Jay-Yoon Lee

✉ lee.jayoon@gmail.com | ⌂ leejayyoon.github.io | in leejayyoon

RESEARCH INTEREST

· Injecting knowledge/constraints into neural models.
· Making the models more coherent, interpretable, and well-performing on low-resource.
· Automatically capturing implicit constraints using energy-based models.
· Broadly interested in structured prediction, multi-task learning, logical reasoning, case-based reasoning, and better representation learning for the aforementioned topics.

EDUCATION

08/2013 - 06/2020	**Computer Science Department, School of computer science, Carnegie Mellon University**

Ph.D. in Computer Science. Advisor: Jaime Carbonell
Ph.D. thesis: "*Injecting output constraints to neural NLP models,*" 2020
Committee: 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**

M.S. in Computer Science. *Transferred to PhD program at the end without completing.*

03/2003 – 05/2008	**Korea Advanced Institute of Science and Technology (KAIST)**

B.S. *Summa Cum Laude* in Electrical Engineering
Undergrad thesis: "*Robust speech recognition using prediction methods based on artificial Neural Networks,*" 2007
Thesis advisor: Chulhoon Park.

WORKING EXPERIENCE

July 2020 – current	**Postdoctoral associate** under Professor Andrew McCallum
	UMass Amherst, Amherst, MA
	Incorporating latent structures or constraints automatically from data utilizing energy-based models and spatial box embeddings.
Oct 2015 – July 2020	**Research Assistant** under Professor Jaime Carbonell
	Carnegie Mellon University, Pittsburgh, PA
	Injecting output constraints into the neural NLP models,
	Part-forecasting project (Boeing sponsored).
June 2012 – Oct 2015	**Research Assistant** under Professor Christos Faloutsos
	Carnegie Mellon University, Pittsburgh, PA
	Anomaly detection algorithm on large graphs (DARPA ADAMS project).
Oct 2019 – Jan 2020	**Research Internship** at Language and Speech Group
	Google AI, New York, NY
	Tag-constrained Transformer for text normalization.
June 2020 – Aug 2020	**Research Internship** at Deep Learning Group
	Microsoft Research, Redmond, WA
	RL for task-oriented dialogue when action is defined on the combinatorially space.
June 2017 – Aug 2017	**Research Internship** at Information and Data Sciences Group
	MSR & Bing, Redmond, WA
	Improving conversation using multiple metrics.
June 2016 – Aug 2016	**Research Internship** at Machine Learning Research Group

	Orcale Labs, Burlington, MA
	Injecting hard-constraints to sequence-to-sequence model for syntactic parsing.
May 2015 – Aug 2015	**Research Internship**
	Yahoo! Labs, Sunnyvale, CA
Sep 2009 – Jun 2011	**Researcher**
	ZEROIN Corporation, Seoul, Korea
	Developed INDEX Management System for National Pension Fund (Java)
Jul 2008 – Aug 2009	**Associate**
	National Information & Credit Evaluation (NICE) Pricing Services, INC., Seoul, Korea
	Evaluation of Equity-Linked securities and SWAP securities.
Oct 2007 – Jan 2008	**Fall Internship**
	Amicus Wireless Technology, Sunnyvale, CA
	Software, hardware tests for WiMAX chipsets.

PUBLICATIONS

Asterix (*) denotes joint first author, i.e., equal contribution.

[1] **Structured energy network as a dynamic loss function. case study. a case study with multi-label classification**.
Openreview preprint, 2021.
Jay-Yoon Lee and Dhruvesh Patel and Purujit Goyal and Andrew McCallum.

[2] **Consistent event-event relation extraction using box embeddings**.
In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. 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 *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 *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 *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 9594–9611, 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 *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 *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 *Artificial Intelligence and Statistics*, pages 571–580. PMLR, 2015.
Lee*, Jay-Yoon and Zaheer, Manzil* and Günnemann, Stephan and Smola, Alex.

[12] **Influence propagation: Patterns, model and a case study**.
In *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.

[13] **Net-ray: Visualizing and mining billion-scale graphs**.
In *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.

[14] **Detecting insider threats in a real corporate database of computer usage activity**.
In *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**.
In *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 embedding can represent asymmetric relationship such as parent-child relationships very well whereas vector embeddings have difficulty.
· 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 taxanomy.
· 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 existing question-answer pairs rather than memorizing everything in the parameter.
· 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 in few-shot learning setup.
· This shows better performance while only fine-tuning 0.6% of the large pre-trained model compared to existing adapter methods for domain transfer which utilize more than 5% of the parameters.
· One paper is under review at a conference.

Case-based reasoning (CBR) for QA and KBQA   2020 – current
Postdoctoral associate   *UMass Amherst*

· By using semi-parametric modeling, CBR learns to refer to existing question-answer pairs rather than memorizing everything in the parameter.
· 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.

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 beteween 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 taks.
· 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 texets is in progress.

Gradient-based inference using output constraints
Research Assistant

2017 – 2019

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.
· One conference paper [8] (**AAAI**), workshop paper[10] (**AKBC**), and one patent (US20180121807A1) got published.

Part forecasting for airplane parts
Research Assistant

2015 – 2020

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
Research Assistant

2012 – 2015

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

Fall 2014	**Teaching Assistant**, *Multimedia Databases and Data Mining*, Carnegie Mellon University
Spring 2015	**Teaching Assistant**, *Machine Learning (Ph.D. level)*, Carnegie Mellon University

RESEARCH MENTOR

Zhiyang Xu
Former MS student

2021 – current

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
current PhD student

2020 – curent

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
Former MS student

2021 – current

UMass Amherst

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

Eunjeong Hwang
Former MS student

2021

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 2018-2019

Former MS, current Ph.D. student *Carnegie Mellon University*

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

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

OTHER EXPERIENCES

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 Universicty, PA
	Team: Jay-Yoon Lee, Mariya Toneva, Avinava Dubey, Ahmed Hefny, Dan Schwartz, and Ying Yang

TECHNICAL STRENGTHS

Major Language	**Python**, Java, MATLAB
Library & Others	Pytorch, Tensorflow, AllenNLP, Wandb, Keras, Git, LaTex