Seminars and meetings with D.L. researchers from academia and industries

Spring Session 2022:
Young Researchers around the World

Organizer: Prof. **Andrea Asperti**

When: all Fridays at 17 p.m.
Where: on line, via Teams

Speakers: Massimo Caccia (Mila, Univ. of Montreal), Giulia Cantini (Helmholtz Zentrum München), Emanuele la Malfa (Univ. of Oxford), Donato Maragno (Univ. of Amsterdam), Stefano Mezza (UNSW Sidney), Antonio Vergari (Univ. of Edinburgh)

Calendar

April 22 2022

speaker: **Emanuele La Malfa**

affiliation: **University of Oxford**

title: Robustness for Natural Language Processing

abstract: In deep learning, a model is robust when it makes consistent decisions over similar inputs. While this notion is not novel in machine learning (e.g., SVM), deep neural networks are highly susceptible of so called "adversarial attacks", which cause models' misclassifications for tiny variations of the input. In recent years, continuous robustness has been adopted as a standard de facto to certify the adversarial robustness of a model, but it's only recently that the notion has been criticized as "flawed" in other domains such as NLP. In natural language, robustness must be aligned with linguistics, so that a model benefits of invariance to inputs that are not "pathological" and indeed linguistically coherent. Despite harder to implement, semantic robustness is the next "big step" towards models that are more reliable and thus enforce trust in humans.

bio: Emanuele La Malfa graduated in Computer Science and Engineering at the Politecnico of Milan. He worked for two years at IBM, and he is now Ph.D. Student at the University of Oxford, where he studies robustness for Natural Language Processing.

April 29 2022

speaker: Donato Maragno

affiliation: University of Amsterdam

title: Mixed-Integer Optimization with Constraint Learning

abstract: We establish a broad methodological foundation for mixed-integer optimization with learned constraints. We propose an end-to-end pipeline for data-driven decision making in which constraints and objectives are directly learned from data using machine learning, and the trained models are embedded in an optimization formulation. We exploit the mixed-integer optimization-representability of many machine learning methods, including linear models, decision trees, ensembles, and multi-layer perceptrons. The consideration of multiple methods allows us to capture various underlying relationships between decisions, contextual variables, and outcomes. We also characterize a decision trust region using the convex hull of the observations, to ensure credible recommendations and avoid extrapolation. We efficiently incorporate this representation using column generation and clustering. In combination with domain-driven constraints and objective terms, the embedded models and trust region define a mixed-integer optimization problem for prescription generation. We implement this framework as a Python package (OptiCL) for practitioners. We demonstrate the method in both chemotherapy optimization and World Food Programme planning. The case studies illustrate the benefit of the framework in generating high-quality prescriptions, the value added by the trust region, the incorporation of multiple machine learning methods, and the inclusion of multiple learned constraints.

May 6 2022

speaker: Giulia Cantini

affiliation: Helmholtz Zentrum München

title: Deep learning models of post-transcriptional gene regulation

abstract: Since the advent of next generation sequencing techniques (NGS), that allow sequencing of a whole human genome at low-cost and short time, researchers have started to dispose of a growing databank of biological data and at the same time, the need for computational approaches able to mine these data has grown. In recent years, deep learning methods have proved to be able to effectively learn relevant information from vast quantities of data in many fields and have been successfully applied also to genomics and other -omics, where they are able to derive novel biological hypotheses.

In this talk I will show how neural networks, in particular convolutional neural networks (CNNs), can be used to learn patterns of interest from biological sequences to help answer fundamental biological questions. I will focus on the specific task of protein-RNA interactions at post-transcriptional level and I will show a case study, giving a practical example of how the output of these models can be decoded and interpreted.

bio: Giulia Cantini graduated in Computer Science and Intelligent Systems at the University of Bologna and is now a 1st year PhD student in the Computational RNA

Biology lab at Helmholtz Center Munich, working tightly with the Institute for Immunology at the Ludwig Maximilian Universität. Here, she studies deep learning methods for deciphering post-transcriptional regulation focussed on RNA-binding proteins (RBPs) in T cells. Her current work involves researching multimodal models to integrate RNA sequence, structure and epitranscriptomic data to improve binding prediction methods.

May 13 2022

speaker: **Antonio Vergari**

affiliation: **University of Edinburgh**

title: From Variational to Deterministic Autoencoders. Or The Joy Of Density Estimation in Latent Spaces

abstract: This talk is a skeptical inquiry on Variational Autoencoders (VAEs) for generative modeling. While offering a theoretically-backed and popular framework, VAEs still pose unanswered theoretical questions and considerable practical challenges to be learned. In this talk I will first introduce an alternative framework that is simpler, easier to train, and deterministic, yet has many of the advantages of VAEs. I will discuss how substituting stochasticity in VAEs with deterministic but regularized AEs (RAEs) can lead to an equally smooth and meaningful latent space. RAEs offer comparable or better samples than VAEs when equipped with a density estimation phase in the learned latent space. Then, I will then show how this density estimation step can greatly improve sample quality when data are images or structured objects such as molecules. Finally, I will connect RAEs to akin models in the past and very recent literature of generative modeling, tracing further open research directions.

bio: Antonio Vergari is a Lecturer (Assistant Professor) in Machine Learning at the University of Edinburgh. His research focuses on efficient and reliable machine learning in the wild, tractable probabilistic modeling and combining learning with complex reasoning. Previously he was postdoc in the StarAI Lab lead by Guy Van den Broeck at UCLA. Before that he did a postdoc at the Max Planck Institute for Intelligent Systems in Tuebingen in the Empirical Inference Department of Bernhard Schoelkopf. He obtained a PhD in Computer Science and Mathematics at the University of Bari, Italy. He likes to tease and challenge the deep generative models community at large on how we desperately need reliable and efficient models nowadays. To this extent, he organized a series of events: the Tractable Probabilistic Modeling Workshop (ICML2019, UAI2021), the Tractable PRobabilistic Inference MEeting (T-PRIME) at NeurIPS 2019 and presented a series of tutorials on complex probabilistic reasoning and models at UAI 2019, AAAI 2020, ECAI 2020, IJCAI 2021. The last effort of the series will be organizing a Dagsthul Seminar on "New Trends in Tractable Probabilistic Inference" in 2022 with Dr. Priyank Jaini, Prof. Kristian Kersting and Prof. Max Welling.

May 20 2022

speaker: Stefano Mezza

affiliation: UNSW Sydney

title: Natural Language Understanding for Open-ended Dialogue Systems

abstract: Spoken Dialogue Systems are becoming increasingly popular due to the advancement of speech recognition and speech synthesis technologies. Although recent dialogue agents have made remarkable progress in their understanding capabilities, we are still far away from achieving truly human-like conversations with machines, with most commercial conversational agents being limited to robotic, task-based interactions. This talk will give a brief overview on how Spoken Dialogue Systems are built and describe their key components. We will then shift our focus on Natural Language Understanding (the comprehension by computers of the structure and meaning of human language) and examine a state-of-the-art neural network architecture to perform this task for open-ended Dialogue Systems.

bio: Stefano Mezza is a PhD student at the University of New South Wales in Sydney, Australia. He got his BSc in Computer Science at the University of Bologna and then specialised in Machine Learning and Natural Language Processing at the Universities of Trento and Edinburgh. He collaborated with various private institutions, such as Amazon Alexa, Emotech Labs and Babylon Health, before deciding to pursue a career in academia at UNSW. His research focuses on Natural Language Understanding and Dialogue Act Tagging, and how to incorporate emotional awareness, consistency and pragmatics in automated dialogue agents.

May 27 2022

speaker: Massimo Caccia

affiliation: Mila, Univ. of Montreal

title: Task-agnostic Continual Reinforcement Learning: When Upper Bounds Cease to be

abstract: Through learning the invariants of multiple tasks and environments, one can hope that the general intelligence seen in animals and humans will eventually emerge in artificial agents. Multi-task learning (MT), in which all tasks and environments can be simultaneously spawned and jointly learned, is often an impractical assumption. For this reason, the field of Continual Learning (CL) has surfaced, in which artificial agents experience tasks in sequence. Another restrictive assumption is task awareness in which an oracle provides task labels and task boundaries, enabling the agent with full observability of the environment state. MT and task awareness are usually considered soft upper bounds for sequential and task-agnostic learning.

We study the challenging setting of task-agnostic continual reinforcement learning (TACRL) in which RL's standard nonstationarities, i.e. in the observation transitions, are compounded with the task-agnostic CL's ones, i.e., the latent sequence of tasks that the agents need to learn. A simple approach to TACRL consists in augmenting a model-free RL algorithm with a recurrent mechanism to handle

partial observability as well as a replay mechanism for CL purposes. We refer to this baseline as replay-based recurrent reinforcement learning (3RL).

We find surprising occurrences of 3RL matching and overcoming the MT and task-awareness soft upper bounds. We further lay out hypotheses that could explain this inflection point of continual and task-agnostic learning research. Our hypotheses are empirically tested in continuous control tasks via a large-scale study of the popular multi-task and continual-learning benchmark Meta-World. By comparing different training regimes and analyzing the statistics of the gradients used to learn, we find some evidence that 3RL's outperformance stems from its ability to quickly infer how new tasks relate with the previous ones, enabling skill transfer.

bio: Massimo Caccia is a fifth-year Ph.D. student at the Quebec Artificial Intelligence Institute (Mila) under the supervision of Laurent Charlin and soon-to-be intern at DeepMind in the Continual Learning team led by Marc'Aurelio Ranzato. Formerly, He has interned at Amazon, ElementAI (now ServiveNow) and Spotify Research.

He is interested in algorithms able to accumulate transferable knowledge or skills enabling better and faster generalization to future tasks. Accordingly, his research topics lie in continual learning and meta-learning. Recently, He developed a particular interest in the idea of composing existing skills to learn new ones quickly. This looks like the natural appeal of continual learning as it can allow curriculum learning to deliver on its promises, as well as propel reinforcement learning and robotics forward. Hence, the current focus on continual RL problems.