EFFICIENT MOTION-BASED TASK LEARNING

Nicholas Malone[1], Aleksandra Faust[1], Brandon Rohrer[2] , John Wood[1], Lydia Tapia[1]

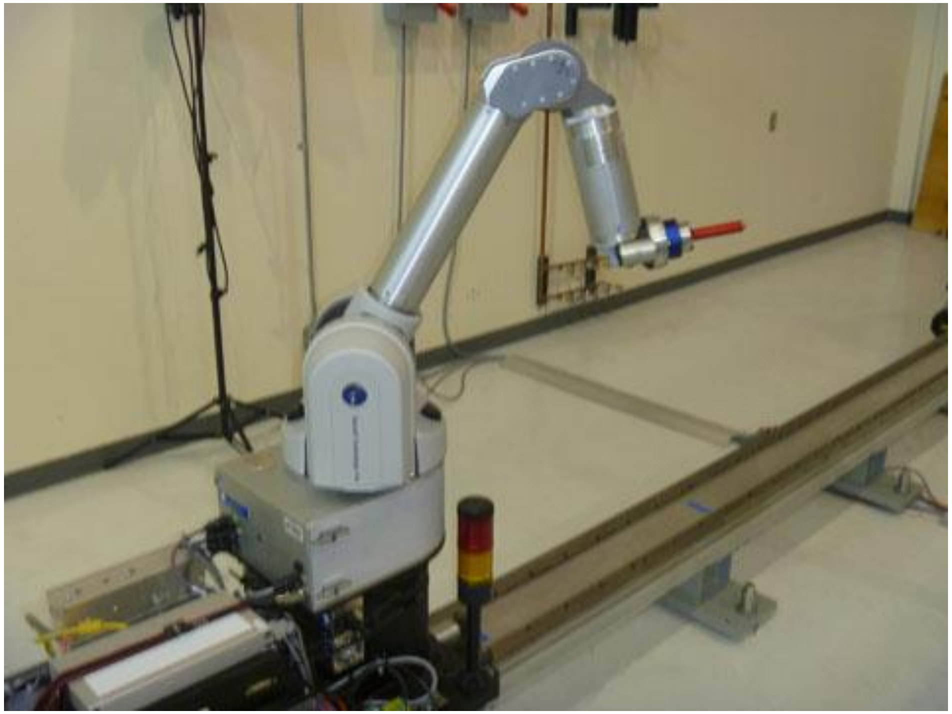
[1]University of New Mexico

[2]Sandia National Laboratories

12 October 2012

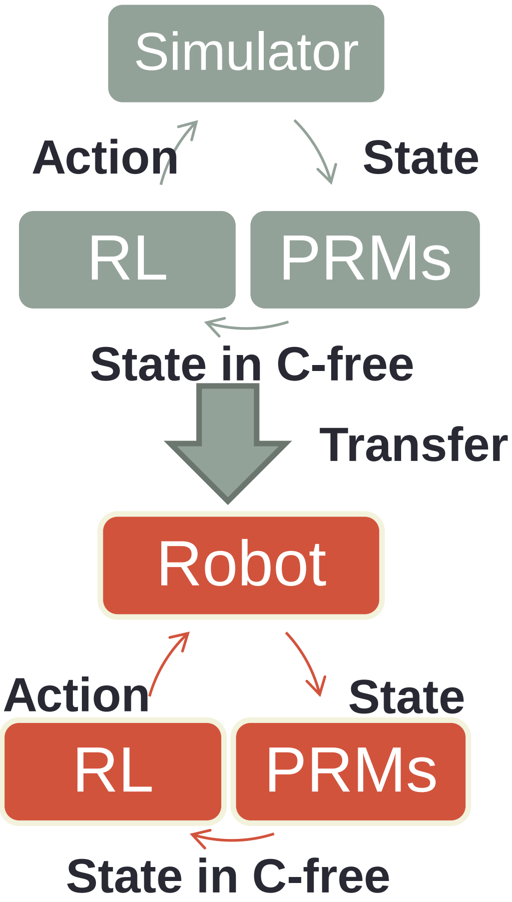
Motivation

- Robotic arm movement in dynamic environment
- Challenges:
  - High dimensionality
  - Environmental noise
  - Hardware imperfections
  - Hardware dynamics changes
  - Task changes over time



Task Learning Framework

- Characteristics
  - Adaptive
  - Safe on physical system
  - Real-time performance
  - Short training phase on hardware
- Based on
  - Probabilistic Roadmaps (PRMs)
  - Reinforcement learning
  - Transfer learning

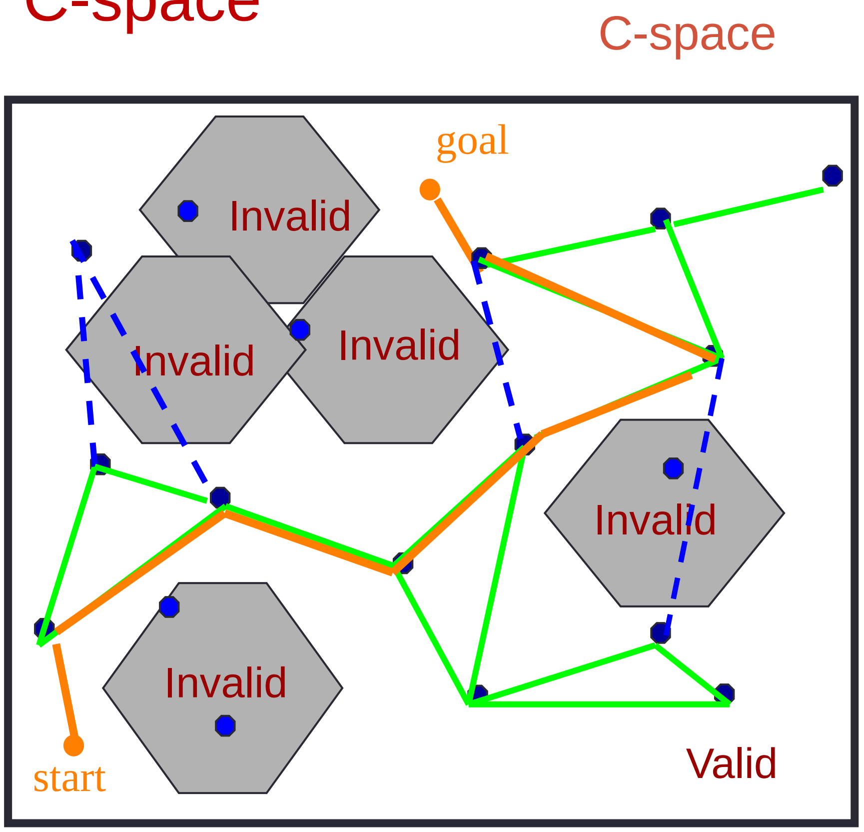


[Malone et al. *ICRA* 2012]

PRM

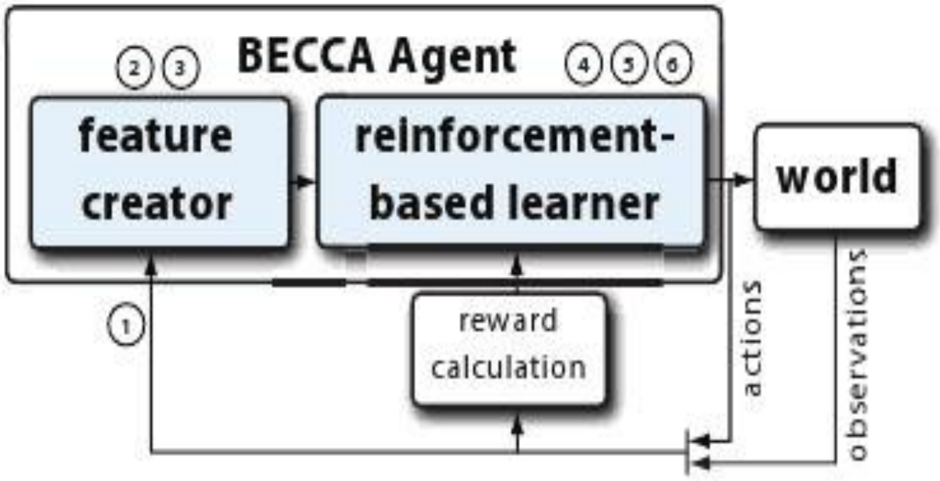
[Kavraki et al. *IEEE Trans. Robot. Automat.* 1996]

- Approximates the topology of C-space
- PRM
  - Nodes
    - Random samples
  - Edges
    - Paths fully in C-free space
    - Produced by a local planner
  - Pre-computed once
- RL receives
  - Current node in PRM
  - Allowed transitions



Reinforcement Learner

- BECCA Agent
  - Unsupervised feature creator
  - Online reinforcement learner
  - Model based
  - Working memory concept
  - Solves higher order Markov decision process



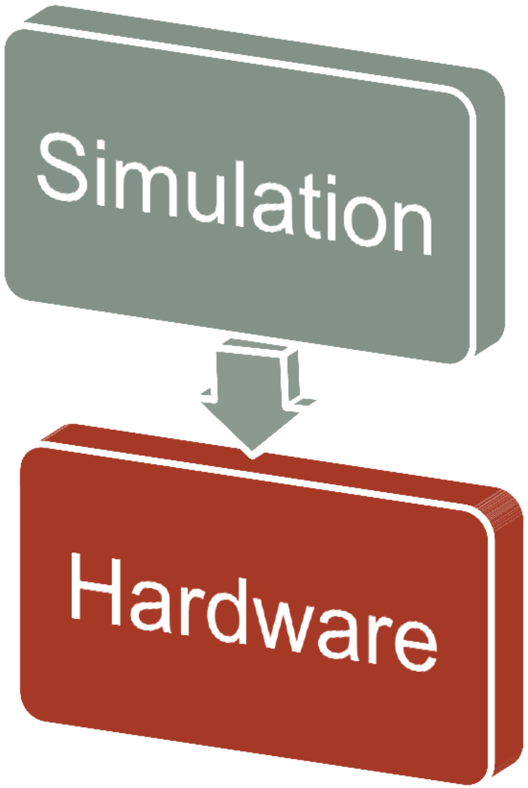
- Prior Applications
  - WAM robot pointing
  - Surveyor SRV-1 robot hide and seek
  - CoroBot hand-eye coordination
  - Visual tracking
  - Visual and audio integration

[Rohrer et al. *BICA* 2010]
[Rohrer *ICDL* 2011]
[Rohrer *AAAI Spring Symposium* 2012]

Transfer Learning

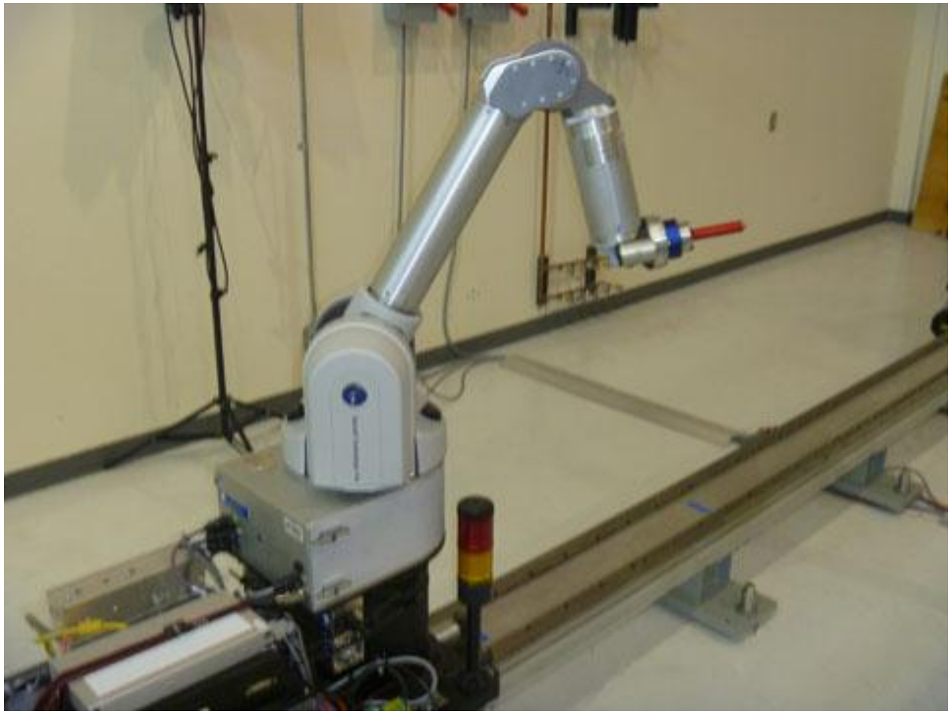
[Taylor and Stone, 2009]

- Transfer from simulation to hardware
  - PRM, RL agent, reward function
- Simulation
  - Many iterations
  - Fast per step
- Task performs well and safe for robot
- Hardware
  - Simultaneous learning and task execution
  - Fewer iterations
  - Steps take longer



Experiments - Methodology

- On Barrett 7DoFs WAM
- Learning performance
  - Cumulative reward
- Transfer learning impact
  - Time saved by the transfer
  - Performance boost from the transfer
- Adaptability
  - Time to recover after environment change



- Pointing task evaluated on both stationary and non-stationary target

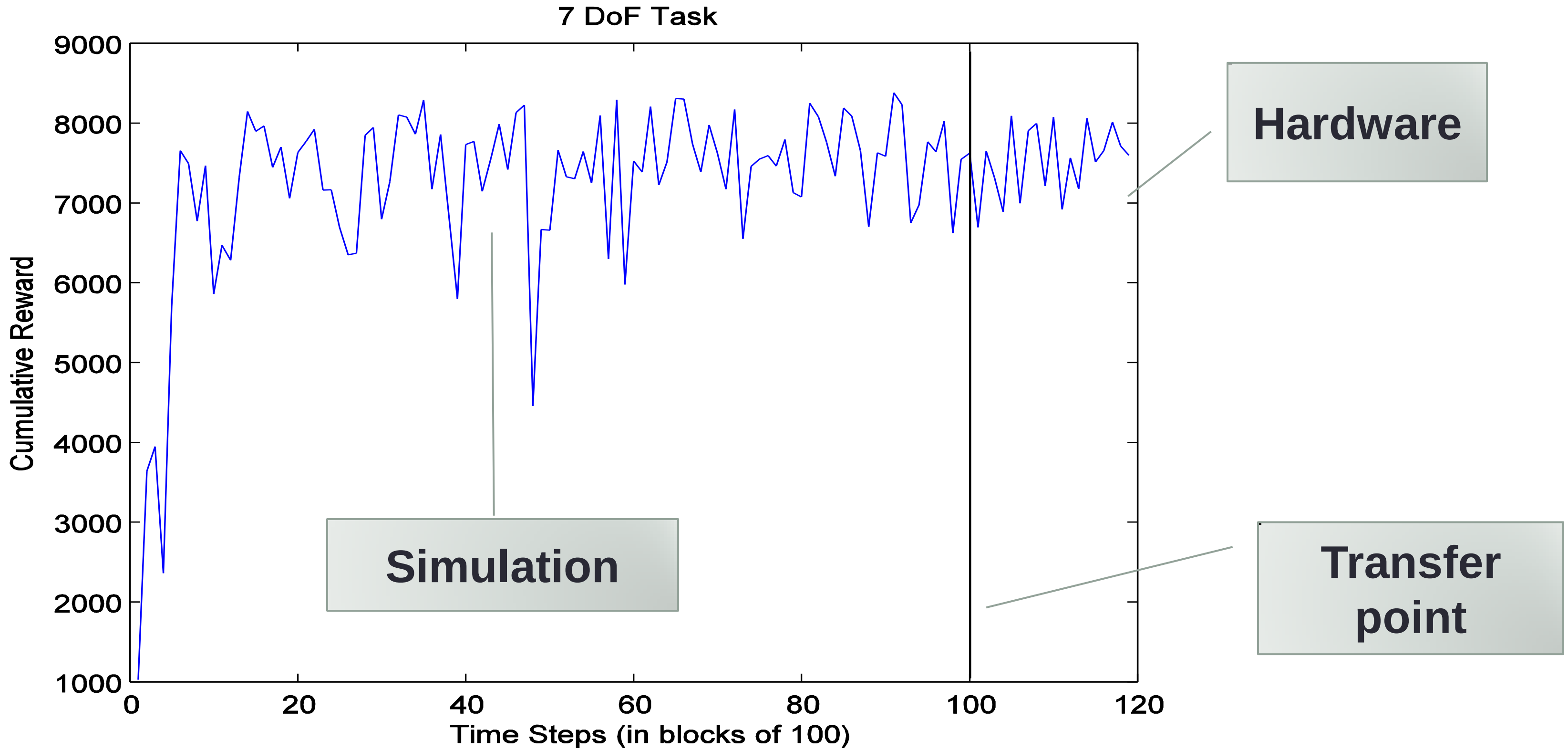
Experiments - Parameters

- Reinforcement learner:
  - Exploration rate: 30%
  - Reward: gradient reward based on the distance from the goal

- PRM
  - 50 nodes
  - Connected to 3 nearest neighbors

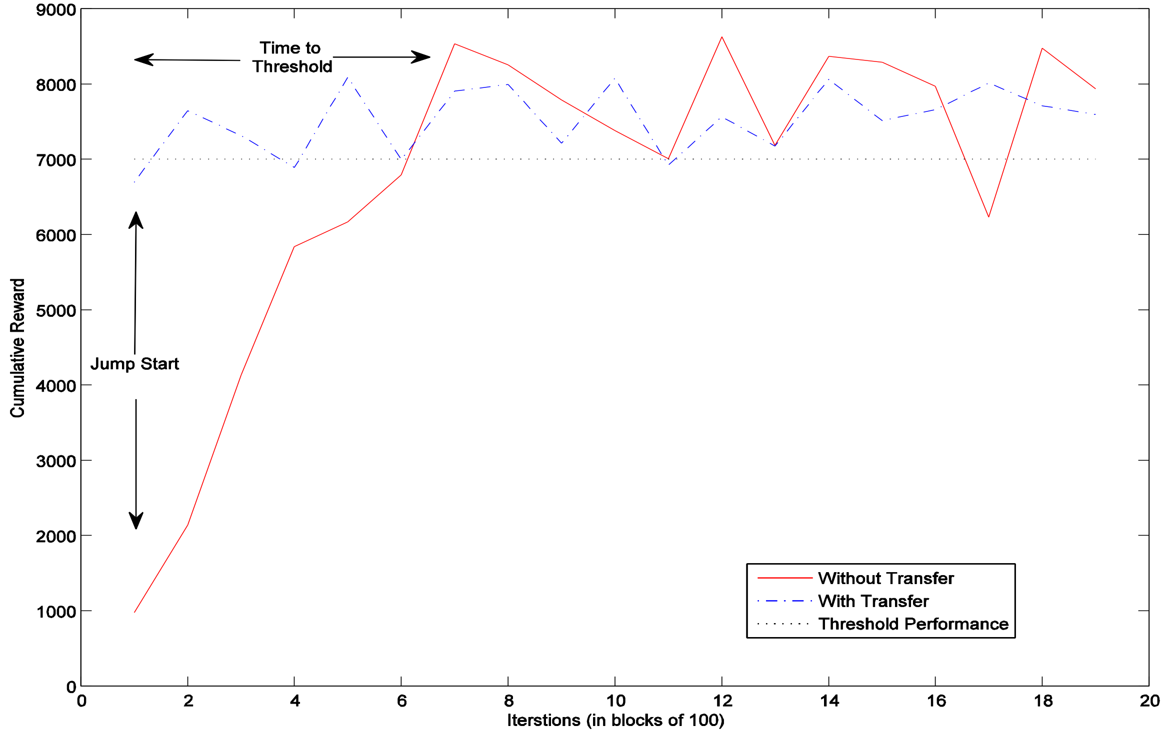
- Local planner
  - Simulation: linear interpolation
  - Hardware: on-board WAM controller

Stationary Task: Learning Performance



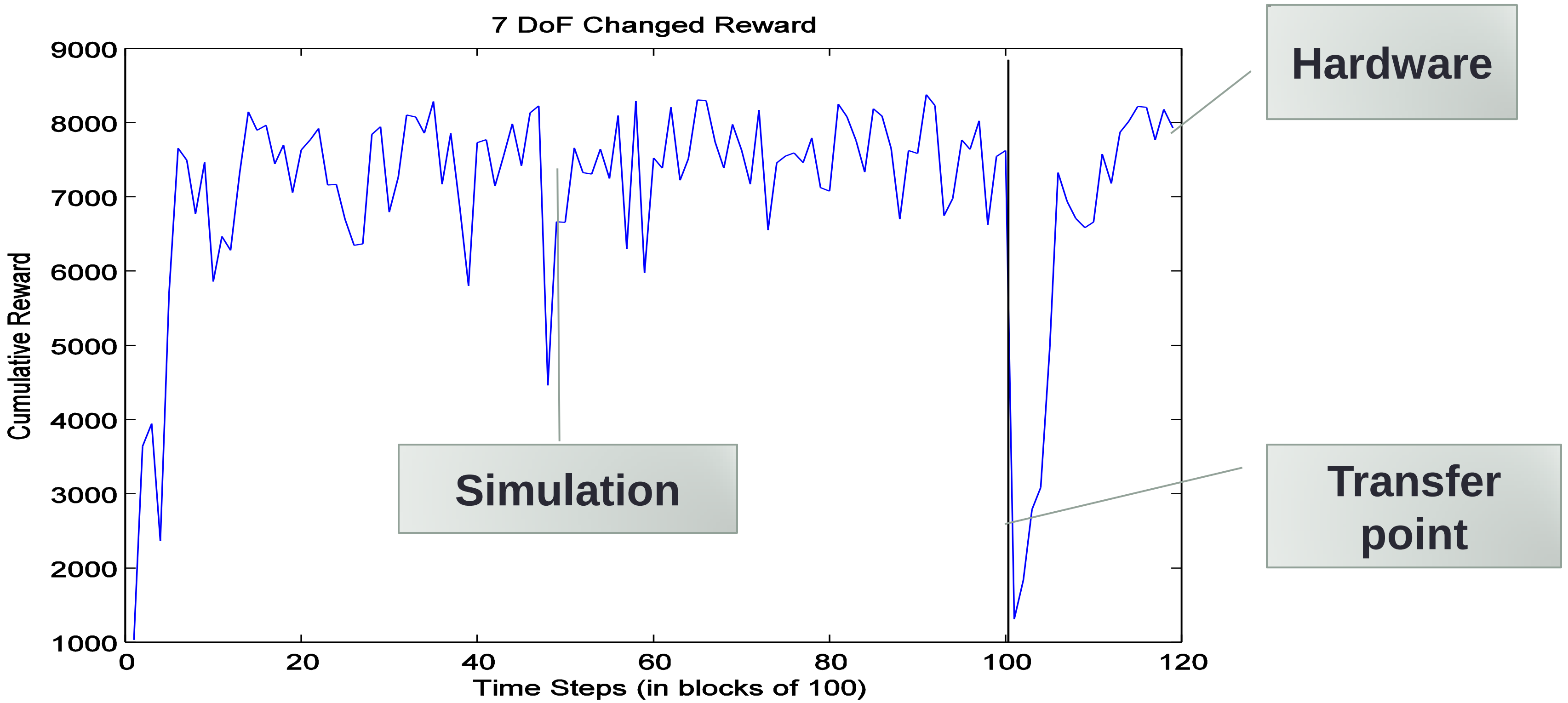
- Cumulative reward before and after the transfer from simulation to hardware
- No loss of learning performance after the transfer

Stationary Task: Transfer Learning Impact



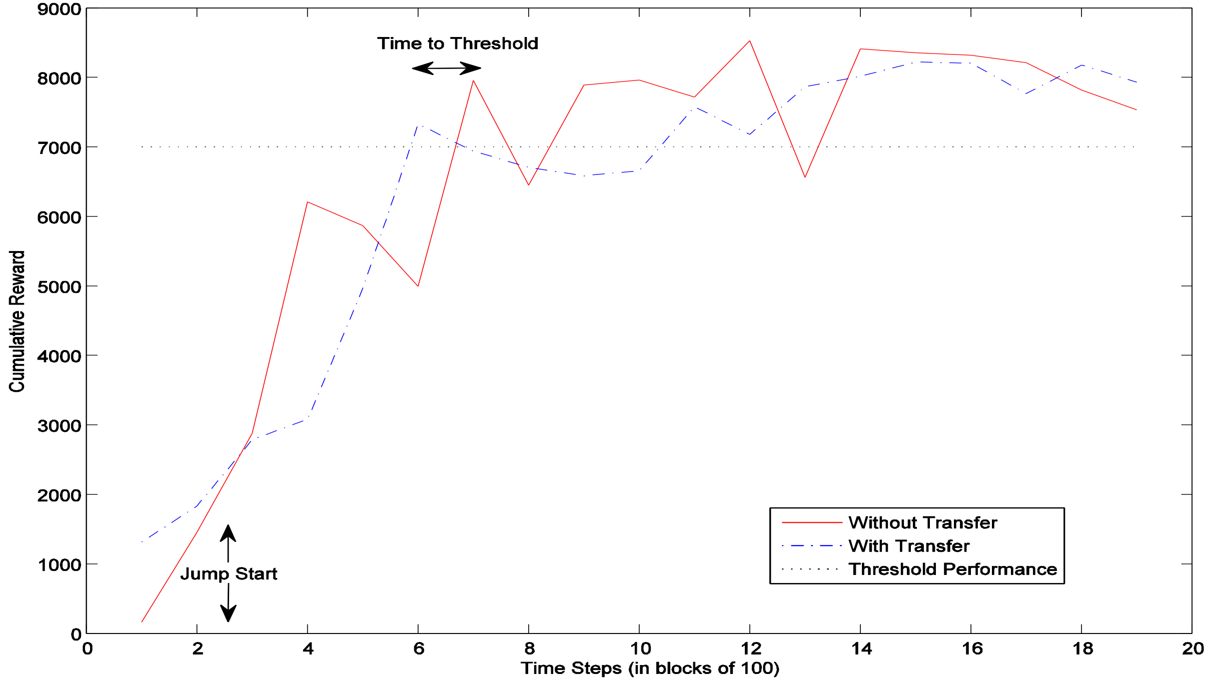
- Jumpstart 5700 reward points
- 500 steps less until threshold
- Significant time savings when training in simulation first

Non-stationary Task: Learning Performance



- Cumulative reward before and after the transfer from simulation to hardware
- Fast recovery after the target moves

Non-stationary Task: Transfer Learning Impact



- Jumpstart 1300 reward points
- 100 steps less until threshold
- Previous training still has benefits even when target changes

Running time: Simulation vs. Hardware

Task	Simulation (min)	Hardware (min)
2000 steps	23	122

- Simulation 5 times faster
- Training in simulation saves 100 minutes over 2000 steps

Efficient Motion-based Task Learning

https://www.cs.unm.edu/amprg/Research/RobotLearning/

Conclusion

- Motion-based task learning framework
  - PRM
  - Online reinforcement learner
  - Transfer learning
- Implemented on Barrett 7DoF WAM
  - Pointing task with stationary target
  - Pointing task with non-stationary target
- Results
  - Adapts to the new changes in the environment
  - Performs well in high dimensional spaces
  - Safe for hardware

Q&A

Thank you!

THE UNIVERSITY *of* NEW MEXICO

Sandia National Laboratories

Backup

Quadrocopter Learning

- Reinforcement learning approach for generating trajectories with minimal residual oscillations (swing-free) for rotorcraft carrying a suspended load that allows the trained agent to be transferred to a variety of models, state and action spaces and produce a number of different trajectories.

- https://www.youtube.com/watch?v=DIs7qJAg91o