

플랫폼독립적 환경에서 한글 입력 방법

신갑재^o, 이형직, 변정용

동국대학교 컴퓨터학전공

jjlove77@wonhyo.dongguk.ac.kr, xingse@mail.dongguk.ac.kr

byunjy@dongguk.ac.kr

A Korean Input Method in Platform_free Environment

Gab-Jae Shin^o, HyoungJick Lee, Jeongyong Byun

Dept. of Computer Science, Dongguk University at Gyeongju

요 약

현행 한글 입력방법은 운영체제 의존적 문자입력 방법을 채택하고 있다. 즉 운영체제가 채택하는 문자입력방식에 의존적이다. 특히 인터넷의 발전으로 세계 어느 곳에서나 인터넷을 통한 정보 검색이나 전자상거래와 같은 일을 할 수 있게 되었다. 정보검색이나 전자상거래와 같은 각종 한글 인터넷 응용은 전 세계 어느 곳에서나 이용 요구가 있다. 여기서 한글의 출력은 대부분 가능하지만 입력은 운영체제의 입력기(IME,AMI)에 의존적이어서 입력을 하려면 한글 입력기를 설치해야 한다. 또한 대부분 일반 이용자는 이를 설치하기 어렵고, 국외에서 이용하려면 플랫폼은 개인 휴대 시스템이 아닐 경우 입력기를 설치할 수 없을 것이다. 이를 편리하게 가능하도록 하기 위하여 해당 한글 웹응용이 직접 한글 입력기능을 제공하여 플랫폼에 의존하지 않고 독립적으로 한글 입력이 가능한 방안을 제시하고 이를 구현하여 가능성을 보였다. 본 논문은 플랫폼 독립적인 환경에서 인터넷을 통한 한글의 입력방법의 필요성과 그것이 가지는 한계점을 검토하고 응용프로그램과의 연동하는 한글 입력방법을 제안하고 구현하였다.

1. 서론

정보 처리 기술과 웹 서비스의 발달로 사람들은 인터넷을 통하여 정보획득을 하는 것이 보편화되었다. 특히 우리나라는 지난 6월 19일 정보통신부[8]가 발표한 자료에 따르면 우리나라 초고속인터넷 보급은 지난해 6월 말 현재 921만명으로 3월 말 878만명보다 43만명이 늘어난 것으로 집계됐다. 921만명의 가입자는 전국 1450만 가구 기준으로 약 64%고, 4700만 인구기준으로 약 20%에 해당하는 수치다. 또한 7월 24일자 한겨레신문[9]에는 한국의 인터넷 보급률은 지난해말 현재 51.5%로 미국(53.9%)에 비해 낮았으나 홍콩(39.3%), 일본(37.8%), 대만(35.1%), 싱가포르(29.9%)보다는 높다고 전

하고 있다. 위의 수치에서 말하듯이 이제 인터넷은 사람들 생활 깊숙이 자리 잡고 있다고 생각해야 될 것이다. 국내에서 인터넷 사용자가 이렇듯 국외에서도 한글사이트에 접속하여 여러 가지 업무를 처리하는 일도 빈번할 것이며, 인터넷을 사용하는 사용자의 대부분은 컴퓨터에 대한 전문적인 지식이 없을 것이다. 한글이 지원되지 않는 플랫폼이라 해도 한글 폰트가 운영체제에 설치되어 있다면 한글사이트를 검색 할 수는 있다. 그러나 정작 업무를 보기 위해 한글을 입력하려면 윈도우 시리즈 같은 경우에는 한글 'IME'를 다운받아서 설치하여야 하고 리눅스 같은 경우에는 'AMI'를 설치하여야 하는데 일반 사용자 입장에서 한글 입력기의 설치 상당히 까다롭다. 물론 개인 소유 PC 같은 경우에는 어떻게든 운영체제에 맞는 한글 입력기를 설치하겠지만,

공항이나 호텔 등과 같은 공공시설에 설치되어 있는 컴퓨터에 한글 입력기를 설치하고 한글을 입력한다는 것은 상당한 전문적인 지식을 요구하며, 입력기 설치 방법을 알고 있다고 해도 개인 소유가 아니기 때문에 맘대로 설치할 수 없다.

또한, 통계청[10]이 지난 6월 27일 발표한 '2001년 연간 전자상거래 통계조사 결과'에 따르면 거래주체별로 기업간(B2B)전자상거래가 108조9천460억원으로 91.6%를 차지했고 기업-정부간(B2G)거래가 7조370억원으로 5.9%, 기업-소비자간(B2C)거래가 2조5천800억원으로 2.2%였다. 위의 통계에서 말하듯이 인터넷을 통한 전자상거래 또한 이미 보편화되었으며 지금도 더욱 발전하고 있는 추세다. 전자상거래 역시 웹페이지에서 거래가 이루어지기 때문에 한글 입력은 운영체제에 한글 입력기를 설치하여야 된다. 그래서 현재 운영체제의 한글입력 지원 여부에 따라 위와 같은 한글 입력기 설치 문제가 발생한다. 그래서 인터넷사용과 전자상거래를 사용하는 컴퓨터에 대한 전문적인 지식이 없는 사용자를 위해서 기존 한글 입력기보다 편한 플랫폼 독립적인 한글입력방법을 제안하고 구현한다.

2. 기존 연구 현황

2.1 운영체제 문자 입력기(IME)

운영체제 문자 입력기는 운영체제에서 사용 가능한 모든 콤포넌트에서 문자 입력을 도와준다. 대표적으로 MS Windows 시리즈의 IME가 있는데, 이 IME는 운영체제에 덧붙여진 형태로 제작되어 있어서 대부분의 MS Windows환경의 응용 소프트웨어에서 작동한다. 또한 MS IME는 언어 확장팩을 다운로드해서 설치하는 방법으로 다국어 입력을 지원하고 있다. 또한 요즘 나온 Windows XP[11]에서는 설치할 때 다국어 부분을 체크하고 제어판에서 세계 각국의 언어를 선택하면 한글뿐만 아니라 세계 모든 나라의 언어를 대부분 입력할 수 있게 만들어져 있다. 하지만 아직 WindowsXP가 널리 보급되지 않아서 WindowsXP 사용자는 별로 없으며 대부분이 Windows98를 사용하고 있는 상황이다. 아직까지 Windows98은 언어 확장팩을 MS 홈페이지에서 다운로드받아야 하는 번거로움이 있다.

또한, 몇 년 전부터 오픈소스 정책으로 많은 각광을 받고 있는 리눅스와 전통적 대형 운영체제인 유닉스에서는 기본 GUI 환경인 X윈도우에서는 특정한 문자 입력기가 제공되지 않고 XIM[7]이라는 프로토콜을 규정

하여 필요한 나라의 문자입력기를 만들어 사용하게 하고 있다. XIM 프로토콜을 구현한 대표적인 입력기로는 'hanIM'과 'AMI'등이 있다. 하지만 XIM프로토콜은 ISO - 2022 규격의 1바이트 문자와 2바이트 문자만 지원하고 유니코드는 지원하지 못하여 자국어만 입력이 가능하다. 물론 자국어가 아닌 언어도 해당 언어의 입력기만 설치하면 입력은 가능하지만 설치하는데는 리눅스에 대한 해박한 지식이 있어야 하며 일반 사용자가 설치하기에는 거의 불가능하다.

2.2 웹에서의 한글 입력

현재 웹에서의 한글 입력은 운영체제에 완전히 의존적이다. 대부분의 검색사이트나 이메일 사이트는 플랫폼 독립적인 한글 입력을 지원하고 있지 않다. MS Windows 시리즈의 익스플로러나 리눅스와 유닉스의 X-Windows 에 많이 쓰이는 네스케이프 역시 특별한 입력기가 없으면 운영체제가 지원하지 않는 언어는 입력이 불가능하며 한글 또한 예외는 아니다.

3. 한글 입력시스템의 설계

3.1 요구 사항

요구되는 운영체제는 자바를 지원하는 모든 운영체제가 된다. 또한 운영체제 독립적인 실행환경을 제공하기 위하여 사용되는 언어인 JAVA2 JDK1.4는 현재 Sun Microsystems사에서 발표한 가장 최신의 자바 언어 개발 도구이다. 아직 운영체제의 웹브라우저가 JAVA2 환경을 제대로 지원하지 못하기 때문에 JDK1.4용 프로그램을 개발하고 실행하기 위해서는 반드시 JRE 1.4가 있어야 하며 유니코드 지원 한글 폰트가 설치되어야 한다. 한글 입력시스템의 개발에 대한 요구사항은 아래와 같다.

첫째, 기존 한글 입력방법 보다 한글 입력시 사용자의 노력을 최소화하여야 한다.

둘째, 본 입력기의 실행만으로 한글입력 및 처리가 가능하여야 된다.

셋째, 자바를 지원하는 모든 운영체제에서 동작을 보장한다.

넷째, 자바 애플릿으로 개발하여 어떤 웹브라우저에서도 동작을 보장한다.

다섯째, 유니코드를 사용하여 세계 모든 나라에서 응용 프로그램의 이용이 가능하게 한다.

여섯째, 본 시스템을 개발하면서 생성된 코드들은 소프

트웨어 공학적 측면에서 재사용성이 높아야 한다.

3.2 웹에서 한글 입력시스템 설계

웹에서 한글 입력을 하려면 실시간으로 입력한 코드를 각 언어에 해당하는 유니코드로 바꾸어야 하는데 이때 자바 애플릿을 사용하여야 한다. 또한 입력을 일일이 서버와 통신하면서 변환하면, 즉 서블릿으로 구현하면 서버에 많은 부하를 가져다 줄 수 있고 플랫폼 독립을 위해서 클라이언트의 소스는 사용하지 않기 위해서 자바 애플릿의 사용은 필수적이다. 자바애플릿을 사용하려면 HTML로 구현된 웹페이지에서 새로운 애플릿창을 띄워야 하는데, 이때 자바 스크립트의 사용이 필요하고 자바 스크립트와 자바 애플릿의 통신[6]이 필요하다.

3.2.1 자바스크립트에서 애플릿 호출

자바 애플릿과 자바스크립트는 웹브라우저에서 서로 함수 호출을 통하여 의사소통을 할 수 있다. 자바 스크립트는 웹 브라우저의 다른 컴포넌트를 제어하는 것과 같은 방법으로 애플릿을 제어 할 수 있다. 즉, <applet> 태그에서 name 에트리뷰트를 이용해서 애플릿의 함수를 호출할 수 있고 매개변수도 교환 가능하다.

3.2.2. 애플릿에서 자바스크립트 호출

자바 애플릿에서 자바 스크립트를 호출하는 방법은 URL과 JLObject를 사용하는 두 가지 방법이 있다. 여기서는 JLObject를 사용하는 방법을 사용하였다. JLObject의 call() 메소드를 이용하면 자바스크립트의 함수를 호출 할 수 있다.

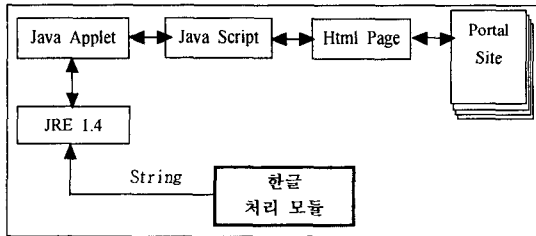


그림 1. 한글 입력 시스템의 개략적 설계

3.3 한글 처리 모듈 설계

한글이나 아랍의 문자들은 기본 문자 요소를 가지고 조합하여 하나의 문자를 만들어 사용하는 모아쓰기 방식을 택하고 있다. 아랍의 문자들은 로마자와 달리 하나

의 코드와 문자가 일 대 일로 대응되지 않는다. 따라서 문자를 조합하여 주는 조합 문자 입력기가 필요하다.

문자를 조합하여주는 소프트웨어를 FA(Finite Automata)라고 하는데, 이것은 언어 접수기(Language Acceptor)와 언어 변환기(Language Transducer)의 두 가지 기능을 가지고 있다. 언어 접수기는 현재 입력하는 문자가 하나의 음절로 인식될 때까지 상태를 파악하여 음절의 구성을 인식하는 부분이고, 언어 변환기는 접수기에서 인식된 음절을 해당코드로 변환하여 주는 역할을 담당한다. 이때 각각 필요한 것은 접수기에서 문자 조합 규칙이 되고, 변환기에서는 코드 변환을 위한 코드 테이블이 된다.

| |
|---|
| - 조성(19개) |
| {'r', 0x1100}, {'R', 0x1101}, {'s', 0x1102}, {'e', 0x1103}, {'E', 0x1104}, {'f', 0x1105}, {'a', 0x1106}, {'q', 0x1107}, {'Q', 0x1108}, {'i', 0x1109}, {'T', 0x110A}, {'d', 0x110B}, {'w', 0x110C}, {'W', 0x110D}, {'c', 0x110E}, {'z', 0x110F}, {'x', 0x1110}, {'v', 0x1111}, {'g', 0x1112} |
| - 중성(14개) |
| {'k', 0x1161}, {'o', 0x1162}, {'i', 0x1163}, {'O', 0x1164}, {'j', 0x1165}, {'p', 0x1166}, {'u', 0x1167}, {'P', 0x1168}, {'h', 0x1169}, {'y', 0x116D}, {'n', 0x116E}, {'b', 0x1172}, {'m', 0x1173}, {'l', 0x1175} |
| - 복중성(7개) |
| {'hk', 0x116A}, {'ho', 0x116B}, {'hl', 0x116C}, {'nj', 0x116F}, {'np', 0x1170}, {'nl', 0x1171}, {'ml', 0x1174} |
| - 중성(16개) |
| {'r', 0x11A8}, {'R', 0x11A9}, {'s', 0x11AB}, {'e', 0x11AE}, {'f', 0x11AF}, {'a', 0x11B7}, {'q', 0x11B8}, {'i', 0x11BA}, {'T', 0x11BB}, {'d', 0x11BC}, {'w', 0x11BD}, {'c', 0x11BE}, {'z', 0x11BF}, {'x', 0x11C0}, {'v', 0x11C1}, {'g', 0x11C2} |
| - 복중성(11개) |
| {'rt', 0x11AA}, {'sw', 0x11AC}, {'sg', 0x11AD}, {'fr', 0x11B0}, {'fa', 0x11B1}, {'fq', 0x11B2}, {'fr', 0x11B3}, {'fx', 0x11B4}, {'fv', 0x11B5}, {'fg', 0x11B6}, {'qr', 0x11B9} |

표 1. 로마자 입력에 따른 유니코드 기반 현대 한글 변환 테이블

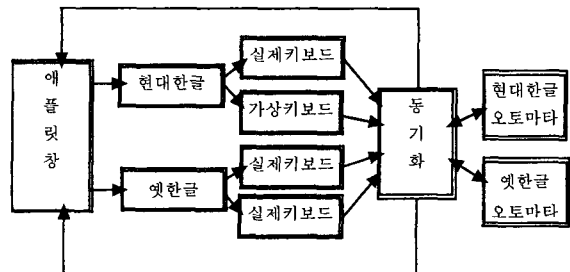


그림 2. 한글 처리 모듈의 세부 설계

4. 한글 입력시스템의 구현

컴퓨터가 영어권 나라에서 만들어 졌기 때문에 비 영어권 나라의 언어를 입력하기는 매우 까다롭다. 특히 한글 입력은 음소의 조합으로 음절을 이루기 때문에 모아쓰기를 하기 위해서는 오토마타[12]가 필요하다. 본 연구에서는 정음형[1] 오토마타를 사용하였다. 또한 현재 국내에서 사용자들이 널리 사용하고 있는 한글 2벌식 자판을 기준으로 구현하였으며 플랫폼 독립을 위해서 유니코드를 사용하였으며 현대 한글은 유니코드[3]의 완성형 11172자를 표현 할 수 있게 하였으며 옛한글은 유니코드 제 1 수준 구현에서는 조합하지 못하기 때문에 조합된 음절을 지원하지 못하고 있다.

4.1 자바 애플릿에서 이벤트 처리

애플릿을 사용하면 다국어 입력을 위해 구현된 모든 클래스 및 라이브러리가 다운로드 된다. 또한 다국어 입력을 하기 위해서는 이벤트에 대한 메소드의 사용이 필수적이다. 실제키보드 입력, 가상키보드 입력 및 여러 가지 입력에 대한 액션처리에 대한 이벤트를 등록하여야 된다.

자바 애플릿은 자바 AWT 이벤트의 사용을 지원하는데 AWT의 이벤트는 2가지 모델[6]이 있다. 여기서 2가지 모델의 비교 설명은 생각하고 본 시스템은 두 번째 모델인 Delegation 모델을 사용하였기 때문에 Delegation 모델만 간단히 설명하겠다.

Delegation 모델은 각 액션 이벤트에 대한 Listener를 Implements해야 한다. 또한 각 액션에 대해 매핑되는 메소드가 따라 정의되어 있으나 내부내용은 구현되어 있지 않기 때문에 사용자는 이 메소드들을 적절히 구현할 필요가 있다.

4.1.1 실제 키보드 이벤트 처리

한국어 입력 프로그램 애플릿 창이 활성화 된 상태에서 문자를 실제 키보드에서 입력하게 되면 keyPressed() 와 keyReleased ()와 keyTyped () 이 세가지 메소드가 키의 한번 눌렀다 띄는 것으로 호출이 되는데 여기서 다국어 입력기는 keyPressed() 메소드만 사용하였다. 키가 눌러지면 keyPressed() 메소드로 키이벤트가 전달 되는데 여기서 현재 눌러진 코드값은 getKeyCode메소드로 Key 값을 전달받을 수 있다. 이 Key 값은 현재 키보드에서 눌러진 로마자의 대문자 코드 값이 저장되어 있는

데 이 대문자 코드를 적절히 가공하여 한글 오토마타로 보내어 조합하게 된다.

```
public void keyPressed(KeyEvent e){
    int key=0;
    key = e.getKeyCode();
    .....
    key 값을 한글 오토마타에서
    조합하여 유니코드로 변환한다.
}
```

4.1.2 가상 키보드 이벤트 처리

한글 플랫폼 환경이 아닌 컴퓨터는 우리가 흔히 사용하고 있는 키보드 한글입력기(IME 2002)가 아닐 수 있다. 그렇기 때문에 가상키보드는 한글 자판에 속달되지 않는 사용자도 사용할 수 있게 하기 위해 필요하다.

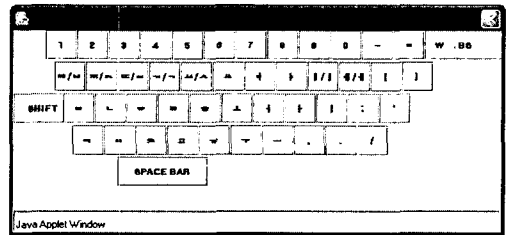


그림 3. 현대한글입력을 위한 가상 키보드

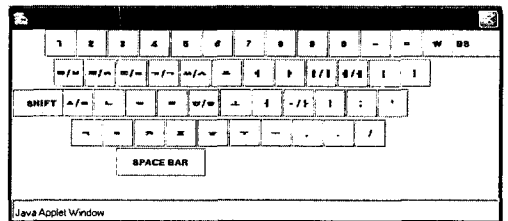


그림 4. 옛한글 입력을 위한 가상 키보드

가상 키보드는 자바 Swing의 JButton을 사용하는데 문자를 가상 키보드에서 입력을 하면 actionPerformed() 메소드가 호출되는데 현재 눌러진 버튼의 값이 getActionCommand() 메소드에 의해서 String으로 전달 받을 수 있다. 이 String 값에는 현재 눌러진 버튼의 값이 넘어오는데 이 값은 한글이기 때문에 한글배열과 비교하여 몇 번째 버튼이 눌렀는지를 검사하여 같은 순서에 있는 로마자배열에서 그 값을 찾아서 한글 오토마타로 보내어 조합하게 된다. 이때 눌러진 한글을 바로 조

합할 수 있으나 실제 키보드의 이벤트 처리와 동기화 시키기 위하여 다시 로마자로 바꾸어 오토마타로 넘겨 준다.

```
public void actionPerformed(ActionEvent ev){
    .....
    .....
    String command = ev.getActionCommand();
    s = applet.multi.handleCharacter(c,state);
    command 값을 한글 오토마타에서 조합하여
    유니코드로 변환한다.
}
```

5. 실행 결과

본 절에서는 다국어 문자 입력기 연구의 결과를 그림과 함께 설명한다.

다국어 입력기에서 운영체제의 사용언어와 관계없이 한국어를 사용할 수 있다. 다음 그림은 영문 윈도우에서 현대 한글을 입력할 수 있음을 보여준다.

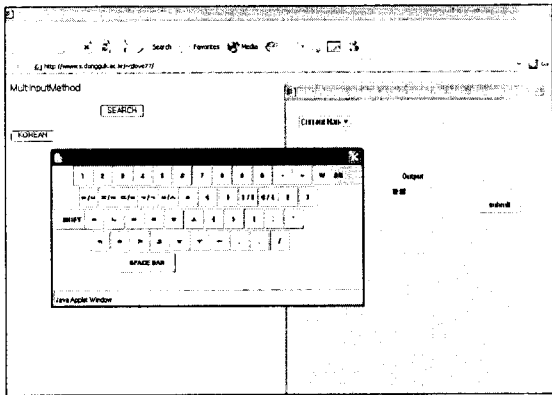


그림 5. 현대 한글 입력 실행 화면

다음 그림은 현대 한글이 입력된 후 옛 한글이 입력된 그림이다. 현대 한글을 입력하는 중에 옛 한글을 입력할 수 있기 때문에 현대한글과 옛 한글은 같은 버퍼로 동기화 시켜야 된다.

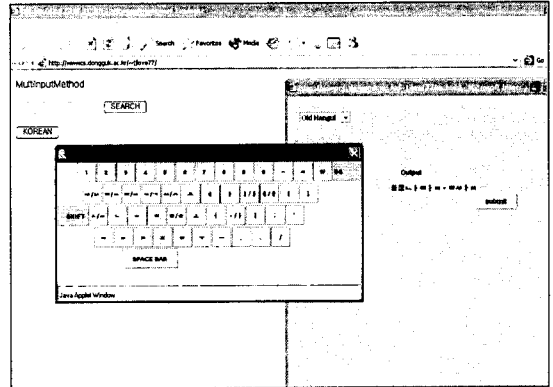


그림 6. 현대 한글 입력 후 옛한글 입력 실행 화면

위의 그림에서 Submit 버튼을 클릭하면 애플릿 창을 호출한 원래 HTML로 구성된 웹페이지로 현재 입력한 문자가 전달된다. 여기서 애플릿 창을 호출한 웹페이지가 'empas' 라고 가정하고 입력된 결과를 검색해 보았다.

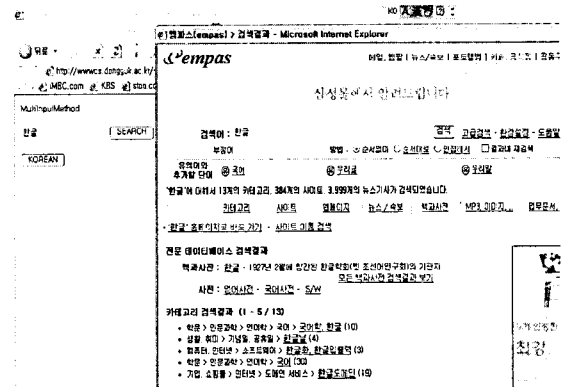


그림 7. <한글> 입력 후 'empas' 창에서 검색 한 실행화면

6. 결론

한글을 지원하지 않는 플랫폼 상에서 한글 웹응용에 한글을 입력하려면 한글 입력기가 필요한데 이 경우 대부분 한글 IME를 설치할 수 없는 시스템 상황이거나 또는 설치 능력이 없는 일반 사용자일 경우가 거의 대부분이다. 이에 대한 문제 해결을 위하여 매우 간편하고 쉽게 한글 입력을 할 수 있는 방안을 현재 보편적으로 많이 사용하고 있는 클라이언트 운영체제인 엠에스 윈도우즈류와 리눅스, 유닉스 등에 자바언어를 이용하여 구현하여 보았다.

기존 운영체제에 입력기를 설치하려면 설치 방법에 대

한 상당한 지식이 필요하지만 이것은 매우 간단하게 처리할 수 있다. 따라서 컴퓨터 지식이 거의 없는 일반사용자에게 본 연구결과는 매우 적합하다. 그리고 한글 검색기나 쇼핑몰 등을 한글을 지원하지 않는 시스템의 경우 자신의 컴퓨터가 아니기 때문에 한글 입력기(IME,AMI) 설치가 불가능할 것이므로 본 연구는 그러한 문제가 발생하지 않기 때문에 매우 적합하고 유용하다. 아래의 표는 실험한 운영체제와 웹브라우저에 대한 실험결과이다.

| 언어 환경 | 운영체제 | 웹브라우저 | |
|-------|------------|-------|-------|
| | | 익스플로러 | 네스케이프 |
| 영 문 | Windows 98 | ○ | ○ |
| | Windows XP | ○ | ○ |
| | Linux | - | ○ |
| 한 글 | Windows 98 | ○ | ○ |
| | Windows XP | ○ | ○ |
| | Linux | ○ | ○ |
| | Unix | - | ○ |

표 2. 본 연구에서 실험한 운영체제와 실험 결과

유니코드의 사용으로 유니코드가 지원하는 언어를 사용하는 세계 어떤 나라에서도 그 나라의 특성에 맞는 언어처리부분만 구현한다면 자국어 입력기를 쉽게 응용하여 구현 할 수 있게 하여 본 시스템의 재사용성을 높였으며 일반 사용자의 한글 입력을 위한 노력을 최소화하게 구현하였다. 리눅스와 윈도우즈 NT와 같은 운영체제에서는 계정이 없으면 해당 PC를 사용할 수 없는 문제점이 있지만, 게스트 계정만 등록해 놓으면 누구나 웹에서의 한글 입력이 가능하게 된다. 아직 이메일 통신은 각 클라이언트의 코드와 폰트문제 등으로 구현하지 못하였지만 위의 문제만 해결한다면 플랫폼 독립적인 이메일 송수신도 가능할 것이며 또한 현재 나날이 발전하고 있는 PDA 분야나 XML 분야로도 응용이 가능하여 연구를 계속 할 가치가 있다. 그러나 아직 웹에서의 HTML 태그의 폼에 일일이 애플릿을 링크해야 되는 문제가 남아있고 웹에서 밖에 입력되지 않는 한계가 있다. 또한 옛 한글의 입력은 유니코드 제 1 수준 구현에서는 조합하지 못하기 때문에 조합된 음절을 지원하지 못하고 있다.

6. 참고문헌

[1] 변정용, "한글정보처리를 위한 라이브러리 개발", 시스템공학연구소 1차년도 최종 보고서, 1995 "

[2] 홍윤표, "한글코드에 관한 연구", 1995"

[3] Unicode Group, "www.unicode.org"

[4] Ken Lunde, "CJKV Information Processing", O'REILLY, 1999

[5] 이현우,김형국,홍성민, "Java Programming Bible Ver.2" (주)영진출판사, 2000"

[6] 최재영,최종명,유재우 "프로그래머를 위한 Java 2™" 홍릉과학출판사, 2000

[7] ncadl.nca.or.kr "그래픽 사용자 인터페이스 한글 처리 표준 : XIM 프로토콜", 1995

[8] 정보통신부 "www.nic.go.kr"

[9] 인터넷 한겨레 신문 "www.hani.co.kr"

[10] 통계청 "www.nso.go.kr"

[11] 윈도우XP 안내 "www.micorsoft.com/piracy/howtotell"

[12] 백승권,변정용 "윈도우즈-95환경에서 DLL로 구현한 정음형한글 오토마타", 동국대학교 전자계산학과, 1997"

[13] 윤지현,변정용 "IIIMP를 이용한 다국어 문자 입력 시스템", 동국대학교 대학원 전자계산학과, 1999"