

# 未知不均一潮流中での航行を考慮した 劣駆動水中ロボットの動作計画と制御

## Motion Planning and Control of Under Actuated Underwater Vehicle in the Unknown and Non-Uniform Sea Flow

川野 洋\*<sup>1</sup>  
Hiroshi KAWANO

\*<sup>1</sup> 日本電信電話株式会社, NTT コミュニケーション科学基礎研究所  
NTT Corporation, NTT Communication Science Laboratories

The development of a practical motion planning and control algorithm for under actuated underwater vehicles has been very problematic for several reasons. First, not only the kinematical characteristics of the motion but also the dynamical characteristics of the underwater vehicle must be considered in the motion planning calculation. Second, it is very difficult to ascertain the exact distribution of the velocity of non-uniform sea flow around obstacles on the seabed before the mission. This paper introduces a motion digitizing method based on an artificial neuron model for MDP based motion planning and a method for making up for the difference between learning and mission environments. The performance of the proposed algorithm is examined by the dynamical simulation of an under-actuated underwater vehicle cruising in the unknown non-uniform sea flow with vortex.

### 1. はじめに

ロボットの制御において、ロボットの動作計画と軌道追従制御を両立させることは重要である。一般に、ロボットは、運動学的に複雑な特性を持っている事が多いので、それを考慮した目標軌道をあらかじめ動作計画によって決定し、決定された目標軌道をロボットの動力学的な特性を考慮した制御アルゴリズムにより追従させることが多い[Craig 1989]。ところが、ロボットの運動学的な特性と動力学的な特性をうまく切り分けることは難しく、動作計画によって決定された軌道が、動力学的には追従が困難であることがしばしば起こる。また、動作計画を行う際に想定されているロボットの動作環境と、実際にロボットが動作する時の動作環境の間に大きな違いがあった場合、動作計画そのものが破綻してしまうという問題もある。これらの問題は、ロボットの中でも特に劣駆動ロボットと呼ばれるものにおいて、重要である。[Kawano 2002]。

劣駆動ロボットとは、ロボットが可能な運動の自由度数よりも、搭載された制御用アクチュエータの自由度数が少ないロボットを指す[Mita 2000]。劣駆動ロボットの例としては、舵と主推進器のみを持った水中ロボット、四輪自動車型移動ロボット、マニピュレータ搭載の人工衛星等が挙げられる。劣駆動ロボットは、アクチュエータの自由度数が低いにも関わらず、その全自由度の位置姿勢が可制御であることから研究対象として注目され、その運動学的特性に対する多くの数学的考察及び分析が行われてきた。これまでの研究によると、劣駆動ロボットの運動学的特性を支配する拘束条件は、多くの場合、積分不可能な複数の方程式によって表現される。そのような拘束条件は、非ホロノミックな拘束と呼ばれ、その拘束条件に従うロボットは、非ホロノミックロボットと呼ばれる。任意の位置姿勢にある非ホロノミックロボットを別の任意の位置姿勢に誘導することは、連続状態フィード

連絡先: 川野 洋, NTT コミュニケーション科学基礎研究所, 神奈川県厚木市森の里宮富 3-1, 電話: 046-240-3185, Fax: 046-240-4716, kawano@avg.brl.ntt.co.jp

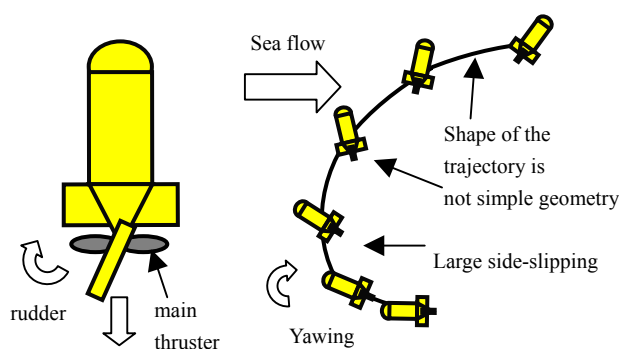


図1 劣駆動水中ロボットとその運動学的特性

バック制御では不可能であることが数学的に証明されており[Nakamura 1993]、その解決のために、非ホロノミックな拘束条件を考慮した動作計画の研究が多数行われてきた。

ところが、これまでの研究で考慮されてきた非ホロノミックな拘束条件が、劣駆動ロボットの運動学的特性を完璧に記述可能なわけではない。例えば、図1に示すような、劣駆動水中ロボットは、4輪自動車型移動ロボットと同様、直接横方向に移動することが出来ず、非ホロノミックな拘束条件に支配されているとされているが、その拘束条件の方程式は、「舵もしくは、方位角制御用車輪の切り角の大きさによって、時間遅れなしにロボットの方角速度が決定される」という前提の元で定義されており、それを出发点として、これまでの動作計画の研究が行われてきた。そのような前提では、最短時間軌道の形状は円弧と直線の組み合わせで構成されることが示されている[Laumond 1994]。しかし、実際の劣駆動水中ロボットでは、流体力学的な理由により、舵の切り角に対する方位角速度の応答が極端に遅く、この前提を適用できない[Kim 2004]。また、実海域には水中ロボットの航行速度に匹敵する速さの潮流が存在し、水中ロボットの軌道形状が大きくゆがめられてしまう。以上の理由により、非ホロノミック

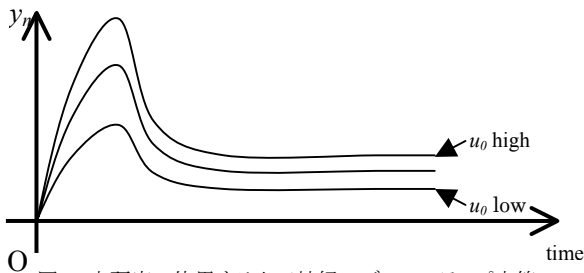


図2 本研究で使用する人工神経モデルのステップ応答

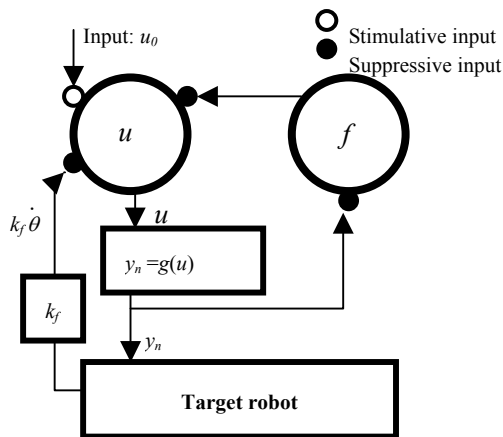


図3 軌道要素実現のための提案制御アルゴリズム

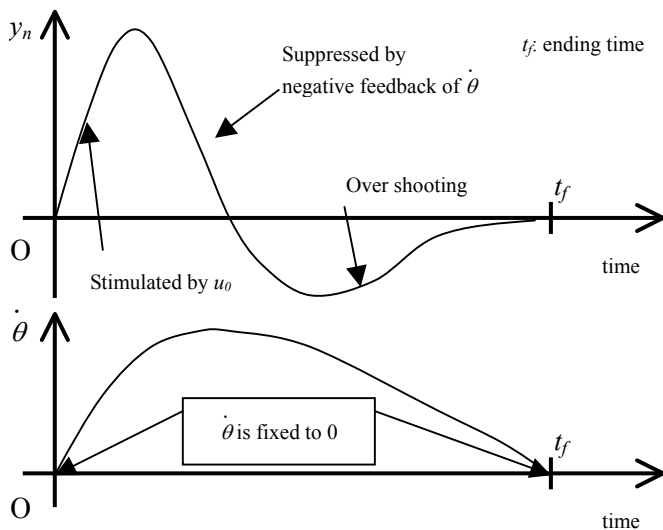


図4 制御対象ロボットへの制御入力と方位角速 $\dot{\phi}$ の時間履歴

な4輪自動車型移動ロボットの軌道計画問題で扱われてきた手法が、実海域で任務行動を行う劣駆動水中ロボットに応用された例は未だない。

こうした現状を踏まえ、筆者らは、劣駆動ロボットの動作計画の研究を、これまでとは異なるアプローチで進めてきた。筆者らは、強化学習を応用し、一定速度の既知な一様流が存在する海中環境における、劣駆動水中ロボットの動作計画及び制御のシミュレーションに成功している[Kawano 2002]。しかし、実際の海中環境は、潮流は一様ではなく、障害物の存在によって不均一になり渦も発生している。さらに、それらの潮流の分布をロボットの任務行動前に知ることは難しい。そこで、本論文は、未知でかつ不均一な潮流が存在する海中環境にて適用可能な、劣駆動ロボットの動作計画および制御を目指し、そのための強化

学習の応用法を提案する。提案手法の根幹をなすのは、人工神経素子モデルの応用による劣駆動水中ロボットの動作分割手法である。提案手法は、劣駆動水中ロボットの運動学的特性と動力学的特性の両方を考慮して、未知な潮流が存在する中でも、離散的マルコフ決定過程に従うようにロボットの動作を分割することが可能である。また、事前の学習によって得られた価値関数を生かして、未知潮流中で適切な航行を可能とする行動選択手法も提案する。本論文では、提案手法を数値シミュレーションによって検証し、その有効性を示す。

## 2. 神経素子モデルと強化学習

### 2.1 神経素子モデル

本研究では、松岡らにより提案された興奮性と疲労性を備えた人工神経素子モデルを利用する[Matsuoka 1987]。本モデルは、動物の脳内にて、周期的な運動のパターンを生成する神経回路の機能を模したものであるとされており、人型ロボットの二足歩行パターン生成にも応用されている[Taga 1995]。定義は以下の通りである。

$$T_r \frac{du}{dt} + u = u_0 - bf, \quad (1)$$

$$y_n = g(u), \quad (2)$$

ここで、 $u$  は、神経素子の興奮度、 $u_0$  は、神経素子への入力パルスの強さ、 $f$  は、神経素子の疲労度、 $y_n$  は、神経素子の発火率である。 $T_r$  は、興奮度を支配する時定数、 $b$  は、疲労度が興奮の抑制に与える影響度を定める定数( $b > 0$ )である。疲労度 $f$ は、以下の方程式に従って変化する。

$$T_a \frac{df}{dt} + f = y_n, \quad (3)$$

$T_a$  は、疲労度を支配する時定数である。本神経素子モデルのステップ応答特性は、図2に示すとおりであり、ハイパスフィルタと同様の応答を示す。入力パルス $u_0$ の値が大きいほど、初期発火率 $y_n$ が高くなる。

### 2.2 強化学習

強化学習は、マルコフ決定過程に従う学習主体が、行動選択を行った結果として環境から受ける報酬値に応じて、最適な行動方策を学習していくものである。本研究では、強化学習の手法として、動的計画法を使用する。動的計画法は、学習主体の行動モデルを、マルコフ状態遷移モデル(MDP モデル)にてモデル化し、獲得されたモデルを利用して最適行動方策を計算する強化学習の基本的な手法である。MDPモデルは、

$$P_{ss'}^a = \Pr \{ s_{t+1} = s' \mid s_t = s, a_t = a \}, \quad (4)$$

$$R_{ss'}^a = E \{ r_{t+1} \mid s_t = s, a_t = a, s_{t+1} = s' \}. \quad (5)$$

のように定義される。ここで、 $P_{ss'}^a$  は、行動 $a$ を選択した結果、行動主体の状態が、離散状態 $s$ から $s'$ に遷移する確率であり、 $t$ は、離散的時間ステップを指す。 $R_{ss'}^a$ は、報酬値 $r$ の期待値を指す。獲得されたMDPモデルを利用し、Bellman方程式を解くことで、状態価値関数 $V^\pi(s)$ を得る。 $V^\pi(s)$ の定義は、

$$V^\pi(s) = E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid s_t = s \right\}, \quad (6)$$

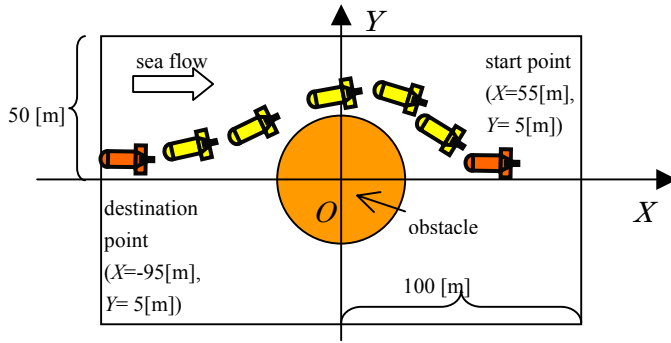


図5 シミュレーションで想定した海中環境

である。各状態  $s$  にて、 $V^{\pi}$  の値は、最適方策に従った場合に受ける報酬の和を表す。本研究では、MDP モデルにおける状態  $s$  を水中ロボットの位置と方位角、行動  $a$  を、次章にて述べる水中ロボットの軌道要素に対応させる。

### 3. 提案手法

#### 3.1 未知外乱を考慮した動作分割手法

強化学習は、離散的状態表現の使用を前提としているので、本来は連続値で表現される水中ロボットの運動をマルコフ決定過程に従うように、離散化表現しなければならない。本研究では、水中ロボットの方位角速度が一定値  $0$  となる点で水中ロボットの軌道を分割する。本研究では、分割された軌道要素をロボットの行動単位として扱う。この分割方法により、水中ロボットの運動を幾何学的に拘束せず、滑らかさを失わないで軌道分割出来る。また、本来は MDP モデルを構成する際に参照すべき変数として、水平面内位置のデカルト座標と、方位角、方位角速度が必要であるが、本手法では、方位角速度を MDP モデルに組み込む必要がない。そのため、MDP モデル内の状態数の削減にも有効である。

以上の手法で分割された軌道要素を、未知の外乱が存在する環境中でも実現するためには、軌道要素の始点と終点においては、方位角速度を  $0$  に安定化し、その他の部分では、旋回運動が行えるよう、方位角速度が  $0$  以外の値をとることが可能な制御アルゴリズムが必要である。本研究では、そのために、2.1 節で述べた人工神経モデルを応用した制御アルゴリズムを提案する。提案アルゴリズムの構成を図 3 に示す。数式での定義は以下の通りである。

$$T_r \frac{du}{dt} + u = u_0 - bf - k_f \dot{\theta} \quad (7)$$

$$y_n = k_y u + y_{const} \quad (8)$$

$$T_a \frac{df}{dt} + f = y_n \quad (9)$$

ここで、 $u_0$  は、軌道要素の始点にて与えられる刺激入力、 $\dot{\theta}$  は、方位角速度、 $k_f$  は、方位角速度  $\dot{\theta}$  の負のフィードバックゲイン ( $k_f > 0$ )、 $y_n$  は、制御対象ロボットへの制御信号であり(ここでは舵の切り角)、 $k_y$  は、比例定数 ( $k_y > 0$ )、 $t_f$  は、軌道要素の終点時刻である。提案アルゴリズムにおいては、神経素子の興奮度の増加は、ロボットからの方位角速度の負のフィードバックにより抑制され、制御信号値が発振を起こさないように各定数が調整される。提案アルゴリズムの動作は以下の流れである。

①軌道要素開始時刻にて、刺激入力  $u_0$  によって興奮を促され

た神経素子は、その興奮度  $u$  に比例した制御信号  $y_n$  を出力する。

②制御対象ロボットは、制御信号を受けて、方位角速度  $\dot{\theta}$  を増加させる。それによって、ロボットは、旋回運動を行う。

③方位角速度  $\dot{\theta}$  の負フィードバックと神経素子内の疲労度  $f$  の上昇により、神経素子内の興奮度  $u$  が下がっていく。それに応じて、制御信号  $y_n$  も下がり、ロボットの旋回速度  $\dot{\theta}$  は  $0$  に安定化されていく。

④軌道要素終点において、方位角速度  $\dot{\theta}$  は  $0$  になる。

提案アルゴリズムにおいては、終点時刻  $t_f$  と、刺激入力  $u_0$  の大きさによって、各軌道要素でのロボットの移動量、方位角変位量を様々な値とすることが出来る。軌道要素の始点時刻と終点時刻での旋回速度を  $0$  にするために、フィードバック制御を行っているので、未知外乱中でも、適切に始点終点での旋回速度を  $0$  に安定化できる。

#### 3.2 未知潮流中での行動選択手法

本研究では、一様流の環境にて、MDP モデルの作成を行う。水中ロボットの運動において、潮流によって影響を受けやすいのは、水中ロボットのデカルト座標位置であることに注目する。本研究では、MDP モデル作成時に、各行動(軌道要素)におけるデカルト座標値と方位角の変位量の平均値  $E\{dX(s, a)\}$ 、 $E\{dY(s, a)\}$ 、 $E\{d\theta(s, a)\}$  を測定し、その値を行動選択時に利用する。MDP 獲得時と異なる速さの潮流が存在した場合の水中ロボットの各行動における変位量  $D_x(s, a)$ 、 $D_y(s, a)$  は、

$$D_x(s, a) = E\{dX(s, a)\} + df_x t_f \quad (10)$$

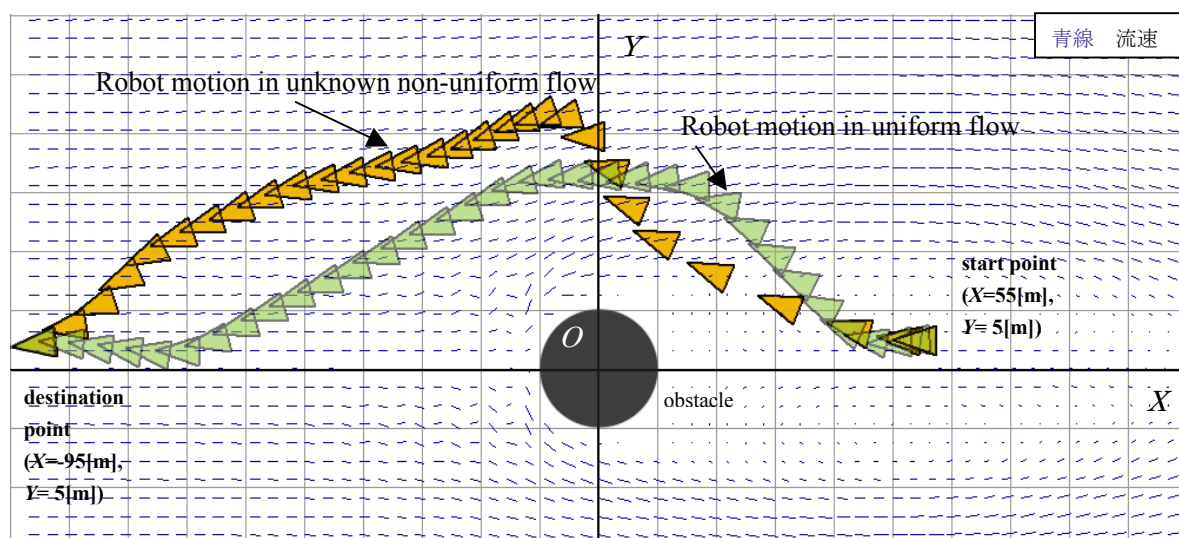
$$D_y(s, a) = E\{dY(s, a)\} + df_y t_f \quad (11)$$

と、記述可能である。ここで、 $df_x$ 、 $df_y$  は、MDP モデル作成時の一様流速度からの、実環境での潮流速度の差分である。 $df_x$ 、 $df_y$  の分布が、水中ロボットの任務環境全体で既知であれば、状態価値関数  $V^{\pi}(s)$  の再計算は容易であるが、そうではないので  $V^{\pi}(s)$  の再計算は行わない。本研究では、行動選択の際に、(10)(11)を利用して、水中ロボットの各行動選択後の状態  $s'$  を推定し、 $V^{\pi}(s')$  を最大化する行動  $a$  を選択する。

### 4. シミュレーション

#### 4.1 海中環境モデル

提案手法の有効性を検証するためにシミュレーションを行った。本シミュレーションで想定した海中環境を、図 5 に示す。水平面内での障害物回避行動を任務行動とし、障害物は、原点  $O$  に直径  $20m$  の円形障害物が存在するものとした。MDP モデルの離散状態  $s$  は、 $X, Y$  軸を  $10m$  単位、方位角  $\theta$  を  $9$  度単位で離散化して構成した。MDP モデルは、海流が障害物の影響を受けていない速さ  $1[m/sec]$  の一様流の状態では  $E\{dX(s, a)\}$ 、 $E\{dY(s, a)\}$ 、 $E\{d\theta(s, a)\}$  を測定し、それを元に作成した。ただし、一様流の条件では、 $E\{dX(s, a)\}$ 、 $E\{dY(s, a)\}$ 、 $E\{d\theta(s, a)\}$  の値は、 $s$  を構成する状態変数のうち、 $\theta$  に支配されるものであり、 $X, Y$  の値には支配されないことを利用して MDP モデルを作成した。実環境モデルとして、レイノルズ数  $100$  の流れをナビエ-ストークス方程式の数値計算によってシミュレートし、その結果得られた定常流分布解を利用した。

図6 シミュレーション結果(レイノルズ数  $Re=100$ )

行動セットとして、 $u_0$  の値を -4.0, -3.2, -2.4, -1.6, -0.8, 0, 0.8, 1.6, 2.4, 3.2, -4.0,  $t_f$  の値を、12.5, 25, 37.5, 50, 62.5 [sec] とする軌道要素を用意した。水中ロボットモデルとして、静水中での旋回半径 60[m] 程度、舵の切り角に対する方位角速度のステップ応答時間は、15 秒程度の運動モデルを作成した [Kim 2003]。航行中推力は一定値 300[N] とした。

#### 4.2 シミュレーション結果

図 6 に、シミュレーション結果を示す。実海中環境では、障害物の後方に渦が発生している。また、障害物両脇の流れは、流速が MDP モデル作成時の 1.5 倍ほどであり、障害物を中心に流れが両側に広がるようになっている。水中ロボットは、自機位置での潮流速度を計測することができるが全体の分布に関する情報は得ることが出来ない。水中ロボットの動作開始位置を、障害物後方の渦中とし、その後の動作を調べた。動作開始直後は、渦の中を航行しているが、流速自体が比較的低いので、旋回半径が大きくなっている ( $X=50\sim 10$ [m])。水中ロボットの行動選択はこのことも考慮して、ちょうど障害物をよけられる行動が選択されている。障害物の脇を通過する際には、速い潮流に流されて、対地前進速度が下がり、一様流条件時のロボット動作と比較して、水中ロボットの軌道が外側に膨らんでいる ( $X=10\sim 10$ [m])。その後も、周りの流れに押されて、水中ロボットの軌道は外側に押されぎみであるが ( $X=-10\sim 40$ [m])、障害物から十分離れたところで、潮流分布が、一葉流に近くなり、これまで、外側に膨らんだ軌道を修正して、目標地点に到達している ( $X=-40\sim 100$ [m])。以上の結果から、提案手法により、渦あり不均一未知潮流が存在する環境での水中ロボットの障害物回避航行が可能になったといえる。

#### 5. 結論

渦が存在する不均一未知潮流分布中で適用可能な、劣駆動水中ロボットの動作計画及び制御手法を提案した。提案手法は、マルコフ決定過程における動的計画法の応用によるものであり、未知外乱が存在する中でも、適切に水中ロボットの運動を分割離散化するための手法と、未知外乱が存在する中での行動選択手法からなる。それらのうち、ロボット動作の分割離散化手法は、興奮性と疲労性を持った人工神経素子モデルを応用したものである。実海中環境のモデルとして、ナビエストークス方程式

の数値計算により、海中障害物周りの渦あり流れを再現し、その中で航行を行う劣駆動水中ロボットのシミュレーションを行った。その結果、海中環境の流れが一様流である条件で動作計画を行った劣駆動水中ロボットに、適切に実環境中で障害物回避航行させることが出来た。今後は、さらに厳しい実海中環境として、高レイノルズ数の非定常渦流れが存在する環境中でも適用可能な、劣駆動ロボットの動作計画および制御手法について研究していく予定である。

#### 参考文献

- [Craig 1989] John J. Craig: *Introduction to Robotics: Mechanics and Control second edition*, Addison-Wesley, 1989.
- [Laumond 1994] J. P. Laumond, P. E. Jacobs, M. Taix and R. M. Murray : A Motion Planner for Nonholonomic Mobile Robots, *IEEE Transactions on Robotics and Automation*, Vol. 10, No. 5, pp.577-593, October, 1994.
- [Mita 2000] 美多 勉: 非線形制御入門-劣駆動ロボットの技能制御入門-, 昭晃堂, 2000年.
- [Nakamura 1993] 中村仁彦: 非ホロノミックロボットシステム 第1回 非ホロノミックなロボットって何?, 日本ロボット学会誌, Vol.11, No.4, pp.41-48, May, 1993.
- [Kawano 2002] H. Kawano and T. Ura: Fast Reinforcement Learning Algorithm for Motion Planning of Non-Holonomic Autonomous Underwater Vehicle in Disturbance, *Proceedings of 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.903-908, October 2002.
- [Kim 2003] K. Kim and T. Ura: Fuel-Optimal Guidance and Tracking Control of AUV under Current Interaction, *Proceedings of ISOPE 2003*, pp.191-196, May, 2003.
- [Matsuoka 1987] K. Matsuoka: Mechanism of Frequency and Pattern Control in the Neural Rhythm Generators, *Biological Cybernetics*, Vol.56, pp345-353, 1987.
- [Taga 1995] G. Taga: A model of the neuro-musculo-skeletal system for human locomotion -II. Real-time adaptability under various constraints-, *Biological Cybernetics*, Vol. 73, pp.113-121, 1995.
- [Sutton 1998] R. S. Sutton and A. G. Barto: *Reinforcement Learning: An Introduction*, MIT Press, 1998.