.

학습에 있어서 training set과 Validation set이 정확도의 차이가 나는 것은 패턴으로 찾아낸 Who-term과 관련된 실험 질의의 수가 비교적 적은 편이고, 질문을 패턴으로 찾아내어 분류할 경우 질의가 사람의 이름과 같은 사실 여부를 묻는 질문보다 정확한 설명을 요구하는 경우가 있기 때문으로 분석된다.

<표 1> 질의 패턴에 대한 학습 정확률

	Training set	Validation set
1	0.918	0.696
2	0.966	0.783
3	0.966	0.869
4	0.918	0.827
5	0.952	0.827
6	0.952	0.827
7	0.947	0.869
8	0.971	0.826
9	0.932	0.696
10	0.923	0.652

Average	0.944	0.787
---------	-------	-------



<그림 2> 질의 패턴의 학습 정확률

4. 실험 및 평가

4.1 질의 확장 시스템

제안한 질의 확장 시스템은 TREC에서 1988년, 1989년, 1990년 AP의 문서 45810건 총 100,164건을 대상 문서로 하고 해당 문서에 대한 TREC-9의 24개 who 질의를 사용한다. 실험에 사용된 모든 질의 및 문서는 POS 태깅, 스태밍, 불용어 제거 등의 전처리 과정을 거친다. 질의와 정답 단어, 구, 문장 혹은 문단 사이의 유사도[10]를 계산하는 방식은 다음 식(3)과 같다.

$$Sim(Q, D) = \sum_{i,j} \alpha * \delta(q_i, d_j), \quad (3)$$

where Q : Queries D : Documents α : weight
 if $\delta(q_i, d_j) = 1$ then $q_i = d_j$ otherwise 0

또한 질의와의 유사도에 따라서 선정된 추가 용어는 확장하기 전 원래 질의에 포함된 질의어와 비교해서 어느 정도로 신뢰할 수 있는가를 반영하는 탐색가중치를 가지게 된다. 기존 연구에서는 주로 원 질의와의 유사도를 추가 질의어의 가중치로 사용하였다[11].

본 연구에서는 제안한 방법에 적용할 수 있는 적합한 가중치를 선택하기 위하여 질의 확장에 추

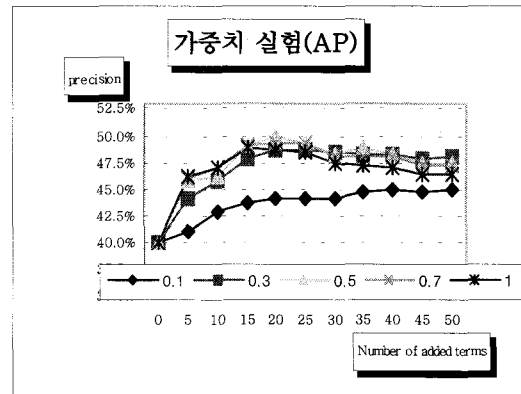
가되는 단어수와 가중치의 변화에 따른 정확률을 실험하였다. <그림 3>은 AP문서에 대해 각각 가중치를 0.1에서 1까지 0.1씩 증가시키면서 정확률의 변화 추이를 측정하여 그 결과를 그래프로 나타내었다.

<그림 3>에 나타난 실험 결과를 보면 확장 용어의 가중치가 0.1에 가깝게 낮을 경우에는 확장 초기에 낮은 성능을 보이다가 점차 향상된다. 반대로 가중치가 1.0에 가까우면 확장 초기에 매우 높은 성능을 보이다가 일정 시점을 넘기면 급격하게 성능이 저하된다. 그러나 가중치를 1.0으로 하였을 때, 즉 초기 질의어와 확장 용어를 대등하게 간주하였을 때에도 일정 수준까지는 성능이 향상됨은 물론이고 높은 성능을 보이기는 하지만 전체적으로 고정 가중치를 0.5 값으로 설정하는 것이 바람직한 것으로 보인다. 적정 수준의 가중치를 사용한 질의 확장은 검색성능 및 효율 면에서 오히려 우월하다는 것을 확인할 수 있었다.

따라서 본 논문에서 제안한 질의 확장 검색에서 원 질의어의 가중치를 1이라고 하였을 때, 확장되는 용어의 가중치는 유사도 0.5로 설정하였다. 또한, 질의 확장시 지나치게 많은 용어를 질의에 추가하게 되면 시스템의 효율이 저하된다는 점을 감안하여 추가 용어의 범위를 20~30개 정도로 한정할 경우 좋은 성능을 얻을 수 있다.

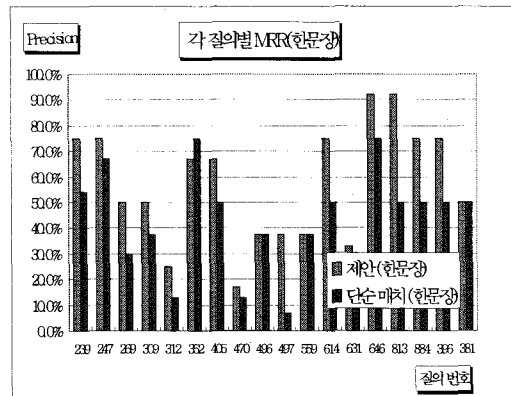
그러므로 수식(1)에서는 질의(Q)와 대상문서(D) 혹은 문단 사이의 유사도는 질의와 대상문서 혹은 문단에서 단어가 일치하는 경우에는 1 그렇지 않은 경우에는 0값으로 나타내었다. α 는 가중치로서 질의어에 존재하는 단어와 일치하는 경우에 대해서는 1.0 을, WordNet을 이용하여 확장된 단어와 일치하는 경우에는 앞서 실험한 타당성에 의해 가중치를 0.5 로 설정하였다.

또한, 정답의 크기에 있어서는, 질의어는 대개 정답과 가까이 나타나기 때문에 예측 질의어가 나타날 범위를 예측 정답을 포함하는 정답 크기(윈도우)로 제한할 수 있고 본 실험을 위해서는 정답 크기의 사이즈를 한 문장 혹은 세 문장단위로 제한하였다.



<그림 3> 확장되는 용어들의 적합한 가중치실험(AP)

그리고 제안한 질의 확장 방법과 단순 매치를 적용한 18개의 질의 처리에 대한 성능을 비교 분석하였다. 평가 척도로는 TREC에서 사용하는 것과 마찬가지로 각 질의에 대해 후보 몇 순위에서 정답이 제시되었는지를 기준으로 정답을 포함하는 순위를 역순한 MRR의 평균을 사용하였다. <그림 4>의 실험 결과에서 확인할 수 있듯이 18개의 각 질의의 정답문의 크기가 한 문장일 때 단순 매치의 경우 평균 정확률은 0.43 제안한 경우 0.57로 나타났다.

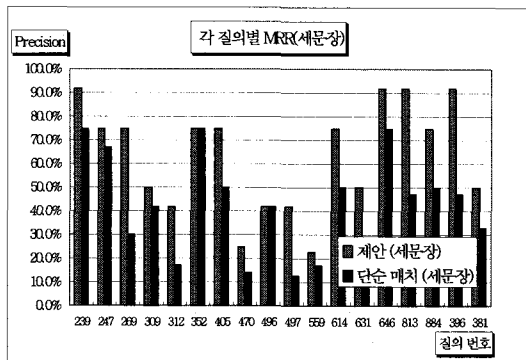


<그림 4> 각 질의별 정확률 (정답 크기:한 문장)

<그림 5>에서는 정답 문장의 크기가 세 문장일 때 단순 매치의 경우 평균 정확률은 0.42인데 비해 제안한 경우가 0.63으로 성능이 향상됨을 보여주었다. 대부분의 질의에 대해서 본 논문에서 제안한 방법이 단순 매치 방법에 비해 정확률 측면

에서 높은 결과치를 보여 주었으며, 정답 크기가 한 문장 단위의 결과보다는 세 문장 단위의 크기에서 보다 정확한 정답을 추출할 수 있었다.

그 이유는 정답이 포함된 구의 크기가 커질수록 포함된 어휘의 수가 증대하고, 질이 확장시 질의 벡터에 추가된 어휘와 매치된 가능성이 높아지기 때문으로 분석된다. 한편, 그림에서 중간 중간 점수가 상대적으로 낮은 부분이 발견되는데 이는 아직 질의문의 분류 방식이 최적화 되지 못했음을 의미하는 것으로 이러한 문제의 해결을 위해서는 보다 자세하고 포괄적인 정답 유형 분류가 필요하다.



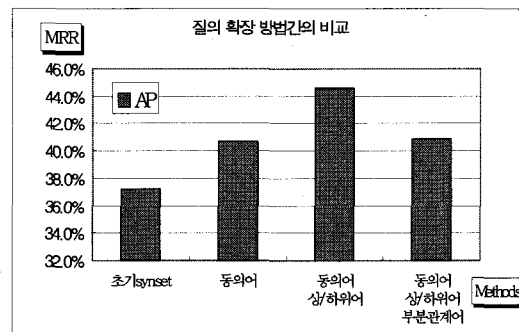
<그림 5> 각 질의별 정확률 (정답 크기:세 문장)

제안한 시스템의 성능을 평가하기 위해 원 synset만을 확장한 경우, 원 synset에 동의어를 함께 확장한 경우, 원 synset에 동의어와 상하위어를 함께 확장한 경우 그리고 마지막으로 원 synset에 동의어와 상하위어 부분 관계어를 함께 확장한 경우에 대해서 성능을 측정하였다. 부분 관계어의 확장에 있어서는 패턴 매치를 위한 탐색 공간의 크기를 고려해 연결 계수를 1로 설정하였다.

실험 결과는 <표 2>에서 나타난 바와 같으며 AP 각각의 문서 집단에 대해 질의 확장 방법을 그래프로 나타내면 <그림 6>과 같다. 이 방식의 확장에 있어서는 초기 synset으로부터 동의어와 상/하위어만을 확장하여 질의 벡터에 추가한 경우가 MRR 측면에서 가장 높은 성능을 보였다.

<표 2> 제안한 시스템의 질의 확장 방법 간의 성능 비교

시스템의 질의 확장 방법	MRR
	AP
초기synset	0.372
동의어	0.407
동의어, 상/하위어	0.446
동의어, 상/하위어, 부분관계어	0.409



<그림 6> 질의 확장 방법간의 비교(AP)

5. 결론

본 논문에서는 질의 개념을 표현하기 위해 동원되는 어휘 집합을 보다 세부적으로 그룹화 할 수 있는 특성을 활용하여, 카테고리 분류를 세분화하고 각 카테고리마다의 개념 커버리지에 기반한 의미적 질의 확장 기법을 제안하였다. 제안한 시스템은 질의 패턴을 학습하여 보다 깊게 세분화된 카테고리로 분류함으로써 유사한 개념을 표현하는 다른 어휘들로 의미 확장을 수행할 수 있도록 하여 정확률을 향상 시켰다.

본 논문에서 제안한 방법의 성능 평가를 위한 실험을 위해서는 TREC 질의 201~893 중 who 질문 117개에 대하여 빈도수를 이용하였다. 구성된 카테고리 개념 커버리지 리스트는 분류 학습에서 일반적으로 이용되는 나이브 베이즈안 분류 학습 알고리즘을 이용하여 학습하였다. 검증을 위해

10-fold 교차 검증을 수행한 결과 학습한 패턴에 대해서는 정확률을 기준으로 94.44%, 학습되지 않은 패턴에 대해서도 78.70%의 정확률을 얻을 수 있었다.

또한, 제한한 QA 시스템에 18개의 질의를 적용하여 성능을 검증한 결과 질의 정답문의 크기가 한 문장일 때 평균 정확률은 0.43에서 0.57로 정답문의 크기가 세 문장일 때 평균 정확률은 0.42에서 0.63으로 성능이 향상됨을 보여주었다.

본 연구에서는 의미적 질의 확장 방법을 제안하였으나 제안한 시스템에서 사용한 질의 어간 유형이 who만을 대상으로 하였기 때문에 when, why와 같은 다른 질의 카테고리에 대한 확장이 필요하고, 학습 문장 패턴에서도 본 연구에서 제안한 패턴 이외의 광범위한 데이터에 대해 실험이 필요하다. 또한 학습 문장 패턴에서 정답 유형이 잘못 할당된 패턴을 교정하거나 중요하지 않은 패턴을 제거하는 등, 학습 데이터를 충분히 정제하고 학습 성능을 높이면 더 좋은 시스템 성능을 얻을 것으로 기대된다.

참 고 문 헌

- [1] M. A. Pasca, S.M. Harabagui, "High Performance Question/Answer", In Proceedings of the 24rd ACM-SIGIR Conference, pp.366-374, 2001.
- [2] R. Mandela, T. Tokunaga, and H. Tanaka, "Combining Multiple Evidence from Different Types of Thesaurus", In Proceedings of the 17th ACM-SIGIR Conference, pp.15-19, 1999.
- [3] E. M. Voorhees, "Query Expansion using Lexical-Semantic Relations", In Proceedings of the 17th ACM-SIGIR Conference, pp. 61-69, 1994.
- [4] G. Miller, "WordNet: A Lexical Database for English", Communications of the ACM 38(11) pp.39-41, 1995.
- [5] D. Moldovan and R. Mihalcea, "Using WordNet and Lexical Operators to Improve Internet Searches", IEEE Internet Computing, pp.34-43, 2000.
- [6] R. Mandela, T. Tokunaga and H. Tanaka, "Combining Multiple Evidence from Different Types of Thesaurus for Query Expansion, In Proceedings of the 22nd Annual International ACM SIGIR Conference, pp.15-19, 1999
- [7] J. Prager, D. Radev, E. Brown, and A. Coden, "The Use of Predictive Annotation for Question Answering in TREC8", In Proceedings of the Eighth Text REtrieval Conference(TREC-8), pp399-411, 2000.
- [8] J. Prager, D. Radev and K. Czuba "Answering What-Is Questions by Virtual Annotation", In Human Language Technology Conference, San Diego, CA, 2001.
- [9] TREC(Text REtrieval Conference) Overview, <http://trec.nist.gov/overview.html>
- [10] R. Gaizauskas and K. Humphreys, "A Combined IR/NLP Approach to QA Against Large Text Collection," In Proceedings of RIAO, pp.1288-1304, 2000.
- [11] G. Salton and C. Buckley, "Term-weighting Approaches in Automatic Text Retrieval," Information Processing and Management 24(5), pp.513-523. 1988.



김 해 정 (Hae-Jung Kim)

- 1987년 2월 : 경북대학교 수학과 (이학사)
- 1989년 2월 : 경북대학교 전자공학과 전산공학전공(공학석사)
- 2005년 2월 : 경북대학교 컴퓨터공학과 (공학박사)
- 2005년 9월 ~ 현재 : 계명대학교 교양학부 초빙전임강사
- 관심분야 : 정보검색, 온톨로지



부 기 동 (Ki-Dong Bu)

- 정회원
- 1983년 2월 : 경북대학교 전자공학과 전자계산기전공(공학사)
- 1986년 2월 : 경북대학교 전자공학과 전산공학전공(공학석사)
- 1996년 8월 : 경북대학교 전자공학과 전산공학전공(공학박사)
- 1988년 3월 ~ 현재 : 경일대학교 컴퓨터공학부 교수
- 관심분야 : 데이터베이스, 정보검색