Andreas Vlachos

CONTACT INFORMATION

One Euston Square
40 Melton Street
London, NW1 2FD, UK

WWW: http://www.cl.cam.ac.uk/~av308
Office: +442035495395
E-mail: a.vlachos@cs.ucl.ac.uk

RESEARCH INTERESTS

Structured prediction, Clustering, Natural language understanding, Active learning.

RESEARCH EXPERIENCE

Research Associate at the University College London, Department of Computer Science
- 01/2014 - now: Automatic fact-checking, in collaboration with the BBC R&D team

Research Associate at the University of Cambridge, Computer Laboratory
- 06/2011 - 12/2013: Semantic Parsing in the EU FP7 SpaceBook project:
  http://www.spacebook-project.eu/

Research Associate at the University of Wisconsin-Madison, Department of Biostatistics
- 02/2010 - 06/2011: Structured Prediction for Biomedical Information Extraction

Research Assistant at the University of Cambridge, Computer Laboratory
- 01/2008 - 12/2009: Developing Multilingual Technologies for Automatic Lexical Acquisition:
  http://www.britishcouncil.org/france-science-alliance-programme.htm
- 11/2004 - 9/2006: Integrating Literature, Experiments and Curation in Drosophila Genomics
  Research: http://www.wiki.cl.cam.ac.uk/rowiki/NaturalLanguage/FlySlip

Research Assistant at the Athens University of Economics and Business
- 2002 - 2003: Design and implementation of a trainable system for end of sentence detection in
  Greek and English texts.
- 2001: Participation in the implementation of the financial terminology dictionary DelosLexicon.

EDUCATION

University of Cambridge, England, UK

PhD in Computer Science, October 2006 - December 2009

- Supervisor: Ted Briscoe
- Title: Semi-supervised learning for biomedical information extraction.

University of Edinburgh, Scotland, UK

MSc in Informatics, October 2003 - September 2004

- Supervisor: Miles Osborne
- Dissertation Topic: Active Learning with Support Vector Machines.

Athens University of Economics and Business, Greece

BSc in Informatics, 1998-2003

TEACHING EXPERIENCE

- Guest lecturer on biomedical text mining (UCam, UCL) and active learning (UWarwick).
- Co-supervisor of 8 MPhil (UCam, UCL) and 1 undergraduate theses (UCam).
- Supervisor for the undergraduate courses in Natural Language Processing, Information Retrieval
  and Probability Theory (UCam).
- Demonstrator for the Information Retrieval and Question Answering practicals (UCam).

OTHER ACADEMIC ACTIVITIES

- Co-organizer of the EACL 2014 workshop "Dialogue in Motion".
- Program Committee member for ACL, IJCAI, AAAI, EACL, BioNLP.

- Reviewer for EMNLP, CoNLL, IJCNLP, Computational Linguistics, Transactions of the Association for Computational Linguistics, Machine Learning, ACM Transactions in Speech and Language Processing, ACM Transactions on Information Systems, Journal of Artificial Intelligence Research, BMC Bioinformatics, Journal of Natural Language Engineering, Journal of Bioinformatics, Computer, Speech and Language.
- Organizer of the Numerical Optimization reading at UCL.
- Co-organizer of the Natural Language Processing reading group at the University of Cambridge.
- Panelist in the EMNLP 2011 Worskhop on Unsupervised Learning in NLP.

INDUSTRIAL LINKS
- Consultancy for QualPact Inc. on biomedical natural language processing applications, resulting in one successful patent.
- Consultancy for Codean Inc. on extracting information from legal documents.

SCHOLARSHIPS - GRANTS

Amazon Web Services in Education, http://aws.amazon.com/education

Large Scale Non-Parametric Semi-Supervised Part-of-Speech Tagging.

Gates Cambridge Scholarship, http://www.gates.scholarships.cam.ac.uk

Tuition fees and living expenses for the PhD at Cambridge.

Student Award Agency for Scotland, www.saas.gov.uk

Tuition fees scholarship for the MSc in Informatics at Edinburgh.

INVITED TALKS

Imitation Learning for Structured Prediction in Natural Language Processing
- Department of Computer Science, University of Wisconsin-Madison, October, 2013.
- Keynote at the 12th Language and Computation Day, University of Essex, September 2013.

The SpaceBook project: Assisting tourists in navigating and exploring a city
- Department of Computer Science, University of Warwick, November, 2013.
- The 32nd SGAI International Conference on Artificial Intelligence, British Computing Society, Cambridge, December 2012.
- Higher School of Economics, National Research University, Faculty of Philology, Moscow, December 2012.

Relation extraction using a handful of labeled examples, a web search engine and imitation learning
- Oxford Computational Linguistics Seminar, University of Oxford, March 2013.

Search-based Structured Prediction applied to Biomedical Event Extraction
- Institute for Language, Cognition and Computation seminar, University of Edinburgh, October 2012.
- Department of Statistics, Imperial College London, June 2012.
- Gatsby Computational Neuroscience Unit, University College London, March 2012.
- Artificial Intelligence and Information Systems seminar at the University of of Illinois at Urbana-Champaign, April 2011.
- Insitute of Informatics and Telecommunication Seminar at the National Center for Scientific Research, "Demokritos", Athens, January 2011.

Unsupervised and Constrained Dirichlet Process Mixture Models for Verb Clustering
- Institute for Language and Speech Processing, Athens, September 2009.
- Dublin Computational Linguistics Research Seminar at Trinity College Dublin, March, 2009.
- Laboratoire d'Informatique de Paris-Nord, November 2008.

Publications   Gubbins, J., Vlachos, A. 2013. Dependency Language Models for Sentence Completion, in *Proceedings of EMNLP (short papers)*, Seattle, WA.

Andreas, J., Vlachos, A., Clark, S. 2013. Semantic Parsing as Machine Translation, in *Proceedings of ACL (short papers)*, Sofia, Bulgaria.

Vlachos, A. 2012. An investigation of imitation learning algorithms for structured prediction, in *Journal of Machine Learning Research W&C Proceedings*, Volume 24, *Proceedings of the 10th European Workshop in Reinforcement Learning*, pages 143-154.

Vlachos, A., Craven, M. 2012. Biomedical event extraction from abstracts and full papers using search-based structured prediction, in *BMC Bioinformatics*, Volume 13, Supplement 11.

Vlachos, A., Craven, M. 2011. Search-based Structured Prediction applied to Biomedical Event Extraction, in *Proceedings of the Fifteenth Conference on Computational Natural Language Learning*, pages 49-57, Portland, OR.

Vlachos, A., Craven, M. 2011. Biomedical Event Extraction from Abstracts and Full Papers using Search-based Structured Prediction, in *Proceedings of BioNLP Shared Task 2011 Workshop*, pages 36-40, Portland, OR.

Vlachos, A. 2011. Evaluating unsupervised learning for natural language processing tasks, in *Proceedings of the First Workshop on Unsupervised Learning in NLP*, pages 35-42, Edinburgh, UK.

Vlachos, A., Ghahramani Z., Briscoe, T. 2010. Active learning for constrained Dirichlet process mixture models, in *Proceedings of the 2010 workshop on GEometrical Models of Natural Language Semantics*, pages 57-61, Uppsala, Sweden.

Vlachos, A., Craven, M. 2010. Detecting Speculative Language using Syntactic Dependencies and Logistic Regression, in *Proceedings of the Fourteenth Conference on Computational Natural Language Learning*, pages 18-25, Uppsala, Sweden.

Vlachos, A. 2010. Two strong baselines for the BioNLP 2009 event extraction task, in *Proceedings of the 2010 Workshop on Biomedical Natural Language Processing*, pages 1-9, Uppsala, Sweden.

Bratières, S., van Gael, J., Vlachos, A., Ghahramani, Z. 2010. Scaling the iHMM: Parallelization versus Hadoop, in *Proceedings of the 2010 IEEE International Conference on Computing and Information Technology*, pages 1235-1240, Bradford, UK.

Van Gael, J., Vlachos, A., Ghahramani, Z. 2009. The infinite HMM for unsupervised PoS tagging, in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 678-687, Singapore.

Vlachos, A., Korhonen, A., Ghahramani, Z. 2009. Unsupervised and Constrained Dirichlet Process Mixture Models for Verb Clustering, in *Proceedings of the Workshop on Geometrical Models of Natural Language Semantics*, pages 74-82, Athens, Greece.

Vlachos, A., Buttery, P., Ó Séaghdha, D., Briscoe, T. 2009. Biomedical Event Extraction without Training Data, in *Proceedings of the BioNLP 2009 Workshop Companion Volume for Shared Task*, pages 37-40, Boulder, CO.

Vlachos, A. 2008. A stopping criterion for active learning, in *Computer, Speech and Language*, Volume 22, Issue 3, July 2008, pages 295-312.

Karamanis, N., Seal, R., Lewin, I., McQuilton, P., Vlachos, A., Gasperin, A., Drysdale, R., Briscoe, T. 2008. Natural Language Processing in aid of FlyBase curators, in *BMC Bioinformatics*, Volume 9:193, 2008.

Vlachos, A., Ghahramani, Z., Korhonen, A. 2008. Dirichlet Process Mixture Models for Verb Clustering, in *ICML workshop on Prior Knowledge for Text and Language*, Helsinki, Finland.

Vlachos, A. 2007. Evaluating and combining biomedical named entity recognition systems, in *Proceedings of the Workshop on Biological, Translational, and Clinical Language Processing*, pages 199-200, Prague, Czech Republic.

Vlachos, A. 2007. Tackling the BioCreative2 Gene Mention task with Conditional Random Fields and syntactic parsing, in *Proceedings of the Second BioCreative Challenge Evaluation Workshop*, pages 85-87, Madrid, Spain.

Korkontzelos, I., Vlachos, A., Lewin, I. 2007. From gene names to actual genes, in *Proceedings of the Annual Meeting of the ISMB BioLINK Special Interest Group on Text Data Mining*, Vienna, Austria.

Vlachos, A. 2006. Active Annotation, in *Proceedings of the Workshop on Adaptive Text Extraction and Mining*, Trento, Italy.

Vlachos, A., Gasperin, C. 2006. Bootstrapping and Evaluating Named Entity Recognition in the Biomedical Domain, in *Proceedings of the HLT-NAACL BioNLP Workshop on Linking Natural Language and Biology*, pagse 138-145, New York, New York.

Vlachos, A., Gasperin, C., Lewin, I., Briscoe, E. J. 2006. Bootstrapping the recognition and anaphoric linking of named entities in Drosophila articles, in *Proceedings of the Pacific Symposium in Biocomputing*, Volume 11, pages 100-111, Maui, HI.

Theodorakis, M., Vlachos, A., Kalamboukis, T. Z., 2004. Using Hierarchical Clustering to Enhance Classification Accuracy, in *Proceedings of 3rd Hellenic Conference in Artificial Intelligence*, Samos, Greece.

OPEN SOURCE SOFTWARE	• Search-based Structured Prediction applied to Biomedical Event Extraction • Cost-sensitive classification with Adaptive Regularization of Weights
PROGRAMMING LANGUAGES	Python, Java, C++, MATLAB.
NATURAL LANGUAGES	Greek, English, French, Russian.

ACADEMIC
REFEREES

Stephen Clark, current post-doctoral supervisor

William Gates Building	*WWW:* http://www.cl.cam.ac.uk/~sc609
15 JJ Thomson Avenue	*Office:* +441223763704
Cambridge, CB3 0FD, UK	*E-mail:* stephen.clark@cl.cam.ac.uk

Mark Craven, former post-doctoral supervisor

6730 Medical Sciences Center	*WWW:* http://www.biostat.wisc.edu/~craven/
1300 University Avenue	*Office:* +16082656181
Madison, Wisconsin 53706, USA	*E-mail:* craven@biostat.wisc.edu

Ted Briscoe, PhD supervisor

William Gates Building
15 JJ Thomson Avenue
Cambridge, CB3 0FD, UK

WWW: `http://www.cl.cam.ac.uk/~ejb1`
Office: +441223334616
E-mail: ted.briscoe@cl.cam.ac.uk