X UNL School

Written by Ronaldo Martins
Tuesday, 31 July 2012 22:16 - Last Updated Tuesday, 07 August 2012 12:11

The UNDL Foundation invites applications for the X UNL School, to take place in the Library of Alexandria, in Alexandria, Egypt, from October 7th to 11th, 2012. The workshop will be dedicated to the development and improvement of the grammatical resources for the UNL framework. The UNDL Foundation will pay the travel and accommodation expenses for the selected candidates not living in Alexandria. Approved candidates will be entitled to sign a contract for the development of the language modules for their respective languages.

IMPORTANT DATES

-   01/09/2012: Deadline for the applications
-   10/09/2012: Notification of accepted candidates
-   30/09/2012: Deadline for the pre-workshop tasks (corpus translation and dictionary preparation)
-   07-11/10/2012: X UNL School
-   15/01/2013: Deadline for the post-workshop tasks (Corpus1000)

REQUISITES

The UNDL Foundation will only consider applications complying strictly with the three requisites below:

1. Candidates must have successfully completed the grammars to UNL-ize and NL-ize the Corpus500;
2. Candidates must have completed CLEA250, CLEA500 and CLEA750; and
3. Candidates must have an university degree in Linguistics, Computer Science or related field.

Corpus500 (instructions available at www.unlweb.net/wiki/Corpus500) is a set of 500 structures that cover the most frequent semantic structures of UNL. CLEA (instructions available at   www.unlw eb.net/wiki/CLEA
) is the Certificate for Language Engineering Aptitude and may be pursued online at VALERIE - the Virtual Learning Environment for UNL.

APPLICATION

In order to apply, candidates must send a CV to r.martins@undlfoundation.org before September 1st, 2012. Candidates must also prove their aptitude for natural language processing in the UNL framework by sending the following files in plain text format with UTF-8

encoding (LID should be replaced by the ISO 639-2 three-character code for the intended language). All the instructions are available at www.unlweb.net/wiki/Corpus500 .

1. LID_corpus500.txt, with the human translation, to the target language, of the sentences of the Corpus 500;
2. LID_ana_dic.txt, with the natural language analysis dictionary used to UNL-ize the translated version of the Corpus 500 (with IAN);
3. LID_ana_tgrammar.txt, with the transformation grammar used to UNL-ize the translated version of the Corpus 500 (with IAN);
4. LID_ana_dgrammar.txt, with the disambiguation grammar used to UNL-ize the translated version of the Corpus 500 (with IAN), if any;
5. LID_gen_dic.txt, with the natural language generation dictionary used to NL-ize the UNL version of the Corpus 500 (with EUGENE);
6. LID_gen_tgrammar.txt, with the transformation grammar used to NL-ize the UNL version of the Corpus 500 (with EUGENE), including the morphological module (inflectional paradigms);

7. LID_gen_dgrammar.txt, with the disambiguation grammar used to NL-ize the UNL version of the Corpus 500 (with EUGENE), if any.

SELECTION

The UNDL Foundation will select 15 candidates, one per language, according to the best F-measure (weighted average of the precision and recall) of the analysis and generation modules. In case two or more candidates provide modules equally good in terms of analysis and generation, the selection process will consider, in this order: 1) previous participation in any UNL School; 2) strongest experience (in terms of UNLdots) in the UNLweb; 3) strongest experience in natural language processing; and 4) highest academic degree.

VENUE

Library of Alexandria, in Alexandria, Egypt.

SUPPORT

The UNDL Foundation will pay the travel and accommodation expenses for the selected candidates not living in Alexandria. These include:

1. a round-trip plane, bus or train ticket from/to Alexandria
2. 7 nights at a hotel in Alexandria
3. 7 per diem of EGP300.00 (total of EGP2,100.00)

PRE-WORKSHOP TASKS (to be completed before 30/09/2012)

Before the workshop, the participants are expected to translate the workshop corpus (50 sentences) from English into their respective languages and to provide the corresponding dictionary entries (for natural language analysis and generation) according to the formalism described at www.unlweb.net/wiki/index.php/Dictionary_Specs .

WORKSHOP ACTIVITIES (7-11/10/2012)

During the workshop, the participants are expected to provide the syntactic and semantic modules of the grammar necessary to generate the workshop corpus from UNL into their native language, and from their native language into UNL. The grammar is expected to comply with the formalism described at www.unlweb.net/wiki/index.php/Grammar_Specs , and will be provided through the UNLdev, a web-based integrated development environment for creating and editing dictionary entries and grammar rules for natural language processing. The UNDL Foundation will provide all the training and support necessary for the accomplishment of the tasks.

POST-WORKSHOP TASKS (to be completed before 15/01/2013)

After the workshop, the participants will be invited to extend their experimental grammars (for both analysis and generation) for the Corpus 1000, which is an extension of the Corpus 500.

FOLLOW-UP

The participants who accomplish the post-workshop tasks before January 15th, 2013, will be invited to sign a contract for the development of the grammar modules for their respective languages. This contract will be based on Corpus50000 and will include:

1. Lemmatization of all word forms appearing in the corpus (USD0.10 per lemma, up to USD1,000.00 total, i.e., 10,000 lemmas)
2. The development of a grammar for the analysis of the 1,000 most frequent syntactic structures of the source language, extracted from the corpus (USD2,000.00)
3. The development of a grammar for the generation of the 1,000 most frequent semantic structures of UNL (USD2,000.00)
4. The participation in the Advanced Level Grammar Workshop in 2013

CERTIFICATION

The UNDL Foundation will issue a Certificate of Participation, upon evaluation, for all the participants.

THE UNL AND THE UNDL FOUNDATION

The UNDL Foundation is a non-profit organization based in Geneva, Switzerland, which has received, from the United Nations, the mandate for implementing the Universal Networking Language (UNL). The UNL is an artificial language that has been used for several different tasks in natural language engineering, such as machine translation, multilingual document generation, summarization, information retrieval and semantic reasoning. It has been, since 1996, a unique initiative to reduce language barriers and strengthen cross-cultural communication in the framework of the UN.

LOCAL ORGANIZATION

[Prof. Dr. Sameh Alansary](#) (University of Alexandria, Library of Alexandria)

FURTHER INFORMATION

For further information, please contact:

  Ronaldo Martins, PhD Language Resources Manager UNDL Foundation r.martins@undlfoundation.org