Named-Entity Recognition (NER) is still a challenging task for languages with low digital resources. The main difficulties arise from the scarcity of annotated corpora and the consequent problematic training of an effective NER pipeline. To abridge this gap, in this paper we target the Persian language that is spoken by a population of over a hundred million people world-wide. We first present and provide ArmanPerosNERCorpus, the first manually-annotated Persian NER corpus. Then, we introduce PersoNER, an NER pipeline for Persian that leverages a word embedding and a sequential max-margin classifier. The experimental results show that the proposed approach is capable of achieving interesting MUC7 and CoNNL scores while outperforming two alternatives based on a CRF and a recurrent neural network.

1 Introduction

Named-Entity Recognition (NER), introduced in the sixth Message Understanding Conference (MUC-6) (Grishman and Sundheim, 1996), concerns the recognition of Named Entities (NE) and numeric expressions in unstructured text. Since 1996, great effort has been devoted to NER as a foundational task for higher-level natural language processing tasks such as summarization, question answering and machine translation.

Shortage of gold standards has initially limited NER investigation to high-resource languages such as English, German and Spanish (Tjong Kim Sang and De Meulder, 2003). Gradually, publicly available encyclopedias have enabled combinations of semi-supervised and distant supervision approaches for other languages (Althobaiti et al., 2015). However, low-resource languages still face a significant scarcity of public repositories. For instance, only 8.8% of Wikipedia articles in Hindi are identified as entity-based articles in Freebase (Al-Rfou et al., 2015). In this work, we aim to enable supervised NER for a low-resource language, namely Persian, by providing the first manually-annotated Persian NE dataset. The Persian language, despite accounting for more than a hundred million speakers around the globe, has been rarely studied for NER (Khormuji and Bazrafkan, 2014) and even text processing (Shamsfard, 2011). In addition, we present PersoNER, a Persian NER pipeline consisting of a word embedding module and a sequential classifier based on the structural support vector machine (Tsochantaridis et al., 2005). The proposed pipeline achieves interesting MUC7 and CoNNL scores and outperforms two alternatives based on a CRF and a recurrent neural network.

2 Related Work

Early research on NER was mostly devoted to handcrafted rule-based systems which are intrinsically language-dependent, and thus laborious to be extended to new languages. As a consequence, recent studies are mainly focused on language-independent machine learning techniques that attempt to learn...
statistical models for NER from data (Nadeau and Sekine, 2007). Moreover, replacement of manually-
annotated gold standards with very large “silver standard” corpora mollifies the scarcity of supervised
data. Silver standards are NE annotated corpora derived from processing Wikipedia’s text and meta-
information alongside entity databases such as Freebase (Nothman et al., 2013; Al-Rfou et al., 2015).

Existing NER approaches mainly divide over two categories: in the first, the task is decoupled into
an initial step of word embedding, where words are mapped to feature vectors, followed by a step of
word/sentence-level classification. The feature vector can be as simple as a binary vector of text features
like ‘word is all uppercased’ or a more complex, real-valued vector capturing semantic and syntactic
aspects of the word. Word2vec (Mikolov et al., 2013), GloVe (Pennington et al., 2014) and Hellinger-
PCA (Lebret and Collobert, 2014) are well-known examples of unsupervised word embeddings applied
successfully to the NER task. For classification, sequential classifiers such as HMMs (Zhou and Su,
2002), CRFs (Lafferty et al., 2001; Finkel et al., 2005) and deep neural networks (Al-Rfou et al., 2015)
have been amongst the most popular choices.

The second category, proposed by (Collobert et al., 2011) and recently followed by many includ-
ing (Mesnil et al., 2013; Mesnil et al., 2015) and others, leverages recurrent neural networks (RNNs) to
deliver end-to-end systems for NER. With this approach, an implicit word embedding is automatically
extracted in the network’s early layers by initializing the training with random values or a preliminary
embedding. In this paper, we apply and compare approaches from both categories.

3 The Proposed Approach

The workflow of PersoNER is illustrated in Figure 1. The steps include data collection, text normaliza-
tion, word embedding and entity classification. In this section, we focus on the two technical modules,
word embedding and classification, while data collection and text normalization are described in Sec-
tion 4.

3.1 Word Embedding

Term-frequency (tf), term-frequency inverse-document-frequency (tf-idf), bag of words (bow) and word
cocurrence are general statistics intended to characterize words in a collection of documents. Out of
them, word co-occurrence statistics have the ability to represent a word by the frequencies of its sur-
rounding words which well aligns with the requirements of NER. Recently, Lebret and Collobert (2014)
have shown that a simple spectral method analogous to PCA can produce word embeddings as useful
as those of neural learning algorithms such as word2vec. Given an unsupervised training corpus and a
vocabulary, \( V \), the co-occurrence matrix, \( C_{|V|\times|D|} \), in (Lebret and Collobert, 2014) is computed as:

\[
C(v_i, d_j) = p(d_j|v_i) = \frac{n(v_i,d_j)}{\sum_d n(v_i,d)}
\]
where \(v_i \in V; i = 1 \ldots |V|\) and \(d_j \in D \subseteq V; j = 1 \ldots |D|\). \(n(v_i,d_j)\) is the count of occurrences of context word \(d_j\) in the neighborhood of reference word \(v_i\). Thus, \(C(v_i,:)\) represents discrete probability distribution \(p(d|v_i)\) and is used to characterize \(v_i\). Since words are represented as discrete distributions, Lebret and Collobert (2014) argue that it is more appropriate to measure their distances in a Hellinger space. Accordingly, \(H(C)\) is the transformation of \(C\) into Hellinger space where the distance between any two discrete probability distributions, \(P\) and \(Q\), is given by:

\[
dist(P,Q) = \frac{1}{\sqrt{2}}||\sqrt{P} - \sqrt{Q}||_2. \tag{2}
\]

Eventually, PCA is applied to reduce the dimensionality of \(H(C) \in \mathbb{R}^{|V| \times |D|}\) to \(h(C) \in \mathbb{R}^{|V| \times m}\), where \(m \ll |D|\).

3.2 Classification

In this subsection, we first briefly introduce sequential labeling as a formal problem and then describe the sequential classifier based on the structural support vector machine.

3.2.1 Sequential Labeling

Sequential labeling predicts a sequence of class labels, \(y = \{y_1, \ldots, y_T\}\), based on a corresponding sequence of measurements, \(x = \{x_1, \ldots, x_T\}\). It is a very common task in NLP for applications such as chunking, POS tagging, slot-filling and NER. A widespread model for sequential labeling is the hidden Markov model (HMM) that factorizes the joint probability of the measurements and the labels, \(p(x,y)\), by arranging the latter in a Markov chain (of order one or above) and conditioning the measurement at frame \(t\) on only the corresponding label. For an HMM of order one, \(p(x,y)\) is expressed as:

\[
p(x,y) = p(y_1) \prod_{t=2}^{T} p(y_t|y_{t-1}) \prod_{t=1}^{T} p(x_t|y_t) \tag{3}
\]

where \(p(y_1)\) is the probability of the initial class, terms \(p(y_t|y_{t-1})\) are the transition probabilities and terms \(p(x_t|y_t)\) are the emission, or measurement, probabilities. By restricting the emission probabilities to the exponential family, i.e., \(p(x_t|y_t) \propto \exp(w^T f(x_t, y_t))\), the logarithm of probability \(p(x,y)\) can be expressed as the score of a generalized linear model:

\[
\ln p(x,y) \propto w^T \phi(x,y) = w_{in} f(y_1) + \sum_{t=2}^{T} w_{tr}^T f(y_t, y_{t-1}) + \sum_{t=1}^{T} w_{em}^T f(x_t, y_t) \tag{4}
\]

where \(w_{in}\), \(w_{tr}\) and \(w_{em}\) are the linear models for assigning a score to the initial classes, transitions and emissions, respectively. Functions \(f(y_1), f(y_t, y_{t-1})\) and \(f(x_t, y_t)\) are arbitrary, fixed “feature” functions of the measurements and the labels.

The generalized linear model in (4) is more suitable for discriminative training than the generative probabilistic model in (3). Notable discriminative approaches are conditional random fields (CRFs) (Lafferty et al., 2001) and structural SVM (Tsochantaridis et al., 2005). In particular, structural SVM has built a very strong reputation for experimental accuracy in NLP tasks (Joachims et al., 2009; Tang et al., 2013; Qu et al., 2014) and for this reason we exploit it in our NER pipeline.

Eventually, given a measurement sequence \(x\) in input, inference of the optimal label sequence can be obtained as:

\[
y = \arg \max_y p(x,y) = \arg \max_y (w^T \phi(x,y)) \tag{5}
\]

This problem can be efficiently solved in \(O(T)\) time by the Viterbi algorithm working in either the linear or logarithmic scale (Rabiner, 1989).
3.2.2 Structural SVM

From a supervised training set of sequences, \( \{X,Y\} = \{x^i, y^i\}, i = 1 \ldots N \), structural SVM finds the model’s parameters, \( w \), by minimizing the usual SVM trade-off between the hinge loss and an \( L^2 \) regularizer (Tsochantaridis et al., 2005). Its learning objective can be expressed as:

\[
\arg\min_{w, \xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi^i \quad \text{s.t.} \\
w^T \phi(x^i, y^i) - w^T \phi(x^i, y) \geq \Delta(y^i, y) - \xi^i, \quad i = 1 \ldots N, \quad \forall y \in \mathcal{Y}
\]  

(6)

In the objective function, the first term is the regularizer while the second term, \( \sum_{i=1}^{N} \xi^i \), is the hinge loss, i.e. a convex upper bound over the total loss on the training set. Hyperparameter \( C \) is an arbitrary, positive coefficient that balances these two terms. In the constraints, \( w^T \phi(x, y) \) computes the generalized linear score for a \((x, y)\) pair. In the case of sequential labeling, such a score is given by Eq. (4).

Eventually, \( \Delta(y^i, y) \) is the loss function chosen to assess the loss over the training set.

For an NER task with \( M \) entity classes, each sequence of length \( T \) adds \( (M + 1)T \) constraints to (6). Due to their exponential number, exhaustive satisfaction of all constraints is infeasible. However, (Tsochantaridis et al., 2005) has shown that it is possible to find \( \epsilon \)-correct solutions with a subset of the constraints of polynomial size consisting of only the “most violated” constraint for each sequence, i.e. the labeling with the highest sum of score and loss:

\[
\xi^i = \max_y \left( -w^T \phi(x^i, y^i) + w^T \phi(x^i, y) + \Delta(y^i, y) \right)
\]

\[
\rightarrow \bar{y}^i = \arg\max_y (w^T \phi(x^i, y) + \Delta(y^i, y))
\]  

(7)

This problem is commonly referred to as “loss-augmented inference” given its resemblance with the common inference of Eq. (5) and is the core of structural SVM. In the case of scores and losses that can be computed frame by frame (such as the 0-1 loss or the Hamming loss), the Viterbi algorithm with appropriate weights can still be used to compute the loss-augmented inference in \( O(T) \) time.

4 Data Collection

In this section, we describe the collection and preprocessing of the Persian corpora. The datasets consist of 1) an unsupervised corpus, called PersoSentencesCorpus, that we use for the word embedding module and 2) a manually named-entity annotated data set of Persian sentences, called ArmanPersoNERCorpus, that we use for supervised classification. Alongside this publication, we release ArmanPersoNERCorpus\(^1\) as the first ever publicly-available Persian NER dataset.

4.1 PersoSentencesCorpus

A very large corpus of documents covering a variety of contexts is required to populate an effective co-occurrence matrix. We fulfill this requirement by accumulating the following three datasets of Persian sentences:

- The Leipzig corpora\(^2\) with 1,000,000 sentences from news crawling and 300,000 from Wikipedia.
- The VOA\(^3\) news dataset with 277,000 sentences.
- The Persian Dependency Treebank\(^4\) with 29,982 sentences (Rasooli et al., 2013).

The aggregated corpus, called PersoSentencesCorpus, holds more than 1.6 million sentences and seems of adequate size to train the co-occurrence matrix.

\(^1\)http://poostchi.info/hanieh/NLP/ArmanPersoNERCorpus.txt
\(^2\)http://corpora2.informatik.uni-leipzig.de/download.html
\(^3\)http://www.ling.ohio-state.edu/~jonsafari/corpora/index.html\#persian
\(^4\)http://dadegan.ir/en/perdt/
4.2 ArmanPersoNERCorpus

To create an NE dataset, in collaboration with ArmanSoft\(^5\), we have decided to manually annotate NEs in a subset of the BijanKhan\(^6\) (Bijankhan et al., 2011) corpus which is the most-established tagged Persian corpus, yet lacking entity annotation. We selected the subset from news sentences since they are the most entity-rich. Before the annotation, a comprehensive manual was designed based on the definition of Sekine’s extended named entities (Sekine, 2007) adapted to the Persian Language. The annotation task was led by an experienced lead annotator who instructed the front-end annotators (two native post-graduate students) and revised their annotations. The guidelines were very clear and we expected minimal subjectivity. We have verified this hypothesis in two ways: by a sample of 500 already annotated NEs chosen randomly, and by another sample of 500 already annotated NEs from the two most semantically-close classes (location and organization). Both samples were revised by three other, independent native annotators and the percentages of corrections have been only 1.8% and 1.9%, respectively.

All NEs have been annotated in IOB format. The annotated dataset, ArmanPersoNERCorpus, contains 250,015 tokens and 7,682 sentences (considering the full-stop as the sentence terminator). It can be used to train NER systems in future research on Persian NER, but it also offers an ideal test set for evaluation of NER systems trained on silver standards. The NEs are categorized into six classes: person, organization (such as banks, ministries, embassies, teams, nationalities, networks and publishers), location (such as cities, villages, rivers, seas, golfs, deserts and mountains), facility (such as schools, universities, research centers, airports, railways, bridges, roads, harbors, stations, hospitals, parks, zoos and cinemas), product (such as books, newspapers, TV shows, movies, airplanes, ships, cars, theories, laws, agreements and religion), and event (such as wars, earthquakes, national holidays, festivals and conferences); other are the remaining tokens. It is worth noting that annotation was not trivial since individual tokens have been categorized according to the context. For instance, “Tokyo” is a different type of entity in sentence “Tokyo\(_{loc}\) is a beautiful city” versus sentence “London\(_{org}\) and Tokyo\(_{org}\) sign flight agreement”. Table 1 summarizes the number of tokens for each entity class in ArmanPersoNERCorpus.

Table 1: Class percentages in ArmanPersoNERCorpus.

<table>
<thead>
<tr>
<th>Entity type</th>
<th>Person</th>
<th>Organization</th>
<th>Location</th>
<th>Facility</th>
<th>Event</th>
<th>Product</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Tokens (NT)</td>
<td>5,215</td>
<td>10,036</td>
<td>4,308</td>
<td>1,485</td>
<td>2,518</td>
<td>1,463</td>
<td>224,990</td>
</tr>
<tr>
<td>Percentage</td>
<td>2.08%</td>
<td>4.01%</td>
<td>1.72%</td>
<td>0.59%</td>
<td>1.00%</td>
<td>0.58%</td>
<td>89.99%</td>
</tr>
<tr>
<td>Number of Unique-Tokens (NUT)</td>
<td>1,829</td>
<td>1,290</td>
<td>832</td>
<td>548</td>
<td>556</td>
<td>634</td>
<td>15,677</td>
</tr>
<tr>
<td>Percentage (NUT/NT)</td>
<td>35.07%</td>
<td>12.85%</td>
<td>19.31%</td>
<td>36.90%</td>
<td>22.08%</td>
<td>43.33%</td>
<td>6.96%</td>
</tr>
</tbody>
</table>

4.3 Text Normalization

As the preprocessing phase, the PersoSentencesCorpus has been normalized and tokenized following the approach proposed in (Feely et al., 2014) that suggests applying a pipeline of useful tools to deal with written Persian. The pipeline starts with PrePer (Seraji, 2013) which maps Arabic specific characters to their Persian Unicode equivalent. In addition, it replaces the full space between a word and its affix with a zero-width-non-joiner character. Then, a Farsi text normalizer (Feely, 2013) omits Arabic and Persian diacritics and unifies variant forms of some Persian characters to a single Unicode representation. Finally,
<table>
<thead>
<tr>
<th>Ezāfe</th>
<th>POS-tag</th>
<th>Inflexion</th>
<th>Token NER-tag</th>
<th>Transliteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>N</td>
<td>اسامه</td>
<td>ب-PERS</td>
<td>Seved</td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>محمود</td>
<td>I-PERS</td>
<td>Mahmoud</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>محدث</td>
<td>I-PERS</td>
<td>Mohadess</td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>مسئول</td>
<td>O</td>
<td>manager</td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>اکتشاف</td>
<td>O</td>
<td>discovery</td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>شرکت</td>
<td>B-ORG</td>
<td>Company</td>
</tr>
<tr>
<td>EZ</td>
<td>AJ</td>
<td>ملی</td>
<td>I-ORG</td>
<td>National</td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>نفت</td>
<td>I-ORG</td>
<td>Oil</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>ایران</td>
<td>I-ORG</td>
<td>Iranian</td>
</tr>
<tr>
<td>O</td>
<td>P</td>
<td>در</td>
<td>O</td>
<td>in</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>مصاحبه</td>
<td>O</td>
<td>interview</td>
</tr>
<tr>
<td>O</td>
<td>P</td>
<td>با</td>
<td>O</td>
<td>with</td>
</tr>
<tr>
<td>EZ</td>
<td>AJ</td>
<td>واحد</td>
<td>B-ORG</td>
<td>Unit</td>
</tr>
<tr>
<td>EZ</td>
<td>AJ</td>
<td>مرکزی</td>
<td>I-ORG</td>
<td>Central</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>خبر</td>
<td>I-ORG</td>
<td>News</td>
</tr>
<tr>
<td>O</td>
<td>P</td>
<td>با</td>
<td>O</td>
<td>with</td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>اعلام</td>
<td>O</td>
<td>declaring</td>
</tr>
<tr>
<td>O</td>
<td>DET</td>
<td>DET</td>
<td>O</td>
<td>this</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>سرگرمی</td>
<td>O</td>
<td>announcement</td>
</tr>
<tr>
<td>O</td>
<td>V</td>
<td>افزودن</td>
<td>O</td>
<td>adds</td>
</tr>
<tr>
<td>O</td>
<td>PUNC</td>
<td>DELM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>هفته</td>
<td>O</td>
<td>diggings</td>
</tr>
<tr>
<td>O</td>
<td>AJ</td>
<td>بیشتر</td>
<td>O</td>
<td>more</td>
</tr>
<tr>
<td>O</td>
<td>P</td>
<td>در</td>
<td>O</td>
<td>in</td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>میدان</td>
<td>B-LOC</td>
<td>field</td>
</tr>
<tr>
<td>EZ</td>
<td>AJ</td>
<td>نقش</td>
<td>I-LOC</td>
<td>Changuleh</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>انتظار</td>
<td>O</td>
<td>expectation</td>
</tr>
<tr>
<td>O</td>
<td>V</td>
<td>داریم</td>
<td>O</td>
<td>have</td>
</tr>
<tr>
<td>EZ</td>
<td>N</td>
<td>ذخیره</td>
<td>O</td>
<td>reservoirs</td>
</tr>
<tr>
<td>O</td>
<td>DET</td>
<td>DET</td>
<td>O</td>
<td>In</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>میدان</td>
<td>O</td>
<td>field</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>افزایش</td>
<td>O</td>
<td>increase</td>
</tr>
<tr>
<td>O</td>
<td>V</td>
<td>پیدا</td>
<td>O</td>
<td>will</td>
</tr>
<tr>
<td>O</td>
<td>PUNC</td>
<td>DELM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: A snapshot of ArmanPersoNERCorpus.
5 Experiments

In this section, we report NER results based on the PersoSentencesCorpus and ArmanPersoNERCorpus datasets. The classification task is challenging given the much lower frequencies of the entity classes versus the non-entity class (other), as shown in Table 1. For this task, we have not used any of the additional linguistic information that is available from the dataset (such as POS tag, inflexion etc).

To calculate the co-occurrence matrix, $C$, we have used a context window of radius 5. The size of the dictionary, $V$, from the PersoSentencesCorpus is $|V| = 49,902$ and that of subset $D$ is $D = 7,099$, obtained by selecting only the words with count greater than 15. The word embedding matrix $h(C)$ has been computed by heuristically setting $m = 300$. For classification, each word has been encoded as a 3-gram that includes the previous and following feature vectors. All the models used for classification share the same word embeddings.

For classification, we have compared the proposed SVM-HMM with a CRF and a deep learning approach based on the Jordan-RNN (Mesnil et al., 2013). For the SVM-HMM we have used structural SVM from (Joachims, 2008) with a Markov chain of order 3 and learning constant $C = 0.5$. The CRF is from the HCRF library (Morency et al., 2010) and is trained with an $L2$ regularizer of weight 100. The Jordan-RNN is a recurrent neural network from (Mesnil et al., 2013) trained with 100 hidden states and initialized using the same features vectors. All parameters were chosen by 3-fold cross-validation over a reasonable range of values. The indices for the three folds are available in the dataset to allow for future result comparison. We have also tried continuous bag of words (Mikolov et al., 2013), skip-grams (Mikolov et al., 2013) and GloVe (Pennington et al., 2014) as embeddings, and the Elman-RNN (Mesnil et al., 2013) as classifier, but results have proved generally less accurate.

Table 2 shows the comparison of the average MUC7 and CoNLL scores from the 3-fold cross-validation for the three classifiers. The MUC7 and CoNLL scores are $F_1$ values adapted to the NER task, with the CoNLL score generally stricter than MUC7 (Nadeau and Sekine, 2007). As shown in Table 2, the scores achieved by the SVM-HMM are higher overall and for all classes but one, with the Jordan-RNN as the second best. To verify statistical significance, we have also run a paired t-test over the results from the six individual classes and confirmed statistical significance of the differences even at $p = 0.02$. The relative ranking between SVM-HMM and the CRF is supported by similar results in the literature, including (Nguyen and Guo, 2007; Tang et al., 2013; Lei et al., 2014), showing that regularized minimum-risk classifiers tend to outperform equivalent models trained under maximum conditional likelihood. The relative ranking between SVM-HMM and the RNN is instead somehow in contrast with the recent results in the literature, and a possible explanation for it is the relatively small size of the dataset compared to the number of free parameters in the models. We plan future comparative experiments with larger corpora to further probe this assumption.

6 Conclusion

In this paper, we have presented and released ArmanPersoNERCorpus, the first manually-annotated Persian NE dataset, and proposed an NER pipeline for the Persian language. The main components
of the pipeline are word embedding by Hellinger PCA and classification by a structural SVM-HMM classifier. Experiments conducted over the ArmanPersoNERCorpus dataset have achieved interesting overall \( F_1 \) scores of 72.59 (MUC7) and 65.13 (CoNNL), higher than those of a CRF and a Jordan-RNN. The released dataset can be used for further development of Persian NER systems and for evaluation of systems trained on silver-standard corpora, and the achieved accuracy will provide a baseline for future comparisons.

References


Buzhou Tang, Hongxin Cao, Yonghui Wu, Min Jiang, and Hua Xu. 2013. Recognizing clinical entities in hospital discharge summaries using structural support vector machines with word representation features. *BMC Medical Informatics and Decision Making*, 13(SUPPL.1).


Copyright of each paper stays with the respective authors (or their employers).

ISBN978-4-87974-702-0
Welcome to COLING 2016 – the 26th International Conference on Computational Linguistics — held in Osaka, Japan! It is the third COLING in Japan after Tokyo (1980) and Kyoto (1994). It is a special pleasure for me to be General Chair (10 years after chairing the joint COLING-ACL 2006 in Sydney) of a COLING held in Japan, a country I love.

COLING is organised under the auspices of the International Committee on Computational Linguistics (ICCL, http://nlp.shef.ac.uk/iccl/index.html). ICCL is a very special committee, with no fixed rules and no funding, whose only function is to make sure that a COLING appears every two years and that it is a good and friendly conference.

I have participated to many COLINGs, since the one in Pisa in 1973. It was a COLING without email! I still remember when Antonio Zampolli (Local chair) received by Hans Karlgren (Program chair) a sketch of the program written by hand, almost unreadable, and asked me (very young at the time) to interpret it. I have seen COLINGs where submissions arrived on paper and many packages were sent around the world to area chairs, to be sent to reviewers, and all the results back again by normal mail. It seems impossible now.

COLING has changed over the years, together with the changes in our field. But it has always been important for ICCL to maintain the COLING “spirit”: we always wanted COLING to be an inclusive and broad conference. We also want to underline that in our field “language” is important and we therefore pay special attention to having papers and workshops focusing on understanding language properties and complexities. Moreover, for us the social part of the conference is as important as the scientific one.

An outstanding competent and dedicated team has worked for the organisation of COLING 2016. I wish to warmly thank, also on behalf of ICCL, all the various Chairs, too many to mention them all here, for the wonderful work they have done. It has been a pleasure and a privilege for me to work together with all of them: they made my work as General chair very easy. But I owe a special thanks to Yuji Matsumoto and Rashmi Prasad, Program chairs, for their hard work in managing so smoothly an impressive number of submissions, many more than we expected. And I wish to express my deepest gratitude to the Local chairs – Eiichiro Sumita, Takenobu Tokunaga and Sadao Kurohashi – who have done a fantastic work with great dedication in all the various phases of the conference organisation, always keeping everything under control. Not an easy task, as I know too well!

I also want to thank the generosity of all the sponsors for their great support to COLING.

Last but not least, I thank the colleagues (so many) who submitted their work to COLING, the organisers of Workshops and Tutorials, the participants (more than 900 at the time of writing) and the many students among them. It is important that many young researchers can attend COLINGs. They show the great interest of our community in COLING.

I hope that you benefit not only from the scientific programme but also from the social parts of COLING. I hope you get from this COLING both new exciting ideas and also new friends.

Enjoy COLING 2016 in Japan!

Nicoletta Calzolari (ICCL, ILC-CNR and ELRA)
Preface: Program Chairs

It is with great pleasure that we welcome you to the 26th International Conference on Computational Linguistics (COLING 2016) in Osaka, Japan! COLING covers a broad spectrum of technical areas related to natural language and computation. This year, we received 1,039 valid submissions (from a total of 1127 submissions), of which we accepted 337 papers (32.4% acceptance rate). 134 papers were selected for oral presentation and 203 papers for poster presentation. No distinction is made in these proceedings between papers presented orally or as a poster, as they were not distinguished qualitatively but rather by judging the best mode for delivering the paper content.

To effectively cover the broad spectrum of topics included in the conference, we have 18 thematic areas, each chaired by two or more area chairs. We are extremely grateful to the area chairs, who led and monitored the reviewing and reviewer discussions, and sent us detailed recommendation reports resulting from the reviewing process, including best paper recommendations. We cannot thank enough the over 800 reviewers who have put in the requisite time and effort to carefully assess the very large number of submissions we received this year. Their dedication and commitment, and willingness to work with us even when there were tight time constraints, made the entire task proceed much more smoothly than we had hoped! Almost all papers were reviewed by at least three reviewers and we are very happy with the highly strong set of papers accepted for presentation. We thank all authors for their submissions describing their very commendable research, and hope that authors of papers we could not accept have nevertheless benefitted from the feedback they received from reviewers.

We have structured the accepted submissions into ten sessions, with multiple thematic areas included in parallel, either for oral presentation or poster presentation. Only one session – the first session – does not have a parallel poster session. We are delighted to have four invited speakers to the conference: Joakim Nivre from Uppsala University: “Universal Dependencies – Dubious Linguistics and Crappy Parsing?”; Reiko Mazuka from RIKEN Brain Science Institute & Duke University: “Getting the Input Right: Refining our Understanding of What Children Hear”; Dina Demner-Fushman from the U.S. National Library of Medicine: “NLP to support clinical tasks and decisions”, and Simone Teufel from University of Cambridge: “A Look at Computational Argumentation and Summarisation from a Text-Understanding Perspective”.

We are extremely grateful to the members of the best paper committee, Tim Baldwin, Vincent Ng, and Hinrich Schütze, who agreed to put in extra time to select the two best papers at the conference. Best paper nominations were collected in a bottom-up fashion, with reviewers first providing their recommendation for each paper, and area chairs then collecting the positive recommendations, and upon their own assessment of the corresponding reviews and papers, selecting some or all to be forwarded to the PC chairs. PC chairs then invited the three experts to form a committee (chaired by the PC chairs) to select the two best papers from this set of nominated papers.

We would like to thank the many members of the organizing committee who have helped us in crucial ways at various stages of organizing the technical program – the General Chair, Nicoletta Calzolari; the Local Chairs, Eiichiro Sumita, Takenobu Tokunaga and Sadao Kurohashi; the Publication Chairs, Hitoshi Isahara and Masao Utiyama; the Publicity Chairs, Srinivas Bangalore, Dekai Wu and Antonio Branco; and the Web Master Akifumi Yoshimoto. Our special thanks go to Swapna Somasundaran for her voluntary help to recruit additional reviewers to handle the much larger than expected submissions to the conference. Last but not the least, we are grateful to the softconf manager, Rich Gerber, for his continuous help with our various questions and needs.

We hope that you enjoy the conference!

Yuji Matsumoto, Nara Institute of Science and Technology, Japan
Rashmi Prasad, University of Wisconsin-Milwaukee, U.S.A.
Welcome to the COLING 2016!

It is a pleasure to welcome you to COLING 2016 organized by the Japanese Association of Natural Language Processing (ANLP) in Osaka. It has been 22 years since Japan last held the conference. While we are meeting here to discuss NLP, there is no substitute for personal contact. Therefore, we have arranged breaks, a reception, an excursion and a delightful banquet to facilitate discussion, collaboration and making connections. We hope and the modern conference venue together with the ambience of western Japan including Osaka, Nara and Kyoto (famous for their nature, culture, history, and food), help to make this an enjoyable experience for all. We hope the conference will result in accelerated growth of NLP.

Organizing a conference always takes a lot of work, and fortunately, we have experienced people from all around the world in attendance at the COLING 2016 site. It is both an honor and a great pleasure to work with them, and we thank them gratefully.

Since the proposal to host COLING was accepted by ICCL in 2014, our world has experienced some drastic changes. Under unfavorable economic conditions in Japan and considering the distance from Europe and America, we had to make a very conservative financial plan for the conference. The sponsorship chairs worked very hard and collected 33 sponsors, which is considerably more than in previous COLINGs.

This year’s conference has attracted a huge number of submissions and has a high level of participation, reflecting the ongoing dynamism in artificial intelligence around the globe. We were both overwhelmed by the numbers of visa applications we had to handle, and at the same time delighted and excited by the tremendous response.

We’d like to end by reporting two special features of COLING 2016: (1) COLING will assist student participants with registration subsidies. Successful applicants for the Student Support Program will receive all-inclusive free registration; (2) the collocation of the first international symposium for young researchers working on Natural Language Processing (YRSNLP) as an official satellite event at COLING 2016.

Welcome, and enjoy the conference!

Eiichiro SUMITA, Takenobu TOKUNAGA, and Sadao KUROHASHI

COLING 2016 Local Chairs
Organisers

**General Chair**
Nicoletta Calzolari

**Programme Chairs**
Yuji Matsumoto
Rashmi Prasad

**Local Chairs**
Eiichiro Sumita
Takenobu Tokunaga
Sadao Kurohashi

**Workshops Chairs**
Key-Sun Choi
Monica Monachini
Yusuke Miyao

**Demo Chair**
Hideo Watanabe

**Tutorial Chairs**
Marcello Federico
Akiko Aizawa

**Publication Chairs**
Hitoshi Isahara
Masao Utiyama

**Sponsorship Chairs**
Satoshi Sekine
Su Jian
Patrick Pantel
Chengqing Zong
Ralf Steinberger

**Publicity Chairs**
Srinivas Bangalore
Dekai Wu
Antonio Branco

**Advisor to the Local Committee**
John Judge

**Web Master**
Akifumi Yoshimoto

**Student Volunteer Coordinator**
Yugo Murawaki

**Program Booklet Coordinator**
Akihiro Tamura
Programme Committee

Programme Chairs
Yuji Matsumoto
Rashmi Prasad

Linguistic Issues in NLP
Tracy Holloway King
Annie Zaenen

Machine Learning for NLP
Trevor Cohn
Sujith Ravi

Computational Psycholinguistics
Vera Demberg
Shravan Vasishth

Morphology, Segmentation, Tagging, Chunking
Daichi Mochihashi
Hinrich Schutze

Syntactic and Semantic Parsing, Grammar Induction
Daisuke Kawahara
Joakim Nivre

Lexical Semantics, Ontologies
Diana Inkpen
Roberto Navigli

Semantic Processing, Distributional Semantics, Compositionality
Chris Biemann
Roser Morante
Stephen Wu

Discourse Relations, Coreference, Pragmatics
Vincent Ng
Bonnie Webber

Natural Language Generation, Summarization
Donia Scott
Hiroya Takamura
Michael White

Paraphrasing, Textual Entailment
Roy Bar-Haim
Kentaro Inui

Sentiment Analysis, Computational Argumentation
Iryna Gurevych
Swapna Somasundaran
Information Retrieval, Information Extraction, Question Answering
Mausam
Marie-Francine Moens

Applications
Tim Baldwin
Maria Liakata

Dialog Processing and Dialog Systems, Multimodal Interfaces
Nina Dethlefs
Simon Keizer
Giuseppe Riccardi

Speech Recognition, Text-To-Speech, Spoken Language Understanding
Florian Metze
Chung-Hsien Wu

Machine Translation
Graham Neubig
Taro Watanabe
Min Zhang

Resources, Software and Tools
Christian Chiarcos
Nancy Ide
James Pustejovsky

Under-resourced Languages
Alexis Palmer
Richard Sproat

Reviewers
A Ramanathan, Aasish Pappu, Abhijit Mishra, Adam Przepiórkowski, Adam Grycner, Adam Meyers, Adam Tsakalidis, Adeline Nazarenko, Adrià de Gispert, Aida Nematzadeh,

AiTi Aw, Akihiro Tamura, Alan Akbik, Alan Ritter, Albert Gatt, Aldo Gangemi, Ales Horak, Alessandro Lenci, Alessandro Mazzei, Alessandro Moschitti,

Alessio Palermo Aprosio, Alex Rudnick, Alexander Clark, Alexander Panchenko, Alexandros Potamianos, Alexis Nasr, Alexis Palmer, Aleš Tamchyna, Alisa Zhila, Amal Zouaq,

Amanda Stent, Amba Kulkarni, Amin Mantrach, Amit Goyal, Amita Misra, Amjad Abu-Jbara, Amy Isard, Amália Mendes, Anastasia Krithara, Anders Björkelund,

Anders Søgaard, Andrea Horbach, Andreas Peldszus, Andreas Scherbakov, Andreas Søeborg Kirkedal, Andreas Vlachos, Andrei Popescu-Belis, Andrew Finch, Anette Frank, Animesh Mukherjee,

Anita Ramm, Anja Belz, Ann Bies, Ann Clifton, Anna Rumshisky, Anne Cocos, Anne-Lyse Minard, Annemarie Friedrich, Anoop Kunchukuttan, Antonio Pareja Lora,
Ryu Iida, Ryuichiro Higashinaka,

Sa-kwang Song, Sabine Bergler, Sabino Miranda-Jiménez, Sadao Kurohashi, Saif Mohammad, Salah Ait-Mokhtar, Sam Wiseman, Sameer Singh, Sampo Pyysalo, Sanda Harabagiu,

Sanghoun Song, Sanjeev Kumar Karn, Sara Rosenthal, Sara Tonelli, Sascha Griffiths, Sascha Rothe, Satoshi Sekine, Saurav Sahay, Sean Chester, Sebastian Ebert,

Sebastian Hellmann, Sebastian Krause, Sebastian Martschat, Sebastian Nordhoff, Sebastian Padó, Sebastian Riedel, Seong-Bae Park, Shachar Mirkin, Sharid Loáiciga, Shashi Narayan,

Shibamouli Lahiri, Shih-Hung Wu, Shinsuke Mori, Shravan Vasishth, Shuly Wintner, Siddharth Patwardhan, Simon Keizer, Simon Mille, Simone Paolo Ponzetto, Simonetta Montemagni,

Sina Zarrieß, Sinno Jialin Pan, Smaranda Muresan, Sophia Ananiadou, Stefan Bott, Stefan L. Frank, Stefan Ultes, Stefano Faralli, Steffen Eger, Stephan Busemann,

Stephan Oepen, Stephen Wan, Stergos Afantenos, Steve DeNeefe, Steven Bethard, Steven Moran, Steven Wilson, Su-Youn Yoon, Sujay Kumar Jauhar, Sumithra Velupillai,

Sunghwan Mac Kim, Suresh Manandhar, Susanne Burger, Suzan Verberne, Svetla Koeva, Svetlana Stoyanchev, Svitlana Vakulenko, Svitlana Volkova,

Tadashi Nomoto, Takenobu Tokunaga, Takuya Matsuzaki, Tal Linzen, Tamás Váradi, Tanel Alumäe, Tara McIntosh, Tatsuya Kawahara, Tejaswini Deoskar, Tengfei Ma,

Tetsuji Nakagawa, Thiago Castro Ferreira, Thien Huu Nguyen, Thierry Hamon, Thomas Meyer, Tim Rocktäschel, Tim Van de Cruys, Timo Baumann, Ting Han, Ting Liu,

Titus von der Malsburg, Tommaso Caselli, tommaso pasini, Tomohide Shibata, Tomoya Iwakura, Tomáš Brychčeň, Tomáš Hercig, Tong Wang, Torsten Zesch, Tracy Holloway King, Travis Goodwin, Tristan Miller, Tsutomu Hirao, Tsvetana Dimitrova,

Uladzimir Sidarenka, Ulrike Pado,

Valentina Presutti, Valia Kordoni, Vanni Zavarella, Vasile Rus, Vasileios Lampos, Vassilina Nikoulina, Verónica Pérez-Rosas, Veselin Stoyanov, Victoria Rosén,

Victoria Yaneva, Viet-An Nguyen, Vincent Ng, Virach Sornlertlamvanich, Vivek Kulkarni, Vivek Kumar Rangarajan Sridhar, Vivek Srikumar, Vladimir Eidelman, Véronique MORICEAU,


Xavier Tannier, Xiangyu Duan, Xiao Ling, Xiaochang Peng, Xiaojun Wan, Xinchi Chen, Xingyi Song, Xu Sun, Xuancong Wang, Xuanjing Huang,
Yadollah Yaghoobzadeh, Yael Netzer, Yajie Miao, Yang Li, Yang Liu, Yannick Versley, Yannis Korkontzelos, Yanshan Wang, Yi-Ting Huang, Yingtao Tian,

Yoan Gutiérrez, Yoav Goldberg, Yogarshi Vyas, Yonatan Bisk, Yoong Keok Lee, Yuan Cao, Yufang Hou, Yuichihiro Matsubayashi, Yulan He, Yulia Tsvetkov, Yun-Nung Chen, Yuta Tsuboi, Yvette Graham,

Zdenka Uresova, Zhenghua Li, Zhiguo Wang, Zhiyuan Chen, Zhongqiang Huang, Zhongyu Wei, Zico Pratama Putra, Ákos Kádár, Özlem Çetinoğlu, Željko Agić

Secondary Reviewers

Abdalghani Abujabal, Rodrigo Agerri, Milan Aggarwal, Domagoj Alagic, Alessio Aprosio Palmero, Nikolay Arefiev, Kartik Asooja, Hanna Bechara, Irshad Bhat, Filip Boltuzic,

Alexey Borisov, Julia Bosque-Gil, Lars Bungum, Alberto Castro-Hernandez, Andrew Cattle, Yong Cheng, Maximu Coavoux, Łukasz Dębowski, Milan Dojinovski, Shichao Dong,

Erick Fonseca, Stella Frank, Miguel Ángel García Cumberras, Matt Gardner, Simon-Pierre Genot, Maite Giménez, Filip Ilievski, Sharmistha Jat, Salud María Jiménez-Zafra, Charles Jochim,

Jenna Kanerva, Mladen Karan, Yasen Kiprov, Bettina Klimek, Julien Kloetzer, Arne Köhn, Alexandros Komninos, Canasai Kruengkrai, Wu Kui, Gorka Labaka,

Po-Ting Lai, Egoitz Laparra, Joseph Lark, Jing Li, Minglei Li, Zhuo Li, Ming Liao, Ching-sheng Lin, Yunfei Long, Jing Ma,

Shuming Ma, Simone Magnolini, Márton Makrai, Sherwin Malmasi, Fernando Martinez Santiago, John McCreac, Zoran Medic, Todor Mihaylov, Tsvetomila Mihaylova, Benjamin Milde,

Pruthwik Mishra, Véronique Moriceau, Alessandro Moschitti, Diego Moussallem, Kay Müller, Jason Narad, Jumana Nassour, Ciro Baron Neto, Minh Le Ngoc, Jong-Hoon Oh,

Tomoko Ohta, Yuyu J. Ong, Sebastián Peña Saldarriaga, Haoruo Peng, Reshmi Gopalakrishna Pillai, Denys Proux, Gustavo Publio, Xinyin Qiu, Roi Reichart, Alexey Romanov,

Salvatore Romeo, Sophie Rosset, Dominic Rout, Andreas Rücklé, Nick Ruiz, Hassan Sajjad, Himanshu Sharma, Shiqi Shen, George Spithourakis, Pei-Hao Su,

Nina Tahmasebi, Partha Talukdar, Liling Tan, Serra Sinem Tekiroglu, Sara Tonelli, Adam Tsakalidis, Kateryna Tymoshenko, Dmitry Ustalov, Antonio Uva, Emiel van Miltenburg,

Chantal van Son, David Vilar Torres, Tim vor der Brück, Bo Wang, Pidong Wang, Shuohang Wang, Yu-Chun Wang, Johannes Welbl, Tsung-Hsien Wen, Michael White,

Chun-Kai Wu, Shuangzhi Wu, Jiacheng Xu, Wang Xuancong, Jian Zhang, Meishan Zhang, Meng Zhang, Nina Zhou, Patrick Ziering
Universal Dependencies is a framework for cross-linguistically consistent treebank annotation that has so far been applied to over 50 languages. It was developed primarily to support multilingual parsing research, but the resources have proven useful for a wide range of studies that were not foreseen originally, including research on language typology. A basic design principle in Universal Dependencies is to give priority to grammatical relations between content words, which are more likely to be parallel across languages, and to treat function words essentially as features of content words. This principle has been criticized both for being incompatible with theoretical linguistics, which tend to treat function words as syntactic heads, and for being suboptimal as a representation for dependency parsing, where higher accuracy is often observed with function words as heads. I will argue that both of these criticisms rest on a misinterpretation of the syntactic representations, and I will show that an alternative interpretation is compatible with both sound linguistics and improved parsing technology.
As models for language learning become increasingly sophisticated, it is essential to pay close attention to the purported input received by learners. This talk presents two examples of phonological input for which a failure to account for relevant factors has led to misleading conclusions. A fully annotated dataset of infant-directed speech is now allowing a more refined analysis of what children actually hear. The first example concerns vowel-duration contrasts (long vs. short) in Japanese. One previous study, working under the assumption that long and short vowels occurred with equal frequency, concluded that the distinction could be learned by a simple distributional model. Our dataset, however, reveals that (a) in reality over 90% of vowels in Japanese are short, and (b) the distribution of long vowel duration is entirely encompassed within that of short vowels. The second example concerns the widely accepted claim that when adults speak to infants (infant-directed speech, IDS), they speak with a slower speech rate than when speaking to adults (adult-directed speech, ADS). Studies supporting this conclusion, however, have consistently failed to account for the fact that IDS utterances are shorter than those of ADS. Our dataset differentiates between utterance-internal speech rate and utterance-final lengthening, and finds taken separately, these values almost identical between IDS and ADS. As it turns out, IDS appeared to have a slower overall rate only because of the greater frequency of utterance-final segments.
Invited talk 3

NLP to support clinical tasks and decisions

Dina Demner-Fushman (U.S. National Library of Medicine)

Clinical decision support (CDS) provides clinicians and patients with information needed to enhance health and health care. Clinical NLP – natural language processing methods to support healthcare by operationalizing clinical information contained in clinical narrative – is an integral part of CDS. Clinical NLP has started in the early 1960s, with several successful applications now integrated in daily care. I will first discuss the successful applications that are already positively impacting clinical practice, as well as publicly available resources, including those developed by our group. Consumer language understanding is an equally important and rapidly evolving part of CDS. In the second part of the talk, I will present our work in understanding consumer health questions. I will conclude with the challenges and opportunities to contribute to these fascinating research areas that have practical implications for our health.
Invited talk 4
A Look at Computational Argumentation and Summarisation from a
Text-Understanding Perspective
Simone Teufel (University of Cambridge)

In the past five years, computational argumentation has emerged as a new, active research field. This field studies all aspects of analysing and generating human argumentation, including argument mining, supportive debating technologies, logical representation of arguments, models of reasoning, and the connection of discourse processing and argumentation. As somebody who is mainly interested in the text-understanding challenges of computational argumentation, I think this new field has the potential to advance (and provide means of evaluating) the text-understanding capabilities of today’s NLP systems.

When humans construct an argument in order to convince others, how do they order and structure the information they want to convey? I will argue that whatever principles are at work, they are almost identical to those needed when summarising a text. Amongst the relations of particular interest are entailment, causal and rhetorical relationships. I will give an overview of currently available (text understanding-based) analysis methods that can inform our understanding of these principles, and I will also say a few words about a proposition-based approach to summarisation we have developed at Cambridge University that has the potential to contribute insights to computational argumentation.
Table of Contents

Boosting for Efficient Model Selection for Syntactic Parsing  
Rachel Bawden and Benoît Crabbé .......................................................... 1

A Universal Framework for Inductive Transfer Parsing across Multi-typed Treebanks  
Jiang Guo, Wanxiang Che, Haifeng Wang and Ting Liu ................................ 12

Grammar induction from (lots of) words alone  
John K Pate and Mark Johnson ............................................................. 23

A Redundancy-Aware Sentence Regression Framework for Extractive Summarization  
Pengjie Ren, Furu Wei, Zhumin CHEN, Jun MA and Ming Zhou ....................... 33

Generating Video Description using Sequence-to-sequence Model with Temporal Attention  
Natsuda Laokulrat, Sang Phan, Noriki Nishida, Raphael Shu, Yo Ehara, Naoaki Okazaki, Yusuke Miyao and Hideki Nakayama ............................................................... 44

An Improved Phrase-based Approach to Annotating and Summarizing Student Course Responses  
Wencan Luo, Fei Liu and Diane Litman ................................................... 53

CATENA: CAusal and TEmporal relation extraction from NAatural language texts  
Paramita Mirza and Sara Tonelli ........................................................... 64

Forecasting Word Model: Twitter-based Influenza Surveillance and Prediction  
Hayate ISO, Shoko WAKAMIYA and Eiji ARAMAKI ...................................... 76

Task-Oriented Intrinsic Evaluation of Semantic Textual Similarity  
Nils Reimers, Philip Beyer and Iryna Gurevych ........................................ 87

Expanding wordnets to new languages with multilingual sense disambiguation  
Mihael Arcan, John Philip McCrae and Paul Buitelaar .................................. 97

A Correlational Encoder Decoder Architecture for Pivot Based Sequence Generation  
Amrita Saha, Mitesh M. Khapra, Sarath Chandar, Janarthanan Rajendran and Kyunghyun Cho ................................................................. 109

Zero-resource Dependency Parsing: Boosting Delexicalized Cross-lingual Transfer with Linguistic Knowledge  
Lauriane Aufrant, Guillaume Wisniewski and François Yvon .......................... 119

Improving historical spelling normalization with bi-directional LSTMs and multi-task learning  
Marcel Bollmann and Anders Søgaard ..................................................... 131

Deceptive Opinion Spam Detection Using Neural Network  
Yafeng Ren and Yue Zhang ....................................................................... 140

Integrating Topic Modeling with Word Embeddings by Mixtures of vMFs  
Ximing Li, Jinjin Chi, Changchun Li, Jihong Ouyang and Bo Fu ..................... 151

Bayesian Language Model based on Mixture of Segmental Contexts for Spontaneous Utterances with Unexpected Words  
Ryu Takeda and Kazunori Komatani ........................................................ 161
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic Syllabification for Manipuri language</td>
<td>Loitongbam Gyanendra Singh, Lenin Laitonjam and Sanasam Ranbir Singh</td>
<td>349</td>
</tr>
<tr>
<td>Learning to Distill: The Essence Vector Modeling Framework</td>
<td>Kuan-Yu Chen, Shih-Hung Liu, Berlin Chen and Hsin-Min Wang</td>
<td>358</td>
</tr>
<tr>
<td>Continuous Expressive Speaking Styles Synthesis based on CVSM and MR-HMM</td>
<td>Jaime Lorenzo-Trueba, Roberto Barra-Chicote, Ascension Gallardo-Antolin, Junichi Yamagishi</td>
<td>369</td>
</tr>
<tr>
<td>An Automatic Prosody Tagger for Spontaneous Speech</td>
<td>Monica Domínguez, Mireia Farrús and Leo Wanner</td>
<td>377</td>
</tr>
<tr>
<td>Frustratingly Easy Neural Domain Adaptation</td>
<td>Young-Bum Kim, Karl Stratos and Ruhi Sarikaya</td>
<td>387</td>
</tr>
<tr>
<td>A House United: Bridging the Script and Lexical Barrier between Hindi and Urdu</td>
<td>Riyaz A. Bhat, Irshad A. Bhat, Naman Jain and Dipti Misra Sharma</td>
<td>397</td>
</tr>
<tr>
<td>Deeper syntax for better semantic parsing</td>
<td>Olivier Michalon, Corentin Ribeyre, Marie Candito and Alexis Nasr</td>
<td>409</td>
</tr>
<tr>
<td>Language Independent Dependency to Constituent Tree Conversion</td>
<td>Young-Suk Lee and Zhiguo Wang</td>
<td>421</td>
</tr>
<tr>
<td>Promoting multisecond expressions in A* TAG parsing</td>
<td>Jakub Waszczuk, Agata Savary and Yannick Parmentier</td>
<td>429</td>
</tr>
<tr>
<td>Incrementally Learning a Dependency Parser to Support Language Documentation in Field Linguistics</td>
<td>Morgan Ulinski, Julia Hirschberg and Owen Rambown</td>
<td>440</td>
</tr>
<tr>
<td>Inducing Multilingual Text Analysis Tools Using Bidirectional Recurrent Neural Networks</td>
<td>Othman ZENNAKI, Nasredine Semmar and Laurent Besacier</td>
<td>450</td>
</tr>
<tr>
<td>Bitext Name Tagging for Cross-lingual Entity Annotation Projection</td>
<td>Dongxu Zhang, Boliang Zhang, Xiaoman Pan, Xiaocheng Feng, Heng Ji and Weiran XU</td>
<td>461</td>
</tr>
<tr>
<td>Determining the Multiword Expression Inventory of a Surprise Language</td>
<td>Bahar Salehi, Paul Cook and Timothy Baldwin</td>
<td>471</td>
</tr>
<tr>
<td>A Hybrid Deep Learning Architecture for Sentiment Analysis</td>
<td>Md Shad Akhtar, Ayush Kumar, Asif Ekbal and Pushpak Bhattacharyya</td>
<td>482</td>
</tr>
<tr>
<td>Word Segmentation in Sanskrit Using Path Constrained Random Walks</td>
<td>Amrith Krishna, Bishal Santra, Pavankumar Satuluri, Sasi Prasanth Bandaru, Bhumí Faldu, Yajuvendra Singh and Pawan Goyal</td>
<td>494</td>
</tr>
<tr>
<td>Mongolian Named Entity Recognition System with Rich Features</td>
<td>Weihua Wang, Feilong Bao and Guanglai Gao</td>
<td>505</td>
</tr>
<tr>
<td>Appraising UMLS Coverage for Summarizing Medical Evidence</td>
<td>Elaheh ShafieiBavani, Mohammad Ebrahimi, Raymond Wong and Fang Chen</td>
<td>513</td>
</tr>
<tr>
<td>Large-scale Multi-class and Hierarchical Product Categorization for an E-commerce Giant</td>
<td>Ali Cevahir and Koji Murakami</td>
<td>525</td>
</tr>
</tbody>
</table>
Product Classification in E-Commerce using Distributional Semantics
Vivek Gupta, Harish Karnick, Ashendra Bansal and Pradhuman Jhala ............................... 536

AttSum: Joint Learning of Focusing and Summarization with Neural Attention
Ziqiang Cao, Wenjie Li, Sujian Li, Furu Wei and Yanran Li .................................................. 547

Using Relevant Public Posts to Enhance News Article Summarization
Chen Li, Zhongyu Wei, Yang Liu, Yang Jin and Fei Huang ..................................................... 557

A Proposition-Based Abstractive Summariser
Yimai Fang, Haoyue Zhu, Ewa Muszyńska, Alexander Kuhnle and Simone Teufel ............ 567

Cross-lingual Learning of an Open-domain Semantic Parser
Kilian Evang and Johan Bos .................................................................................................. 579

A subtree-based factorization of dependency parsing
Qiuye Zhao and Qun Liu ...................................................................................................... 589

K-SRL: Instance-based Learning for Semantic Role Labeling
Alan Akbik and Yunyao Li .................................................................................................. 599

Keystroke dynamics as signal for shallow syntactic parsing
Barbara Plank ....................................................................................................................... 609

A Bayesian model for joint word alignment and part-of-speech transfer
Robert Östling ..................................................................................................................... 620

Splitting compounds with ngrams
Naomi Tachikawa Shapiro .................................................................................................. 630

GAKE: Graph Aware Knowledge Embedding
Jun Feng, Minlie Huang, Yang Yang and Xiaoyan Zhu ...................................................... 641

Ranking Responses Oriented to Conversational Relevance in Chat-bots
Bowen Wu, Baoxun Wang and Hui Xue ............................................................................. 652

Probabilistic Prototype Model for Serendipitous Property Mining
Taesung Lee, Seung-won Hwang and Zhongyuan Wang .................................................. 663

Identifying Cross-Cultural Differences in Word Usage
Aparna Garimella, Rada Mihalcea and James Pennebaker .................................................. 674

Reading-Time Annotations for "Balanced Corpus of Contemporary Written Japanese"
Masayuki Asahara, Hajime Ono and Edson T. Miyamoto ................................................ 684

"How Bullying is this Message?: A Psychometric Thermometer for Bullying"
Parma Nand, Rivindu Perera and Abhijeet Kasture ............................................................ 695

Learning grammatical categories using paradigmatic representations: Substitute words for language acquisition
Mehmet Ali Yatbaz, Volkan Cirik, Aylin Küntay and Deniz Yuret ....................................... 707

Understanding the Lexical Simplification Needs of Non-Native Speakers of English
Gustavo Paetzold and Lucia Specia ..................................................................................... 717
How Interlocutors Coordinate with each other within Emotional Segments?
  Firoj Alam, Shammar Absar Chowdhury, Morena Danieli and Giuseppe Riccardi .............. 728

Advancing Linguistic Features and Insights by Label-informed Feature Grouping: An Exploration in the Context of Native Language Identification
  Serhiy Bykh and Detmar Meurers ......................................................... 739

Modeling Diachronic Change in Scientific Writing with Information Density
  Raphael Rubino, Stefania Degaetano-Ortlieb, Elke Teich and Josef van Genabith ............ 750

Different Contexts Lead to Different Word Embeddings
  Wenpeng Hu, Jiajun Zhang and Nan Zheng ............................................. 762

Machine Learning for Metrical Analysis of English Poetry
  Manex Agirrezabal, Íñaki Alegria and Mans Hulden .................................... 772

Automated speech-unit delimitation in spoken learner English
  Russell Moore, Andrew Caines, Calbert Graham and Paula Buttery ......................... 782

Learning to Identify Sentence Parallelism in Student Essays
  Wei Song, Tong Liu, Ruiji Fu, Lizhen Liu, Hanshi Wang and Ting Liu ....................... 794

Evaluating anaphora and coreference resolution to improve automatic keyphrase extraction
  Marco Basaldella, Giorgia Chiaradia and Carlo Tasso .................................. 804

Retrieving Occurrences of Grammatical Constructions
  Anna Ehrlemark, Richard Johansson and Benjamin Lyngfelt ................................ 815

Automatic Extraction of Learner Errors in ESL Sentences Using Linguistically Enhanced Alignments
  Mariano Felice, Christopher Bryant and Ted Briscoe .................................... 825

Contrasting Vertical and Horizontal Transmission of Typological Features
  Kenji Yamauchi and Yugo Murawaki ..................................................... 836

How Regular is Japanese Loanword Adaptation? A Computational Study
  Lingshuang Mao and Mans Hulden ....................................................... 847

Using Linguistic Data for English and Spanish Verb-Noun Combination Identification
  Uxoa Íñurrieta, Arantzaz Diaz de Iñarraza, Gorka Labaka, Kepa Sarasola, Itziar Aduriz and John Carroll ................................................................. 857

Analyzing Gender Bias in Student Evaluations
  Andamlak Terkik, Emily Prud’hommeaux, Cecilia Ovesdotter Alm, Christopher Homan and Scott Franklin ................................................................. 868

Adverse Drug Reaction Classification With Deep Neural Networks
  Trung Huynh, Yulan He, Alistair Willis and Stefan Rueger ............................... 877

Chinese Preposition Selection for Grammatical Error Diagnosis
  Hen-Hsen Huang, Yen-Chi Shao and Hsin-Hsi Chen ..................................... 888

Extending the Use of Adaptor Grammars for Unsupervised Morphological Segmentation of Unseen Languages
  Ramy Eskander, Owen Rambow and Tianchun Yang ....................................... 900
CharNER: Character-Level Named Entity Recognition
Onur Kuru, Ozan Arkan Can and Deniz Yuret .......................................................... 911

A Neural Model for Part-of-Speech Tagging in Historical Texts
Christian Hardmeier .............................................................................................. 922

Extracting Discriminative Keyphrases with Learned Semantic Hierarchies
Yunli Wang, Yong Jin, Xiaodan Zhu and Cyril Goutte .......................................... 932

Hashtag Recommendation Using End-To-End Memory Networks with Hierarchical Attention
Haoran Huang, Qi Zhang, Yeyun Gong and Xuanjing Huang .............................. 943

Automatic Labelling of Topics with Neural Embeddings
Shraey Bhatia, Jey Han Lau and Timothy Baldwin ............................................ 953

Memory-Bounded Left-Corner Unsupervised Grammar Induction on Child-Directed Input
Cory Shain, William Bryce, Lifeng Jin, Victoria Krakovna, Finale Doshi-Velez, Timothy Miller, William Schuler and Lane Schwartz ................................................................. 964

‘Calling on the classical phone’: a distributional model of adjective-noun errors in learners’ English
Aurélie Herbelot and Ekaterina Kochmar ........................................................... 976

Are Cohesive Features Relevant for Text Readability Evaluation?
Amalia Todirascu, Thomas Francois, Delphine Bernhard, Nuria Gala and Anne-Laure Ligozat 987

Named Entity Recognition for Linguistic Rapid Response in Low-Resource Languages: Sorani Kurdish and Tajik
Patrick Littell, Kartik Goyal, David R. Mortensen, Alexa Little, Chris Dyer and Lori Levin ...... 998

Multilingual Supervision of Semantic Annotation
Peter Exner, Marcus Klang and Pierre Nugues .................................................. 1007

Siamese Convolutional Networks for Cognate Identification
Taraka Rama ........................................................................................................ 1018

Exploring Differential Topic Models for Comparative Summarization of Scientific Papers
Lei He, Wei Li and Hai Zhuge ................................................................................ 1028

Bridging the gap between extractive and abstractive summaries: Creation and evaluation of coherent extracts from heterogeneous sources
Darina Benikova, Margot Mieskes, Christian M. Meyer and Iryna Gurevych .......... 1039

Chinese Poetry Generation with Planning based Neural Network
Zhe Wang, Wei He, Hua Wu, Haiyang Wu, Wei Li, Haifeng Wang and Enhong Chen ....... 1051

Predicting sentential semantic compatibility for aggregation in text-to-text generation
Victor Chenal and Jackie Chi Kit Cheung ............................................................... 1061

Sequential Clustering and Contextual Importance Measures for Incremental Update Summarization
Markus Zopf, Eneldo Loza Mencía and Johannes Fürnkranz .................................. 1071

Natural Language Generation through Character-based RNNs with Finite-state Prior Knowledge
Raghav Goyal, Marc Dymetman and Eric Gaussier ............................................. 1083
A Hybrid Approach to Generation of Missing Abstracts in Biomedical Literature  
Suchet Chachra, Asma Ben Abacha, Sonya Shooshan, Laritza Rodriguez and Dina Demner-Fushman ................................................................. 1093

Imitation learning for language generation from unaligned data  
Gerasimos Lampouras and Andreas Vlachos .................................................. 1101

Product Review Summarization by Exploiting Phrase Properties  
Naitong Yu, Minlie Huang, Yuanjuan Shi and Xiaoyan Zhu .............................. 1113

Generating Questions and Multiple-Choice Answers using Semantic Analysis of Texts  
Jun Araki, Dheeraj Rajagopal, Sreecharan Sankaranarayanan, Susan Holm, Yukari Yamakawa and Teruko Mitamura .................................................. 1125

Evaluation Strategies for Computational Construction Grammars  
Tania Marques and Katrien Beuls ................................................................. 1137

Building a Monolingual Parallel Corpus for Text Simplification Using Sentence Similarity Based on Alignment between Word Embeddings  
Tomoyuki Kajiwara and Mamoru Komachi ................................................... 1147

Word2Vec vs DBnary: Augmenting METEOR using Vector Representations or Lexical Resources?  
Christophe Servan, Alexandre Berard, zied elloumi, Hervé Blanchon and Laurent Besacier . 1159

Broad Twitter Corpus: A Diverse Named Entity Recognition Resource  
Leon Derczynski, Kalina Bontcheva and Ian Roberts ...................................... 1169

Semantic overfitting: what ‘world’ do we consider when evaluating disambiguation of text?  
Filip Ilievski, Marten Postma and Piek Vossen .............................................. 1180

Extraction of Keywords of Novelties From Patent Claims  
Shoko Suzuki and Hiromichi Takatsuka ......................................................... 1192

Leveraging Multilingual Training for Limited Resource Event Extraction  
Andrew Hsi, Yiming Yang, Jaime Carbonell and Ruochen Xu ......................... 1201

LILI: A Simple Language Independent Approach for Language Identification  
Mohamed Al-Badrashiny and Mona Diab ....................................................... 1211

High Accuracy Rule-based Question Classification using Question Syntax and Semantics  
Harish Tayyar Madabushi and Mark Lee ...................................................... 1220

Incorporating Label Dependency for Answer Quality Tagging in Community Question Answering via CNN-LSTM-CRF  
Yang Xiang, Xiaqiang Zhou, Qingcai Chen, Zhihui Zheng, Buzhou Tang, Xiaolong Wang and Yang Qin .............................................................. 1231

Semantically Motivated Hebrew Verb-Noun Multi-Word Expressions Identification  
Chaya Liebeskind and Yaakov HaCohen-Kerner ............................................ 1242

Semantic Relation Classification via Hierarchical Recurrent Neural Network with Attention  
Minguang Xiao and Cong Liu ........................................................................ 1254

A Unified Architecture for Semantic Role Labeling and Relation Classification  
Jiang Guo, Wanxiang Che, Haifeng Wang, Ting Liu and Jun Xu ....................... 1264

xxvii
Facing the most difficult case of Semantic Role Labeling: A collaboration of word embeddings and co-training
Quynh Ngoc Thi Do, Steven Bethard and Marie-Francine Moens ........................................... 1275

Predictability of Distributional Semantics in Derivational Word Formation
Sebastian Padó, Aurélie Herbelot, Max Kisselew and Jan Šnajder ........................................... 1285

Survey on the Use of Typological Information in Natural Language Processing
Helen O’Horan, Yevgeni Berzak, Ivan Vulic, Roi Reichart and Anna Korhonen ......................... 1297

From phonemes to images: levels of representation in a recurrent neural model of visually-grounded language learning
Lieke Gelderloos and Grzegorz Chrupała ................................................................. 1309

Linguistic features for Hindi light verb construction identification
Ashwini Vaidya, Sumeet Agarwal and Martha Palmer .................................................... 1320

Cross-lingual Transfer of Correlations between Parts of Speech and Gaze Features
Maria Barrett, Frank Keller and Anders Søgaard ...................................................... 1330

Sentence Similarity Learning by Lexical Decomposition and Composition
Zhiguo Wang, Haitao Mi and Abraham Ittycheriah ...................................................... 1340

Chinese Hypernym-Hyponym Extraction from User Generated Categories
Chengyu Wang and Xiaofeng He ................................................................. 1350

Dynamic Generative model for Diachronic Sense Emergence Detection
Martin Emms and Arun kumar Jayapal ............................................................. 1362

Semi-supervised Word Sense Disambiguation with Neural Models
Dayu Yuan, Julian Richardson, Ryan Doherty, Colin Evans and Eric Altendorf ................. 1374

Fast Gated Neural Domain Adaptation: Language Model as a Case Study
Jian Zhang, Xiaofeng Wu, Andy Way and Qun Liu ...................................................... 1386

Machine Translation Evaluation for Arabic using Morphologically-enriched Embeddings
Francisco Guzmán, Houda Bouamor, Ramy Baly and Nizar Habash ......................... 1398

Ensemble Learning for Multi-Source Neural Machine Translation
Ekaterina Garmash and Christof Monz ............................................................ 1409

Phrase-based Machine Translation using Multiple Preordering Candidates
Yusuke Oda, Taku Kudo, Tetsuji Nakagawa and Taro Watanabe ........................................ 1419

Hand in Glove: Deep Feature Fusion Network Architectures for Answer Quality Prediction in Community Question Answering
Sai Praneeth Suggu, Kushwanth Naga Goutham, Manoj K. Chinnakotla and Manish Shrivastava ................................................................. 1429

Learning Event Expressions via Bilingual Structure Projection
Fangyuan Li, Ruihong Huang, Deyi Xiong and Min Zhang ........................................... 1441

Global Inference to Chinese Temporal Relation Extraction
Peifeng Li, Qiaoming Zhu, Guodong Zhou and Hongling Wang .................................... 1451
Improved relation classification by deep recurrent neural networks with data augmentation
Yan Xu, Ran Jia, Lili Mou, Ge Li, Yunchuan Chen, Yangyang Lu and Zhi Jin .......................... 1461

Relation Extraction with Multi-instance Multi-label Convolutional Neural Networks
Xiaotian Jiang, Quan Wang, Peng Li and Bin Wang ......................................................... 1471

Named Entity Disambiguation for little known referents: a topic-based approach
Andrea Glaser and Jonas Kuhn ......................................................................................... 1481

Building RDF Content for Data-to-Text Generation
Laura Perez-Beltrachini, Rania Sayed and Claire Gardent .............................................. 1493

Parallel Sentence Compression
Julia Ive and François Yvon ............................................................................................. 1503

An Unsupervised Multi-Document Summarization Framework Based on Neural Document Model
Shulei Ma, Zhi-Hong Deng and Yunlun Yang ................................................................. 1514

From OpenCCG to AI Planning: Detecting Infeasible Edges in Sentence Generation
Maximilian Schwenger, Alvaro Torralba, Joerg Hoffmann, David M. Howcroft and Vera Demberg ..................................................................................................................... 1524

The Next Step for Multi-Document Summarization: A Heterogeneous Multi-Genre Corpus Built with a Novel Construction Approach
Markus Zopf, Maxime Peyrard and Judith Eckle-Kohler .................................................. 1535

SentiHood: Targeted Aspect Based Sentiment Analysis Dataset for Urban Neighbourhoods
Marzieh Saeidi, Guillaume Bouchard, Maria Liakata and Sebastian Riedel ...................... 1546

On the Impact of Seed Words on Sentiment Polarity Lexicon Induction
Dame Jovanoski, Veno Pachovski and Preslav Nakov ....................................................... 1557

Evaluating Argumentative and Narrative Essays using Graphs
Swapna Somasundaran, Brian Riordan, Binod Gyawali and Su-Youn Yoon ....................... 1568

Selective Co-occurrences for Word-Emotion Association
Ameeta Agrawal and Aijun An .......................................................................................... 1579

Weighted Neural Bag-of-n-grams Model: New Baselines for Text Classification
Bofang Li, Zhe Zhao, Tao Liu, Puwei Wang and Xiaoyong Du ............................................ 1591

A Deeper Look into Sarcastic Tweets Using Deep Convolutional Neural Networks
Soujanya Poria, Erik Cambria, Devamanyu Hazarika and Prateek Vij .............................. 1601

Exploring Distributional Representations and Machine Translation for Aspect-based Cross-lingual Sentiment Classification
Jeremy Barnes, Patrik Lambert and Toni Badia .................................................................... 1613

A Bilingual Attention Network for Code-switched Emotion Prediction
Zhongqing Wang, Yue Zhang, Sophia Lee, Shoushan Li and Guodong Zhou .................... 1624

UTCNN: a Deep Learning Model of Stance Classification on Social Media Text
Wei-Fan Chen and Lun-Wei Ku ....................................................................................... 1635

The Role of Intrinsic Motivation in Artificial Language Emergence: a Case Study on Colour
Miquel Cornudella, Thierry Poibeau and Remi van Trijp ................................................. 1646
Predicting the Evocation Relation between Lexicalized Concepts
Yoshihiko Hayashi .......................................................... 1657

Collecting and Exploring Everyday Language for Predicting Psycholinguistic Properties of Words
Gustavo Paetzold and Lucia Specia ..................................... 1669

Using Argument Mining to Assess the Argumentation Quality of Essays
Henning Wachsmuth, Khalid Al Khatib and Benno Stein .......... 1680

Grammatical Templates: Improving Text Difficulty Evaluation for Language Learners
Shuhan Wang and Erik Andersen ...................................... 1692

Still not there? Comparing Traditional Sequence-to-Sequence Models to Encoder-Decoder Neural Networks on Monotone String Translation Tasks
Carsten Schnober, Steffen Eger, Erik-Lân Do Dinh and Iryna Gurevych ...................................................... 1703

Towards Time-Aware Knowledge Graph Completion
Tingsong Jiang, Tianyu Liu, Tao Ge, Lei Sha, Baobao Chang, Sujian Li and Zhifang Sui .............. 1715

Learning to Weight Translations using Ordinal Linear Regression and Query-generated Training Data for Ad-hoc Retrieval with Long Queries
Javid Dadashkarimi, Masoud Jalili Sabet and Azadeh Shakery ................................................................. 1725

Neural Attention for Learning to Rank Questions in Community Question Answering
Salvatore Romeo, Giovanni Da San Martino, Alberto Barrón-Cedeño, Alessandro Moschitti, Yonatan Belinkov, Wei-Ning Hsu, Yu Zhang, Mitra Mohtarami and James Glass ......................... 1734

Simple Question Answering by Attentive Convolutional Neural Network
Wenpeng Yin, Mo Yu, Bing Xiang, Bowen Zhou and Hinrich Schütze .......................................................... 1746

Recurrent Dropout without Memory Loss
Stanislau Semeniuta, Aliaksei Severyn and Erhardt Barth ........ 1757

Modeling topic dependencies in semantically coherent text spans with copulas
Georgios Balikas, Hesam Amoualian, Marianne Clausel, Eric Gaussier and Massih R Amini . 1767

Consensus Attention-based Neural Networks for Chinese Reading Comprehension
Yiming Cui, Ting Liu, Zhipeng Chen, Shijin Wang and Guoping Hu .............................................................. 1777

Semantic Annotation Aggregation with Conditional Crowdsourcing Models and Word Embeddings
Paul Felt, Eric Ringger and Kevin Seppi ............................ 1787

Interactive-Predictive Machine Translation based on Syntactic Constraints of Prefix
Na Ye, Guiping Zhang and Dongfeng Cai ............................ 1797

Topic-Informed Neural Machine Translation
Jian Zhang, Liangyou Li, Andy Way and Qun Liu .................. 1807

A Distribution-based Model to Learn Bilingual Word Embeddings
Hailong Cao, Tiejun Zhao, Shu ZHANG and Yao Meng .......... 1818

Pre-Translation for Neural Machine Translation
Jan Niehues, Eunah Cho, Thanh-Le Ha and Alex Waibel ......... 1828

Direct vs. indirect evaluation of distributional thesauri
Vincent Claveau and Ewa Kijak ...................................... 1837

xxx
D-GloVe: A Feasible Least Squares Model for Estimating Word Embedding Densities
Shoaib Jameel and Steven Schockaert ................................................................. 1849

Predicting human similarity judgments with distributional models: The value of word associations.
Simon De Deyne, Amy Perfors and Daniel J Navarro ........................................... 1861

Distributional Hypernym Generation by Jointly Learning Clusters and Projections
Josuke Yamane, Tomoya Takatani, Hitoshi Yamada, Makoto Miwa and Yutaka Sasaki .... 1871

Incremental Fine-grained Information Status Classification Using Attention-based LSTMs
Yufang Hou ........................................................................................................... 1880

Detection, Disambiguation and Argument Identification of Discourse Connectives in Chinese Discourse Parsing
Yong-Siang Shih and Hsin-Hsi Chen ....................................................................... 1891

Multi-view and multi-task training of RST discourse parsers
Chloé Braud, Barbara Plank and Anders Søgaard .................................................. 1903

Implicit Discourse Relation Recognition with Context-aware Character-enhanced Embeddings
Lianhui Qin, Zhisong Zhang and Hai Zhao ............................................................ 1914

Measuring Non-cooperation in Dialogue
Brian Plüss and Paul Piwek .................................................................................. 1925

Representation and Learning of Temporal Relations
Leon Derczynski .................................................................................................... 1937

Revisiting the Evaluation for Cross Document Event Coreference
Shyam Upadhyay, Nitish Gupta, Christos Christodoulopoulos and Dan Roth ................ 1949

Modeling Discourse Segments in Lyrics Using Repeated Patterns
Kento Watanabe, Yuichiroh Matsubayashi, Naho Orita, Naoaki Okazaki, Kentaro Inui, Satoru Fukayama, Tomoyasu Nakano, Jordan Smith and Masataka Goto .............................. 1959

Multi-level Gated Recurrent Neural Network for dialog act classification
Wei Li and Yunfang Wu .......................................................................................... 1970

Multimodal Mood Classification - A Case Study of Differences in Hindi and Western Songs
Braja Gopal Patra, Dipankar Das and Sivaji Bandyopadhyay ...................................... 1980

Detecting Context Dependent Messages in a Conversational Environment
Chaozhuo Li, Yu Wu, Wei Wu, Chen Xing, Zhoujun Li and Ming Zhou ....................... 1990

Joint Inference for Mode Identification in Tutorial Dialogues
Deepak Venugopal and Vasile Rus ........................................................................ 2000

Dialogue Act Classification in Domain-Independent Conversations Using a Deep Recurrent Neural Network
Hamed Khanpour, Nishitha Guntakandla and Rodney Nielsen ................................. 2012

Non-sentential Question Resolution using Sequence to Sequence Learning
Vineet Kumar and Sachindra Joshi ......................................................................... 2022

Context-aware Natural Language Generation for Spoken Dialogue Systems
Hao Zhou, Minlie Huang and xiaoyan zhu ............................................................. 2032

xxxi
Weakly-supervised text-to-speech alignment confidence measure
Guillaume Serrière, Christophe Cerisara, Dominique Fohr and Odile Mella.................2042

Domainless Adaptation by Constrained Decoding on a Schema Lattice
Young-Bum Kim, Karl Stratos and Ruhi Sarikaya............................................2051

Sub-Word Similarity based Search for Embeddings: Inducing Rare-Word Embeddings for Word Similarity Tasks and Language Modelling
Mittul Singh, Clayton Greenberg, Youssef Oualil and Dietrich Klakow.......................2061

Semi-automatic Detection of Cross-lingual Marketing Blunders based on Pragmatic Label Propagation in Wiktionary
Christian M. Meyer, Judith Eckle-Kohler and Iryna Gurevych...............................2071

Ambient Search: A Document Retrieval System for Speech Streams
Benjamin Milde, Jonas Wacker, Stefan Radomski, Max Mühlhäuser and Chris Biemann......2082

Semi-supervised Gender Classification with Joint Textual and Social Modeling
Shoushan Li, Bin Dai, Zhengxian Gong and Guodong Zhou ....................................2092

Predicting proficiency levels in learner writings by transferring a linguistic complexity model from expert-written coursebooks
Ildikó Pilán, Elena Volodina and Torsten Zesch...................................................2101

User Classification with Multiple Textual Perspectives
Dong Zhang, Shoushan Li, Hongling Wang and Guodong Zhou ...............................2112

Says Who...? Identification of Expert versus Layman Critics’ Reviews of Documentary Films
Ming Jiang and Jana Diesner...............................................................................2122

Knowledge-Driven Event Embedding for Stock Prediction
Xiao Ding, Yue Zhang, Ting Liu and Junwen Duan..................................................2133

Distributed Representations for Building Profiles of Users and Items from Text Reviews
Wenliang Chen, Zhenjie Zhang, Zhenghua Li and Min Zhang.................................2143

Improving Statistical Machine Translation with Selectional Preferences
Haiqing Tang, Deyi Xiong, Min Zhang and Zhengxian Gong.....................................2154

Hierarchical Permutation Complexity for Word Order Evaluation
Miloš Stanojević and Khalil Sima’an.................................................................2164

Interactive Attention for Neural Machine Translation
Fandong Meng, Zhengdong Lu, Hang Li and Qun Liu.............................................2174

Get Semantic With Me! The Usefulness of Different Feature Types for Short-Answer Grading
Ulrike Pado........................................................................................................2186

Automatically Processing Tweets from Gang-Involved Youth: Towards Detecting Loss and Aggression
Terra Blevins, Robert Kwiatkowsk, Jamie MacBeth, Kathleen McKeown, Desmond Patton and Owen Rambow.........................................................2196

Content-based Influence Modeling for Opinion Behavior Prediction
Chengyao Chen, Zhitao Wang, Yu Lei and Wenjie Li.............................................2207
Data-driven learning of symbolic constraints for a log-linear model in a phonological setting
Gabriel Doyle and Roger Levy ................................................................. 2217

Chinese Tense Labelling and Causal Analysis
Hen-Hsen Huang, Chang-Rui Yang and Hsin-Hsi Chen ......................... 2227

Exploring Topic Discriminating Power of Words in Latent Dirichlet Allocation
Yang Kai, Cai Yi, Chen Zhenhong, Leung Ho-fung and LAU Raymond ........ 2238

Textual Entailment with Structured Attentions and Composition
Kai Zhao, Liang Huang and Mingbo Ma .................................................... 2248

plWordNet 3.0 – a Comprehensive Lexical-Semantic Resource
Marek Maziarz, Maciej Piasecki, Ewa Rudnicka, Stan Szpakowicz and Paweł Kędzia 2259

Time-Independent and Language-Independent Extraction of Multiword Expressions From Twitter
Nikhil Londhe, Rohini Srihari and Vishrawas Gopalakrishnan .................. 2269

Incremental Global Event Extraction
Alex Judea and Michael Strube ............................................................... 2279

Hierarchical Memory Networks for Answer Selection on Unknown Words
jiaming xu, Jing Shi, Yiqun Yao, Suncong Zheng, Bo Xu and Bo Xu .............. 2290

Revisiting Taxonomy Induction over Wikipedia
Amit Gupta, Francesco Piccinno, Mikhail Kozhevnikov, Marius Pasca and Daniele Pighin 2300

Joint Learning of Local and Global Features for Entity Linking via Neural Networks
Thien Huu Nguyen, Nicolas Faucoglia, Mariano Rodriguez Muro, Oktie Hassanzadeh, Alfio Massimiliano Gliozzo and Mohammad Sadoghi ........................................ 2310

Structured Aspect Extraction
Omer Gunes, Tim Furche and Giorgio Orsi ............................................. 2321

Robust Text Classification for Sparsely Labelled Data Using Multi-level Embeddings
Simon Baker, Douwe Kiela and Anna Korhonen .................................... 2333

Mathematical Information Retrieval based on Type Embeddings and Query Expansion
Yiannos Stathopoulos and Simone Teufel .............................................. 2344

Text Retrieval by Term Co-occurrences in a Query-based Vector Space
Eriks Sneiders ......................................................................................... 2356

Pairwise Relation Classification with Mirror Instances and a Combined Convolutional Neural Network
Jianfei Yu and Jing Jiang ......................................................................... 2366

FastHybrid: A Hybrid Model for Efficient Answer Selection
Lidan Wang, Ming Tan and Jiawei Han ...................................................... 2378

Extracting Spatial Entities and Relations in Korean Text
Bogyum Kim and Jae Sung Lee ............................................................... 2389

Hybrid Question Answering over Knowledge Base and Free Text
kun xu, Yansong Feng, Songfang Huang and Dongyan Zhao ...................... 2397
Improved Word Embeddings with Implicit Structure Information
Jie Shen and Cong Liu .......................................................... 2408

Word Embeddings and Convolutional Neural Network for Arabic Sentiment Classification
Abdelghani Dahou, Shengwu Xiong, Junwei Zhou, Mohamed Houcine Haddoud and Pengfei Duan .......................................................... 2418

Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts
Xingyou Wang, Weijie Jiang and Zhiyong Luo ........................................ 2428

Stance Classification in Rumours as a Sequential Task Exploiting the Tree Structure of Social Media Conversations
Arkaitz Zubiaga, Elena Kochkina, Maria Liakata, Rob Procter and Michal Lukasik ........................................ 2438

Tweet Sarcasm Detection Using Deep Neural Network
Meishan Zhang, Yue Zhang and Guohong Fu ........................................ 2449

Agreement and Disagreement: Comparison of Points of View in the Political Domain
Stefano Menini and Sara Tonelli .......................................................... 2461

Targeted Sentiment to Understand Student Comments
Charles Welch and Rada Mihalcea ...................................................... 2471

Towards Sub-Word Level Compositions for Sentiment Analysis of Hindi-English Code Mixed Text
Aditya Joshi, Ameya Prabhu, Manish Shrivastava and Vasudeva Varma ........................................ 2482

Distance Metric Learning for Aspect Phrase Grouping
Shufeng Xiong, Yue Zhang, Donghong Ji and Yinxia Lou ........................................ 2492

Constraint-Based Question Answering with Knowledge Graph
Junwei Bao, Nan Duan, Zhao Yan, Ming Zhou and Tiejun Zhao ........................................ 2503

Selecting Sentences versus Selecting Tree Constituents for Automatic Question Ranking
Alberto Barrón-Cedeño, Giovanni Da San Martino, Salvatore Romeo and Alessandro Moschitti ........................................ 2515

Attention-Based Convolutional Neural Network for Semantic Relation Extraction
yatian shen and Xuanjing Huang .......................................................... 2526

Table Filling Multi-Task Recurrent Neural Network for Joint Entity and Relation Extraction
Pankaj Gupta, Hinrich Schütze and Bernt Andrassy ........................................ 2537

Bilingual Autoencoders with Global Descriptors for Modeling Parallel Sentences
Biao Zhang, Deyi Xiong, jinsong su, Hong Duan and Min Zhang ........................................ 2548

Multi-Engine and Multi-Alignment Based Automatic Post-Editing and its Impact on Translation Productivity
Santanu Pal, Sudip Kumar Naskar and Josef van Genabith ........................................ 2559

Measuring the Effect of Conversational Aspects on Machine Translation Quality
Marlies van der Wees, Arianna Bisazza and Christof Monz ........................................ 2571

Enriching Phrase Tables for Statistical Machine Translation Using Mixed Embeddings
Peyman Passban, Qun Liu and Andy Way ........................................ 2582
Anecdote Recognition and Recommendation
Wei Song, Ruiji Fu, Lizhen Liu, Hanshi Wang and Ting Liu .................................................. 2592

Training Data Enrichment for Infrequent Discourse Relations
Kailang Jiang, Giuseppe Carenini and Raymond Ng ................................................................. 2603

Inferring Discourse Relations from PDTB-style Discourse Labels for Argumentative Revision Classification
Fan Zhang, Diane Litman and Katherine Forbes-Riley .............................................................. 2615

Capturing Pragmatic Knowledge in Article Usage Prediction using LSTMs
Jad Kabbara, Yulan Feng and Jackie Chi Kit Cheung ................................................................. 2625

Aspect Based Sentiment Analysis using Sentiment Flow with Local and Non-local Neighbor Information
Shubham Pateria ......................................................................................................................... 2635

Two-View Label Propagation to Semi-supervised Reader Emotion Classification
Shoushan Li, Jian Xu, Dong Zhang and Guodong Zhou ............................................................ 2647

A Joint Sentiment-Target-Stance Model for Stance Classification in Tweets
Javid Ebrahimi, Dejing Dou and Daniel Lowd ............................................................................ 2656

SenticNet 4: A Semantic Resource for Sentiment Analysis Based on Conceptual Primitives
Erik Cambria, Soujanya Poria, Rajiv Bajpai and Bjoern Schuller ............................................. 2666

Joint Embedding of Hierarchical Categories and Entities for Concept Categorization and Dataless Classification
Yuezhang(Music) Li, Ronghuo Zheng, Tian Tian, Zhiting Hu, Rahul Iyer and Katia Sycara . 2678

Latent Topic Embedding
Di Jiang, Lei Shi, Rongzhong Lian and Hua Wu ........................................................................ 2689

Neural-based Noise Filtering from Word Embeddings
Kim Anh Nguyen, Sabine Schulte im Walde and Ngoc Thang Vu ........................................ 2699

Integrating Distributional and Lexical Information for Semantic Classification of Words using MRMF
Rosa Tsegaye Aga, Lucas Drumond, Christian Wartena and Lars Schmidt-Thieme ............ 2708

Semi Supervised Preposition-Sense Disambiguation using Multilingual Data
Hila Gonen and Yoav Goldberg ................................................................................................. 2718

Monday mornings are my fave :) #not Exploring the Automatic Recognition of Irony in English tweets
Cynthia Van Hee, Els Lefever and Veronique Hoste ................................................................. 2730

CNN- and LSTM-based Claim Classification in Online User Comments
Chinnappa Guggilla, Tristan Miller and Iryna Gurevych .......................................................... 2740

Experiments in Idiom Recognition
Jing Peng and Anna Feldman .................................................................................................... 2752

An Empirical Evaluation of various Deep Learning Architectures for Bi-Sequence Classification Tasks
Anirban Laha and Vikas Raykar ................................................................................................ 2762

Learning Succinct Models: Pipelined Compression with L1-Regularization, Hashing, Elias-Fano Indices, and Quantization
Hajime Senuma and Akiko Aizawa ............................................................................................ 2774
Bad Company—Neighborhoods in Neural Embedding Spaces Considered Harmful
Johannes Hellrich and Udo Hahn ................................................................. 2785
Implementing a Reverse Dictionary, based on word definitions, using a Node-Graph Architecture
Sushrut Thorat and Varad Choudhari ....................................................... 2797
Guillem Collell and Marie-Francine Moens .............................................. 2807
On the contribution of word embeddings to temporal relation classification
Paramita Mirza and Sara Tonelli ............................................................. 2818
Modeling Context-sensitive Selectional Preference with Distributed Representations
Naoya Inoue, Yuichiroh Matsubayashi, Masayuki Ono, Naoaki Okazaki and Kentaro Inui 2829
Exploring the value space of attributes: Unsupervised bidirectional clustering of adjectives in German
Wiebke Petersen and Oliver Hellwig ......................................................... 2839
Distributional Inclusion Hypothesis for Tensor-based Composition
Dimitri Kartsaklis and Mehrnoosh Sadrzadeh .................................. 2849
Parameter estimation of Japanese predicate argument structure analysis model using eye gaze information
Ryosuke Maki, Hitoshi Nishikawa and Takenobu Tokunaga ......................... 2861
Reading and Thinking: Re-read LSTM Unit for Textual Entailment Recognition
Lei Sha, Baobao Chang, Zhifang Sui and Sujian Li .................................. 2870
A Paraphrase and Semantic Similarity Detection System for User Generated Short-Text Content on Microblogs
Kuntal Dey, Ritvik Shrivastava and Saroj Kaushik .................................. 2880
Modeling Extractive Sentence Intersection via Subtree Entailment
Omer Levy, Ido Dagan, Gabriel Stanovsky, Judith Eckle-Kohler and Iryna Gurevych 2891
Context-Sensitive Inference Rule Discovery: A Graph-Based Method
Xianpei Han and Le Sun ................................................................. 2902
Modelling Sentence Pairs with Tree-structured Attentive Encoder
Yao Zhou, Cong Liu and Yan Pan .......................................................... 2912
Neural Paraphrase Generation with Stacked Residual LSTM Networks
aaditya prakash, Sadid A. Hasan, Kathy Lee, Vivek Datla, Ashequl Qadir, Joey Liu and Oladimeji Farri .......................................................... 2923
English-Chinese Knowledge Base Translation with Neural Network
Xiaocheng Feng, Duyu Tang, Bing Qin and Ting Liu ................................ 2935
Keyphrase Annotation with Graph Co-Ranking
Adrien Bougouin, Florian Boudin and Beatrice Daille ................................ 2945
What’s in an Explanation? Characterizing Knowledge and Inference Requirements for Elementary Science Exams
Peter Jansen, Niranjan Balasubramanian, Mihai Surdeanu and Peter Clark .... 2956
“All I know about politics is what I read in Twitter”: Weakly Supervised Models for Extracting Politicians’ Stances From Twitter
Kristen Johnson and Dan Goldwasser ................................................................. 2966

Leveraging Multiple Domains for Sentiment Classification
Fan Yang, Arjun Mukherjee and Yifan Zhang .................................................. 2978

Political News Sentiment Analysis for Under-resourced Languages
Patrik F. Bakken, Terje A. Bratlie, Cristina Marco and Jon Atle Gulla .................. 2989

Fast Inference for Interactive Models of Text
Jeffrey Lund, Paul Felt, Kevin Seppi and Eric Ringger ...................................... 2997

Combining Heterogeneous User Generated Data to Sense Well-being
Adam Tsakalidis, Maria Liakata, Theo Damoulas, Brigitte Jellinek, Weisi Guo and Alexandra Cristea .............................................................. 3007

Hashtag Recommendation with Topical Attention-Based LSTM
Yang Li, Ting Liu, Jing Jiang and Liang Zhang .................................................. 3019

Better call Saul: Flexible Programming for Learning and Inference in NLP
Parisa Kordjamshidi, Daniel Khashabi, Christos Christodoulopoulos, Bhargav Mangipudi, Sameer Singh and Dan Roth ........................................ 3030

Crowdsourcing Complex Language Resources: Playing to Annotate Dependency Syntax
Bruno Guillaume, Karén Fort and Nicolas Lefebvre ........................................... 3041

Borrow a Little from your Rich Cousin: Using Embeddings and Polarities of English Words for Multilingual Sentiment Classification
Prerana Singhal and Pushpak Bhattacharyya .................................................... 3053

A Character-Aware Encoder for Neural Machine Translation
Zhen Yang, Wei Chen, Feng Wang and Bo Xu ................................................... 3063

Convolution-Enhanced Bilingual Recursive Neural Network for Bilingual Semantic Modeling
jinsong su, Biao Zhang, Deyi Xiong, Ruochen Li and Jianmin Yin ......................... 3071

Improving Attention Modeling with Implicit Distortion and Fertility for Machine Translation
Shi Feng, Shujie Liu, Nan Yang, Mu Li, Ming Zhou and Kenny Q. Zhu ............... 3082

Neural Machine Translation with Supervised Attention
Lemao Liu, Masao Utiyama, Andrew Finch and Eiichiro Sumita ......................... 3093

Lightly Supervised Quality Estimation
Matthias Sperber, Graham Neubig, Jan Niehues, Sebastian Stüker and Alex Waibel .......................................................... 3103

Improving Translation Selection with Supersenses
Haiqing Tang, Deyi Xiong, Oier Lopez de Lacalle and Eneko Agirre .................... 3114

Is all that Glitters in Machine Translation Quality Estimation really Gold?
Yvette Graham, Timothy Baldwin, Meghan Dowling, Maria Eskevich, Teresa Lynn and Lamia Tounsi ............................................................. 3124

Connecting Phrase based Statistical Machine Translation Adaptation
Rui Wang, Hai Zhao, Bao-Liang Lu, Masao Utiyama and Eiichiro Sumita ............ 3135
Fast Collocation-Based Bayesian HMM Word Alignment
Philip Schulz and Wilker Aziz .................................................. 3146

Learning to translate from graded and negative relevance information
Laura Jehl and Stefan Riezler .................................................. 3156

Universal Reordering via Linguistic Typology
Joachim Daiber, Miloš Stanojević and Khalil Sima’an .................. 3167

A Deep Fusion Model for Domain Adaptation in Phrase-based MT
Nadir Durrani, Hassan Sajjad, Shafiq Joty and Ahmed Abdelali .... 3177

Inducing Bilingual Lexica From Non-Parallel Data With Earth Mover’s Distance Regularization
Meng Zhang, Yang Liu, Huanbo Luan, Yiqun Liu and Maosong Sun ............................................................................ 3188

What Makes Word-level Neural Machine Translation Hard: A Case Study on English-German Translation
Fabian Hirschmann, Jinseok Nam and Johannes Fürnkranz ............ 3199

Improving Word Alignment of Rare Words with Word Embeddings
Masoud Jalili Sabet, Heshaam Faili and Gholamreza Haffari ........ 3209

Measuring the Information Content of Financial News
Ching-Yun Chang, Yue Zhang, Zhiyang Teng, Zahn Bozanic and Bin Ke ............................................................................. 3216

Automatic Generation and Classification of Minimal Meaningful Propositions in Educational Systems
Andreea Godea, Florin Bulgarov and Rodney Nielsen ..................... 3226

First Story Detection using Entities and Relations
Nikolaos Panagiotou, Cem Akkaya, Kostas Tsioutsoulikis, Vana Kalogeraki and Dimitrios Gunopulos .................................................. 3237

Textual complexity as a predictor of difficulty of listening items in language proficiency tests
Anastassia Loukina, Su-Youn Yoon, Jennifer Sakano, Youhua Wei and Kathy Sheehan .............................................................. 3245

Renfen HU, Jiayong Chen and Kuang-hua Chen .............................. 3254

Joint Inference for Event Coreference Resolution
Jing Lu, Deepak Venugopal, Vibhav Gogate and Vincent Ng ........ 3264

Event Detection with Burst Information Networks
Tao Ge, Lei Cui, Baobao Chang, Zhifang Sui and Ming Zhou ........... 3276

Corpus Fusion for Emotion Classification
Suyang Zhu, Shoushan Li, Ying Chen and Guodong Zhou ............... 3287

Effective LSTMs for Target-Dependent Sentiment Classification
Duyu Tang, Bing Qin, Xiaocheng Feng and Ting Liu ......................... 3298

Towards assessing depth of argumentation
Manfred Stede ............................................................................ 3308

Video Event Detection by Exploiting Word Dependencies from Image Captions
Sang Phan, Yusuke Miyao, Duy-Dinh Le and Shin’ichi Satoh ............. 3318

xxxviii
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicting Restaurant Consumption Level through Social Media Footprints</td>
<td>Yang Xiao, Yuan Wang, Hangyu Mao and Zhen Xiao</td>
<td>3328</td>
</tr>
<tr>
<td>A Novel Fast Framework for Topic Labeling Based on Similarity-preserved Hashing</td>
<td>Xian-Ling Mao, Yi-Jing Hao, Qiang Zhou, Wen-Qing Yuan, Liner Yang and Heyan Huang</td>
<td>3339</td>
</tr>
<tr>
<td>Sequence to Backward and Forward Sequences: A Content-Introducing Approach to Generative Short-Text Conversation</td>
<td>Lili Mou, Yiping Song, Rui Yan, Ge Li, Lu Zhang and Zhi Jin</td>
<td>3349</td>
</tr>
<tr>
<td>Disfluent but effective? A quantitative study of disfluencies and conversational moves in team discourse</td>
<td>Felix Gervits, Kathleen Eberhard and Matthias Scheutz</td>
<td>3359</td>
</tr>
<tr>
<td>A Neural Network Approach for Knowledge-Driven Response Generation</td>
<td>Pavlos Vougiouklis, Jonathon Hare and Elena Simperl</td>
<td>3370</td>
</tr>
<tr>
<td>PersoNER: Persian Named-Entity Recognition</td>
<td>Hanieh Poostchi, Ehsan Zare Borzeshi, Mohammad Abdous and Massimo Piccardi</td>
<td>3381</td>
</tr>
<tr>
<td>Efficient Data Selection for Bilingual Terminology Extraction from Comparable Corpora</td>
<td>Amir Hazem and Emmanuel Morin</td>
<td>3401</td>
</tr>
<tr>
<td>TweetGeo - A Tool for Collecting, Processing and Analysing Geo-encoded Linguistic Data</td>
<td>Nikola Ljubešić, Tanja Samardzic and Curdin Derungs</td>
<td>3412</td>
</tr>
<tr>
<td>Extending WordNet with Fine-Grained Collocational Information via Supervised Distributional Learning</td>
<td>Luis Espinosa Anke, Jose Camacho-Collados, Sara Rodríguez-Fernández, Horacio Saggion and Leo Wanner</td>
<td>3422</td>
</tr>
<tr>
<td>A News Editorial Corpus for Mining Argumentation Strategies</td>
<td>Khalid Al Khatib, Henning Wachsmuth, Johannes Kiesel, Matthias Hagen and Benno Stein</td>
<td>3433</td>
</tr>
<tr>
<td>Universal Dependencies for Turkish</td>
<td>Umut Sulubacak, Memduh Gokirmak, Francis Tyers, Çağrı Çöltekin, Joakim Nivre and Gülsen Eryiğit</td>
<td>3444</td>
</tr>
<tr>
<td>Creating Resources for Dialectal Arabic from a Single Annotation: A Case Study on Egyptian and Levantine</td>
<td>Ramy Eskander, Nizar Habash, Owen Rambow and Arfath Pasha</td>
<td>3455</td>
</tr>
<tr>
<td>Multilingual Aliasing for Auto-Generating Proposition Banks</td>
<td>Alan Akbik, Xinyu Guan and Yunyao Li</td>
<td>3466</td>
</tr>
<tr>
<td>PanPhon: A Resource for Mapping IPA Segments to Articulatory Feature Vectors</td>
<td>David R. Mortensen, Patrick Littell, Akash Bharadwaj, Kartik Goyal, Chris Dyer and Lori Levin</td>
<td>3475</td>
</tr>
<tr>
<td>Text Classification Improved by Integrating Bidirectional LSTM with Two-dimensional Max Pooling</td>
<td>Peng Zhou, Zhenyu Qi, Suncong Zheng, Jiaming Xu, Hongyun Bao and Bo Xu</td>
<td>3485</td>
</tr>
</tbody>
</table>
More is not always better: balancing sense distributions for all-words Word Sense Disambiguation
Marten Postma, Ruben Izquierdo Bevia and Piek Vossen .................................................. 3496

Language classification from bilingual word embedding graphs
Steffen Eger, Armin Hoenen and Alexander Mehler .......................................................... 3507

Word Embeddings, Analogies, and Machine Learning: Beyond king - man + woman = queen
Aleksandr Drozd, Anna Gladkova and Satoshi Matsuoka .................................................... 3519

Semantic Tagging with Deep Residual Networks
Johannes Bjerva, Barbara Plank and Johan Bos ................................................................. 3531

A Supervised Approach for Enriching the Relational Structure of Frame Semantics in FrameNet
Shafqat Mumtaz Virk, Philippe Muller and Juliette Conrath .................................................. 3542

Reddit Temporal N-gram Corpus and its Applications on Paraphrase and Semantic Similarity in Social
Media using a Topic-based Latent Semantic Analysis
Anh Dang, Abidalrahman Moh’d, Aminul Islam, Rosane Minghim, Michael Smit and Evangelos
Milios ................................................................................................................................... 3553

Dictionaries as Networks: Identifying the graph structure of Ogden’s Basic English
Camilo Garrido and Claudio Gutierrez .................................................................................. 3565

Structured Generative Models of Continuous Features for Word Sense Induction
Alexandros Komninos and Suresh Manandhar ................................................................... 3577
Conference Program

Tuesday, December 13, 2016

09:00–09:30 Opening:

09:30–10:30 Invited talk 1: Joakim Nivre (Uppsala University)

11:00–12:30 Session 1-A: Syntactic and Semantic Parsing, Grammar Induction I

- Boosting for Efficient Model Selection for Syntactic Parsing
  Rachel Bawden and Benoît Crabbé

- A Universal Framework for Inductive Transfer Parsing across Multi-typed Tree-banks
  Jiang Guo, Wanxiang Che, Haifeng Wang and Ting Liu

- Grammar induction from (lots of) words alone
  John K Pate and Mark Johnson

11:00–12:30 Session 1-B: Natural Language Generation, Summarization I

- A Redundancy-Aware Sentence Regression Framework for Extractive Summarization
  Pengjie Ren, Furu Wei, Zhumin CHEN, Jun MA and Ming Zhou

- Generating Video Description using Sequence-to-sequence Model with Temporal Attention
  Natsuda Laokulrat, Sang Phan, Noriki Nishida, Raphael Shu, Yo Ehara, Naoaki Okazaki, Yusuke Miyao and Hideki Nakayama

- An Improved Phrase-based Approach to Annotating and Summarizing Student Course Responses
  Wencan Luo, Fei Liu and Diane Litman
Tuesday, December 13, 2016 (continued)

11:00–12:30  Session 1-C: Applications I

*CATENA: CAusal and TEmporal relation extraction from NAtural language texts*
Paramita Mirza and Sara Tonelli

*Forecasting Word Model: Twitter-based Influenza Surveillance and Prediction*
Hayate ISO, Shoko WAKAMIYA and Eiji ARAMAKI

*Task-Oriented Intrinsic Evaluation of Semantic Textual Similarity*
Nils Reimers, Philip Beyer and Iryna Gurevych

11:00–12:30  Session 1-D: Resources, Software, Tools and Under-resourced languages I

*Expanding wordnets to new languages with multilingual sense disambiguation*
Mihael Arcan, John Philip McCrae and Paul Buitelaar

*A Correlational Encoder Decoder Architecture for Pivot Based Sequence Generation*
Amrita Saha, Mitesh M. Khapra, Sarath Chandar, Janarthanan Rajendran and Kyunghyun Cho

*Zero-resource Dependency Parsing: Boosting Delexicalized Cross-lingual Transfer with Linguistic Knowledge*
Lauriane Aufrant, Guillaume Wisniewski and François Yvon

12:30–14:00  Lunch break
Tuesday, December 13, 2016 (continued)

14:00–16:00 Session 2-A: Machine Learning for NLP I

*Improving historical spelling normalization with bi-directional LSTMs and multi-task learning*
Marcel Bollmann and Anders Søgaard

*Deceptive Opinion Spam Detection Using Neural Network*
Yafeng Ren and Yue Zhang

*Integrating Topic Modeling with Word Embeddings by Mixtures of vMFs*
Ximing Li, Jinjin Chi, Changchun Li, Jihong Ouyang and Bo Fu

*Bayesian Language Model based on Mixture of Segmental Contexts for Spontaneous Utterances with Unexpected Words*
Ryu Takeda and Kazunori Komatani

14:00–16:00 Session 2-B: Morphology, Segmentation Tagging, Chunking I

*Label Embedding for Zero-shot Fine-grained Named Entity Typing*
Yukun Ma, Erik Cambria and SA GAO

*The Role of Context in Neural Morphological Disambiguation*
Qinlan Shen, Daniel Clothiaux, Emily Tagtow, Patrick Littell and Chris Dyer

*Asynchronous Parallel Learning for Neural Networks and Structured Models with Dense Features*
Xu Sun

*An Empirical Exploration of Skip Connections for Sequential Tagging*
Huijia Wu, Jiajun Zhang and Chengqing Zong
Tuesday, December 13, 2016 (continued)

14:00–16:00  Session 2-C: Natural Language Generation, Summarization II

*Exploring Text Links for Coherent Multi-Document Summarization*
Xun Wang, Masaaki Nishino, Tsutomu Hirao, Katsuhito Sudoh and Masaaki Nagata

*Syntactic realization with data-driven neural tree grammars*
Brian McMahan and Matthew Stone

*Abstractive News Summarization based on Event Semantic Link Network*
Wei Li, Lei He and Hai Zhuge

*A General Optimization Framework for Multi-Document Summarization Using Genetic Algorithms and Swarm Intelligence*
Maxime Peyrard and Judith Eckle-Kohler

14:00–15:30  Session 2-D: Speech Recognition, Text-To-Speech, Spoken Language Understanding

*Exploiting Sentence and Context Representations in Deep Neural Models for Spoken Language Understanding*
Lina M. Rojas Barahona, Milica Gasic, Nikola Mrkšić, Pei-Hao Su, Stefan Ultes, Tsung-Hsien Wen and Steve Young

*Predictive Incremental Parsing Helps Language Modeling*
Arne Köhn and Timo Baumann

*A Neural Attention Model for Disfluency Detection*
Shaolei Wang, Wanxiang Che and Ting Liu
14:00–16:00 Session 2-P: Poster Session 1

**Morphology, Segmentation, Tagging, Chunking**

*Detecting Sentence Boundaries in Sanskrit Texts*
Oliver Hellwig

*Consistent Word Segmentation, Part-of-Speech Tagging and Dependency Labelling Annotation for Chinese Language*
Mo Shen, Wingmui Li, HyunJeong Choe, Chenhui Chu, Daisuke Kawahara and Sadao Kurohashi

*Attending to Characters in Neural Sequence Labeling Models*
Marek Rei, Gamal Crichton and Sampo Pyysalo

*A Word Labeling Approach to Thai Sentence Boundary Detection and POS Tagging*
Nina Zhou, AiTi Aw, Nattadaporn Lertcheva and Xuancong Wang

*Assigning Fine-grained PoS Tags based on High-precision Coarse-grained Tagging*
Tobias Horsmann and Torsten Zesch

*Data-Driven Morphological Analysis and Disambiguation for Morphologically Rich Languages and Universal Dependencies*
Amir More and Reut Tsarfaty

*Automatic Syllabification for Manipuri language*
Loitongbam Gyanendro Singh, Lenin Laitonjam and Sanasam Ranbir Singh
Speech Recognition, Text-To-Speech, Spoken Language Understanding

*Learning to Distill: The Essence Vector Modeling Framework*
Kuan-Yu Chen, Shih-Hung Liu, Berlin Chen and Hsin-Min Wang

*Continuous Expressive Speaking Styles Synthesis based on CVSM and MR-HMM*
Jaime Lorenzo-Trueba, Roberto Barra-Chicote, Ascension Gallardo-Antolin, Junichi Yamagishi and Juan M Montero

*An Automatic Prosody Tagger for Spontaneous Speech*
Monica Dominguez, Mireia Farrús and Leo Wanner

*Frustratingly Easy Neural Domain Adaptation*
Young-Bum Kim, Karl Stratos and Ruhi Sarikaya

Syntactic and Semantic Parsing, Grammar Induction

*A House United: Bridging the Script and Lexical Barrier between Hindi and Urdu*
Riyaz A. Bhat, Irshad A. Bhat, Naman Jain and Dipti Misra Sharma

*Deeper syntax for better semantic parsing*
Olivier Michalon, Corentin Ribeyre, Marie Candito and Alexis Nasr

*Language Independent Dependency to Constituent Tree Conversion*
Young-Suk Lee and Zhiguo Wang

*Promoting multiword expressions in A* TAG parsing*
Jakub Waszczuk, Agata Savary and Yannick Parmentier
Under-resourced Languages

*Incrementally Learning a Dependency Parser to Support Language Documentation in Field Linguistics*
Morgan Ulinski, Julia Hirschberg and Owen Rambow

*Inducing Multilingual Text Analysis Tools Using Bidirectional Recurrent Neural Networks*
Othman ZENNAKI, Nasredine Semmar and Laurent Besacier

*Bitext Name Tagging for Cross-lingual Entity Annotation Projection*
Dongxu Zhang, Boliang Zhang, Xiaoman Pan, Xiaocheng Feng, Heng Ji and Weiran XU

*Determining the Multiword Expression Inventory of a Surprise Language*
Bahar Salehi, Paul Cook and Timothy Baldwin

*A Hybrid Deep Learning Architecture for Sentiment Analysis*
Md Shad Akhtar, Ayush Kumar, Asif Ekbal and Pushpak Bhattacharyya

*Word Segmentation in Sanskrit Using Path Constrained Random Walks*
Amrith Krishna, Bishal Santra, Pavankumar Satuluri, Sasi Prasanth Bandaru, Bhumi Faldu, Yajuvendra Singh and Pawan Goyal

*Mongolian Named Entity Recognition System with Rich Features*
Weihua Wang, Feilong Bao and Guanglai Gao

Applications

*Appraising UMLS Coverage for Summarizing Medical Evidence*
Elaheh ShafieiBavani, Mohammad Ebrahimi, Raymond Wong and Fang Chen

*Large-scale Multi-class and Hierarchical Product Categorization for an E-commerce Giant*
Ali Cevahir and Koji Murakami

*Product Classification in E-Commerce using Distributional Semantics*
Vivek Gupta, Harish Karnick, Ashendra Bansal and Pradhuman Jhala

16:00–16:30  coffee break
Session 3-A: Natural Language Generation, Summarization III

*AttSum: Joint Learning of Focusing and Summarization with Neural Attention*
Ziqiang Cao, Wenjie Li, Sujian Li, Furu Wei and Yanran Li

*Using Relevant Public Posts to Enhance News Article Summarization*
Chen Li, Zhongyu Wei, Yang Liu, Yang Jin and Fei Huang

*A Proposition-Based Abstractive Summariser*
Yimai Fang, Haoyue Zhu, Ewa Muszyńska, Alexander Kuhnle and Simone Teufel

Session 3-B: Syntactic and Semantic Parsing, Grammar Induction II

*Cross-lingual Learning of an Open-domain Semantic Parser*
Kilian Evang and Johan Bos

*A subtree-based factorization of dependency parsing*
Qiuye Zhao and Qun Liu

*K-SRL: Instance-based Learning for Semantic Role Labeling*
Alan Akbik and Yunyao Li

Session 3-C: Morphology, Segmentation Tagging, Chunking II

*Keystroke dynamics as signal for shallow syntactic parsing*
Barbara Plank

*A Bayesian model for joint word alignment and part-of-speech transfer*
Robert Östling

*Splitting compounds with ngrams*
Naomi Tachikawa Shapiro
Tuesday, December 13, 2016 (continued)

16:30–18:00  Session 3-D: Applications II

*GAKE: Graph Aware Knowledge Embedding*
Jun Feng, Minlie Huang, Yang Yang and Xiaoyan Zhu

*Ranking Responses Oriented to Conversational Relevance in Chat-bots*
Bowen Wu, Baoxun Wang and Hui Xue

*Probabilistic Prototype Model for Serendipitous Property Mining*
Taesung Lee, Seung-won Hwang and Zhongyuan Wang

16:30–18:00  Session 3-P: Poster Session 2

**Computational Psycholinguistics**

*Identifying Cross-Cultural Differences in Word Usage*
Aparna Garimella, Rada Mihalcea and James Pennebaker

*Reading-Time Annotations for “Balanced Corpus of Contemporary Written Japanese”*
Masayuki Asahara, Hajime Ono and Edson T. Miyamoto

"How Bullying is this Message?": A Psychometric Thermometer for Bullying
Parma Nand, Rivindu Perera and Abhijeet Kasture

*Learning grammatical categories using paradigmatic representations: Substitute words for language acquisition*
Mehmet Ali Yatbaz, Volkan Cirik, Aylin Künay and Deniz Yuret

*Understanding the Lexical Simplification Needs of Non-Native Speakers of English*
Gustavo Paetzold and Lucia Specia

*How Interlocutors Coordinate with each other within Emotional Segments?*
Firoj Alam, Shammur Absar Chowdhury, Morena Danieli and Giuseppe Riccardi
Linguistic Issues in NLP

Advancing Linguistic Features and Insights by Label-informed Feature Grouping: An Exploration in the Context of Native Language Identification
Serhiy Bykh and Detmar Meurers

Modeling Diachronic Change in Scientific Writing with Information Density
Raphael Rubino, Stefania Degaetano-Ortlieb, Elke Teich and Josef van Genabith

Different Contexts Lead to Different Word Embeddings
Wenpeng Hu, Jiajun Zhang and Nan Zheng

Machine Learning for Metrical Analysis of English Poetry
Manex Agirrezabal, Iñaki Alegria and Mans Hulden

Automated speech-unit delimitation in spoken learner English
Russell Moore, Andrew Caines, Calbert Graham and Paula Buttery

Learning to Identify Sentence Parallelism in Student Essays
Wei Song, Tong Liu, Ruiji Fu, Lizhen Liu, Hanshi Wang and Ting Liu

Evaluating anaphora and coreference resolution to improve automatic keyphrase extraction
Marco Basaldella, Giorgia Chiaradia and Carlo Tasso

Retrieving Occurrences of Grammatical Constructions
Anna Ehrlemark, Richard Johansson and Benjamin Lyngfelt

Automatic Extraction of Learner Errors in ESL Sentences Using Linguistically Enhanced Alignments
Mariano Felice, Christopher Bryant and Ted Briscoe

Contrasting Vertical and Horizontal Transmission of Typological Features
Kenji Yamauchi and Yugo Murawaki

How Regular is Japanese Loanword Adaptation? A Computational Study
Lingshuang Mao and Mans Hulden

Using Linguistic Data for English and Spanish Verb-Noun Combination Identification
Uxoa Iñurrieta, Arantza Diaz de Ilarraza, Gorka Labaka, Kepa Sarasola, Itziar Aduriz and John Carroll
Applications

Analyzing Gender Bias in Student Evaluations
Andamlak Terkik, Emily Prud’hommeaux, Cecilia Ovesdotter Alm, Christopher Homan and Scott Franklin

Adverse Drug Reaction Classification With Deep Neural Networks
Trung Huynh, Yulan He, Alistair Willis and Stefan Rueger

Chinese Preposition Selection for Grammatical Error Diagnosis
Hen-Hsen Huang, Yen-Chi Shao and Hsin-Hsi Chen

Wednesday, December 14, 2016

09:00–10:00 Invited talk 2: Reiko Mazuka (RIKEN Brain Science Institute & Duke University)

10:00–10:30 coffee break

10:30–12:00 Session 4-A: Morphology, Segmentation, Tagging, Chunking III

Extending the Use of Adaptor Grammars for Unsupervised Morphological Segmentation of Unseen Languages
Ramy Eskander, Owen Rambow and Tianchun Yang

CharNER: Character-Level Named Entity Recognition
Onur Kuru, Ozan Arkan Can and Deniz Yuret

A Neural Model for Part-of-Speech Tagging in Historical Texts
Christian Hardmeier
Wednesday, December 14, 2016 (continued)

10:30–12:00  **Session 4-B: Applications III**

*Extracting Discriminative Keyphrases with Learned Semantic Hierarchies*
Yunli Wang, Yong Jin, Xiaodan Zhu and Cyril Goutte

*Hashtag Recommendation Using End-To-End Memory Networks with Hierarchical Attention*
Haoran Huang, Qi Zhang, Yeyun Gong and Xuanjing Huang

*Automatic Labelling of Topics with Neural Embeddings*
Shraey Bhatia, Jey Han Lau and Timothy Baldwin

10:30–12:00  **Session 4-C: Computational Psycholinguistics and Linguistic Issues in NLP I**

*Memory-Bounded Left-Corner Unsupervised Grammar Induction on Child-Directed Input*
Cory Shain, William Bryce, Lifeng Jin, Victoria Krakovna, Finale Doshi-Velez, Timothy Miller, William Schuler and Lane Schwartz

*‘Calling on the classical phone’: a distributional model of adjective-noun errors in learners’ English*
Aurélie Herbelot and Ekaterina Kochmar

*Are Cohesive Features Relevant for Text Readability Evaluation?*
Amalia Todirascu, Thomas Francois, Delphine Bernhard, Nuria Gala and Anne-Laure Ligozat

10:30–12:00  **Session 4-D: Resources, Software, Tools and Under-resourced languages II**

*Named Entity Recognition for Linguistic Rapid Response in Low-Resource Languages: Sorani Kurdish and Tajik*
Patrick Littell, Kartik Goyal, David R. Mortensen, Alexa Little, Chris Dyer and Lori Levin

*Multilingual Supervision of Semantic Annotation*
Peter Exner, Marcus Klang and Pierre Nugues

*Siamese Convolutional Networks for Cognate Identification*
Taraka Rama
Wednesday, December 14, 2016 (continued)

10:30–12:00  Session 4-P: Poster Session 3

Natural Language Generation, Summarization

*Exploring Differential Topic Models for Comparative Summarization of Scientific Papers*
Lei He, Wei Li and Hai Zhuge

*Bridging the gap between extractive and abstractive summaries: Creation and evaluation of coherent extracts from heterogeneous sources*
Darina Benikova, Margot Mieskes, Christian M. Meyer and Iryna Gurevych

*Chinese Poetry Generation with Planning based Neural Network*
Zhe Wang, Wei He, Hua Wu, Haiyang Wu, Wei Li, Haifeng Wang and Enhong Chen

*Predicting sentential semantic compatibility for aggregation in text-to-text generation*
Victor Chenal and Jackie Chi Kit Cheung

*Sequential Clustering and Contextual Importance Measures for Incremental Update Summarization*
Markus Zopf, Eneldo Loza Mencía and Johannes Fürnkranz

*Natural Language Generation through Character-based RNNs with Finite-state Prior Knowledge*
Raghav Goyal, Marc Dymetman and Eric Gaussier

*A Hybrid Approach to Generation of Missing Abstracts in Biomedical Literature*
Suchet Chachra, Asma Ben Abacha, Sonya Shooshan, Laritza Rodriguez and Dina Demner-Fushman

*Imitation learning for language generation from unaligned data*
Gerasimos Lampouras and Andreas Vlachos

*Product Review Summarization by Exploiting Phrase Properties*
Naitong Yu, Minlie Huang, Yuanyuan Shi and xiaoyan zhu

*Generating Questions and Multiple-Choice Answers using Semantic Analysis of Texts*
Jun Araki, Dheeraj Rajagopal, Sreecharan Sankaranarayanan, Susan Holm, Yukari Yamakawa and Teruko Mitamura
Resources, Software and Tools

Evaluation Strategies for Computational Construction Grammars
Tania Marques and Katrien Beuls

Building a Monolingual Parallel Corpus for Text Simplification Using Sentence Similarity Based on Alignment between Word Embeddings
Tomoyuki Kajiwara and Mamoru Komachi

Word2Vec vs DBnary: Augmenting METEOR using Vector Representations or Lexical Resources?
Christophe Servan, Alexandre Berard, zied ellioumi, Hervé Blanchon and Laurent Besacier

Broad Twitter Corpus: A Diverse Named Entity Recognition Resource
Leon Derczynski, Kalina Bontcheva and Ian Roberts

Semantic overfitting: what 'world' do we consider when evaluating disambiguation of text?
Filip Ilievski, Marten Postma and Piek Vossen

Information Retrieval, Information Extraction, Question Answering

Extraction of Keywords of Novelties From Patent Claims
Shoko Suzuki and Hiromichi Takatsuka

Leveraging Multilingual Training for Limited Resource Event Extraction
Andrew Hsi, Yiming Yang, Jaime Carbonell and Ruochen Xu

LILI: A Simple Language Independent Approach for Language Identification
Mohamed Al-Badrashiny and Mona Diab

High Accuracy Rule-based Question Classification using Question Syntax and Semantics
Harish Tayyar Madabushi and Mark Lee

Incorporating Label Dependency for Answer Quality Tagging in Community Question Answering via CNN-LSTM-CRF
Yang Xiang, Xiaojing Zhou, Qingcai Chen, Zhihui Zheng, Buzhou Tang, Xiaolong Wang and Yang Qin

Semantically Motivated Hebrew Verb-Noun Multi-Word Expressions Identification
Chaya Liebeskind and Yaakov HaCohen-Kerner
Thursday, December 15, 2016

09:00–10:00  Invited talk 3: Dina Demner-Fushman (U.S. National Library of Medicine)

10:00–10:30  coffee break

10:30–12:30  Session 5-A: Semantic Processing, Distributional Semantics, Compositionality I

Semantic Relation Classification via Hierarchical Recurrent Neural Network with Attention
Minguang Xiao and Cong Liu

A Unified Architecture for Semantic Role Labeling and Relation Classification
Jiang Guo, Wanxiang Che, Haifeng Wang, Ting Liu and Jun Xu

Facing the most difficult case of Semantic Role Labeling: A collaboration of word embeddings and co-training
Quynh Ngoc Thi Do, Steven Bethard and Marie-Francine Moens

Predictability of Distributional Semantics in Derivational Word Formation
Sebastian Padó, Aurélie Herbelot, Max Kisselew and Jan Šnajder

10:30–12:30  Session 5-B: Computational Psycholinguistics and Linguistic Issues in NLP II

Survey on the Use of Typological Information in Natural Language Processing
Helen O’Horan, Yevgeni Berzak, Ivan Vulic, Roi Reichart and Anna Korhonen

From phonemes to images: levels of representation in a recurrent neural model of visually-grounded language learning
Lieke Gelderloos and Grzegorz Chrupała

Linguistic features for Hindi light verb construction identification
Ashwini Vaidya, Sumeet Agarwal and Martha Pulmer

Cross-lingual Transfer of Correlations between Parts of Speech and Gaze Features
Maria Barrett, Frank Keller and Anders Søgaard
Thursday, December 15, 2016 (continued)

10:30–12:30 Session 5-C: Lexical Semantics, Ontologies & Paraphrasing, Textual Entailment I

Sentence Similarity Learning by Lexical Decomposition and Composition
Zhiguo Wang, Haitao Mi and Abraham Ittycheriah

Chinese Hypernym-Hyponym Extraction from User Generated Categories
Chengyu Wang and Xiaofeng He

Dynamic Generative model for Diachronic Sense Emergence Detection
Martin Emms and Arun Kumar Jayapal

Semi-supervised Word Sense Disambiguation with Neural Models
Dayu Yuan, Julian Richardson, Ryan Doherty, Colin Evans and Eric Altendorf

10:30–12:30 Session 5-D: Machine Translation I

Fast Gated Neural Domain Adaptation: Language Model as a Case Study
Jian Zhang, Xiaofeng Wu, Andy Way and Qun Liu

Machine Translation Evaluation for Arabic using Morphologically-enriched Embeddings
Francisco Guzmán, Houda Bouamor, Ramy Baly and Nizar Habash

Ensemble Learning for Multi-Source Neural Machine Translation
Ekaterina Garmash and Christof Monz

Phrase-based Machine Translation using Multiple Preordering Candidates
Yusuke Oda, Taku Kudo, Tetsuji Nakagawa and Taro Watanabe
Thursday, December 15, 2016 (continued)

10:30–12:30  Session 5-P: Poster Session 4

Information Retrieval, Information Extraction, Question Answering

*Hand in Glove: Deep Feature Fusion Network Architectures for Answer Quality Prediction in Community Question Answering*
Sai Praneeth Suggu, Kushwanth Naga Goutham, Manoj K. Chinnakotla and Manish Shrivastava

*Learning Event Expressions via Bilingual Structure Projection*
Fangyuan Li, Ruihong Huang, Deyi Xiong and Min Zhang

*Global Inference to Chinese Temporal Relation Extraction*
Peifeng Li, Qiaoming Zhu, Guodong Zhou and Hongling Wang

*Improved relation classification by deep recurrent neural networks with data augmentation*
Yan Xu, Ran Jia, Lili Mou, Ge Li, Yunchuan Chen, Yangyang Lu and Zhi Jin

*Relation Extraction with Multi-instance Multi-label Convolutional Neural Networks*
Xiaotian Jiang, Quan Wang, Peng Li and Bin Wang

*Named Entity Disambiguation for little known referents: a topic-based approach*
Andrea Glaser and Jonas Kuhn

Natural Language Generation, Summarization

*Building RDF Content for Data-to-Text Generation*
Laura Perez-Beltrachini, Rania SAYED and Claire Gardent

*Parallel Sentence Compression*
Julia Ive and François Yvon

*An Unsupervised Multi-Document Summarization Framework Based on Neural Document Model*
Shulei Ma, Zhi-Hong Deng and Yunlun Yang

*From OpenCCG to AI Planning: Detecting Infeasible Edges in Sentence Generation*
Maximilian Schwenger, Alvaro Torralba, Joerg Hoffmann, David M. Howcroft and Vera Demberg
The Next Step for Multi-Document Summarization: A Heterogeneous Multi-Genre Corpus Built with a Novel Construction Approach
Markus Zopf, Maxime Peyrard and Judith Eckle-Kohler

Sentiment Analysis and Computational Argumentation

SentiHood: Targeted Aspect Based Sentiment Analysis Dataset for Urban Neighbourhoods
Marzieh Saeidi, Guillaume Bouchard, Maria Liakata and Sebastian Riedel

On the Impact of Seed Words on Sentiment Polarity Lexicon Induction
Dame Jovanoski, Veno Pachovski and Preslav Nakov

Evaluating Argumentative and Narrative Essays using Graphs
Swapna Somasundaran, Brian Riordan, Binod Gyawali and Su-Youn Yoon

Selective Co-occurrences for Word-Emotion Association
Ameeta Agrawal and Aijun An

Weighted Neural Bag-of-n-grams Model: New Baselines for Text Classification
Bofang Li, Zhe Zhao, Tao Liu, Puwei Wang and Xiaoyong Du

A Deeper Look into Sarcastic Tweets Using Deep Convolutional Neural Networks
Soujanya Poria, Erik Cambria, Devamanyu Hazarika and Prateek Vij

Exploring Distributional Representations and Machine Translation for Aspect-based Cross-lingual Sentiment Classification.
Jeremy Barnes, Patrik Lambert and Toni Badia

A Bilingual Attention Network for Code-switched Emotion Prediction
Zhongqing Wang, Yue Zhang, Sophia Lee, Shoushan Li and Guodong Zhou

UTCNN: a Deep Learning Model of Stance Classification on Social Media Text
Wei-Fan Chen and Lun-Wei Ku
Thursday, December 15, 2016 (continued)

**Computational Psycholinguistics**

*The Role of Intrinsic Motivation in Artificial Language Emergence: a Case Study on Colour*
Miquel Cornudella, Thierry Poibeau and Remi van Trijp

*Predicting the Evocation Relation between Lexicalized Concepts*
Yoshihiko Hayashi

*Collecting and Exploring Everyday Language for Predicting Psycholinguistic Properties of Words*
Gustavo Paetzold and Lucia Specia

**Applications**

*Using Argument Mining to Assess the Argumentation Quality of Essays*
Henning Wachsmuth, Khalid Al Khatib and Benno Stein

*Grammatical Templates: Improving Text Difficulty Evaluation for Language Learners*
Shuhan Wang and Erik Andersen

*Still not there? Comparing Traditional Sequence-to-Sequence Models to Encoder-Decoder Neural Networks on Monotone String Translation Tasks*
Carsten Schnober, Steffen Eger, Erik-Lân Do Dinh and Iryna Gurevych

12:30–14:00  *Lunch break*
Thursday, December 15, 2016 (continued)

14:00–16:00  Session 6-A: Information Retrieval, Information Extraction, Question Answering I

Towards Time-Aware Knowledge Graph Completion
Tingsong Jiang, Tianyu Liu, Tao Ge, Lei Sha, Baobao Chang, Sujian Li and Zhifang Sui

Learning to Weight Translations using Ordinal Linear Regression and Query-generated Training Data for Ad-hoc Retrieval with Long Queries
Javid Dadashkarimi, Masoud Jalili Sabet and Azadeh Shakery

Neural Attention for Learning to Rank Questions in Community Question Answering
Salvatore Romeo, Giovanni Da San Martino, Alberto Barrón-Cedeño, Alessandro Moschitti, Yonatan Belinkov, Wei-Ning Hsu, Yu Zhang, Mitra Mohtarami and James Glass

Simple Question Answering by Attentive Convolutional Neural Network
Wenpeng Yin, Mo Yu, Bing Xiang, Bowen Zhou and Hinrich Schütze

14:00–16:00  Session 6-B: Machine Learning for NLP II

Recurrent Dropout without Memory Loss
Stanislau Semeniuta, Aliaksei Severyn and Erhardt Barth

Modeling topic dependencies in semantically coherent text spans with copulas
Georgios Balikas, Hesam Amoualian, Marianne Clausel, Eric Gaussier and Massih R Amini

Consensus Attention-based Neural Networks for Chinese Reading Comprehension
Yiming Cui, Ting Liu, Zhipeng Chen, Shijin Wang and Guoping Hu

Semantic Annotation Aggregation with Conditional Crowdsourcing Models and Word Embeddings
Paul Felt, Eric Ringger and Kevin Seppi
Thursday, December 15, 2016 (continued)

14:00–16:00  Session 6-C: Machine Translation II

Interactive-Predictive Machine Translation based on Syntactic Constraints of Prefix
Na Ye, Guiping Zhang and Dongfeng Cai

Topic-Informed Neural Machine Translation
Jian Zhang, Liangyou Li, Andy Way and Qun Liu

A Distribution-based Model to Learn Bilingual Word Embeddings
Hailong Cao, Tiejun Zhao, Shu ZHANG and Yao Meng

Pre-Translation for Neural Machine Translation
Jan Niehues, Eunah Cho, Thanh-Le Ha and Alex Waibel

14:00–16:00  Session 6-D: Semantic Processing, Distributional Semantics, Compositionality II

Direct vs. indirect evaluation of distributional thesauri
Vincent Claveau and Ewa Kijak

D-GloVe: A Feasible Least Squares Model for Estimating Word Embedding Densities
Shoaib Jameel and Steven Schockaert

Predicting human similarity judgments with distributional models: The value of word associations.
Simon De Deyne, Amy Perfors and Daniel J Navarro

Distributional Hypernym Generation by Jointly Learning Clusters and Projections
Josuke Yamane, Tomoya Takatani, Hitoshi Yamada, Makoto Miwa and Yutaka Sasaki
Thursday, December 15, 2016 (continued)

14:00–16:00  Session 6-P: Poster Session 5

Discourse Relations, Coreference, Pragmatics

**Incremental Fine-grained Information Status Classification Using Attention-based LSTMs**
Yufang Hou

**Detection, Disambiguation and Argument Identification of Discourse Connectives in Chinese Discourse Parsing**
Yong-Siang Shih and Hsin-Hsi Chen

**Multi-view and multi-task training of RST discourse parsers**
Chloé Braud, Barbara Plank and Anders Søgaard

**Implicit Discourse Relation Recognition with Context-aware Character-enhanced Embeddings**
Lianhui Qin, Zhisong Zhang and Hai Zhao

**Measuring Non-cooperation in Dialogue**
Brian Plüss and Paul Piwek

**Representation and Learning of Temporal Relations**
Leon Derczynski

**Revisiting the Evaluation for Cross Document Event Coreference**
Shyam Upadhyay, Nitish Gupta, Christos Christodoulopoulos and Dan Roth

**Modeling Discourse Segments in Lyrics Using Repeated Patterns**
Kento Watanabe, Yuichiroh Matsubayashi, Naho Orita, Naoaki Okazaki, Kentaro Inui, Satoru Fukayama, Tomoyasu Nakano, Jordan Smith and Masataka Goto
Dialog Processing and Dialog Systems, Multimodal Interfaces

*Multi-level Gated Recurrent Neural Network for dialog act classification*  
Wei Li and Yunfang Wu

*Multimodal Mood Classification - A Case Study of Differences in Hindi and Western Songs*  
Braja Gopal Patra, Dipankar Das and Sivaji Bandyopadhyay

*Detecting Context Dependent Messages in a Conversational Environment*  
Chaozhuo Li, Yu Wu, Wei Wu, Chen Xing, Zhoujun Li and Ming Zhou

*Joint Inference for Mode Identification in Tutorial Dialogues*  
Deepak Venugopal and Vasile Rus

*Dialogue Act Classification in Domain-Independent Conversations Using a Deep Recurrent Neural Network*  
Hamed Khanpour, Nishitha Guntakandla and Rodney Nielsen

*Non-sentential Question Resolution using Sequence to Sequence Learning*  
Vineet Kumar and Sachindra Joshi

*Context-aware Natural Language Generation for Spoken Dialogue Systems*  
Hao Zhou, Minlie Huang and xiaoyan zhu

Speech Recognition, Text-To-Speech, Spoken Language Understanding

*Weakly-supervised text-to-speech alignment confidence measure*  
Guillaume Serrière, Christophe Cerisara, Dominique Fohr and Odile Mella

*Domainless Adaptation by Constrained Decoding on a Schema Lattice*  
Young-Bum Kim, Karl Stratos and Ruhi Sarikaya

*Sub-Word Similarity based Search for Embeddings: Inducing Rare-Word Embeddings for Word Similarity Tasks and Language Modelling*  
Mittul Singh, Clayton Greenberg, Youssef Oualil and Dietrich Klakow
Applications

Semi-automatic Detection of Cross-lingual Marketing Blunders based on Pragmatic Label Propagation in Wiktionary
Christian M. Meyer, Judith Eckle-Kohler and Iryna Gurevych

Ambient Search: A Document Retrieval System for Speech Streams
Benjamin Milde, Jonas Wacker, Stefan Radomski, Max Mühlhäuser and Chris Biemann

Semi-supervised Gender Classification with Joint Textual and Social Modeling
Shoushan Li, Bin Dai, Zhengxian Gong and Guodong Zhou

Predicting proficiency levels in learner writings by transferring a linguistic complexity model from expert-written coursebooks
Ildikó Pilán, Elena Volodina and Torsten Zesch

User Classification with Multiple Textual Perspectives
Dong Zhang, Shoushan Li, Hongling Wang and Guodong Zhou

Says Who…? Identification of Expert versus Layman Critics’ Reviews of Documentary Films
Ming Jiang and Jana Diesner

Knowledge-Driven Event Embedding for Stock Prediction
Xiao Ding, Yue Zhang, Ting Liu and Junwen Duan

Distributed Representations for Building Profiles of Users and Items from Text Reviews
Wenliang Chen, Zhenjie Zhang, Zhenghua Li and Min Zhang

16:00–16:30 coffee break
Thursday, December 15, 2016 (continued)

16:30–18:00  Session 7-A: Machine Translation III

Improving Statistical Machine Translation with Selectional Preferences
Haiqing Tang, Deyi Xiong, Min Zhang and Zhengxian Gong

Hierarchical Permutation Complexity for Word Order Evaluation
Miloš Stanojević and Khalil Sima’an

Interactive Attention for Neural Machine Translation
Fandong Meng, Zhengdong Lu, Hang Li and Qun Liu

16:30–18:00  Session 7-B: Applications IV

Get Semantic With Me! The Usefulness of Different Feature Types for Short-Answer Grading
Ulrike Pado

Automatically Processing Tweets from Gang-Involved Youth: Towards Detecting Loss and Aggression
Terra Blevins, Robert Kwiatkowski, Jamie MacBeth, Kathleen McKeown, Desmond Patton and Owen Rambow

Content-based Influence Modeling for Opinion Behavior Prediction
Chengyao Chen, Zhitao Wang, Yu Lei and Wenjie Li

16:30–18:00  Session 7-C: Computational Psycholinguistics and Linguistic Issues in NLP III

Data-driven learning of symbolic constraints for a log-linear model in a phonological setting
Gabriel Doyle and Roger Levy

Chinese Tense Labelling and Causal Analysis
Hen-Hsen Huang, Chang-Rui Yang and Hsin-Hsi Chen

Exploring Topic Discriminating Power of Words in Latent Dirichlet Allocation
Yang Kai, Cai Yi, Chen Zhenhong, Leung Ho-fung and LAU Raymond
Thursday, December 15, 2016 (continued)

16:30–18:00  Session 7-D: Lexical Semantics, Ontologies & Paraphrasing, Textual Entailment II

Textual Entailment with Structured Attentions and Composition
Kai Zhao, Liang Huang and Mingbo Ma

plWordNet 3.0 – a Comprehensive Lexical-Semantic Resource
Marek Maziarz, Maciej Piasecki, Ewa Rudnicka, Stan Szpakowicz and Paweł Kędzia

Time-Independent and Language-Independent Extraction of Multiword Expressions From Twitter
Nikhil Londhe, Rohini Srihari and Vishrawas Gopalakrishnan

16:30–18:00  Session 7-P: Poster Session 6

Information Retrieval, Information Extraction, Question Answering

Incremental Global Event Extraction
Alex Judea and Michael Strube

Hierarchical Memory Networks for Answer Selection on Unknown Words
jiaming xu, Jing Shi, Yiqun Yao, Suncong Zheng, Bo Xu and Bo Xu

Revisiting Taxonomy Induction over Wikipedia
Amit Gupta, Francesco Piccinno, Mikhail Kozhevnikov, Marius Pasca and Daniele Pighin

Joint Learning of Local and Global Features for Entity Linking via Neural Networks
Thien Huu Nguyen, Nicolas Faucoglia, Mariano Rodriguez Muro, Oktie Hassanzadeh, Alfio Massimiliano Gliozzo and Mohammad Sadoghi

Structured Aspect Extraction
Omer Gunes, Tim Furche and Giorgio Orsi

Robust Text Classification for Sparsely Labelled Data Using Multi-level Embeddings
Simon Baker, Douwe Kiela and Anna Korhonen

Mathematical Information Retrieval based on Type Embeddings and Query Expansion
Yiannos Stathopoulos and Simone Teufel
Text Retrieval by Term Co-occurrences in a Query-based Vector Space
Eriks Sneiders

Pairwise Relation Classification with Mirror Instances and a Combined Convolutional Neural Network
Jianfei Yu and Jing Jiang

FastHybrid: A Hybrid Model for Efficient Answer Selection
Lidan Wang, Ming Tan and Jiawei Han

Extracting Spatial Entities and Relations in Korean Text
Bogyum Kim and Jae Sung Lee

Hybrid Question Answering over Knowledge Base and Free Text
kun xu, Yansong Feng, Songfang Huang and Dongyan Zhao

Improved Word Embeddings with Implicit Structure Information
Jie Shen and Cong Liu

Sentiment Analysis, Computational Argumentation

Word Embeddings and Convolutional Neural Network for Arabic Sentiment Classification
Abdelghani Dahou, Shengwu Xiong, Junwei Zhou, Mohamed Houcine Haddoud and Pengfei Duan

Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts
Xingyou Wang, Weijie Jiang and Zhiyong Luo

Stance Classification in Rumours as a Sequential Task Exploiting the Tree Structure of Social Media Conversations
Arkaitz Zubiaga, Elena Kochkina, Maria Liakata, Rob Procter and Michal Lukasik

Tweet Sarcasm Detection Using Deep Neural Network
Meishan Zhang, Yue Zhang and Guohong Fu

Agreement and Disagreement: Comparison of Points of View in the Political Domain
Stefano Menini and Sara Tonelli

Targeted Sentiment to Understand Student Comments
Charles Welch and Rada Mihalcea
Thursday, December 15, 2016 (continued)

Towards Sub-Word Level Compositions for Sentiment Analysis of Hindi-English Code Mixed Text
Aditya Joshi, Ameya Prabhu, Manish Shrivastava and Vasudeva Varma

Distance Metric Learning for Aspect Phrase Grouping
Shufeng Xiong, Yue Zhang, Donghong Ji and Yinxia Lou

Friday, December 16, 2016

09:00–10:00 Invited talk 4: Simone Teufel (University of Cambridge)

10:00–10:30 coffee break

10:30–12:30 Session 8-A: Information Retrieval, Information Extraction, Question Answering II

Constraint-Based Question Answering with Knowledge Graph
Junwei Bao, Nan Duan, Zhao Yan, Ming Zhou and Tiejun Zhao

Selecting Sentences versus Selecting Tree Constituents for Automatic Question Ranking
Alberto Barrón-Cedeño, Giovanni Da San Martino, Salvatore Romeo and Alessandro Moschitti

Attention-Based Convolutional Neural Network for Semantic Relation Extraction
yatian shen and Xuanjing Huang

Table Filling Multi-Task Recurrent Neural Network for Joint Entity and Relation Extraction
Pankaj Gupta, Hinrich Schütze and Bernt Andrassy
Friday, December 16, 2016 (continued)

10:30–12:30  Session 8-B: Machine Translation IV

Bilingual Autoencoders with Global Descriptors for Modeling Parallel Sentences
Biao Zhang, Deyi Xiong, Jinsong Su, Hong Duan and Min Zhang

Multi-Engine and Multi-Alignment Based Automatic Post-Editing and its Impact on Translation Productivity
Santanu Pal, Sudip Kumar Naskar and Josef van Genabith

Measuring the Effect of Conversational Aspects on Machine Translation Quality
Marlies van der Wees, Arianna Bisazza and Christof Monz

Enriching Phrase Tables for Statistical Machine Translation Using Mixed Embeddings
Peyman Passban, Qun Liu and Andy Way

10:30–12:30  Session 8-C: Discourse Relations, Coreference, Pragmatics

Anecdote Recognition and Recommendation
Wei Song, Ruiji Fu, Lizhen Liu, Hanshi Wang and Ting Liu

Training Data Enrichment for Infrequent Discourse Relations
Kailang Jiang, Giuseppe Carenini and Raymond Ng

Inferring Discourse Relations from PDTB-style Discourse Labels for Argumentative Revision Classification
Fan Zhang, Diane Litman and Katherine Forbes-Riley

Capturing Pragmatic Knowledge in Article Usage Prediction using LSTMs
Jad Kabbara, Yulan Feng and Jackie Chi Kit Cheung
10:30–12:30  Session 8-D: Sentiment Analysis, Computational Argumentation I

Aspect Based Sentiment Analysis using Sentiment Flow with Local and Non-local Neighbor Information
Shubham Pateria

Two-View Label Propagation to Semi-supervised Reader Emotion Classification
Shoushan Li, Jian Xu, Dong Zhang and Guodong Zhou

A Joint Sentiment-Target-Stance Model for Stance Classification in Tweets
Javid Ebrahimi, Dejing Dou and Daniel Lowd

SenticNet 4: A Semantic Resource for Sentiment Analysis Based on Conceptual Primitives
Erik Cambria, Soujanya Poria, Rajiv Bajpai and Bjoern Schuller

10:30–12:30  Session 8-P: Poster Session 7

Machine Learning for NLP

Joint Embedding of Hierarchical Categories and Entities for Concept Categorization and Dataless Classification
Yuezhang(Music) Li, Ronghuo Zheng, Tian Tian, Zhiting Hu, Rahul Iyer and Katia Sycara

Latent Topic Embedding
Di Jiang, Lei Shi, Rongzhong Lian and Hua Wu

Neural-based Noise Filtering from Word Embeddings
Kim Anh Nguyen, Sabine Schulte im Walde and Ngoc Thang Vu

Integrating Distributional and Lexical Information for Semantic Classification of Words using MRMF
Rosa Tsegaye Aga, Lucas Drumond, Christian Wartena and Lars Schmidt-Thieme

Semi Supervised Preposition-Sense Disambiguation using Multilingual Data
Hila Gonen and Yoav Goldberg

Monday mornings are my fave :) #not Exploring the Automatic Recognition of Irony in English tweets
Cynthia Van Hee, Els Lefever and Veronique Hoste
CNN- and LSTM-based Claim Classification in Online User Comments
Chinnappa Guggilla, Tristan Miller and Iryna Gurevych

Experiments in Idiom Recognition
Jing Peng and Anna Feldman

An Empirical Evaluation of various Deep Learning Architectures for Bi-Sequence Classification Tasks
Anirban Laha and Vikas Raykar

Learning Succinct Models: Pipelined Compression with L1-Regularization, Hashing, Elias-Fano Indices, and Quantization
Hajime Senuma and Akiko Aizawa

Semantic Processing, Distributional Semantics, Compositionality

Bad Company—Neighborhoods in Neural Embedding Spaces Considered Harmful
Johannes Hellrich and Udo Hahn

Implementing a Reverse Dictionary, based on word definitions, using a Node-Graph Architecture
Sushrut Thorat and Varad Choudhari

Guillem Collell and Marie-Francine Moens

On the contribution of word embeddings to temporal relation classification
Paramita Mirza and Sara Tonelli

Modeling Context-sensitive Selectional Preference with Distributed Representations
Naoya Inoue, Yuichiroh Matsubayashi, Masayuki Ono, Naoaki Okazaki and Kentaro Inui

Exploring the value space of attributes: Unsupervised bidirectional clustering of adjectives in German
Wiebke Petersen and Oliver Hellwig

Distributional Inclusion Hypothesis for Tensor-based Composition
Dimitri Kartsaklis and Mehrnoosh Sadrzadeh

Parameter estimation of Japanese predicate argument structure analysis model using eye gaze information
Ryosuke Maki, Hitoshi Nishikawa and Takenobu Tokunaga

lxxi
Paraphrasing, Textual Entailment

*Reading and Thinking: Re-read LSTM Unit for Textual Entailment Recognition*
Lei Sha, Baobao Chang, Zhifang Sui and Sujian Li

*A Paraphrase and Semantic Similarity Detection System for User Generated Short-Text Content on Microblogs*
Kuntal Dey, Ritvik Shrivastava and Saroj Kaushik

*Modeling Extractive Sentence Intersection via Subtree Entailment*
Omer Levy, Ido Dagan, Gabriel Stanovsky, Judith Eckle-Kohler and Iryna Gurevych

*Context-Sensitive Inference Rule Discovery: A Graph-Based Method*
Xianpei Han and Le Sun

*Modelling Sentence Pairs with Tree-structured Attentive Encoder*
Yao Zhou, Cong Liu and Yan Pan

*Neural Paraphrase Generation with Stacked Residual LSTM Networks*
aaditya prakash, Sadid A. Hasan, Kathy Lee, Vivek Datla, Ashequl Qadir, Joey Liu and Oladimeji Farri

12:30–14:00 *Lunch break*

14:00–15:30 Session 9-A: Information Retrieval, Information Extraction, Question Answering III

*English-Chinese Knowledge Base Translation with Neural Network*
Xiaocheng Feng, Duyu Tang, Bing Qin and Ting Liu

*Keyphrase Annotation with Graph Co-Ranking*
Adrien Bougouin, Florian Boudin and Beatrice Daille

*What’s in an Explanation? Characterizing Knowledge and Inference Requirements for Elementary Science Exams*
Peter Jansen, Niranjan Balasubramanian, Mihai Surdeanu and Peter Clark
14:00–15:30 Session 9-B: Sentiment Analysis, Computational Argumentation II

“All I know about politics is what I read in Twitter”: Weakly Supervised Models for Extracting Politicians’ Stances From Twitter
Kristen Johnson and Dan Goldwasser

Leveraging Multiple Domains for Sentiment Classification
Fan Yang, Arjun Mukherjee and Yifan Zhang

Political News Sentiment Analysis for Under-resourced Languages
Patrik F. Bakken, Terje A. Bratlie, Cristina Marco and Jon Atle Gulla

14:00–15:30 Session 9-C: Applications V

Fast Inference for Interactive Models of Text
Jeffrey Lund, Paul Felt, Kevin Seppi and Eric Ringger

Combining Heterogeneous User Generated Data to Sense Well-being
Adam Tsakalidis, Maria Liakata, Theo Damoulas, Brigitte Jellinek, Weisi Guo and Alexandra Cristea

Hashtag Recommendation with Topical Attention-Based LSTM
Yang Li, Ting Liu, Jing Jiang and Liang Zhang

14:00–15:30 Session 9-D: Resources, Software, Tools & Under-resourced languages III

Better call Saul: Flexible Programming for Learning and Inference in NLP
Parisa Kordjamshidi, Daniel Khashabi, Christos Christodoulopoulos, Bhargav Mangipudi, Sameer Singh and Dan Roth

Crowdsourcing Complex Language Resources: Playing to Annotate Dependency Syntax
Bruno Guillaume, Karèn Fort and Nicolas Lefebvre

Borrow a Little from your Rich Cousin: Using Embeddings and Polarities of English Words for Multilingual Sentiment Classification
Prerana Singhal and Pushpak Bhattacharyya
14:00–15:30 Session 9-P: Poster Session 8

Machine Translation

A Character-Aware Encoder for Neural Machine Translation
Zhen Yang, Wei Chen, Feng Wang and Bo Xu

Convolution-Enhanced Bilingual Recursive Neural Network for Bilingual Semantic Modeling
jinsong su, Biao Zhang, Deyi Xiong, Ruochen Li and Jianmin Yin

Improving Attention Modeling with Implicit Distortion and Fertility for Machine Translation
Shi Feng, Shujie Liu, Nan Yang, Mu Li, Ming Zhou and Kenny Q. Zhu

Neural Machine Translation with Supervised Attention
Lemao Liu, Masao Utiyama, Andrew Finch and Eiichiro Sumita

Lightly Supervised Quality Estimation
Matthias Sperber, Graham Neubig, Jan Niehues, Sebastian Stüker and Alex Waibel

Improving Translation Selection with Supersenses
Haiqing Tang, Deyi Xiong, Oier Lopez de Lacalle and Eneko Agirre

Is all that Glitters in Machine Translation Quality Estimation really Gold?
Yvette Graham, Timothy Baldwin, Meghan Dowling, Maria Eskevich, Teresa Lynn and Lamia Tounsi

Connecting Phrase based Statistical Machine Translation Adaptation
Rui Wang, Hai Zhao, Bao-Liang Lu, Masao Utiyama and Eiichiro Sumita

Fast Collocation-Based Bayesian HMM Word Alignment
Philip Schulz and Wilker Aziz

Learning to translate from graded and negative relevance information
Laura Jehl and Stefan Riezler

Universal Reordering via Linguistic Typology
Joachim Daiber, Miloš Stanojević and Khalil Sima’an
A Deep Fusion Model for Domain Adaptation in Phrase-based MT
Nadir Durrani, Hassan Sajjad, Shafiq Joty and Ahmed Abdelali

Inducing Bilingual Lexica From Non-Parallel Data With Earth Mover’s Distance Regularization
Meng Zhang, Yang Liu, Huanbo Luan, Yiqun Liu and Maosong Sun

What Makes Word-level Neural Machine Translation Hard: A Case Study on English-German Translation
Fabian Hirschmann, Jinseok Nam and Johannes Fürnkranz

Improving Word Alignment of Rare Words with Word Embeddings
Masoud Jalili Sabet, Heshaam Faili and Gholamreza Haffari

Applications

Measuring the Information Content of Financial News
Ching-Yun Chang, Yue Zhang, Zhiyang Teng, Zahn Bozanic and Bin Ke

Automatic Generation and Classification of Minimal Meaningful Propositions in Educational Systems
Andreea Godea, Florin Bulgarov and Rodney Nielsen

First Story Detection using Entities and Relations
Nikolaos Panagiotou, Cem Akkaya, Kostas Tsioutsiouliklis, Vana Kalogeraki and Dimitrios Gunopulos

Textual complexity as a predictor of difficulty of listening items in language proficiency tests
Anastassia Loukina, Su-Youn Yoon, Jennifer Sakano, Youhua Wei and Kathy Sheehan

Renfen HU, Jiayong Chen and Kuang-hua Chen

15:30–16:00 coffee break
Friday, December 16, 2016 (continued)

16:00–17:00 Session 10-A: Information Retrieval, Information Extraction, Question Answering IV

Joint Inference for Event Coreference Resolution
Jing Lu, Deepak Venugopal, Vibhav Gogate and Vincent Ng

Event Detection with Burst Information Networks
Tao Ge, Lei Cui, Baobao Chang, Zhifang Sui and Ming Zhou

16:00–17:30 Session 10-B: Sentiment Analysis, Computational Argumentation III

Corpus Fusion for Emotion Classification
Suyang Zhu, Shoushan Li, Ying Chen and Guodong Zhou

Effective LSTMs for Target-Dependent Sentiment Classification
Duyu Tang, Bing Qin, Xiaocheng Feng and Ting Liu

Towards assessing depth of argumentation
Manfred Stede

16:00–17:30 Session 10-C: Applications VI

Video Event Detection by Exploiting Word Dependencies from Image Captions
Sang Phan, Yusuke Miyao, Duy-Dinh Le and Shin’ichi Satoh

Predicting Restaurant Consumption Level through Social Media Footprints
Yang Xiao, Yuan Wang, Hangyu Mao and Zhen Xiao

A Novel Fast Framework for Topic Labeling Based on Similarity-preserved Hashing
Xian-Ling Mao, Yi-Jing Hao, Qiang Zhou, Wen-Qing Yuan, Liner Yang and Heyan Huang
16:00–17:30  Session 10-D: Dialog Processing and Dialog Systems, Multimodal Interfaces

Sequence to Backward and Forward Sequences: A Content-Introducing Approach to Generative Short-Text Conversation
Lili Mou, Yiping Song, Rui Yan, Ge Li, Lu Zhang and Zhi Jin

Disfluent but effective? A quantitative study of disfluencies and conversational moves in team discourse
Felix Gervits, Kathleen Eberhard and Matthias Scheutz

A Neural Network Approach for Knowledge-Driven Response Generation
Pavlos Vougiouklis, Jonathon Hare and Elena Simperl

16:00–17:30  Session 10-P: Poster Session 9

Resources, Software and Tools

PersoNER: Persian Named-Entity Recognition
Hanieh Poostchi, Ehsan Zare Borzesi, Mohammad Abdous and Massimo Piccardi

OCR++: A Robust Framework For Information Extraction from Scholarly Articles
Mayank Singh, Barnopriyo Barua, Priyank Palod, Manvi Garg, Sidhartha Satapathy, Samuel Bushi, Kumar Ayush, Krishna Sai Rohith, Tulasi Gamidi, Pawan Goyal and Animesh Mukherjee

Efficient Data Selection for Bilingual Terminology Extraction from Comparable Corpora
Amir Hazem and Emmanuel Morin

TweetGeo - A Tool for Collecting, Processing and Analysing Geo-encoded Linguistic Data
Nikola Ljubešić, Tanja Samardzic and Curdin Derungs

Extending WordNet with Fine-Grained Collocational Information via Supervised Distributional Learning
Luis Espinosa Anke, Jose Camacho-Collados, Sara Rodríguez-Fernández, Horacio Saggion and Leo Wanner

A News Editorial Corpus for Mining Argumentation Strategies
Khalid Al Khatib, Henning Wachsmuth, Johannes Kiesel, Matthias Hagen and Benno Stein

Universal Dependencies for Turkish
Umut Sulubacak, Memduh Gokirmak, Francis Tyers, Çağrı Çöltekin, Joakim Nivre and Gülşen Eryiğit
Creating Resources for Dialectal Arabic from a Single Annotation: A Case Study on Egyptian and Levantine
Ramy Eskander, Nizar Habash, Owen Rambow and Arfath Pasha

Multilingual Aliasing for Auto-Generating Proposition Banks
Alan Akbik, Xinyu Guan and Yunyao Li

PanPhon: A Resource for Mapping IPA Segments to Articulatory Feature Vectors
David R. Mortensen, Patrick Littell, Akash Bharadwaj, Kartik Goyal, Chris Dyer and Lori Levin

Semantic Processing, Distributional Semantics, Compositionality

Text Classification Improved by Integrating Bidirectional LSTM with Two-dimensional Max Pooling
Peng Zhou, Zhenyu Qi, Suncong Zheng, Jiaming Xu, Hongyun Bao and Bo Xu

More is not always better: balancing sense distributions for all-words Word Sense Disambiguation
Marten Postma, Ruben Izquierdo Bevia and Piek Vossen

Language classification from bilingual word embedding graphs
Steffen Eger, Armin Hoenen and Alexander Mehler

Word Embeddings, Analogies, and Machine Learning: Beyond king - man + woman = queen
Aleksandr Drozd, Anna Gladkova and Satoshi Matsuoka

Semantic Tagging with Deep Residual Networks
Johannes Bjerva, Barbara Plank and Johan Bos
Lexical Semantics, Ontologies

A Supervised Approach for Enriching the Relational Structure of Frame Semantics in FrameNet
Shafqat Mumtaz Virk, Philippe Muller and Juliette Conrath

Reddit Temporal N-gram Corpus and its Applications on Paraphrase and Semantic Similarity in Social Media using a Topic-based Latent Semantic Analysis
Anh Dang, Abidalrahman Moh’d, Aminul Islam, Rosane Minghim, Michael Smit and Evangelos Milios

Dictionaries as Networks: Identifying the graph structure of Ogden's Basic English
Camilo Garrido and Claudio Gutierrez

Structured Generative Models of Continuous Features for Word Sense Induction
Alexandros Komninos and Suresh Manandhar

17:45–18:30 Closing