

# 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

保井俊祐<sup>†</sup> 三堀邦彦<sup>††</sup> 田中隆博<sup>†††</sup> 藤坂尚登<sup>†</sup> 神尾武司<sup>†</sup>  
<sup>†</sup> 広島市立大学情報科学部 <sup>††</sup> 拓殖大学工学部 <sup>†††</sup> 海上保安大学校

## 1. はじめに

船舶では、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極めて重要である。我々は、航路の適切さや操船行動の相互作用を議論するためのツールとして、多船航路探索のためのマルチエージェント強化学習システム(MARLS)[1]の開発に取り組んできた。本稿では、舵角と速力を行動にもつ MARLS を構築し、性能を評価する。

## 2. 多船航路探索用 MARLS

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、多船航路探索は極めて難しい問題となる。我々は、衝突回避ルールである「航法[2]」および衝突の危険がなければ針路を回復すべきであるという観点から「目的地指向性」を行動選択制限として取り入れることで、学習効率の高い MARLS を構築してきた[1]。

航法に従えば、横切り衝突状況(Fig1.(a))では、他船 $j$ を右手に見る船舶 $k$ が針路を右に変更して衝突を避けなければならない。MARLS では、Fig.1(b)に示すように船舶 $j$ の周囲に船舶 $k$ にのみ有効な衝突判定領域(C-area)を配置する。さらに、両船舶が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( $x=L_S/2$ )に到達したときの船舶間の位置関係を予測する。その結果、自船 $k$ の $y$ 座標が $y_k > -D_{LAS}$ を満足する場合、回避行動が不十分であると判断し、舵角 $\delta_k$ を $\delta_k \geq 0$ に制限することで衝突の発生を抑制する[3]。他の衝突状況も同様である。一方、目的地指向性の行動選択制限では、自船 $k$ の船首角がゴール方向から大きく外れている場合、 $\delta_k \geq 0$  または  $\delta_k \leq 0$  に制限される。つまり、我々が提案する MARLS では、行動選択制限によって他船との衝突回避およびゴールへの到達に対する探索の方向性がある程度与えられるため、分散学習環境下であっても高い学習効率が得られることになる。

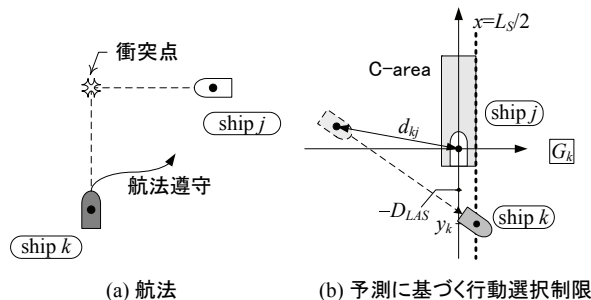


Fig.1 横切りにおける航法と行動選択制限

## 3. 行動表現の拡張

幅狭海域では速力の変更に先行して針路の変更のみで衝突を回避する傾向がある。そのため、従来法では速

力を標準船速 $V_S$ に固定し、舵角 $\delta$ のみを行動としてきた。例えば、舵角 $\delta$ は次式のように定義される。

$$\delta \in \{0, 10, -10, 20, -20\} \text{ (deg.)} \quad (1)$$

一方、東京湾などで得られる AIS(船舶自動識別装置)データを解析すると、速力は必ずしも一定ではない。つまり、MARLS による獲得航路の実用的価値を高めるには、速力も行動として取り入れる必要がある。そこで本稿では、速力を行動として取り扱うために、2つの方法を検討する。

第1の方法では、速力 $V$ を次式で定義する。

$$V \in \{V_S, 0.5V_S, 1.2V_S\} \text{ (knot)} \quad (2)$$

第2の方法では、速力 $V$ を直接行動とするのではなく、加速度 $A$ を行動とすることで、速力の変更を可能とする。ただし、速力 $V$ を状態変数に含める必要がある。

$$A \in \{0, -a, a\} \text{ (knot/s)} \quad (3)$$

## 4. 計算機実験およびまとめ

文献[1]にならい、 $42L_S \times 42L_S$ の海域に同型6船舶(船体長 $L_S$ )が存在する問題を想定した。また、船舶パラメータは海上保安大学校の練習船こじまの値を使用した。従来法をM0、舵角と速力を行動とする方法をM1、舵角と加速度を行動とする方法をM2として、各方法で30回の学習を試行した。ただし、 $V_S=12\text{knot}$ 、 $a=0.01\text{knot/s}$ とした。

Table 1 に学習成功試行回数( $N_{SLT}$ )、学習成功時の平均エピソード数( $N_{EPS}$ )、航路獲得に成功した試行回数( $N_{GET}$ )、獲得航路長の平均( $D_{ave}$ )、最短( $D_{min}$ )、最長( $D_{max}$ )を示す。この結果より、舵角と速力の両方を行動とする場合、学習成功に要するエピソード数の増加が発生するものの、平均獲得航路長については M1 では同程度となり、M2 では改善がみられた。

Table 1 学習成功時における性能比較

|    | $N_{SLT}$ | $N_{EPS}$ | $N_{GET}$ | $D_{ave}(m)$ | $D_{min}(m)$ | $D_{max}(m)$ |
|----|-----------|-----------|-----------|--------------|--------------|--------------|
| M0 | 30        | 47995     | 30        | 28870        | 28175        | 30798        |
| M1 | 30        | 54293     | 27        | 28910        | 28405        | 31146        |
| M2 | 30        | 65787     | 29        | 28548        | 28234        | 28907        |

## 謝辞

本研究は JSPS 科研費 JP16K00309 の助成を受けたものです。

## 参考文献

- [1] T. Kamio, et al.: Effects of Prior Knowledge on Multi-Agent Reinforcement Learning System to Find Courses of Ships, AIIIPS, vol.12, no2, pp18-23, 2010.
- [2] 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).
- [3] 神尾, 他: 強化学習ベース多船航路探索法による衝突回避ルールの評価, 電子情報通信学会 2015 年総合大会 講演論文集, p175, 2015.