

令和 2 年度 防衛装備庁
安全保障技術研究推進制度

研究成果報告書
極少数の人間とAIの協働による
課題対処に関する基礎研究

令和 3 年 5 月

三菱重工業株式会社

本報告書は、防衛装備庁の安全保障技術研究推進制度による委託業務として、三菱重工業株式会社が実施した令和2年度「極少数の人間とAIの協働による課題対処に関する基礎研究」の成果を取りまとめたものです。

0. 研究背景

農業や製造業などの社会活動をはじめ、災害対応等の非定常的な社会基盤維持においてもAIに期待される役割が増大している。特に労働力の減少が見込まれる日本では、少数の人間によって多数の事象に対処するため、人と多数のロボット群等のAI エージェント（以下AI 群と呼ぶ）がチームワークを発揮し目的に指向して協働することが欠かせない。図0-1に、数に対して数で対処する必要がある問題（ここではスウォーム対処と呼ぶ）の例を示す。少数の人間とAI 群の連携に資する技術は、図0-1のような幅広い分野への波及も期待できるなど社会的価値の非常に高いものといえる。



図 0-1 社会的に考えられるスウォーム対処の例

ここでは、社会的要請が高く、数的劣勢環境が想定され、かつ以下の特徴を有する難易度の高い適用先の1つとして警備問題を想定する。高度な問題対処を想定することにより、その成果は他の適用先へも容易に適用が可能と考えられる。警備問題は、以下のような特徴を有する。

- ・ 対処目標の意図が不明、もしくは敵対的行動をとる
- ・ 対処目標の装備等の機能・性能や、取り締まる側のセンサで探知可能な範囲の外の目標有無等、取り締まる側は対処目標の情報を全て観測することはできず、不完全な情報に基づく意思決定が求められる
- ・ 他のエージェントからもたらされる逐次的な情報から高度な事象進展を予測して、その結果に応じて人と各AIエージェントの役割/タスクを適切に配分する必要がある

大量の対処目標（アンチ・スウォーム）を考慮したとき、少数の人間の能力のみでこれらに対処することは困難と考えられる。また、実力行使を伴う警備問題は人命や法規的判断に関わる事項であることを踏まえれば、フレーム問題の観点からも無人技術のみで対処することも困難と考えられる。従って、多数の対処目標が存在する警備問題では、人と多数のAI群が同一環境中で協働して対処することが望まれる。

時々刻々と変化する現実世界において人と多数のAI 群が同一環境で協働するためには、少なくとも大きく2点の問題点が存在する。1つ目は、多数の意思決定主体が同一環境中で目的を一として効果的に行動するには、個々のエージェントの局所的戦略と大域目標達成のための戦略を整合させる必要があり、そのための有効なメカニズムが明らかでない点である。2つ目は、人とAI 群が行う意思決定と行動の意図を相互に共有することが容易ではない点である。そのため、実用化されている自律システムのほとんどは、人と離隔した環境で協働するか、人によって遠隔操作を行うか、人とは異なる役割のみを予め明示的に割り当てて行うなど、人同士に比べ協働は限定的である。

本研究では、動的に変化する切迫した状況下において対処すべき対象が多数あり、少数の人間のみで対処することが困難な場面に対して人とAI 群がそれぞれ、同じ性能の身体性を持つエージェントにより対処を試みる状況を想定する。このような状況の中、少数の人間とAI エージェント群が連携し協働する有人・無人連携によって、高度で複雑な意思決定と事態対処を実現する技術は、現状の技術開発の延長線上にはなく、本研究のような取り組みによる基礎技術の獲得が必須であり、本研究で得られる基礎技術からの発展が可能である。

1. 委託業務の目的

(1) 研究課題の最終目標

本研究では、人が圧倒的に数的劣勢な状況においても、複数の人工知能（AI）エージェントとチームワークを発揮して協働することにより、多数の相手に対処可能とすることを目的とした基礎研究を実施する。

このため、3年間の業務において、特に人とAIの協調のための組織的枠組みを検討するため、エージェント間の協調を基本的な研究課題に据え、一般化した問題領域として提案されている追跡問題を例題として設定する。追跡問題は一般に複数のPredator（捕食者）が協調してPrey（被食者）を追跡し、接触や取り囲むことで捕らえることを目的とする。この例題に対し、人とAIという異質なエージェントが協働してシステム全体としての目的を達成するという分散協調問題の解決方法として、マルチエージェントシステムにおける合意形成方法として代表的な契約ネット・プロトコルやオークションによる交渉等を適用し、人とAIエージェントがチームワークを発揮することで数的劣勢であっても対処が可能であることを概念実証（Proof of Concept）する。

追跡問題の研究例における古典的なエージェント数はPredatorが4、Preyが1であるが、本研究ではこれを大規模化して多数のAIエージェントと人との間の協働に取り組むこととする。具体的には、ある行動モデルに従って行動する彼側のAIエージェント群=Prey[10^2 オーダー]に対し、人とAIエージェントからなる我側の群=Predatorを制御することで、彼の群を任意の方向に誘導・制御するか、捕獲することを目的とした問題とする。追跡問題のような競争環境では、競争関係にあるPredatorとPreyが相互に適応戦略を追求するため、本研究のようにPredatorの効用を高めてこれを評価しようとする場合、競争相手であるPreyにも高いレベルの戦略が必要となる。本研究では、こうした環境において自己の適応戦略を決定する際に自己の状況把握と相手の戦略把握が必要となるレベルの戦略を要する以下の特徴を有する問題設定とする。

- ・ 常統的な純粋戦略ナッシュ均衡が存在せず(相手の戦略によって自分の最適戦略が異なる)複数の戦略選択肢が存在する
- ・ 同時手番ゲーム(リアルタイムに双方が行動)
- ・ 不完備情報ゲーム(相手の戦略の利得等が不明)
- ・ 情報の不完全性(エージェント同士が他の存在を観測できる範囲に制約を設ける)

しかしながら、こうした戦略のレベルに対して定量的な目標値を設定することは困難であるため、Prey側にも人間を介在させて評価を実施することによりPrey側に“人間程度”の戦略レベルが備わっていることを担保した上で最終的評価を行う。

また、目標とすべき機能・性能の定量的な定義そのものが困難なことから、適切な評価方法の検討が必要である。本研究の実施を通じて、人とAIエージェントの協調行動を評価するための方法論を検討し、評価方法を設計する。

本研究の最終目標は次の2点となる。

- ・ 彼のAIエージェント群=Preyに対し、人とAIエージェントからなる我側の群=Predatorを制御することで、彼の群を任意の方向に誘導・制御あるいは捕獲することを試みるPredator-Preyのシミュレーション環境を構築し、これを用いて人とAIエージェントの協調行動について評価を実施する。
- ・ 評価によって、AIエージェントが人と良好な協調行動を取れたことを確認する。

(2) 最終目標を実現するために克服又は解明すべき要素課題

(1)で示した最終目標を実現するために克服又は解明すべき要素課題は以下の通りである。

① 人とAIの各エージェントが個々の局所的な戦略と大域目標達成のための戦略を整合させるためのアーキテクチャ

最終目標で述べた特徴を有するPreyとPredatorによる競争環境において、Predatorが効果的

に対処するためには、状況の変化に応じてPreyに対する適応戦略を次々に更新する必要がある。そのためには、Predatorを構成する人とAIがチームワークを発揮し、大域目標（組織としての目標）達成に向け組織的に行動することが必要となる。この際、個々のエージェントが観測できる範囲が制約される、情報の不完全性を考慮する必要がある。不完全性の高い状況では、Predator側が認識する状況が次々に変化するため、大域目標達成のための戦略よりも、個々のエージェントの局所的な戦略を尊重することが望ましく、一方で不完全性が低い状況では、各局所戦略が衝突しないよう大域目標達成のための戦略として整合させることが望ましいと考えられる。人と多数のAIが協働する際のこうした戦略整合のための合意形成方法は明らかではなく、解明すべき要素課題である。

② 時々刻々と変化する動的な環境において人とAIエージェント間で意図や意向をタイムリーに共有するヒューマン・スウォーム・インタラクション（HSI）の方法

人—AI群の間の相互の合意形成及び課題対処を実現するためには、人とAIでは認知、判断、行動の方法やそのスピードが全く異なることを考慮し、双方向のインタラクションをスムーズに実現する必要がある。特に、「人からAIエージェント群への指示方法」、「各AIエージェントから人への情報提示方法」及び「チーム全体の課題対処状況の提示方法」の3点については確立した手法が存在せず、HSIを実現するために解明すべき要素課題である。

（3）要素課題に対する攻略手段と実施項目及びそのための体制

人がAI群と協働するためのAIの構築方法及びそれらAI群と人のインターフェイス（HSI）のあり方や構築方法は確立されていない。更に人間系を対象に含む研究では、いかに優れた技術であっても実際に人が受け入れられない限りは価値がないため、実証性が重要であり、可能な限り迅速にプロトタイピングして評価するサイクルを繰り返すことが望ましい。そのため、本研究では、(2)の要素課題に反復的に実証しながら取り組むこととし、以下の①～⑤を実施する。

①. シミュレーション環境の構築

追跡問題の具体的な例題設定（ルール）を検討する。Predator及びPreyが対戦において果たすべき目標やそれぞれの利得、Preyの捕捉条件や性能差などを複数検討し設定する。またこれらの設定はイタレーションの過程でより高いPredatorの戦略を要するように見直しを行う。また、この例題設定のシミュレーション環境を構築する。OS、ミドルウェア及びオープンソースライブラリ等の調査、選定及び導入を行うとともに、シミュレーションの実行に必要な十分な計算機や開発環境の整備などハードウェアの整備を行う。これらソフトウェア基盤及びハードウェアを用いたシミュレーション環境の構築に向け、システム要求仕様を設定し、シミュレーション環境を構築する。また、マルチエージェントの合意形成のために行う契約ネット・プロトコルやオークション等によるエージェント間交渉をソフトウェアとして実装するための基本的枠組みを構成する。構成したシミュレーション環境は、研究期間中、イタレーション毎の評価結果を反映するために必要な改修を行い維持する。併せて、定量的、客観的な評価の方法及び基準等についても検討する。

②. AIエージェントの構築

人とAIの各エージェントが個々の局所的な戦略と大域目標達成のための戦略を整合させるためのアーキテクチャとして、契約ネット・プロトコルやオークション等の、マルチエージェント合意形成において代表的な方法を選択・適用しプロトタイピングを行う。このためには以下の項目を考慮してAIエージェントを構築する必要がある。これらについて、AI理論やゲーム理論を背景とした検討を行うため外部の専門家によるアドバイスを受けることとする。

②-1. 人とAIが効果的に分担可能なタスクの仕様の設定

エージェント間の合意形成のための交渉において用いる「タスク」の仕様を設定する必要がある。AIは確定的な情報の蓄積と分析に優れ、一方の人は不確実性のある環境における情勢判断などの抽象的、大局的な状況認知に長け、過去の経験や推測に基づいて未知の状況でも臨機応変に行動できる点に優位性があると考えられる。この仮説に基づき人とAIそれぞれの優位点を生かすため、タスクは、AIが得意とする不確実性の少ない短期的かつ局所的判断に基づくタスクと、人が得意とするそれらを大局的に統制するタスクに分けることとする。こうした階層的なタスクの仕様についてプロトタイピングを通じて具体化する。

②-2. 各エージェントにおける適切なタスクの条件、タスクの価値基準の設定

契約ネット・プロトコルやオークション等による交渉では、人及びAIの各エージェントは、必要に応じてタスクの提示と、独自の価値基準（例えばコスト、所要時間、確実性など）に従って入札、落札等を行う。これはエージェント間の交渉を特徴づける重要な要素であり、適切なタスクの条件と価値基準を設定することでPredatorの中で望ましいタスクの割り当てを実現し、チームとしてPreyに対する適応戦略を実行できるようになると考えられる。こうしたタスクの条件と価値基準の設定について、プロトタイピングを通じて具体化する。

③. HSIプロトタイプ作成

AIエージェント群に対する指示やAIエージェント群からの情報を表示・操作可能なHSIプロトタイプを仮作する。実施項目として以下のようなものがある。これらについても、AI理論やゲーム理論の観点から外部の専門家によるアドバイスを受けることとする。

③-1. 人からAIエージェント群全体に対する効果的な指示方法の解明

人の意図や意向を考慮してAIエージェントが行動を変化させる協調については未解明の課題がある。例えば、1人の人が100個のAIエージェントに対し全ての行動を逐一実時間で指示することはできないが、「全員直進」と指示することは簡単である。こうした人とAIエージェントの差異を考慮した、人からAIエージェント群全体に対する指示方法を具体化し、実際に評価可能なプロトタイピングを行う。本項目は②-1「人とAIが効果的に分担可能なタスクの仕様の設定」と密接な関係にあるが、ここでは特に②-1で設定したタスクの仕様に対する表示操作系の設計を検討する。

③-2. AIエージェント群の将来予測結果の提示方法の解明

Predatorを構成する評価者の状況判断とPreyの戦略把握に資する将来予測情報として、外挿に基づく個々のPreyの移動傾向及びPredatorを構成するAIエージェントの将来行動情報をHSIプロトタイプに提供可能とする。Preyに関する情報取得により行動の予測結果等を適切に提示することができれば、Predatorは適応戦略をとりやすくなると考えられる。一方で、こうした予測結果は評価者のイメージに合わない場合に錯誤を招く要因となる可能性もあり、評価者の能力によっては過信頼を起こす恐れもある。こうした課題も踏まえて仮作と評価を繰り返し、人と協働可能とする上で効果的なAIエージェント群の将来予測結果の提示方法をプロトタイピングする。

③-3. タスクのマーケット状況等の効果的な人への提示方法の解明

AIエージェント群の間でタスク競合が起きるなど、好ましくない状況が生起している場合や、人間の指示が全体から見て明らかに得策でない場合には、全体行動の改善を促すためにAIエージェント群から人に対してフィードバックを行うことが有効である。本研究では、人とAIエージェントが相互に契約ネット・プロトコルやオークション等を通じてタスクを契約募集し、その入札や落札等を経てタスク配分を行っていくことから、これらのタスクの取引状況（マーケット状況）を直感的に人に対して提示する方法を検討する。

④. トレーニングの実施

①～③で構築した環境を使用して人のトレーニングを行うとともに、②及び③について見直す。特に人とインターフェースについてはプロトタイピングを通じた評価が必須であるため、②～④は繰り返し実施する。トレーニングは設計者によって行い、自らトレーニングを通じて必要と考える機能を設計に反映する。

⑤. フィールド評価の実施

構築したシミュレーション環境を用いて協働による効果の評価を行う。評価の対象者は①～④の設計に関わらなかった者とする。評価にあたっては、一定の慣熟操作の後に複数回行い、例えば航空機パイロットのワークロード評価において一般的に用いられるクーパー・ハーパー・レーティング等を活用し、協働の効果を安全性、効率性、快適性等の異なる軸で定量的に評価できる手法を考慮する。レーティングによる評価とともにコメントを取得し、改善を要する項目に対する具体的な改善方法を検討するためのフィードバックとする。

また協働の効果を評価するにあたって、競争相手側であるPrey側に“人間程度”の戦略レベルが備わっていることを担保するため、最終評価までにPrey側にも人間を介在させて評価を実施する。

⑥. プロジェクトの総合的推進

プロジェクト全体の連携を密としつつ円滑に運営していくため、運営委員会や技術検討会の開催等、参画各機関の連携・調整にあたる。

2. 研究開始時に設定した研究目標の達成度

以下、1.（1）項の最終目標に対する達成度を記載する。なお、1.（3）項の各実施項目に対応する全期間を通じた研究方法及び成果については3章で詳説する。

【最終目標1】

“彼のAIエージェント群=Preyに対し、人とAIエージェントからなる我側の群=Predatorを制御することで、彼の群を任意の方向に誘導・制御あるいは捕獲することを試みるPredator-Preyのシミュレーション環境を構築し、これを用いて人とAIエージェントの協調行動について評価を実施する。”

1.（2）で提示した要素課題を有しかつ現実問題へ拡張性のあるPredator-Preyの例題設定を検討し、本研究では鬼ごっこの一種である“しっぽ鬼”をチーム対戦形式で実施した。ここで、人とAIエージェントが協働する群（以降、チームAと呼ぶ）の対戦相手として、すべて人間が操作する人エージェントのみで構成する群（以降、チームBと呼ぶ）と同条件で対戦させて評価することにより、その勝敗によってチームAが“人間の集団程度”の戦略レベルが備わったうえで有効な協調行動が獲得できたことを確認可能とした。

以上の例題設定を受け、ネットワークを介して100人規模の人間が接続して対戦評価可能なしっぽ鬼のシミュレーション環境を構築した。加えて、チームA内においてマルチエージェント合意形成を図るエージェント間交渉の枠組みを検討しソフトウェアとして実装した。

また本研究では研究期間内にどの程度の性能を達成できるかを目標設定し演繹的に機能を実装していくようなウォーターフォール型で開発することは困難と考え、プロトタイプ構築と評価を繰り返し改善していくアジャイル的な方法を用いた。シミュレーション環境の構築と人とAIエージェントの協調行動の評価のサイクルをスパイラルと呼び、3か年の取組において計3回のスパイラルを実施した。

以上のことから、最終目標1を達成したと考える。

【最終目標2】

“評価によって、AIエージェントが人と良好な協調行動を取れたことを確認する。”

図2-1に、各スパイラル期間中に行ったフィールド評価において、チームAの勝率と、チームA、Bの各チームが相手チームのエージェントを捕獲した合計数で算出した捕獲率の変遷を示す。

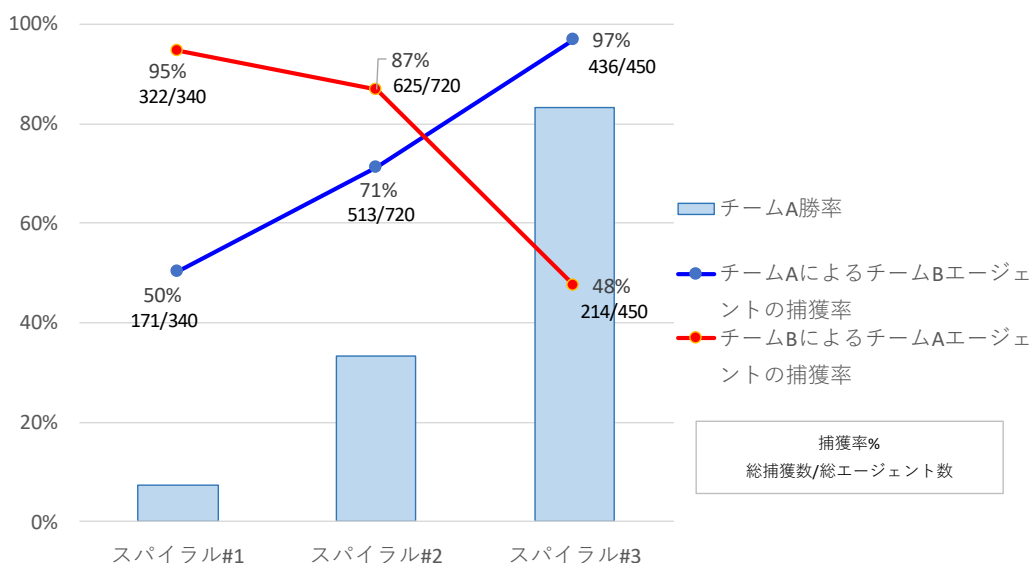


図 2- 1 スパイラル毎のチーム A 勝率とエージェント捕獲率の変遷

スパイラル#1では、構築したしっぽ鬼のシミュレーション環境を用いて、搜索、捕獲、離脱の3種類の基本的なAIエージェントの協調タスク仕様を設定し、少なくとも1対1の状況では人を相手にしても対等に渡り合える（捕獲されず、隙あれば捕獲する）程度の性能を具備してフィールド評価を行った。しかし、チームAの勝率は7%であり、スパイラル#1ではチームBがトータルで95%のチームAエージェント（人またはAI）を捕獲した一方、チームAは50%のチームBエージェントしか捕獲できず、全く歯が立たないと言っていい状況であった。チームAのAIエージェントは少なくとも1対1では人に負けない性能を有していたため、この結果は、例題環境として設定したしっぽ鬼というゲームがエージェント間の協力が重要な問題設定となっていることを裏付けている。

スパイラル#2では、スパイラル#1においてチームBの人々がフィールド評価においてどのように連携したか観察し、分析した結果に基づき、AIエージェントの捕獲タスクに対して連携捕獲機能を追加した。捕獲タスクに対して足止役や捕獲役といった役割を定義するとともに、契約された捕獲タスク内で動的にタスク配分を割り当て可能とした。さらに、AIエージェントが実行中のタスクおよび役割を可視化する役割表明機能を追加したことで、協働する人にとってAIエージェントの意図を把握可能とした。その結果、チームAによるチームBエージェントの捕獲率が71%と改善したものの、チームBによるチームAエージェントの捕獲率は87%と依然高い水準にあり、チームAの勝率は33%にとどまった。

スパイラル#3では、スパイラル#2のフィールド評価で観察された以下の2点に対する対策を行った。

- (1) チームBの巧みな分散逃避によってチームAのAIエージェントが誘引され、分離し、チームBが局所的に数的優勢なエリアに誘導される状況があった。こうした状況に対応するために用意していた離脱タスクは、人が指示した位置に半強制的に移動させることが可能であったが、エージェントの背後警戒が疎かになり意図した効果を発揮できなかった。
（対策）離脱タスクを削除し、人が指示した位置に警戒しながら集合し、陣形を組む集合タスクに置き換えた。
- (2) 捕獲タスクは、3体のエージェントがサブチームを構成して捕獲対象となるエージェントの捕獲を試みるが、捕獲対象のエージェントが逃避する場合、3体のエージェントが捕獲対象との予想会合点位置に向かって移動するという比較的単純な追跡方法であった。捕獲対象が逃避を継続した場合、捕獲を試みる3体のエージェントは延々と追跡を行い、状況が収束しないばかりか、チームBが局所的に数的優勢なエリアに誘導される一因となった。
（対策）逃避する捕獲対象をボロノイ領域分割に基づき味方同士で囲い込むように移動する追跡方法に改善した。

この結果スパイラル#3では、チームAはスパイラル#3中に行ったフィールド評価全体で97%のチームBエージェントを捕獲し、一方チームBによるチームAエージェントの捕獲率は48%に留まる好成績となった。チームAの勝率は83%に達し、チームAに“人間の集団程度”の戦略レベルが備わっていることを示すことができたと言える。

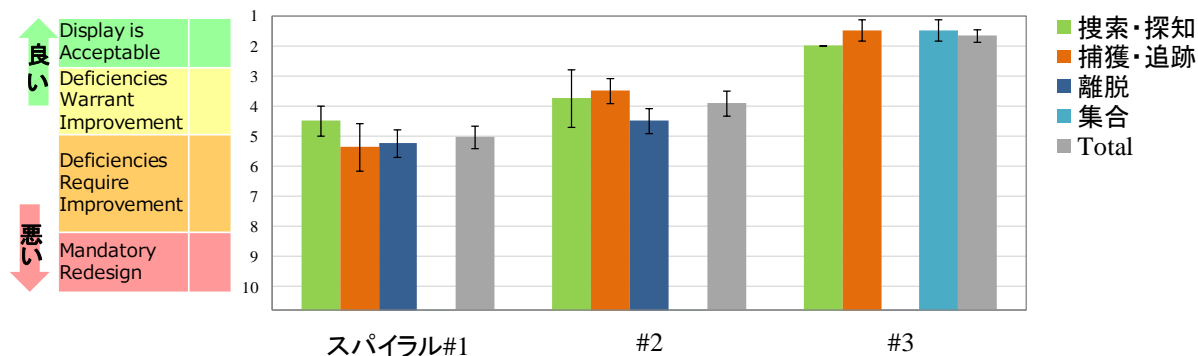


図 2- 2 レーティング結果の変遷

フィールド評価では、チームAにおいてAI群と協働を行った参加者から、各評価の実施後にHSIプロトタイプに関するコメントをヒアリングした。ヒアリングでは、口述あるいは記述により自由な形式でコメントを取得するとともに、レーティングも取得した。レーティングは、航空機の飛行性能や操作性（操縦性）を評価する指標であるクーパー・ハーパー・レーティングを基に、無人システムの操作性や状況監視等への適用を想定して提案されているMCH-UVD（Modified Cooper Harper scales for assessing Unmanned Vehicle Displays）を用いて、レーティング・スケールとして定量化した。図2- 2 にレーティング結果の変遷を示す。スパイラル#3では全てのタスクについて“Display is Acceptable”のレーティングを得た。

以上から、評価によって、AIエージェントが人と良好な協調行動を取れたことが確認できたといえ、最終目標2を達成したと考える。

さらに、スパイラル#2において人間と強化学習エージェントの協調に関する目途付けを行って、捕獲タスクにおける連携捕獲機能に深層強化学習によって得た強化学習モデルを適用し、スパイラル#3にてフィールド評価結果との比較を行った。その結果、人間が設計して作りこんだものと遜色ないモデルを強化学習によって発見的に作成可能であることを確認したほか、可読性のないブラックボックス方策を持つ強化学習モデルであっても、適切にタスクと役割を定義し、AIがこれから行おうとするタスクあるいは役割を宣言することで人と協調が可能であることを示した。

3. 委託業務における研究の方法及び成果

3. 1 本研究の進め方

1. (3) で述べたように、本研究で取り組む、人がAI群と協働するためのAIの構築方法及びそれらAI群と人のインターフェイス (HSI) のあり方や構築方法は確立されていない。更に人間系を対象に含む研究では、いかに優れた技術であっても実際に人が受け入れられない限りは価値がないため、実証性が重要である。人の感性に受け入れられるインターフェイスを設計段階から綿密に検討して完璧に製作することは困難であり、ヒューマン・インターフェイスに関する標準化の分野でも、ISO13407で規定された人間中心設計プロセスにて、問題点を改善することに主眼を置いて以下の手順を繰り返すという方法論について述べているのみである。

1. 利用の状況の理解と明確化
2. ユーザや組織の要求事項の明確化
3. 設計による解決策の作成
4. 設計を要求に照らして評価

本研究においても、ヒューマン・インターフェイスを含む人とAI群の協働の在り方について可能な限り迅速にプロトタイピングして評価するサイクルを繰り返すことが望ましいと考え、要素課題に反復的に実証しながら取り組むこととした。そのため、まず評価を行うプラットフォームとしての①シミュレーション環境の構築を行った上で、②AIエージェントの構築、③HSIプロトタイプの実施、④トレーニングの実施、⑤フィールド評価の実施を繰り返し行うこととした。図3.1-1 に研究実施フローを示す。繰り返しは短期と長期の2種類のサイクルで行った。短期のサイクルとは、②及び③でAIエージェント及びHSIプロトタイプを作成し、研究実施者自ら④において試行して②及び③について短いサイクルで見直すループ(図中のイタレーション・ループ)をいう。一方、長期のループとは②～④を通じて構築したAIエージェント及びHSIプロトタイプを用いて④フィールド評価を行って評価結果のフィードバックを行う比較的長いサイクルのループ(図中のスパイラル・ループ)を指し、本研究では初年度に①を実施し、2年目にスパイラル1回、3年目にはスパイラルを2回実施した。次節以降において、各取り組み内容を詳説する。

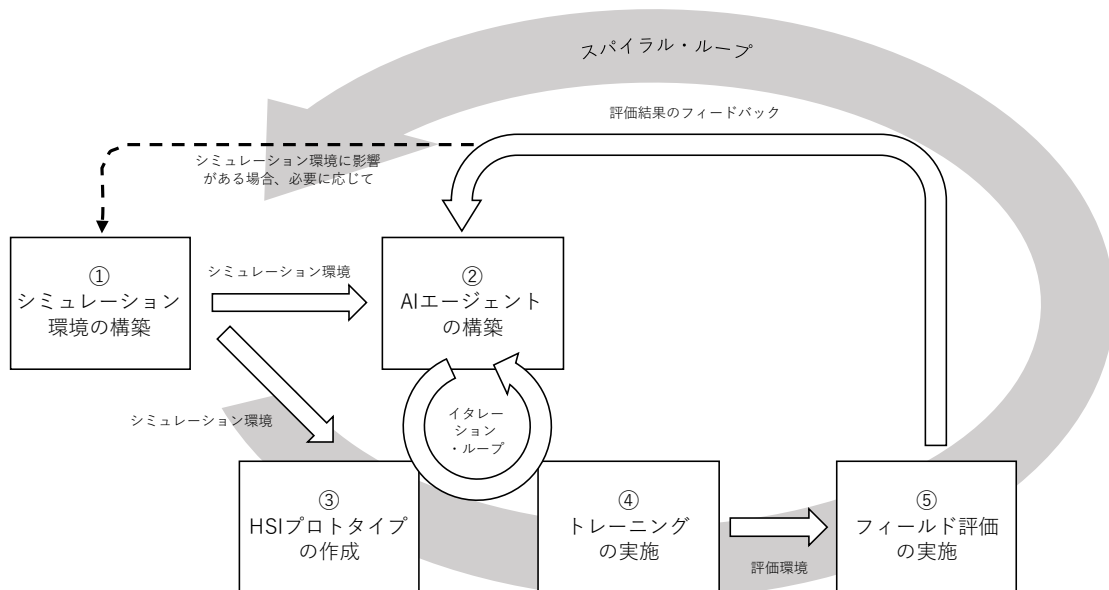


図 3.1- 1 研究実施フロー

3. 2 シミュレーション環境の構築

1. (3)の実施項目①に対応する例題設定検討を示す。さらに、製作したシミュレーション環境のシステム構成、エージェント間交渉の枠組み及び人とAIエージェントの協調行動の評価方法及び基準について検討した成果を示す。

3. 2. 1 例題設定検討

例題設定検討として、1. 例題設定条件の再確認、2. 先行研究調査、3. 取り組む例題環境の検討、4. ゲーム環境設定およびルールの設定結果、5. 人とAIエージェントの協調行動の評価方法及び基準 を実施した結果を以下に示す。

3. 2. 1. 1 例題設定条件の再確認

1. 項(1) に記載の例題設定に関する記述を踏まえ、例題設定条件を以下の(a)～(e)の様に整理した。また、評価方法を考慮し(f)の条件を加えた。

(a) 群れによるPredator-Prey

『ある行動モデルに従って行動する彼側のAIエージェント群=Prey[10^2 オーダー]に対し、人とAIエージェントからなる我側の群=Predatorを制御することで、彼の群を任意の方向に誘導・制御するか、捕獲することを目的とした問題』

『Prey側にも人間を介在させて評価を実施することによりPrey側に“人間程度”の戦略レベルが備わっていることを担保した上で最終的評価を行う』

(b) 複数の戦略

『常統的な純粋戦略ナッシュ均衡が存在せず(相手の戦略によって自分の最適戦略が異なる)複数の戦略選択肢が存在する』

(c) 不完備情報ゲーム

相手の戦略の利得等が不明

(d) 同時手番ゲーム

リアルタイムに双方が行動

(e) 情報の不完全性

エージェント同士が他の存在を観測できる範囲に制約を設ける

(f) チーム形式の対称戦

人間のみのチームと同条件で対決し、勝つことによって有効な協調行動が獲得できたことの結果の1つとする

3. 2. 1. 2 先行研究調査

既存研究において3. 2. 1. 1項の例題設定条件を満足する例題設定がないか調査を行った。

初めに古典的なPredator-Preyの先行研究事例について以下に述べる。

Predator-Prey は、1985年 Bendaらにより“複数の知的エージェントが協力して1つの解を求める際に、エージェントの間に存在する組織構造が問題解決の効率にどのような影響を与えるか” [1] [3]について評価する目的で「追跡問題」として提案された。

Bendaらによって提案された追跡問題の定義は以下の通りである。また、追跡問題の概要を図3. 2. 1. 2- 1に示す

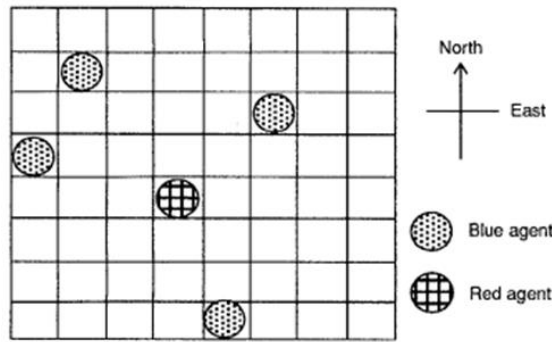


図 3. 2. 1. 2- 1 追跡問題の概要 [1] [3]

- ・ 2次元グリッド空間
- ・ 各エージェントは上下左右の移動及び静止という行動が選択可能
- ・ 同時手番ゲーム(3. 2. 1. 1項 (d)に相当)
- ・ 各エージェントの視界及び通信が制限されている(3. 2. 1. 1項 (e)に相当)
- ・ 4つのBlue agent(Predator)がRed agent(Prey)を四方から取り囲むことがタスク

1989年、Gasserらは終盤の問題解決の戦略として「predatorはpreyの四方のエリアからpreyの距離を狭めれば必ず捕獲できる」ことに着目し、四方のエリアに対するPredatorの割当問題としてアプローチした[2][3]。このアプローチは追跡問題について取り得る最適戦略が少ないことを示唆している。Gasserらによるアプローチの概要を図3. 2. 1. 2- 2に示す。

以上から、本研究において古典的なPredator-Preyから拡張が必要となる3. 2. 1. 1項の例題設定条件は以下のとおりである。

- 群れによる Predator-Prey
- 複数の戦略
- 不完備情報ゲーム
- チーム形式の対称戦

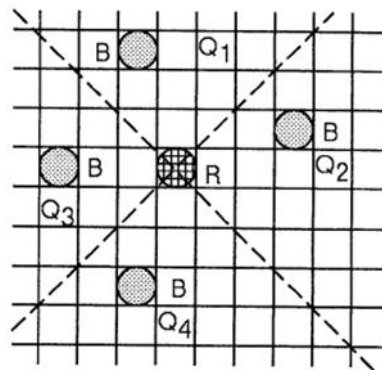


図3. 2. 1. 2- 2 Gasserらによるアプローチの概要 [2] [3]

次に、ゲーム理論において一般的に用いられる35種類の標準問題 (List of games in game theory[4]) のうち、これらの例題設定条件を満足する標準問題が存在するかどうかを調査した。調査にあたってList of games in game theory[4]で各問題の特徴項目と、3. 項の例題設定条件の対応を整理した結果を表3. 2. 1. 2- 1 に示す。調査の結果、所要の条件を全て同時に満たす標準問題はなかった。

同様に、流用可能性のあるオープンソースのビデオゲームを対象にList of open-source video games[5]を整理し調査した。しかし、List中のゲームはいずれも娯楽を目的としており、現状のままで同様の条件を満たすゲームは存在しないことが分かった。

表 3.2.1.2- 1 ゲーム理論における標準問題についての整理項目

項目	取り得る値	例題設定条件との対応
Number of players	1, 2, ..., N	N の場合、(a) 群れによる Predator-Prey の要素があると した。
Strategies per player	1, 2, ..., variable, infinite	共にvariable, infiniteの場合、(b) 複数の戦略 及び(c) 不完備情報ゲーム の要素があると した。
No. of pure strategy Nash equilibria	1, 2, ..., variable, infinite	
Sequential	Yes, No, N/A	Noの場合、(d) 同時手番ゲーム の要素があると した。
Perfect information	Yes, No, N/A	Noの場合、(e) 情報の不完全性 に適合すると した。

ゲーム理論で取り扱われる一般的な標準問題及びオープンソースのビデオゲームに本研究の目的に合致するものがないため、複数の戦略を対象にした既存の取り組み例に流用可能な例題がないか調査を行った。ここでは、類似の取り組みとして有力な2件の調査結果について以下にまとめる。

まず、StarCraftを利用した研究について述べる。StarCraftとは資源収集、ユニット生成、攻撃及び偵察等の戦術を同時進行して行うリアルタイムストラテジーゲームである。以下の様に、プレイヤーの初期戦略が三すくみの構造になっており、偵察によって相手の戦略を理解することが重要な要素の1つである。

- ・ 自陣を防衛する初期戦略は攻撃的な初期戦略に対して強い
- ・ 自軍を拡張する初期戦略は防衛的な初期戦略に対して強い
- ・ 敵陣を攻撃する初期戦略は拡張的な初期戦略に対して強い

アメリカ人工知能学会AAAI (Association for the Advancement of Artificial Intelligence) が主催するAIIDE (Artificial Intelligence and Interactive Digital Entertainment) では2010年からStarCraft : BroodwarにおけるAIの強さを競う競技会 [6]を開催しており、2018年はSamsung (SAIDA) チームが優勝した。この競技会では、競技参加者が自身のプログラムを用いてゲーム内の情報を取得したり、コマンドを送ったりするための専用のI/Fが用意されている。

また、DeepMindからはStarCraft IIのAI開発用環境がリリースされている。しかし、StarCraft及びStarCraft IIは娯楽を目的とした商用ゲームであり、用意されたI/F以上の環境側改修はできず、本研究に用いるには拡張性に乏しい。例えば、あるチーム内の個々のキャラクターを別々の人が操作するようなI/Fは想定されていないため、本研究には適さない。

次に、人とAIの協働に関する取り組み例として、DeepMindが2018年7月に発表したCapture the Flag (旗取りゲーム) [7]について述べる。図3.2.1.2- 3に Capture the Flagの様子を示す。人と機械学習を行うエージェントの協調に関する研究として、FPS(First Person shooter)ゲームであるQuake III Arena(1999年発売)のゲームエンジンを利用した「旗取りゲーム」を採用している。40人の人間とAIをランダムに組み合わせ、チーム戦を実施した取り組みが発表されている。図3.2.1.2- 4に人のプレイの様子を示す。本研究の例題として採用することを考えた場合、Quake III Arenaのゲームエンジン自体はオープンソースであるがGPLライセンスのため商用利用に懸念があるうえ、エンジンとしては古く、再利用可能なリソースに乏しい。また、DeepMindは9月にゲームエンジンUnityとの本格提携を発表[8]しており、今後Unityへの移行が想定されることから、Quake III Arenaを用いたCapture the Flagの環境は今後陳腐化することが予想され、本研究には適さない。



図 3.2.1.2- 3 Capture the Flag の様子[7]

図 3.2.1.2- 4 人のプレイの様子[7]

3. 2. 1. 3 取り組む例題環境の検討

3. 2. 1. 2 項で示した先行研究調査より、既存の環境では 3. 2. 1. 1 項の例題設定条件にすべて合致するものはないため、新たにゲーム環境を整備する必要がある。新たにゲーム環境を整備するにあたって、初めに例題を選定する必要がある。その例題検討での留意事項について以下のように整理した。

- ・ 3. 2. 1. 1 項で示した例題設定条件 (a) ~ (f)
- ・ 人の長所 (適応力 , 認知能力 . . .) と AI の長所 (即応力 , 正確性 . . .) が生きる協調要素がある例題であること (以下、(g) とする)
- ・ 人間の参加者がルールを容易に理解でき、納得感があること (以下、(h) とする)
- ・ 現実問題への拡張を想起すること (以下、(i) とする)

以上の留意事項について適合確認を行い例題について検討した。表 3.2.1.3- 1 に例題検討過程を示す。

表 3.2.1.3- 1 例題検討過程

例題名称	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	点数*
両チームに鬼がいる鬼ごっこ	◎	○	○	◎	◎	◎	○	○	○	13
しっぽ鬼	◎	◎	○	◎	◎	◎	◎	◎	○	16
けいどろ	◎	◎	○	◎	◎	△	◎	◎	○	14
アメフト	△	◎	○	◎	◎	△	◎	○	○	11
サッカー	△	◎	○	◎	◎	◎	◎	○	○	13
騎馬戦	◎	○	○	◎	◎	◎	◎	△	○	13
旗取ゲーム	△	○	○	◎	◎	◎	◎	◎	○	13
不審船対処問題	○	○	◎	◎	◎	△	◎	△	◎	12
機甲戦(戦車)	○	◎	◎	◎	◎	○	○	△	◎	13
空対空戦闘	△	◎	◎	◎	◎	○	○	○	◎	12

*)◎:適合(2点), ○:条件付き適合(1点), △:適合させる場合無理が生じる(0点)

以上の適合確認の結果、しっぽ鬼が高い適合性があると判断した。しっぽ鬼とは鬼ごっこから派生した遊びであり、以下のように説明されている。

“数人（全体の20パーセント程度）の鬼を決め、鬼以外は帽子や縄跳びの縄などをズボンと体の間に挟み半分以上を外に出した状態（これをしっぽに見立てる）で逃げ回り、鬼にこれを取られると、退場となる。制限時間内に全員が退場すると鬼側の勝ちとなり、逆に1人でも生き残ると鬼側の負けとなる。鬼を無くして全員でしっぽを取り合うように行うこともある。この場合、制限時間内に最も多くのしっぽを取った者が勝ちとなる。この場合もしっぽを取られた者はその場で退場となる。しっぽを取られても退場とせず、しっぽを取ったら付けても良いとするものもある。チーム戦として、2チーム以上に分けて行うこともある。基本的なルールは同じで、この場合はしっぽを取られた人が少ないチームが勝ちとなる。” [9]

しっぽ鬼は鬼ごっこの派生であるため「(a) 群れによるPredator-Prey」および「(d) 同時手番ゲーム」の条件に適合する。加えて、「(f) チーム形式の対称戦」の条件から本研究では2チームに分けたチーム戦のしっぽ鬼を整備することとする。

また、しっぽ鬼の個人戦略としては以下のように「(b) 複数の戦略」が存在する。

- ・ 相手を追いかける
- ・ 相手から逃げる
- ・ しっぽを取られないように動く

状況を観察し、相手は誰を目標に追いかけているのか、もしくは誰から逃げているのかといった相手の戦略理解が必要であり、「(c) 不完備情報ゲーム」であると言える。

さらに、しっぽ鬼の特徴として、しっぽは人の後ろに取り付けられ人間の「(e) 情報の不完全性」をゲーム要素としていることが挙げられる。これにより相手が他人に気を取られている内にしっぽを取りに行くといった味方同士の協調要素（本項の留意事項 (g) 相当）も発生する。

以上より、チーム戦のしっぽ鬼を例題としてゲーム環境を整備することとした。

3. 2. 1. 4 ゲーム環境設定およびルールの設定結果

チーム戦によるしっぽ鬼について、ゲームの環境及びルールを図3.2.1.4- 1に示す。

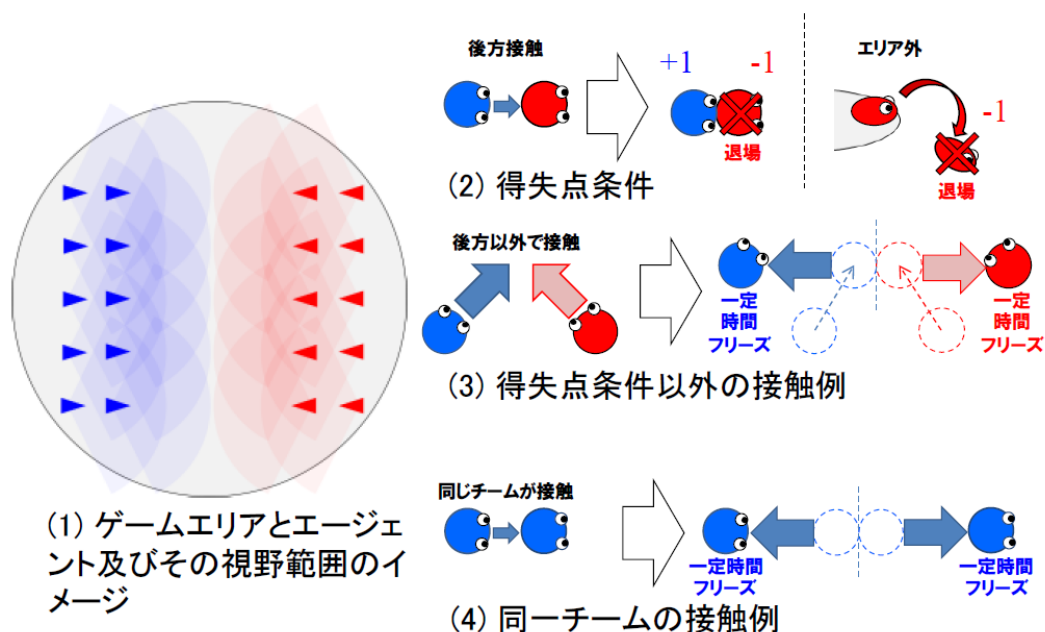


図 3.2.1.4- 1 しっぽ鬼のゲームの環境及びルール

- ・ ゲームエリアは角部などの特異点が生じないように、円形エリアとした。
- ・ 各エージェントの一人称の視野は前方への扇形状に制限されるが、同じチームのエージェントの位置と向き及び同じチームのエージェントの誰かが観測している相手チームのエージェントの位置と向きの情報は共有される。
- ・ 以下のように、人とAIの混成チーム(チームA)と全て人のチーム(チームB)に分ける。
 - チームA: 人は1体のエージェントを直接操作しつつ、味方のAIに指示が出せるものとする。エージェントの色は青とした。
 - チームB: 人は1体のエージェントの直接操作のみとする。エージェントの色は赤とした。
- ・ 各チームを構成するエージェント数は可変とする。
- ・ 得失点条件は、自己のエージェントが相手チームエージェントの後方に接触したとき味方チームへ点が入るものとし、相手チームエージェントに自己のエージェントの後方が接触されたとき自己のエージェントはゲームから退場し味方チームの点数が減るものとする。自己のエージェントがエリア外に出たらゲームから退場とし、味方チームの点数が減る。
- ・ 相手チームのエージェントと前面同士が接触した場合あるいは自チームのエージェント同士が接触した場合、一定距離移動し一定時間硬直(フリーズ)する。
- ・ 全てのエージェントの運動能力は同一とする。エージェントの選択可能な行動は前後左右斜めを含む移動、左右2種の回転運動及び移動しない場合の組み合わせとする。

以上のルールで規定されるしっぽ鬼に関して特筆すべきは、全てのエージェントの身体的能力は同一のため、局所的に1対1となった場合、互いに後方を狙ってぐるぐると周回し、逃走する相手をいつまでたっても追いつけない状況になり収束しない。このゲームにおける基本的な捕獲戦術は、局所的に数的優勢の状況を作り出し、相手にとにかく接触して一定時間フリーズの状態に持ち込んだ(以後、これを行うエージェントを足止め役と呼ぶ)うえで、その間に他のチームメイトが捕獲する(同捕獲役と呼ぶ)という役割分担のチームプレイが必要な問題設定となっている。

3. 2. 2 シミュレータ環境のシステム構成

例題設定であるしっぽ鬼を実現するシミュレーション環境のハードウェア及びソフトウェアの各構成について以下に示す。

3. 2. 2. 1 ハードウェア構成

各スパイラルのフィールド評価において、フィールド評価参加者によるエージェント操作入力及びコミュニケーションに使用するゲームパッド及びヘッドセットを整備した。また、フィールド評価用ソフトウェアは、一般的なOA用パソコンで動作するように構築したことにより、フィールド評価参加者の人数分のパソコンを新たに整備する必要が無いように考慮した。フィールド評価は全て社内で行ったため、各社員が普段利用しているパソコンを用いて参加した。整備したハードウェア構成は以下の通りである。

- ・ F310r フィールド評価用ゲームパッド：フィールド評価参加人数分

主にフィールド評価時の操作用としてフィールド評価参加者に貸与した。外観を図 3.2.2.1- 1 に示す。

【主な仕様】

コントローラ入出力：A、B、X、Y ボタン、LB、RB ボタン、左、右アナログトリガー、スタート、バックボタン、クリック可能なアナログミニジョイスティック x2 本、8 方向 D パッド、ホームボタン、スポーツモード

接続 I/F:USB



図 3.2.2.1- 1 F310r フィールド評価用ゲームパッド外観

- ・ HS-HP20UBK フィールド評価用ヘッドセット：フィールド評価参加人数分

主にフィールド評価時の音声通信用としてフィールド評価参加者に貸与した。外観を図 3.2.2.1- 2 に示す。

【主な仕様】

ヘッドホンタイプ：オーバーヘッド両耳用

接続 I/F:USB



図 3.2.2.1- 2 HS-HP20UBK フィールド評価用ヘッドセット外観

- ・ フィールド評価に用いたパソコン(参考)

フィールド評価参加者が社内で普段の OA 作業で使用する社有パソコンであり、シミュレーション環境実行に使用した。フィールド評価参加者の業務内容によって仕様が異なるが、目安としてフィールド評価参加者の所属部署で一般的な OA 用パソコンの仕様を示す。特段に高性能な仕様ではなく、ごく一般的なパソコンである。

【最も広くフィールド評価参加者に利用されているパソコンの主な仕様】

Model: Dell Optiplex5050 スモールフォームファクタ

CPU: Corei7 7700 4core/8thread 3.6GHz

MEMORY: 16GB

STRAGE: HDD 500GB

3. 2. 2. 2 ソフトウェア構成

ソフトウェアの概念図を図3.3.2.2- 1 に示す。2チームともに100体までのエージェントを登場させることができ、登場エージェントは人による操作かAIかを選択できる。人による操作が選択された場合、操作を担当するフィールド評価参加者が使用するパソコンが操作端末として割り当てられる。また、AIの処理はホストパソコンが一括して行う。

シミュレーション環境を構築するうえで、本研究のコアとなる例題環境仕様に注力するため、開発環境やプログラム部品は一般的な民生品を活用した。本ソフトウェアはWindows OSで動作し、Unityをゲームエンジンとして利用している。Unityはプログラム開発環境であるMicrosoft Visual Studioとの連携が可能であり、Windows OSで動作するアプリケーションのプログラム部品を提供するライブラリ群を利用できる。また、UnityにはUnity Assetと呼ばれる機能拡張を行うためのパッケージ群があり、それらをプログラミング言語 C#スクリプトによって制御している。前項で記述したように、契約ネット・プロトコルに関連したエージェント間のメッセージはライブラリwebsocket-sharpを利用してHTTP + WebSocket(TCP/IP)で通信しており、他のデータ通信は既製品でありオンラインゲーム開発で広く用いられているUnity Asset Mirrorを利用して通信している。

本ソフトウェアの構成概念図を図3.2.2.2- 2 に示す。本ソフトウェアのネットワーク構成を表3.2.2.2- 1、本ソフトウェアで使用する開発環境/ライブラリ/Unity Assetの構成一覧を表3.2.2.2- 2 に示す。

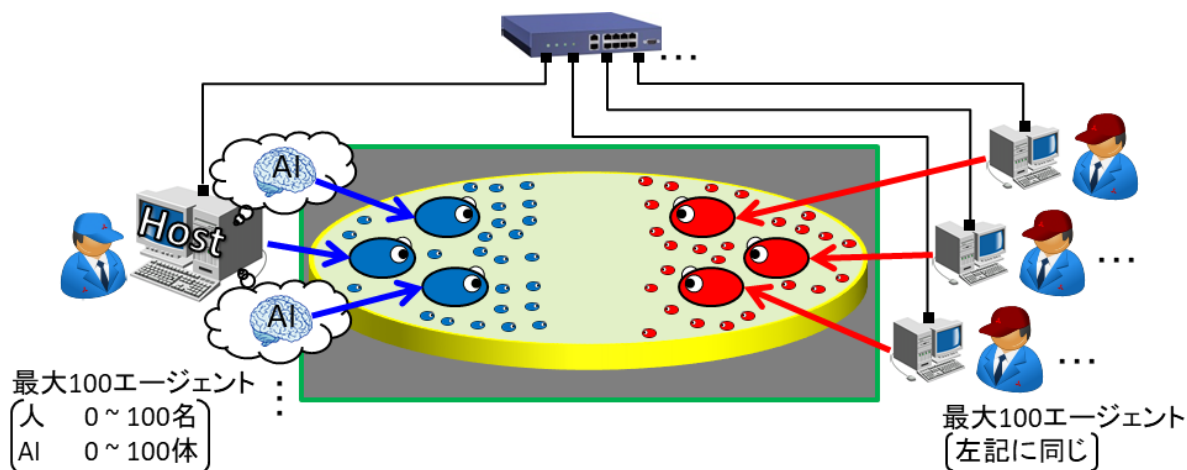


図 3.2.2.2- 1 シミュレータ環境のソフトウェア概念図

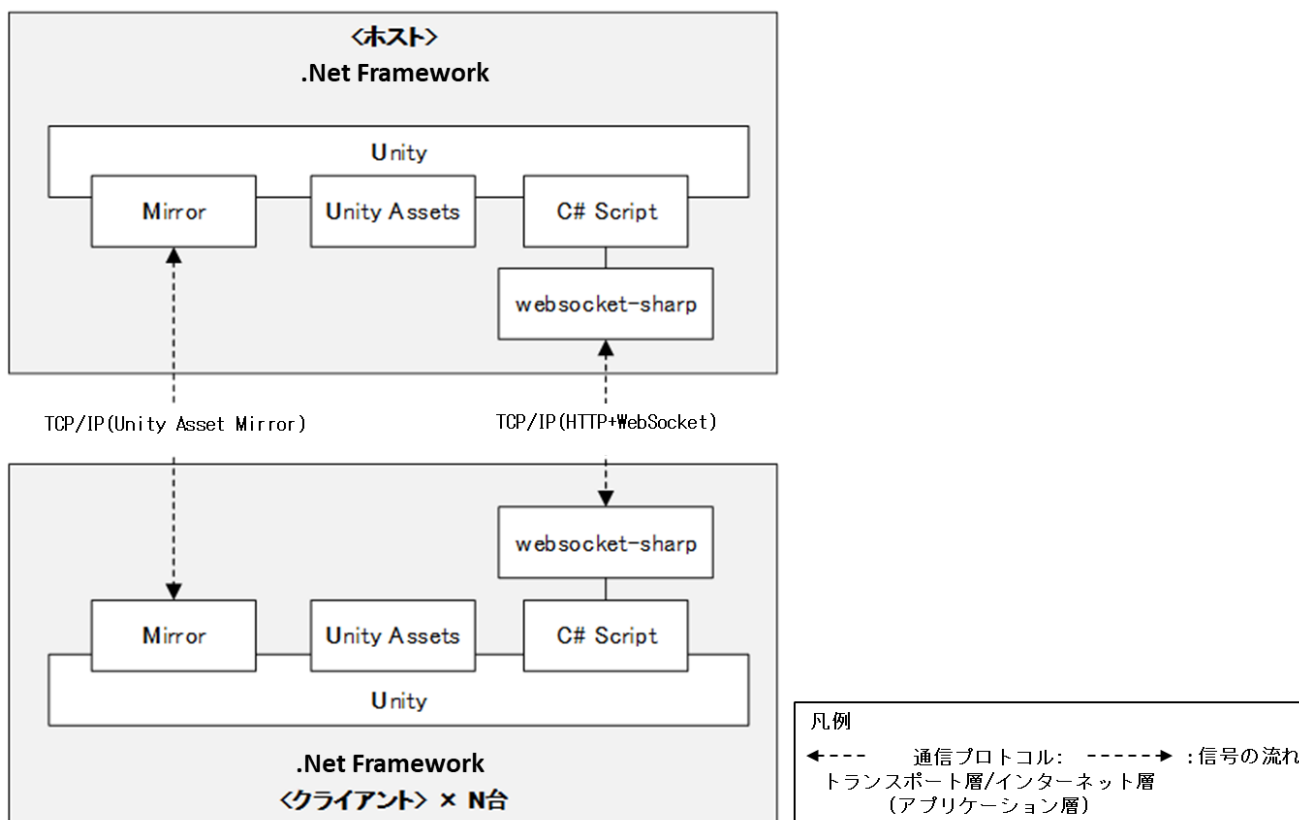


図 3.2.2.2- 2 ソフトウェア構成概念図

表 3.2.2.2- 1 ネットワーク構成

構成	通信プロトコル		ポート番号	備考
	アプリケーション層	トランスポート層/インターネット層		
シミュレーションデータ通信	Unity Asset Mirror	TCP/IP	8081	以下のデータ通信を行う。 ・ 3D オブジェクト毎の位置・姿勢データ ・ 上記以外のゲームスコア等の同期データ ・ 音声チャットデータ
契約ネット・プロトコル通信	HTTP + WebSocket	TCP/IP	8082	契約ネット・プロトコルに関連したエージェント間のメッセージ通信を行う。

表 3.2.2.2- 2 使用した開発環境/ライブラリ/Unity Asset の構成一覧

分類	名称	バージョン	機能/用途	ライセンス
開発環境	Unity	v2019.1.14f1	ゲームエンジン及びその開発環境	商用
開発環境	Microsoft Visual Studio Professional	2017	プログラミング言語 C# スクリプトの開発環境	商用
ライブラリ	Microsoft .NET Framework Developer Pack	v4.7.1	Windows OS で動作するアプリケーションの基盤ライブラリ群	—
ライブラリ	websocket-sharp	2020/03/06 時点の公式リポジトリ最新版	Unity でも利用可能なリアルタイム性の高い通信を実現できるライブラリ、エージェント間のメッセージ通信に使用	MIT License
Unity Asset	Behavior Designer	v1.6.4	AI エージェントのビヘイビアツリー開発に使用	税別 \$80
Unity Asset	Mirror	v3.17.4	Unity のゲームシーン等を通信によって複数台のパソコンへ共有可能にする Asset	MIT License
Unity Asset	Dissonance Voice Chat	v6.4.5	ゲーム内のボイスチャットを可能にする Asset	税別 \$75
Unity Asset	Dissonance for Mirror Networking	v6.4.5	Dissonance Voice Chat の音声データを Mirror による通信で共有可能にする Asset	—
Unity Asset	Smooth Sync	v3.22	オブジェクト位置をネットワーク上でスムーズかつ正確に同期するための補間または外挿処理を可能にする Asset	税別 \$15
Unity Asset	FairyGUI	v3.5.1	Unity 上での画面開発に使用	MIT License
Unity Asset	DOTween	v1.2.335	エージェント硬直を知らせる点滅表示アニメーション付与に使用	—
Unity Asset	SimpleSelect	v1.1.1	人が統合情報画面から指示を出す AI エージェントを選択する機能に使用	税別 \$5
Unity Asset	Network Lobby	v2.1	ゲーム開始前に参加者が待機するロビー画面の開発に使用	—
Unity Asset	Unity ParticlePack	v1.4	エージェント退場時の爆発エフェクトに使用	—
Unity Asset	Unity Standard Assets	v1.1.5	人が操作する人型エージェント（オブジェクト名：Ethan）を使用	—
Unity Asset	Unity Space Robot Kyle	v1.0	AI が操作するロボット型エージェント（オブジェクト名：Kyle）を使用	—

3. 2. 3 エージェント間交渉の枠組み

3. 2. 3. 1 背景

例題設定であるしっぽ鬼では互いのエージェントの能力は同一であるため、1vs1では状況が進展せず味方同士のチームワークが重要となる設定としている。効果的なチームワークのための要素の1つとして、相手の行動を含む状況を観測しながらチームとしての目標（大域目標）を達成するために適応戦略を動的に更新する能力が求められる。一方、本研究ではAIエージェントは自律的に行動し、それぞれの局所的な状況に基づく個々の戦略（局所的な戦略）を有している。例えば、大域目標が特定の場所に到達することであっても、差し迫った危険があるエージェントはそれを避けることを優先すべきである。一方、差し迫った危険にさらされていないエージェントは大域目標達成のため、目標位置に移動する必要がある。

少数の人間がAIエージェントに指示を与える中央集権型のアプローチは、エージェントの数が増えるにしたがって非効率になる。また現実的には、人間の認知能力の限界やネットワークの帯域の問題から、エージェントの数が増えるにしたがって各エージェントが扱える情報の偏りが発生する。このスケーラビリティの問題と情報の不均衡について、個々の局所的な戦略と大域目標達成のための戦略を整合させるため分散制御型のアプローチを採用する。

本研究で提案するエージェント間交渉の仕組みとして契約ネット・プロトコル (Contract Net Protocol, CNP) 及びビヘイビアツリー (Behavior Tree, BT) の2つの手法を組み合わせることとした。以降にBT及びCNPの説明を行い、次に2つの手法を組み合わせさせたエージェント間交渉ソフトウェアと、人間がエージェント間交渉に加わるための工夫点について述べる。

3. 2. 3. 2 ビヘイビアツリー

BT[10]は自律エージェントのふるまいをツリー状に表現する手法で、階層的な意思決定を可視的に定義することが出来る。ゲーム業界で考案され、ゲーム開発コミュニティに広く採用されている。BTは一般的に、NPC (Non-Playable Character) をモデル化するために、スクリプトやFSM (Finite State Machine) に代わって使用されている。

図3.2.3.2-1にBTの表現例としてPick & Placeタスクを示す。Pick & Placeタスクでは対象物 (Ball) を発見し、対象物を掴んで、その後目標位置に置くという一連のふるまいを定義している。一般的にBTは左から右へと評価をするため、左に配置されているふるまいが優先順位の高いものとみなされる。また、BTはモジュール性が高く、各ふるまいはSub-BTとしての展開を設定できる。図3.2.3.2-2にPick BallのSub-BTの例を示し、代表的なBTノードの説明を表3.2.3.2-1に示す。Pick Ballのふるまいでは、対象物が近くになれば近づき、その後対象物を掴んでいなければ対象物を掴むというふるまいに分解して定義している。

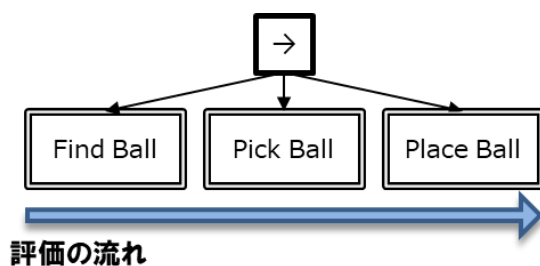


図 3.2.3.2-1 BT の表現例 (Pick & Place)

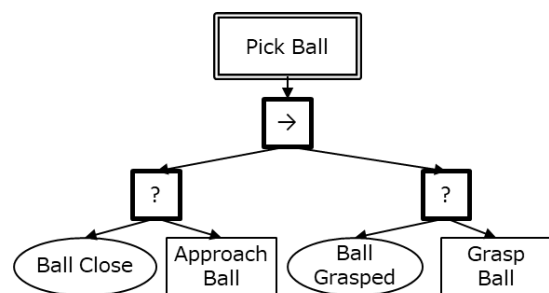






図 3.2.3.2-2 Sub-BT の展開例 (Pick Ball)

表 3.2.3.2- 1 代表的な BT ノードの説明

Node Type	Symbol	Succeeds	Fails	Running
Sequence		If all children succeed	If one child fails	If one child returns running
Selector		If one child succeeds	If all children fail	If one child returns running
Action		Upon Completion	If impossible to complete	During completion
Condition		If true	If false	Never

上記のように、BT はリアルタイムで制約のある複雑なアプリケーションの開発に対して以下の様な特徴がある。

【BT のメリット】

- 人間が理解しやすい
- モジュール性があり、再利用性が高い
- 階層構造をもつ
- 変更箇所による影響範囲がわかりやすい

【BT のデメリット】

- BT エンジンの実装は複雑
- 左から右へ全ての条件を評価するため、計算コストが高い

以上のメリットにあるように BT は人間が理解しやすいため、人間が瞬時にエージェントのアクションの意図と目的を理解するのに大きく寄与する。また、デメリットについても Unity 及び Behavior Designer といった既製品を利用することで複雑な BT エンジンの実装を省略することができ、計算コストも実用上問題ないレベルであることを確認した。

3. 2. 3. 3 契約ネット・プロトコル

CNP[11]は複雑なタスクを独立したタスクに分解し、それらを複数のエージェントに割り当てるためのプロトコルであり、現実の契約交渉に類似している。

CNPにおける契約交渉の流れを図3.2.3.3- 1に示す。CNPではエージェント（マネージャ）が初めにタスク告示メッセージを送信する。次に、告示されたタスクを実行できるエージェントは、そのタスクに対する入札メッセージをマネージャに送信する。マネージャは受け取った入札メッセージを評価し、適切なエージェントを選択し、そのエージェントに落札メッセージを送信する。落札メッセージを受け取ったエージェントは落札者と呼ばれる。最後に、マネージャは落札者から実行結果の詳細を示す実施報告を受け取る。

CNPの重要な特徴はマネージャ及び落札者の相互選択性である。すなわち、契約を締結する際にマネージャがどのエージェントを落札者として選定するか、入札を行おうとするエージェントは入札行為を行うか否か、各々の価値基準に基づいて意思決定をすることができる。この相互選択性は、本研究の自律エージェントに対して分散制御型のアプローチを採用するという趣旨に適合している。

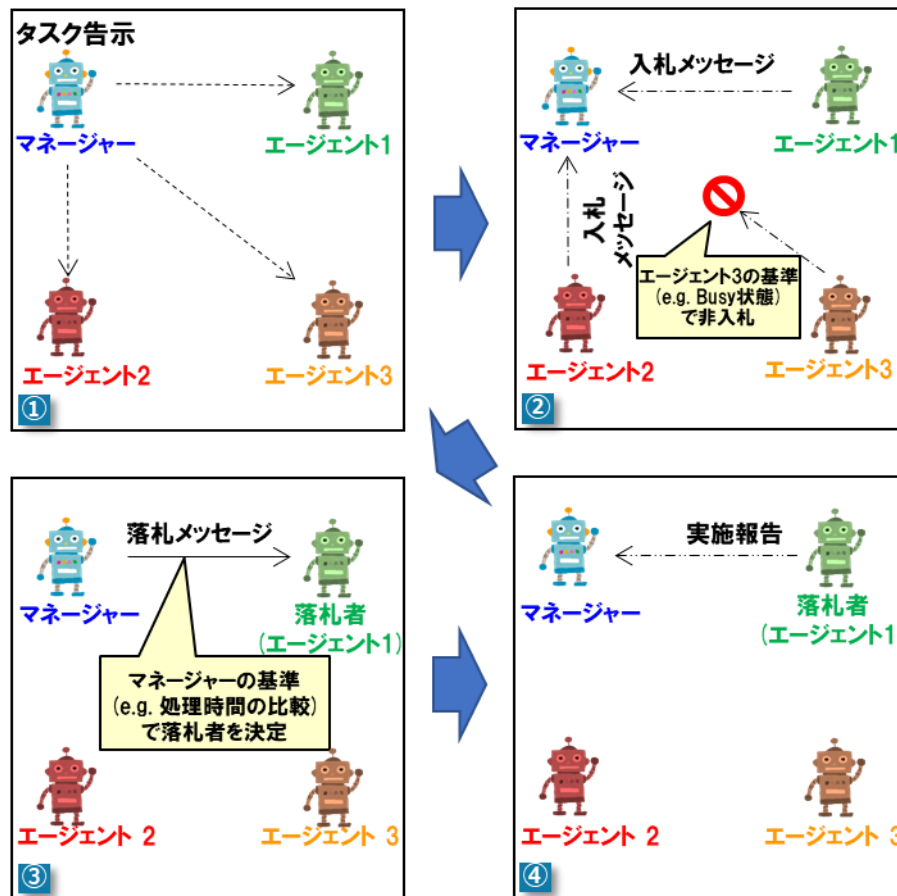


図 3.2.3.3- 1 CNP による交渉の流れ

3. 2. 3. 4 エージェント間交渉ソフトウェア

自律エージェントの実施するタスクの分類として協調タスクと個別タスクの2つを定義した。協調タスクは、大域目標達成のために、エージェントが協調して実施するタスクとして定義し、CNPによる契約によって駆動される。一方、個別タスクは個々のエージェントの置かれている状況に応じた局所的戦略の達成のために実施するタスクとして定義し、個々のエージェントの裁量によって実行される。さらに、人間とAIの協調を考える上では、人間に理解しやすい形で協調タスク及び個別タスクが定義されている必要がある。

ここでショベルカーとダンプカーの協調例を挙げ、エージェント間交渉ソフトウェアの概念を図3.2.3.4- 1に示す。個別タスクには自己の安全を守るSafety Taskと局所的な目標達成のためのSingle Taskを設定している。Safety Taskは「衝突防止」といった基本的なタスクを定義しており、Single Taskについては個々の能力や状況に応じてショベルカーは「土砂切削」、ダンプカーは「移動」を定義する。協調タスクではショベルカーは土砂を「ダンプへ積載」し、ダンプカーは「積載位置で待機」するよう定義する。ショベルカーとダンプカーの組み合わせや積載位置の決定といった協調に必要な事項はCNPによって交渉され、それぞれに協調タスクが割り当てられる。

以上のようにBTを用いて各タスクを人間にも理解しやすいように階層的に定義し、協調タスクについてはCNPによる契約で駆動する枠組みがエージェント交渉の骨子である。

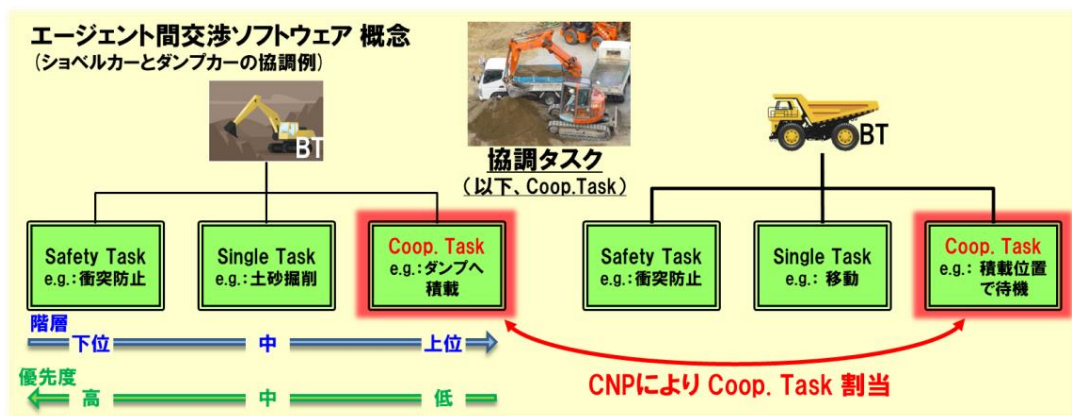


図 3. 2. 3. 4- 1 エージェント間交渉ソフトウェアの概念図

3. 2. 3. 5 人がエージェント間交渉に加わるための工夫点

前項ではBTの導入によってAIエージェント側の意図や戦略の把握を可能にし、CNPの導入によって情報伝達を可能にする枠組みを説明した。一方、一般的にAIと人間には身体性/認知能力に大きな差異がある。AIは高速かつ正確な応答が可能であり、人間は状況の認知/適応能力に優れている。そのため、人間がエージェント間交渉に加わる場合、人間の応答性を考慮した交渉が必要となる上、極力AIエージェントの迅速な応答を阻害しないための仕組みが必要となる。加えて、時々刻々と状況が変化する環境において、戦略を見直し、新たなタスクの再割り当てが可能である必要もある。本課題を解決するため、CNPに対し、以下に示す工夫を行った。

全てのエージェントはマネージャAI及び落札者AIの機能を有するとし、マネージャAIのフローチャートを図3. 2. 3. 5- 1に、落札者AIのフローチャートを図3. 2. 3. 5- 2に示す。

マネージャAIの工夫点は以下の通りである。

- ・ 工夫点 1：マネージャAIは協調タスクを選択するとき、協調タスクの種類に応じてそれぞれの入札期限(T_Max)を定義可能とする。特に人間の入札を想定する協調タスクはT_Maxを十分長く設定することで、人間の応答性を考慮した交渉を可能にする。
- ・ 工夫点 2：人間が参加することを考慮し、落札メッセージの前に落札の可能性が無いエージェントは早く開放するよう落選メッセージを送ることでAIエージェントの迅速な応答を阻害しないようにする。
- ・ 工夫点 3：協調タスクが進捗しないと想定される場合、マネージャ側は協調タスク中止を判断できるようにし、タスクの再割り当てを可能にすることで人間及びAIエージェントの戦略を柔軟に変更可能にする。

落札者AIの工夫点は以下のとおりである。

- ・ 工夫点 1：協調タスクに参加しない場合、自身とマネージャを必要以上に待たせず解放するために非入札メッセージを送信することで、人間の応答性を考慮した交渉を可能にする。
- ・ 工夫点 2：マネージャAIの工夫点 2と同様に、落選メッセージを受領した場合には自身が別の協調タスクに入札できるよう解放することでAIエージェントの迅速な応答を阻害しないようにする。
- ・ 工夫点 3：マネージャAIの工夫点 3と同様に、協調タスクが進捗しないと想定される場合、落札者側も協調タスク中止を判断できるようにし、別の協調タスクへの入札を可能にすることで人間及びAIエージェントの戦略を柔軟に変更可能にする。

以上の工夫により、人間がエージェント間交渉に加わった場合にも、人間及びAIエー

ジェントの身体性/認知能力のそれぞれの優位点を生かす仕組みとした。この内容は特願2019-158885号として特許出願中である。

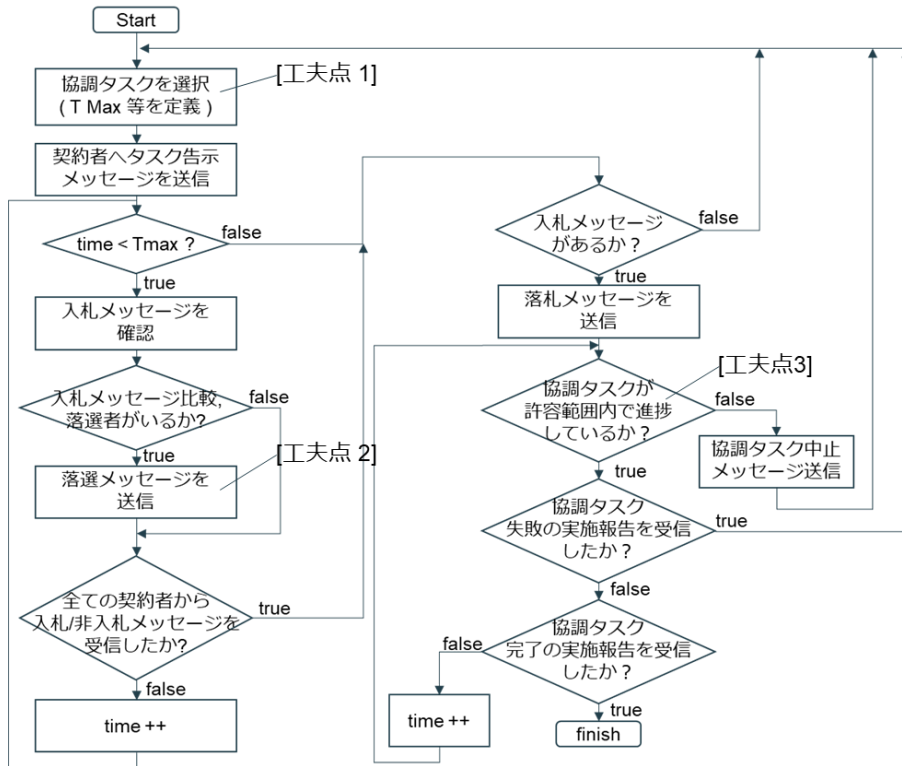


図 3.2.3.5- 1 マネージャ AI のフローチャート (特願 2019-158885 号)

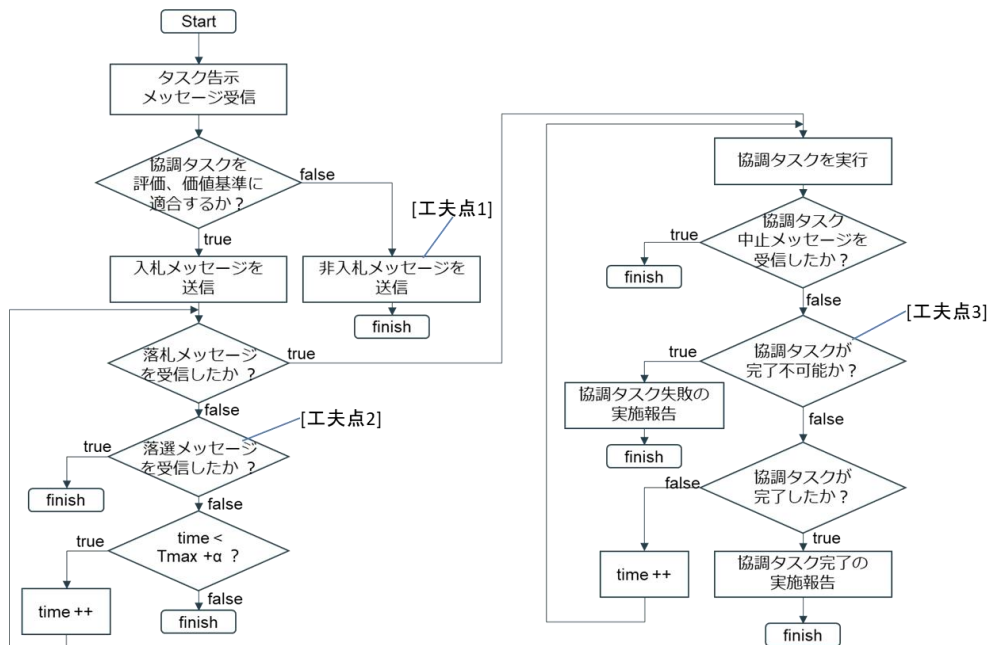


図 3.2.3.5- 2 落札者 AI のフローチャート (特願 2019-158885 号)

3. 2. 4 人とAIエージェントの協調行動の評価方法及び基準

本項では、人とAIエージェントの協調行動の評価方法及び基準に関する検討結果を示す。

人とAIエージェントの協調行動を実現するアーキテクチャ及びHSIに求められる機能、性能を定量的に定義することは困難であるが、しっぽ鬼のような勝敗で優劣が決まる競争的なゲームまたはミッションにおいては、最終的には、ゲームまたはミッションの成功率、即ち、相手のチームに勝利すること、勝率が高いことが有効な指標となる。したがって、チームAが相手チームに高い勝率で勝利することを示すことができれば、そのチームのアーキテクチャとHSIについて、その機能や性能の有効性を間接的に示すことができる。また、チーム戦によるしっぽ鬼では“制限時間内に相手を捕まえた数が多いチーム”が勝利を収めるため、勝率のみならず、チームメンバの生存率または捕獲率によって、どの程度優位にゲームまたはミッションを進めることができたのか、その度合いを測ることも重要と考える。生存率はゲーム終了時点においてチーム内の全エージェント数に対する生存エージェント数の比率である。一方、捕獲率はゲーム終了時点において相手チームの全エージェント数に対する相手チームの被捕獲エージェント数の比率であり、 $(\text{捕獲率}) = 1 - (\text{相手チームの生存率})$ の関係により生存率と捕獲率は同様の意味合いがある。

これらの指標は、チームAの能力と相手チームの能力との相対関係により決まるため、チームAが高い能力を有することを示すためには、相手チームも高い能力を有する必要がある点に注意が必要である。しかしながら、相手チームとして、どの程度の能力を保有すれば良いのか、定量的な目標値を設定することは困難である。そこで、相手チームのメンバは全て人とし、“人間の能力を有する”構成とした。このチームBにチームAが勝利することで、人間のチーム以上の能力を発揮できるアーキテクチャ及びHSIであることを示すことができる。

人とAIエージェントが協働する上では、人はAIエージェントとのインタラクションを何らかのインターフェースを介して行う必要があるとあり、当該インターフェースの機能・性能、即ち、効率的かつ効果的にインタラクションが行えるかなどの評価が人とAIエージェントの協調行動の評価として有効であると考えられる。これには、特定のタスクに対するシステムの操作性を評価するための指標であるクーパー・ハーパー・レーティング (CHR) [12] が有効と考える。この指標は本来、航空機の飛行性能や操作性（操縦性）を評価する指標として使用されているものであるが、B. Donmez [13] らによって、無人システム向けの操作性や状況監視等への適用を想定した改良が提案されている。本研究においても、B. Donmezらの手法“MCH-UVD”を適用することとした。

以上より、人とAIエージェントの協調行動を評価する基準となる指標として、以下の3点を用いることとした。また、対戦に参加した人から意見を得ることにより、定性的な傾向の確認が可能と考えることから、積極的にアンケートやヒアリングを行うこととした。

- ・ 勝率
- ・ 生存率(*1)
- ・ MCH-UVD に基づくHSIのレーティング結果

*1: 3. 3項 ~ 3. 5項に示す、各スパイラルの結果の説明においては捕獲率ではなく生存率のみを示す。

3. 3 スパイラル#1

3. 3. 1 AIエージェントの構築

3. 3. 1. 1 人とAIが効果的に分担可能なタスクの仕様の設定

本項では、人とAIエージェントが効果的に分担可能なタスクの仕様について、検討結果を示す。

しっぽ鬼では、以下のルールを設定している。

- ・ 両チーム共に、相手の背中を捕り合う。
- ・ ゲームの点数が多い方が勝利する。
- ・ 相手の背中を捕った（接触した）場合に、点数が得られると共に、背中を捕られたエージェントはゲームから退場する。
- ・ 味方同士のエージェントが接触した場合、または相手チームのエージェントの背中以外に接触した場合、一定距離、後方へ跳ね飛ばされた後、一定時間動くことができなくなる。
- ・ 自分の視野内の相手のみ自分から観測可能である。
- ・ 味方が観測した相手の位置、進行方向は、どの味方にも共有される。

本ルール下において、しっぽ鬼に参加する各エージェントが実施する行動、即ちエージェントが行うべきタスクとしては、以下の選択肢が考えられる。これらのタスクは、しっぽ鬼の大域目標である“相手チームに勝利すること”を達成するために、各エージェントが誰の指示も受けることなく、各自の判断に基づき実施すべきものであると言える。

- ①相手を探して見つける。
- ②相手を追いかけて、背中を触る
- ③相手から逃げる
- ④ゲームエリアの端を避ける（ゲームエリアから外れ、退場することを避けるため）
- ⑤味方を避ける（味方と衝突して動けなくなることを避けるため）

また、各自の判断で実施するタスク以外に、例えば“相手を探して見つける”タスクを分担して、重複を最小化し色々な場所を探しにいくことができれば、効率的に相手を探することができる。さらに、相手を見つけた際に、“相手を追いかけて捕まえる”タスクを、味方と協力して実施すれば、より捕獲が成功する確率も上がるはずである。明らかに劣勢な状況においては、誰かの号令と共に、一斉に逃げ出し、再度作戦を練り直すことも有効であると考えられる。

このように上記のうち、①～③のタスクは、味方と協力して行うことでより効果的な作戦が可能となると言える。これらのタスクは、大域判断に基づき、各エージェントを統制して行うことで効果を発揮することから、大局的に統制するタスクであるとも言える。一方で④、⑤は他のエージェントの行動や情報に依存することなく、自身の判断のみで実行可能なタスクであり、不確実性の少ない短期的かつ局所的判断に基づくタスクといえる。

以上より、不確実性の少ない短期的かつ局所的判断に基づくタスク（3. 2. 3. 4項で個別タスクとして定義）と、人が得意とするそれらを大局的に統制するタスク（3. 2. 3. 4項で協調タスクとして定義）を、表3.3.1.1- 1の通り整理を行った。

表 3.3.1.1- 1 エージェントのタスク一覧

タスク分類	タスク名	タスク概要	取る行動
個別タスク	搜索（個別）	相手エージェントを搜索するタスク	ゲームエリア内をランダムに移動する。
	捕獲（個別）	相手エージェントを追いかけて捕獲するタスク	いずれかの味方エージェントが発見した相手エージェントへ移動し、背後に回り込み、背中に触れる。
	衝突回避	味方エージェントとの衝突を避けるタスク	味方エージェントとの衝突を避けられる移動方向に変更する。
	エリア逸脱回避	ゲームエリア外に出ることを回避するタスク	ゲームエリアの中央方向へ移動する。
協調タスク	搜索	任意のエリアで、味方エージェントと協力して相手エージェントを搜索するタスク	人から指示されたエリアに移動し、エリア到着後は、エリア内をランダムに搜索する。
	捕獲	味方エージェントと協力して相手エージェントを追いかけて捕獲するタスク	相手エージェントを発見したエージェントから指示された相手エージェントの背後に回り込んで背中を触る。
	離脱	集団で相手エージェントから任意のエリアに回避・離脱するタスク	搜索と同様。

タスクのうち、個別タスクは、各エージェントが自律的な行動として実施すべきタスクであるため、BTに基づき実行する仕組みとしている。

一方、協調タスクは、誰かがタスクのマネージャとなり、協調を行うことが求められる。そこで、誰がマネージャとなって協調を行うと効果的かを次のように検討した。

まず搜索については、相手の位置が分からない状況で主に行うタスクであり、不確実性が高く、何人で探すか、どの位置を探すか等、大局的な判断を要するタスクと考える。また離脱についても、各エージェントが観測する局所的判断のみならず、ゲームエリア全体で起きている各所の状況も照らし合わせて判断する必要があり、大局的な判断に位置づけられるタスクと考える。一方捕獲タスクは、各エージェントが相手エージェントを見つけたことを契機として発生する可能性が高いタスクであり、迅速かつ局所的な状況も踏まえ、捕獲可否や味方への支援要請などを行うべきタスクであると考え。1. 項 (3) ②-1 では、不確実性の少ない短期的かつ局所的判断に基づくタスクはAIに優位性があり、大局的に統制するタスクは人の方が優れているとの仮説について述べた。この仮説に従い、大局的な判断を要する搜索タスク及び離脱タスクについては、人がマネージャとなり、人からAIエージェントにタスクの依頼を行うこととした。一方、捕獲タスクについては、AIエージェントがマネージャとなり、当該AIエージェントから他のAIエージェントにタスクの依頼を行うものとした。なお、これら、タスクの依頼は、3. 2. 3. 3 項に示すCNPによる交渉において、タスク告示という形で他のエージェントへ告示され、タスクに参加できる状況に置かれたエージェントが、入札、落札という契約行為を経て、協調タスクとして実行される。

これらのAIエージェントは、人とAIエージェントが協働するチームのメンバーとしての用途の他、当該チームのトレーニング時における対戦相手としても求められる。ただ

し、厳密にどのような対戦相手がトレーニング時に適しているか定義を行うことは難しいため、スパイラル#1では、対戦相手のAIエージェントは、大局的な統制による協調タスクを有せず、個別タスクのみを実施するエージェントとしてモデル化を行うこととした。

3. 3. 1. 2 エージェントにおける適切なタスクの条件及び価値基準の設定

本項では、3. 3. 1. 1項で検討したタスク仕様にに基づき、各エージェントへ望ましいタスクの割り当てを行うための条件及び価値基準について、検討した結果を示す。

しっぽ鬼のような競争環境では、各エージェントが時々刻々変化する状況の中で、様々なタスクを臨機応変に実行してゲームを優位に進めていく必要がある。そこで、どのような状況で、どのタスクを実施すべきであるか、即ち各タスクを実施する条件とタスクの優先順位が重要となる。

また、協調タスクにおいては、タスクの告示を受けたエージェントが、各自が実施する個別タスクとの優先順位も踏まえ、入札を行うか否か判断する必要がある他、入札者が多数の場合、タスクを告示したエージェントが、多数の入札者の中からどのような基準でタスクを実際に担当するエージェント（落札者）を選定するかということも重要となる。

そこで、まず個別タスクの実施条件及び協調タスクのタスクを告示する条件や、タスク告示を受けたエージェントが入札する価値基準、タスク告示を行ったエージェントが落札者を選ぶ価値基準について、表3. 3. 1. 2- 1及び表3. 3. 1. 2- 2の通り、整理を行った。本表の条件等は、実際に想定される状況に基づき、仮定したものである。

表 3. 3. 1. 2- 1 個別タスクの実施条件

個別タスク名	タスクを実施する条件
搜索	いずれの味方エージェントも、相手エージェントを発見していない場合
捕獲	いずれかの味方エージェントが相手エージェントを発見した場合
エリア逸脱回避	エージェントがゲームエリアの端から一定距離以内となった場合
衝突回避	味方エージェントが自身の移動ルート上且つ近距離に存在した場合

表 3.3.1.2- 2 協調タスクの告示条件、入札／落札者選定の価値基準


協調タスク名	タスクを告示する条件	入札する価値基準	落札者を選定する際の価値基準
探索	人がタスク告示した場合	入札期間に、他に優先度の高い協調タスクを実施中でない	指示されたエリアまでの移動距離が近い順に、入札エージェント数がタスクの上限エージェント数に達するまで、落札できるものとする。
捕獲	相手を発見した場合 (発見したエージェントがタスク告示)	探索タスクと同様	相手エージェントまでの移動距離が近い順に、入札エージェント数がタスクの上限エージェント数に達するまで、落札できるものとする。
離脱	探索タスクと同様	探索タスクと同様	探索タスクと同様

次に、個別タスクと協調タスクの間の優先順位について、ゲーム中に想定される状況に基づき、検討を実施した。

ゲーム中はいかなる時にも味方エージェントを不利な状況に追い込まないようにするため、ゲームエリアからの逸脱や、味方エージェントとの衝突は避ける必要がある。また離脱タスクは、数的に劣勢になっているエージェント群を任意のエリアに退避させる際に有用なタスクであり、探索タスクや捕獲タスクを実施中のエージェントも従った方が効果的であることが予想される。これらのタスクのように、安全性を確保するために必要なタスクは、他のタスクに対し優先的に実施する必要があると考える。また、あるエージェントが探索中に相手エージェントを見つけた場合、発見不確実な探索を継続するよりも、すでに発見した相手エージェントを追いかけて捕獲を行う方が効果的と考えられるため、追跡・捕獲を行うタスクは、探索タスクよりも優先度を高くした方が効果的と考える。

以上の経緯より、各タスクの優先順位を表3.3.1.2- 3の通り設定した。

表 3.3.1.2- 3 タスクの優先順位

優先順位	タスク
優先度高  優先度低	1 エリア逸脱回避
	2 衝突回避
	3 離脱 (協調タスク)
	4 捕獲 (協調タスク)
	5 探索 (協調タスク)
	6 捕獲 (個別タスク)
	7 探索 (個別タスク)

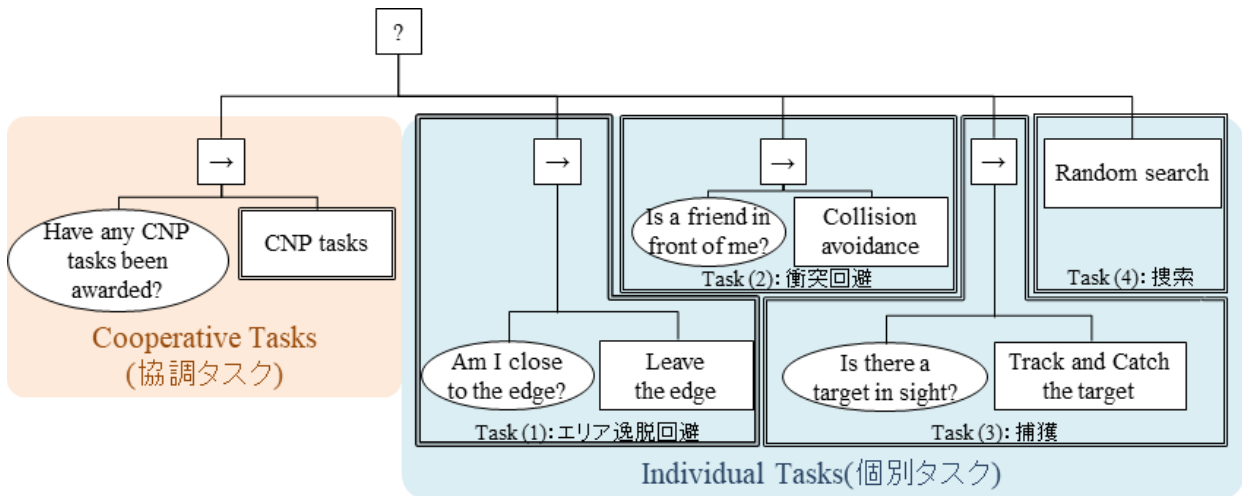
以上の検討結果を踏まえ、スパイラル#1に使用するAIエージェントの行動モデルとして、3.2.3項の枠組み及び3.3.1.1項のタスク仕様も踏まえ、図3.3.1.2-1に示すモデルを設計・試作した。

本モデルは、BTに基づき構築されており、図の左側から条件判定処理が実行されるため、左側に記載した処理から順番に(優先的に)実行される。表3.3.1.2-3タスクの優先順位の優先順位に示す通り、個別タスクのうち、エリア逸脱回避や衝突回避については、どのタスクを実施中であっても、優先的に実施する必要がある。また、エリア逸脱回避と衝突回避以外は、協調タスクの方が、個別タスクよりも優先的に実施する必要

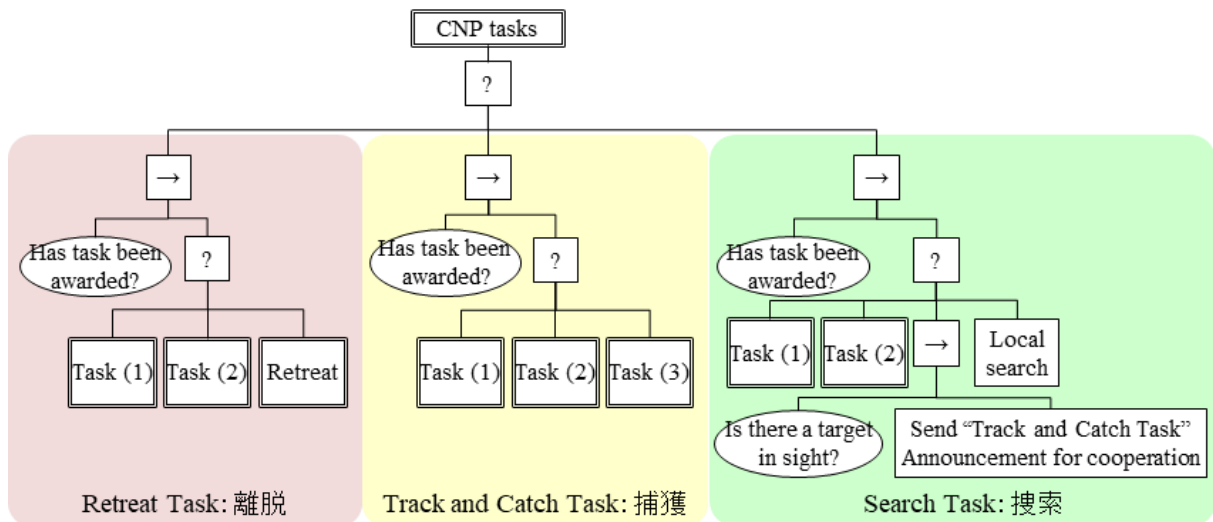
がある。以上より、以下を考慮し、BTを構築した。

- ・協調タスクのBTは、個別タスクのBTよりも左側に配置し、優先的に実行させる構造とする。
- ・個別タスクのBTでは、エリア逸脱回避、衝突回避のタスクを他の個別タスクよりも左側に配置し、優先的に実施させる構造とする。
- ・協調タスクのBTの中に、エリア逸脱回避、衝突回避のタスクを含めると共に、エリア逸脱回避、衝突回避のタスクを各協調タスクよりも左側に配置し、優先的に実施させる構造とする。

最も優先すべき個別タスクであるエリア逸脱回避と衝突回避のタスクを、協調タスク内に再利用して包括した点が工夫点であり、BTでは、個別タスクとして設計した各処理の木構造をそのまま協調タスクの一部として再利用可能である。このような再利用性に優れる点がBTを採用したことの優位点である。



(i) 行動モデル全体



(ii) 協調タスクに関する行動モデル

図 3.3.1.2- 1 AI エージェントの行動モデル

3. 3. 2 HSIプロトタイプの実作

本項では、HSIプロトタイプに関する設計結果及びHSIプロトタイプの実作結果について示す。

HSIプロトタイプの実作に当たっては、まず、本研究と類似性がある、多数のロボットや無人機を人が操作・監視するインターフェースの先行事例の調査を行った。以下に、調査結果として、2件の先行事例を示す。

(1) Research Environment for Supervisory Control of Heterogeneous Unmanned vehicle (RESCHU) [14]

1人のオペレータが6機の無人偵察機を制御して脅威の偵察を行うためのインターフェースであり、偵察対象として複数の目標と脅威地点が設定される。RESCHUの画面を図3.3.2-1に示す。画面は、主に5つの領域から構成されており、各々の用途を表3.3.2-1に示す。

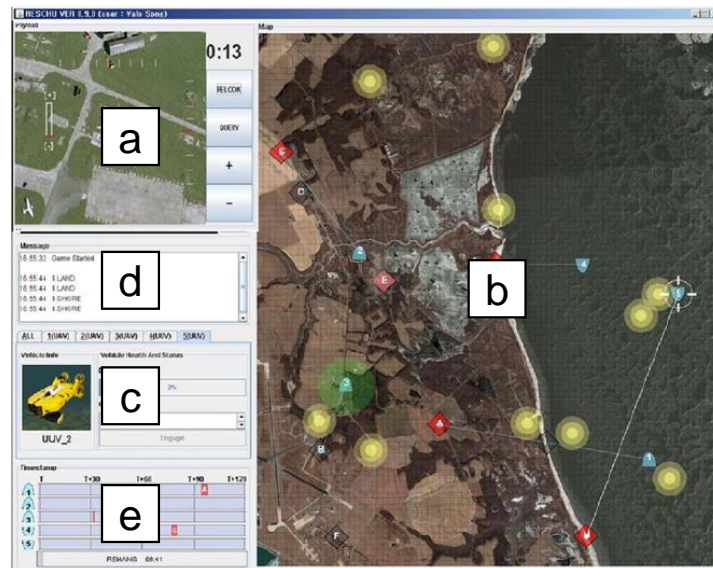


図 3.3.2- 1 RESCHU のインターフェース

表 3.3.2- 1 RESCHU のインターフェースにおける各画面の用途

記号	画面名称	用途
a	Live Camera Feed	無人機のメインセンサーの画像を表示し、周囲の状況を確認する。
b	Map Area	無人機や偵察対象の位置、フィールドの地形等の情報を俯瞰表示し、位置関係等を確認する。
c	Unmanned Vehicle Status	無人機のステータス（搭載センサーのステータスや残燃料量等）を確認する。
d	Communication Channel	無人機との通信状況を確認する。
e	Mission Timeline	偵察対象への到着時刻など、ミッションの推移を表示し、ミッション達成状況を確認する。

(2) Multi-robot Control System(MrCS) [15]

RoboCup Rescueで使用されたインターフェースであり、最大24機のロボットを制御し被災者を発見することがミッションの目的である。MrCSの画面を図3.3.2-2に示す。画面は、主に4つの領域から構成されており、各々の用途を表3.3.2-2に示す。

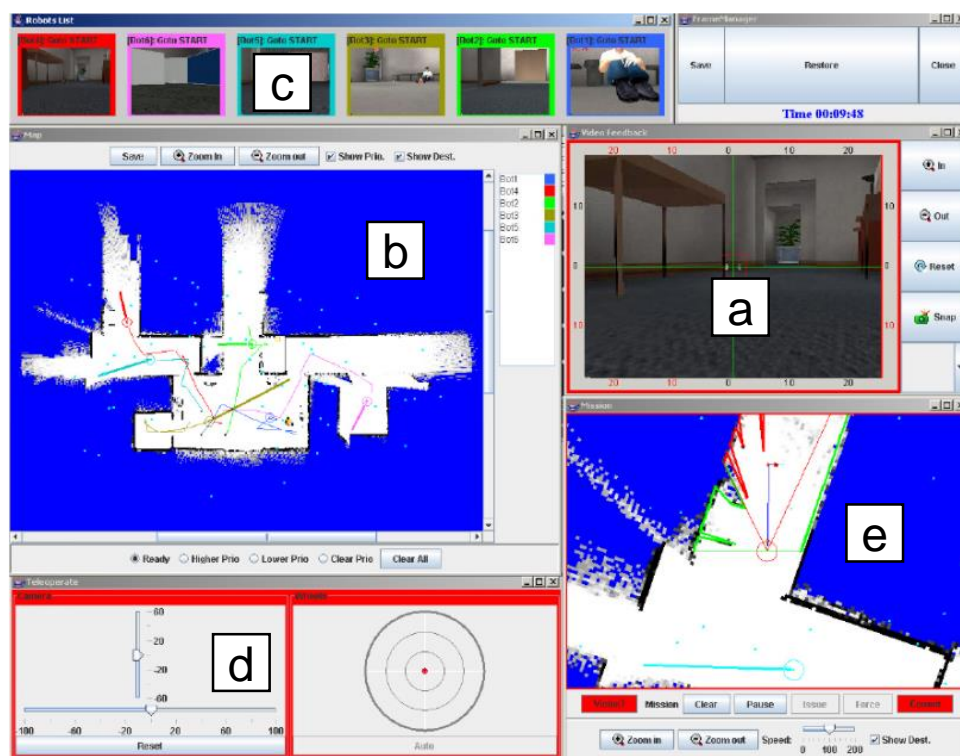


図 3.3.2-2 MrCS のインターフェース

表 3.3.2-2 MrCS のインターフェースにおける各画面の用途

記号	画面名称	用途
a	Video Feed	操作対象のロボットの選択メインセンサーの画像を表示し、周囲の状況を確認する。
b	Map	マップ上におけるロボットの位置や移動ルートを確認表示し、位置関係等を確認する。
c	Robot List	操作対象のロボットの選択を行うと共に、各ロボットのメインセンサーの画像を表示し、各ロボット及び周辺の状況を確認する。
d	Teleoperate	操作対象を操作する際の操作量を設定する。
e	Mission	被災者の発見状況等を表示し、ミッションの進捗状況を確認する。

いずれの先行事例も、1人の人が遠隔地にある多数のエージェント(ここではロボット)を指揮統制するアプリケーションであり、同一環境において人とAI群が協働する本研究とは異なる部分もあるが、動的環境における人とAI群間のインターフェースの事例として有用性が高いと考える。したがって、スパイラル#1最初のHSIプロトタイプとして、これらの先行事例を参考にすることとした。いずれの先行事例においても、以下の5つの要素からなる画面構成としており、しっぽ鬼におけるHSIプロトタイプは、これらの要素を含んだ画面構成とした。

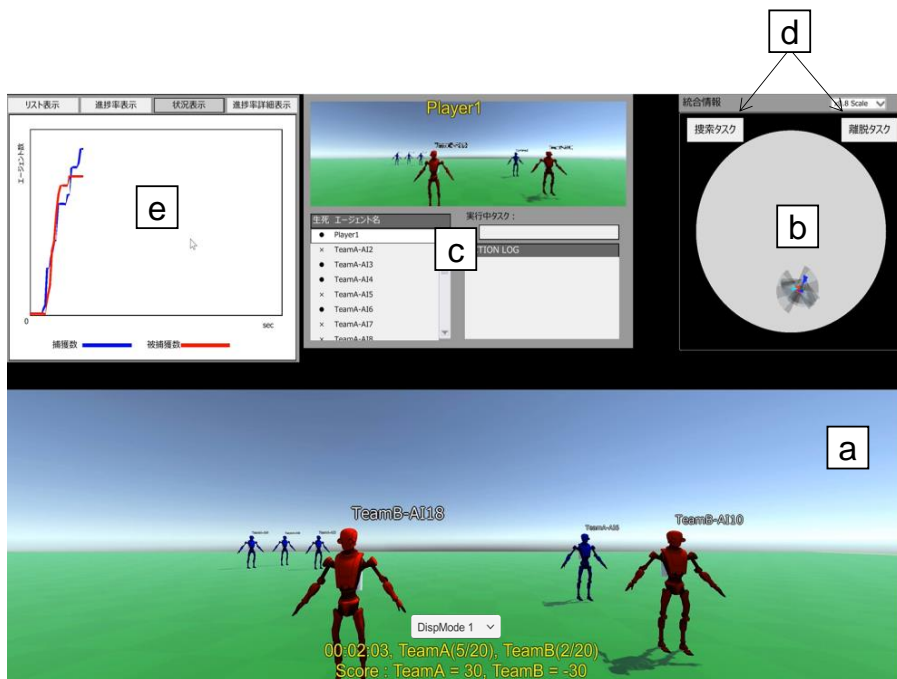
- ・操作対象のロボット／無人機のメインセンサーの映像・画像を表示し、オペレータ視点で周囲の状況を確認する画面
- ・ミッションの対象フィールドのマップ及びマップ上のロボット／無人機の位置や移動経路を確認するための画面
- ・ロボット／無人機のステータスを確認する画面
- ・ロボット／無人機に指示をする／操作を行う画面
- ・ロボット／無人機のミッションの達成状況・進捗状況を確認する画面

上記画面構成に基づき、作成したしっぽ鬼のHSIプロトタイプを図3.3.2-3に示す。また、各サブ画面の表示内容と用途を表3.3.2-3に示す。本HSIプロトタイプでは、3.3.2.3項にて述べるタスクのマーケット状況等の効果的な人への提示方法の確認のため、以下の2つの表示形式を用意し、人とAIエージェントのチームに参加する人が、どちらの表示形式を使用するか選択できるようにした。

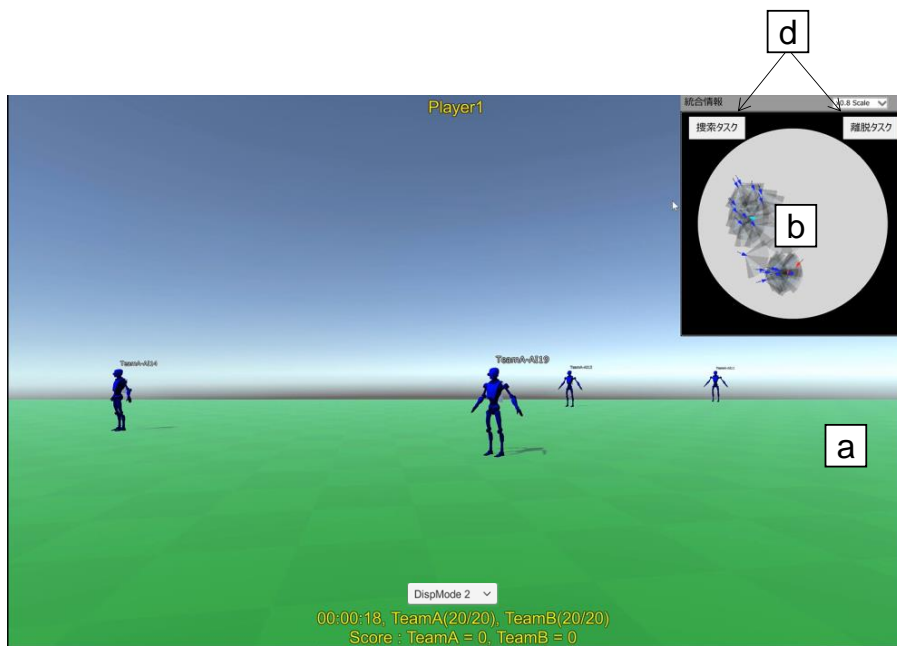
- ・DisplayMode 1: 全サブ画面を表示する形式
- ・DisplayMode 2: a, b, dのサブ画面を表示する形式

また、人が自身のエージェントを操作するインターフェースとして、ゲームコントローラまたはキーボードを使用することとした。

使用可能なインターフェースを表3.3.2-4にまとめる。



(i) DisplayMode 1



(ii) DisplayMode 2

図 3.3.2- 3 しっぽ鬼の画面構成

表 3.3.2- 3 しっぽ鬼のインターフェースにおける各サブ画面の表示内容と用途

記号	表示内容	用途
a	操作対象のエージェントの一人称視点	操作対象エージェントの外界視界を表示し、エージェントの周囲の状況を確認するために使用する。
b	統合情報表示	味方エージェント及び発見した相手エージェントの現在位置を俯瞰的に表示し、味方と相手の位置関係や味方の将来行動予測を把握するために使用する。
c	選択した味方エージェントのステータス表示	選択したエージェントの外界視界、ステータスを確認するために使用する。
d	契約ネットプロトコルメッセージ確認/送信	協調タスクのエージェント間交渉に必要なタスク告示を行うために使用する。 細部は、3. 3. 2. 1項にて述べる。
e	ミッション及びエージェント間交渉の達成状況・進捗状況の情報表示	味方と相手のエージェント数の推移や、協調タスクの契約状況、協調タスクに参加するエージェントのタスク進捗状況を確認するために使用する。 細部は3. 3. 2. 3項にて述べる。

表 3.3.2- 4 使用可能なインターフェース
(○印は使用可能であることを意味する)

インターフェース	所属チーム		備考	
	チームA	チームB (参考)		
画面	a	○	○	
	b	○	○	
	c	○*1	—	
	d	○	—	
	e	○*1	—	
ゲームコントローラ	○	○	エージェントのアクションは、 3. 2. 1. 4項 参照	
キーボード	○	○		
マウス	○	—		
ヘッドセット	—	○*2		

*1: Display Mode 1 の場合のみ

*2: メンバ同士で音声チャットによるコミュニケーションが可能

3. 3. 2. 1 人からAIエージェント群全体に対する効果的な指示方法の解明

本項では、HSIプロトタイプの問題である、人からAIエージェント群全体に対する効果的な指示方法の解明の一貫として、人からAIエージェント群に対して、CNPに基づくタスク告示や契約交渉の流れに基づき、表示操作系の設計を行った結果及びHSIプロトタイプの内容を示す。

CNPにおける契約交渉の流れを図3.2.3.3- 1 に示したが、マネージャが複数のエージェントにタスクの実行を依頼したい場合、そのすべてのエージェントに対し、タスクの告示を行う必要がある。この時、マネージャがエージェント1体1体に対して、タスクを告示する場合、エージェント数の増加に伴い、操作する回数も比例増加してしまうという課題がある。また、タスクに関する情報量が多いほど、操作する回数も増加することとなる。したがって、極力、エージェント数、情報量に依存しないインターフェースとする必要がある。

上記経緯より、以下の設計方針に基づき、人からAIエージェントへの指示を行うインターフェースを設計することとした。

- ・同時に複数のAIエージェントにタスク告示を行う方法とする。
- ・タスクに関する情報（入札期限、落札上限エージェント数）は、予めデフォルト値を規定しておくことで省力化する。
- ・極力、1画面のみで、操作が完結するようにする。

同時に複数のAIエージェントにタスクを告示するためには、告示対象とするAIエージェントを一度に複数指定できることが重要となるが、これは、リアルタイムストラテジーゲームにおいて広く使用される、マウスによる範囲選択で複数エージェントを指定する方法とした。

タスクの告示時にエージェントに伝達するタスクの情報や、契約に関する情報としては、以下の情報がある。本情報の中で、告示対象とするタスク及び検索エリアの情報以外は、予めデフォルト値を設定することで、最小限の情報入力としつつ、人が必要に応じて適宜変更可能とした。

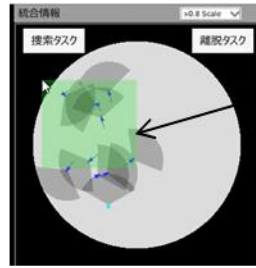
- 告示対象とするタスク
- 告示するタスクに必要な情報
 - ・検索エリアの位置及び範囲、または離脱先のエリアの位置及び範囲
 - ・タスクの実施期間
 - ・入札期限
 - ・落札上限エージェント数

また、これらの操作に必要なインターフェースは、統合情報表示画面に集約することで、複数の画面に渡って操作を要しないよう配慮した。

さらに、落札者を選定する際は、予め合理的な落札の優先順位を規定することで人の判断を介する必要をなくし、契約プロセスに要する時間を短くした。

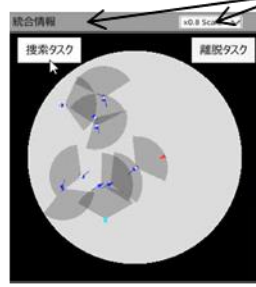
図3.3.2.1- 1 に、実際に設計及びプロトタイプを行ったタスク告示から契約までの流れを示す。本HSIプロトタイプにより、対象エージェント数に関わらず最短で、わずか4回の操作（手順1～4まで）でタスクの契約に至ることが可能である。

手順1



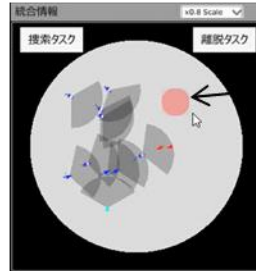
タスク告示を行うエージェントの範囲を選択する。

手順2



告示するタスク（検索タスクまたは離脱タスク）を選択する。

手順3



検索エリアまたは離脱先のエリアの位置及び範囲を指定する。

手順4

タスク告示メッセージ送信 [検索]

検索時間 [sec]

検索中心

検索範囲

契約エージェント数

入札期限 [sec]

タスクの実行及び契約に必要な情報を確認し、送信ボタンを押す。
（表示される左記のポップアップ画面にて操作する。）

手順5

AGENT	契約番号	契約状況	タスク概要
TeamA-A16	Player1-CONTRACT-1	Award	検索
TeamA-A14	Player1-CONTRACT-1	Award	検索
TeamA-A13	Player1-CONTRACT-1	Award	検索
TeamA-A17	Player1-CONTRACT-1	Award	検索
TeamA-A19	Player1-CONTRACT-1	Award	検索

表 3.3.2.3-1 に示す画面にて、契約結果を確認する。

図 3.3.2.1- 1 設計及びプロトタイプを行ったタスク告示から契約までの流れ

3. 3. 2. 2 AIエージェント群の将来予測結果の提示方法の解明

本項では、AIエージェント群の将来予測結果の提示方法の解明の一貫として実施した、対戦相手の移動傾向情報及びAIエージェントの将来行動情報を提示する際の内容及び提示方法に関する検討結果及びプロトタイプ結果を示す。

しっぽ鬼における行動の選択肢（相手を探す、相手を追いかける等）を3. 3. 1. 1項 ①～⑤にて示したが、これらの行動は、ゲームエリア上のある位置への移動や、相手エージェントへの移動が主な行動となる。そのため、AIエージェントの将来行動情報として、次にどこに向かおうとしているかが重要と考える。また、これらの行動情報は、対戦相手の位置や分布及び自分が協働しようとするAIエージェントとの相対的な位置関係を踏まえた上で表示しなければ、各エージェントが何を行おうとしているか正確に把握することは困難と考える。

また、対戦相手の移動傾向として移動履歴を表示することで、その相手のゲームに対する取組姿勢や戦略を読み取ることができ、自チームの戦略に活用できると考える。例えば、相手が、積極的に自チームの集団に向かって来る移動履歴を示していれば、この相手が積極的に捕獲に来るタイプであると読み取れ、そのような相手は積極的に対処する必要がある。また、相手が自チームの集団には近づかず逃げ回る移動履歴を示していれば、この相手が積極的に捕獲に参加しないタイプであると読み取れ、積極的に追いかけて、相手の出方を伺うことが有効と思われる。

以上の経緯から、HSIプロトタイプでは、AIエージェント群の将来予測結果として、自チームのAIエージェントの移動目標位置とそこまでの予想経路を、対戦相手の移動傾向情報として、対戦相手の移動履歴を統合表示画面に表示することとした。

HSIプロトタイプとして試作した将来行動情報例を図3. 3. 2. 2- 1に示す。

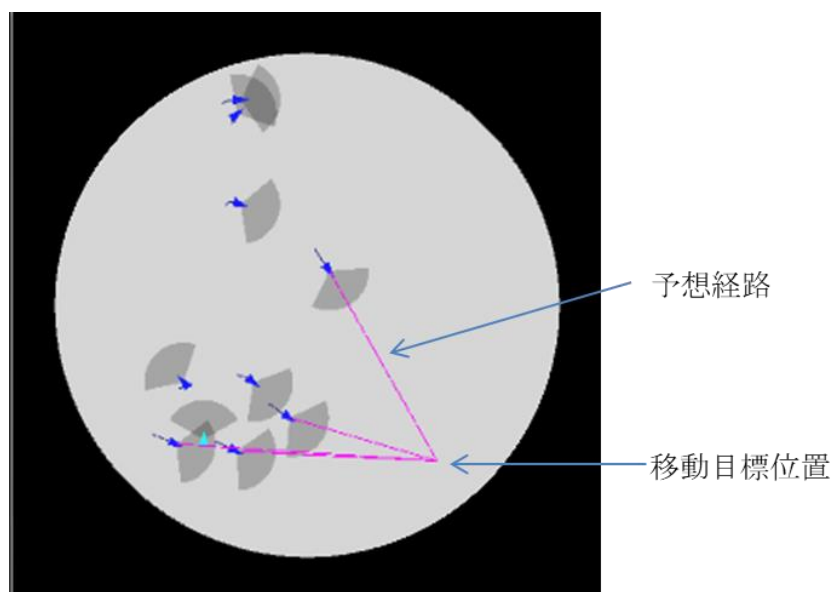


図 3. 3. 2. 2- 1 探索タスクにおける将来行動情報表示例

3. 3. 2. 3 タスクのマーケット状況等の効果的な人への提示方法の解明

本項では、エージェント間交渉における、タスクのマーケット状況等の効果的な人への提示方法の解明の一貫として実施した、エージェント間交渉における協調タスクの取引状況等を直観的に人に提示する方法に関する検討結果及びプロトタイプ結果を示す。

本情報の提示に際し、まず、協調タスクの取引状況を把握するための情報について整理を行った。協調タスクについては、図3.2.3.3-1に示す流れで契約を行うため、そのプロセスが正常に行われているか確認するためには、以下の状況を確認することが有効と考える。

- ・ 協調タスクが告示されたか。
- ・ 協調タスクに対し、各エージェントが入札をしているか。
- ・ 落札者の選定を実施したか。
- ・ 実施報告を受けたか。

また、落札者の選定から実施報告を受けるまでに、各エージェントがタスクを順調に実施しているか、次のタスクをいつ頃実施できる状態となるか把握することも、人とAIエージェントが協働する上では有用と考える。なお、協調タスクにおける進捗の把握に際しては、複数のエージェントが協力して行うものであるため、1人1人のエージェントの進捗のみならず、集団全体としての進捗も重要であることを考慮する必要がある。

以上より、取引状況等を提示するための情報として、以下の2種類の情報を提示することとした。

(1) 契約プロセスの進捗状況

(2) 契約後、各エージェントにおけるタスクの進捗状況

次に、情報の表示方法について、検討を実施した。

人が(1)の契約プロセスの進捗状況を確認する必要性が生じるのは、取引が進捗しない等の契約プロセスに何らかの問題が発生した場合が想定され、人は各エージェントの入札状況等、多数の情報を確認する必要がある。表示方法としては、リスト形式、図形式、グラフ形式など、様々な手法があるが、表3.3.2.3-1に示す通り、多数の取引状況の動静と詳細を一度に確認可能なリスト形式にて表示することとした。このリストは、各エージェントから入札結果や落札者選定の結果、タスクを実施するエージェントを表示するリストであり、通常は、契約や入札のプロセスが行われる都度、リストが増減することを確認することで、プロセスが何らかの形で進捗していることを把握でき、且つ、必要な場合にその細部の内容を確認できる。

また、(2)のタスクの進捗状況は、全タスクで同一の指標で進捗状況を定量的に把握できるよう、パーセンテージを単位とするグラフ形式で表示することとした。本表示に際しては、タスク全体としての進捗率の他、タスクに参加するエージェントごとの進捗も確認できるようにしている。本画面を表3.3.2.3-1に示す。

表 3.3.2.3- 1 契約プロセスの進捗状況を確認する画面



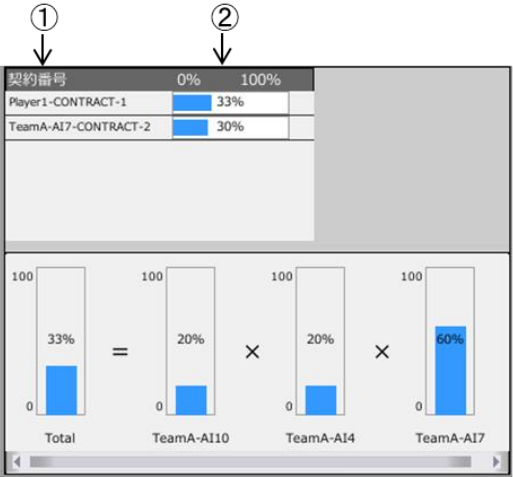
項目	内容																																									
画面表示	<div style="text-align: center;"> ① ② ③ ④  </div> <table border="1" style="margin: 10px auto; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #cccccc;">AGENT</th> <th style="background-color: #cccccc;">契約番号</th> <th style="background-color: #cccccc;">契約状況</th> <th style="background-color: #cccccc;">タスク概要</th> </tr> </thead> <tbody> <tr><td>TeamA-A18</td><td>TeamA-A17-CONTRACT-2</td><td>Bid</td><td>捕獲</td></tr> <tr><td>TeamA-A12</td><td>TeamA-A17-CONTRACT-2</td><td>Bid</td><td>捕獲</td></tr> <tr><td>TeamA-A14</td><td>TeamA-A17-CONTRACT-2</td><td>Bid</td><td>捕獲</td></tr> <tr><td>TeamA-A110</td><td>TeamA-A17-CONTRACT-2</td><td>Bid</td><td>捕獲</td></tr> <tr><td>TeamA-A16</td><td>TeamA-A17-CONTRACT-2</td><td>Bid</td><td>捕獲</td></tr> <tr><td>TeamA-A15</td><td>TeamA-A17-CONTRACT-2</td><td>Bid</td><td>捕獲</td></tr> <tr><td>TeamA-A19</td><td>TeamA-A17-CONTRACT-2</td><td>Bid</td><td>捕獲</td></tr> <tr><td>TeamA-A17</td><td>Player1-CONTRACT-1</td><td>Award</td><td>搜索</td></tr> <tr><td>TeamA-A13</td><td>TeamA-A17-CONTRACT-2</td><td>Bid</td><td>捕獲</td></tr> </tbody> </table>		AGENT	契約番号	契約状況	タスク概要	TeamA-A18	TeamA-A17-CONTRACT-2	Bid	捕獲	TeamA-A12	TeamA-A17-CONTRACT-2	Bid	捕獲	TeamA-A14	TeamA-A17-CONTRACT-2	Bid	捕獲	TeamA-A110	TeamA-A17-CONTRACT-2	Bid	捕獲	TeamA-A16	TeamA-A17-CONTRACT-2	Bid	捕獲	TeamA-A15	TeamA-A17-CONTRACT-2	Bid	捕獲	TeamA-A19	TeamA-A17-CONTRACT-2	Bid	捕獲	TeamA-A17	Player1-CONTRACT-1	Award	搜索	TeamA-A13	TeamA-A17-CONTRACT-2	Bid	捕獲
AGENT	契約番号	契約状況	タスク概要																																							
TeamA-A18	TeamA-A17-CONTRACT-2	Bid	捕獲																																							
TeamA-A12	TeamA-A17-CONTRACT-2	Bid	捕獲																																							
TeamA-A14	TeamA-A17-CONTRACT-2	Bid	捕獲																																							
TeamA-A110	TeamA-A17-CONTRACT-2	Bid	捕獲																																							
TeamA-A16	TeamA-A17-CONTRACT-2	Bid	捕獲																																							
TeamA-A15	TeamA-A17-CONTRACT-2	Bid	捕獲																																							
TeamA-A19	TeamA-A17-CONTRACT-2	Bid	捕獲																																							
TeamA-A17	Player1-CONTRACT-1	Award	搜索																																							
TeamA-A13	TeamA-A17-CONTRACT-2	Bid	捕獲																																							
仕様詳細	①AGENT	入札／落札したエージェント名称を示す。																																								
	②契約番号	契約番号を示す。 契約番号は、“タスク告示を行ったエージェント名称”+CONTRACT-“連番”																																								
	③契約状況	告示されたタスクに対する契約状況を示す。 ・Bid : 入札 ・Award : 落札																																								
	④タスク概要	告示対象の協調タスク（搜索、捕獲、離脱）を示す。																																								

表 3.3.2.3- 2 タスクの進捗状況を確認する画面

項目	内容									
画面表示	 <p>① ② ③</p> <p>タスクの進捗率</p> <table border="1"> <thead> <tr> <th>契約番号</th> <th>0%</th> <th>100%</th> </tr> </thead> <tbody> <tr> <td>Player1-CONTRACT-1</td> <td>45%</td> <td></td> </tr> <tr> <td>TeamA-AI7-CONTRACT-2</td> <td>71%</td> <td></td> </tr> </tbody> </table> <p>タスク告示総数： 25 件 入札待ち契約： 0 件 実行中契約： 0 件 実施完了契約： 10 件 内、成功契約： 4 件</p>	契約番号	0%	100%	Player1-CONTRACT-1	45%		TeamA-AI7-CONTRACT-2	71%	
契約番号	0%	100%								
Player1-CONTRACT-1	45%									
TeamA-AI7-CONTRACT-2	71%									
(2) 詳細表示	 <p>① ② ④</p> <table border="1"> <thead> <tr> <th>契約番号</th> <th>0%</th> <th>100%</th> </tr> </thead> <tbody> <tr> <td>Player1-CONTRACT-1</td> <td>33%</td> <td></td> </tr> <tr> <td>TeamA-AI7-CONTRACT-2</td> <td>30%</td> <td></td> </tr> </tbody> </table> <p>33% = 20% × 20% × 60%</p> <p>Total TeamA-AI10 TeamA-AI4 TeamA-AI7</p>	契約番号	0%	100%	Player1-CONTRACT-1	33%		TeamA-AI7-CONTRACT-2	30%	
契約番号	0%	100%								
Player1-CONTRACT-1	33%									
TeamA-AI7-CONTRACT-2	30%									
仕様詳細	<p>①契約番号 契約番号を示す。</p> <p>②進捗率 契約に参加する全エージェントのタスク進捗率総計を示す。</p> <p>③契約状況内訳 タスク告示総数、入札待ち契約数など、内訳数を示す。</p> <p>④進捗率詳細 契約に参加するエージェントごとのタスク進捗率を示す。</p>									

3. 3. 3 トレーニングの実施

構築したシミュレーション環境、AIエージェント及びHSIを使用して設計者自身がトレーニングを行うとともに、トレーニングを通じて必要と考える機能を設計に反映するイタレーションとして、以下の作業を実施した。

- ・ 研究従事者によるシミュレーション環境の変更及び動作確認
- ・ 上記確認結果に基づき、シミュレーション環境に反映すべき改修仕様の検討

本作業では、主にゲームの戦略性の確認及び人とAIの協調が行えるかという観点から動作確認及び変更を実施した。

例えば、ゲームの戦略性の観点では登場するエージェント数に応じてフィールドの大きさやエージェントの速度の調整を行った。シミュレーション環境でフィールドが小さすぎる場合やエージェントの移動速度が速すぎる場合、短時間での判断及び操作が求められるゲームになってしまい、人がAIと協調する余裕が無くなる他、戦略性が低下してしまう。一方、フィールドが大きすぎる場合やエージェントの移動速度が遅すぎる場合、ゲーム時間が長くなることで、フィールド評価参加者のストレスが過重になってしまう。フィールドの大きさやエージェントの速度の最適値を事前に決めることは困難であるため、動作確認を行う過程で調整を行った。

また、人とAIの協調が行えるかという観点ではHSIの画面配置について再検討を行った。その結果、HSIの画面配置では特定のAIエージェントを指定するための操作と指示内容のボタンを極力近く配置し、操作に時間を要しないように変更を行った。また、操作対象のエージェントの一人称視点とゲームの全体状況を把握するための統合情報画面の大きさの比率などは人によって好みや思考の違いがあると考え可変にする仕組みを実装した。

トレーニングの一環としてゲームの戦略性及び人とAIの協調が行えるかという観点からの確認を目的として、実験補助者を募り小グループ実験を実施した。小グループ実験の詳細は以下の通りである。

日時 : 2019年10月16日(水) 15:00~16:00

参加者 : 研究者4名、実験補助者6名

目的 : シミュレーション環境の改善点や研究の方向性について有意な意見を聴取する

結果 : 合計37項目のゲームの戦略性や人とAIの協調方法に対する改善意見を収集、主要コメントは以下の通り

- 取組みの趣旨やゲームの目的が分かりづらい。
- ゲームに戦略性をあまり感じられない。
- プルダウンからタスクを選んで入札、AI エージェントを画面でリストから選んで落札送信というのは時間がかかりすぎて使いづらい。

以上の小グループ実験の結果を踏まえ、フィールド評価では実験開始前に事前説明を行った上で、シミュレーション環境の慣熟を目的とした模擬対戦を実施した後に本実験を行う等、取組みの趣旨やゲームの目的をフィールド評価参加者に周知した。また、ゲームやパラメータを見直し戦略性を高め、3. 3. 2. 1項に示すように人とAIの協調では最小限の情報入力で完了するように改善した。

3. 3. 4 フィールド評価の実施

試作したAIエージェント及びHSIプロトタイプを評価するため、フィールド評価を実施した。フィールド評価参加者は、全て社内エンジニアのうち、3. 3. 1項、3. 3. 2項、3. 3. 3項のAIエージェント構築、HSIプロトタイプ製作及びトレーニングに関わっていない者とした。フィールド評価当日の流れの典型例を図3. 2. 3. 4- 1に示す。最初に試験目的や設定などの概要をチームA、チームBそれぞれの参加者に対して別々に説明を行うブリーフィングを実施し、試験評価と同条件ながら評価には用いない慣熟試験を実施したうえで、慣熟後にヒアリングまたはアンケート記入の時間を設けた。慣熟試験実施後、評価として行う対戦（フィールド評価）を実施した。フィールド評価では様々な設定で評価を行う一方、変動要素を最小化するために同一のAIエージェント及びHSIプロトタイプを用いた試験を同一日に、同一条件で少なくとも3回ずつ行った。評価を繰り返す行うにあたり、評価における気づき事項などを詳細かつ正確に収集するため、評価と評価の間に細かくアンケート記入の時間を設け、全ての評価が完了した後でチームA、チームBそれぞれ別々にヒアリングまたはアンケート回収を行った。またチームAに対してはHSIに対するレーティングを取得した。

■ フィールド評価当日の流れ（典型例）



※：研究進捗に伴い慣熟が増加したこと、新型コロナウイルス感染防止対策として分散実施あるいは周知及び個別集約とする等、柔軟な形で実施した。

図 3.3.4- 1 フィールド評価当日の流れ（典型例）

フィールド評価で評価する内容は次のとおりである。

【評価する内容】

- ・ 試作したアーキテクチャとHSIの基本性能
 - ・ 協働するエージェント数の増加に対するアーキテクチャとHSIのスケラビリティ
- フィールド評価の実施時期及び各評価において実施した内容を表3. 3. 4- 1 に示す。

表 3.3.4- 1 フィールド評価の実施時期及び主な実施内容

実施回	実施日	実施内容
1	2019/12/24	試作したアーキテクチャとHSIの基本性能確認のため、10vs10の対戦評価を実施（以降、Trial.1-1という）
2	2020/01/24	1回目のフィールド評価結果を受け、AIエージェントのアルゴリズムを改良した上で、再度1回目と同様、10vs10の対戦評価を実施（以降、Trial.1-2という）
3	2020/03/13	<ul style="list-style-type: none"> ・2回目のフィールド評価結果を受け、AIエージェントのアルゴリズムを改良した上で、再度1回目同様、10vs10の対戦評価を実施（以降、Trial.1-3という） ・また、対戦エージェント数を増やした20vs20の対戦評価を行い、アーキテクチャとHSIのエージェント数の変化によるスケーラビリティを確認（以降、Trial.2という）

以降では、フィールド評価の実施方法、評価結果及び評価結果に対する考察及びスパイラル#2に向けた取組について記載する。

3. 3. 4. 1 評価方法

本項では、フィールド評価の評価方法について記載する。

フィールド評価では、チームAとチームBの対戦シミュレーションを行い、3. 2. 4項の検討結果に基づき、勝率、生存率、MCH-UVD によるレーティング結果について比較を行った。

対戦においては、まず初めに、試作したアーキテクチャとHSIの基本性能を確認するために、両チームともエージェント数を10とし、対戦を実施した。本対戦は、計3回に渡って実施し、対戦ごとに、AIエージェントのアルゴリズムの改良を行った。

また、エージェント数の増加に伴い、アーキテクチャとHSIへの影響を確認するため、エージェント数を20に増やした対戦として、Trial. 2を設定した。

各Trialにおける、対戦人数及び使用したAIエージェントのアルゴリズムの条件を表3. 3. 4. 1- に示す。

表 3. 3. 4. 1- 1 対戦人数及び使用した AI エージェントのアルゴリズムの条件

Trial		条件	
		対戦人数	AI エージェントのアルゴリズムのレベル
1	1-1	チーム A: 人エージェント : 1 体 AI エージェント : 9 体 チーム B: 人エージェント : 10 体	1vs1 の状況で、ほとんどの場合、人に負ける
	1-2	Trial. 1-1 と同様	1vs1 の状況で、人に負ける場合あり
	1-3	Trial. 1-1 と同様	1vs1 の状況では、人に負けない
2		チーム A: 人エージェント : 1 体 AI エージェント : 19 体 チーム B: 人エージェント : 20 体	Trial. 1-3 と同様

評価に際しては、評価結果が、特定の個人の技量に依存することを避けるため、チームAには2人の人に参加してもらい、各対戦にはそのうちの1名がチームAの人エージェントを操作して3試合(1回の試行で2名×3試合、計6試合を行う)ずつ対戦し、2人分の評価結果を取得した。

各試行の前には、両チーム共に、ウォームアップとして、AIエージェントのみのチームとの3回の練習対戦を実施した。

3. 3. 4. 2 評価結果

まず、アーキテクチャとHSIの基本性能の評価結果として、Trial. 1における結果を示す。

アーキテクチャとHSIの基本性能の評価結果として、Trial. 1におけるチームAの勝率を表3. 3. 4. 1-、生存率を図3. 3. 4. 2- 1に示す。

AIエージェントのアルゴリズムの改良により、Trial. 1-3において、チームAは最終的に1試合で勝利しており、また、チームBの生存者数も、Trial. 1-1~1-3にかけて、徐々に減少していることが分かる。

しかしながら、Trial. 1-3にて使用したAIエージェントは、1vs1では人に負けない強さを有していたにも関わらず、ほとんどの試合でチームBに敗北しており、勝利の鍵がエージェント同士のチームワークであることを物語っている。

また、各試行後に取得した、チームA参加者2名によるMCH-UVD のレーティング結果を図3. 3. 4. 2- 2に示す。なお、本レーティング結果は、2名の評価結果のうち、レーティングが低かった方を記載している。また、この時使用したレーティングの判断基準を表3. 3. 4. 2- に示す。本基準は、B. Donmezらの手法“MCH-UVD”による。図3. 3. 4. 2- 2より、AIエージェントの改良によりレーティング結果が向上しているが、Trial. 1-3のレーティング結果取得時に、チームAの2名から以下の改善要望が提示されており、更なるレーティング結果の向上に向け、スパイラル#2では、これらの改善を行うことを検討した。

- ・自身のエージェントを操作しながら、味方のAIエージェントへタスク指示を出す余裕が少ないため、操作方法、指示の出し方等を改善する必要がある。
- ・味方のAIエージェントとの協働をより効果的に行うため、AIエージェントの行動や状態をより把握しやすくする必要がある。
- ・AIエージェントのアルゴリズムを改良し、相手チームを追いかける役と相手の針路を妨害する役などの役割を分担できるようにするなど、チームワークを向上させる必要がある。

表 3. 3. 4. 2- 1 試験 1 の結果

Trial	チーム A の結果		
	勝利数	敗北数	勝率 [%]
1-1	0	6	0
1-2	0	6	0
1-3	1	5	19

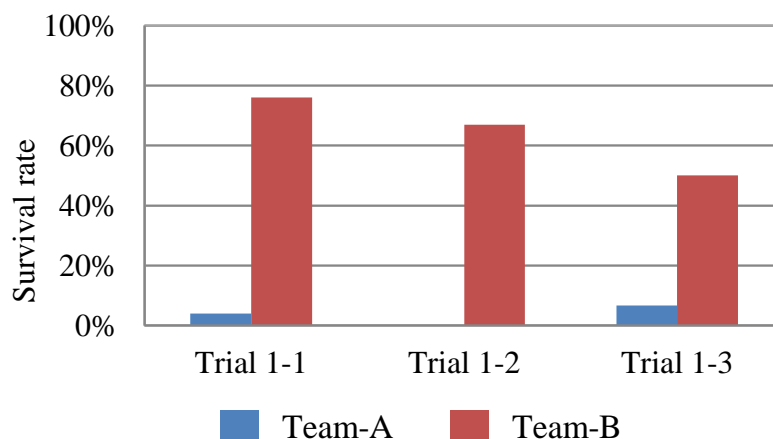


図 3. 3. 4. 2- 1 Trial. 1 における各チームの生存率

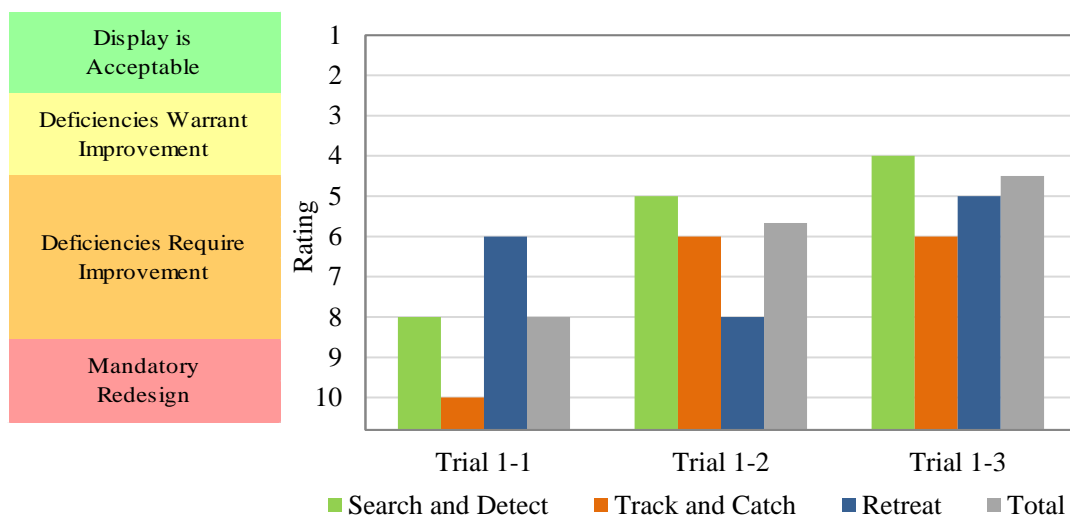


図 3.3.4.2- 2 HSI に対するタスクごとのレーティング結果

表 3.3.4.2- 2 使用したレーティングの判断基準

Rating	評価		評価の考え方
1	Display is Acceptable	非常に良好	表示系の変更は不要である
2		良好	パフォーマンスに影響を与えない小さな問題がある
3	Deficiencies Warrant Improvement	柔軟性に欠ける	ある程度の意思決定を支援するが、複数の意思決定オプションが生成されなかったりカスタマイズできない
4		判断支援不足	複数の意思決定をするための必要情報が提供されていない
5	Deficiencies Require Improvement	非効率なアクション	意思決定の実行に、複数ステップ必要となる。または、適切なタイミングで実行するの十分な余裕がない
6		情報が未集約	タスクに必要な情報が集約されていない。重要な情報の関係を判別、検索および導出する必要がある
7		タスク切替支援が不十分	再開されたタスクのステータスや、注意がタスクから離れている間に発生した変更は強調表示されない
8		注意指示が貧弱	複雑な画面構成により情報の取得が妨げられる。注意が必要場面では、適切なタイミングでオペレータの注意を引くことができない
9	Mandatory Redesign	行動が困難	タスクを実行するのが困難で誤用されやすい
10		情報取得が困難	ディスプレイに重要な情報が欠けている、重要な情報が見つからない、または検索時間が情報を無意味にする。

また、エージェント数の増加に伴う、アーキテクチャとHSIへの影響に対する評価結果として、表3.3.4.2- 1 に勝率を、図3.3.4.2- 3 に生存率、図3.3.4.2- 4 にHSIのレーティング結果を示す。

表3.3.4.2- 3 より、エージェント数の増加に伴う勝率への影響の変化は見られなかった。一方、図3.3.4.2- 3 より、チームAの生存率は増加したが、チームBは減少する傾向となった。また、図3.3.4.2- 4 より、エージェント数の増加に伴うレーティング結果への影響はなかった。

表 3.3.4.2- 3 エージェント数の増加に伴う生存率への影響

Trial	チーム A の結果		
	勝利数	敗北数	勝率 [%]
1-3	1	5	19
2	1	5	19

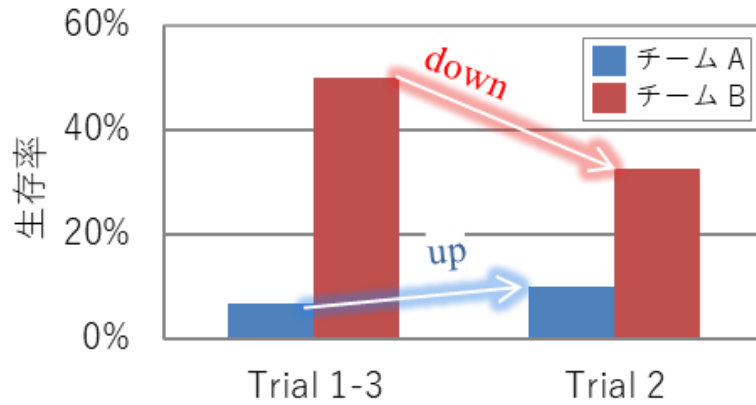


図 3.3.4.2- 3 エージェント数の増加に伴う生存率への影響

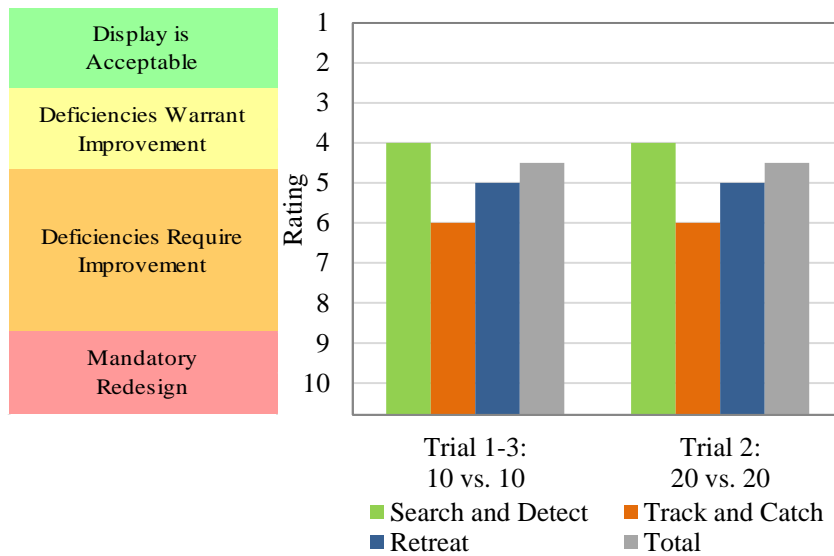


図 3.3.4.2- 4 HSI に対するタスクごとのレーティング結果

3. 3. 4. 3 評価結果に対する考察及びスパイラル#2に向けた取組について

本項では、3. 3. 4. 2 項の評価結果について考察すると共に、今後の課題、取組について示す。

まず、アーキテクチャとHSIの基本性能について考察を行う。

Trial.1の3回の対戦における傾向（表3.3.4.2-1、図3.3.4.2-1 参照）から、個々のAIエージェントの1vs1の状況における強さが、HSIのレーティング結果を向上させており、このことはAIエージェントの個々の能力が、人とのAIエージェント間が協調する上で重要な特徴であることを示唆している。即ち、AIエージェントが人と効果的に協働するためには、AIエージェントが最低限有すべき能力の基準というものがあると考えられる。具体的な能力として、ゲームエリアから逸脱しないことや、人の支援がなくとも、1vs1の対戦にて相手に負けることがないように注意して行動する等、個として十分な生存性を持っていることが必須だと考える。

人とAIエージェントのチームワークの資とするため、チームBがどのようにチームワークを発揮していたか分析した結果、チームBは2人以上で以下の役割を配分し、効率的に相手を捕まえていたことが分かった。

- ・ 足止役： 相手と正面を向いている人や相手の進行方向にいる人が相手にぶつかり、相手を止める。
- ・ 捕獲役： 相手の背後にいる人が、停止している相手を捕まえる。

また、各自がその場の状況に応じて、自らが果たすべき役割を即座に理解し、積極的に自分の役割を“宣言する”ことで、仲間と意思の疎通を行っていることが確認された。この意思疎通方法により、最小限のコミュニケーション量で仲間と役割を共有し、効果的なチームワークが発揮できていたと考える。

一方、チームAでは、以下の2つの理由から、協力して相手を捕獲することが困難であった。

- ・ AIエージェントは協調タスクを配分することはできたが、捕獲タスクにおいてAIエージェント間での役割配分がなく、AIエージェント同士でチームBのようなチームワークを発揮できなかった。
- ・ 人が直感的に各AIエージェントの状態やタスク実行状況を把握することが困難であったため、人がAIエージェントと協力して行動することが容易ではなかった。

チームBが実施した協調及びコミュニケーション手法や、特に積極的に自分の役割を“宣言する”方法は、1. 項（2）②にて述べた「AIエージェントから人への情報提示方法」として、非常に参考になるものと考えられる。スパイラル#2では、チームAにおいても、チームBと同様の協調及びコミュニケーションを実現することで、どの程度効果的に人とAIエージェントが協働できるようになるか検討した。

HSIに関しては、チームAの2人は両名とも、“DisplayMode 2”を使用した。これは、3. 3. 2. 3 項で検討したタスクのマーケット状況等の情報の表示が直感的でなく、動的な切迫環境では確認の必要性が低かったことと、確認している余裕がなかったとのコメントが得られた。また、タスクの進捗状況は、統合情報表示上に表示されるAIエージェントの将来行動情報を用いてある程度把握できることから、本表示上にタスクの進捗状況も合わせて表示した方が、より直感的に人がAIエージェントの状態等を確認できると考える。

次に、アーキテクチャとHSIのスクレーパビリティについて考察を行った。表3.3.4.2-3、図3.3.4.2-3及び図3.3.4.2-4の結果から、AIエージェント数の増加に関して以下の傾向が得られた。

- ・ 勝率は、AIエージェント数が増加しても、変化しない。
- ・ レーティング結果は、AIエージェント数が増加しても、変化しない。

- ・ チームAの生存率は増加したが、チームBは低下した。

チームBの生存率の低下の原因として、以下の2つが考えられる。

- ① チームAのAIエージェント数の増加により、チームBの人1人当たりの応答能力が相対的に低下した。
- ② Trial. 2に新規で参加した11人は、Trial. 1に参加していなかったため、エージェントの操作技術に十分に習熟していなかった。

①及び②の仮説について考察を行うため、Trial. 2における、チームBの各プレイヤーの成績を総捕獲数の多い順に並べた結果を図3.3.4.3-1に示す。図の左縦軸は総捕獲数、右縦軸は平均生存率、横軸は各プレイヤーの名前（ラベル）で、赤枠のものはTrial. 2からの新たな参加者である。ここで、捕獲数は各人の攻勢積極性を表し、生存率は防御性を表している。たとえば、メンバA、B、Cは、生存率は低いが、捕獲数は多く、高い攻撃積極性を示したプレイヤー集団と言える。一方、メンバM、P、Qは、捕獲数、生存率共に低く、攻撃に消極的且つ防御性も低いプレイヤー集団と言える。また、メンバK、L、S、Tは、捕獲数は少ないが、生存率が高く、防御性重視のプレイヤー集団と言える。図3.3.4.3-1より、Trial. 2に新たに参加した11人は、他の9人と同様、高い攻撃積極性を示したプレイヤーや攻撃に消極的且つ防御性も低いプレイヤー、防御性重視のプレイヤーの集団に万遍なく含まれており、Trial. 1に参加していた9人と同等の攻撃積極性と防御性の傾向を有していたと言える。つまり、②の仮説は当てはまらず、構築したアーキテクチャ及びHSIは、協働するエージェント数が増えるほど、人のみの協働よりも相対的に優れた成績を示す可能性があると考えられる。ただし、スパイラル#1では、10vs10と20vs20の2ケースのみの実験に留まるため、スパイラル#2以降、より多人数による評価において、エージェント数の変化による感度を確認していく必要があると考える。

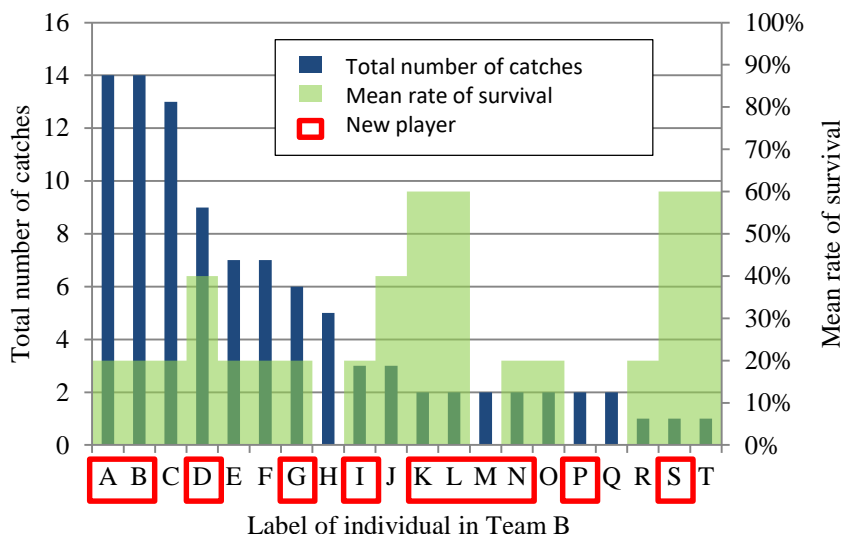


図 3.3.4.3- 1 チーム B の各プレイヤーの成績

スパイラル#1のフィールド評価では表3.3.4.2-1に示す通り、10vs10 (Trial. 1)及び20vs20 (Trial. 2)の何れにおいても人とAIが協働するチームAの勝率は19%にとどまった。しかし図3.3.4.2-3の通り、全て人で構成されるチームBの生存率はTrial. 1-3で50%、Trial. 2で33%と人数が増えるにつれて減少傾向である。つまり人間の数で言えばたった1人のチームAを相手に、Trial. 1-3ではチームBを構成する10人のうち半数の5人、Trial. 2では20人のうち6割以上となる約13.5人が犠牲になる計算となる。この結果は、人間の人数から言えばチームAは圧倒的劣勢な状況であるが、複数のAIエージェントと協働することにより多数の人間に対抗できる可能性を示唆している。

3. 4 スパイラル#2

3. 4. 1 AIエージェントの構築

3. 3. 4. 3項に記載した通り、スパイラル#1の結果から導出されたAIエージェントの改善として、スパイラル#2では役割配分機能を実装した。

役割配分機能については、スパイラル#1における人間の協調行動を参考とし、協調タスク内の捕獲タスク(以下、連携捕獲タスクと呼び、個別タスク内の捕獲タスクと区別する)に着目した。AIエージェントの協調タスクに関する行動モデルを図3. 4. 1- 1に示し、本項で対象とするBTノードを赤枠で示している。しっぽ鬼の例題設定においては全てのエージェントの身体的能力は同一のため、対戦相手を捕まえるためには少なくとも2体のエージェントが必要となり、それぞれが「足止役」と「捕獲役」の役割を果たす必要があることを3. 2. 1. 4項で記載した。この2つの役割に加え、足止を実行したエージェントがフリーズし無防備となるため、これを保護するための「サポート役」の役割を追加して3つの役割に分解するものとし、3. 4. 1. 1項に連携捕獲タスクに関する優先順位の変更点を、3. 4. 1. 2項に連携捕獲タスクにおける役割配分機能の仕様について説明する。

以上に加え、3. 4. 1. 3項では人間と強化学習エージェントの協調に関する目途付けを紹介する。人間とAIエージェントが協調するためには、AIの動作ロジックに説明可能性があり、人間のAIに対する信頼性が適切なレベルに保たれている必要がある。近年盛んに研究が行われている機械学習の研究では目覚ましい性能が報告されている一方、その複雑なアルゴリズムから説明可能性が低いとされている。本研究では、スパイラル#1で定義したタスク仕様に基づいて設計した行動モデルを骨格として、連携捕獲タスクにおける役割分担毎に期待される行動を行う強化学習AIモデルを獲得し、その強化学習AIモデルと人間が適切に協調できるかを確認した。スパイラル#2では人間との協調の可能性について目途付けをし、スパイラル#3にてフィールド評価との比較実験として検証することとした。

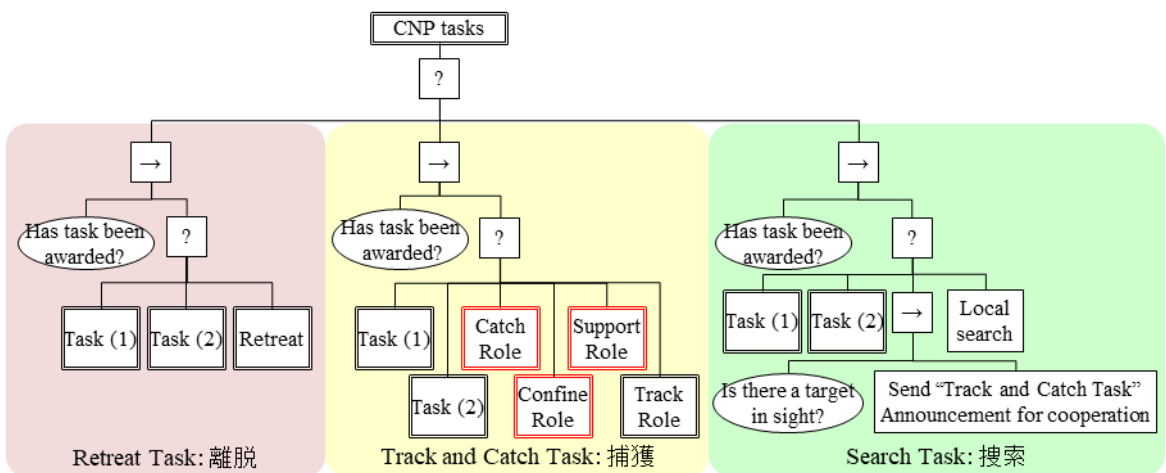


図 3. 4. 1- 1 AI エージェントの協調タスクに関する行動モデル
(赤枠箇所はスパイラル#1 からの差分を示す)


3. 4. 1. 1 エージェントにおける適切なタスクの条件及び価値基準の設定

本項では、連携捕獲タスクに関してスパイラル#1からの変更点について記載する。

スパイラル#1では、検索タスクおよび離脱タスクについてしっぽ鬼における大局的なタスク指示と位置付けて人間からエージェント群へ指示をするものと定義し、一方、連携捕獲タスクに関してはAIエージェント間のみで契約される局所的なタスクとしていた。

スパイラル#2では連携捕獲タスクについて、人が積極的にAIエージェント群と協力して行えるよう、人からAIエージェント群に対して連携捕獲タスクを指示（告示）できるよう変更した。人からAIエージェント群への連携捕獲タスクの指示はAIエージェント間のみでの連携捕獲タスクよりも優先度が高いものと定義した。表3. 4. 1. 1- 1 にタスクの優先順位を示す。人が連携捕獲タスクを指示した場合、人エージェントは連携捕獲を実施する単位である3エージェントの一員となっており、2体のAIエージェントと連携捕獲を実施することになる。2体のAIエージェントは、人が操作するエージェントが3. 5. 1. 2項にて後述する連携捕獲タスクのある役割を、3体のAIエージェントが分担するのと同じ判定フローに基づき実行中であるとみなして他の役割を実行するように実装した。つまり、チームAの人は連携捕獲タスクに参加するAIエージェントが提示するタスク/役割シンボル(3. 3. 2. 1項に記載する)を通じて、自身が連携捕獲タスクに参加しているエージェントの1体として期待されている役割を果たす必要がある。

表 3. 4. 1. 1- 1 タスクの優先順位(スパイラル#1 からの変更分は赤字)

優先順位	タスク
 優先度高	1 エリア逸脱回避
	2 衝突回避
	3 離脱（協調タスク）
	4 連携捕獲（協調タスク） （人間からの指示）
	5 連携捕獲（協調タスク） （AIからの指示）
	6 検索（協調タスク）
	7 捕獲（個別タスク）
	8 検索（個別タスク）
優先度低	

3. 4. 1. 2 人とAIが効果的に分担可能なタスクの仕様の設定

本項では連携捕獲タスクの役割配分機能について記載する。

始めに、連携捕獲タスクを実行する状況において、移動する捕獲対象に到達するまでの時間とその位置を見積もることは重要である。例えば、あるエージェントが捕獲対象を足止めできたとしても、捕獲対象がフリーズしている間に味方のエージェントが捕獲対象に到達できない場合は足止が有効でないことは自明である。

捕獲対象に到達するまでの時間は、捕獲対象の移動方向と速度が近い将来には変化しないと仮定して計算した。連携捕獲タスクを実行するエージェントの位置を \mathbf{P}_B とし、エージェントの最大速度を V_B とする。捕獲対象の位置を \mathbf{P}_R とし、現在の速度ベクトルを \mathbf{V}_R とすると、捕獲対象への到達時間 T_a は式(1)によって求められる。

$$T_a = \frac{-\mathbf{R} \cdot \mathbf{V}_R - \sqrt{(\mathbf{R} \cdot \mathbf{V}_R)^2 - |\mathbf{R}|^2(|\mathbf{V}_R|^2 - V_B^2)}}{|\mathbf{V}_R|^2 - V_B^2} \quad (1)$$

ここで、 \mathbf{R} は捕獲を実行するエージェントから捕獲対象への相対位置ベクトル $\mathbf{R} = \mathbf{P}_R - \mathbf{P}_B$ である。

到達時間 T_a の実数解が存在する場合、到達位置 \mathbf{P}_r は式(2)で表される。

$$\mathbf{P}_r = \mathbf{P}_R + T_a \mathbf{V}_R \quad (2)$$

ここで、到達位置 \mathbf{P}_r がゲームエリア外である場合には、到達時刻 T_a は、解がないものとして扱われる。

以上で求められた到達時間を用いて、図3.4.1.2-1に示す通り各役割を実行する条件を決定した。

まず、捕獲対象が連携タスクに参加しているAIエージェント(以下、連携エージェントと呼ぶ)に対して反方位に逃げ続ける場合、連携エージェントは到達時刻の解をもたないため、追跡役を例外として割り当てる。追跡役のエージェントは最大速度で捕獲対象の現在位置に移動するというふるまいをする。

次に、捕獲対象の状態が評価され、フリーズ中である場合は連携エージェントは即時に捕獲を試みる。連携エージェントの中で、到達時間が最小であり、かつ到達時間が最大許容時間 T_{max} より小さい場合に、捕獲役を実行する。一方、捕獲対象がフリーズしていない場合、連携エージェントは捕獲対象の足止めを試みる。足止役を実行するエージェントが捕獲目標をフリーズ状態にすることに成功した後に、そのフリーズ状態が継続している間に味方のエージェントが捕獲することを想定する。つまり、連携エージェントの中で到達時間が最小であり、それが最大許容時間 T_{max} より小さい場合には、足止役を実行する。なお、最大許容時間 T_{max} にはエージェントのフリーズ状態の継続時間よりも短い時間を変数として設定した。このとき、他のエージェントの到着時間が最大許容時間 T_{max} より大きい場合には、捕獲役が割り当てられる。これは、十分近くに連携捕獲可能な味方エージェントがないので、単独で目標をとらえようと試みるふるまいである。最後に、捕獲役又は足止役を持たないエージェントはサポート役が割り当てられる。

各役割の仕様を以下に記載する。

- ・ 捕獲役：捕獲対象を捕獲する役割。この役割を持つエージェントは捕獲対象の背後に向かって移動した後、捕獲対象の背面に接触するように移動する。
- ・ 足止役：捕獲対象を足止する役割、この役割を持つエージェントは捕獲対象の姿勢に関係なく、現在の捕獲対象の位置に移動する。
- ・ サポート役：ほかの連携エージェントをサポートする役割。足止役の周囲に捕獲役以外のエージェントが1体いれば、足止役に近づきやすくなるため、本研究ではこの役割を持つエージェントは、足止役と同じく捕獲対象の姿勢に関係なく現在の捕獲対象の位置に移動するふるまいとした。

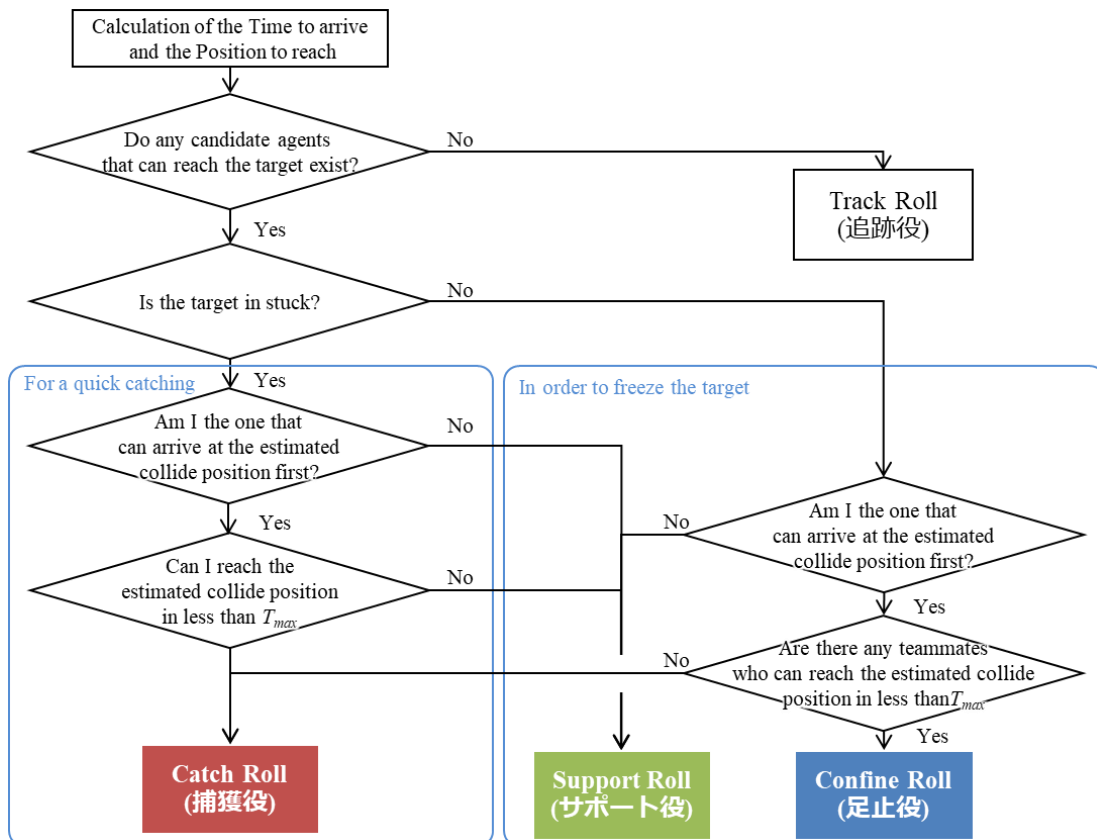


図 3. 4. 1. 2- 1 連携捕獲タスクにおける各役割の実行条件

3. 4. 1. 3 人間と強化学習エージェントの協調に関する目途付け

強化学習によって獲得されたAIモデルと人間が協調できるかを評価するために、スパイラル#2では限定された条件において強化学習を実施し、連携行動を獲得することができるか目途付けを行った。

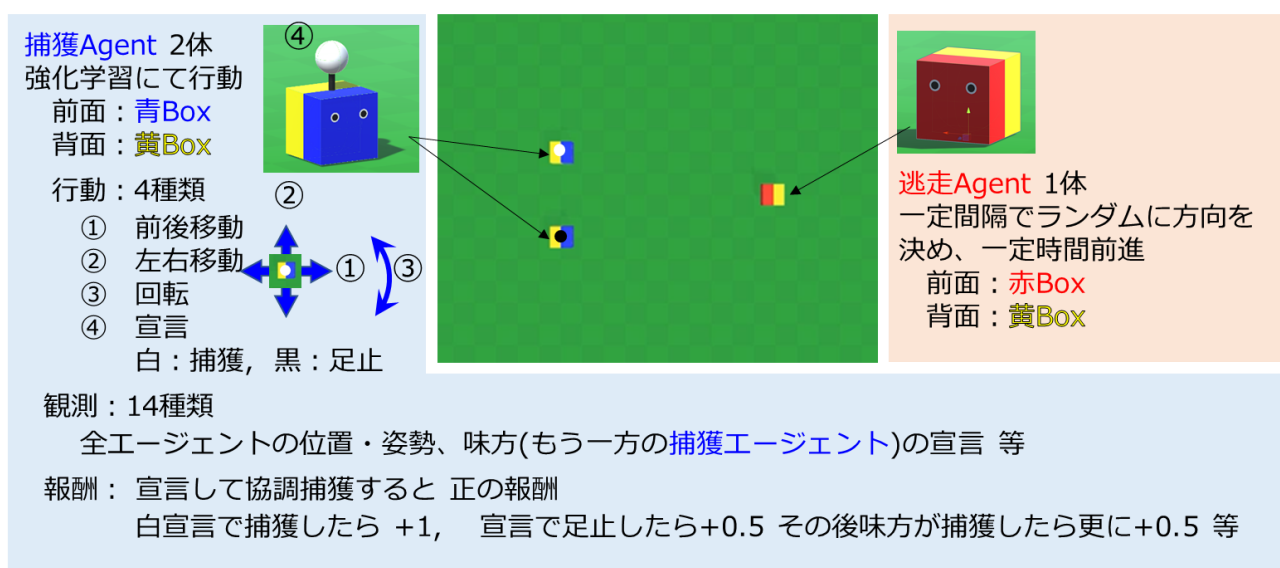
強化学習とは、意思決定主体であるエージェントがある環境の中で試行錯誤を繰り返すことにより、望ましい行動を獲得させる機械学習の一手法である。“エージェントと環境は離散的な時間ステップ t で連続して相互作用を行う。ある時刻 t において、エージェントは環境から状態 s_t を受け取り、それに基づいて行動 a_t を選択する。これにより、エージェントは環境から報酬 r_{t+1} を受け取り、環境の状態が s_{t+1} に変化したことを知る。エージェントには環境の状態に応じて何らかの行動を返す関数が実装されており、この関数を方策 (policy) と呼ぶ。一連の行動の結果として受け取る報酬を最大化するような方策を見つけることが強化学習の主な目的となる。” [16]

強化学習の目途付けとして、本研究の“しっぽ鬼”の開発環境に採用したUnityで学習環境を作成し、Unityのツールキットの1つであるML-Agents[17]を利用してエージェントの強化学習を実施した。学習環境及びエージェントの設定を図3. 4. 1. 3- 1に示す。

学習する対象は、しっぽ鬼の問題設定を最小化したものとして、捕獲エージェント2体が連携して逃走エージェント1体を捕まえる状況を抽出した。学習環境及びルールは3. 2. 1. 4項に記述したしっぽ鬼のものに準拠し、捕獲エージェントの前面が逃走エージェントの背面に接触したら学習エピソードが終了し、捕獲エージェント同士を含む他の接触は一定距離移動した後、一定時間フリーズすることとした。

エージェントの強化学習アルゴリズムはPPO(Proximal Policy Optimization) [18]を採用した。PPOは2017年にopenAIから発表された強化学習手法であり、TRPO(Trust Region Policy Optimization) [19]を改善するために提案された。TRPOでは、方策の更新時にKLダイバージェンスの大きさが一定値以下となるようにすることで方策が大きく変化することを制限するのに対し、PPOでは更新前の方策の出力確率と更新後の方策の出力確率の比が一定の範囲内になるように制限するというシンプルなアイデアを用いた。これにより、TRPOが抱えていた、実装が複雑、Actor-Critic モデルで方策関数と価値関数のパラメータを共有するのが難しいなどの課題を解決している。この実装しやすさと優れたパフォーマンスからML-agentsのデフォルトアルゴリズムの1つとなっている。

図 3. 4. 1. 3- 1 学習環境及びエージェントの設定



捕獲エージェントの行動は、前後移動、左右移動、姿勢回転及び宣言の4種類の離散値出力とした。ここで宣言とは -1 及び $+1$ の離散値が定義されており、学習環境上では -1 が白色、 $+1$ が黒色で表示し、後述する報酬によって -1 では捕獲する動機が、 $+1$ では足止めをする動機が発生するように設定されている。宣言は人間との協調を考えたときに「AIエージェントが何をしようとしているか宣言しないと、人間は協調できないのではないか」という仮説を立て追加したものである。

捕獲エージェントへの観測は、以下の14種類を定義した。ここで捕獲エージェントの連携を企図して、味方の宣言は観測として常に入力されているものとした。加えて、逃走エージェントに関する観測については、捕獲エージェント2体のどちらの視野にも逃走エージェントを含めていない場合は不観測とし、しっぽ鬼の設定に則ったものとしている。

- ・ 自身(捕獲エージェント)の位置・姿勢及びフリーズしている場合のフリーズ残り時間 (4種類)
- ・ 味方の捕獲エージェントの位置・姿勢及びフリーズしている場合のフリーズ残り時間 (4種類)
- ・ 味方の捕獲エージェントの宣言 (1種類)
- ・ 逃走エージェントの位置・姿勢及びフリーズしている場合のフリーズ残り時間 (4種類、捕獲エージェントが共に逃走エージェントを視野に入れていない場合は逃走エージェントを最後に観測した時点の値を代入)
- ・ 逃走エージェントを最後に観測した時点からの経過時間 (1種類)

捕獲エージェントへの報酬は宣言によって捕獲と足止の動機が発生するようにした。具体的には、宣言が -1 (白、捕獲)のときに捕獲すると $+1$ の報酬を付与する。また、宣言が $+1$ (黒、足止)のときには逃走エージェントをフリーズさせると $+0.5$ の報酬を付与し、かつその後味方の捕獲エージェントがフリーズ中の逃走エージェントを捕獲したら $+0.5$ の報酬を追加した。その他、極力早く捕獲するためにステップが経過するごとに -0.001 の微小な負の報酬を付与するといった、学習モデルが最適な行動を出力するための報酬関数を複数定義している。

捕獲対象となる逃走エージェントは一定間隔でランダムに進行方向を決め、進行方向に対して姿勢回転及び前進するロジックを実装し、捕獲エージェントの学習中に再生することとした。

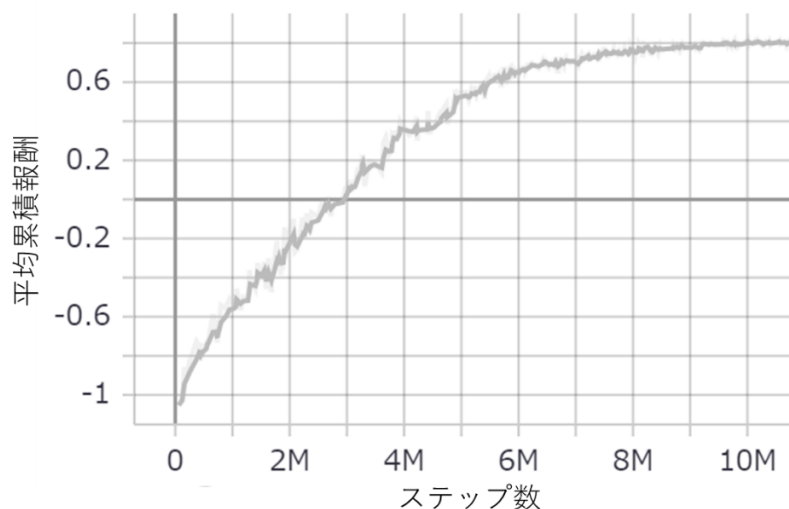


図3.4.1.3- 2 学習目途付けにおける報酬推移

以上の条件で学習した結果を図3.4.1.3-2に示す。平均累積報酬は1エピソード中の2エージェントがそれぞれ獲得した累積報酬の平均値となっており、連携が成功している場合は、 $1 - 0.001 \times (\text{1エピソードの経過ステップ})$ が平均累積報酬となる。ステップが経過する毎に1に漸近していくことから、2体の捕獲エージェントが足止及び捕獲の連携を素早く行えていることがわかる。本学習結果の獲得モデルを再生した結果、逃走エージェントに対して捕獲エージェントが挟み込むように別れ、逃走エージェントに対して正対している捕獲エージェントは宣言 +1（黒、足止）、逃走エージェントの背後に位置している捕獲エージェントは宣言 -1（白、捕獲）で連携して捕獲している様子が観察された。

本研究の主題である人間とAIの協調の目途付けをすべく、捕獲エージェントの1体を人間操作に変えて確認したところ、捕獲エージェント同士の衝突が多少発生してしまうものの、すべての確認ケースにおいて逃走エージェントを捕獲できた。よって人間と強化学習AIは協調できる目途があると判断した。

一方、逃走エージェントを人間が操作し、捕獲エージェント2体は獲得したモデルを再生したケースを確認したところ、人間が逃走エージェントを停止させる、後進させるといった学習時と異なる操作をした場合、捕獲エージェントの連携行動が機能せず、結果として捕獲できないケースがあることが判明した。これは、学習中の逃走エージェントが姿勢回転及び前進するのみといった確定的な動きしかしないため、学習していない状況化ではうまく対処できないことが考えられる。この課題への対処として、捕獲目標の逃走エージェントの行動に多様性を持たせる必要があると判断し、スパイラル#3ではセルフプレイ強化学習に取り組むこととした。セルフプレイ強化学習とは、対戦型の学習環境において、対戦する双方のチームが交互に学習を行う強化学習手法のひとつである。セルフプレイ強化学習を取り入れることで、逃走エージェントの学習が進むごとに様々な行動を試行錯誤し捕獲対象としての多様性が自動的に生まれるものと考えた。スパイラル#3における強化学習によるエージェントモデルの獲得については、3.5.1.3項に記述する。

3. 4. 2 HSIプロトタイプの実成

スパイラル#2のHSIプロトタイプでは、1. (3) ③-1. 人からAIエージェント群全体に対する効果的な指示方法の解明として、人からAIエージェントへ捕獲タスクを指示する操作を追加した。この操作の追加に関して、3. 5. 2. 1項に示す。また、1.

(3) ③-2. AIエージェント群の将来予測結果の提示方法の解明としてAIエージェントの実行中タスク/役割をシンボルとして表明し、加えて捕獲対象をターゲットラインによって視認させやすくする改修を行った。この改修に関して、3. 4. 2. 2項に示す。

HSIプロトタイプの実成の実施事項のうち、1. (3) ③-3. タスクのマーケット状況等の効果的な人への提示方法の解明については、3. 3. 2. 3項に記載した、スパイラル#1のHSIプロトタイプと同様であるため本項における記載は省略する。

3. 4. 2. 1 人からAIエージェント群全体に対する効果的な指示方法の解明

人とAIエージェントが協力して捕獲を行う状況において、連携捕獲タスクを契約する場合に以下の2つの方法が想定される。

- ① 人からAIエージェントにタスク告示をし、AIエージェントが応答する方法
- ② AIエージェントから人にタスク告示が行われ、人間が応答する方法

①に関しては、AIエージェントからの応答に対して、自動的にタスク契約の締結が結ばれる仕組みを設ければ連携可能である。このとき、タスク契約を締結するための落札メッセージは捕獲対象に対して近いAIエージェント2体に対して自動的に送信するものとした。②に関しては、人はAIよりも応答に時間を要すること及びAIエージェントが多数であることから、しっぽ鬼の状況においてメリットが少ない。よって連携捕獲タスク機能は人からAIエージェント及びAIエージェント間のみで行えるものとし、AIエージェントから人へは指示を行わせないこととした。

このとき、人間からAIエージェントに連携捕獲タスクの指示を出す手順は、人間が捕獲を要すると判断したときには時間的切迫度が高いことから、捕獲対象を選択した後にゲームコントローラのボタンを押すと指示が送信される簡便な操作とした。このとき、捕獲対象は人間が操作するエージェントに最も近いエージェントが自動的に選択され、要すれば近い順に捕獲対象を変更できるようにした。捕獲対象選択カーソルを図3.4.2.1-1に示す。

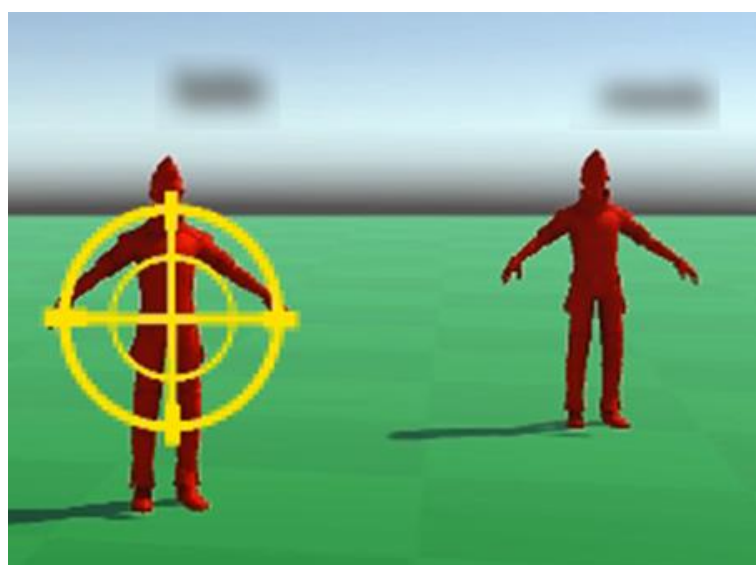


図 3.4.2.1- 1 捕獲対象選択カーソル
(黄色円形のマーク、左のエージェントが捕獲対象選択されている)

3. 4. 2. 2 AIエージェント群の将来予測結果の提示方法の解明

3. 4. 2. 1 項に示した通り、連携捕獲タスクは人からAIエージェントに指示可能とした。ここで、連携捕獲タスクが人とAIエージェントで契約締結された以降の役割配分に関して、以下の2つの方法を検討した。

- ① AIエージェントが宣言した役割・行動に基づき、人が今の状況からすべき役割を決める方法
- ② 人が宣言した役割・行動に基づき、AIエージェントが今の状況からすべき役割を決める方法

以上の2つの方法について以下3点の考察を行った。

- ・ 人は捕獲目標と会合できる時刻を正確に計算できないため、“どの役割を担えば良いか判断できない”。
- ・ 人が宣言を行う場合、宣言する操作を行う時間的余裕がない上、誤操作により、正確に宣言できない可能性がある。
- ・ ただ、AIエージェントが宣言する役割を見ることで、人に期待されている役割を判断することは可能と考える。

以上の考察から①の方法を採用し、AIエージェントが人へ役割を宣言するのみの役割配分を実施した。AIエージェントの強みは、配置状況から各エージェントがどの役割を担うべきかを瞬時に判断できることであり、AIエージェントが捕獲対象と役割を宣言することで、人は自分が担うべき役割が分かり、行動しやすくなる効果が期待できる。

人がAIエージェントと協働するために、AIエージェントのタスクまたは役割を宣言させる方法としての“タスク/役割シンボル”とAIエージェントがどのエージェントを捕獲対象にしているかを示す“ターゲットライン”を追加した。図3. 4. 2- 1 にチームAの一人称視点を示し、表3. 4. 2- 1 にタスク/役割シンボルの種類、表3. 4. 2- 2 にターゲットラインの種類について示す。

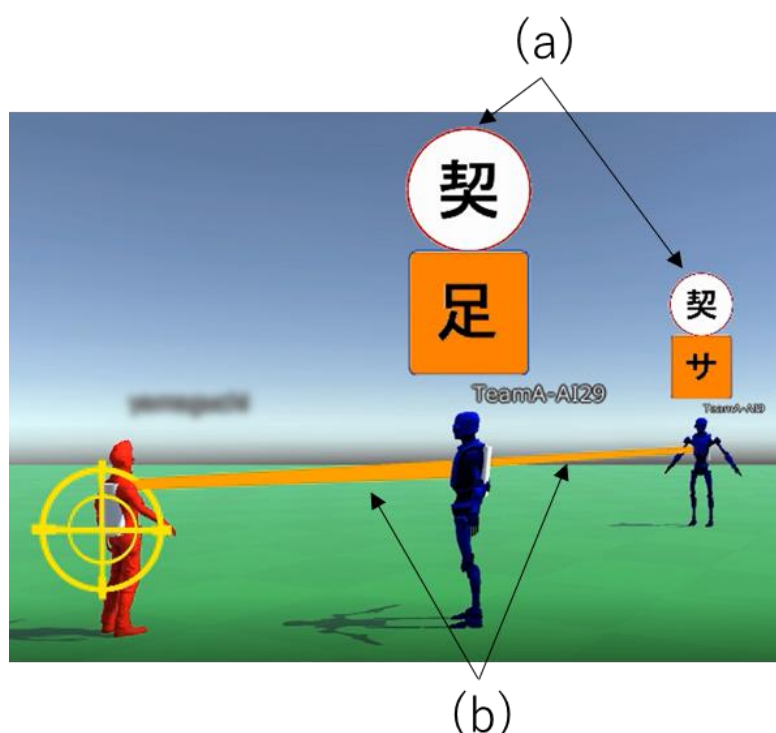





図 3. 4. 2- 1 チーム A の一人称視点
(a) タスク/役割シンボル, (b) ターゲットライン

表 3.4.2.2- 1 タスク/役割シンボルの種類

協調タスク実行	離脱 (協調タスク)	捕獲-捕獲役 (協調タスク)
 ※協調タスクの場合にのみ上部に表示		 ※AI間は白色、人間からの指示の場合は橙色
捕獲-足止役 (協調タスク)	捕獲-サポート役 (協調タスク)	捕獲-追跡役 (協調タスク)
 ※AI間は白色、人間からの指示の場合は橙色	 ※AI間は白色、人間からの指示の場合は橙色	 ※AI間は白色、人間からの指示の場合は橙色
搜索 (協調タスク)	エリア逸脱回避 (個別タスク)	衝突回避 (個別タスク)
		
捕獲-回り込み (個別タスク)	捕獲-追跡 (個別タスク)	搜索 (個別タスク)
		

表 3.4.2.2- 2 ターゲットラインの種類

ターゲットライン種別	色	説明
	水色	個別タスクの捕獲対象を示すターゲットライン
	マゼンダ	AI間の協調タスクの捕獲対象を示すターゲットライン
	橙色	人間からの指示された協調タスクの捕獲対象を示すターゲットライン

3. 4. 3 トレーニングの実施

3. 4. 1 項及び3. 4. 2 項の検討結果に対する有効性の確認と検討結果へのフィードバックのため、5vs5 程度の小グループ実験を3回実施した。各小グループ実験の詳細を表3. 4. 3- に示す。

表 3. 4. 3- 1 小グループ実験の詳細

日時	参加人数 (内、研究従事者数)	トレーニング結果から改修した代表的な項目
2020年 8月21日	6名 (3名)	連携エージェントの捕獲役が、捕獲対象のフリーズによる移動距離も考慮して背後に向かうよう改修
2020年 8月27日	6名 (3名)	捕獲対象1体に対して、複数の連携捕獲タスクが重複しないよう改修
2020年 9月10日	7名 (3名)	移動ロジックの違いにより、AIエージェントの速度が人エージェントに比べ、若干が速いことが判明、移動速度が全く一緒になるように修正

3. 4. 4 フィールド評価の実施

3. 4. 4. 1 評価方法

本項では、スパイラル#2のフィールド評価の評価方法について記載する。

2020年9月17日及び2020年10月23日にフィールド評価を2回実施した。スパイラル#1に引き続き、チームAとチームBの対戦シミュレーションを行い、3. 2. 4項の検討結果に基づき、勝率、生存率、MCH-UVDによるレーティング結果について取得した。

また、スパイラル#2以降はさらなるエージェント数の増加の影響を確認するため、10vs10、20vs20の対戦に加え、30vs30の対戦を追加した。試験ケースが増えたことによる、フィールド評価参加者の待機時間を極小化するために10vs10及び20vs20の対戦については平行して実施することとした。フィールド評価1回分の実施スケジュールを表3. 4. 4. 1- 1に示す。

評価に際しては、評価結果が、特定の個人の技量に依存することを避けるため、チームAには2人の人に参加してもらい、各対戦にはそのうちの1名がチームAの人エージェントを操作して3試合(1回の試行で2名×3試合、計6試合を行う)ずつ対戦し、2人分の評価結果を取得した。

各試行の前には、両チーム共にウォームアップとして、AIエージェントのみのチームとの3回の慣熟対戦を実施した。

表 3. 4. 4. 1- 1 フィールド評価1回分の実施スケジュール

時間目安	実施内容 1	実施内容 2
13:15-13:30(15分)	研究従事者が各自席に回り参加者点呼	-
13:30-13:50(20分)	慣熟対戦(両チーム)(3対戦)	-
13:50-14:05(15分)	慣熟後アンケート記入(各自)+休憩	-
14:05-14:35(30分)	チームA参加者① vs チームB 20人 (3対戦)	チームA参加者② vs チームB 10人 (3対戦)
14:35-14:50(15分)	対戦後アンケート記入(各自)+休憩	-
14:50-15:20(30分)	チームA参加者① vs チームB 10人 (3対戦)	チームA参加者② vs チームB 20人 (3対戦)
15:20-15:35(15分)	対戦後アンケート記入(各自)+休憩	-
15:35-16:05(30分)	チームA参加者① vs チームB 30人 (3対戦)	-
16:05-16:35(30分)	チームA参加者② vs チームB 30人 (3対戦)	-
16:35-16:50(15分)	MCH-UVDレーティング評価(チームA)	対戦後アンケート記入(チームB全員)

3. 4. 4. 2 評価結果

スパイラル#2のフィールド評価の結果として、チームAの勝率を表3. 4. 4. 2- に、各チームの生存率を図3. 4. 4. 2- に示す。

表 3. 4. 4. 2- 1 スパイラル#2におけるチームAの勝率の結果

種別	チームAの結果		
	勝利数	敗北数	勝率 [%]
10vs10	7	5	58
20vs20	4	8	33
30vs30	1	11	8

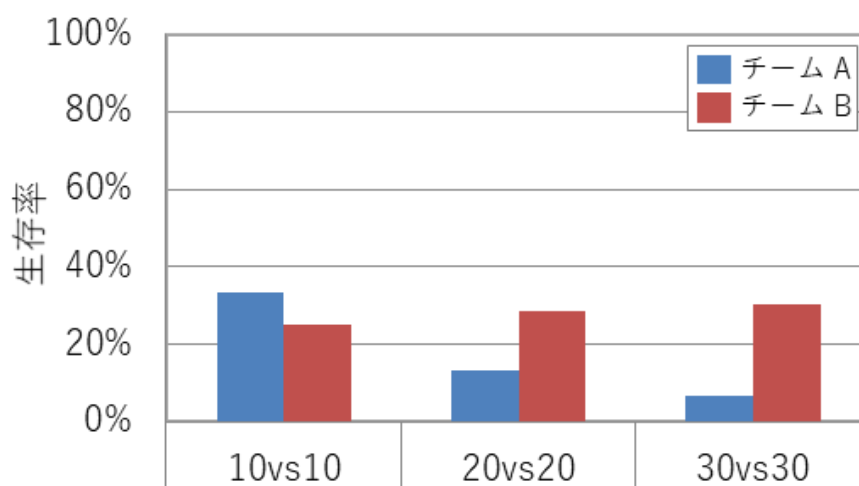


図 3.4.4.2- 1 スパイラル#2 における各チームの生存率

10vs10についてはチームAの生存率がチームBの生存率上回り、連携捕獲及びHSI改善の効果が窺える。しかし20vs20のチームAの生存率はチームBよりも低く、初めて実施した30vs30についてはチームAの生存率が非常に低い（6.4%）結果となった。

また、各試行後に取得した、チームA参加者によるMCH-UVD のレーティング結果を図 3.4.4.2- に示す。なお、レーティング結果は各回2名による評価の平均を記載した。

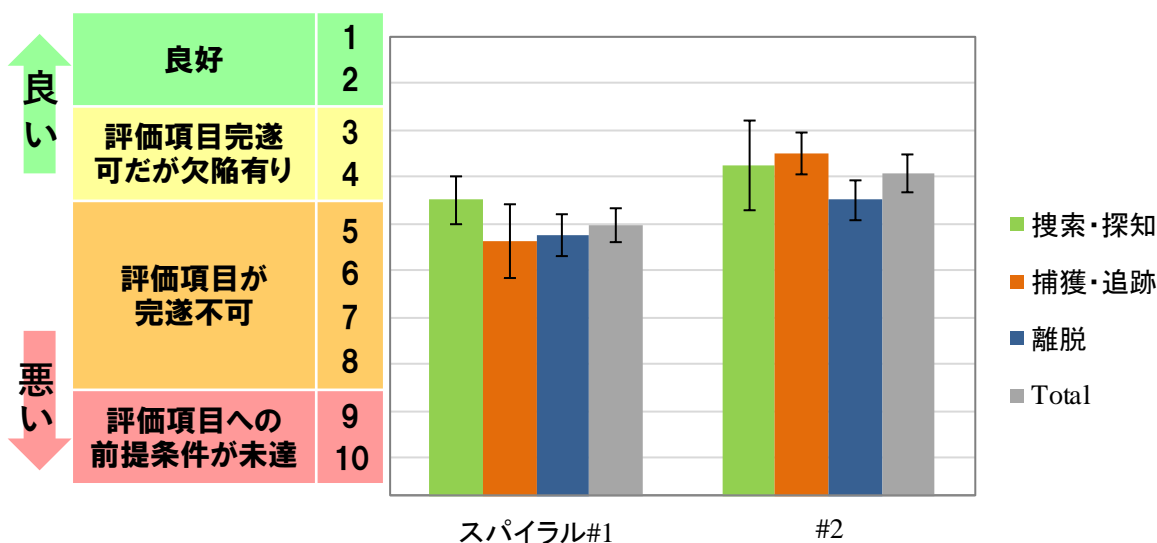


図 3.4.4.2- 2 HSI に対するタスクごとのレーティング結果

レーティングを平均的に見た場合、改善傾向はあるが十分とはいえない結果となった。

3.4.4.3 評価結果に対する考察及びスパイラル#3に向けた取組について

本項では、3.4.4.2項の評価結果について考察すると共に、今後の課題、取組について示す。3.4.4.2項に示したように、チームAの生存率は期待したような高い値にはならなかった。原因を考察するため、典型例について個別の対戦状況の分析を行った。図3.4.4.3- 1に、ある10vs10の対戦時のエージェント数の推移例を示す。この対戦は対戦終了時点でチームAの生存数8となりチームAが大差で勝利したケースである。この評価においてチームA参加者は、「開始後すぐに味方全員を1カ所に集め、全員で同じ行動をとり、常に数的有利の場面を作るようにした。」とコメントしている。グラフを見てもチームAは常に数的に優位である。180秒付近に両者のエージェント数が近づく

が、2エージェント以上の差を保っており、それ以降は差が広がっていき、チームBのエージェントを全て捕獲していることが分かる。

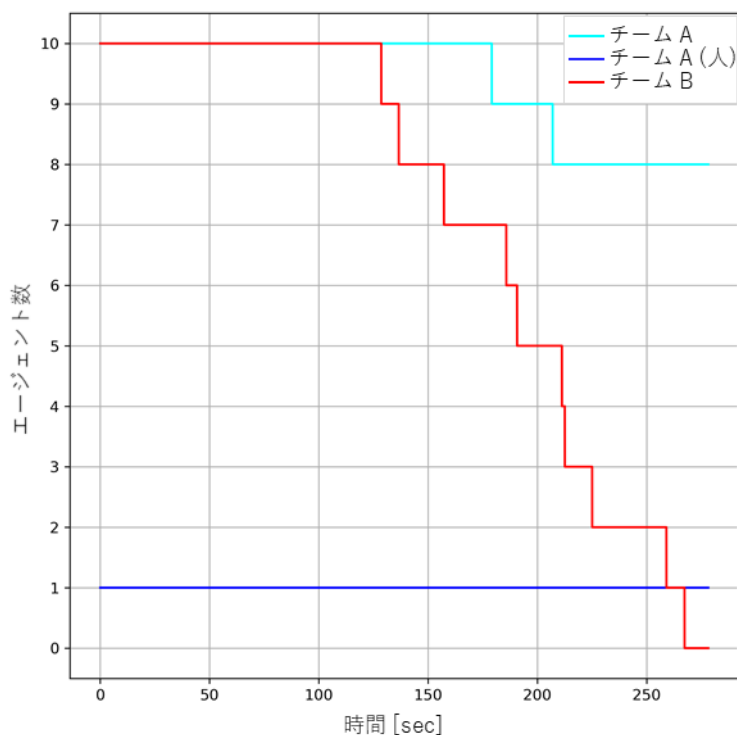


図 3. 4. 4. 3- 1 10vs10 対戦時のエージェント数の推移例
(8vs0 でチーム A 勝利)

次に図3. 4. 4. 3- 2 に同様に、ある30vs30の対戦時のエージェント数の推移例を示す。この対戦はチームAが負けたケースである。チームA参加者は、「10vs10 (図3. 4. 4. 3- 1 のケース) と同じようになるべく固まって行動させるが、全体の数が多いため、2班に分けた。」とコメントしている。図3. 4. 4. 3- 2によれば、300秒付近まではチームAが数的に優位を保っているにも関わらず、その後逆転されていることがわかる。チームAの減少カーブはほぼ一定である一方、チームBは230秒付近を境に減少に歯止めがかかっている。チームA参加者のコメントにあるように、AI群を2班に分けたことが局所的に数的劣勢な状況を作り出した可能性が高い。

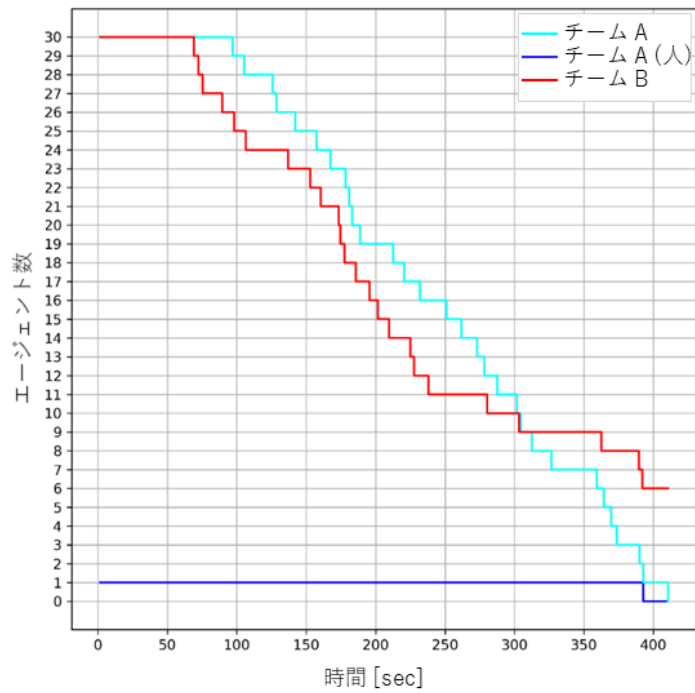


図3.4.4.3- 2 30vs30対戦時のエージェント数の推移例
(0vs6でチームB勝利)

続いて図3.4.4.3- 3 に、別の30vs30の対戦時のエージェント数の推移例を示す。この対戦は対戦終了時点でチームAの生存数23となり、スパイラル#2のフィールド評価における12回の30vs30対戦において、唯一チームAが勝利ケースである。チームA参加者は、「序盤はAIが最も群れているグループに人も加わり、終盤はなるべくAIのグループ同士が合流するように敵を誘導しようとして行動した。」とコメントしている。図3.4.4.3- 3 によれば、チームAエージェント数は120秒付近以降ほとんど減少していない。全体的なエージェント数に差が大きくなると形勢逆転は困難となることが分かる。参加者のコメントからも、局所的な数的優勢を保つことが重要であることが分かる。

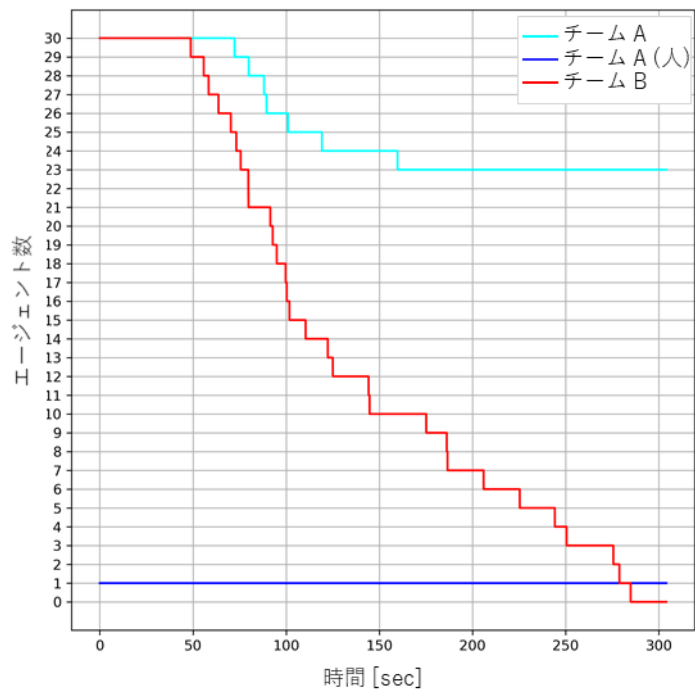


図3.4.4.3- 3 30vs30対戦時のエージェント数の推移例
(23vs0でチームA勝利)

さらに細かくAIエージェントの行動を観察すると、図3.4.4.3-4に示すような「1つのチームBエージェントを多数のチームAエージェントが数珠つなぎで延々と追跡する状況」が多数回生起していることが分かった。このような状況は、捕獲対象とされたチームBの操作者が、自身が捕獲対象となったことをチームAエージェント群の挙動から察知し、周囲に十分な数の味方がいない時に生起する。この後チームBの巧みな分散逃避によってチームAのAIエージェントが誘引され、分離し、チームBが局所的に数的優勢なエリアに誘導される状況があった。さらに、こうした状況に対応するために用意していた離脱タスクは、人が指示した位置に半強制的に移動させることが可能であったが、エージェントの背後警戒が疎かになり意図した効果を発揮できなかった。

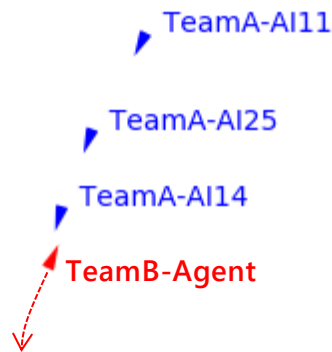


図 3.4.4.3-4 チーム B エージェント（赤）逃避時の追跡状況

捕獲タスクは、3体のエージェントがサブチームを構成して捕獲対象となるエージェントの捕獲を試みるが、捕獲対象が逃避する場合、3体のエージェントが捕獲対象との現在位置に向かって移動するという比較的単純な追跡方法であった。捕獲対象が逃避を継続した場合、捕獲を試みる3体のエージェントは延々と追跡を行い、状況が収束しないばかりか、チームBが局所的に数的優勢なエリアに誘導される一因となった。

そこで、スパイラル#3において以下の対策を行うこととした。

- ・ 離脱タスクを見直し、人が指示した位置に自身の背後警戒を行いながら移動して集合し、陣形を組む集合タスクに置き換える
- ・ 逃避する捕獲対象を味方同士で囲い込んで逃避先領域を限定するように追跡方法を改善する

3. 5 スパイラル#3

3. 5. 1 AIエージェントの構築

3. 5. 1. 1 人とAIが効果的に分担可能なタスクの仕様の設定

本項では、スパイラル#3にて実施した連携捕獲タスクの追跡役ロジック及び集合タスクの仕様について示す。

スパイラル#2のフィールド評価コメントを受け、スパイラル#2以前の捕獲対象の現在位置を目標に移動する追跡役のロジックでは、AIエージェントが1列に並ぶような状況になる可能性が高く有効ではないことがわかる。逃げ続ける捕獲対象への対策として、3体の追跡役は捕獲対象の逃走可能な領域を制限することが有効である。Huangらは、複数の追跡者が1体の逃走者を追い詰める問題に対して、“*area-minimization strategy*” [20]を提案している。“*area-minimization strategy*”を適用したシナリオ結果を図3.5.1.1- 1に引用する。“*area-minimization strategy*”は、追跡者と逃走者との間の共有Voronoi境界の中間点に向かって移動することにより、追跡者は有限時間で回避者を捕獲できることが証明されている。

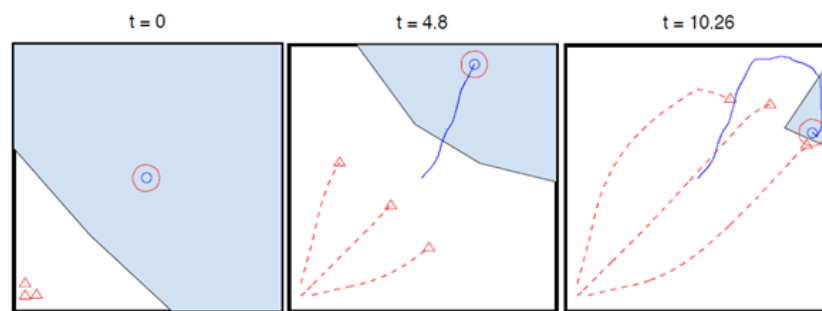


図 3.5.1.1- 1 *area-minimization strategy*を利用して3体の追跡者(三角、点線)が逃走者(丸、実線)を追い詰めているシナリオ例 [20]

次に、ゲーム中盤以降において局所的に数的優位を確立するために、スパイラル#3から追加した集合タスクの仕様について示す。集合タスクには“集合前”と“集合後”の2つのフェーズがあり、“集合前”フェーズでは集合役のみが実行され、人から指示された集合位置から一定半径内の整列円内に到達したら集合後にフェーズが移行する。

“集合後”フェーズには整列役、停止役及び反撃役がある。集合タスク(集合後)のイメージと各エリアについて図3.5.1.1- 1に示す。また、各役割の詳細は表3.5.1.1- 1に示す。

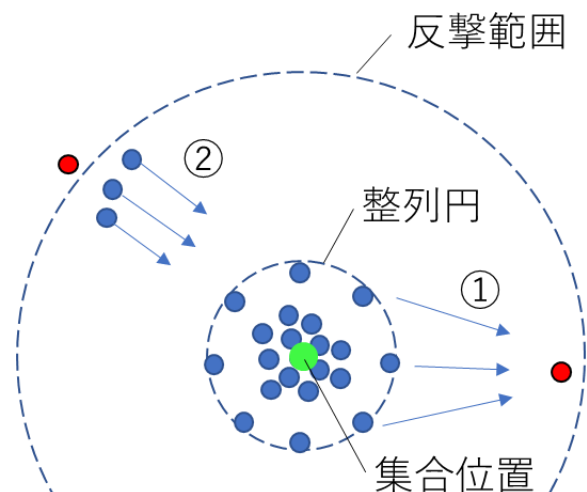


図 3.5.1.1- 2 集合タスク(集合後)のイメージと各エリア
青丸はチームAエージェント、赤丸はチームBエージェントを示す
① は反撃開始、②は反撃中止のイメージを示す

表 3.5.1.1- 1 集合タスクの各役割

役割	説明
集合役	集合前フェーズにおいて、人から指示された集合位置に移動する。移動する際には、一定距離内に相手チームがいる場合には相手チームエージェントの中で一番近いエージェントに正対し、一定距離内に相手チームがいない場合には右回りに回転して周囲を警戒する。移動方向の一定範囲内にフリーズしている相手チームエージェントがいる場合には個別タスクの捕獲を実行する。整列円内に入った場合には集合役をやめ、集合後フェーズに移行する
整列役	整列円上に8個の整列スロットが存在し、集合後フェーズにおいて整列スロットに空きがあり、かつ反撃役を実施していない場合は空いている整列スロットの位置で円の外側に向かって整列する。整列後はその位置で待機する。
停止役	集合後フェーズにおいて整列スロットに空きがなく、かつ反撃役を実施していない場合は整列円の内部で待機する。
反撃役	集合後フェーズにおいて、反撃範囲内に相手チームが侵入した場合には3体のAIエージェントが個別タスクの捕獲を実行することで対処する。捕獲対象の相手チームエージェントを捕獲した場合及び捕獲対象が反撃範囲の外に出た場合には実施をやめる。

ここでスパイラル#2以前の離脱タスクの違いについては、集合前には近づいてくる相手チームのエージェントに対して正対し、足止された場合にも後続のエージェントが捕獲をするロジックになっている。集合後には反撃範囲直径内に近づく相手チームエージェントに対して反撃をするために、ただ指示位置に向かうだけの離脱タスクに比べて、相手チームに脅威性を与えながら集合するタスクとなっている。

3. 5. 1. 2 エージェントにおける適切なタスクの条件及び価値基準の設定


本項では、集合タスク及びAI間の連携捕獲の実行条件に関してスパイラル#2からの変更点について記載する。

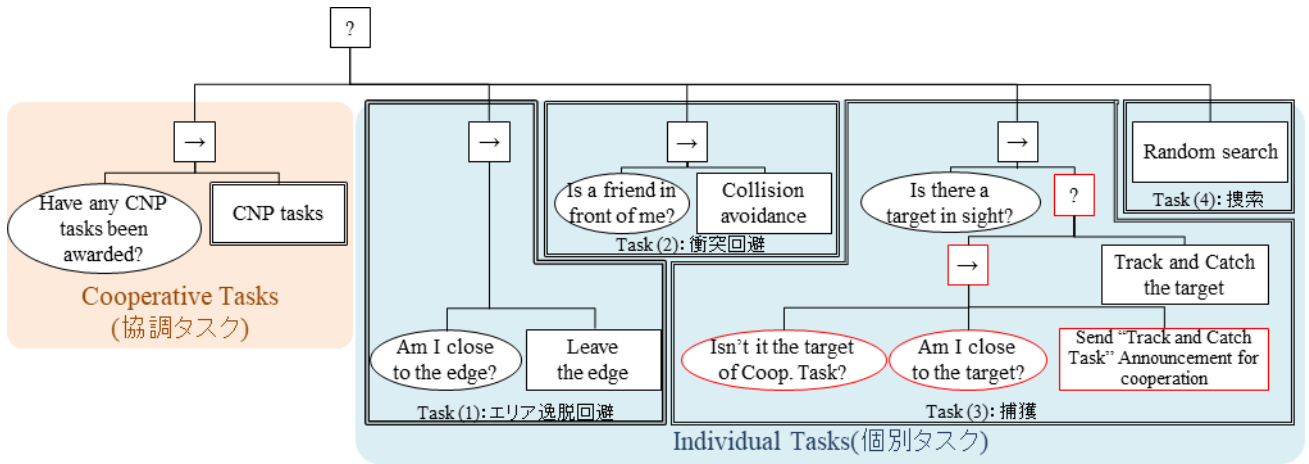
スパイラル#1では、離脱タスク及び搜索タスクはしっぽ鬼における大局的なタスク指示と位置付け、人間からエージェント群へ指示をするものと定義していた。離脱タスクは協調タスクの中では最も高優先度のタスクであり、離脱タスクが契約された場合には、エリア逸脱回避及び味方同士の衝突回避をしつつ、指示された位置に向かって直線的に前進するタスクとしていた。一方、スパイラル#1及び#2のフィールド評価を受け、離脱タスクは直線的に前進するだけのタスクということもあり対戦相手の人間から意図が読まれ離脱できずにAIエージェントが捕獲されてしまう可能性が高いことが判明した。そこで、離脱タスクを集合タスクとして見直すこととした。このとき、集合タスクは人間からエージェント群へ指示をする大局的なタスクという定義は引き継ぐものとした。表3.5.1.2-1にタスクの優先順位を示す。

加えて、AI間の連携捕獲についてAIが連携捕獲タスク行う条件を見直した。スパイラル#2までは、図3.3.1.2-1に示したように搜索タスクにおいて捕獲対象を発見したAIエージェントがAI間の連携捕獲タスク指示を行う条件としていた。この条件であると、AIエージェント間の連携捕獲タスクの発生回数が少なくなると考え、スパイラル#3では以下のふるまいに変更した。図3.5.1.2-1にAIエージェントの行動モデルを示す。

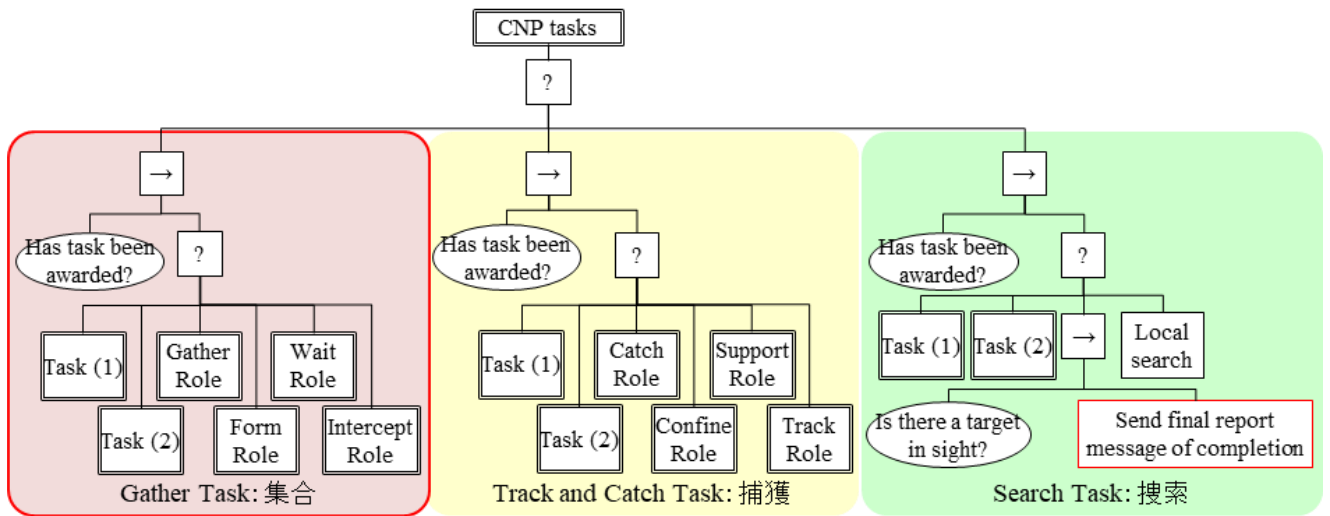
- ・ 搜索タスクで対戦相手を視野に入れたら、搜索タスク参加者全員のタスクを終了（タスク完遂の実施報告）させる
- ・ 個別タスクの捕獲タスク内において、以下の2つの条件を満たしたら連携捕獲タスクを指示する
 - 捕獲対象が連携捕獲タスクに既に割当たっていないか
 - 自エージェントが捕獲対象に最も近いか

表 3.5.1.2-1 タスクの優先順位(スパイラル#2 からの変更分は赤字)

優先順位		タスク
	優先度高	1 エリア逸脱回避
		2 衝突回避
		3 集合（協調タスク）
		4 連携捕獲（協調タスク） （人間からの指示）
		5 連携捕獲（協調タスク） （AIからの指示）
		6 搜索（協調タスク）
		7 捕獲（個別タスク）
	優先度低	8 搜索（個別タスク）



(i) 行動モデル全体



(ii) 協調タスクに関する行動モデル

図 3.5.1.2- 1 AI エージェントの行動モデル
(赤枠箇所はスパイラル#2 からの差分を示す)

3. 5. 1. 3 強化学習によるエージェントモデルの獲得

本項では3. 4. 1. 3項の考察に基づき、セルフプレイ強化学習を実施し、強化学習によるエージェントモデルを獲得した取り組みを示す。

始めに、セルフプレイ強化学習を実施するために、逃走エージェントを強化学習可能な設定とした。逃走エージェントの強化学習アルゴリズムはPP0を使用し、観測や行動の定義は3. 4. 1. 3項のスパイラル#2における捕獲エージェントと同等とした。逃走エージェントと捕獲エージェントが一定距離離れた場合、一定ステップ経過した場合及び捕獲エージェントに逃走エージェントが捕獲された場合にエピソードが終了する。逃走エージェントの報酬は一定距離離れた場合に逃げ切ったものとみなし+1の報酬を付与し、捕獲された場合には-1の報酬を付与した。また、極力、長く逃走するためにステップが経過するごとに+0.001の微小な正の報酬を付与した。

捕獲エージェントは、連携捕獲タスクの基本単位が3エージェントであることから、3体登場させることとし、アルゴリズム、行動及び報酬は3. 4. 1. 3項のスパイラル#2の実施内容と同等とした。観測は自エージェントからの相対位置及び相対姿勢を使用するようにし、以下の16種類に見直した。

- ・ 自身のフリーズ残り時間 (1種類)
- ・ 味方の捕獲エージェント2体の相対位置・相対姿勢及びフリーズしている場合のフリーズ残り時間 (8種類)
- ・ 味方の捕獲エージェント2体の宣言 (2種類)
- ・ 逃走エージェントの相対位置・相対姿勢及びフリーズしている場合のフリーズ残り時間 (4種類、捕獲エージェントが共に逃走エージェントを視野に入れていない場合は逃走エージェントを最後に観測した時点の値を代入)
- ・ 逃走エージェントを最後に観測した時点からの経過時間 (1種類)

セルフプレイ強化学習では学習するチームをどのような比率でスワップするかを決定する必要がある。この連携捕獲の学習環境では、逃走エージェントは逃げるのみの比較的易しい問題であるとともに、今回獲得したい学習モデルは捕獲エージェントであるため、捕獲エージェントチームの学習頻度を大きく設定して試行錯誤した結果、捕獲エージェント：逃走エージェントの学習頻度を8:1と設定した。

各チームの平均累積報酬の推移を図3.5.1.3- 1に示す。セルフプレイでは対戦相手が強くなっていくため、学習中の報酬では学習の進捗具合を評価できない。よって、ELOレーティング[21]等の強さを示す指標を用いて学習の進捗具合を評価することが多い。各エージェントの学習中のELOレーティングの推移を図3.5.1.3- に示す。ELOレーティングはレーティング値が過大となるインフレーションまたは過小となるデフレーションが発生しやすいという特徴があり、図3.5.1.3- でもインフレーションが発生しているものの、各チームが過去のモデルよりも強いモデルを獲得している証左であり、学習が適切に行われていることがわかる。3. 4. 1. 3項と同様に捕獲エージェント及び逃走エージェントをそれぞれ人間操作に変えて実行したところ、どちらのケースにおいても捕獲エージェントは逃走エージェントを取り囲み適切に足止及び捕獲を宣言しつつ連携している様子が確認でき、3. 5. 4. 3項の比較実験において本項の強化学習モデルを使用することとした。

セルフプレイ強化学習は、徐々に強いエージェントを相手に学習するため、最初から強いエージェントを相手にするよりも学習が進みやすいというメリットがあるが、本項のようにチームによって報酬関数等が違う非対称な環境を扱うとき、学習をスワップさせる学習頻度が適切でない場合に以下の課題が発生する。

- ・ 例えば本学習の逃走エージェントのように相手エージェントがいない方向に移動するのみといった、一方のチームの扱う問題が簡単である場合、そのチームだけが急激に強くなってしまい、他方のチームの学習進捗具合と乖離し学習が進まなくなる。
- ・ 問題設定によっては、初めに一方のチームだけが学習すべき状況がある。

以上の課題に対処するため、各チームのレーティングが同程度で推移しながら学習

可能とするように自動で学習頻度を調整する発明をし、特許出願を行った。

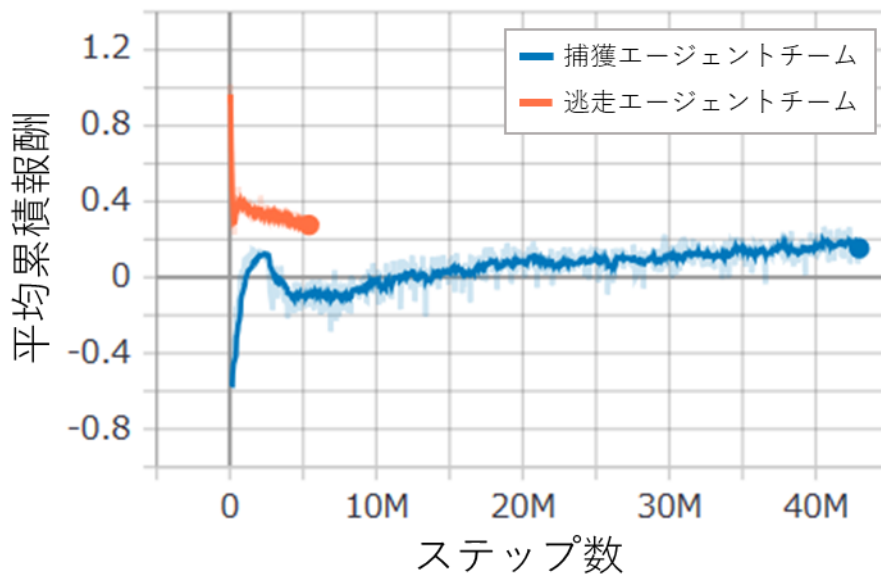


図 3.5.1.3- 1 平均累積報酬の推移

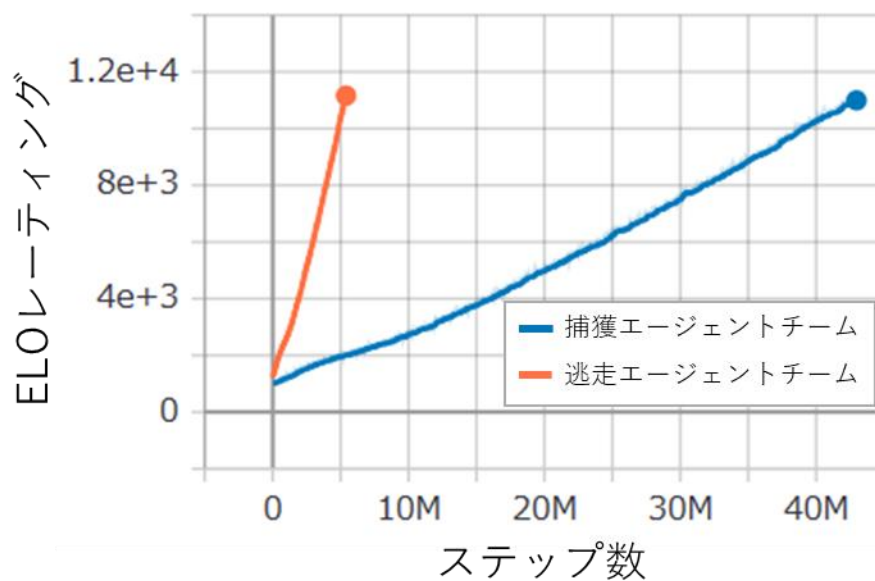


図3.5.1.3- 2 ELOレーティングの推移

3. 5. 2 HSIプロトタイプの実装

スパイラル#3のHSIプロトタイプでは、1. (3) ③-1. 人からAIエージェント群全体に対する効果的な指示方法の解明として、追加した集合タスクの適切な指示方法を検討し、実装を行った。この実装に関して、3. 5. 2. 1項に示す。また、1. (3) ③-2. AIエージェント群の将来予測結果の提示方法の解明としてAIエージェントの実行中タスク/役割をシンボルが直感的に理解できるように、ピクトグラムに変更する改修を行った。この改修に関して、3. 5. 2. 2項に示す。

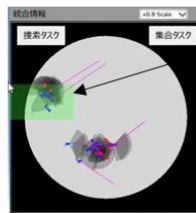
HSIプロトタイプの実装事項のうち、1. (3) ③-3. タスクのマーケット状況等の効果的な人への提示方法の解明については、3. 3. 2. 3項に記載した、スパイラル#1のHSIプロトタイプと同様であるため本項における記載は省略する。

3. 5. 2. 1 人からAIエージェント群全体に対する効果的な指示方法の解明

本項では集合タスクの指示に関し操作系の改修を行った結果及びHSIプロトタイプの内容を示す。

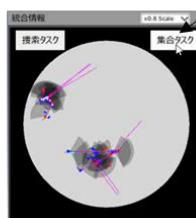
基本方針は、3. 3. 2. 1項に示すスパイラル#1の方針を踏襲したが、集合タスクの指示に関してはゲーム中盤の切迫度が高い状況が想定されることから、さらに入力回数を減らすように検討し、スパイラル#1では4回の操作が必要であったものを、3回の操作での指示を可能にした。その結果、契約エージェント数に関しては選択したAIエージェントの数で自動的に設定し送信されるようにした。図3. 5. 2- に、実際に設計及びプロトタイプを行った集合タスクのタスク告示から契約までの流れを示す。

手順1



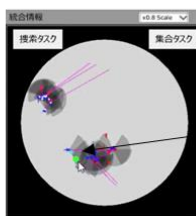
タスク告示を行うエージェントの範囲を選択する。

手順2



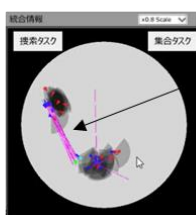
集合タスクを選択する。

手順3



集合位置(緑色の丸)を指定する。
自動的に契約エージェント数が決定され、さらに指示が送信される。

結果



集合位置へ移動開始する。
(マゼンダの予想経路が表示される)

図 3. 5. 2- 1 集合タスクのタスク告示から契約までの流れ

3. 5. 2. 2 AIエージェント群の将来予測結果の提示方法の解明

3. 4. 2. 2項に示したタスク/役割シンボルに関して、直感的にわかりやすいようピクトグラムに変更した。変更したタスク/役割シンボルの種類に関して表3.5.2.2-1に示す。スパイラル#2 表3.4.2.2-1の協調タスク実行を示す“契”のシンボルを廃止し、協調タスクに関しては丸型のシンボルに統一することで個別タスクと区別することとした。また、集合タスクでは各フェーズにシンボルを割り当てることとし、個別タスクの捕獲については回り込み中及び追跡中の区別をしないこととした。

表 3.5.2.2-1 タスク/役割シンボルの種類

集合-集合役 (協調タスク)	集合-整列役 (協調タスク)	集合-停止役 (協調タスク)
		
集合-反撃役 (協調タスク)	捕獲-捕獲役 (協調タスク)	捕獲-足止役 (協調タスク)
	 ※AI間は白色、人間からの指示の場合は橙色	 ※AI間は白色、人間からの指示の場合は橙色
捕獲-サポート役 (協調タスク)	捕獲-追跡役 (協調タスク)	搜索 (協調タスク)
 ※AI間は白色、人間からの指示の場合は橙色	 ※AI間は白色、人間からの指示の場合は橙色	
エリア逸脱回避 (個別タスク)	衝突回避 (個別タスク)	捕獲 (個別タスク)
		
搜索 (個別タスク)		
		

3. 5. 3 トレーニングの実施

3. 5. 1 項及び3. 5. 2 項の検討結果に対する有効性の確認と検討結果へのフィードバックのため、4vs4及び10vs10の小グループ実験を2回実施した。各小グループ実験の詳細を表3. 5. 3- 1に示す。

表 3. 5. 3- 1 小グループ実験の詳細

日時	参加人数 (内、研究従事者数)	確認事項
2020年 10月14日	5名 (3名)	連携捕獲タスクの追跡役が捕獲対象を囲い込むように動作することを確認した。本改修を担当した研究従事者はチームAとして対戦に勝つことができたものの、他の研究従事者ではチームAとして対戦に勝つことができなかったため、引き続きトレーニングの実施が必要と判断した。
2020年 12月17日	12名 (3名)	<ul style="list-style-type: none">・ 人からの連携捕獲タスク送信を多用して実験した。連携捕獲タスクの捕獲役が、捕獲対象に対して足止めされないことを確認し、10vs10の対戦においてチームAが全勝した。・ 集合タスク指示を多用して実験した。集合タスクの挙動が意図通りであることを確認し、10vs10の対戦においてチームAが全勝した。

3. 5. 4 フィールド評価の実施

3. 5. 4. 1 評価方法

3. 4. 4. 1 項に示したスパイラル#2の評価方法と同様に、スパイラル#3のフィールド評価を2021年1月7日に実施した。スパイラル#3の評価結果を3. 5. 4. 2項に示す。

2021年1月7日の評価終了後、チームA参加者1名のみで3. 5. 1. 3項で獲得した強化学習モデルを連携捕獲タスクと置き換えて30vs30の対戦評価を3回実施した。強化学習エージェントモデルとの協調評価結果を3. 5. 4. 3項に示す。

また、2021年2月4日に追加実験として、チームAの人エージェント操作者として同時に2名参加したケースを10vs10、20vs20及び30vs30の条件で各3回対戦評価を実施した。加えて、チームAが数的劣勢環境のケースとして10vs20及び20vs30の条件かつチームAには2人の人に参加してもらい、各対戦にはそのうちの1名がチームAの人エージェントを操作して3試合(1回の試行で2名×3試合、計6試合を行う)ずつ対戦し、2人分の評価結果を取得した。この追加実験の結果を3. 5. 4. 4項に示す。

3. 5. 4. 2 評価結果

スパイラル#3のフィールド評価の結果として、チームAの勝率を表3. 5. 4. 2- に、各チームの生存率を図3. 4. 4. 2- に示す。勝率は明らかに向上し、20vs20及び30vs30ではチームAが全てのチームBエージェントを捕獲して勝率100%となった。チームAの生存率でも10vs10で40%、20vs20及び30vs30で約60%となった。チームA参加者から「味方が散らばりやすかったので集合機能を使った。」や、「味方が分散している時は適宜集合タスクを用いて有利な状況を作ることを意識した。」といったコメントが得られており、スパイラル#3で追加した集合タスクにより、スパイラル#2で生起していたような、チームAのAIエージェントが分散して局所的に数的劣勢な状況に誘引される状況を抑制できた。また、チームBの人は劣勢になると回避傾向となり、簡単にはなかなか捕獲されない。例えばいくつかの対戦では、チームBの1人が領域の縁に留まり、誰からも見つからずに最後まで生き残ろうとする事例があった。チームAのAIエージェントは観測中のチームBエージェントが1体もない場合は個々にランダム検索(個別タスクの検索)を行うが、エージェントを発見するのは困難である。チームAに人エージェントがいる場合には、対戦開始からその時点までの経緯を踏まえて怪しい箇所を推測することができ、そこへ協調タスクとしての検索タスクを告示することで効果的に発見できる。チームA参加者の1人は「三回目の最後の1人がずっと端にいたので検索が困難だった。人間が指示を出さないと端にいる相手チームを見つけるのは大変かもしれない。」とコメントしている。人とAI群それぞれの優位点を活かした協働が行われた1つの証左といえる。

一方、10vs10では勝率が50%に留まったが、チームAの生存率がチームBを上回っており、チームAが優勢だったことが分かる。エージェント総数が少ない場合、チームA内における人エージェントの相対的な割合が増加するため、人が積極的に直接的な捕獲行動に関わることが求められるが、その能力がAIエージェントと比べて劣っていることが原因として考えられる。そこで、人とAI群が効果的に協働を行った結果としてチームBに勝利したわけではなく、単にAI群が優れているという懸念を払しょくするため、追加実験としてチームAに人エージェントを参加させない条件でも対戦評価を行った。その結果を表3. 5. 4. 2- 及び図3. 4. 4. 2- に「30vs30(チームA:人エージェントなし)」として示した。勝率はチームAに人エージェントがいなくても100%であり、AI群が優れていることは間違いないが、図3. 4. 4. 2- の生存率では人がいる場合に比べてチームAの生存率が低下しており、チームAの人エージェントがいることでさらに良い結果を出していることが分かる。これらのことから、人とAI群が効果的に協働していたと考えられる。

表 3.5.4.2- 1 スパイラル#3 におけるチーム A の勝率の結果

種別	チーム A の結果		
	勝利数	敗北数	勝率 [%]
10vs10	3	3	50
20vs20	6	0	100
30vs30	6	0	100
30vs30 (チーム A : 人エージェントなし)	3	0	100

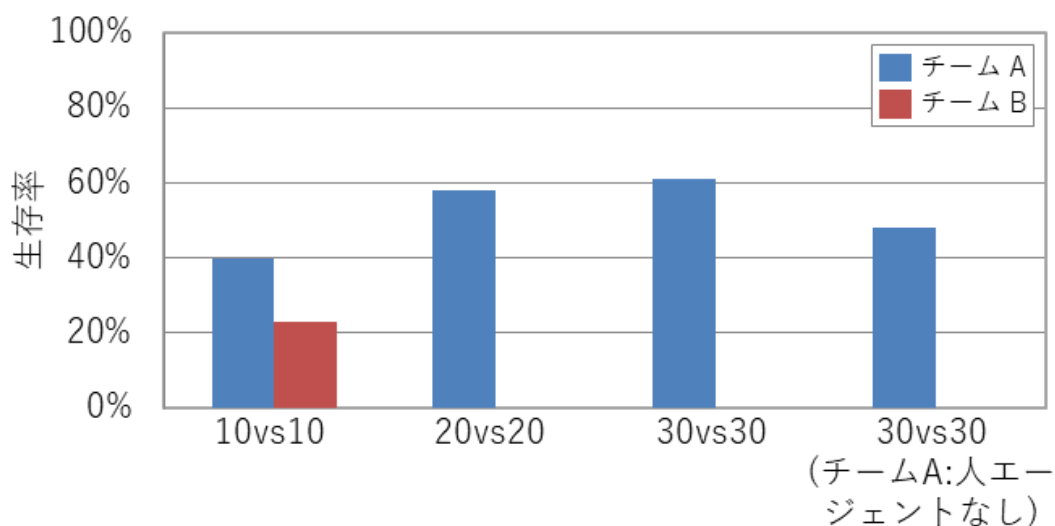


図 3.5.4.2- 1 スパイラル#3 における各チームの生存率

また、各試行後に取得した、チームA参加者によるMCH-UVD のレーティング結果を図 3.5.4.2- 1 に示す。レーティング結果は各回2名による評価の平均を記載した。スパイラル#3では全ての項目について“Display is Acceptable”以上の評価を得たことが分かる。スパイラル#2と#3では、“検索・探知”及び“捕獲・追跡”に関する表示、操作について変更をしていないが、「AIエージェントの性能向上」及び「AIエージェント群が分散しなくなった」というコメントと共にレーティング評価が向上しており、MCH-UVD のレーティングが表示、操作といった直接的なインターフェースに限らず、タスク実行性などの動作を含む包括的な評価となっていることがうかがえる。従って、2名のチームA参加者が共に高いレーティングを与えたことは、チームAの人にとってAI群が「タスク遂行で協働できるチームメイト」と評価したといえる。

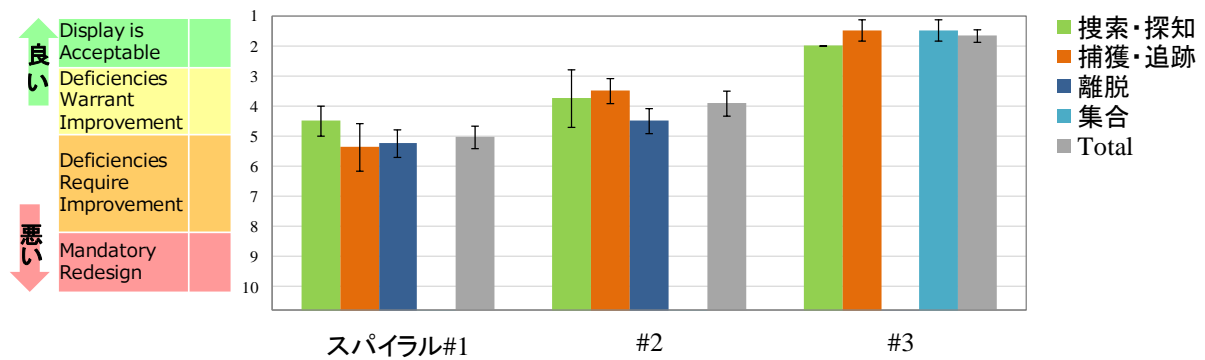


図 3.5.4.2- 2 スパイラル#1 から#3 までのレーティング結果の変遷

以上により、チームAに“人間の集団程度”の戦略レベルが備わっており、本HSIプロトタイプによって人とAI群が良好な協調行動を取れたことが確認できたといえる。

3.5.4.3 強化学習エージェントモデルとの協調評価結果

AIエージェントの連携捕獲タスクを3.5.1.3項で獲得した強化学習モデルで置き換え、30vs30の条件で評価した結果を示す。以下、本項の評価を学習評価と呼び、3.5.4.2項に示すスパイラル#3フィールド評価結果（チームAに人がいないケースを除く）を#3評価と呼び、区別する。

学習評価では3戦中3勝（勝率 100%）の結果となり、#3評価（30vs30）と同様となった。生存率を図3.5.4.3-1に示す。チームAの生存率は36%となり#3評価（30vs30）の61%の方が高い結果となったため、強化学習エージェントモデルの性能は連携捕獲タスクに及ばないものの遜色のない結果と言える。さらに定性的評価としての学習評価終了後アンケートにおいても、以下のコメントが得られている。

- ・ 少し苦戦している感じもしたが、連携捕獲タスクとの大きな変化は感じなかった。一方、同じ敵に何回もぶつかるなど、挙動不審なAIが何体かいた。

以上より、人間と強化学習エージェントモデルとの協働について将来性及び課題は以下の通りである。

【将来性】

- ・ 強化学習を用いることで、人間が設計して作りこんだものと遜色のないモデルを発見的に作成可能であり、設計期間の短縮が見込まれる
- ・ 強化学習と宣言の組み合わせによって人間とも連携しやすい見通しがある

【課題】

- ・ 生存率に関しては連携捕獲タスクを上回るものではない
- ・ アンケートコメントにより、人間が疑問に思う強化学習モデルの行動があった

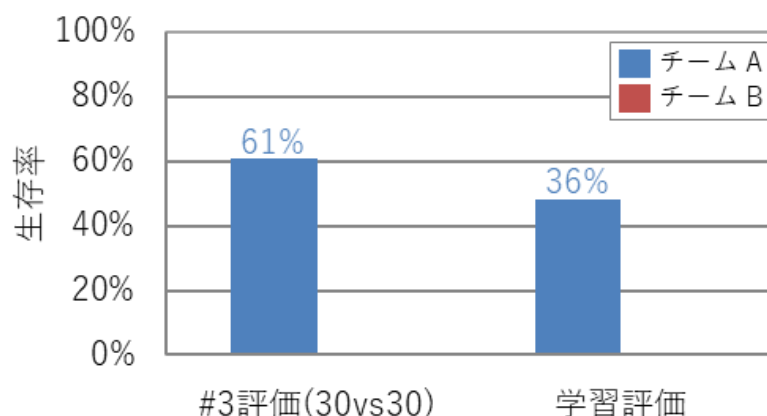


図 3.5.4.3- 1 連携捕獲タスクと強化学習エージェントモデルの生存率比較

3.5.4.4 追加実験結果

本項では追加実験として実施したチームAの人エージェントが複数いるケースとチームAが数的劣勢状況のケースについて実験結果を示す。

チームAの人エージェントが複数いるケースとして、チームA参加者を2人同じ対戦評価に参加させた。以降、このケースの評価を“人複数評価”と呼ぶ。ここで、チームA参加者同士はチャット機能で会話可能な条件とし、どちらの参加者からもAIエージェントへ指示可能とした。#3評価及び人複数評価の勝率の比較結果を表3.5.4.4- 1に示し、生存率の比較を図3.5.4.4- 1に示す。定性的評価としてアンケートコメントを以下に示す。

- ・ ゲーム開始時の人エージェントの位置によって、指示を出すAIエージェントを上下に分けて近い方に指示を出すといった分担をしたところ、特に不便はなかった。
- ・ 人エージェントは極力集合して離れないように立ち回った。

以上より、アンケートではネガティブな評価がなかったものの、生存率の点では全体的に成績が下がっている。チャット機能でチームA参加者2名の意思疎通が可能にはなっているものの、スパイラル#3までに作成したシステムは1人の人間とAI群の協調を根底に設計しており、人が複数になった場合の効率的な意思疎通の方法を考慮していないことが原因の1つと考えられる。

表 3.5.4.4- 1 #3 評価及び人複数評価の勝率の比較結果

	#3 評価の結果			複数人評価の結果		
	勝利数	敗北数	勝率 [%]	勝利数	敗北数	勝率 [%]
10vs10	3	3	50	1	2	33
20vs20	6	0	100	3	0	100
30vs30	6	0	100	3	0	100

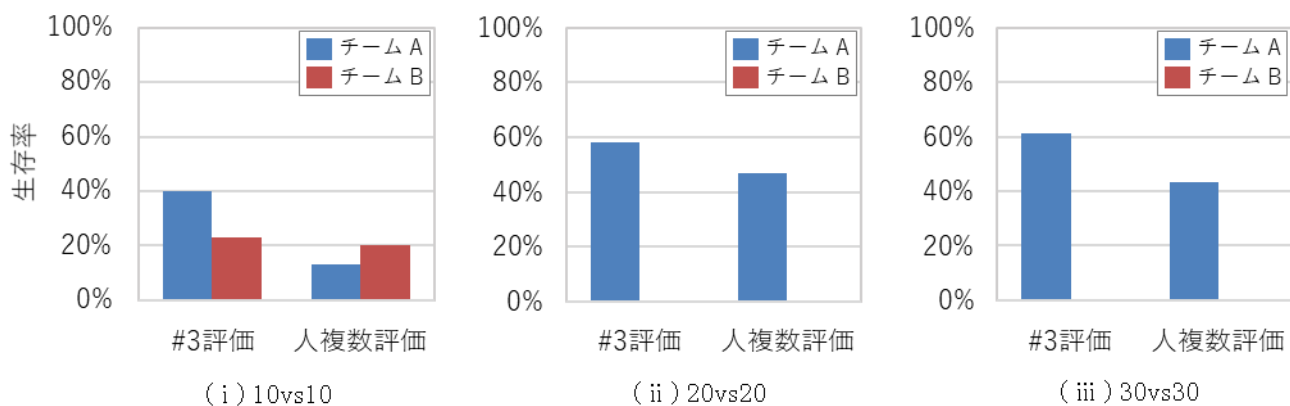


図 3.5.4.4- 1 #3 評価及び人複数評価の生存率の比較結果

次に、チームAが数的劣勢な状況の挑戦的なケースとして、10vs20及び20vs30の条件でチームAについて2人の参加者によりそれぞれ3回の対戦評価を実施した。以降このケースを“数的劣勢評価”と呼ぶ。数的劣勢評価における勝率は10vs20及び20vs30ともに、0勝6敗の勝率0%であった。#3評価及び人数的劣勢評価の勝率の比較結果を表3.5.4.4- 2表3.5.4.4- 1に示し、数的劣勢評価の生存率を図3.5.4.4- 2に示す。以上の結果から、しっぽ鬼の例題設定環境ではエージェント数が勝敗に大きく影響することがわかる。10vs20よりも20vs30の方がチームBの生存率が低くなっていることから、各チームのエージェント数の比率を考慮する必要があり、11vs20, 12vs20, 13vs20 …といったチームAを徐々に増やしていくような条件で対戦評価が可能であれば、チームAとチームBの勝敗が拮抗するようなエージェント数の比率が求められていた可能性がある。

表 3.5.4.4- 2 #3 評価及び数的劣勢評価の勝率の比較結果

	#3 評価の結果			数的劣勢評価の結果		
	勝利数	敗北数	勝率 [%]	勝利数	敗北数	勝率 [%]
10vs10	3	3	50	-	-	-
10vs20	-	-	-	0	6	0
20vs20	6	0	100	-	-	-
20vs30	-	-	-	0	6	0
30vs30	6	0	100	-	-	-

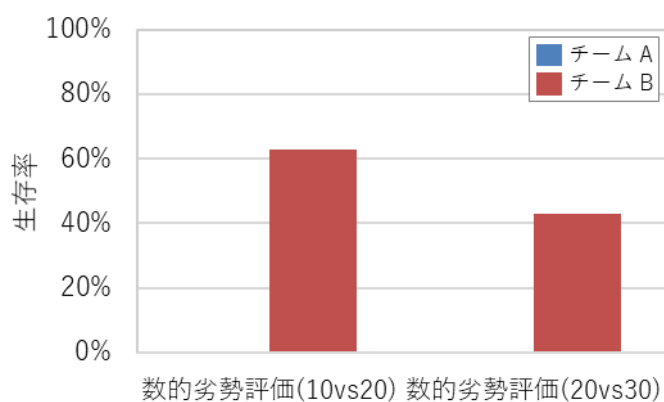


図 3.5.4.4- 2 数的劣勢評価の生存率

3. 5. 4. 5 スパイラル#3の考察

スパイラル#3では、スパイラル#2で抽出した以下の改善点を対策したことにより、人のみの集団であるチームBを凌駕する人とAI群が協働するチームAを構築したことが実証できた。

- ・ 離脱タスクを見直して、AI群を安全に集合させる集合タスクとした
- ・ 連携捕獲タスクの追跡役ロジックを改善し、捕獲対象の逃避先領域を限定するように連携しながら追跡を行うようにした
- ・ タスク/役割シンボルを直感的にわかりやすいようピクトグラムに変更し HSI を改善した

さらに、チームAに人エージェントを参加させない条件でも対戦評価を行い、チームAに人がいる場合の方がチームAの生存率が高く、コメントやレーティングといった定性評価の観点からもチームAの人とAI群が効果的に協働を行ったことが示された。定量的、定性的いずれの観点からも人とAI群が協働するチームAがAI群のみのチームA、人エージェントのみのチームBと比較して最も優れていた。

また、スパイラル#2で目途付けを行って追加した強化学習エージェントモデルとの協調評価では、人間が設計して作りこんだ連携捕獲タスクの成績を上回ることはなかったものの、十分な成績（勝率100%）を示し、強化学習の将来性を感じさせる結果となった。可読性のないブラックボックス方策を持つ強化学習モデルは、人との協働において大きな制約があると考えられるが、本研究で提案したようにチームとして実施するタスクを個別タスクと協調タスクとして定義し、役割分担に分解したうえでその範囲内の行動を学習で獲得させ、実行時にはどのタスクや役割を行っているかを表明させることにより、人と協働するうえで効果があることを示したことは大きな成果であると考えられる。一方、アンケートコメントとして人間が疑問に思う行動があったことなど、さらに改善すべき課題も存在することを示した。

最後に追加実験として、チームAの人エージェントが複数いるケースとチームAが数的劣勢状況のケースについて紹介した。

チームAに2人の人間が参加したケースでは、アンケートコメントでは否定的な内容はなかったが、勝率及び生存率の観点では1人の場合と比べて若干成績が低下した。スパイラル#3までに作成したシステムはチームAに複数の人エージェントが参加することが可能な作りとなっているものの、1人の人間とAI群の協調を根底に設計しており、人が複数になった場合の人同士の効率的な意思疎通の方法は今後の課題である。むしろ、人同士の協調に資するシステム側の支援がほとんどなくとも、人のみのチームBを凌駕する結果となったことは評価に値すると考える。

また、チームAがほとんどのケースで勝率100%という結果となったことを受けて実施した、チームAが数的劣勢状況からスタートする挑戦的なケースでは、評価を行った10vs20及び20vs30の条件では一気に勝率0%となり、例題設定が“数が正義”という狙った問題設定となっていることが裏付けられた結果となった。評価に大勢の人を動員するというリソース負荷の高い実験であったため本研究では行うことができなかったが、各チームのエージェント数の比率を変えていき、どこで拮抗するかによってチームAとチームBの能力差を定量的に示すことができる可能性がある。

参考文献

- [1] M. Benda, V. Jagannathan, et al. “On Optimal Cooperation of Knowledge Sources,” Technical Report BCS-G2010-28, Boeing AI Center, 1985.
- [2] Less Gasser, Nicholas Rouquette, et al. “Representing and Using Organizational Knowledge in Distributed AI Systems,” Distributed Artificial Intelligence, Volume II, 1989, pp. 55-78.
- [3] 大沢 英一, 沼岡 千里, et al. “分散人工知能における標準的小問題,” コンピュータソフトウェア, Vol.10, No.3, 1993, pp. 3-19.
- [4] M. Shor, “game theory .net - Games,” <<http://www.gametheory.net/Dictionary/games/>>, 2016, (2019.5.24閲覧).
- [5] Curlie.org: Open source games, <https://curlie.org/Computers/Open_Source/Software/Games> (2019.5.24閲覧).
- [6] David Churchill, “AIIDE StarCraft AI Competition - A History of Starcraft AI Competitions (and UAlbertaBot),” <<https://www.cs.mun.ca/~dchurchill/starcraftaicomp/history.shtml>>, 2016, (2019.5.24閲覧).
- [7] M. Jaderberg, W. Czarnecki, et al. “Capture the Flag: the emergence of complex cooperative agents,” <<https://deepmind.com/blog/capture-the-flag/>> , 2018, (2019.5.24閲覧).
- [8] Unity Blog: Unity と DeepMind の提携で加速する AI 研究, <<https://blogs.unity3d.com/jp/2018/09/26/unity-and-deepmind-partner-to-advance-ai-research/>> , 2018, (2019.5.24閲覧).
- [9] ウィキペディア日本語版:しっぽ鬼, <<https://ja.wikipedia.org/wiki/しっぽ鬼>> (2019.5.24閲覧).
- [10] M. Colledanchise, and P. Ogren, “Use of BTs in Robotics and AI, in Behavior Trees in Robotics and AI: An Introduction,” 1st Ed., CRC Press, 2018.
- [11] R. G. Smith, “The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver,” In IEEE Transactions on Computers, Vol. C-29, No. 12, December 1980, pp. 1104-1113.
- [12] G. E. Cooper and R. P. Harper, “The use of pilot rating in the evaluation of aircraft handling qualities,” NASA Technical Note, NASA TN D-5153, Apr. 1969.
- [13] B. Donmez, M. L. Cummings, H. D. Graham, S. Brzezinski, “Modified Cooper Harper Scales for Assessing Unmanned Vehicle Displays,” In Proceedings of the 10th Performance Metrics for Intelligent Systems Workshop, 2010, pp. 235-242.
- [14] C. E. Nehme, “Modeling human supervisory control in heterogeneous unmanned vehicle systems,” in Department of Aeronautics and Stronautics, Cambridge, MA: MIT, 2009.
- [15] S. Balakirsky, S. Carpin, et al. “Towards heterogeneous robot teams for disaster mitigation: Results and performance metrics from robocup rescue,” Journal of Field Robotics, 24, 2007, pp. 943-967.
- [16] R. S. Sutton and A. G. Barto, “Reinforcement Learning,” The MIT Press, 1998. 三上 貞芳, 皆川 雅章 訳, “強化学習,” 森北出版, 2000, pp.55-56.
- [17] A. Juliani, V. Berges, et al. “Unity: A general platform for intelligent agents,” arXiv preprint arXiv:1809.02627, 2018
- [18] J. Schulman, F. Wolski, et al. “Proximal policy optimization algorithms,” arXiv preprint arXiv:1707.06347, 2017.
- [19] J. Schulman, S. Levine, et al. “Trust region policy optimization,” arXiv preprint arXiv:1502.05477, 2015.

- [20] H. Huang, W. Zhang, et al. "Guaranteed decentralized pursuit-evasion in the plane with multiple pursuers," Proceedings of IEEE Conference on Decision and Control, 2011, pp. 4835-4840.
- [21] A. E. ELO, "The Rating of Chess Players, Past and Present," B.T. Batsford Ltd., London, 1978.

4. 委託業務全体の成果

4. 1 計画時に想定していなかった成果（副次的成果）や、目標を超える成果

(1) タスク分解と役割配分

図0-1「社会的に考えられるスウォーム対処の例」に示したようなチームプレイを必要とする対処問題（ここではスウォーム対処と呼ぶ）は、自動車の運転のように多数の無作為で一般的な人の集合による協調とは本質的に異なり、事態対処を目的として訓練を積む特定のプロフェッショナル集団によって対処が行われる。こうしたスウォーム対処では、従前から個別タスクと協調タスクの仕様設計と協調タスクのタスク分解による役割の定義は事前に行われている。例えばTambeは、複雑で動的な環境における不確実性を克服し、チームが首尾一貫して行動できるようにするチームワークの一般的なモデルとしてSTEAMを提案した[1]。STEAMでは、チームワークの明示的なモデルとして、エージェントの組織階層と協調タスクを設計する。その際に重要な概念が役割である。役割とは、チームの全体的な活動のための個々あるいは部分的なチーム（サブチーム）が担う一連の行動セットに対する抽象的な仕様を示す。役割を各エージェントに割り当てるためには、タスクを適切なレベルの粒度に分解する必要がある。StoneらはRoboCupサッカーチームを例に、同種エージェントによるマルチエージェントタスクを柔軟な役割に分解するチームワーク構造例を提示した[2]。Stoneらの提案では、タスクを抽象化して表現した「役割」を一連の手続きとして事前に定義し、エージェント間でこれを「ロッカールーム合意」として共有しておくことにより、限定的な通信によって実用的な協調が実現可能なことを示した。しかし、これらはAI群の構築に関する取り組み例であり、人とAI群の協調を対象としたものではない。

本研究では、AIエージェントの構築は「人とAIが効果的に分担可能なタスクの仕様の設定」及び「各エージェントにおける適切なタスクの条件、タスクの価値基準の設定」という手順でBTとCNPを用いて可視的に実施した。端的に言えばエージェントの方策をプリプログラムしたといえる。人手によってAIの動作を設計するプリプログラムとしたのは、人とAI群が協働するためには、Tambe及びStoneらの先行研究を踏まえれば、AIエージェントがこれから行おうとする一連の行動が説明可能で人にとって理解可能でなければならないということが念頭にあったためである。AIの動作設計では、例題設定において考えるタスクの仕様を「一つのエージェントだけで遂行可能なタスク」を個別タスク、「複数のエージェントによって協力して実行する必要があるタスク」を協調タスクとして分類して設定し、そのタスクをいつ、どのように実行するかを設計した。スパイラルを重ねる中でタスク仕様は何度も見直しを行い、その過程で特に複数エージェント間での協力を要する協調タスクの設計の流れを次のように整理した。

- (i) 人がAI群と協力して行おうとするミッションを完遂するために、どのようなタスクがあるのかを整理する
 - ・ タスクは単一エージェントによって実行可能な個別タスク（縁から落ちない等）と、複数のエージェントによって実行することが必要な協調タスク（Aが相手を足止めし、Bが脆弱な状態になるAのしっぽの安全を確保し、その間にCが相手のしっぽを取る、等）に分けて定義する
 - ・ 協調タスクには、そのタスクを完遂するために必要な複数の役割（チームの全体的な活動のための個々あるいは部分的なチーム（サブチーム）が担う一連の行動セット）が存在する
- (ii) 協調タスクを役割に分解する（これをタスク分解と呼ぶ）
- (iii) 協調タスクを実行するサブチームの構成には、エージェント間交渉及び合意形成手法を利用できる（本研究ではCNPを利用）

- (iv) サブチームが構成され、予め協調する人がタスクと役割について理解していれば（コンセンサスがあれば）、AIが強化学習モデルのようにブラックボックスであっても、AIがこれから行おうとするタスクあるいは役割を宣言することで協調が可能である

AI群同士の協調にフォーカスしたTambeのSTEAM及びStoneらの成功例を参考に、同様の検討方法によって人同士のチームプレイを凌駕する実用性を有する人とAI群の協調が実現できたことは、今後の人とAI群のチームワークの実現に向けて汎用的な協調タスク設計に向けたLessons Learnedを提供できたものとする。

(2) 人と協働可能な強化学習モデル

研究代表者は、4.1(1)に示したように、人とAI群が協調するためには、人にとってAIエージェントの行動が説明可能で理解可能でなければならないとの考えから、近年隆盛著しい機械学習による行動モデルの獲得、特に深層強化学習を用いた価値あるいは方策を有するエージェントの（人との協働に限った場合の）有用性には懐疑的であった。

しかし、3.5項で示したように、協調タスクのタスク分解と役割の定義を行った上で、明確化されたタスクと役割の範疇で強化学習を行うことにより、人にとって意図が理解可能な強化学習エージェントを構築することが可能であることを示した。本研究では、人手によってAIの動作を設計した場合と比較して、被捕獲率及び参加者のコメントにおいて強化学習エージェントが劣る結果となったが、捕獲率はどちらも100%であり、大きく差のある結果ではない。これは、強化学習によっても同等の性能を獲得できる可能性を示唆している。現時点では強化学習にあたってハイパー・パラメータの調整やエージェントの入出力や報酬の設計に様々な勘所が存在し、容易に有用なモデルが得られるわけではないが、強化学習によって人と協働が可能で、かつ人手による設計にかかる期間を大幅に短縮できる可能性を示したことは、当初計画にはなかった望外の成果と言える。

参考文献

- [1] M. Tambe, “Towards Flexible Teamwork,” *Journal of Artificial Intelligence Research*, Vol. 7, No. 1, 1997, pp. 83-124.
- [2] P. Stone, and M. Veloso, “Task decomposition, dynamic role assignment, and low-bandwidth communication for real-time strategic teamwork,” *Artificial Intelligence*, Vol. 110, No. 2, 1999, pp. 241-273.

4.2 研究課題の発展性（間接的成果を含む）

本研究の成果は、少数の人間とAI群が連携し協働する有人・無人連携を行うアプリケーションに幅広く応用が可能である。例えば、図0-1に示した資源探査等のエリア探索をはじめ、追い込み問題としての農林水産業及び畜産業における応用、捜索救難及び消防等の災害対処、警察機関が行う警備行動から防衛対処など、現状では人によるプロフェッショナル集団が行っているミッションを無人化または省力化する場合が直接的な応用先として考えられる。特に、場合によっては人命危険に及ぶような重大な法規的判断を伴う事態対処を要するような対象問題では、完全無人化が困難で人による関与が必須となるため、適用性が高い。

さらに4.1項に述べたように、人との連携を前提とするAI群の設計法及び強化学習適用法としては、本論文で対象とした例題のように競争的な環境に関わる問題に限らず、多くの分野に示唆を与え発展に寄与し得ると考える。例えば現状のファクトリー・オートメーションや物流における無人システムは、人の作業員とは隔離した環境で稼働させるか、人とは異なる役割のみを予め明示的に割り当てていることがほとんどである。タスク分解と役割を再定義し、自動化システムが役割を宣言するようにすることで、作業員とシームレスにタスクを遂行するような発展的適用が考えられる。

4. 3 研究成果の発表・発信に関する活動

防衛装備庁が行う研究開発の取組を講演やポスター、展示を通じて紹介する技術シンポジウム2019（2019年11月12～13日開催）において、オープンイノベーションをテーマに開催されたパネルディスカッションにパネリストとして登壇し、本研究を含めた企業における研究開発について産業界の立場から情報を発信した。

また、防衛大学校 先端学術推進機構の統合先端科学研究中間発表会（2020年10月28日開催、研究課題：武力紛争の人道化に向けた遠隔操縦ロボット用統合状況把握システムの開発）に招へい頂き、人工知能の倫理面含む広範な課題について議論を行った。

5. プロジェクトの総合的推進

5. 1 研究実施体制とマネジメント

運営委員会の開催及び外部アドバイザーの意見聴取を行った。

運営委員会及び外部アドバイザー意見聴取に関する開催状況を表5.1- 1に記載し、各結果について5. 1. 1項及び5. 1. 2項に示す。

表5.1- 1 運営委員会及び外部アドバイザー意見聴取の開催状況 (1/2)

年月日	開催名称	実施場所	備考
2019/1/24	金出教授意見聴取 (FY2018 #1)	株式会社i-TC (金出教授ご自宅)	-
2019/2/4	運営委員会 (FY2018 #1)	TV会議	-
2019/2/5	松尾教授意見聴取 (FY2018 #1)	産業技術大学院大学	-
2019/3/8	運営委員会 (FY2018 #2)	TV会議	-
2019/3/13	松尾教授意見聴取 (FY2018 #2)	産業技術大学院大学	-
2019/4/8	運営委員会 (FY2019 #1)	TV会議	-
2019/5/31	運営委員会 (FY2019 #2)	TV会議	-
2019/6/27	運営委員会 (FY2019 #3)	TV会議	-
2019/7/2	金出教授意見聴取 (FY2019 #1)	安保ホール	-
2019/7/19	松尾教授意見聴取 (FY2019 #1)	産業技術大学院大学	-
2019/7/29	運営委員会 (FY2019 #4)	TV会議	-
2019/8/30	運営委員会 (FY2019 #5)	TV会議	-
2019/10/4	運営委員会 (FY2019 #6)	TV会議	-
2019/10/18	松尾教授意見聴取 (FY2019 #2)	産業技術大学院大学	-
2019/11/6	運営委員会 (FY2019 #7)	TV会議	-
2019/11/18	金出教授意見聴取 (FY2019 #2)	安保ホール	-
2019/12/6	運営委員会 (FY2019 #8)	TV会議	-
2020/1/17	運営委員会 (FY2019 #9)	TV会議	-
2020/1/22	松尾教授意見聴取 (FY2019 #3)	産業技術大学院大学	-
2020/2/14	運営委員会 (FY2019 #10)	TV会議	-
2020/3/18	松尾教授意見聴取 (FY2019 #4)	電話会議	-
2020/3/19	運営委員会 (FY2019 #11)	TV会議	-
2020/3/27	金出教授意見聴取 (FY2019 #3)	Web会議	-

表5.1- 1 運営委員会及び外部アドバイザー意見聴取の開催状況 (2/2)

年月日	開催名称	実施場所	備考
2020/6/15	運営委員会 (FY2020 #1)	Web会議	-
2020/7/9	松尾教授意見聴取 (FY2020 #1)	Web会議	-
2020/7/15	運営委員会 (FY2020 #2)	Web会議	-
2020/8/3	金出教授意見聴取 (FY2020 #1)	Web会議	
2020/8/24	運営委員会 (FY2020 #3)	Web会議	-
2020/9/18	運営委員会 (FY2020 #4)	Web会議	-
2020/9/28	松尾教授意見聴取 (FY2020 #2)	Web会議	-
2020/10/19	金出教授意見聴取 (FY2020 #2)	Web会議	-
2020/10/23	運営委員会 (FY2020 #5)	Web会議	-
2020/11/16	松尾教授意見聴取 (FY2020 #3)	Web会議	-
2020/11/20	運営委員会 (FY2020 #6)	Web会議	-
2020/12/25	運営委員会 (FY2020 #7)	Web会議	-
2021/1/22	運営委員会 (FY2020 #8)	Web会議	-
2021/1/25	金出教授意見聴取 (FY2020 #3)	Web会議	-
2021/1/25	松尾教授意見聴取 (FY2020 #4)	Web会議	-
2021/2/26	運営委員会 (FY2020 #9)	Web会議	-

5. 1. 1 運営委員会

研究実施者が社内の複数拠点に跨って所在していることから、毎月約1回の定例会議として、三菱重工業株式会社の丸の内本社、相模原工場、名古屋誘導推進システム製作所の3拠点をTV会議もしくはWeb会議で接続し運営委員会を開催した。

研究開始当初は必要に応じ不定期に行っていた社内調整に代わり、2018年度は運営委員会を2回実施した。これに引き続き2019年度中は11回、2020年度中は9回実施した。運営委員会ではプロジェクト全体の進捗状況を確認し情報の横通しを行うことで、各拠点の連携・調整にあたった。

別拠点の研究実施者との調整には、2019年度までは社内のTV会議システムを用いたほか、フィールド評価へ参加者として参画させるなど、運営委員会のほかにも極力面着の機会を設けて各拠点の研究実施者との緊密な連携に努めた。2020年度は新型コロナウイルスの感染拡大に伴い面着による会議の実施を回避したが、新型コロナウイルス対策として社内に新たに整備されたWeb会議システムを活用し、動画やプログラム実行画面の共有などを通じて面着会議と遜色のない運営委員会を開催した。

5. 1. 2 外部アドバイザー意見聴取

採択条件であった『状況判断、相手の戦略把握に基づくゲーム理論的なアプローチに関する実施内容を追加し、「AIと人の協働」に関する理論的な掘り下げをすること。このため、AI理論やゲーム理論の専門家を分担者又は外部アドバイザーとして加えること。』に従い、以下2名の外部アドバイザーとの技術コンサルタント契約を締結し、意見聴取を実施した。

金出 武雄 教授

研究分野：コンピュータビジョン、ロボット工学、人工知能、マルチメディア

狙い：研究課題解明方針、研究の方向性に対する世界的実績に基づく指導

略歴：

1974年	京都大学 大学院工学研究科 博士課程修了
1974-1976年	京都大学 工学部 助手
1976-1980年	京都大学 工学部 助教授
1980-1985年	カーネギーメロン大学 ロボティクス研究所および計算機科学科 高等研究員
1985-1994年	カーネギーメロン大学 ロボティクス研究所及び計算機科学科 教授
1992-2001年	カーネギーメロン大学 ロボティクス研究所長
1993-1998年	カーネギーメロン大学 ワイタカー冠教授
1998年-	カーネギーメロン大学 ワイタカー冠全学教授
2004-2010年	産業技術総合研究所 デジタルヒューマン研究センター長
2006-2012年	カーネギーメロン大学 生活の質工学センター長
2014年-	大阪大学 産業科学研究所 特任教授
2014年-	奈良先端科学技術大学院大学 情報科学研究科客員教授
2015年-	産業技術総合研究所 名誉フェロー
2016年-	理化学研究所 革新知能統合研究センター 特別顧問
2017年-	京都大学 高等研究院 招聘特別教授

松尾 徳朗 教授

研究分野：コンベンション経営学、情報経済学、応用情報学、材料情報学

狙い：AI理論、ゲーム理論についてのコメント聴取

略歴：

2006年	名古屋工業大学大学院工学研究科博士後期課程 修了
2006-2012年	山形大学大学院理工学研究科 准教授
2010-2011年	カリフォルニア大学アーバイン校 客員研究員
2011-2014年	名古屋工業大学 プロジェクト教授
2011-2013年	上海大学 客員研究員
2011年-2018年	セントラルミシガン大学SEITI* リサーチフェロー *: Software Engineering Information Technology Institute
2012年-2020年	産業技術大学院大学 情報アーキテクチャ専攻 教授
2015年-2017年	インドネシアBina Nusantara University 客員教授
2015年-	名古屋工業大学 プロジェクト教授
2016年-	ネバダ大学ラスベガス校 客員教授
2017年-	AIIT*人工知能とサービス科学研究所長 *: Advanced Institute of Industrial Technology (産業技術大学院大学)
2018年-2020年	中国澳門市立澳門城市大学數據科學研究院招聘教授
2019年-	インドネシアBina Nusantara University 招聘教授
2020年-	台湾Asia University(亜州大学) 客員教授
2020年-	名古屋工業大学 客員教授
2020年-	東京都立産業技術大学院大学産業技術研究科 教授

金出教授には計7回、松尾教授には計11回、数か月単位での研究進捗をご報告し、適時非常に効果的なご意見を頂いた。両先生においては大変ご多忙の中、毎回予定時間を超過するほど熱心にご指導を賜り、感謝申し上げたい。

当初は極力面着の機会を設け有意義な意見を聴取すべく努めたが、2020年3月以降は新型コロナウイルスの感染拡大に伴い、面着による意見聴取の実施を回避し、Web会議システムを活用した。Web会議システム上において、動画やプログラム実行画面の共有などを通じて面着会議と遜色のない意見聴取を開催した。

多数回に及ぶご指導において聴取した外部アドバイザー意見は、「A：実施する」、「B：実施検討する」、「C：実施済みもしくは実施しない」及び「D：対象外(コメントのみ)」の4種類に分類し、処置方針を明確にした。外部アドバイザーから聴取したコメントの中から、処置が必要な分類 Aのコメントを処置結果とともに 表5.1.2-1 に示す。

また、最終回の外部アドバイザー意見聴取（両先生ともに2021/2/25 に実施）では本研究全期間の成果をご説明し以下のように高評価をいただいた。

- ・ 金出教授：約3年間、本研究の報告を聞いていて興味深いものと思った。研究成果について適切な理解を得るために、公表先の媒体毎に対象とする読者を意識して、上手にアピールすること。
- ・ 松尾教授：既に良い結果は得られているため、残りの期間でチャレンジングな追加実験に挑戦してもよいと思う。例えば、チームAの人エージェントが複数になった場合に新しい技術課題が見いだせないだろうか。本研究成果により、国益に資する具体的なフィールドはいくつも存在するので、さらに大規模な研究プロジェクトの提案と大型の研究費の獲得による研究継続およびプロダクトの提供が強く望まれる。

表5.1.2-1 外部アドバイザーから受領したコメント(抜粋) (1 / 4)

No.	指摘者	分類	コメント	処置結果
1	金出教授 (FY2018 #1)	HSI	情報の入力方法について、AI へはデータリスト、人間には画像というようにI/Fが異なる方式を使うことは問題ないと思う。しかし、人間へのデータの提示方法(画面の大きさ、視点の位置及びエージェント表示の抽象化/記号化等)は人の“ゲームの強さ”に密接した要素であると思われるので、HSI 検討時に十分に検討すること。	スパイラル#1において複数のHSI画面を試作し評価した。スパイラル#2ではAIエージェントのタスクシンボル等を追加し、スパイラル#3ではシンボルをピクトグラムに変更するなどの改善を行った。
2	金出教授 (FY2018 #1)	評価方法	一般的にゲームがうまい人は熟考せずに反射的に操作していると思われるが、即応的な挙動ではAIが勝るため、ゲーム内の時間に関わる設定項目は十分に検討し、協調要素で勝ち負けが分かったことが分かるような設定になれば望ましい。	スパイラル#1にてエージェント速度の調整を実施し、人間の操作として無理な設定ではないことを確認した。
3	松尾教授 (FY2018 #1)	例題設定	こうしたリアルタイムな対戦形式の例題に対し人とAIの混合チームとして協調を目指す設定は既存研究としてはあまり見たことがなく興味深い。②複数の戦略、③不完備情報の設定が重要であり、取りうる戦略が多いほど現実問題に近いものになる。	②複数の戦略、③不完備情報の設定を有する「しっぽ鬼」を環境に採用した。
4	松尾教授 (FY2018 #2)	AI	協調行動の種類については様々考えられる。まずはゲームで考えられる戦略を列挙し、協調行動を定義する。その後に、協調行動の選択・実行切り替え方法を検討するやり方が良いのではないかな。	スパイラル#1で協調行動を定義し、スパイラル#3で一部の協調行動の見直しを行った。
5	金出教授 (FY2019 #1)	評価方法	スパイラル#1のフィールド評価は、現実的に実施できる選択として数十人程度で、1人の人が1エージェントを直接操作する前提で問題ないと考え。人の都合がつけられるのであれば、フィールド評価を行う際、例えば2vs2→4vs4→6vs6…と段階的にエージェント数を変化させ、ゲームのパフォーマンスを評価するのも良いだろう。	スパイラル#1のフィールド評価では10vs10及び20vs20で1人の人が1エージェントを直接操作する条件で実施した。スパイラル#2及び#3においては30vs30の条件も追加した。
6	松尾教授 (FY2019 #1)	AI	現段階では堅実にAIエージェントのロジックをヒューリスティックなルールで実装することで良いだろう。将来の発展性として人間が犯しやすいミスを誘うような仕組みを考えるのも面白いと考える。	AIエージェントのロジックはBTを使用してヒューリスティックなルールを実装し、発展性も確保した実装とした。

表5.1.2-1 外部アドバイザーから受領したコメント(抜粋) (2 / 4)

No.	指摘者	分類	コメント	処置結果
7	松尾教授 (FY2019 #1)	AI	タスクとして相手に関する情報を探る行動があると考え。例えば、価値の低いエージェントを前線に出し不観測の領域を観測するための行動をさせるタスクなどがあるだろう。	人が AI エージェントへ「検索」指示をすることにより不観測領域を観測させることを可能にした。
8	金出教授 (FY2019 #2)	評価方法	評価指標として、双方のエージェントの生存数の時歴も有効だと考える。	双方エージェントの生存数の時歴をログに残すようにし、評価指標として取り上げた。
9	松尾教授 (FY2019 #3)	評価方法	チーム B の参加者はそれぞれのゲーム状況及び自身の行動の変化を忘れてしまう恐れがあるので、ゲームの合間にアンケートやメモを記入してもらうようにすること	フィールド評価参加者へ各トライアルの合間に記入するメモ用紙を配布し、最終アンケート記入の参考とするとともに、回収し分析時の参考とした。
10	松尾教授 (FY2019 #3)	評価方法	アンケートを取るときは社会科学の側面を考慮する必要がある。どういう情報を与えるとどのように考え方が変わったか分かるようにアンケートを取っておくと良い。そのためには、その人の背景情報を事前にアンケートで取得すること。	フィールド評価の際、事前アンケートを実施し、「普段からゲームをするか？」といった背景情報を取得した。
11	松尾教授 (FY2019 #4)	対外発表	対外発表先として、AAI2020 (9th International Congress on Advanced Applied Informatics)はどうか。	対外発表先として AAI2020 にて口頭発表を実施した。
12	金出教授 (FY2019 #3)	AI	AI エージェント同士での連携を強化学習によって獲得する取り組みは実施すべき。但し、AI エージェント間の連携が強くなることで、「研究の本来の目的である『人と AI の協働』に資するものか」という観点でも考えておく必要がある。	スパイラル#2 及び#3 において AI エージェント同士での連携を強化学習によって獲得する取り組みを実施し、スパイラル#3 では人間と強化学習モデルを用いた AI エージェントの協調を評価した。
13	松尾教授 (FY2020 #1)	HSI	AI の宣言を視覚的に追加するのはいいと思う。一方、チーム A に複数の人間が参加することを想定した場合、人から人へのコミュニケーション方法と、AI から人へのコミュニケーション方法が異なるためミスコミュニケーションが発生する可能性がある。チーム A に複数の人間が参加する場合には、以上のことを配慮する必要があるのではないか。	スパイラル#3 にてチーム A の人エージェントを 2 体にして追加評価を行い、課題を抽出した。

表5.1.2-1 外部アドバイザーから受領したコメント(抜粋) (3 / 4)

No.	指摘者	分類	コメント	処置結果
14	松尾教授 (FY2020 #1)	対外 発表	AAI 2020 の査読結果が良かったことから、論文誌 IJSCAI(International Journal of Smart Computing and Artificial Intelligence)への推薦もあり得ると考える。論文誌投稿の推薦があれば投稿してはどうか。	論文誌 IJSCAI へ論文投稿を実施した。
15	金出教授 (FY2020 #1)	HSI	ヒューマン・インターフェースの基本原則は、不要なことを表示しないということなので、協調タスクを示す“契”のシンボルは不要ではないか。	スパイラル#3にて、“契”のシンボルを削除し、ピクトグラム化かつ色を変化させることで表示をわかりやすくした。
16	松尾教授 (FY2020 #2)	AI	逃げ回っているチーム B に AI エージェントが一方向から犬追いしており、チーム A の戦力が無駄遣いになってしまっていることがある。 例えば、局所的な戦闘において、少数の追加エージェントの手助けがあれば、チーム A が優位に立てるにも拘わらず、助けに行かないことで局所的に不利になっているケースがあるのではないか。	スパイラル#3にて、ボロノイ領域を用いた追跡ロジックを実装した。加えて、集合タスクを実装して局所的にエージェント数で優越することを可能とした。
17	松尾教授 (FY2020 #2)	AI	様子を見ている捕獲対象(人)がいた場合に、AI エージェントが牽制を行う行動を取るなど、複数の戦略が切り替えられるようにするような改善も必要ではないか。	スパイラル#3にて、集合タスクの指示をすると AI エージェントが防御的な行動をとり、集合タスクを解除すると AI エージェントが攻撃的になるといった、複数の戦略を切り替えることを可能とした。
18	金出教授 (FY2020 #2)	HSI	人が指示する捕獲タスクを考えると、エージェント総数が増加することで、人間が AI エージェントに影響を与えられる割合が下がってくる可能性がある。チーム A の操作が全般的ストラテジの指示だけならば、戦績は全般的ストラテジの良し悪しで決まり、エージェント総数には依らないはず。 スパイラル#2 の現状は全般的なストラテジの指示が出来てないのではないか。	スパイラル#3にて、集合タスクによりゲーム中盤においても全般的なストラテジの指示を可能にした。
19	金出教授 (FY2020 #2)	HSI	乱戦になった場合、仕切り直して集合するような指示が出せれば有効ではないか。	スパイラル#3にて、集合タスクにより乱戦になった場合にも状況を立て直すことを可能とした。

表5.1.2-1 外部アドバイザーから受領したコメント(抜粋) (4 / 4)

No.	指摘者	分類	コメント	処置結果
20	金出教授 (FY2020 #2)	投稿論文	AAI2020 の論文について、主張したい内容が不明確である。何を最も訴えたいのか、ポイントを絞って分かりやすく記述すべきである。例えば効果的な協調のために BT+CNP という組み合わせが重要なのか、効果的な協調を評価するために単に BT+CNP をツールとして使っているだけなのか等を整理して論文の主張を組み立てること。	当該論文を作成した時点では、効果的な協調のために BT+CNP という組み合わせが重要と考えていたが、研究を続けている過程で、特定の手法を採用したことは単にツールとして選択したに過ぎず、本質的には以下が本研究の成果であると再整理した。 ・可視的な AI 行動定義 (BT 等) と、エージェント間交渉及び合意形成手法 (CNP 等) を組み合わせれば、人と AI 群の協調を促進可能であること。 さらに副次的成果を 4. 1 項のように整理した。今後、論文発表等あたってはこれらの整理結果を反映していく。
21	金出教授 (FY2020 #3)	投稿論文	No. 20 のコメントについて仮に後者である場合、ツールの使い方など、適用方法に意味がある場合には論文になるが、ツールの適用方法を細部まで説明しないとイケない内容では、良い論文とは言えない。 本研究でもツールの使い方などの詳説を必要とせず理解してもらえような成果があるのではないか。例えば、「人間に対してある情報をどのような見せ方をすれば良いかが分かった」などのことは言えないだろうか。	No. 20 のコメントと当方処置結果も念頭に置き、2020 年度投稿分 (人工知能学会論文誌速報論文) の論文を執筆した。 他にも、AI エージェントが提示する役割宣言シンボル表示の改善効果や、ボロノイ領域分割に基づいて捕獲対象を追い詰めていく機能による効果など、細かい粒度で論文としてまとめられる内容は幾つか考えられる。スコープを絞った方が理解も得やすいと考えられるので、成果公表にあたって考慮していく。

5. 2 経費の効率的執行

研究の実施においては、社内既存設備を最大限活用した。特に、フィールド評価実施時には社内ネットワークインフラ及びOA作業用のPCを最大限利用することで予算を最小限に留めた。

設備備品については、シミュレータ開発に必要となるシミュレータ開発用PC及び開発環境ゲームエンジンUnity、強化学習の実行及び解析に必要となるシミュレータ計算機など、真に研究実施に必須となる設備備品の購入に限定した。

また、当社複数拠点にまたがる研究分担者との会議及び外部アドバイザーの意見聴取については極力Web会議を使用し出張旅費を制限するなど効率的な予算執行を徹底した。

6. まとめ、今後の予定

本研究では、動的に変化する切迫した対処すべき対象が多数あり、少数の人間のみで対処することが困難な場面を想定し、人が多数のAIエージェントとチームワークを発揮して協働することにより対処可能とすることを目的とした基礎研究を実施した。

初年度には、対象問題として必要な要素課題を有し現実問題へ拡張性のあるPredator-Preyの例題設定としてチーム形式の”しっぽ鬼“を選定し、シミュレーション環境を構築して人とAIエージェントの協調行動の評価を可能とした。以後はAIエージェント及びHSIプロトタイプ of 構築と評価を繰り返して改善していくアジャイル的な方法により実施し、R2年度とR3年度の2か年に3回のサイクル（これをスパイラルと呼ぶ）を繰り返した。チームワークを実現するためのタスク仕様と役割を定義して、可読性の高いBehavior Treeにより定義し、契約ネット・プロトコルで分散的にタスクと役割を分担できるアーキテクチャを有したプロトタイプ・システムを構築した。スパイラル#1では、人とAI群が協働するチームは全て人のチームに勝率7%と全く歯が立たなかったが、タスク仕様をはじめとする試作を繰り返し改善し、スパイラル#2で勝率33%、スパイラル#3で勝率83%に達し、人が同一の身体性を持つAI群と協働して、“人間の集団程度”の戦略レベル発揮できることを実証した。また、フィールド評価においてAI群との協働を評価したチームA参加者から取得したレーティングは、最終段階で全ての指標でAcceptable以上の評価を得た。

さらに、スパイラル#2において人間と強化学習エージェントの協調に関する目途付けを行って、捕獲タスクにおける連携捕獲機能に深層強化学習によって得た強化学習モデルを適用し、スパイラル#3にてフィールド評価結果との比較を行った。その結果、人間が設計して作りこんだものと遜色ないモデルを強化学習によって発見的に獲得可能であることを確認したほか、可読性のないブラックボックス方策を持つ強化学習モデルであっても、適切にタスクと役割を定義し、AIがこれから行おうとするタスクあるいは役割を宣言することで人と協働が可能であることを示した。

スパイラル#1、#2の結果に基づいて行った改善点は、全てが終わった後に振り返ってみれば特段に画期的な内容ではなく自然な対応に思えるが、研究開始時点で演繹的にこれらの改善内容を導くことはできなかった。本研究のように事前に要求機能・性能を定量的に示すことが困難な内容の場合、プロトタイプと評価を繰り返し実施していくことが有効であることが裏付けられたといえる。特に、多数の人によって実際に生起した連携行動の実験データを取得し、人とAI群が協働するチームのタスク分解と役割定義及びAIの行動に反映した流れは、多くの応用先に対して参考になるものと考えられる。フィールド評価は人によるエージェントの操作者を同時に多数参加させる必要があり、要員確保の準備や事後のデータ集約等の運営が運営の負荷が高かったが、本研究において特色ある最も貴重な試みとなったと考える。

今後は、本研究によって得た人と協働することを前提としたAIエージェントの設計にあたっての知見、特にBehavior Treeなどの可読性の高いAI手法と契約ネット・プロトコルなどのマルチエージェントシステムにおける合意形成手法を組み合わせたアーキテクチャと、タスク分解と役割の定義、人と協働可能な強化学習モデルの構成方法を活用した実用化に取り組む。当社は、プロフェッショナル集団が行う様々なミッションをサポートするシステムを広範に提供しており、特に人との関りが不可欠な製品における省力化、省人化に取り組んでいく。

7. 研究発表、知的財産権等の状況

(1) 研究発表等の状況

種別	件数
学術論文	2
学会発表	1
展示・講演	該当なし
雑誌・図書	該当なし
プレス	該当なし
その他	該当なし

(2) 知的財産権等の状況

発明の名称	発明者	出願登録区分	出願番号 (出願日)	出願区分	出願国	登録番号 (登録日)
無人機協調システム、無人機協調処理方法及びプログラム	松波 夏樹、 唐鎌 聡太郎	出願	特願2019-158885号 (2019年8月30日)	国内	日本	-
学習装置、学習方法及び学習プログラム	唐鎌 聡太郎、 松波 夏樹	出願	特願2021-083430号 (2021年5月17日)	国内	日本	-

(3) その他特記事項

該当なし