Hyper English Word Study Dictionary with Indo-European roots

Yasunori Matsuo   s1070195

Supervised by Prof. Nobuyoshi Asai

Abstract

Bound dictionaries have been directly transfered to virtual form. Neither include the concept of learning. Grouping English word based on Indo-European Roots is one effective methods of learning, but the previous English-Japanese Hyper Dictionary (EJHD) [1] has the problem that the user does not understand how to use the English word. The problem was solved by adding examples. But by this function, new problem develops. The EJHD needs a function of explaining fine nuances of synonyms from Indo-European roots. Consequently, I suggest a necessary function to upgrade the EJHD. The EJHD enhances a user ability by changing the web design and displaying many special characters using the Unicode (UTF-8). By transition from Windows to Linux, The stability of EJHD system improves. It operates on non-selecting a platform used Java technology including Servlet and JSP.

1 Introduction

1.1 Background

Now, many kinds of the English-Japanese dictionaries are on the Internet. Most virtual dictionaries have been directly transfered from bound dictionaries. They have the advantage of enhancing the searching speed. But that is inadequate for learning English words. To understand English, it is necessary to understand the grammar, culture and words of English. But these cannot be memorized all at once. Consequently, when learning English words, the most effective learning method is using Indo-European roots.

1.1.1 What is Indo-European roots?

Proto-Indo-European, which is the original source of Indo-European languages containing English, does not exist as a real language in history. Proto-Indo-European is a virtual language. It was created by the comparative linguistics and developed rapidly in the 19th century Europe. It uses the theoretical re-composition following the regular change in sounds and rhyme found in the Indo-Europeans. *Indo-European-Roots* were structured by some stem pattern groups of the Proto-Indo-Eropeans. This language's derivatives are from about twelve linguistic groups: English, German, Dutch, French, Italian,

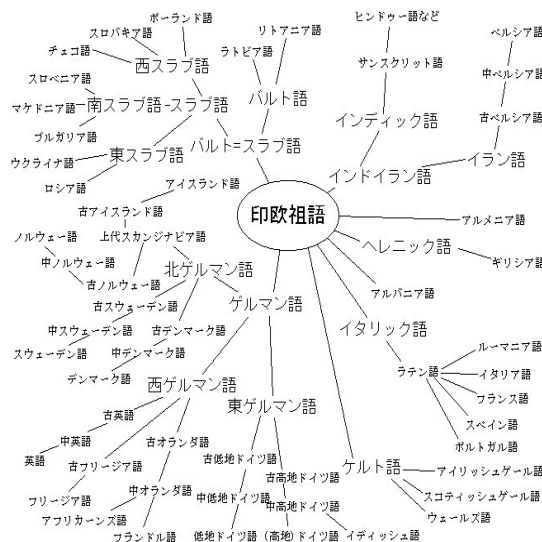Greek, Russian, Sanskrit, and so on (see Figure1). [2]

Figure 1: derivation

1.1.2 What is English-Japanese Hyper Dictionary (EJHD)?

The English-Japanese Hyper Dictionary is a dictionary which is made using Hyper Text Markup Language (HTML) on the World Wide Web (WWW).

- The function of the EJHD before upgrading are as follows:

  1. referencing from the front of English words
  2. referencing from the front of Indo-European roots
  3. indicating the Indo-European roots list
  4. referencing from the front of middle change words
  5. indicating the middle change words list
  6. referencing from the front of prefixes
  7. indicating the prefixes list
  8. referencing from the front of stems
  9. indicating the stems list

10. indicating the synonyms list

11. the number of registered English words is 10700

12. the number of registered Indo-European roots is 1348

13. the number of registered middle change words is 181

14. the number of registered prefixes is 71

15. the number of registered stems is 96

16. the number of registered synonyms is 196

1.1.3 Advantage of using English-Japanese Hyper Dictionary

We can learn English word more efficiently by using Indo-European roots.

Indo-European Roots are useful to distinguish words. For example, the Japanese word *utagau* means doubt and suspect in English.

According to *The American Heritage Dictionary of Indo-European Roots Second Edition (2nd ed.)(AHD 2nd ed.)*, The Indo-European Root of "doubt" is dwo-. [2] According to the *AHD 2nd ed.*, **dwo- means two**. "doubt" is derived from dividing A and B.

Indo-European Roots of "suspect" is spek- and upo. [2] According to the *AHD 2nd ed.*, **spek- means to observe** and **upo means under, up from under, over**.

These Indo-European Roots are not equivalent and the difference of meaning is obvious.

- "dwo-" corresponds these modern English words:

  1. **between:** in or into the space separating two or more points, objects, people, etc

  2. **bicameral:** having two main parts, such as the Senate and the House of Representatives in the US, and the House of Commons and the House of Lords in Britain

  3. **biceps:** the large muscle at the front of the top part of the arm

  4. **bicycle:** a road vehicle with two wheels that you ride by pushing the PEDALS with your feet

  5. **binary:** using only 0 and 1 as a system of numbers

  These words are associated with dwo-'s meaning, "two."

- "spek-" corresponds these modern English words:

1. **aspect:** a particular part or feature of a situation, an idea, a problem, etc.; a way in which it may be considered

2. **special:** not ordinary or usual; different from what is normal

3. **telescope:** a piece of equipment shaped like a tube, containing LENSES, that you look through to make objects that are far away appear larger and nearer

These words are associated with spek-'s meaning, "to observe"

- "upo" corresponds these modern English words:

1. **above:** at or to a higher place or position than sth/sb

2. **resurgent:** becoming stronger or more popular again

3. **subject:** used to refer to a particular time in the past or future

4. **succeed:** to achieve sth that you have been trying to do or get; to have the result or effect that was intended

These words are associated with upo's meaning, "under, up from under, over"

1.1.4 About The *American Heritage Dictionary of the English Language* (AHD) and The *Oxford Advanced Learner's Dictionary* (OALD)

The *AHD* [3] is one of the prime reference dictionaries for magazines, newspapers, and Internet content providers and it gets very high evaluation at Amazon.com which is the famous shopping web site in America. AHD's special consultants include authorities on anthropology, architecture, cinema, and law, plus military science, music, religion, and sports, which is reflected in an impressively comprehensive coverage of the arts, culture and technology. The 1350 entries of Indo-European Roots are more than double the number contained in the Appendix of Indo-European roots of the *AHD*. [3]

The *OALD* has eighty thousand references and contains information on British and American English. [4]

1.2 Purpose

The purpose of this research is to make a system for learning English word more efficiently by adding examples and changing the web design. The system should be low cost and operate on non-selecting a platform. These requirements are fulfilled by Java2 SDK Standard Edition (J2SE), Java Server Pages (JSP), Servlet Container, Web server, and MySQL. Java is a programming language developed by Sun Microsystems, Inc.

2 Explanation of Technology

2.1 Develop and Server Environment

OS	RedHat7.3J
Web Server	Apache1.3.27
Java	J2SDK1.4.1-01
Servlet Container	Tomcat4.0.6
RDB Server	MySQL3.23.49
JDBC Driver	mm.mysql

2.2 About Apache

Apache, which derives from "a patchy," is the Web server software, Apache was developed by The National Center for Supercomputing Applications (NCSA). NCSA is a leader in defining the future's high-performance computing infrastructure for scientists and for society. The Apache group is improving it based on httpd1.3. It runs with almost all computers which are UNIX, Linux, Mac, IBM/PC compatible machine and so on. [5]

2.3 About Tomcat

Tomcat is the servlet container that is used in the official Reference Implementation for the Java Servlet and Java Server Pages technologies. [6]

2.4 About MySQL

The MySQL is the most popular open source database in the world. It is free under the GNU General Public License (GPL) and delivers high performance in multi users and multi threads. Its database server architecture makes it extremely fast and easy to customize. [7]

2.5 About JDBC Driver

The driver of connection Servlet and MySQL is Java Database Connectivity (JDBC), which is a unified interface. It lets developers working with the Java programming language easily build programs and applets that interact with MySQL and connect all corporate data, even in a heterogeneous environment.

2.6 About Servlet and JSP

Servlet is the Java Program which operates on the server. JSP is one of the mechanism to achieve a Web site. It makes dynamic generated Web page displays, and Java version of Active Server Pages (ASP). It is one of the mechanism with Internet Information Service (IIS).

2.7 The Difference between Servlet and JSP

Servlet outputs HTML, but JSP implant in HTML.

```
************< Servlet Sample>************
```

```
import java.io.*;
import javax.servlet.*;
import javax.servlet.http.*;

public class hello extends HttpServlet{
  public void doGet(
   HttpServletRequest request,
    HttpServletResponse response)
     throws ServletException,IOException{
      response.setContentType(
        "text/html;charset=EUC-JP");
      PrintWriter out=response.getWriter();
      out.println("<html>");
      out.println("<head><title>
           Hello World</title></head>");
      out.println("<body>");
      out.println("<p>Hello World!!<p>");
      out.println("</body>");
      out.println("</html>");
      }
}
**************< JSP Sample>**************
<%@ page contentType="text/html;
charset=EUC-JP" %>
<html>
<head><title>Hello World</title></head>
<body>
<% out.print("<p>Hello World!!<p>"); %>
</body>
</html>
************< End of Sample>************
```

Servlet need to compile before execution but JSP need not to.

2.8 About Java Beans

Java Beans is a Java class created by specific rules like calling the other Java program easily. Java Beans needs to have the non-argument constructor and Accessor Method, and implementing Serializable Interface. [8]

2.9 Why is not it CGI but Servlet and JSP

Servlet and JSP has many advantages:

1. Basically, ASP does not execute without the Microsoft Windows environment, but Servlet and JSP are able to run regardless of the operating system.

2. Servlet and JSP can construct a system at a moderate price.

3. CGI (Perl), ASP and PHP use the system called interpreter type, but Servlet and JSP are the compile languages so the process efficiency is better.

4. Servlet and JSP is divided into a design and logic territory, so compartmentalization of the designer and programmer is very clear. Servlet is in charge of a logic and JSP is in charge of a design.

3 Setup

3.1 About Red Hat Package Manager (RPM)

RPM is a software to manage the package developed by Red Hat Software Inc. It is a standard in Linux distribution of Red Hat series, distributed on The GNU General Public License (GPL). It has two functions. First, is the query function of searching a package or specific file with database. Second, it is able to execute installation, upgrade and uninstallation easily with check function of necessary file presence for a package. [9]

3.2 Installation and Configuration

3.2.1 JAVA

First, launch the executable file you downloaded, j2sdk-1_4_1_(version number)-linux-i586-rpm.bin, by using the following commands from the directory in which it is located:

% chmod a+x j2sdk-1_4_1_(version number)-linux-i586-rpm.bin
% ./j2sdk-1_4_1_(version number)-linux-i586-rpm.bin

Second, Run the rpm command to install the packages that comprise the Java 2 SDK:

% rpm -iv j2sdk-1_4_1_(version number)-linux-i586.rpm

Third, configurate the PATH:

% setenv PATH "/usr/java/j2sdk1.4.1_01/bin"

3.2.2 Apache

First, melt the file you downloaded, apache_1.3.27.tar.gz, by using the following commands:

% gtar zxvf apache_1.3.27.tar.gz

Second, install the apache by using the following commands:

% ./configure –prefix=/usr –with-layout=RedHat
–enable-module=all –enable-shared=max
–disable-module=auth_dbm –with-perl=/usr/bin/perl
–enable-suexec –suexec-docroot=/home/httpd/html
–suexec-caller=nobody

% make
% make install

Third, copy the file, mod_webapp.so, supplied with the Apache Jakarta Project to /usr/libexec. Add the following messages to /etc/httpd/conf/httpd.conf for cooperating with tomcat:

LoadModule webapp_module libexec/mod_webapp.so
WebAppConnection conn warp localhost:8008
WebAppDeploy hidic conn /hidic

3.2.3 Tomcat

Melt the file, jakarta-tomcat-4.0.6-LE-jdk14.tar.gz, to /usr/tomcat4 by using the following commands:

% gtar zxvf jakarta-tomcat-4.0.6-LE-jdk14.tar.gz

3.2.4 MySQL

Run the rpm command to install the packages

% rpm -ivh mysql-server-(version number)
% rpm -ivh mysql-(version number)
% rpm -ivh mysqlclient9-(version number)
% rpm -ivh mysql-devel-(version number)

4 Research Contents

4.1 Transition from Windows to Linux

Improved EJHD's system is constructed by free software, for example RedHat, Apache, J2SDK, Tomcat and MySQL. By transition from Windows to Linux, the stability of EJHD's system improves.

4.2 Adding Examples and Changing Web Page Design

A See Figure2

4.3 Display of special characters with Unicode(UTF-8)

The special characters were replace by images on previous EJHD. The user can resize the word's size with browser setting, but the size of images cannot change. New EJHD displays many special characters used the Unicode (UTF-8) as word. Replaced words are $\bar{a}$, $\bar{e}$, $\bar{i}$, $\bar{o}$, $\bar{u}$, g^w *and* k^w.

5 Conclusion

In this research, the problem that users do not understand how to use English words was solved by adding examples. However by this function, a new problem develops. The EJHD needs a function of explaining a fine nuance

Figure 2: New Design and Examples

between synonyms from Indo-European roots. Consequently, I suggest a necessary function to upgrade the EJHD. The EJHD enhances a user ability by changing the web design and displaying many special characters using the Unicode (UTF-8). By transition from Windows to Linux, The stability of EJHD's system improves. It operates on non-selecting a platform used Java technology including Servlet and JSP.

Acknowledgments

References

[1] N. Asai laboratory at University of Aizu, Hyper English Dictionary,
http://mfntsv1.u-aizu.ac.jp/hidic/hdhome.html

[2] C. Watkins, *The American Heritage Dictionary of Indo-European Roots Second Edition*,
Houghton Mifflin Company, 2000.

[3] Margery S. Berube, *The American Heritage Dictionary of The English Language 4th Edition*,
Houghton Mifflin Company, 2000.

[4] A. S. Hornby and Sally Wehmeier, *Oxford Advanced Learner's Dictionary*, Oxford University Press, 2002.

[5] ASCII Corporation, "ASCII Digital Dictionary"
http://yougo.ascii24.com/gh/45/004590.html

[6] Apache Software Foundation,
"The Jakarta Site — Apache Tomcat"
http://jakarta.apache.org/tomcat/index.html

[7] MyNA, My SQL Reference,
http://www.mysql.gr.jp/jpdoc/3.23.x/manual.ja_toc.html
(16 Febuary 2002)

[8] Yoshihiro Yamada, *10   JSP/
        *, Shoei Inc, 2000.

[9] ASCII Corporation, "ASCII Digital Dictionary"
http://yougo.ascii24.com/gh/39/003919.html

[10] S. Yoshimura, "A study about the hyper-dictionary used Indo-European roots information", Master's thesis in the department of Science and Engineering at University of Tsukuba, 1997.

[11] Yasuhiko Ikebe,

                        ,
                3   27-30, 99-102, 161-166,
2001.