

# Advances in Argument Mining

**Chris Reed**

Centre for Argument Technology  
University of Dundee, UK  
c.a.reed@dundee.ac.uk

**Katarzyna Budzynska**

Institute for Philosophy & Sociology  
Polish Academy of Sciences  
Warsaw, Poland  
budzynska.argdiap@gmail.com

## 1 Description

*This course aims to introduce students to an exciting and dynamic area that has witnessed remarkable growth over the past 36 months. Argument mining builds on opinion mining, sentiment analysis and related to tasks to automatically extract not just what people think, but why they hold the opinions they do. From being largely beyond the state of the art barely five years ago, there are now many hundreds of papers on the topic, millions of dollars of commercial and research investment, and the 6th ACL workshop on the topic will be in Florence in 2019. The tutors have delivered tutorials on argument mining at ACL 2016, at IJCAI 2016 and at ESSLLI 2017; for ACL 2019, we have developed a tutorial that provides a synthesis of the major advances in the area over the past three years.*

Argument and debate form cornerstones of civilised society and of intellectual life. Processes of argumentation run our governments, structure scientific endeavour and frame religious belief. Recognising and understanding argument are central to decision-making and professional activity in all walks of life, which is why we place them at the centre of academic pedagogy and practice; it's why such a premium is placed upon these skills; and it's why rationality is one of the very defining notions of what it is to be human.

As our understanding of how arguments are assembled, are interpreted and have impact has improved, so it has become possible to frame computational questions about how it might be possible for machines to model and replicate the processes involved in identifying, reconstructing, interpreting and evaluating reasoning expressed in natural language arguments. This, then, is argument mining: identifying that an argument is present, de-

composing an argument into its constituent parts, determining how those parts are connected and structured – and how they connect with other arguments and argument parts – and finally, evaluating the quality of those connections. Thorough overviews are provided in, e.g., (Stede and Schneider, 2018; Lippi and Torroni, 2016) with a wide range of more detailed themes covered elsewhere, such as premise-conclusion recovery (Stab and Gurevych, 2017), types of argumentation pattern (Walton et al., 2008), relationships between semantic and argumentative structures (Becker et al., 2017), how ethos of speakers interacts with argument structure (Duthie and Budzynska, 2018) and automated assessment of argument persuasiveness (Carlile et al., 2018).

**Growth.** From just a handful of papers in print around 2010, argument mining has grown rapidly with Google Scholar now reporting around 2,000 articles mentioning the topic in their title. ACL, EMNLP and NAACL have included over 50 articles on argument mining in the past three years alone, in addition to the 92 articles published at the ACL workshop series on Argument Mining (co-founded by Reed). Argument mining has been building momentum both within the ACL community and further afield, with both specialist conferences (such as Computational Models of Argument, COMMA, co-chaired in 2018 by Budzynska) and generalist conferences (such as IJCAI) devoting increasing time to papers, workshops and graduate-level training on argument mining. The creation and publication of datasets has been an important contributor to the vitality of the field, with papers at LREC and in LRE increasing the breadth of this foundational aspect to the field (Abbott et al., 2016). By the same token, new SEMEVAL tasks have also started to set the goalposts and shape robust comparative evalua-

tions (<https://competitions.codalab.org/competitions/17327>).

**Challenge.** Argument mining is a particularly challenging task and is an exciting domain in which to work because it is increasingly clear that deep learning and distributional techniques alone are not delivering the same kind of successes that have been enjoyed in some other areas of NLP. Many labs working with algorithms for argument mining are finding that hybrid approaches that integrate rule-based and statistical methods are more likely to deliver the strongest performance. As a result, recent argument mining has been pushing the boundaries of approaches to NLP in general.

**Argument mining in the press.** The past year has seen a rapidly accelerating public profile for argument technologies in general, with Reed commissioned to produce articles that have appeared in Newsweek ([arg.tech/newsweek](http://arg.tech/newsweek)) and on the BBC ([arg.tech/bbcnews](http://arg.tech/bbcnews)), and media events such as IBM’s Project Debater launch (e.g., [www.wired.com/story/now-the-computer-can-argue-with-you](http://www.wired.com/story/now-the-computer-can-argue-with-you)). The BBC too has commissioned technology that includes the first live deployment of argument mining ([www.bbc.co.uk/taster/pilots/evidence-toolkit-moral-maze](http://www.bbc.co.uk/taster/pilots/evidence-toolkit-moral-maze)) in supporting identification of fake news. The tutorial will make use of these high-profile applications of argument mining to contextualise and motivate the topics covered.

## 2 Type

As argument mining has been covered at an ACL tutorial previously, in 2016. This tutorial is classified as ‘Introductory’ and the syllabus is designed to minimise prerequisites. We aim, however, to focus heavily on results from the past three years during which time significant progress has been made.

## 3 Outline

The tutorial is structured in two parts, each of which mixes lecturing with practical work. In the first part, we will cover theory of argument structure from (i) the basics in argumentation theory; through (ii) recent results in computational models of argument; to (iii) the latest (as yet largely unpublished) techniques that allow modelling of

dialogical argumentation. To consolidate understanding of the material in each of these 20 minute blocks, the first part concludes with a 30 minute practical session in which students will get an opportunity to apply the theory to an example drawn from a real-life setting.

In the second part, we will cover techniques for argument mining from (i) straightforward application of machine learning techniques; through (ii) use of BiLSTMs in particular for exploiting sequence structure latent in argument presentation; to (iii) the development of hybrid approaches to argument mining. We will again encourage deep understanding from the students through a short practical implementation exercise making use of R.

The outline syllabus runs thus:

### Part A: Foundations

- **A1** (20 mins). Theory of argument structure – linked, convergent, serial, divergent, rebut, undercut – indicators – enthymemes – logos, ethos, pathos.
- **A2** (20 mins). Semantic types – argumentation schemes – ADU segmentation – datasets, corpora and shared tasks.
- **A3** (20 mins). Argument in dialogue – Inference Anchoring Theory – reported speech – complex and implicit speech acts.
- **A4** (30 mins). Practical session: Analysing natural argument.

*Break*

### Part B: Applications

- **B1** (20 mins). Simple machine learning. IOB schema for segmentation – classifiers for argnonarg – classifiers for premise-conclusion.
- **B2** (20 mins). Advanced machine learning. Word embeddings for argumentation schemes – BiLSTM models for argumentation sequence patterns.
- **B3** (20 mins). Hybrid approaches. Argument structure parsing – illocutionary structure parsing – dialogical priors.
- **B4** (30 mins). Practical session: Argument mining in R.

All materials will be made available at a dedicated tutorial website as they were for our ACL2016, IJCAI 2016 and ESSLLI 2017 tutorials.

This website will be located at <http://arg.tech/acl2019tut>.

#### 4 Prerequisites

The tutorial is intended to be accessible to most ACL attendees, so has straightforward prerequisites:

- Basic familiarity with supervised machine learning techniques and the way they are employed and assessed
- Experience of using R will be an advantage, but is not required, for practical session B4.

Attendees are expected to have or to be working towards a PhD in computational linguistics or a closely cognate area, but no previous experience of academic investigation of argument is expected.

#### 5 Tutors

**Katarzyna Budzynska** (Computational Ethos Lab, Polish Academy of Sciences, [budzynska.argdiap@gmail.com](mailto:budzynska.argdiap@gmail.com), [www.computationalethos.org](http://www.computationalethos.org)).

Katarzyna is an associate professor in philosophy at the National Polish Academy of Sciences and an associate professor in computing at the University of Dundee (UK). Her work focuses on communication structures of argumentation, dialogue and ethos. She has published two books and over 80 peer-reviewed papers including articles in journals such as *Artificial Intelligence*, *Association for Computing Machinery (ACM TOIT)* and *Synthese*. Katarzyna founded a national movement, the Polish School of Argumentation, and sits in the steering committees of a new initiative, the European Conference on Argumentation ([ecargument.org](http://ecargument.org)), and the ArgDiaP Association for Argumentation, Dialogue and Persuasion ([argdiap.pl](http://argdiap.pl)). Most recently, she established her research group – the Computational Ethos Lab which develops innovative technologies to process the use of ethos in natural language in order to predict the results of presidential elections, detect trolls and cyber-bullies in social media, and uncover potential terrorist threats. With Villata, she delivered

a tutorial on Argument Mining at IJCAI 2016, and with Reed an extensive week-long course at the 29th European Summer School in Logic, Language, and Information (ESSLLI2017).

**Chris Reed** (Centre for Argument Technology, University of Dundee, [c.a.reed@dundee.ac.uk](mailto:c.a.reed@dundee.ac.uk), [www.arg.tech](http://www.arg.tech)). Chris is Full Professor of Computer Science and Philosophy at the University of Dundee, where he heads the Centre for Argument Technology. Chris has been working at the overlap between argumentation theory and artificial intelligence for over twenty years, has won over £6m of funding from government and commercial sources and has over 200 peer-reviewed papers in the area (including papers in ACL, COLING, IJCAI, ECAI and AAAI) and five books. He has also been instrumental in the development of the Argument Interchange Format, an international standard for computational work in the area; he is spear-heading the major engineering effort behind the Argument Web; and he was a founding editor of the Journal of Argument & Computation. He was co-organiser of COMMA 2014, of the first ACL workshop on Argumentation Mining in 2014, was the chair of the third workshop on Argument Mining with ACL in 2016, and has recently won funding for a £1m project on the topic in collaboration with IBM. With Gurevych, Stein and Slonim, he delivered a tutorial on Argument Mining at ACL 2016 which was extremely well attended, and followed that with a course at ESSLLI 2017 with Budzynska.

#### Acknowledgments

Our work in developing this tutorial has been supported in part in the UK by EPSRC under grant EP/N014781/1.

#### References

- Rob Abbott, Brian Ecker, Pranav Anand, and Marilyn A. Walker. 2016. Internet Argument Corpus 2.0: An SQL schema for dialogic social media and the corpora to go with it. In *Language Resources and Evaluation Conference (LREC)*, Portorož, Slovenia.
- Maria Becker, Alexis Palmer, and Anette Frank. 2017. Semantic clause types and modality as features for argument analysis. *Journal of Argument and Computation*, 8(2):95–112.
- Winston Carlile, Nishant Gurrapadi, Zixuan Ke, and Vincent Ng. 2018. Give me more feedback: Annotating argument persuasiveness and related attributes

in student essays. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 621–631. Association for Computational Linguistics.

Rory Duthie and Katarzyna Budzynska. 2018. A deep modular rnn approach for ethos mining. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, pages 4041–4047.

Marco Lippi and Paolo Torrioni. 2016. Argumentation mining: State of the art and emerging trends. *ACM Trans. Internet Technology*, 16(2):10:1–10:25.

Christian Stab and Iryna Gurevych. 2017. Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3):619–659.

Manfred Stede and Jodi Schneider. 2018. *Argumentation Mining*. Synthesis Lectures on Human Language Technologies. Morgan Claypool.

Douglas Walton, Chris Reed, and Fabrizio Macagno. 2008. *Argumentation Schemes*. Cambridge University Press.