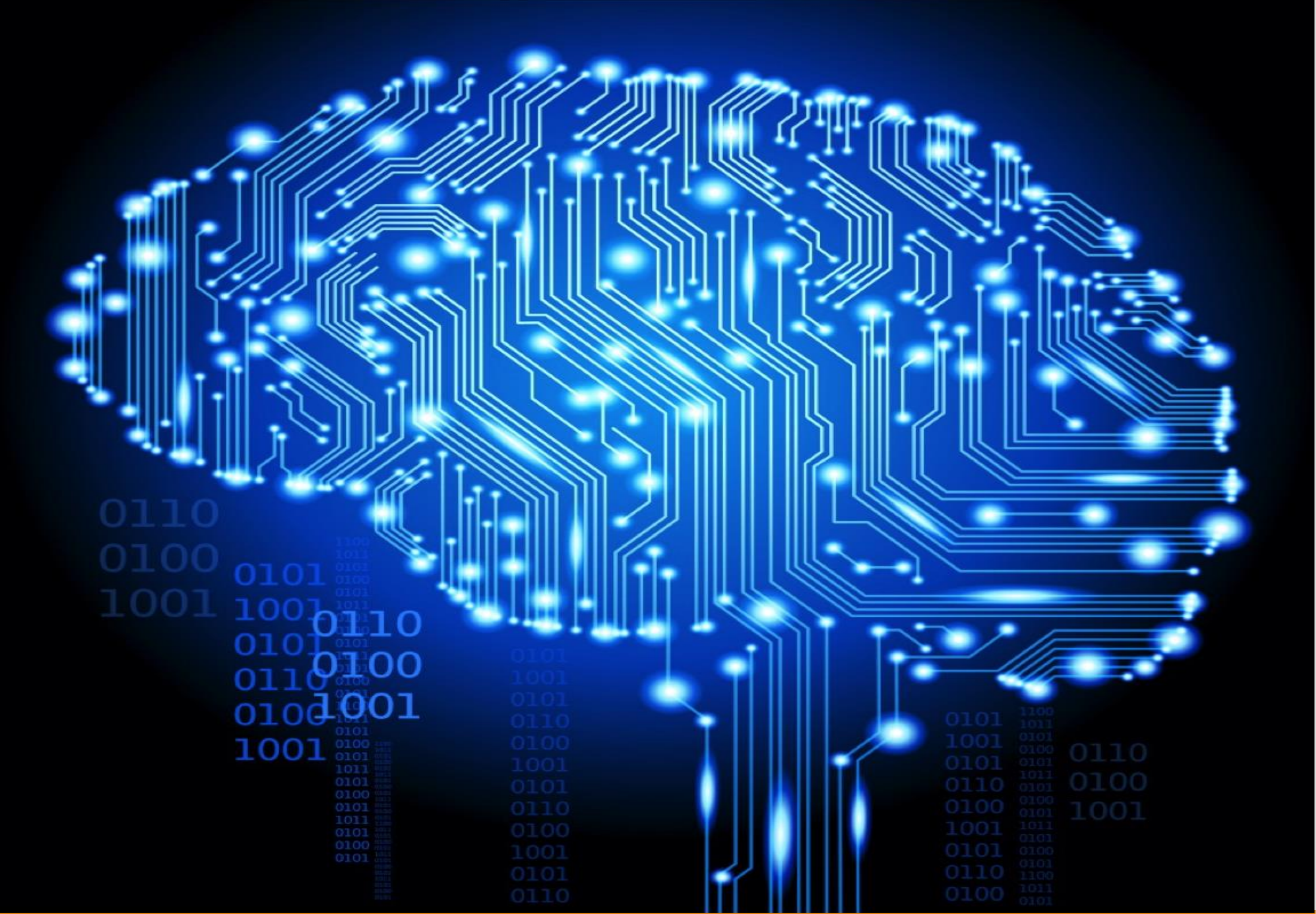


Lecture 14: Exotic Deep Learning

Deep Learning @ UvA

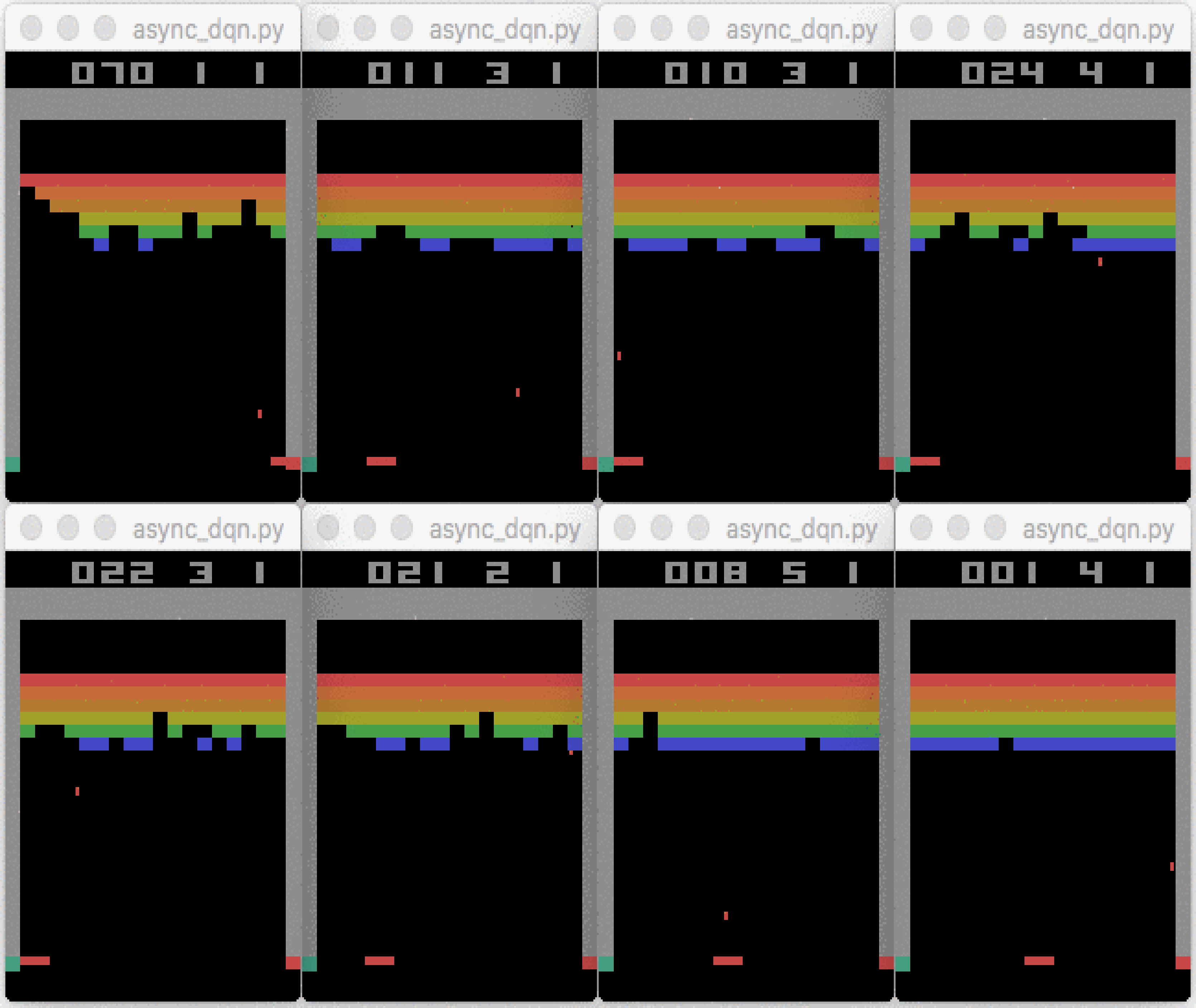
Previous Lecture

- Memory networks
- How to evaluate memory networks

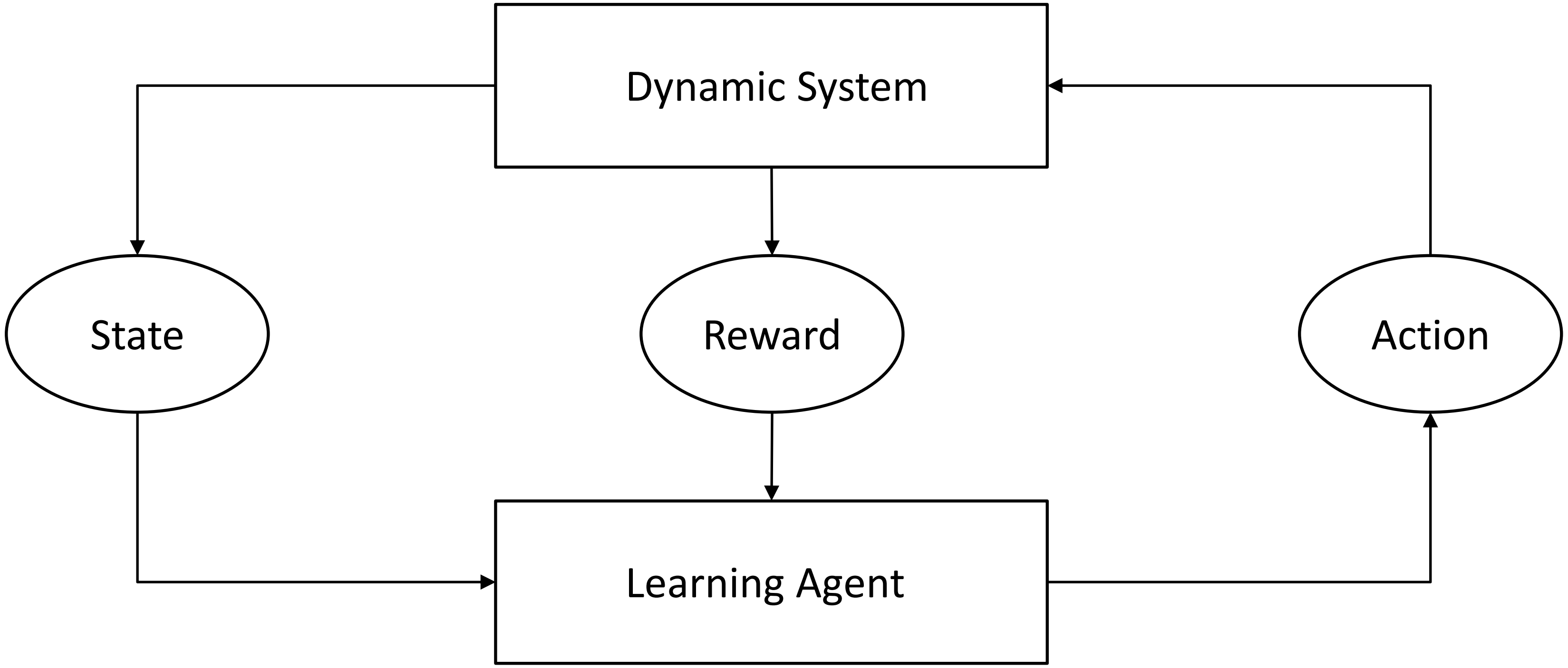
Lecture Overview

- Reinforcement Learning
- One-shot learning
- Dynamic filters

Reinforcement Learning

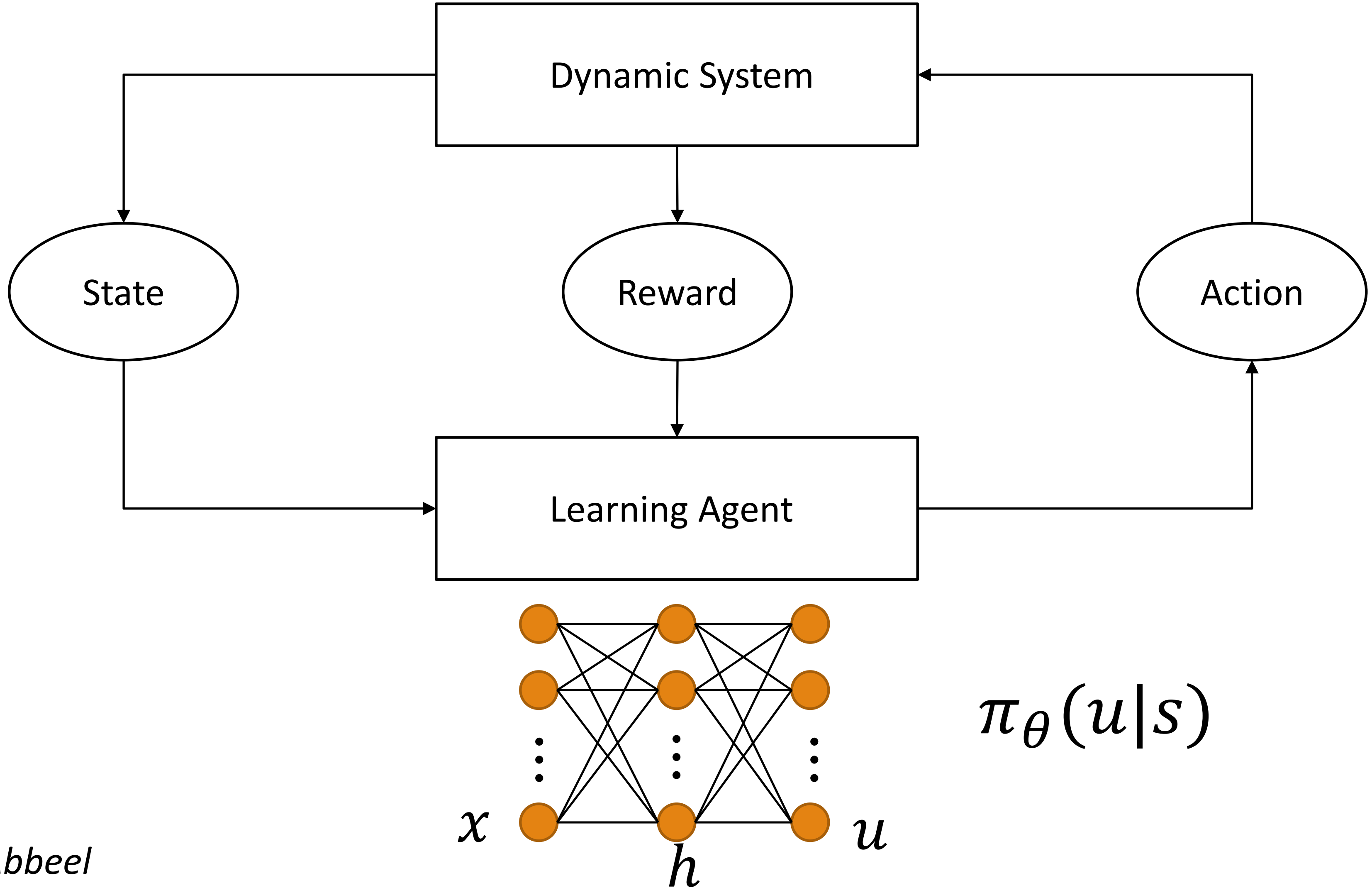


Reinforcement Learning



Slides inspired by P. Abbeel

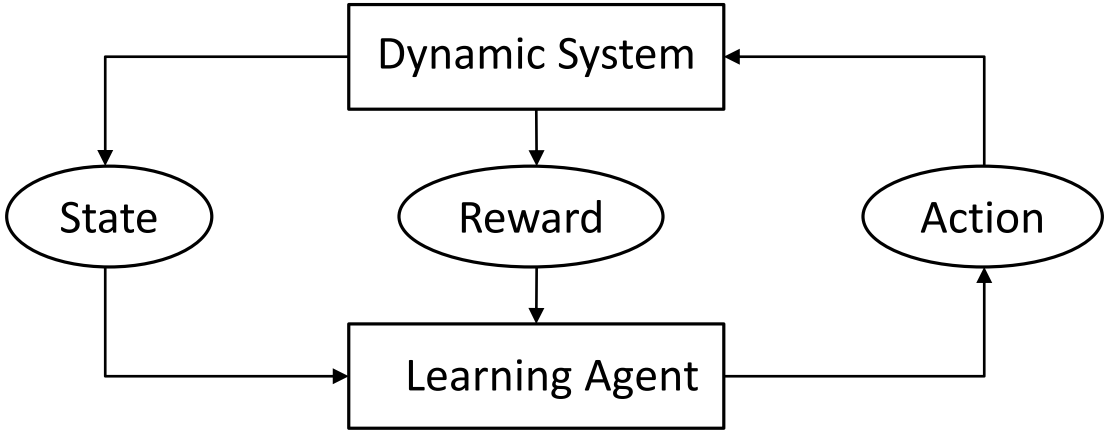
Policy optimization

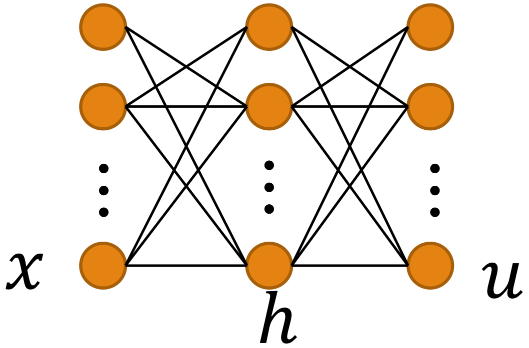


$$\pi_\theta(u|s)$$

Slides inspired by P. Abbeel

Policy optimization





$$\pi_\theta(u|s)$$

- Train learning agent for the optimal policy $\pi_\theta(u|s)$ given states s and possible actions u
- The policy class can be either deterministic or stochastic

Slides inspired by P. Abbeel

Benefits of Policy Optimization

- Often defining the policy $\pi_\theta(u|s)$ is easier than defining the Q-function
- Moreover, computing the Q-value is often too expensive
  - Hard to solve $\arg\max_{\alpha} Q_\theta(s, a)$
  - Especially when having continuous or high-dimensional action spaces
- Have a deep network optimize the $\pi_\theta(u|s)$

Slides inspired by P. Abbeel

Deep Reinforcement Learning

- Regards r_t at time t
- Actions π taken according to a policy $\pi = \mathrm{P}(a|s)$
- An action-value function

$$Q(s,a) = \max_{\pi} \mathbb{E}[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots | s_t = s, a_t = a, \pi]$$

where γ is a discount factor of the future rewards
  - Future rewards should not be as important, because we do not know the future
- Use a non-linear function approximator to model the action value function

$$Q^*(s,a) \approx Q(s,a;\theta)$$

Deep Reinforcement Learning

- Non-linear function approximator?
  - Deep Networks
- Input is as raw as possible, e.g. image frame
  - Or perhaps several frames
- Output is the best possible action out of a set of actions for maximizing future reward
- **Important:** no need anymore to compute the actual value of the action-value function and take the maximum: $\arg\max_{\alpha} Q_\theta(s, a)$
  - The network returns directly the optimal action

Extra Tricks

- Replay memory/Experience replay
  - Store memories $< s, a, r, s' >$
  - Train using random stored memories instead of the latest memory transition
  - Breaks the temporal dependencies – SGD works well if samples are roughly independent
- Skipping frames
  - Saves time and computation
  - Anyways, from one frame to the other there is often very little difference
- ε-greedy behavioral policy with annealed temperature during training
  - Select random action (instead of optimal) with probability ε
  - In the beginning of training our model is bad, no reason to trust the "optimal" action
- Alternatively: Exploration vs exploitation
  - early stages $\equiv$ strong exploration
  - late stages $\equiv$ strong exploitation

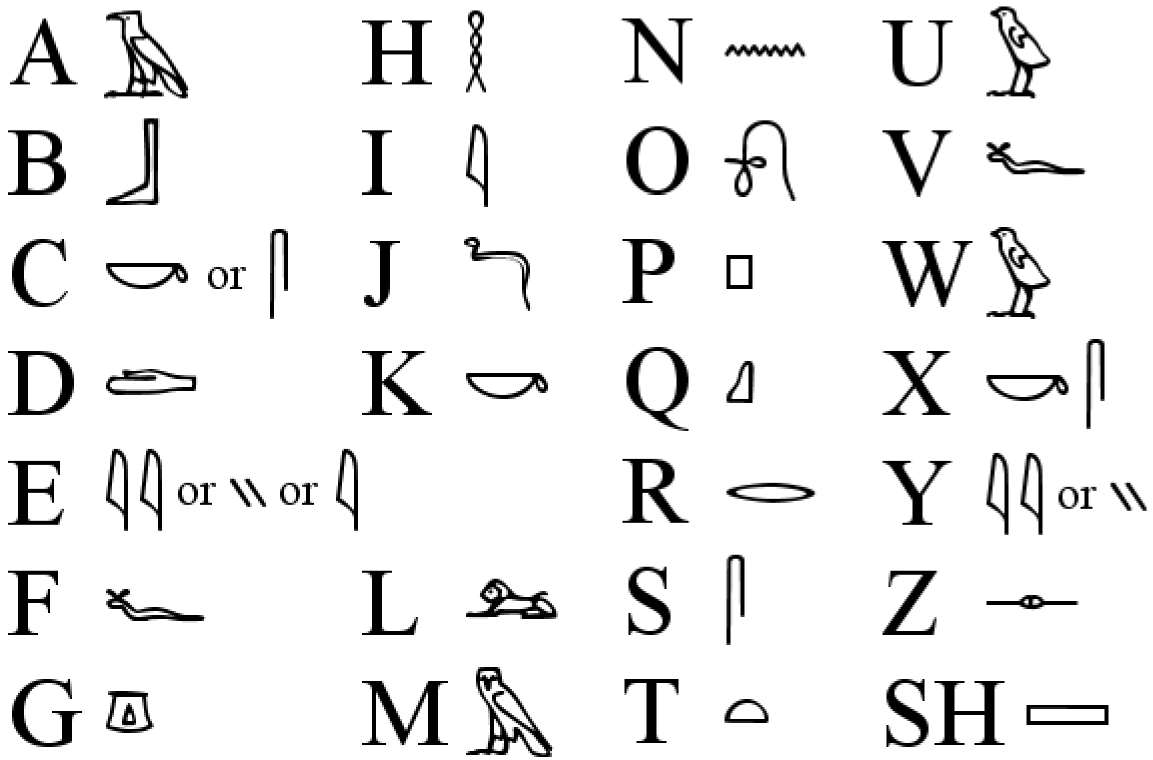
One-Shot Learning

Why One-Shot?

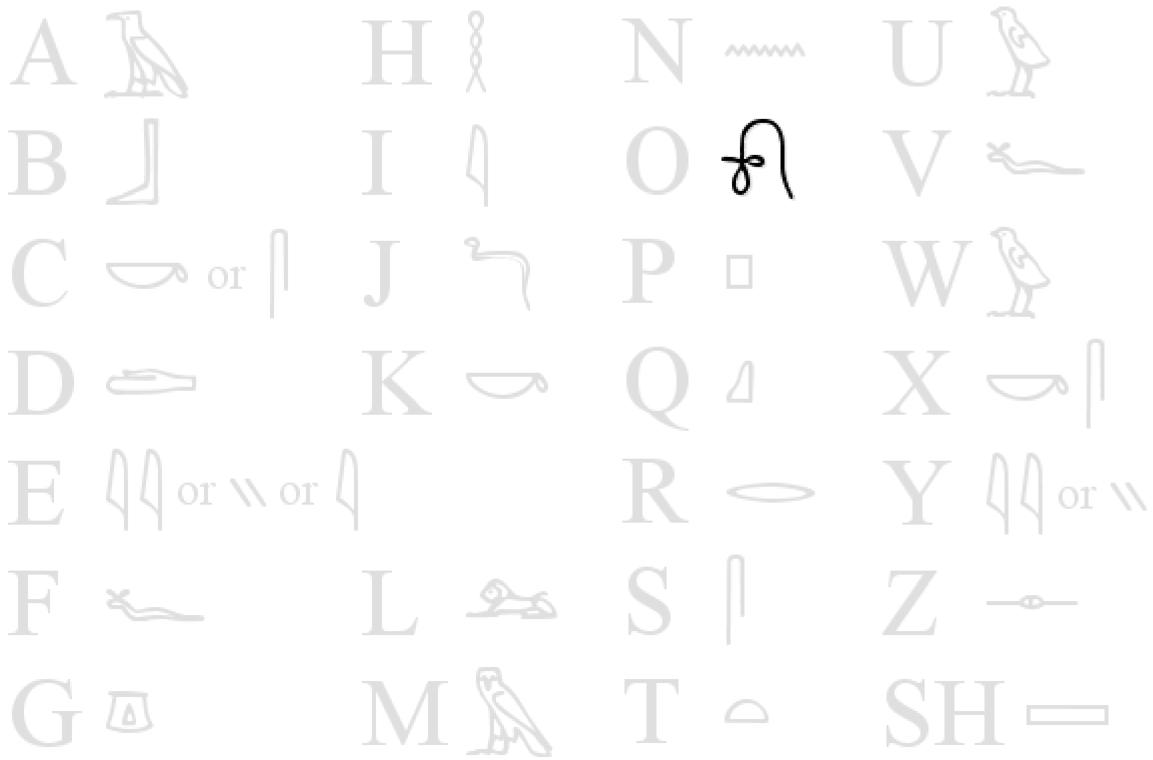
- Nowadays available datasets are huge
  - Imagenet: 16 million images
  - MSCOCO: $> 200{,}000$ question answer-pairs
- However, can we rely on huge datasets for everything?
- Also, …

Slides inspired by P. Abbeel

Why One-Shot?



Why One-Shot?



Why One-Shot?

Why One-Shot?

Why On-Shot?

- Nowadays available datasets are huge
  - Imagenet: 16 million images
  - MSCOCO: $> 200{,}000$ question answer-pairs
- However, can we rely on huge datasets for everything?
- Also, … humans can learn from one example
  - why not deep networks also?

Slides inspired by P. Abbeel

How to do one-shot?

- Siamese networks-instance matching
  - String geometric requirements
  - Hard when target should be matched on a conceptual and not on pixel level
  - *Siamese Instance Search for Tracking, Tao et al.*
- Matching networks
  - *Matching Networks for One-Shot Learning, Vinyals et al.*
- Memory networks
  - *One-shot learning with Memory-Augmented Neural Networks, Santoro et al.*
- Dynamic filter generation
  - Instead of fixing filters after training, learn to generate filter based on input

More exotic stuff

- Dropping layers
- Searching for Exotic Particles in High-Energy Physics with Deep Learning
- Neural programming
- Reinforcement Learning for SLAM navigation
- Gaussian Process and Deep Networks
- And many more stuff …
- So, just explore the possibilities

Summary

- Reinforcement Learning
- One-shot learning
- Dynamic filters
- More exotic deep learning is yet to come

MSc theses

Autonomous Drone Navigation

- We have a quadcopter in the QUVA Lab, which one can fly using the joystick. We want now to teach the quadcopter to learn to fly itself, without handcrafting its navigation software
- Related concepts
  - Supervised learning
  - Reinforcement learning
- Extra requirements
  - Experience with drone and mobile programming
- Contact: Efstratios Gavves (egavves@uva.nl)

Generative Models and Adversarial Networks

- Recently there has been a significant increase of interest for generative networks, also because of the emergence of Adversarial Networks. However, we still lack understanding of adversarial networks and how they precisely fit within the generative model framework. This topic is more open-ended and will be defined upon discussion with the candidate, after reading the very recent literature
- Related concepts
  - Adversarial networks
  - Variational Autoencoders
  - Generative Models
- Extra requirements
  - More interested for the theoretical aspect of deep learning
- Contact: Efstratios Gavves (egavves@uva.nl)

Recurrent Siamese Trackers

- In the recent literature a new type of tracking framework has emerged that casts tracking as an instance search problem. This framework, however, has the disadvantage that ignores the obvious temporal coherence between the frames of the tracked object over time. As such integration with recurrent temporal models is a promising direction.
- Related concepts
  - Instance search
  - Siamese networks
  - Recurrent networks
- Contact: Efstratios Gavves (egavves@uva.nl)

Deep Knowledge Memory Networks

- Training memory networks for answering sophisticated questions, modelling dialogues entails two components. First, learning how to read, write and access memories. Second, learning to use relevant memories for answering a question. Currently, memory networks rely on the same training set for these two training aspects. However, this does not necessarily need to be the case. In this thesis we will investigate how to use external knowledge datasets for sophisticated question answering and dialogue modelling.
- Related concepts
  - Knowledge databases
  - Memory networks
  - Dialogue models
- Contact: Efstratios Gavves (egavves@uva.nl), E. Kanoulas (e.kanoulas@uva.nl)

Reading material & references

- https://sites.google.com/site/deeplearningsummerschool2016/speakers