Learning Control at Intermediate Reynolds Numbers

Experiments with perching UAVs and flapping-winged flight

Russ Tedrake

Assistant Professor
MIT Computer Science and Artificial Intelligence Lab

IROS 2008 Learning Workshop
September 22, 2008

Massachusetts Institute of Technology

CSAIL

The Burden of Learning Control

- Reinforcement learning has the potential to generate high-performance <u>nonlinear</u>, adaptive control policies for complicated systems...

The Burden of Learning Control

- Reinforcement learning has the potential to generate high-performance <u>nonlinear</u>, adaptive control policies for complicated systems...

  but our techniques *must* be validated vs. more traditional control techniques.

The Burden of Learning Control

- Reinforcement learning has the potential to generate high-performance nonlinear, adaptive control policies for complicated systems...

  but our techniques *must* be validated vs. more traditional control techniques.

- There are many problems in robotics where traditional nonlinear control offers relatively few solutions

The Burden of Learning Control

- Reinforcement learning has the potential to generate high-performance nonlinear, adaptive control policies for complicated systems...

  but our techniques *must* be validated vs. more traditional control techniques.

- There are many problems in robotics where traditional nonlinear control offers relatively few solutions
  - Underactuated systems, many stochastic systems, ...

The Burden of Learning Control

- Reinforcement learning has the potential to generate high-performance nonlinear, adaptive control policies for complicated systems...

  but our techniques *must* be validated vs. more traditional control techniques.

- There are many problems in robotics where traditional nonlinear control offers relatively few solutions
  - Underactuated systems, many stochastic systems, ...
  - Learning control can quickly become the state-of-the-art

The Burden of Learning Control

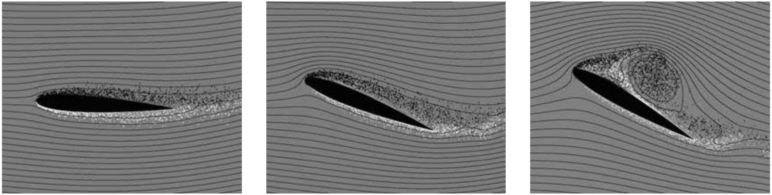
- Reinforcement learning has the potential to generate high-performance <u>nonlinear</u>, adaptive control policies for complicated systems...

  but our techniques *must* be validated vs. more traditional control techniques.

- There are many problems in robotics where traditional nonlinear control offers relatively few solutions
  - Underactuated systems, many stochastic systems, ...
  - Learning control can quickly become the state-of-the-art

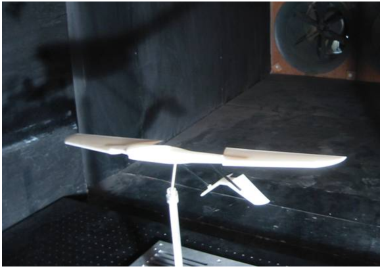
Control in complicated fluid dynamics

Example: Landing on a perch



The State-of-the-Art in Perching

Approach 1: Morphing plane



Approach 2: Over-powered hover

- Both sacrifice performance to use linear control on modified vehicles (can't compete w/ birds!)
- Can learning nonlinear control produce superior performance on existing vehicles?

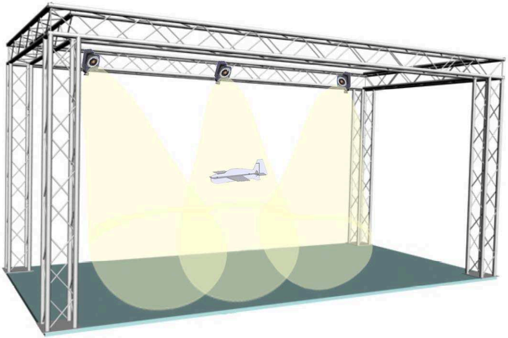
Technical challenges for control

- Dynamics quickly get too complicated for conventional control design
  - Fluid dynamics are time-varying and very nonlinear
  - CFD simulations in these regimes can take days to compute
  - Severe lack of compact (design accessible) models
- Limited control authority
  - Flow is only partially observable
  - Stalls result in intermittent losses of control authority
  - "Underactuated control" - control actions have long-term consequences

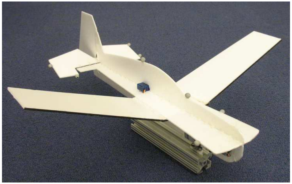
Learning Control for Perching

- Accurate Navier-Stokes simulations takes days to compute, but...
- Model-based Reinforcement Learning:
  1. Learn *approximate* model of unsteady dynamics from real flight data
  2. Formulate the goal of control as the long-term optimization of a scalar cost
  3. Offline model-based numerical optimal control
  4. Online model-free optimal control ("learning")

Experiment Design

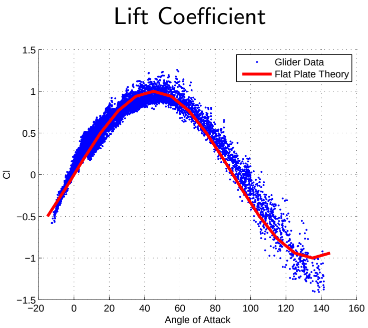


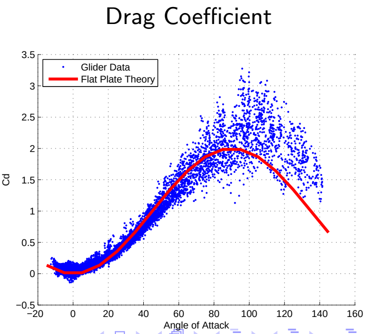
- Glider (no propellor)
- Dihedral (passive roll stability)
- Offboard sensing and control



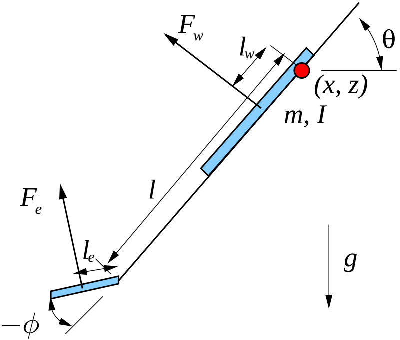
System Identification

- Nonlinear rigid-body vehicle model
- Linear (w/ delay) actuator model
- Real flight data
  - Very high angle-of-attack regimes
  - Surprisingly good match to theory
  - Vortex shedding





A dynamic model



- Planar dynamics
- Aerodynamics from model
- State: $\mathbf{x} = [x, y, \theta, \phi, \dot{x}, \dot{y}, \dot{\theta}]$
- Only actuator is the elevator angle, $\mathbf{u} = \dot{\phi}$

Feedback Control Design

- Optimal feedback control formulation:

$$J^\pi(\mathbf{x}) = \min\left[(\mathbf{x}-\mathbf{x}^d)^T\mathbf{Q}(\mathbf{x}-\mathbf{x}^d), \quad J^\pi(\mathbf{x}')\right],$$
$$\mathbf{x}' = f(\mathbf{x}, \pi(\mathbf{x})))$$

  - $\mathbf{x}$ is the estimated state (from motion capture)
  - π is the feedback policy, commanding elevator angle as a function of state
  - f is the identified system model
  - $\mathbf{Q}$ is a positive definite cost matrix
- Discretize dynamics on a mesh over state space
- Optimized Dynamic Programming algorithm approximates the optimal policy, $\pi^*(\mathbf{x})$, from model

Glider Perching

- Enters motion capture @ 6 m/s.
- Perch is < 3.5 m away.
- Entire trajectory @ 1 second.

Requires Separation!

Flow visualization (very preliminary)

Flow visualization (very preliminary)

Transition to Outdoors: Gust Response

- Current experiments are in "still air"
- Outdoor flights require robustness to
  - ambient wind conditions
  - persistent flow structures (e.g., around the perch)
  - short-term "gust" disturbances

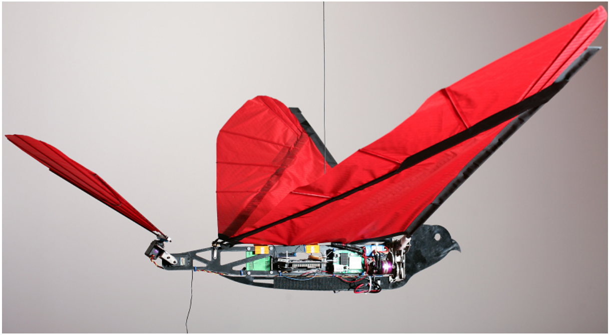
- Capture statistical environment model
- Instrument motion capture arena with known aerodynamic disturbances, obstacles, and wind

- Already performed detailed LTV controllability analysis with different actuator configurations

The Intermediate Reynolds Number regime

- Reynolds number (Re) - dimensionless quantity that correlates with the resulting kinematics of the fluid
  - Low Re - viscousity dominates, flow is laminar
  - High Re - turbulence
  - Intermediate Re - complicated, but structured flow (eg, vortex shedding). Glider perching example is Re 50,000 down to Re 15,000.
- At Intermediate Re:
  - Lots of interesting control problems
  - Almost no good control solutions

Bird-scale flapping flight



Autonomous Flapping-Winged Flight

Motivation

- Bird-scale flapping vehicles will not surpass the speed or efficiency of fixed-wing aircraft for steady-level flight in still air
  - Propellors produce thrust very efficiently
  - Aircraft airfoils can be highly optimized (for speed or efficiency)

Motivation

- Bird-scale flapping vehicles will not surpass the speed or efficiency of fixed-wing aircraft for steady-level flight in still air
  - Propellors produce thrust very efficiently
  - Aircraft airfoils can be highly optimized (for speed or efficiency)
- But looking more closely...

Efficient flying machines

- An albatross can fly for hours (or even days) without flapping, even migrating upwind (exploiting gradients in the shear layer)

- Butterflies migrate thousands of kilometers, carried by the wind

Super-maneuverability

- Peregrine falcons have been clocked at 240+ mph in dives, and have the agility to snatch moving prey

- Bats have been documented...
  - Catching prey on their wings
  - Manuevering through thick rain forests at high speeds
  - Making high speed 180 degree turns
  - ...

Manipulating the Air

- Birds far surpass the performance of our best engineered systems (especially UAVs) in metrics of efficiency, acceleration, and maneuverability.

Manipulating the Air

- Birds far surpass the performance of our best engineered systems (especially UAVs) in metrics of efficiency, acceleration, and maneuverability.
- The secret:

Manipulating the Air

- Birds far surpass the performance of our best engineered systems (especially UAVs) in metrics of efficiency, acceleration, and maneuverability.

- The secret:

  Birds (and fish, ...) exploit unsteady aerodynamics at intermediate Re

Manipulating the Air

- Birds far surpass the performance of our best engineered systems (especially UAVs) in metrics of efficiency, acceleration, and maneuverability.
- The secret:

  Birds (and fish, ...) exploit unsteady aerodynamics at intermediate Re

- A *manipulation* problem
  - Requires unconventional mechanical and control designs
  - Once you start thinking of bird flight as manipulating the air, it becomes harder to appreciate fixed-wing flight

Example: Efficient swimming upstream

from George Lauder's Lab at Harvard (Liao et al, 2003)

Prospects for machine learning

- Key observations about fluid-body interactions at intermediate Re
  - Considerable previous work in system identification permits the use of approximate models
  - Won't always be able to discretize the state space
  - Relatively compact policies (few parameters) can generate a large repetoire of behaviors

Prospects for machine learning

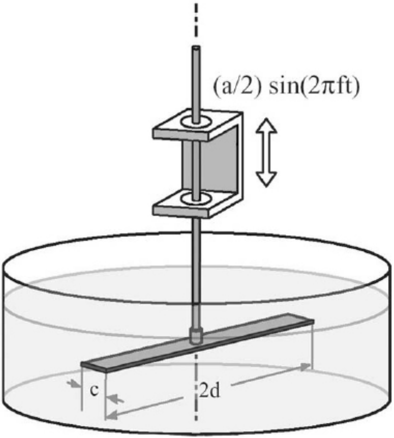
- Key observations about fluid-body interactions at intermediate Re
  - Considerable previous work in system identification permits the use of approximate models
  - Won't always be able to discretize the state space
  - Relatively compact policies (few parameters) can generate a large repetoire of behaviors

- Formal analysis of the policy gradient algorithms reveals:
  - Performance (via SNR) degrades with the number of control parameters
  - Performance is (locally) invariant to the complexity of the plant dynamics

The Heaving Foil

work with Jun Zhang (NYU Courant)

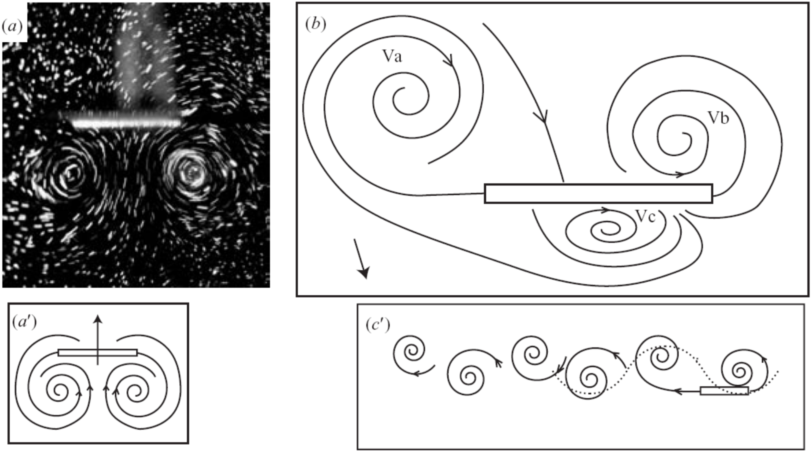
The Heaving Foil



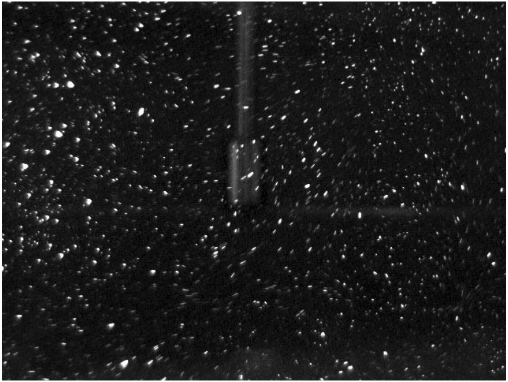
[Vandenberghe et al., 2006]

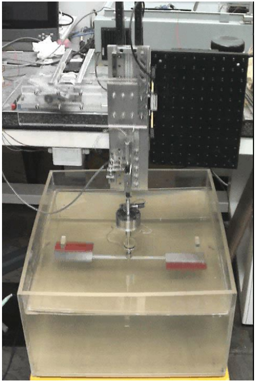
- Rigid, symmetric wing
- Driven vertically
- Free to rotate horizontally

Symmetry breaking leads to forward flight

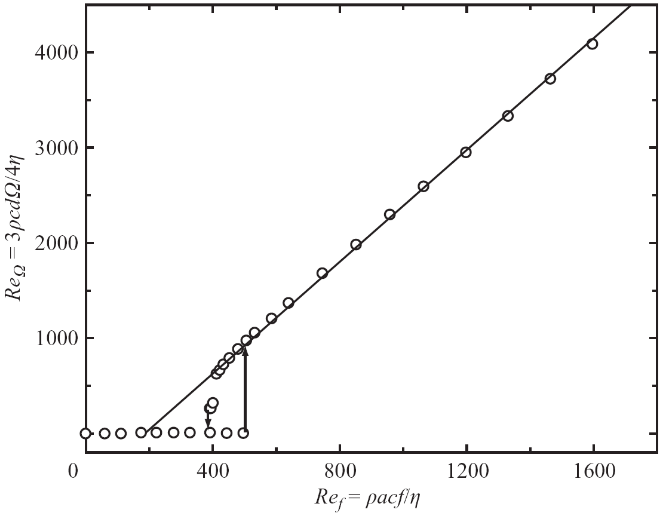


Flow visualization





Effect of flapping frequency



The control problem

- Previous work only used sinusoidal trajectories
- Optimize stroke form to maximize the "efficiency" of forward flight
  - Add vertical load cell (measures $F_z(t)$)
  - Dimensionless cost of transport:

$$c_{mt} = \frac{\int_T |F_z(t)\dot{z}(t)|dt}{mg \int_T \dot{x}(t)dt}$$

The control problem

- Previous work only used sinusoidal trajectories
- Optimize stroke form to maximize the "efficiency" of forward flight
  - Add vertical load cell (measures $F_z(t)$)
  - Dimensionless cost of transport:

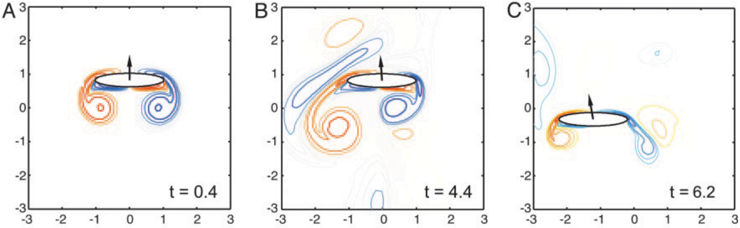
$$c_{mt} = \frac{\int_T |F_z(t)\dot{z}(t)| dt}{mg \int_T \dot{x}(t) dt}$$

  - Fortunately

$$\min c_{mt} = \min \frac{\int_T |F_z(t)\dot{z}(t)| dt}{\int_T \dot{x}(t) dt}$$

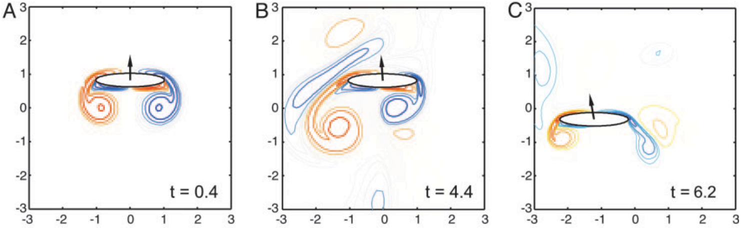
Prospects for optimization

- CFD model[Alben and Shelley, 2005]



Prospects for optimization

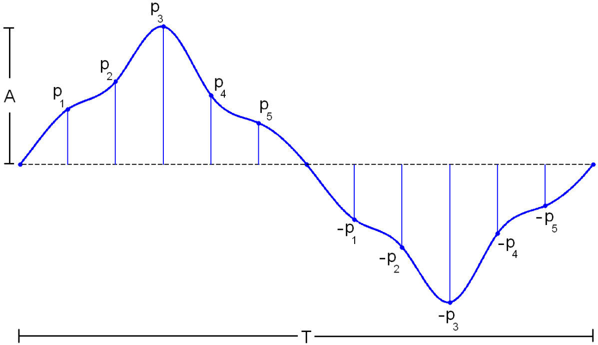
- CFD model[Alben and Shelley, 2005]



- Takes approximately 36 hours to simulate 30 flaps

Experimental Optimization

- Can we perform the optimization directly in the fluid?
- Direct policy search
  - Needs to be robust to noisy evaluations
  - Minimize number of trials required



Optimized policy gradient

- The basic algorithm (weight perturbation):
  - Perturb the control parameters, $\mathbf{p}$ by some amount $\mathbf{z}$ from $N(0, \sigma)$
  - Perform the update:

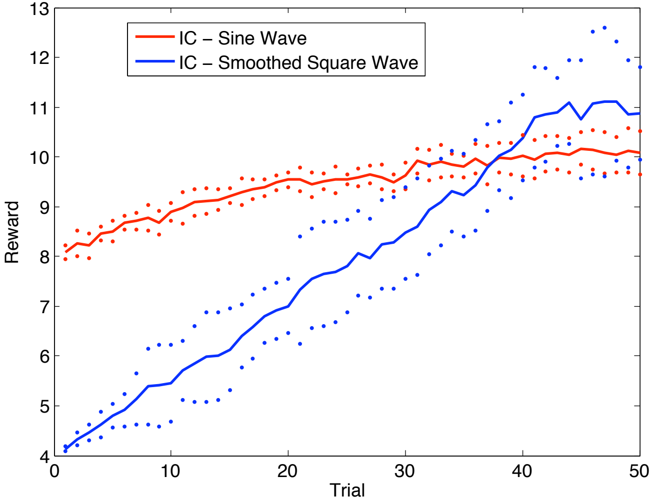
$$\Delta \mathbf{p} = -\eta(c_{mt}(\mathbf{p} + \mathbf{z}) - c_{mt}(\mathbf{p}))\mathbf{z}$$

  - Strong performance guarantees

$$E[\Delta \mathbf{p}] \propto -\frac{\partial c_{mt}}{\partial \mathbf{p}}.$$

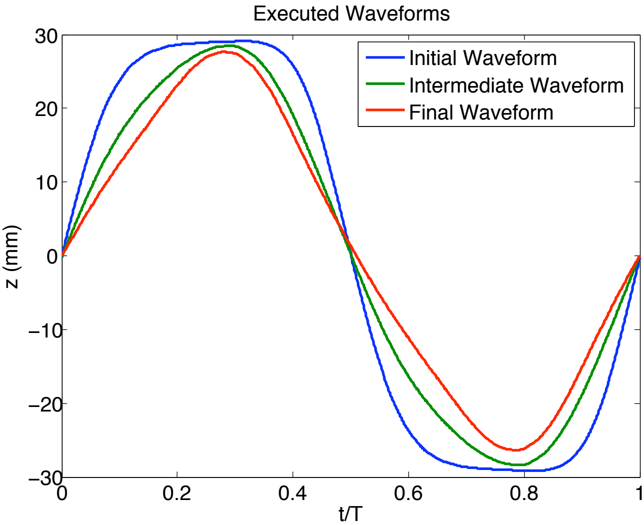
  - Poor performance (requires many trials)
- SNR optimized policy gradient [Roberts and Tedrake, 2009]

Learning results



learns in about 10,000 flaps (@ 15 minutes)

Learning results (cont.)



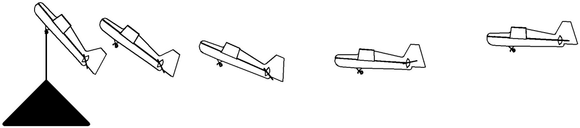
A dynamic explanation

- Forward speed is linear in flapping frequency
  - from experiments
  - statement about average speed
- Drag forces quadratic in speed ($F \propto \rho S V^2$)
- Triangle wave obtains highest average speed w/ minimal drag

Implications

- Enabling tool for experimental fluid dynamicists
- Suggests that motor learning algorithms could produce efficient control solutions in fluids
- Suggests that we can use this to control robotic birds
- Exciting prospects for online learning in changing environments

Summary



- Nonlinear, underactuated control w/ imperfect models via machine learning control (birds don't solve Navier-Stokes)
- Allows our machines to exploit unsteady flow effects
- Soon, robotic birds will:
  - Fly efficiently and autonomously
  - Outperform fixed-wing vehicles in maneuverability

Acknowledgements

- Students:
  - John Roberts
  - Rick Cory
  - Zack Jackowski
  - Steve Proulx
- Funding:
  - MIT, MIT CSAIL
  - NEC corporation
  - Lincoln Labs ACC
  - Microsoft

References

Alben, S. and Shelley, M. (2005).
Coherent locomotion as an attracting state for a free flapping body.
Proceedings of the National Academy of Sciences, 102(32):11163–11166.

Roberts, J. W. and Tedrake, R. (2009).
Signal-to-noise ratio analysis of policy gradient algorithms.
In *To appear in Advances of Neural Information Processing Systems (NIPS) 21*, page 8.

Vandenberghe, N., Childress, S., and Zhang, J. (2006).
On unidirectional flight of a free flapping wing.
Physics of Fluids, 18.