Quantifying Generalization in Reinforcement Learning

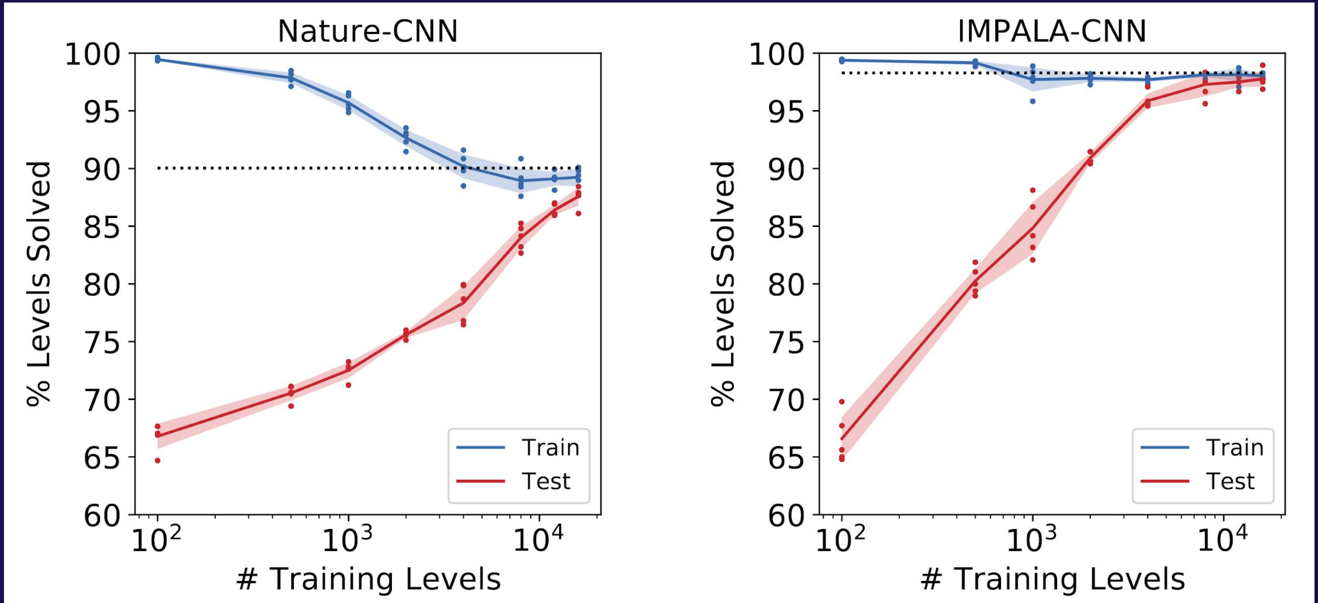
Karl Cobbe, Oleg Klimov, Chris Hesse, Taehoon Kim, John Schulman

OpenAI

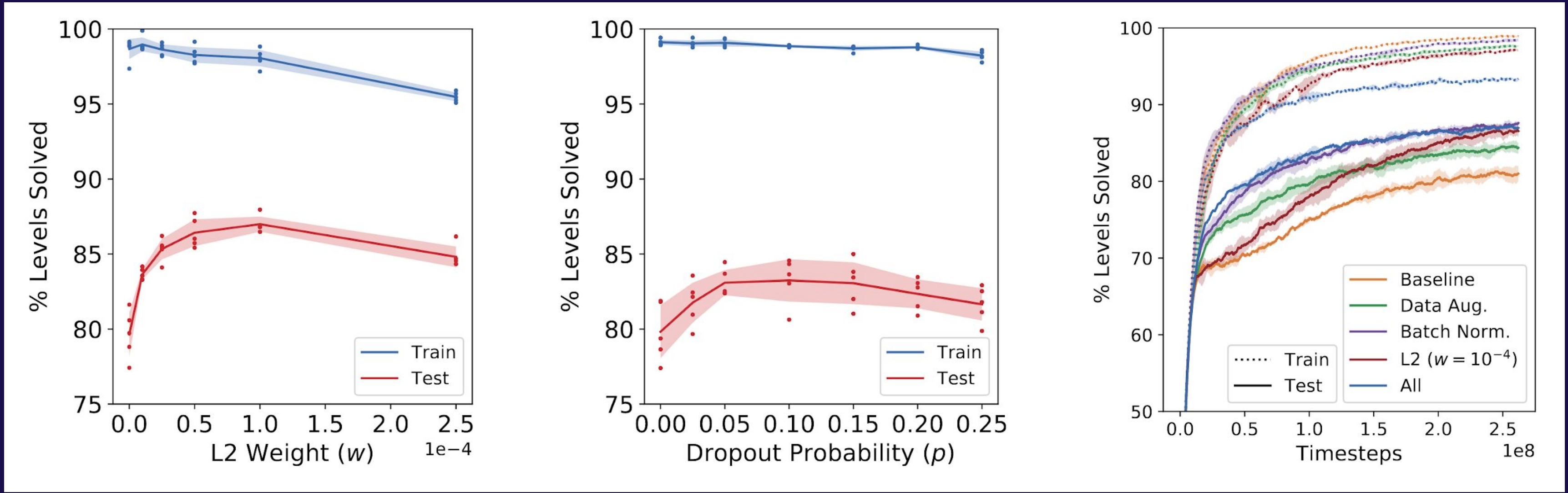
- Many common deep RL benchmarks ignore generalization
- CoinRun: a procedurally generated environment with distinct train/test sets
- Deep architectures, dropout, L2 regularization improve generalization
- Better benchmarks lead to better architectures and algorithms

The CoinRun Environment

- Level generation conditioned on difficulty
- Level diversity provides natural curriculum
- Can construct training sets of arbitrary size



- Larger training sets lead to better generalization
- Dotted line represents agent trained with unrestricted levels
- Deeper architecture generalizes better (3 vs. 15 conv layers)



- Train on 500 fixed levels
- L2, dropout, batch norm and data augmentation improve generalization
- Increasing policy stochasticity improves generalization, but slows training

Takeaways

- Agents are capable of overfitting to a large number of specific environments
- Deeper architectures and regularization reduce overfitting
- Lessons learned from CoinRun should apply in more complex environments

Thanks for listening!

- Come to poster Pacific Ballroom #32 tonight!
- Code: https://github.com/openai/coinrun
- Special thanks to my co-authors for their contributions to the environment and the paper:
    - John Schulman
    - Oleg Klimov
    - Chris Hesse
    - Taehoon Kim