

強化学習によるマイクロペリスタルティックポンプの動作シーケンスの獲得 Operation sequence acquisition for micro peristaltic pump by reinforcement learning

山梨大 [○]阿部岳晃, 大原伸介, 浮田芳昭

Univ. of Yamanashi, [○]Takaaki Abe, Shinsuke Ohhara, Yoshiaki Ukita

E-mail: yukita@yamanashi.ac.jp

マイクロペリスタルティックポンプは、マイクロ流体デバイス内のマイクロ、ナノリットルスケールの流体ポンピングに用いられている^{1,2}。このポンプは、ダイヤフラムによる流路の開閉によって生じる流体の移動を利用したもので、複数のダイヤフラムを連動させることでポンピングを行う。そのため、ダイヤフラムの動作シーケンスはポンプの流量を決定する最も重要な要素である。しかし、ダイヤフラムの動作シーケンスの決定は設計者の経験に依存している。そこで本研究では、強化学習を用いたペリスタルティックポンプの動作シーケンスの最適化を検討した。

Fig.1 にポンプのモデルを示す。ポンプの流量はマイクロビーズの移動量から測定した。このシステムにおける Markov decision process (MDP) の構成要素を次のように定義した。

状態 s : ダイヤフラムの現在フェーズ

行動 a : ダイヤフラムの次のフェーズ

報酬 $R_{(s,a)}$: 状態 s で行動 a を実行した場合の流量

ここで、ダイヤフラムの状態は確率的に遷移せず、遷移先 $s' = a$ と仮定したため遷移関数は省略した。この仮定をもとに Bellman Equation から状態 s における行動 a の価値 $V_{(s,a)}$ の式(1)を得る。

$$V_{(s,a)} = R_{(s,a)} + \gamma \max_{a'} V_{(s',a')} \quad (1)$$

ここで、 γ は割引率である。状態 s で行動 a を実行した場合の流量をすべての組み合わせでサンプリングし、それを報酬関数 $R_{(s,a)}$ として行動価値を算出した。各状態で価値が最大の行動をとると、(000)→(100)→(011)の3フェーズのシーケンスに収束する結果となった。このシーケンスと、従来の一般的なシーケンス²の流量を測定した結果を Fig.2

に示す。算出されたシーケンスは、従来の3フェーズ及び6フェーズシーケンスに対して流量が80%以上向上した。したがって、ペリスタルティックポンプの動作に対して強化学習のアルゴリズムが適応可能であり、ダイヤフラムの動作シーケンスの決定に有効であることが示された。

1. C. K. Byun et al.,: Electrophoresis,35, p.245, (2014)

2. L.-S. Jang et al.,: Biomed Microdevices, 9, p.185, (2007)

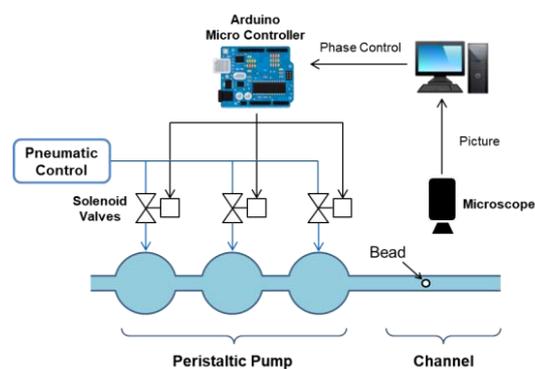


Fig.1 Learning system for pump sequence

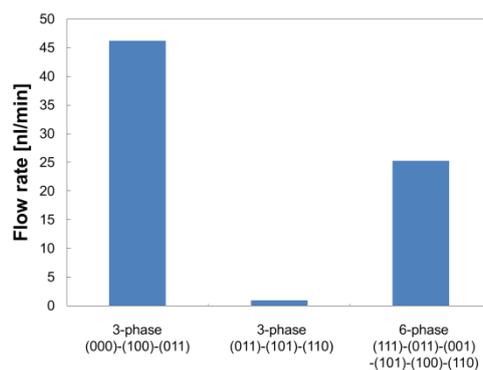


Fig.2 Comparison of flow rate