

S⁴ Simulation System 活用事例



強化学習とは？ 機械学習との違いや製造業における具体例

Interview

強化学習とは

強化学習の概要、教師あり学習との違いを教えてください。

松下 強化学習は機械学習の分野の1つです。機械学習の中にはいくつかのタスクの枠組みがありますが、多くの方が典型的にイメージするのは教師あり学習ではないでしょうか。教師あり学習ではあらかじめ答え（正解ラベル）が付与された問題（特徴量）を大量に学習させて、新たな問題が入力されたときに答えを出力できるAIを作成することを目指します。

強化学習も何らかの正解に向けて判断をするAIを作るという意味では似ているのですが、正解に対する考え方に特徴があります。強化学習では、AI（エージェント）に何らかの選択をさせたときにその行動に対する評価（報酬）を与えて、評価が大きくなるような行動の仕方を学習させていきます。教師あり学習の正解ラベルと異なり、報酬はある1つの行動に対して即座に与えられる必要はなく、状況に応じて何度かの行動を行った結果に対して与えられる形でかまいません。

例えば将棋の次の一手を考える場合、教師あり学習的な発想で盤面状況（特徴量）に対して最適な一手（正解ラベル）を人間が付与して学習させるというのは非常に困難です。素人には一見悪手に見える手が、勝利につながる最善手であるというのは珍しいことではありません。トッププロ同士の対局では解説のプロ棋士でもその一手が良い手なのかどうかの判断に悩むこともあります。一方で、局面が進んだある盤面に詰みが成立しているかどうかは、ルールさえわかれば簡単に判断して情報として与えることが可能です。この場合、強化学習で最終盤面での勝敗を報酬として設定することで、短期的な形勢判断に囚われず最終的な勝利に向けて打ち手を選ぶという学習を行うことができます。

現時点で手に入っている情報だけで判断ができるという場合は教師あり学習を適用するケースが多いと思いますが、時間経過とともに変化する環境に応じて適切な判断を繰り返さないと良い結果が得られないというような問題設定は強化学習の使いどころです。囲碁や将棋のようなターン制のゲームや、この後ご紹介する生産工程の自動化はまさに強化学習向きの課題です。



NTTデータ数理システム
シミュレーション&マイニング部
松下 亮祐

強化学習・シミュレーションに関する技術課題に対して、論文調査からモデル開発やプログラミングまでを一貫して担当。お客様の課題に対して、手法を実装し考察を繰り返す、実験系（試行錯誤をしながら課題を追及すること）の仕事を中心に活動。

製造業の生産工程の強化学習の適用事例

製造業で強化学習を活用のご相談は増えていますか？

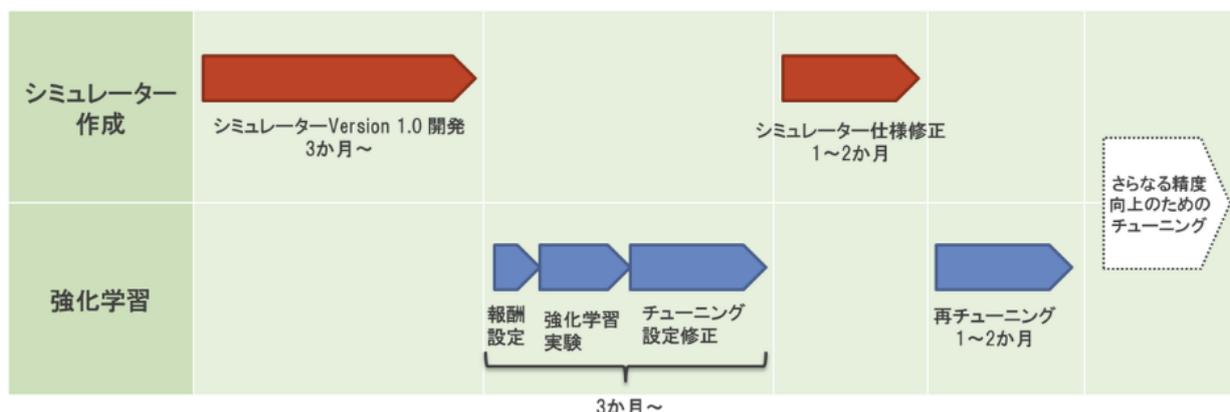
松下 インダストリー4.0 やFA(Factory Automation)の流れで生産工程を自動化したいというご相談は多く、強化学習のご提案をする機会も増えています。今でも将棋や囲碁の例を引き合いに、手法として強化学習を候補にあげながらご相談いただくことも多いです。やはり、プロの棋士達と互角かそれを超える強さを示したのは社会的にも大きなインパクトだったのだなと感じます。

生産工程を例として、具体的な強化学習活用プロジェクトの進め方について

松下 多くの強化学習活用プロジェクトは第一段階でシミュレーターを作るところから始まります。強化学習は観測されたデータを用いた学習も可能ですが、データが網羅されていないと学習の効率が落ちます。また、強化学習は、試行を繰り返して学習をしていくので試行回数を多くするには、効率が悪いです。現実には選択しないような選択肢を選択することで、その選択肢をとらないことも学習させる必要があります。そこで現実世界ではなく、シミュレーション上でさまざまな試行を繰り返すことで、効率よく学習させることができます。生産工程の場合は、工場の生産ラインのシミュレーションモデルを実装します。

実際にシミュレーションモデルを構築する際には、生産ラインの仕様を伺ってシミュレーションモデルに落とし込んでいきますが、各処理工程や工程間の移動に要する時間などのデータも必要になります。このシミュレーションモデルの構築は比較的ハードルが高い作業で、教師つきデータを用意すればスタートできる教師あり学習との相違点になるかなと思います。

シミュレーターが完成したらそれをブラックボックスな環境として強化学習モデルの学習を開始します。工場の生産工程などの場合は、納期通りに生産が完了することやスループットが向上することをうまい具合に報酬として定量化することによって学習を実現します。



開発工程イメージ

シミュレーションモデルの開発に用いるツールについて教えてください

松下 シミュレーションにおいては当社のシミュレーションソフト S⁴ Simulation System を用います。

S⁴ Simulation System にはシミュレーターとしての機能や生産ラインをモデリングするためのパーツや API が豊富に用意されているため、スクラッチで実装するのに比べて効率よく実装できます。(手前みそですが) 実装の工数などが大幅削減できて便利だなと思っています。強化学習でアプローチする際は、一般的に生産工程に限らず物理シミュレーションや社会シミュレーションなど、既存のシミュレーターが対応していればそれを使うのがベターですが、ない場合はシミュレーションプログラムをスクラッチで実装することもあります。

強化学習のアルゴリズムはかつては Chainer のライブラリを使用し学習は DQN (ディープラーニングと強化学習を組み合わせた手法) を用いて学習していました。最近ではライブラリが PyTorch に引き継がれているので、PFRL (PyTorch ベースの強化学習ライブラリ) を使用します。

S⁴ Simulation System は Python ベースのシミュレーターです。Python プログラムから API として呼び出すこともできますし、S⁴ Simulation System から Python の豊富なライブラリを呼び出すこともできます。強化学習をはじめ、Python の様々な機械学習ライブラリと連携させる場合に S⁴ Simulation System は親和性が高いシミュレーターですので、おすすめできます。

今後について

今後チャレンジしていきたい課題はありますか？

松下 技術的な観点ですと…、手掛けてきた多くの問題がシングルエージェントシミュレーションモデルで表現可能であったため、今後はマルチエージェントの強化学習にチャレンジしてみたいです。例えば製造業だとそれぞれタスク(目的)が違う複数のロボットを強化学習で行う場合には、マルチエージェントシミュレーションになります。それに限らず製造業での生産工程に関して、より実績を積んで皆様が気軽に強化学習を使えるように技術ナレッジを積み上げていきたいと思っています。