
Robust Multi-Agent Reinforcement Learning for Self-Adaptive Systems

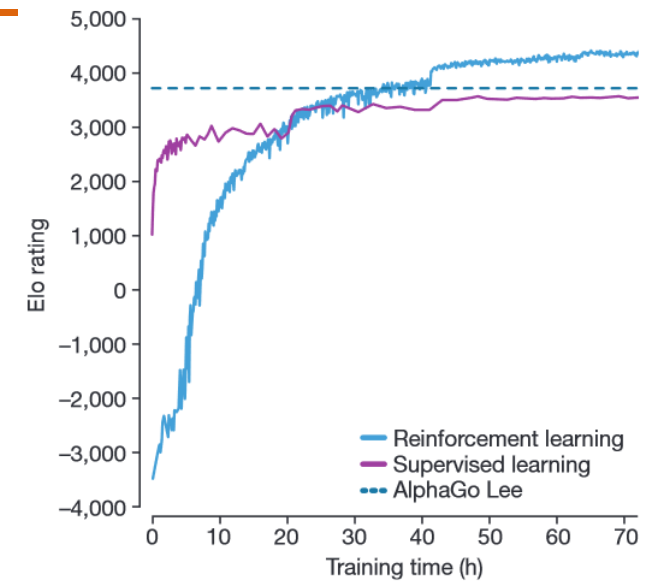
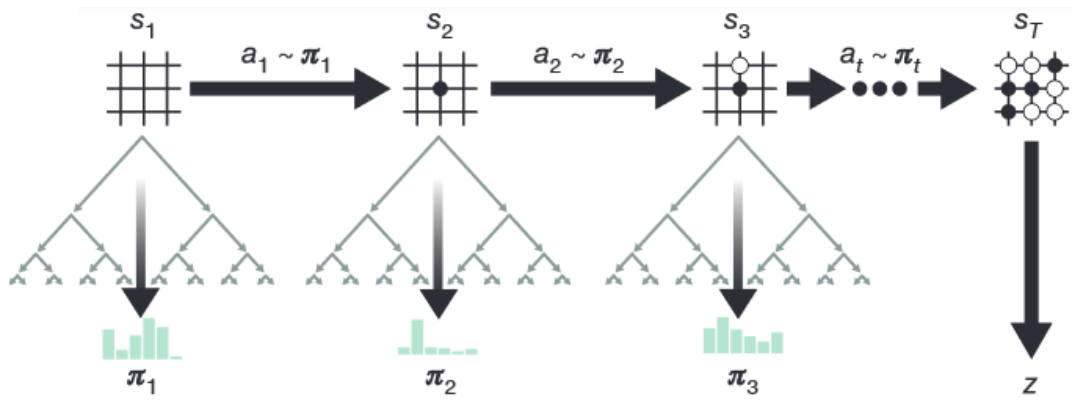
Christian M. Adriano

Prof. Dr. Holger Giese

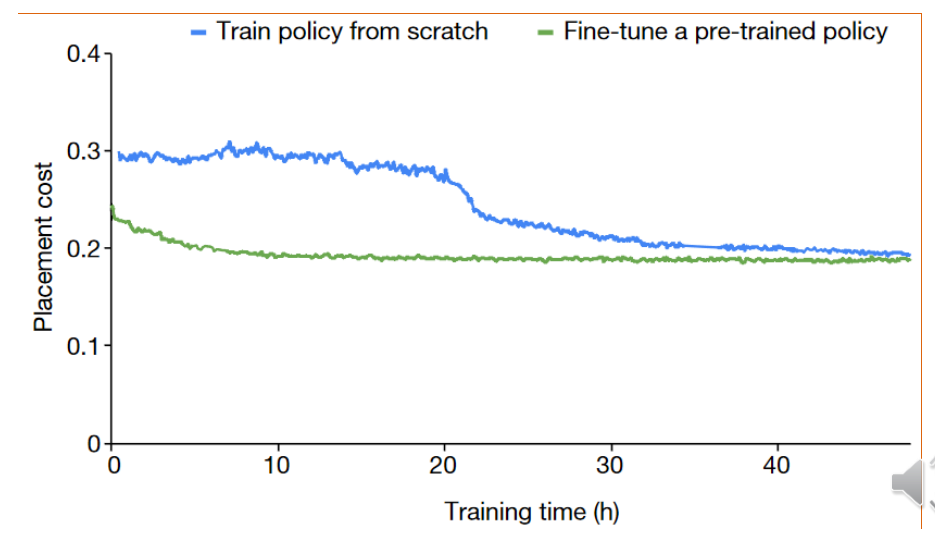
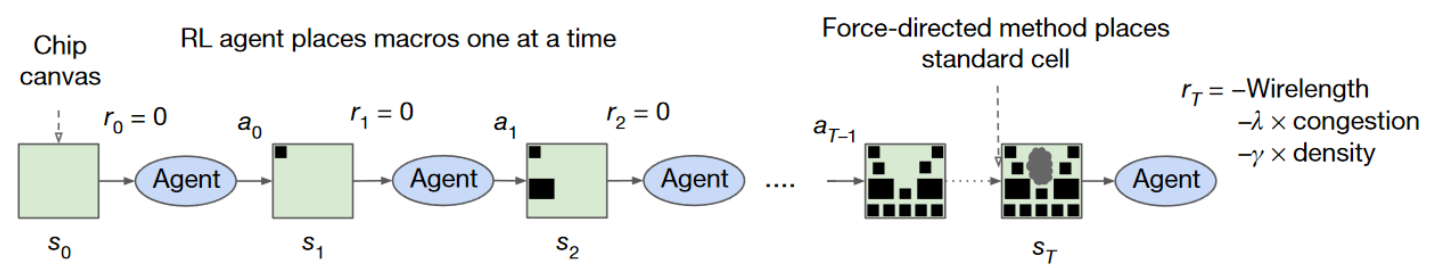
System Analysis and Modeling Group

Context – The Progress of Reinforcement Learning

Alpha Go-Zero learned to win without supervision, only by self-play **[DeepMind 2017]**



RL agent designed optimal layouts for TPU chip circuit **[Google 2021]**

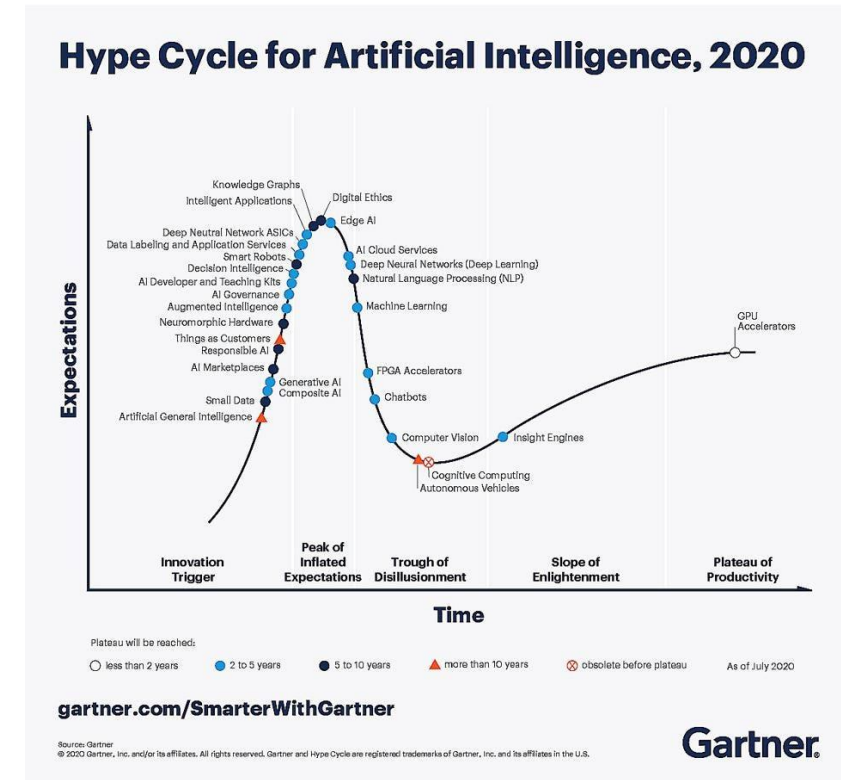


However

AI systems are not being deployed

- **55%** of companies surveyed haven't deployed a machine learning model [**Algorithmia 2020**]
- **72%** that began AI pilots before 2019 haven't deployed a single system yet [**Capgemini 2020**]

Why? Current models cannot adapt to more complex and evolving realities - adversarial environment



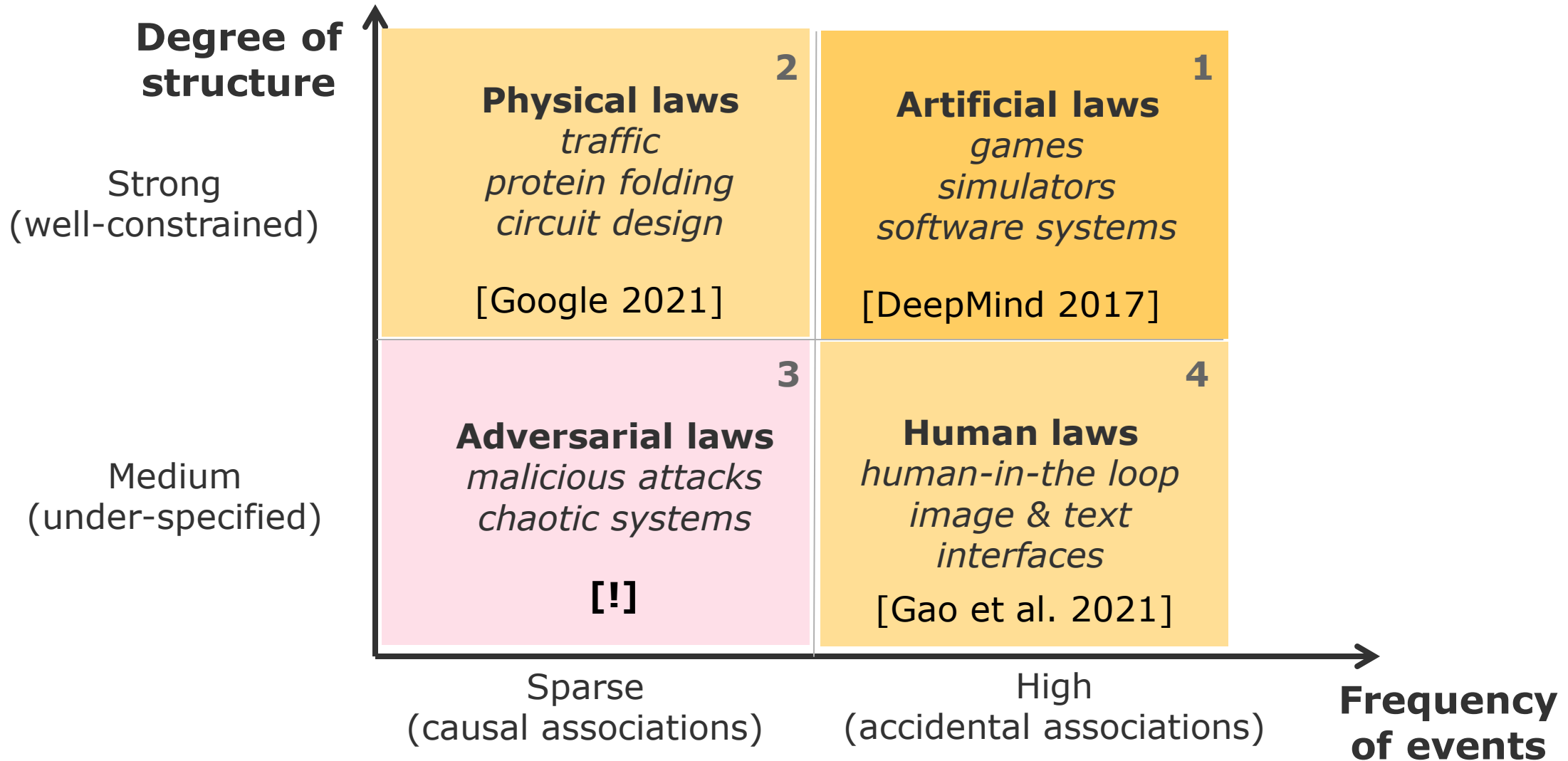
[Gartner 2020]

Problem? Lack of Robustness in AI Systems

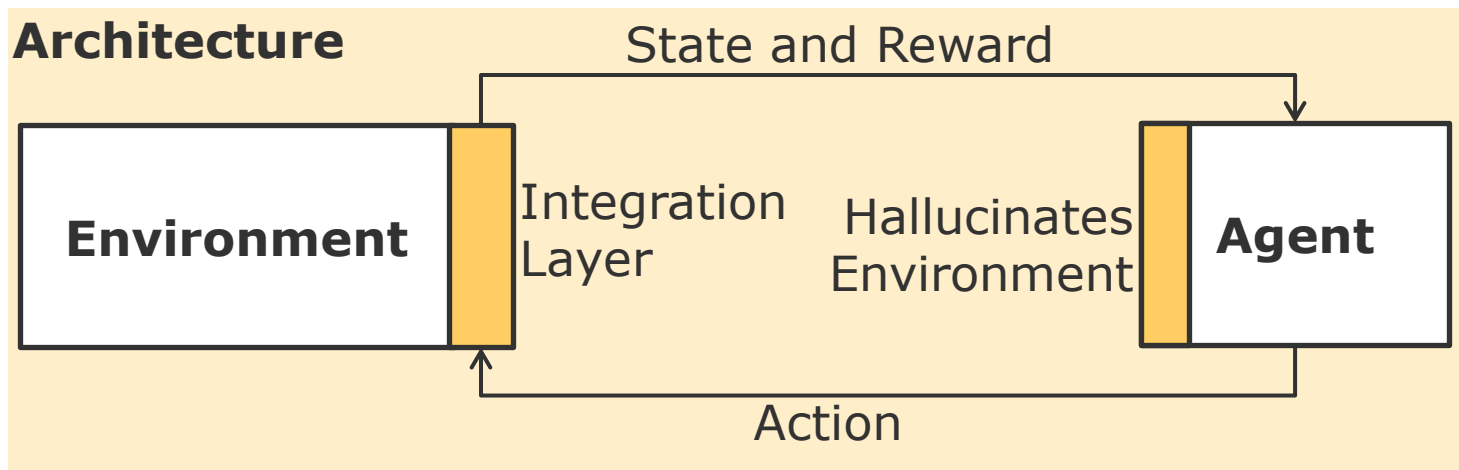
[Jordan 2019], [D'Amour et al. 2020]



Nature of the Problem - Structure vs Frequency



Architecture, Case Study and Project Goals



Project Goals & Techniques

1. Reward sparsity
(large state space)

*off-fine & online RL
robust MDPs*

2. Distribution shift
(unsafe actions)

*malicious
multi-agents*

3. Concept drift
(accidental associations)

*causal discovery
reward shaping*

4. Domain shift
(unsafe states)

*domain adaptation
meta-learning*

Incremental



Case Study

Platform: E-Commerce for online shops

States: Component failure modes

Actions:

- Self-Repair
- Self-Optimization [Vogel et al. 2018]

Technology stack:

PyTorch/Tensor Flow, Open AI Gym, GraphQL, etc.

References



- [**Algorithmia 2020**] The State of Enterprise ML, https://info.algorithmia.com/hubfs/2019/Whitepapers/The-State-of-Enterprise-ML-2020/Algorithmia_2020_State_of_Enterprise_ML.pdf
- [**Capgemini 2020**] The AI-Powered Enterprise, <https://www.capgemini.com/gb-en/research/the-ai-powered-enterprise/>
- [**Chen et al. 2021**] Chen, L., et al., 2021, Decision Transformer: Reinforcement Learning via Sequence Modeling, <https://arxiv.org/pdf/2106.01345.pdf>
- [**D'Amour et al. 2020**] D'Amour, A. et al., 2020, Underspecification presents challenges for credibility in modern machine learning, <https://arxiv.org/pdf/2011.03395.pdf>
- [**DeepMind 2017**] Silver, D., et al., 2017, Mastering the game of Go without human knowledge, *Nature*, pp. 354-371, doi:10.1038/nature24270. <https://deepmind.com/blog/article/alphago-zero-starting-scratch>
- [**Gao et al. 2021**] Gao, C., et al., 2021, Advances and Challenges in Conversational Recommender Systems: A Survey, <https://arxiv.org/pdf/2101.09459.pdf>
- [**Gartner 2020**] <https://www.forbes.com/sites/louiscolombus/2020/10/04/whats-new-in-gartners-hype-cycle-for-ai-2020/?sh=723a6e57335c>
- [**Google 2021**] Mirhoseini, A., et al., 2021, A graph placement methodology for fast chip design. *Nature*, <https://doi.org/10.1038>
- [**Janer et al. 2021**] Janer, M., et al., 2021, Reinforcement learning as one big sequence modeling problem, <https://arxiv.org/pdf/2106.02039.pdf>
- [**Jordan 2019**] Jordan, M., 2019, *Artificial Intelligence—The Revolution Hasn't Happened Yet*, MIT Press, <https://hdsr.mitpress.mit.edu/pub/wot7mkc1/release/9>
- [**Vogel et al. 2018**] Vogel, T., et al., 2018, mRUBiS: an exemplar for model-based architectural self-healing and self-optimization, in *SEAMS '18*, pp. 101–107. <https://doi.org/10.1145/3194133.3194161>
- [**Sutton & Barto 2018**] Sutton, R. & Barto, A., 2018, *Reinforcement Learning, An Introduction*, MIT Press