CoNLL 2017

Proceedings of the

CoNLL SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection

August 3–4, 2017
Vancouver, Canada

Support:

©2017 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-945626-69-2

Preface

This volume contains the system description papers associated with the CoNLL-SIGMORPHON shared task in morphological reinflection held at CoNLL 2017. This is the first time a CoNLL shared task has directly addressed the learning of morphology from examples—a fundamental task in NLP where good solutions promise to benefit many downstream tasks. Moreover, models that learn complex morphological patterns from example data are also of significant linguistic interest.

To support the task, we collected and curated data from 52 languages, forming a typologically and genealogically diverse data set against which to evaluate performance of the systems. We divided the learning challenge into two sub-tasks: (1) learning to inflect nouns, adjectives, and verbs from their lemmata (citation forms) into a desired target form, and (2) completing partially filled inflection tables or paradigms. Both of these tasks have been discussed in the NLP and linguistics literature. Participants were further asked to complete each sub-task under a variety of different data conditions.

A total of 12 teams with members from 15 institutions participated in the shared task with a total of 27 system submissions. Of these, 11 submitted system description papers, which are included here. Consistent with last year's SIGMORPHON 2016 shared task results, neural network models performed very well in each data condition, including with a very low-resource training set. Another noteworthy aspect of the results is formed by the various biasing and data augmentation solutions that the different teams exploited to yield good performance with scarce examples.

The creation of several components in the shared task received support from DARPA I20 in the program Low Resource Languages for Emergent Incidents (LORELEI). We wish to thank Google for sponsoring an award given to the strongest overall system(s) and the organizers of CoNLL 2017 for their help. We also want to thank the participants and other members of the community who often provided thoughtful commentary on the data and the task itself.

We hope the data sets, which are now available, will serve as a useful resource to develop further techniques and research into morphological learning.

Mans Hulden, on behalf of the shared task organizers
June 2017
Boulder, CO

Organizers:

Mans Hulden (chair)	University of Colorado
Ryan Cotterell	Johns Hopkins University
Jason Eisner	Johns Hopkins University
Manaal Faruqui	Google
Christo Kirov	Johns Hopkins University
Sandra Kübler	Indiana University
John Sylak-Glassman	Johns Hopkins University
Ekaterina Vylomova	University of Melbourne
Géraldine Walther	University of Zurich
Patrick Xia	Johns Hopkins University
David Yarowsky	Johns Hopkins University

Table of Contents

CoNLL-SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection in 52 Languages
Ryan Cotterell, Christo Kirov, John Sylak-Glassman, Géraldine Walther, Ekaterina Vylomova, Patrick Xia, Manaal Faruqui, Sandra Kübler, David Yarowsky, Jason Eisner and Mans Hulden 1

Training Data Augmentation for Low-Resource Morphological Inflection
Toms Bergmanis, Katharina Kann, Hinrich Schütze and Sharon Goldwater 31

The LMU System for the CoNLL-SIGMORPHON 2017 Shared Task on Universal Morphological Reinflection
Katharina Kann and Hinrich Schütze 40

Align and Copy: UZH at SIGMORPHON 2017 Shared Task for Morphological Reinflection
Peter Makarov, Tatiana Ruzsics and Simon Clematide 49

Morphological Inflection Generation with Multi-space Variational Encoder-Decoders
Chunting Zhou and Graham Neubig 58

ISI at the SIGMORPHON 2017 Shared Task on Morphological Reinflection
Abhisek Chakrabarty and Utpal Garain 66

Experiments on Morphological Reinflection: CoNLL-2017 Shared Task
Akhilesh Sudhakar and Anil Kumar Singh 71

If you can't beat them, join them: the University of Alberta system description
Garrett Nicolai, Bradley Hauer, Mohammad Motallebi, Saeed Najafi and Grzegorz Kondrak ... 79

Character Sequence-to-Sequence Model with Global Attention for Universal Morphological Reinflection
Qile Zhu, Yanjun Li and Xiaolin Li 85

Data Augmentation for Morphological Reinflection
Miikka Silfverberg, Adam Wiemerslage, Ling Liu and Lingshuang Jack Mao 90

Seq2seq for Morphological Reinflection: When Deep Learning Fails
Hajime Senuma and Akiko Aizawa 100

SU-RUG at the CoNLL-SIGMORPHON 2017 shared task: Morphological Inflection with Attentional Sequence-to-Sequence Models
Robert Östling and Johannes Bjerva 110

Conference Program

Thursday, August 3rd, 2017

11:00–11:30 *CoNLL-SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection in 52 Languages*
Ryan Cotterell, Christo Kirov, John Sylak-Glassman, Géraldine Walther, Ekaterina Vylomova, Patrick Xia, Manaal Faruqui, Sandra Kübler, David Yarowsky, Jason Eisner and Mans Hulden

11:30–12:30: Poster session: shared task systems

Training Data Augmentation for Low-Resource Morphological Inflection
Toms Bergmanis, Katharina Kann, Hinrich Schütze and Sharon Goldwater

The LMU System for the CoNLL-SIGMORPHON 2017 Shared Task on Universal Morphological Reinflection
Katharina Kann and Hinrich Schütze

Align and Copy: UZH at SIGMORPHON 2017 Shared Task for Morphological Reinflection
Peter Makarov, Tatiana Ruzsics and Simon Clematide

Morphological Inflection Generation with Multi-space Variational Encoder-Decoders
Chunting Zhou and Graham Neubig

ISI at the SIGMORPHON 2017 Shared Task on Morphological Reinflection
Abhisek Chakrabarty and Utpal Garain

Experiments on Morphological Reinflection: CoNLL-2017 Shared Task
Akhilesh Sudhakar and Anil Kumar Singh

If you can't beat them, join them: the University of Alberta system description
Garrett Nicolai, Bradley Hauer, Mohammad Motallebi, Saeed Najafi and Grzegorz Kondrak

Character Sequence-to-Sequence Model with Global Attention for Universal Morphological Reinflection
Qile Zhu, Yanjun Li and Xiaolin Li

Data Augmentation for Morphological Reinflection
Miikka Silfverberg, Adam Wiemerslage, Ling Liu and Lingshuang Jack Mao

Seq2seq for Morphological Reinflection: When Deep Learning Fails
Hajime Senuma and Akiko Aizawa

SU-RUG at the CoNLL-SIGMORPHON 2017 shared task: Morphological Inflection with Attentional Sequence-to-Sequence Models
Robert Östling and Johannes Bjerva