Decision making 101

Peter Latham
Gatsby Computational Neuroscience Unit
UCL

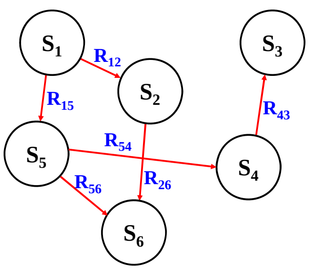
December 4, 2013

Why is it hard to make good decisions?

Because we have to maximize future rewards.

 or  ?

There is a whole field – reinforcement learning – devoted to answering this question.

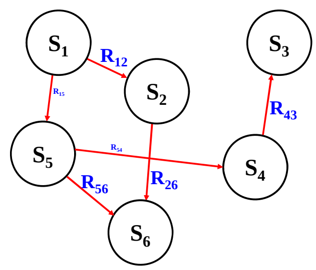


States

Actions

Rewards

There is a whole field – reinforcement learning – devoted to answering this question.

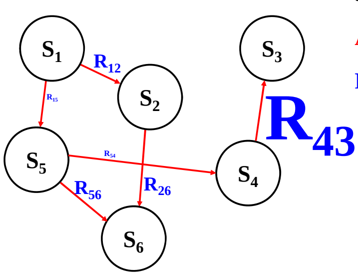


States

Actions

Rewards

There is a whole field – reinforcement learning – devoted to answering this question.

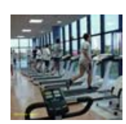


States

Actions

Rewards

State:
- hungry
- haven't had a date in 6 months

Actions:  

high immediate reward
low future reward, on average

low immediate reward
high future reward, on average

State:
- not hungry
- haven't had a date in 6 months

Actions:



~~high~~ low immediate reward
low future reward, on average

low immediate reward
high future reward, on average

- Lots of states
- Lots of possible actions
- **Future** rewards have to be estimated
- Uncertain rewards

This is a really hard problem!

But the stakes are high:

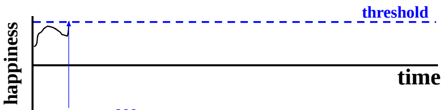
If you don't choose the right actions, you don't reproduce.

And as a species you die out.

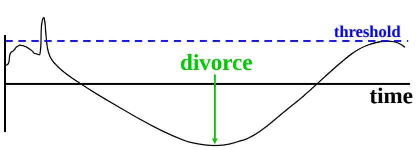
We're going to talk about this in a simplified setting:

- There are two possible actions.
- We have to decide **when** to take the action.

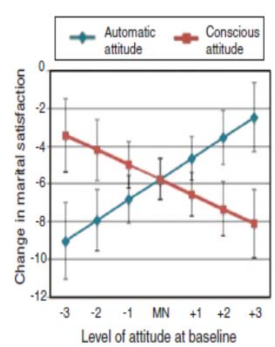
Do you propose marriage now and risk a lifetime of misery and suffering, or wait a little longer and risk [fill in the blank] running off with your best friend?



Do you propose marriage now and risk a lifetime of misery and suffering, or wait a little longer and risk [fill in the blank] running off with your best friend?



The problem: we can't **directly** use out current state to make decisions; we have to take into account future happiness.

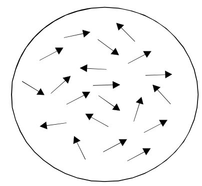


Science 29 November 2013: 1119-1120

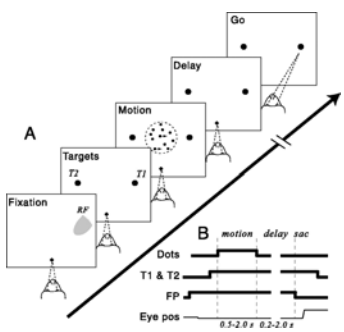
To get a handle on decision-making, we study a much simpler problem: the famous

random dot kinematogram

Task: look at the dots and guess which way they're moving.



Experimental setup



Shadlen and Newsome (2001)

Experimental setup

dots on   make saccade as soon as possible
maximize reward/unit time

time

Huk & Shadlen, *J Neurosci* 25:10420-10436 (2005)

This has many of the features of more realistic tasks:

1. Subjects have to integrate noisy evidence over time.
   - the longer you wait, the more information you have about direction.

2. Subjects have to decide when to make a saccade.
   - make a saccade too soon and you risk being wrong
   - wait too long and you throw away valuable reward

Questions:

1. Where in the brain is information about the direction of motion stored?

2. What algorithms do subjects use to make a decision?

3. Are subjects optimal?

4. How is the task learned?

5. What does this task tell us about decision-making in general?

Questions:

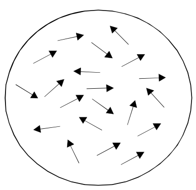
1. Where in the brain is information about the direction of motion stored?
2. What algorithms do subjects use to make a decision?
3. Are subjects optimal?
4. How is the task learned?
5. What does this task tell us about decision-making in general?

Things you have to know

Coherence: fraction of dots moving in the same direction

high coherence: task is easy
low coherence: task is hard

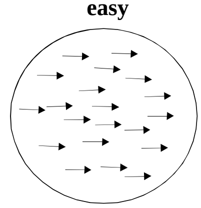
hard

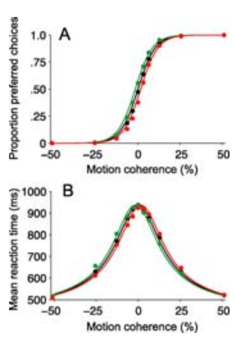


Coherence: fraction of dots moving in the same direction

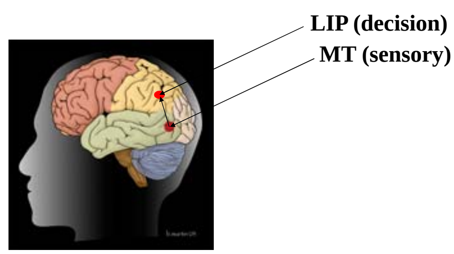
high coherence: task is easy
low coherence: task is hard

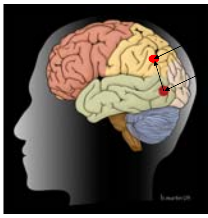
easy





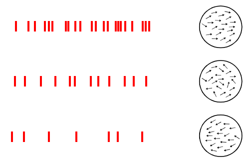
Huk & Shadlen (2005)



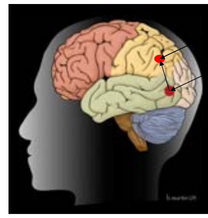


LIP (decision)
MT (sensory)

MT neuron:

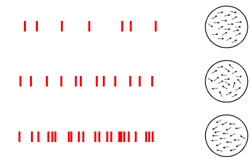


right preferring neuron

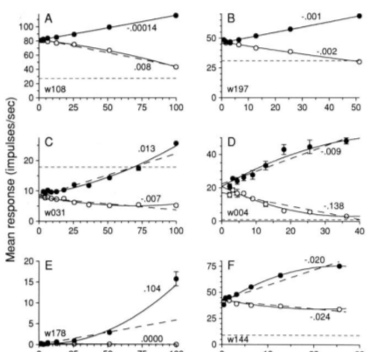


LIP (decision)
MT (sensory)

MT neuron:

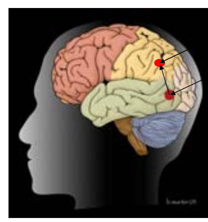


left preferring neuron



coherence (%)

Britten, Shadlen, Newsome, Movshon (1993)

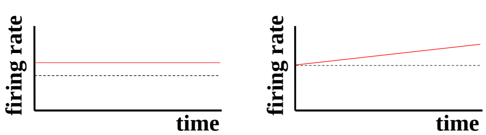


LIP (decision)
MT (sensory)

LIP accumulates MT spikes

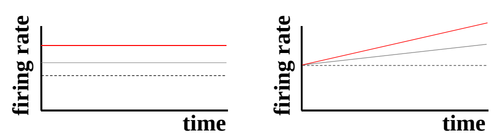
MT firing rate

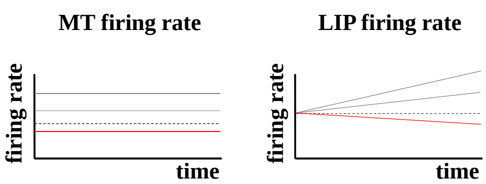
LIP firing rate



MT firing rate

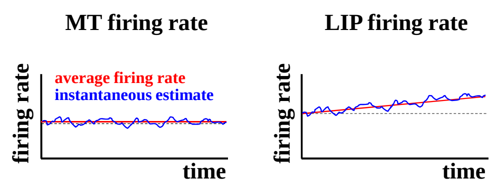
LIP firing rate



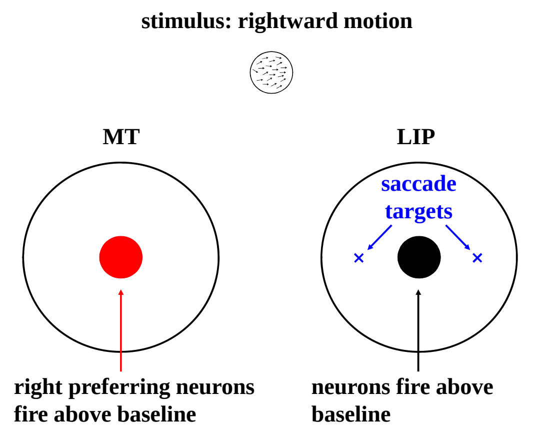


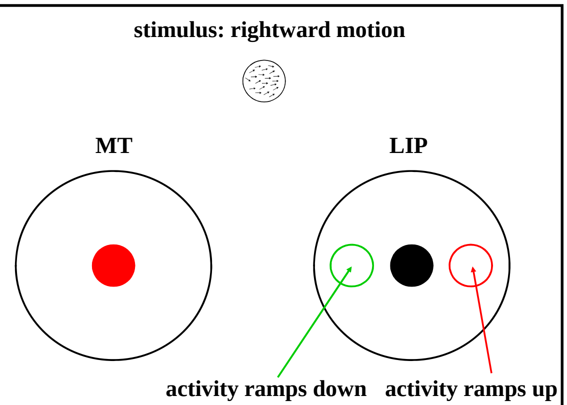
slope is proportional
to MT firing rate
relative to baseline

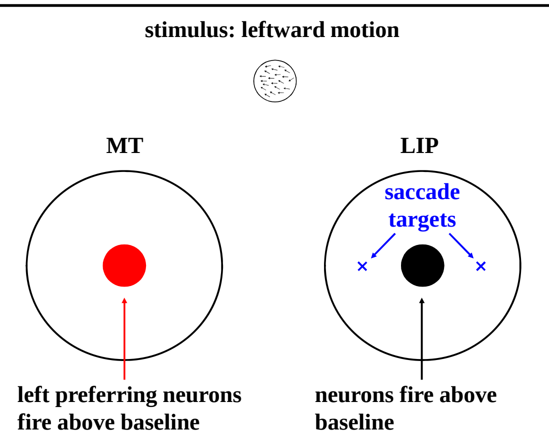
LIP neurons count MT
spikes relative to baseline



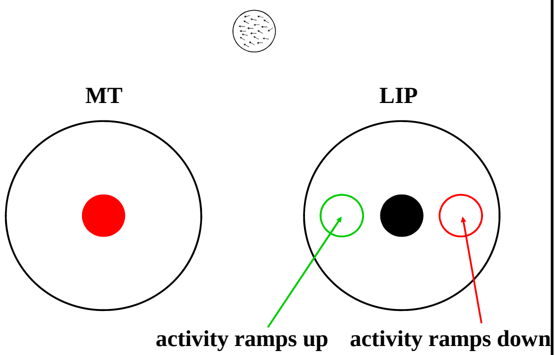
What's really going on in MT and LIP

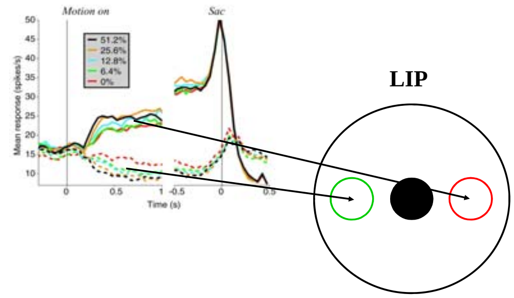




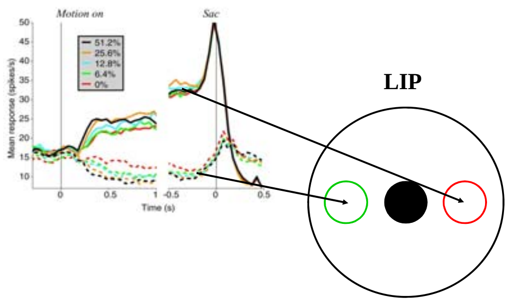


stimulus: leftward motion



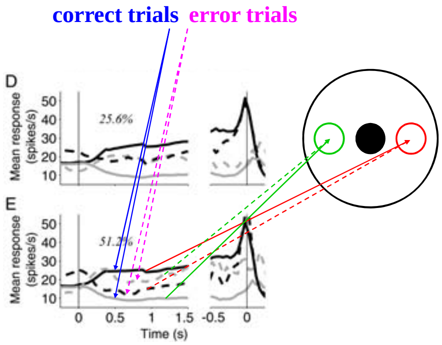


Shadlen and Newsome (2001)



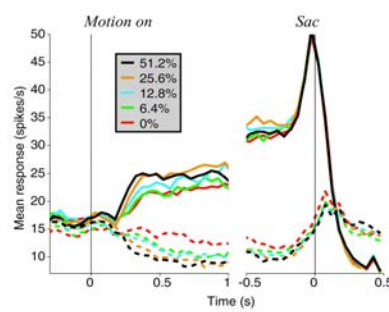
Shadlen and Newsome (2001)

We see a signal in LIP, but what is the evidence that LIP is really involved in decision making?



Shadlen and Newsome (2001)

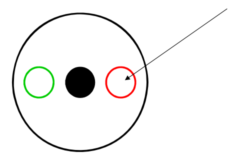
Saccade is made at a threshold

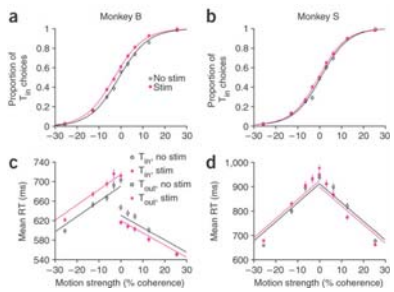


Shadlen and Newsome (2001)

"**Microstimulation** of macaque area LIP affects decision-making in a motion discrimination task"

Hanks, Ditterich & Shadlen
Nature Neurosci. 9:682-689 (2006)



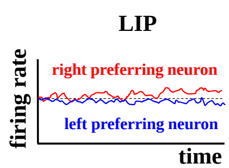


The evidence that LIP is involved in decision making.

1. error trials
2. threshold
3. microstimulation

The story so far:

1. LIP is heavily involved in a simple decision-making task.
2. It appears to accumulate evidence, and when the evidence is strong enough, a decision is made. **It is not known whether the decision is made in LIP, or somewhere else!**
3. But to make a decision, we need more than just evidence, we also need value.



Based on these rates, you calculate that the dots are moving to the right with probability 0.9.

Which do you choose, right or left?

Reward for choosing right correctly: £1
Reward for choosing left correctly: £1,000,000

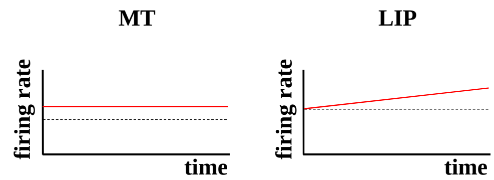
Expected reward for choosing right: 90p
Expected reward for choosing left: £100,000

Does LIP represent probability or value?

Mike Shadlen: probability
Paul Glimcher: value

The truth: neither!

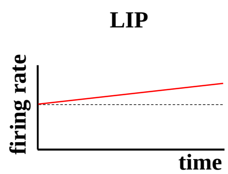
Let’s take a look at evidence for the two possibilities, starting with probability.



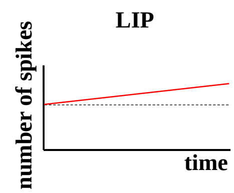
slope is proportional to MT firing rate relative to baseline

LIP neurons count MT spikes relative to baseline

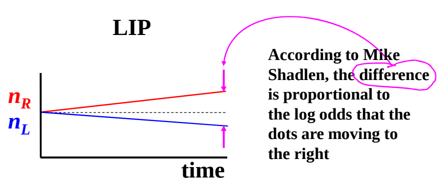
The instantaneous firing rate in LIP is proportional to the number of spikes emitted by MT above baseline.



The instantaneous firing rate in LIP is proportional to the number of spikes emitted by MT above baseline.



The instantaneous firing rate in LIP is proportional to the number of spikes emitted by MT above baseline.



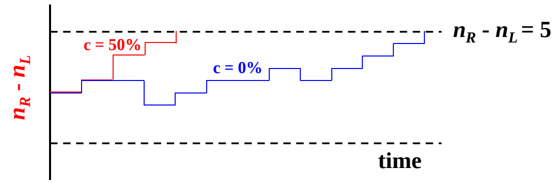
$$\log\left[\frac{P(\text{right})}{P(\text{left})}\right] = \text{constant} \times (n_R - n_L)$$

This is incredibly convenient, because LIP neurons give you an instantaneous readout of the probability that the dots are moving to the right.

This is exactly what you need to make a decision!

$$\log\left[\frac{P(\text{right})}{P(\text{left})}\right] = \text{constant} \times (n_R - n_L)$$

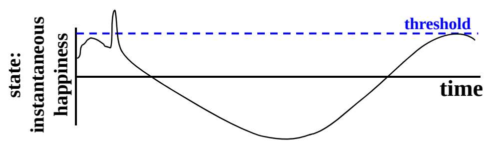
But there's a serious problem: the "constant" isn't constant.



$$\log\left[\frac{P(\text{right})}{P(\text{left})}\right] = \text{high} \qquad \log\left[\frac{P(\text{right})}{P(\text{left})}\right] = \text{zero}$$

This is a classic hard problem faced by the brain:
- getting rid of "nuisance parameters"
- disentangling causes

The marriage problem: what's causing happiness



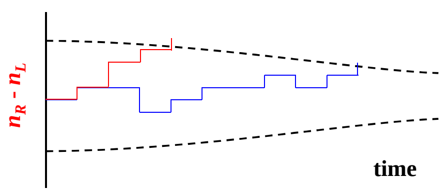
Summary so far:

1. **The job faced by the brain is to look at LIP spikes and decide when to make a decision.**

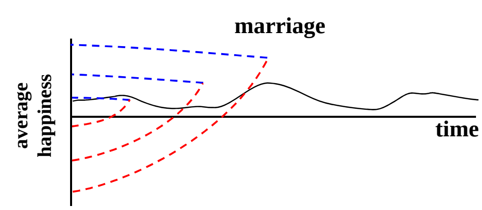
2. **LIP reports $n_R - n_L$, which is a signal that tells us something about the probability that the dots are moving to the right. However, we don't know coherence, so $n_R - n_L$ doesn't tell us directly about the probability.**

3. **LIP is not coding directly for probability.**

We can, though, get information about coherence just by looking at the dots. This, plus a lot of math, leads us to a collapsing bound as the optimal strategy.

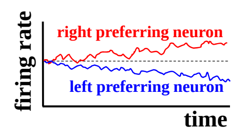


A collapsing bound is probably reasonably optimal for a broad class of problems.



However, probability isn't everything – you also need to know about value.

LIP



Based on these rates, you calculate that the dots are moving to the right with probability 0.9.

Which do you choose, right or left?

Reward for choosing right correctly: £1
Reward for choosing left correctly: £1,000,000

Expected reward for choosing right: 90p
Expected reward for choosing left: £100,000

According to Paul Glimcher* LIP codes for value.

The experiment:

targets appears

central fixation appears

*Platt & Glimcher *Nature* 400:233-238 (1999)

According to Paul Glimcher* LIP codes for value.

The experiment:

targets appears

central fixation changes color

*Platt & Glimcher *Nature* 400:233-238 (1999)

According to Paul Glimcher* LIP codes for value.

The experiment:

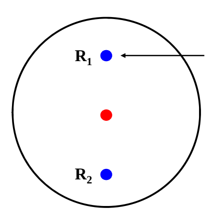
targets appears

central fixation changes color

*Platt & Glimcher *Nature* 400:233-238 (1999)

According to Paul Glimcher* LIP codes for value.

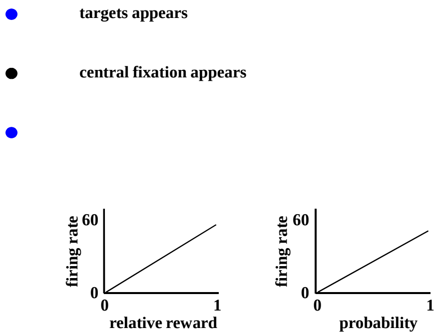
The experiment:

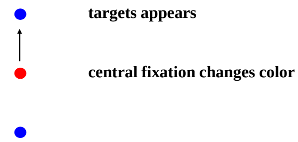


- recorded from neurons in LIP whose receptive field contained the target.
- the amount of the reward or the reward probability is varied in blocks.
- firing rate versus relative reward or probability is measured.

relative reward = $\frac{R_1}{R_1 + R_2}$

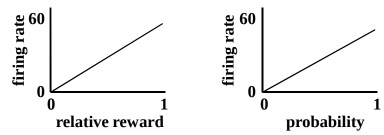
*Platt & Glimcher *Nature* 400:233-238 (1999)



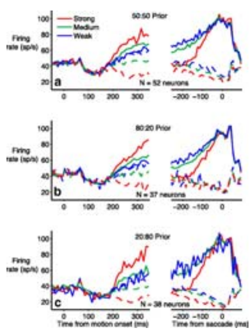


no integration of evidence!

that makes it hard to compare with the random dots experiment, which was all about acummalating evidence.

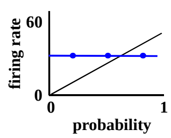


The experiment has been done with random dots and unequal prior probability of a reward.



Hanks T D et al. J. Neurosci. 2011;31:6339-6352

Compare to Platt and Glimcher:



LIP is not coding for value

Summary

1. Making decisions is hard: it requires estimating future rewards in an uncertain world with many states and many possible actions.

2. By looking at LIP in a very simple task, we tapped into the neural basis of decision-making.

3. Although the task was simple, it contained features of real world problems:
   - information is noisy
   - there is a nuisance parameter (coherence)
   - the brain has to accumulate, and represent, information over time.

Summary

4. We saw that LIP accumulates evidence over time, and when it reaches a threshold (in terms of firing rate), a decision is made.

5. Where the decision is made is an open question – it might be in LIP, or it might be an area monitoring LIP.

6. LIP doesn't encoded either value or probability. It may just be representing something about the desirability of future actions.

Summary

7. A key element (which I didn't talk about) is learning, for which there is a whole field (reinforcement learning). For simple tasks it does a good job explaining behavior; for more complicated tasks it does not do as well.

8. Finally, in decision-making, as with many areas of neuroscience, there are more questions than answers.