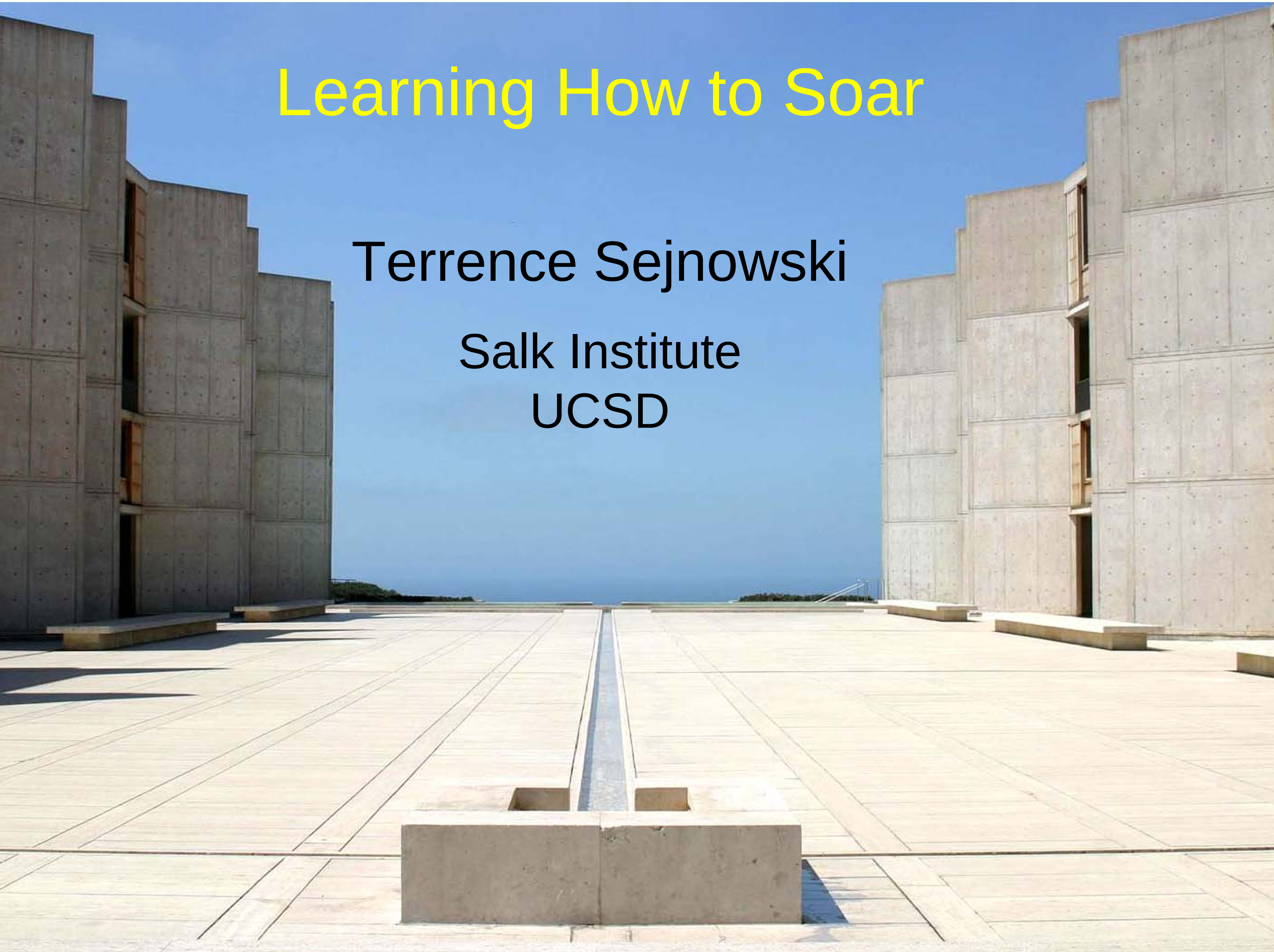
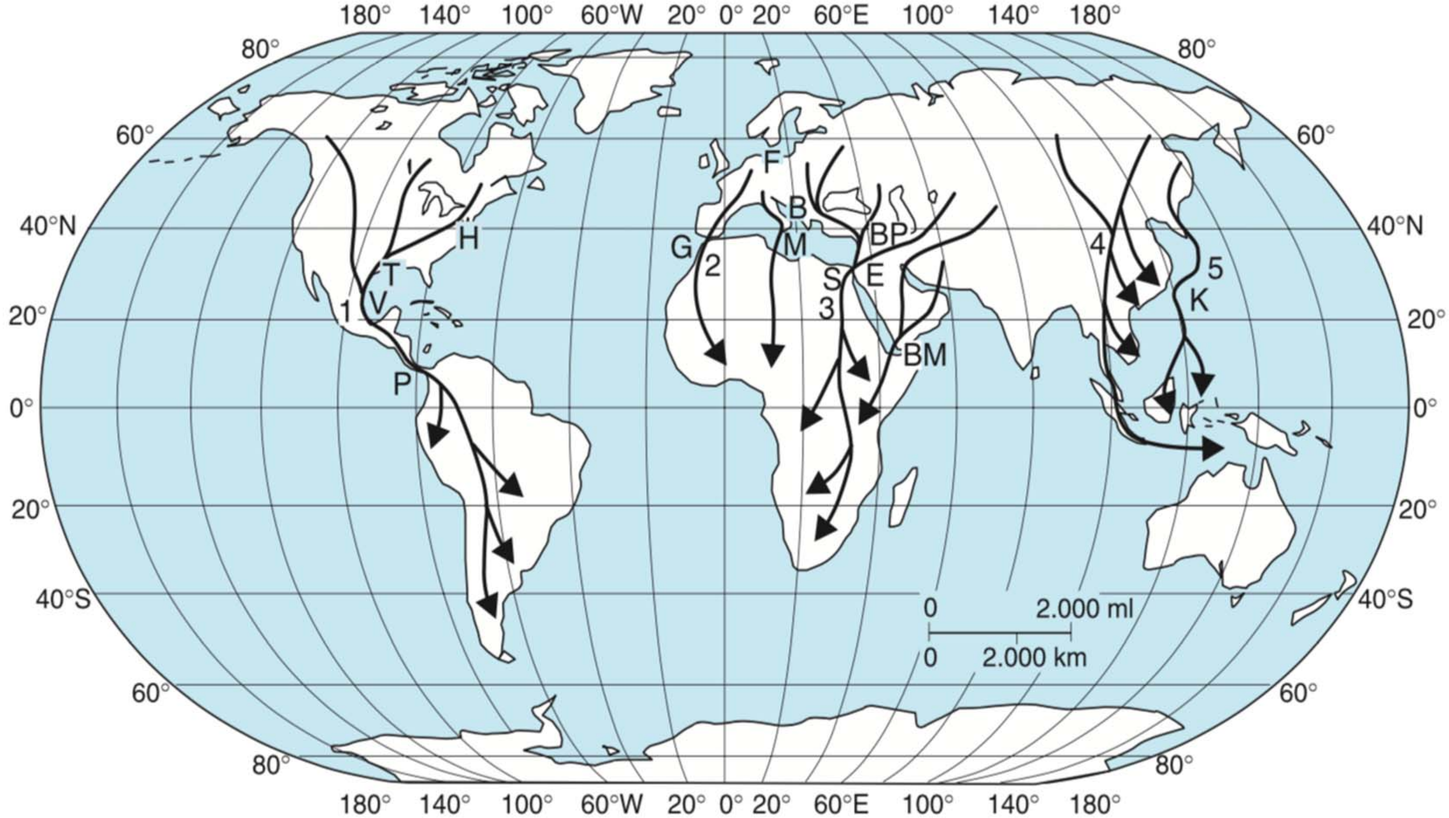
Learning How to Soar

Terrence Sejnowski

Salk Institute
UCSD

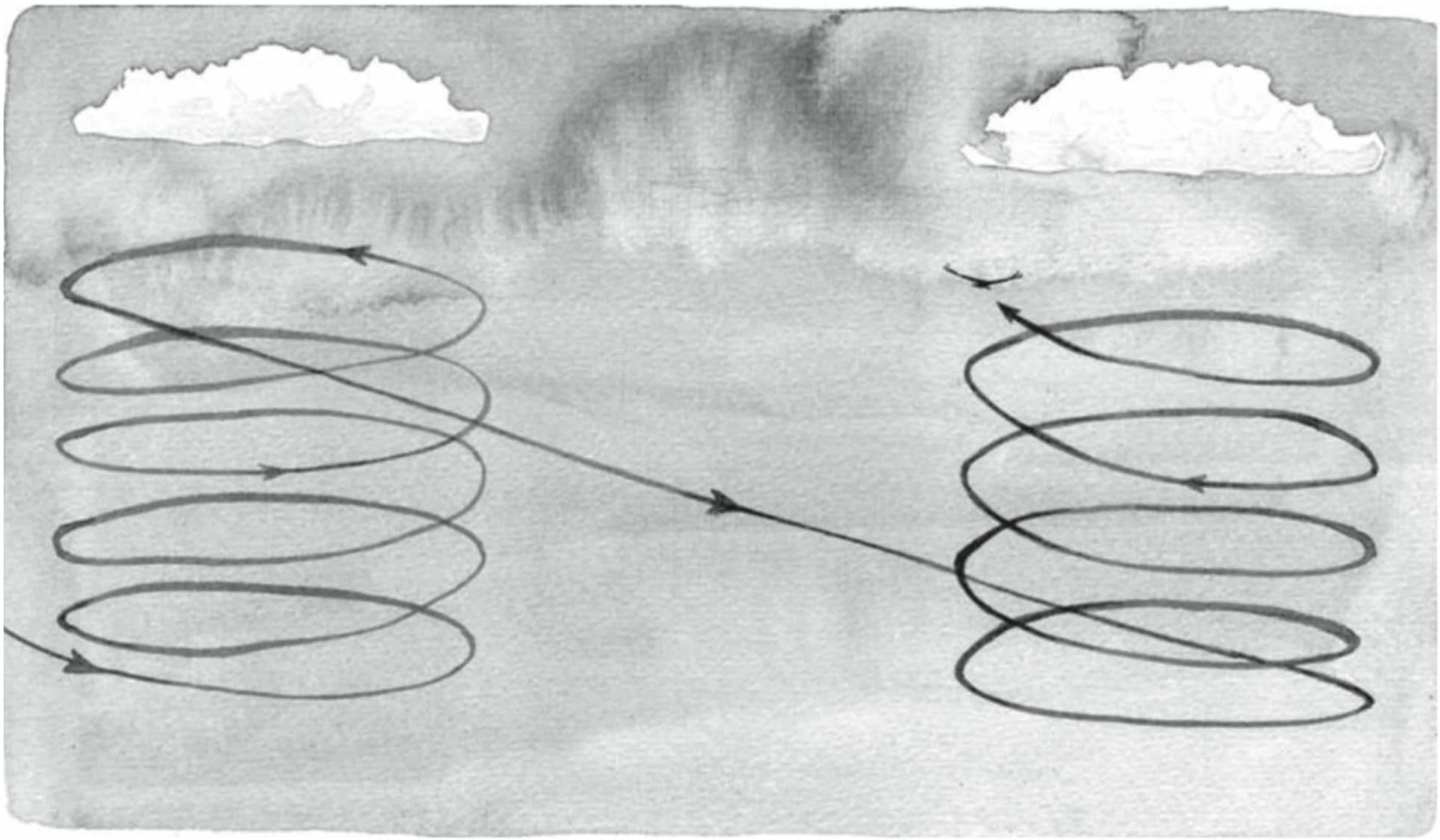


Bird Migration

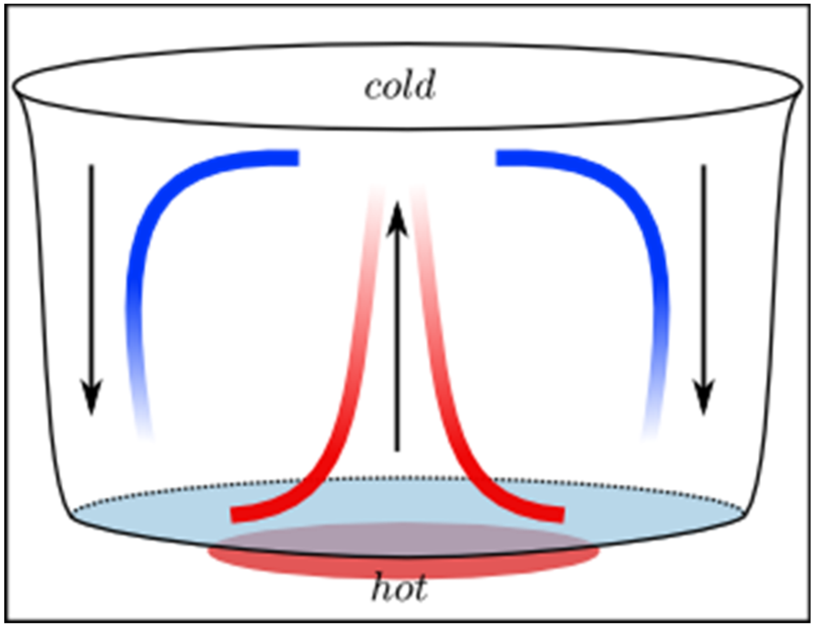


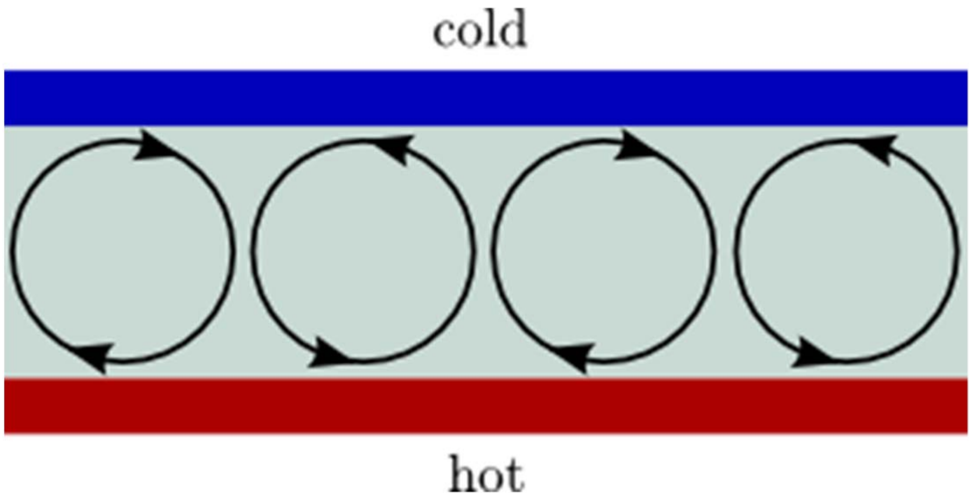
Migration Ecology of Birds, Ian Newton

Thermal Soaring



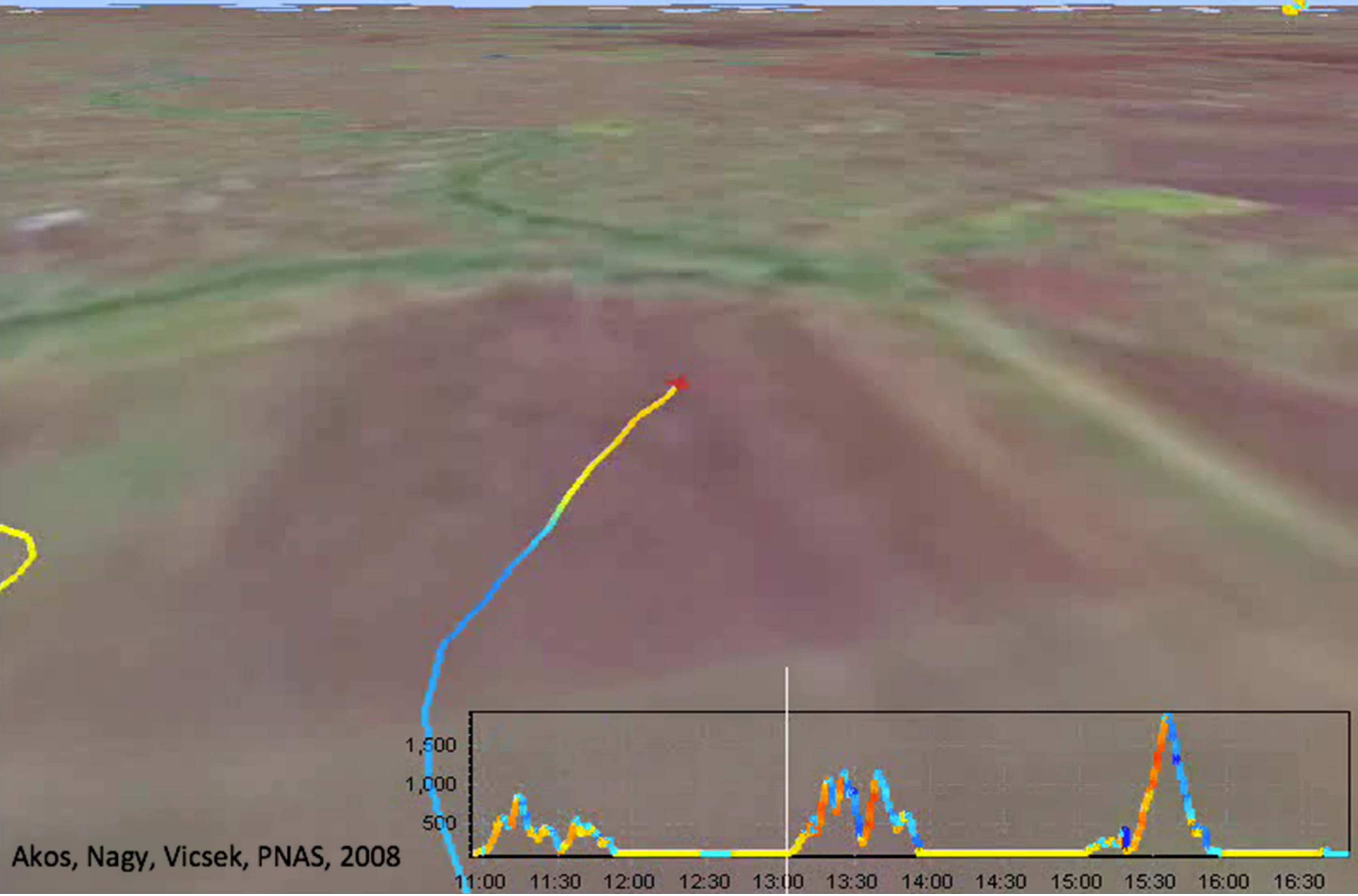
Rayleigh-Bénard Convection





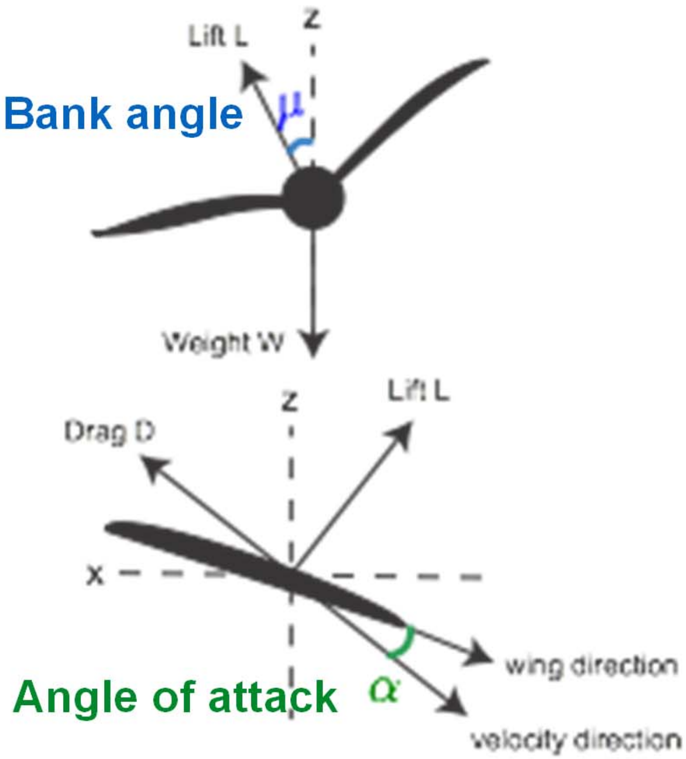
Atmospheric Turbulence

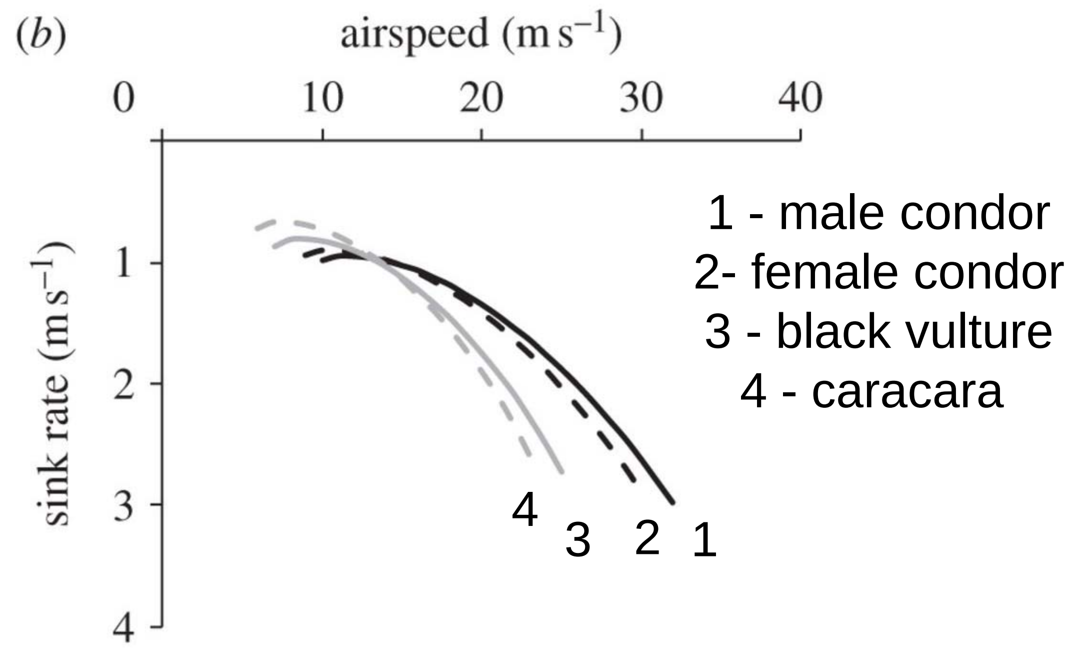
Tracking a Falcon with GPS



Humans Soar Too

Glider Aerodynamics





Control over bank angle and angle of attack

Shephard & Lambertucci, 2013

How do Birds Find and Navigate Thermals?

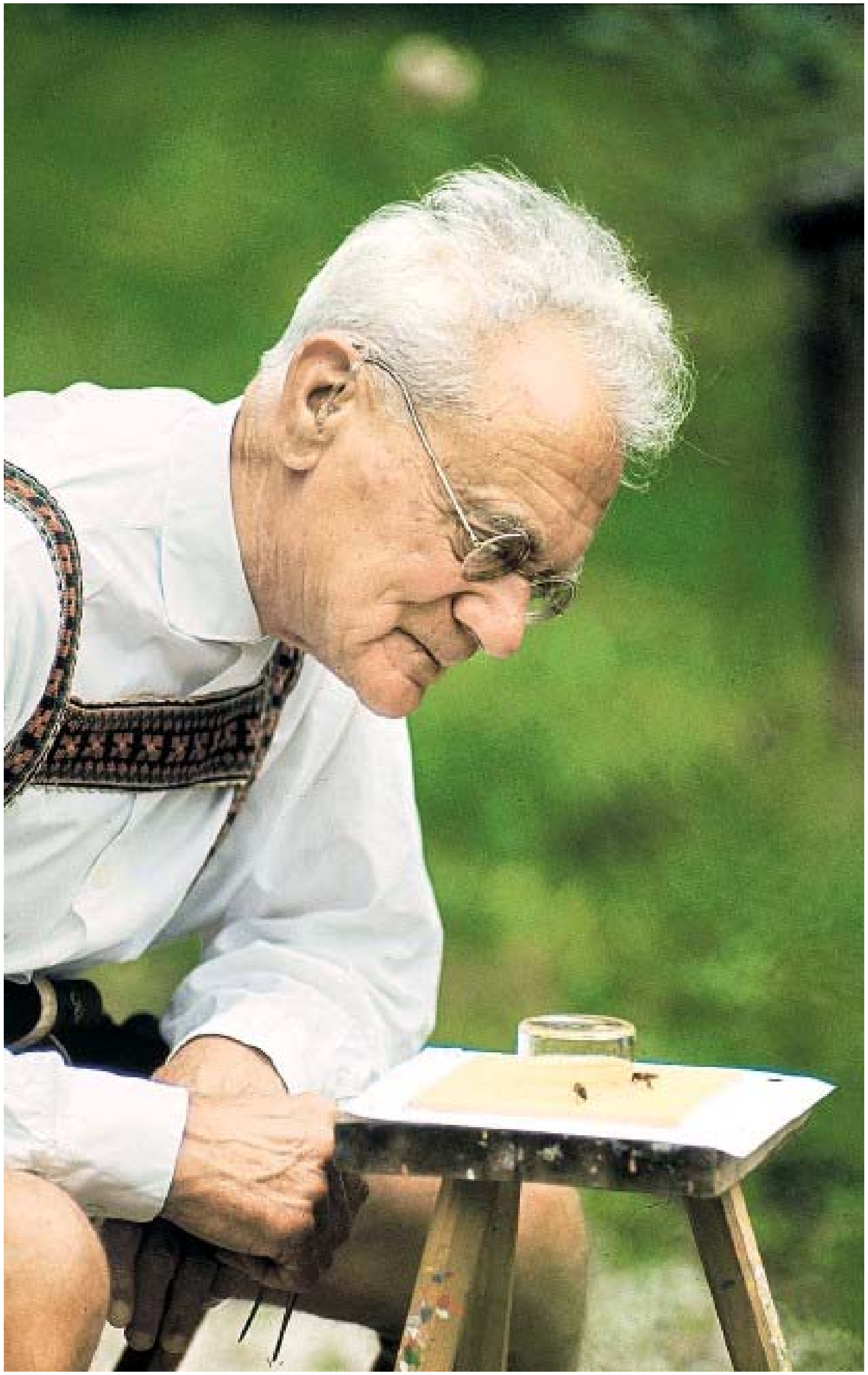
- What quantities do birds sense?
- Vertical velocities, temperature, gradients, etc?
- How should the bird respond to these cues?

Experiments are hard to control and strategies are difficult to infer from limited data

Physics simulations are complex and there are many variables.

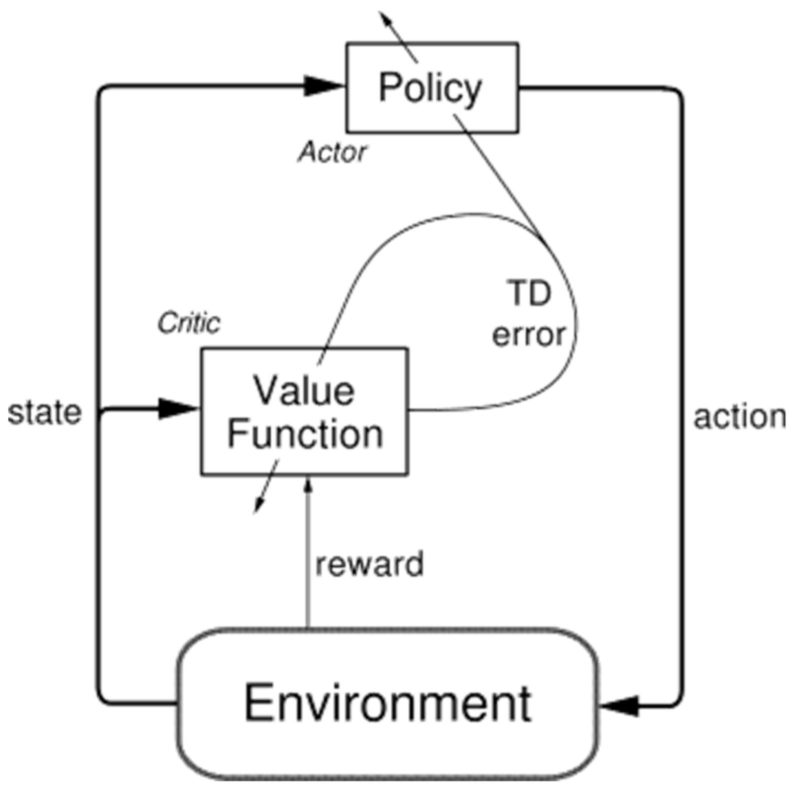
What should an optimal agent sense?


Time is Honey



Karl von Frisch

Temporal Difference Learning



TD - error :

$$\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t)$$

Actions are determined by preferences :

$$\pi_t(s,a) = \Pr\left\{a_t = a \middle| s_t = s\right\} = \frac{e^{p(s,a)}}{\sum_b e^{p(s,b)}},$$

Update the preferences :

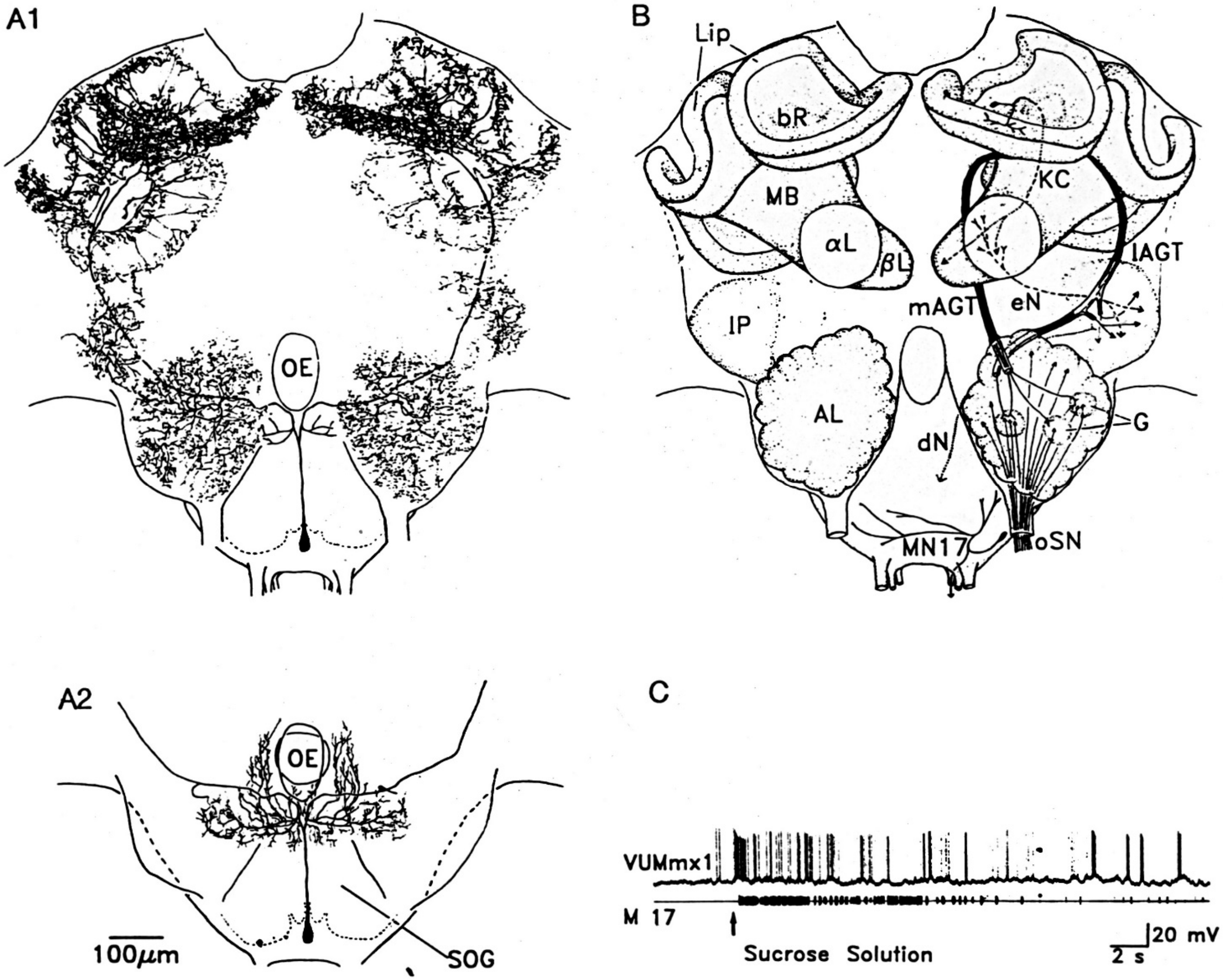
$$p(s_t, a_t) \leftarrow p(s_t, a_t) + \beta\delta_t$$

Sutton and Barto, 1988

The value function update :

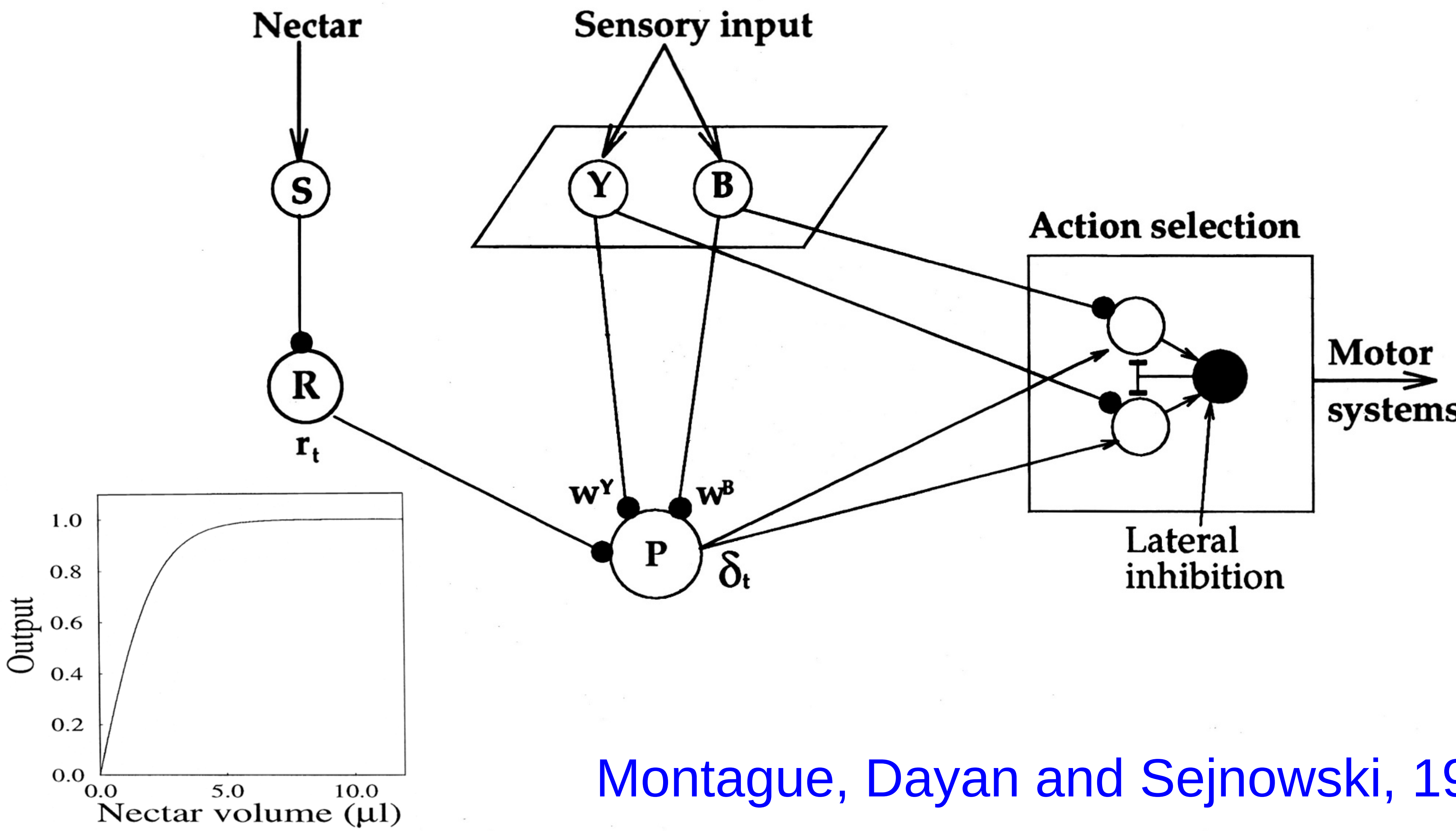
$$V(s_t) \leftarrow V(s_t) + \alpha\delta_t$$

VUMmx1 - Octopamine

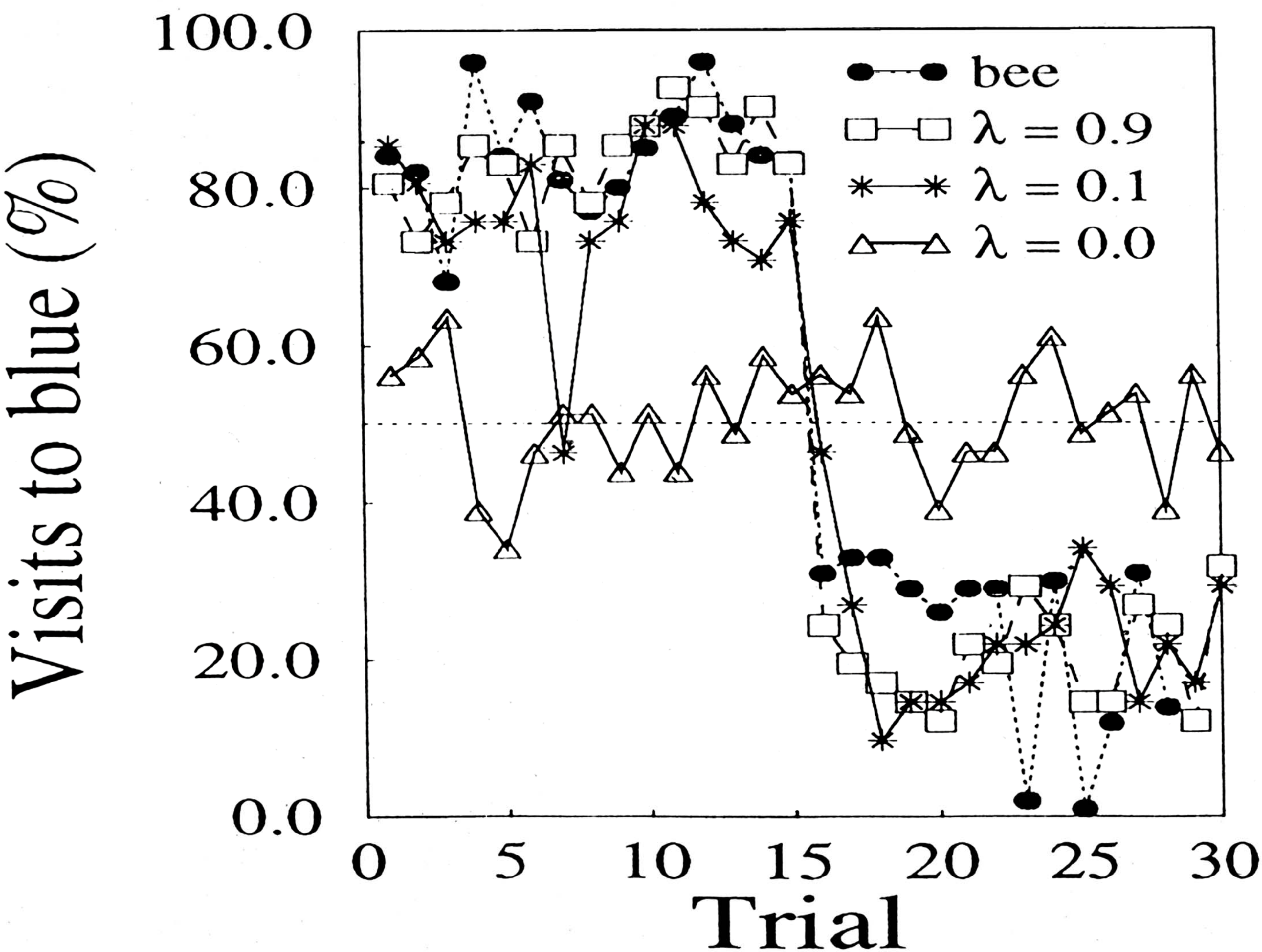


Hammer and Menzel, 1997

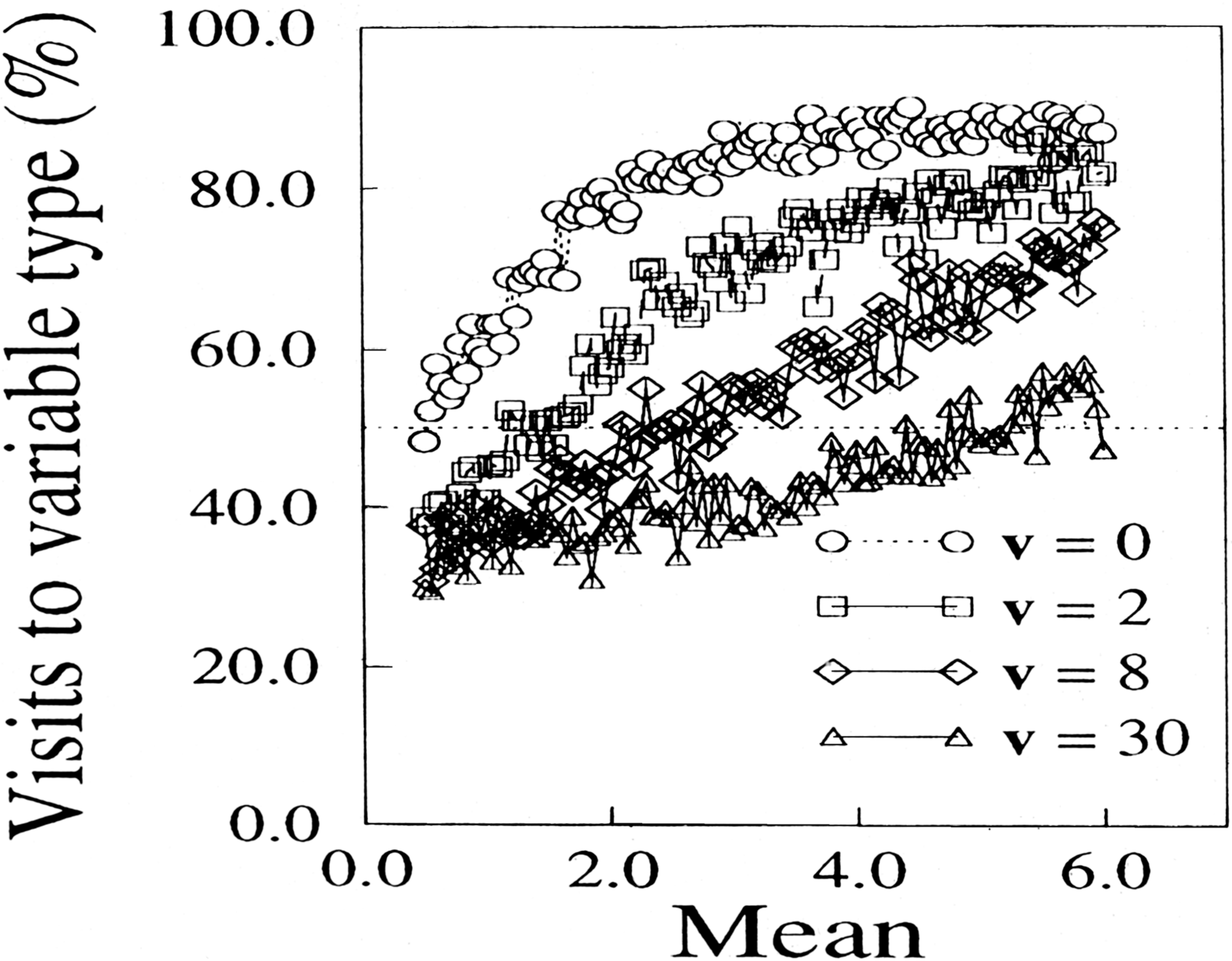
Temporal Difference Learning

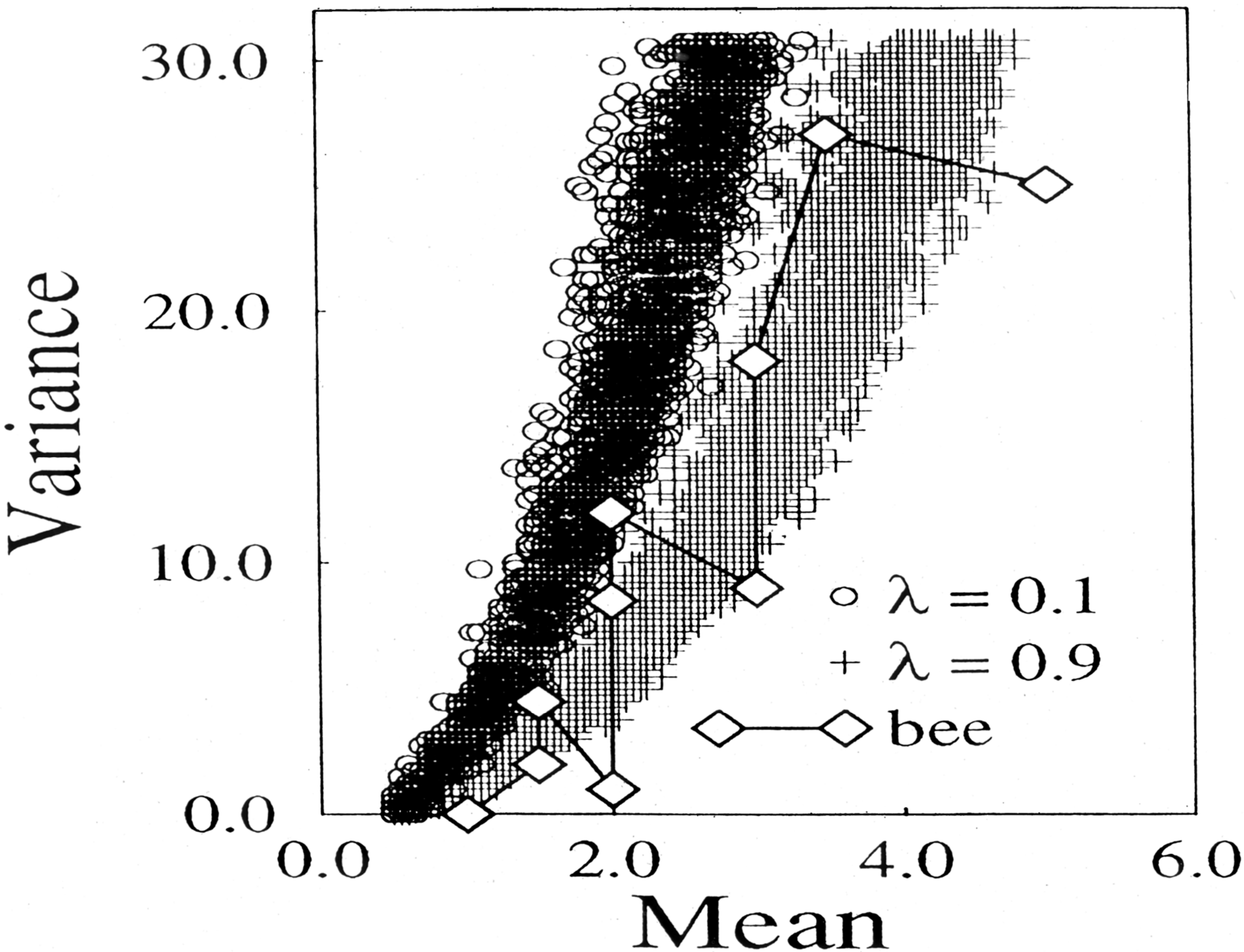


Montague, Dayan and Sejnowski, 1994



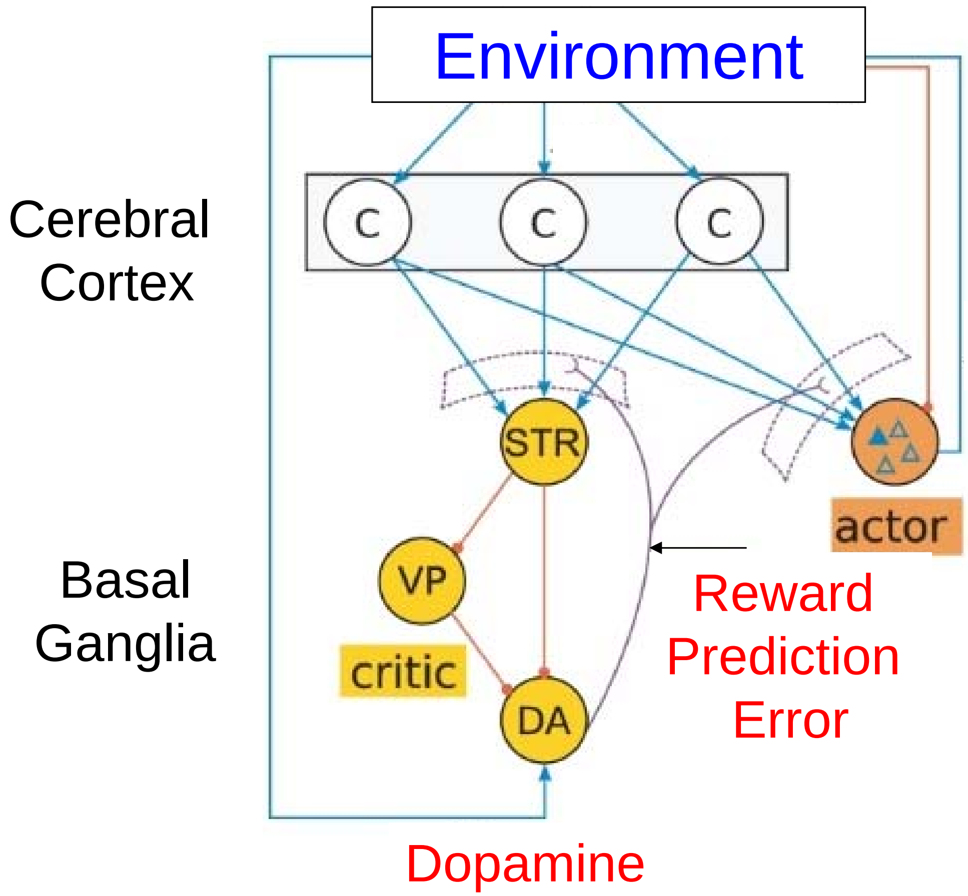
Visits to blue (%)
100.0
80.0
60.0
40.0
20.0
0.0
0
5
10
15
20
25
30
Trial
bee
λ = 0.9
λ = 0.1
λ = 0.0


Visits to variable type (%)
100.0
80.0
60.0
40.0
20.0
0.0
0.0
2.0
4.0
6.0
Mean
v = 0
v = 2
v = 8
v = 30

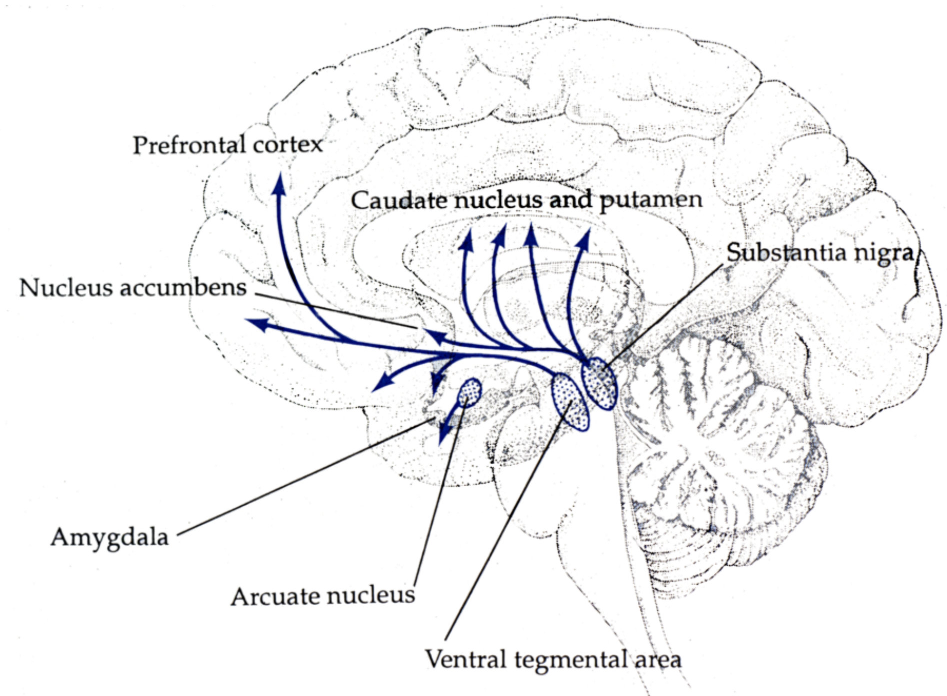


Variance
30.0
20.0
10.0
0.0
0.0
2.0
4.0
6.0
Mean
○ λ = 0.1
+ λ = 0.9
bee

Actor Critic Model



Dopamine Neurons

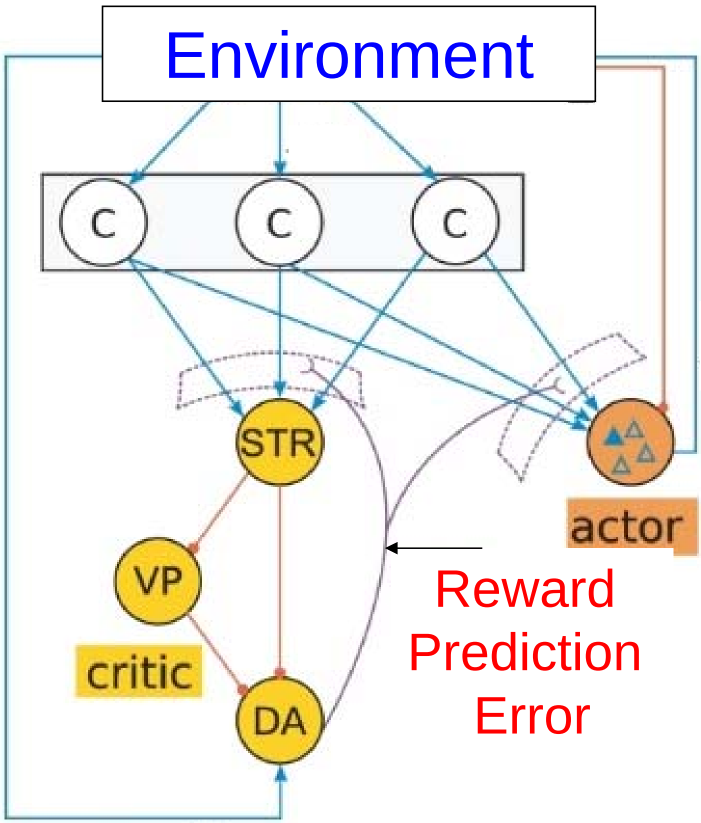


Montague, Dayan and Sejnowski, 1996

Temporal Difference Learning

Cerebral Cortex

Basal Ganglia



Dopamine

Go Defeat, 2017

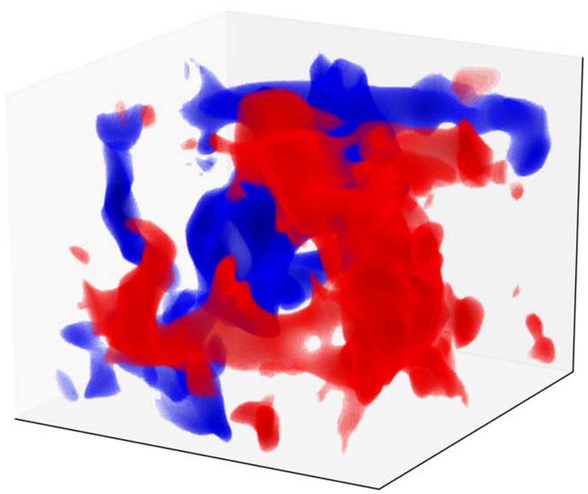
DeepMind

What Do Thermals Look Like?

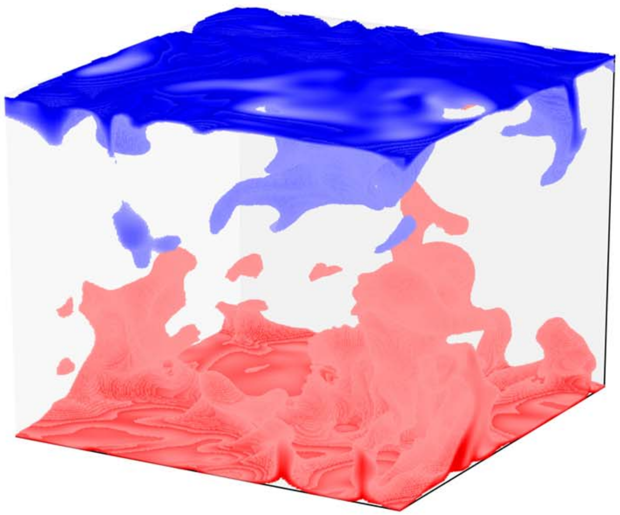
Rayleigh-Benard convection

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} = -\nabla P + \left(\frac{\mathrm{Pr}}{\mathrm{Ra}}\right)^{1/2} \nabla^2 \boldsymbol{u} + \boxed{\theta \hat{\boldsymbol{z}}}$$

$$\frac{\partial \theta}{\partial t} + \boldsymbol{u} \cdot \nabla \theta = \frac{1}{(\mathrm{Pr}\,\mathrm{Ra})^{1/2}} \nabla^2 \theta \,,$$



Vertical velocity field

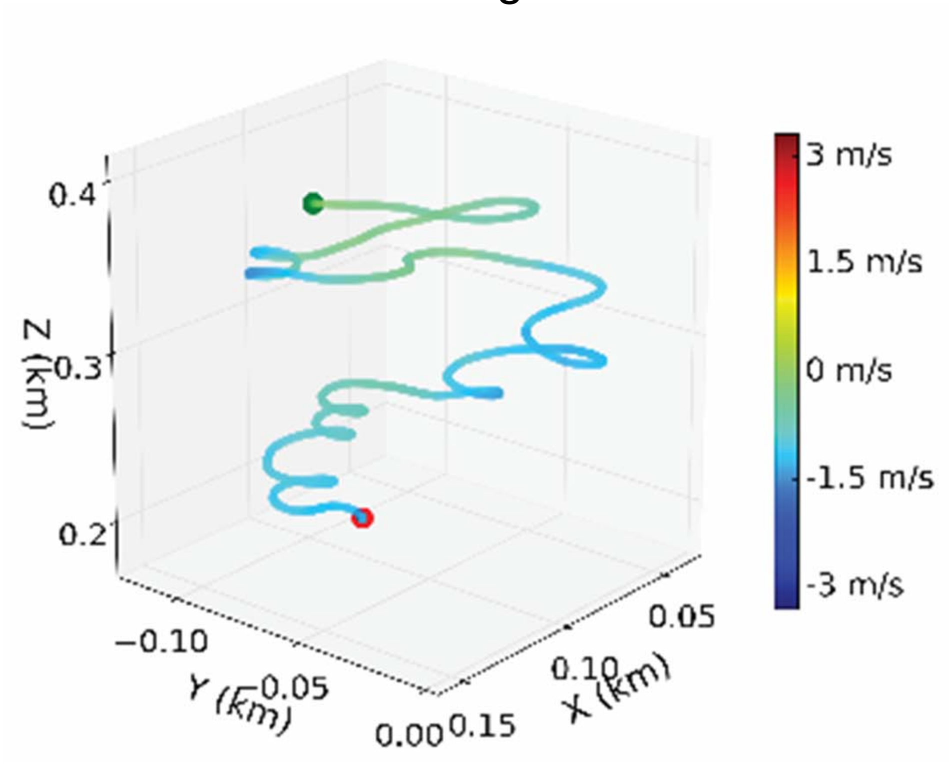


Temperature field

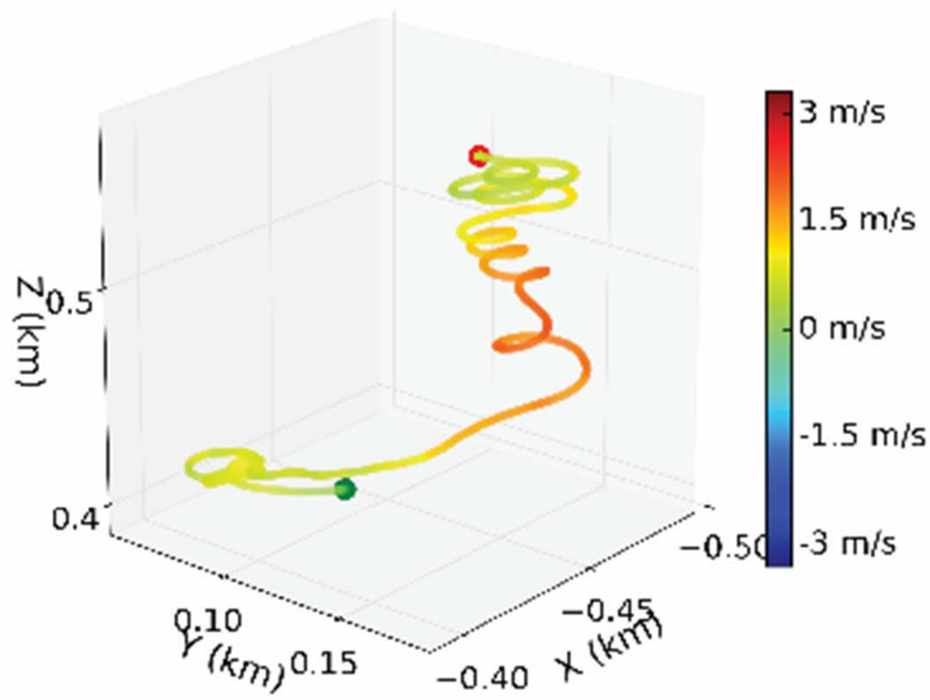
Reddy, Vergassola, Sejnowski, 2017

Sink or Soar?

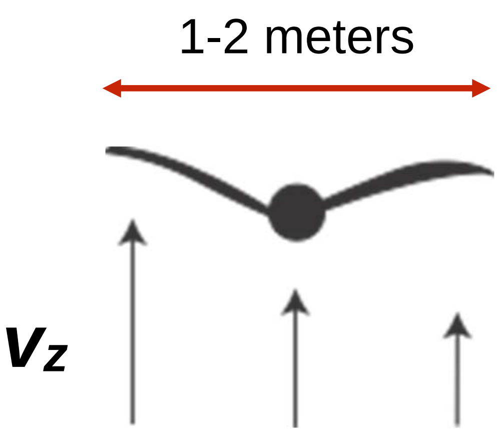
Pre-training

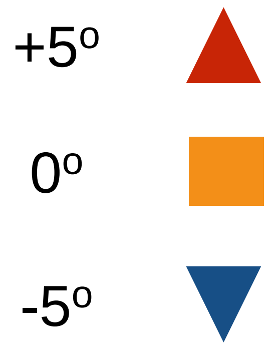


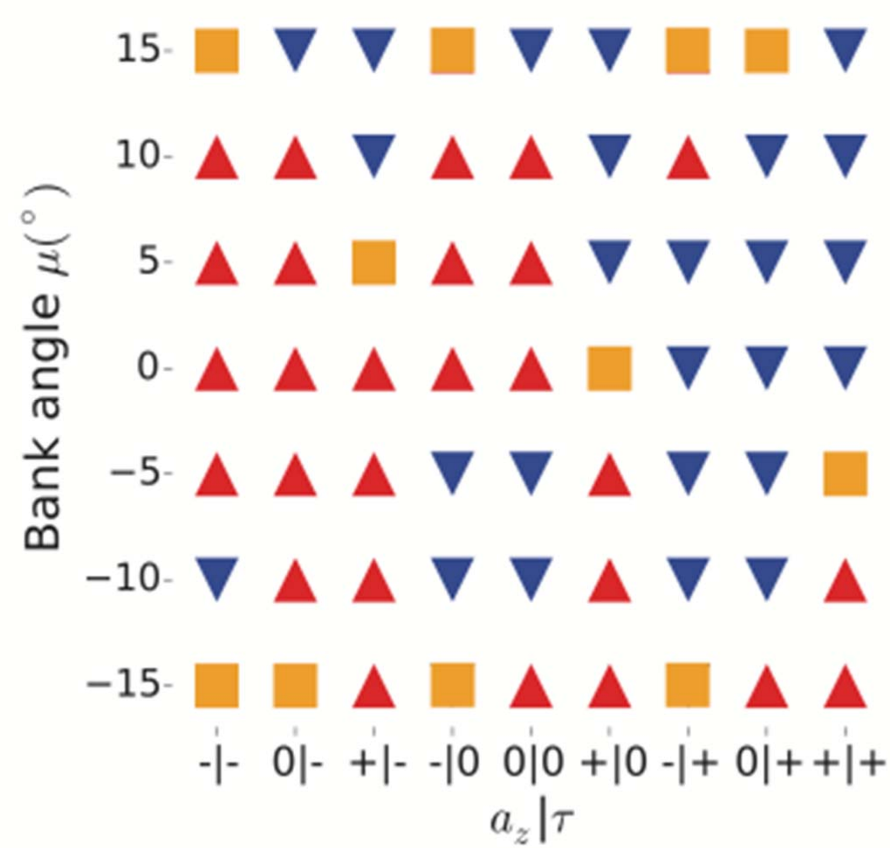
Post-training



Learned Policy



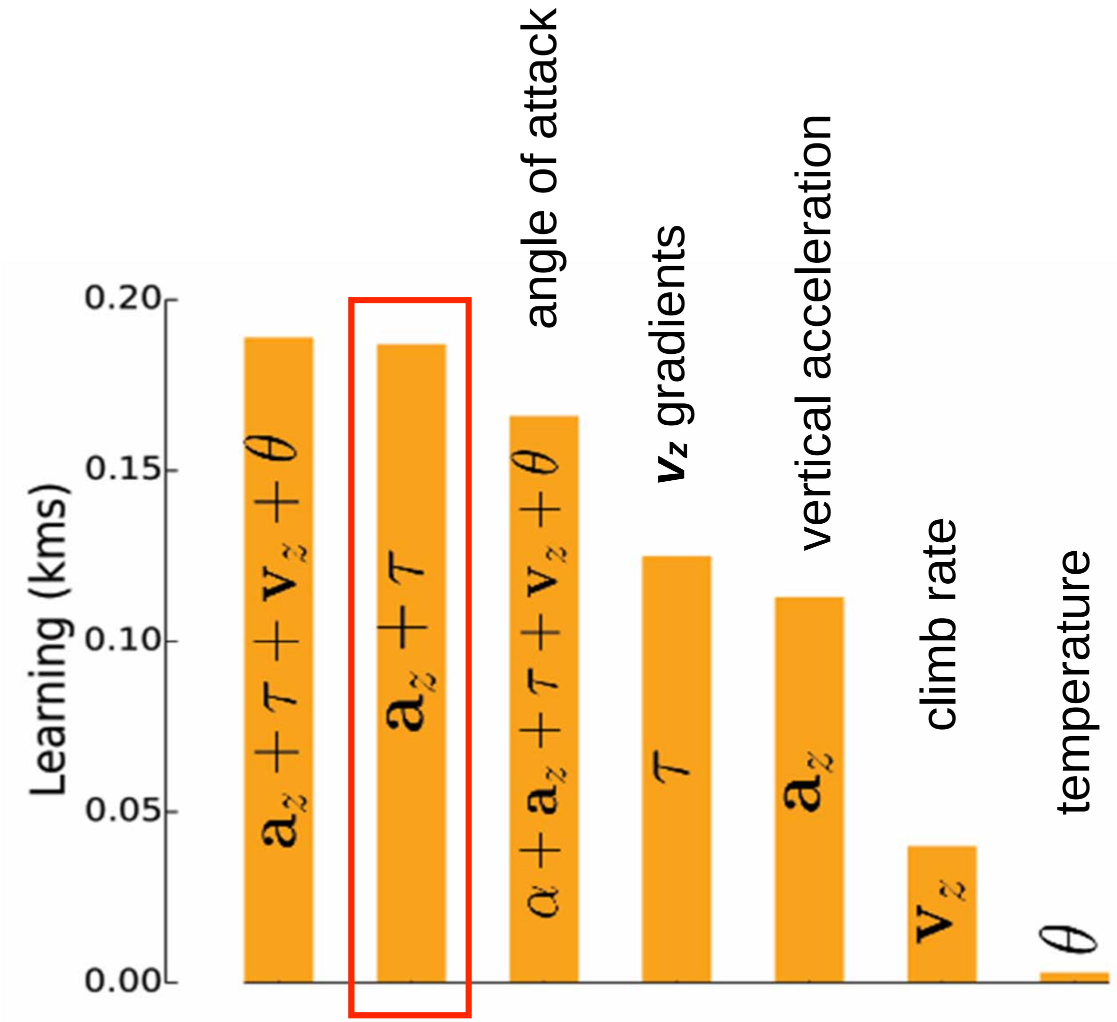




τ Vertical velocity gradient

a_z Vertical acceleration

Conclusions



a_z and
v_z gradients
across wings
are useful

control over
angle of attack
is not useful

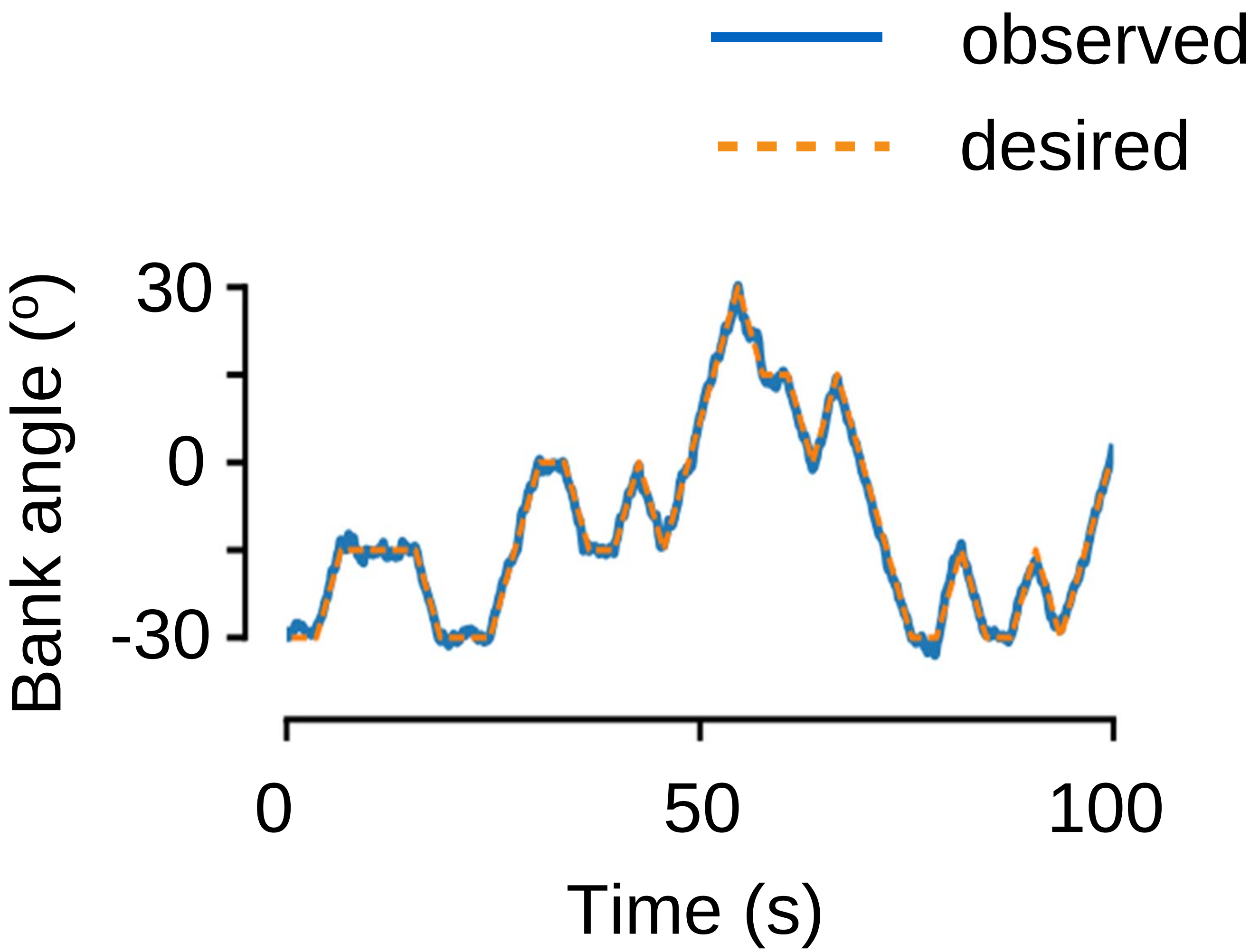
Field Experiments

GoPro Glider

Field Experiments

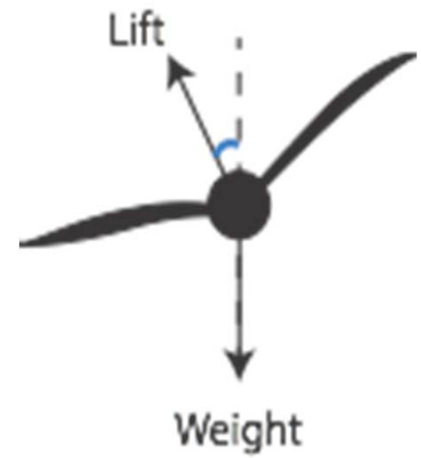
Gautam Reddy

Field Experiments



Measuring the Vertical Wind Velocity

GPS and barometer measurement give vertical ground velocity

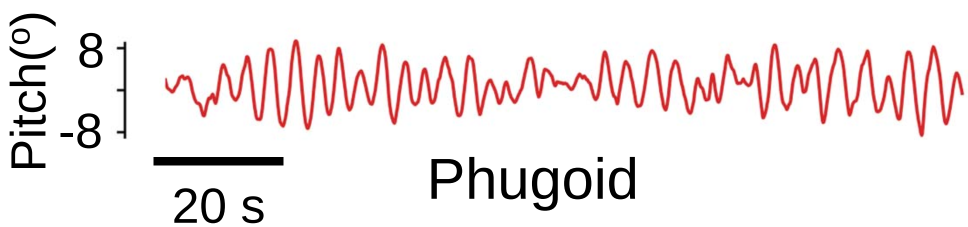


We need to estimate ***wind*** velocity

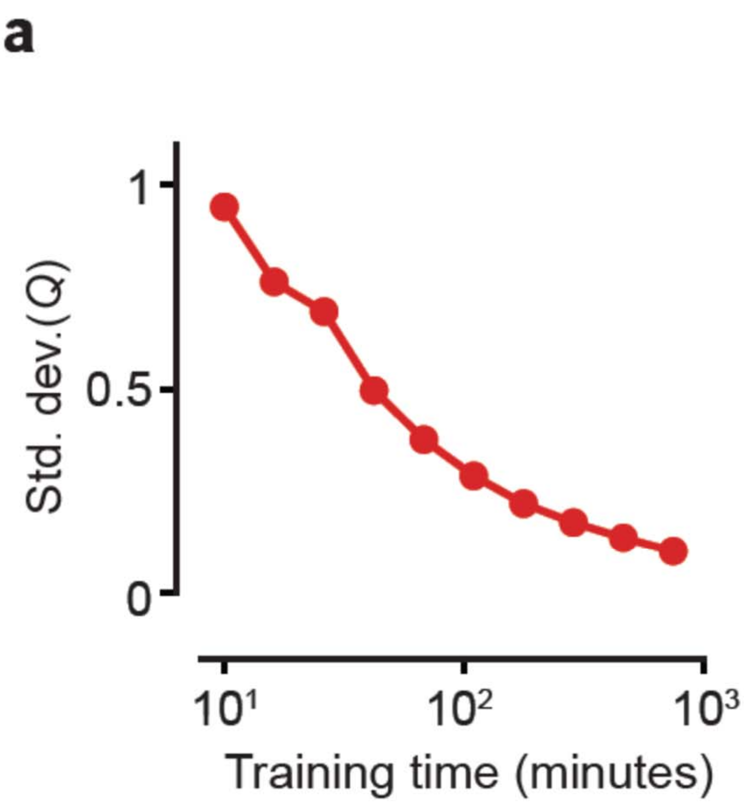
GPS/baro

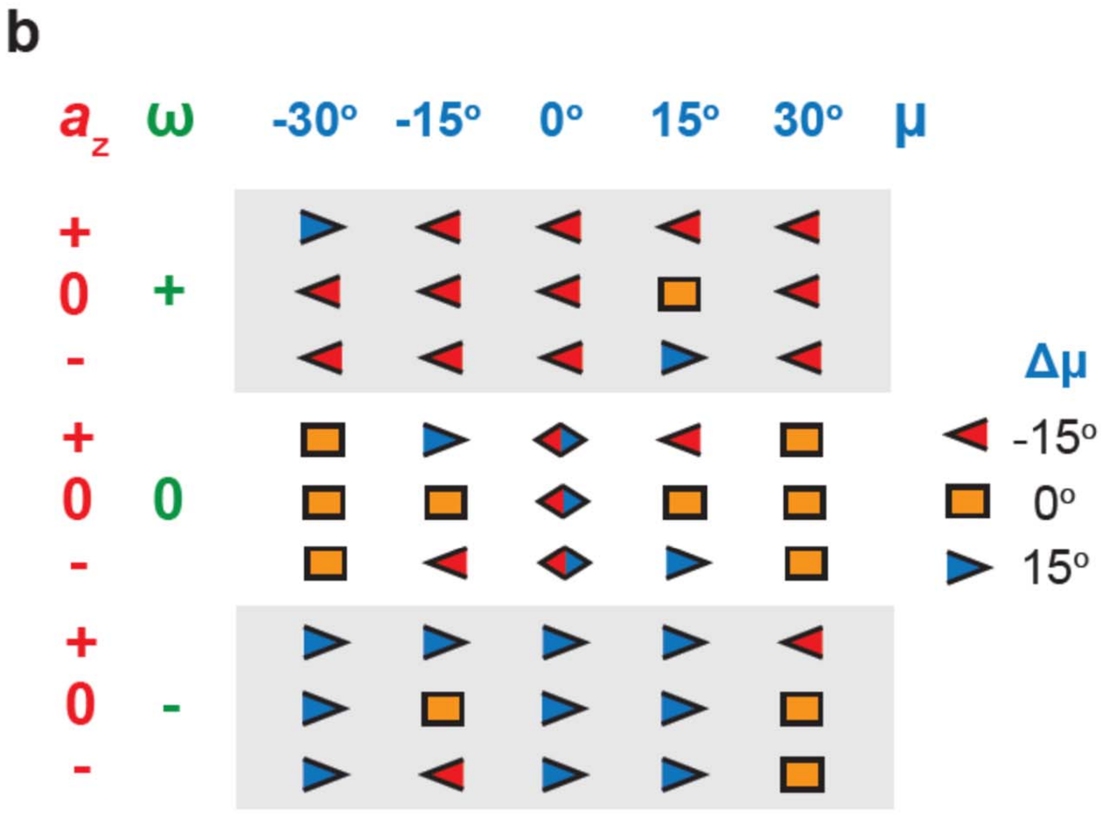
ground vel. = wind vel. + *glider's air vel.*

modeling



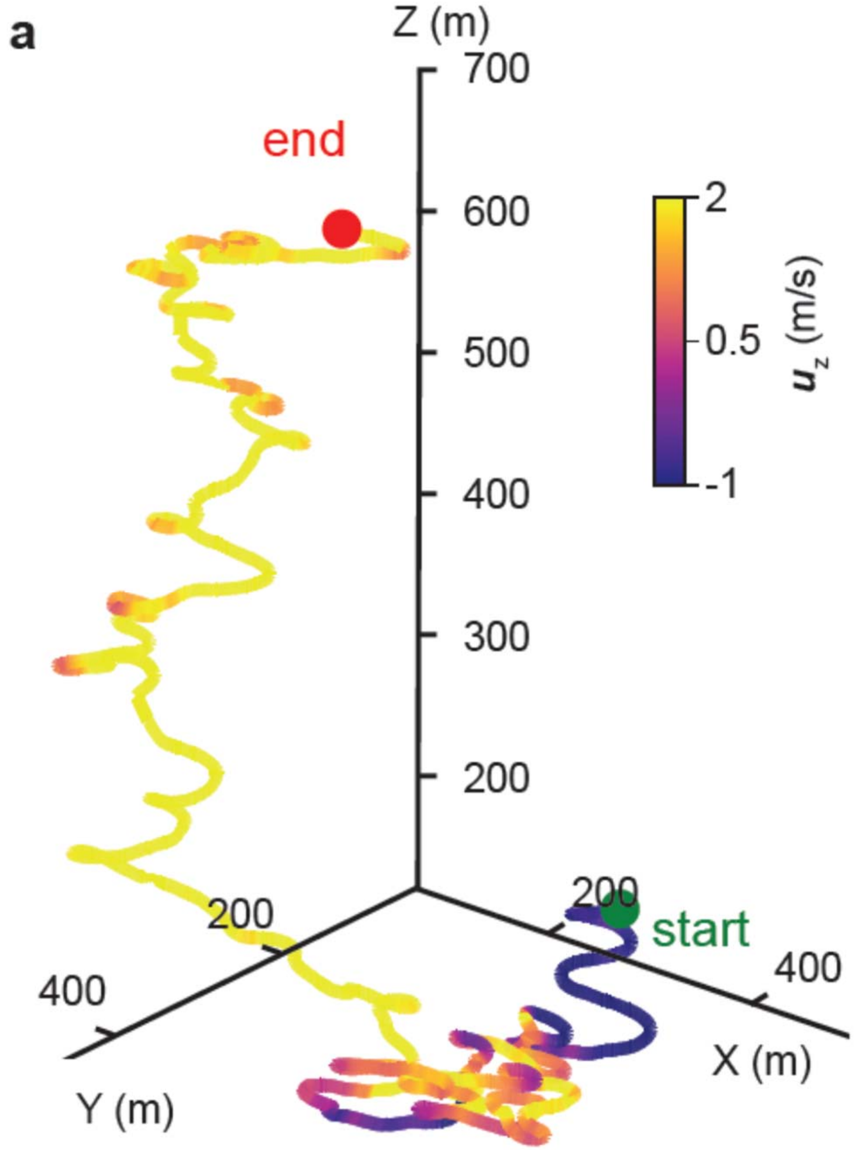
Training a Glider in the Field





Reddy, Vergassola, Sejnowski, 2018

Training a Glider in the Field



Field Experiments

artificial intelligence **meets** human intelligence

THE DEEP LEARNING REVOLUTION

TERRENCE J. SEJNOWSKI

Thank You

Peter Dayan

Read Montague

Gautam Reddy

Massimo Vergassola

John Doyle

SIMONS FOUNDATION