

INLG 2012

Proceedings of the Seventh International Natural Language Generation Conference

Program Chairs:

Barbara Di Eugenio and Susan McRoy

Generation Challenge Chairs:

Albert Gatt, Anja Belz, Alexander Koller, and Kristina Striegnitz

30 May 2012 – 1 June 2012
Starved Rock State Park
Utica, IL USA

Endorsed by the ACL Special Interest Group on Natural Language Generation (SIGGEN)

©2012 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 978-1-937284-23-7

Introduction

Welcome to the Seventh International Natural Language Generation Conference (INLG 2012). INLG 2012 is the biennial meeting of the ACL Special Interest Group on Natural Language Generation (SIGGEN). The INLG conference provides the premier forum for the discussion, dissemination, and archiving of research and results in the field of Natural Language Generation. Previous INLG conferences have been held in Ireland, the USA, Australia, the UK and Israel. Prior to 2000, INLG meetings were held as international workshops with a history stretching back to 1983. In 2012, INLG is being co-hosted by the University of Illinois-Chicago and the University of Wisconsin-Milwaukee; and held at Starved Rock State Park in Utica, IL, USA.

The INLG 2012 program consists of presentations of substantial, original, and previously unpublished results on all topics related to natural language generation. This year we received 27 submissions (12 full papers, 9 short papers and 6 demo proposals) from 10 different countries from around the world. As in previous years, each submission was reviewed by at least three members of an international program committee of leading researchers in the field. Based on these reviews 8 submissions were accepted as full papers, 10 as short papers, and 3 demos (1 paper was withdrawn). The accepted papers are of the highest quality and cover all of the major aspects of natural language generation.

This year, the conference program includes two keynote speakers. Kathleen McCoy, Professor of Computer and Information Sciences at the University of Delaware, will speak on *Natural Language Generation and Assistive Technologies*. James Lester, Professor of Computer Science at North Carolina State University, will speak on *Expressive NLG for Next-Generation Learning Environments: Language, Affect, and Narrative*. This year we are also delighted to host the 2012 Generation Challenges organized by Anja Belz, Albert Gatt, Alexander Koller, and Kristina Striegnitz. This is a part of INLG that has been growing in importance over the last several conferences and is a great addition to the event.

The organizing committee would like to offer their thanks to our invited speakers for agreeing to join us, to the program committee for their dedicated work, and, most of all, to the authors of all submitted papers. A conference like INLG would not happen without much help from many quarters: the organizers of the last two INLG conferences for sharing their wisdom, specifically, Mike White (OSU) for INLG 2008 and Ielka van der Sluis (Groningen) for INLG 2010; the SIGGEN board for allowing us to host the conference and for their assistance; Lin Chen (UIC), the INLG 2012 webmaster; Priscilla Rasmussen at ACL for her enormous help, always extremely prompt and offered with cheerfulness; Rich Gerber and Paolo Gai from SoftConf for setting up and supporting the START submission site for INLG 2012; Lorie Miller from AgoraNet for her assistance with registration; Margie VandeWyngaerde from Starved Rock Lodge and Conference Center for her help at every step of the way. We have also received sponsorship from CogenTex, for which we are extremely grateful. Finally, we would like to welcome you to Starved Rock and hope that you have an enjoyable and inspiring visit.

Barbara Di Eugenio and Susan McRoy
INLG 2012 Program Co-Chairs

Organizers:

Barbara Di Eugenio, University of Illinois at Chicago
Susan McRoy, University of Wisconsin–Milwaukee

Program Committee:

John Bateman, University of Bremen, Germany
Anja Belz, University of Brighton, UK
Stephan Busemann, DFKI GmbH, Germany
Giuseppe Carenini, The University of British Columbia, Canada
Nathalie Colineau, CSIRO, Australia
Robert Dale, Macquarie University, Australia
Reva Freedman, Northern Illinois University, US
Claire Gardent, CNRS/LORIA, France
Albert Gatt, University of Malta, Malta
Nancy Green, University of North Carolina at Greensboro, US
Helmut Horacek, Saarland University, Germany
Alistair Knott, University of Otago, New Zealand
Oliver Lemon, Heriot Watt University, Edinburgh, UK
Chris Mellish, The University of Aberdeen, UK
Margaret Mitchell, The University of Aberdeen, UK
Elena Not, Fondazione Bruno Kessler, Italy
Jon Oberlander, The University of Edinburgh, UK
Cecile Paris, CSIRO ICT Centre, Australia
Paul Piwek, the Open University, UK
Rashmi Prasad, University of Wisconsin—Milwaukee, US
Ehud Reiter, University of Aberdeen, UK
Graeme Ritchie, University of Aberdeen, UK
Donia Scott, University of Sussex Falmer, UK
James Shaw, Thomson Legal and Regulatory, US
Manfred Stede, Universitaet Potsdam, Germany
Amanda Stent, AT& T Labs, US
Matthew Stone, Rutgers, US
Kristina Striegnitz, Union College, US
Michael Strube, EML Research, Germany
Mariet Theune, University of Twente, The Netherlands
Takenobu Tokunaga, Tokyo Institute of Technology, Japan
Ielka van der sluis, University of Groningen, Netherlands
Keith vanderLinden, Calvin College, US
Sebastian Varges, University of Potsdam, Germany
Leo Wanner, Universitat Pompeu Fabra, Spain
Michael White, The Ohio State University, US
Tie-jun Zhao, Harbin Institute of Technology, China

Invited Speakers:

Kathleen McCoy, University of Delaware
James Lester, North Carolina State University

Table of Contents

| | |
|--|----|
| <i>Natural Language Generation and Assistive Technologies</i> | |
| Kathleen McCoy | 1 |
| <i>Expressive NLG for Next-Generation Learning Environments: Language, Affect, and Narrative</i> | |
| James Lester | 2 |
| <i>Learning Preferences for Referring Expression Generation: Effects of Domain, Language and Algorithm</i> | |
| Koolen Ruud, Krahmer Emiel and Theune Mariët | 3 |
| <i>Referring in Installments: A Corpus Study of Spoken Object References in an Interactive Virtual Environment</i> | |
| Kristina Striegnitz, Hendrik Buschmeier and Stefan Kopp | 12 |
| <i>MinkApp: Generating Spatio-temporal Summaries for Nature Conservation Volunteers</i> | |
| Nava Tintarev, Yolanda Melero, Somayajulu Sripada, Elizabeth Tait, Rene Van Der Wal and Chris Mellish | 17 |
| <i>Towards a Surface Realization-Oriented Corpus Annotation</i> | |
| Leo Wanner, Simon Mille and Bernd Bohnet | 22 |
| <i>Generation for Grammar Engineering</i> | |
| Claire Gardent and German Kruszewski | 31 |
| <i>Perceptions of Alignment and Personality in Generated Dialogue</i> | |
| Alastair Gill, Carsten Brockmann and Jon Oberlander | 40 |
| <i>Optimising Incremental Generation for Spoken Dialogue Systems: Reducing the Need for Fillers</i> | |
| Nina Dethlefs, Helen Hastie, Verena Rieser and Oliver Lemon | 49 |
| <i>Linguist’s Assistant: A Multi-Lingual Natural Language Generator based on Linguistic Universals, Typologies, and Primitives</i> | |
| Tod Allman, Stephen Beale and Richard Denton | 59 |
| <i>“Hidden semantics”: what can we learn from the names in an ontology?</i> | |
| Allan Third | 67 |
| <i>On generating coherent multilingual descriptions of museum objects from Semantic Web ontologies</i> | |
| Dana Dannélls | 76 |
| <i>Extractive email thread summarization: Can we do better than He Said She Said?</i> | |
| Pablo Duboue | 85 |
| <i>Rich Morphology Generation Using Statistical Machine Translation</i> | |
| Ahmed El Kholi and Nizar Habash | 90 |

| | |
|--|-----|
| <i>Reformulating student contributions in tutorial dialogue</i> | |
| Pamela Jordan, Sandra Katz, Patricia Albacete, Michael Ford and Christine Wilson | 95 |
| <i>Working with Clinicians to Improve a Patient-Information NLG System</i> | |
| Saad Mahamood and Ehud Reiter | 100 |
| <i>Sign Language Generation with Expert Systems and CCG</i> | |
| Alessandro Mazzei | 105 |
| <i>Planning Accessible Explanations for Entailments in OWL Ontologies</i> | |
| Tu Anh T. Nguyen, Richard Power, Paul Piwek and Sandra Williams | 110 |
| <i>Interactive Natural Language Query Construction for Report Generation</i> | |
| Fred Popowich, Milan Mosny and David Lindberg | 115 |
| <i>Blogging birds: Generating narratives about reintroduced species to promote public engagement</i> | |
| Advaith Siddharthan, Matthew Green, Kees van Deemter, Chris Mellish and Rene van der Wal | 120 |
| <i>Natural Language Generation for a Smart Biology Textbook</i> | |
| Eva Banik, Eric Kow, Nikhil Dinesh, Vinay Chaudri and Umangi Oza | 125 |
| <i>Generating Natural Language Summaries for Multimedia</i> | |
| Duo Ding, Florian Metze, Shourabh Rawat, Peter Schulam and Susanne Burger | 128 |
| <i>Midge: Generating Descriptions of Images</i> | |
| Margaret Mitchell, Xufeng Han and Jeff Hayes | 131 |
| <i>Preface to Papers for GenChal</i> | |
| Generation Challenge | 134 |
| <i>The Surface Realisation Task: Recent Developments and Future Plans</i> | |
| Anja Belz, Bernd Bohnet, Simon Mille, Leo Wanner and Michael White | 136 |
| <i>KBGen – Text Generation from Knowledge Bases as a New Shared Task</i> | |
| Eva Banik, Claire Gardent, Donia Scott, Nikhil Dinesh and Fennie Liang | 141 |
| <i>Content Selection From Semantic Web Data</i> | |
| Nadjet Bouayad-Agha, Gerard Casamayor, Leo Wanner and Chris Mellish | 146 |
| <i>Shared Task Proposal: Syntactic Paraphrase Ranking</i> | |
| Michael White | 150 |

Conference Program

Invited Speakers

Invited Talk I, Wednesday, May 30

1:45-2:45 *Natural Language Generation and Assistive Technologies*
Kathleen McCoy

Invited Talk II, Friday, June 1

9:00-10:00 *Expressive NLG for Next-Generation Learning Environments: Language, Affect, and Narrative*
James Lester

General Program

Lunch, Wednesday, May 30, (12:00-1:30pm)

Opening (1:30-1:45)

Invited Talk I (1:45-2:45)

Break (2:45-3:00)

Presentations, Wednesday, May 30

3:00-3:30 *Learning Preferences for Referring Expression Generation: Effects of Domain, Language and Algorithm*
Koolen Ruud, Krahmer Emiel and Theune Mariët

3:30-4:00 *Referring in Installments: A Corpus Study of Spoken Object References in an Interactive Virtual Environment*
Kristina Striegnitz, Hendrik Buschmeier and Stefan Kopp

4:00-4:30 *MinkApp: Generating Spatio-temporal Summaries for Nature Conservation Volunteers*
Nava Tintarev, Yolanda Melero, Somayajulu Sripada, Elizabeth Tait, Rene Van Der Wal and Chris Mellish

No Day Set (continued)

Dinner, Wednesday, May 30 (7:00pm)

Breakfast, Thursday, May 31 (8:30-9:30am)

Presentations, Thursday, May 31

9:30-10:00 *Towards a Surface Realization-Oriented Corpus Annotation*
Leo Wanner, Simon Mille and Bernd Bohnet

10:00-10:30 *Generation for Grammar Engineering*
Claire Gardent and German Kruszewski

Break (10:30-10:45)

Panel: Advances in Natural Language Generation (10:45-12:00)

Lunch (12:00-1:30pm)

Generation Challenge Session, Thursday, May 31 (1:30-3:00pm)

Break (3:00-3:30pm)

Presentations, Thursday, May 31

3:30-4:00 *Perceptions of Alignment and Personality in Generated Dialogue*
Alastair Gill, Carsten Brockmann and Jon Oberlander

4:00-4:30 *Optimising Incremental Generation for Spoken Dialogue Systems: Reducing the Need for Fillers*
Nina Dethlefs, Helen Hastie, Verena Rieser and Oliver Lemon

No Day Set (continued)

Poster Session, Thursday, May 31 (4:30-6:30pm)

Dinner, Thursday, May 31 (7:30pm)

Breakfast, Friday June 1(8:00-9:00am)

Invited Talk II (9:00-10:00am)

Break (10:00-10:20am)

Presentations, Friday, June 1

10:20-10:50 *Linguist's Assistant: A Multi-Lingual Natural Language Generator based on Linguistic Universals, Typologies, and Primitives*
Tod Allman, Stephen Beale and Richard Denton

10:50-11:20 *"Hidden semantics": what can we learn from the names in an ontology?*
Allan Third

11:20-11:50 *On generating coherent multilingual descriptions of museum objects from Semantic Web ontologies*
Dana Dannélls

Closing, Friday, June 1 (11:50-12:00)

Section for Poster and Demo Session

Posters, Thursday, May 31 (4:30-6:30)

Extractive email thread summarization: Can we do better than He Said She Said?
Pablo Duboue

Rich Morphology Generation Using Statistical Machine Translation
Ahmed El Kholly and Nizar Habash

Reformulating student contributions in tutorial dialogue
Pamela Jordan, Sandra Katz, Patricia Albacete, Michael Ford and Christine Wilson

No Day Set (continued)

Working with Clinicians to Improve a Patient-Information NLG System
Saad Mahamood and Ehud Reiter

Sign Language Generation with Expert Systems and CCG
Alessandro Mazzei

Planning Accessible Explanations for Entailments in OWL Ontologies
Tu Anh T. Nguyen, Richard Power, Paul Piwek and Sandra Williams

Interactive Natural Language Query Construction for Report Generation
Fred Popowich, Milan Mosny and David Lindberg

Blogging birds: Generating narratives about reintroduced species to promote public engagement
Advaith Siddharthan, Matthew Green, Kees van Deemter, Chris Mellish and Rene van der Wal

Demonstrations, Thursday, May 31 (4:30-6:30)

Natural Language Generation for a Smart Biology Textbook
Eva Banik, Eric Kow, Nikhil Dinesh, Vinay Chaudri and Umangi Oza

Generating Natural Language Summaries for Multimedia
Duo Ding, Florian Metze, Shourabh Rawat, Peter Schulam and Susanne Burger

Midge: Generating Descriptions of Images
Margaret Mitchell, Xufeng Han and Jeff Hayes

No Day Set (continued)

Section for Generation Challenge Session

GenChal Session, Thursday, May 31 (1:30-3:00)

- 1:30-1:40 *Preface to Papers for GenChal*
Generation Challenge
- 1:40-2:15 *The Surface Realisation Task: Recent Developments and Future Plans*
Anja Belz, Bernd Bohnet, Simon Mille, Leo Wanner and Michael White
- 2:15-2:30 *KBGen – Text Generation from Knowledge Bases as a New Shared Task*
Eva Banik, Claire Gardent, Donia Scott, Nikhil Dinesh and Fennie Liang
- 2:30-2:45 *Content Selection From Semantic Web Data*
Nadjet Bouayad-Agha, Gerard Casamayor, Leo Wanner and Chris Mellish
- 2:45-3:00 *Shared Task Proposal: Syntactic Paraphrase Ranking*
Michael White

