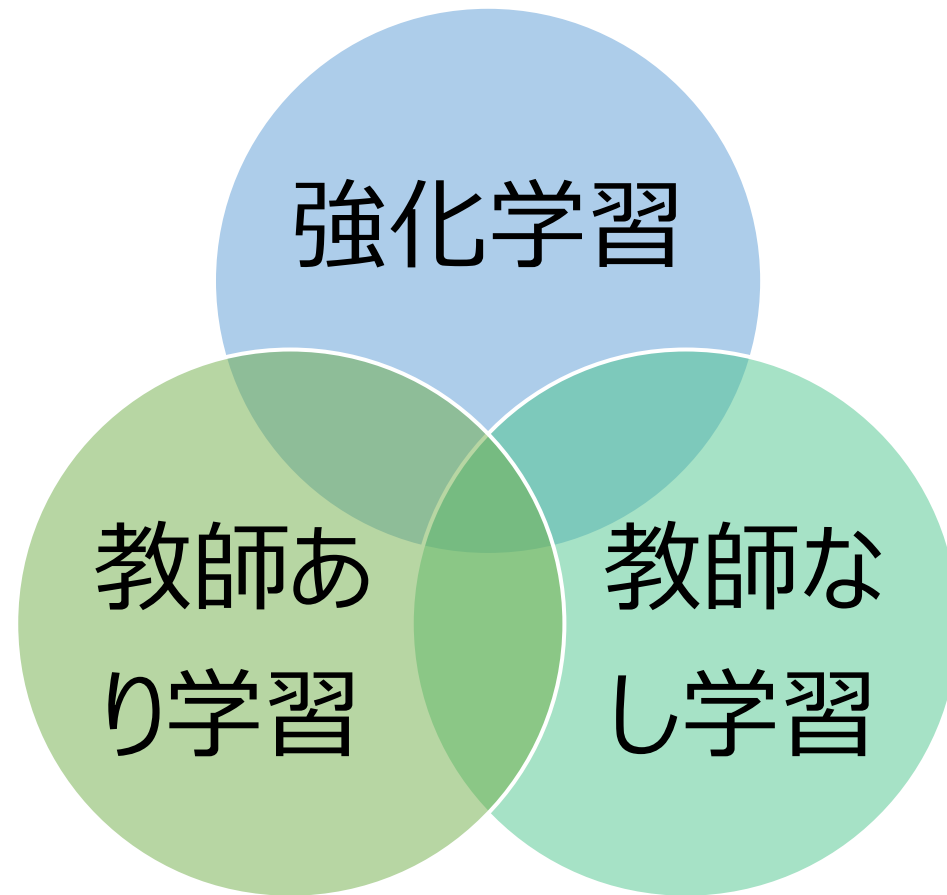


自然言語処理 — 強化学習入門 —

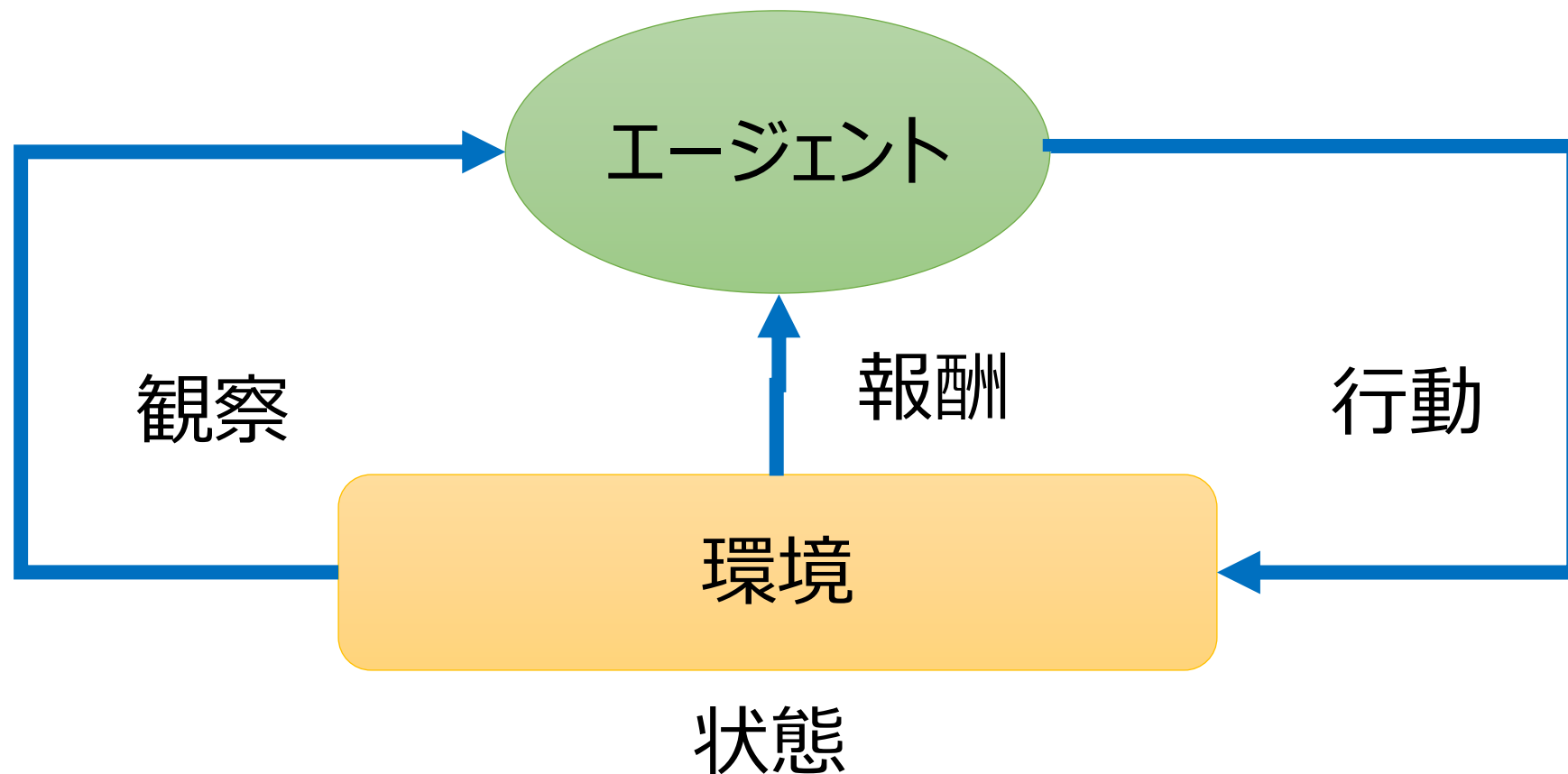
<https://satoyoshiharu.github.io/nlp/>

機械学習

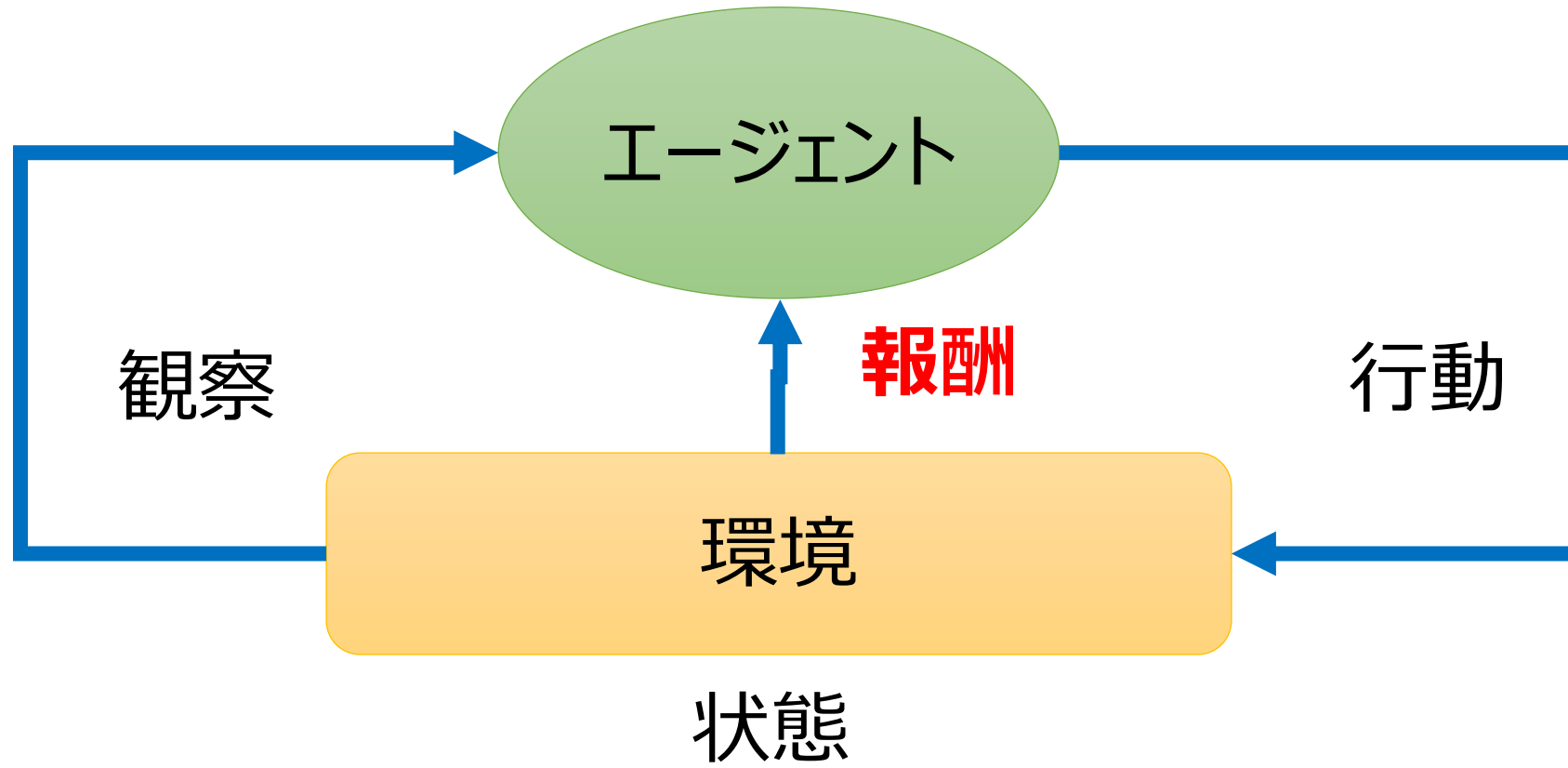


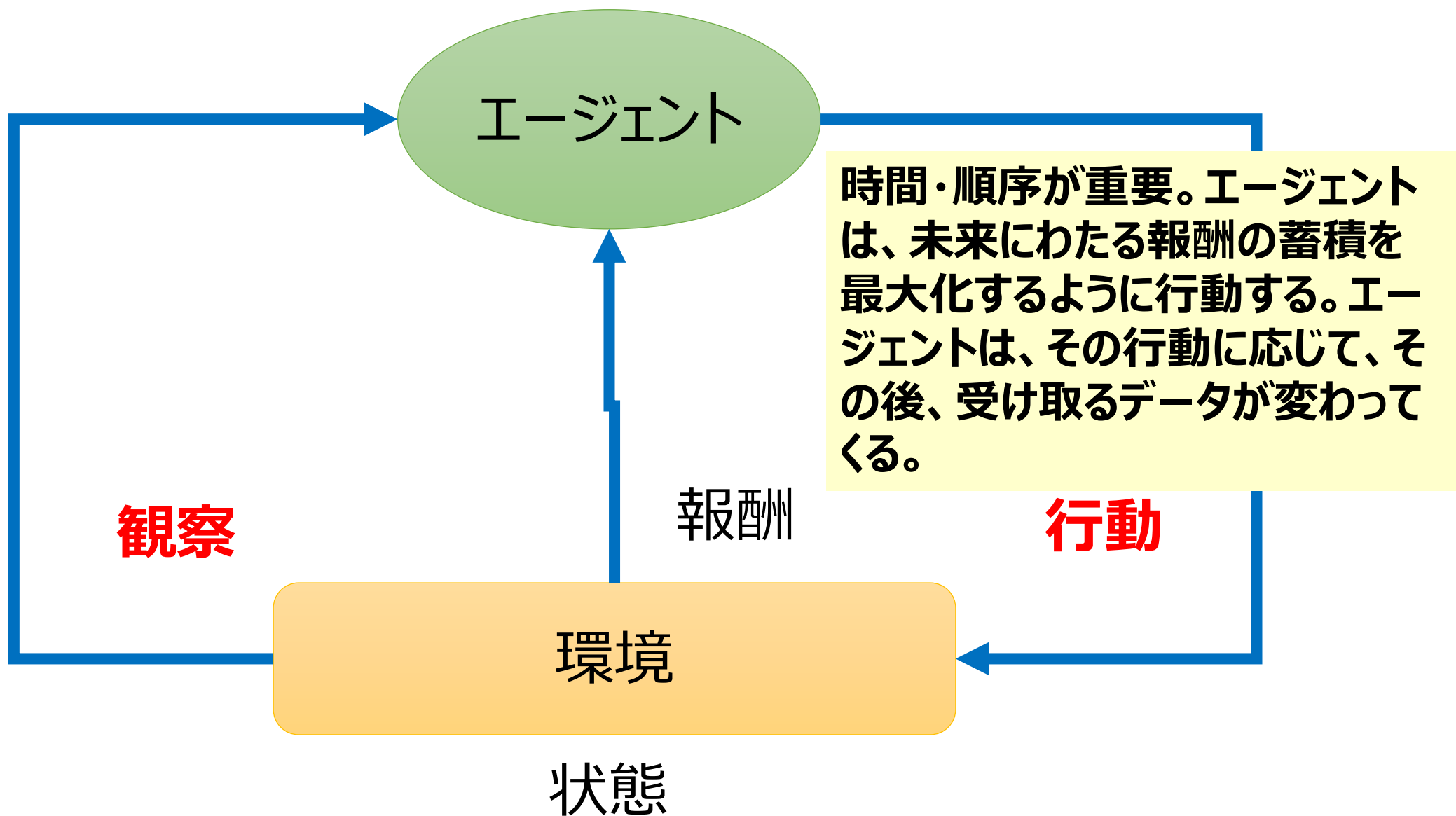
強化学習の基本的な考え方

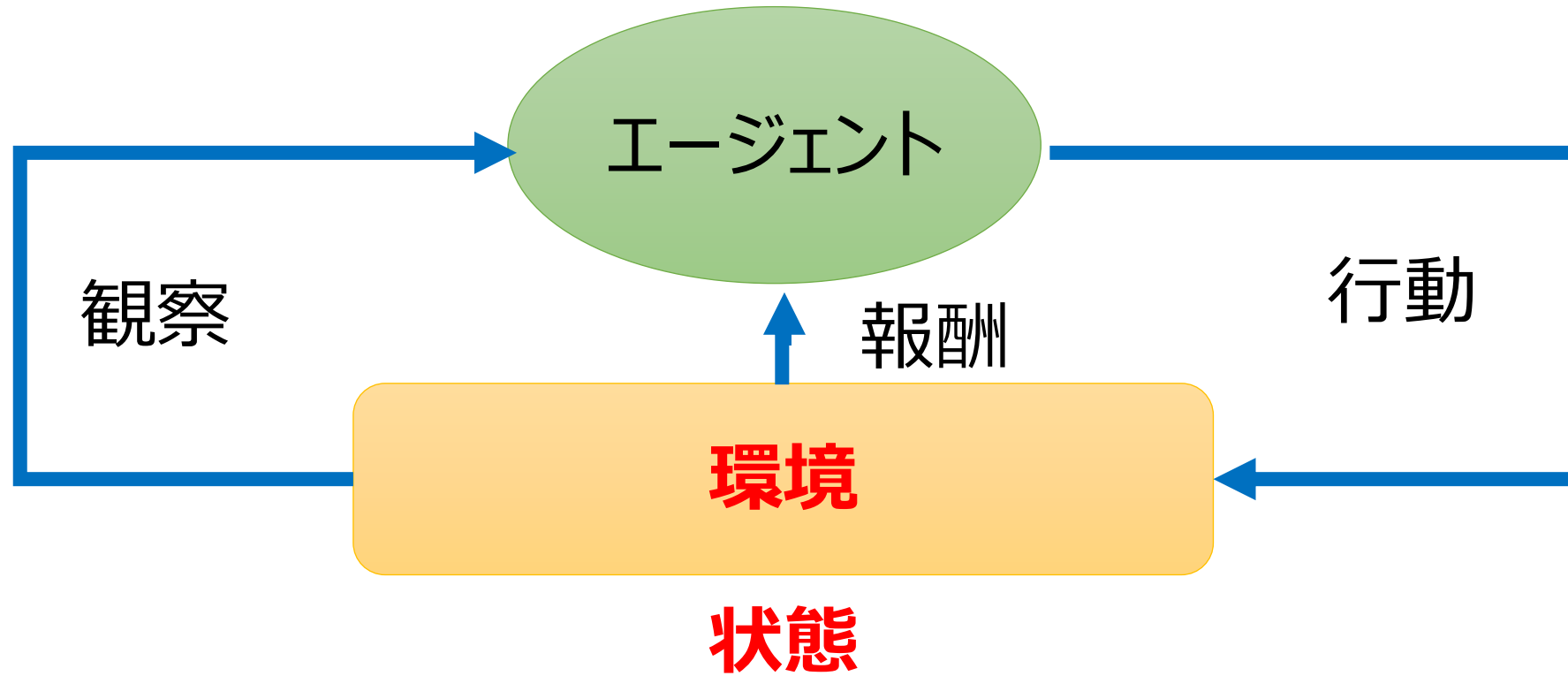
エージェントが環境に働きかけて学習する



教師なし、報酬信号（ある時刻でのスカラー値）だけがある。即時の報酬でなく、未来にわたる報酬の蓄積を気にする。それは後でわかる。例：棒が倒れたかそうでないか、ゲームに勝ったか負けたか。







- 状態が、履歴から独立して、直前の状態だけで決まる場合、マルコフ的という ($P[S_{t+1}|S_t] = P[S_{t+1}|S_1, \dots, S_t]$)。強化学習は、マルコフ状態を前提としたマルコフ決定理論として定式化されている。
- 多くのエージェントは環境の知識(モデル)を持たない。ゲームへの応用ではゲームの規則さえ知らずにいいパフォーマンスを出す。膨大なプレイデータで見様見真似している。

強化学習の構成要素

ポリシー（方策）

- エージェントの行動を決める

価値

- ある状態ないし行動がどれだけいいか

モデル

- エージェントが環境を表現する

エージェントのタイプ

状態の価値の漸化式的な関係（ベルマン方程式）を利用する

- 代表的なアルゴリズム：Deep Q-Learning
- 応用例：Googleのレコメンド
- 日本語の解説記事は、圧倒的にこちらが多い。

ポリシーを直接最適化する

- 代表的なアルゴリズム：ポリシー・グラディエント法
- 応用例：AlphaGo
- 価値ベースよりこっちのほうが多いという人が多い。

Actor Critic（価値とポリシー両方を最適化する）

- 研究の主流はいまここみたい。

強化学習はニューラルネットを利用

ニューラ
ルネット

価値ネットワークやポ
リシーネットワークとし
て、関数を近似する

強化学
習

経験から学ぶという
考え方・道具の枠組
み

強化学習全般に関する入門的資料：おすすめは以下の二つ

- [A friendly introduction to deep reinforcement learning, Q-networks and policy gradients, by Serrano.Academy](#)
 - 英語、YouTube動画
 - 内容が入門的。ベルマン方程式、価値(Q)ネットワーク、ポリシーネットワーク、ポリシー・グラディエントに関して、数式なしで、アニメーションで説明していて、非常にわかりやすい。
- [作りながら学ぶ強化学習 -初歩からPyTorchによる深層強化学習まで、by 小川雄太郎](#)
 - コードサンプルをいじりながらの概念解説なので、数式も頭に入りやすい。
 - 第4回～第8回：Windows+Chrome+TryJupyterだとWindowsへのffmpegのインストールが面倒。Google Colab（多分Ubuntu）でやったほうが楽。

ポリシー・グラディエントに関する入門的解説

- [Deep Reinforcement Learning: Pong from Pixels](#)、by Andrej Karpathy
 - 英語
 - ポリシー・グラディエントに関し、わかりやすいと評判のブログ解説
 - 教師あり学習の変形という説明なので、ニューラルネット方面からは入りやすい
 - ポリシー・グラディエント定理の説明もわかりやすい
 - YouTube版：<https://www.youtube.com/watch?v=tqrcjHuNdmQ>

Policy Gradient サンプルコード

- 動作確認済み
 - [作りながら学ぶ！強化学習、第5回](#)
 - [pg-pong.py](#) by Andrej Karpathy のPython3版 -> https://colab.research.google.com/drive/1BTiGjP_FD0PdYfazpn61nWPC0YJuZPOE?usp=sharing
 - ROMエラーは、以下で回避
 - `%pip uninstall atari-py`
 - `%pip install atari-py==0.2.5`

Deep Q-Learningに関する入門的解説

- [ゼロからDeepまで学ぶ強化学習](#)
 - Deep Q-Learningの考え方、実装テクについて、わかりやすい。

Deep Q-Learning サンプルコード

- 動作確認済み

- [PyTorch「強化学習（DQN）チュートリアル」サンプルコード](#)
- [作りながら学ぶ！強化学習、第6,7,8回](#)

- 未確認

- [作りながら学ぶ！強化学習、第9,10回](#)
- [作りながら学ぶ！強化学習、第15回](#)
- [【入門】Q学習の解説とpythonでの実装 ～シンプルな迷路問題を例に～](#)
- https://github.com/icoxfog417/chainer_pong

強化学習に関する解説資料（研究者向き）

- [Tutorial: Deep Reinforcement Learning](#), by David Silver、Google DeepMind
 - 英語の66ページスライド
 - 実際のシステム例の説明は高度、ある程度なじんだあとでないと難しい。それ以外、前半から半ばまでは、入門者向きで、コンパクトにわかりやすくまとめられている。
- [ゼロから始める深層強化学習](#), by Preferred Networkの人
 - 前半、前田さんの部分：数式を使った説明なので、最初はわかりにくい。ある程度概念になじんだ後で読むと、要するに・・・みたいな解釈の部分が、納得できる。
- [Reinforcement Learning: An Introduction](#) , by Sutton & Barto
 - 英語、フリーで読める
 - この分野の古典、標準教科書
- [Introduction to Reinforcement Learning with David Silver](#)
 - 英語
 - Suttonさんの標準教科書相当のトピックをDavid Silverさんが再構築した講義録
- [深層強化学習アルゴリズムまとめ](#)
 - 網羅的な整理だが、細っつい。自分で課題を解く段階になったら、参照すべき。