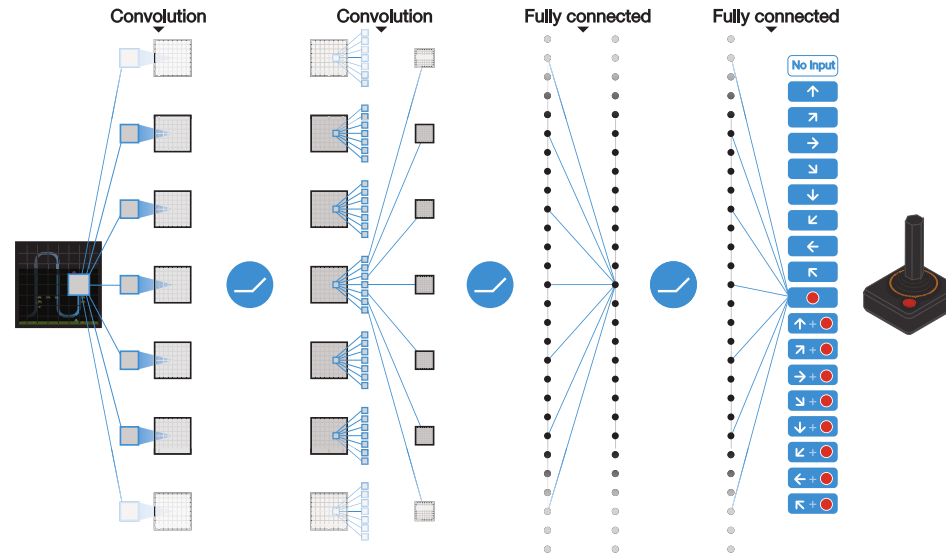


# 1 複雑な状況をどう理解して解決するのか？

---

- 強化学習というニューラルネットワークモデルがあるわけではない
- 動的で複雑な環境に対処 → **強化学習** + DL → 一般人工知能への礎
  - DQN ATARIのビデオゲーム, <https://www.nature.com/articles/nature14236>
  - AlphaGo 囲碁, <https://www.nature.com/articles/nature16961>
  - AlphaGoZero 囲碁, <https://www.nature.com/articles/nature24270>

# 2 Deep Q Network



Mnih et al. (2015) より

- **Q 学習** Q learning に DNN を採用
- CNN が LeNet, LeCun et al. (1998) 同様であったように, 強化学習 RL も昔からの技術 Sutton and Barto (1998)
- ではなぜ, 今になって囲碁や自動運転に応用できるようになったのか?
  - ⇒ コンピュータの能力, データ規模, アルゴリズムの改良, エコシステム(ArXiv, Linux, Git, ROS, AMT, TensorFlow)

# 3 強化学習

---

- 強化学習 ⇒ 意思決定
  - エージェント agent が 行動(行為) action をする
  - 行動によって 状態 が変化する
  - 環境 から与えられる 報酬 によって 目標 が決定
- 深層学習: ⇒ 表現, 表象
  - 教師信号として目標が与えられる
  - 目標を達成するために外部状況の 表現 を獲得

強化学習 + 深層学習 = 人工知能

- 強化学習 ⇒ 目標の設定
- 深層学習 ⇒ 内部表象の獲得機構を提供

## 4 用語の整理

---

- 教師信号なし **報酬信号** reward signal
- 遅延フィードバック
- **価値** Value
- **行為** Action
- **状態** State
- TD 学習
  - **Sarsa**
  - **Q 学習**
  - **アクタークリティック**
- 報酬  $R_t$ : **スカラ値**
  - 時刻  $t$  でエージェントの行った行動を評価する指標
  - エージェントは**累積報酬** cumulative reward の最大化する
  - 報酬仮説: **目標は累積期待報酬の最大化として記述可能**

# 5 デモ

---

- ブロック崩し: <https://www.youtube.com/watch?v=V1eYniJ0Rnk>
- スペースインベーター: <https://www.youtube.com/watch?v=W2CAghUiofY>
- OpenMind selfplay: <https://www.youtube.com/watch?v=OBcjhp4KSgQ>

## 6 教科書, 参考文献など

---

- An Introduction to Reinforcement Learning, Sutton and Barto, 1998, MIT Press, 1998
  - <http://incompleteideas.net/book/the-book.html>, Sutton and Barto (1998) (2018年第2版出版予定)
  - 翻訳 強化学習 <https://www.amazon.co.jp/dp/4627826613>
- Algorithms for Reinforcement Learning, Szepesvari, Morgan and Claypool, 2010
  - <https://sites.ualberta.ca/~szepesva/papers/RLAlgsInMDPs.pdf>
- デービッド・シルバーの講義 <http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching.html>
- ジョン・シュルマンのビデオ講義 <https://www.youtube.com/watch?v=oPGVsoBonLM>
- 「これからの強化学習」 <https://www.amazon.co.jp/dp/4627880316/>

# 文献

---

LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. 1998. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86: 2278–2324. <https://doi.org/10.1109/5.726791>.

Mnih, Volodymyr, Korya Kavukchuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, et al. 2015. "Human-Level Control Through Deep Reinforcement Learning." *Nature* 518: 529–33. <https://doi.org/10.1038/nature14236>.

Sutton, Richard S., and Andrew G. Barto. 1998. *Reinforcement Learning*. Cambridge, MA: MIT Press.