

Bogacz -

Predictive coding

Biological learning beyond BP

UKRI - Brain Network Dynamics Unit

Generative model of the world

- Relying on local computations
- and local plasticity

Developments

- see slide of papers, list
- Whittington & Bogacz 2017 "Approx BP"
- Prospective Config 2022
 - ↳ Alt credit assignment

Prospective Config

~ Predictive Coding (PC)

Reduced Interference (RI)

So you can cause collateral damage in a fully connected¹¹ model
i.e. destroy a common path used by another

PC: looks at total config, tries to avoid error from RI

Even str connection on other connec

Energy machine

This sounds similar to something I've seen before...

Neuron activity and weights are not directly "attached"

Nodes let the activity converge (Relaxation?) and then
adjust weights to minimize energy.

i.e. "Energy in springs"

Neural Implementation

Salmon Problem

~ spring distance is analogous to prediction error

sim

— not sure how big a difference there is in how
it BPs, the err, like does it go all at once?

Q: How big are the steps

Target Alignment

Note:

it kinda

avg^s over all err?

or because the
tensor is through
multiple layers

w/ larger nets

Q: why did you use linear activation func?

a: simplicity

fashion MNIST

Train most batches

Limited Training set

- Last few slides have just seen performance graphs

Concept Drifting

Continual & Reinforcement learning

Q: Why is the Variance higher in the perf of the PC
(I_t is high)

A: Doesn't know

Q: missed this one

Q: Dropout? Network Modification

Q: Key problem? what did he say?

Something about speed?

A: "It is slower." wait how is it local, he has to go through multiple layers. to be local it has to be able to update all neurons @ once.

Q: Adversarial attacks?

A: "can only speculate"

Q: Bio plausibility

A: "Natural splicing interpretation"

Boerlin et al. 2013 # Reading # P2

Inference of latent state

Inf during motor learning

BP Algo

Or this might explain the diff

Energy machine

\hat{x} : predictive activity

Node Dynamics

\bar{i} : current layer

\hat{j} : previous layer

$$\dot{x}_i^{(\ell)} = -\frac{1}{\tau} \left[x_i^{(\ell)} - \sum_{\hat{j}} v_{i\hat{j}}^{(\ell-1)} f(x_{\hat{j}}^{(\ell-1)}) \right]^2$$

$$= -\frac{1}{\tau} \left(x_i^{(\ell)} - \sum_{\hat{j}} v_{i\hat{j}}^{(\ell-1)} f(x_{\hat{j}}^{(\ell-1)}) \right)$$

$$+ \sum_{\hat{j}} x_i^{(\ell+1)} - \sum_{\hat{j}} v_{i\hat{j}}^{(\ell)} f(x_{\hat{j}}^{(\ell)}) \quad v_{i\hat{j}}^2 f'(x_{\hat{j}}^{(\ell)})$$

$$\dot{x}_i^{(\ell)} = -\epsilon_i^{(\ell)} + f'(x_i^{(\ell)}) \sum_{\hat{j}} v_{i\hat{j}}^{(\ell)} \epsilon_{\hat{j}}^{(\ell+1)}$$

Predictive Coding Network

see list of reading

Local Plasticity

- Basically uses hebbian

Err in output layer

Feature not bug

Song et al. 2022 # reading

~ Equilibrium prop # topic "Inferring Neural activity..."

avoiding neural activity shift

Plasticity - stability

others, see lit.

can be used together

Spiking interpretation

Berlin, Machens, Denner (2013)

- cont. variable corresponding to spikes (x) is count

$s(t)$ spike

$$\frac{dx}{dt} = -\lambda x + s(t)$$



- can't decide x , only on/off spike
- how do we det spike

add energy

$$s_i(t) = 1 \text{ iff it reduces } E$$

"leads to integrated
fire eq"

- some how we can get STDP
- might be another paper

Q: missed this one

Q: Bayesian & generative models (?)

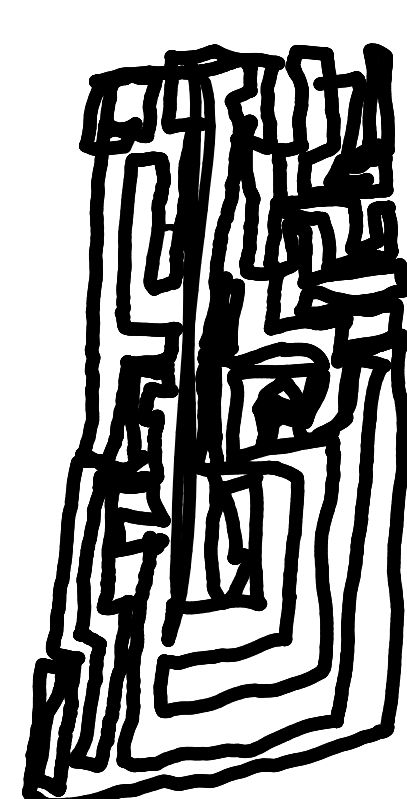
Q: Does it diverge

A: no, stable, quadratic eq requires min

Q: Reward structure

A: RL &

Dopaminergic neurons in BG



Q

