

□ 사례 발표 □

한글관련 제품 소개

김 두 현[†] 정 내 권^{††} 김 근^{†††} 권 강 현^{††††} 박 흥 호^{†††††}

◆ 목 차 ◆

- | | |
|------------------------|----------------|
| 1. 훈 글 | 3. 훈민정음 오피스 98 |
| 2. 마이크로소프트 워드 -한글의 세계화 | 4. 나모 웹에디터 |

본 고에서는 본 특집호의 주제인 한글공학과 관련하여 현재 상용화되어 시판되고 있는 제품 중 특히 문서 편집과 관련한 제품들로서 이들중 대표적이라 할 수 있는 훈글, 마이크로소프트 워드, 훈민정음, 나모 웹에디터에 대하여 소개한다. 본 고의 내용은 각 사의 관련자들로부터 원고를 넘겨받아 정리한 것으로 각 제품의 홍보보다는 한글공학적인 기능을 소개하는 데에 초점이 맞추어져 있다.

1. 훈 글

(한글과 컴퓨터 정 내권 이사)

1.1 개 요

워드프로세서에서 한글 처리 기술은 키보드를 통한 입력, 화면/프린터를 통한 출력이라는 원초적인 것에서부터 고도의 자연어처리까지 매우 다양한 분야가 존재한다. 워드프로세서는 다른 어떤 종류의 소프트웨어보다 한글 처리 기술의 혜택을 많이 누릴 수 있기 때문에, 한글 처리가 중요한 경쟁 요소로 인식되고 있다.

훈글은 그 첫 버전부터 한글의 기본적인 입출력을 넘어선 지능적인 처리에 지속적인 관심을 기울여 왔는데, 크게 한글 처리의 근간이 되는 코드 체계와, 지능적인 문자열 처리, 자연어 처리의 세 가지 계층으로 구분해 살펴보기로 한다.

1.2 훈글의 코드 체계

훈글은 독자적인 코드 체계를 기반으로 구현되어 있다. 현재의 코드 체계가 확립된 시점은 도스용 훈글 2.0이 발표된 92년이다. 훈글 2.0은 설계 단계에서부터 국학 연구 분야에 대한 지원을 가장 중요한 기능 중 하나로 고려하였는데, 당시는 물론 현재까지도 이러한 요구 조건을 충족할 수 있는 코드 체계가 존재하지 않기 때문에 독자적인 코드 체계를 정의하는 것이 불가피하였다.

훈글의 코드 체계는 표현 가능한 모든 현대 한글, 발견된 문헌에 등장한 모든 옛글자, 특수한 분야에 사용되는 매우 낮은 빈도의 한자도 모두 지원하는 것을 목표로 하였다. 또한 세계의 모든 문자를 한 문서에서 동시 표현하기 위해 2바이트의 일정한 크기로 모든 문자를 표현한다. 기본 구조에 있어 유니코드와 매우 흡사함을 알 수 있을 것이다.

하나의 소프트웨어 엔진에서 모든 나라의 문자를 동시 처리해야 할 필요성이 늘어감에 따라, 또 윈도우즈 NT의 커널이 유니코드를 기반으로 구

† 정회원 : 한국전자통신연구소 책임연구원

†† 정회원 : 한글과 컴퓨터 이사

††† 정회원 : 한국마이크로소프트 이사

†††† 정회원 : 삼성전자 부장

††††† 정회원 : 나모 인터랙티브 대표이사

현되어 있음을 감안할 때 유니코드가 아스키코드에 이어 다음 세대의 표준이 될 것임은 명백하다.

그러나 유니코드에는 현대 한글의 불완전 음절(초성+중성, 중성+중성과 같은 음절을 뜻하며, 발음 등을 설명하는 문장에서는 실제 쓰인다. 따라서 일반적으로 알고 있듯 현대 한글 음절의 수는 11,172자가 아니라 엄밀하게는 12,319자이다.)과 옛글자가 정의되어 있지 않은데, 옛글자의 경우 자모를 조합하여 n 개의 코드로 한 음절을 표현하는 스킴이 정의되어 있으나 구현의 복잡함으로 인해 실용성이 없다고 할 수 있다.

한자 역시 한글 코드보다 4000자 가량 많으나 한/중/일 3국에서 공통적으로 쓰이는 것을 위주로 선정되었기 때문에 한글 코드에는 있으나 유니코드에 없는 한자가 800여 자이다. 이는 뒤집어 말하면 중국, 일본에서는 쓰지 않는, 우리만 쓰는 한자가 800여 자가 있다는 뜻이 된다.

이처럼 유니코드에 없는 문자들이 국학 연구 분야에서는 매우 중요한 의미를 가지는데, 한글의 차기 버전에서는 국제 표준의 흐름에 맞춰 유니코드를 채용할 것이지만, 없는 문자에 대해서는 사용자 정의 영역 등의 방법을 통해 보완할 것이다.

1.3 지능적인 문자열 처리

마땅한 용어가 떠오르지 않아 “지능적인 문자열 처리”라고 이름을 붙인 기능들로써는 한영 자동 전환, 빠른 교정, 검색/치환에서의 조사 변환 등을 들 수 있다.

한영 자동 전환은 한글과 영문이 섞인 문장을 한/영 전환키를 누르지 않고 한 흐름에 입력할 수 있도록 입력되는 문자의 시퀀스를 통해 자판 상태를 자동으로 판단하는 기능을 뜻한다. 예를 들어 “gksmf teamdms”이라고 타이핑했을 때 “한글 team은”이라고 입력되는데, 이때 한/영 키를 전혀 누르지 않는다.

빠른 교정은 틀리기 쉬운 유형과 그에 대한 올바른 유형을 테이블에 등록해 놓고 이후 그와 같은 입력 패턴이 발견되면 자동으로 교정하는 기능을 말한다. 예를 들어

“ㄴ수있다” (오류) -> “ㄴ 수 있다” (교정)

과 같이 등록해 놓으면 “먹을수있다”라고 입력했을 때 “먹을 수 있다”로 자동으로 교정된다.

검색/치환에서 조사 변환은 받침 유무에 따라 조사가 변화하는 우리말의 특성을 고려한 기능으로, 예를 들어 문서 중 ‘식품’을 찾아 ‘먹거리’로 일괄 변환할 경우 다음과 같이 조사 부분도 자동으로 변환하는 기능을 뜻한다.

식품이 -> 먹거리가

식품은 -> 먹거리는

지금까지 살펴본 기능은 지능적인 문자열 처리로 분류될만한, 상상할 수 있는 많은 것 중에서 한글에 구현되어 있는 몇 가지를 살펴본 것이며 다양한 응용이 가능하다.

이 분야의 관건은 결국 얼마나 많은 경우에 대한 세심한 배려를 하느냐로 결론지을 수 있는데, 예를 들어 “자판틀(을) 선택하세요”와 같은 메시지를 내는 소프트웨어는 무성의한 느낌을 준다. 소스코드에는 “%s를(을) 선택하세요”로 되어 있을 테니 프로그래머는 못 느낄 것이다.

지능적인 문자열 처리는 종류가 많고 난이도가 다양하기 때문에 자연어 처리와 경계가 모호한 경우가 많다. 굳이 분리한 것은 필자의 주관적 견해에 따른 것일 뿐임을 다시 밝힌다.

1.4 자연어 처리

아직은 국내 상용 소프트웨어의 자연어 처리 기술은 맞춤법 검사에 한정되어 있고, 그나마도 완성도에 있어서도 큰 편차를 보일 정도로 성숙하지 않았다. 그러나 이 분야가 향후 한글 처리의 핵심이라고 할 수 있다.

맞춤법 검사기가 안고 있는 태생적인 한계는

맞는 내용을 틀렸다고 오인하는 것과, 틀린 내용을 감지하지 못하고 지나치는 두 가지로 분류할 수 있는데, 두가지는 서로 상충되는 부분이 있기 때문에 어느 선에서의 타협이 필요하다. 근본적인 해결책은 분석 대상을 문장 전체로 넓혀 컨텍스트를 완전하게 고려하는 것이다.

한글의 맞춤법 검사 기능은 오랜 시간동안 축적된 필드의 피드백을 바탕으로, 다양한 휴리스틱을 적용하여 이 문제를 최소화하고 있다.

2.5 앞으로의 발전 방향

앞서 강조한 대로 향후 한글 처리의 핵심은 자연어 처리 분야이며, 한글은 이 분야의 기능 강화에 주력할 것이다. 맞춤법 검사의 범위를 문장 단위로 넓힌 문장 검사 기능, “문서의 제목만 골라 별도의 파일로 만들어라”와 같은 자연어 문장을 인식하여 해당하는 동작을 수행하는 자연어 명령어 인식 기능 등을 현재 개발 중에 있다. 또한 이러한 기능을 컴포넌트화 함으로써 한글뿐만이 아닌 필요한 모든 소프트웨어가 우수한 자연어 처리 엔진의 혜택을 받을 수 있도록 할 것이다.

2. 마이크로소프트 워드

- 한글의 세계화 -

(한국마이크로소프트 김 근 이사)

2.1 개요

한국 마이크로소프트는 자체 연구 개발 인력 70여명과 함께 새로운 한글 공학 관련 기술 개발에 주력하고 있다. 특히 한글 공학팀에서는 필요한 한글 기술을 개발하고 있으며, 한글 엔진을 개발하여 마이크로소프트의 각종 제품에 탑재하고 있다. 워드프로세서는 우리의 글을 표현하는 가장 기본적인 도구이기 때문에 매우 신중한 노력과 개발인원이 투자되고 있다. 마이크로소프트의 한

글 공학 관련 기능들을 살펴본다.

2.2 한글 공학 관련 기술

2.2.1 한글 맞춤법 검사기

일반 사용자들도 쉽게 사용하고 있는 맞춤법 검사 기능은 매우 기본적인 한글 공학의 기반 기술이다. 마이크로소프트는 한글 윈도우 3.1 개발 이후, 한글 맞춤법 검사기능 개발에 노력하였으며, 1994년 마이크로소프트 워드 5.0에서부터 맞춤법 검사기능을 제공하기 시작하였다. 이 맞춤법 검사기는 한글을 연구하는 학자들과 마이크로소프트의 개발자들이 함께 노력하여 개발하였으며 현재 한글 공학팀의 표준 기술로서 마이크로소프트의 응용 프로그램에 제공되고 있다. 마이크로소프트 워드97에서는 윈도우95의 멀티태스킹 기능을 이용하여 사용자가 입력하는 과정을 두 개의 프로세스로 개발하여 사용자가 입력하는 순간에 자동적으로 맞춤법 검사가 이루어지도록 백그라운드에서 맞춤법을 검사하는 기술을 개발하여 한글 맞춤법 검사의 새로운 장을 연 바 있다. 다른 마이크로소프트의 기술들이 소프트웨어 개발 업계에 영향을 미친 것처럼, 마이크로소프트의 한글 공학팀에서 개발한 이러한 맞춤법 검사 기능은 다른 국내 업체들에게도 유사한 기능을 개발하도록 유도하게 되었다.

이미 전문지인 마이크로소프트웨어는 1996년 6월호 기사에서 마이크로소프트의 한글 맞춤법 검사기가 가장 정확하며 안정적이라는 기사를 게재한 바 있으며 국내 워드프로세서 업체들과 경쟁을 통하여 보다 우수한 한글 공학 기술 개발에 주력하고 있다.

2.2.2 찾기과 바꾸기 기능

워드프로세서에서 찾기와 바꾸기 기능은 매우 편리한 기능이다. 하지만 바뀐 명사에 따라서 조사 역시 변경되어야 하기 때문에, 조사를 고려하지 않았던 예전 워드프로세서 사용자들은 문서가

인쇄된 다음, 틀린 조사를 발견하여 곤란을 겪은 경우가 많았다. 그러나 마이크로소프트 워드 97에서는 이러한 문제점을 해결한 조사 자동 변환 기능을 제공해 주고 있다. 이 기능 역시 우리말의 실사와 허사를 구분하는 기술을 기반으로 구현되고 있다.

2.2.3 문서 찾기 기능

마이크로소프트 오피스의 모든 제품에는 문서 찾기 기능이 제공되고 있다. 이러한 문서 찾기는 미리 그 색인을 작성해 놓으면 그 검색 속도를 더욱 빠르게 할 수 있다. 이 기술을 빨리 찾기(FindFast)라고 부르고 있는데, 색인을 작성할 때 실사와 허사를 구별하는 기술 구현에 의해서 필요한 명사들을 주된 핵심 색인 단어로 추출하여 보다 빠른 검색을 가능케 하고 있다. 이 기술은 Index Server라는 Windows NT용 웹 검색 서버에도 제공되고 있어 웹 서버 사용자들이 무료로 사용할 수 있다.

2.2.4 글꼴 관련 기술

마이크로소프트는 윈도우용 운영체제를 개발하면서 필요한 글꼴을 국내 개발 업체들과 함께 검토해 왔으며, 한글 글꼴 기술들을 공개한 바 있다. 이 트루타입 폰트 기술은 윈도우용 인쇄 시장을 개척하고 있으며, 윈도우용 전자 출판용 소프트웨어가 계속 발전하고 사용될 수 있는 기반 기술이 되고 있다. 트루타입 폰트는 기존의 특정 소프트웨어에서만 사용하기 위해서 몇백 메가의 공간을 점유하는 글꼴과는 달리 윈도우용 모든 소프트웨어에서 공통적으로 사용할 수 있기 때문에 사용자들의 편의성을 도모할 수 있다.

2.2.5 유니코드 지원

90년대 초반의 컴퓨터 환경의 주제는 어떠한 코드를 사용하는가라는 점이었다. 국가전산망에서 이미 사용되고 있는 Unix의 제어문자 영역인 C1 영역을 피해서 설정된 한글 2,350글자의 완성형과 이를 고려하지 않고, 순수한 표현 가능한 한글의

조합인 11,172글자를 표현할 수 있는 조합형 가운데 어떤 것이 좋고 나쁜가 또는 기술이 문화에 미치는 영향에 대한 토론이 이루어졌었다. 마이크로소프트는 조합 가능한 모든 한글인 11,172글자와 고어 자모음을 표현하면서도 컴퓨터 환경에서 효과적으로 한글을 사용할 수 있으며, 정부에서 표준으로 승인한 Unicode 2.0을 표준 코드로 운영체계에 탑재하고 있다. 단일언어로서 최대 영역을 확보한 Unicode 2.0 코드는 우리가 사용하고 있는 프로그램들안에 이미 사용되고 있으며, 사용자들이 새로운 기술을 배울 필요없이 자연스럽게 적용되고 있다. Microsoft Windows 98/NT, Microsoft Internet Explorer, Netscape Navigator, Sun Java, Microsoft Office 97(마이크로소프트 워드 97 포함) 제품은 이미 Unicode 2.0을 내부적으로 사용하고 있다. Unicode의 장점은 한글 처리의 자연스러움과 한글의 세계화이다.

조합 가능한 모든 한글이 하나의 테이블로 구현됨에 따라서 실사와 허사를 구분하여 역순 사전을 제작하거나, 허사 구분을 통하여 필기체 인식, 음성 인식 기술의 기초를 마련한다는 장점이 있다. 또한 한글 영역이 전 세계 모든 프로그램에 공통적으로 적용되게 됨에 따라서, 한글 문서를 전 세계로 보내거나 국내에서 제작된 우수한 프로그램을 쉽게 전 세계로 수출할 수 있는 기반이 마련된 것이다.

2.2.6 오피스 길잡이(도움말 기능)

컴퓨터 사용자들이 가장 어려움을 겪는 부분은 자세한 사용 방법이다. 이를 위해서 마이크로소프트는 가정교사가 직접 질문하고 답변하는 형식으로 가능 설명을 제공하는 오피스 길잡이 기술을 제공하고 있다. 이 기술은 우리말로 입력된 각종 문장들을 분석하여 사용자가 어떠한 내용을 원하는지 파악하여 적절한 도움말을 추천한다. 가로방향 문서 인쇄는 어떻게 하니? 라고 직접 입력하면 해당 도움말을 보여준다. 추후 대화 응답식

컴퓨터에서 구현될 기본적인 기능을 문자 응답 방식으로 구현하고 있다.

2.3 향후 전망/비전

2.3.1 최고의 한글 워드프로세서

1999년에 발표될 마이크로소프트 워드2000은 최고의 한글 워드프로세서로서, 표현 가능한 고어 130여 만자와 표현 가능한 한글 11,172 글자를 모두 표현할 수 있도록 개발 되고 있다. 복잡한 메뉴를 어려워 하는 사용자를 위해서 간단한 메뉴와 전문가를 위한 메뉴를 한꺼번에 사용할 수 있도록 지원하고 있으며, 기존의 워드프로세서 사용자를 위하여 완벽한 문서 변환기와 도움말 프로그램 개발에 도전하고 있다. 마이크로소프트는 워드2000이 최고의 한글 워드프로세서의 위치를 담당할 수 있도록 개발 노력을 계속할 것이다.

2.3.2 다국어 지원 환경

Unicode의 지원을 통하여 마이크로소프트 워드 2000에서는 영문 윈도우 시스템이나 일본어 윈도우 시스템에서도 한글을 입력할 수 있도록 개발되며, 메뉴와 도움말 또한 원하는 언어로 설정할 수 있게 되어 있다. 따라서 전 세계 어디에서나 한글 문서를 입력할 수 있으며, 한글 문서를 전 세계 어디로 보낼 수 있게 되었다. 즉 한글의 세계화를 통하여 한글의 편리함과 장점을 알릴 수 있다.

2.3.3 지속적인 한글 공학 연구 투자

마이크로소프트는 Windows CE처럼 다양한 플랫폼에서도 한글을 보다 쉽게 입력할 수 있도록 웬 인식 기술 개발에 투자하고 있다. 특히 미래 컴퓨터 환경에 가장 큰 영향을 미치게 될 음성 인식과 대화형 시스템 개발을 위하여 막대한 개발 비용을 투자하고 있으며 지속적인 노력을 경주하고 있다. 문화와 예술은 세월이 흘러도 그 가치는 언제나 불변하지만, 기술의 세계는 결코 그렇지 않다. 몇 년 전의 컴퓨터 기술은 곧 퇴보를 의미하는 것이다. 마이크로소프트는 편리한 한글

구현 환경을 제공하고 구현하여 개발자들이 한글 문제가 아니라 진정한 소프트웨어 개발에 전념할 수 있도록 도구와 환경을 제공하고 있다.

3. 훈민정음 오피스 98

(삼성전자 권 강현 부장)

3.1 개요

훈민정음은 1992년 10월 9일 한글날 국내 최초로 발표된 윈도우용 한글 워드프로세서이다. 훈민정음은 윈도우를 기반으로한 스타일 가이드 채용 및 타 응용 프로그램과의 자원 공유등 윈도우 표준 환경 지원을 모토로 개발하고 있으며 조합형 한글 및 다국어 지원을 위한 기반 구조도 갖추어 놓고 있다. 문서 호환성 부분을 강화하여 MS-워드, 아래아 한글, 행망 문서, HTML, RTF등 대부분의 문서를 데이터 손실없이 훈민정음 문서 내로 읽어주고 있으며, MS-워드, HTML, RTF 등으로 저장할 수 있다. 차기 버전에서는 아래아 한글 문서 형식으로 저장까지 할 예정이다.

3.2 훈민정음에서의 한글처리

현재 훈민정음 7.0에서는 모든 현대 한글(11,172자)을 표현할 수 있으며, 16,000여자의 확장 한자와 KS-5657에서 규정하고 있는 한글 고어(1,677자)를 지원하고 있다. 내부적으로 유니코드, 조합형 코드 등 다양한 코드도 접목되도록 설계되어 있다. 일반적인 한글 처리인 한글 맞춤법, 한영 오타 자동 변환, 자동 교정 기능 등을 제공하고있으며, 타 제품과 차별되는 기능에 대해 좀더 상세하게 설명하기로 한다.

훈민정음에서의 한국어 정보 처리 기능은 요소 기술과 응용 기술로 구분한다.

요소 기술로는 한국어 형태소 분석기와 한국어 으뜸꼴 인식기이며, 응용 기술로는 으뜸꼴 찾고 바꾸기가 제공되고 있으며, 차기 버전에서 제공될 한국어 문헌 용례 색인 시스템, 한국어 전자사전,

한국어 문법/맞춤법 검사기, 한국어 문서 자동 요약 시스템 등이 있다.

3.2.1 형태소 분석기

가장 기본적인 요소인 형태소 분석기는 단어에 대한 형태소를 해석하는 것으로 형태소간의 결합 관계를 검사하여 각각 형태소로 분석한 후 분석된 결과를 가지고 여러 응용 분야에서 활용한다.

3.2.2 으뜸꼴 인식기

으뜸꼴 인식기는 문자열이 아닌 형태소 단위로 단어를 인식하는 것으로 조사 결합에서 체언을 해석하고, 어미 결합에서 용언 기본형을 분석하는 시스템이다. 이것을 이용하면 기존의 문자열을 찾는 방식으로는 찾기 어려운 용언 활용형을 아주 효과적으로 탐색할 수 있다.

예) 학교에, 학교, 학교를 => 으뜸꼴: [학교]
 예) 예쁜, 예뻐서, 예쁩니다, 예뻤다 => 으뜸꼴: [예쁘다]

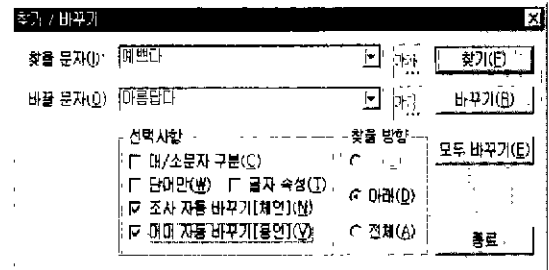
훈민정음에서 유일하게 처리하는 으뜸꼴 찾기/바꾸기는 일반적인 문자열 탐색으로는 처리하기 어려운 용언 활용형을 효과적으로 찾을 수 있으며, 바꿀 단어의 활용형에 적절한 형태로 교정하여 변환시킨다. [조사 자동 바꾸기]는 조사와 결합한 체언을 바꿀 경우 체언에 결합한 조사의 당-홀소리 결합 관계를 검사하여 적합한 형태로 바꾼다. [어미 자동 바꾸기]는 어미와 결합한 용언을 불규칙 활용형, 당-홀소리 제약, 모음조화 제약, 동사-형용사 결합 제한 등등 여러 가지 형태적인 결합 관계를 적합한 형태로 바꾼다. 또한 시제 형태, 존대 형태, 종결 형태, 모음조화, 당-홀소리, 용언 어간 변격 관계 등을 교정할 수 있다.

예) 으뜸꼴 [예쁘다]를 찾기하면 예쁜, 예뻐서, 예쁩니다, 예뻤다 등을 찾을 수 있다.
 예) [조사 자동 교정]: 학교를 대학으로 바꾸면 학교를 => 대학을로 바뀌 준다.
 예) [어미 자동 교정]: 예쁜, 예뻐서, 예쁩니다 의 으뜸꼴인 예쁘다를 =>[아름답다]로 찾

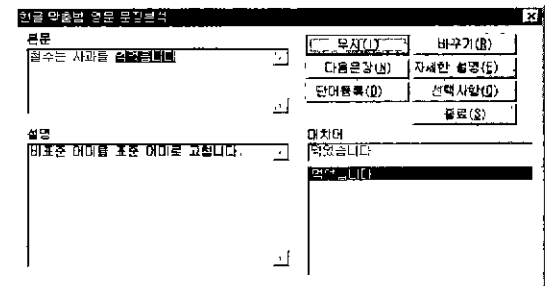
기/바꾸기하면 예쁜은 아름다운 으로 예뻐서는 아름다워서, 예쁩니다는 아름답습니다로 자동으로 바뀌어 준다.

3.2.3 맞춤법 검사기

훈민정음 맞춤법 검사기에서는 맞춤법 오류에 대한 학습을 할 수 있도록 맞춤법 내용이 제시해 준다. 한국어 맞춤법 검사기는 형태소 분석을 통하여 맞춤법을 검사하는 것으로 맞춤법 검사기와 맞춤법 학습기로 구성되었으며, 기존의 검사기에는 없는 검사기와 학습기를 연동시켜 교정 단어에 대한 맞춤법 내용을 학습할 수 있다.



(그림 1) 찾기/바꾸기 대화상자-으뜸꼴로 찾기,조사 자동 교정,어미 자동 교정



(그림 2) 맞춤법 검사기 화면 비표준 어미 오류

- 1) 표준어, 외래어, 사이 시옷, -이/-히 검사, 자판 사용에서 습관적인 오류 검사
- 2) 복합명사 띄기, 보조용언 띄기, 띄어쓰기, 반복되는 단어, 비한글(숫자, 한자, 영문자)과 조사결합 검사

- 3) 컴퓨터 용어, 시사 용어, 행정 용어 순화, 일본어투 용어 순화, 생활 외래어 순화
- 4) 사용자 사전 관리 기능
- 5) 교정에 대한 맞춤법 학습기 도움말

교정 예1) 표준어 처리
 [교정 단어] 먹었습니다
 [교정 해설] 비표준 어미를 표준어미로 고칩니다.[오류 단어] 먹었습니다

교정 예2) 순화용어 처리
 [오류 단어] 오뎡
 [교정 단어] 꼬치, 꼬치안주
 [교정 해설] (행정) 순화한 용어만 사용하십시오.

3.3 향후 훈민정음에서의 한글 정보 처리

98년말 발표될 훈민정음의 차기 버전(훈민 워드 98)은 맨 먼저 불완전 한글을 포함한 모든 한글을 표현할 수 있도록 개발중이다. 물론 고어 입력기도 제공될 예정이다. 그 후 모든 언어를 표현할 수 있는 다국어 버전을 출시할 예정이며, 이는 영문 OS 상에서도 한글 및 다국어를 자유자재로 표현할 수 있다. 여기에서는 한글 정보 처리에 관한 내용을 언급하고자 한다.

3.3.1 한국어 문법 및 맞춤법 검사기

한국어 문법 및 맞춤법 검사기에서는 맞춤법과 병행하여 문법 요소를 처리하는 동시에 관련된 문법 내용을 학습할 수 있도록 문법 학습기를 도움말 형태로 제공한다. 또한 교정에 대한 낱말 풀이나 사용 용례를 학습할 수 있도록 [한국어 전자사전]과 연동시켜서 학습 기능을 강화시켰다.

[바른글 모음집]에서는 혼동하여 틀리기 쉬운 단어나 문장을 도움말 형태로 제공하고, [순화용어 모음집]에서는 행정 용어, 생활 외래어, 일본어투 용어 등을 도움말 형태로 제공한다.

- 1) 이중 표준어 교정, 문장 부호 일치 검사, 혼동어 검사, 사용자 글쓰기 습관 검사
- 2) 교정에 대한 문법 학습기 도움말
- 3) 바른글 모음집
- 4) 순화 용어 모음집

3.3.2 한국어 전자사전

한국어 전자사전은 으뜸꼴 인식기와 문헌 용례 색인 시스템을 이용하여 제작한 것으로 표제어, 품사, 뜻풀이, 유사어, 관련어와 낱말 사용 용례에 대한 정보를 수록하였으며 맞춤법 검사기에서 필요할 때마다 사용하거나 그냥 실행하도록 되어 있다.

- 1) 낱말 사용 예제를 제시 - 실제로 사용하는 문학 작품에서 예제를 제시
- 2) 한국 단편 소설에서 뽑은 낱말 용례를 이용하여 낱말에 대한 표현법을 효과적으로 익힐 수 있다.
- 3) 용례 자세히 보기 기능을 지원하여, 필요한 작품 본문을 자세하게 볼 수 있다.
- 4) 학생 뿐만 아니라 작문을 자주 하는 사람에게도 저술용으로 매우 효과적이다.
- 5) 1994.12. 문서처리기의 기능에 대한 연구에서 김병선교수 지적한 것으로, 어떤 특정한 어휘의 쓰임세를 검토해 보거나 문체를 교정하려고 할 때 필요하다.

3.3.3 한국어 문서 자동 요약 시스템

한국어 문서 자동 요약 시스템은 주어진 문서에서 핵심적인 내용을 요약하는 것으로 문서 작업량이 많고 비슷한 종류의 문서가 많은 경우에 신속하게 문서 내용을 파악할 수 있어서 매우 효과적이다. 기본적으로 형태소 분석과 으뜸꼴 인식을 통하여 핵심어를 선별하고, 각각 문장과 문단의 중요도를 점수로 매겨서 주요 내용을 처리한다. 중요도를 매기는 방법은 핵심어의 빈도, 문장 길이, 문장에서 단어 수, 문단 길이, 문단에서

잡던 “워드스타”라는 워드프로세서도 글자 모양이나 문단 모양을 바꾸기 위해서 일일이 점을 찍고 시작하는 명령어를 본문 중간 중간에 삽입함으로써, 인쇄할 때 비로소 글자 모양과 문단 모양이 가미된 결과를 얻을 수 있었다. 상황은 조금 다르지만 지금 홈페이지 작성을 위하여 일반 사용자들에게 HTML 태그를 가르치는 일이 바로 이런 상황과 비슷하다. 하지만 좋은 비주얼(visual) 웹 에디터 프로그램을 사용하면, 태그를 전혀 모르더라도 훌륭한 웹 문서를 만들 수 있다.

HTML 편집의 반 이상에 표 편집 기능을 사용하는데, 복잡한 표 편집 기능은 비주얼 에디터를 사용하면 시간과 노력면에서 엄청난 작업량을 줄여준다. 이미지 맵 편집 기능도 비주얼 에디터가 아니면 사실상 불가능한 작업이다. 하지만 하나의 태그에 여러 단계의 속성(attribute)을 지정해야 하는 HTML 태그의 특성상 편리한 비주얼 에디터를 만드는 데는 어려움이 많이 따른다. 애초에 view-only인 웹 브라우저를 전제로 만들어지는 HTML 스펙 또한 위지웍 에디터로 구현하기 곤혹스러운 부분이 많은 것을 인정하지 않을 수 없다. 결국 이러한 난제들은 고도의 편집 엔진 기술과 사용자 인터페이스(user-interface) 고안으로 피해야 한다. 여기에 하루가 멀다하고 새롭게 선보이는 인터넷 신기술과 1년도 채 못 되어 규약을 개정하고 추가하는 W3C(월드와이드웹 협회)의 HTML 스펙을 발빠르게 구현해 나가야 하므로, 우수한 개발 능력과 경험을 겸비해야만 경쟁력 있는 비주얼 에디터를 계속 내놓을 수 있다.

4.2 나모 웹에디터의 한글 공학적인 특징

외국산 소프트웨어를 사용하다 보면, 본문에 한글을 입력하였을 때 깨져서 나오는 경우나 일부 한글이 깨지는 경우, 특정 대화 상자에서 글자 크기가 맞지 않아 넘쳐서 일부 기능을 사용할 수 없는 경우, 아주 편리한 기능이지만 한글은 사용

할 수 없는 경우 등을 자주 맞닥뜨리게 된다.

나모 웹에디터는 순수 국내 기술로 개발한 패키지 소프트웨어이므로, 메뉴나 대화 상자, 도움말, 설명서, 고객센터 등이 모두 우리말이다. 영문판과 일본어판도 함께 출시되어 있지만, 국내 환경에 가장 적합하게 개발된 제품이다. 입력과 동시에 잘못 쓰인 낱말을 바로 잡아 주는 자동 교정(auto-correct) 기능은 띄어쓰기와 조사/어미 감별, 표준말 등을 즉석에서 바로잡아 준다.

인코딩 정보가 잘못되어 한글이 완전히 깨져서 저장된 문서도 나모 웹에디터는 [도구 - 깨진 기호를 글자로] 메뉴를 선택하면 전체 내용을 즉각 복구해 준다. 외국에서 만든 웹 저작 도구는 인코딩 오류가 발생할 확률이 거의 없지만, 두 바이트 문자를 사용하는 나라에서는 자주 일어나는 치명적 문제이다. 그래서 아예 하나의 웹 사이트에 대하여 일괄적으로 인코딩 종류를 바로잡는 기능도 제공한다.

클립아트 갤러리나 새 문서 템플릿도 우리의 생활과 문서 습관에 맞는 내용을 미리 담고 있다. 서양과 우리의 문서 습관 중에서 가장 큰 차이가 “줄 간격”이라고 본다. 알파벳은 글자 위아래에 빈 공간이 있으므로 별도의 줄 간격이 없어도 가독성 있는 페이지가 된다. 하지만 한글은 모든 글자가 위아래 빈 공간 없이 꼭 차 있으므로, 별도의 줄 간격을 지정하지 않으면 가독성이 크게 떨어지는 페이지가 된다. 웹 브라우저로 볼 때 영문 페이지에 비하여 한글 페이지가 읽기에 답답한 것은, 바로 이 줄 간격이 없기 때문이다. 나모 웹에디터에서는 W3C에서 최근에 발표한 CSS(계층적 스타일 시트) 규약을 지원함으로써, 웹 페이지에 자유자재로 줄 간격을 지정할 수 있도록 했다. CSS 규약은 넷스케이프나 인터넷 익스플로러에서도 4.x판에서 비로소 제대로 지원하기 시작한 규약이며, MS 프론트페이지 98보다 먼저 지원하기 시작했으며, 나모 웹에디터가 더 충실하게 구현하고 있다.

4.3 나모 웹에디터의 계획

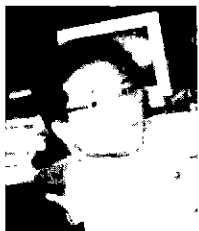
앞으로 나모 웹에디터가 계획하고 있는 새로운 기능들은, 현재 둘로 나뉘어져 있는 웹에디터와 사이트매니저를 하나로 합치는 일과 유니코드를 지원하고, 템플릿 위저드를 추가하며, 웹 페이지 작성에 꼭 필요한 그래픽 처리 기능을 추가하는 것이다. 즉 편집 기능은 더욱 강력하게 만들고, 편의 기능은 더욱 간편하게 구현함으로써, 인터넷 시대에 걸맞은 세계적인 웹에디터로 키우는 일이다. 우리의 목표는 우선 세계 2위의 웹 편집기다. MS가 차지하지 못하는 30 - 40%의 시장에서 1위를 하는 것이 일차 목표이다. 나모 웹에디터 2.1 판만으로도 세계 시장에서 알아주는 편리하고 강력한 홈페이지 저작 도구이며, 앞으로 더욱 향상된 모습으로 국산 패키지 소프트웨어에 대한 자부심을 드릴 수 있을 것으로 믿는다.



김 두 현

1985년 서울대학교 컴퓨터공학과 (학사)
 1987년 KAIST 전산학과 (석사)
 1987년-현재 한국전자통신연구소 책임연구원
 1991년-1993년 미 스탠포드연구소 객원 연구원

1993년 정보처리기술사 취득
 관심분야 : 분산 실시간 멀티미디어처리, 멀티미디어 시스템 성능평가, 멀티미디어 통신



정 내 권

아래아 한글 1.52, 2.0, 2.1, 2.5 개발 주역
 윈도우즈용 아래아 한글 3.0 개발 주역
 아래아 한글 프로 96 개발 주역
 아래아 한글 오피스 96 개발 주역

아래아 한글 일본판 개발 주역
 한컴홈 97 개발 주역
 한컴오피스 97 개발 주역
 현재 아래아 한글 차기버전(내부코드명 5.0) 개발 중
 99년 하반기 개발완료 예정 아래아 한글

김 근

1985년 서울대학교 사범대학 영어과
 1987년 The University of Chicago MBA 취득
 1987년-1990년 Samsung Hewlett Packard 영업담당

1990년-1992년 HP Asia Region Marketing Center Sales Development Manager
 1992년-1995년 HP Korea Computer System Organization Commercial Systems Marketing Manager
 1995년-1996년 HP Asia Region Financial Service Industry Business Unit Manager
 1996년-1997년 HP Korea Computer Systems Organization Marketing Director
 1997년-현재 (주)마이크로소프트 Marketing Unit 담당이사



권 광 현

1980년 한양대학교 전자공학과
 1998년 University of Illinois at Urbana-Champaign상경대학 MBA과정(휴학중)
 1980년- 삼성그룹입사 및 육군입대 (ROTC)

1982년- MIS 개발실 시스템연구원(IMS DB/DC)
 1984년- Application Systems 개발실 주임연구원(Unix-IBM Interface, System Integration)
 1989년- 워크스테이션 개발실 선임연구원(X Windows/Motif)
 1993년- 응용소프트웨어(패키지분야) 개발팀장/수석연구원(훈민정음 개발외)
 1998년- 비즈니스 소프트웨어 그룹장(개발영업 총괄부장)
 1980년-현재 삼성전자 미디어서비스사업팀 부장



박 홍 호

부산대학교 사범대학 국어교육과 부산남일고 교사
 한글 문화원 연구원
 (주)한글과컴퓨터 개발이사
 현재 (주)나모 인터랙티브 대표 이사

관심분야 : 한글공학, 자연어 처리, 업무용 응용 소프트웨어