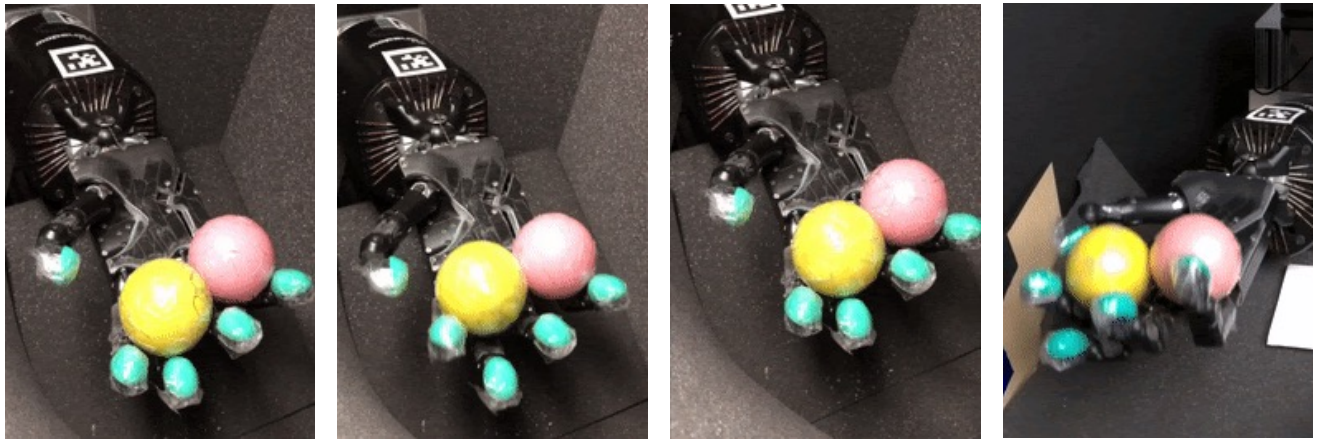
Summary on Model-based RL

Christopher Mutschler

Real-world systems

MPC (CEM-like) + Ensemble of Probabilistic Models

- Online Planning (MPC)
- Trajectory sampling for uncertainty propagation
- Ensembles of probabilistic neural networks for modeling



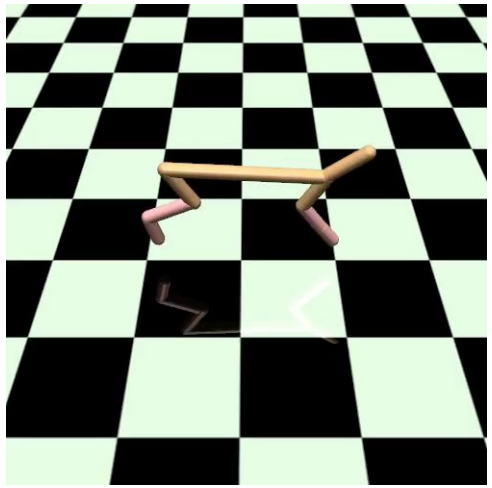
Training progress on the ShadowHand hardware.
From left to right: 0-0.25 hours, 0.25-0.5 hours, 0.5-1.5 hours, ~2 hours.

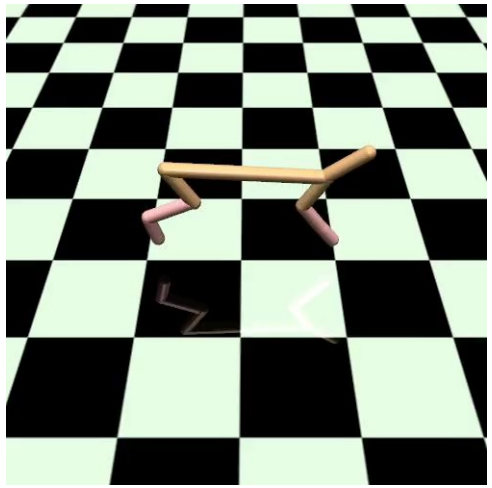
A. Nagabandi et al.: Deep dynamics models for learning dexterous manipulation. Conf. Robot Learning. 2020.
https://bair.berkeley.edu/blog/2019/09/30/deep-dynamics/

Real-world systems

Sample efficiency benchmarks

- Example Domain: MuJoCo HalfCheetah





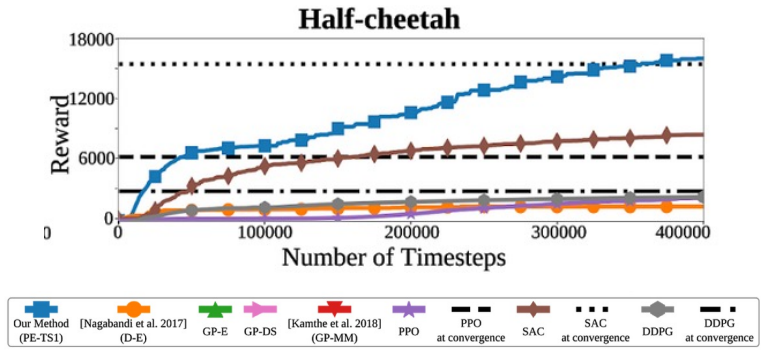
http://ai.berkeley.edu/lecture_slides.html

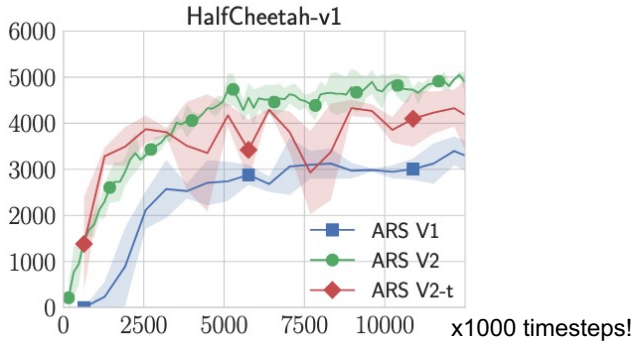
Real-world systems

Sample efficiency benchmarks

- Example Domain: MuJoCo HalfCheetah

FAIL!	Model-based RL (e.g., PILCO)
	↑
$k \cdot 10^5$	Model-based Deep RL (e.g., PETS)
	↑
$k \cdot 10^6$	Value-based Off-Policy Methods (e.g., SAC)
	↑
$k/2 \cdot 10^7$	Policy Gradient Methods (e.g., TRPO)
	↑
$k/2 \cdot 10^8$	Fully Online Methods (e.g., A3C)
	↑
$k \cdot 10^8$	Gradient-free Methods (e.g., ARS)





References

- https://sites.google.com/view/mbrl-tutorial
- Fazeli et al. (2019). See, feel, act: Hierarchical learning for complex manipulation skills with multisensory fusion. Science Robotics, 4(26).
- Fisac et al. (2019). A General Safety Framework for Learning-Based Control in Uncertain Robotic Systems. IEEE Transactions on Automatic Control.
- Sadigh et al. (2016). Planning for autonomous cars that leverage effects on human actions. RSS 2016.
- Silver et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484.
- Segler, Preuss, & Waller (2018). Planning chemical syntheses with deep neural networks and symbolic AI. Nature, 555(7698).
- Salas & Powell (2013). Benchmarking a scalable approximate dynamic programming algorithm for stochastic control of multidimensional energy storage problems.
- Warren Powell's 2017 ECSO tutorial, "A Unified Framework for Optimization under Uncertainty"
- https://de.mathworks.com/help/ident/ug/modeling-a-vehicle-dynamics-system.html
- M. I. Palmqvist et al.: Model predictive control for autonomous driving of a truck. KTH Royal Institute of Technology School of Electrical Engineering. 2016.
- https://de.wikipedia.org/wiki/Deep_Learning
- https://en.wikipedia.org/wiki/Long_short-term_memory
- Ebert, Finn, et al. (2018); Finn & Levine (2017); Finn, Goodfellow, & Levine (2016)
- https://bair.berkeley.edu/blog/2018/11/30/visual-rl/
- Lange et al.: Batch reinforcement learning. In Reinforcement learning (pp. 45-73). Springer, Berlin, Heidelberg. 2012.
- Kaiser et al.: Model-based reinforcement learning for atari. arXiv preprint arXiv:1903.00374. 2019.
- https://worldmodels.github.io
- Ha & Schmidhuber: World Models. NeurIPS 2018.
- Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.
- Janner et al.: When to Trust Your Model: Model-Based Policy Optimization. NeurIPS 2019.
- Andrychowicz, OpenAI: Marcin, et al. "Learning dexterous in-hand manipulation." The International Journal of Robotics Research 39.1 (2020): 3-20.
- https://deepmind.com/research/case-studies/alphago-the-story-so-far

References (cont'd)

- https://deepmind.com/blog/article/muzero-mastering-go-chess-shogi-and-atari-without-rules
- Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." nature 529.7587 (2016): 484-489.
- Silver, David, et al. "Mastering the game of go without human knowledge." nature 550.7676 (2017): 354-359.
- https://deepmind.com/blog/article/alphago-zero-starting-scratch
- Silver, David, et al. "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play." Science 362.6419 (2018): 1140-1144.
- Schrittwieser, Julian, et al. "Mastering atari, go, chess and shogi by planning with a learned model." Nature 588.7839 (2020): 604-609.
- https://www.youtube.com/watch?v=tNAIHEse7Ms
- Todorov and Li (2005). A generalized iterative LQG method.
- Jürgen Schmidhuber: Curious model-building control systems. IJCNN.1991.
- Pathak et al.: Curiosity-driven exploration by self-supervised prediction. ICML. 2017.
- https://www.inovex.de/blog/uncertainty-quantification-deep-learning
- Janner et al (2019). When to Trust Your Model: Model-Based Policy Optimization.
- Chuan Guo et al.: On Calibration of Model Neural Networks. ICML. 2017.
- https://de.wikipedia.org/wiki/Model_Predictive_Control
- Williams et al.: Information Theoretic MPC for Model-Based Reinforcement Learning. ICRA. 2017.
- Kurtland Chua et al.: Deep reinforcement learning in a handful of trials using probabilistic dynamics models. NIPS 2018.
- Marc Deisenroth et al.: PILCO: A model-based and data-efficient approach to policy search. ICML. 2011.
- https://www.youtube.com/watch?v=XiigTGKZfks
- Nitish Srivastava et al.: Dropout: a simple way to prevent neural networks from overfitting. JMLR. 2014.
- A. Nagabandi et al.: Deep dynamics models for learning dexterous manipulation. Conf. Robot Learning. 2020. https://bair.berkeley.edu/blog/2019/09/30/deep-dynamics/
- http://ai.berkeley.edu/lecture_slides.html

References

Further reading & watching:

- Marc Deisenroth: The Role of Uncertainty in Model-based Reinforcement Learning. Workshop on Uncertainty Propagation in Composite Models, Munich, 2019. https://deisenroth.cc/talks/2019-10-10-munich.pdf
- Igor Mordatch and Jessica Hamrick: Tutorial on Model-Based Methods in Reinforcement Learning. Presented at International Conference on Machine Learning (ICML) 2020. https://sites.google.com/view/mbrl-tutorial
- Sergey Levine: CS 285 at UC Berkeley – Deep Reinforcement Learning. Lectures “Lecture 10: Model-based Planning”, “Lecture 11: Model-based Reinforcement Learning”, and “Lecture 11: Model-based Policy Learning”, http://rail.eecs.berkeley.edu/deeprlcourse/