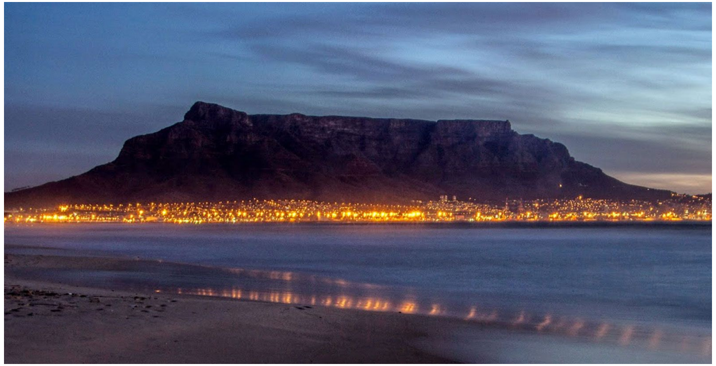
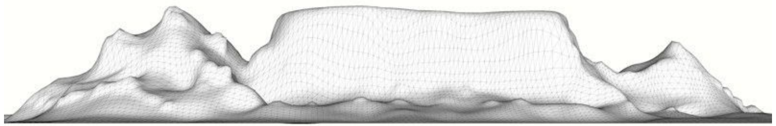
Deep Learning
The AI revolution and its frontiers

Nando de Freitas, Matt Hoffman, Ziyu Wang, Brendan Shillingford, Caglar Gulcehre, Misha Denil, Scott Reed, Serkan Cabi, Tobias Pfaff, Tom Paine, Yannis Assael, Yusuf Aytar, Yutian Chen,, Sergio Gomez, David Budden, Natasha Jaques

DeepMind

Hard coding has its limits





The machine learning way

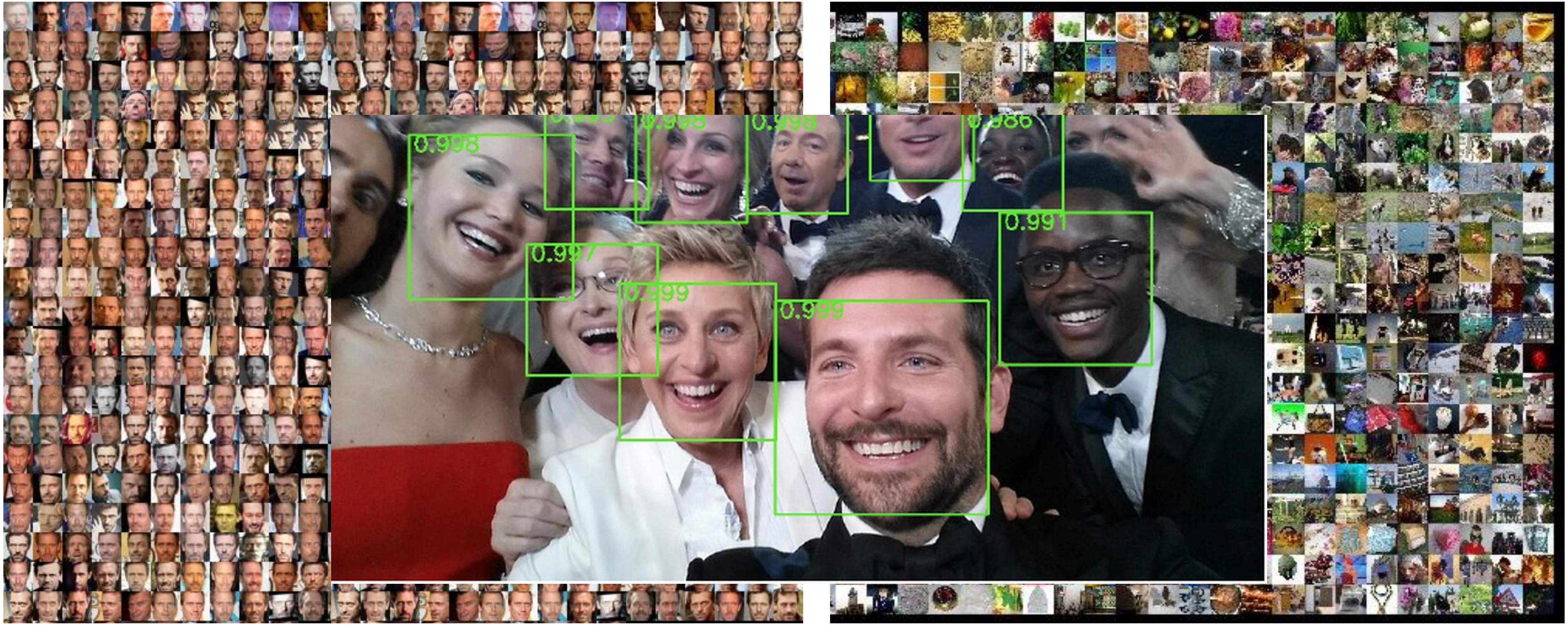
Faces

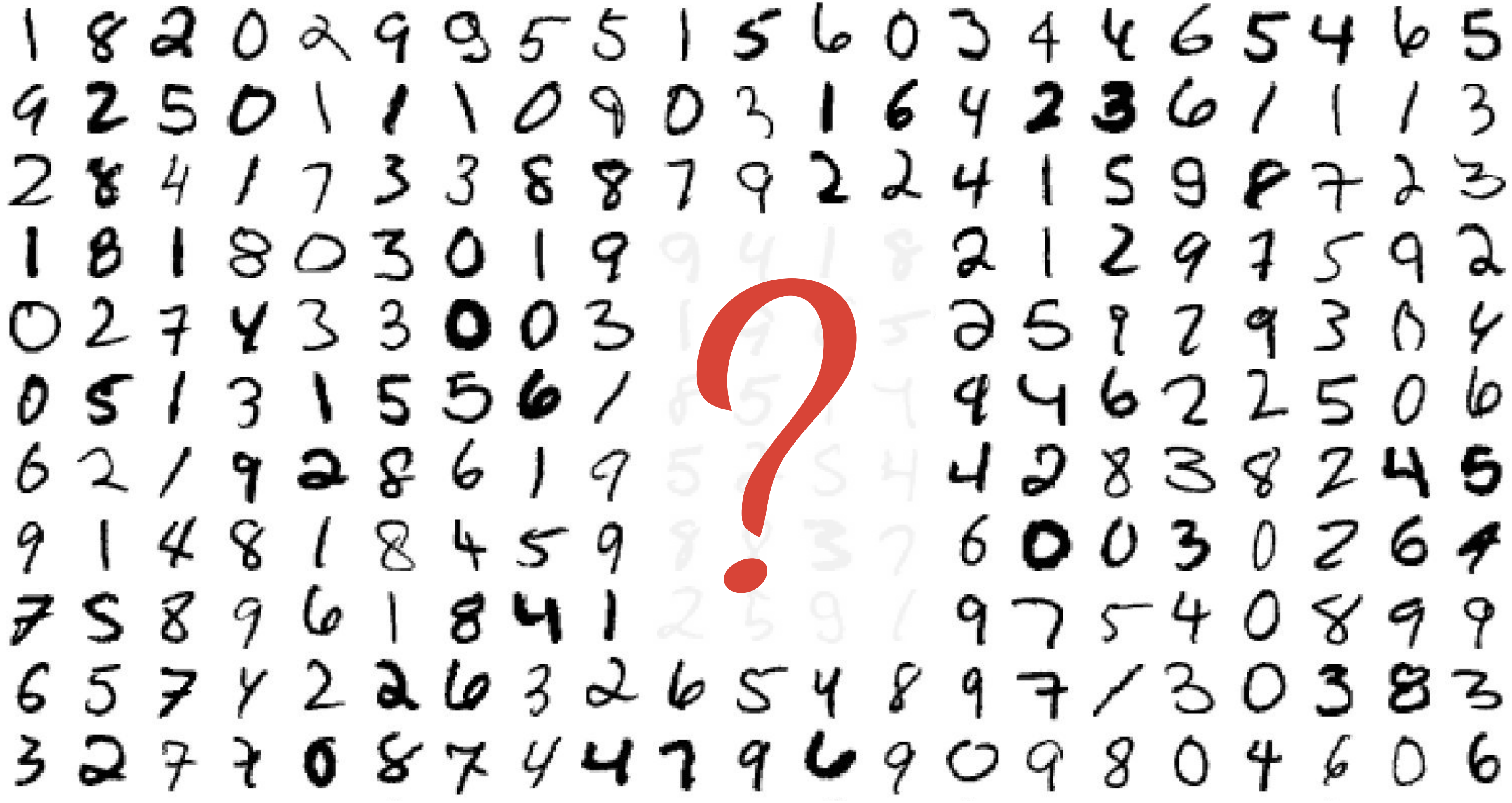
Not faces

The machine learning way

Faces

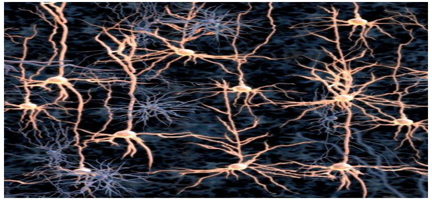
Not faces

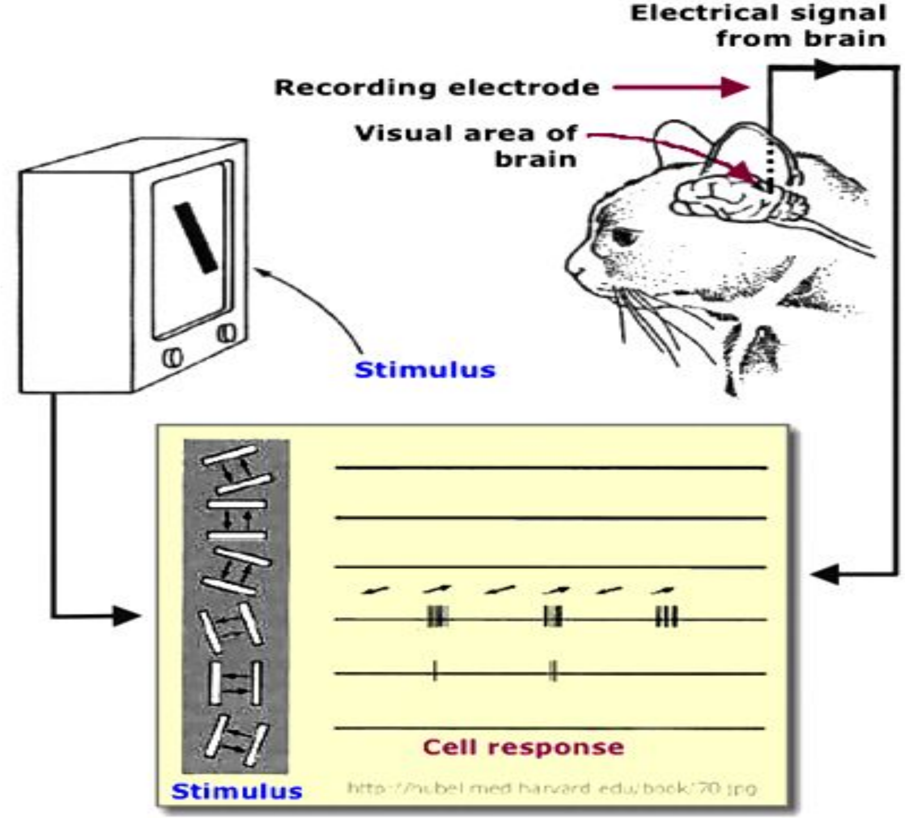


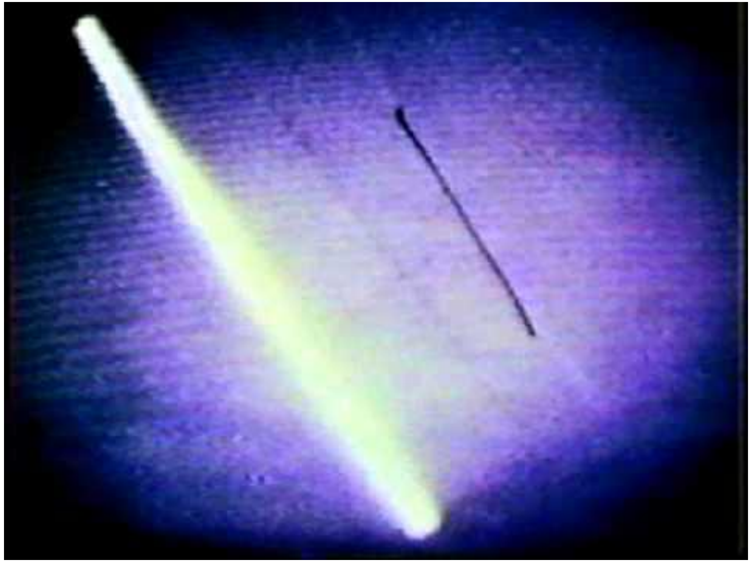


MNIST = Mixed National Institute of Standards and Technology - Download the dataset at http://yann.lecun.com/exdb/mnist/

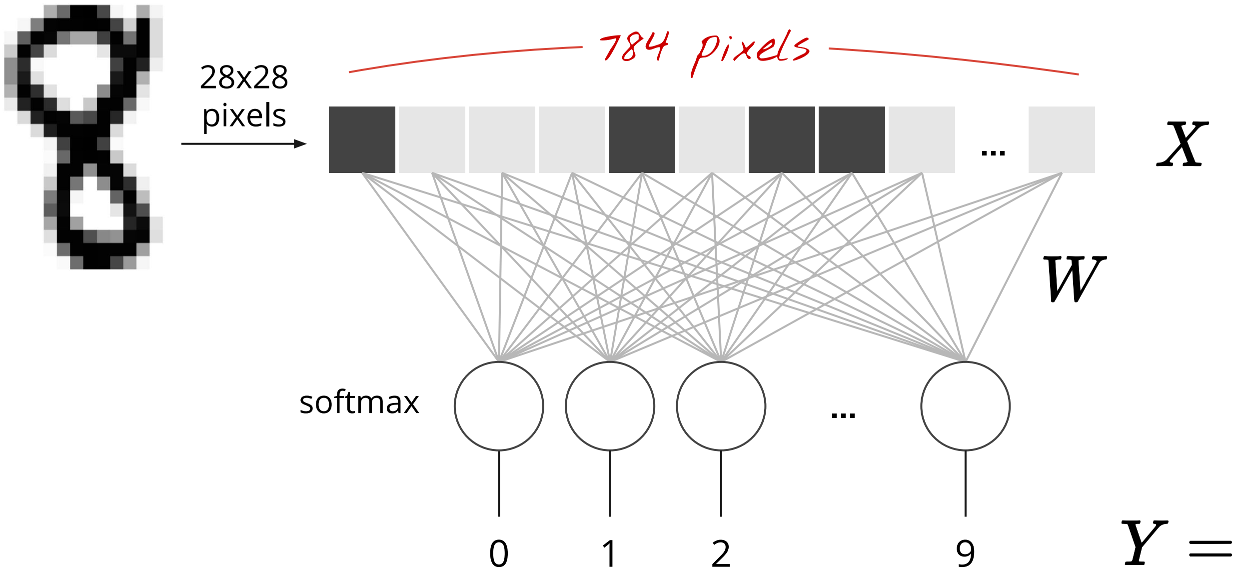
Hubel & Wiesel







A simple neural net for classifying images of digits



$$Y = softmax(\overbrace{X.W + b}^{z})$$

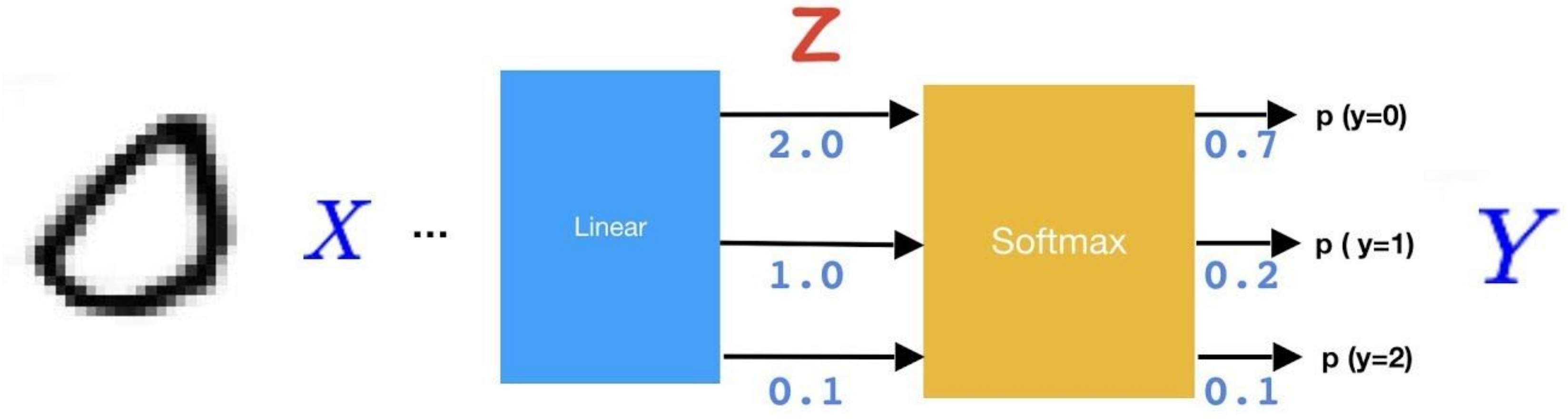
Computed probabilities Y

0.1	0.0	0.0	0.2	0.0	0.1	0.1	0.0	**0.4**	0.1
0	1	2	3	4	5	6	7	8	9

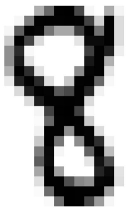
$$softmax(z_i) = \frac{exp(z_i)}{\sum_j exp(z_j)}$$

[Modified slide by from Martin Gorner]

The modular (layer-wise) view of neural nets



Supervised learning



0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	**1**	0

Actual probabilities, "one-hot" encoded

Cross entropy: $-\sum Y_i' . log(Y_i)$

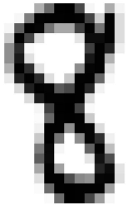
$$Y = softmax(X.W + b)$$

Predicted probabilities Y

0	1	2	3	4	5	6	7	8	9
0.1	0.0	0.0	0.2	0.0	0.1	0.1	0.0	**0.4**	0.1

[Modified slide by from Martin Gorner]

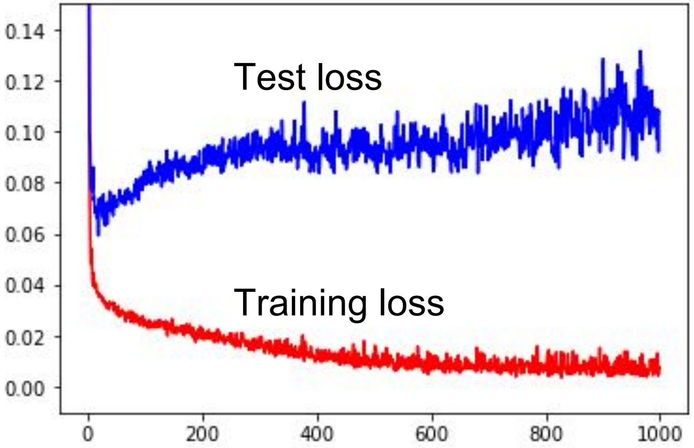
Supervised learning



0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	1	0

Actual probabilities, "one-hot" encoded

Cross entropy: $-\sum Y_i' . log(Y_i)$



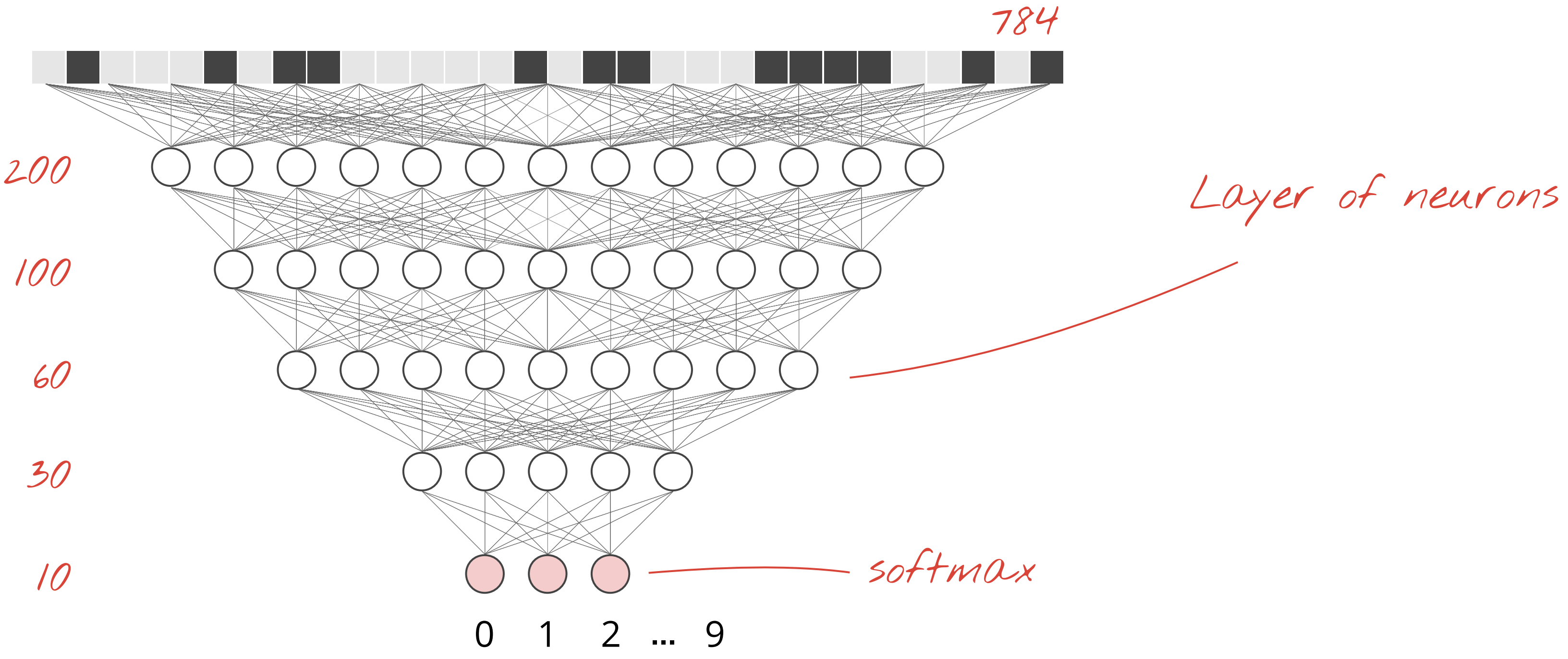
$$Y = softmax(X.W + b)$$

Predicted probabilities Y

0.1	0.0	0.0	0.2	0.0	0.1	0.1	0.0	0.4	0.1
0	1	2	3	4	5	6	7	8	9

[Modified slide by from Martin Gorner]

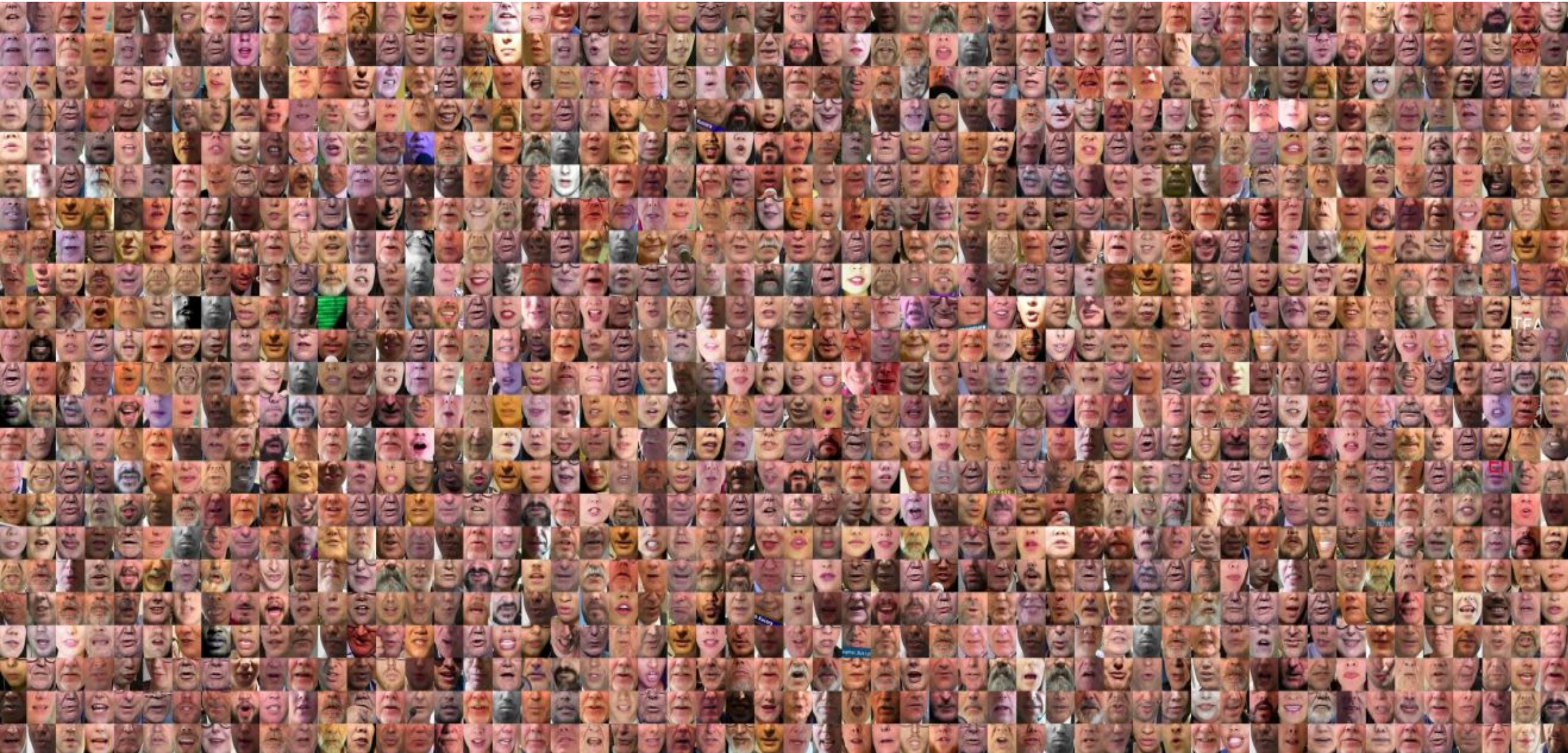
Want better results? Go bigger and train end-2-end !



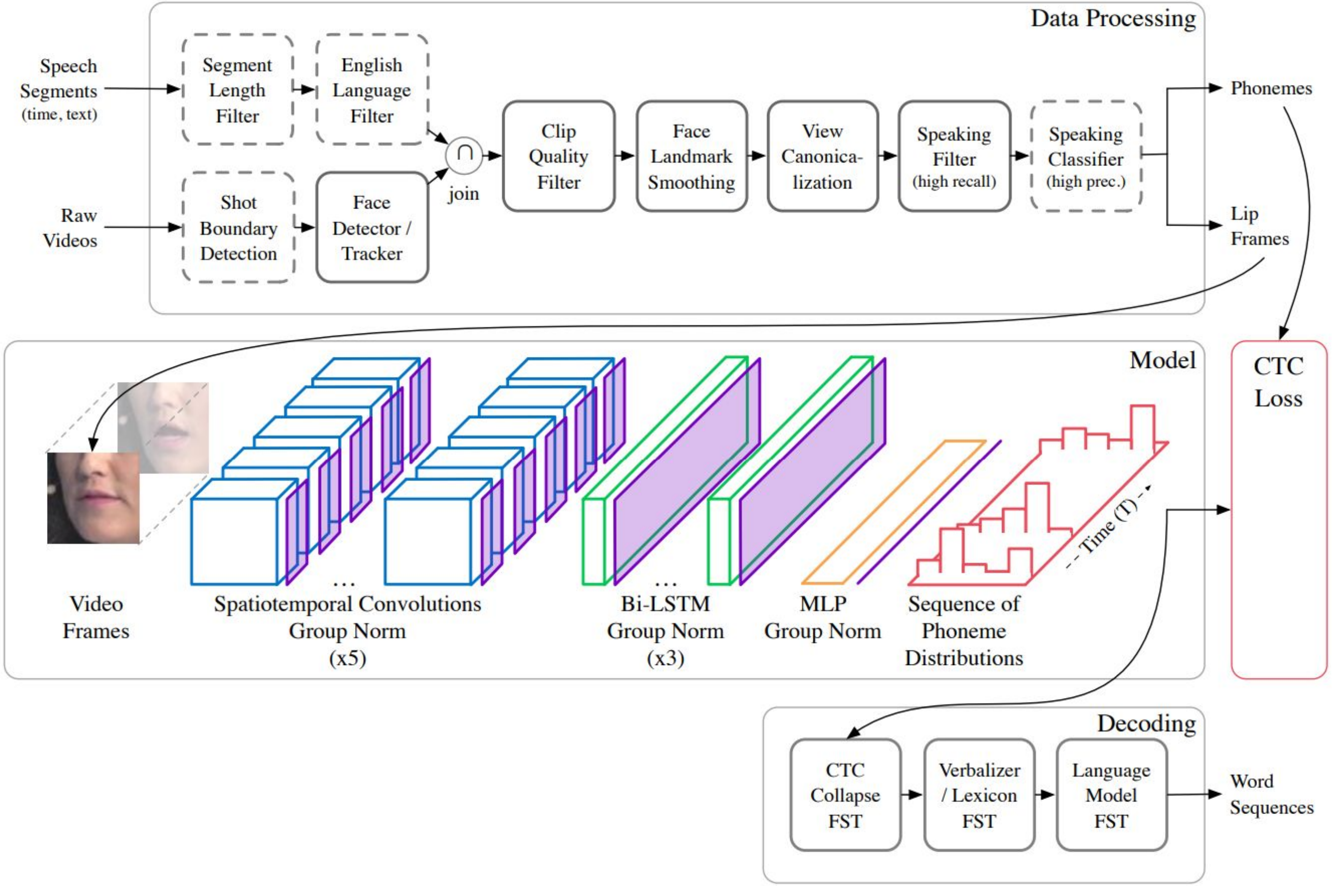
Case study: Visual speech recognition

- Improving speech recognition in noisy environments.
- Silent interfaces.
- Helping patients with aphonia.
- Resolving multi-talker simultaneous speech.

Step 1: Create a massive labelled dataset



Step 2: Train a large deep network



Step 3: Evaluate performance

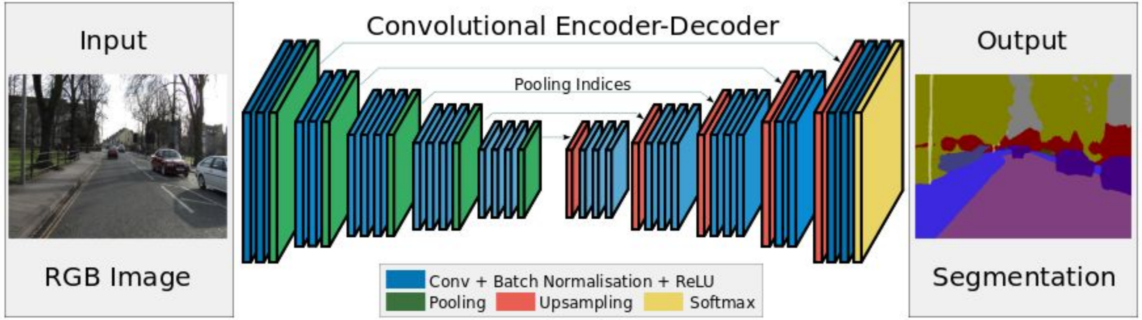
Ground truth:

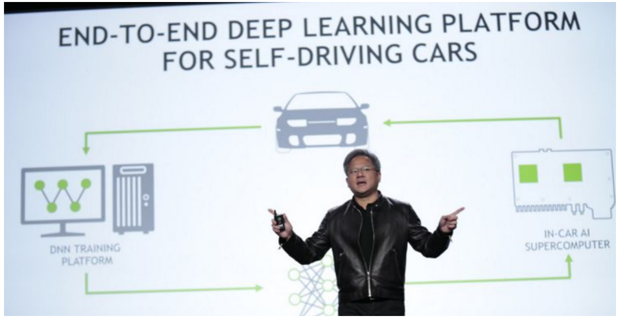
I am making this video for all my friends and family who may not know what I'm going through having done their research or even people that are actually going through what I'm going through

Prediction:

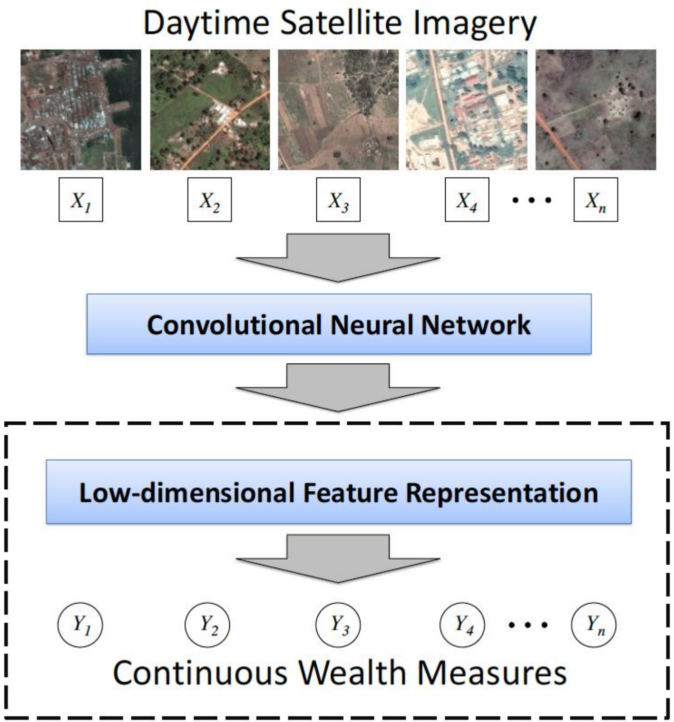
and making this video for all my friends and family you may not know what I'm going to do a little research or even people that actually going through what I'm going through

Word Error Rate: 22.86%

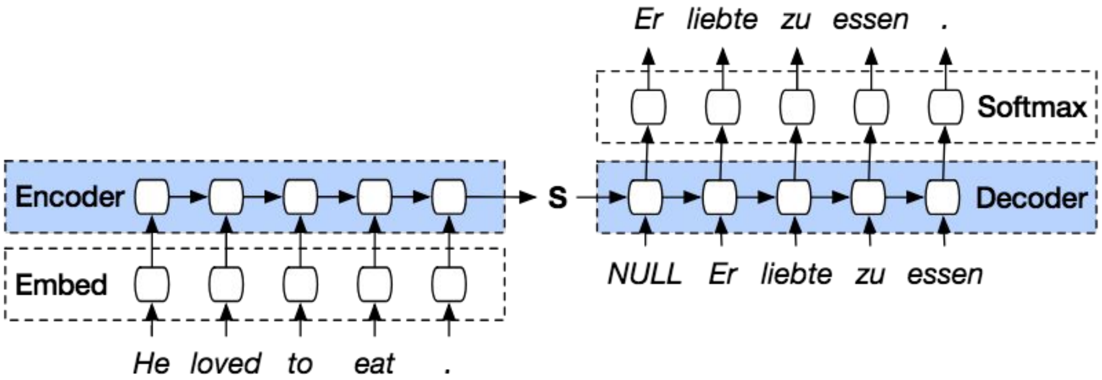




Thousands of applications



[Stanford]

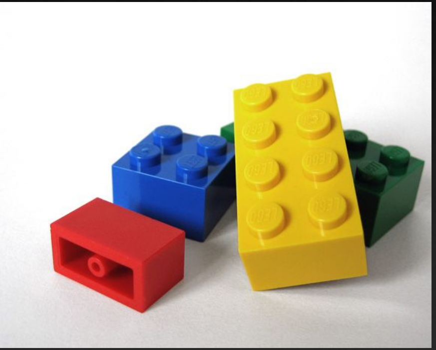


The key AI revolution ingredients

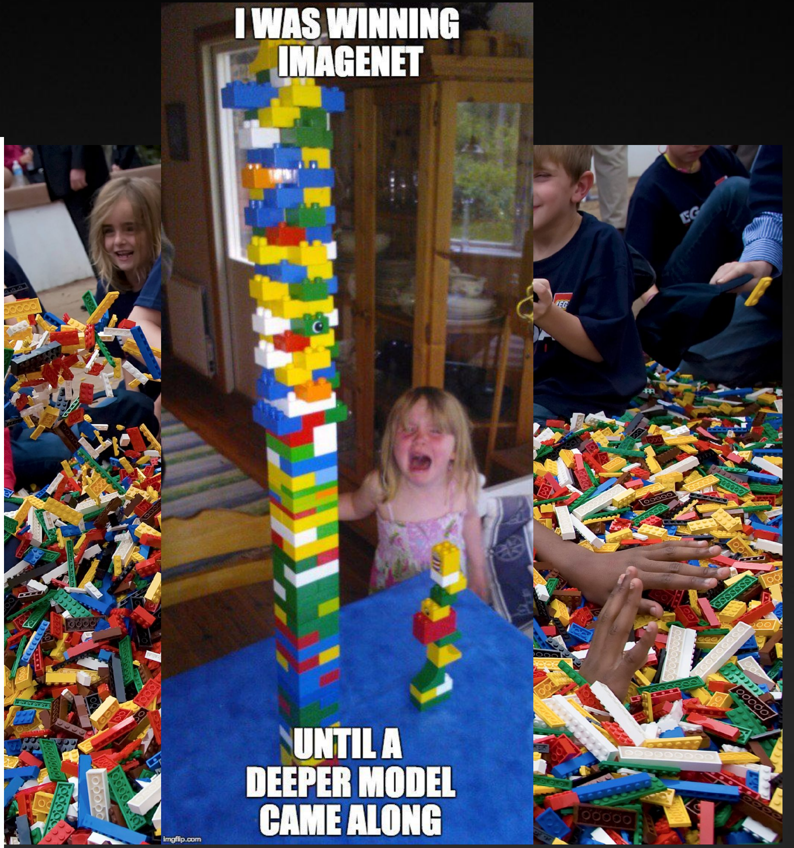
1. Science
2. Data
3. Computing
4. Software

Deep learning research is like playing with lego

Not like this

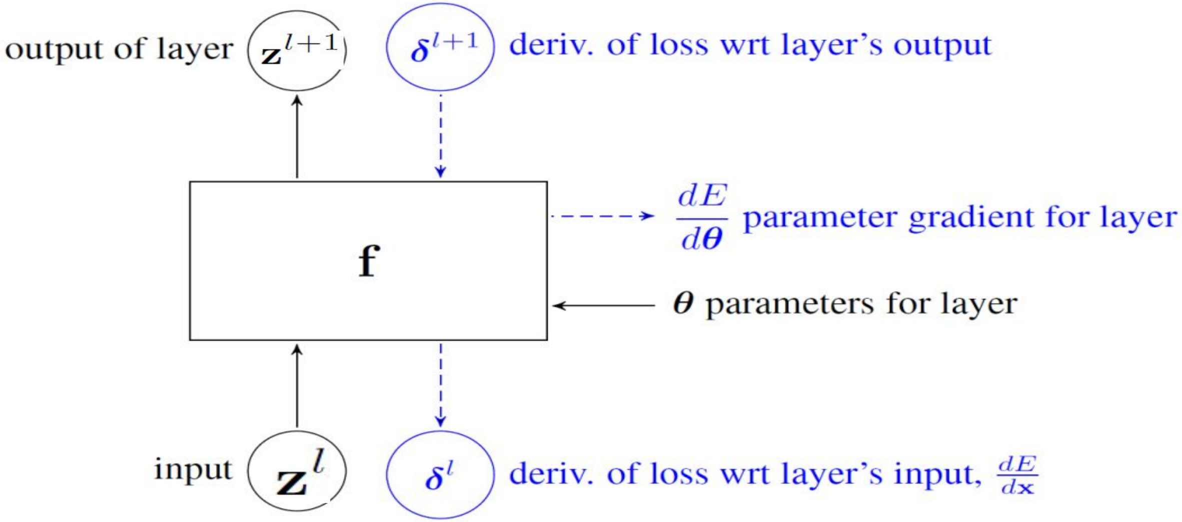


But rather like this



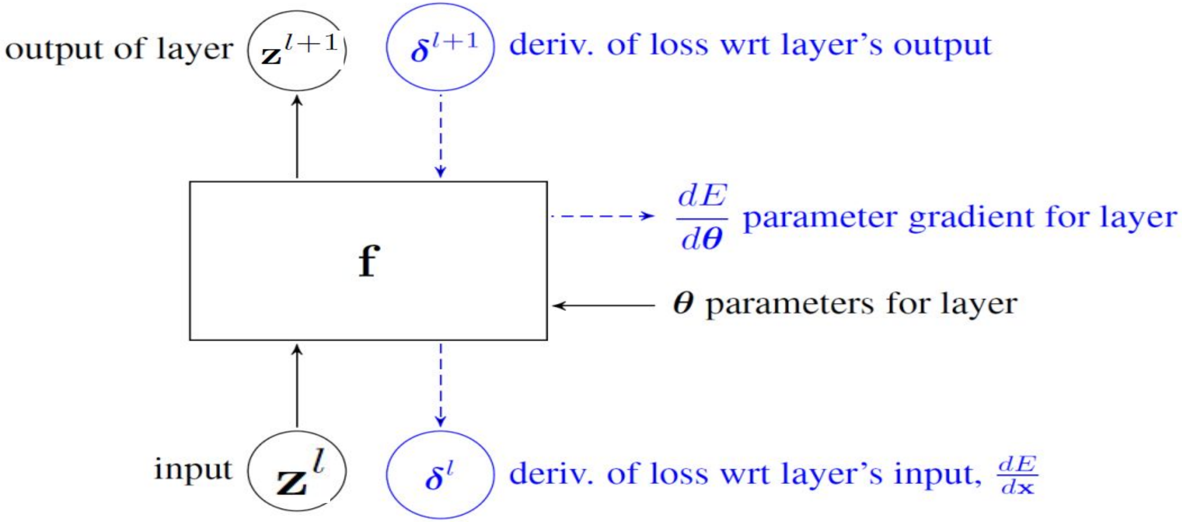
Combinatorial re-use is robust and amazing for creativity

Modularity backprop



$$\mathbf{z}^{l+1} = \mathbf{f}^{l}(\mathbf{z}^{l}; \boldsymbol{\theta}^{l})$$

Modularity backprop

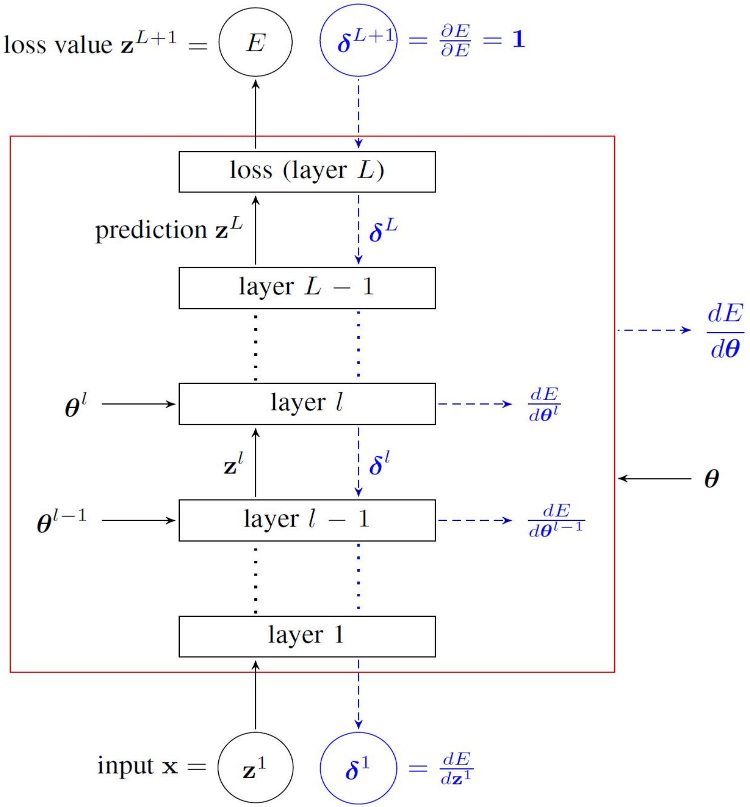


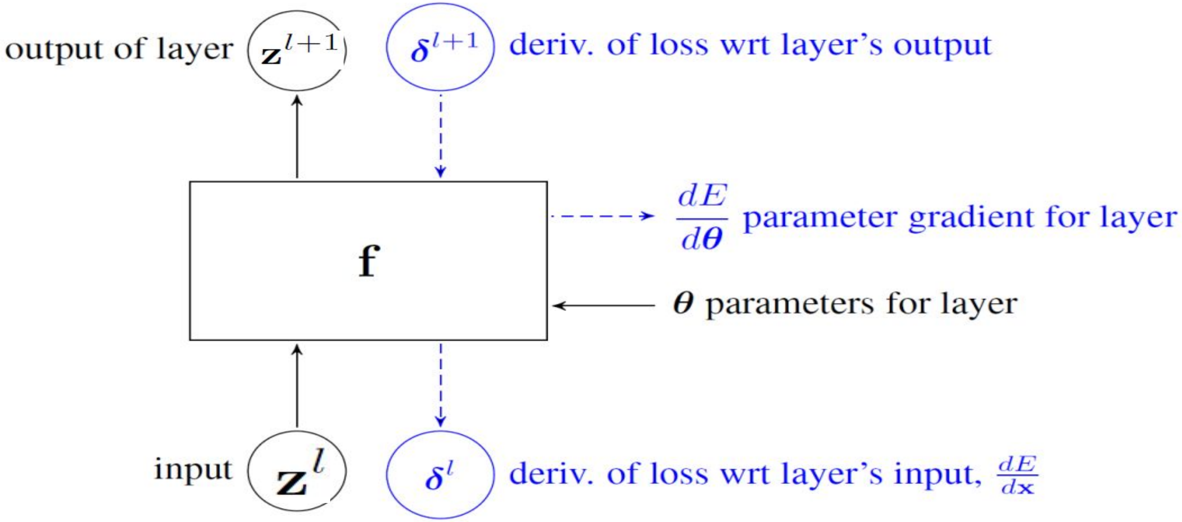
$$\mathbf{z}^{l+1} = \mathbf{f}^{l}(\mathbf{z}^{l}; \boldsymbol{\theta}^{l})$$

$$\boldsymbol{\delta}^{l} := \frac{\partial E}{\partial \mathbf{z}^{l}} = \frac{\partial E}{\partial \mathbf{z}^{l+1}} \frac{\partial \mathbf{z}^{l+1}}{\partial \mathbf{z}^{l}} = \boldsymbol{\delta}^{l+1} \frac{\partial \mathbf{f}^{l}(\mathbf{z}^{l}; \boldsymbol{\theta}^{l})}{\partial \mathbf{z}^{l}}$$

$$\frac{\partial E}{\partial \boldsymbol{\theta}^{l}} = \frac{\partial E}{\partial \mathbf{z}^{l+1}} \frac{\partial \mathbf{z}^{l+1}}{\partial \boldsymbol{\theta}^{l}} = \boldsymbol{\delta}^{l+1} \frac{\partial \mathbf{f}^{l}(\mathbf{z}^{l}; \boldsymbol{\theta}^{l})}{\partial \boldsymbol{\theta}^{l}}$$

Modularity backprop



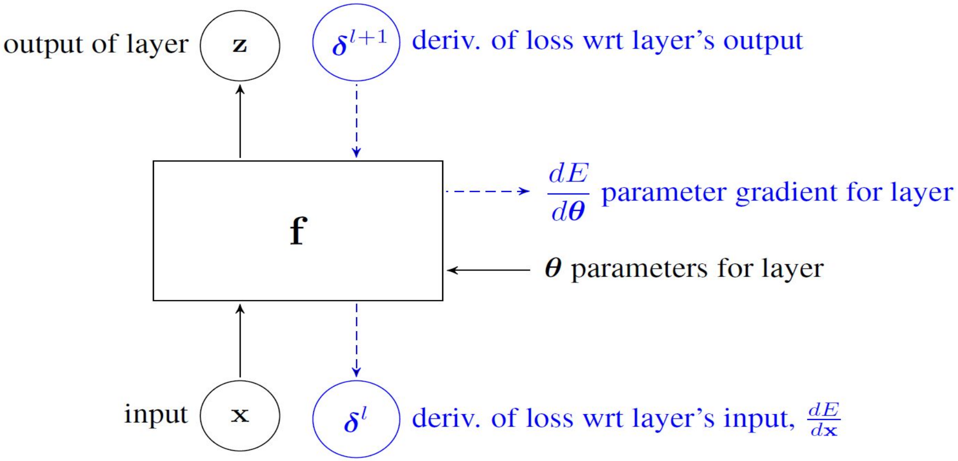


$$\mathbf{z}^{l+1} = \mathbf{f}^l(\mathbf{z}^l; \boldsymbol{\theta}^l)$$

$$\boldsymbol{\delta}^l := \frac{\partial E}{\partial \mathbf{z}^l} = \frac{\partial E}{\partial \mathbf{z}^{l+1}} \frac{\partial \mathbf{z}^{l+1}}{\partial \mathbf{z}^l} = \boldsymbol{\delta}^{l+1} \frac{\partial \mathbf{f}^l(\mathbf{z}^l; \boldsymbol{\theta}^l)}{\partial \mathbf{z}^l}$$

$$\frac{\partial E}{\partial \boldsymbol{\theta}^l} = \frac{\partial E}{\partial \mathbf{z}^{l+1}} \frac{\partial \mathbf{z}^{l+1}}{\partial \boldsymbol{\theta}^l} = \boldsymbol{\delta}^{l+1} \frac{\partial \mathbf{f}^l(\mathbf{z}^l; \boldsymbol{\theta}^l)}{\partial \boldsymbol{\theta}^l}$$

Example: Linear layer

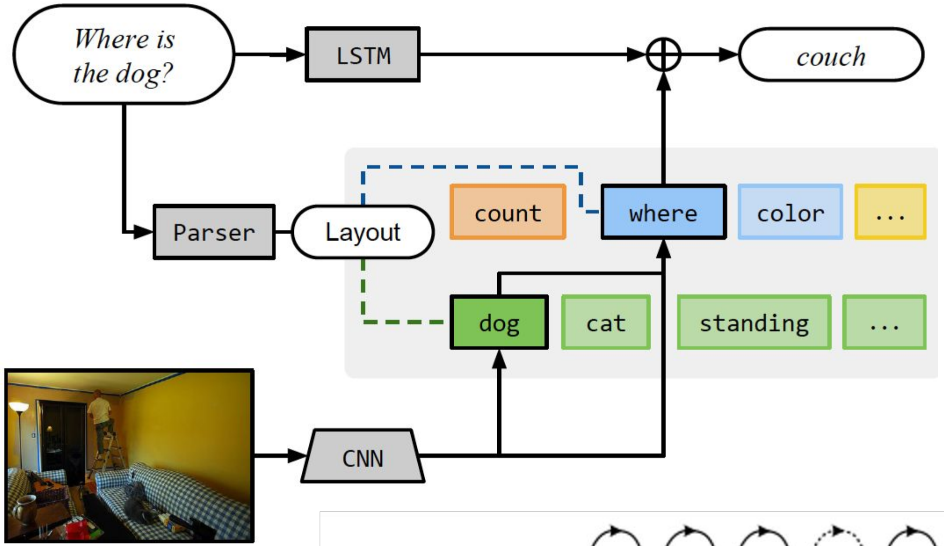


$$z_j = f_j(\mathbf{x}; \boldsymbol{\theta}_j) = \sum_i x_i \theta_{ij}$$

$$\delta_i^l = \sum_j \delta_j^{l+1} \frac{\partial f_j(\mathbf{x}; \boldsymbol{\theta}_j)}{\partial x_i} = \sum_j \delta_j^{l+1} \theta_{ij}$$

$$\frac{\partial E}{\partial \theta_{ij}} = \sum_j \delta_j^{l+1} \frac{\partial f_j(\mathbf{x}; \boldsymbol{\theta}_j)}{\partial \theta_{ij}} = \delta_j^{l+1} x_i$$

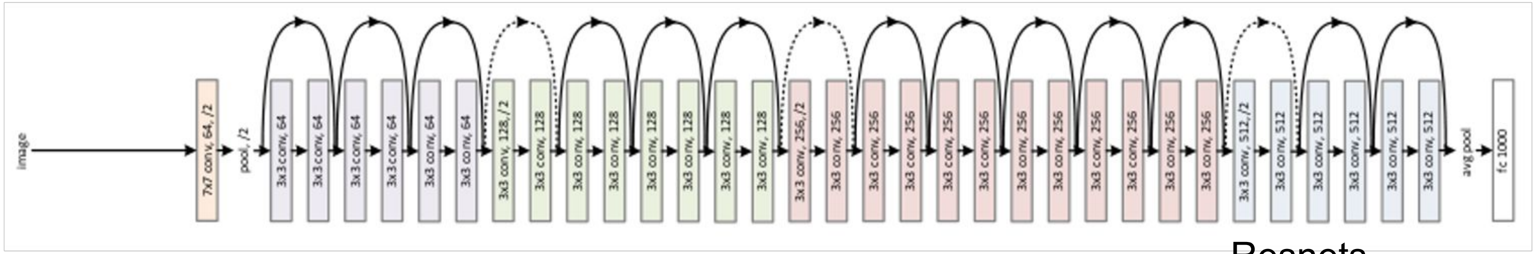
Another view as to why Deep Learning is so successful: Deep nets represent algorithms learned from data



Jacob Andreas

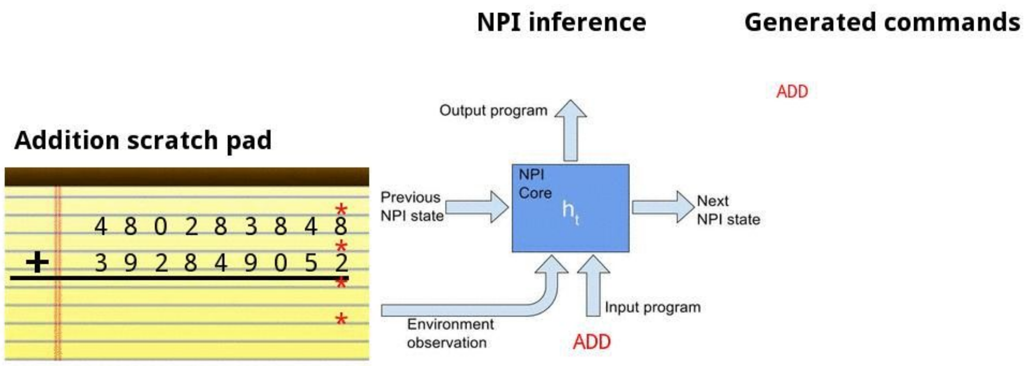
```
: BUBBLE ( a1 ... an n-1 -- one pass )
    DUP IF >R
        { observe D0 D-1 -> permute D-1 D0 R0 }
        1- BUBBLE R>
        \ ** Alternative sketch **
        \ { observe D0 D-1 -> choose NOP SWAP }
        \ R> SWAP >R 1- BUBBLE R>
    ELSE
        DROP
    THEN
;
```

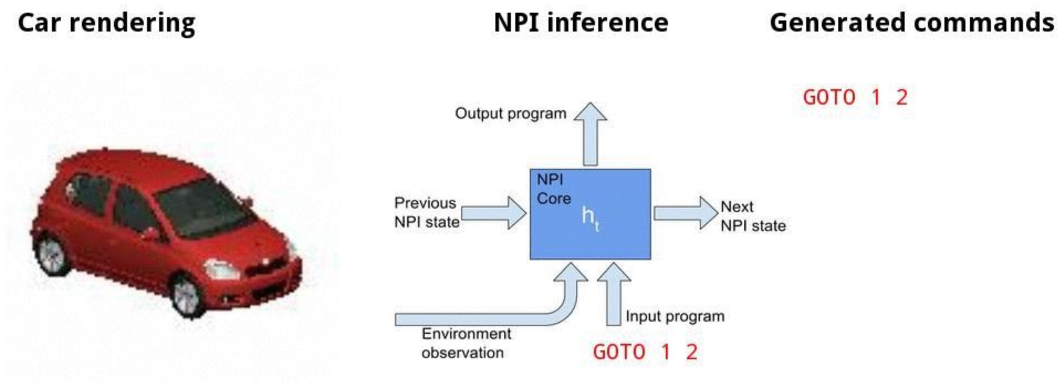
Differentiable Forth Interpreter
Sebastian Riedel et al



Resnets

Neural-Programmer Interpreters (NPI)

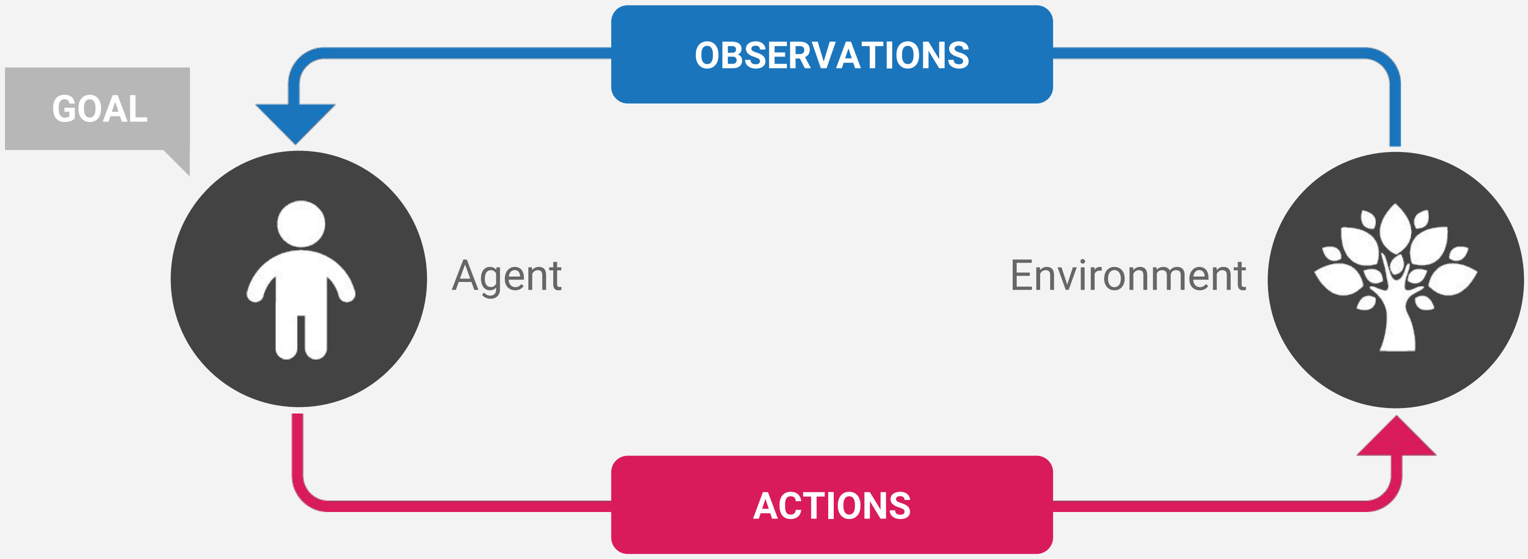




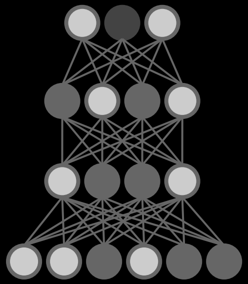
Frontiers that excite me

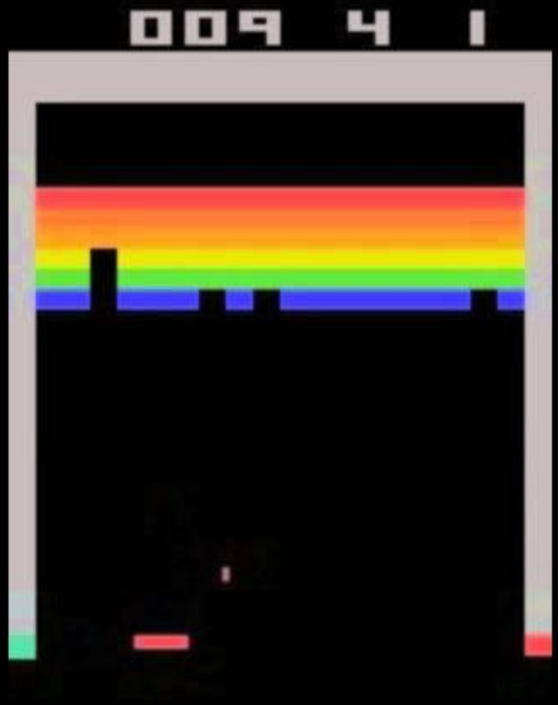
1. Reinforcement learning
2. Meta learning
3. Imitation
4. Robotics
5. Concepts and abstraction
6. Awareness and consciousness
7. Causal reasoning

Reinforcement Learning Framework



Atari with deep RL

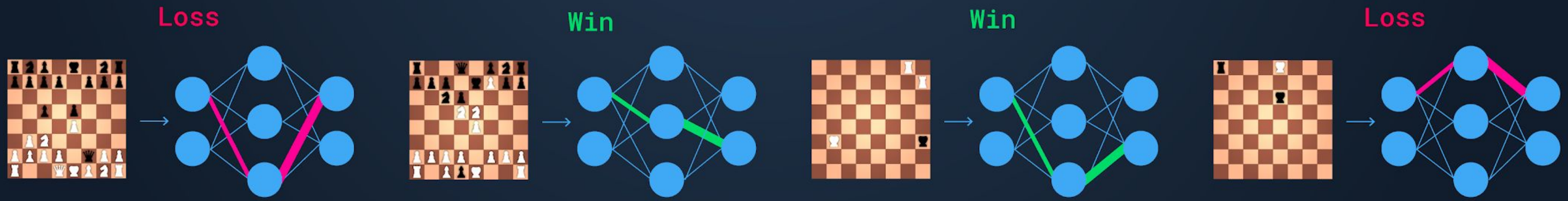




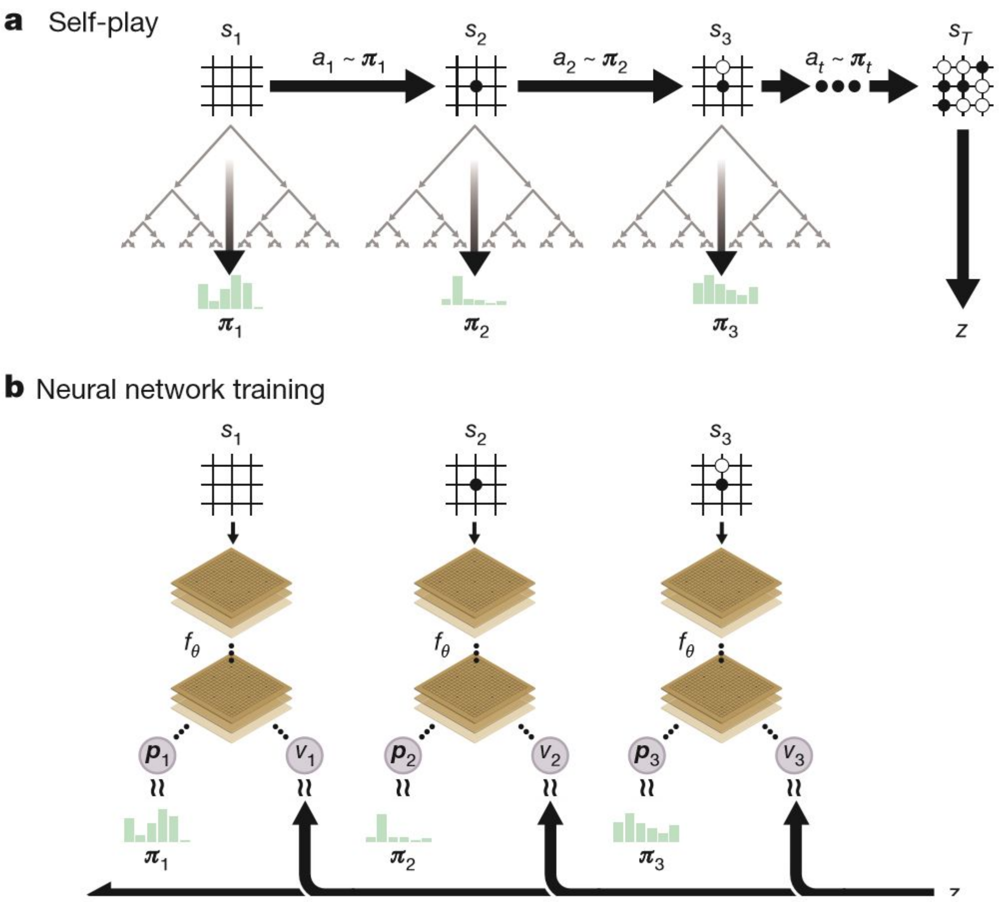
AlphaZero

Starts off playing completely randomly

Learns from experience using reinforcement learning



Alphazero



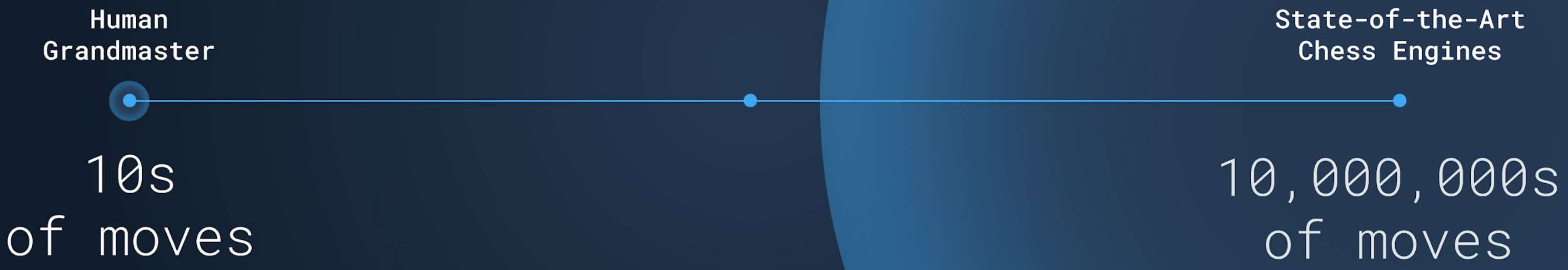
a Self-play
s_1
$a_1 \sim \pi_1$
s_2
$a_2 \sim \pi_2$
s_3
$a_t \sim \pi_t$
s_T
π_1
π_2
π_3
z
b Neural network training
s_1
s_2
s_3
f_θ
f_θ
f_θ
p_1
v_1
p_2
v_2
p_3
v_3
π_1
π_2
π_3
z

Amount of Search per Decision

Human
Grandmaster

10s
of moves

Amount of Search per Decision



Amount of Search per Decision

Human Grandmaster

10s of moves

AlphaZero

10,000s of moves

State-of-the-Art Chess Engines

10,000,000s of moves

Mastering many games with a general RL algorithm

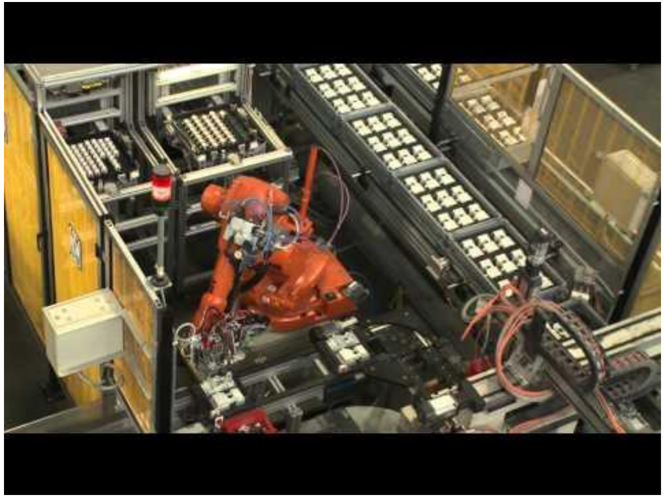
Chess

Shogi

Go

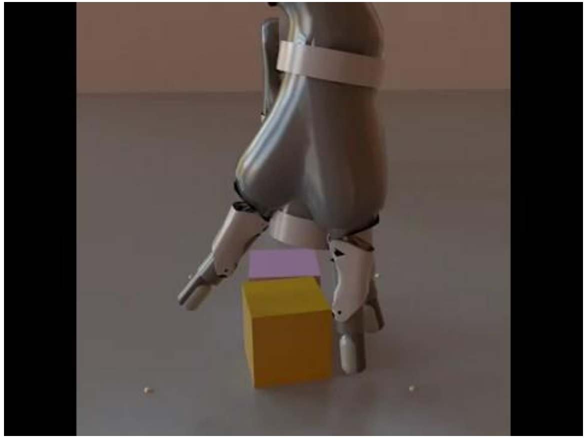
Why not just use robots?

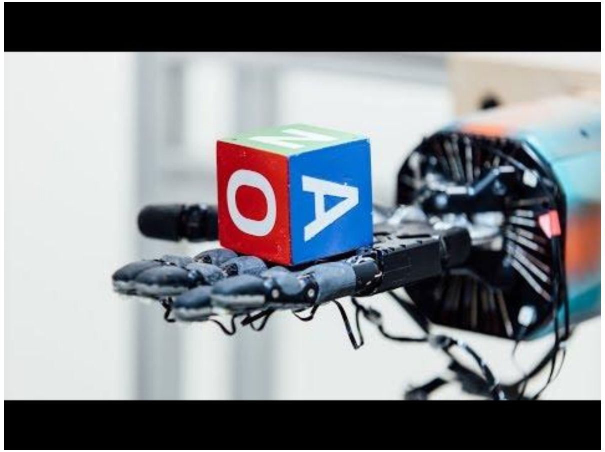
- **Industrial Automation**
  - Perfect repeatability, exceedingly-strong and dangerous robots.
  - Structured workspace, intensive human engineering, very costly to adjust.
  - Limited to some materials.
- **State-of-the-Art: DARPA Robotics Challenge**
  - Fragile stack of hand-crafted estimators and controllers.



Thanks Tom Erez

Deep learning for manipulation has begun!





Imitation

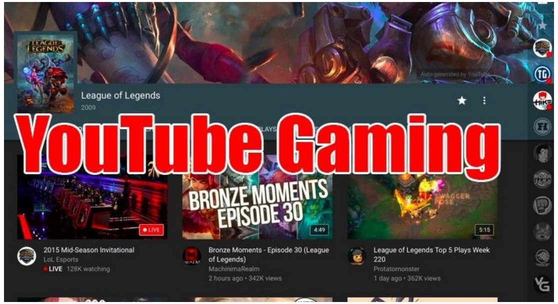
1. Helps us address exploration in RL.
2. It forces us to address third-person transfer, awareness of self and others, concepts and abstraction.
3. It manifests also in communities of agents and it is essential to culture.
4. It results in many useful apps.

Imitating people is important

- Models of translation, speech, ...
- Universal collaboration: Machines and humans working together
- Models for psychologists, sociologists, kinesiologists, economists, …
- The monetary value of behavioural models of people is vast (trillions)

Learning by watching YouTube

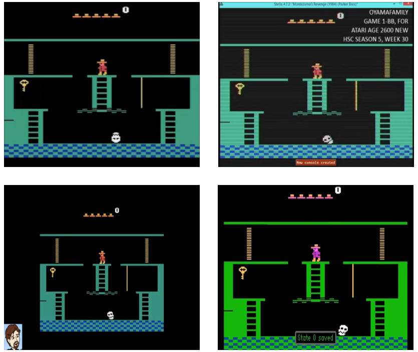
People learn many tasks by watching online videos
Despite huge gaps in visual appearance,
sensing modalities,
body differences, etc..



Challenges

Learning by watching Youtube

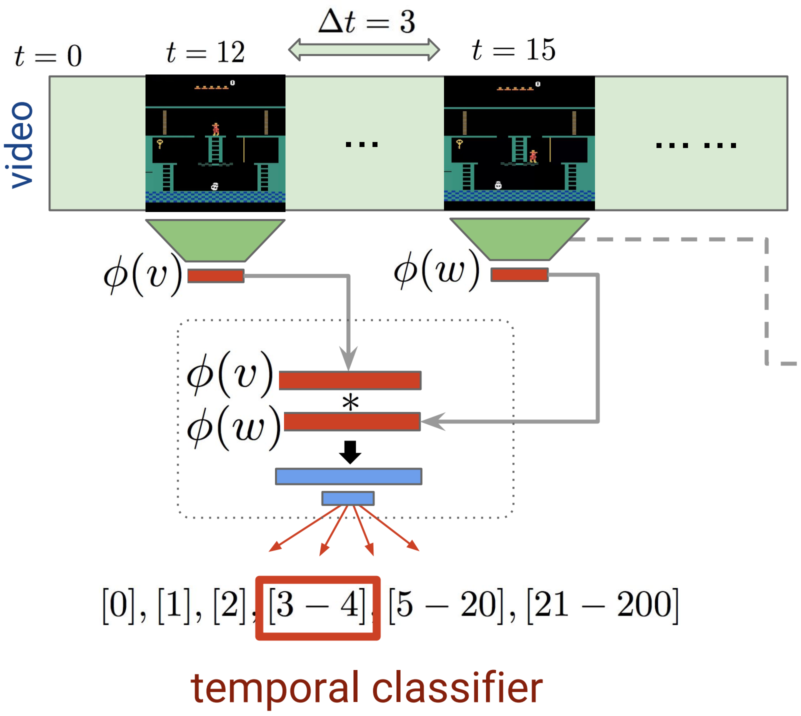
Domain Gap

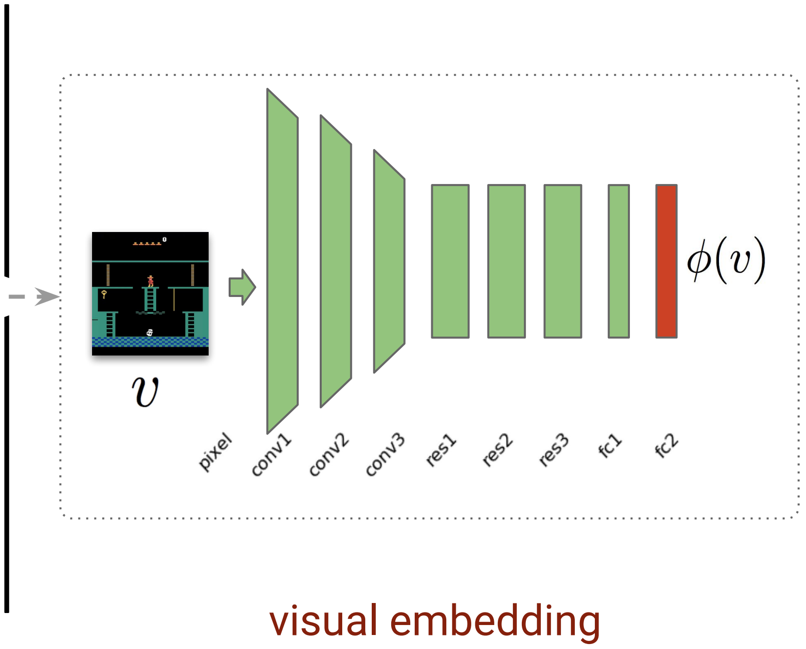


No Actions

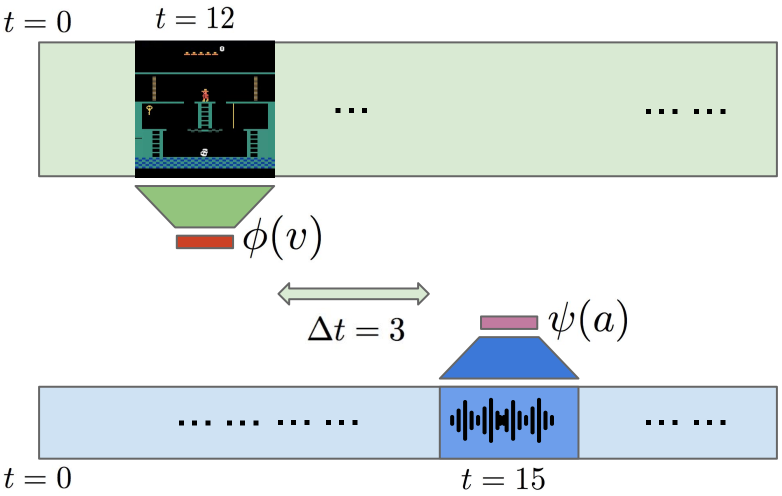
No Rewards

Temporal distance classification (TDC)

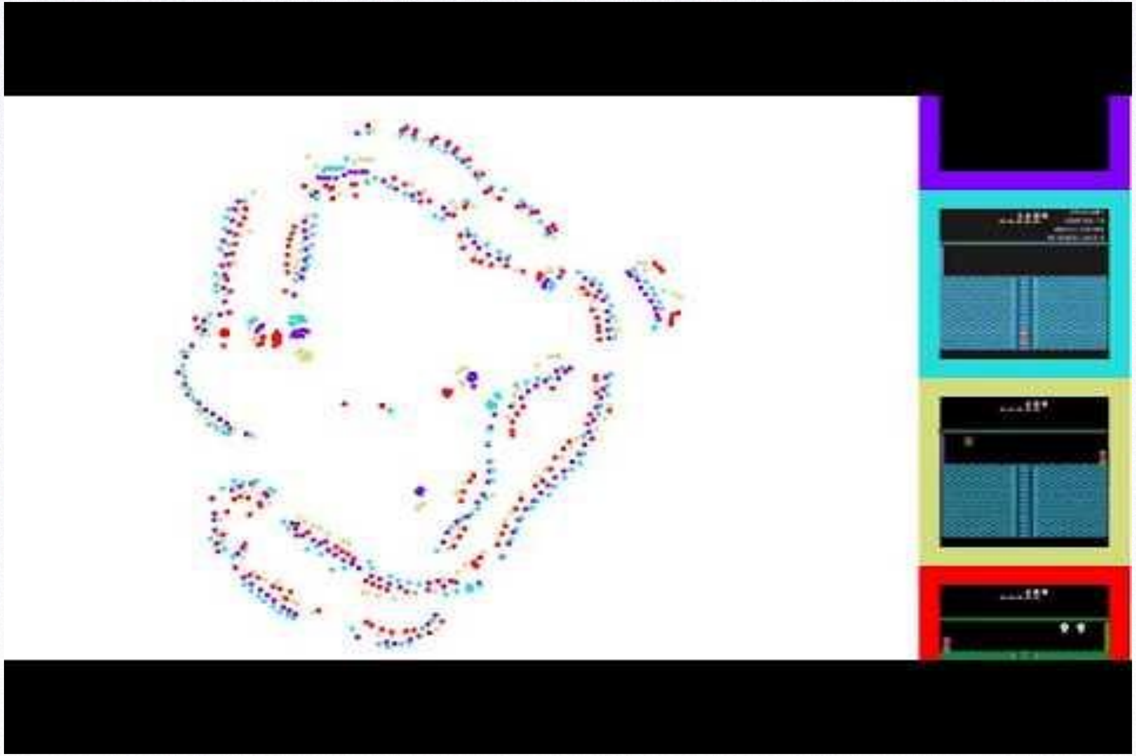




Cross-modal distance classification (CMC)

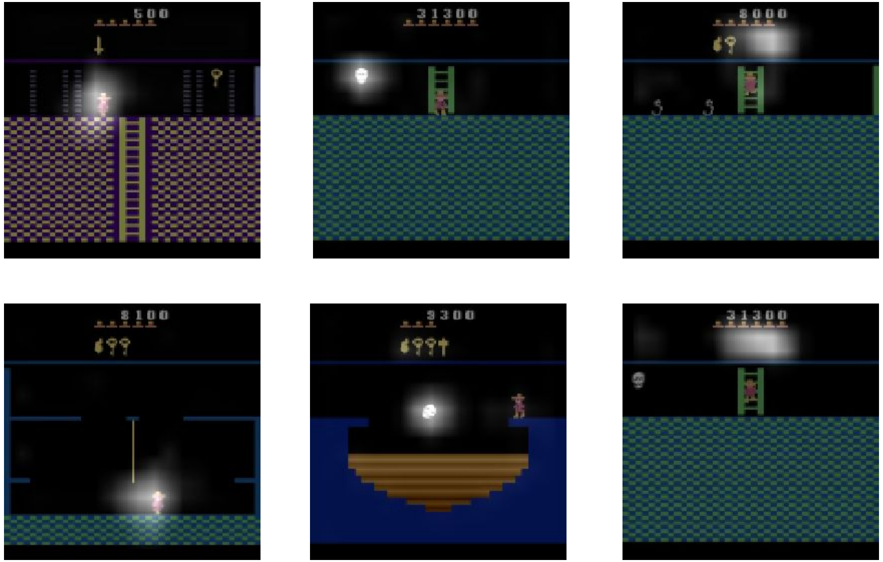


Closing the domain gap



What does the embedding focus on?

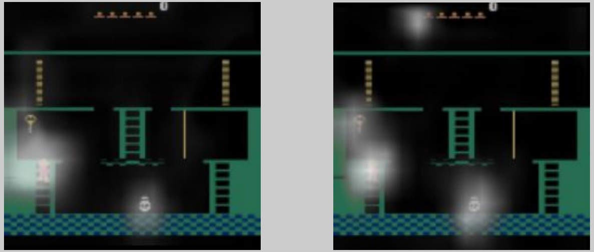
Neuron Activations and how multimodality helps in learning concepts



Neuron #46
Player

Neuron #8
Enemies

Neuron #39
Inventory

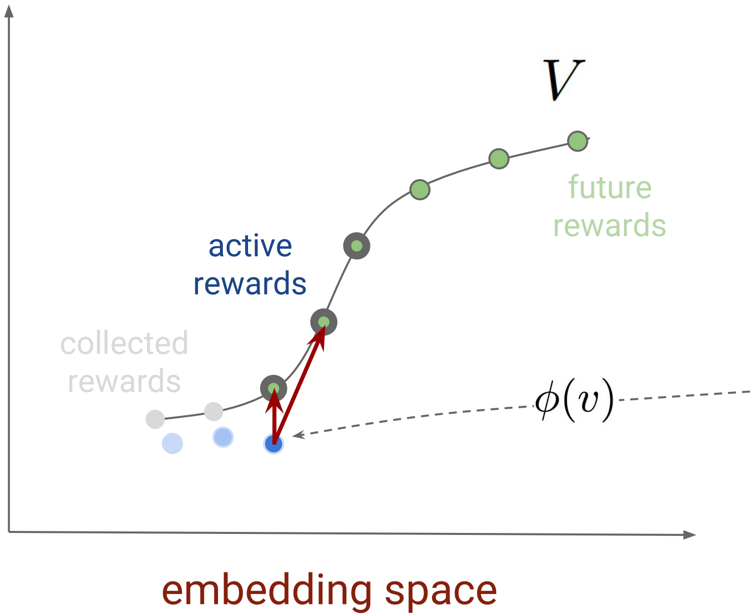


Visual only

Cross-modal

Imitation through RL via checkpoints

IMPALA Agent (or any other RL Method)



Similarity to checkpoint on expert trajectory:

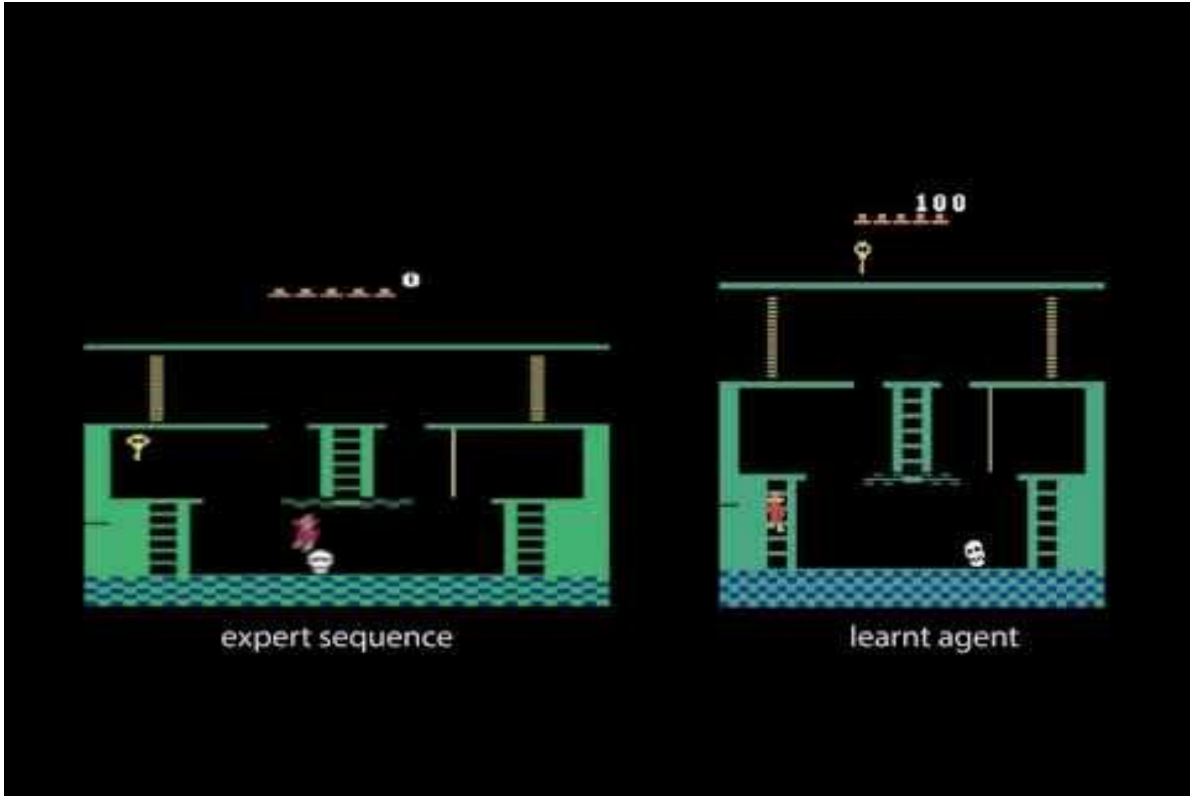
$$\bar{\phi}(v_{\text{agent}}) \cdot \bar{\phi}(v_{\text{checkpoint}})$$

$$r_{\text{imitation}} = \begin{cases} 0.5 & \text{if } \bar{\phi}(v_{\text{agent}}) \cdot \bar{\phi}(v_{\text{checkpoint}}) > \gamma \\ 0.0 & \text{otherwise} \end{cases}$$

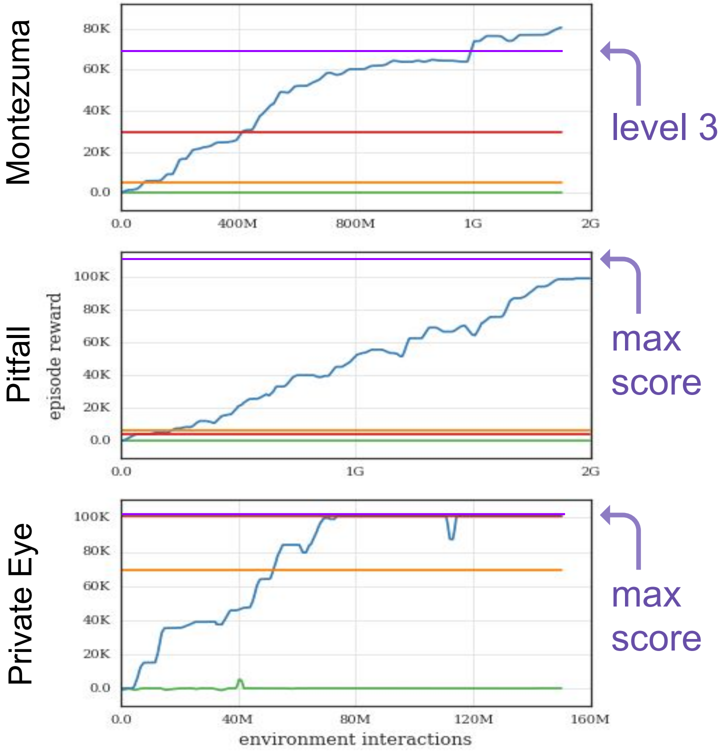
observation

v

Imitation through RL to treat variation

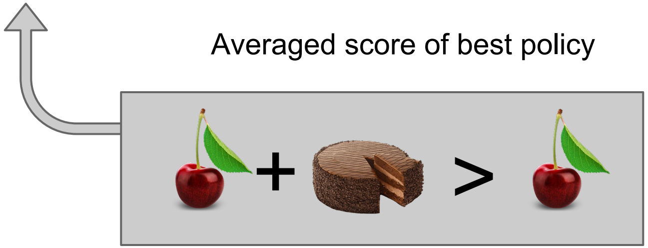


Results on hard-exploration Atari games

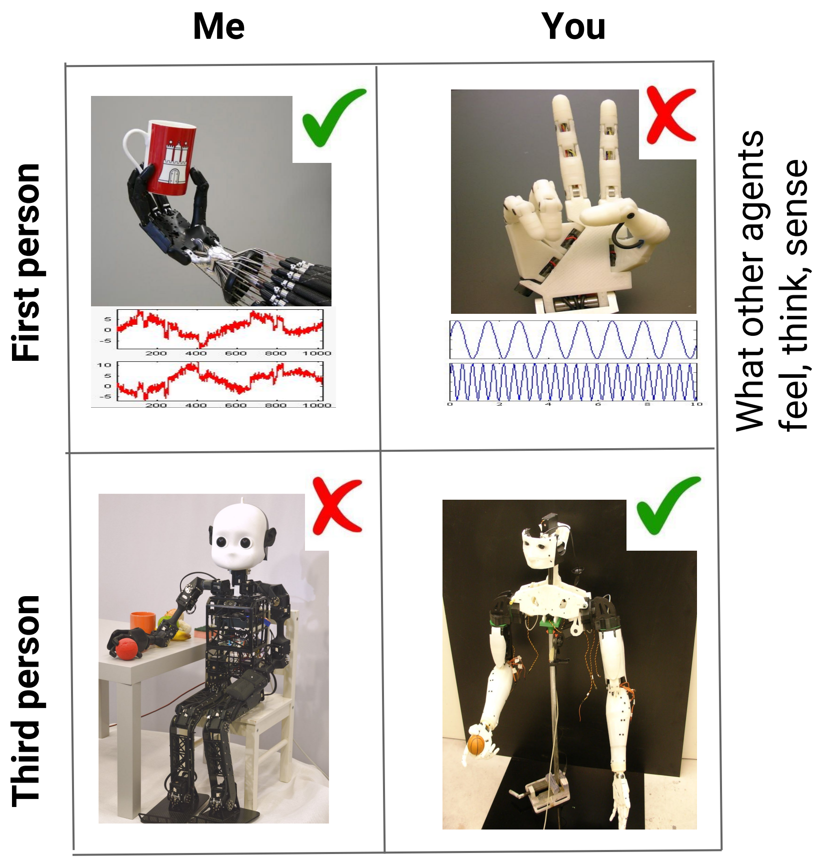


	Montezuma	Pitfall!	Private Eye
Pure RL	~ 2,500	~ 0	~ 50
Avg. Human	4,743	6,464	69,571
DQfD (NIPS 2018)	29,384	3,997	**100,747**
Ours	**58,175**	**74,323**	98,763

Averaged score of best policy



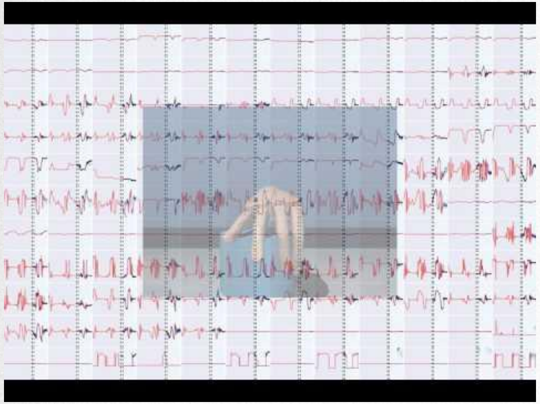
Awareness, theory of mind and consciousness



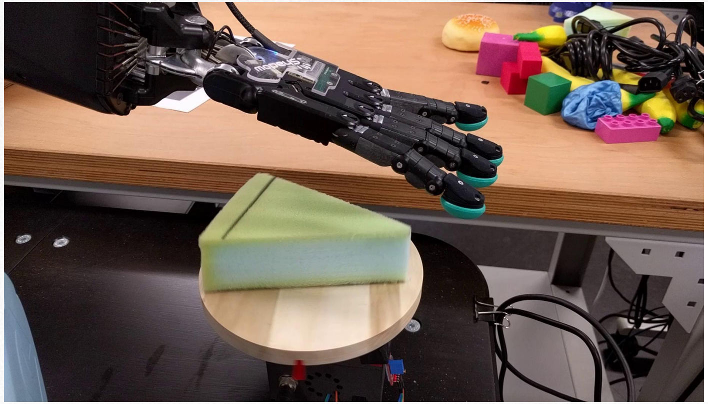
1. Knowledge of the self would help with disentangling and representation learning.
2. It is important to learn to identify other agents, their behaviour, and intentions.
3. An intelligent machine must know what it knows and what it does not know. It must have mental, body, and environment awareness.
4. Awareness provides a framework not only for imitation but also for teaching and empathy.

Awareness, theory of mind and consciousness

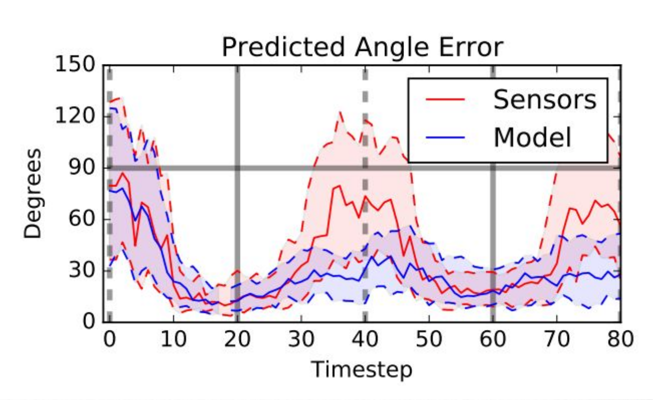
1. **Make many body predictions with deep state space models**



2. **Learn awareness models by playing (RL)**



3. **Obtain dynamic, persistent, wholistic representations of the world**



Learning **slow** to learn **fast**

- Infants are endowed with systems of core knowledge for reasoning about objects, actions, number, space, and social interactions [eg E. Spelke].
- The slow learning process of evolution led to the emergence of components that enable fast and varied forms of learning.

Few-shot meta learning

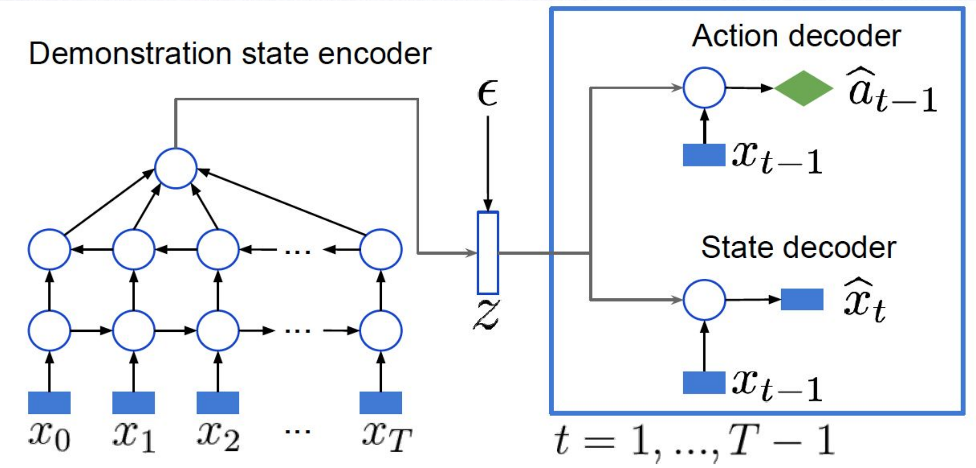
Traditional (supervised) ML

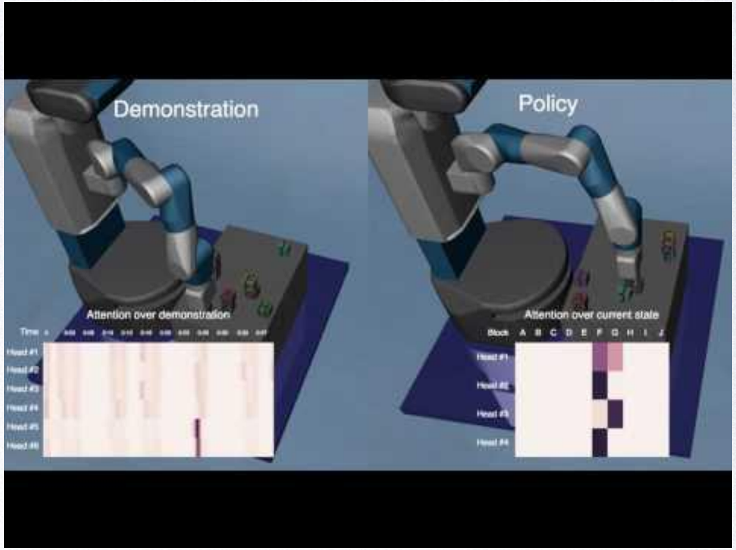
- Train a function approximator with lots of data
- Deploy a (eg) classifier that expects input data to generate predictions

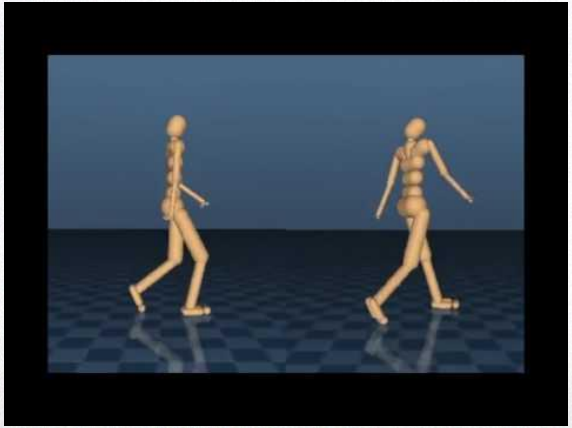
Unleashing the few-shot learning machines

- Train a learning model with lots of data, and with the capacity to learn
- Deploy a model that needs a few data (demos, labeled samples) to rapidly adapt to a novel task

Conditional policies for one-shot imitation learning







Ziyu Wang, Josh Merel, Scott Reed, Greg Wayne, NdF, Nicolas Heess (2017)

Yan Duan, Marcin Andrychowicz, Bradly Stadie, Jonathan Ho, Jonas Schneider, Ilya Sutskever, Pieter Abbeel, Wojciech Zaremba (2017)

Causal reasoning

- How do we learn models of the world to achieve counterfactual reasoning?
- One option: Agent in simulation must learn to cause the simulated environment to imitate the real environment, it can then intervene, determine causation, and use it to make predictions. E.g., it can learn that the rooster doesn’t cause the sun to rise. We need to neuralize this.

Many other frontiers
This is an exciting time to do research in AI

- Abstraction, concepts, relations, objects, programs, architectures
- Selecting tasks for self-supervision automatically
- Continual knowledge representation in an agent
- Grounded language understanding (and lifting)
- The emotional/motivational system
- Solid, flexible, software frameworks
- Module invention (convolution, normalization, resnets, attention, LSTMs)
- Ethics and governance (bias, liability, ownership, explainability, labor, …)

Thank you!

DeepMind