

Today's Topics:

1. ELRA - Language Resources Catalogue - Update (info@elda.org)
2. Proposal for Creation of ACL Special Interest Group (SIG) for MT (Alon Lavie)
3. Vacancy POSTDOC (2 years), Statistical NLP (Khalil Sima'an)

----- Forwarded message -----

From: info@elda.org
To: destinataires inconnus;
Date: Mon, 15 Dec 2008 16:09:54 +0100
Subject: [Mt-list] ELRA - Language Resources Catalogue - Update
Our apologies if you have received multiple copies of this announcement.

ELRA - Language Resources Catalogue - Update

New!

A more robust search engine has been integrated to our catalogues. You may now use either the new standard search from the catalogues' main page or the extended search controlled thanks to three drop-down lists ("Match", "Format" and "Sort"). Direct access to the controlled search is given below:
- ELRA Catalogue: <http://catalog.elra.info/search.php>
- Universal Catalogue: <http://universal.elra.info/search.php>

ELRA is happy to announce that 1 new WordNet database and 2 new Speech Resources are now available in its catalogue:

ELRA-M0047 Czech WordNet

The Czech WordNet captures nouns, verbs, adjectives, and partly adverbs, and contains 28,201 word senses (synsets). Every synset encodes the equivalence relation between several literals (at least one is present), having a unique meaning, belonging to one and the same part of speech, and expressing the same lexical meaning. Each Czech synset is related to the corresponding synset in the Princeton WordNet 2.0. via its identification number ID. There is at least one language-internal relation between a synset and another synset in the database.
For more information, see: http://catalog.elra.info/product_info.php?products_id=1089

ELRA-S0294 CHIEDE Corpus: a spontaneous child language corpus of Spanish

The spontaneous child language corpus, CHIEDE, consists of 58,163 words, in 30 texts, with 7 hours and 53 minutes of recordings and 59 child participants. About a third of the whole corpus is formed by child language and the remaining two thirds by adult speech. The main feature of CHIEDE is the interactions spontaneity: texts are recordings of communicative situations in their natural context.
For more information, see: http://catalog.elra.info/product_info.php?products_id=1090

ELRA-S0295 LILA Korean database

The LILA Korean database comprises 1,000 Korean speakers (500 males and 500 females) recorded over the Korean mobile telephone network. Each speaker uttered around 60 read and spontaneous items.
For more information, see: http://catalog.elra.info/product_info.php?products_id=1091

For more information on the catalogue, please contact Valérie Mapelli <mailto:mapelli@elda.org>

Visit our On-line Catalogue: <http://catalog.elra.info>

Visit the Universal Catalogue: <http://universal.elra.info>

Archives of ELRA Language Resources Catalogue Updates: <http://www.elra.info/LRs-Announcements.html>

----- Forwarded message -----

From: Alon Lavie <alavie@cs.cmu.edu>
To: NIST MT List <mt_list@nist.gov>, EAMT MT List <mt-list@eamt.org>, SIGParse List <sigparse-list@lists.andrew.cmu.edu>
Date: Tue, 06 Jan 2009 12:12:30 -0500
Subject: [Mt-list] Proposal for Creation of ACL Special Interest Group (SIG) for MT
Dear friends and colleagues,

We are contacting you to recruit your support for an initiative to create a new special interest group (SIG) for Machine Translation within the Association for Computational Linguistics (ACL). MT is already a very active area of research within ACL, with significant representation within the various

ACL meetings and conferences in the form of workshops, tutorials and tracks within the main conference research programs. The WMT-08 workshop at ACL-08 in Columbus this year, for example, attracted an attendance of around 150 people, and was by far the most popular and highly attended satellite event at the conference.

The main rationale for establishing a new ACL SIG for MT is to create a central focal point for planning and coordinating this growing set of MT activities within ACL. A secondary objective is to create a central entity within ACL that can explore opportunities for greater cooperation and collaboration with various other MT organizations, such as the regional MT associations in Europe (EAMT), Asia (AAMT) and North America (AMTA), as well as the growing commercial MT industry. ACL SIGs that organize large events have some additional privileges. One important advantage is that such SIGs receive preference in requests for scheduling SIG-sponsored annual workshops at specific ACL conferences. This should help us coordinate and schedule future MT workshops at ACL conferences. SIGs also have some financial privileges. They are allowed to have a "shadow account" at ACL, which can provide the SIG funds for special expenses at SIG-organized events, which are not normally covered by the main ACL organization.

The organizational structure that we are proposing for the ACL SIG-MT would consist of a SIG President, Secretary, and three Regional SIG officers (one for each of the three ACL regional chapters). The primary responsibilities of the regional officers would be to promote SIG-MT activities in their region and to coordinate activities with the ACL regional chapters and with the regional MT associations (EAMT, AAMT and AMTA). Elections for the officer positions would be held every three years. A founding board of officers will run the operations of the SIG in the first year, until the first elections can be held. Philipp Koehn, David Chiang and Chris Callison-Burch have agreed to serve in the positions of President, Secretary and Regional Officer for North America (respectively) during this period. We hope to select the other two regional officers as soon as possible.

We have already floated the idea of establishing a SIG-MT for ACL with several members of the ACL exec, and have received positive responses. In order to follow the formal procedure for proposing a new SIG, however, a formal proposal needs to be submitted to the ACL exec, with expressed support from at least 25 ACL members. The proposal must also contain a proposed SIG constitution, consistent with ACL's general SIG guidelines. A draft of such a constitution (based on constitutions of other ACL SIGs) is attached to this email message.

If you participated (and registered) in any ACL conference during 2008, you are considered a member of ACL in good standing. To all such ACL members - we would appreciate an explicit response of support (or disapproval) from each of you, along with any comments or suggestions regarding the above proposal. If possible, please respond within the next week or so, as we would like to then proceed with formally proposing the SIG to the ACL exec.

Thanks,

Philipp Koehn
David Chiang
Chris Callison-Burch
Alon Lavie

SIG-MT Constitution

I. Statement of Purpose

The purpose of the Association for Computational Linguistics (ACL) Special Interest Group on Machine Translation (SIG-MT) shall be:

- To promote interest in the research area of Machine Translation (MT).
- To provide members of the ACL having a special interest in MT with a means for exchanging news and views on recent research developments and other matters of interest to research in MT.
- To coordinate and facilitate the organization of annual research workshops and meetings in the area of MT.
- To coordinate and expand cooperation and collaboration of the ACL MT research community with other MT associations, organizations, such as the regional MT associations AMTA, EAMT and AAMT, and with commercial MT companies.
- To sponsor other meetings and workshops related to MT research that appear to be timely and worthwhile.

II. Membership

Membership is open to any ACL member who professes an interest in Machine Translation. Membership will be terminated on request, or automatically if no reply is received within six months of posting a membership list update request.

III. Elected Officers

The elected officers of the SIG shall consist of a President, a Secretary and three Regional Information Officers, one each for the geographical regions of Europe, Asia and the Americas. The term of all elected officers of the SIG shall be three years.

The duties of the President shall be:

- To have primary executive authority over actions and activities of the SIG.
- To prepare a written report on the activities of the SIG for the Executive Committee of the ACL, for presentation to the ACL at its Annual Business Meeting.
- To designate such other officers as may be needed from time to time to fulfill the goals of the SIG, in particular a returning officer as specified in Sections IV and VI below.

The duties of the Secretary shall be:

- To maintain a membership roster of the SIG.
- To monitor and keep track of any funds awarded to the SIG by the ACL; to collect and manage any dues that may be required by the organization; and to present a written annual report on the SIG finances to the Executive Committee of the ACL.
- To maintain, in consultation with the President and the Regional Information Officers, a document stating the aims of the group and the names and affiliations of the current officers, suitable for sending to prospective members and other inquirers.

The duties of the Regional Information Officers shall be:

- To communicate with members of the SIG in their regional area.
- To answer inquiries about the SIG.
- To act as representatives of the SIG in communications with other MT organizations in their region.
- Any additional agreed upon specific duties assigned to them by the SIG President.

IV. Election of Officers

All officers of the SIG shall be elected by a vote of its membership. Elections shall normally take place at least one month before the expiration of current terms of office. In the event of a casual vacancy, an extraordinary election shall be held for an officer to complete the current term of office. Candidates must be nominated by two members, must accept nomination, and must be members in good standing of both ACL and SIG-MT. Votes may be entered by electronic or physical mail, and votes arriving by the notified closing date will be counted by a returning officer appointed by the President and agreed to by all candidates. Nominations shall be called at least three months prior to the election, and shall close two months prior to the election. All members shall be notified of the election date and the nominations at least four weeks prior to the close of voting, and results shall be notified to the members within four weeks after the close of voting. Notification is considered delivered when dispatched by the medium and to the address last notified by a member to the secretary.

V. Dues

SIG-MT membership is currently free of any membership dues. Subject to the approval of the Executive Committee of the ACL, the SIG may decide to levy annual membership dues, where the rate is determined by the SIG-MT Executive Committee.

VI. Referenda

A change to this constitution, or a motion on any other matter, may be initiated by five SIG members (petitioners) in writing or by agreement of the President and Secretary (on the suggestion of a member or otherwise). The motions will be put to a vote of the membership within six months of receipt of a petition, if the petitioners so request. Otherwise, constitutional amendments and other motions will be put to the vote in conjunction with regular elections. Members will be given notification at least three months in advance of any referendum, and further motions or amendments proposed by two members shall be accepted up to two months before the close of voting. Votes may be entered by electronic or physical mail, and votes arriving by the notified closing date will be counted by a returning officer appointed by the President and agreed to by all petitioners. All members shall be notified of the date of the referendum and the proposed motions and amendments at least four weeks prior to the close of voting, and results shall be notified to the members within four weeks of the close of voting. Any change to this constitution requires a two-thirds majority of the SIG-MT membership and is subject to approval by ACL. Any other matter requires a simple majority of those

actually voting. Notification is considered delivered when dispatched by the medium and to the address last notified by a member to the secretary.

----- Forwarded message -----

From: "Khalil Sima'an" <k.simaan@uva.nl>

To: "mt-list@eamt.org" <mt-list@eamt.org>

Date: Thu, 08 Jan 2009 12:24:00 +0100

Subject: [Mt-list] Vacancy POSTDOC (2 years), Statistical NLP

Detail to be found on: Vacancy postdoc researcher, 2 years, computational linguistics/NLP

Vacancy

The Institute for Logic, Language and Computation, University of Amsterdam, has a vacancy for a Postdoctoral Researcher

1.0 fte (38h per week)

For both internal and external candidates

Summary

- * Position: POSTDOC researcher
- * Duration: 2 years (full time, 38hrs per week)
- * Salary (gross per month): minimum Euro 2379 and maximum Euro 3755 in the first year.
- * The Collective Employment Agreement of the Dutch Universities is applicable.
- * Last date for application: February 15, 2009.

Project

Priors for the Estimation of Probabilistic Grammars from Incomplete Natural Language Data
VIDI project Sima'an [2007-2011]

Abstract: The project aims at exploring the utility of available prior knowledge in specifying the model (hypothesis space) and the constraints on the objective estimation function for unsupervised estimation of probabilistic grammars for Statistical Machine Translation and Statistical Parsing. This project addresses problems of unsupervised estimation in the face of different kinds of incomplete data (parallel and monolingual corpora):

1. A raw text corpus,
2. A partially annotated corpus possibly with negative examples, and
3. A raw in-domain corpus and an out-of-domain annotated corpus.

Corpus annotation is e.g., word-alignment, syntactic or semantic structure.

Example topics studied within the project (see project related publications):

- * Statistical machine translation with hierarchical/syntactic structure, e.g., statistical estimation [Mylonakis and Sima'an 2007 and 2008], lexical syntactic structure [Hassan, Way and Sima'an 2007 and 2008]

- * Statistical machine translation with morphological structure -- Work in progress 2009,
- * Inducing hidden compositional structure for machine translation (possibly using syntactic structure) -- Work in progress 2009.

For more detail see: Webpage

Requirements

We seek a cooperative yet independent researcher who is passionate about empirical research, statistical methods and NLP/CL.

The following are the main requirements:

- * obtained a Ph.D. recently (or will obtain it in the very near future, in which case testimonial from the PhD thesis Supervisor is needed)
- * a strong background, experience in and affinity with the largest subset of the following areas:

1. Computational Linguistics/Natural Language Processing working with statistical methods
2. Machine Learning Applied to NLP with good knowledge of unsupervised statistical methods.
3. Statistical Machine Translation.
4. Unsupervised statistical methods for statistical Parsing.

Appointment

The appointment will be on a temporary basis for a maximum period of 2 years. Based on a full-time appointment (38 hours per week) the gross monthly salary will range from € 2379,- to, at the most, € 3755,- , in the first year depending on expertise and experience.

The Collective Employment Agreement of the Dutch Universities is applicable.

Application

To Dr Khalil Sima'an by email (k.simaan@uva.nl with subject heading: VIDDI-POSTDOC).

The following documents (in pdf):

1. Letter expressing interest
2. Curriculum Vitae
3. List of publications (possibly including a downloadable pdf of your thesis)
4. Three letters of reference sent directly by the referees to Dr Khalil Sima'an
5. Optional: Research plan (1 page)

Last date for application: February 15, 2009.

For inquiries contact

Dr Khalil Sima'an.
Language and Computation
Computer Science
University of Amsterdam

--

Khalil Sima'an

Institute for Logic, Language and Computation (ILLC)
University of Amsterdam
Tel#: +31-20-525 6573
Fax#: +31-20-525 5206
<http://staff.science.uva.nl/~simaan>

Address:
Plantage Muidersgracht 24
1018 TV Amsterdam
The Netherlands