Introduction

Background: Pilco

Roadmap

Conclusion

Control of an inverted double pendulum

using Machine Learning and Camera Feedback

Nicholas Capel

November 21, 2016

Contents

1 Introduction

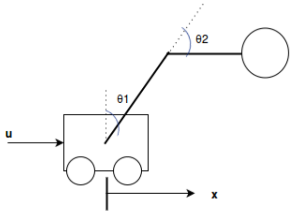
2 Background: Pilco

3 Roadmap

4 Conclusion

The inverted double pendulum problem

Figure: Diagram of inverted double pendulum



$$\mathbf{x} = [\dot{x}_t, x_t, \dot{\theta}_{1(t)}, \theta_{1(t)}, \dot{\theta}_{2(t)}, \theta_{2(t)}]^t$$

Discover controller

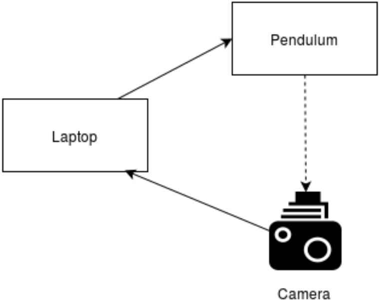
$\mathbf{x} \longmapsto \pi(x) = u$

$min(\sum_{t=0}^{T} \mathbb{E}[c(x_t)])$

- Task 1: Swing up $(\theta_1 = \pi,\ \theta_2 = 0)$
- Task 2: Stabilize $(\theta_1 = 0,\ \theta_2 = 0)$

Experimental setup

Figure: Diagram of experimental setup



- $\mathbf{x}$ vector generated using camera feedback

Motivation

Comparison of Approaches

Classical Approach

- 1st order approximation breaks down
- Model-based approach

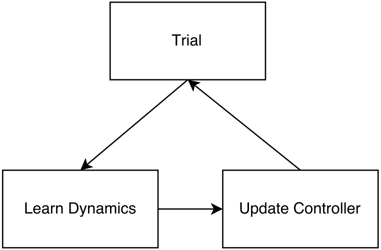
PILCO Approach

- Can learn model for entire state space
- Data-based approach

- Video-based feedback
  - State uncertainty
  - Delay
- This has been done before
  - But no realistic simulation (noise + uncertainty)

Overview of Pilco

Figure: PILCO algorithm



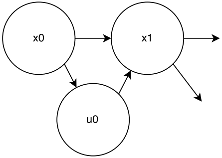
Overview of Pilco

Learning Dynamics Model

- Fundamental assumption that the next state $x_{t+1} = f(x_t, u_t)$
- Models this transition using a Gaussian process $p(x_t | x_{t-1}, u_{t-1}) = N(x_t, \Sigma_t)$

Optimising policy

Figure: Rollout



- Compute Cost Function $J = \sum_{t=0}^{T} \mathbb{E}[c(x_t)]$
- Gradient Descent $(\frac{\partial J}{\partial \theta})$ over policy parameters to find policy that minimizes J

Computer simulation experiments

- Computer simulations explore how the PILCO algorithm performs with noise and time delay
- Pendulum initialized in upright position

PILCO toolbox

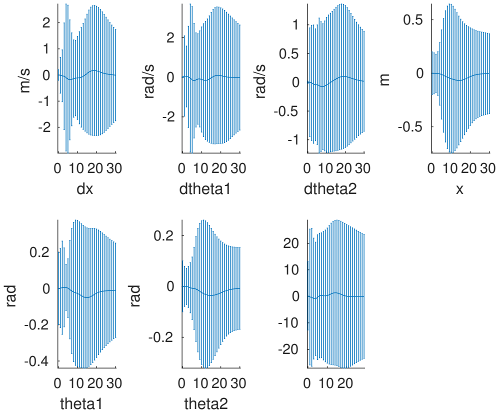
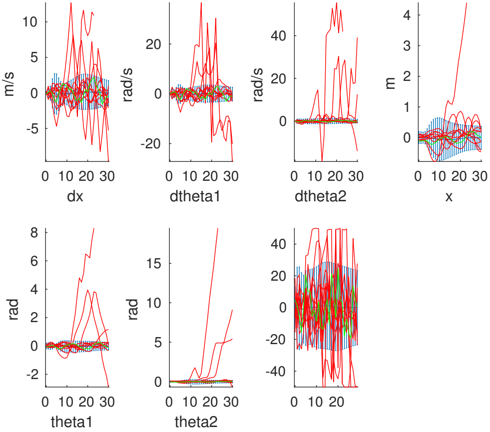


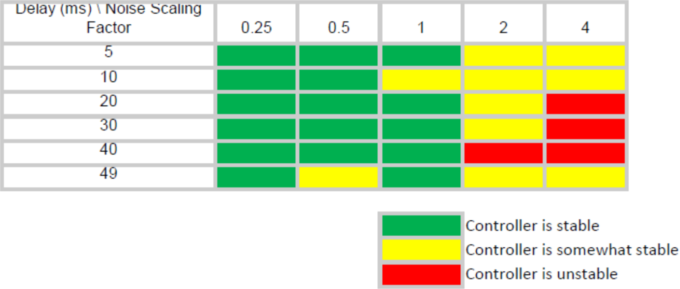
Figure: Computer Simulation Results for single experiment

Figure: Computer Simulation Results for single experiment with rollouts



Preliminary Results

Figure: Stability of Controller for various noise and delay levels



Noise scaling factor

Observation noise at each coordinate modelled as a Gaussian with variance scaled by this factor

Analytical Handle on control problem

- Linearize systems of equations of the inverted double pendulum about equilibrium
- Introduce controller with delay
- Introduce noise to the readings

Physical Experiments

- Explore maximum time delay for transmission and processing of a single camera frame
  - Expected delay of approximately 30ms
  - Test time delay via a perturbance to the double pendulum system
- Run PILCO algorithm with real system for the stabilization task
- Run PILCO algorithm with real system for the swingup and stabilization task

Conclusion and Outlook

- It seems that the algorithm is more sensitive to input noise than to delays
- Run repeated computer simulations to confirm results
- Perform analysis of inverted double pendulum problem
- Perform physical experiments on camera and cart system