

Jay-Yoon Lee

✉ lee.jayoon@gmail.com | 🏠 leejayyoon.github.io | 🌐 leejayyoon

RESEARCH INTEREST

- Injecting *knowledge as constraints* into neural models.
- Making the models more *coherent*, *interpretable*, and *controllable*.
- Resolving *low-resource* problem by incorporating prior knowledge such as constraints.
- Automatically capturing implicit constraints using *energy-based models*.
- **Keywords:** NLP, Constraint injection, Structured Prediction, Energy-based models, Multi-task & Transfer learning, Domain adaptation, Logical reasoning, and better representation learning for the aforementioned topics.

EDUCATION

08/2013 - 08/2020	Manning College of Information & Computer Sciences, University of Massachusetts, Amherst <i>Postdoctoral Associate</i> LEARNING LATENT CONSTRAINTS AUTOMATICALLY FROM DATA UTILIZING ENERGY-BASED MODELS AND SPATIAL BOX REPRESENTATIONS. <ul style="list-style-type: none">– Served as a research mentor for 7 PhD students.– Supervised research projects for 20 masters students, including projects with IBM and Facebook.
08/2013 - 08/2020	Computer Science Department, School of computer science, Carnegie Mellon University <i>Ph.D. in Computer Science.</i> Advisor: Jaime Carbonell <u>Ph.D. thesis:</u> INJECTING OUTPUT CONSTRAINTS TO NEURAL NLP MODELS, 2020 <u>Committee:</u> Jaime Carbonell, William Cohen, Graham Neubig, Yulia Tsvetkov, Dan Roth.
08/2011 – 05/2013	Language Technology Institute & Lane Center for Computational Biology, School of computer science, Carnegie Mellon University <i>M.S. in Computer Science. Transferred to PhD program at the end without completing.</i>
03/2003 – 05/2008	Korea Advanced Institute of Science and Technology (KAIST) <i>B.S., Summa Cum Laude, in Electrical Engineering</i> <u>Undergrad thesis:</u> ROBUST SPEECH RECOGNITION USING PREDICTION METHODS BASED ON ARTIFICIAL NEURAL NETWORKS, 2007 <u>Thesis advisor:</u> Chulhoon Park.

WORK EXPERIENCES

July 2020 – current	Postdoctoral associate under Professor Andrew McCallum <i>UMass Amherst, Amherst, MA</i>
Oct 2015 – July 2020	Research Assistant under Professor Jaime Carbonell <i>Carnegie Mellon University, Pittsburgh, PA</i> Injecting output constraints into the neural NLP models for low-resource, domain adaptation.; Part-forecasting, price prediction project (Boeing sponsored).
June 2012 – Oct 2015	Research Assistant under Professor Christos Faloutsos <i>Carnegie Mellon University, Pittsburgh, PA</i> Anomaly detection algorithm on large graphs (DARPA ADAMS project).
Oct 2019 – Jan 2020	Research Internship at Language and Speech Group <i>Google AI, New York, NY</i> Tag-constrained Transformer for text normalization.
June – Aug 2017 ,2019	Research Internship at Information and Data Sciences & Deep Learning Group <i>Microsoft Research, Redmond, WA</i>

	Improving dialogue model using multiple non-differentiable metrics. Task-oriented dialogue when action is defined on the combinatorially space.
June 2016 – Aug 2016	Research Internship at Machine Learning Research Group <i>Oracle Labs</i> , Burlington, MA Injecting hard-constraints to sequence-to-sequence model for syntactic parsing. (Patent published)
Sep 2009 – Jun 2011	Researcher <i>ZEROIN Corporation</i> , Seoul, Korea Developed INDEX Management System for National Pension Fund (Java)
Jul 2008 – Aug 2009	Associate <i>National Information & Credit Evaluation (NICE) Pricing Services, INC.</i> , Seoul, Korea Evaluation of Equity-Linked securities and SWAP securities.
Oct 2007 – Jan 2008	Fall Internship <i>Amicus Wireless Technology</i> , Sunnyvale, CA Software, hardware tests for WiMAX chipsets.

PUBLICATIONS

Asterix () denotes joint first author, i.e., equal contribution. Paper titles are linked to the pdf file.*

- [1] **Structured Energy Network as a dynamic loss function: A case study with multi-label Classification.**
Openreview preprint, 2021.
Lee, Jay-Yoon and Patel, Dhruvesh and Goyal, Purujit and McCallum, Andrew.
- [2] **Event-Event Relation Extraction using Probabilistic Box Embedding.**
In *ACL (Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics)*, 2022.
Hwang, Eunjeong and **Lee, Jay-Yoon** and Yang, Tianyi and Patel, Dhruvesh and Zhang, Dongxu and McCallum, Andrew.
- [3] **Modeling Label Space Interactions in Multi-label Classification using Box Embeddings.**
In *ICLR (International Conference on Learning Representations)*, 2022.
Patel, Dhruvesh and Dangati, Pavitra and **Lee, Jay-Yoon** and Boratko, Michael and McCallum, Andrew.
- [4] **Improved Latent Tree Induction with Distant Supervision via Span Constraints.**
In *EMNLP (Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing)*, 2021.
Xu, Zhiyang and Drozdov, Andrew and **Lee, Jay-Yoon** and O’Gorman, Tim and Rongali, Subendhu and Finkbeiner, Dylan and Suresh, Shilpa and Iyyer, Mohit and McCallum, Andrew.
- [5] **Case-based Reasoning for Natural Language Queries over Knowledge Bases.**
In *EMNLP (Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing)*, 2021.
Das, Rajarshi and Zaheer, Manzil and Thai, Dung and Godbole, Ameya and Perez, Ethan and **Lee, Jay-Yoon** and Tan, Lizhen and Polymenakos, Lazaros and McCallum, Andrew.
- [6] **StructSum: Summarization via Structured Representations.**
In *EACL (Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume)*, pages 2575–2585, 2021.
Balachandran, Vidhisha and Pagnoni, Artidoro and **Lee, Jay-Yoon** and Rajagopal, Dheeraj and Carbonell, Jaime G and Tsvetkov, Yulia.
- [7] **Semi-supervised learning on meta structure: Multi-task tagging and parsing in low-resource scenarios.**
In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 8344–8351, 2020.
Lee*, Jay-Yoon and Lim*, KyungTae and Carbonell, Jaime and Poibeau, Thierry.

- [8] **Gradient-based inference for networks with output constraints.**
In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 4147–4154, 2019.
Lee, Jay-Yoon and Mehta, Sanket Vaibhav and Wick, Michael and Tristan, Jean-Baptiste and Carbonell, Jaime.
- [9] **Towards Semi-Supervised Learning for Deep Semantic Role Labeling.**
In *EMNLP (Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing)*, pages 4958–4963, 2018.
Lee*, Jay-Yoon and Mehta*, Sanket Vaibhav and Carbonell, Jaime G.
- [10] **Enforcing Output Constraints via SGD: A Step Towards Neural Lagrangian Relaxation.**
In *AKBC@NIPS*, 2017.
Jay Yoon Lee and Michael L. Wick and Jean-Baptiste Tristan and Jaime G. Carbonell.
- [11] **Preferential attachment in graphs with affinities.**
In *AISTATS (Artificial Intelligence and Statistics)*, pages 571–580. PMLR, 2015.
Lee*, Jay-Yoon and Zaheer*, Manzil and Günnemann, Stephan and Smola, Alex.
- [12] **Net-ray: Visualizing and mining billion-scale graphs.**
In *PAKDD (Pacific-Asia Conference on Knowledge Discovery and Data Mining)*, pages 348–361. Springer, 2014.
Kang, U and **Lee, Jay-Yoon** and Koutra, Danai and Faloutsos, Christos.
- [13] **Influence propagation: Patterns, model and a case study.**
In *PAKDD (Pacific-Asia Conference on Knowledge Discovery and Data Mining)*, pages 386–397. Springer, 2014.
Lin, Yibin and Raza, Agha Ali and **Lee, Jay-Yoon** and Koutra, Danai and Rosenfeld, Roni and Faloutsos, Christos.
- [14] **Detecting insider threats in a real corporate database of computer usage activity.**
In *KDD (Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining)*, pages 1393–1401, 2013.
Senator, Ted E and Henry, G and Goldberg, Alex Memory and Young, William T and Rees, Brad and Pierce, Robert and Huang, Daniel and Reardon, Matthew and **Lee, Jay-Yoon** and Koutra, Danai and others.
- [15] **Fast anomaly detection despite the duplicates (Poster).**
In *WWW (Proceedings of the 22nd International Conference on World Wide Web)*, pages 195–196, 2013.
Lee, Jay-Yoon and Kang, U and Koutra, Danai and Faloutsos, Christos.

PATENTS PUBLISHED

[US20180121807A1] **When output units must obey hard constraints** May 3, 2018

- Gradient-based inference searches over continuous model-weight space to produce constraint-satisfying output.

RESEARCH EXPERIENCE

Structured Energy network as A Loss (SEAL) 2021 – current
Postdoctoral associate *UMass Amherst*

- Proposed SEAL framework that utilizes structured energy network as a trainable loss function (*loss-net*) that can teach the standard discriminative network (*task-net*).
- The energy network, loss-net, can learn dependencies in the output space and thus can distill this information to task-net.
- Proposed dynamic loss-net-learning framework that can adapt to task-net distribution in training time.
- Built framework for SEAL on top of allenNLP that supports independent optimizers and schedulers for loss-net and task-net.
- One conference paper submission (under review).
- The paper showed notable improvements for multi-label classification (MLC) task, for both feature- and text-based MLC, without providing any prior dependencies between labels.

Probabilistic box embeddings for capturing constraints automatically

2020 – current

Postdoctoral associate

UMass Amherst

- Previous works have shown that spatial box embeddings can represent asymmetric relationship such as parent-child relationships very well whereas vector embeddings often fail in representing them.
- In ICLR22 [3], with an application of multi-label classification, we showed that box embeddings can capture label taxonomy automatically and perform better than vector embeddings. The paper further showed that it can perform on par with a model that utilizes the true taxonomy.
- In ACL22 [2], we showed that box representations can model event-event relations more logically where symmetry, transitivity, and conjunctive constraints are much well enforced compared to vector-representation models.
- Two papers [3, 2] published at **ICLR** and **ACL**.

Case-based reasoning (CBR) for QA and KBQA

2020 – current

Postdoctoral associate

UMass Amherst

- By using semi-parametric modeling, CBR learns to refer to known (labeled) question-answer cases rather than memorizing all the information of training data in the parameter space.
- CBR approach showed the state-of-the-art performance in Knowledge-Based Question Answering (KBQA) [5].
- CBR approach especially showed its strength when new examples from unseen entities, relations were provided.
- One paper has been published in **EMNLP** and additional research paper on QA with natural texts is in progress.

Extracting a subgraph from large-scale commonsense graph using box embeddings

2021 – current

Postdoctoral associate

UMass Amherst

- Commonsense graphs are often too big to use graph neural networks on top of it and for that reason, people apply heuristic approach to prune the graph which could lead to missing important structures of graph.
- Collaboration with IBM research in extracting a subgraph from large commonsense graph and applying it to commonsense QA and text games.

Meta adapters

2020 – 2021

Postdoctoral associate

UMass Amherst

- Proposed a meta-learning framework that utilizes parameter-efficient training framework for adapters (parameter-efficient learning method for large pre-trained models such as BERT) in few-shot learning setup.
- This shows better performance while only fine-tuning 0.6% of the BERT compared to existing adapter methods for domain transfer which utilize more than 5% of the BERT parameters.
- One paper is under review at a conference.

Semi-supervised learning with output constraints for neural NLP model

2017 – current

Research Assistant, Postdoctoral associate

Carnegie Mellon University & UMass Amherst

- This approach showed a significant improvements in low-resource and also showed on par or better performance on top of state-of-the-art models [9, 7].
- In [9], syntactic constraint from parse tree was applied to semantic role labeling (SRL).
- In [7], coherence (agreement constraint) was promoted between multi-view models, thus called Co-meta, using unlabeled dataset.
- During postdoc, I have extended this research toward distant supervision on unsupervised syntactic parsing model [4] where easy-to-gather span constraints are injected as a constraint to parsing tasks.
- Currently working on injecting constraint that the output of two similar tasks, abstract meaning representation (AMR) and information extraction (IE), have to be consistent utilizing unlabeled data. This showed better performance than the state-of-the-art model and is being prepared for publication.
- Three papers [9, 7, 4] have been published in **AAAI** and **EMNLP** and additional research on Question-Answering (QA) with natural texts is in progress.

Gradient-based inference using output constraints

2017 – 2019

Research Assistant

Carnegie Mellon University

- Most of the sequence-prediction tasks such as sequence generation and sequence tagging requires combinatorial search in exponential search in order to enforce global constraints on the sequence.
- Instead of conducting “discrete” combinatorial search, this project proposes to conduct gradient-based inference on “continuous” model parameters which produces output.

- This approach showed better and faster performance than combinatorial search on several tasks: transducer with arbitrary constraints, syntactic parsing with structural constraints, and SRL with syntactic constraints.
- This approach was also useful for *low-resource problem* and *domain-adapation*.
- One conference paper [8] (**AAAI**), workshop paper[10] (**AKBC**), and one patent (US20180121807A1) got published.

Part forecasting and price prediction for airplane parts

2015 – 2020

Research Assistant

Carnegie Mellon University

- This was part of Boeing project which had a sponsorship of \$15m between 2015-2021.
- Part forecasting had several regression task: predicting part longevity, part price, and part inventory.
- This resulted in significant amount of profit for Boeing (unable to disclose the amount).

Detecting anomalies in large-scale graphs

2012 – 2015

Research Assistant

Carnegie Mellon University

- Funding was provided by the U.S. Army Research Office (ARO) and Defense Advanced Research Projects Agency (DARPA) under Contract Number W911NF-11-C-0088.
- Given a billion-scale graph, inspecting this enormous graph for an outlier requires novel data structures and scalable analysis tools.
- Published two conference papers in **KDD**, **PAKDD** and one poster at **WWW**.

TEACHING EXPERIENCE

Spring 2022	Project mentor , <i>Industry Mentorship Program course.</i> , UMass Amherst
Spring 2021	Ph.D. evaluator , <i>Industry Mentorship Program course.</i> , UMass Amherst
Spring 2015	Teaching Assistant , <i>Machine Learning (Ph.D. level)</i> , Carnegie Mellon University
Fall 2014	Teaching Assistant , <i>Multimedia Databases and Data Mining</i> , Carnegie Mellon University

RESEARCH MENTOR EXPERIENCE

I was fortunate to mentor and collaborate with various Ph.D. and MS students. The following are selected to demonstrate my research mentor experiences.

Zhiyang Xu

2021 – current

Former MS student

UMass Amherst

- Conducted semi-supervised learning research that injects constraints utilizing unlabeled datasets.
- This collaboration led to a publication in EMNLP21 [4] and in preparation for another paper. The collaboration is ongoing even after he joined Ph.D. program at the Virginia Tech.

Dhruvesh Patel

2020 – current

current PhD student

UMass Amherst

- Conducted large-scope research on automatically capturing constraints which led to two conference publications [3, 2] and one submission under review.

Purujit Goyal

2021 – current

Former MS student

UMass Amherst

- Conducted research on capturing latent output dependency using structured energy network. This led to a preprint [1] under submission.
- The collaboration is ongoing even after he joined Openreview.

Eunjeong Hwang

2021

Former MS student

UMass Amherst

- Conducted research on making event-event relation extraction to capture logical constraint using box representations.
- This collaboration led to publication in ACL22 short paper [2] and she will be joining Ph.D. program at the University of British of Columbia.

Sanket Mehta*Former MS, current Ph.D. student*

2018-2019

Carnegie Mellon University

- Conducted research on injecting hard output constraints to neural NLP models [9, 8].

Vidhisha Balachandran*Former MS, current LTI Ph.D. student*

2018-2019

Carnegie Mellon University

- Conducted research on injecting a prior of extracted paragraph structure to the abstract document summarization [6].

PROFESSIONAL SERVICE

2019-2022

Reviewer, *ICML, ICLR, NeurIPS, ACL and EMNLP***HONORS AND AWARDS**

Mar 2003 - Feb 2008

Scholarship in Science & Technology, Korea Student Foundation

Mar 2005 - Feb 2008

Scholarship in Mathematics, Korea Foundation for Advanced Studies

Aug 2002

Prize of encouragement, 5th Nationwide High School Mathematics Competition, Korea University

Aug 2016

Best Presenter (\$3k prize), Global Top Talent Forum, Hyundai Motor Group, San Diego, CA

June 2016

1st place (\$5k team prize), Qualcomm NeuroHackathon, Carnegie Mellon University, PA**TECHNICAL SKILLS**

Major Language

Python, Java, MATLAB, C++, SQL

Library & OthersPytorch, Tensorflow, AllenNLP, **Weights & Bias**, Keras, Git, LaTeX