

目標3 2050年までに、AIとロボットの共進化により、自ら学習・行動し人と共生するロボットを実現

月面探査／拠点構築のための自己再生型 AI ロボット

研究開発項目

## 2. 分散型・Plug and Play 可能な AI

### 2023 年度までの進捗状況

#### 1. 概要

本研究開発項目では、「変幻自在」なモジュラー・マルチエージェントなロボットシステムに組み込まれ、自己成長する AI システムの実現を目指します。ロボットの動作を生成し制御するための AI として深層強化学習が研究され成果をあげてきていますが、現時点での研究成果は、単一身体ロボットや、個別のタスク学習での実装が中心です。組み換え可能な再構成型ロボット、および異種形態のロボット群に適用するためには、学習成果を Plug and Play (転用、再利用、再構築) 可能とする手法の確立が必要であり、「階層型強化学習」を発展させることが有力なアプローチとなります。この手法を活用する本研究開発課題においては、特に組立てタスクを用いて開発される AI 技術の評価します。

#### 2. これまでの主な成果

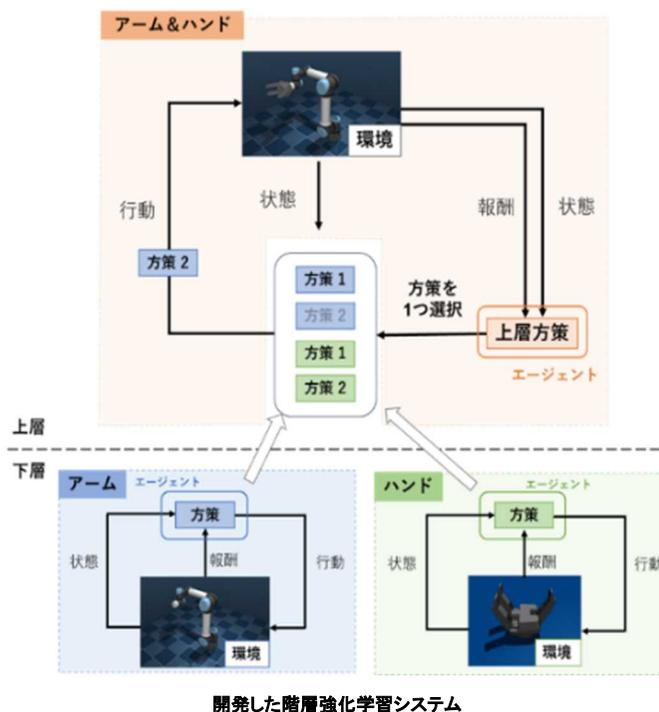
令和 5 年度は、前年度に整備した卓上マニピュレータのシミュレーション環境において階層型強化学習を実装し、目標とした学習性能を達成しました。

#### 課題1: 階層型強化学習による分散型 AI の研究開発とロボットへの実装

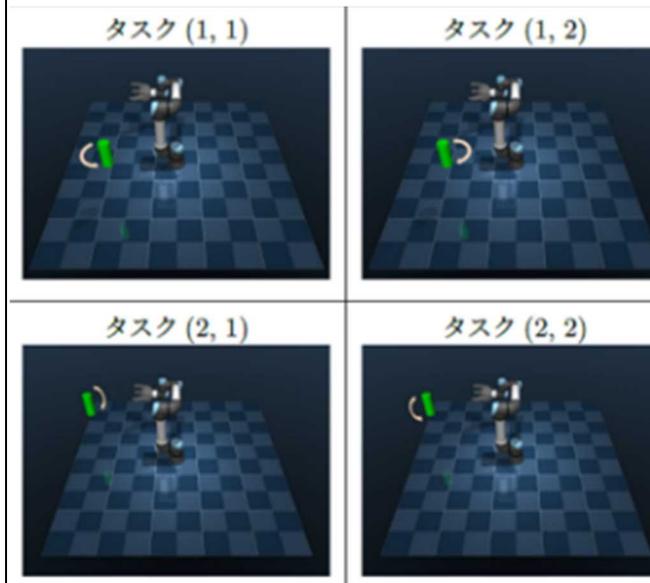
本年度の成果として、ハンドモジュールとアームモジュールを有するロボットシステムを対象としました。ハンドロボットとアームロボットが結合し異なるロボットの形態に変化した際の学習性能についての検証をおこないました。つまり、一つのハンド・アームロボットにおいて、ハンドロボットのみでの学習データ、アームロボットのみでの学習データのそれぞれを活用することにより、はじめからハンド・アームロボット一体のロボットで学習する場合と比較しました。結果として、学習データを活用した方が大幅に少ない試行回数によって、マニピュレーション動作方針を獲得することを達成しました。

まず次の図に示すような階層的な強化学習システムを構築し、アームモジュール、ハンドモジュールそれぞれに対して下位方針を獲得しました。具体的には、ハンド下位方針は物体把持とバルブの回転操作を学習し、アームモジュールはバルブへのリーチング動作を獲得しました。

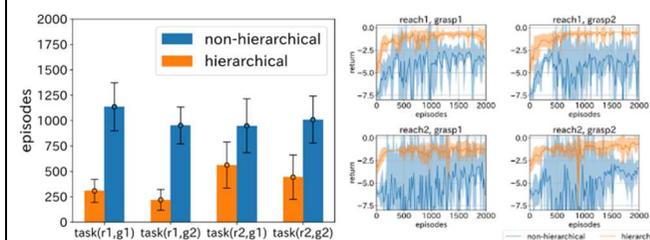
タスクとしては 4 種類のバルブ操作課題を考え、上位方針が下位方針モジュールの切り替えを学習することにより階層強化学習の学習性



能を評価しました。評価方法としては、階層化によりリポジトリに下位方針として蓄積された学習成果を活用することによる学習結果と、活用しない場合とを比較したところ、青色のバーで示された、リポジトリの蓄積を用いない場合に要した学習試行回数に比べて、それぞれのタスクについて、いずれも階層強化学習手法を用いた場合(オレンジ色のバーで表示)のほうが少ない学習回数によってバルブ操作課題が達成されています。4 つのタスクを平均して、階層強化学習を用いた場合は、リポジトリの蓄積を用いない場合に比べて大幅に少ない学習試行で方針が獲得可能となりました。加えて、階層強化学習によりリポジトリに蓄積された学習成果を活用した方(オレンジ色)が、活用しない場合(青線)に比べ、学習曲線の立ち上がり早く、かつ、異なるサンプルデータ系列を用いた場合においても安定して学習が行えていることも確認できました。



4つの異なるバルブ操作タスクによる学習性能評価



(左)各タスク達成に要した学習試行回数 (右)各タスク学習時の学習曲線

#### 3. 今後の展開

上記の課題を克服する要素技術を開発し、月面でのモジュラーロボットによる組立タスクを想定した各マイルストーンタスクにおいて本 AI 技術を応用した実証実験を行っていきます。