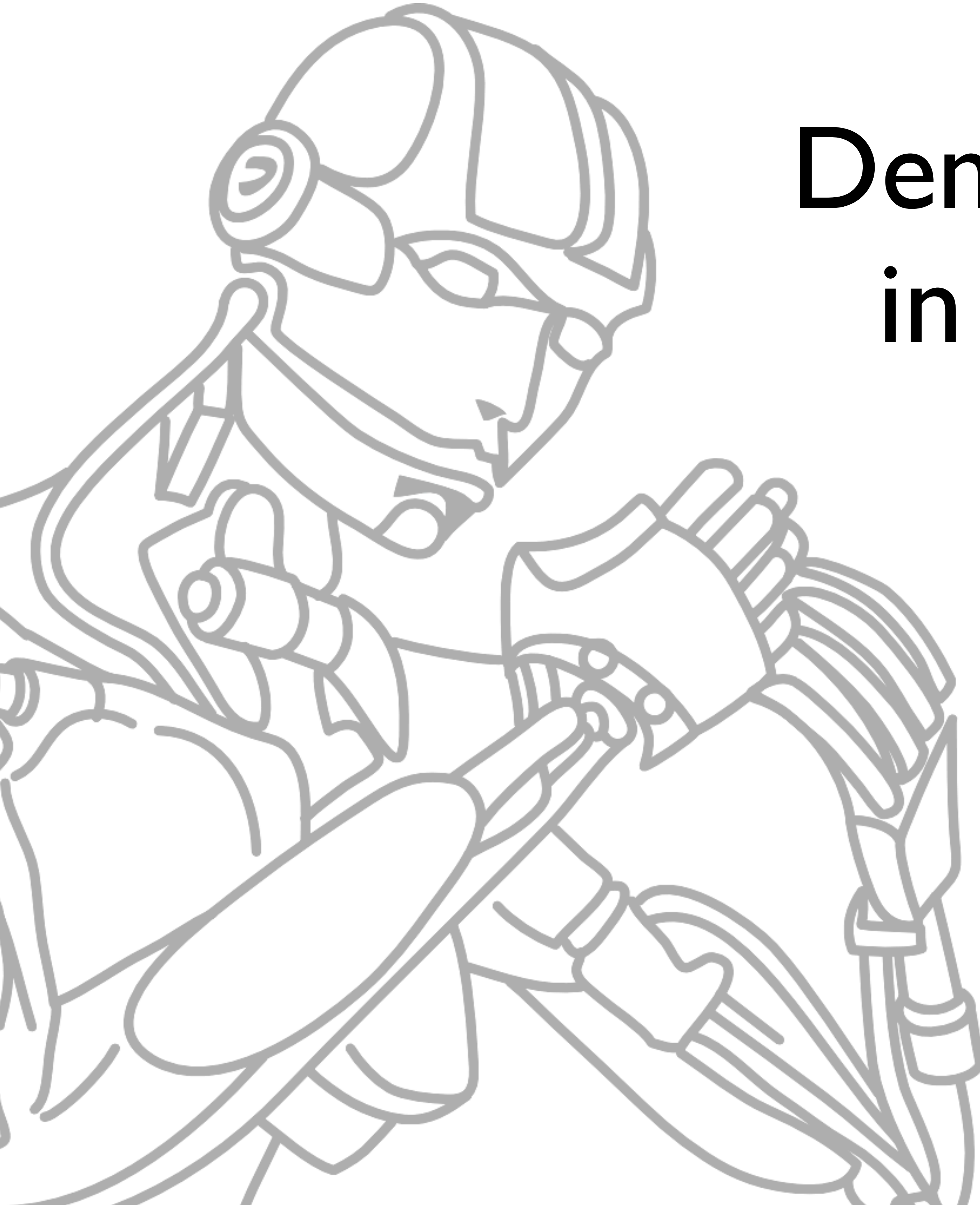


Generalizing Demonstrated Actions in Manipulation tasks

Oliver Kroemer

Renaud Detry

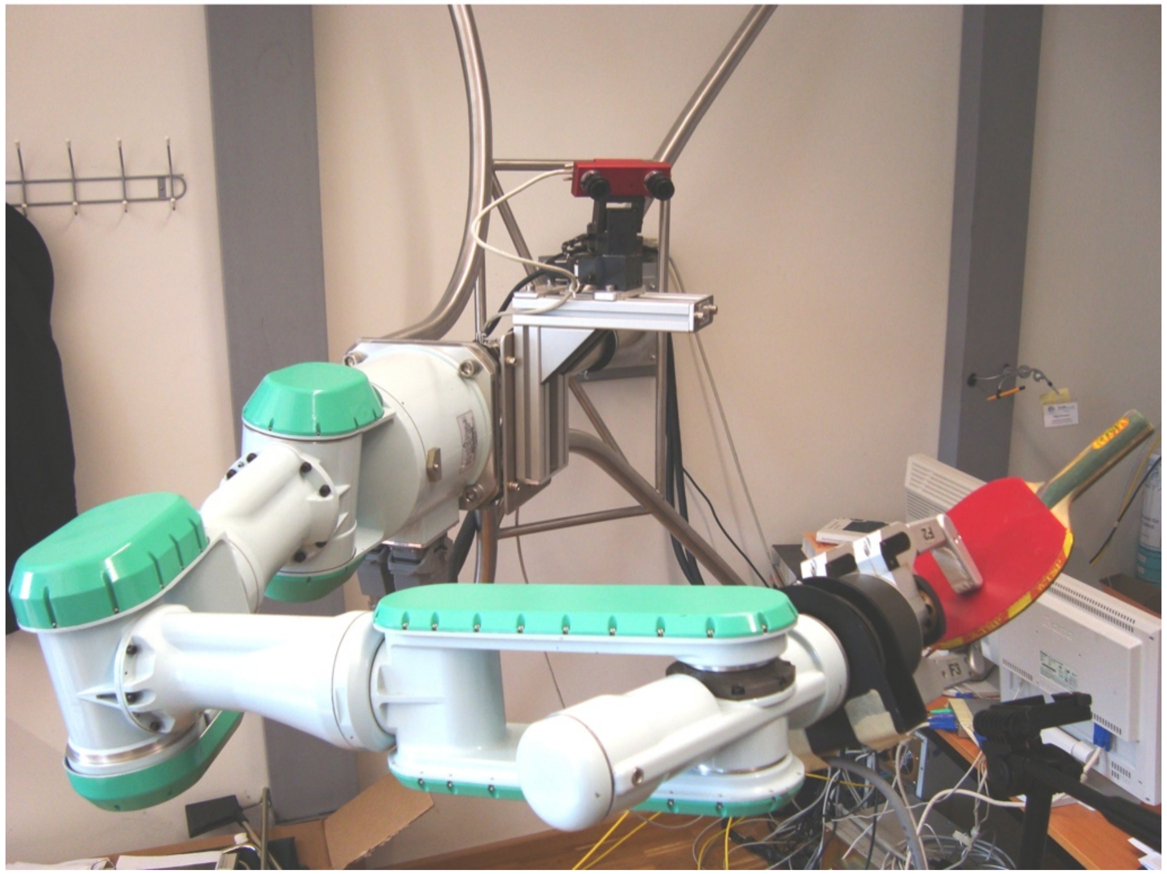
Justus Piater

Jan Peters

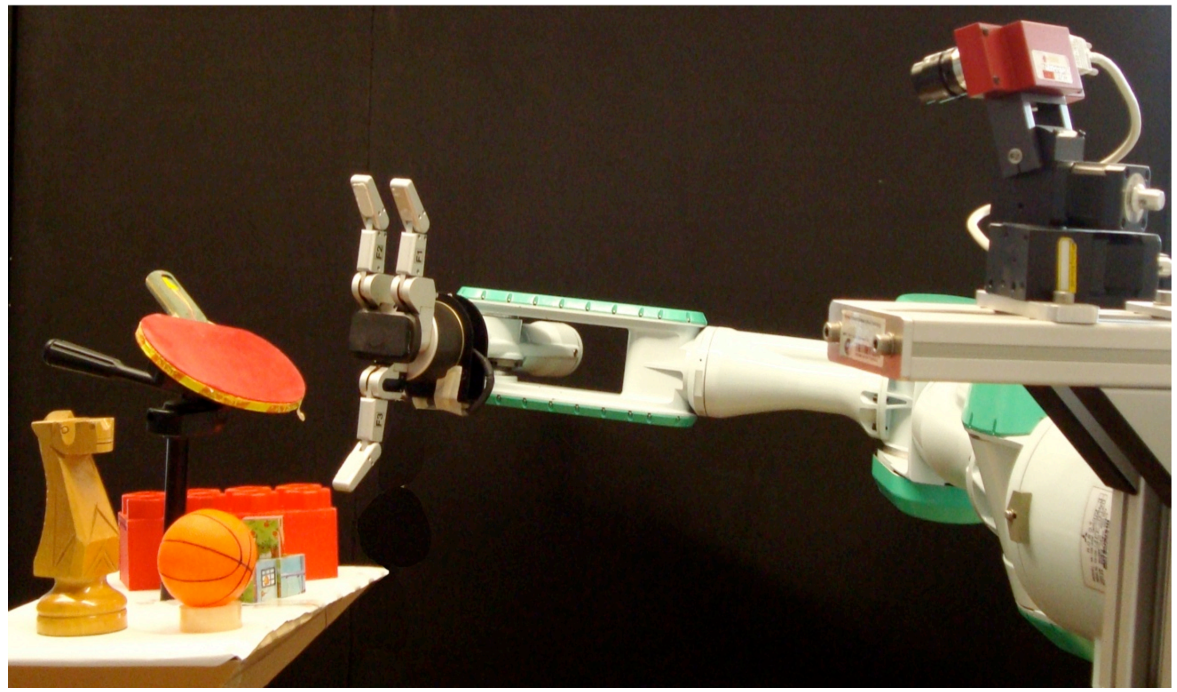
Max Planck Institute for Biological Cybernetics
University of Liege

Motivation

- Humans are good at grasping
  - Easier to demonstrate movements than code them
  - Transfer by imitation learning
  - Give robots positive traits
- How to generalize?
  - Task's variables change
  - Task's objects change



Aims



- Suitable representation of movements
  - Generalize to new object locations and orientations
  - Robust and adaptive motor primitives
- Learn adaptation of movement to new object
  - Use robot's experiences to improve grasp
  - Optimize grasp for new object

Outline

I. Introduction

II. Generalizing Grasping Movements

III. Generalizing to New Objects

IV. Robot Experiments

V. Conclusion

Dynamical Systems Motor Primitives (DMPs)

- Provide adaptive movement encodings
- Transformed System:

Spring and Damper External Force

$$\ddot{y} = \alpha_z(\beta_z \tau^{-2}(g - y) - \tau^{-1}\dot{y}) + a\tau^{-2} f(x)$$

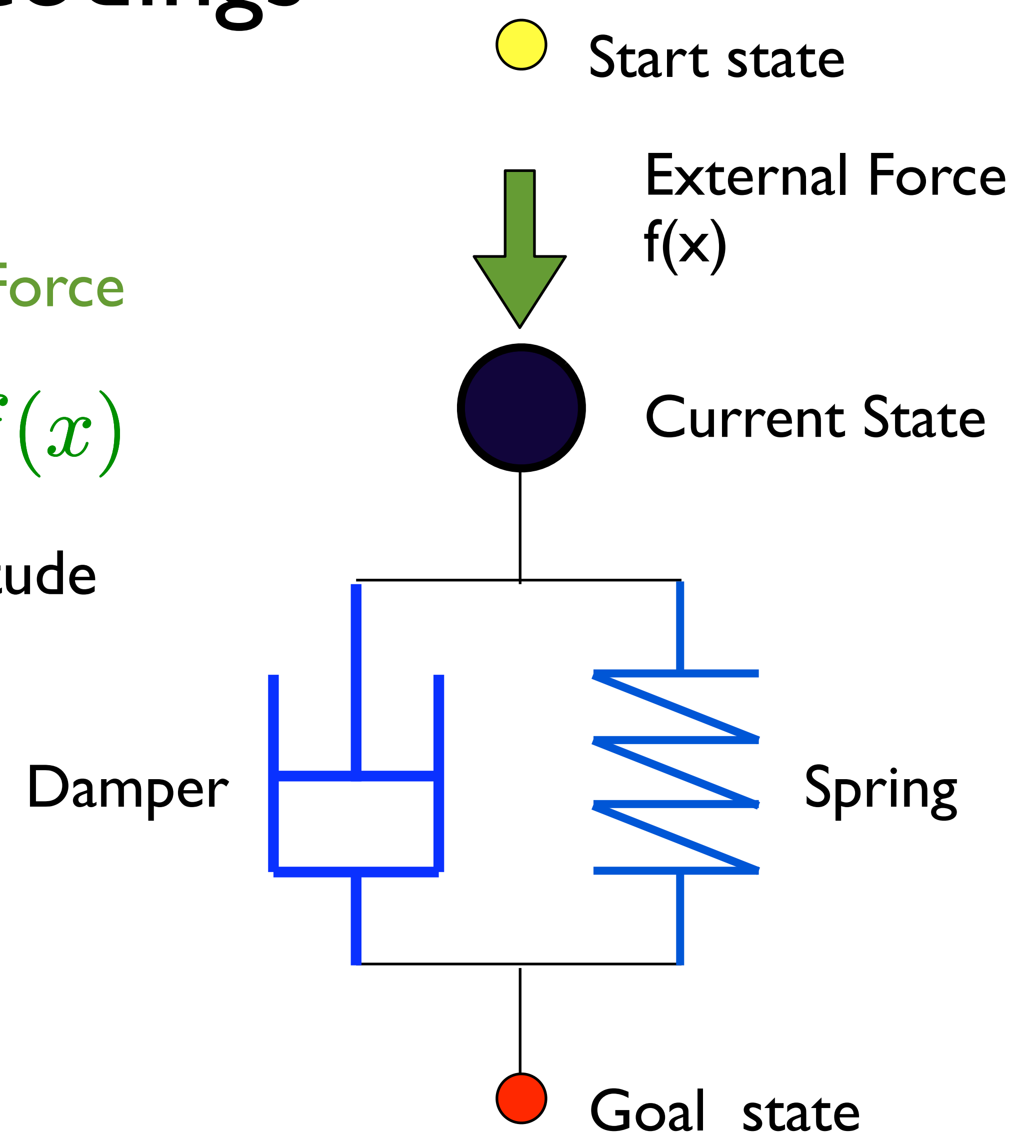
Goal State Current State Amplitude

- Canonical System:

$$\dot{x} = -\tau x$$

Time Constant Canonical State

  - Phase shifts from 1 to 0
  - Synchronizes DMPs



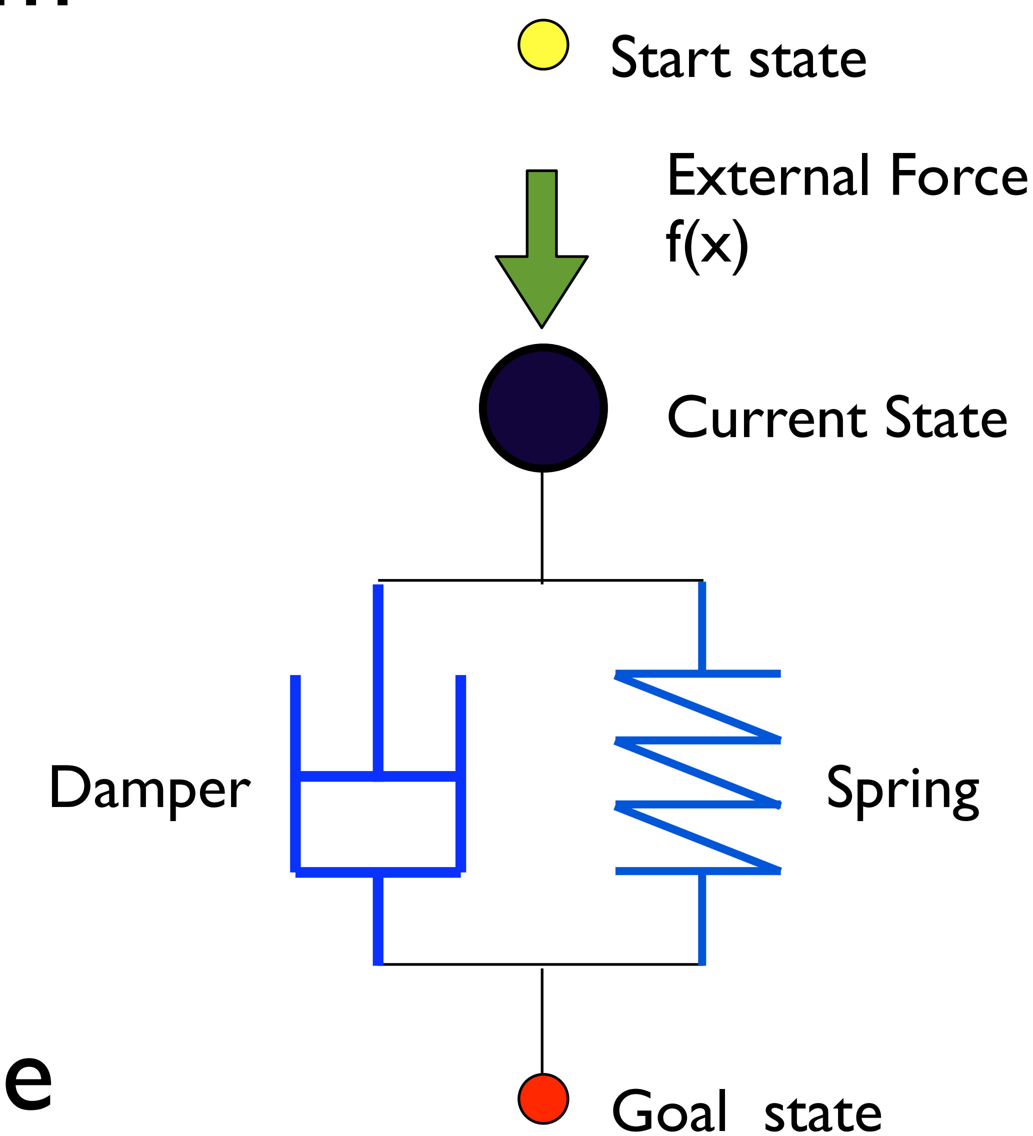
Dynamical Systems Motor Primitives (DMPs)

- Critically damped passive system
- Arbitrary trajectory with force

$$f\left(x\right)=\frac{\sum_{j=1}^{M}\psi_j\left(x\right)w_jx}{\sum_{i=1}^{M}\psi_i\left(x\right)}$$

  - Easy to learn by imitation
- DMPs are inherently stable
- Amplitude set to maintain shape

$$a=g-y_0$$

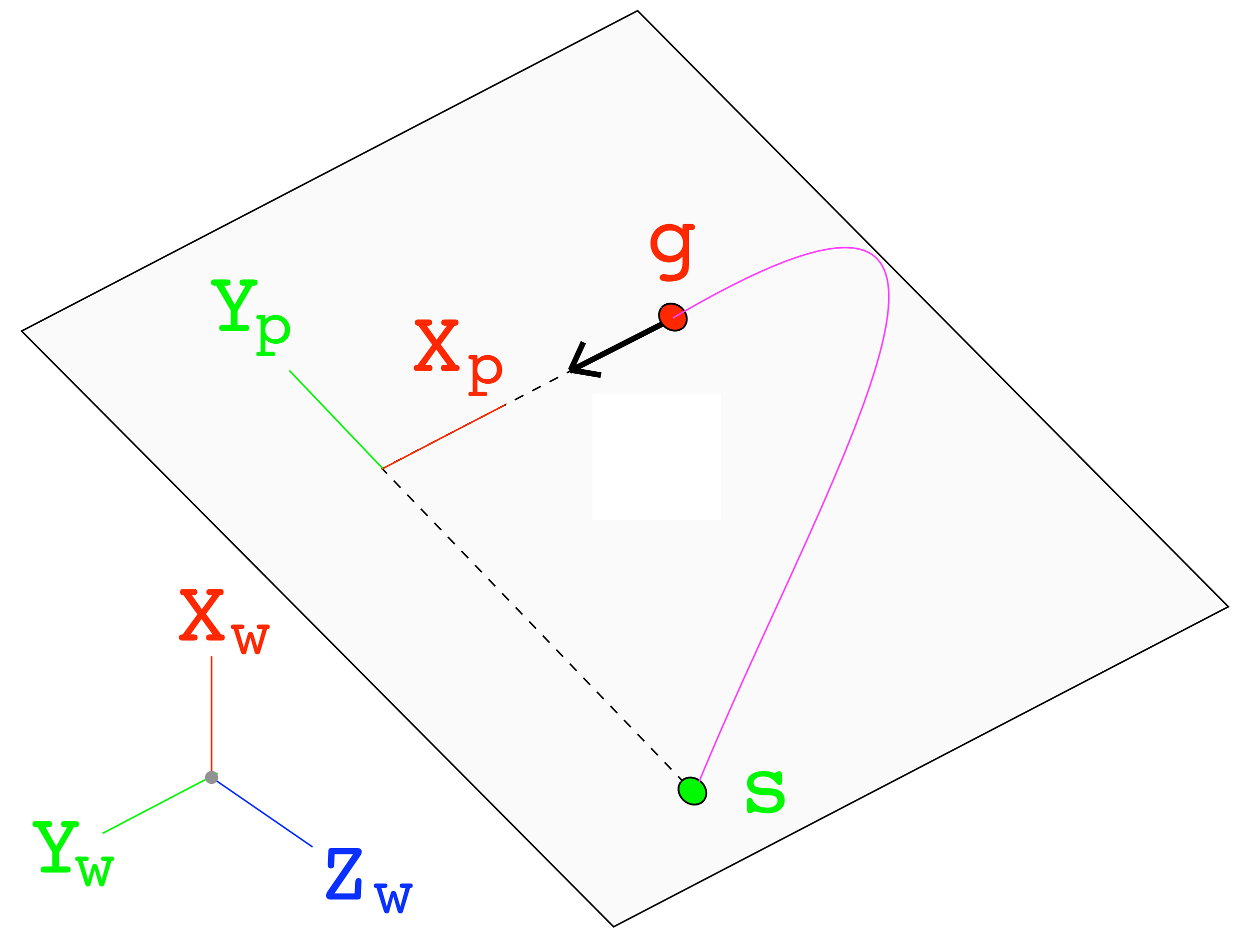


Dynamical Motor Primitives for Grasping

- Stable and always reach goal if possible
- Adapt to different goal points
- Straightforward to learn from demonstration
- Allow for preshaping the hand
  - Synchronize fingers with reaching motion
  - Grasps are generally more stable and controlled
  - Finger goals obtained from target object’s geometry
- Foundation of robust and adaptive representations

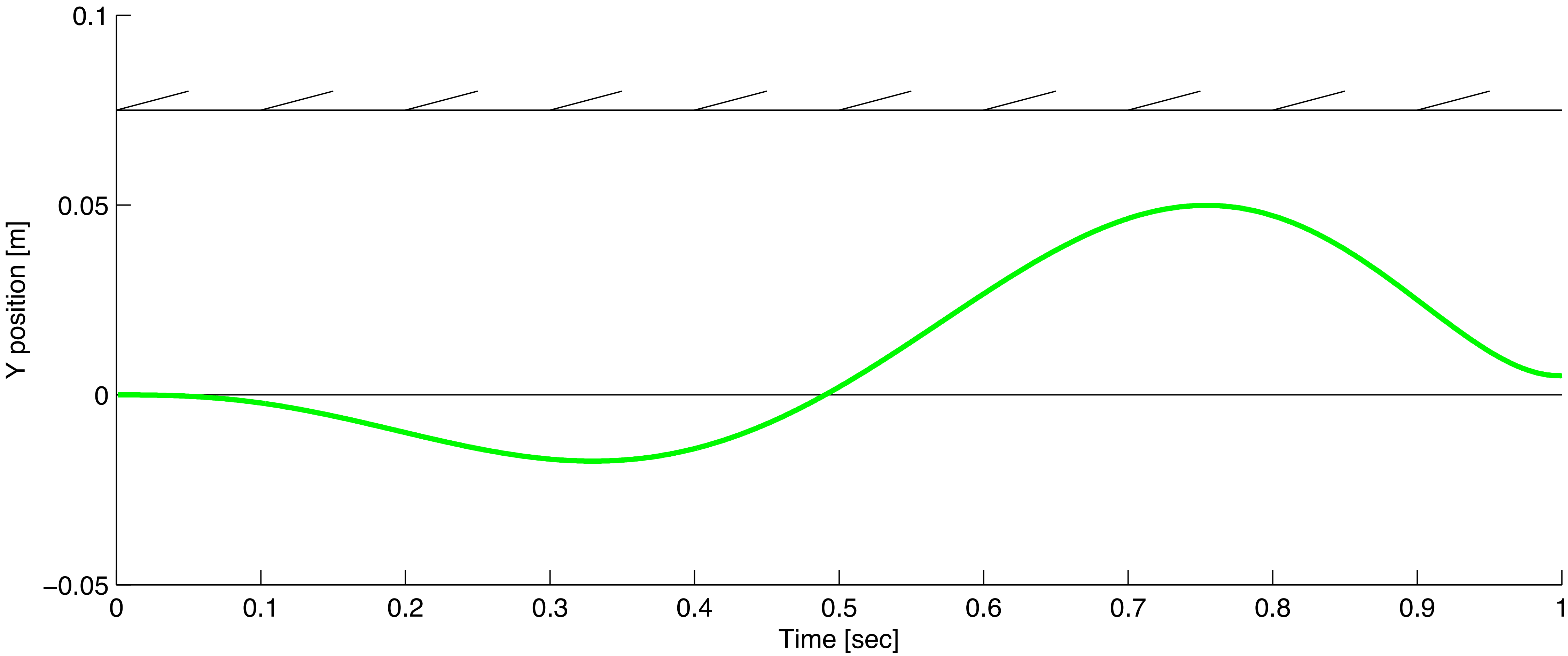
Generalizing Movements to New Grasps

- Generalize to new grasps
  - Depends on task's coordinate system
- Need specific approach direction
- Task-specific frame
- Need new amplitude:

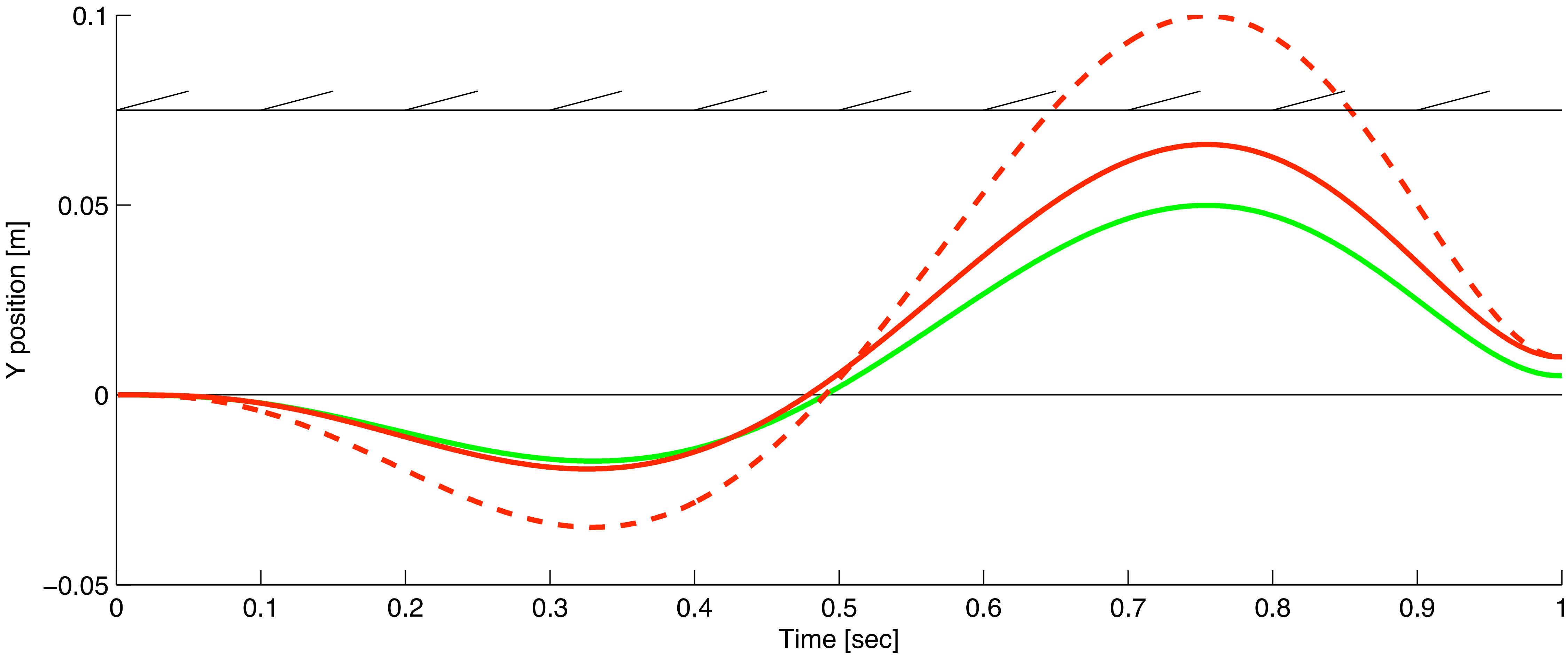


$$a = \|\eta(g - y_0) + (1-\eta)(g_T - y_{0T})\|$$

Conservative DMP Generalization



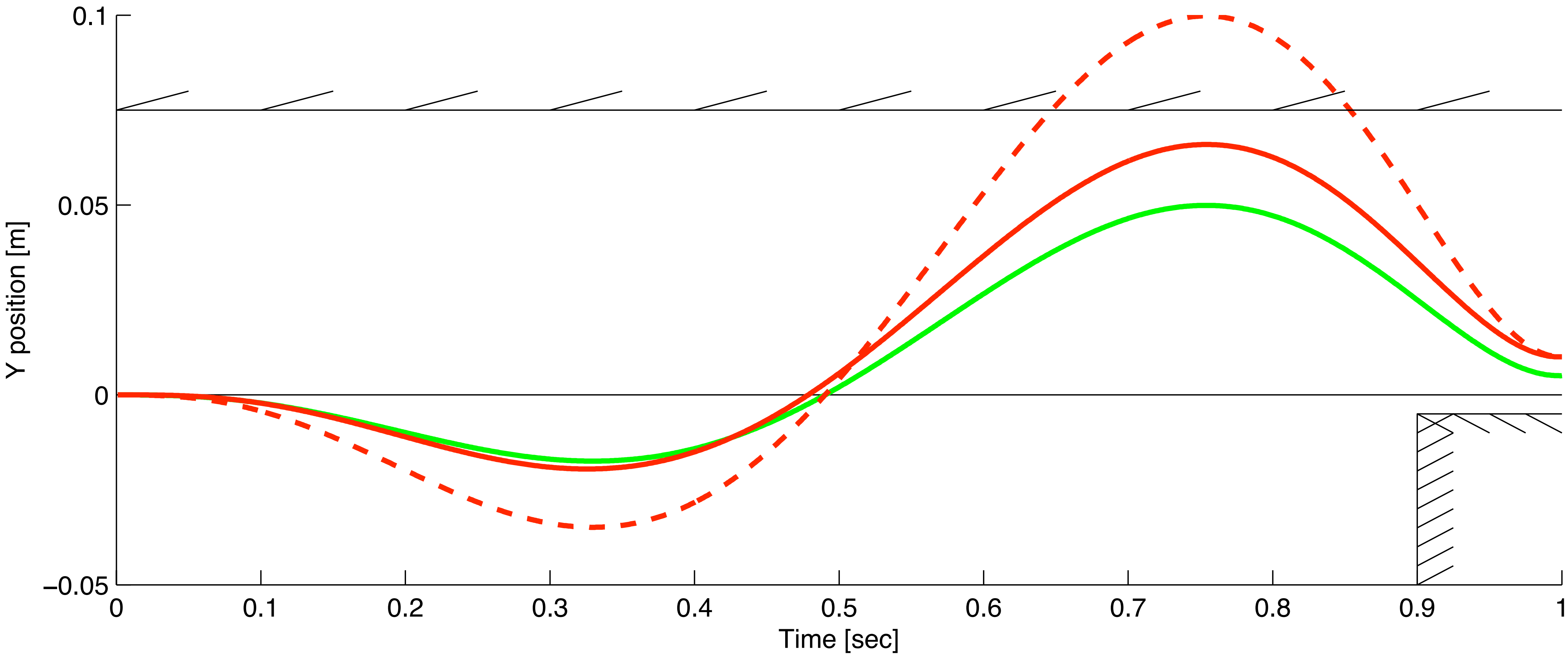
Conservative DMP Generalization



Dashed line: $a = g - y_0$

Solid line: $a = \|\eta(g - y_0) + (1 - \eta)(g_T - y_{0T})\|$

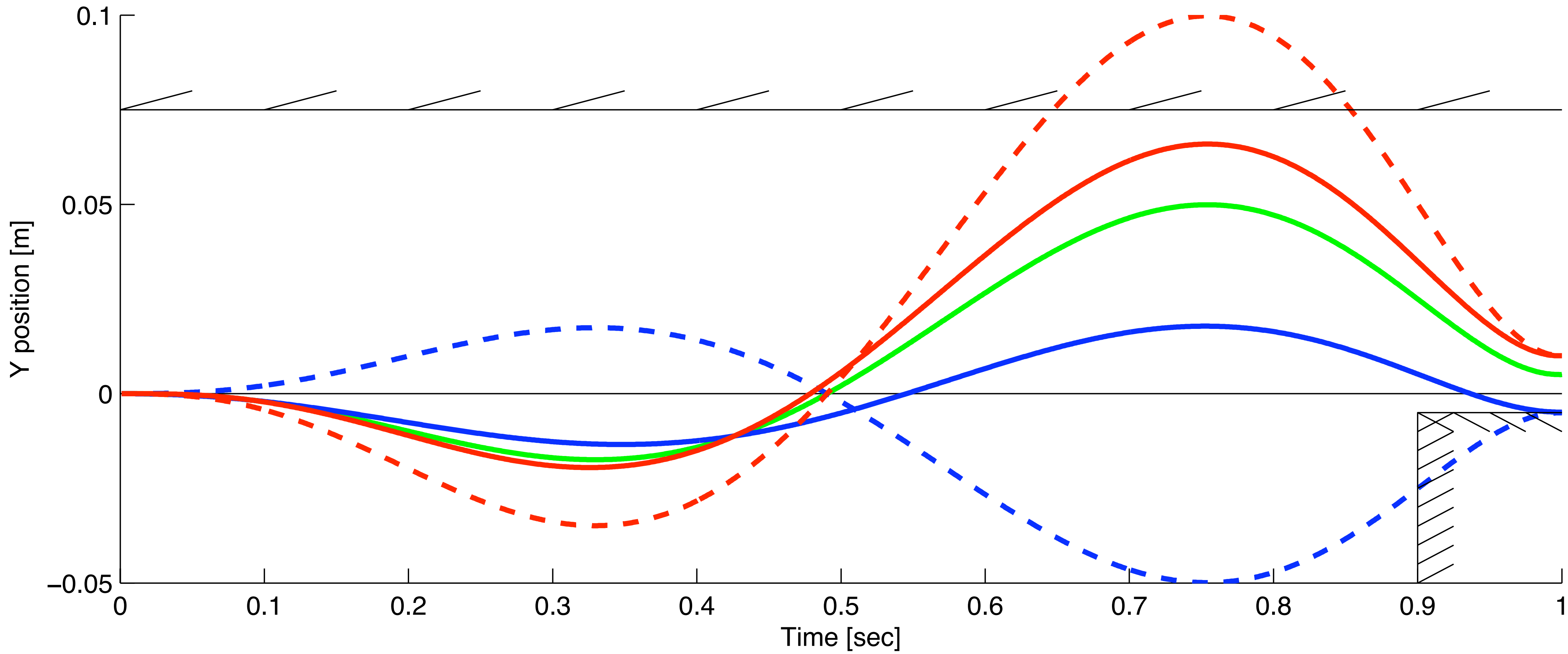
Conservative DMP Generalization



$$a = g - y_0$$

$$a = \| \eta(g - y_0) + (1 - \eta)(g_T - y_{0T}) \|$$

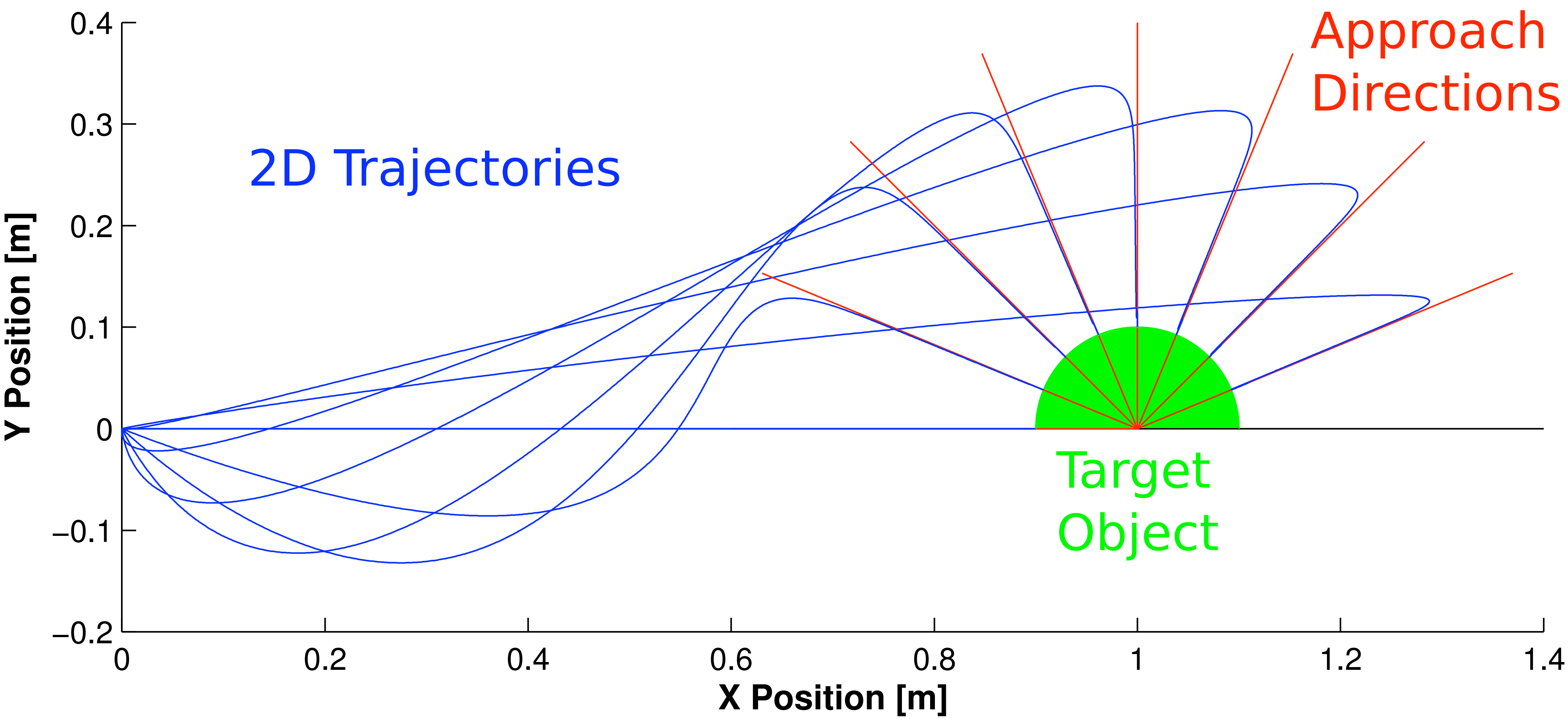
Conservative DMP Generalization



$$a = g - y_0$$

$$a = \|\eta(g - y_0) + (1 - \eta)(g_T - y_{0T})\|$$

Generalized Trajectories



Outline

I. Introduction

II. Generalizing Grasping Movements

III. Generalizing to New Objects

IV. Robot Experiments

V. Conclusion

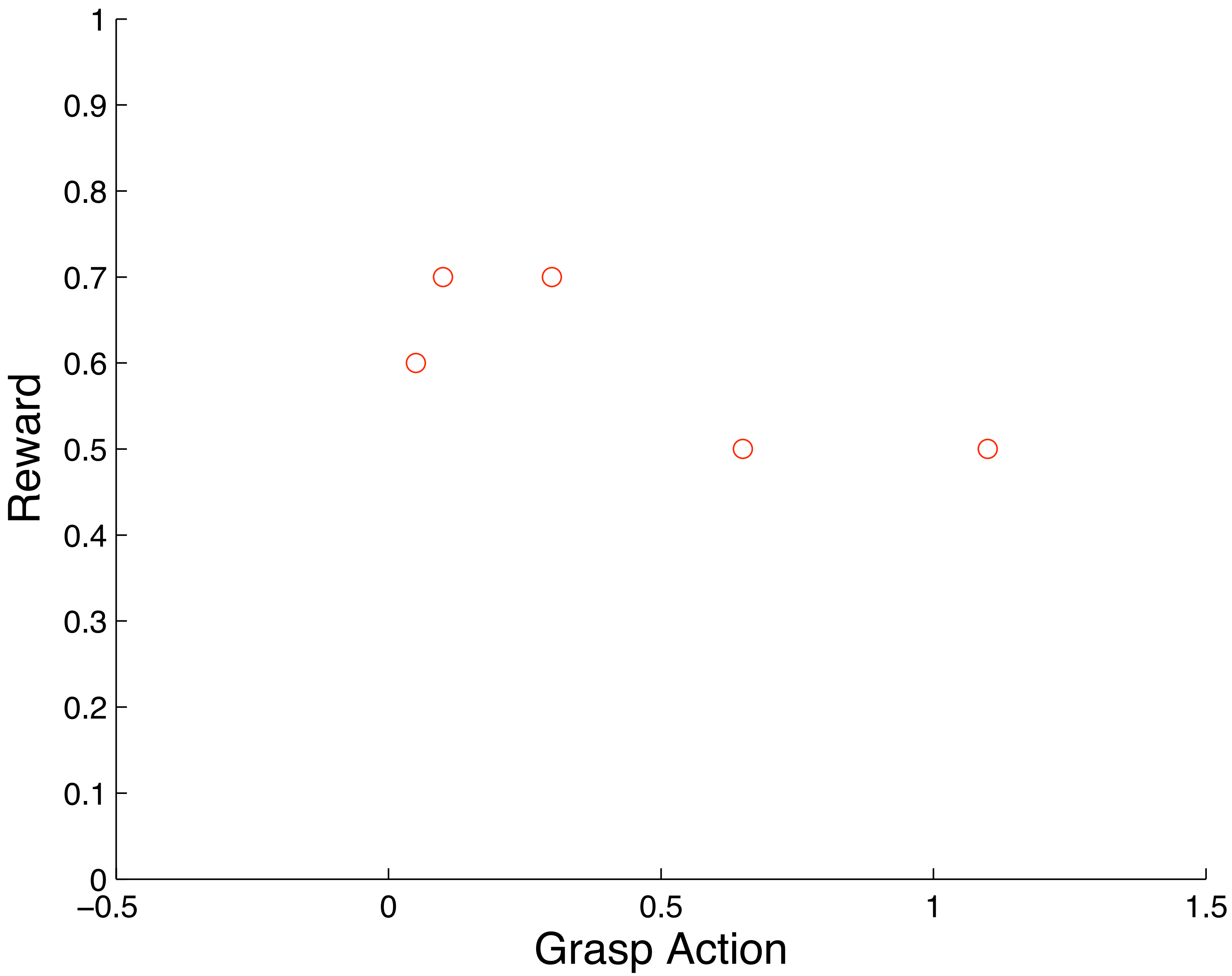
Adapting to New Objects

- Can predict/acquire new grasps on new objects
  - Shape similarity, morphing, wand, etc.
- Lose grasp quality due to change of object
  - Local changes: different geometry, texture, friction, etc.
  - Global changes: different center of mass, support plane, etc.
- Aim to quickly regain lost quality
  - Automatically improve grasps using robot's experiences
  - Optimize grasp using Continuum-Armed Bandits

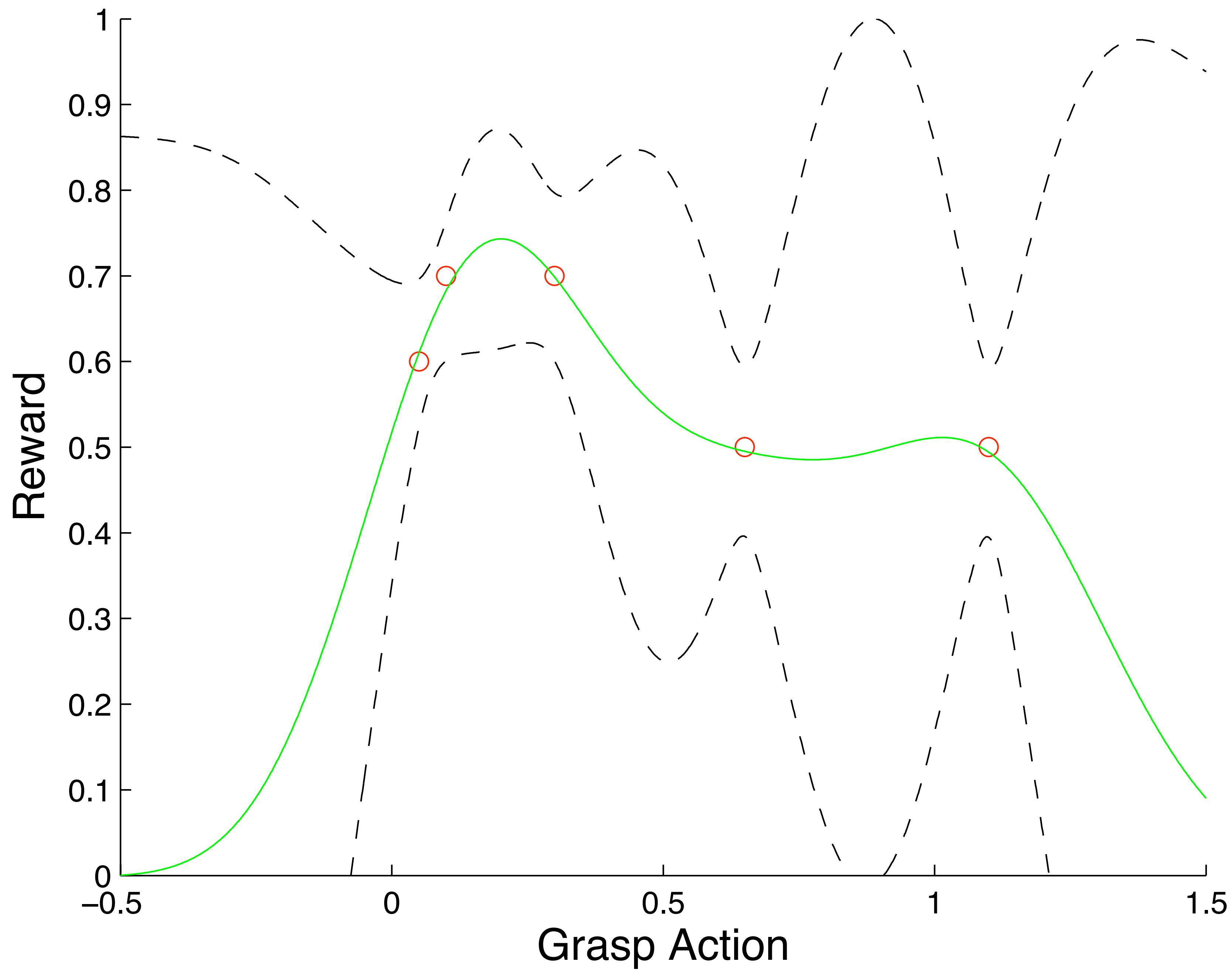
Continuum-Armed Bandits

- Problem interpreted from RL as a "bandits" problem
  - Action = Choose 6D grasp location
  - Reward = Quality of resulting grasp
- Continuum-Armed Bandit variant
- Upper Confidence Bound (UCB) Policies
  - Incorporate expected reward and confidence bound
  - Successful in discrete bandits scenario
  - Current implementations rely on discretizing space!

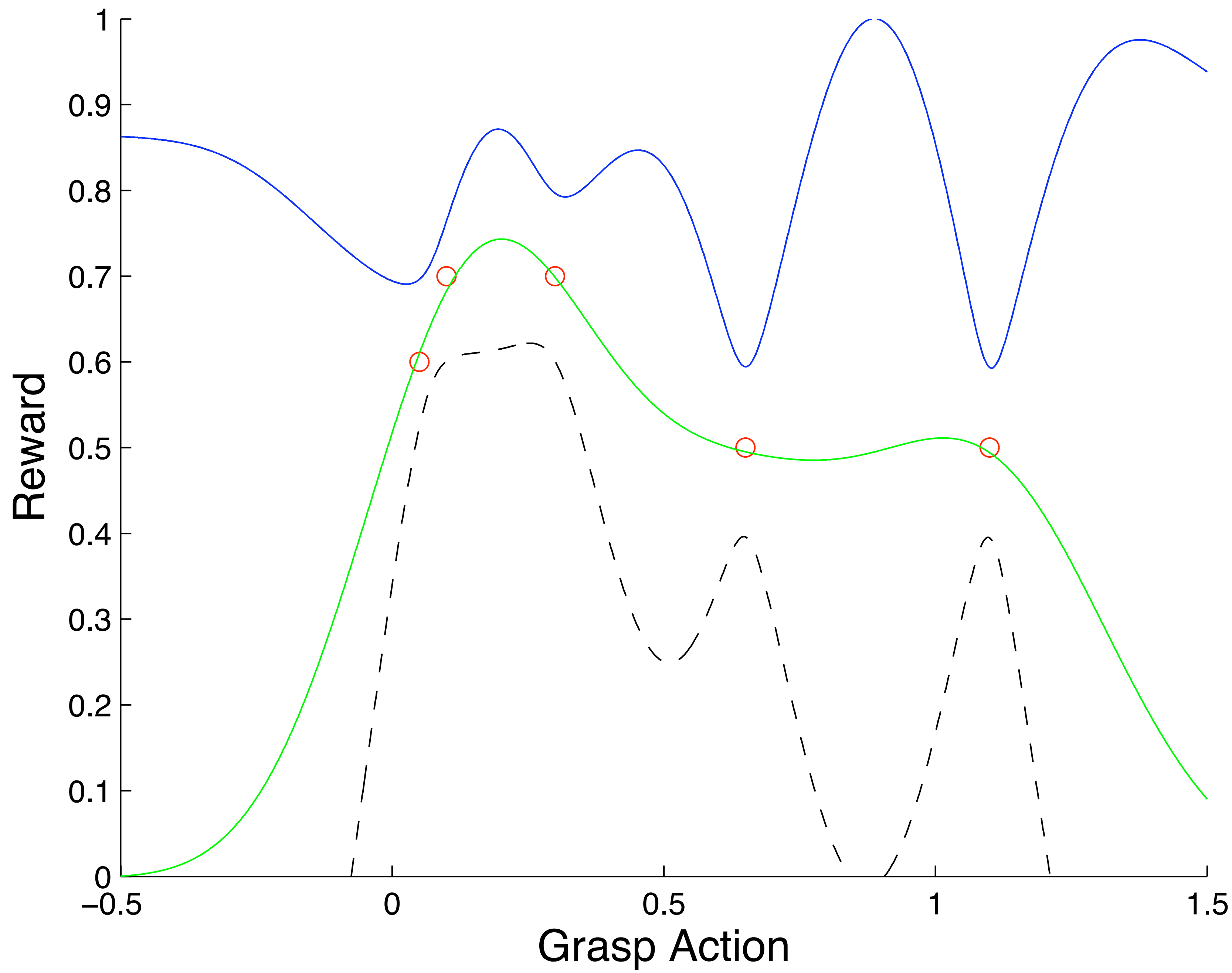
Initial Situation



Gaussian Process Model



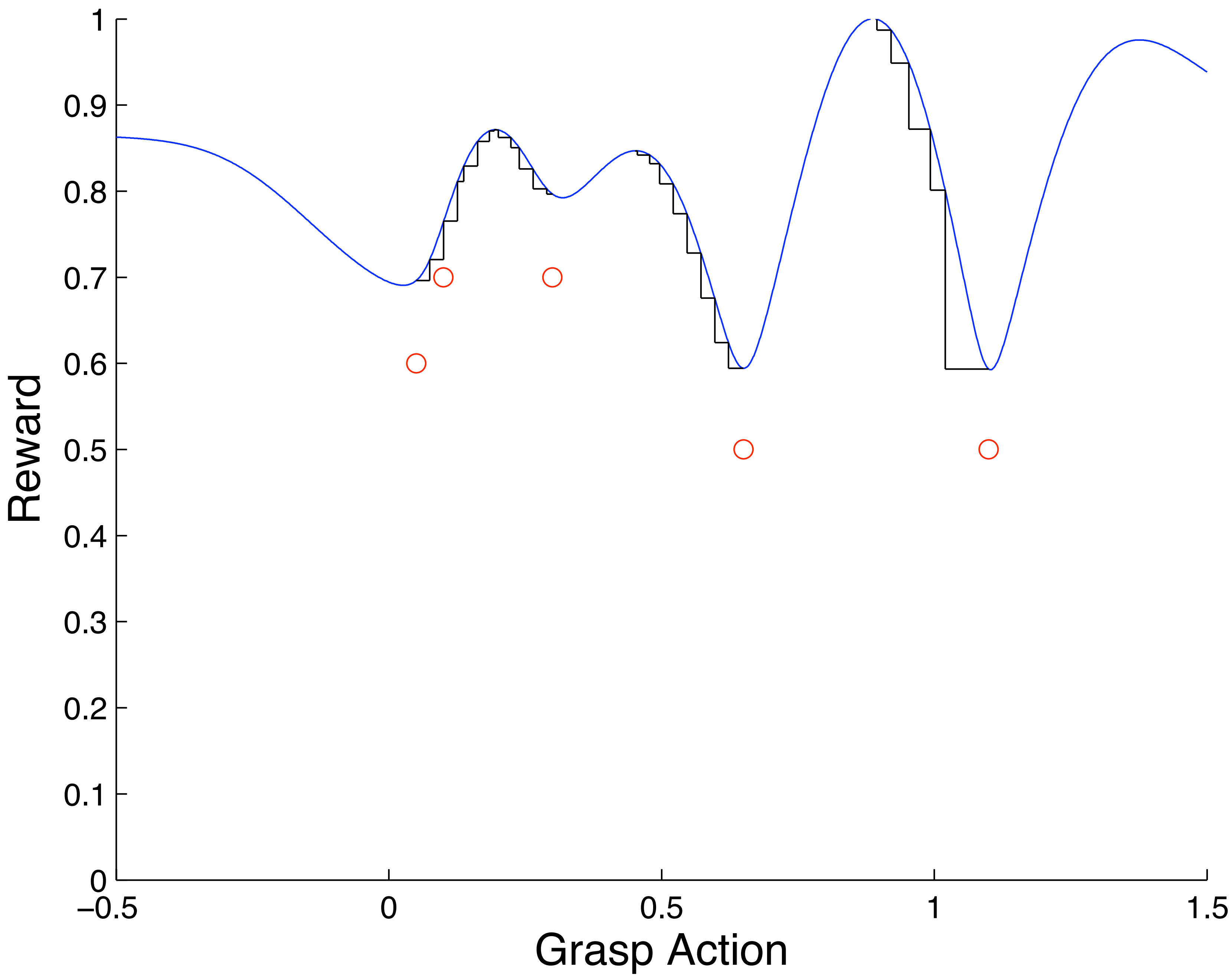
UCB Merit Function



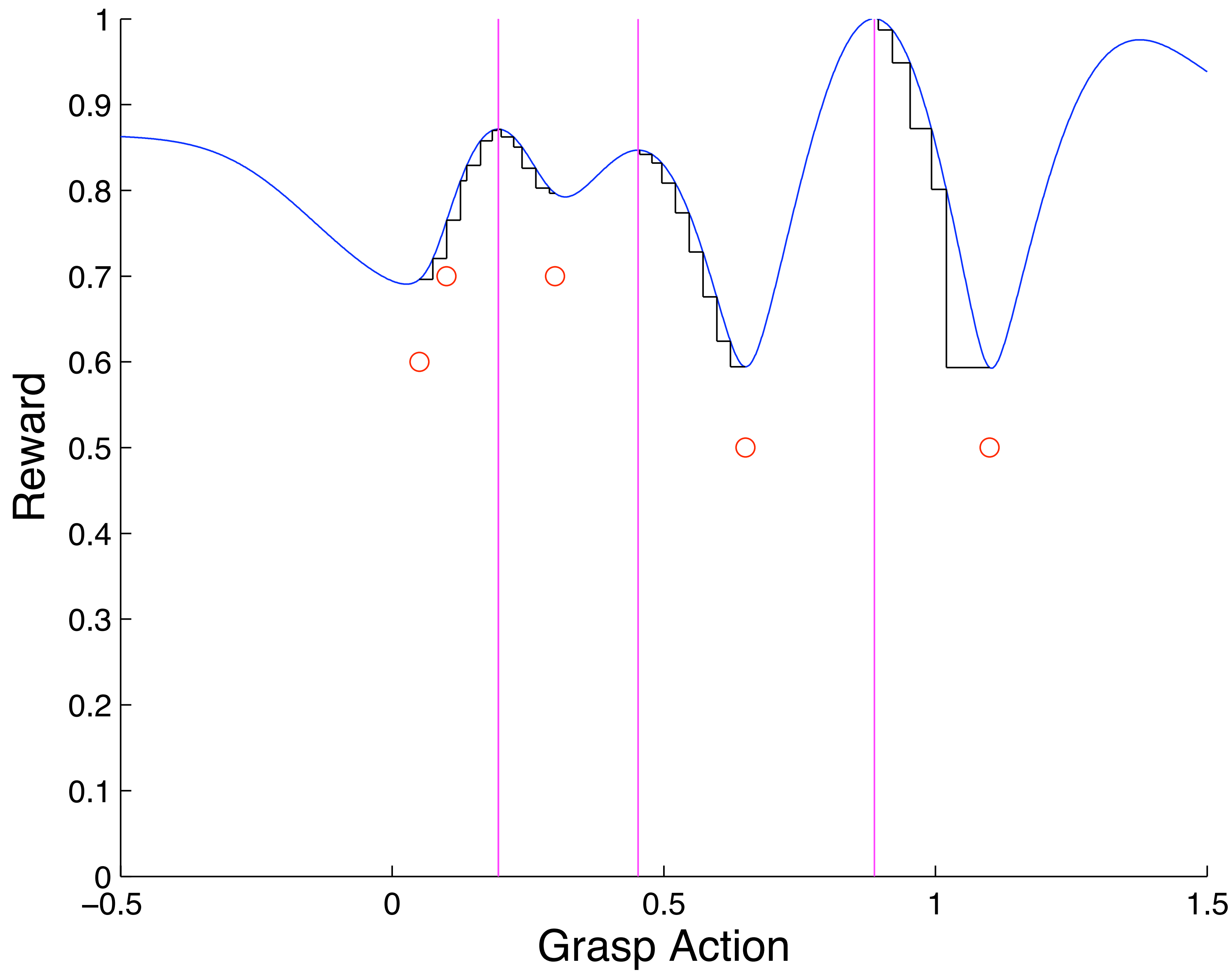
Choosing the Next Action

- Infinite number of possible actions
- Focus on a small set of actions
  - Want to find candidates for maximum
  - Focus on Local Maxima close to the data points!
- Mean Shift inspired method to find local maxima
  - Iterative procedure to find local maximum of a point
  - Initialize iterative procedure with all previous points
- Back to our example...

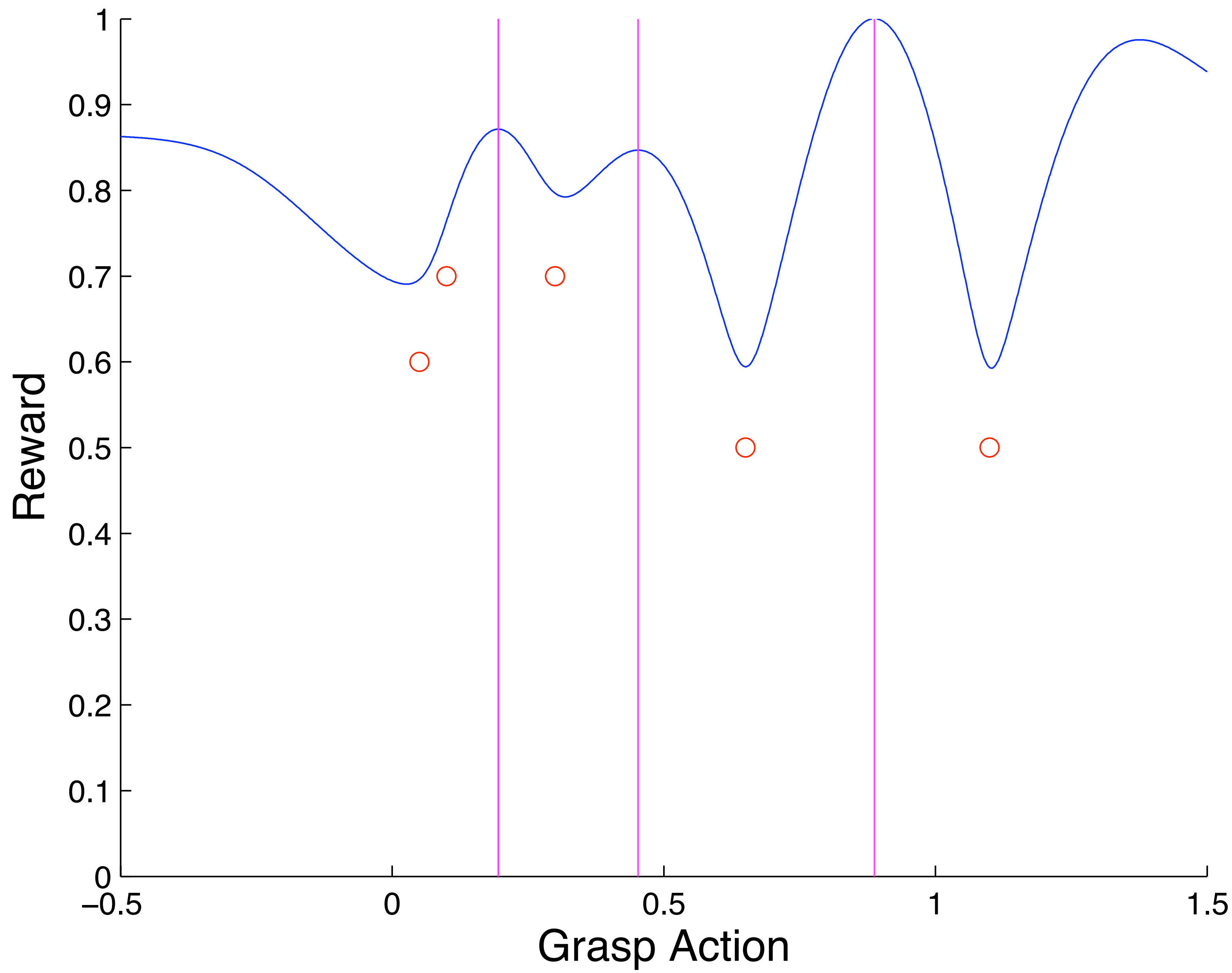
Detecting Maxima



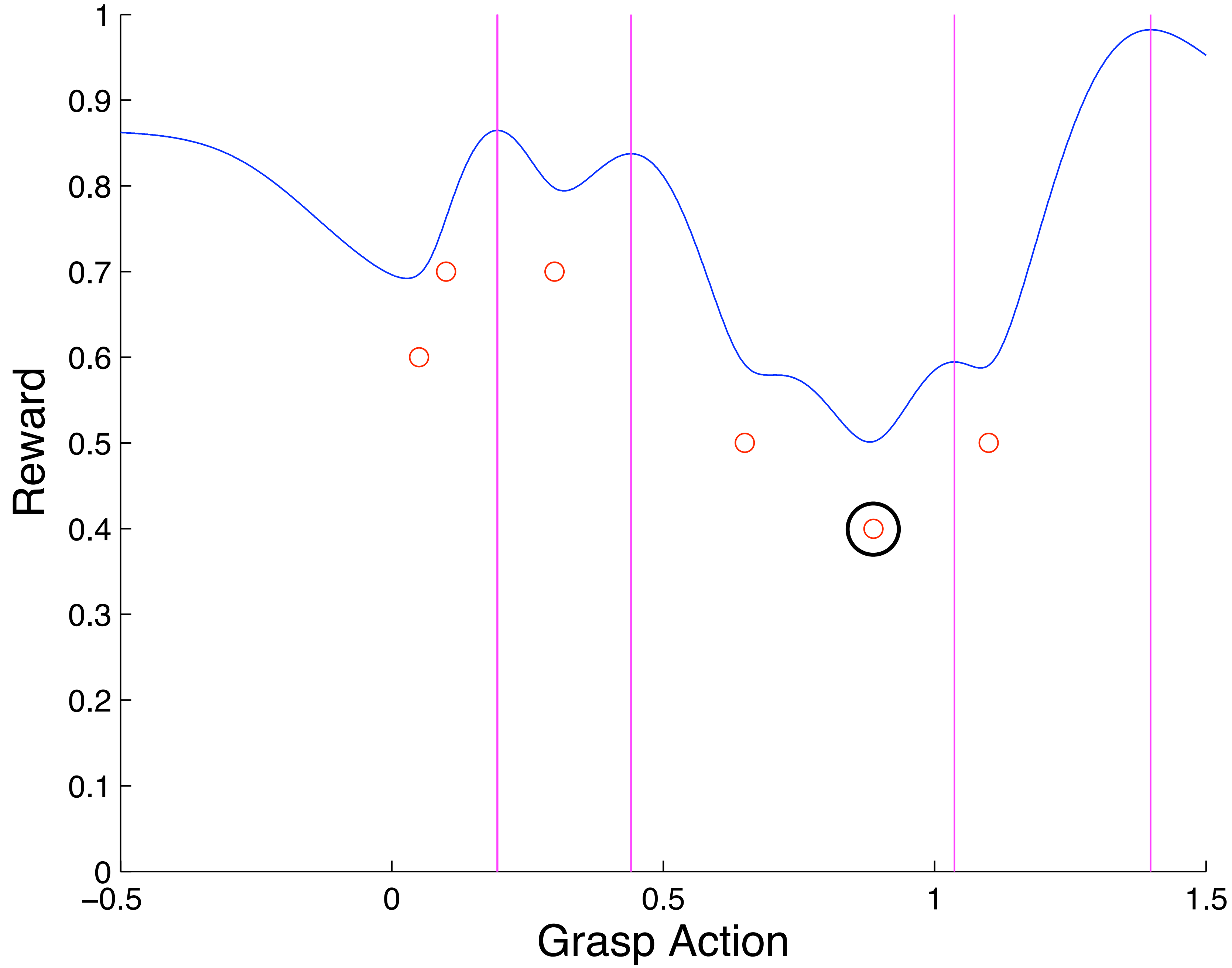
Detecting Maxima



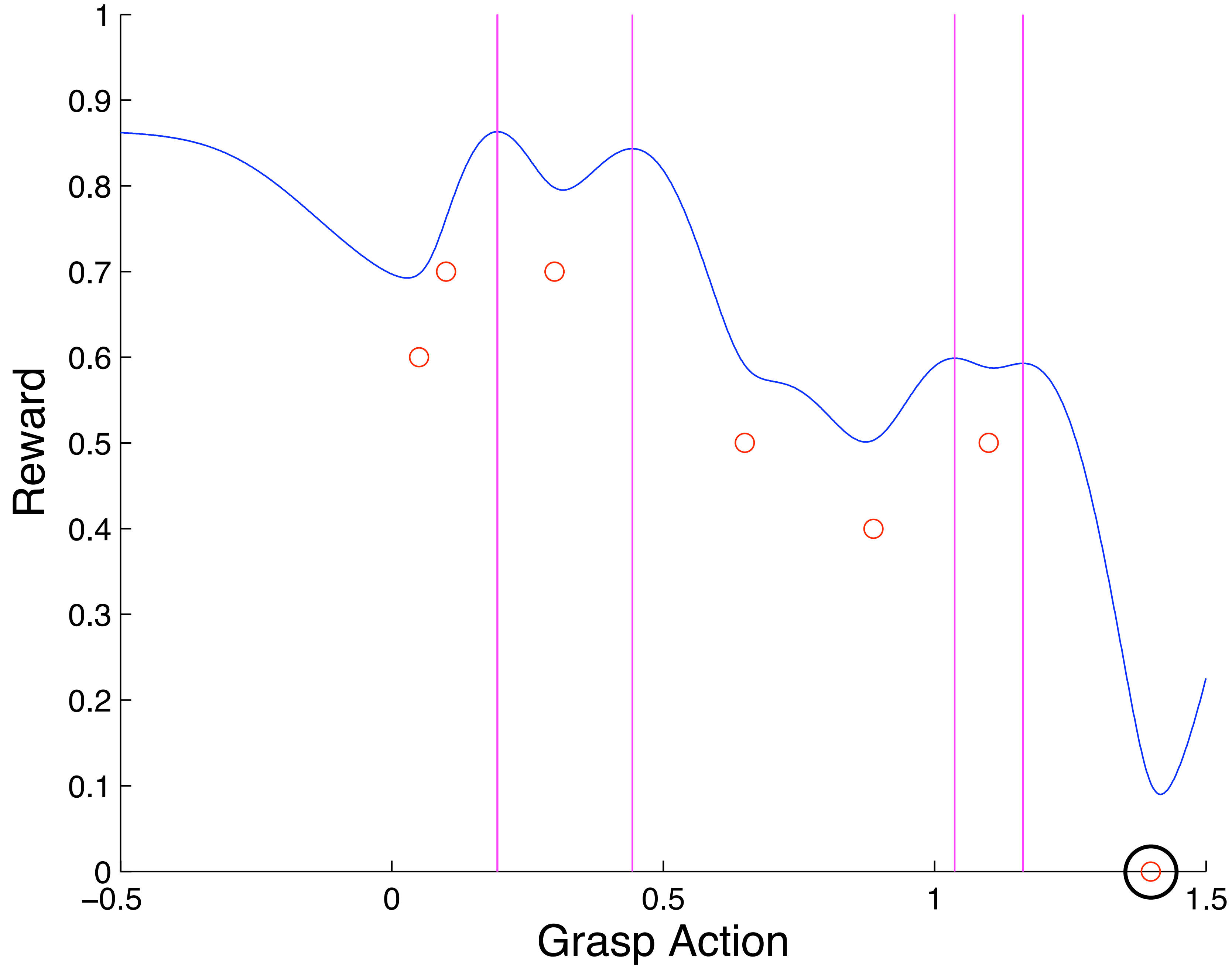
Evaluate Candidate



Another Trial ...



Another Trial ...



Outline

I. Introduction

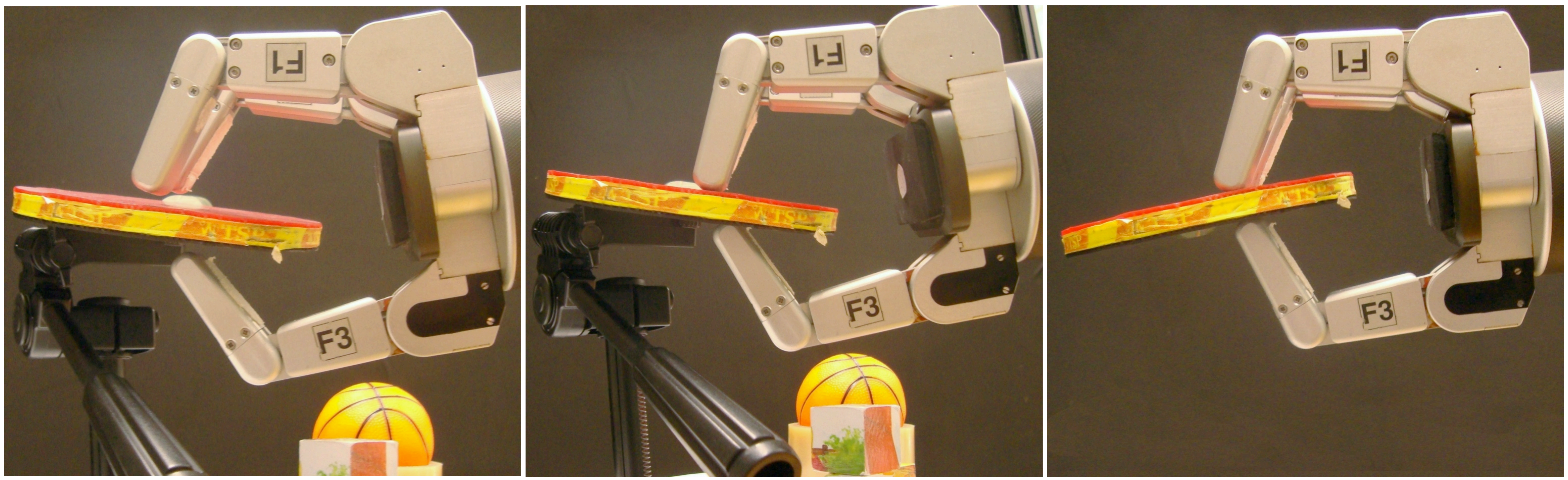
II. Generalizing Grasping Movements

III. Generalizing to New Objects

IV. Robot Experiments

V. Conclusion

Experimental Procedure



- Determine pose of object
- Approach object with preshaping of hand
- Grasp and lift object
- Reward given for little finger movement during lifting

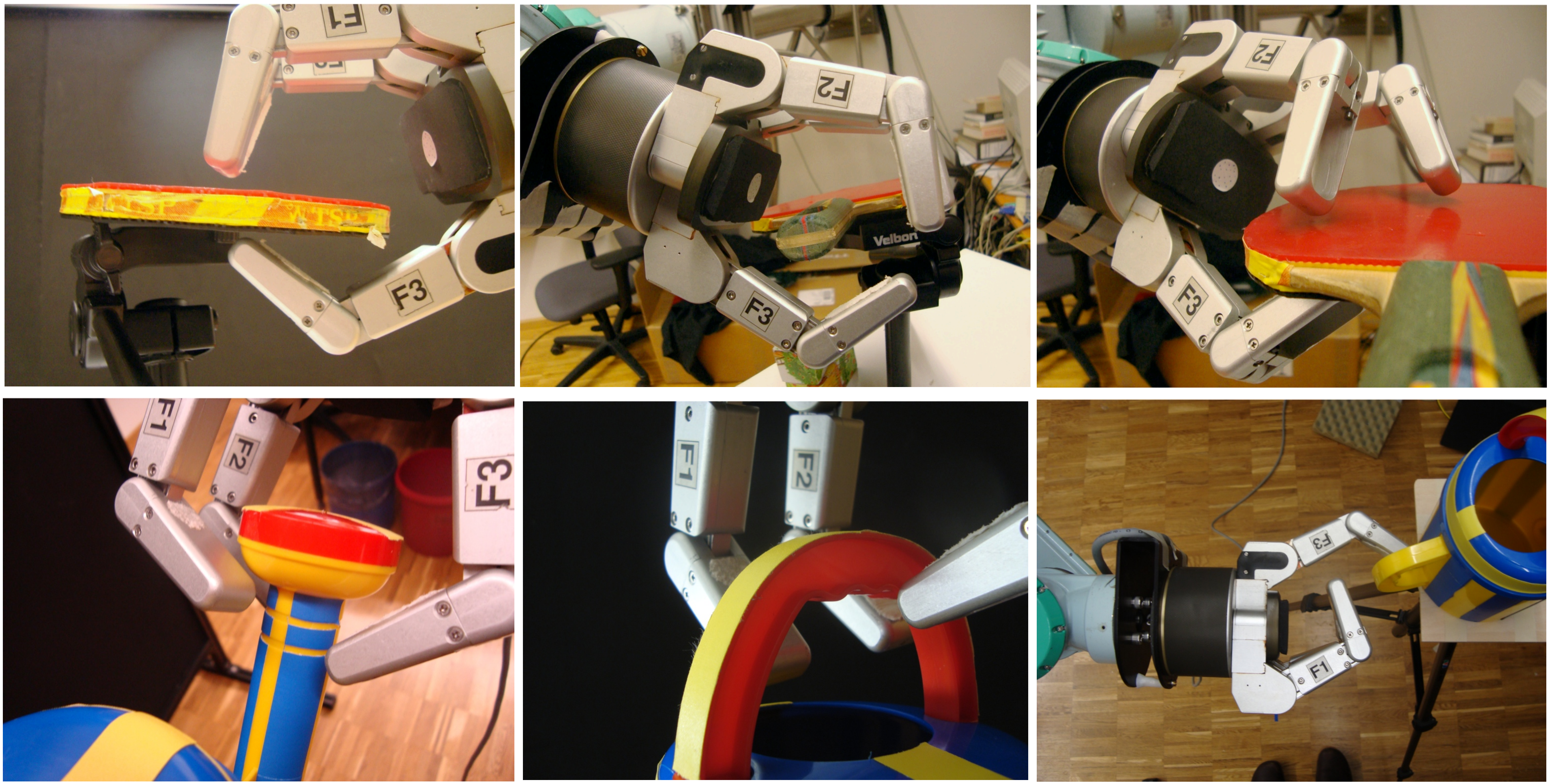
Initialization and Generalization

- VICON system used to acquire grasping movements
  - Applied markers only to fingers and hand
  - Acquired only a single grasping movement
  - Grasping movement demonstrated on a ball
- Robot learned to grasp:

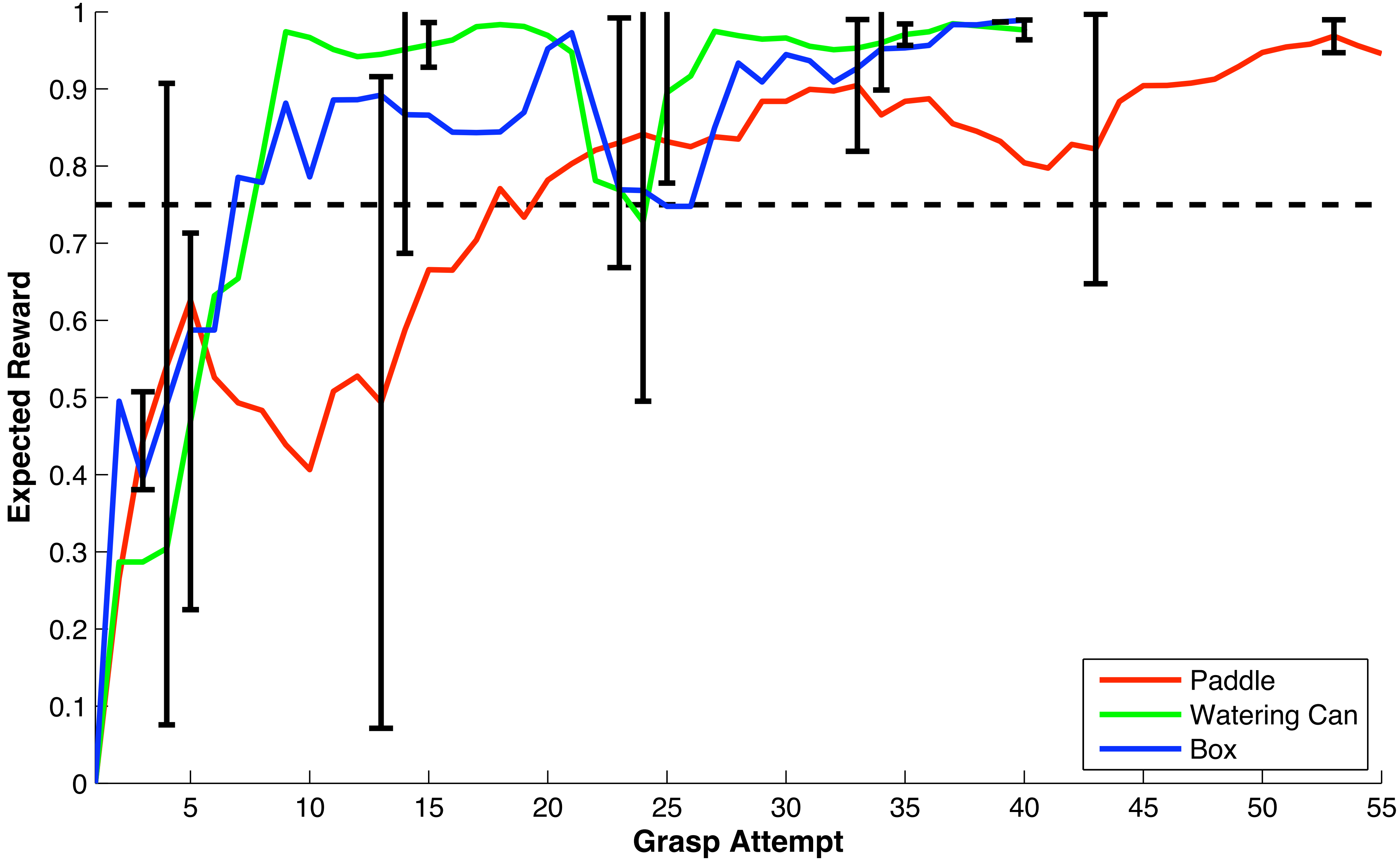
Range of Motion


- Only needed a single demonstration
- Adapt to different approach directions
- Right-handed robot

Preshaping of the Hand



Adapting to Objects



Outline

I. Introduction

II. Generalizing Grasping Movements

III. Generalizing to New Objects

IV. Robot Experiments

V. Conclusion

Summary

- Imitation learning is intuitive for programming robots
- Need to generalize demonstrated actions
- Used task-specific motor primitives as adaptive action representations
- Optimize grasps for new objects using the continuum-armed bandits framework
- Future work
    - Apply concepts to wider range of actions