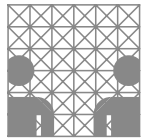
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

MIN Faculty
Department of Informatics



Improving Imitation Learning with Reinforcement Learning

Niklas Fiedler

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics
Technical Aspects of Multimodal Systems

November 26, 2019

Outline

1. Introduction
   - Motivation
2. Imitation Learning
   - Demonstration Methods
   - Behavioral Cloning
   - Inverse Reinforcement Learning
3. Combining Reinforcement Learning and Imitation Learning
   - BC Application
   - IRL Application
4. Conclusion

Goal

- Imitate expert behavior
- Improve learning by including knowledge given by demonstration
- Learn expert policies

$\rightarrow$ Make use of expert demonstrations

Motivation
Humans are Awesome

https://rejectedprincesses.tumblr.com/post/150495232038/chynara-madinkulova-long-hair-and-aida-akmatova

Learning from experts is natural behavior

[Haw50], `https://www.wakecounseling.com/therapy-blog/play-therapy`

Method to learn a behavior based on a demonstration

Various forms of demonstration.

Two prominent methods of implementation:

1. Behavioral Cloning
2. Inverse Reinforcement Learning

Method to learn a behavior based on a demonstration

Various forms of demonstration.

Two prominent methods of implementation:

1. Behavioral Cloning
2. Inverse Reinforcement Learning

Imitation Learning

Method to learn a behavior based on a demonstration

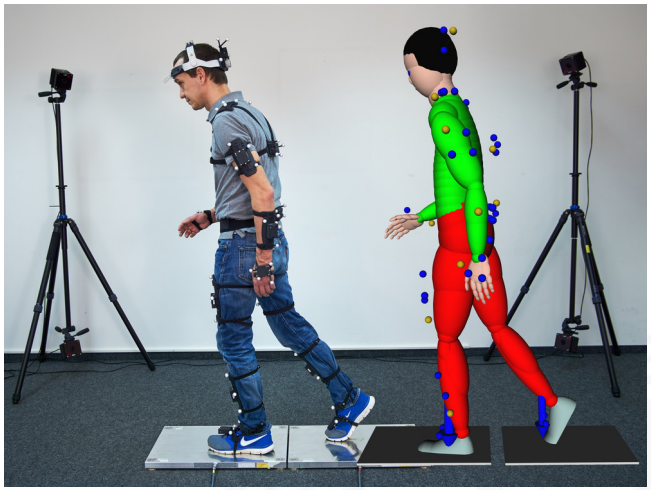
Various forms of demonstration.

Two prominent methods of implementation:

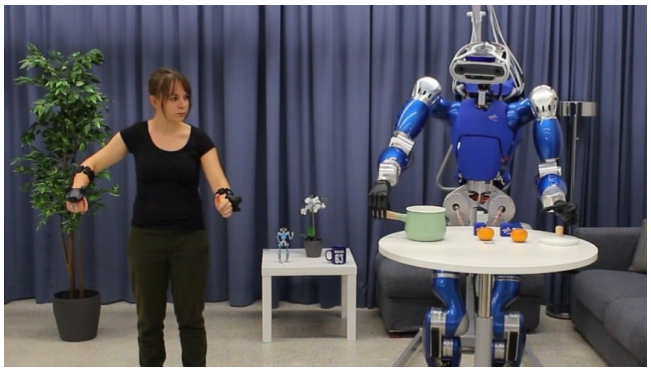
1. Behavioral Cloning
2. Inverse Reinforcement Learning

Demonstration Methods

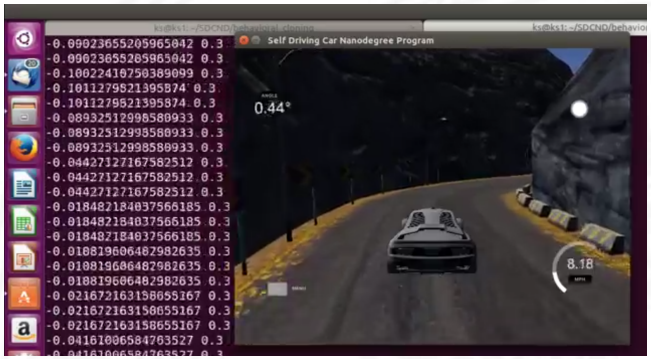
Virtual/Augumented Reality



Tracking of Human Motions



Teleoperation



Video Stream

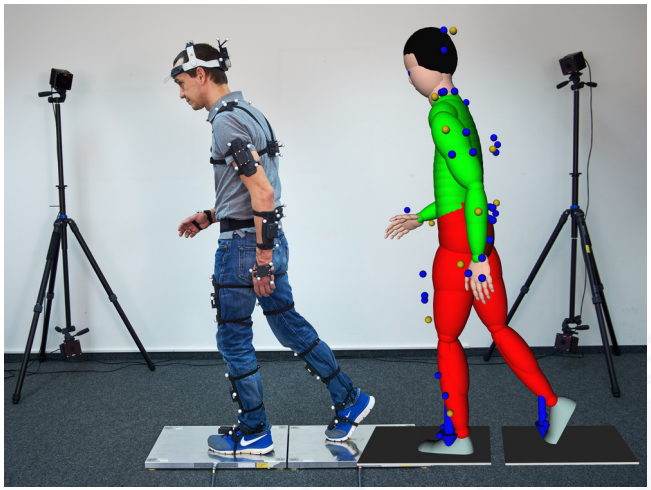
s3.ap-south-1.amazonaws.com/kidobotikz.sprw/master/assets/images/blog/blog-2018110811630.jpg

siamagazin.com/bimanual-teleoperation-of-a-compliant-whole-body-controlled-humanoid-robot/

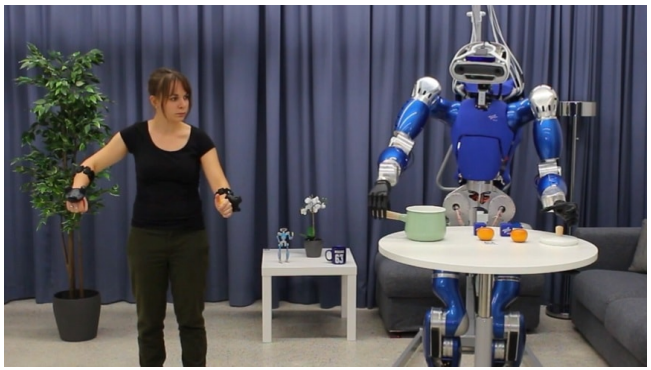
https://ar-tracking.com/applications/motion-capture/ https://www.youtube.com/watch?v=5BTIE_fhReo

Demonstration Methods

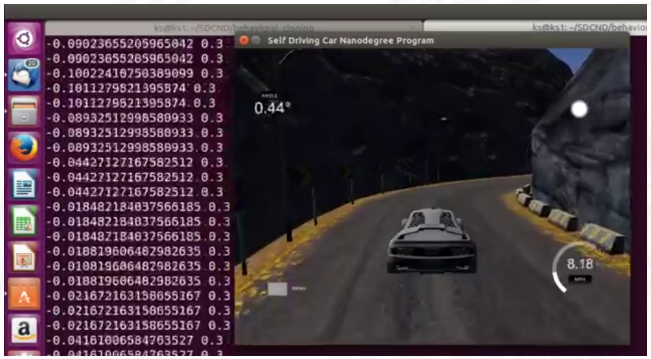
Virtual/Augumented Reality



Tracking of Human Motions



Teleoperation



Video Stream

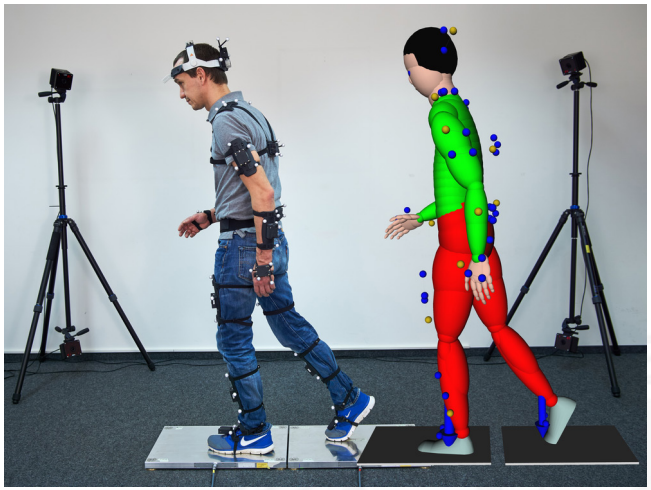
s3.ap-south-1.amazonaws.com/kidobotikz.sprw/master/assets/images/blog/blog-2018110811630.jpg

siamagazin.com/bimanual-teleoperation-of-a-compliant-whole-body-controlled-humanoid-robot/

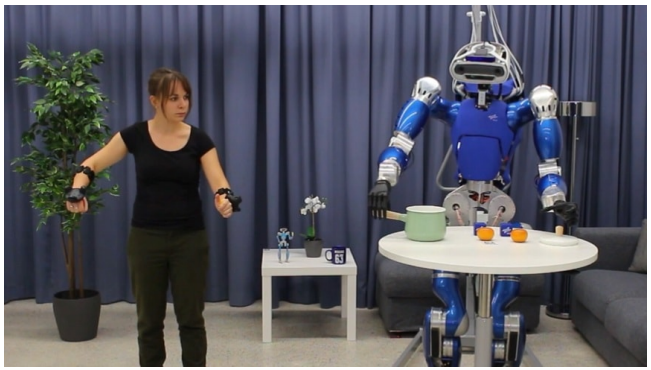
https://ar-tracking.com/applications/motion-capture/ https://www.youtube.com/watch?v=5BTIE_fhReo

Demonstration Methods

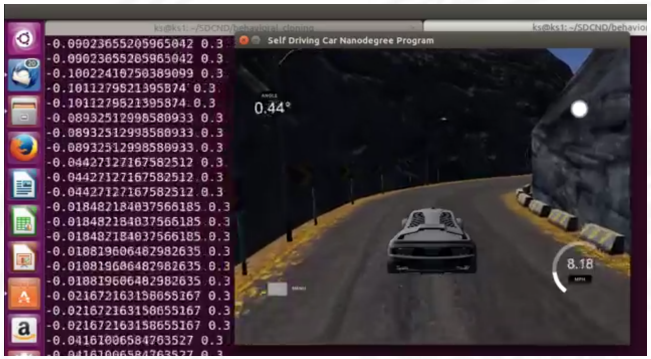
Virtual/Augumented Reality



Tracking of Human Motions



Teleoperation



Video Stream

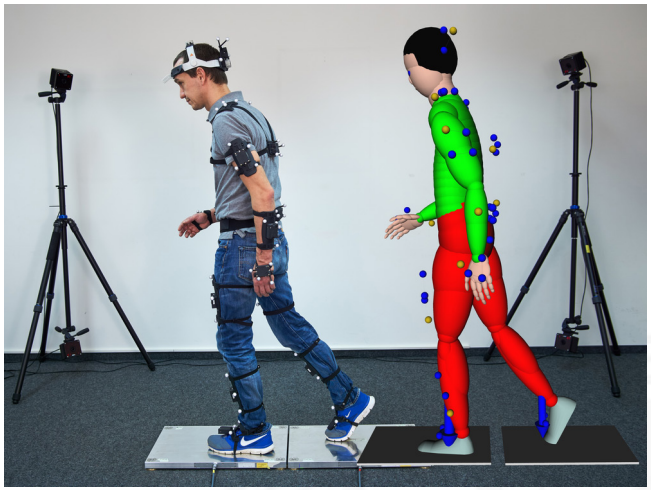
s3.ap-south-1.amazonaws.com/kidobotikz.sprw/master/assets/images/blog/blog-2018110811630.jpg

siamagazin.com/bimanual-teleoperation-of-a-compliant-whole-body-controlled-humanoid-robot/

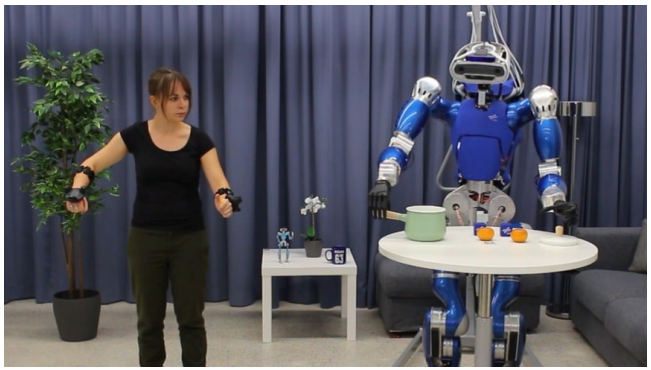
https://ar-tracking.com/applications/motion-capture/ https://www.youtube.com/watch?v=5BTIE_fhReo

Demonstration Methods

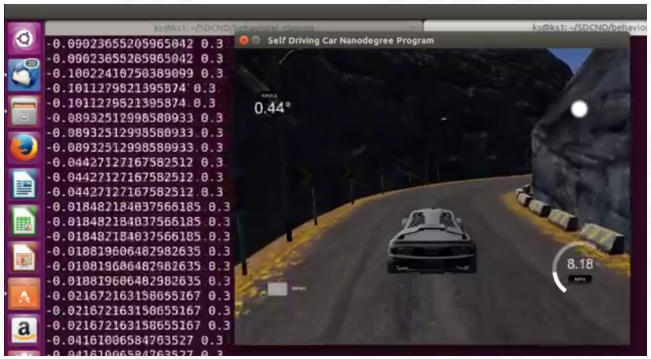
Virtual/Augumented Reality



Tracking of Human Motions



Teleoperation



Video Stream

s3.ap-south-1.amazonaws.com/kidobotikz.sprw/master/assets/images/blog/blog-2018110811630.jpg

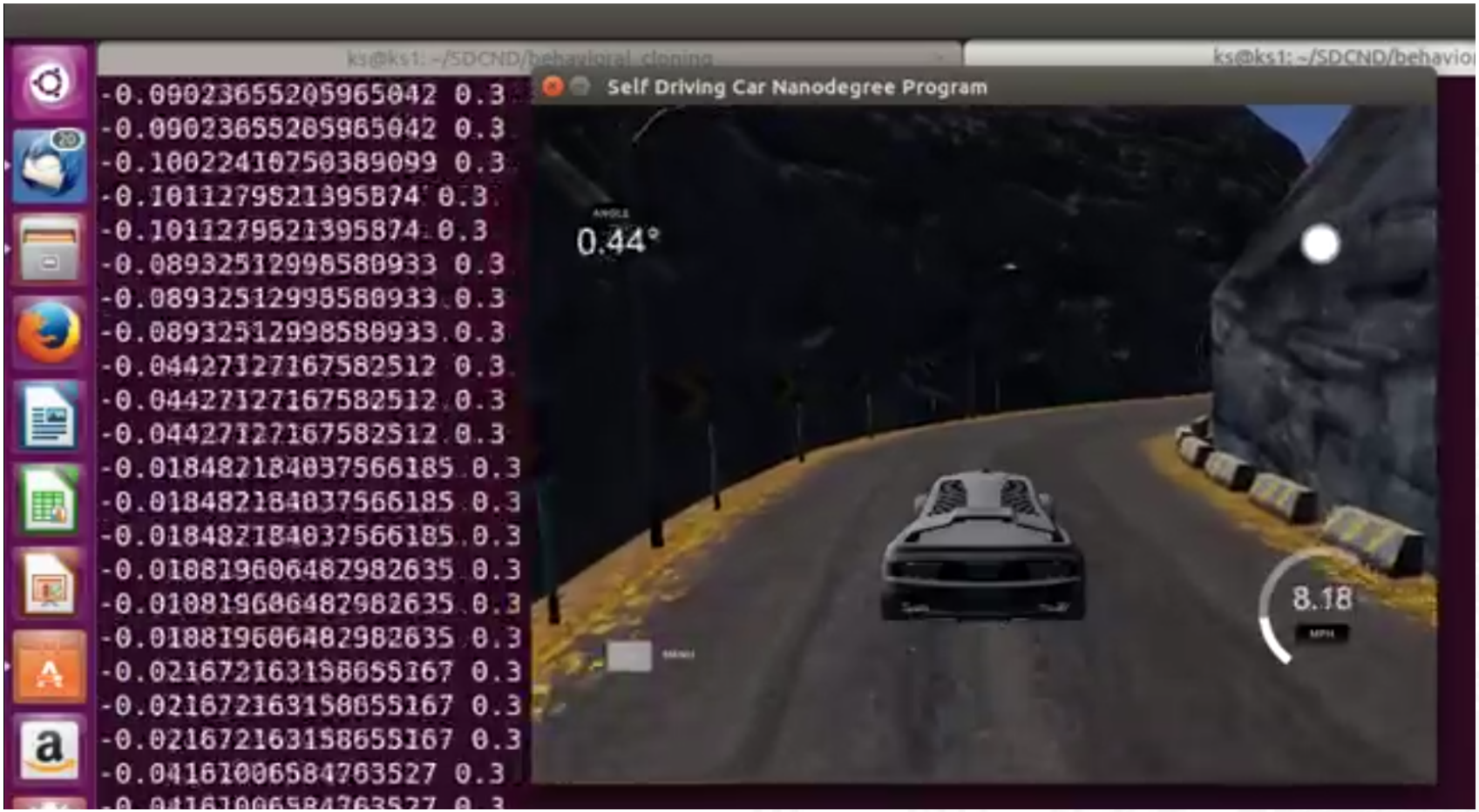
siamagazin.com/bimanual-teleoperation-of-a-compliant-whole-body-controlled-humanoid-robot/

https://ar-tracking.com/applications/motion-capture/ https://www.youtube.com/watch?v=5BTIE_fhReo

Behavioral Cloning

- Training a direct link between demonstrated input and output
- Large amounts of training data necessary
- Poor generalization

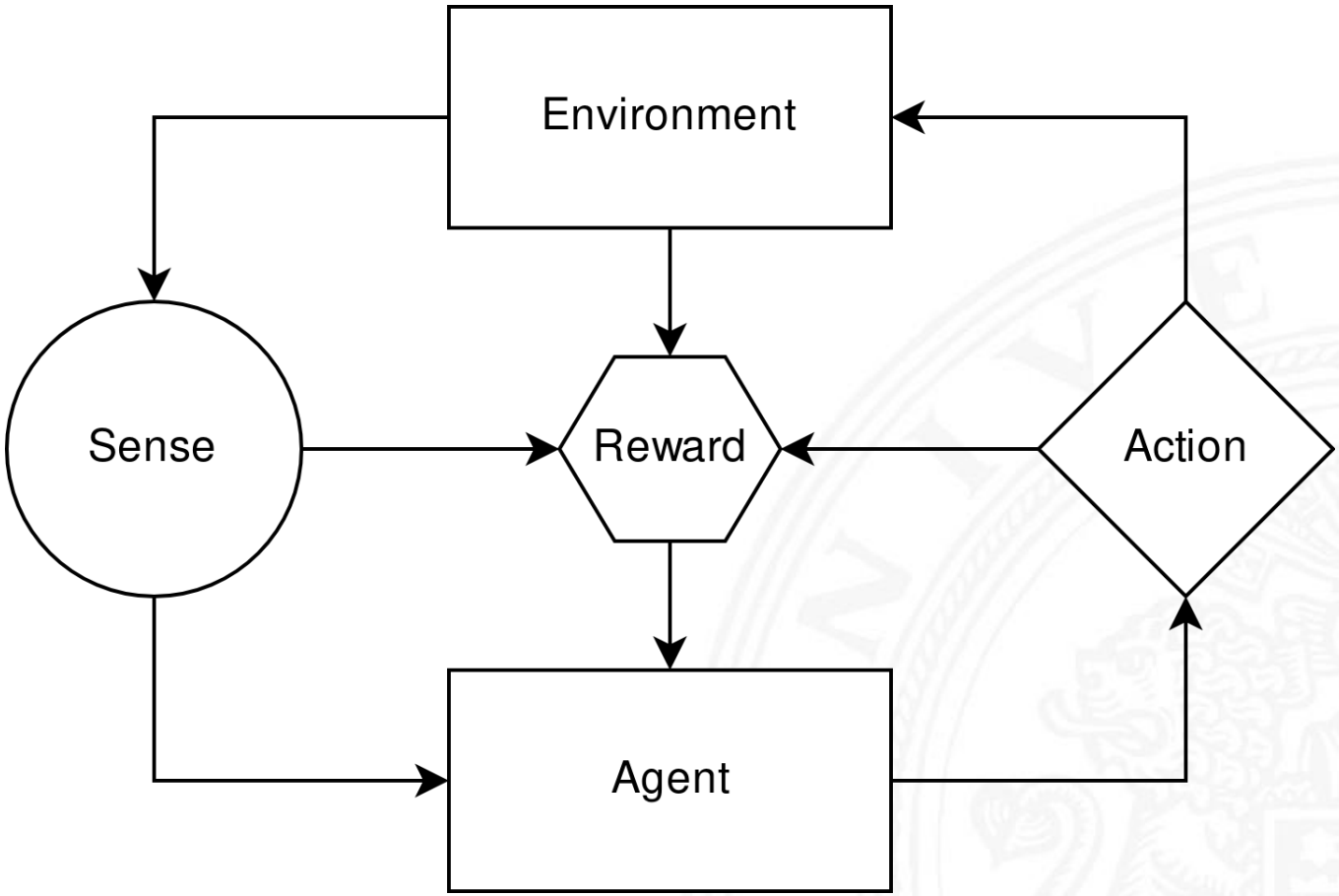
Behavioral Cloning
Video



https://www.youtube.com/watch?v=5BTIE_fhReo

Inverse Reinforcement Learning

Reinforcement Learning



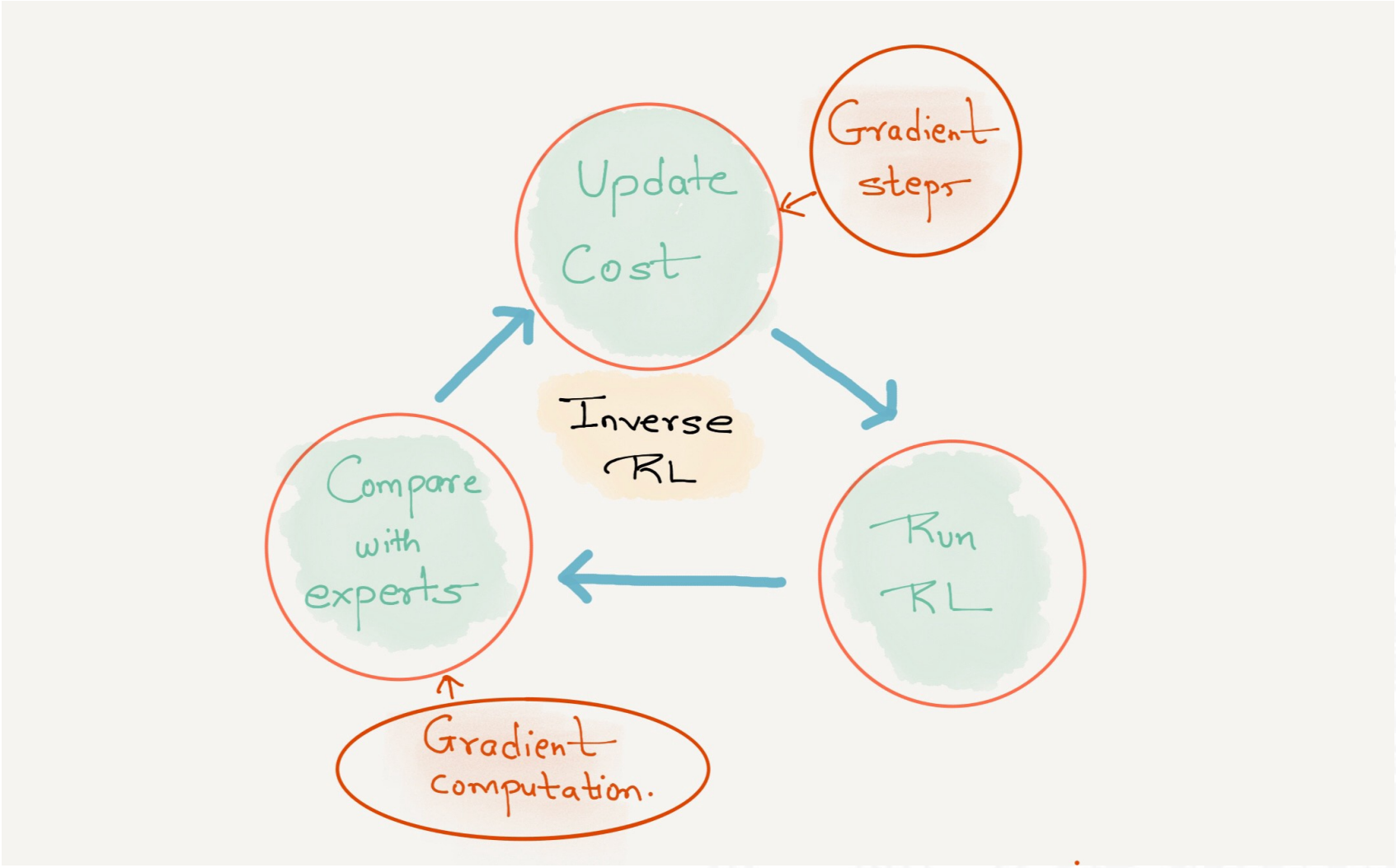
Inverse Reinforcement Learning

Reinforcement Learning vs. Inversed Reinforcement Learning

	RL	**IRL**
given	(partially observed) reward function $\mathcal{R}$	policy π or history sampled from that policy
searching	optimal policy π for given reward	reward function $\mathcal{R}$ for which given behavior is optimal

https://thinkingwires.com/posts/2018-02-13-irl-tutorial-1.html

Inverse Reinforcement Learning



https://medium.com/@sanketgujar95/generative-adversarial-imitation-learning-266f45634e60

Behavioral Cloning

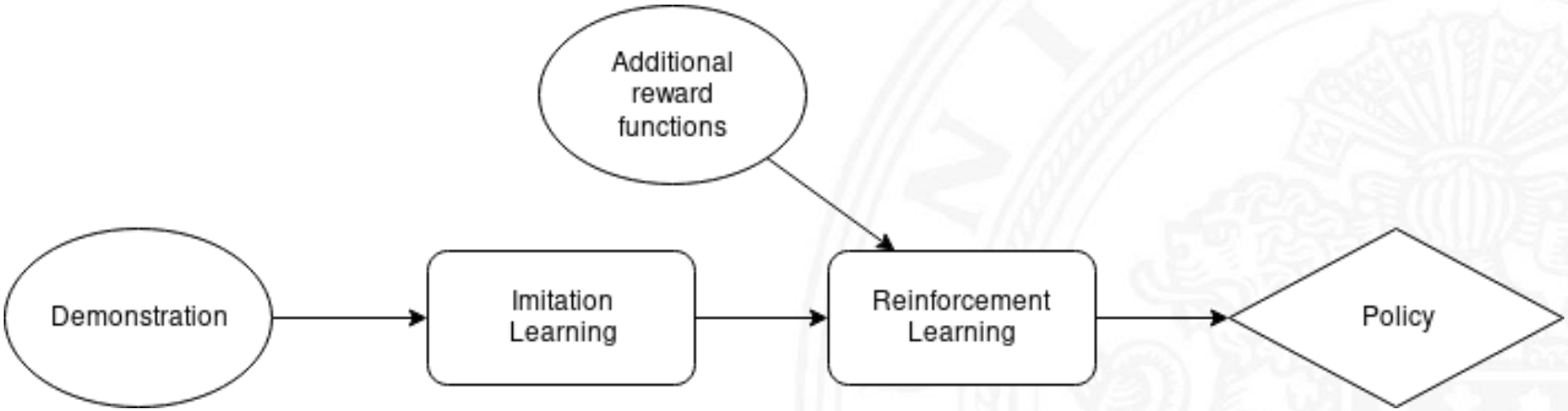
- Weak generalization
- Relatively low computational effort

Inversed Reinforcement Learning

- Strong generalization
- Large computational effort
- Complex structure

Combining Reinforcement Learning and Imitation Learning

- Reducing the impact of shortcomings of both methods
- Applications should outperform demonstrators after RL applications
- Accelerated training process
- Extending the capabilities learned with imitation learning



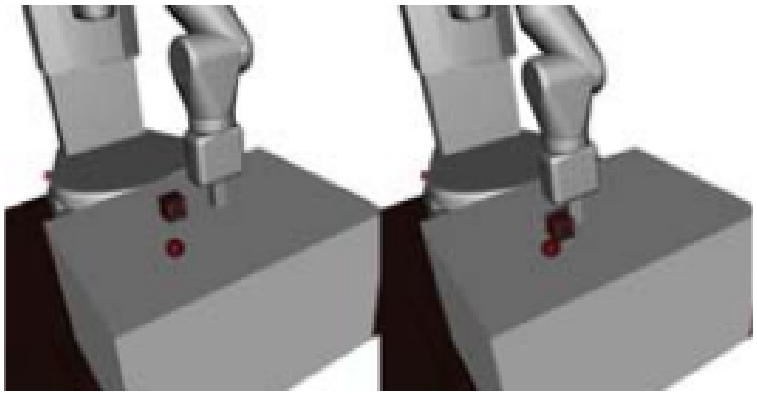
Overcoming Exploration in Reinforcement Learning with Demonstrations

Ashvin Nair[12], Bob McGrew[1], Marcin Andrychowicz[1], Wojciech Zaremba[1] and Pieter Abbeel[12]

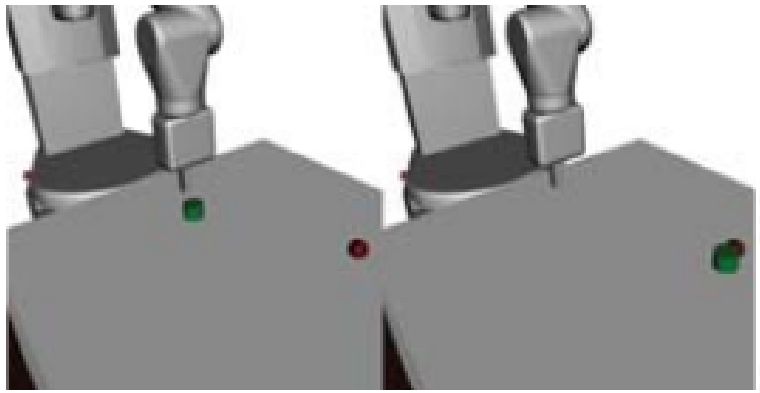
2018 IEEE International Conference on Robotics and Automation (ICRA)

[1]OpenAI
[2]University of California, Berkeley

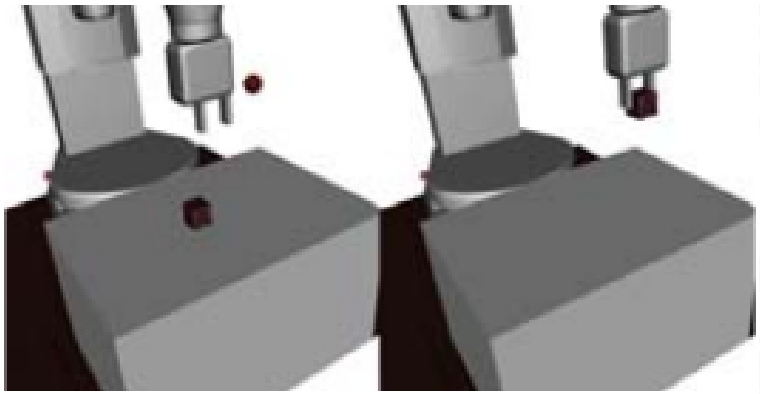
BC Application
Goal



Pushing

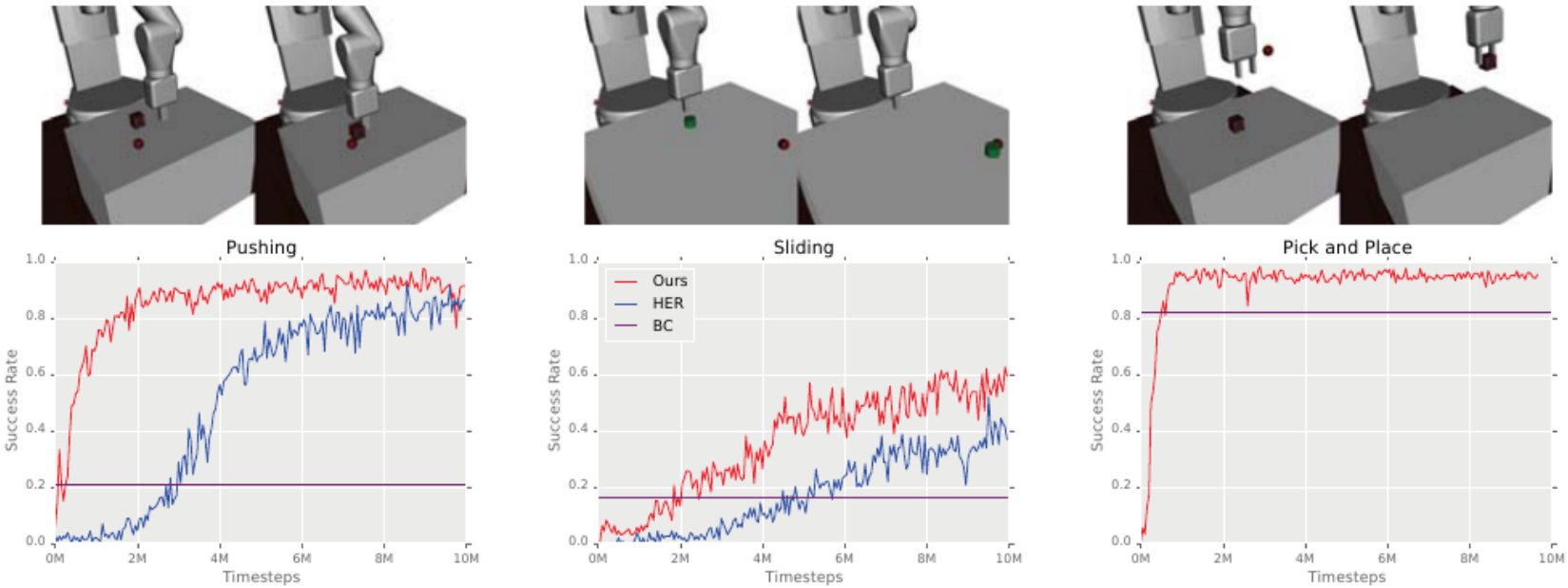


Sliding



Pick and Place

[NMA+18]


Pushing
Sliding
Pick and Place
Ours
HER
BC
Success Rate
Timesteps
0.0
0.2
0.4
0.6
0.8
1.0
0M
2M
4M
6M
8M
10M

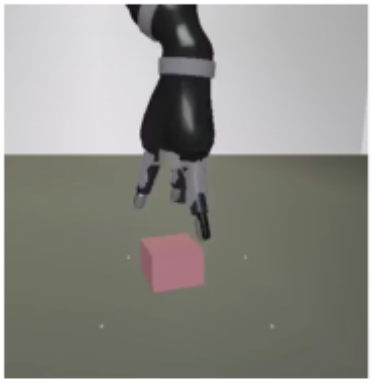
[NMA+18]

Reinforcement and Imitation Learning for Diverse Visuomotor Skills

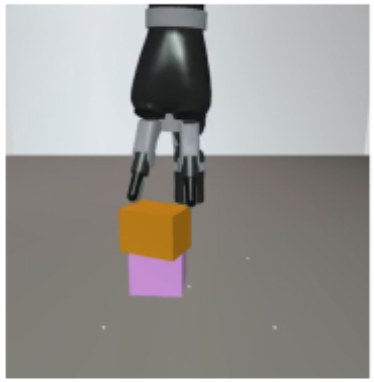
Yuke Zhu[1], Ziyu Wang[2], Josh Merel[2], Andrei Rusu[2], Tom Erez[2], Serkan Cabi[2], Saran Tunyasuvunakool[2], Janos Kramar[2], Raia Hadsell[2], Nando de Freitas[2] and Nicolas Heess[2]

[1]Computer Science Department, Stanford University
[2]OpenAI

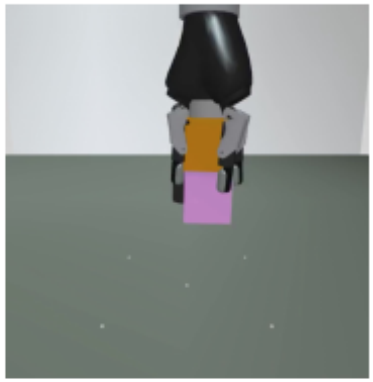
IRL Application
Goal



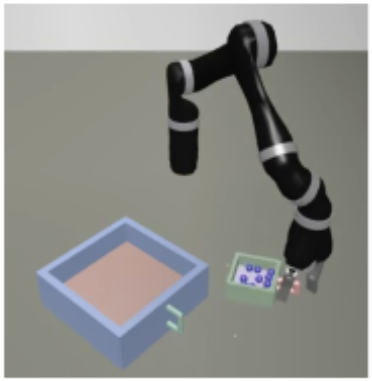
block lifting



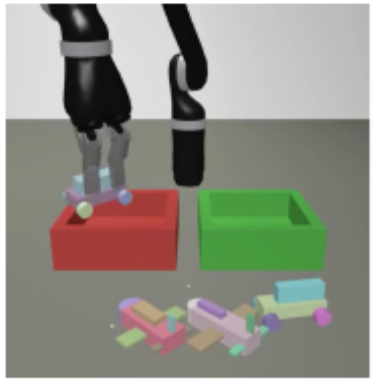
block stacking



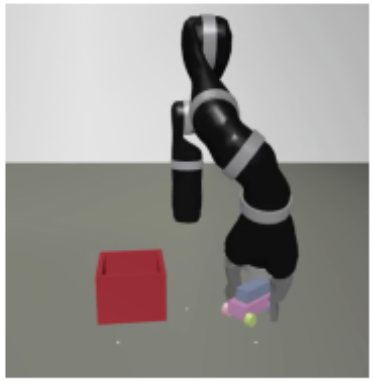
clearing table with blocks



pouring liquid



order fulfillment



clearing table with a box

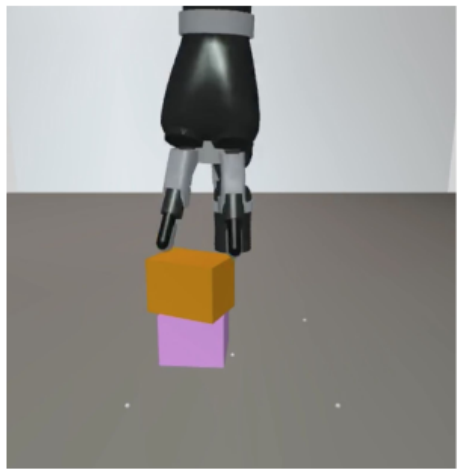
[ZWM+18]

IRL Application
Method

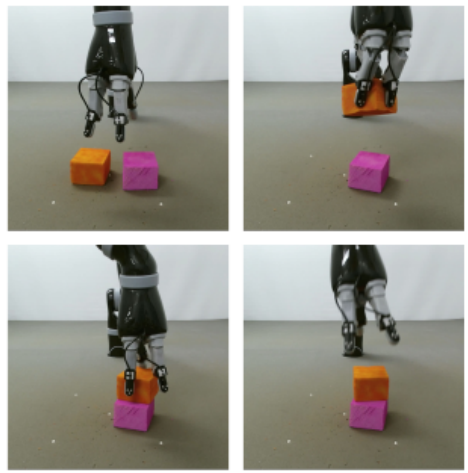
Collecting demonstrations → **Training in simulation** → **Running on real robot**

3D motion controller

physics engine



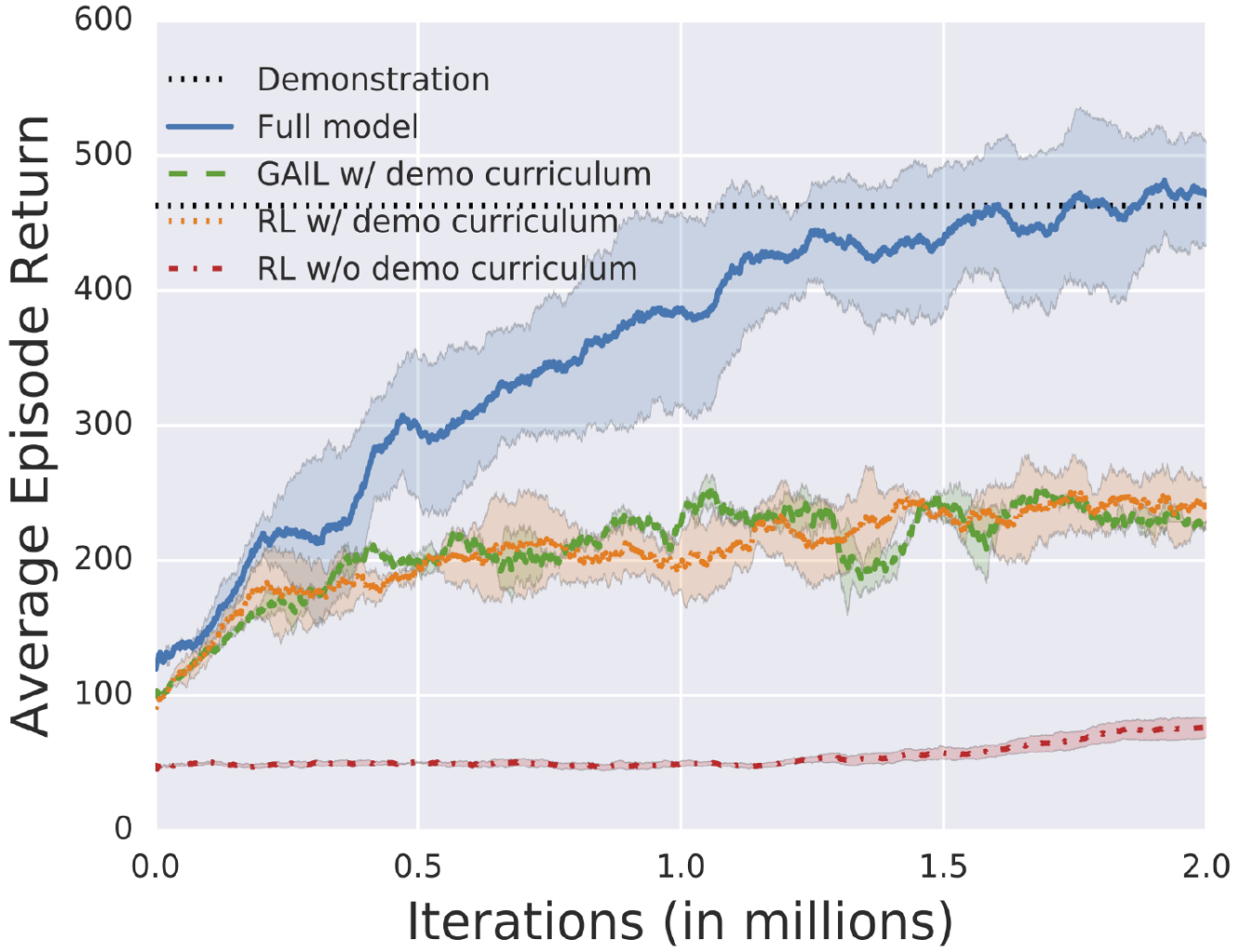
real environment



[ZWM+18]

IRL Example

Results - Block stacking



[ZWM+18]

Combining Reinforcement Learning and Imitation Learning

Comparison

BC Approach

- Behavioral Cloning
- Simulation only
- **Goal**: improve training performance and task complexity

IRL Approach

- Inversed Reinforcement Learning
- Policies transferred to real robot
- **Goal**: improve result performance and task complexity

Conclusion

- Behavioral cloning is a convenient option to directly mimic experts behavior
- Inverse reinforcement learning is able to learn expert policies
- More complex reinforcement learning tasks can be realized
- When combined with reinforcement learning, demonstrators can be outperformed

Useful Links

GAIL explained in a blog post:
`https://medium.com/@sanketgujar95/generative-adversarial-imitation-learning-266f45634e60`

Behavioral Cloning explained in a blog post:
`https://medium.com/@ksakmann/behavioral-cloning-make-a-car-drive-like-yourself-dc6021152713`

Source code and model of behavioral cloning based self-driving car:
`https://github.com/ksakmann/CarND-BehavioralCloning`

References

[Haw50] TH Hawkins, *Opening of milk bottles by birds*, Nature **165** (1950), no. 4194, 435–436.

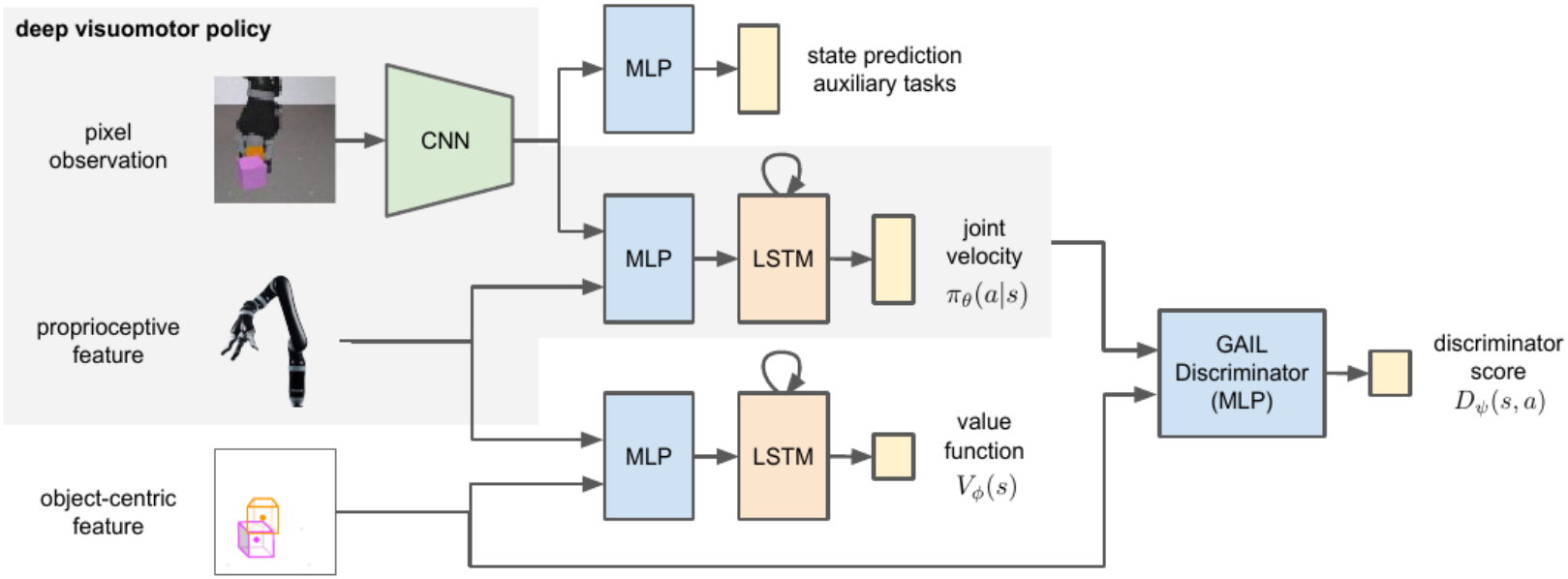
[NMA+18] Ashvin Nair, Bob McGrew, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel, *Overcoming exploration in reinforcement learning with demonstrations*, 2018 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2018, pp. 6292–6299.

[ZWM+18] Yuke Zhu, Ziyu Wang, Josh Merel, Andrei Rusu, Tom Erez, Serkan Cabi, Saran Tunyasuvunakool, János Kramár, Raia Hadsell, Nando de Freitas, et al., *Reinforcement and imitation learning for diverse visuomotor skills*, arXiv preprint arXiv:1802.09564 (2018).

Prediction in Collaboration

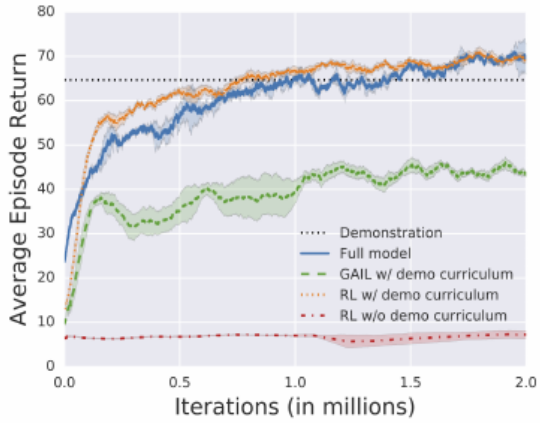
https://www.kuka.com/-/media/kuka-corporate/images/industries/case-studies/schwingenmontage/flexfellow_mrk_header.jpg

IRL Application
Network Structure

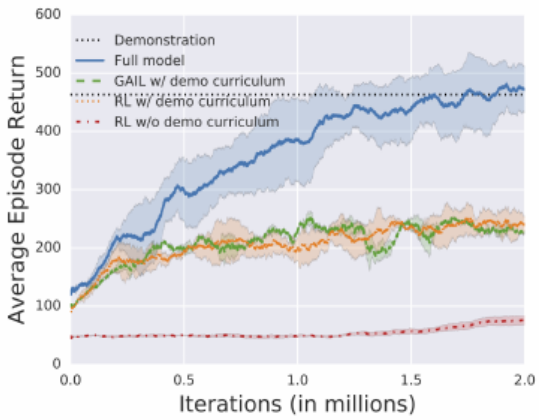


[ZWM+18]

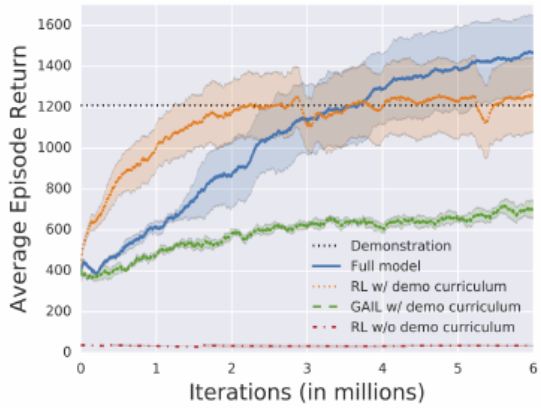
IRL Application
Full Results



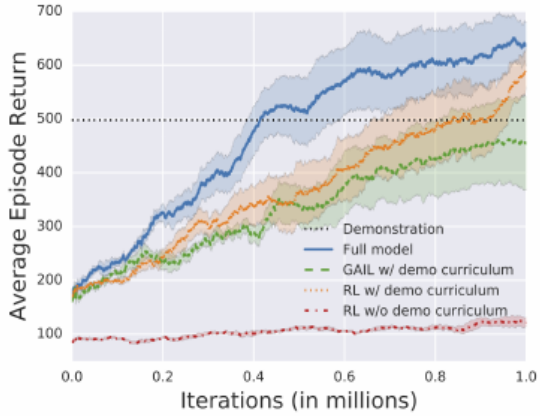
(a) Block lifting



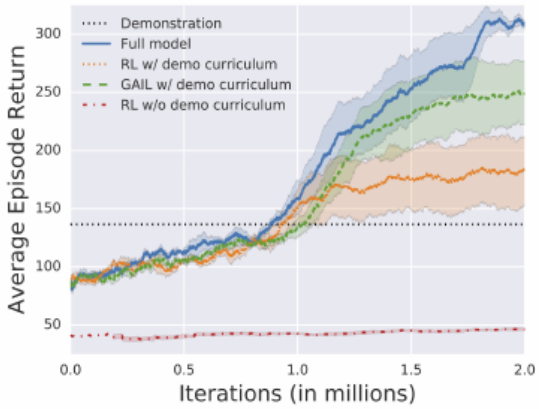
(b) Block stacking



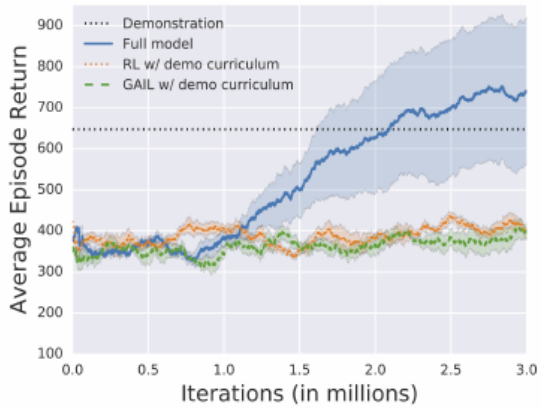
(c) Clearing table with blocks



(d) Clearing table with a box



(e) Pouring liquid



(f) Order fulfillment

[ZWM+18]