

Lecture 1: Introduction

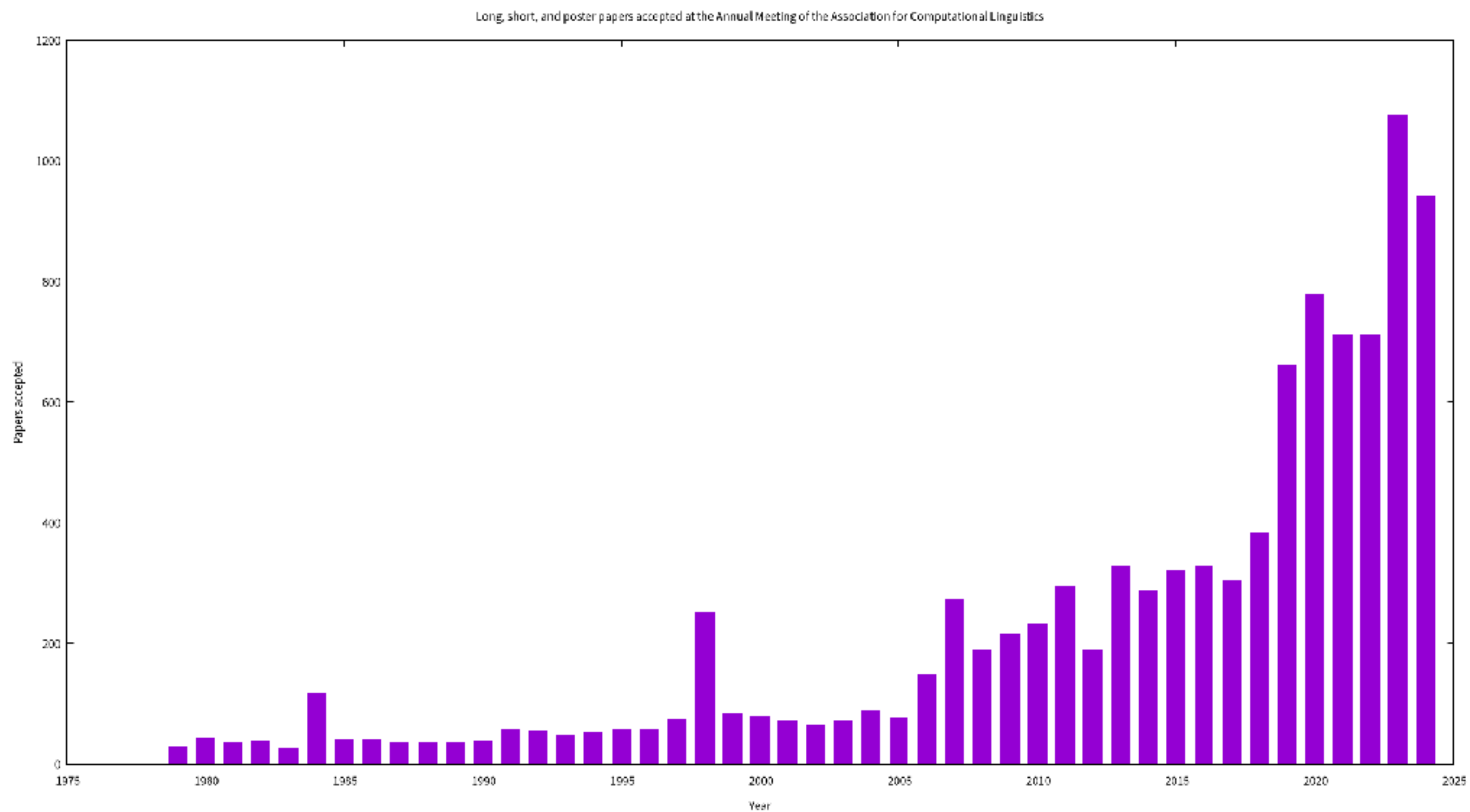
COMP0087 2024/25

14 January 2025

Pontus Saito Stenetorp

- University College London:
 - * Professor of Natural Language Processing
 - * Natural Language Processing group leader
 - * Deputy Director for the Centre for Artificial Intelligence
- National Institute of Informatics:
 - * Research and Development [Centre] for Large Language Models (LLMC)
 - * Specially Appointed Researcher
 - * Hiring research engineers (<https://llmc.nii.ac.jp>)
- p.stenetorp@cs.ucl.ac.uk
- <https://pontus.stenetorp.se>

Module content



Learning goals:

1. Foundational "building blocks"

- Token representations, recurrent neural networks, transformers, etc.

2. Why language is difficult

- Failure cases, relationship to thought, etc.

3. Practical, empirical understanding of building NLP systems

- Analysis, empirical experiments, etc.
- Group project

Module logistics

Delivery format:

- Recorded lectures
- Reading list
- "Flipped" sessions
- Invited lectures
- Examination: Group project (100%)

- Pontus Saito Stenetorp
 - * Module leader
 - * Responsibilities:
 - Overall module delivery
 - Lectures
 - "Flipped" sessions

- Vasileios "Bill" Lamos
 - * Module co-leader
 - * Responsibilities:
 - Group project management
 - Teaching assistant management
 - Marking management

- Teaching assistants:
 - * Ten in total
 - * Responsibilities:
 - Group project supervision
 - Marking and co-marking

Communication:

- *Largest* postgraduate module
- General, non-personal, queries:
 - * Please, pretty please use the Moodle forum
- Personal queries:
 - * comp0087@cs.ucl.ac.uk
- General, project-related queries:
 - * Moodle forum
- Specific, project-related queries:
 - * Assigned teaching assistant
- Administrative logistics (registration, etc.):
 - * cs.pgt-students@ucl.ac.uk

Extenuating circumstances:

- Not mandatory, but I would appreciate talking to me

Pontus' office hours:

- Mondays (not reading week): 09:00-11:00
- Booking: a.winters@ucl.ac.uk (please do **not** Cc me)
- Topics do not have to be limited to the module

Generative models for coursework:

- Permitted (although not *personally* encouraged)
- *Cite* when used
- Equally liable for own *and* generated content
- See UCL's "Engaging with Generative AI in your education and assessment"

Student representatives (reps):

- One per programme
- Student task #1: Elect and inform me of your student reps

Module requirements

Requirements: Maths

- Self check:
 - * Matrix-matrix multiplication?
 - * Chain rule?
 - * Conditional probability?
 - * Multi-variate function optimisation?
- Remedy:
 - * "Deep Learning" by Goodfellow et al. (2016), Chapters 2 to 4 (49 pages):
 - <https://www.deeplearningbook.org/contents/ml.html>

Requirements: Machine learning

- Self check:

- * Gradients for a three-layer, multi-layer perceptron (MLP)?

- * Loss-curve analysis?

- * Regularisation?

- * Overfitting?

- Remedy:

- * "Deep Learning" by Goodfellow et al. (2016), Chapter 5 (65 pages):

- <https://www.deeplearningbook.org/contents/ml.html>

Requirements: Programming and computer skills

- Self check:

- * Write multi-file, 2,000-lines project?
- * Read multi-file, 10,000-lines project?
- * Package install/management?
- * IEEE 754?
- * Shell scripting?
- * Version control?
- * LaTeX?

- Remedy:

- * "Python Crash Course: A Hands-On, Project-Based Introduction to Programming" by Matthes (2023)
 - <https://nostarch.com/python-crash-course-3rd-edition>
- * "Missing Semester of Your CS Education":
 - <https://missing.csail.mit.edu>
- * "Learn LaTeX in 30 minutes":
 - https://www.overleaf.com/learn/latex/Learn_LaTeX_in_30_minutes

Group project

Coursework (Group):

- 100% of module mark:
 - * Mandatory bi-weekly progress meetings
 - * Five, bi-weekly, single-page progress reports
 - * Eight-page final report
 - Not including references
- Groups of **six** or **seven**
 - * Exceptions will be rare
- Group formation due 17 January by midnight
 - * Start looking **now**
 - * Decide on group name
 - * Post UCL e-mails in dedicated Moodle forum

Student-designed project:

- Requirement:
 - * **Must** involve language
 - * Empirical investigation
- Designed by students, with the **help** of us
 - * Ultimate responsibility and control is with **you**
- Outcomes **may** be publishable
 - * But **no** guarantees
- Best part of the module! (?)

Default project:

- "A Decomposable Attention Model for Natural Language Inference" by Parikh et al. (2016)
 - * "Proto-transformer"
 - * <https://aclanthology.org/D16-1244>
- Reimplement **without** libraries
- Evaluate against other models for a number of datasets and tasks

Project timeline sketch:

- Week 1 to 2: Project design and literature review
- Week 3 to 4: Data collection, experimental design, initial implementation
- Week 5 to 6: Experiments not working out: Revisions, revisions, revisions, ...
- Week 7 to 8: Additional experiments and analysis
- Week 9 to 10: Report writing

Advice:

- Minimum viable project
- Align with group interests and teaching assistant expertise
- Design to be "split"

Computational resources:

- **Limited**: Design your projects with this in mind!
- UCL Computer Science:
 - * <https://tsg.cs.ucl.ac.uk/gpus>
 - * ~30 x NVIDIA RTX 4060 cards (16GB)
 - * ~25 x NVIDIA RTX 3090 cards (24GB)
- Google Colab:
 - * <https://colab.research.google.com>
 - * Web "notebooks" with free GPUs and TPUs (LLM access?)
 - * Can pay for better service
- GPU server rental
- co:here credits
 - * Application process announced shortly

Tasks:

- High-level categories:
 - * Classification
 - * Generation
- Which tasks are you familiar with?
- Listing: <https://paperswithcode.com/datasets?mod=texts>

Dataset **restrictions**

Teaching assistants

Yao Lu:

- Last-year PhD student
- Interests:
 - * Model robustness
 - * Large language model training
- Favourite paper:
 - * "Fantastically Ordered Prompts and Where to Find Them: Overcoming Few-Shot Prompt Order Sensitivity" by Lu et al. (2022)
 - * <https://aclanthology.org/2022.acl-long.556>

Noah Siegel:

- 2nd-year PhD student/Senior research engineer at Google DeepMind
- Interests:
 - * Safety: Scalable oversight, faithfulness, and explanations
- Favourite paper:
 - * "The Probabilities Also Matter: A More Faithful Metric for Faithfulness of Free-Text Explanations in Large Language Models" by Siegel et al. (2024)
 - * <https://aclanthology.org/2024.acl-short.49>

Wiem Ben Rim:

- 2nd-year PhD student
- Interests:
 - * Factuality and safety
 - * Retrieval augmented generation
 - * Applications of NLP
- Favourite paper:
 - * "Behavioral Testing of Knowledge Graph Embedding Models" by Rim et al. (2021)
 - * https://openreview.net/forum?id=3_2B2MliB8V

Karen Hambardzumyan:

- 2nd-year PhD student at UCL and Facebook
- Interests:
 - * Mechanistic interpretability
 - * Multi-agent systems
- Favourite paper:
 - * "WARP: Word-level Adversarial ReProgramming" by Hambardzumyane et al. (2021)
 - * <https://aclanthology.org/2021.acl-long.381>

Hossein A Rahmani ("Saeed"):

- 3rd-year PhD student
- Interests:
 - * Evaluations and benchmarks
 - * Self-Improvement
- Favourite paper:
 - * "Large Language Models Cannot Self-Correct Reasoning Yet" by Huang et al. (2024)
 - * <https://arxiv.org/abs/2310.01798>

Lovish Madaan:

- 2nd-year PhD student at UCL and Facebook (GenAI)
- Interests:
 - * Generalisation behaviours in reinforcement learning
 - * Better/harder evaluation
- Favourite paper:
 - * "Amortizing intractable inference in large language models" by Hu et al. (2024)
 - * <https://arxiv.org/abs/2310.04363>

Edan Toledo:

- 1st-year PhD student at UCL and Facebook
- Interests:
 - * General search methods
 - * Agentic systems
 - * Meta-learning
 - * Reinforcement learning
- Favourite paper:
 - * "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play" by Silver et al. (2018)
 - * <https://www.science.org/doi/10.1126/science.aar6404>

Eduardo Sánchez:

- 3rd-year PhD student at UCL and Facebook
- Interests:
 - * Low-resource languages
 - * Multilinguality
- Favourite paper:
 - * "Linguini: A benchmark for language-agnostic linguistic reasoning" by Sánchez et al. (2024)
 - * <https://arxiv.org/abs/2409.12126>

Yuxuan Shu:

- 3rd-year PhD student
- Interests:
 - * Time series forecasting
 - * Sample generation
- Favourite paper:
 - * "Unsupervised hard Negative Augmentation for
contrastive learning" by Shu and Lamos (2024)
 - * <https://arxiv.org/abs/2401.02594>

Jiayi Wang:

- 3rd-year PhD student
- Interests:
 - * Inclusive and efficient multilingual large language models
- Favourite paper:
 - * "Multilingual Pretraining Using a Large Corpus Machine-Translated from a Single Source Language" by Wang et al. (2023)
 - * <https://arxiv.org/abs/2410.23956>

Literature review

Scientific quality:

- Hard problem
- Field specific
- Where was it published?
 - * Conference: Peer review (high)
 - * Journal: Peer review (high)
 - * Workshop: Peer review (medium to low)
 - * Pre-print: "Does this person work in academia?" (low)
 - * Whitepaper: "We are X" (none)
 - * Web page: "I can post online" (none)
- There are exceptions, but be careful initially

Sources for reading:

- Association for Computational Linguistics (ACL) Anthology
 - * Open access, free for everyone
 - * My "top" conferences for quality:
 1. Empirical Methods in Natural Language Processing (EMNLP)
 2. Annual Meeting of the ACL (ACL)
 3. Annual Conference of the Nations of the Americas Chapter of the ACL (NAACL)
 - * Transactions of the Association for Computational Linguistics (TACL)
- "General" machine learning venues:
 - * International Conference on Learning Representations (ICLR)
 - * Conference on Neural Information Processing Systems (NeurIPS)
- Computation and Language subject on arXiv:
 - * <https://arxiv.org/list/cs.CL/recent>

Upcoming: "Reading a paper" lecture

Questions and/or clarifications?

Chatbot Arena:

- How do we evaluate **strong** models?
- Go to: <https://lmarena.ai>:
 - * Try to **break** strong models:
 - What kind of inputs are difficult?
 - Try to find patterns
 - Discuss with your neighbours
 - * Last year: "What is the most northern point of South Sweden?"
 - "The most northern point of [Göteborg] is [Töre in the County of Norrbotten]" (gpt-4-0613)

- Student tasks:
 1. Elect and inform me of your student reps
 2. Form groups by 17 January (Friday) 23:59 (Europe/London)

Thank you for your attention, questions?