Integrated Reinforcement Learning

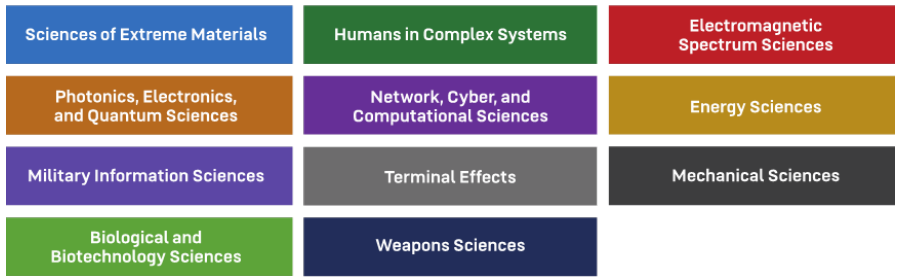
Derya Cansever

US Army Research Office

Army Research Office

Army's primary collaborative link to the World-Wide Scientific Community

ARO has 45 Programs focused on Science Organized in 11 Competencies (areas)



Multi-Agent Network Control Program

The objective of the Multi-Agent Network Control program is to establish the physical, mathematical and information processing foundations for the control of complex dynamic networks with possibly multiple controllers that may operate using different information sets.

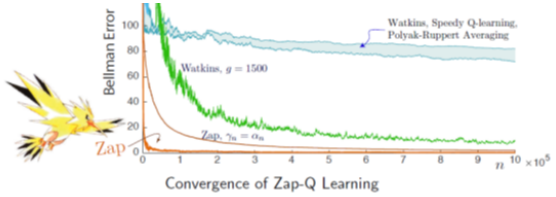
- Distributed and Time-Varying Control of Networked Systems
- Data Driven Control and learning
- Control of Quantum Systems and novel applications of control theory

ARO Grant Types

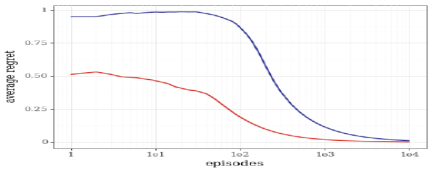
Award Type	Target	Funding
Single Investigator (SI)	Single-laboratory projects	~$141K/year for ~3.4 years avg*
Short Term Innovative Research (STIR)	Very high-risk pilot projects	$60K for 9 mo.
Early Career Awards (formerly Young Investigator Program)	Early-career PIs	$120K/year for 3 years
Conferences / Workshops / Symposia	Academic State of Science	$10K–$30K
Presidential Early Career Award for Scientists and Engineers (PECASE)	Promising future leaders	$200K/year for 5 years
Defense University Research Instrumentation Program (DURIP)	Instrumentation	$200K/year average FY22
Multidisciplinary University Research Initiative (MURI)	Large multidisciplinary programs	~$1.25M/year up to 5 years
Historically Black College/University and Minority Institution (HBCU/MI)	Minority serving institutions	~$140K/year for 3 years
Small Business Technology Transfer (STTR)	Multi-phase awards bridging academia & industry	$150K (6 mo.) to $1M (24 mo.)
Small Business Innovative Research (SBIR)	Multi-phase research for industry transition	$150K (6 mo.) to $1M (24 mo.)

Reasons for Celebration of Reinforcement Learning

- Performance way beyond human capabilities with an algorithm based solely on reinforcement learning, without human data, guidance or domain knowledge beyond game rules.



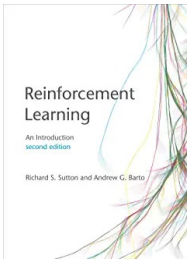
- On solid theoretical foundation
- Significant progress in RL algorithm performance.

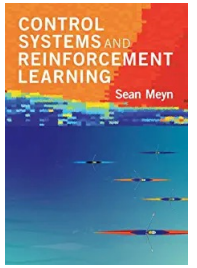


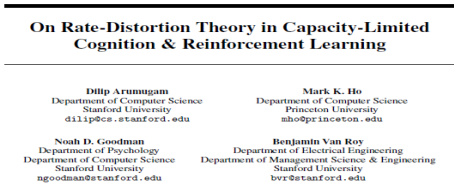
- Encouraging progress in optimal data selection for RL.

Mastering the game of Go without human knowledge

David Silver[1]*, Julian Schrittwieser[1]*, Karen Simonyan[1]*, Ioannis Antonoglou[1], Aja Huang[1], Arthur Guez[1], Thomas Hubert[1], Lucas Baker[1], Matthew Lai[1], Adrian Bolton[1], Yutian Chen[1], Timothy Lillicrap[1], Fan Hui[1], Laurent Sifre[1], George van den Driessche[1], Thore Graepel[1] & Demis Hassabis[1]







On Rate-Distortion Theory in Capacity-Limited Cognition & Reinforcement Learning

Dilip Arumugam
Department of Computer Science
Stanford University
dilip@cs.stanford.edu

Mark K. Ho
Department of Computer Science
Princeton University
mho@princeton.edu

Noah D. Goodman
Department of Psychology
Department of Computer Science
Stanford University
ngoodman@stanford.edu

Benjamin Van Roy
Department of Electrical Engineering
Department of Management Science & Engineering
Stanford University
bvr@stanford.edu

And many more amazing fundamental contributions and applications of RL

A Reason for More Work in Reinforcement Learning

Human Brain: 20 W

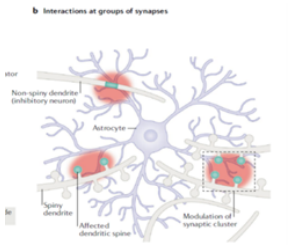
Google AlphaZero TPUs: 1 MW

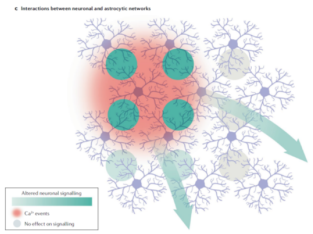
- 4-5 order of magnitudes difference in power requirements.
- Does the difference in performance commensurate with power consumption?
- Study of the brain may inspire further advances in RL.

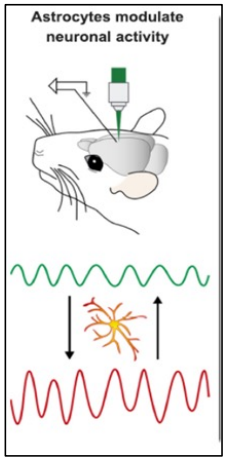
Novel Mechanisms of Neuro-Glia Bio-Computation and Reinforcement Learning

ARO initiated MURI Topic performed by

Massachusetts Institute of Technology · University of Minnesota · Caltech

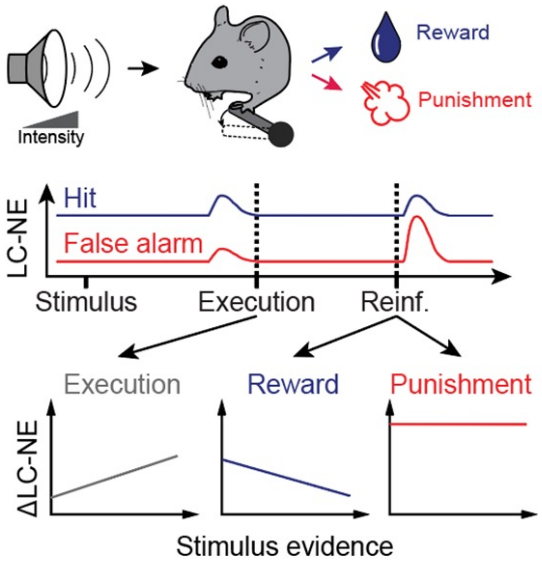






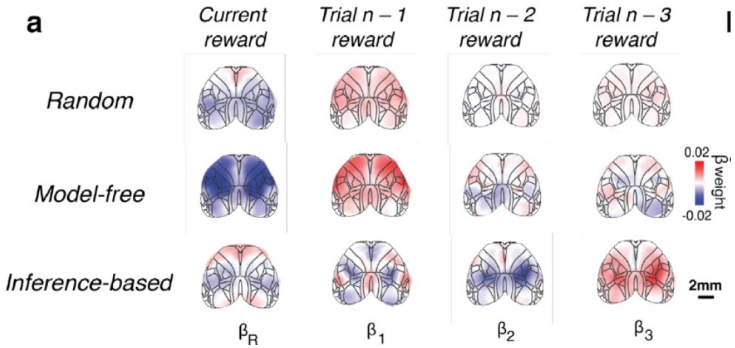
- Astrocytes are believed to be essential to RL, through temporal calcium dynamics, and their interaction with synapses, neurons and neuro-modulatory systems

Experiments with Rodents

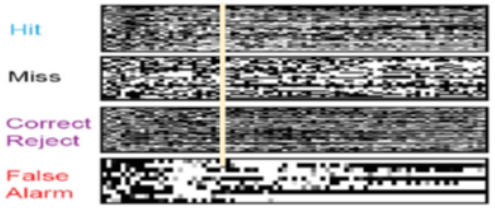


Observations

- Rodents exhibit a mixture of both model-free and inference (model) based strategies in RL.
- Both strategies may co-exist in the same task.
- Inference-based behavior increases with training.
- Learning occurs at different parts of the brain simultaneously.



- Astrocyte signals consistent with control and coordination functions of neuron activities



In nature, RL appears to be distributed, hierarchical, multi-mode and integrated.