

目標航路を行動とする強化学習ベース多船航路探索法における探索能力の向上に関する研究

寺口 凌 神尾 武司 藤坂 尚登

広島市立大学大学院 情報科学研究科 システム工学専攻

1. はじめに

船舶運航において安全性と効率性を勘案した航路を事前に選定することは依然として重要である。我々は『多船航路探索用マルチエージェント強化学習システム (Multi-Agent Reinforcement Learning System: MARLS)』の開発に取り組んできた[1]。そして最近、目標航路を行動とする多船航路探索用 MARLS (TC-MARLS)を用いて、衝突回避タイミングを明確化する方法を開発した[2]。これにより、操船者に有用な知識を提供することが期待されるが、学習初期の航路探索能力の低さが問題となっている。そこで本研究では、基本的な探索強化の方法に対する TC-MARLS の性能を検証する。

2. 目標航路を行動とする多船航路探索用 MARLS (MARLS using Target Courses:TC-MARLS)

TC-MARLS[2]は強化学習の1つである Q 学習をベースとしており、基本的には航法[3]および目的地指向性を満足するように設計されている。目標航路とは、Fig.1 のように『回避、回避量の抑制、針路回復』といった役割ごとに細分化して定義される。TC-MARLS における船舶は、状況に応じて現在の進路の維持または変更を選択する。維持ならば、現在の目標航路をそのまま追従する。変更ならば、維持以外の目標航路を選択し、追従する。TC-MARLS はそれ以前の MARLS[1]と比べ、学習コストが格段に削減されたが、学習初期段階から各船舶が容易にゴールに到達できるため、探索が十分に行われない。そこで本研究では、探索を強化させるための基本的な 2 つの方法に対して、TC-MARLS の性能を検証する。

3. 探索強化のための基本的な方法

3.1 目標航路の追加による探索の強化

Fig.1 に示されるように、目標航路は維持を表す航路 A と役割を表す航路の集合として定義される。本研究では回避用航路 (Fig.1(a) の B~D) と回避量の抑制航路 (Fig.1(b) の B~D) を 3 種類から 6 種類に増やす。つまり、目標航路のバリエーションを増やすことで航路探索の強化を試みる。

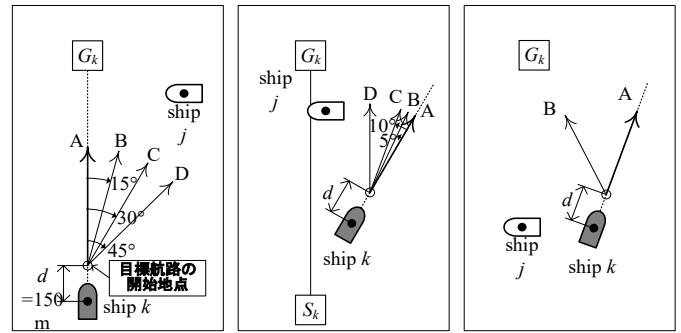
3.2 方策による探索の強化

TC-MARLS は学習初期から正の報酬を得やすい性質から、探索と知識利用のバランスが後者に傾きすぎていると考えられる。そこで本研究では、 ϵ -greedy 方策における ϵ 値を上げることで探索の強化を試みる。本手法では、最初の 5000 エピソードまで ϵ 値を ϵ_0 で固定し、そこからエピソードの増加に応じて ϵ 値を下げていく。そこで ϵ_0 を大きくすることで航路探索を強化する。

4. 計算機実験

テスト問題を $42L_S \times 42L_S$ (船体長 $L_S=107$ [m]) の海域に 6 隻の船舶が存在する多船航路決定問題とする (Fig. 2)。議論の簡単化のため、船舶のスタートおよびゴール以外のパラメータは全エージェントで共通とする。20000 エピソード連続してエピソード成功率が 80% 以上であった場合、学習に成功したと判断する。TC-MARLS[2]、目標航路を追加した TC-MARLS、方策による探索強化を施した TC-MARLS をそれぞれ TC、TC-A、TC-P と

し、表 1 に学習効率性と航路効率性を示す。 N_{SLT} は学習成功試行回数、 N_{EPS} は学習成功時の平均エピソード数、 N_{GET} は航路獲得に成功した試行回数、 L_{ave} は獲得航路の総航路長の平均、 N_s は実際に使用された状態数である。



(a)回避(B~D) (b)回避量の抑制(B~D) (c)針路回復(B)

Fig.1:航法による目標航路の定義

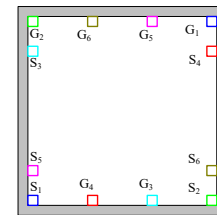


Fig.2:テスト問題 (6 船問題)

表 1:性能比較

	N_{SLT}	N_{EPS}	N_{GET}	L_{ave} [m]	N_s
TC	30	25462	30	26826	848
TC-A	30	25402	30	26814	1270
TC-P	29	24838	29	26731	980

N_s を比較すると、TC よりも TC-A、TC-P の方が大きい。探索範囲が広がったことがわかる。 N_{EPS} から TC と比較して TC-A、TC-P が学習効率を向上させたといえる。 L_{ave} から TC、TC-A、TC-P の順に航路効率が上がっていることがわかる。

5. まとめ

今回の研究により、基本的な探索強化の方法であっても TC-MARLS の性能改善に有効であることが確認された。今後の課題として、TC-A と TC-P の併用により更なる性能改善やより効果的な探索強化方法を提案することである。

参考文献

- [1] T. Kamio, et al.: Effects of Prior Knowledge on Multi-Agent Reinforcement Learning System to Find Courses of Ships, AJIIPS, vol.12, no2, pp.18-23, 2010.
- [2] T. Kamio, et al. "Identification of Avoidance Starting Points by Reinforcement Learning-Based Multi-Ship Course Search Method with Target Courses as Actions," Proc. of NOLTA, pp.589-592, Dec.2022.
- [3] 海上保安庁警備救難部航行安全課(監修), "海上交通三法及び関係法令," 成山堂書店, 東京, 2003.
- [4] Richard.S.Sutton ほか, 三上貞芳・皆川雅章訳 (2000) "強化学習", 森北出版社, pp.3-5.