UNL Explorer

Hiroshi Uchida[1]   Meiying Zhu[1]   Khan Md. Anwarus Salam[1,2]

(1) UNDL Foundation, Tokyo, Japan.
(2) The University of Electro-Communications, Tokyo, Japan.

`uchida@undl.org, zhu@undl.org, khan@undltokyo.org`

ABSTRACT

Universal Networking Language (UNL) is a language for computer to represent knowledge and information described in natural languages. Universal Words (UWs) constitute the vocabulary of UNL. The UNL Explorer is a web based application, which combines all the components of UNL system to be accessible online. The users of UNL Explorer are not only researchers and linguists who are interested to work with UNL technologies, but also general people who want to communicate free from language barriers. This paper describes the features of UNL Explorer. In brief, UNL Explorer provides many powerful features such as multilingual context search, multilingual communication such as UNL Talk and multilingual dictionary. Using multilingual context search users can retrieve documents in any language. Moreover, UNL Explorer shows the documents in various languages such as English, Japanese and more than 40 languages. Users can also access the UWs based multilingual dictionaries. UNL Society members can contribute online for updating their language dictionary entries.

KEYWORDS: Multilingual Context Search; Machine Translation; Multilingual Dictionary; Universal Networking Language (UNL); Ontology;

1   Introduction

Universal Networking Language (UNL) is a language for computer to represent knowledge and information described in natural languages. Universal Words (UWs) constitute the vocabulary of UNL. UW is a word for constructing UNL expressions (UNL Graph). So keys to the information in UNL documents are UWs. UWs are stored in the UW dictionary.

UNL Explorer is a web based application, which combines all the components of UNL system to be accessible online. The users of UNL Explorer are not only researchers and linguists who are interested to work with UNL technologies, but also general people who want to communicate free from language barriers. This paper describes the features of UNL Explorer.

In brief, UNL Explorer provides many powerful features such as multilingual context search, multilingual communication such as UNL Talk, multilingual dictionary and UNL Ontology. Using multilingual context search users can retrieve documents in any language. UNL Explorer provides very promising solution for search. Moreover, UNL Explorer shows the documents in various languages such as English, Japanese and more than 40 languages. Users can also access the UWs based multilingual dictionaries. UNL Society members can contribute online for updating their language dictionary entries.

2   BACKGROUND

2.1   Universal Networking Language (UNL)

UNL initiative was originally launched in 1996 as a project of the Institute of Advanced Studies of the United Nations University (UNU/IAS)[1]. Describing the detail technical information UNL book was first published in 1999 (Uchida et. al. 1999). In 2001, the United Nation University set up the UNDL Foundation[2], to be responsible for the development and management of the UNL project. In 2005, a new technical manual of UNL was published (Uchida et. al. 2005), which defined UNL as an knowledge and information representation language for computer. UNL has all the components to represent knowledge described in natural languages. UWs constitute the vocabulary of UNL and each concept that natural languages have is represented as unique UW. A UW of UNL is defined in the following format:

$$<uw> =:: <headword>[<constraint list>]$$

Here, English words or phrases are used for headword, because of easy understanding for the people in the world. UW can be a word, a compound word, a phrase or a sentence. Universal Words (UWs) constitute the vocabulary of UNL. UW is a word for constructing UNL expressions (UNL Graph). So keys to the information in UNL documents are UWs. UWs are stored in the UW dictionary. UWs are inter-linked with other UWs using "relations" to form the UNL expressions of sentences. These relations specify the role of each word in a sentence. Using "attributes" it can express the subjectivity of author. Currently, UWs are available for many languages such as Arabic, Bengali, Chinese, English, French, Indonesian, Italian, Japanese, Mongolian, Russian, Spanish and so forth.

---

[1]http://www.ias.unu.edu/
[2]http://www.undl.org/

2.2 UNL Ontology

UNL Ontology is a semantic network with hyper nodes. It contains UW System which describes the hierarchy of the UWs in lattice structure, all possible semantic co-occurrence relations between each UWs and UWs definition in UNL. With the property inheritance based on UW System, possible relations between UWs can be deductively inferred from their upper UWs and this inference mechanism reduces the number of binary relation descriptions of the UNL Ontology. In the topmost level UWs are divided into four categories: adverbial concept, attributive concept, nominal concept and predicative concept.

3 UNL Explorer

UNL Explorer[3] is a web based application, which combines all the components of UNL system to be accessible online. In brief, UNL Explorer provides many powerful features such as multilingual context search, multilingual communication such as UNL Talk, multilingual dictionary and UNL Ontology.

Multilingual context search enable the people to retrieve desired documents by the natural language query of any language. Each document is provided in UNL by automatically analysing the original document, together with the original document. The queries are converted into UNL. Then the system try to match this UNL graph with the existing UNL documents with inference. Knowledge which is necessary to make inference are provided in UNL Ontology, especially in UWs definition.

Retrieved documents or any other documents can be shown in various languages such as English, Japanese and more than 40 languages, by automatically generating each language sentences from UNL expressions.

This function allows the users to translate any documents in different languages. This allows the users to communicate across language barriers by UNL Talk. This option can be very useful for communicating in cross-cultural communication. To realize this function the system need the UW dictionaries for many languages, which defines the correspondence between UWs and each languages words.

This data can be access as multilingual dictionary. UNL Society members can contribute online for updating their language dictionary entries. Figure1 shows the UNL Explorer screen shot with explanations of the options.

FIGURE 1 –UNL Explorer Homepage screen shot

---

[3]http://www.undl.org/unlexp/

3.1 Multilingual Context Search

Multilingual context search enable the people to retrieve desired documents by the natural language query of any language. UNL Explorer provides multilingual search facility for UNL documents. For this each document is provided in UNL by automatically analysing the original document, together with the original document. The queries are converted into UNL. Then the system try to match this UNL graph with the existing UNL documents with inference. Knowledge which is necessary to make inference is provided in UNL Ontology, especially in UWs definition. To perform multilingual context search user can write the search query in text box and click the Search button.
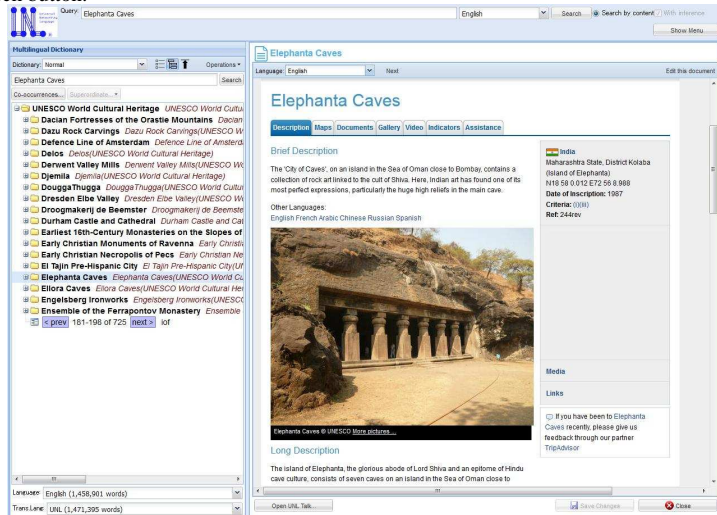

FIGURE 2 –UNL Explorer screen shot showing the UNESCO document on "Elephanta Caves"

Retrieved documents or any other documents can be shown in various languages such as English, Japanese and more than 40 languages, by automatically generating each language sentences from UNL expressions. Search Query: Write the Keyword or Content to search from UNL Information and Knowledge Management System. The UNL Explorer will show the results in UNL or in a desired natural language by Deconverting the UNL expressions of the information using the UNL Deconverter. In background UNL Explorer translates using UNL Enconverter and Deconverter. Both UNL EnConverter and Deconverter support different languages such as Chinese, English, Japanese and so forth.

3.2 UNL Talk

UNL Talk allows the users to communicate across language barriers by UNL Talk. This option can be very useful for communicating in cross-cultural communication. To realize this function the system need the UW dictionaries for many languages, which defines the correspondence between UWs and each languages words. Figure 3 shows the screen shot of UNL Talk where the

authors are communicating using Bengali and English. Each users can see the messages in their mother language.
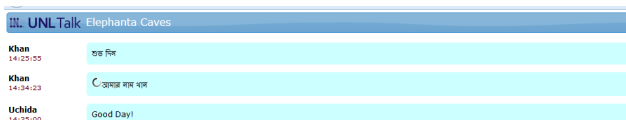

FIGURE 3 –Screen shot of UNL Talk where users communicating in Bengali and English

3.3 Multilingual Dictionary

One of the most unique features of UNL Explorer is the multilingual dictionary which is available for more than 40 languages. This multilingual dictionary is based on UWs dictionaries for many languages, which defines the correspondence between UWs and each languages words. Users can use the dictionary side by side for any of these language pair.
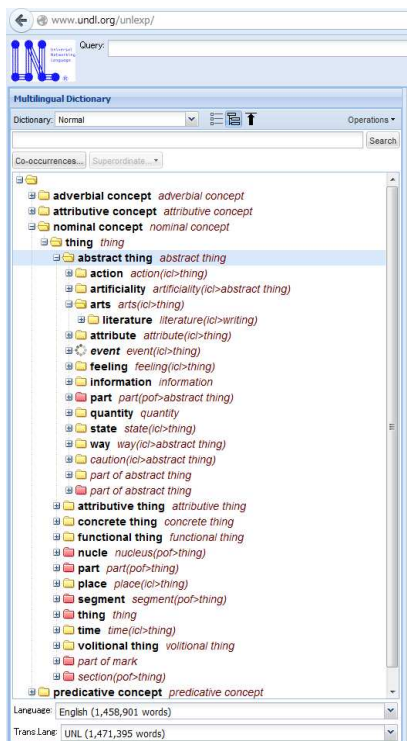

FIGURE 4– Multilingual dictionary frame showing English-UNL

UNL Explorer users can browse the multilingual dictionary which comes in the left side, which we refer as "multilingual dictionary frame" as shown in Figure 4. This tool also provides the search facility for UWs dictionary. Users can search the meaning of a word in any languages. Users can choose their desired language pairs from the options given in the downside of the multilingual dictionary frame. Figure 4 shows the screen shot of Universal Words frame. It displays the UWs system hierarchy (a lattice structure) in a plain tree form. Information can be navigated through the UWs system and users are also able to know the position of each concept of a UW in the conceptual hierarchy at the same time.

UNL Explorer also provides an advanced search facility for discovering UWs relations from UNL Ontology. This option allows user to get the semantic co-occurrence of any UWs. Users can also check incoming and outgoing relationships of each UWs using this facility. This UNL Ontology search mechanism is accessible for computer program using UNL Explorer API. However, to use this API, user need to be a UNL society member by signing an agreement with UNDL Foundation.

UNL Society members can contribute online for updating their language dictionary entries. From 'Universal Word' Properties menu, users can browse the word semantics. It is possible to edit the UWs dictionaries online. UNL society members can also download the UWs dictionaries.

To ensure every language speakers can create the correct UWs dictionary entry, UNL Explorer provide the explanation of UWs in different natural languages (Khan et. al., 2011). It is a novel contribution for auto generating the UWs explanations from the semantic background provided by UNL Ontology.

Conclusion

In this paper, we described the features of online based UNL Explorer. For making the people freely communicate with each other in their mother language, UNL technology is very promising. UNL Explorer provides useful features such as multilingual context search, multilingual communication using UNL Talk and multilingual dictionary for general users. UNL Explorer also provides API for researchers and application developers to use UNL technologies such as UNL Enconverter, DeConverter, UWs Dictionary, UNL Ontology based on corpus.

References

H. Uchida, M. Zhu, T. Della Senta. "A gift for a millenium". Tokyo: IAS/UNU. 1999.

H. Uchida, M. Zhu, and T. Della Senta. "The Universal Networking Language", 2nd ed. UNDL Foundation, 2005.

Khan Md. Anwarus Salam, Hiroshi Uchida and Tetsuro Nishino. "How to Develop Universal Vocabularies Using Automatic Generation of the Meaning of Each Word", 7th International Conference on Natural Language Processing and Knowledge Engineering (NLPKE'11), Tokushima, Japan. ISBN: 978-1-61284-729-0. Page 243 – 246. 2011.