Vol.2018-ICS-191 No.7
2018/3/26

合理的でない放流者からため池の生態系を保全するための QS-Learning に基づく見回りルートの学習機構の試作

杉山 慶多[1,a)]　福田 直樹[2]

概要：本論文では，合理的でない放流者からため池の生態系を保全するための，QS-Learning に基づく見回りルートの学習機構の試作について述べる．試作した学習機構では，ユーザがパラメータを設定することにより，ため池の見回りルートを生成することが可能である．

1. はじめに

在来種が多く生息するため池などの淡水域には，多くの絶滅危惧種が生育・生息している．オオクチバスやブルーギルなどの侵略的外来魚により淡水域の環境が悪化している．生物多用性の観点から，これらの外来種から在来種を守る必要があるという議論がある．池干し駆除などの方法により，生態系を再生することは可能であるとされるが，これには多大な労力と費用がかかる．生態系が保たれたままのため池も各地に残されていると考えられ，残された生態系を保全することは一つの可能な戦略であると考えられる．環境省は，「入れない」，「捨てない」，「拡げない」の外来種被害予防三原則 [2] の普及啓発を実施している．残された生態系を保全するためには，ため池に外来種を持ち込ませないことが重要であると考えられる．

ため池に外来種が持ち込まれる要因として，ペットとして飼育していた外来種を家庭で飼いきれなくなり，放流してしまうことが考えられる．これらの放流者の多くは一般人であると考えられ，テロリストなどの犯罪者のようには，十分な合理性持たない場合もあると考えられる．ここでの合理性をもつとは，普段どこでどのようにため池等が警備されているかを事前に知っており，その中で，警備に見つからないようにそのリスクを最小化して行動することを意味する．放流者の多くは複数回放流することはなく，1 度放流するだけであるとすれば，放流者に仮に合理的な判断が可能であったとしても，十分な観測と戦略を持つことができないかもしれない．ゲーム理論的なアプローチであるセキュリティゲームを解くことによるアプローチでは，対象となる「環境への攻撃者」が合理的な行動を行うことを仮定しているため，その行動が合理性に乏しい場合を考えた場合には，必ずしもこのアプローチによる解決は容易ではないと考えられる．

本研究では，このような合理的でない放流者からため池の生態系を保全する問題に対して，我々は，このような条件下で効率良くため池の見回りをするためのルートを生成することを目指す．本論文では，合理的でない放流者からため池の生態系を保全するための，QS-Learning に基づく見回りルートの学習機構の試作について述べる．

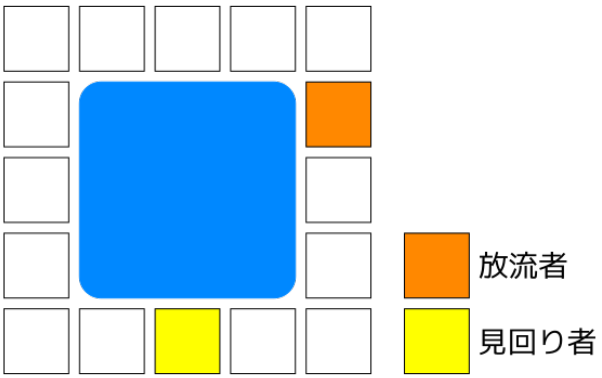

図 1　シミュレータにおける環境の表現の例
Fig. 1　An example representation of the environment on the simulator

2. 環境設定とプレイヤーの行動

図 1 に，本研究であつかうため池と放流者とプレイヤー（見回り者）のモデルを示す．本モデルでは，ため池を囲むマス目上にプレイヤーが存在し，プレイヤーは，隣接するマス目に移動することが可能である．放流者は，外来種を放流することを目的に，あらかじめ指定した確率でマス目

[1] 静岡大学情報学部
Faculty of Informatics, Shizuoka University
[2] 静岡大学学術院情報学領域
School of Informatics, Academic Institute, Shizuoka University
[a)] cs15050@s.inf.shizuoka.ac.jp

© 2018 Information Processing Society of Japan

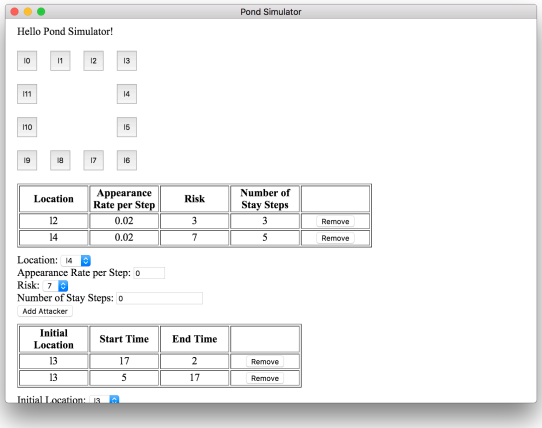

図 2 試作機構のユーザインタフェースの概観
Fig. 2 An overview of the UI of our prototype mechanism

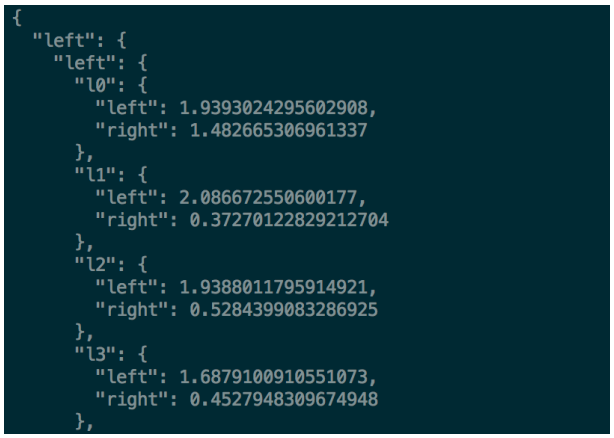

図 3 学習結果 (Q-table) の出力例
Fig. 3 An example output of learned results(Q-table)

に出現し，放流が完了するまで一定ステップ間滞在する．マス目ごとに，放流しやすい場所，しにくい場所，どちらでもない場所があり，これがその場に必要な滞在時間にも影響する．見回り者は，図 1 中では一人としているが，時間帯ごとに人数を増減させることもできる．見回り者が放流者を認識することができる範囲は，現在見回り者がいるマス目から見渡せる場所のみとする．つまり，見回り者は，マス目のどこに放流者が出現しているかを，それを見渡せる場所に行くまでは知ることはできない．見回り者は，各ステップでマス目を移動し，放流者が放流する前に放流者を発見すると，放流者が放流しようとしていた生物の危険度に応じて報酬を得ることができる．放流者が放流してしまった場合，放流者が放流した生物の危険度に応じてマイナスの報酬を得る．マス目には放流者の他に放流者と紛らわしい行動をするものが存在する場合があり，見回り者はそのマス目を調べる行動をすることで，放流者かどうかを判別することが可能である．放流者は，多くの場合放流しやすい場所から放流するが，しばしば，そうでない場所から放流することがある．見回り者がそのような穴場となる場所を見つけることも，一つの課題である．

3. 学習機構の試作

図 2 に，試作した学習機構の概観を示す．試作した学習機構は，Electron 1.8.2[*1] を使用して試作している．本学習機構を使用するユーザは，ここで示したユーザインタフェースなどを用いて，学習のための各種パラメータを設定することが可能である．放流者については，放流者の出現場所と出現確率，放流する生物の危険性を設定することが可能である．同様に，紛らわしい人の出現場所と出現確率を設定することが可能である．見回り者については，見回り者の各時間帯における人数や初期位置を設定することが可能である．この手順に従って必要なパラメータを設定した状態で Start Learning を押下することで，学習機構に基づいて見回りルート案を生成する．

本学習機構において，見回りルートの学習には，QS-Learning [1] を用いる．QS-Learning を用いることにより，状態数が多い場合であっても，通常の Q-Learning よりも学習の収束するに時間がかかる状況を改善できると考えられる．

4. 実行例

図 3 に，学習機構から得られた学習結果の例を示す．この学習結果に基づいて，実際の見回りルートが得られる．図 3 で表示されているのは，これは，見回り者が意思決定に使用する Q-Table の値の一例を JSON 形式で出力したものであり，それぞれ，オブジェクトのルートから 2 つ前の行動，1 つ前の行動，現在位置，次の行動，報酬を表している．

5. おわりに

本論文では，合理的でない放流者からため池の生態系を保全するための，QS-Learning に基づく見回りルートの学習機構の試作について述べた．試作した学習機構では，ユーザがパラメータを設定することにより，ため池の見回りルートを生成することが可能である．今後の課題として，攻撃者の性格などを含めるなど，より現実的な学習機構となるよう，手法を改善する必要がある．また，学習結果を評価する仕組みも必要とされる．

参考文献

[1] Rosenfeld, A., et al.: Leveraging Human Knowledge in Tabular Reinforcement Learning: A Study of Human Subjects, IJCAI (2017) pages 3823-3830.
[2] 環境省：平成 29 年版 環境・循環型社会・生物多様性白書 page 328, 入手先 ⟨http://www.env.go.jp/policy/hakusyo/h29/pdf.html⟩

[*1] https://electronjs.org/

© 2018 Information Processing Society of Japan

(2018.03.01).

© 2018 Information Processing Society of Japan