

進捗レポート (18T0006, 小室 光広)

今週の進捗サマリ

- テーマ再定義
- 実装技術選定

1 進捗詳細

1.1 テーマ再設定

従来より考えていた NLP タスクをベンチマークとするユーザー意図抽出モデル (Part of Leica model) は収束が難しいと判断した。修士論文を執筆する上で確実に時間が足りないと予想されるためである。したがって、本修士論文で扱うテーマは、従来の NLP タスクではなく、より簡単な Open AI Gym 環境における強化学習フレームワークの提案とする。

本研究では、従来の強化学習における局所解脱出性や、学習収束速度向上を目的として、強化学習フレームワークにおける Agent に Durable value をもたせることで試みる。本研究では以下に示す 2 つのモデルを設定した。本研究では Sync モデ

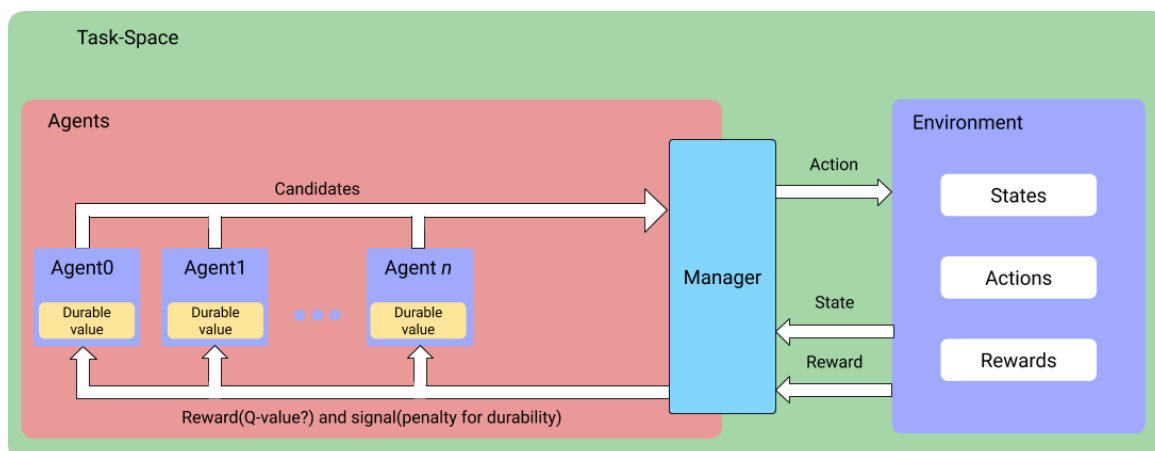


図 1: Sync モデル (仮称)

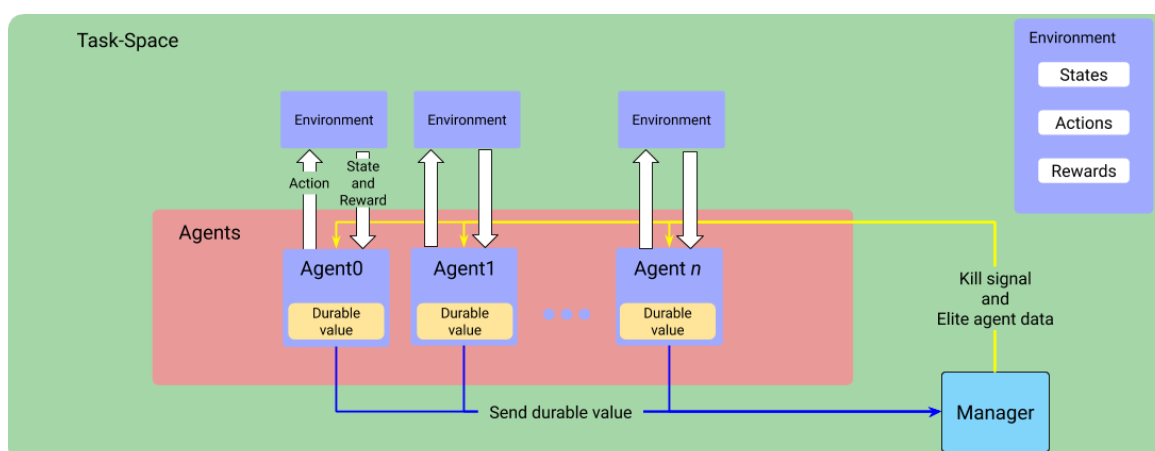


図 2: Async モデル (仮称)

ルを始めに対象として扱い、時間があれば Async モデルの検討を行う。

以下、Sync モデルについて詳細を述べる。Sync モデルにおいて各 Agent は Durable value を持っており、例えば各エピソードで同じ動作をすることや、想定しない動作に収束する局所解状態を維持することが見られた場合、その Agent が持つ

Durable value は減少する。ある閾値を超えて減少した場合その Agent は Kill または Swap される。Swap はその時点である評価手法で優秀と判断された Agent のオブジェクトをそのままコピーすることで実現する。

以上の Sync モデルにおいて、Agent は CNN 等のニューラルネットワークを使用し、DQN による Open AI Gym Atari によるゲームでその精度値を評価する。

1.2 実装技術選定

モデルを実装するに当たり、実装で必要となる以下のライブラリを調査した。

- TF-agent
- keras-rl
- Chainer-rl
- PyTorch

まず TF-agent は TensorFlow 2.0 beta 以降に追加となった強化学習向けライブラリである。本ライブラリに Environment Suite として Open AI Gym 環境も含まれており、比較的扱いやすい (TensorFlow 1.13.x 以降 Eager execution がデフォルトとなり、Session の概念が消えたため) ため実装も比較的しやすいと考えられる。一方、ライブラリは beta 版であり、今後正式リリース時に挙動が変化することも容易に考えられる。

次に Keras-rl である。このライブラリは Keras をベースとし、比較的わかりやすいコードでネットワーク定義などができる。一方、本研究において必要となる学習ループ中の処理について自前で実装する場合、`keras.model.fit()` によりラップされているため、ソースコードを読み、クラスを継承した上で実装しなければならず、時間的制約が問題となる。

Chainer-rl は Chainer を強化学習向けに適用した新たなライブラリである。国産機械学習ライブラリであり、文献等も多く存在するが、これも学習ループがラップされているため実装上の時間的制約で断念した。

最後に PyTorch である。本ライブラリは本研究室井川氏の研究実装において使用された実績があり、Facebook が作る OSS 機械学習ライブラリである。本ライブラリでは学習ループを自前で用意する必要があり、本研究実装に都合が良いと考えられる。加えて TensorBoard による可視化も対応することからデータ分析が容易になるという利点もある。

以上より、本研究では PyTorch により実装を行っていく。実装にあたっては適宜ライブラリ調査及び井川氏からの助言、コード提供などにより行う。

2 今後の予定

今後は、まず強化学習における Hello World 扱いとなる CartPole-v0 を PyTorch により実装する。その後 Atari の適度な難易度の問題をベンチマークとし、本提案モデルの実装を行う。実験は 9 月終了を目指し、10 月以降修士論文の執筆及び AROB2020 参加のための Abstract 提出を目指す。