

Róbert Csordás

e-mail: robert@idsia.ch

tel: +41 76 4961457

website: <https://robertcsordas.github.io/>

EDUCATION

IDSIA <http://idsia.ch>, Lugano, Switzerland

PhD student

2018 April - 2023 September

Supervised by Prof. Jürgen Schmidhuber. Worked on systematic generalization.

Budapest University of Technology and Economics, Budapest, Hungary

Electrical Engineering. MSc (grad. 2015) and BSc (grad. 2012). Grade: excellent.

WORK

EXPERIENCE

IDSIA <http://idsia.ch>

2023 October -

Postdoctoral Researcher

Lugano, Switzerland

Working on systematic generalization.

DeepMind - <https://www.deepmind.com>

2022 June - 2022 October

Research Scientist Intern

London, United Kingdom

I worked on graph neural networks, improving generalization on algorithmic problems and external memory for Transformers.

AIMotive (formerly AdasWorks) - <https://aimotive.com>

2015-2018

AI Research Scientist

Budapest, Hungary

Worked on deep neural networks for self driving cars.

- Monocular depth prediction using neural networks.
- Neural stereo matching - predicting robust depth map with neural network.
- Recurrent network research - Convolutional LSTMs for stabilizing detections, free space detection, etc.
- Object detection, semantic segmentation

Hungarian Academy of Sciences - Institute for Computer Science and

Control - <https://www.sztaki.hu>

2015

Software Engineer

Budapest, Hungary

Worked on classical computer vision projects. For example:

- Detecting objects thrown over the fence; detecting human leaving a car.
- Autonomous forklift control system.

Innomed Medical Inc. - <http://innomed.hu>

2007 - 2015

Embedded Software/Hardware Engineer

Budapest, Hungary

- Designed the software architecture of Linux based patient monitor (C++, QT).
- Maintained the software of the InnoCare-S patient monitor (C++).
- Wrote low level hardware drivers for InnoCare-T12.

PUBLICATIONS

Róbert Csordás, Piotr Piękos, Kazuki Irie, Jürgen Schmidhuber: **SwitchHead: Accelerating Transformers with Mixture-of-Experts Attention** - We present an MoE attention that can match the performance of parameter-matched dense models.
arXiv preprint <https://arxiv.org/abs/2312.07987>

Róbert Csordás, Kazuki Irie, Jürgen Schmidhuber: **Approximating Two-Layer Feedforward Networks for Efficient Transformers** - We present an improved

MoE that can match the performance of parameter-matched dense models.
EMNLP Findings 2023 <https://arxiv.org/abs/2310.10837>

Kazuki Irie*, Róbert Csordás*, Jürgen Schmidhuber: **Topological Neural Discrete Representation Learning à la Kohonen** - We show that VQ used in VQ-VAEs is a special case of SOMs, which are more robust and converge faster.
arXiv preprint <https://arxiv.org/abs/2302.07950>

Anian Ruoss, Grégoire Delétang, Tim Genewein, Jordi Grau-Moya, Róbert Csordás, Mehdi Bennani, Shane Legg, Joel Veness: **Randomized Positional Encodings Boost Length Generalization of Transformers** - We propose randomized, ordered positional encodings to improving length generalization on algorithmic tasks.
ACL 2023 <https://arxiv.org/abs/2305.16843>

Róbert Csordás, Kazuki Irie, Jürgen Schmidhuber: **CTL++: Evaluating Generalization on Never-Seen Compositional Patterns of Known Functions, and Compatibility of Neural Representations** - We extend the CTL dataset to test systematicity and show how NNs develop incompatible representations and fail to generalize.
EMNLP 2022 <https://arxiv.org/abs/2210.06350>

Borja Ibarz, Vitaly Kurin, George Papamakarios, Kyriacos Nikiforou, Mehdi Bennani, Róbert Csordás, Andrew Dudzik, Matko Bošnjak, Alex Vitvitskyi, Yulia Rubanova, Andreea Deac, Beatrice Bevilacqua, Yaroslav Ganin, Charles Blundell, Petar Veličković: **A Generalist Neural Algorithmic Learner** - We show that graph neural networks are capable of learning many algorithms together and they can generalize to larger problem instances.
LoG 2022 <https://arxiv.org/abs/2209.11142>

Kazuki Irie*, Róbert Csordás*, Jürgen Schmidhuber: **The Dual Form of Neural Networks Revisited: Connecting Test Time Predictions to Training Patterns via Spotlights of Attention** - We investigate dual form representations of NNs to get insights into how their behaviour depends on the training samples.
ICML 2022 <https://arxiv.org/abs/2202.05798>

Kazuki Irie, Imanol Schlag, Róbert Csordás, Jürgen Schmidhuber: **A Modern Self-Referential Weight Matrix That Learns to Modify Itself**
ICML 2022 <https://arxiv.org/abs/2202.05780>

Róbert Csordás, Kazuki Irie, Jürgen Schmidhuber: **The Neural Data Router: Adaptive Control Flow in Transformers Improves Systematic Generalization** - We propose to improve data routing in Transformers by gating and geometric attention, achieving systematic generalization on algorithmic tasks.
ICLR 2022 <https://arxiv.org/abs/2110.07732>

Róbert Csordás, Kazuki Irie, Jürgen Schmidhuber: **The Devil is in the Detail: Simple Tricks Improve Systematic Generalization of Transformers** - We significantly improve the systematic generalization of Transformers on a variety of systematic generalization datasets using simple tricks.
EMNLP 2021 <https://arxiv.org/abs/2108.12284>

Kazuki Irie, Imanol Schlag, Róbert Csordás, Jürgen Schmidhuber: **Going Beyond Linear Transformers with Recurrent Fast Weight Programmers** - We explore the recurrent Fast Weight Programmers (FWPs), which exhibit advantageous

properties of both Transformers and RNNs.

NeurIPS 2021

<https://arxiv.org/abs/2106.06295>

Róbert Csordás, Sjoerd van Steenkiste, Jürgen Schmidhuber: **Are Neural Nets Modular? Inspecting Functional Modularity Through Differentiable Weight Masks** - We develop a method for analyzing emerging functional modularity in neural networks based on differentiable weight masks and use it to point out important issues in current-day neural networks.

ICLR 2021

<https://openreview.net/forum?id=7uVcpu-gMD>

Róbert Csordás, Jürgen Schmidhuber: **Improving Differentiable Neural Computers Through Memory Masking, De-allocation, and Link Distribution Sharpness Control** - Addresses 3 different issues with the original DNC architecture. Also proposes a new, better content-based lookup mechanism.

ICLR 2019

<https://openreview.net/forum?id=HyGEM3C9KQ>

Róbert Csordás, László Havasi, and Tamás Szirányi: **Detecting objects thrown over fence in outdoor scenes** - A new technique for detecting objects thrown over a critical area of interest in a video sequence made by a monocular camera.

VISAPP 2015

<http://goo.gl/ZDkk4g>

WORKSHOP PAPERS

Kazuki Irie, Imanol Schlag, Róbert Csordás, Jürgen Schmidhuber: **Improving Baselines in the Wild**

NeurIPS 2021 DistShift

<https://openreview.net/forum?id=9vx0rkNTs1x>

HIGH SCHOOL PUBLICATIONS

CallTheTux - Development of CallTheTux, a universal GSM stack for Linux.

Petnica Papers, 2007

<https://goo.gl/QTcy5U>

RealVM - Development of a new type of virtual machine which would allow parallel execution and fast switching between different operating systems.

Petnica Papers, 2006

<https://goo.gl/8TNhf5>

PrologAPI - Enabling the usage of Prolog constructs from C++.

Petnica Papers, 2005

<https://goo.gl/KpV3sF>

PATENTS

Róbert Csordás, Ágnes Kis-Benedek, Balázs Szalkai: **Method and Apparatus for Generating a Displacement Map of an Input Dataset Pair** - A neural network based method for fast and robust stereo matching for depth map generation.

US10380753

<https://pimg-fpiw.uspto.gov/fdd/53/807/103/0.pdf>

TECHNICAL STRENGTHS

Python, PyTorch, TensorFlow, C, C++, CUDA, OpenCV, Algorithms, Linux, JavaScript, Bash, Matlab, Assembly

OTHER SKILLS

Machine learning frameworks: PyTorch, JAX, TensorFlow, Torch

Parallel programming: CUDA, numba

Electronics: KiCAD, Eagle, PIC, PIC32, AVR, AVR32, ARM, X MOS, Xilinx

Databases: MySQL, MongoDB, Sphinx search

JavaScript technologies: NodeJS, jQuery

Mobile development: Android, iOS (Swift)

Operating systems: Linux, OS X, Windows

Markup languages: L^AT_EX, XML, Markdown

**HOBBY
PROJECTS**

MobileECG II - <https://github.com/robertcsordas/MobileECG-II> 2014 - 2016
Open source Holter ECG. Designed the schematic diagram and the firmware.

engineerjs.com - <http://engineerjs.com> 2013 - 2015
Extendable online computing environment for engineers, with physical quantity, complex numbers and linear algebra support.

LANGUAGES

Hungarian (native); English, Serbian (fluent); Italian (intermediate); German (beginner)