

Final Report: Summer School on Natural Language Generation, Summarisation , and Dialogue

Roman Kutlak, Artemis Parvizi, Kees van Deemter

August 5, 2015

1 Description

The objective of this summer school was to introduce participants to the concepts and research questions in natural language generation (NLG), summarisation and dialogue systems. Although these three areas produce natural language, their distinct communities seldom interact because each community relies on different methods and because the inputs to each kind of system are different. There is, however, considerable overlap in the kinds of problems that need to be considered, from selecting the right content to evaluating the systems. We believe that focusing on the similarities of the different areas can stimulate “cross-pollination” of research. For example, most summarisation techniques could benefit from deeper semantic processing as performed in dialogue systems. Similarly, many NLG systems could benefit from techniques used by dialogue systems to substantially improve the generated output.

2 Attendees

Forty three participants (excluding the lecturers and organisers) from 15 countries attended the summer school. Most participants were PhD students, about ten participants were post-doctoral researchers, and three participants were from industry. Participants had varied background including NLG, summarisation, human-computer interaction, dialogue systems, linguistics and machine learning. The following participants were from SICSA institutions:

1. **Rodrigo de Oliveira**, University of Aberdeen
2. **Stephanie Inglis**, University of Aberdeen
3. **Xiao Li**, University of Aberdeen
4. **Blessing Mbipom**, Robert Gordon University
5. **Milan Markovic**, University of Aberdeen
6. **Yanchao Yu**, Heriot-Watt University

3 Summer School's Outcomes and Impact

NLG, summarisation and dialogue systems have wide applicability in data summarisation, human-computer interaction and e-health care. The main outcomes of the summer school were to stimulate the NLG community and to link it more strongly with other research fields.

The website <http://nlgsummer.github.io> also provides slides from the lectures, and recordings will be added shortly. We hope this resource will boost research in NLG and related disciplines for years to come. Several projects also started to be discussed during the summer school; examples include NLG for Tibetan, and a project combining NLG and research on readability. The University of Aberdeen is one of the world's most active centres of NLG research and this summer school was a significant contribution to the wider NLG community. Given the number of participants and their diversity, we believe that we made a good progress towards these goals.

4 Lecturers

Gerard Casamayor (Universitat Pompeu Fabra)

Robert Dale (ARRIA NLG)

Thomas François (Université catholique de Louvain)

Albert Gatt (University of Malta)

Yannis Konstas (University of Edinburgh)

Paul Piwek (Open University)

Ehud Reiter (University of Aberdeen, ARRIA NLG)

Advaith Siddharthan (University of Aberdeen)

Yaji Sripada (University of Aberdeen, ARRIA NLG)

Kees van Deemter (University of Aberdeen)

5 Summary of the Lectures and Lab Sessions

Course Summary

1. Introduction to NLG: Ehud Reiter (*University of Aberdeen, ARRIA NLG*)

This course introduces the basic concepts of NLG, including document planning, microplanning, and realisation. It will also include examples of real NLG systems, including both what they do and how they work. It will also briefly discuss system-building issues such as knowledge acquisition and evaluation. The course will conclude with a quick summary of commercial NLG activity.

2. NLG in detail

- **Content Determination:** *Gerard Casamayor (Universitat Pompeu Fabra)*
This talk summarises existing approaches to content determination for NLG, as well as touches on the closely related topics of text ordering and structuring. It discusses why content determination is hard and what sorts of (hand-crafted or learned) models can be used to inform it. It is often useful to regard content determination as a search problem, and we take this approach in order to compare the different methods that have been used.
- **Micro-planning:** *Albert Gatt (University of Malta)*
This session is devoted to discussing different micro-planning tasks. In particular, we will review some of the “classic” sub-tasks that microplanners have often been designed to perform, notably: (a) lexicalisation, the task of choosing the right words or lemmas to express the contents of the message; (b) aggregation, the task of merging distinct representations into a single, more concise representation; (c) referring expression generation, the task of selecting the content (and, to some extent, the form) of referential noun phrases in text.
- **Surface Realisation:** *Albert Gatt (University of Malta)*
In this session, we discuss the domain of realisation with reference to a number of different languages. The purpose of this is mainly to delineate the problem: different languages will make different demands on a realiser, and in some cases, syntactic choices will have repercussions for microplanning decisions. Next, we look at an overview of different realisers, starting with some classic rule-based examples, such as RealPro and KPML (Bateman, 1997). These are contrasted with recent approaches, such as HALOGEN (Langkilde-Geary and Knight, 2002) or OpenCCG (White et al, 2007), where the aim is to minimise the rule-based component while allowing syntactic choices to be made probabilistically, usually through the use of n-gram based language models.

3. Learning to generate: Concept-to-text generation using machine learning: *Yannis Konstas (University of Edinburgh)*

In this tutorial, we explore systems in NLG that learn the well-known pipeline modules of content selection, microplanning and surface realisation, automatically from data. We visit methods that model each step into a probabilistic model or other weighted function, and learn their parameters by optimising a text output-related objective metric (e.g., BLEU, METEOR scores). Generation is then viewed as a common search problem, which entails finding the best combination of parameters given the trained model and an input. We also compare systems that optimise each module in isolation, as well as jointly.

4. Evaluation: *Ehud Reiter (University of Aberdeen, ARRIA NLG)*

This course discusses different ways to evaluate NLG systems, including task-based, human-based, and metric-based evaluation. We also summarise practice in each of these types of evaluations, and present the design and outcomes of several real NLG evaluations.

5. Introduction to Summarisation: *Advaith Siddharthan (University of Aberdeen)*

This tutorial covers a range of text summarisation approaches described in the literature for extractive and abstractive summarisation, including models for sentence selection in extractive summarisation based on various statistical definitions of “topic”, abstractive summarisation through aggregation and deletion, attempts at microplanning (e.g., generating referring expressions), the use of template based generation, and issues of text planning or sentence ordering.

6. Introduction to Dialogue Systems: *Paul Piwek (Open University)*

The aim of this tutorial is to provide an overview of the key ideas at the heart of most dialogue systems and to describe some active research topics in this area. In order to get a systematic understanding of a range of dialogue systems, we use the notion of a dialogue game. Several systems are described in terms of the dialogue game that they play. This allows us to compare and contrast these systems and distinguish four broad categories: reactive, fixed-task, selected-task and joint-task dialogue systems. The tutorial includes a discussion of dialogue system evaluation and the use of machine learning to improve dialogue system performance. We conclude by considering two current research topics: non-cooperation and incrementality.

7. Readability: *Thomas François (Université catholique de Louvain)*

In this readability module, we first outline the main tendencies in the field, with a focus on recent work that applies NLP techniques to readability. We also describe the usual methodology framework applied to design a readability model and discuss some of the choices available within this framework. A summary of the evaluation techniques used in the field concludes the first part of the module.

In the second part, we discuss some of the main challenges and issues in the field as we see them, such as the challenge of collecting large datasets of difficulty-annotated texts to train modern statistical algorithms, the issue of cross-domain generalisation, or the adaptation of readability methods to lower levels (sentence or word). This module is concluded with some perspectives for future research in the field.

8. Cognitive Modelling: the case of reference *Kees van Deemter (University of Aberdeen)*

In this course we look at natural language generation from a completely different perspective: computational cognitive modelling. We give an overview of the aims of Computational Cognitive Modelling (CCM) and the types of CCMs that have been developed so far. The main topic of lecture talk is production of referring expressions as viewed by computational linguists as well as psycholinguists and how computational models can tell us something about human reference production.

Finally, we return to NLG as a whole, asking whether referring expression generation is somehow special. Tentatively, we argue that all of NLG can be seen as an exercise in Computational Cognitive Modelling, though a more utilitarian view (which emphasises the generation of useful referring expressions) is valid as well.

9. The New Science of Information Delivery: *Robert Dale (ARRIA NLG)*

Natural Language Generation is a means to an end. That end is the delivery of information, and the great thing about NLG is that it provides a way of automating the delivery of the right information to the right people at the right time in the right way. But to really do that well, we need to understand the task of information delivery, and we need to understand how a variety of scientific disciplines provide underpinnings for how we can do that task well. This talk aims to stand back and look at the bigger picture: what it would mean to have a science of information delivery, with NLG as a key player along with a number of other technologies.

Open Lab

Yaji Sripada (University of Aberdeen, ARRIA NLG)

The Open Lab sessions are an opportunity to test your learning at the summer school. You will build an NLG system end-to-end using the software modules we provide. You can work alone or as a team. The idea is not to push you into writing code round-the-clock to build loads of functionality. Instead you will be encouraged to focus on learning first-hand the design choices and trade-offs involved in building an NLG application. If you are new to computer programming but willing to learn, help will be available in these supervised lab sessions to get you started with your first program!

Poster/Demo Session

At this session the participants can showcase the applications that they have developed in the open lab sessions and receive feedback from the lecturers and other participants. Also, there is an opportunity for all the participants to bring along a poster of the work they are currently involved in. It will be a good networking event for all of us.

Evening lecture

Paul Piwek (Open University) Towards machines that mean what they say

In this talk, I will examine what it involves for a speaker to mean what they say. Traditional well-established accounts of meaning focus on (1) the speaker grasping the truth conditions of what they say and (2) having the right intentions. I will discuss some limitations of this view and then proceed to outline an alternative approach which is inspired by the work of the American philosopher Robert Brandom. The proposal shifts the focus from the individual speaker and their intentions to the language games that speakers engage in with other language users. I will describe first steps towards formalising such theory of meaning and the prospects of using this work to develop machines that mean what they say.

6 Schedule

Day \ Time	Monday	Tuesday	Wednesday	Thursday	Friday
9-10:30	Intro to NLG Ehud Reiter	Machine Learning Yannis Konstas	Intro to Summarisation Advaith Siddharthan	System Evaluation Ehud Reiter	Cognitive Modelling Kees van Daeleer
10:30-10:45	Break	Break	Break	Break	Break
10:45-12:15	NLG in Detail (Content Determination) Chris Mellish	Intro to Dialogue Systems Paul Piwek	Intro to Summarisation Advaith Siddharthan	Readability Thomas François	The New Science of Information Delivery Robert Dale
12:15-13:00	Lunch	Lunch	Lunch	Lunch	Lunch
13:00-14:30	NLG in Detail (Micro Planning) Albert Gatt	Intro to Dialogue Systems Paul Piwek	NLG in Practice Yaji Srivada	Readability Thomas François	Poster/Demo Session
14:30-14:45	Break	Break	Break	Break	Break
14:45-16:15	NLG in Detail (Surface Realisation) Albert Gatt	Open Lab Yaji Srivada	Open Lab Yaji Srivada	Open Lab Yaji Srivada	Discussion
16:15-16:30	Break				
16:30-18:00	Open Lab Yaji Srivada				
18:00-18:30	Social Event	Break	Break		The end
18:30-19:30		Evening Lecture Paul Piwek	Evening Lecture Graham Ritchie	Evening off	