

Does Learning Mean a Decrease in Entropy?

Paul Smaldino

NCASO Final Project

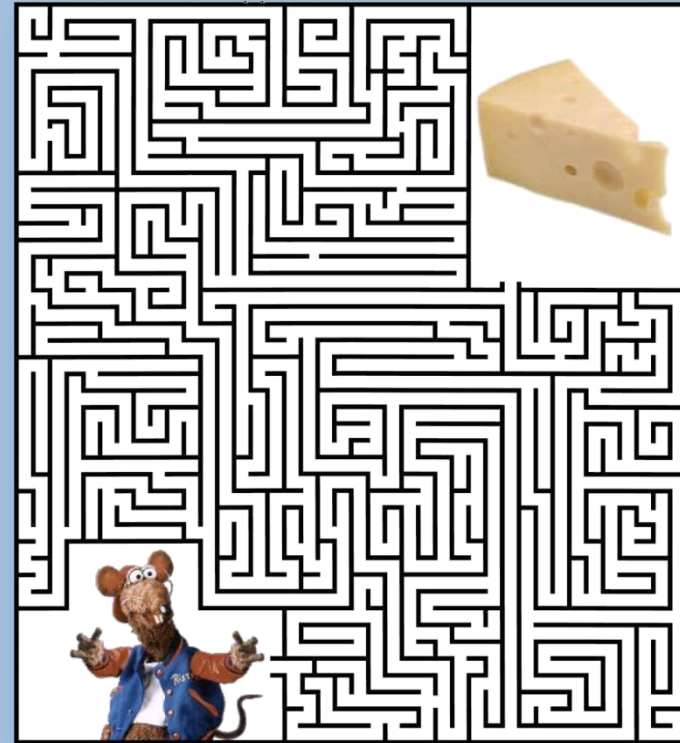
March 13, 2009

Spatial Exploration

Where am I? How did I get here?

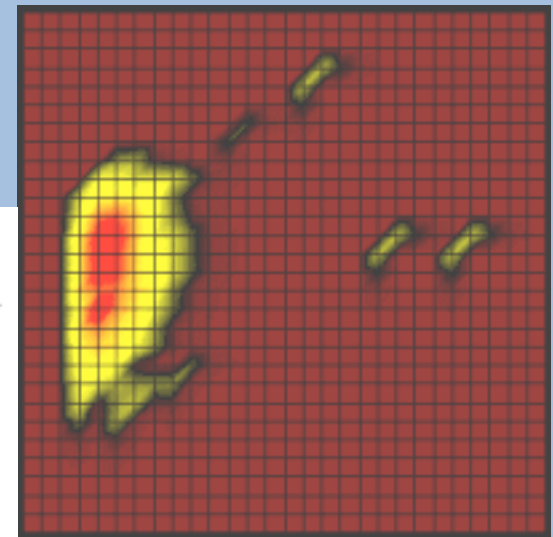
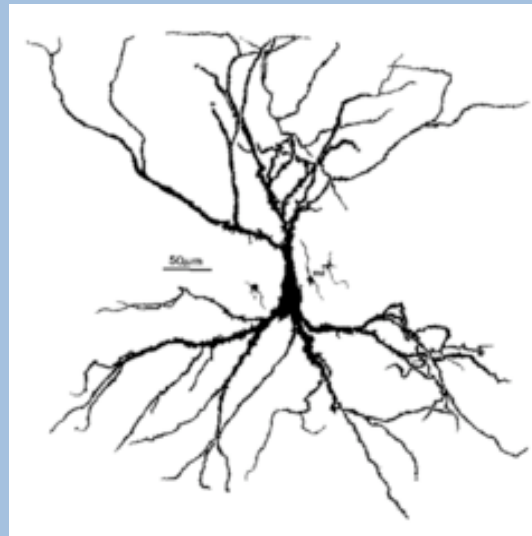
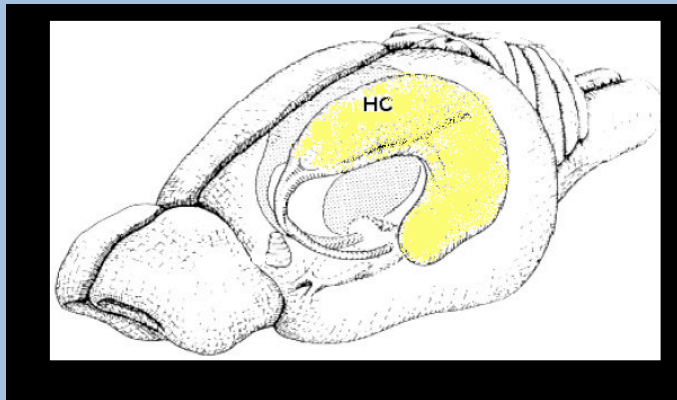
Where am I going?

- Animal behavior
 - Foraging
 - Territoriality
- Cognitive Science
 - Consolidation of learning and memory
 - “More is up.”

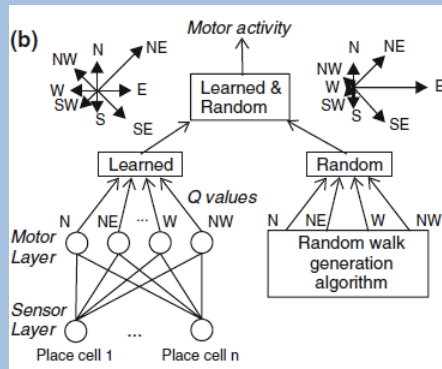


Hippocampal Place Cells

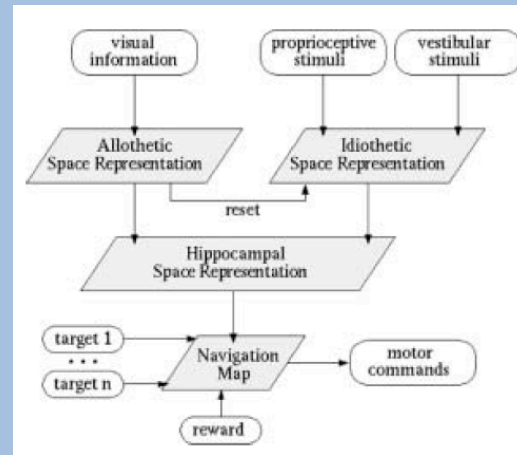
- Pyramidal neurons in CA1 and CA3
- Maximal firing rate at preferential location



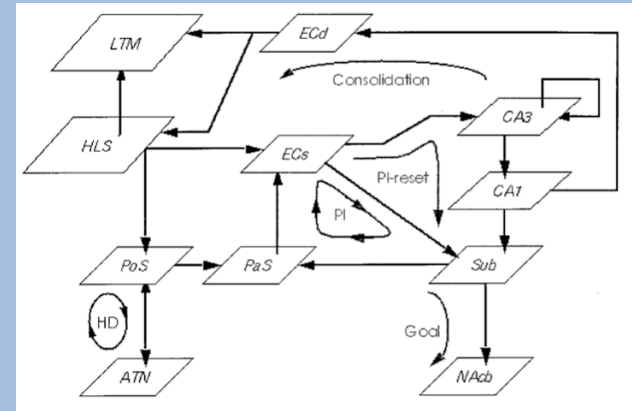
Many Computational Models



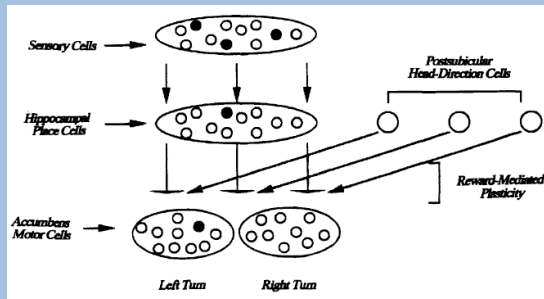
Tamosiunaite et al., 2008



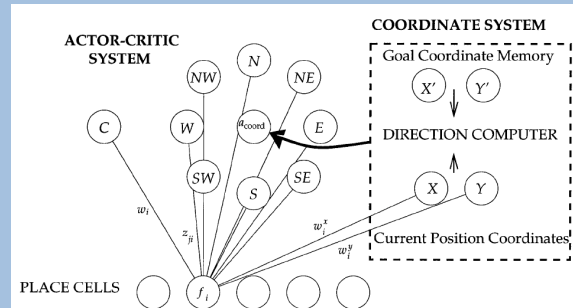
Arleo & Gerstner, 2000



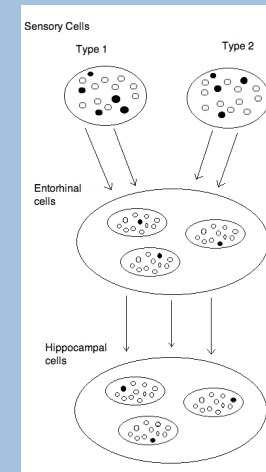
Redish & Touretzky, 1998



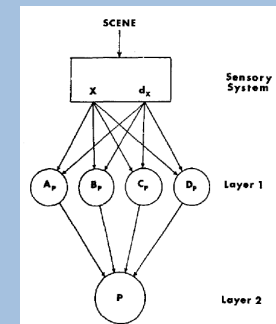
Brown & Sharp, 1995



Foster et al., 2000



Sharp, 1991



Zipser, 1985

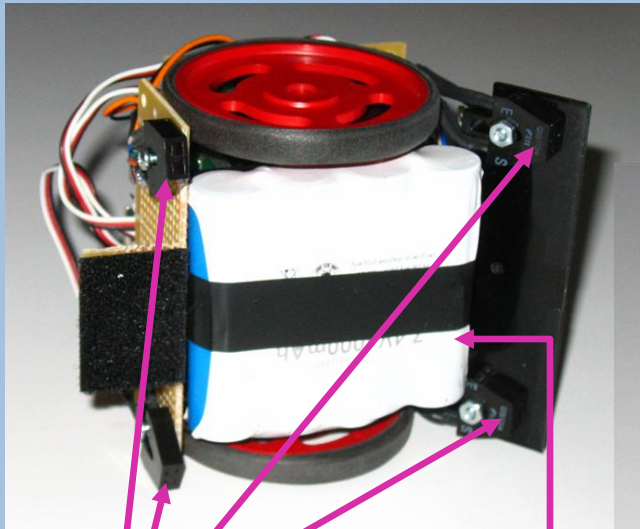
Learning and Entropy

- Entropy is a measure of “uncertainty”
- If learning reduces uncertainty, entropy should decrease.
- If entropy steadily declines: Woo!
- If something else happens: Interesting.

- We can also check out some other measures and see what comes up.

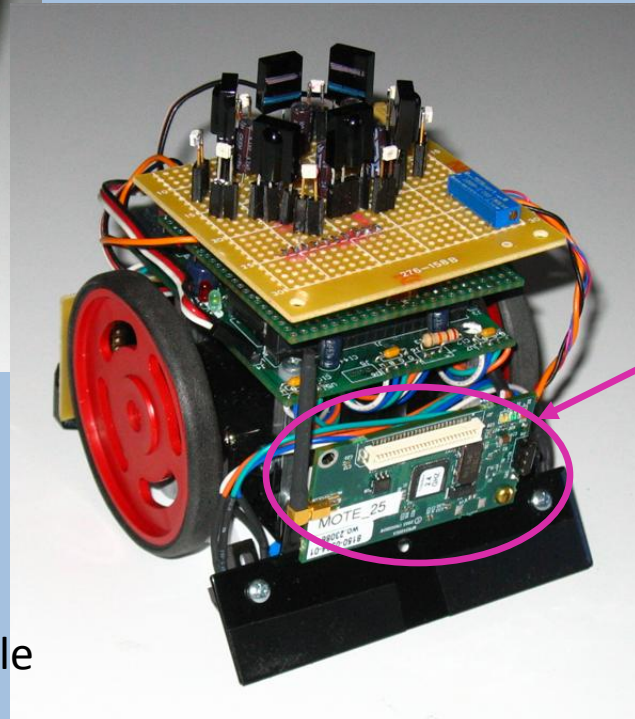
ROBOTS!!!

Sumo robot platform

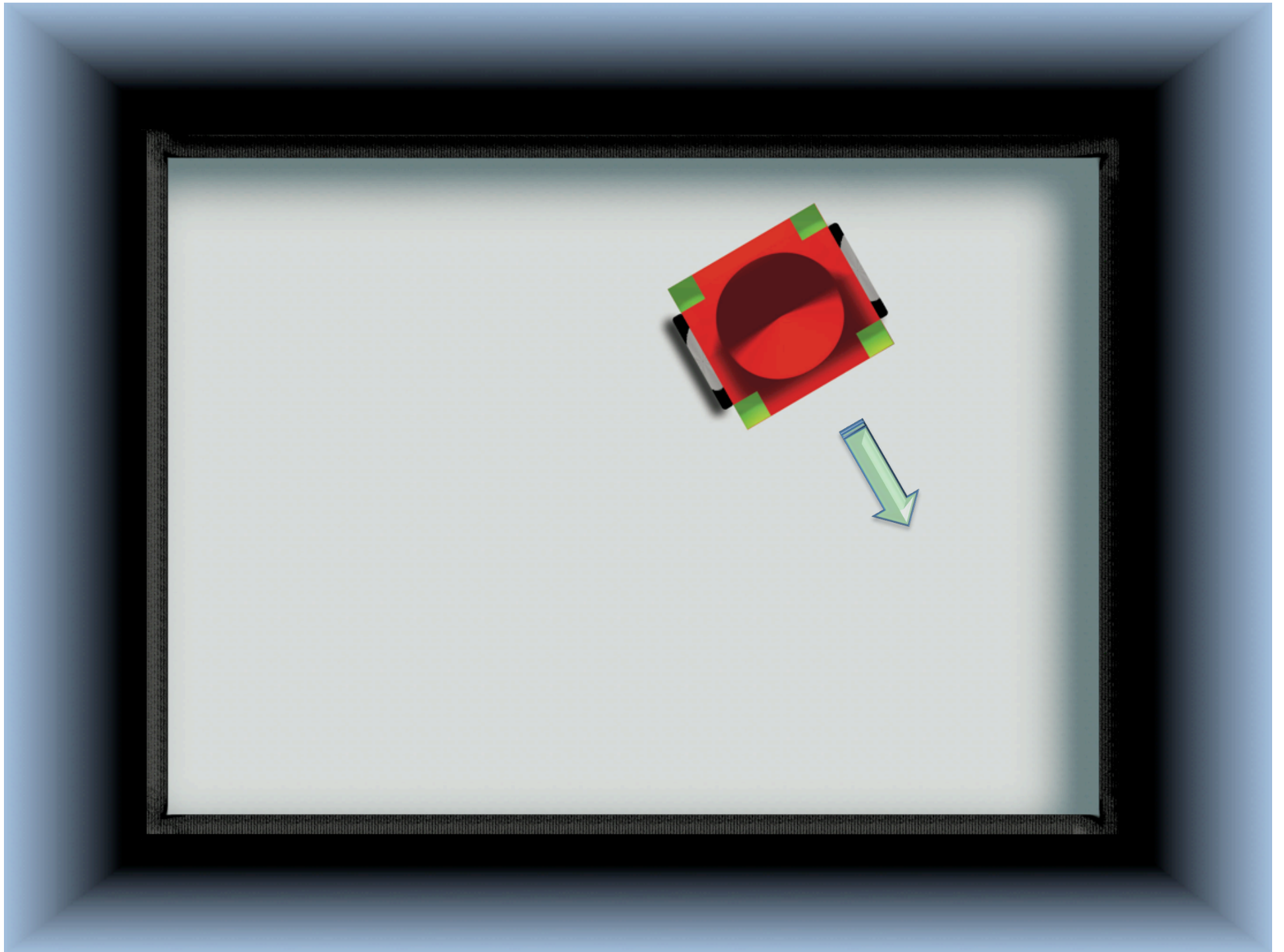


Infrared "Wall" sensors
(4 corners)

Rechargeable/Replaceable
battery pack



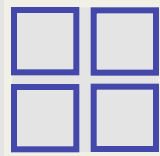
Mote (RF comm.)



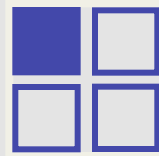
Edge Sensor States

Edge sensor data: [MSb ... LSb] = [RB LB RF LF]

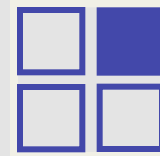
Valid
Combinations



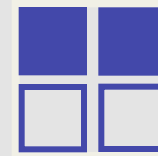
0: [0000]



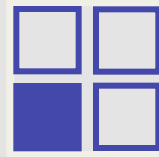
1: [0001]



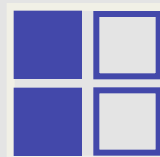
2: [0010]



3: [0011]



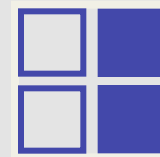
4: [0100]



5: [0000]



8: [1000]

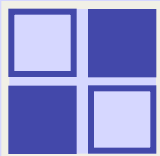


10: [1010]

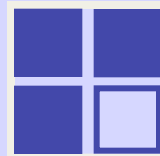


12: [1100]

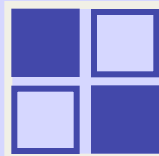
“Less” Valid Combinations



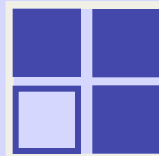
6: [0110]



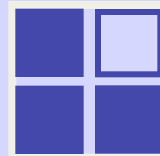
7: [0111]



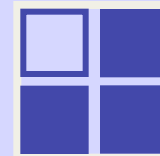
9: [1001]



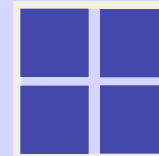
11: [1011]



13: [1101]



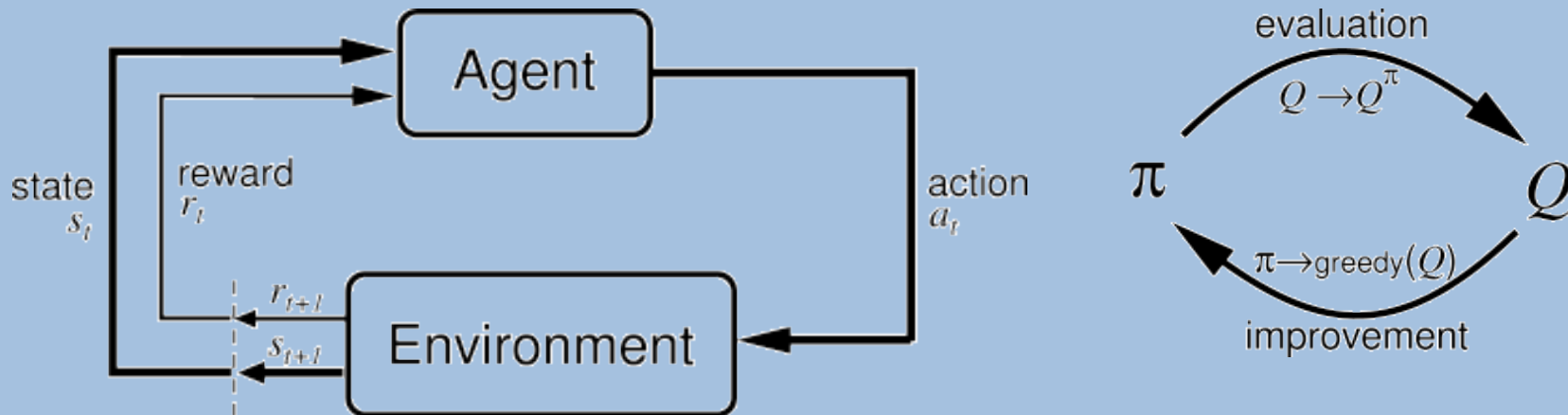
14: [1110]



15: [1111]

Reinforcement Learning

- A Model Framework for Learning
 - Common method of unsupervised learning
 - Provides discrete structure for analysis



Monte Carlo RL

Action set:

- ('F', 'L', 'R', 'A', 'N')
- Forward, Left 90°, Right 90°, About-Face (L 180°), No Move

Reward scheme:

- “Points” given for amount of time in forward travel
(1sec forward = 1 pt)
- 2 Points for 'L' or 'R' that results in a clear path previously blocked
- 1 Point for 'A' that results in a clear path previously blocked
- 0 Points else

Monte Carlo RL

RL Policy:

- All (s,a) pairs begin equiprobable
- “Exploitation”:
 - Once a positive reward is given for (s,a), policy is greedy--always choose action with highest previous return
- “Exploration”:
 - System works in “episodes”; at beginning of each episode, agent explores by picking a random action regardless of state

Sequence of Events:

- Get current state
- Get next action from policy and execute
- Get reward resultant from action and state
- Update policy with new rewards

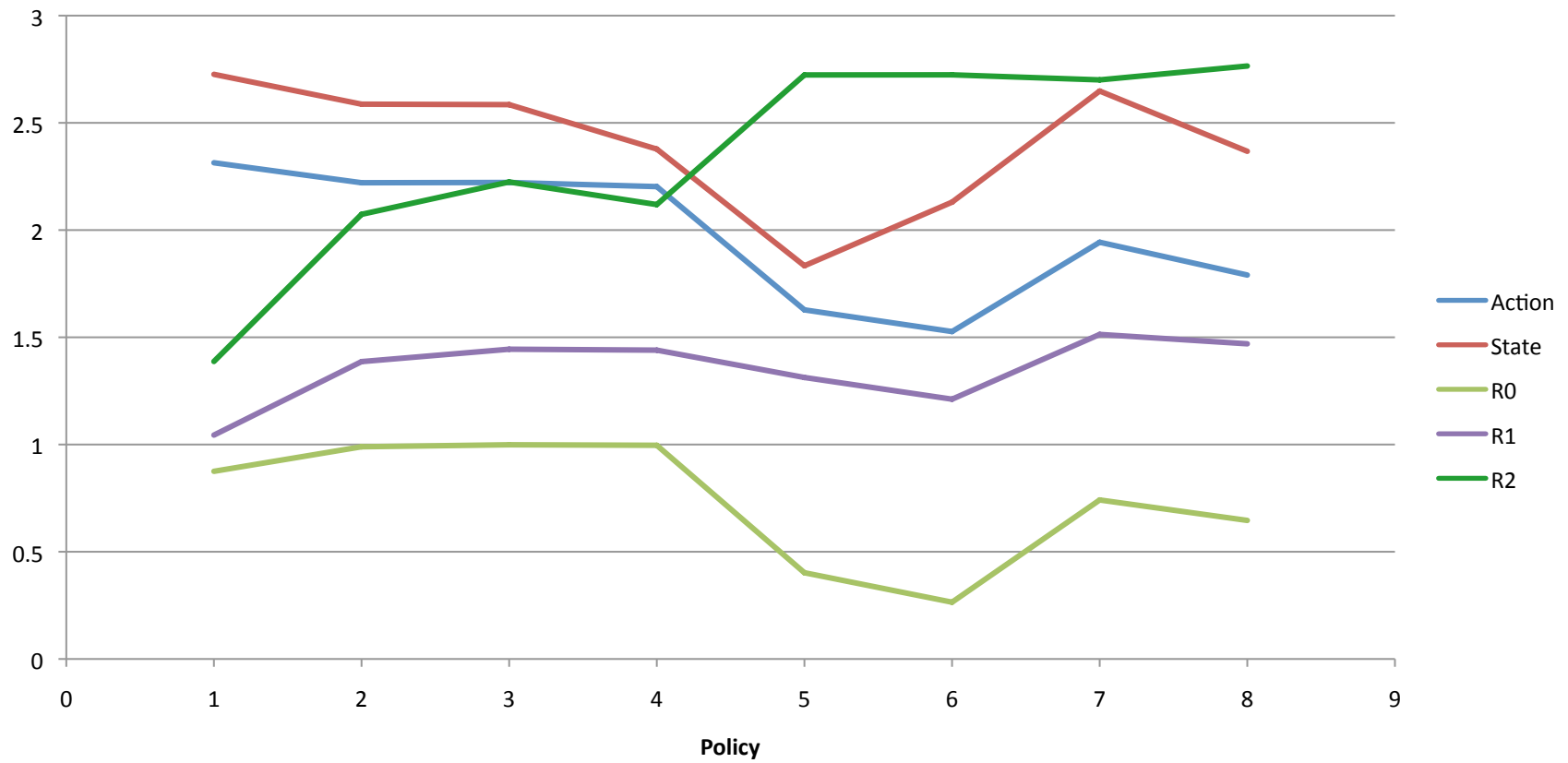
Data Used

- One trial of robot learning (500 timesteps)
- Each policy in process subsequently run independently (no learning) for 200 timesteps
- Frequency of observed actions, states, and rewards was to estimate measures of information

The Policies

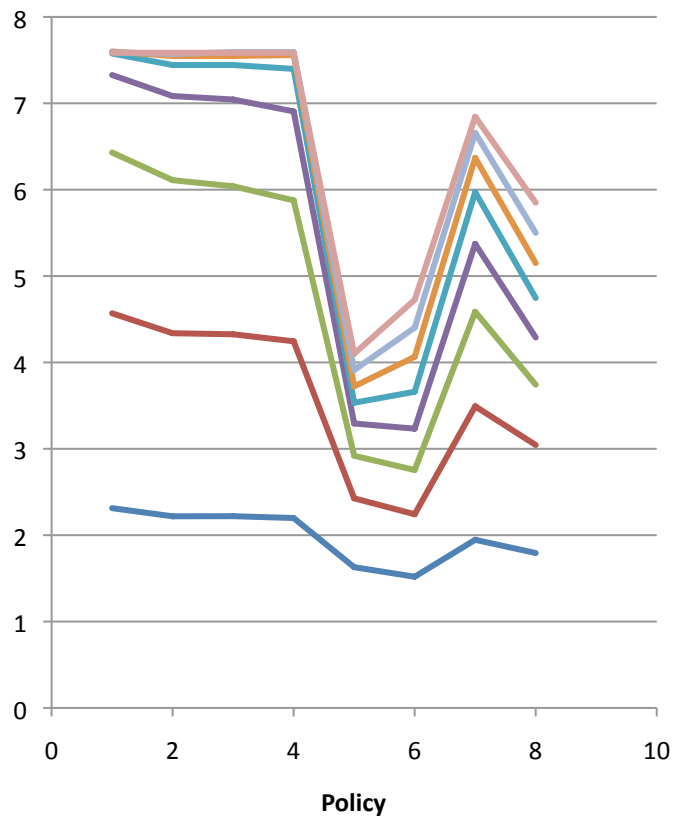
Pol#	PF	P[0]	P[1]	P[2]	P[3]	P[4]	P[5]	P[8]	P[10]	P[12]
1	6	RAND	RAND	RAND	RAND	RAND	RAND	RAND	RAND	RAND
2	2	F	RAND	RAND	RAND	RAND	RAND	RAND	RAND	RAND
3	3	F	RAND	RAND	A	RAND	RAND	RAND	RAND	RAND
4	2	F	RAND	A	A	RAND	RAND	RAND	RAND	RAND
5	1	F	R	A	A	RAND	RAND	RAND	RAND	RAND
6	52	F	R	A	A	F	RAND	RAND	RAND	RAND
7	26	F	R	L	A	F	RAND	RAND	RAND	RAND
8	410	F	R	L	A	F	RAND	F	RAND	RAND

Entropy

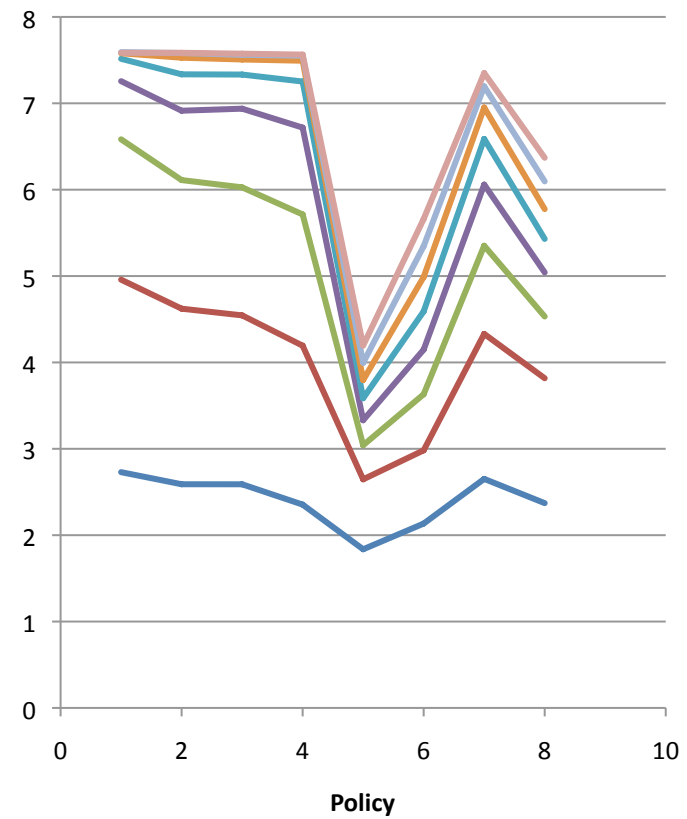


Block Entropy

Block Entropy - Action



Block Entropy - State

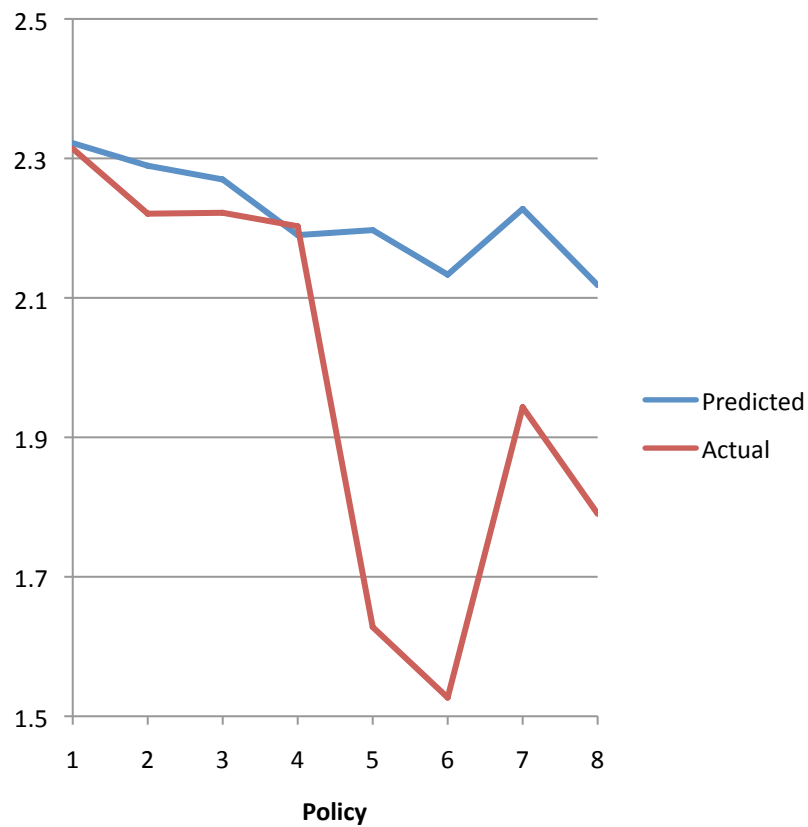


Action Entropy and State Distribution

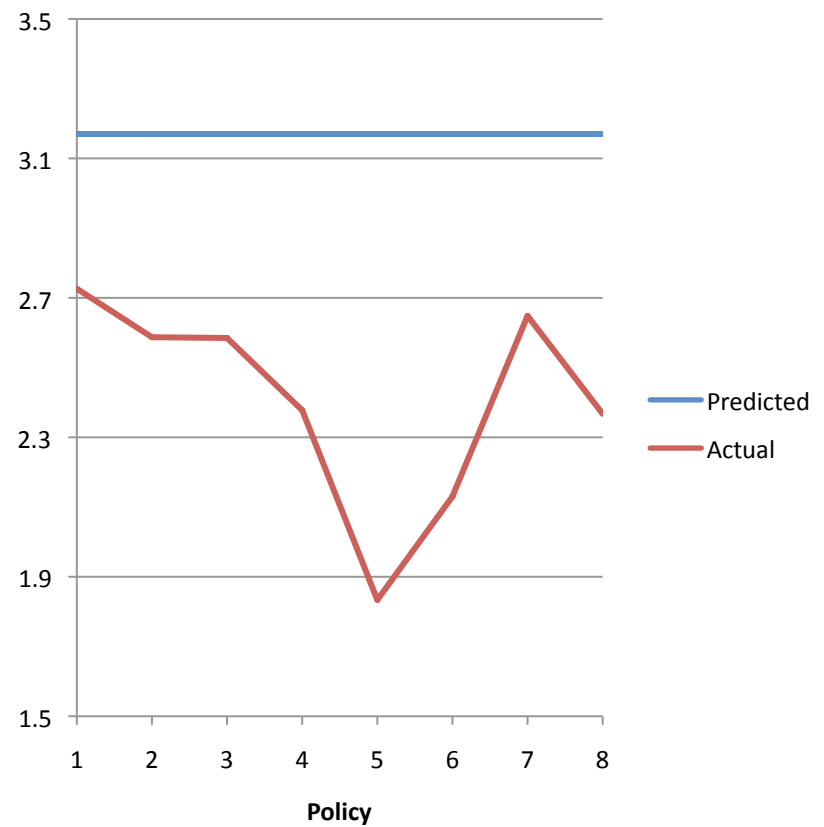
- Can we learn something about the distribution of states from entropy of the actions?
 - If states occur randomly, then we can predict the action entropy simply by looking at the policy.
 - If the action entropy is different, what does this tell us about the state transitions?

Learning About State Distribution

Action Entropy



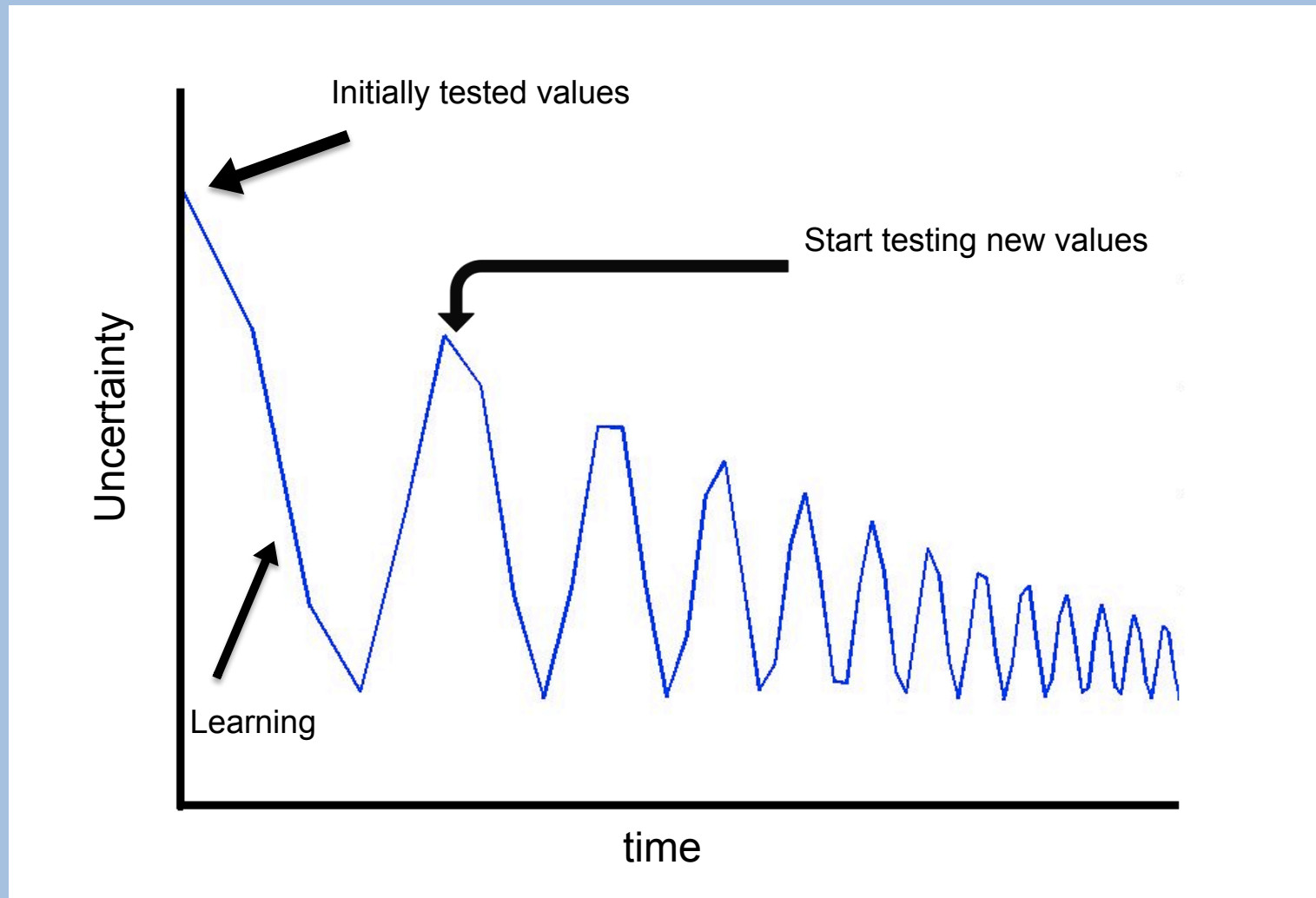
State Entropy



The Policies

Pol#	PF	P[0]	P[1]	P[2]	P[3]	P[4]	P[5]	P[8]	P[10]	P[12]
1	6	RAND	RAND	RAND	RAND	RAND	RAND	RAND	RAND	RAND
2	2	F	RAND	RAND	RAND	RAND	RAND	RAND	RAND	RAND
3	3	F	RAND	RAND	A	RAND	RAND	RAND	RAND	RAND
4	2	F	RAND	A	A	RAND	RAND	RAND	RAND	RAND
5	1	F	R	A	A	RAND	RAND	RAND	RAND	RAND
6	52	F	R	A	A	F	RAND	RAND	RAND	RAND
7	26	F	R	L	A	F	RAND	RAND	RAND	RAND
8	410	F	R	L	A	F	RAND	F	RAND	RAND

How Might Biological Learning Occur?



Conclusions?

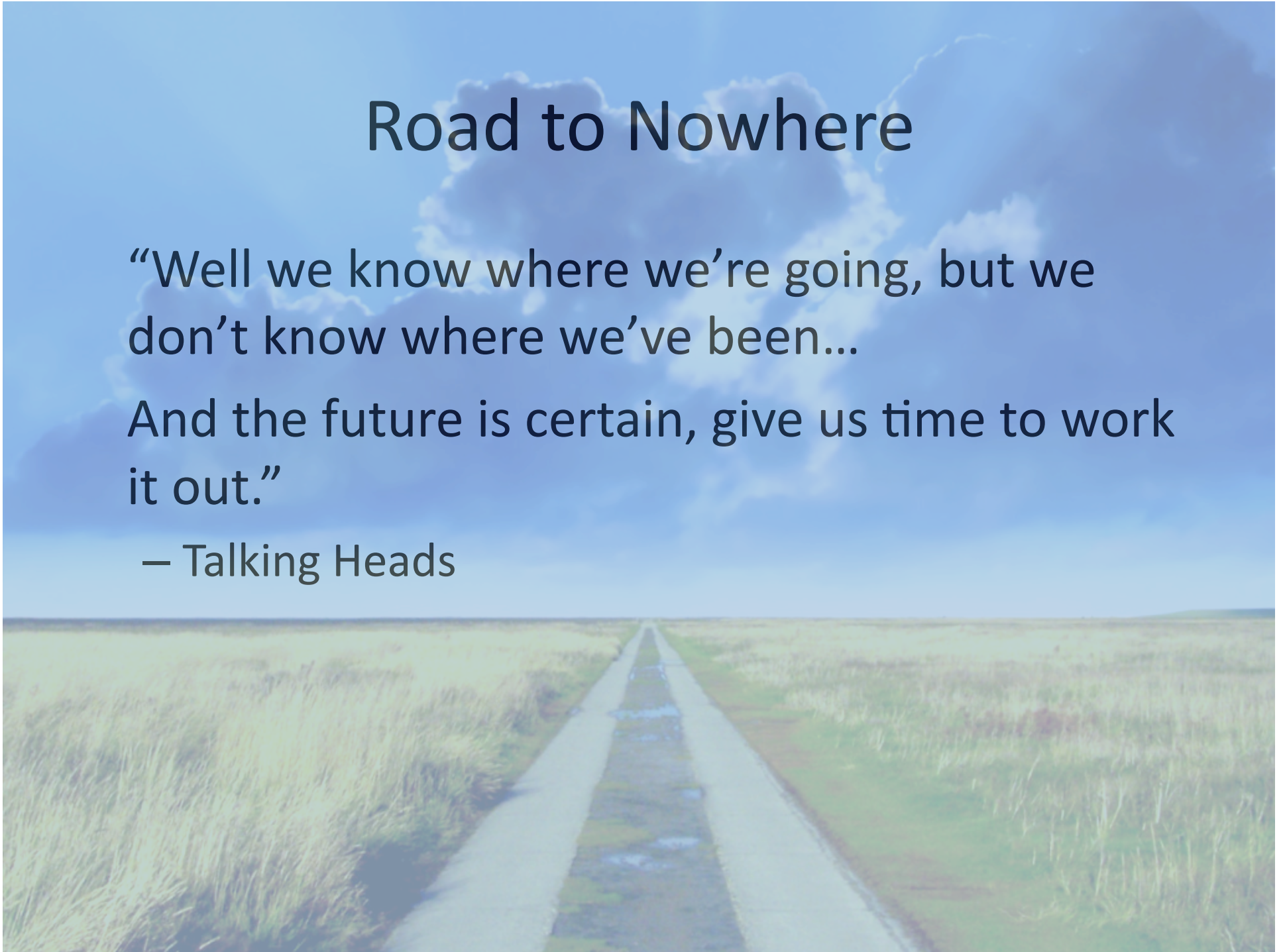
- Nothing definite.
- Maybe a way to study natural (i.e. biological) learning?
 - What can we learn about the internal state given the observed behavior?

Road to Nowhere

“Well we know where we’re going, but we don’t know where we’ve been...

And the future is certain, give us time to work it out.”

– Talking Heads



Thanks!

