DAY 2 - NIPS 2018 - MAIN CONFERENCE I 04/12/18

Talk 1 — Public Policy <=> ML - E. Felten (Princeton)

- Have to work with decision makers => (Politics is not a search for truth)
  - -> all Qs are decidable => regardless of facts, time, etc.
- Majority voting => indiv. logically consistent <=> aggregate not necessarily!
  - ↳ result: legislators are not logically consistent ⊕ does not care about facts
  - ↳ debate about facts happens on lower level, legislators simply can't
- World outside simplistic model => lives are complicated (seemingly inconsistent
- Consulting -> decision makers = Generalists -> context switching constantly
  - -> what info does policy maker want?
    - ↳ Not just too many facts
    - ↳ Not just dictate decision -> 'who elected you?'

    BE USEFUL!
    ↓
    knowledge (You) <-> Preferences (Him/HER)

  - -> Get their knowledge and preferences
    - => ASK QUESTIONS!
  - -> structure decision space

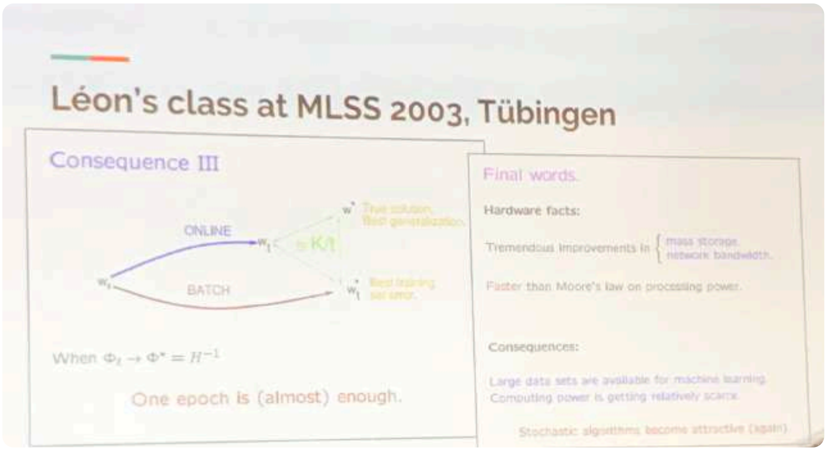
  Multi-round protocol! -> Ask for feedback! <- build trust

- Need to create viable career paths!
  - ↳ build incentives to communicate => respect for this activity from within community
    - ↳ right opportunities

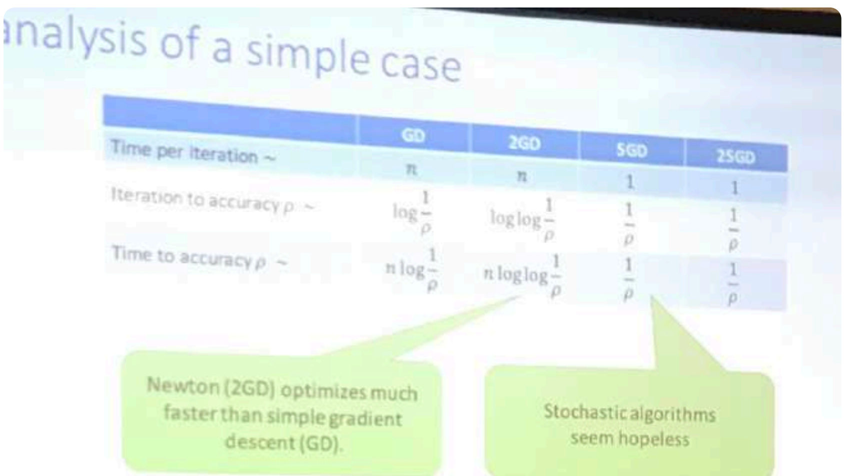
NOTES:
- Great talk => Entertaining and deep at same time
- Humor and not judging any side
- Recommendations could have been more specific
  - -> How to build incentives to either comm. or became decision maker? => Need for role models!

Test of Time Award – Bottou, Bousquet ⇒ Trade-offs of Large Scale Training

- Estimation vs. Approx. vs. Optimization Error ⇒ SGD
- 2007 → SVM / Kernel Machines ⟶ Which resources are limited?!
  (also cond. random fields)
  ↳ Advocated SGD not Quadratic Pr.
  ↳ Motivated by duplicate observations and nested computations



⇒ SPEND LESS TIME ON MORE DATA



→ Base example: Smooth + convex loss
↳ only optimize until excess error
⇒ SGD dominates!

- Afterwards: – Variance reduction in SGD
  – Implicit regularization by early stopping
- Takeaways: – Truth does not follow fashion ⇒ experiments → scientific, not validity tests
  – Use of mathematics ⇒ can't trust blindly!
- 3 problems: – Role of over-parametrization / generalization
  – Role of compositionality ⇒ Hierarchy of skills
  – Causality in AI

P. Track 1 – RL and Neuro ⇒ ① Baldi – Neuronal Capacity

- $C(\mathcal{F}) = \log_2 |\mathcal{F}|$ → McCulloch-Pitts Neurons
  ↳ linear threshold gates (LTG) ⇒ $|\mathcal{F}|$: finite

class of approx. fct. → number of boolean fcts. can be implemented
→ number of bits that can be stored in network

value of fcts. capable

- $0.5 n^2 \leq C(LTG(n)) \leq n^2$

$$= n^2(1+o(1))$$

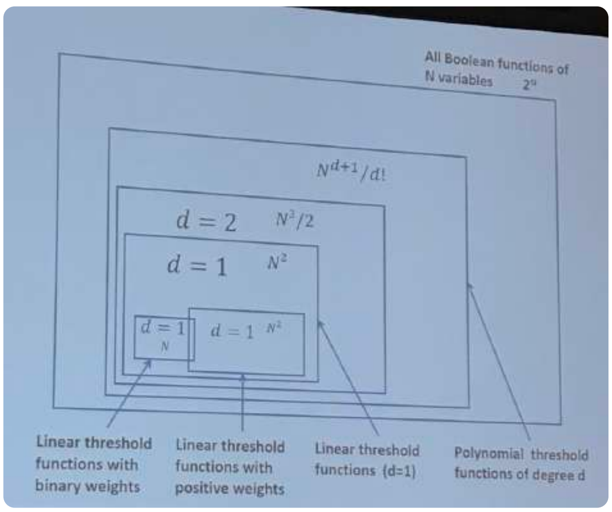
LINEAR

- $\binom{n}{d+1} \leq C(PTG(n,d)) \leq \frac{n^{d+1}}{d!}$

$$= \frac{n^{d+1}}{d!}(1+o(1))$$

POLYNOMIAL

→ proved in this paper!

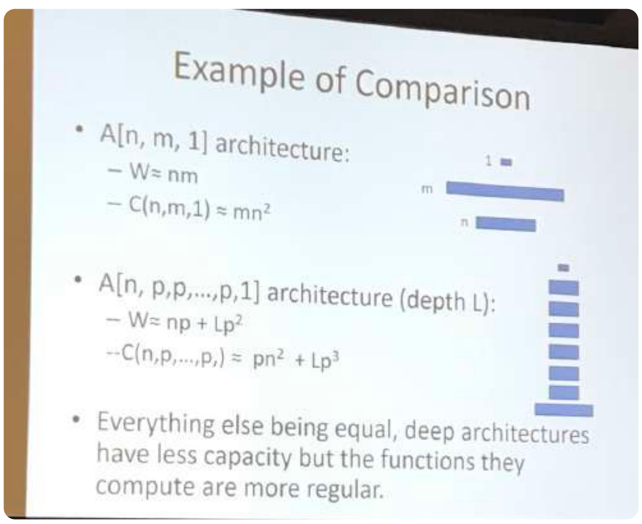


→ for single neurons ⇒ extension to RNNs or MLPs:

$$C(n_1, \ldots, n_L) \asymp \sum_{k=1}^{L-1} \min(n_1, n_2, \ldots, n_k)\, n_k n_{k+1}$$

number units layer 1

↳ use results to compare architectures!

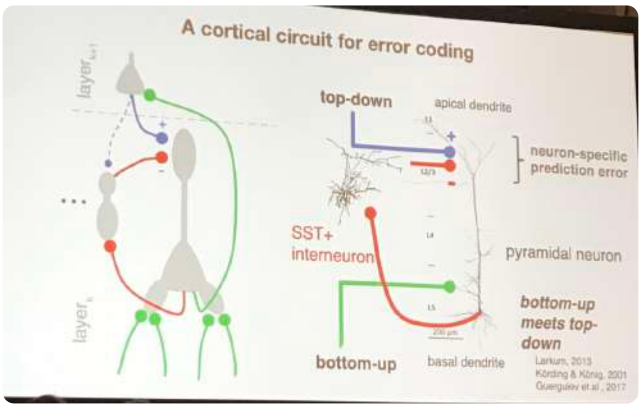


→ deeper architectures can't compute more functions ⇒ but functions have nicer properties ↳ ('structural regularization')

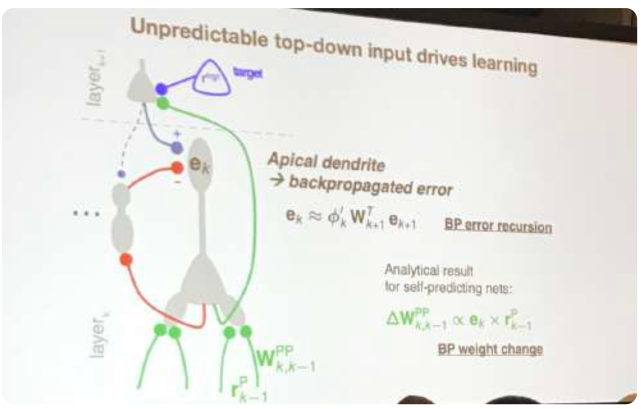
P. Track 2 – RL and Neuro ⇒ ② Dendritic cortical circuits ≈ Backprop

- Backprop: ① forward pass → ② error comp. → ③ backward pass
- Wish list:
  1. Cont.-time neuronal dynamics → not ↑↓
  2. Feedback connectivity
  3. No weight transport → W^T multiplication → Lillicrap!
  4. Local plausible plasticity rule

- Apical – top-down
  Basal – bottom-up } Compartmental → SST+ interneurons
  ↓
  lateral inhibitory prediction

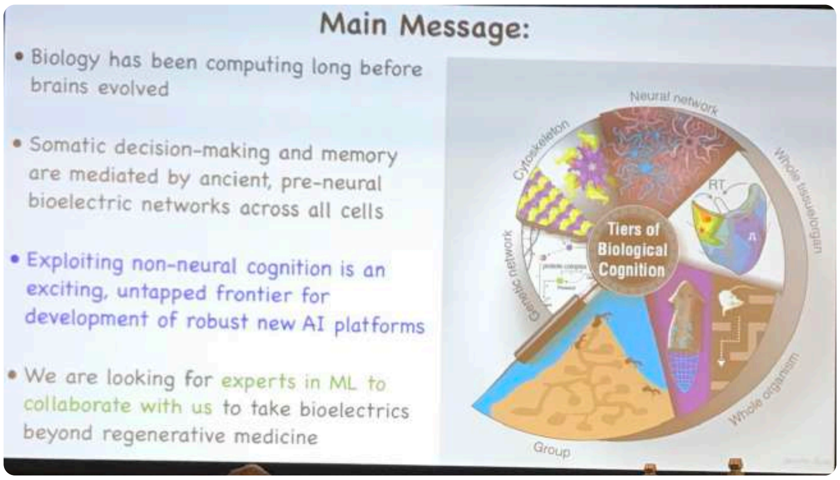


- Chiu et al (2018) → pyramidal PFC
- Form of distillation approach!
- Test on MNIST → approx. performance of backprop!



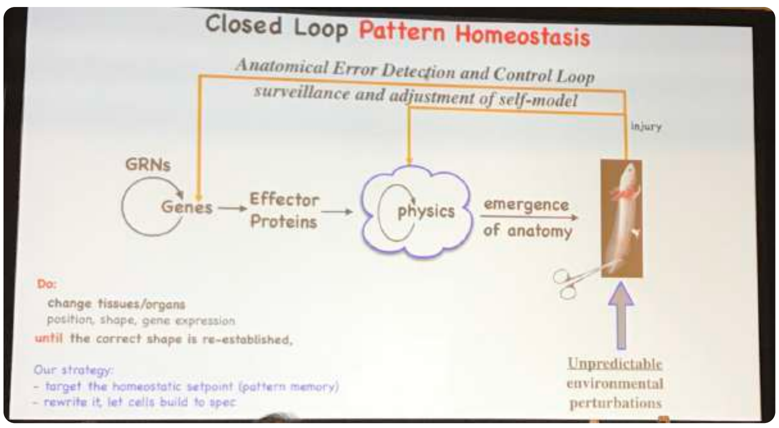
⇒ within-layer dendritic inhibition
↳ Apical dendrite activity
↳ backprop error!

TALK – M. Levin – Bioelectric Computation beyond the NS



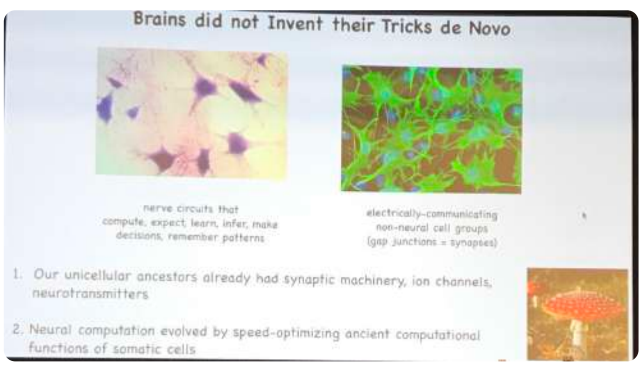
• Brain - body plasticity:
metamorphosis → caterpillar's memories are stable through hardware transition!
↳ behavioral programs ↕ body (re)generation

• Embryogenesis → development ⇒ self-assembly → guided
→ open loop system too simple ⇒ plasticity ⊕ regeneration
↳ knowing when to stop! ⇒ does not regrow more than needed
↳ COMPUTATION PROBLEM!



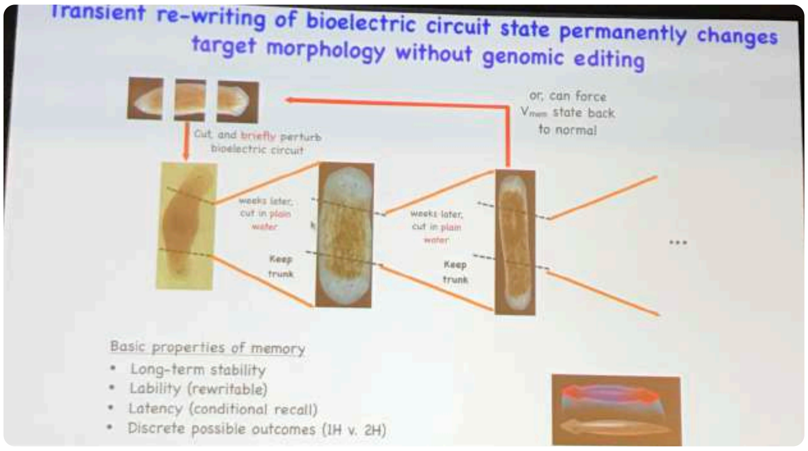
• biology → still at hardware level
↳ reverse engineering ⇔ control

• Pattern homeostasis ⇒ goal integration, robustness
→ like brain ⇒ somatic tissues form bioelectric networks making decisions



→ Membrane Voltage Patterns:
* Normal / endogenous
* Pathological ⇒ cancer cells
Abnormal electr. activity
↳ Manipulation ⇔ Reprogramming
⇒ true physiological change / no genome editing!

⇒ Software control beyond genome!



• Use bioelectric signals to trigger cascades / anatomical subroutines!

↳ "compiles"

• Future ⇒ Go beyond neural inspiration → all somatic cells!

Neural Ordinary Differential Eq.

• Adaptive solvers > Euler approx. ⟷ Resnet ⇒ Continuous Time Dynamics!

• Neural Networks ⟺ ODEs

$$L(\theta) = L\left(\int_{t_0}^{t_1} f(z(t), t, \theta)\, dt\right) \Rightarrow \text{How to diff.?}$$

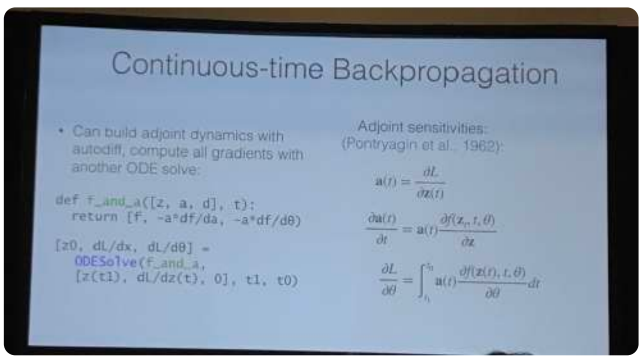
↳ smooth dynamic change with depth

↳ less parameters required

↳ No fixed number of layers!

⇒ Solver determines depth conf.!

↳ Can't control time complexity! at train time!



⇒ But at test time: Can tradeoff accuracy of approx. vs. speed.

• Can extend to SDEs and Poisson Point Process