WILKER FERREIRA AZIZ

Institute for Logic, Language and Computation
University of Amsterdam
Science Park 107, F2.11
Amsterdam, 1098 XG, Netherlands

w.aziz@uva.nl
https://wilkeraziz.github.io
https://probabll.github.io
Tel: +31 (0) 631991009

## Research

My research concerns teaching computers to learn efficiently about problems as complex as those involving understanding and manipulating human languages. To that end, I rely on strong results from probability theory, Bayesian statistics, formal languages and automata, optimisation, and linguistics. I develop machine learning models as well as techniques to estimate such models efficiently from data. Because the world only exposes partial supervision to its workings, my research has a strong component of unsupervised learning and other forms of latent variable modelling. I have a deep appreciation for Bayesian thinking and I am a firm believer in the role of explanations in science.

## Selected Publications

- *Interpretable Neural Predictions with Differentiable Binary Variables*.
  Joost Bastings, Wilker Aziz and Ivan Titov. In *ACL*, 2019.
  https://www.aclweb.org/anthology/P19-1284/
- *Block Neural Autoregressive Flow*.
  Nicola De Cao, Wilker Aziz and Ivan Titov. In *UAI*, 2019.
  http://auai.org/uai2019/proceedings/papers/511.pdf
- *Question Answering by Reasoning Across Documents with Graph Convolutional Networks*.
  Nicola De Cao, Wilker Aziz and Ivan Titov. In *NAACL*, 2019.
  https://www.aclweb.org/anthology/N19-1240/
- *A Stochastic Decoder for Neural Machine Translation*.
  Philip Schulz, Wilker Aziz and Trevor Cohn. In *ACL*, 2018.
  https://www.aclweb.org/anthology/P18-1115/
- *Fast collocation-based Bayesian HMM word alignment*.
  Philip Schulz and Wilker Aziz. In *COLING*, 2016.
  https://www.aclweb.org/anthology/C16-1296/
- *Exact Decoding for Phrase-Based Statistical Machine Translation*.
  Wilker Aziz, Marc Dymetman and Lucia Specia. In *EMNLP*, 2014.
  https://www.aclweb.org/anthology/D14-1131/

For a complete list refer to https://wilkeraziz.github.io/publications.html.

## Education

2010–2014 **Ph.D. Computational Linguistics**
Research Institute in Information and Language Processing – University of Wolverhampton
**Thesis:** Exact Sampling and Optimisation in Statistical Machine Translation
**Supervisors:** Prof. Dr. Lucia Specia, Dr. Marc Dymetman, Prof. Dr. Ruslan Mitkov
**Summary:** I introduce an approach to exact optimisation and sampling based on a form of adaptive rejection sampling which addresses challenges in global optimisation and unbiased sampling in high-dimensional discrete spaces. In this view, an intractable goal distribution is upperbounded by a tractable proxy distribution which is then incrementally refined to be closer to the goal.

2005–2010 **B.Sc. Computer Engineering**
Escola de Engenharia de São Carlos - Universidade Estadual de São Paulo (USP)
**Monograph:** Lexical Substitution for Statistical Machine Translation
**Summary:** I propose a context model based on word co-occurrence and supervised learning to rank for cross-language lexical substitution.

## Employment

01/2019–present **Assistant Professor, Institute for Logic, Language and Computation, Universiteit van Amsterdam, Netherlands**
**Summary:** My research group focuses on probabilistic models of language with a strong focus on unsupervised learning and latent variable modelling.

01/2015–12/2018 **Research Associate, Institute for Logic, Language and Computation, Universiteit van Amsterdam, Netherlands**
**Summary:** I joined the Statistical Language Processing and Learning Lab led by Professor Khalil Sima'an in January 2015 where I worked on several aspects of machine translation (e.g. word alignment, word reordering, and morphological analysis and generation) and paraphrasing employing log-linear, Bayesian, and deep generative models.

11/2013–12/2014 **Research Associate, Department of Computer Science, University of Sheffield, UK**
**Summary:** My work was funded by EPSRC under the MODIST (MOdelling DIscourse in Statistical Translation) project led by Prof. Dr. Lucia Specia. Discourse information typically requires nonlocal forms of parameterisation. I developed better decoding algorithms for SMT aiming at incorporating global features, particularly, I worked on a lazy incorporation of nonlocal parameterisation using a form of adaptive rejection sampling.

08/2013–12/2013 **Internship, Xerox Research Centre Europe (XRCE), Grenoble, France**
**Summary:** I worked with the Machine Learning for Document Access and Translation group under supervision of Dr. Marc Dymetman and Dr. Sriram Venkatapathy on developing an exact decoder/sampler for phrase-based SMT.

03/2009–02/2010 **Internship, Xerox Research Centre Europe (XRCE), Grenoble, France**
**Summary:** I worked with the Cross-Language Technologies group under supervision of Dr. Marc Dymetman and Dr. Lucia Specia on the use of context models and textual entailment to improve statistical machine translation coverage and quality.

## Services

I often serve on the program committee of ML/CL/NLP/MT conferences:

- CL/NLP conferences: ACL, EMNLP, NAACL, Coling, IJCNLP, CoNLL
- ML conferences: NeurIPS (top reviewer 2019), ICML, ICLR
- MT conferences: WMT, EAMT, AMTA, MT Summit

and I review for CL/NLP/MT journals: Transactions of ACL (TACL), Natural Language Engineering, Computer Speech and Language, and Machine Translation. I serve (or have served) as area chair for *Machine Learning for Natural Language Processing* for *SEM 2019 and ACL 2020.

## Patents

U.S. Patent Application Filing: SAMPLING AND OPTIMIZATION IN PHRASED-BASED MACHINE TRANSLATION USING AN ENRICHED LANGUAGE MODEL REPRESENTATION
Inventor(s): Marc Dymetman; Wilker Aziz; Sriram Venkatapathy
U.S. Ser. No.: 13/750,338. Filed on: 01/25/2013

U.S. Patent Application Filing: DYNAMIC BI-PHRASES FOR STATISTICAL MACHINE TRANSLATION
Inventor(s): Marc Dymetman; Wilker Aziz; Nicola Cancedda; Jean-Marc Coursimault; Vassilina Nikoulina; Lucia Specia.
U.S. Ser. No.: 12/780,040. Filed on: 05/20/2010

## Teaching

I am highly committed to education and I believe in empowering students with knowledge in a healthy manner. Since 2015, I coordinate, design, and implement MSc and BSc courses offered at UvA covering topics such as:

- probabilistic graphical models and Bayesian methods for NLP
- approximate probabilistic inference: Markov chain Monte Carlo sampling and variational inference
- weighted automata and grammars, semirings, and deductive systems
- statistical and neural approaches to natural language processing

I mostly design my own materials and value depth and quality.

| | |
|---|---|
| **Course** | *Deep Learning for Natural Language Processing* |
| **Role** | Lecturer (2019)<br>offered in collaboration with Christof Monz (IvI) |
| **Programme** | Master's of AI (UvA) |
| **Description** | The course covers advanced supervised and unsupervised learning techniques in natural language processing with a focus on statistical learning powered by deep neural networks. |
| **Course** | *Basic Probability Theory* |
| **Role** | Coordinator (2019) |
| **Programme** | Master's of Logic (UvA) |
| **URL** | https://basicprobability.github.io |
| **Description** | The course covers the basics of combinatorics, axiomatic probability theory, discrete and continuous random variables, and maximum likelihood estimation. |
| **Course** | *Unsupervised Language Learning* |
| **Role** | Lecturer (2018)<br>offered in collaboration with Ekaterina Shutova (ILLC) |
| **Programme** | Master's of AI (UvA) |
| **URL** | https://uva-slpl.github.io/ull/ |
| **Description** | The course covers advanced unsupervised learning techniques in natural language processing with a focus on meaning representation (including deep generative models of word and sentence representation). |
| **Course** | *Natural Language Processing 2* |
| **Role** | Lecturer (2015-2017) |
| **Programme** | Master's of AI (UvA) |
| **URL** | https://uva-slpl.github.io/nlp2/ |
| **Description** | The course covers structure prediction problems related to translation (e.g. unsupervised alignment, synchronous grammar induction, statistical and neural MT). |
| **Course** | *Natural Language Models and Interfaces* |
| **Role** | Coordinator (2018-present) |
| **Programme** | Bachelor's of AI (UvA) |
| **URL** | https://uva-slpl.github.io/nlmi/ |
| **Description** | The course covers some of the essential techniques in natural language processing with a focus on language modelling and word representation. |

For a complete list of courses and projects including available material and project outcomes, please refer to https://wilkeraziz.github.io/teaching.html. For invited lectures and talks, please refer to https://wilkeraziz.github.io/talks.html.

I have recently developed a tutorial on variational inference and deep generative models for NLP audiences. This is a joint effort with Philip Schulz, and we have been taking this tutorial (or parts of it) to diverse audiences at universities (e.g. Melbourne, Monash, Macquarie, Amsterdam, Heidelberg, and Instituto Superior Técnico in Lisbon), companies (e.g. Amazon, Naver Labs, Yandex), and international conferences (ACL 2018). For more information check our schedule and available resources: https://vitutorial.github.io.

## Supervision

I see supervision as a very special form of training where the goal is to transfer not only skills, but also an attitude towards research and towards people. Above all, I respect my students' personal lives and incentivise them to do the same. My experience with one-to-one supervision includes PhD, MSc, and BSc students.

| | ONGOING | ROLE |
|---|---|---|
| PHD | Bryan Eikema (2019–present; UvA) | supervisor |
| TOPIC | *Data-Efficient Probabilistic Neural Machine Translation* | |
| PHD | Nicola De Cao (2019–present; UvA) | co-supervisor |
| TOPIC | *Role of Knowledge Bases and Logical Reasoning in Question Answering* | |
| PHD | Joost Bastings (2017–present; UvA) | co-supervisor |
| THESIS | *Incorporating Linguistic Structure in Neural Machine Translation* | |
| MSC | Lina Murady (2019–present; UvA) | supervisor |
| TOPIC | *Interpretable Probabilistic Text Classifiers* | |
| MSC | Mathijs Pieters (2019–present; UvA) | supervisor |
| TOPIC | *Global Explanations in Text Classification* | |
| MSC | Eelco van der Wel (2019–present; UvA) | supervisor |
| TOPIC | *Controllable Text Generation* | |
| MSC | Ruben Gerritse (2019–present; UvA) | supervisor |
| TOPIC | *Semi-Supervised Neural Machine Translation* | |

| | COMPLETE | ROLE |
|---|---|---|
| PHD | Philip Schulz (2015–2018; UvA) | co-supervisor |
| THESIS | *Latent Variable Models for Machine Translation and How to Learn Them* | |
| PHD | Joachim Daiber (2016–2018; UvA) | co-supervisor |
| THESIS | *Typologically Robust Statistical Machine Translation* | |
| PHD | Miloš Stanojević (2015–2017; UvA) | co-supervisor |
| THESIS | *Permutation Forests for Modeling Word Order in Machine Translation* | |
| MSC | Nuno Mota (2018–2019; UvA) | supervisor |
| THESIS | *Textual (Generalised) Any-Shot Learning - The Case of Relation Classification* | |
| MSC | Akash Raj (2019; UvA) | supervisor |
| THESIS | *Semi-supervised Morphological Reinflection using Rectified Random Variables* | |
| MSC | Dhruba (2019; UvA) | supervisor |
| THESIS | *Supervised Neural Disease Normalization* | |
| MSC | Daan van Stigt (2018–2019; UvA) | supervisor |
| THESIS | *Neural Language Models with Latent Syntax* | |
| MSC | Tom Pelsmaeker (2017–2018; UvA) | supervisor |
| THESIS | *Effective Estimation of Deep Generative Models of Language* | |
| MSC | Bryan Eikema (2017–2018; UvA) | supervisor |
| THESIS | *Semi-Supervised Learning for Neural Machine Translation* | |
| MSC | Sanders Bijl de Vroe (2017; UU) | supervisor |
| THESIS | *Character-level Neural Architectures for Jointly Predicting Word Alignments and Word-internal Structure in Morphologically Complex Languages* | |

For more information and links to theses, please refer to https://wilkeraziz.github.io/people.html.

## Personal

I am a native Portuguese speaker who most of the time speaks English (fluent) or Italian (intermediate), and once upon a time knew some French. I play the guitar and try to keep Brazilian bossa alive, though on a very small scale. I love pets (all of them!), but I find dogs objectively superior to cats.