

영작문 자동 채점 시스템에서의 중복 보고 오류 제거를 통한 성능 향상

이 현 아^{*} · 김 지 은^{**} · 이 공 주^{***}

요 약

영어 작문 자동 채점 시스템은 수험자가 작성한 영작문을 사람의 개입 없이 시스템이 처리하여 점수나 피드백을 줄 수 있는 시스템이다. 본 논문에서 개발한 영작문 자동 채점 시스템은 단어 단계, 구문 단계, 의미 단계의 세 단계 처리 과정을 통해서 오류를 탐지하고 탐지된 오류 개수를 기반으로 채점 점수를 결정한다. 이와 같이 독립적인 세 단계의 처리 과정에서 오류를 탐지하므로 동일한 오류 현상에 대해 서로 다른 단계에서 서로 다른 이름으로 오류를 탐지하는 경우가 발생할 수 있다. 이는 결과적으로 전체 시스템의 채점 점수의 정확도를 저하시키는 원인이 된다. 본 논문에서는 동일한 오류 현상에 대해 서로 다른 단계에서 탐지된 오류 쌍을 '중복 보고 오류'라고 정의한다. 본 논문에서는 중복 보고 오류를 찾는 방법을 제안하고 중복 보고 오류 제거를 통해 영작문 자동 채점 시스템의 채점 점수를 향상시킬 수 있음을 보인다.

키워드 : 영작문 자동 채점 시스템, 중복 보고 오류, 채점 점수, 성능 향상

Accuracy Improvement of an Automated Scoring System through Removing Duplicatedly Reported Errors

Hyun Ah Lee^{*} · Jee Eun Kim^{**} · Kong Joo Lee^{***}

ABSTRACT

The purpose of developing an automated scoring system for English composition is to score English writing tests and to give diagnostic feedback to the test-takers without human's efforts. The system developed through our research detects grammatical errors of a single sentence on morphological, syntactic and semantic stages, respectively, and those errors are calculated into the final score. The error detecting stages are independent from one another, which causes duplicating the identical errors with different labels at different stages. These duplicated errors become a hindering factor to calculating an accurate score. This paper presents a solution to detecting the duplicated errors and improving an accuracy in calculating the final score by eliminating one of the errors.

Keywords : An Automated Scoring System, English Composition, Duplicated Errors, Calculating A Score

1. 서 론

일반적인 영작문 자동 채점 시스템은 학생이 작성한 영작문을 분석하여 적절한 피드백과 함께 점수를 부여하는 시스템이다. 이와 같은 시스템들은 주로 여러 문장이나 여러 단락으로 구성된 에세이에 대한 자동 채점 시스템들이다 [4-7].

본 논문에서 다루고 있는 영작문 자동 채점 시스템[1]은 영어를 제 2 외국어로 접한 지 얼마 되지 않은 학생들을 대상으로 에세이보다는 개별 문장에 대한 기본 문법을 세밀하

게 채점하여 피드백을 주는 것을 목표로 한다. 수험자가 작성한 영어 문장을 자동 채점하는 과정은 (그림 1)과 같다. 우선 교사는 영어 작문 문제와 동시에 영어 작문에 대한 정답집합을 시스템에 제공한다. 수험자는 제시된 문제를 읽고 영어 문장을 작성하여 시스템에 입력한다. 영어 작문 자동 채점 시스템은 수험자의 영어 작문 결과와 교사가 제공한 정답집합을 비교하여 채점을 수행한다. 채점이 완료되면 (그림 1)의 '채점결과'와 같은 정보가 수험자에게 제공된다.

본 논문에서 개발한 영어 작문 자동 채점 시스템은 크게 세 단계의 처리 과정을 수행한다. 첫 번째 단계에서는 수험자가 입력한 영작문 문장을 분석하여 단어 오류가 있는지를 검사한다. 철자 오류, 띄어쓰기 오류, 복수형 활용 형태나 동사 활용 형태 오류 등이 이에 해당한다. 두 번째 단계에서는 수험자의 영작문에 구문 오류가 있는지를 검사한다. 주

* 본 연구는 금오공과대학교 학술연구비에 의하여 연구된 논문임.

† 종신회원: 금오공과대학교 컴퓨터공학부 조교수

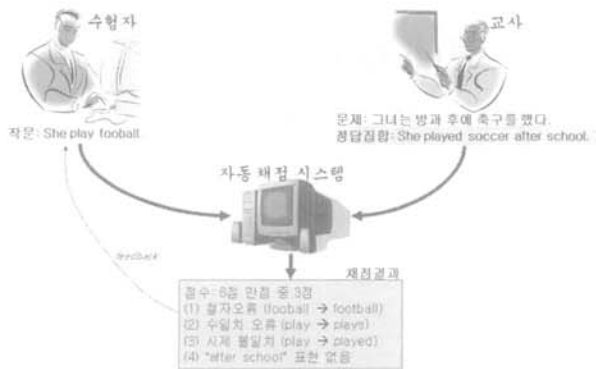
** 정 회 원: 한국의국어대학교 영어학과 조교수

*** 정 회 원: 충남대학교 전기정보통신공학부 부교수(교신저자)

논문접수: 2008년 12월 29일

수정일: 1차 2009년 2월 5일

심사완료: 2009년 2월 5일



(그림 1) 영작문 자동 채점 시스템의 자동 채점 과정

어와 동사 간의 수 일치 오류가 대표적인 구문 오류이다. 세 번째 단계에서는 정답 문장과의 비교를 통해 의미적 차이를 검사하는 매핑 오류 처리 과정이다. 첫 번째 단계와 두 번째 단계는 수험자의 입력 문장만을 분석하여 그 안에 존재할 수 있는 오류를 탐지해 내는 것인 반면, 세 번째 단계에서의 오류 처리는 정답 문장과의 매핑/비교를 통해 차이를 계산하여 얻어진다. (그림 1)의 채점결과 중 (1)은 단어 수준의 오류이며 (2)는 구문 단계의 오류이고 (3)과 (4)는 정답 문장과의 비교를 통해 얻어진 매핑 오류이다.

이와 같이 수험자의 영작문은 세 단계의 분석 과정을 거치면서 각 단계에서 발견한 오류 개수가 최종 채점 점수로 환산된다. 그렇기 때문에 탐지한 오류의 정확한 개수는 곧 영작문 자동 채점 시스템의 채점 점수의 정확도에 영향을 미친다. 단어 단계, 구문 단계, 매핑 단계의 세 단계는 각각 처리하는 방법과 사용하는 정보가 상이하다. 그렇기 때문에 각 단계의 처리는 다른 단계와는 독립적으로 이루어지고 이로 인하여 다음과 같은 문제점이 발생할 수 있다. 다음의 예를 살펴보도록 하자.

(예제 1)	정답문장: The Earth is bigger than the moon. 학생문장: The Earth is small than the Moon.
오류1:	MODIFIER_COMP_ERROR 4-7 (구문 단계의 오류)
오류2:	LEXICAL_ERROR 4 bigger (매핑 단계의 오류)

(예제 1)에서 오류1은 학생문장의 4~7번째 단어에 해당하는 부분에서 비교급 사용 오류(MODIFIER_COMP_ERROR)가 발생했음을 보고하고 있다. 즉, 'small than the Moon'에서 'than'이 사용된 비교구문에서는 원형 'small'이 아닌 비교급을 사용해야 함을 알려주고 있다. 반면 오류2는 학생문장 4번째 단어의 어휘 'small'이 정답문과 동일하지 않음을 보고하고 있으며 정답문장과의 비교를 통해 'bigger'로 수정되어야 함을 알려주고 있다. 오류1은 구문 단계에서 인식된 오류이며 오류2는 매핑 단계에서 인식된 오류로써, 이 두 오류는 비록 같은 단어에서 발생했지만 두 개의 서로 다른 오류로써 수험자에게 보고되어야 한다.

(예제 2)	정답문장: She is too weak to carry the bag. 학생문장: She is too weak to carry the her bag.
오류1:	EXTRA_DET_ERROR 7-9 (구문 단계의 오류)
오류2:	UNNECESSARY_NODE_ERROR 8 (her) (매핑 단계의 오류)

(예제 2)에서 오류1은 학생문장의 7~9번째 단어 부분 'the her bag'에서 불필요한 관사가 사용되고 있음을 보고하는 오류이다. 구문 단계의 처리에서는 정답문장과의 비교를 수행하지 않기 때문에 'the'와 'her'가 동시에 사용되어서는 안 된다는 것을 탐지할 뿐, 어떤 것을 제거해야 하는지 알 수 없다. 반면 오류2는 정답문장과의 비교를 통해서 학생문장의 8번째 단어 'her'가 불필요한 단어임을 보고하고 있다. 사실 이 두 오류는 하나의 동일한 오류에 대해 각 단계의 처리 과정에서 서로 다른 이름으로 보고하고 있는 것이며, 정확한 채점 점수를 얻기 위해서나 수험자에게 정확한 피드백을 보고하기 위해서 하나의 오류는 제거되어야 한다.

앞에서 살펴본 바와 같이 (예제 1)과 (예제 2)는 구문 단계의 오류와 매핑 단계의 오류가 각각 한 개씩 발생한 경우이다. (예제 1)의 경우에는 각 단계의 오류가 서로 독립적인 오류로써 보고되어야 하는 반면 (예제 2)의 경우에는 각 단계의 오류가 동일한 현상에 대한 서로 다른 이름의 오류이므로 하나는 제거되어야 한다. 본 논문에서는 이와 같이 서로 다른 단계에서 오류로 인식되었으나 사실 하나의 오류는 제거되어야 하는 오류 쌍을 '중복 보고 오류'라고 부르기로 한다.

본 논문에서 사용하고 있는 영작문 자동 시스템에서는 단어 수준의 오류가 16개, 구문 단계의 오류가 46개, 매핑 단계의 오류가 14개 정의되어 있다. 정의되어 있는 오류 개수가 많고, 수험자가 영작문 시 범하는 오류가 매우 다양하기 때문에 어떤 오류 쌍들이 중복 보고 오류로 발생할 수 있는지를 추정하는 것은 매우 어렵다. 본 논문에서는 점별 상호 정보(pointwise mutual information)와 상대 빈도 수(relative frequency count)를 이용하여 중복 보고 오류로 간주될 수 있는 오류 쌍들을 우선 찾아보고자 한다. 중복 보고 오류 후보들을 찾은 후에 결정 트리(decision tree)를 이용하여 중복 보고 오류를 확정할 수 있는 규칙을 학습해 본다. 이를 이용하여 중복 보고 오류를 제거하고 이를 기반으로 최종 채점 점수를 산출해 본다. 실험을 통해 영작문 자동 채점 시스템의 채점 정확도를 높일 수 있음을 보인다. 본 논문에서 소개하고 있는 중복 보고 오류 발견 알고리즘은 완전 자동이 아닌 인간보조(human-aided)가 필요한 방법이다.

본 논문은 다음과 같이 구성되어 있다. 2장에서 관련 연구를 간단히 살펴보고, 3장에서 중복 보고 오류를 결정할 수 있는 방법을 제안한다. 4장에서 실험결과를 제시하며 마지막으로 5장에서 결론을 맺는다.

2. 관련 연구

영작문 자동 채점 시스템의 전체 개요와 정의되어 있는 단계별 오류, 그리고 채점 규칙을 소개한다. 자세한 내용은 [1]을 참조하기 바란다.

본 논문에서 개발한 영작문 자동 채점 시스템의 전체 구성은 (그림 2)와 같다. 자동 채점 과정은 우선, 교사가 제공한 정답집합에 대해 형태소분석, 구문분석을 수행하여 단순화된 의존구조를 만들어 놓는다. 수험자가 영작문 답안을 작성하면 이에 대해 형태소 분석, 구문 분석을 동일하게 수

행하는데, 수험자가 작성한 문장에는 오류가 있을 수 있기 때문에 오류 발생을 허용하면서 분석한다. 수험자의 문장도 구문 분석의 결과를 토대로 단순화된 의존구조로 만든다. 두 개의 의존구조를 비교하면서 매핑 오류를 찾아내고 정답

과의 유사도를 계산한다. 본 논문의 영작문 자동 채점 시스템이 다루고 있는 단어 오류는 <표 1>과 같다. 구문 오류는 <표 2>와 같다.

<표 1> 단어 단계의 오류

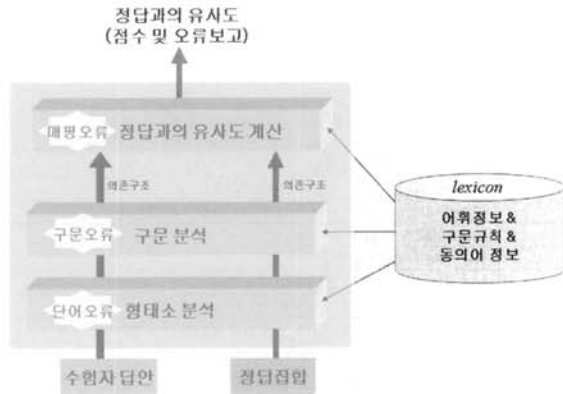
오류 코드	오류 설명	예 제
1 FIRST_WORD_CASE_ERROR	문장시작단어 CASE 오류	"the apple is red."
2 NON_FIRST_WORD_UPPER_CASE_ERROR	문장중간단어 대문자 사용 오류	"I can DO it"
3 WORD_CASE_USAGE_ERROR	단어의 대소문자 사용 오류	"I am from korea."
4 FINAL_PUNCTUATION_ERROR	문장 종결 부호 사용 오류	Can you do that.
5 GENERAL_PUNCTUATION_ERROR	문장 구두점 사용 오류	He will? be a scientist.
6 HYPHEN_ERROR	하이픈 사용 오류	grand-father → grandfather
7 APOSTROPHE_ERROR	어포스트로피 사용 오류	I'am
8 WORD_SPACE_ERROR	단어의 띄어쓰기 오류	every one → everyone
9 NOUN_PLURAL_FORM_ERROR	명사 복수형 활용 오류	boxes childs
10 NOUN_COUNTABLE_FORM_ERROR	명사 수 정보 오류	sugars
11 VERB_INFLECTION_FORM_ERROR	동사 활용 오류	watches turning
12 ADV_ADJ_ER_EST_FORM_ERROR	비교급/최상급 활용 오류	prettier beautifulst
13 MISSPELLING_ERROR	철자 오류	palys
14 MISSPELLING_SPACE_ERROR	철자 오류(붙여쓰기 오류)	thesedays → these days
15 CONFUSABLE_WORD_ERROR	혼동되는 단어 사용 오류	accept/except advise/advice
16 WORD_FORMALITY_ERROR	단어 형식 오류	wanna gonna

<표 2> 구문 단계의 오류

오류 코드	오류 설명	예 제
1 DET_NOUN_AGR_ERROR	관사-명사 수 일치 오류	a sheep / a bits
2 SUBJ_VERB_AGR_ERROR	주어-동사 수 일치 오류	He have gone.
3 VERB_COMPL_AGR_ERROR	Be동사 보어 일치 오류	He is the tallest students.
4 DET_NOUN_CV_ERROR	관사 타입 오류	a apple / an yacht
5 DET_FORM_ERROR	관사 형태 오류	for him wife
6 EXTRA_PREP_ERROR	전치사 추가 오류	She went to there.
7 EXTRA_CONJ_ERROR	접속사 추가 오류	Although he is poor, but he is happy.
8 EXTRA_ADV_ERROR	부사 추가 오류	You had better than go home now.
9 EXTRA_DET_ERROR	관사 추가 오류	the Jane swimming
10 MODIFIER_COMP_ERROR	비교급 오류	more prettier / more pretty
11 MODIFIER_SUPER_ERROR	최상급 오류	most successful the most tallest student
12 ADJ_SUBCAT_ERROR	형용사 타입 오류	an asleep baby
13 VERB_SUBCAT_ERROR	동사 타입 오류	He lived being 90.
14 NOUN_SUBCAT_ERROR	명사 타입 오류	the book that I gave it to him
15 RELCL_SUBJ_PRON_ERROR	주격 관계대명사 오류	That is the man whom is nice.
16 RELCL_OBJ_PRON_ERROR	목적격 관계대명사 오류	That is the man who I have seen before.
17 RELCL_SUBJ_ERROR	관계대명사 오류	She is the teacher who she came.
18 RELCL_PRON_MISSING_ERROR	관계대명사 오류	She is a teacher come to my school last week.
19 VERB_FORM_ERROR	동사 형태 오류	He spoken French.
20 VERB_ING_FORM_ERROR	ING 형태 오류	be knowing/needing
21 VERB_VOICE_FORM_ERROR	동사 수동태 오류	He got pull.
22 VERB_INF_FORM_ERROR	동사 원형 오류	He can speaks English.
23 VERB_PERFECT_FORM_ERROR	동사 완료형 오류	He has went home.
24 VERB_TO_INF_FORM_ERROR	To부정사 오류	He wants help many people.
25 ADV_TYPE_ERROR	부사 타입 오류	he is much (more) famous than his father.
26 PREP_TYPE_ERROR	전치사 타입 오류	It is better to me to go home.
27 PRON_TYPE_ERROR	대명사 타입 오류	It is different from he opinion.
28 RELCL_PRON_TYPE_ERROR	관계대명사 타입 오류	I will order the same what you eat.
29 AUX_TYPE_ERROR	조동사 타입 오류	I would better go home.
30 NOUN_NUMBER_ERROR	명사 수 오류	He is more intelligent than any other students.
31 DET_NOUN_COUNT_ERROR	한정사-명사 일치 오류	much leaves, many furniture
32 AUX_MISSING_ERROR	조동사 생략 오류	He worked hard so that he won the prize.
33 CONJ_MISSING_ERROR	접속사 생략 오류	The price get lower, people buy more.
34 VERB_MISSING_ERROR	동사 생략 오류 (be동사)	I would happy if I were a bird.
35 NEG_ERROR	부정구문 오류	They don't have been there.
36 ADV_NEG_ERROR	부정문의 부사 사용 오류	I don't like it, too.
37 ADJ_ORDER_ERROR	형용사 어순 오류	special something
38 SUBJ_VERB_WORD_ORDER_ERROR	주어-동사 어순 오류	How I can go to the school?
39 VERB_ADV_ORDER_ERROR	부사 어순 오류 (빈도 부사)	She make always me happy.
40 ADVP_POSITION_ERROR	부사구 위치 오류	He is more popular in Korea than any other actor.
41 LEXICALIZED_ERROR	숙어 오류	You had better go to home now.
42 WHWORD_ERROR	의문대명사 오류	Which she said is true.
43 SUBCONJ_ERROR	상속접속사 오류	I hope if I were a bird.
44 COND_CLAUSE_ERROR	조건문 오류	I wish I was a bird.
45 IDIOM_PREP_ERROR	관용 표현 내의 전치사 오류	It is different with that.
46 PUNC_ERROR	문장 부호 오류	When the final SENT PUNC is missing

<표 3> 매핑 단계의 오류

오류 코드	오류 설명	예 제
1 OBLIGATORY_NODE_MISSING_ERROR	필수격 성분 생략 오류	정답: You look tired. 오류: You tired.
2 OPTIONAL_NODE_MISSING_ERROR	선택격 성분 생략 오류	정답: You had better go home now. 오류: You had better go home.
3 UNNECESSARY_NODE_ERROR	불필요한 단어 추가 오류	정답: That is why I called you. 오류: That is because why I call you.
4 LEXICAL_ERROR	단어 선택 오류	정답: This city is where my grandfather was born. 오류: This place is where my grandfather was born.
5 MOOD_UNMATCHED_ERROR	mood 불일치 오류 (명령문, 의문문, 병서문과 같은 서법 불일치)	정답: Which do you prefer, winter or summer? 오류: You like a summer or winter.
6 MODAL_UNMATCHED_ERROR	modal 불일치 오류 (가능/필요/의무 불일치)	정답: I could buy the computer. 오류: I bought the computer.
7 TENSE_UNMATCHED_ERROR	시제 불일치 오류 (현재/과거/미래 불일치)	정답: He is more famous than his father. 오류: He was more famous than his father
8 VOICE_UNMATCHED_ERROR	태 불일치 오류 (수동태/능동태 불일치)	정답: French is spoken in France. 오류: French is speaking by France
9 NEGA_UNMATCHED_ERROR	부정어 사용 불일치 오류	정답: I have not been there. 오류: I have been to there.
10 ASPECT_UNMATCHED_ERROR	aspect 불일치 오류 (진행형, 완료형 불일치)	정답: She went there. 오류: She was going there.
11 FORM_UNMATCHED_ERROR	form 불일치 오류	정답: I sometimes heard Mary playing the piano. 오류: I sometimes heard Mary's playing the piano.
12 AGR_UNMATCHED_ERROR	단수/복수 불일치 오류	정답: I read the books. 오류: I read the book.
13 DET_UNMATCHED_ERROR	관사 불일치 오류	정답: I read the book. 오류: I read a book.
14 IDIOM_UNMATCHED_ERROR	숙어표현 불일치 오류	정답: I had better go home. 오류: I had better now go to home.



(그림 2) 영작문 자동 채점 시스템의 전체 개요

매핑 오류는 <표 3>과 같다.

본 논문의 영작문 자동 채점 시스템에서는 단어 오류, 구문 오류, 매핑 오류의 세 단계로 나누어 점수를 계산하였으며 이 세 단계의 점수를 합산하여 최종 점수화하였다. 각 단계마다 0~2점 사이의 점수가 할당되어 있기 때문에 최종 점수의 총점은 6점이다. 점수를 부여하는 기준은 현직 중학교 교사들과의 협의하에 <표 4>와 같은 기준표를 얻어 사용하였다. 최종 채점 점수가 각 단계에서 얻어진 오류 개수를 기반으로 이루어지기 때문에 각 단계별 오류 개수를 정확히 얻어내는 것이 매우 중요하다.

<표 4> 점수 계산을 위한 기준표

점수	단어 오류	구문 오류	매핑 오류
2	없음	없음	모두 적절
1	1~2개	1개 이하	1~2개 틀림
0	3개 이상	2개 이상	3개 이상 틀림

3. 중복 보고 오류

본 논문에서 다루는 영작문 자동 채점 시스템은 탐지된 오류를 다음과 같은 형식으로 보고한다.

오류_이름 | 오류_범위 | 부가_정보

각각의 오류 보고는 세 필드로 구성되는데 각 필드는 ' ' 로 구분된다. 첫 번째 필드에는 탐지된 오류 이름으로 <표 1~3>의 오류 코드가 사용된다. 두 번째 필드에는 수험자 문장에서 해당 오류가 발생한 오류 범위를 나타낸다. 오류 범위 "3,5-7"는 학생이 입력한 문장 중 3번째 단어와 5번째 부터 7번째 단어까지에서 해당 오류가 발생했음을 의미한다. 주로 구문 단계에서 탐지된 오류는 구절단위이기 때문에 여러 단어에 걸친 오류 범위를 갖는다. 세 번째 필드는 오류 종류에 따라 부가 정보 값이 있을 수도 있으며 없을 수도 있다. 주로 정답문장과의 비교를 통해 해당 오류가 어떻게 수정되어야 하는지를 알려주는 오류수정제안 스트링이 나타난다.

3.1 중복 보고 오류 정의

본 논문에서는 두 개의 오류가 다음의 세 가지 조건을 동시에 만족할 때 '중복 보고 오류'라고 한다.

- (조건 1) 두 개의 오류 범위 중 공유되는 부분이 존재한다.
- (조건 2) 두 개의 오류는 서로 다른 처리 단계에서 탐지된 오류이다.

(조건 3) 두 개의 오류는 동일한 오류 현상을 다루고 있다.

여기서 (조건 1)은 두 오류의 처리 범위에서 서로 공통적으로 다루는 단어가 한 단어 이상 존재해야 함을 의미한다.

단, 매핑 오류 중 'OBLIGATORY_NODE_MISSING_ERROR'(필수적 성분 생략 오류)와 'OPTIONAL_NODE_MISSING_ERROR'(선택적 성분 생략 오류)의 경우에는 정답문장에는 쓰였으나 학생이 작성한 문장에는 나타나지 않은 단어에 대한 오류 보고이기 때문에 오류의 지적 범위가 학생문장의 위치가 아닌 정답문장의 위치로 나타난다. (예제3)에서 오류 5와 오류6이 이에 해당하며 오류 범위 '(7)'과 '(8)'은 각각 학생문장의 위치가 아닌 정답문장에서의 위치 7번째 단어인 'to'와 8번째 단어인 'our'가 생략되었음을 의미한다. 이처럼 오류 범위가 정답문장의 위치를 나타낼 경우에는 학생답안문의 모든 단어와 공유된다고 간주한다. 즉, (예제 3)의 오류5와 오류6은 오류1~4까지와 모두 다 공유되는 부분이 존재한다고 간주한다.

정답문장: She is a teacher who came to our school last week. (예제 3) 학생문장: She is a teacher who come school last weak.
오류1: CONFUSABLE_WORD_ERROR[9]week (단어 단계의 오류)
오류2: SUBJ_VERB_AGR_ERROR[3-7] (구문 단계의 오류)
오류3: VERB_SUBCAT_ERROR[6-7] (구문 단계의 오류)
오류4: TENSE_UNMATCHED_ERROR[6]came[past] (매핑 단계의 오류)
오류5: OPTIONAL_NODE_MISSING_ERROR[7]to (매핑 단계의 오류)
오류6: OPTIONAL_NODE_MISSING_ERROR[8]our (매핑 단계의 오류)

(예제 3)에서 오류1과 오류2는 서로 공유되는 오류 범위가 없기 때문에 (조건 1)을 만족하지 못하여 중복 보고 오류가 아니다. 또한 오류2와 오류3은 공유되는 오류 범위(6-7 단어)는 존재하나 두 오류가 모두 구문 단계의 오류이기 때문에 역시 중복 보고 오류가 아니다. 또한 오류2와 오류4는 (조건 1)과 (조건 2)는 모두 만족하지만 (조건 3)을 만족하지 못하기 때문에 중복 보고 오류가 아니다. 오류2는 주어와 서술어의 수가 일치되지 않음을 지적하는 오류이며 오류4는 시제 선택이 잘못된 오류임을 지적하고 있어 같은 현상이 아니기 때문이다¹⁾. 오류3과 오류5는 (조건 1)과 (조건 2)를 만족한다. 또한 오류3과 오류5는 'come'이라는 동사가 직접 목적어를 취할 수 없으며 'to'와 같은 전치사가 이끄는 부사구를 취해야 한다는 동일한 오류 현상을 지적하고 있어 (조건 3) 역시 만족한다. 즉 이 두 오류는 중복 보고 오류이다.

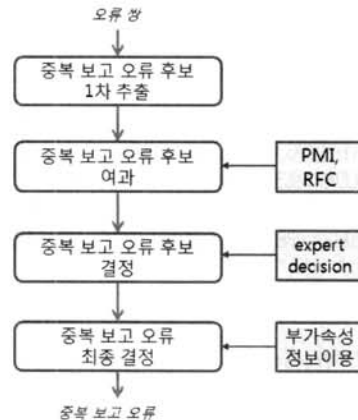
3.2 중복 보고 오류 발견

중복 보고 오류는 3.1절에서 언급한 바와 같이 세 가지 조건을 모두 만족해야 한다. 중복 보고 오류를 찾기 위한 과정은 (그림 3)과 같다. 첫째 수험자가 작성한 영작문 문장 14,892개에 대해 영작문 자동 채점 시스템의 자동 채점 결과를 얻는다. 이 결과로부터 3.1절에서 언급한 (조건 1)과 (조건 2)를 만족하는 중복 보고 오류 후보를 추출한다 (중복 보고 오류 후보 1차 추출). 이렇게 추출한 오류 쌍들 중 점별 상호정보(pointwise mutual information)와 상대빈도 수

<표 5> (조건 1)과 (조건 2)을 만족하는 오류 쌍 (일부)

1	CONFUSABLE_WORD_ERROR[9]week	OPTIONAL_NODE_MISSING_ERROR[7]to
2	CONFUSABLE_WORD_ERROR[9]week	OPTIONAL_NODE_MISSING_ERROR[8]our
3	SUBJ_VERB_AGR_ERROR[3-7]	TENSE_UNMATCHED_ERROR[6]came[past]
4	SUBJ_VERB_AGR_ERROR[3-7]	OPTIONAL_NODE_MISSING_ERROR[7]to
5	SUBJ_VERB_AGR_ERROR[3-7]	OPTIONAL_NODE_MISSING_ERROR[8]our
6	VERB_SUBCAT_ERROR[6-7]	TENSE_UNMATCHED_ERROR[6]came[past]
7	VERB_SUBCAT_ERROR[6-7]	OPTIONAL_NODE_MISSING_ERROR[7]to
8	VERB_SUBCAT_ERROR[6-7]	OPTIONAL_NODE_MISSING_ERROR[8]our

(relative frequency count)를 이용하여 매우 드물게 발생하는 오류 쌍이나 실제 중복 보고 오류로써 무의미한 오류 쌍들을 여과한다. 세 번째 단계인 중복 보고 오류 후보 결정은 영작문을 채점할 수 있는 전문가(expert)의 결정에 의해 수동으로 이루어진다. 마지막으로 부가 속성 정보를 이용하여 중복 보고 오류를 최종적으로 결정한다. 다음 절부터 각 단계에 대한 자세한 설명을 다룬다.



(그림 3) 중복 보고 오류 결정 과정

3.2.1 중복 보고 오류의 후보 1차 추출

수험자의 영작문에 대한 채점 결과로부터 (조건 1)과 (조건 2)를 만족하는 오류 쌍들을 추출한다. (예제3)으로부터 (조건 1)과 (조건 2)를 만족하는 오류 쌍들은 다음의 8개를 추출할 수 있었다.

전체 14,892개의 문장에 대해 (조건1)과 (조건2)를 만족하는 오류 후보 1차 추출의 결과는 모두 150,419개의 오류 쌍들이 추출되었다.

3.2.2 중복 보고 오류의 후보 여과

3.2.1절에서 추출한 오류 쌍들을 각 오류 종류별로 그 빈도수를 계산한다. 150,419개의 오류 쌍들은 657개의 오류 종류별로 정리할 수 있었다. <표 6>은 구해진 오류 종류별 오류 쌍들의 빈도수 일부를 보여준다. 추출된 오류 쌍들 중에서 점별 상호정보(pointwise mutual information)와 상대빈도 수(relative frequency count)를 이용하여 의미 없는 오류 쌍을 여과한다 [8].

1) 'come'을 'came'으로만 고쳐 작성할 경우 오류2와 오류4는 한꺼번에 수정이 될 수 있어 이 두 오류를 동일한 오류로 간주할 수도 있으나, 수험자에게 수험자의 실수를 정확히 피드백 주기 위해서 본 논문에서는 이 두 오류를 서로 다른 오류로 취급한다.

<표 6> 오류 종류별로 추출한 중복 보고 오류 후보 (일부)

오류1	오류1의 발생빈도	오류2	오류2의 발생빈도	같이 발생한 빈도
AUX_MISSING_ERROR	201	TENSE_UNMATCHED_ERROR	3744	127
CONFUSABLE_WORD_ERROR	69	OPTIONAL_NODE_MISSING_ERROR	17979	54
SUBJ_VERB_AGR_ERROR	1226	TENSE_UNMATCHED_ERROR	3744	364
SUBJ_VERB_AGR_ERROR	1226	OPTIONAL_NODE_MISSING_ERROR	17979	1424
VERB_SUBCAT_ERROR	289	TENSE_UNMATCHED_ERROR	3744	82
VERB_SUBCAT_ERROR	289	OPTIONAL_NODE_MISSING_ERROR	17979	287
			

<표 7> 중복 보고 오류 후보의 부류

A 중복 보고 오류임 (207개)	(DET_NOUN_CV_ERROR, DET_UNMATCHED_ERROR) (EXTRA_DET_ERROR, DET_UNMATCHED_ERROR) (MODIFIER_COMP_ERROR, FORM_UNMATCHED_ERROR) (VERB_PERFECT_FORM_ERROR, ASPECT_UNMATCHED_ERROR) (MISSPELLING_ERROR, LEXICAL_ERROR) ...
B 중복 보고 오류 아님 (47개)	(SUBJ_VERB_AGR_ERROR, TENSE_UNMATCHED_ERROR) (AUX_MISSING_ERROR, UNNECESSARY_NODE_ERROR) (CONJ_MISSING_ERROR, DET_UNMATCHED_ERROR) (CONJ_MISSING_ERROR, AGR_UNMATCHED_ERROR) (MISSPELLING_ERROR, OPTIONAL_NODE_MISSING_ERROR) (DET_NOUN_CV_ERROR, OPTIONAL_NODE_MISSING_ERROR) ...
C 중복 보고 오류인 경우와 아닌 경우 혼합 (44개)	(VERB_FORM_ERROR, ASPECT_UNMATCHED_ERROR) (VERB_ING_FORM_ERROR, TENSE_UNMATCHED_ERROR) (CONJ_MISSING_ERROR, OPTIONAL_NODE_MISSING_ERROR) (VERB_SUBCAT_ERROR, OPTIONAL_NODE_MISSING_ERROR) (EXTRA_PREP_ERROR, UNNECESSARY_NODE_ERROR) ...

로 영작문을 채점할 수 있는 전문가에 의해 수동으로 결정된다. 전문가는 111개 오류 쌍의 29,588개의 오류 후보에 대해 중복 오류인지 아닌지를 수동으로 판단한다. 수동으로 판단한 결과로부터 111개의 오류 쌍 종류는 세 개의 부류로 나눌 수 있었다 (<표 7>). 첫 번째 부류 A는 (조건 1~2)을 만족하는 모든 오류 후보 쌍이 중복 보고 오류인 경우이고, 두 번째 부류 B는 (조건 1~2)을 만족하는 모든 오류 후보 쌍이 중복 보고 오류가 아닌 경우이며, 세 번째 부류 C는 경우에 따라서 중복 보고 오류인 경우도 있고 그렇지 않은 경우도 있는 것들이다.

A부류의 오류 쌍인 경우에는 중복 보고 오류이기 때문에 (조건 1)과 (조건 2)을 만족하는 오류 쌍에 대해 두 오류 중 하나를 제거해도 무방하다. B부류의 오류 쌍은 중복 보고 오류가 아니기 때문에 각각 두 개의 독립적인 오류로 간주해야 한다. C부류의 오류 쌍인 경우에는 부가적인 정보를 이용하여 중복 오류인지 그렇지 않은지를 결정해야 한다. C부류에 대한 결정 방법은 3.2.4 절에서 다룬다.

$$PMI(E_1, E_2) = \log \frac{P(E_1, E_2)}{P(E_1)P(E_2)} \quad (1)$$

$$RFC(E_1, E_2) = \frac{freq(E_1, E_2)}{freq} \quad (2)$$

절별 상호정보 PMI (수식 1)는 두 오류가 동시에 발생하는 정도를 나타내는 수치이고, 상대빈도 수 RFC (수식 2)는 전체 오류 후보 쌍의 평균 발생 빈도에 대한 상대적 빈도를 의미한다. (수식 2)에서 \overline{freq} 는 모든 후보 쌍의 빈도수에 대한 평균이다. 여과 조건은 (수식 3) 같다.

$$PMI(E_1, E_2) \times RFC(E_1, E_2) \geq k \quad (3)$$

임계값 k 를 넘는 오류 쌍만을 남긴다. 본 논문에서 k 는 0.4를 사용하였으며, 결과로 남는 오류 쌍의 종류는 모두 111가지였다.

3.2.3 중복 보고 오류의 후보 결정

3.2.2 절에서 여과한 오류는 모두 111개의 오류 종류에 대해 총 29,588개의 오류 쌍이 발생하였다. 각 오류 종류당 평균 278개의 오류 쌍이 발생하였다. (조건 3)의 부합 여부, 즉 동일한 현상을 다루고 있는 오류인지 아닌지는 최종적으

3.2.4 부가 정보를 이용한 중복 보고 오류 결정

C부류에 대한 부가적인 정보를 추출하기 위해 C부류에 속하는 오류 후보들을 살펴보자. 다음의 예들을 살펴보자.

(예제4) 정답문장: I don't know why she went there. 학생문장: I don't know why she go to their.	
오류1: CONFUSABLE_WORD_ERROR 8 there (단어 단계의 오류)	
오류2: SUBJ_VERB_AGR_ERROR 6 went 1S 2S 1P 2P 3P->3S (구문 단계의 오류)	
오류3: EXTRA_PREP_ERROR 6 -8 (구문 단계의 오류)	
오류4: UNNECESSARY_NODE_ERROR 7 (to) (매핑 단계의 오류)	
오류5: TENSE_UNMATCHED_ERROR 6 went (매핑 단계의 오류)	
[pres inf->past]	

(예제5) 정답문장: Would you like to come? 학생문장: you go to home?	
오류1: FIRST_WORD_CASE_ERROR 1 (단어 단계의 오류)	
오류2: EXTRA_PREP_ERROR 3-4 (구문 단계의 오류)	
오류3: OBLIGATORY_NODE_MISSING_ERROR (매핑 단계의 오류)	
(1,3) Would like	
오류4: UNNECESSARY_NODE_ERROR 4 (매핑 단계의 오류)	
(home)	
오류5: LEXICAL_ERROR 2 come (매핑 단계의 오류)	

(예제 4)과 (예제 5)에서는 동일하게 'EXTRA_PREP_ERROR'와 'UNNECESSARY_NODE_ERROR'가 중복 보고 오류 후보로 추출된다. 그러나 (예제 4)의 경우에는 중복 보고 오류이지만, (예제 5)의 경우에는 중복 보고 오류가 아니

<표 8> 중복 보고 오류를 결정하기 위한 부가 정보

E1, E2의 오류범위 중 공유되는 길이 (shared_length)
E1, E2의 오류범위 중 공유되지 않는 길이 (non_shared_length)
E1오류의 오류수정제한 스트링의 품사와 부가정보 (E1.POS)
E1오류의 오류수정제한 스트링의 품사와 부가정보 (E2.POS)
E1오류와 E2오류의 오류수정제한 스트링의 스트링 유사성 (edit_distance)
E1오류의 오류 발생 위치 (E1.position)
E1오류의 오류 발생 위치 (E2.position)
E1오류의 발생 위치-E2오류의 발생 위치 (diff_error_position)

다. 이 두 경우의 차이는 (예제 4)의 경우에는 불필요하다고 지목된 단어가 전치사 'to'인데 비해 (예제 5)의 경우에는 명사 'home'이 불필요한 단어로 지적되고 있다.

(예제 6) 정답문장: We enjoyed listening to music. 학생문장: We enjoy listen to music
오류1: VERB_ING_FORM_ERROR[2-3] (구문 단계의 오류)
오류2: TENSE_UNMATCHED_ERROR[2] (매핑 단계의 오류) enjoyed[past]
오류3: TENSE_UNMATCHED_ERROR[3] (매핑 단계의 오류) listening[ing]

(예제 7) 정답문장: I look forward to seeing the concert tonight. 학생문장: I look forward to see concert tonight
오류1: VERB_ING_FORM_ERROR[2-6] (구문 단계의 오류)
오류2: TENSE_UNMATCHED_ERROR[5] (매핑 단계의 오류) seeing[ing]
오류3: DET_UNMATCHED_ERROR[6] (매핑 단계의 오류) the concert

(예제 6)과 (예제 7)의 경우에는 모두 'VERB_ING_FORM_ERROR'와 'TENSE_UNMATCHED_ERROR'가 중복 보고 오류 후보로 추출되었다. (예제 6)의 경우에는 한 문장 내에서 동일한 종류의 중복 보고 오류 후보가 2개나 발생하였다 (오류 1-오류 2, 오류 1-오류 3). (예제 6)과 (예제 7) 두 경우 모두 부가정보에 'ing'가 사용된 'TENSE_UNMATCHED_ERROR'가 포함된 쌍이 중복 보고 오류이다.

본 논문에서는 중복 보고 오류 후보로부터 최종적으로 중복 보고 오류를 결정하기 위하여 <표 8>과 같은 부가 정보를 추가하였다. 오류 쌍 후보 (E1, E2)에 대해 <표 8>의 8가지 부가 속성을 추출하였다. 결정트리(decision tree)를 이용하여 각 오류 종류별로 필요한 속성집합과 중복 보고 오류를 결정하기 위한 규칙을 추출하였다.

결정트리는 CART [3]를 이용하여 최종 중복 보고 오류 임을 결정하기 위해 필요한 속성 집합과 결정 규칙을 학습하였다. CART 학습에 사용된 속성 정보는 <표 8>과 같으며, <표 9>에 학습된 규칙의 일부를 보인다.

4. 중복 보고 오류 제거 성능에 대한 실험

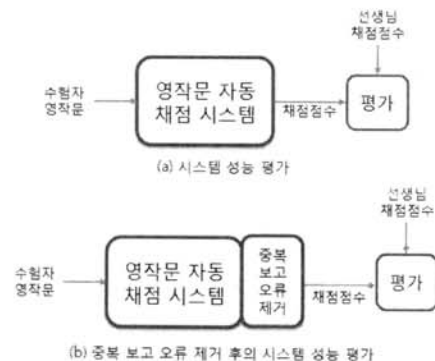
영작문 자동 채점 시스템에서 중복 보고 오류를 제거하는 과정은 다음과 같다. 수험자 영작문에 대해 영작문 자동 채점 시스템이 출력해 주는 오류 보고에 대해 (조건 1~3)을

<표 9> 결정트리를 이용하여 학습한 중복 보고 오류 결정 규칙 (일부)

E1=CONJ_MISSING_ERROR, E2=OPTIONAL_NODE_MISSING_ERROR if E2.POS = 'conj' and E2.position = 1 then 중복_보고_오류=TRUE
E1=EXTRA_PREP_ERROR, E2=UNNECESSARY_NODE_ERROR if E2.POS='prep' then 중복_보고_오류=TRUE
E1=VERB_SUBCAT_ERROR, E2=OPTIONAL_NODE_MISSING_ERROR if diff_error_position<=3 and E2.POS={prep, adv} then 중복_보고_오류
E1=VERB_ING_FORM_ERROR E2=TENSE_UNMATCHED_ERROR if E2.POS='verb-ing' then 중복_보고_오류=TRUE
E1=ADV_ADJ_ER_EST_FORM_ERROR, E2=OPTIONAL_NODE_MISSING_ERROR if diff_error_position<=3 and E2.POS=adv then 중복_보고_오류=TRUE
E1=ADV_ADJ_ER_EST_FORM_ERROR, E2=FORM_UNMATCHED_ERROR if diff_error_position<=0 and E2.POS={adj, adv, adj-comparative, adv-comparative} then 중복_보고_오류=TRUE
E1=MISSPELLING_SPACE_ERROR, E2=UNNECESSARY_ERROR if diff_error_position<=0 and non_shared_length=0 and edit_distance<=3 and E2.POS='det' then 중복_보고_오류=TRUE

모두 만족하는 오류 쌍을 추출한다. 중복 보고 오류 발생 시 두 오류 중 하나는 제거되어야 한다. 제거되는 오류는 오류의 우선순위를 기반으로 처리하였다. 첫째 (형태소 단계 오류, 매핑 단계 오류) 쌍에 대해 매핑 단계의 오류를 제거하고, 두 번째로 (구문 단계 오류, 매핑 단계 오류) 쌍에 대해 매핑 단계의 오류를 제거한다. 마지막으로 (형태소 단계 오류, 구문 단계 오류) 쌍이 남아 있는 경우 형태소 단계의 오류를 제거한다. 구문 단계의 오류는 구문 전체를 파악하여 탐지된 오류이기 때문에 가장 높은 우선순위를 갖도록 한 반면 매핑 단계의 오류는 단순히 정답문장과의 단어 대 단어의 매핑으로만 탐지된 오류이기 때문에 가장 낮은 우선순위를 갖도록 하였다.

중복 보고 오류를 제거한 후의 시스템 성능에 대한 평가를 수행해 보았다. 평가 방법은 (그림 4)와 같다. 기본적인 시스템의 성능 평가는 (그림 4)의 (a)와 같이 수험자가 작성한 영작문에 대해 시스템의 채점 점수와 선생님의 채점 점



(그림 4) 시스템 성능 평가 방법

수를 비교하여 그 차이를 시스템의 채점 정확도로 계산한다. 2장에서 설명한 바와 같이 영작문의 채점은 0~6점의 스케일을 갖는다. 본 논문에서 제안한 중복 보고 오류를 제거하고 난 후의 성능 평가는 (그림 4)의 (b)와 같이 수행된다.

시스템의 성능은 선생님의 채점 점수와 시스템의 최종 채점 점수의 일치도(exact agreement)[2]로 계산하였다. 학습에 사용된 14,892문장에 대한 성능 평가 결과는 <표 10>과 같다.

<표 10> 영작문 자동 채점 시스템의 채점 점수의 정확도

점수 일치도	중복 보고 오류 제거 전의 시스템 성능	중복 보고 오류 제거 후의 시스템 성능
정확 일치	48.66%	52.02%
1점 이내 일치	79.86%	83.91%

5. 결 론

본 논문에서는 영작문 자동 채점 시스템에서 발생할 수 있는 중복 보고를 처리하는 방법에 대해 소개하였다. 본 논문의 영작문 자동 채점 시스템은 수험생이 작성한 영작문에 대해 형태소 분석, 구문 분석 과정을 통해 단어 오류나 구문 오류를 탐지해 낸다. 교사가 제공한 답안 문장과의 비교를 통해 의미적 차이를 오류로 탐지해 낸다. 이와 같이 세 단계의 독립적인 처리 과정을 거치다 보니 동일한 오류 현상을 중복해서 탐지해 내는 문제가 발생하였다. 수험자의 답안에서 발생할 수 있는 오류를 모두 다 예측할 수 없기 때문에 중복으로 발생할 수 있는 오류를 예측하는 것 또한 쉽지 않다. 본 논문에서는 실제 수험자들이 작성한 영작문 문장으로부터 중복 보고 오류를 찾아내는 체계적인 방법을 제안하였다. 이렇게 찾아낸 중복 보고 오류를 제거함으로써 전체 영작문 자동 채점 시스템의 정확도를 향상시킬 수 있었다.

참 고 문 헌

[1] 김지은, 이공주, 진경애, 단문형의 영작문 자동 채점 시스템 구축, 정보처리학회논문지B 제14-B권 제3호, pp.223-230, 2007.

[2] Attali, Yigal, Burstein, J. Automated Essay Scoring With e-rater® V.2. *The Journal of Technology, Learning, and Assessment*, Vol4, No.3, Feb., 2006.

[3] Breiman, L., Friedman, J., and Stone, Charles J. *Classification and Regression Trees* Belmont, Calif.: Wadsworth, 1984.

[4] Burstein, J., Daniel Marcu and Kevin Knight, Finding the WRITE Stuff: Automatic Identification of Discourse Structure in Student Essays, *IEEE Intelligent Systems*, Vol.18, Issue1, pp.32-39, 2003.

[5] Burstein, J., Martin Chodorow and Claudia Leacock. Criterion Online Essay Evaluation: An Application for Automated Evaluation of Student Essays. In *Proceedings of the Fifteenth Annual Conference on*

Innovative Applications of Artificial Intelligence, 2003.

[6] Burstein, Jill and Derrick Higgins, Advanced Capabilities for Evaluation Student Writing: Detecting Off-Topic Essays Without Topic-Specific Training, In *Proceedings of the International Conference on Artificial Intelligence in Education*, pp.112-119, July, 2005.

[7] Higgins, D., Burstein, J., Marcu, D., & Gentile, C. Evaluating multiple aspects of coherence in student essays (PDF). In *Proceedings of the Annual Meeting of HLT/NAACL*, 2004.

[8] Su, Keh-Yih, Ming-Wen Wu and Jing-Shin Chang. A Corpus-based Approach to Automatic Compound Extraction, In *Proceedings of the ACL 94*. pp.242-247, 1994.



이 현 아

e-mail : halee@kumoh.ac.kr

1996년 연세대학교 컴퓨터과학과(학사)
 1998년 한국과학기술원 전산학과(공학석사)
 2004년 한국과학기술원 전산학과(공학박사)
 2000년~2004년 (주)다음소프트 언어연구팀장

2004년~현 재 금오공과대학교 컴퓨터공학부 조교수
 관심분야 : 자연언어처리, 정보검색, 지식공학, 기계번역



김 지 은

e-mail : jeeunk@hufs.ac.kr

1985년 한국외국어대학교 영어과(학사)
 1989년 Georgetown University 언어학과(석사)
 1993년 Georgetown University 언어학과(박사)

1993년~1994년 대학 강사
 1995년~2002년 한국마이크로소프트(유) 연구원
 2003년, 2006년~2007년 한국외국어대학교 영어학부 강사
 2008년~현 재 한국외국어대학교 영어대학 영어학과 조교수
 관심분야 : 자연어처리, 전산 언어학, 코퍼스 언어학, 형태론



이 공 주

e-mail : kjoolee@cnu.ac.kr

1992년 서강대학교 전자계산학과(학사)
 1994년 한국과학기술원 전산학과(공학석사)
 1998년 한국과학기술원 전산학과(공학박사)
 1998년~2003년 한국마이크로소프트(유) 연구원

2003년 이화여자대학교 컴퓨터학과 대우전임강사
 2004년 경인여자대학 전산정보과 전임강사
 2005년~현 재 충남대학교 전기정보통신공학부 부교수
 관심분야 : 자연언어처리, 자연어인터페이스, 기계번역, 정보검색