

<인지과학의 응용 및 실용성 분야>

한국어 철자/문법 검사기와 웹 기반 언어 학습시스템의 통합 환경 구축

남현숙*, 김상훈**, 김지원**, 권현주***, 정유진****, 권혁철*

*부산대학교 정보컴퓨터공학과 인공지능연구실, **(주)나라인포테크
부산 광무 여자 중학교, *부산대학교 컴퓨터 및 정보 통신 연구소

*hsnam@solge.cs.pusan.ac.kr

Integration of Korean Spelling/Grammar Checker with Web-Based Language Learning System

HyeonSook Nam, SangHun Kim, Jiwon Kim, HyunJoo Kwon, HyuckChul Kwon
AI Lab., Pusan National university, Nara Infotech,

요약

'바른 우리글 쓰기' 학습시스템은 우리말 작문 학습에 필요한 지식을 익히도록 하고자 만들었다. 이 논문의 궁극적인 목표는 한국어 철자/문법 검사기와 국어 학습시스템을 연계하여 학습자가 적극적으로 학습에 참여하도록 동기를 부여하는 능동적인 언어 학습시스템의 구현이다. 흔히 웹에 기반한 학습시스템에서 학습자는 일방적으로 정보를 제공받거나 독립적인 위치에서 자율적인 학습을 해나가야 하므로 끊임없는 흥미와 관심이 없이는 효율적인 성과를 거두기 어렵다. 이와 같은 약점을 보완하려는 맥락에서 이 학습 시스템은 학습자가 스스로 학습한 후 그 결과를 평가 문제나 한국어 철자/문법 검사기를 통해 즉각적으로 확인해 보고 부족한 내용을 공부할 수 있도록 한다.

1. 서론

교육용 웹사이트는 단순한 정보를 제공하는 사이트와는 달리 정보를 제공함에 있어서 효과적이며 정확하게 전달하기 위해서 지식을 가공하며 적절하게 난이도를 조절해야 한다[2]. 같은 맥락에서 초기 학습시스템이 학습 정보를 정확하게 구성하고 전달하는 데에 초점을 두었다면, 오늘날 대부분 학습사이트는 학습자의 수준과 관심 정도를 고려하여 학습자가 비루하지 않도록 미적인 요소를 가미하는 데에 비중을 두고 있다. 그러나 웹 상에 무수하게 존재하는 수많은 학습사이트 중에서 하나의 사이트가 명맥을 유지해 나가면 학습사이트마다 본래의 목적과 고유의 독창성을 지니는 것이 무엇보다 중요하다.

'바른 우리글 쓰기' 학습시스템은 우리가 글을 쓸 때 필요한 우리 글에 대한 지식을 전달하려고 우리 나라 사람이 자주 틀리는 오류를 중심으로 예제에 기반한 글쓰기 학습을 진행해 나가는 데에 초점을 둔다. 이 학습시스

템은 한국어 철자/문법 검사기와 연계를 꾀함으로써 학습자의 자발적인 학습 활동을 유도하는, 자연어처리 시스템과 언어학습시스템이 통합했다는 점에서 다른 시스템과 구별된다. 학습 시스템과 한국어 철자/문법 검사기를 연결하는 데에 결정적인 역할을 하는 요소는 철자/문법 검사기의 도움말 사전이다. 도움말은 현재까지 약 8,500개가 구축되어 있다.

이 논문에서는 대용량 문서인 도움말 사전의 언어 학습시스템에 어떻게 활용하고 있으며 학습자가 쓴 문장에서 한국어 철자/문법 검사기가 찾아낸 글쓰기 오류에 대한 학습 내용을 어떻게 연결하고 있는 지에 대해 중점적으로 다룬다.

이 논문은 다음과 같이 전개된다.

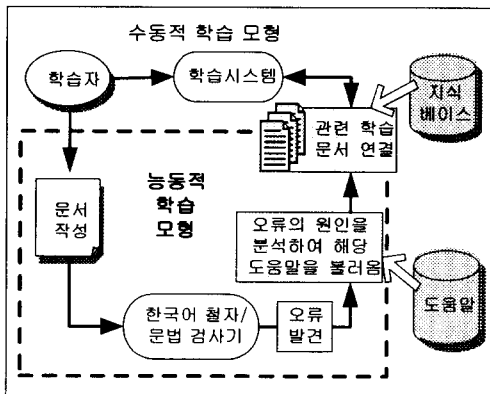
- ① 본 학습시스템에 구현한 수동적·능동적 학습 모형에 대해 소개한다.
- ② 도움말의 구성과 학습시스템의 메뉴 구

성을 비교 설명한다.

- ③ '바른 우리글 쓰기' 학습시스템의 전체 구조 속에서 학습시스템과 한국어 철자/문법 검사기와 도움말의 상호 작용을 예를 통해 설명한다.
- ④ 능동적 학습을 위해 앞으로 향상해 나아가 할 기술과 학습문서 구조와 그 전망에 대해 알아 본다.

2. 능동적 학습 모형 구현

바른 우리글 쓰기 학습시스템은 학습자가 원하는 학습 정보를 찾아서 학습하는 수동적 학습 모형과 우리말 철자/문법 검사기를 통해 학습자가 실력을 스스로 평가해 보고 문제를 해결하는 능동적 학습 모형을 동시에 만족하도록 설계한 시스템이다.



<그림 1 학습시스템 전체 구조>

한국어 철자/문법 검사기는 학습자가 어느 단계에서든지 접속하여 문장을 써 보고 그에 따른 결과를 즉시 확인해 보고 필요한 학습문서를 찾아가서 학습하도록 도와주는 일종의 자율평가시스템¹⁾이다.

이 논문에서 특히 능동적 학습 모형으로서 위 그림에서 점선으로 표시된 영역 내의 학습 줄거리를 다룬다.

¹ Self-testing systems can provide immediate, automatic, directed feedback to students enabling flexible self-assessment at any time during a course [Hartley & Medley 96, p.224].

학습자가 글을 쓸 때 필요한 우리말 지식을 먼저 알고자 한다면 학습 시스템에서 주어진 학습 줄거리를 따라 학습한다. 학습자가 능동적으로 글쓰기 학습에 참여하고자 한다면 학습시스템에 연결되어 있는 철자/문법 검사기에서 문장을 직접 써 본다. 학습자는 한국어 철자/문법 검사기를 통해 글쓰기 오류를 발견하여 수정하거나 그에 따른 도움말을 봄으로써 자신이 범한 오류에 대한 원인을 분석하고 연계되어 있는 학습시스템으로 들어가서 풀이를 보면서 새로운 지식을 습득한다. 이 때 학습자가 더 상세한 지식을 얻고자 한다면 도움말을 매개로 하여 학습시스템의 개별 학습 페이지에 들어가서 심화 학습을 할 수 있다.

3. 한국어 철자/문법 검사기와 학습시스템의 연계

3.1 웹 기반 한국어 철자/문법 검사기

인공지능연구실에서 수년간 개발해 온 한국어 철자/문법 검사기를 한국어 학습시스템에 활용하기 위해 웹 환경에 활용하였다. 이 철자/문법 검사기는 단순한 철자 오류뿐만 아니라 문장 단위에서 문맥상 틀린 단어를 찾아서 고쳐주는 기능을 한다. 다시 말해, 형태, 통사, 의미, 문체와 관련이 있는 다양한 오류를 검증하여 바른 대치어를 생성해 준다. 이러한 각 오류는 해당 도움말을 가지고 있다. 오류의 수만큼이나 다양하게 축적된 도움말 정보는 한국어 학습시스템에 백과사전적 지식을 제공해 준다.

이로써 한국어 철자/문법 검사기는 인터넷에 접속만 하면 누구든지 자신의 어휘력이나 문장력을 점검해 보고 관련 학습을 할 수 있게 되었다.

3.2 도움말과 학습시스템 메뉴 분류

한국어 철자/문법 검사기의 도움말은 우리나라 사람이 자주 틀리는 오류를 언어학적 원리(형태, 문법, 의미, 문체)와 언어정책적 원리(한글 맞춤법, 표준어 규정, 외래어 표기법)에 따라 분류하였다.

학습시스템의 내용은 언어 분석 단위에 따라 맞춤법, 어휘, 문장, 문단 학습의 체계로 짜여져 있다.

아래 표에서 보는 바와 같이 도움말이나 학습 시스템은 글을 쓰는 사람이 가장 필요로 하는 우리말 정보를 제공하기 위해 일반적으로 틀리기 쉬운 예를 중심으로 구성하였다. 도움말의 유형과 학습 시스템의 학습 메뉴는 내용과 구조 면에서 일관성을 지닌다.

<표 1. 도움말 분류 유형과 학습 메뉴>

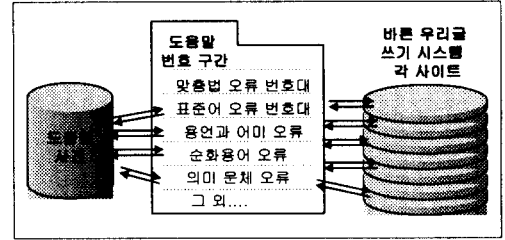
한국어 철자/문법 검사기 도움말 유형	‘바른 우리글 쓰기’ 학습시스템 메뉴
맞춤법 오류 띄어쓰기 오류	맞춤법 다지기 학습 전체 하위 메뉴
유사 발음 오류 표준어 오류, 순화 용어 오류, 외래어 표기 오류	어휘 다듬기 -> 발음이 비슷한 말, 자주 틀리 는 표준어, 외래어와 우 리말 순화
문법 관련 도움말: 용언과 어미, 체언과 조사 관련 도움말	문장 다듬기 -> 문장성 분 호응관계, 피동형과 사동형 문장
의미와 문체 관련 오류, 문장 구조 및 길이 관련 오류, 좋은 문장과 나쁜 문장 도움말	문장 다듬기 -> 문장성분 호응관계, 번역투 표현 문단 다듬기 -> 지시어의 바른 사용, 접속어의 바른 사용

본 학습시스템의 메뉴는 우리말 학습의 기본인 맞춤법 학습(‘맞춤법 다지기’)에서 ‘어휘 학습(‘어휘 다듬기’), 문장 학습(‘문장 다듬기’), 문단 학습(‘문단 다듬기’)으로 나아가는 구조를 취한다. 무한한 웹 공간에서 정보를 조직적으로 구성하여 체계적으로 전달하기 위해서는 잘 계획된 ‘라벨링 시스템(planned labeling system)²⁾을 취해야 한다. 그리하여 학습자가 제목만 보고도 학습문서와 내용을 미리 짐작하여 가상 공간 내에서 길을 잃지 않고 학습을 마치도록 도와준다.

²⁾ (...) users will often assume that consistent, systematic labeling systems do in fact the full extent of the domain that they represent. (...)After one quick look, the user understands how this system

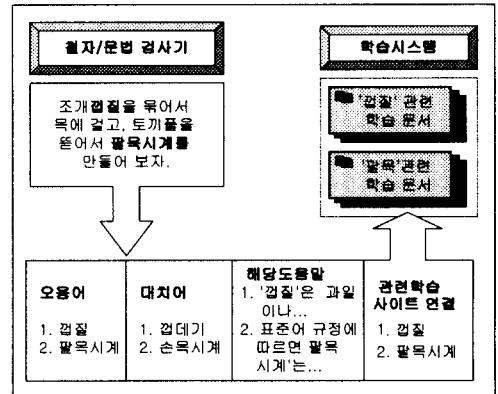
3.3 오류 유형에 따른 학습 문서 연결

도움말은 각 오류 유형별로 고유한 번호 구간에 연결되어 있으며 학습시스템의 각 학습 사이트는 URL을 가지고 있다.



<그림2. 도움말 사전과 학습사이트 연계>

한국어 철자/문법검사기를 통해 학습자가 입력한 문장에서 오류가 검색되면 해당하는 도움말을 제시하고 각 도움말 번호 구간에 일치하는 학습 문서의 URL을 지정하여 관련 학습 사이트가 지정된다. 학습자는 추천하는 사이트를 방문하여 필요한 학습 내용을 익힌다.



<그림 3 글쓰기 오류와 관련 학습문서 연결>

위의 그림에서 보는 바와 같이, 학습자가 ‘검질’과 ‘검데기’를 혼동함과 동시에 ‘손목시계’와 ‘팔목시계’ 중 어느 것이 바른 말인지 모르고 썼다고 가정하자. ‘검질과 ‘검데기’는 두 단어의 뜻이 비슷한 데서 생긴 오류이고, ‘손목시계’와 ‘팔목시계’는 현행 표준어 규정을 잘

works (...)[Rosenfeld & Morville 98, p.75].

물라서 범한 오류이다. 이 때, 한국어 철자 검사기는 의미가 유사한 단어와 표준어 규정 관련 도움말을 제시하며 이 도움말의 오류 유형에 적합한 학습시스템의 문서로 연결하여 관련 학습 정보를 제공한다. 여기서는 ‘겹질’과 ‘겹데기’와 유사한 의미 때문에 사람들이 혼동할 때 알아야 할 학습정보를 제공하는 ‘어휘 다듬기’의 ‘뜻 구별해야 할 말’의 사이트와 표준어 규정 관련 학습사이트를 연결해 준다.

어휘 단계의 오류 및 학습의 진행 과정과 마찬가지로 문장이나 문단과 관련이 있는 학습도 같은 방법으로 나아간다.

- 학습자가 쓴 문장에서 문장성분(주어, 목적어, 용언)이 서로 일치하지 않는다.
- 한국어 문법 검사기는 구문분석을 통해 문장의 통사적 오류를 발견한다.
- 오류에 대한 대치어를 생성하고 해당 도움말을 제시한다.
- 도움말과 관련이 있는 학습 문서를 도움말 번호 구간과 연결되어 있는 URL을 가지고 연결한다.
- 학습자는 학습시스템에서 있는 학습 과정을 필요한 내용을 익힌다.

그러나 학습자가 원하는 학습사이트를 더욱 상세하게 제공하려면 도움말 번호와 학습 사이트의 URL을 구분자로 하여 찾아 주는 방법만으로는 충분하지 않다. 이에 대한 해결책과 이 논문에서 나타나는 문제점은 결론에서 살펴본다.

4. 결론 및 향후 과제

이 논문에서는 한국어 철자/문법 검사기와 학습시스템을 연계하여 자연어처리시스템과 학습시스템이 통합한 시스템을 구현하는 단계를 소개하였다. 도움말 번호를 매개 단위로 학습 사이트의 URL을 연결하는 작업에서 앞으로 해야 할 과제가 많이 남아 있다.

학습자가 범한 글쓰기 오류를 더 세밀하게 분석하여 도움말 번호가 아닌 핵심단어를 구분자로 하여 더 구체적인 학습 사이트를 지정해 줄 수 있는 작업을 수행해야 한다.

가상학습환경에서는 교실학습에서처럼 교사가 모든 학습과정을 통제할 수 없으므로 학습자가 주도적으로 학습자료를 적절하게 이용할 줄 알아야 한다. 그러다보니 학습자가 스스로 학습하려는 굳은 의지가 없이는 쉽게 싫

증을 낼 수 있다. 따라서 잠재적인 학습자의 수준을 파악하여 지식 정도, 연령, 관심 정도에 맞는 학습 줄거리 구성이 시급하다.

직접적으로 접촉할 수 없는 학습자와 학습자, 학습자와 관리자가 자유롭게 토론을 벌이고 정보를 교환할 수 있도록 자동 FAQ와 같은 지능적인 게시판을 만들 수 있는 기술을 향상해야 한다.

이 세 가지 과제가 소위, 인공지능에 더 가까이 접근해 있는 ‘바른 우리글 쓰기’ 학습시스템을 구축하기 위해 실현해야 할 일로 남는다.

참고 문헌

- 1) 신수범, 유인환, 이철현, 컴퓨터 교육 내용에 따른 원격 컴퓨터 교육 모형 탐색, 한국정보과학회, 1999 가을 학술발표 논문집, pp. 685-687
- 2) 유인환, 신수범, 이철현, 이태욱, 자기주도적 학습을 위한 웹 기반 교육환경의 평가 기준 방안, 한국정보과학회'99 봄 학술발표논문집 vol.26 n.1 pp. 697-699
- 3) A.D. Marshall and S.Hurley, Interactive Hypermedia for the World Wide Web, 1996 ACM, pp. 1-5
- 4) D.Jones, Computing by Distance Education: Problems and Solutions, 1996 ACM, pp. 139-146
- 5) L.Losenfeld & P.Morville, Information Architecture for the World Wide Web, O'REILLY, 1998, USA.
- 6) Pamela B.Lawhead et al., The Web and Distance Learning: What is appropriate and what is not, Working Group Reports and Supplemental Proceedings, SIGCSE/SIGCUE IticSE '97, pp. 27-37
- 7) R.Brien, Apport des Sciences Cognitives A La Technologie Educative, Educa Technologiques Vol.1, 1993.2, Canada
- 8) S. Hartley & D. Medley, Enhancing Teaching using the Internet: Report of the Working Group on The World Wide Web as an Interactive Teaching Resource, 1996 ACM pp. 218-228