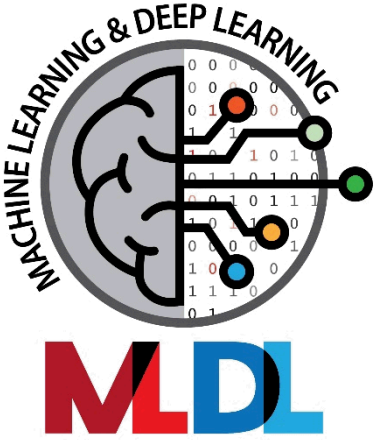
SAND2017-8470C



MLDL

Machine Learning and Deep Learning Conference 2017

In Situ Training Environment for Autonomous Cyber*

Mike Stickland - Emulytics™ Initiatives (5824)

Kasimir Gabert – Cyber Initiatives (5838)

John Jacobellis – Systems Security Research (5827)

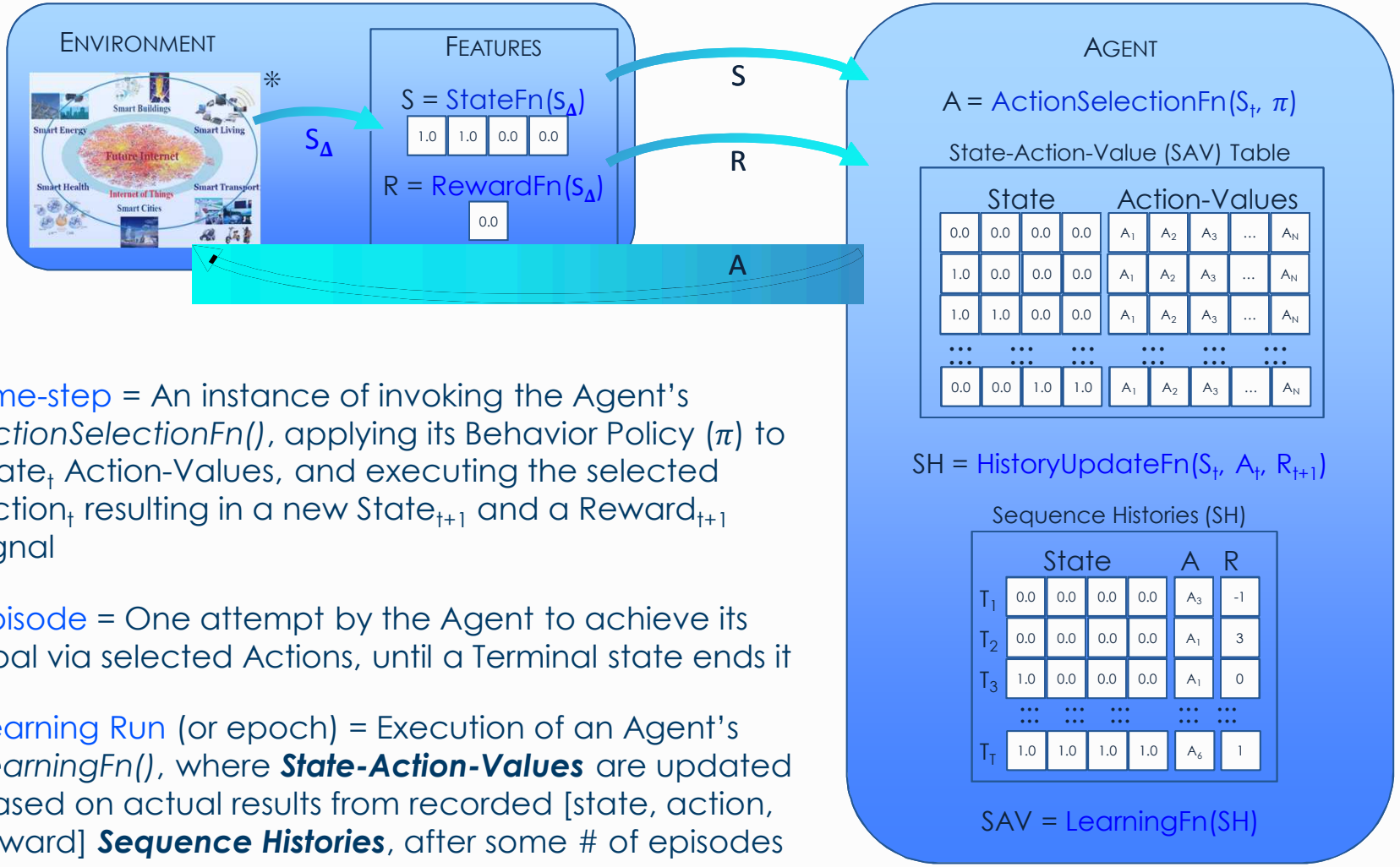
* Funding provided by Sandia National Labs, via 5820's FY17 Proof of Concept Tournament

First Things

- Our Definition of Autonomous Cyber:
  - Emulated or Physical Cyber Environments (real-world)
  - Cyber Agents are Self-learning (from experience, via trial-and-error) and their Actions can Effect the State of the Cyber Environment (i.e. Reinforcement Learning)

- Our Hypothesis: Experimentation & Training of real-world Autonomous Cyber systems is different, than for ex. Self-Driving Vehicles, because the Cyber Environment…
  - will need to be reset to a known state before each training episode
  - will need to be programmatically perturbed (with changes to network topology, device configurations, etc.) between some training episodes

Reinforcement Learning



Time-step = An instance of invoking the Agent's *ActionSelectionFn()*, applying its Behavior Policy (π) to $State_t$ Action-Values, and executing the selected $Action_t$ resulting in a new $State_{t+1}$ and a $Reward_{t+1}$ signal

Episode = One attempt by the Agent to achieve its goal via selected Actions, until a Terminal state ends it

Learning Run (or epoch) = Execution of an Agent's *LearningFn()*, where ***State-Action-Values*** are updated based on actual results from recorded [state, action, reward] ***Sequence Histories***, after some # of episodes

* Image credit to Ameer Nasrallah

Demonstrated Proof-of-Concept

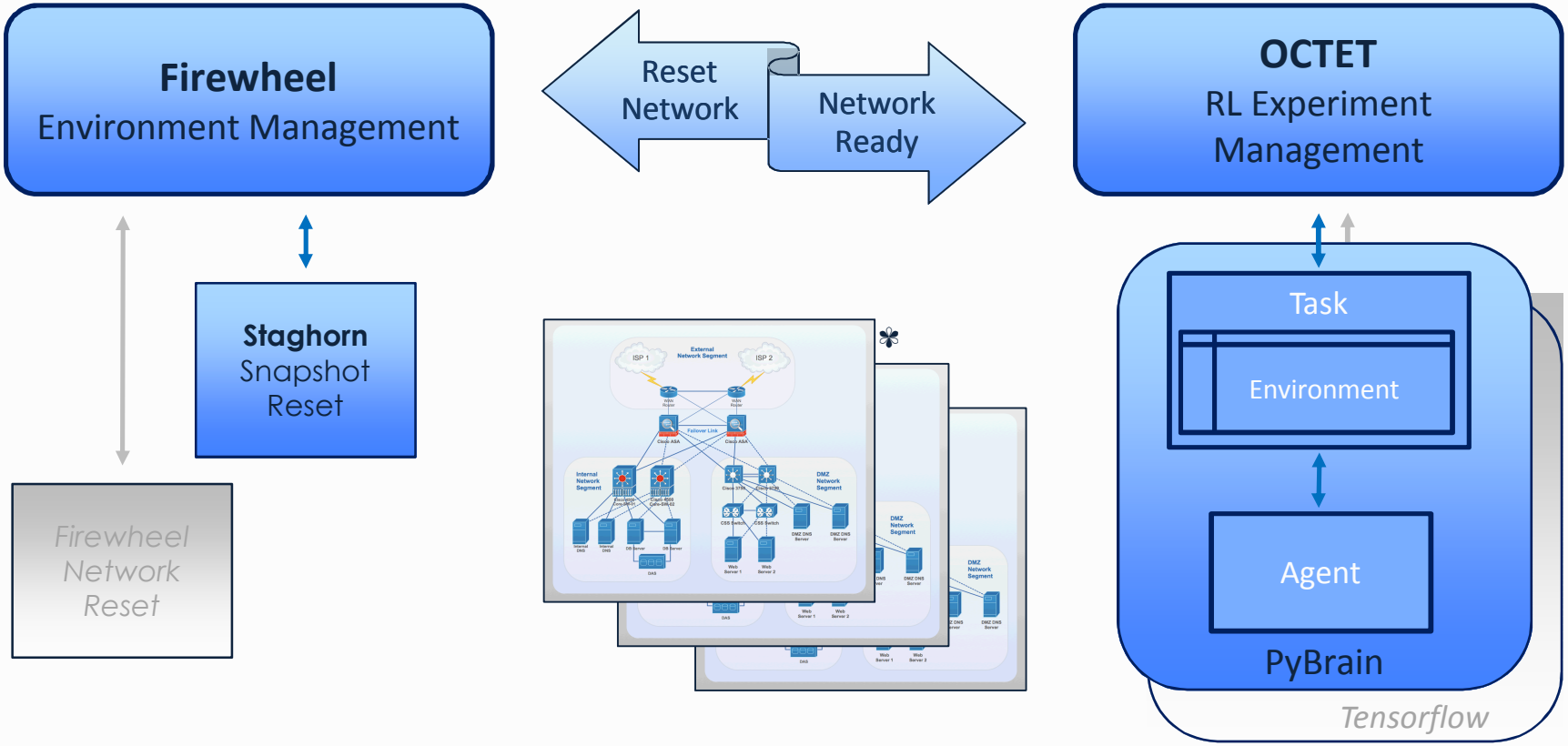
- An integrated, automated platform for training autonomous cyber agents, in situ, where learning episodes can be run repeatedly (many times over and over) in automatically reset, realistic cyber network environments

- Integrates three of Sandia's existing, unique, and advanced Emulytics™ technologies (Firewheel, Staghorn, and OCTET) and an open source reinforcement learning library (PyBrain✧)

Firewheel	Deployment & provisioning of emulated Cyber Network Environments (rapidly, with high-fidelity, at scale)
Staghorn	System-wide Snapshot/modify/resume (very fast, Firewheel add-on)
OCTET	Emulation of Cyber Threat Actors (automated, orchestrated attacks)

✧ Tom Schaul, Justin Bayer, Daan Wierstra, Sun Yi, Martin Felder, Frank Sehnke, Thomas Rückstieß, Jürgen Schmidhuber (2010). PyBrain. Journal of Machine Learning Research, pp. 743--746, vol. 11, 2010. Available from: http://pybrain.org

Autonomous Cyber Training Platform†



† Proof-of-Concept Version 0.1 ✾ Image credit to CS Odessa

Next Things

- Conduct further proof-of-concept activities
  - Demonstrate target policy convergence for (simple) learning tasks
  - Develop time/resource estimates for running experiments using various network scales, learning task complexities, etc.

- Identify potential partners
  - Researchers interested in using this capability / collaboration
  - Sponsors willing to fund platform extension, maturation
  - Funding for fundamental and applied Autonomous Cyber R&D

- Discover/define the science of Autonomous Cyber
  - e.g. Principles of Representing Cyberspace as Environment Features for Reinforcement Learning Algorithms (i.e. Feature Engineering)
    - for Operational Red Team, Cyber Defense, and Assessment tasks, etc.

Email: mgstick@sandia.gov