Artificial Intelligence Reading Club

Chapter06: Deep Reinforcement Learning as Foundation for Artificial General Intelligence

Hao ZHAN

haozhan1993@gmail.com

11/09/2020

Table of Contents

1.Introduction

AGI

A fundamental distinction between Artificial General Intelligence (AGI) and "conventional" Artificial Intelligence (AI) is that AGI focuses on the study of systems that can perform tasks successfully across different problem domains, while AI typically pertains to domain-specific expert systems.

1.Introduction

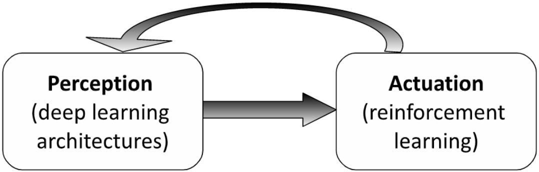
General problem-solving ability is one that humans naturally exhibit.

AGI

A related capability is generalization, which allows mammals to effectively associate causes perceived in their environment with regularities observed in the past.

Another critical human skill involves decision making under uncertainty, tightly coupled with generalization since the latter facilitates broad situation inference

Bipartite AGI architecture comprising of a perception and control/actuation subsystem



- Receive observations from the environment
- Infer current state of the world

- Map inferred states to actions
- Critique these actions based on prospect of rewards

Structured

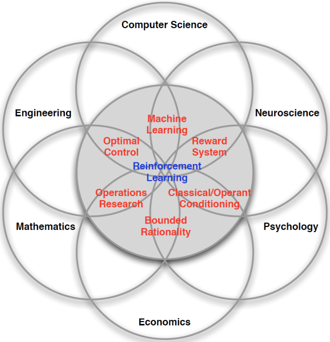
6.2 Review deep learningarchitectures and motivate their role in designing perception engines.

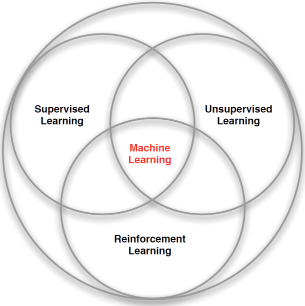
6.3 Review reinforcement learning and outlines how it can be merged with deep learning architectures.

6.4 discusses the scalability implications of designing AGI systems using emerging neuromorphic technology, while in

6.5 conclusions are drawn and future outlook is discussed.

2.Reinforcement Learning





2.Reinforcement Learning

What makes reinforcement learning different from other machine learning paradigms?

- There is no supervisor, only a reward signal
- Feedback is delayed, not instantaneous
- Time really matters (sequential, non i.i.d data)
- Agent's actions affect the subsequent data it receives

2.Reinforcement Learning

What makes reinforcement learning different from other machine learning paradigms?

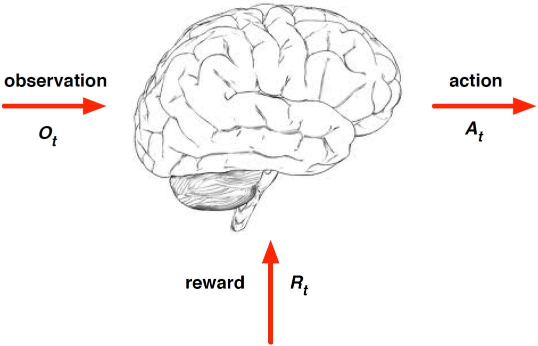
Fly stunt manoeuvres in a helicopter
+ve reward for following desired trajectory
-ve reward for crashing

Defeat the world champion at Backgammon
+-ve reward for winning/losing a game

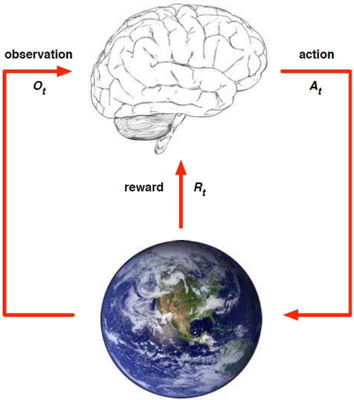
Manage an investment portfolio
+ve reward for each dollar in bank

Control a power station
+ve reward for producing power
-ve reward for exceeding safety thresholds

2.Reinforcement Learning



2.Reinforcement Learning



- At each step t the agent:
  - Executes action A_t
  - Receives observation O_t
  - Receives scalar reward R_t
- The environment:
  - Receives action A_t
  - Emits observation O_{t+1}
  - Emits scalar reward R_{t+1}
- t increments at env. step

2.Reinforcement Learning

- The history is the sequence of observations, actions, rewards

$$H_t = O_1, R_1, A_1, ..., A_{t-1}, O_t, R_t$$

- i.e. all observable variables up to time t
- i.e. the sensorimotor stream of a robot or embodied agent
- What happens next depends on the history:
  - The agent selects actions
  - The environment selects observations/rewards
- State is the information used to determine what happens next
- Formally, state is a function of the history:

$$S_t = f(H_t)$$

2.Reinforcement Learning

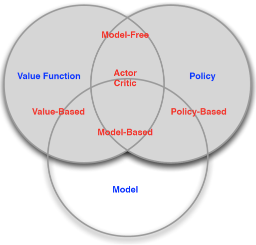
- A model predicts what the environment will do next
- $\mathcal{P}$ predicts the next state
- $\mathcal{R}$ predicts the next (immediate) reward, e.g.

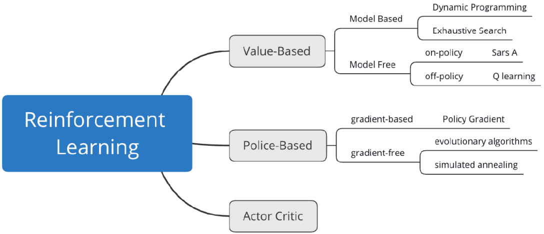
$$\mathcal{P}_{ss'}^{a} = \mathbb{P}[S_{t+1} = s' \mid S_t = s, A_t = a]$$
$$\mathcal{R}_{s}^{a} = \mathbb{E}\left[R_{t+1} \mid S_t = s, A_t = a\right]$$

2.Reinforcement Learning

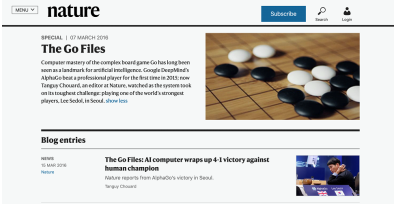
- Value Based
  - No Policy (Implicit)
  - Value Function
- Policy Based
  - Policy
  - No Value Function
- Actor Critic
  - Policy
  - Value Function

- Model Free
  - Policy and/or Value Function
  - No Model
- Model Based
  - Policy and/or Value Function
  - Model





2.Reinforcement Learning


MENU nature Subscribe Search Login

SPECIAL | 07 MARCH 2016

The Go Files

Computer mastery of the complex board game Go has long been seen as a landmark for artificial intelligence. Google DeepMind's AlphaGo beat a professional player for the first time in 2015; now Tanguy Chouard, an editor at Nature, watched as the system took on its toughest challenge: playing one of the world's strongest players, Lee Sedol, in Seoul. show less

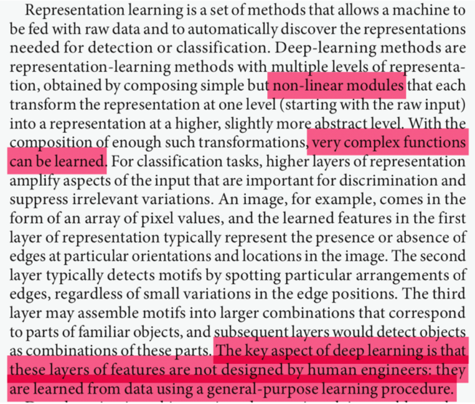
Blog entries

NEWS
15 MAR 2016
Nature

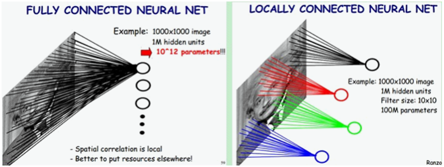
The Go Files: AI computer wraps up 4-1 victory against human champion

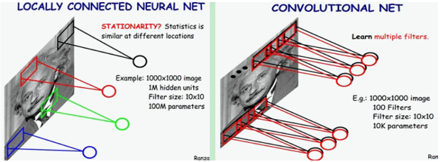
Nature reports from AlphaGo's victory in Seoul.

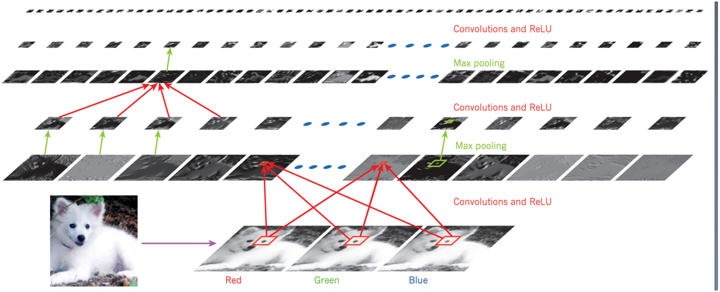
Tanguy Chouard


Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification. Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned. For classification tasks, higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations. An image, for example, comes in the form of an array of pixel values, and the learned features in the first layer of representation typically represent the presence or absence of edges at particular orientations and locations in the image. The second layer typically detects motifs by spotting particular arrangements of edges, regardless of small variations in the edge positions. The third layer may assemble motifs into larger combinations that correspond to parts of familiar objects, and subsequent layers would detect objects as combinations of these parts. The key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure.

2.Reinforcement Learning







Thank you for your time!