Imanol Schlag

CONTACT
INFORMATION

IDSIA: The Swiss AI Lab
D4.05, USI East-Campus
Via la Santa 1, 6962 Lugano, Switzerland

phone: +41 79 347 52 42
e-mail: imanol.schlag@gmail.com

RESEARCH
INTERESTS

I am passionate about exploring the intersection of Machine Learning and Artificial Intelligence through Connectionist Models. My research focus is on creating neural networks that exhibit improved generalization capabilities. My expertise lies in the area of language modelling, where I have published novel model architectures that address the limitations of Transformer models and effectively integrated attention and recurrence in scalable ways. At present, I am dedicated to advancing the capabilities of large language models by pursuing innovative approaches beyond mere scaling, aimed at unlocking more sophisticated behaviour.

CURRENT
ACADEMIC
APPOINTMENTS

Doctoral Assistant, IDSIA - The Swiss AI Lab   **September 2016 to present**
Istituto Dalle Molle di Studi sull'Intelligenza Artificiale
Università della Svizzera italiana
Faculty of Informatics

EDUCATION

Università della Svizzera italiana, Lugano, Switzerland

PhD, Artificial Intelligence and Machine Learning, candidate

- Adviser: Professor Jürgen Schmidhuber
- Area of Study: Artificial Intelligence and Machine Learning

University of St Andrews, St Andrews, Scotland

MSc, Artificial Intelligence, August 2016

- With Distinction
- Thesis Topic: *Face Recognition from Ancient Roman Coins*
- Adviser: Professor Ognjen Arandjelović

University of Applied Sciences and Arts Northwestern Switzerland, Brugg, Switzerland

BSc, Computer Science, August 2015

- With specialisation in *Information Processing and Visualization*
- Thesis Topic: *Face Similarity - Finding Lookalikes from Images*

Swiss Armed Forces Special Forces Training Center, Isone, Switzerland

- Basic Training, 2010
- Non-commissioned Officer School, 2011

PUBLICATIONS

I. Schlag, S. Sukhbaatar, A. Celikyilmaz, W. Yih, J. Weston, J. Schmidhuber, X. Li. Large Language Model Programs. *In preparation*.

D. Hutchins*, **I. Schlag***, Y. Wu, E. Dyer, B. Neyshabur. Block Recurrent Transformer. Neural Information Processing Systems (NeurIPS), 2022.

A. Lewkowycz, A. Andreassen, D. Dohan, E. Dyer, H. Michalewski, V. Ramasesh, A. Slone, C. Anil, **I. Schlag**, T. Gutman-Solo, Y. Wu, B. Neyshabur, G. Gur-Ari, V. Misra. Solving Quantitative Reasoning Problems with Language Models. Neural Information Processing Systems (NeurIPS), 2022.

I. **Schlag**, J. Schmidhuber. Augmenting Classic Algorithms with Neural Components for Strong Generalisation on Ambiguous and High-Dimensional Data. Advances in Programming Languages and Neurosymbolic Systems Workshop (NeurIPS), 2021.

K. Irie, **I. Schlag**, R. Csordás, J. Schmidhuber. A Modern Self-Referential Weight Matrix That Learns to Modify Itself. Deep RL Workshop (NeurIPS), 2021.

K. Irie, **I. Schlag**, R. Csordás, J. Schmidhuber. Improving Baselines in the Wild. Workshop on Distribution Shifts: Connecting Methods and Applications (NeurIPS), 2021.

K. Irie*, **I. Schlag***, R. Csordás, J. Schmidhuber. Going Beyond Linear Transformers With Recurrent Fast Weight Programmers. Neural Information Processing Systems (NeurIPS), 2021.

I. **Schlag***, K. Irie*, J. Schmidhuber. Linear Transformers are Secretly Fast Weight Programmers. In Proc. Int. Conf. on Machine Learning (ICML), 2021.

I. **Schlag**, T. Munkhdalai, J. Schmidhuber. Learning Associative Inference Using Fast Weight Memory. In Int. Conf. on Learning Representations (ICLR), 2021.

I. **Schlag**, P. Smolensky, R. Fernandez, N. Jojic, J. Schmidhuber, J. Gao. Enhancing the Transformer With Explicit Relational Encoding for Math Problem Solving. Preprint arXiv: 1910.06611, 2019.

I. **Schlag** and J. Schmidhuber. Learning to Reason with Third-Order Tensor Products. Neural Information Processing Systems (NeurIPS), 2018.

I. **Schlag** and J. Schmidhuber. Gated Fast Weights for On-The-Fly Neural Program Generation. Workshop on Meta-Learning (NeurIPS), 2017.

I. **Schlag** and O. Arandjelovic. Ancient Roman Coin Recognition in the Wild Using Deep Learning Based Recognition of Artistically Depicted Face Profiles. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2017.

TEACHING EXPERIENCE	**Università della Svizzera italiana**, Lugano, Switzerland	**2017 to 2021**

Teaching assistant

- Machine Learning Fall 17/18
- Deep Learning Lab Fall 17/18, Fall 18/19, Fall 19/20
- Graph Deep Learning Spring 20/21

Course development

- Assisted the development of the first version of the Deep Learning Lab for Fall 17/18

Swiss Armed Forces, KSK, Gren Bat 30/2, Isone, Switzerland   **2012 to 2019**

Military instructor and squad leader in the Special Forces Command (KSK)

- A yearly and mandatory 4 week repetition course

PROFESSIONAL EXPERIENCE	**Meta AI**, Menlo Park, California, USA	**July to December 2022**

Research internship at FAIR with Xian Li and Jason Weston

Google Research, Mountain View, California, USA   **September to February 2022**

Research internship in the Blueshift Team with Behnam Neyshabur (remote)

Microsoft Research, Redmond, Washington, USA   **June to September 2019**

Research internship with Paul Smolensky

Basler Kantonalbank, Basel, Switzerland　　　　　　　　**September 2006 to June 2010**
　　　　Apprentice in informatics

Awards　　　　　**NVAIL Pioneering Research Award**
　　　　　　• For *Learning to Reason with Third-Order Tensor Products*. Received at NeurIPS, 2018.

　　　　　　University of St Andrews
　　　　　　• Medal for the best dissertation in Computer Science, 2016

Reviewing　　　　NeurIPS 19/20/21, ICML 20/21, ICLR 20/21/22