

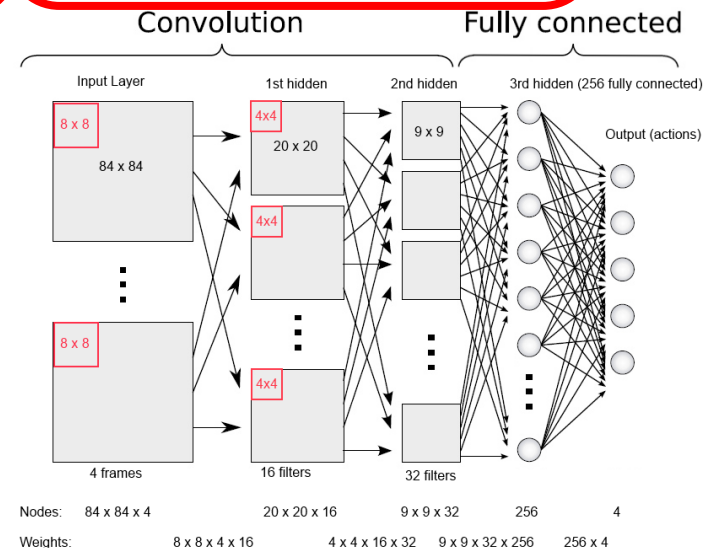
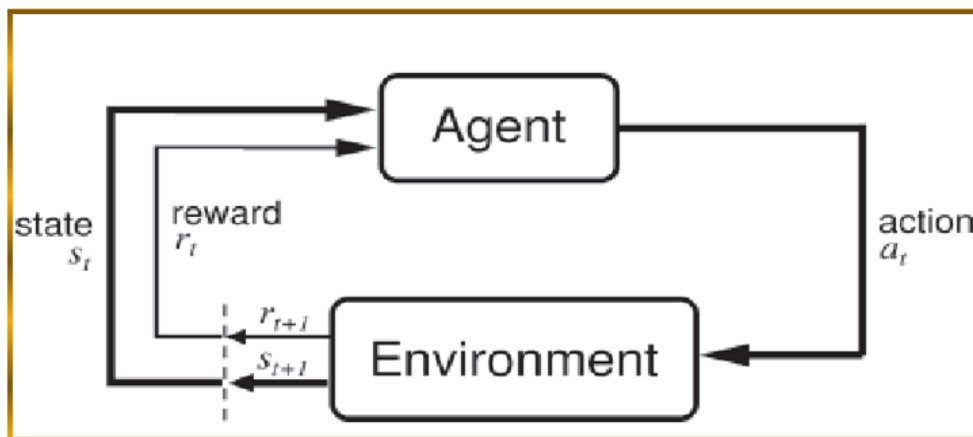
# Replicating DeepMind

Computational neuroscience project by  
Ardi, Ilya, Kristjan and Taivo

Tartu, Estonia, 2014/05/25



# Playing Atari with Deep Reinforcement Learning



# Results

	<b>B. Rider</b>	<b>Breakout</b>	<b>Enduro</b>	<b>Pong</b>	<b>Q*bert</b>	<b>Seaquest</b>	<b>S. Invaders</b>
<b>Random</b>	354	1.2	0	−20.4	157	110	179
<b>Sarsa [3]</b>	996	5.2	129	−19	614	665	271
<b>Contingency [4]</b>	1743	6	159	−17	960	723	268
<b>DQN</b>	<b>4092</b>	<b>168</b>	<b>470</b>	<b>20</b>	<b>1952</b>	<b>1705</b>	<b>581</b>
<b>Human</b>	7456	31	368	−3	18900	28010	3690

# People associated with DeepMind

- Demis Hassabis
  - Poker player, computer game designer and neuroscientist
- Shane Legg
  - Universal Intelligence: A Definition of Machine Intelligence
    - „Intelligence measures an agent’s ability to achieve goals in a wide range of environments.“
- Jaan Tallinn
  - Investor of DeepMind



# Structure of the DeepMind

- Play lots of games
  - Play lots of frames
    - Get a frame and **pre-process it**

# Pre-processing

- Image
  1.  $210 \times 160$  RGB
  2.  $110 \times 84$  gray-scale
  3.  $84 \times 84$
- Every 4<sup>th</sup> frame from the emulator
- State contains 4 frames
- Reward is binarized

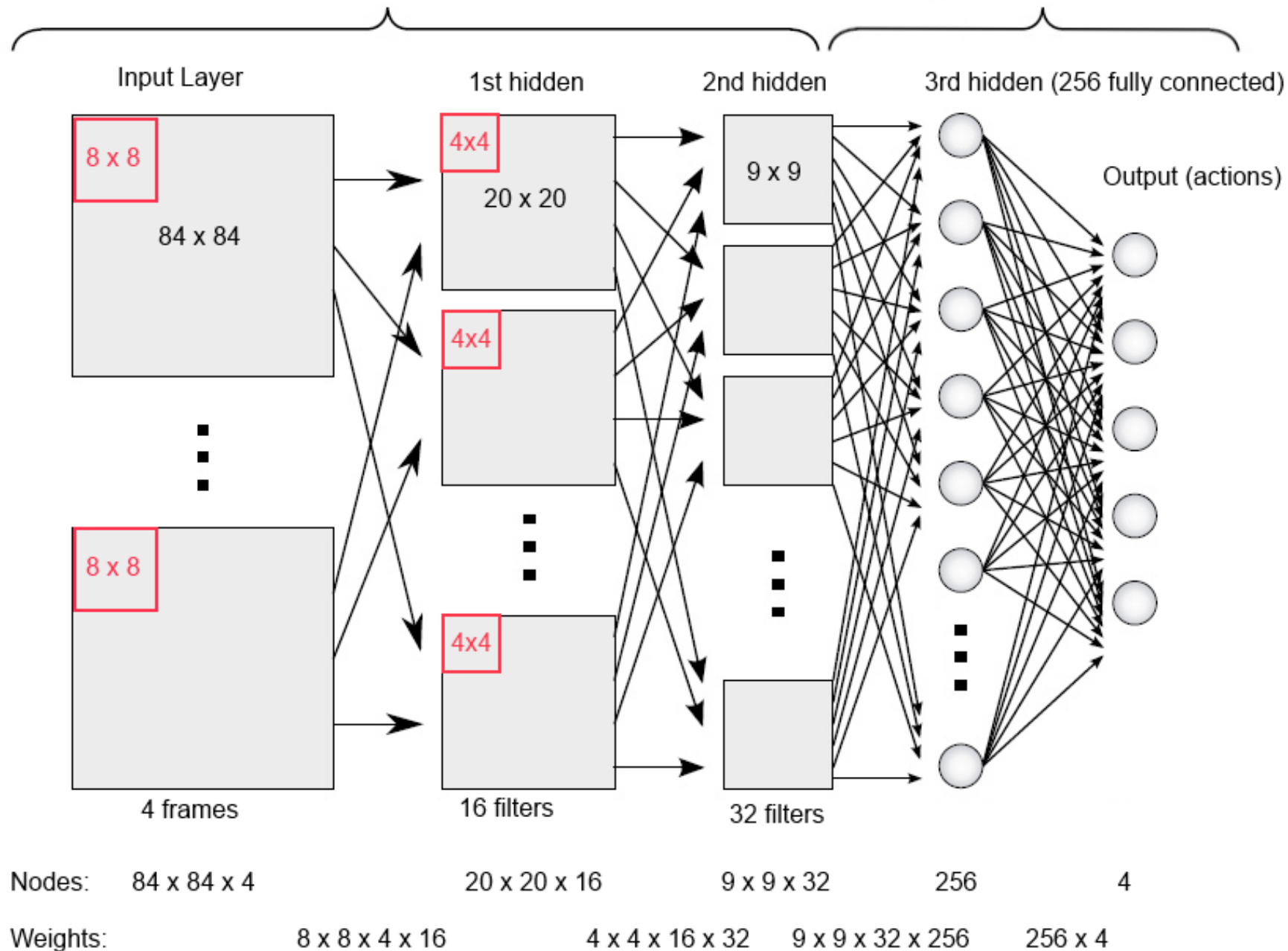
# Structure of the DeepMind

- Play lots of games
  - Play lots of frames
    - Get a frame and pre-process it
    - Choose action using **Deep Neural Net (DNN)** and current state
      - Do it randomly with probability  $\epsilon$



# Convolution

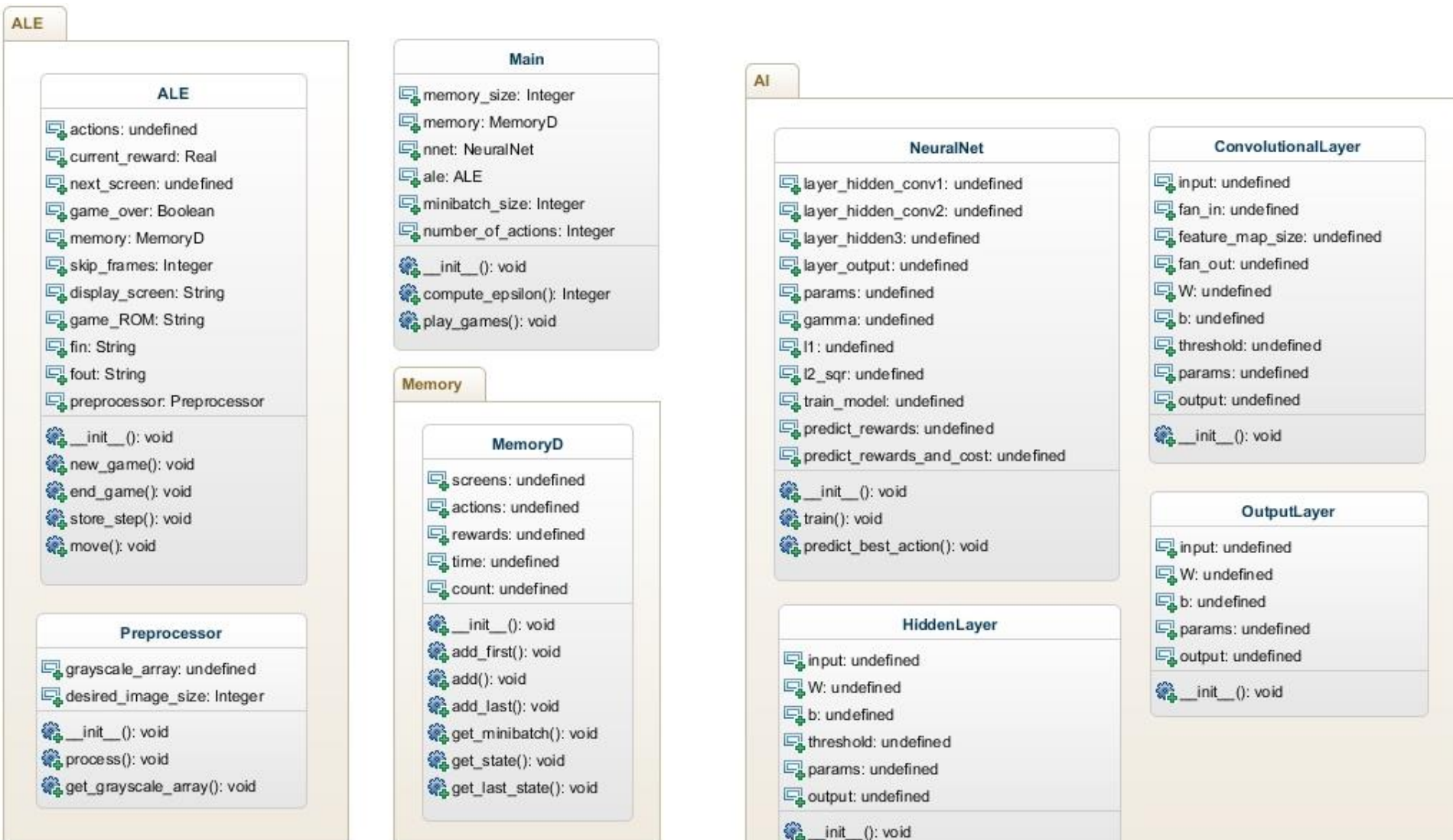
# Fully connected



# Structure of the DeepMind

- Play lots of games
  - Play lots of frames
    - Get a frame and pre-process it
    - Choose action using Deep Neural Net (DNN) and current state
      - Do it randomly with probability  $\epsilon$
    - Observe reward and next image and save everything to the memory
    - Train the network
      - 32 random transitions: 32 x  $(state, reward, action, state)$
      - Q-learning
        - $Reward = r_t + \gamma \cdot DNN(State_{t+1})$
      - Gradient descent

# Did we achieve our goal? Almost yes!



# What next?

- Need to understand and implement one algorithm
- Need to convert some stuff from CPU to GPU
- Need to parallelize some of the code
- Probably need to optimize some parameters

# Replicating DeepMind

Computational neuroscience project by  
Ardi, Ilya, Kristjan and Taivo

Tartu, Estonia, 2014/05/25