

The **Universal Networking Language** (UNL) is an artificial language for representing, describing, summarizing, refining, storing and disseminating information in a natural-language-independent format. It is a kind of mark-up language which represents not the formatting but the core information of a text. As HTML annotations can be realized differently in the context of different applications, machines, displays, etc., so UNL expressions can have different realizations in different human languages.

History

The UNL Programme started in 1996, as an initiative of the Institute of Advanced Studies of the United Nations University
in Tokyo, Japan. In January 2001, the United Nations University set up an autonomous organization, the
UNDL Foundation
, to be responsible for the development and management of the UNL Programme. The Foundation, a non-profit international organisation, has an independent identity from the United Nations University, although it has special links with the UN. It inherited from the UNU/IAS the mandate of implementing the UNL Programme so that it can fulfil its mission. Its headquarters are based in Geneva, Switzerland.

From the very beginning, a consortium of university departments from all regions of the world has been engaged in developing the UNL. That's the UNL Society, a global-scale network of R&D teams, involving several specialists in computer science and linguistics, who are at work creating the linguistic resources and developing the web structure of the UNL System. The UNDL Foundation provides technological support and co-ordinates the implementation of the Programme.

The Programme has already crossed important milestones. The overall architecture of the UNL System has been developed with a set of basic software and tools necessary for its functioning. These are being tested and improved. A vast amount of linguistic resources from the various native languages already under development has been accumulated in the last few years. Moreover, the technical infrastructure for expanding these resources is already in place, thus facilitating the participation of many more languages in the UNL system from now on. A growing number of scientific papers and academic dissertations on the UNL are being published every

year.

The most visible accomplishment so far is the recognition by the Patent Co-operation Treaty (PCT) of the innovative character and industrial applicability of the UNL, which was obtained in May 2002 through the World Intellectual Property Organisation (WIPO). Acquiring the patent for the UNL is a completely novel achievement within the United Nations.

Scope and Goals

The UNL is an effort to achieve a simple basis for representing the most central aspects of information and meaning in a human-language-independent form. As a knowledge representation language, the UNL aims at coding, storing, disseminating and retrieving information independently of the original language in which it was expressed. In this sense, the UNL seeks to provide the tools for overcoming the language barrier in a systematic way.

At first glance, the UNL seems to be an “interlingua”, a sort of pivot-language to which the source texts are converted before being translated into the target languages. It can, in fact, be used for such a purpose, but its primary objective is to serve as an infrastructure for handling knowledge rather than individual languages.

In the UNL approach, there are two basic different movements: UNL-ization and NL-ization. UNL-ization is the process of representing/mapping/analysing the information conveyed by natural language utterances into UNL; NL-ization, conversely, is the process of realizing/manifesting/generating a natural language document out of a UNL graph. These processes are completely independent. For the time being, the NL-ization process has been already fully automatic, whereas the UNL-ization process is still mostly human, even though machine-aided.

Currently, the main goal of the UNL-ization process has been to map the information that is verbally elicited in the surface structure of written texts into a language-independent and machine-tractable database. This means that the UNL representation has not been committed to replicate the lexical and the syntactic choices of the original, but focuses in representing, in a non-ambiguous format, one of its possible readings, preferably the most conventional one. In this sense, the UNL representation has been an interpretation rather than a translation of a given text.

Indeed, it is important to note that at this point in time it would be foolish to state it possible to represent the “full” meaning of any word, sentence or text for any language. Subtleties of

intention and interpretation make the “full meaning”, whatever concept we might have of it, too variable and subjective for any systematic treatment. The UNL avoids the pitfalls of trying to represent the “full meaning” of sentences or texts, targeting instead the “core” or “consensual” meaning that is most often attributed to them. In this sense, much of the subtlety of poetry, metaphor, figurative language, inuendo and other complex, indirect communicative behaviours is beyond the current scope and goals of the UNL. Instead, the UNL targets direct communicative behaviour and literal meanings as a tangible, concrete basis for much or most of human communication in practical, day-to-day settings.

This is the main reason why UNL has not been exactly a interlingua-based machine translation project, even though machine translation is one of the possible and more obvious and promising uses of UNL. The main problem is that the practice of translation has been normally restricted to the notion of "fidelity" (or faithfulness), i.e., any translated version of a text is expected to be a replica, in another language, of the content and of the form of the original. This transfer process, however, is “all too human”, as Nietzsche said, to be replicated by the currently existing technology, which is not prepared to deal with several linguistic and cultural phenomena, such as vagueness, ambiguities, metaphors, presuppositions, ellipses, implicatures and so on. This does not mean that natural language automatic processing, and therefore machine translation, is impracticable; it just means that it is not possible yet to do that completely without humans or in the same way humans do. The results, in any case, are likely to be different from the ones produced by humans. Several techniques (rule-based, memory-based, corpus-based) have been proposed to decrease the role of humans in natural language analyses tasks, but the results, even though already promising, are not of publishing-quality yet, and require substantial human revision.

In addition to translation, the UNL has been exploited for several other different tasks in natural language engineering, such as multilingual document generation, summarization, text simplification, information retrieval and semantic reasoning. Indeed, in UNL-based applications there is no need for the source language to be different from the target one: an English text may be represented in UNL in order to be generated, once again, in English, as a summarized, a simplified or a simply rephrased version of the original.

Finally, it should also be stressed that UNL, differently from other auxiliary languages (such as Esperanto, Interlingua, Volapük, Ido and others), is not intended to be a human language. We do not expect people to speak UNL or to communicate "in" UNL; we do expect them to use UNL and to communicate "through" UNL, but in the same unconscious, invisible and spontaneous way they do with other declarative and procedural languages which are pervasive in everyday applications. As no one is required to know HTML to browse the Internet or even to create websites, everyone would be able to UNL-ize documents and to extract out of them the information needed without any knowledge of UNL. UNL is therefore a formal language

designed for computers, not for humans. Like other logical systems, it seeks to provide the linguistic and semiotic infrastructure for computers (and not for humans) to handle what is meant by natural languages

[read more about UNL in the UNL wiki]