Integrated Reinforcement Learning

Derya Cansever
US Army Research Office
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

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

And many more amazing fundamental contributions and applications of RL
A Reason for More Work in Reinforcement Learning

- 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

MIT  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
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.