

科学研究費助成事業 研究成果報告書

令和元年6月10日現在

機関番号：25403

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K00309

研究課題名(和文) マルチエージェント強化学習による実航路の改善

研究課題名(英文) Multi-Agent Reinforcement Learning System to Improve Actual Ships' Courses

研究代表者

神尾 武司 (Kamio, Takeshi)

広島市立大学・情報科学研究科・講師

研究者番号：20316136

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：本研究では、多船航路探索用マルチエージェント強化学習システム(MARLS)の海上交通アセスメントツールとしての価値を高めるために、参照航路に対して「不要箇所の検出」および「衝突発生(ニアミス)の原因究明」を可能とする2種類の多船航路探索用MARLSを新たに開発した。前者では、航路改善率に基づいて不要箇所を検出し、これに代わる改善航路を探索する方法を採用した。後者では、ニアミス回避開始限界点をニアミスの原因として推定し、それ以降の改善航路を探索する方法を採用した。計算機実験では、提案MARLSの性能を用途に応じて検証した。さらに、提案MARLSが実航路に対しても適用可能であることを確認した。

研究成果の学術的意義や社会的意義

本研究では多船航路探索用MARLSの海上交通アセスメントツールとしての価値を高めるために、実航路に対して「不要箇所の検出」および「衝突発生(ニアミス)の原因究明」が可能な多船航路探索用MARLSを開発した。これにより、船舶工学以外の分野にも高度な海上交通アセスメントツールの重要性が啓蒙できる。この点に大きな学術的意義がある。

日本において海上物流の安全性と効率性の確保は産業界を支える必須条件である。さらに、港則法及び海上交通安全法の一部を改正する法律の施行(22年7月1日)に見られるように海上交通アセスメントツールの重要性は増している。本研究はこのような社会的要請に合致する。

研究成果の概要(英文)：We developed multi-agent reinforcement learning system (MARLS) to search ships' courses. In this study, we have proposed two new kinds of MARLSs to increase the value of MARLS as a marine traffic assessment tool. The first MARLS can detect unnecessary parts of reference courses by the course improvement rate and improve the courses. The second MARLS can detect near-misses in reference courses, predict the limited positions to avoid the near-misses as the causes of the near-misses, and improve the courses. Finally, we have evaluated the performance of our proposed MARLS by numerical experiments and also have confirmed that our proposed MARLS can be used for the actual courses.

研究分野：知能情報学

キーワード：マルチエージェントシステム 強化学習 多船航路探索 海上安全 行動選択制限 参照航路 航路改善 ニアミス回避

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

研究代表者はマルチエージェント強化学習システム(Multi-Agent Reinforcement Learning System: MARLS)による多船航路探索に関する研究に取り組んできた。本研究課題申請時の多船航路探索用 MARLS は、航法を先験的知識として導入することで現実世界でも適用可能な航路を獲得できるという特徴を有していたため、実問題に対する強化学習の成功事例として評価を受けてきた。このような背景のもと、MARLS の性能改善を目的として学習効率や航路効率性の改善については研究を数多く実施してきたが、有益な利用方法については本研究課題申請時まであまり議論してこなかった。その結果、多船航路探索用 MARLS に対してツールとしてのアピール不足が指摘されてきた。

そこで本研究課題では、多船航路探索用 MARLS のツールとしてのアピール不足の解消を動機として、実航路の改善を想定した多船航路探索用 MARLS の開発を目指すこととした。

2. 研究の目的

本研究のベースとなる多船航路探索用 MARLS は各船舶が選択した行動系列だけでなく、衝突状況判定およびその解除がどのようなタイミングで行われたかを同時に記録できるため、海上交通アセスメントツールとしての基本性能を有している。しかしながら、基本 MARLS は提示された各船舶のスタートとゴールを結ぶ最適な航路を探索することを目的としているため、ツールとしての利用範囲が限定されている。

そこで本研究では、海上交通アセスメントツールとして価値の高い応用と考えられる「無衝突航路における不要箇所を検出」および「衝突発生(ニアミス)の原因究明」を実行可能な『実航路の改善を想定した多船航路探索用 MARLS の開発』を最終的な目的とする。

3. 研究の方法

本研究は、研究代表者が既に提案していた多船航路探索用 MARLS(以後、基本 MARLS と呼ぶ)をベースとして行われた。以下では、基本 MARLS について、システム概要、処理フロー、さらに最大の特徴である行動選択制限を述べた後、本研究の研究方法について述べる。

(1) 基本 MARLS

① システム概要

基本 MARLS は航法遵守という条件下で各船舶のスタートとゴールを結ぶ最適な航路を探索すること(つまり、多船航路決定)を目的としたシステムであり、代表的な強化学習法として知られる Q 学習を航路探索エンジンとして採用している。

一般に、強化学習を用いて多船航路決定問題を解く場合、船舶群は分散学習環境下のマルチエージェントとしてモデル化されるため、極めて難しい問題となる。そこで研究代表者は、航法を先験的知識として基本 MARLS に取り入れた。具体的には、他船との衝突状況を航法と照らし合わせて観測し、その結果に基づいて各エージェントの行動選択を制限する。これにより、状態空間の爆発問題や同時学習問題といったマルチエージェント学習における深刻な問題の発生を抑制すると同時に、現実世界でも適用可能な航路を獲得できる枠組みを構築している。

② 処理フロー

現時点において、航法に基づく衝突状況判定を受けていないエージェントは、自船 k の周囲に存在する他船 j を検知し、両者の位置関係を衝突状況判定モデルと比較することで現在の衝突状況を判定する。一方、既に衝突状況判定を受けているエージェントは、他船 j との衝突回避が確定した場合、現在の衝突状況判定が解除される。さらに、各エージェントは衝突状況判定の結果から航法または目的地指向性に基づく行動選択制限を受けながら、強化学習による航路探索を行う。これら一連の処理を通じて、航法を遵守する最適な航路の探索が可能となる。

③ 行動選択制限

行動選択制限(Limited Action Selection: LAS)は多船航路探索用 MARLS における最大の特徴である。以下では、基本 MARLS に組み込まれた航法に基づく行動選択制限 LAS_{NR} および目的地指向性に基づく行動選択制限 LAS_{GO} について概説する。

まず、 LAS_{NR} について説明する。航法が定める典型的な衝突状況の内、行会いと横切りにおいて、自船 k が他船 j を回避する義務がある場合、自船 k は右に舵を切って他船 j を回避する必要がある。 LAS_{NR} はこの自船 k に対して右に舵を切る(つまり、舵角 δ_k を $\delta_k \geq 0$ とする)ように行動を制限する。ただし、必要以上の右回頭を避けるため、Fig.1 に示す方法をとる。例えば、行会いの場合(Fig.1(a)), 各船が船首角と速力を変えずに直進すると仮定して、自船 k が衝突回避達成ラインに到達したときの船舶間の位置関係を予測する。このとき、自船 k の x 座標が $x_k \leq -D_{LAS}$ であれば、十分な回避であると判断し、 LAS_{NR} を一時的に解除する。同様に、横切りでは Fig.1(b)に従う。

次に、 LAS_{GO} について説明する。 LAS_{GO} は船首角とゴール方向の差が大きく、衝突の危険がない船舶は船首をゴールに向けるべきであるという考えに基づくものであり、他船回避からの針路回復という意味でも一般的なルールである。例えば、Fig.2(a)のように自船 k の針路が左に逸れた場合($\theta < A_{Gk} < \pi$)、目的地指向性に基づいて右に舵を切る(つまり、舵角 δ_k を $\delta_k \geq 0$ とする)ように行動を制限する。同様に、針路が右に逸れた場合($-\pi < A_{Gk} < -\theta$)、Fig.2(b)に従う。

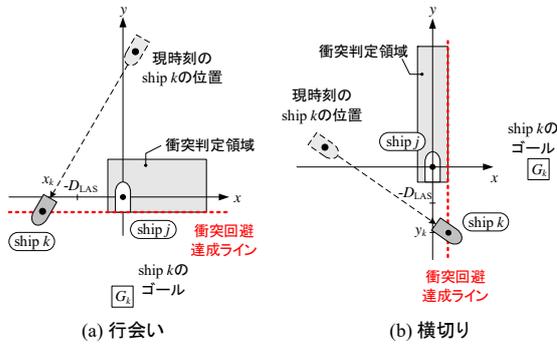


Fig.1 航法に基づく行動選択制限

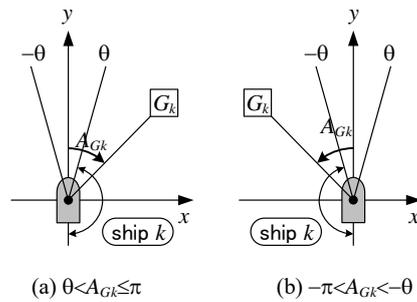


Fig.2 目的地指向性に基づく行動選択制限

(2) 研究方法

本研究の目的は、「無衝突航路における不要箇所を検出」および「衝突発生(ニアミス)の原因究明」を実行可能な『実航路の改善を想定した多船航路探索用 MARLS の開発』である。これを達成するためには、実航路を参照航路として MARLS に与えることを前提とし、参照航路における適切な部分の検出方法とそれ以外の部分の改善方法を構築する必要がある。そこで本研究では、用途に応じた 2 種類の MARLS の開発を以下の方針に基づいて実施する。

第 1 の MARLS は『無衝突航路における不要箇所の検出と航路改善』を用途とする。一般に、無衝突航路における不要箇所は航路効率が低い部分と考えられるため、部分航路に対する航路効率の評価式を定義し、これに基づく不要箇所の検出について検討する。また、不要箇所に対する航路改善については学習効率にも配慮した方法を検討する。

第 2 の MARLS は『ニアミス航路におけるニアミスの原因究明と航路改善』を用途とする。この MARLS では、他船回避開始の遅延によってニアミスが発生すると仮定し、ニアミスの発生原因として回避開始限界点を推定する方法について検討する。また、推定された回避開始限界点以降の参照航路を不要箇所として、学習効率にも配慮した航路改善方法を検討する。

4. 研究成果

(1) 無衝突航路における不要箇所の検出と航路改善

① 手法の概要

参照航路が無衝突航路である場合、その不要箇所を検出し、さらに航路を改善することを目的として開発した多船航路探索用 MARLS の処理は以下の 2 つに大別される。

- I. 参照航路における改善開始点および暫定航路の検出
- II. 暫定航路に基づく航路改善

第 1 処理において、船舶 k の参照航路に対する改善開始点と暫定航路を検出しようとする場合、それ以外の船舶は参照航路に完全追従する。つまり、船舶 k のみを学習エージェントとする Single Agent (SA) モードの MARLS を使用する。

具体的な処理は以下の通りである。船舶 k の参照航路をゴール G からスタート S まで移動するチェックポイントを考える。Fig.3 に示すように、チェックポイントが位置 P_C に移動した直後、強化学習の学習結果(つまり、 Q 値)を初期化する。その上で、船舶 k は S から P_C まで参照航路を完全追従し、 P_C から G までの航路を SA モードの MARLS で短期間探索する。さらに、探索航路に対する航路改善率 η を次式に従って算出する。

$$\eta = (L_{RC} - L_{SC}) / L_{RC} \quad (1)$$

ただし、 L_{RC} と L_{SC} はそれぞれ参照航路と探索航路における P_C から G までの航路長である。

上記の処理において、最大の航路改善率 η_{max} が基準値を超える場合、 η_{max} に対応する P_C が改善開始点 P_{SPI} として検出され、探索航路は暫定航路として保存される。これと同時に、 P_{SPI} から G までの参照航路は不要箇所として検出可能となる。

一般的に、第 1 処理では複数の船舶で暫定航路が検出される。しかし、これらは SA モードの MARLS で検出されるため、暫定航路間で干渉が発生する可能性がある。この問題を解消するために、第 2 処理では Multi Agent (MA) モードの MARLS を使用する。

第 2 処理は以下の通りである。処理の開始に先立ち、 Q 値を初期化する。さらに、改善開始点 P_{SPI} が検出されなかった船舶は S から G までの参照航路を完全追従し、検出された船舶も S から P_{SPI} までは参照航路を完全追従する。つまり、 P_{SPI} 以降が航路改善の対象となる。ここで、暫定航路を利用して効率的に改善航路を探索するために、暫定航路を G から P_{SPI} まで移動するエンドポイントを考える。ただし、エンドポイントは全船舶に共通の時刻によって指定される。

Fig.4 に示すように、エンドポイントが位置 P_E に移動したとき、 P_{SPI} から P_E までは暫定航路を完全追従し、 P_E から G までの航路を MA モードの MARLS で短期間探索する。 P_E からの探索が終了すれば、エンドポイントをさらに P_{SPI} に接近させて同様の短期間探索を繰り返す。全てのエンドポイントが P_{SPI} に到達した後は、MARLS の学習終了条件を満足するまで航路探索が実行され、その結果として改善航路が得られることになる。

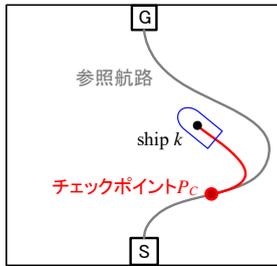


Fig.3 改善開始点の検出

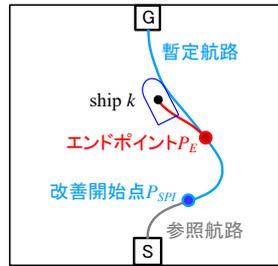
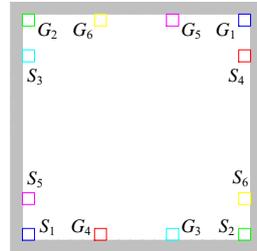


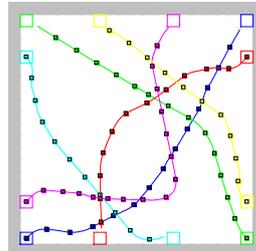
Fig.4 暫定航路に基づく航路改善

Table1 獲得航路長の比較

	L_{ave} [m]	L_{min} [m]	L_{max} [m]
PM	29130	28567	29799
BM	29559	28667	31549



(a) テスト問題



(b) 参照航路

Fig.5 テスト問題と参照航路

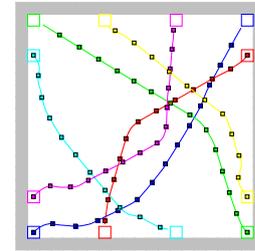


Fig.6 提案 MARLS による獲得航路

② 性能評価

性能評価のために行った計算機実験の結果を示す。テスト問題は $42L_S \times 42L_S$ (L_S は 107m) の海域に 6 船舶が存在する多船航路決定問題である (Fig.5(a) 参照)。議論の簡単化のため、船舶に関するパラメータはスタート S およびゴール G 以外は全て共通とした。さらに、参照航路として基本 MARLS で獲得された航路を使用した (Fig.5(b) 参照)。また、参照航路の総航路長は 30041m であり、航路上のマーカーは 60s 毎の船位を表す。

30 回の試行実験を通じて得られた結果を Table1 と Fig.6 に示す。Table1 は基本 MARLS (BM) と提案 MARLS (PM) による獲得航路の総航路長に関する比較であり、 L_{ave} , L_{min} , L_{max} はそれぞれ平均, 最小, 最大を表す。一方, Fig.6 は提案 MARLS で得られた典型的な航路である。

Table1 より, 提案 MARLS は常に参照航路長 30041m を下回る航路を獲得できることから, 参照航路を安定して改善できることが確かめられた。さらに, 基本 MARLS にみられる獲得航路長のばらつきも提案 MARLS では低減されていることが確認された。さらに, 提案 MARLS の獲得航路 (Fig.6) と参照航路 (Fig.5(b)) を比較すれば, 参照航路の修正箇所が一目瞭然であり, 現実の操船者に対して航路改善のための提案が可能となることが期待できる。

(2) ニアミス航路におけるニアミスの原因究明と航路改善

① 手法の概要

本研究では, ニアミスが他船回避開始の遅延によって発生すると仮定し, ニアミスの原因究明はニアミス回避開始限界点の推定と等価であると考え。この考えのもと, ニアミスを含む参照航路に対して, ニアミスの回避に必要な回避開始限界点を推定し, さらに航路を改善することを目的として開発した多船航路探索用 MARLS の処理は以下の 3 つに大別される。

- I. 参照航路におけるニアミス発生点の検出
- II. 参照航路におけるニアミス回避開始限界点の推定および暫定航路の検出
- III. 暫定航路に基づく航路改善

第 1 処理について述べる。現実の操船者には他船が侵入すると危険を感じる領域が存在する。本研究では, この領域を Fig.7(a) に示すパーソナルエリア (PA) によって近似的に表現する。さらに, 航法によって他船を回避すべき船舶が指定されることから, 船舶 k が船舶 j を避ける必要がある状況で, 船舶 k が船舶 j の PA に侵入した場合, 船舶 k によってニアミスが発生したと判断し, PA への侵入点をニアミス発生点 P_N として検出する (Fig.7(b) 参照)。

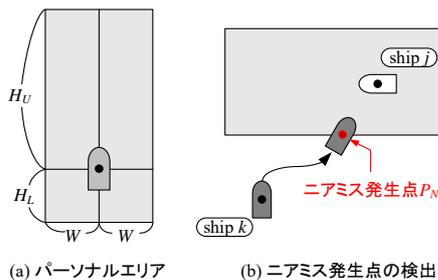
続いて, 第 2 処理について述べる。船舶 k に対してニアミス発生点 P_N が検出された場合, P_N 以前の参照航路にニアミスの回避に必要な回避開始限界点が存在するはずである。本手法では, 船舶 k のみを学習エージェントとし, それ以外の船舶は参照航路に完全追従する SA モードの MARLS を使用して, 参照航路を P_N から遡ることにより回避開始限界点を推定する。

具体的な処理は以下の通りである。船舶 k の参照航路を P_N からスタート S まで移動するチェックポイントを考える。Fig.8 に示すように, チェックポイントが位置 P_c に移動した直後, 強化学習の学習結果 (つまり, Q 値) を初期化する。その上で, S から P_c までは参照航路を完全追従し, P_c からゴール G までの航路を SA モードの MARLS で短期間探索する。これと同時に, 最新 N 回の航路探索エピソードに対するパーソナルエリア侵入率 η を算出する。侵入率 η に基づき, P_c におけるニアミス回避の可能性が十分と判断されれば, P_c をニアミス回避開始限界点の推定値 (以後, ニアミス回避開始点と呼ぶ) P_P として検出し, P_c において発見された最も効率的な航路を暫定航路として保存する。一方, P_c におけるニアミス回避の可能性が不十分であれば,

チェックポイントを S に向けて移動させ、上記の処理を繰り返す。

第 1, 2 処理を通じて、複数の船舶で暫定航路が検出される可能性がある。しかし、これらは SA モードで検出されるため、暫定航路間で干渉が発生する可能性がある。この問題を解消するために、第 3 処理では Multi Agent (MA) モードの MARLS を使用する。

第 3 処理は以下の通りである。処理の開始に先立ち、Q 値を初期化する。さらに、ニアミス回避開始点 P_p が検出されなかった船舶は S から G までの参照航路を完全追従し、検出された船舶も S から P_p までは参照航路を完全追従する。つまり、 P_p 以降が航路改善の対象となる。本手法では、効率的に改善航路を探索するために、暫定航路に追従しながら航路探索を行うことを可能とした行動選択制限 LAS_{TC} を利用する。Fig.9 に示すように、 LAS_{TC} では時刻 t に位置 $O(t)$ にある船舶に対して、暫定航路の位置 $P(t+T)$ を一時的なゴールとして提示し、 LAS_{GO} に倣って舵角を制限する。さらに、速力が暫定航路の位置 $P(t)$ におけるスピードと一致するように速力増分を調整する。ただし、 $O(t)$ と $P(t)$ の距離が基準値を超えた場合、 LAS_{TC} は解除される。さらに、 LAS_{TC} よりも LAS_{NR} と LAS_{GO} が優先される。 LAS_{TC} を導入した MARLS による航路探索は学習終了条件を満足するまで実行され、その結果として改善航路が得られることになる。



(a) パーソナルエリア (b) ニアミス発生点の検出

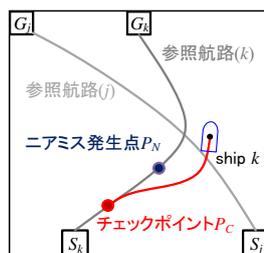


Fig.8 回避開始限界点の推定

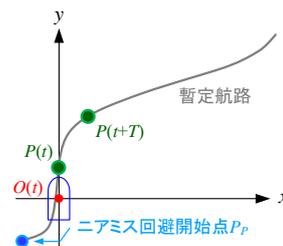


Fig.9 暫定航路追従用行動選択制限(LAS_{TC})

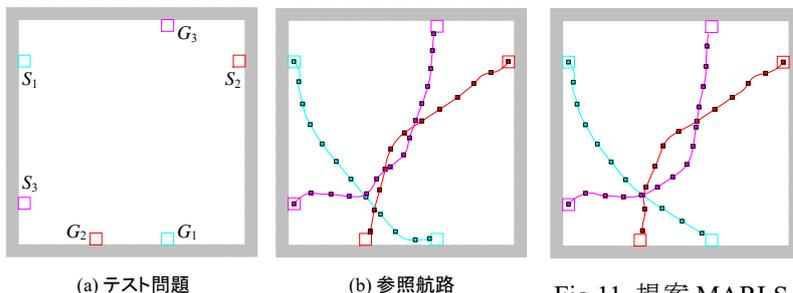
Fig.7 パーソナルエリアと利用法

② 性能評価

性能評価のために行った計算機実験の結果を示す。テスト問題は $42L_S \times 42L_S$ (L_S は 107m) の海域に 3 船舶が存在する多船航路決定問題である (Fig.10(a) 参照)。議論の簡単化のため、船舶に関するパラメータはスタート S およびゴール G 以外は全て共通とした。さらに、参照航路として基本 MARLS で獲得された航路を使用した (Fig.10(b) 参照)。また、参照航路の総航路長は 13209m であり、航路上のマーカーは 60s 毎の船位を表す。

30 回の試行実験を通じて得られた結果を Table2 と Fig.11 に示す。Table2 は第 3 船舶に対して提案 MARLS (PM) が検出したニアミス回避開始点を参照航路の時刻で示したものであり、 T_{ave} , T_{min} , T_{max} はそれぞれ平均、最小、最大を表す。また、PM の 3 倍以上の計算コストを費やして検出したニアミス回避開始点の時刻は常に $T_{true}=240s$ であったため、これを真値とみなす。一方、Fig.11 は PM で得られた典型的な航路である。

Table2 より、計算コストを抑えた PM であっても、真値 T_{true} に近い時刻が検出されることが確認された。さらに、PM の獲得航路 (Fig.11) と参照航路 (Fig.10(b)) を比較すれば、参照航路の修正箇所が一目瞭然であり、現実の操船者に対して回避開始点をニアミスの原因として提示できるだけでなく、航路改善のための提案が可能となることが期待できる。



(a) テスト問題 (b) 参照航路

Fig.10 テスト問題と参照航路

Fig.11 提案 MARLS による獲得航路

Table2 提案 MARLS が検出した回避開始点の時刻

T_{ave} [s]	T_{min} [s]	T_{max} [s]
213	208	240

(3) 実航路への適用

東京湾近郊に設置した AIS (自動船舶識別装置) で得られた実航路データを提案 MARLS に与えた結果、その用途に合わせた改善航路が獲得されることを確認した。その一例として、Fig.12 にニアミスを含む実航路と提案 MARLS による獲得航路を示す。ニアミスを適切に回避する航路が得られていることが確認できる。

(4) 今後の展望

本研究では、実航路を参照航路として与えることを前提とし、「無衝突航路における不要箇所の検出」および「衝突発生 (ニアミス) の原因究明」を実行可能な『多船航路探索用 MARLS による実航路改善法』の枠組みを完成させた。これにより、多船航路探索用 MARLS の海上交通

アセスメントツールとしての価値が一層高まると期待される。しかしながら、実航路の規模や複雑さによっては、その用途に関わらず、必要な計算コストが非常に大きくなる可能性がある。

今後は、多船航路探索用 MARLS の計算コストの低減を図るための研究を進める予定である。

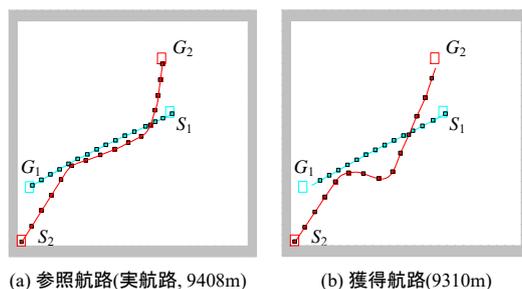


Fig.12 実航路(ニアミス)と獲得航路

5. 主な発表論文等

〔国際会議論文〕(計 2 件)

- ① Takahiro TOMIHARA, Takeshi KAMIO, Takahiro TANAKA, Kunihiko MITSUBORI, and Hisato FUJISAKA, “Enhancement of Reference Course Tracking Capability in Multi-Agent Reinforcement Learning System to Search Ships’ Courses,” Proc. of International Symposium on Nonlinear Theory and Its Applications (NOLTA2018), 査読有, pp.444-447, 2018.
- ② Takeshi KAMIO, Takahiro TANAKA, Kunihiko MITSUBORI, and Hisato FUJISAKA, “Influence of Reference Courses on Reinforcement Learning to Search Ships’ Courses,” Proc. of International Symposium on Nonlinear Theory and Its Applications (NOLTA2017), 査読有, pp.490-493, 2017.

〔研究会技術報告〕(計 3 件)

- ① 富原崇寛, 神尾武司, 田中隆博, 三堀邦彦, 藤坂尚登, “強化学習ベース多船航路探索法によるニアミス航路の改善,” 信学技報, 査読無, NLP2018-127, pp.17-22, 2019.
- ② 神尾武司, 田中隆博, 三堀邦彦, 藤坂尚登, “強化学習ベース多船航路探索法に参照航路の更新が与える影響,” 信学技報, 査読無, NLP2017-63, pp.75-80, 2017.
- ③ 神尾武司, 保井俊祐, 田中隆博, 三堀邦彦, 藤坂尚登, “参照航路が強化学習ベース多船航路探索法に与える影響,” 信学技報, 査読無, NLP2016-108, pp.13-18, 2017.

〔大会論文/予稿〕(計 2 件)

- ① 富原崇寛, 節家 将, 神尾武司, 田中隆博, 三堀邦彦, 藤坂尚登, “強化学習ベース多船航路探索における Q 学習と Profit Sharing の比較,” 電子情報通信学会 2019 年総合大会講演論文集, 査読無, p.101, 2019.
- ② 保井俊祐, 三堀邦彦, 田中隆博, 藤坂尚登, 神尾武司, “舵角と速力を行動にもつ強化学習ベース多船航路探索法,” 電子情報通信学会 2017 年総合大会, 情報・システムソサイエティ特別企画学生ポスターセッション予稿集, 査読無, ISS-P-78, p.78, 2017.

6. 研究組織

(1)研究協力者

研究協力者氏名：田中 隆博

ローマ字氏名：(TANAKA, takahiro)

研究協力者氏名：三堀 邦彦

ローマ字氏名：(MITSUBORI, kunihiko)

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。