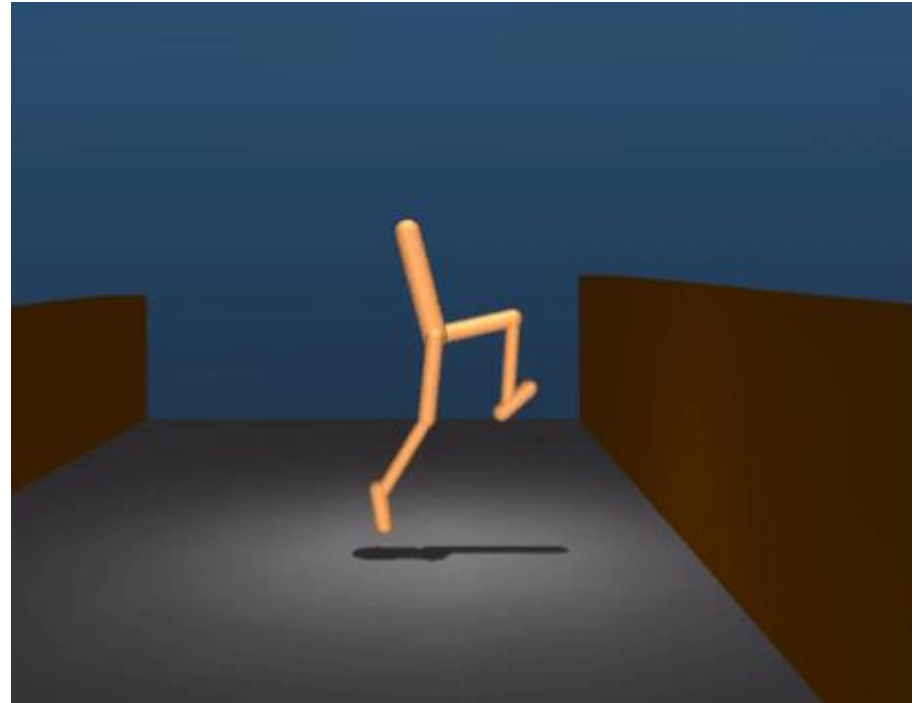


Deep Reinforcement Learning

- ゲームや囲碁では人間を上回るパフォーマンス
e.g. AlphaGo Deep-Q-Network
- 効率的な動きの学習
- ロボットでの応用
→最近の流行り



Q学習について

- 強化学習では、その行動の結果に応じて**報酬**をもらうことによって、実行した行動について自己評価しながら学習する仕組み
- この自己評価の計算を「**Q関数**」を用いて学習する

ここでQ学習には、「**学習率**」と「**割引率**」というパラメータを自分で設定してあげる必要があります。それぞれざっくりと以下のようになります。

- **学習率** (α): 報酬に対してどれだけ敏感に反応するか
- **割引率** (γ): 過去の経験と将来の予測、それぞれの重要度を調整

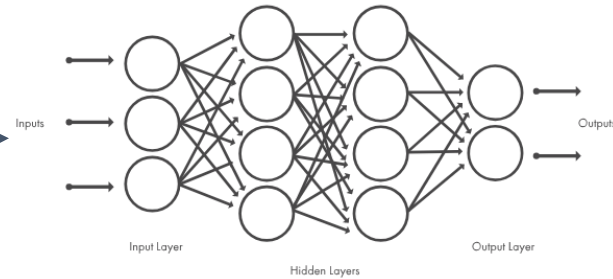
$$Q^*(s_t, a_t) = \mathbb{E}_{\pi^*} \left[r_t + \sum_{k=t+1}^{\infty} \gamma^{k-t} (r_k + \alpha \mathcal{H}_k^{\pi^*}) \right]$$



DQN - 深層Q学習

ディープラーニングを使って、学習率や割引率といったパラメータも自動的に計算できるようになる

$$Q^*(s_t, a_t) = \mathbb{E}_{\pi^*} \left[r_t + \sum_{k=t+1}^{\infty} \gamma^{k-t} (r_k + \alpha \mathcal{H}_k^{\pi^*}) \right]$$



Q関数



ニューラルネットワーク

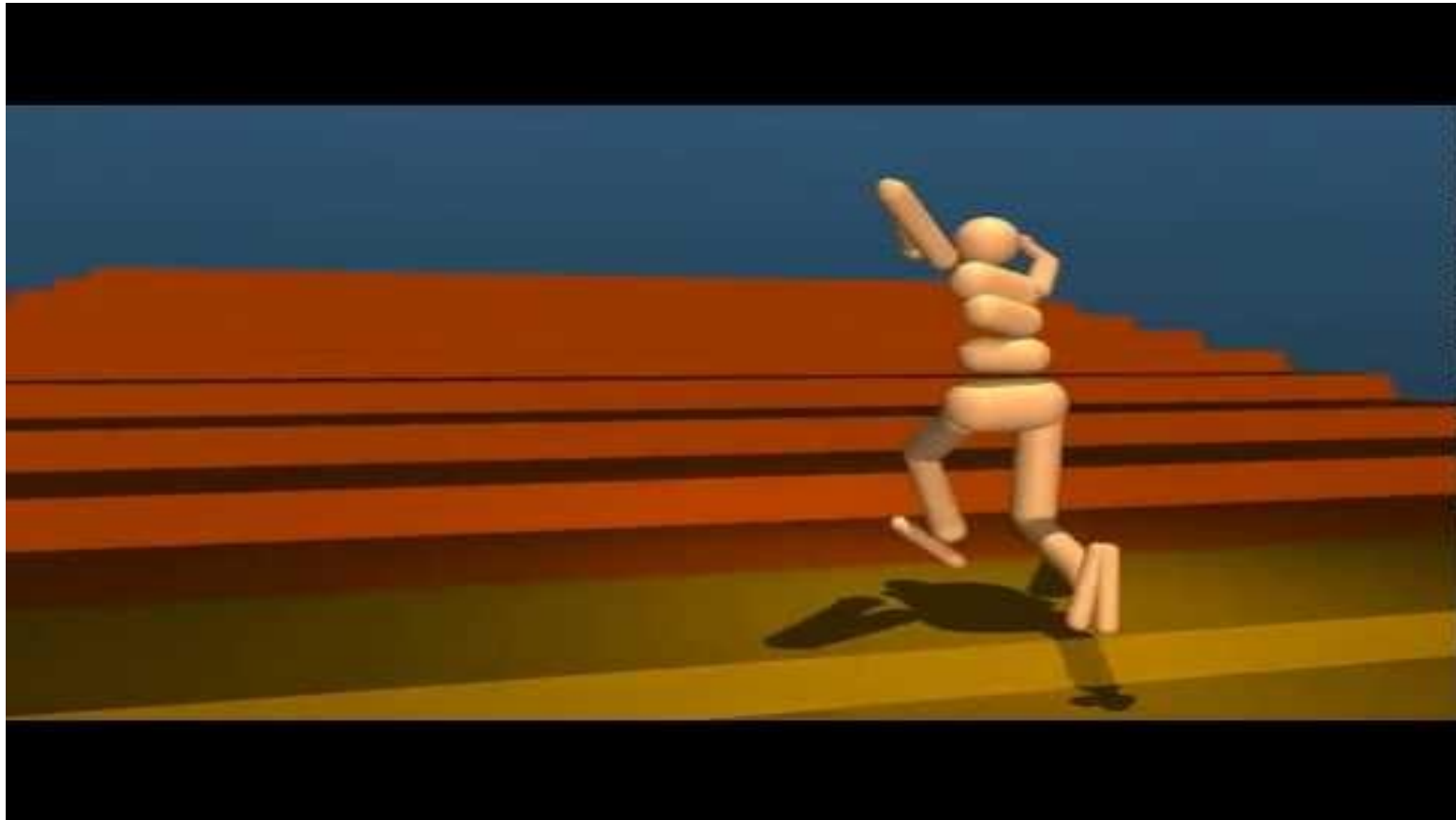
Deep Reinforcement Learning

- ゲームでの応用



Deep Reinforcement Learning

- ロボットでの応用



Deep Reinforcement Learning

- ロボットでの応用



Deep Reinforcement Learning

- ロボットでの応用



Deep Reinforcement Learning

- 服を着る



Deep Reinforcement Learning

- 把持

