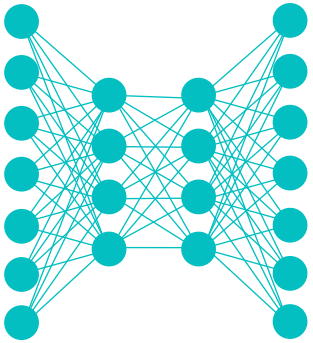
Lecture Notes for

Neural Networks and Machine Learning

Deep-Q Demo



Logistics and Agenda

- Logistics
  - Final paper, OK?
- Agenda
  - Student Presentations
    - SAC
    - AlphaFold
  - Deep Q-Learning Demonstration

Paper Presentation

Jürgen Schmidhuber
@SchmidhuberAI

Big news #AlphaFold uses #DeepLearning for #ProteinFolding prediction. This approach was pioneered by Sepp Hochreiter et al. in 2007 when compute was 1000 times more expensive than today. Their LSTM was orders of magnitude faster than the competitors.

PubMed.gov

Fast model-based protein homology detection without alignment - Pub...

pubmed.ncbi.nlm.nih.gov

Deep Q-Learning Reinforcement Learning

M. Lapan Implementation for Frozen Lake and Atari!

$$Q(s,a) \leftarrow (1-\alpha) \cdot Q(s,a) + \alpha \cdot [r_{s,a} + \gamma \max_{a' \in A} Q(s',a')]$$

$$\mathscr{L} = \Big[Q(s,a) - [r_{s,a} + \gamma \max_{a' \in A} Q^*(s',a')] \Big]^2$$

from current network params

from older network params (better stability)

$$\mathscr{L} = \big[Q(s,a) - [r_{s,a}] \big]^2$$

if no next state (env is done)

```
Follow Along:
08a_Basics_Of_Reinforcement_Learning.ipynb
```

World Models

The Problem

World Models

Can agents learn inside of their own dreams?

DAVID HA
Google Brain
Tokyo, Japan

JÜRGEN SCHMIDHUBER
NNAISENSE
Swiss AI Lab, IDSIA (USI & SUPSI)

March 27
2018

NIPS 2018
Paper

YouTube
Talk

Download
PDF

https://worldmodels.github.io

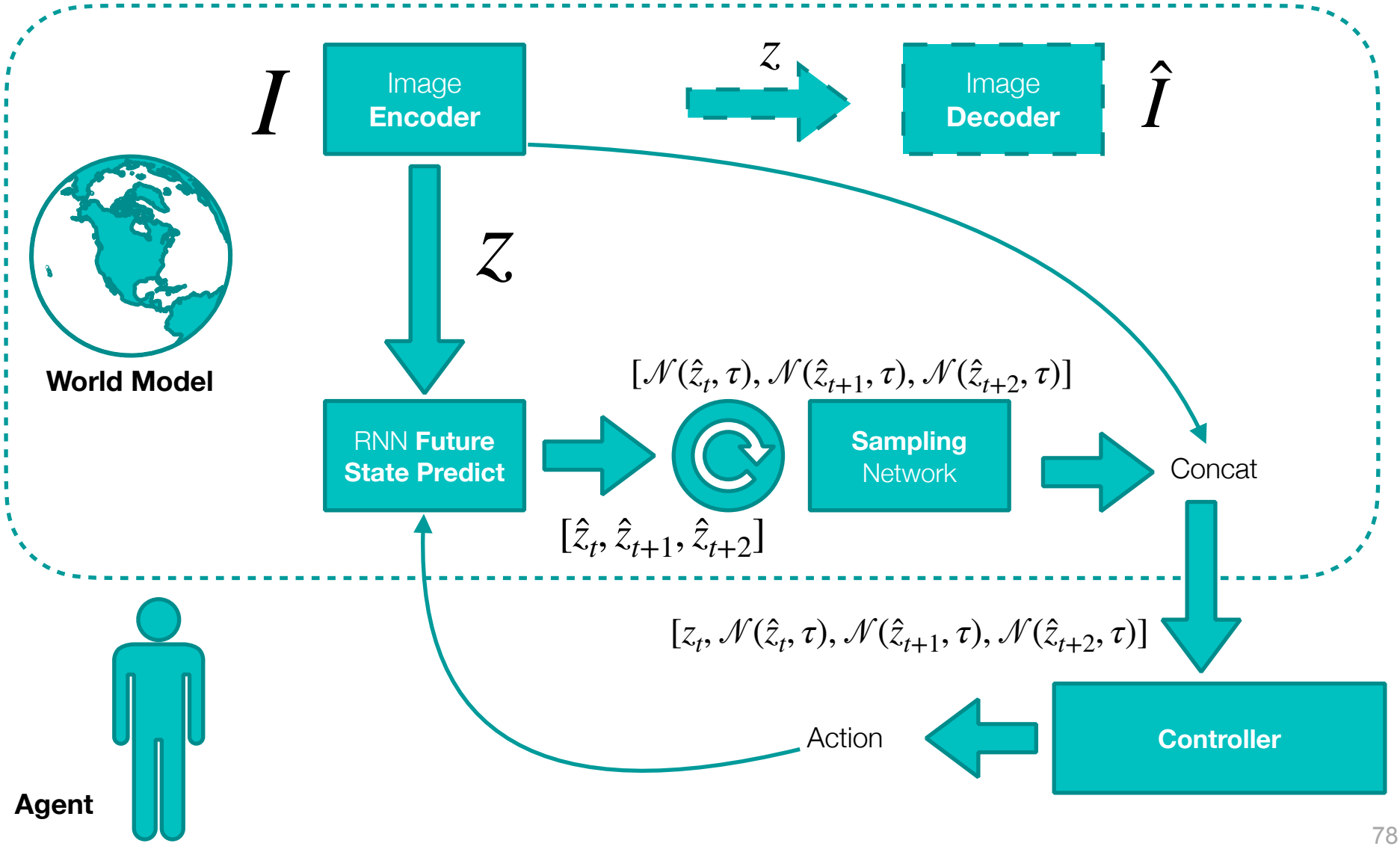
A Motivation

Agents can dream! What a time to be alive!

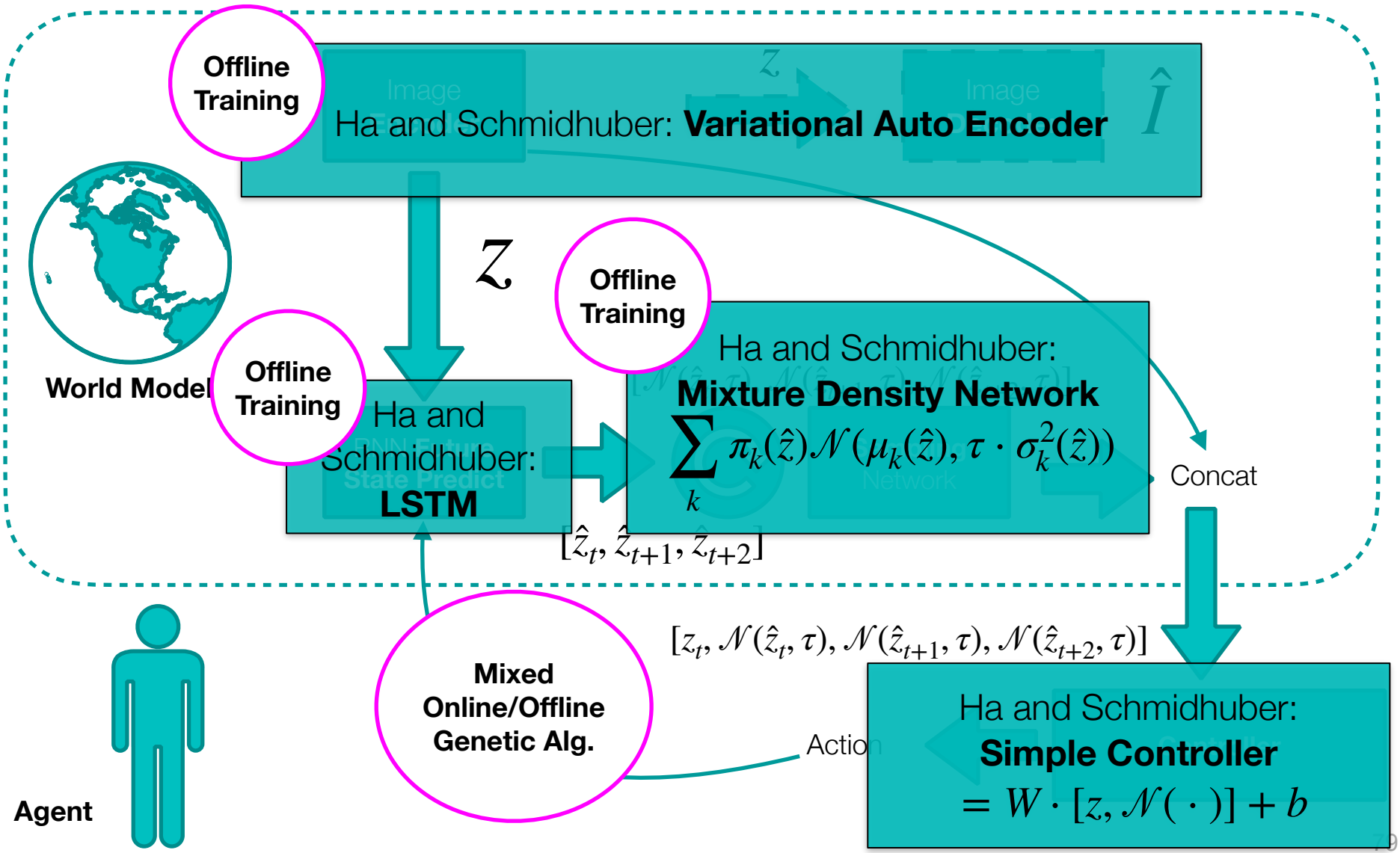
And academia can dream about driving the hype train!

Maybe we should be more careful about the way we describe what an agent does… because they don't dream. That's fluff.

The Main Idea



Implementation

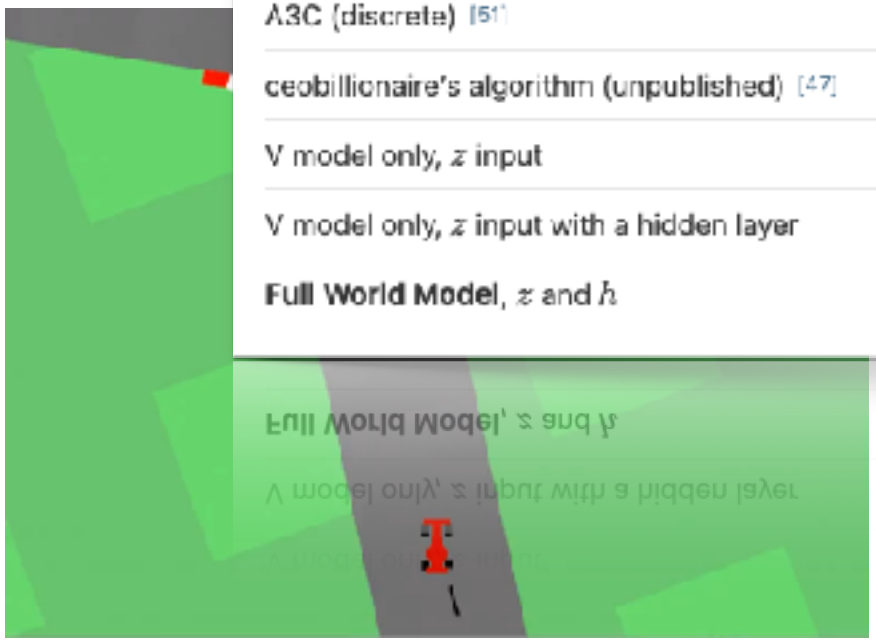


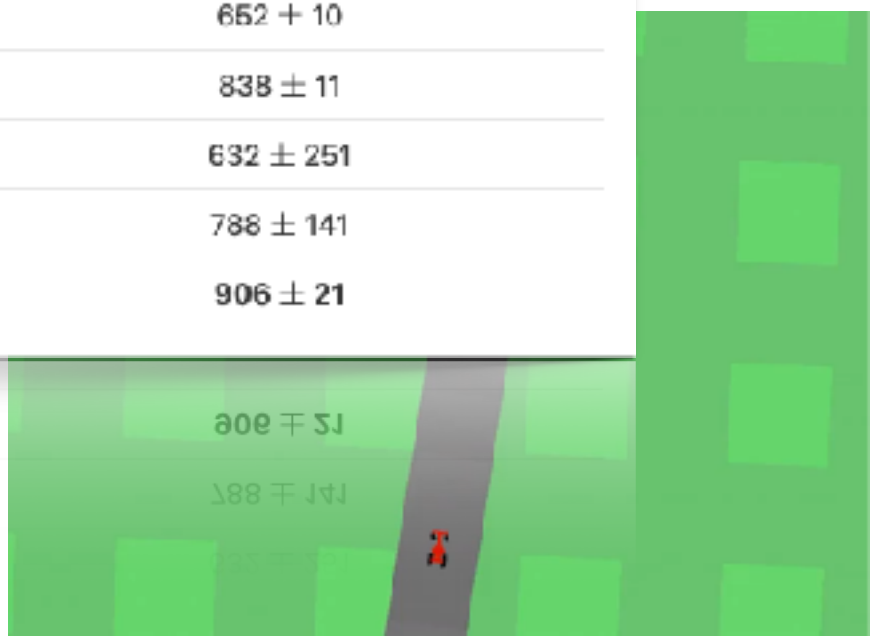
An Example, Racing

- Schmidhuber and Ha Methods:
  - Collect 10,000 rollouts from a random policy.
  - Train VAE (V) to encode each frame
  - Train
  - Evol
    cum

Model	Parameter Count
VAE	4,348,647
	422,368
	867

Method	Average Score over 100 Random Tracks
DQN [53]	343 ± 18
A3C (continuous) [52]	591 + 45
A3C (discrete) [51]	652 + 10
ceobillionaire's algorithm (unpublished) [47]	838 ± 11
V model only, z input	632 ± 251
V model only, z input with a hidden layer	788 ± 141
Full World Model, z and h	**906 ± 21**



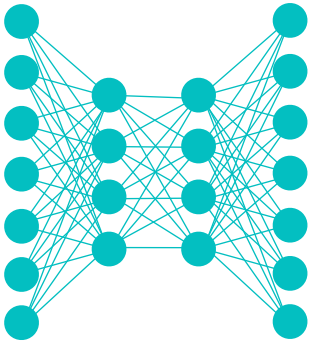


Only use VAE Encoding https://worldmodels.github.io Full World Model

Lecture Notes for

Neural Networks and Machine Learning

Deep Q Learning

Next Time:
World Models

Reading: None