Improvement of a Hyper Dictionary user-friendliness

Yoshikatsu Matsumoto   s1070199

Supervised by Prof. Nobuyoshi Asai

1 Abstract

A dictionary which called the Hyper Dictionary uses Indo-European roots. A user-friendly side has not been improved in previous research, though various functions have so far been improved by University of Aizu undergraduate students. Because links are not yet attached to the examples, links are not attached to English words used in a semantic column for explanations, although the columns are blank, the titles of columns are displayed, and so on. These factors were improved by adding links to English words with example sentences. Links were added to reference English words in a semantic column which enables consulting a word only by a click of a mouse. As a result, conjugated forms of English words made it easier to understand them intuitively. These improvements were able to greatly increase the usability of the dictionary.

2 Introduction

2.1 What is the Hyper Dictionary?

The Hyper Dictionary is an English word study dictionary using Indo-European roots. It has been improved by many University of Aizu undergraduate students until now[14]. This year it is being improved by four senior students. The following section shows the improved functions with each researcher's name in parentheses.

2.1.1 Improving points of a Hyper Dictionary this year

- Addition of pronunciation (Ishikawa)
- Classification of English words (Ootake)
- Addition of suffix searching (Oowada)
- Addition of links to English words in example sentences (Matsumoto)

2.2 What is the Indo-European roots?

Indo-European roots are an imaginative language created by comparative linguistics. The group of languages spoken in west Ireland, the area from eastern Ireland to India and the area from north Scandinavia to southern Italy are called Indo-Eurorean. Comparative linguists have investigated the similar features of languages in this group. Those languages were found to come from a common ancestry from the investigation, and are called Proto-Indo-European. The language is roughly divided into the 12 groups: Anatolian, Indo-Iranian, Greek(Hellenic), Italic, Celtic, Germanic, Armenian, Tocharian, Balto-Slavic, Albanian, and other branches [see Figure 1]. Indo-European roots is an imaginative language which divides the roots of these languages into several patterns.[2]
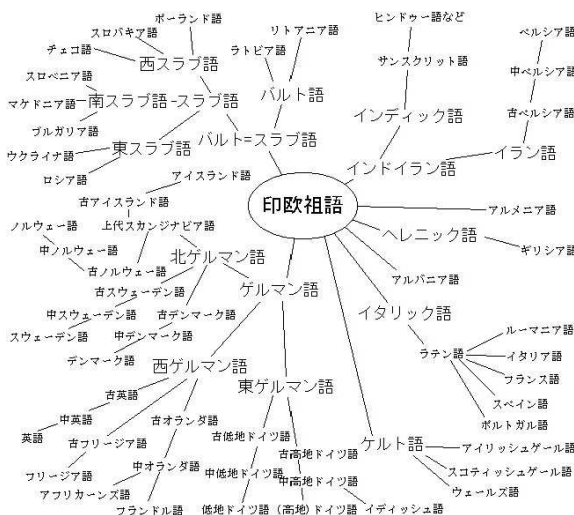

Figure 1: Derivation

2.2.1 What is the American Heritage Dictionary?

The American Heritage Dictionary (AHD) has been compiled by various authorities of cultural anthropology, construction, movie, law, military affairs, music, religion, and sports, etc. It is used by people whose native language is English, and by translators whose native language is not English. It is a dictionary similar to KOU-JIEN, which is used in Japan.[1]

2.3 Background

Now, dictionaries are offered in various forms, such as Internet medium, hard copy and digital mediums. Dictionaries were originaly offered only in paper medium. However, Internet and CD-ROM dictionaries have recently increased with the spread of personal computers. Today, since CD-ROM and Internet dictionaries are spreading, the only advantage of paper dictionaries is their portability. But present internet dictionaries have only a little advantage in comparison with hard copy dictionaries; because there are many Internet dictionaries, it

is possible to use a dictionary of paper medium by the Internet. Another advantage is that time is shortened when consulting a dictionary, and there is no restriction such as the number of letters and pages, and, if clicked, it is possible to look up a linked word, which can be accessed by anyone by the internet. There are some internet dictionaries that already have these feature, such as EIJIROU, which has many example sentences. Some Internet dictionaries also include the pronunciation of English words. If the act of consulting a dictionary is means to just examine the definitions of English words, such Internet dictionaries do not have good study efficiency, because only the definitions are studied without the semantic background of English words. Although various methods to learn English exist, the method using Indo-European roots has been very efficient and effective.[10]

2.4 Purpose

Various dictionaries are now offered, such as on-line, hard copy and digital mediums. One Internet dictionary, called the Hyper Dictionary, is an English word study dictionary which uses Indo-European roots. Although it has already been developed by many University of Aizu undergraduate students until now[14], the focus of this research will be on improving the Hyper Dictionary by making it more user-friendly. Currently, the Hyper Dictionary is not user-friendly. Because links are not yet attached to the examples, links are not attached to English words used in a semantic column for explanations, although the columns are blank, the titles of columns are displayed, and so on. These factors reduce the user-friendliness of a dictionary clearly. So one goal of this research was to improve this aspect of the Hyper Dictionary. Moreover, another major goal is to build the system by adding more English words, example sentences, and Indo-European roots, so learners can study even more English words.

3 Technology of building a server

3.1 Composition of server

- Operating System: RedHatLinux9.0

- Web Server: Apache2.0.40

- Servlet Container: Tomcat4.1.27

- JSP API version: 2.3

- Servlet API Version: 1.2

- Reletional Database Server: MySQL4.0.16

- Java: JavaTM 2 SDK 1.4.2_02

- JDBC Driver: mm.mysql

- Tomcat Connector: mod_jk2

3.2 Explanation of technology

3.2.1 About Linux

Linux is a popular and free Unix-like operating system. This operating system can be customized freely and widely, so it is currently very useful and popular in the field of computing. It's prototype was developed by Linus Torvalds. Linux was widely opened to the public as open source software, and many volunteer developers have participated in developing it.

3.2.2 About Apache

Apache is a Web Server which started being developed in 1995. The Web server of The National Center for Supercomputing Aplication's httpd called NCSA httpd 1.3 is the foundation. In the same way as Linux, it developed by volunteer programmers around the world. The name of Apache has come from a mixture of correction and the Advanced Capability patch. Presently, Apache is the most often used Web server in the world.

3.2.3 About Tomcat

Tomcat was developed as sub project of the Jakarta project. It is an Application Server in which is possible to process a Java Servlet/JSP. Tomcat has function as a Web server independently, but it is normally used as a plug in of APACHE and IIS. Many different servlet containers are available today. The most popular one is recognized as the official servlet/JSP container.

3.2.4 About Mysql

Mysql is Relational Database Management System(RDBMS). It operates by the multithread and multiuser, and it has strong and hi-speed character. It is the most popular RDBMS in Europe and America.

3.2.5 About Java DataBase Connectivity(JDBC) Connector

A JDBC Connector is the driver which is accessed from a Java program to RDBMS. It issues the commands of SQL language to make it possible to operate the database.

3.2.6 About Java Server Pages(JSP)

JSP is a technology which generates Web pages dynamically on the Web server and transmits them to the client. It has an advantage in that it is not necessary to remember a particular script language, and it is possible to use Java Language.

3.2.7 About Java Servlet

Java Servlet is a module-ized program which is executed on the Web server. Servlet is described with Java language, so it does not depend on the particular operating system and hardware. It differs from other CGI programs and the like. Because if a program is called once, it is permanently stationed in memory. So it makes high speed processing possible.

3.2.8 About Java Beans

Java Beans is a style which combines a module-ized Java program, and Application software is built. It can do sharing with two or more developers easily, and reuse of source code.

3.2.9 About Java Technology

The software which was developed with Java does not depend on a particular operating system and hardware. The degree of flexibility of Java is its largest merit of Java.

4 Research Contents

4.1 Addition of link to the English word in example sentence

Last year, new functions were added to the Hyper Dictionary which can display example sentences of each English word. But it displayed merely example sentences. Therefore, when understanding of an English word in an example sentence is not impossible, it was necessary to take the trouble to type the English word in the text box. Clearly, this act reduced usability. I wrote a program that divides each word by the program and attaches a link to each word, when an example is displayed. But by improving this function, a problem was encountered, which is dealing with English words with defferent meaning in spite of having the same spelling, for example, the word "like." When this word is classified roughly, it has two meanings. The first meaning of "like" is to enjoy something, or to think that someone or something is nice. Here, "like" is used in a sentence similar to "She's nice. I like her." A second meaning of "like" is similar to someone or something. This "like" is used in a sentence similar to "She's wearing a dress like mine." These meanings similar to "like(1)," "like(2)" exist in the database of the Hyper Dictionary. And the paper dictionary also has similar to "like1," "like2." When an English word is used in an example sentence, the program is confused about which meaning is most appropriate. Thus, the database of Hyper Dictionary was redeveloped to improve understanding. The last portion of the English words was added (1), (2), and (3) in the

Database. But it is not possible to use (1), (2) and (3) in the example sentences. When these English words are indicated, it must be indicated after deleting (1), (2) and (3). A program was written to actualize it with Java [see Figures 2 and 3].
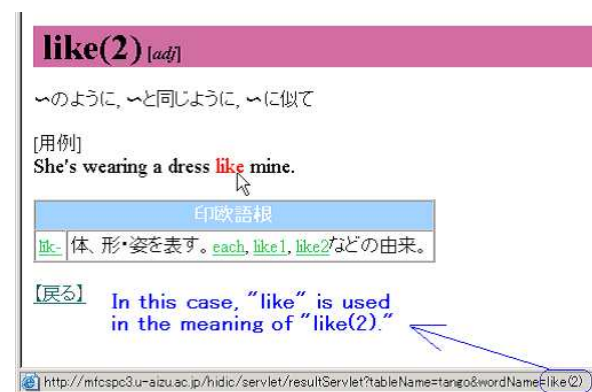

Figure 2: like(1)

Figure 3: like(2)

4.2 Addition of link to reference English word in semantic column

In the semantic column of words, there are times when other English words are used in order to explain a particular English word. In this case, it is the same as the example sentence. Other English words are used in order to explain an English word displayed merely as example sentences. This also reduced usability.

4.3 Addition of the conjugated form of an English word

Pronunciation is important in the conjugation of a conjugated form. Spelling is not so important. It was difficult

to recognize a conjugated form because the Hyper Dictionary does not have a technique to display pronunciation.

4.4 Deletion the unnecessary notations in a Web Page

In the previous Hyper Dictionary, although the column was blank, the title of a column was displayed. But if only the title of a column is displayed, the user has trouble [see Figure 4 and 5].

Figure 4: Previous Hyper Dicitonary

Figure 5: Present Hyper Dictionary

4.5 Correction of the already-known System Bug

There was a misconverted bug in the dictionary until now. This bug appeared when English words and Indo-European roots had both been used for the explanation in the semantic column of the Indo-European roots. All of the English words being used for the explanation are misconverted in "(unknown)," which is displayed. This bug does not appear when words such as a prefix, stem, and so on, were used. The program was modified, and this bug was corrected [see Figure 6 and 7].

Figure 6: Before modifying

Figure 7: After modifying

5 Conclusion

In this research, various functions were added to the Hyper Dictionary. Appropriate links were added to English words in example sentences, and links were added to reference English words in the semantic column. Conjugated forms of English words were also added and unnecessary notations were deleted in the Web Page. The links added to the English words in example sentences and reference English words in the semantic column enables consulting a word only by a click of a mouse. These expansions was able to increase the usability of the Hyper Dictionary substantially. Adding the conjugated forms of English words made it easier to understand a conjugated form intuitively. This also improved usability. Deleting unnecessary notations in the Web Page resulted in a simpler Webpage.

Acknowledgement

References

[1] Margery S. Berube, *The AMERICAN HERITAGE dictionary OF THE ENGLISH LANGUAGE fourth edition*, Houghton Mifflin Company, 2000

[2] C. Watkins, *The AMERICAN HERITAGE dictionary of Indo-European Roots Second Edition*, Houghton Mifflin Company, 2000

[3] N. Asai laboratory at University of Aizu, Hyper English Dictionary with Indo-European roots Ver. 3.00,
http://hidic.u-aizu.ac.jp/

[4] *TAISHUKAN'S Unabridged Genius English-Japanese Dictionary second edition*, Taishukan Inc, 2002

[5] *NEW SHOGAKUKAN Random House ENGLISH-JAPANESE DICTIONARY second edition*, Shogakukan Inc, 2003

[6] Margery S. Berube, *KENKYUSHA'S NEW ENGLISH-JAPANESE DICTIONARY*, KENKYUSYA

[7] MyNA, My SQL Reference,
http://www.mysql.gr.jp/jpdoc/4.0/manual.ja_toc.html

[8] JavaTM 2 Platform, Standard Edition, 1.4.0 API,
http://java.sun.com/j2se/1.4/ja/docs/ja/api/index.html

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

[10] Mutsu Imano, Atsushi Takayasu, Kazuhiro Toda, Rei Nishikawa, Hiroshi Hujimura, *Servlet/JSP Programming*, SOFTBANK Publishing, 2003

[11] EIJIROU on the web,
http://www.alc.co.jp/

[12] The EXCEED dictionary in infoseek,
http://jiten.www.infoseek.co.jp/

[13] The English-Japanese, Japanese-English dictionary in excite,
http://www.excite.co.jp/dictionary/

[14] Past graduation theses about the Hyper Dictionary,
(1)Scientific English (Atsushi Iwatsuki:s1052015)
(2)Upgrading the English-Japanese Hyper-Dictionary (Hideo Sugiura:s1061091)
(3)Upgrading the Indo-European-Root-Based Hyper English Dictionary (Yoshihiro Hoshino:s1061138)
(4)Improving Portability and Usability of the Indo-European-Root-Based English-Japanese Hyper Dictionary (Kumi Miyata:s1061148)
(5)Synonym Usage viewed from Pre-Historic Roots (Chie Yamanaka:s1061155)
(6)Hyper English Word Study Dictionary with Indo-European Roots and Interim Transitional Words (Yasuto Horikawa:s1070190)
(7)Hyper English Word Study Dictionary with Indo-European roots (Yasunori Matsuo:s1070195)
(8)Development of English Word-Study Tool with the Indo-European Roots (Takumi Sekikawa:s1070131)