# KEVIN YANG
# CURRICULUM VITAE

425-786-7398 yangk@berkeley.edu

## EDUCATION

**Ph.D. in Artificial Intelligence** *2019-*
*UC Berkeley*
Advisor: Prof. Dan Klein

**M.Eng., Computer Science** *2018-2019*
*MIT*
Advisor: Prof. Regina Barzilay
Thesis: *Are Learned Molecular Representations Ready for Prime Time?*

**B.S, Computer Science and Mathematics** *2015-2019*
*MIT*
Double major; GPA 5.0/5.0

## RESEARCH INTERESTS

I am broadly interested in planning, control, and alignment methods for natural language generation and/or structured settings, especially as they apply to long-form generation tasks such as creative writing.

## PUBLICATIONS

**RLCD: Reinforcement Learning from Contrast Distillation for Language Model Alignment**
*Kevin Yang, Dan Klein, Asli Celikyilmaz, Nanyun Peng, Yuandong Tian*
*arxiv; to be submitted to ICLR 2024*
We propose a new method for simulating preference data in RLHF alignment pipelines based on generating preference pairs from two contrasting prompts, with strong downstream performance on three diverse alignment tasks and multiple LLaMA model scales.

**DOC: Generating Longer Stories with Recursive Reprompting and Revision**
*Kevin Yang, Dan Klein, Nanyun Peng, Yuandong Tian*
*ACL 2023*
We improve coherence in several-thousand-word-long stories by constructing a more detailed outline and improving the generator's ability to stay faithful to that outline. Humans prefer DOC to our previous Re$^3$ system by a wide margin in both automatic and interactive generation.

**Modular Visual Question Answering via Code Generation**
*Sanjay Subramanian, Medhini Narasimhan, Kushal Khangaonkar, Kevin Yang et al.*
*ACL 2023*
We improve performance on visual question answering tasks requiring multi-step reasoning by synthesizing programs to compose logical reasoning steps, including calling sub-tools.

**PREADD: Prefix-Adaptive Decoding for Controlled Text Generation**
*Jonny Pei, Kevin Yang, Dan Klein*
*Findings of ACL 2023*

We propose a simple, no-training-required approach for modulating the control strength exerted through prompting, by contrasting the logit distributions induced by two contrasting prompts. We achieve strong results on toxic output mitigation, bias reduction, and sentiment control.

**Predicting Compound Activity from Phenotypic Profiles and Chemical Structures**
*Nikita Moshkov, Tim Becker, Kevin Yang, Peter Horvath et al.*
*Nature Communications 2023*
We investigate the effectiveness of three different data sources—chemical structures, imaging data, and gene-expression profiles—for predicting compound activity in laboratory assays.

**Re$^3$: Generating Longer Stories with Recursive Reprompting and Revision**
*Kevin Yang, Yuandong Tian, Nanyun Peng, Dan Klein*
*EMNLP 2022*
We generate plot-coherent 2000+ word stories using a structured system which repeatedly re-prompts a language model based on the plan and previous story, followed by revising for coherence, premise relevance, and factual consistency.

**Automated Crossword Solving**
*Eric Wallace*, Nicholas Tomlin*, Albert Xu*, Kevin Yang* et al.*
*ACL 2022*
We create a system for automatically solving crossword puzzles, and achieve superhuman performance for the first time.

**Addressing Resource and Privacy Constraints in Semantic Parsing Through Data Augmentation**
*Kevin Yang, Olivia Deng, Charles Chen, Richard Shin et al.*
*Findings of ACL 2022*
We propose a data augmentation scheme for low-resource semantic parsing in complex realistic environments, which simultaneously maintains user privacy.

**Multi-Objective Optimization by Learning Space Partitions**
*Yiyang Zhao, Linnan Wang, Kevin Yang, Tianjun Zhang et al.*
*ICLR 2022*
We propose a space-partitioning search algorithm for finding the Pareto frontier in multi-objective optimization problems.

**Learning Space Partitions for Path Planning**
*Kevin Yang*, Tianjun Zhang*, Chris Cummins, Brandon Cui et al.*
*NeurIPS 2021*
We propose a path planning method inspired by a theoretical analysis of search space partitioning, and show strong performance on difficult multimodal, long-horizon path planning problems.

**FUDGE: Controlled Text Generation with Future Discriminators**
*Kevin Yang, Dan Klein*
*NAACL 2021*
We propose a simple, flexible, and highly effective method for controlling generation toward desired attributes using lightweight classifiers.

**A Streaming Approach for Efficient Batched Beam Search**
*Kevin Yang, Violet Yao, John DeNero, Dan Klein*
*EMNLP 2020*
We propose an efficient batching strategy for variable-length decoding on GPU architectures, demonstrating substantial speedups over existing fixed-width and variable-width beam searches.

**Improving Molecular Design by Stochastic Iterative Target Augmentation**
*Kevin Yang, Wengong Jin, Kyle Swanson, Regina Barzilay, Tommi Jaakkola*
*ICML 2020*
We use a simple and theoretically motivated self-training approach guided by an external property predictor to substantially improve over state-of-the-art approaches in molecular design.

**Uncertainty Quantification Using Neural Networks for Molecular Property Prediction**
*Lior Hirschfeld, Kyle Swanson, Kevin Yang, Regina Barzilay, Tommi Jaakkola*
*JCIM 2020*
We comprehensively evaluate and compare several approaches for uncertainty estimation in neural models on molecular property prediction tasks.

**A Deep Learning Approach to Antibiotic Discovery**
*Jonathan Stokes, Kevin Yang, Kyle Swanson, Wengong Jin et al.*
*Cell 2020*
We use computational property prediction models to screen drug databases for potential antibiotic activity, and discover previously unknown antibiotics with novel mechanisms of action which are effective even against bacteria which are resistant to commonly used antibiotics.

**Analyzing Learned Molecular Representations for Property Prediction**
*Kevin Yang, Kyle Swanson, Wengong Jin, Connor Coley et al.*
*JCIM 2019*
We introduce a new variant of message-passing neural networks, demonstrating consistently strong performance that significantly improves over existing baselines on many datasets. We also carefully benchmark models on both public and proprietary industry datasets.

**Learning Multimodal Graph-to-Graph Translation for Molecular Optimization**
*Wengong Jin, Kevin Yang, Regina Barzilay, Tommi Jaakkola*
*ICLR 2019*
We introduce an encoder-decoder architecture for molecular optimization that operates directly on the molecular graph, substantially outperforming string-based baselines as well as pre-existing state of the art.

## PROFESSIONAL ACTIVITIES

**Visiting Researcher, Meta AI** *2022-2023*
Part-time work on methods for advancing AI story generation: topics include long-range factual consistency, improving planning and outlining, and RLHF alignment methods.

**Research Intern, Microsoft Research (Semantic Machines)** *2021*
12-week summer internship working on multiturn semantic parsing with a particular focus on data privacy.

**Teaching Assistant, MIT** *2018*
Course: Introduction to Inference

**Research Intern, ASAPP** *2018*
12-week summer internship researching and productionizing natural language processing models on difficult real-world text datasets.

**Equities Research Intern, DE Shaw** *2017*

11-week summer internship in equities research group. I developed a model predicting existence of hidden liquidity in equities exchanges, with strong results on real-world data.

**Software Engineering Intern, Google** *2016*
12-week summer internship focusing on improving search ad efficiency and optimization.

## AWARDS AND HONORS

| | |
|---|---|
| NSF Graduate Fellowship | *2019* |
| Putnam Top 200 | *2016-2018* |
| International Linguistics Olympiad Gold Medal (5th place) | *2015* |
| USA Math Olympiad Honorable Mention (14th place) | *2013-2015* |

## LANGUAGES

Proficient in Python; experienced in Java, C, R, HTML, Javascript. Also proficient in PyTorch. Natively fluent in Mandarin Chinese, advanced Japanese (JLPT N3), advanced Spanish.

## REFERENCES

**Dan Klein**, Professor
Computer Science Division, UC Berkeley
510-642-3214 klein@berkeley.edu

**Regina Barzilay**, Professor
CSAIL, MIT
617-258-5706 regina@csail.mit.edu