Dude, Where's My MountainCar?

A Course Project

Valentin Tiriac
with Paul Liu and Saidur Rahman

Overview

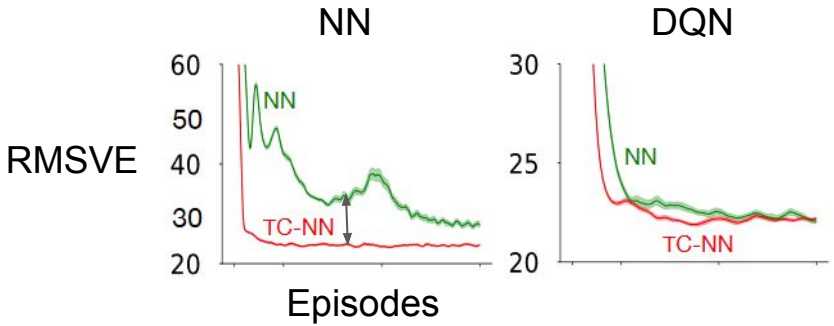
1. What’s the problem?
2. Review some previous work
3. What we’re doing about it

What’s the problem?

Why doesn’t using NN function approximation with Q-learning “just work”?

- Experience replay 🤮
- Target networks 🤮🤮
- Feature engineering 🤮🤮🤮

What's the problem?



(Ghiassian et. al 2020)

1. What's the problem?
2. Review some previous work
3. What we're doing about it

Previous work

The Utility of Sparse Representations for Control in Reinforcement Learning

Vincent Liu[1], Raksha Kumaraswamy[1], Lei Le[2], Martha White[1]
[1]Department of Computing Science, University of Alberta, Edmonton, Canada
{vliu1, kumarasw, whitem}@ualberta.ca
[2]Department of Computer Science, Indiana University Bloomington, Indiana, USA

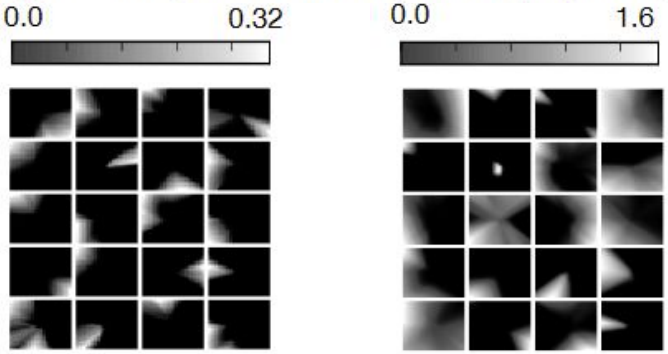
Improving Performance in Reinforcement Learning by Breaking Generalization in Neural Networks

Sina Ghiassian, Banafsheh Rafiee, Yat Long Lo, Adam White
Reinforcement Learning and Artificial Intelligence Laboratory, University of Alberta and
Alberta Machine Intelligence Institute (AMII)

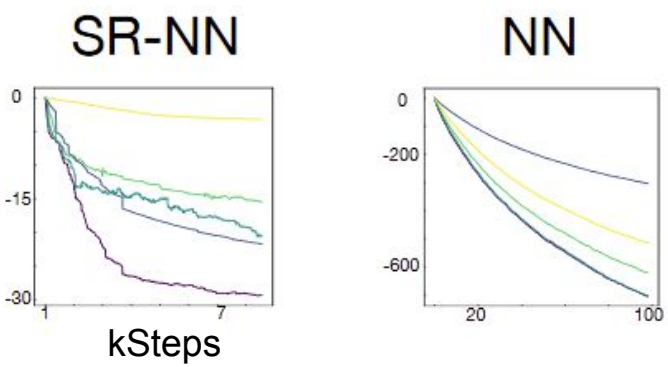
The Utility of Sparse Representations for Control in Reinforcement Learning

Vincent Liu[1], Raksha Kumaraswamy[1], Lei Le[2], Martha White[1]
[1]Department of Computing Science, University of Alberta, Edmonton, Canada
{vliu1, kumarasw, whitem}@ualberta.ca
[2]Department of Computer Science, Indiana University Bloomington, Indiana, USA

Activations for some units

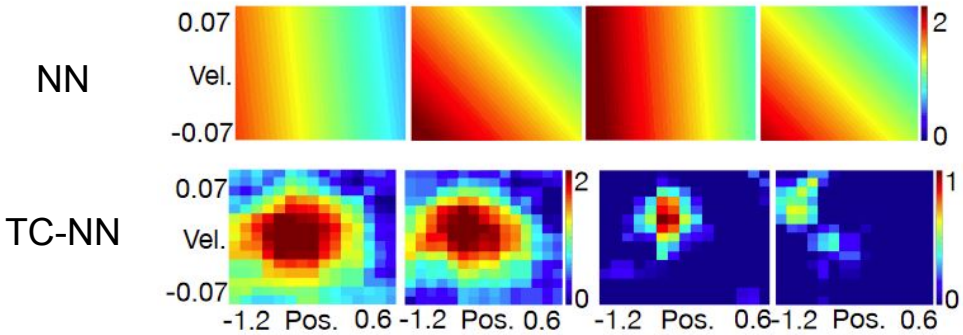


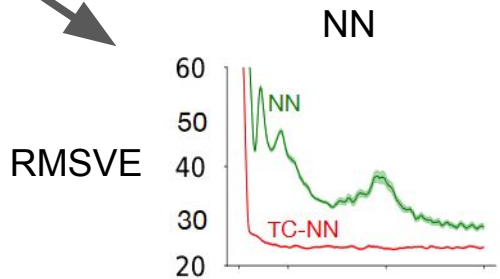
Value for five states



Improving Performance in Reinforcement Learning by Breaking Generalization in Neural Networks

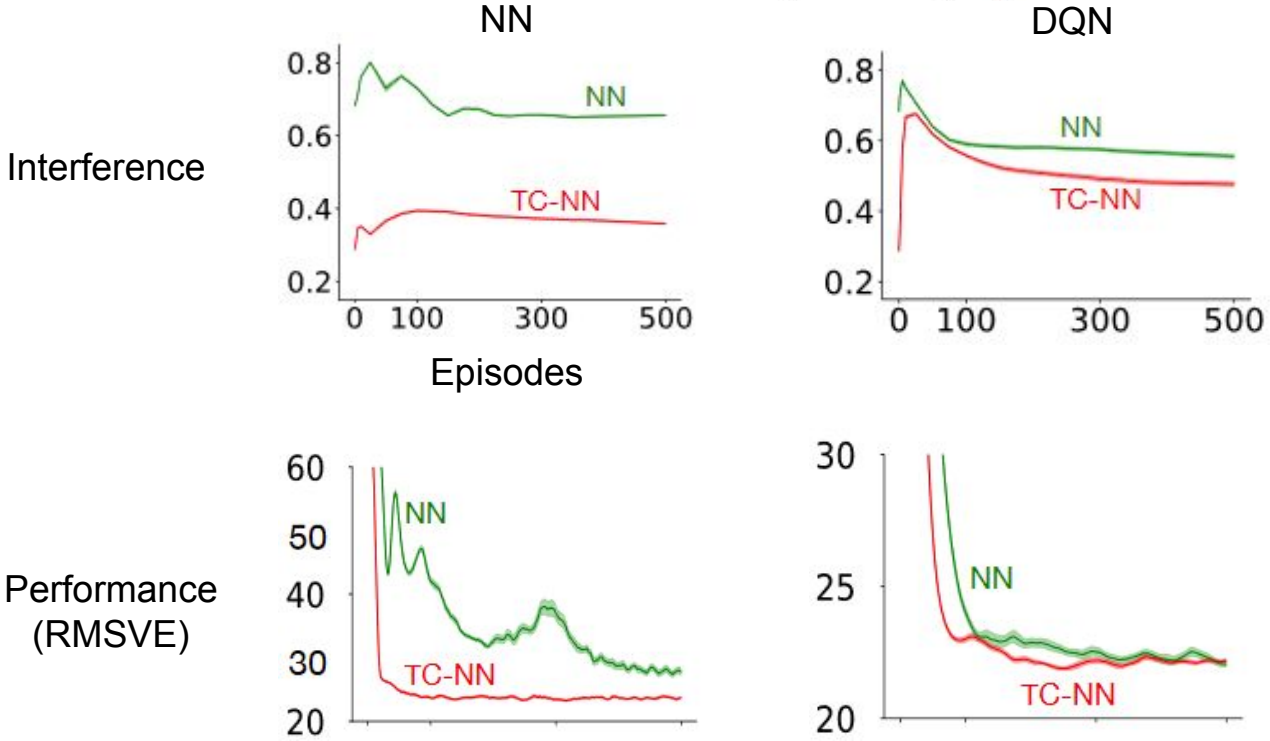
Sina Ghiassian, Banafsheh Rafiee, Yat Long Lo, Adam White
Reinforcement Learning and Artificial Intelligence Laboratory, University of Alberta and
Alberta Machine Intelligence Institute (AMII)





Improving Performance in Reinforcement Learning by Breaking Generalization in Neural Networks

Sina Ghiassian, Banafsheh Rafiee, Yat Long Lo, Adam White
Reinforcement Learning and Artificial Intelligence Laboratory, University of Alberta and
Alberta Machine Intelligence Institute (AMII)



What's "interference"?

$$\mathrm{PI}(S_i, S_j) = \frac{\nabla_{\boldsymbol{w}} \left[\hat{v}(\boldsymbol{w}, S_i)\right]^{\mathsf{T}} \nabla_{\boldsymbol{w}} \left[\hat{v}(\boldsymbol{w}, S_j)\right]}{||\nabla_{\boldsymbol{w}} \hat{v}(\boldsymbol{w}, S_i)||_2 \times ||\nabla_{\boldsymbol{w}} \hat{v}(\boldsymbol{w}, S_j)||_2}$$

After learning (st, at)... ...how does the error at (si, ai) change?

$$PI(\boldsymbol{\theta}_t; (s_t, a_t), (s_i, a_i)) := J(\boldsymbol{\theta}_{t+1}; s_i, a_i) - J(\boldsymbol{\theta}_t; s_i, a_i)$$

$$\begin{aligned} & PI(\boldsymbol{\theta}_t; (s_t, a_t), (s_i, a_i)) \\ & \approx (\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_t) \frac{\partial J(\boldsymbol{\theta}_t; s_i, a_i)}{\partial \boldsymbol{\theta}_t} \\ & = 2\alpha [U_t - Q_{\boldsymbol{\theta}_t}(s_t, a_t)][Q^{\pi}(s_i, a_i) - Q_{\boldsymbol{\theta}_t}(s_i, a_i)] \nabla_{\boldsymbol{\theta}_t} Q_{\boldsymbol{\theta}_t}(s_t, a_t)^{\mathsf{T}} \nabla_{\boldsymbol{\theta}_t} Q_{\boldsymbol{\theta}_t}(s_i, a_i) \end{aligned}$$

(Liu 2019)

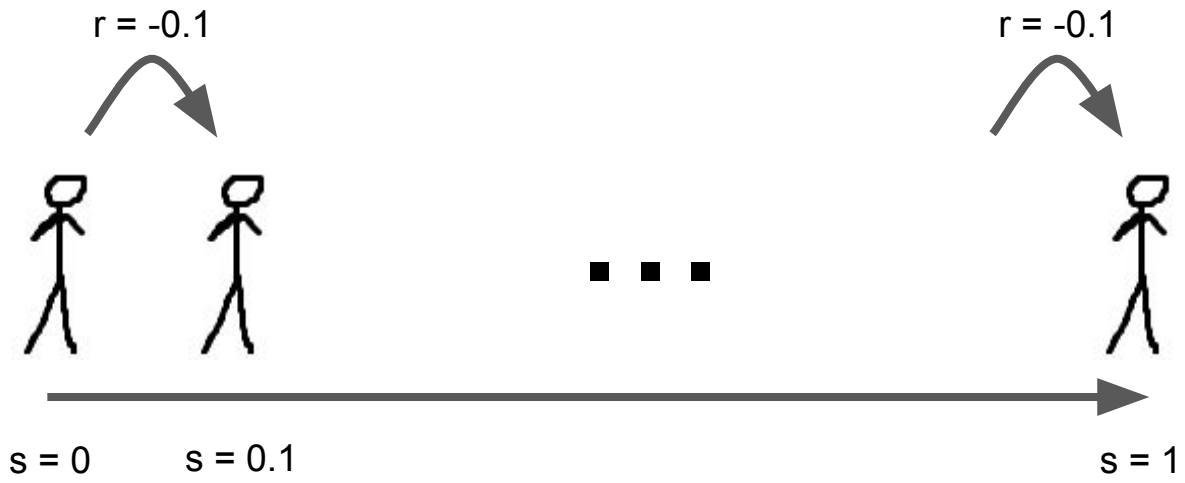
So far we know...

- There’s a problem with NNs as FAs in RL
  - But workarounds exist (ER, preprocessing, sparsity)
- Some potential causes have been studied
  - But we want to dig deeper

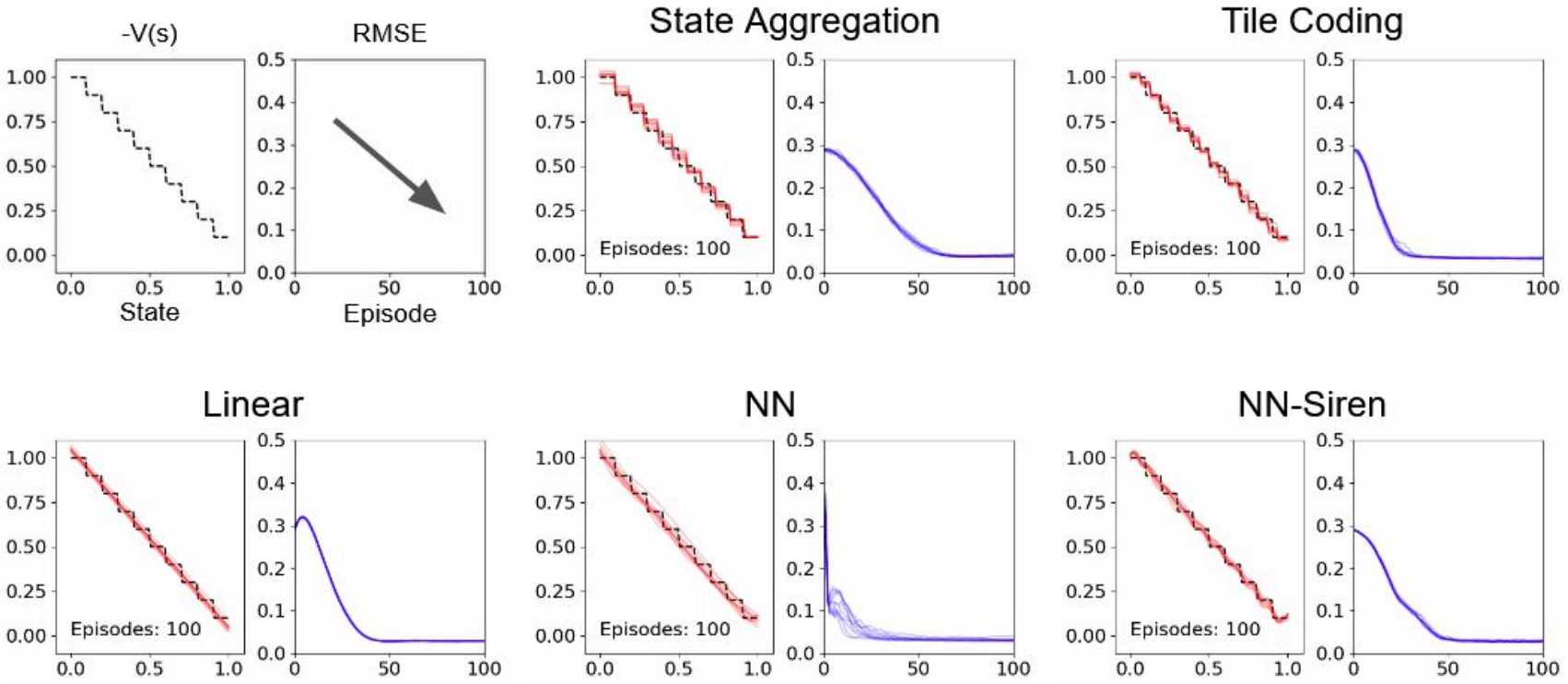
1. What's the problem?
2. Review some previous work
3. What we're doing about it

Some results on a tiny environment

“1D” Environment



“1D” Environment Learned Value Functions

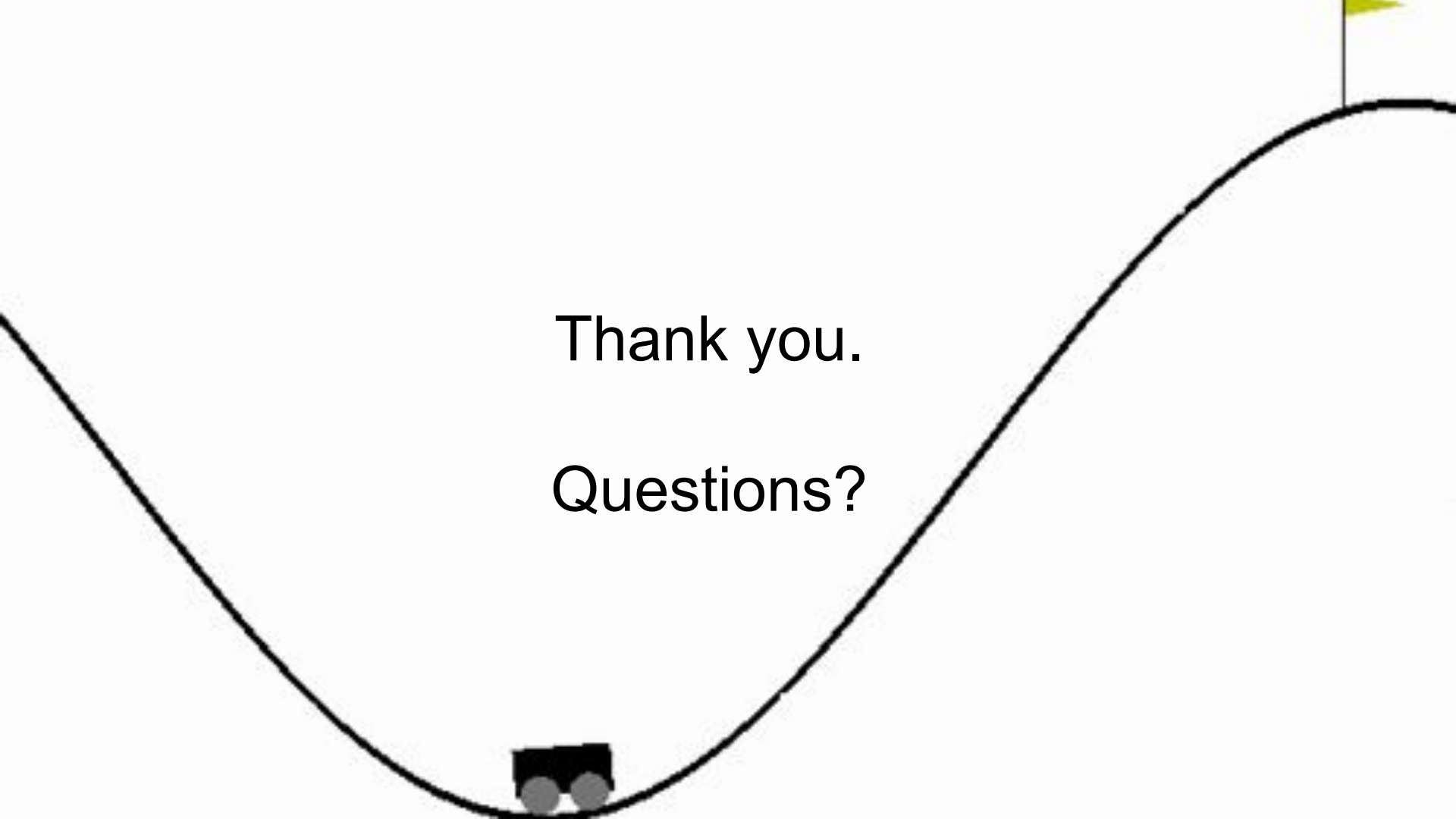


Next Steps

- Deeper look into interference
- Check TC-Linear vs TC-NN
- Comparison of capacity and non-IID training for TC-Linear vs. TC-NN vs. NN
- Check if other NN architectures can break incorrect generalization directly
- Setting: prediction problem + MountainCar

Wrap Up

- Figure out why NNs don’t work with Q-learning
- Previous research looked at some possible causes
- We’re looking deeper into it


Thank you.
Questions?

Vincent Liu, Master’s Thesis (2019)
https://era.library.ualberta.ca/items/b4cd1257-69ae-4349-9de6-3feed2648eb1/view/d301ebee-7c64-4027-9411-ed0ef19d6e8f/Liu_Vincent_201909_MSc.pdf

The Utility of Sparse Representations for Control in Reinforcement Learning
Vincent Liu, Raksha Kumaraswamy, Lei Le, Martha White (2019)
https://arxiv.org/abs/1811.06626

Improving Performance in Reinforcement Learning by Breaking Generalization in Neural Networks
Sina Ghiassian, Banafsheh Rafiee, Yat Long Lo, Adam White (2020)
https://arxiv.org/abs/2003.07417