Reinforcement Learning and Model Predictive Control

RL: optimizes policy π_θ for given cost using data

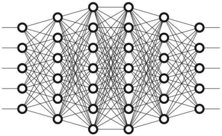
- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

Reinforcement Learning and Model Predictive Control

RL: optimizes policy π_θ for given cost using data

- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

Why do we dislike RL?



"Machine Learning is alchemy", Ali Rahimi, 2018

Reinforcement Learning and Model Predictive Control

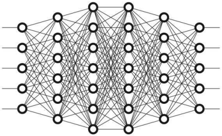
RL: optimizes policy π_θ for given cost using data

- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

MPC: creates "optimal" policy π_θ from model, cost, constraints

- Solid math... everywhere
- Safe & explainable
- Policy is as good as the model is

Why do we dislike RL?



"Machine Learning is alchemy", Ali Rahimi, 2018

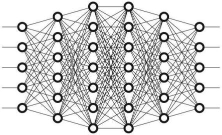
RL: optimizes policy π_θ for given cost using data

- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

MPC: creates "optimal" policy π_θ from model, cost, constraints

- Solid math... everywhere
- Safe & explainable
- Policy is as good as the model is

Why do we dislike RL?



"Machine Learning is alchemy", Ali Rahimi, 2018

How do we get a model?

- First-principles: $\mathbf{x}_+ = \mathbf{f}_\theta(\mathbf{x}, \mathbf{u})$
- SYSID (fit to data)
- Robust MPC for uncertainties

Reinforcement Learning and Model Predictive Control

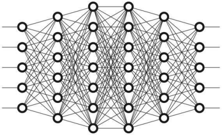
RL: optimizes policy $\boldsymbol{\pi}_{\boldsymbol{\theta}}$ for given cost using data

- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

MPC: creates “optimal” policy $\boldsymbol{\pi}_{\boldsymbol{\theta}}$ from model, cost, constraints

- Solid math... everywhere
- Safe & explainable
- Policy is as good as the model is

Why do we dislike RL?



“Machine Learning is alchemy”, Ali Rahimi, 2018

How do we get a model?

- First-principles: $\mathbf{x}_{+} = \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{u})$
- SYSID (fit to data)
- Robust MPC for uncertainties

Can we combine RL and MPC? Why? How?

Reinforcement Learning and Model Predictive Control

RL: optimizes policy π_θ for given cost using data

- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

MPC: creates "optimal" policy π_θ from model, cost, constraints

- Solid math... everywhere
- Safe & explainable
- Policy is as good as the model is

Can we combine RL and MPC? Why?

RL: big toolbox to "tune" parameters θ from data (tools need adaptation though)

MPC: policy approximation π_θ that:

- Uses prior knowledge
- Ensures safety, formal results, explainable

RL: optimizes policy π_θ for given cost using data

- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

MPC: creates "optimal" policy π_θ from model, cost, constraints

- Solid math... everywhere
- Safe & explainable
- Policy is as good as the model is

Can we combine RL and MPC? Why?

RL: big toolbox to "tune" parameters θ from data (tools need adaptation though)

MPC: policy approximation π_θ that:

- Uses prior knowledge
- Ensures safety, formal results, explainable

Some fun observations

- RL may "tune" cost and constraints in MPC ($\sim$ learn constraints tightening)
- RL may "tune" SYSID model
- SYSID comes back in its "set-membership version", RL tunes uncertainty set

RL: optimizes policy π_θ for given cost using data

- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

MPC: creates "optimal" policy π_θ from model, cost, constraints

- Solid math... everywhere
- Safe & explainable
- Policy is as good as the model is

Can we combine RL and MPC? Why?

RL: big toolbox to "tune" parameters θ from data (tools need adaptation though)

MPC: policy approximation π_θ that:

- Uses prior knowledge
- Ensures safety, formal results, explainable

Some fun observations

- RL may "tune" cost and constraints in MPC ($\sim$ learn constraints tightening)
- RL may "tune" SYSID model
- SYSID comes back in its "set-membership version", RL tunes uncertainty set

Where is this going?

Reinforcement Learning and Model Predictive Control

RL: optimizes policy π_θ for given cost using data

- Solid math for the tuning
- DNNs support approximations
- Safety? Explainability?

MPC: creates "optimal" policy π_θ from model, cost, constraints

- Solid math... everywhere
- Safe & explainable
- Policy is as good as the model is

1. Safe Reinforcement Learning with stability guarantees using min-max Robust NMPC, S. Gros, M. Zanon,Transaction on Automatic Control, (submitted)
2. Reinforcement Learning for mixed-integer problems with MPC-based function approximation, S. Gros, M. Zanon, IFAC 2020 (submitted)
3. Learning Real-Time Iteration NMPC, V. Kungurstev, M. Zanon, S. Gros, IFAC 2020 (submitted)
4. Safe Reinforcement Learning via projection on a safe set: how to achieve optimality? S. Gros, M. Zanon, IFAC 2020 (submitted)
5. Towards Safe Reinforcement Learning Using NMPC and Policy Gradients: Part I - Stochastic case, S. Gros, M. Zanon, Transaction on Automatic Control (submitted, arxiv.org/abs/1906.04057)
6. Towards Safe Reinforcement Learning Using NMPC and Policy Gradients: Part II - Deterministic Case, S. Gros, M. Zanon, Transaction on Automatic Control (submitted, arxiv.org/abs/1906.04034)
7. Safe Reinforcement Learning Using Robust MPC, S. Gros, M. Zanon, Transaction on Automatic Control, 2019 (submitted, arxiv.org/abs/1906.04005)
8. Practical Reinforcement Learning of Stabilizing Economic MPC, M. Zanon, S. Gros, A. Bemporad, European Control Conference 2019
9. Data-driven Economic NMPC using Reinforcement Learning, S. Gros, M. Zanon, Transaction on Automatic Control, 2019