EMNLP-IJCNLP 2019

Multilingual Surface Realisation

Proceedings of the Second Workshop

November 3, 2019
Hong Kong, China

©2019 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 978-1-950737-76-5

Preface

The Second Workshop on Multilingual Surface Realisation (MSR 2019) was held as part of EMNLP 2019 in Hong Kong on 3 November 2019. The MSR workshops aim to bring together researchers interested in surface-oriented Natural Language Generation problems such as word order determination, inflection, functional word determination, etc. A central part of the MSR workshops is an evolving shared task on surface realisation (SR). Following a pilot task in 2011 for English only, the SR shared task went multilingual from 2018, continuing to include both a shallow track (generating from full universal dependency structures) and a deep track (generating from underspecified UD structures). Workshop and shared task are endorsed by the ACL Special Interest Group on Natural Language Generation (SIGGEN).

The 2019 edition of the SR task (SR'19) offered 11 different languages (up from 10 in SR'18) and attracted 33 team registrations from 17 countries (up from 21 registrations for SR'18). 14 teams submitted systems to SR'19 (up from 8 in SR'18), with two teams withdrawing post submission. Nine teams participated in the Shallow Track only, one in the Deep Track only, and two teams took part in both. All submitting teams submitted a system for English, four teams submitted for English only, four teams submitted for all 11 languages, and four teams submitted for between three and 9 languages.

For English, we evaluated 12 Shallow Track systems and four Deep Track systems in human evaluations of readability and meaning similarity (to reference sentences). Not only did we have multiple Deep Track systems (compared to just one in 2018), but the best Deep Track system actually performed equally well or better than most Shallow Track systems for both readability and meaning similarity. Moreover, the best Shallow Track systems are beginning to close the gap to human toplines, in particular for English and Spanish. In terms of progress, the success of the Deep Track systems represents the biggest leap forward from SR'18, while it looks likely that the shallow systems will catch up with human toplines in the near future. The SR tasks have clearly demonstrated that generation from structured meaning representations can be done with impressive success by current neural methods.

MSR 2019 was pleased to host two invited talks, one by Claire Gardent of Nancy University, and one by the artist Maurice Benayoun, also known as MoBen or 莫奔, who is based in Hong Kong. In addition to papers related to the SR'19 shared task, we accepted one paper on wider surface realisation. Given the increased interest and progress we are able to report for SR'19, we plan to continue with a third shared task in 2020, as part of which we plan to investigate ways of linking up to earlier stages of automatic language generation.

We gratefully acknowledge the hard work put in by the SR'19 participating teams, reviewers and local organisers, and more generally, the creativity and enthusiasm generated by participants in the MSR workshops and SR tasks which is of course what keeps them both going.

Organizers:

Simon Mille, Pompeu Fabra University, Spain
Anja Belz, University of Brighton, UK
Bernd Bohnet, Google Research, UK
Yvette Graham, ADAPT Center, Dublin City University, Ireland
Leo Wanner, ICREA and Pompeu Fabra University, Spain

Program Committee:

Jose Maria Alonso, University of Santiago de Compostela, Spain
Miguel Ballesteros, IBM Research, USA
Alberto Bugarín, University of Santiago de Compostela, Spain
Claire Gardent, CNRS, LORIA, France
Kim Gerdes, Sorbonne Nouvelle, France
Yannis Konstas, Heriot Watt University, UK
Emiel Krahmer, Tilburg University, The Netherlands
David McDonald, SIFT, USA
Ryan McDonald, Google Research, USA
Shashi Narayan, University of Edinburgh, UK
Alexis Nasr, University of Aix Marseille, France
Joakim Nivre, Uppsala University, Sweden
Jekaterina Novikova, Heriot Watt University, UK
Stephan Oepen, University of Oslo, Norway
Emily Pitler, Google Research, USA
Ehud Reiter, Aberdeen University, UK
Horacio Saggion, Pompeu Fabra University, Spain
Kees Van Deemter, Utrecht University, The Netherlands
Michael White, Ohio State University, USA
Sina Zarrieß, University of Bielefeld, Germany

Additional Reviewers:

Valerio Basile, Torino University, Italy
Laura Pérez Mayos, Pompeu Fabra University, Spain

Invited Speakers:

Claire Gardent, CNRS-LORIA, France
Maurice Benayoun (a.k.a. MoBen), City University of Hong Kong, China

Table of Contents

The Second Multilingual Surface Realisation Shared Task (SR'19): Overview and Evaluation Results
Simon Mille, Anja Belz, Bernd Bohnet, Yvette Graham and Leo Wanner 1

Learning to Order Graph Elements with Application to Multilingual Surface Realization
Wenchao Du and Alan W Black 18

DepDist: Surface realization via regex and learned dependency-distance tolerance
William Dyer 25

BME-UW at SRST-2019: Surface realization with Interpreted Regular Tree Grammars
Ádám Kovács, Evelin Ács, Judit Ács, Andras Kornai and Gábor Recski 35

Realizing Universal Dependencies Structures
Guy Lapalme 41

IMSurReal: IMS at the Surface Realization Shared Task 2019
Xiang Yu, Agnieszka Falenska, Marina Haid, Ngoc Thang Vu and Jonas Kuhn 50

Surface Realization Shared Task 2019 (MSR19): The Team 6 Approach
Thiago Castro Ferreira and Emiel Krahmer 59

The Concordia NLG Surface Realizer at SRST 2019
Farhood Farahnak, Laya Rafiee, Leila Kosseim and Thomas Fevens 63

The OSU/Facebook Realizer for SRST 2019: Seq2Seq Inflection and Serialized Tree2Tree Linearization
Kartikeya Upasani, David King, Jinfeng Rao, Anusha Balakrishnan and Michael White 68

Improving Language Generation from Feature-Rich Tree-Structured Data with Relational Graph Convolutional Encoders
Xudong Hong, Ernie Chang and Vera Demberg 75

The DipInfoUniTo Realizer at SRST'19: Learning to Rank and Deep Morphology Prediction for Multilingual Surface Realization
Alessandro Mazzei and Valerio Basile 81

LORIA / Lorraine University at Multilingual Surface Realisation 2019
Anastasia Shimorina and Claire Gardent 88

Back-Translation as Strategy to Tackle the Lack of Corpus in Natural Language Generation from Semantic Representations
Marco Antonio Sobrevilla Cabezudo, Simon Mille and Thiago Pardo 94

Conference Program

Sunday, November 3, 2019

8:45–9:00 *Opening*

Invited talk

9:00–10:00 *Invited Talk by Claire Gardent*

SR'19 Overview and results

10:00–10:30 *The Second Multilingual Surface Realisation Shared Task (SR'19): Overview and Evaluation Results*
Simon Mille, Anja Belz, Bernd Bohnet, Yvette Graham and Leo Wanner

10:30–11:00 *Coffee break*

Oral Presentations

11:00–11:25 *Learning to Order Graph Elements with Application to Multilingual Surface Realization*
Wenchao Du and Alan W Black

11:25–11:50 *DepDist: Surface realization via regex and learned dependency-distance tolerance*
William Dyer

11:50–12:15 *BME-UW at SRST-2019: Surface realization with Interpreted Regular Tree Grammars*
Ádám Kovács, Evelin Ács, Judit Ács, Andras Kornai and Gábor Recski

12:15–12:40 *Realizing Universal Dependencies Structures*
Guy Lapalme

12:40–14:00 *Lunch break*

Invited talk

14:00–15:00 *Invited Talk by Maurice Benayoun (a.k.a. MoBen)*

Oral Presentation

15:00–15:30 *IMSurReal: IMS at the Surface Realization Shared Task 2019*
Xiang Yu, Agnieszka Falenska, Marina Haid, Ngoc Thang Vu and Jonas Kuhn

Poster Session (including break)

15:30–17:00 *Surface Realization Shared Task 2019 (MSR19): The Team 6 Approach*
Thiago Castro Ferreira and Emiel Krahmer

15:30–17:00 *The Concordia NLG Surface Realizer at SRST 2019*
Farhood Farahnak, Laya Rafiee, Leila Kosseim and Thomas Fevens

15:30–17:00 *The OSU/Facebook Realizer for SRST 2019: Seq2Seq Inflection and Serialized Tree2Tree Linearization*
Kartikeya Upasani, David King, Jinfeng Rao, Anusha Balakrishnan and Michael White

15:30–17:00 *Improving Language Generation from Feature-Rich Tree-Structured Data with Relational Graph Convolutional Encoders*
Xudong Hong, Ernie Chang and Vera Demberg

Sunday, November 3, 2019 (continued)

Poster Session (continued)

15:30–17:00 *The DipInfoUniTo Realizer at SRST'19: Learning to Rank and Deep Morphology Prediction for Multilingual Surface Realization*
Alessandro Mazzei and Valerio Basile

15:30–17:00 *LORIA / Lorraine University at Multilingual Surface Realisation 2019*
Anastasia Shimorina and Claire Gardent

15:30–17:00 *Back-Translation as Strategy to Tackle the Lack of Corpus in Natural Language Generation from Semantic Representations*
Marco Antonio Sobrevilla Cabezudo, Simon Mille and Thiago Pardo

17:00–18:00 *Round table*