

Towards Robust Motion Planning for Synthetic Cells in a Circulatory System

Thomas A. Berrueta, Ana Pervan, and Todd D. Murphey

Abstract—The immune system protects the body by recognizing and reacting to foreign invaders, known as antigens. When white blood cells find an antigen, they multiply and communicate among themselves to collectively carry out an immune response. Recent advances in the design of active synthetic cells and active Brownian particles have shown that artificial microscopic agents are cable of rudimentary communication, computation and locomotion in response to stimuli. We consider the problem of designing robust exploration strategies for collectives of microscopic agents in a circulatory system. In such an environment, agents are not guaranteed to have perfect sensing, reliable communications or be capable of complex locomotion. Here, we present preliminary efforts in generating locomotion strategies for ensembles of microscopic agents without onboard computation. By leveraging advances in embodied computation and active matter, along with recent results in decentralized coverage algorithms, we propose a robust algorithm for control of synthetic cell swarms.

I. INTRODUCTION

Advances in active matter have enabled the construction of increasingly capable robotic collectives [1], [2]. Ensembles of agents of sizes on the order of a single human cell have incredible potential for a variety of applications. These include autonomous monitoring of oil and gas conduits [3], eletrophysiological recordings with neural dust motes [4], and minimally invasive clinical procedures and drug delivery [5]. In this work, we investigate the use of synthetic cells to imitate some functionality of white blood cells in the human immune system.

The immune system protects the body by recognizing and responding to antigens, which are harmful agents such as viruses, bacteria, and extraneous toxins [6]. When white blood cells find a target, they multiply and swarm the region of interest in a collective effort to neutralize the antigen [7]. In this work, we seek to enable similar functionality in groups of synthetic cells with limited sensing, computation and actuation capabilities.

Synthetic cells are microscopic devices ($\sim 100\mu m$) that may contain simple circuits for performing rudimentary calculations along with limited sensing and non-volatile memory [8]. In addition to computational capabilities, synthetic cells can include ferromagnetic cores to interact either with each other or with external magnetic fields. Synthetic cells are also capable of simple locomotion, similar to that of multflagellar bacteria, by generating localized bubble jets that propel them in a given direction [9]. Despite their small size and limitations, by exploiting synthetic cells' capabilities we propose that robust complex collective behaviors can be achieved.

Authors are with the Neuroscience and Robotics Laboratory (NxR) at the Department of Mechanical Engineering, Northwestern University, Evanston, IL. Email: {tberrueta, anapervan}@u.northwestern.edu, t-murphey@northwestern.edu.

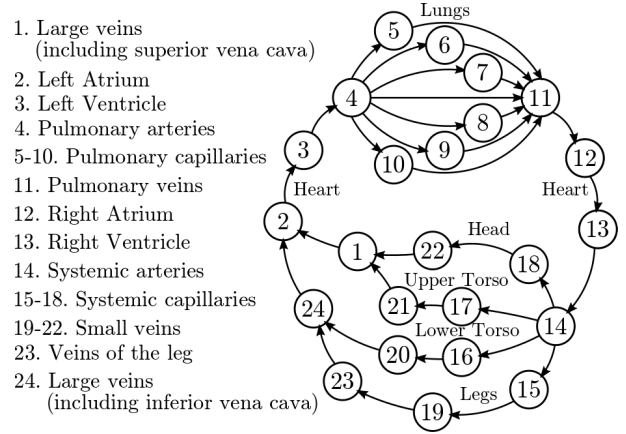


Fig. 1. Simplified model of the human circulatory system designed to preserve its fundamental structural properties [10].

In this manuscript, we formalize the problem statement of antigen localization with a collective of synthetic cells, and briefly describe distributed algorithmic strategies for solving this problem while complying with the limitations endemic to synthetic cells.

II. PROBLEM STATEMENT

The goal for the synthetic cell swarm is to approximately mimic the human body immune response by localizing an antigen in a circulatory system.

Environment: We consider a simplified human circulatory system (see Fig. 1) consisting of 24 discrete states [10]. Each one of these states corresponds to a major structural component of the circulatory system. Naturally, the graph is cyclic and consequently there is no corresponding steady state distribution.

Assumptions: Prior to formally specifying the problem we must state a series of assumptions that we make about the capabilities of synthetic cells. First, we will assume that synthetic cells are capable of sensing which state of the circulatory graph (Fig. 1) they are currently in. This can be done with external signaling using electromagnetic fields, or internally with luminescent or chemical landmarks by stimulating photodiodes or chemiresistors on the synthetic cells [3]. Second, we assume that agents have the ability to leverage their locomotion to choose transitions between states at forks in the graph, albeit nondeterministically. This is physically feasible given the directed locomotion demonstrated in [9]. Third, we assume that each synthetic cell is capable of storing its individual control policy internally through the implementation of logical operators and memristors. Finally, we also make the assumption that synthetic cells are capable of detecting a given antigen and bonding themselves to it with

some positive probability, either through chemical reactions or ferromagnetic bonding.

Formal Specification: We will model this problem as a multi-agent Markov decision process (MDP) over a finite time horizon, H . We specify the multi-agent MDP as a 5-tuple, $\langle C, S, A, T, L \rangle$. The finite set C is the set of all agents considered, corresponding to distinct synthetic cells in the system. The state space S is the space depicted by the directed graph of Fig. 1, and the action space A is the set of all possible transitions between the elements of S . For an agent $c \in C$, at each round $t \in [H]$, we denote its individual states and actions as $s_{c,t}$ and $a_{c,t}$, respectively. Let P_S represent a probability distribution over the state-space, then the transition model T is the following mapping $T : S \times A \rightarrow P_S$ which describes the dynamics of agents on the graph [11]. We will additionally assume that these dynamics are independent between agents and that the scale of circulatory system components are sufficiently large compared to individual synthetic cells, such that there are no interactive effects between agents occupying the same states. The final component of our MDP specification is the loss function L , which will be described in detail in the next section.

III. DISTRIBUTED COVERAGE POLICIES

To solve the antigen localization problem, we propose an approach in which a set of agents generate distributed coverage over the circulatory system based on precomputed control policies requiring no onboard computation other than the ability to execute an action stored in memory. Given that an agent can cover the environment and bind itself to an antigen with a positive probability, localization of the antigen is simply a consequence of coverage. The primary components of the proposed solution require finding a suitable coverage metric, and generating distributed policies that can be encoded in the finite non-volatile memory of synthetic cells.

Promoting Distributed Coverage: Recent work based on insights from ergodic theory has considered the problem of generating coverage deterministically in continuous state spaces [12], [13]. We apply these insights to generate a loss function based on a modified finite state space ergodic coverage metric. The loss L_t that we will consider is the Kullback-Leibler divergence between a goal distribution over the state-space, p , and an empirical time-averaged state distribution q_t up to round t between all agents,

$$L_t = D_{KL}(p||q_t) = \mathbb{E}_p[\ln(p) - \ln(q_t)]. \quad (1)$$

The distribution q_t is comprised of individual agent components such that $q_t = \frac{1}{|C|} \sum_{c \in C} q_{c,t}$, where each subdistribution is given by

$$q_{c,t} = \left\{ \frac{1}{t} \sum_{i=1}^t \mathbb{1}_s(i) \mid \forall s \in S \right\}. \quad (2)$$

Given that we are considering this process over a finite time horizon H , the total loss is then L_H . Using this loss function we are able to specify the task of reconstructing a state space distribution as a function of agent behavior. There is no requirement for the distribution p to be uniformly random, allowing one to simply encode a belief over the antigen

location. Additionally, since this loss function can be fully distributed, the resulting policies can be as well.

Precomputing Coverage Policies: Since synthetic cells are incapable of online computation, the problem demands a strict separation of policy learning and policy execution. To this end, we must learn policies in representations explicitly amenable to encoding in non-volatile memory. Prior work in the generation of low-complexity control policies has shown that it is possible in principle for a system of synthetic cells to achieve a central objective from precomputed policies embedded in memory [14]. While the authors in the cited manuscript construct their policies using concepts from hybrid control theory to find optimal system actions over a given time horizon, we will adapt their insights to a reinforcement learning framework.

Particularly, since our state space is not very large, we can apply a simple ε -greedy Q-learning approach to optimize agent actions while exploring the space of policies [15]. In this context, the Q-function will in fact be a vector-valued function in $\mathbb{R}^{|C|}$ consisting of the individual Q-functions of each agent. The policy used during learning will be the following

$$\pi(s) = \begin{cases} a & \text{with } p(\varepsilon)/|A| \\ \operatorname{argmin}_a Q(s, a) & \text{with } p(1 - \varepsilon), \end{cases} \quad (3)$$

where the ε is a parameter mixing greedy exploitation with random action exploration. Here, we update the Q-function according to the following update law with learning rate α , and discount γ parameters

$$Q(s_t, a) \leftarrow (1 - \alpha)Q(s_t, a) + \alpha [L_t + \gamma \min_{a'} Q(s_{t+1}, a')]. \quad (4)$$

After the Q-function is resolved, each agent's individual policy can be expressed in the following form

$$\pi_c^*(s) = \operatorname{argmin}_a Q_c^*(s, a). \quad (5)$$

However, due to the finite-dimensional nature of the circulatory system model, this can be boiled down to a table of actions to execute in each given state. Thus, as long as synthetic cells have enough memory to encode their respective policy tables, they will be able to generate coverage without any onboard computation by executing their policies in open loop.

IV. PRELIMINARY RESULTS

In an effort to test the proposed algorithm, we implemented the model circulatory system environment and simulated synthetic cells as agents walking on a graph. To test the robustness of our computation-free approach to distributed coverage we consider two types of environment: deterministic and non-deterministic. Under ideal conditions synthetic cells would have the ability to locomote on the graph by deterministically choosing edges to take between nodes. However, current state-of-the-art synthetic cells are incapable of deterministic locomotion. Hence, we must consider stochastic environments to simultaneously consider practical locomotor limitations of synthetic cells. and assess the robustness of our learned policies to inherent process noise.

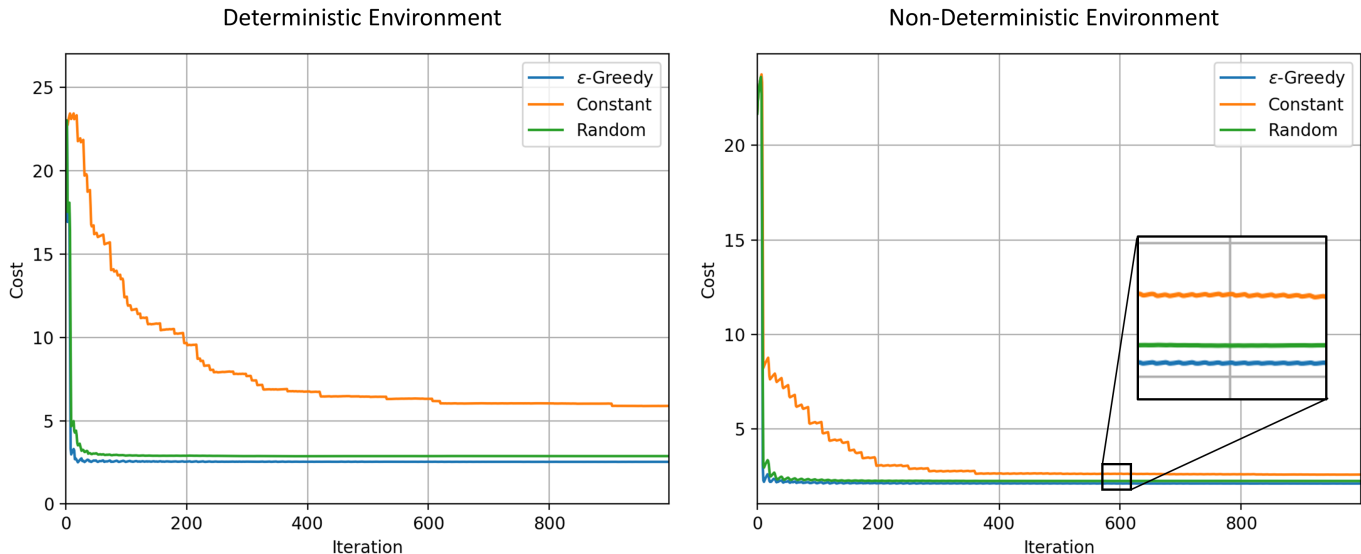


Fig. 2. Preliminary results comparing the performance of our learned coverage policies in both deterministic and nondeterministic models of the environment. The learned policies outperformed uniformly random and constant policies in both environments despite being fundamentally static.

To this end, we conduct a series of simulations where policies are learned using the ϵ -greedy method outlined in the previous section. Once the policies are learned, they are greedily executed. Thus, policies are static and executed without computations outside of state estimation. In addition to the learned policies, we implemented a constant policy and an uniformly random policy to compare against. In constant policies, all agents execute the same fixed strategy chosen independent of the goal distribution—this policy should be interpreted as a worst-case scenario. Uniformly random policies are such that each agent independently chooses a subsequent edge in the graph uniformly at random.

The goal distribution for all simulations was the following $p(S_{22}) = p(S_{18}) = 0.33$, $p(S_{19}) = p(S_{15}) = 0.16$, with the remaining probability mass evenly distributed between other states. To perform optimally, 66% of agents must be sent towards the head, and 33% must be sent towards the legs at the fork in state 14 of Fig. 1. This distribution was chosen such that policies achieving coverage could outperform a uniformly random policy. Results from the simulations are shown in Fig. 2, where we track the average loss over 50 trials for 1000 iterations. In both environments the learned policies outperform constant and uniform policies. Even under stochastic transition models, the learned, static, distributed policy performs better than the uniformly randomized policy.

V. DISCUSSION & FUTURE WORK

In this short manuscript we outline a proposed strategy toward generating motion plans for synthetic cells in a circulatory system, and present preliminary results. Particularly, we focus on the problem of generating coverage with respect to a desired state distribution. However, we note that such a distribution-based specification of the task can be easily manipulated to encode other behaviors through the desired goal distribution. Regarding the feasibility of the proposed procedure, the amount of memory required to encode the coverage

policies is not very large. The proposed circulatory system description is comprised of 24 states and 34 transitions, which would lead to a policy table with 816 elements in the worst-case scenario. However, much of the information encoded in such a table would be redundant given that there are many states in the graph arranged in linear chains. Moreover, with some basic compression procedures the amount of memory required to encode the policy table could be significantly reduced depending on the desired distribution.

The primary challenge in deploying motion planning algorithms on real synthetic cells lies in their limited locomotion capabilities. Current state of the art synthetic cells lack the control authority needed to deterministically select a path at a junction or fork on the road. To try to address these shortcomings in the state-of-the-art, we show that our policies can be robust to stochasticity in the environment despite being fundamentally deterministic.

As a final note, since synthetic cells can be manufactured in batches of very large quantities it may not be desirable to give each cell an individual policy. Instead, the set C could describe difference classes of synthetic cells. Then, as long as the number of synthetic cells belonging to each class is the same, the empirical time-averaged distribution in Eq. (1) will be the same.

In future work we will focus on developing modified strategies that can scale well with the size of the underlying state space. A promising path forward is to represent state space distributions in a Fourier space where the coefficients needed to describe the distribution may be much fewer than the number of states. Moreover, as a result of a future collaboration with the authors of [8], we hope to explore some of the outlined algorithmic strategies in an experimental system of synthetic cells undergoing flow in an environment similar to the circulatory system.

REFERENCES

- [1] F. A. Lavergne, H. Wendehenne, T. Bäuerle, and C. Bechinger, "Group formation and cohesion of active particles with visual perception-dependent motility," *Science*, vol. 364, no. 6435, pp. 70–74, 2019.
- [2] P. Liu, A. L. Cottrill, D. Kozawa, V. B. Koman, D. Parviz, A. T. Liu, J. Yang, T. Q. Tran, M. H. Wong, S. Wang, and M. S. Strano, "Emerging trends in 2d nanotechnology that are redefining our understanding of "nanocomposites";" *Nano Today*, 2018.
- [3] V. B. Koman, P. Liu, D. Kozawa, A. T. Liu, A. L. Cottrill, Y. Son, J. A. Lebron, and M. S. Strano, "Colloidal nanoelectronic state machines based on 2d materials for aerosolizable electronics," *Nature Nanotechnology*, vol. 13, pp. 819–827, Sep. 2018.
- [4] D. Seo, R. M. Neely, K. Shen, U. Singhal, E. Alon, J. M. Rabaey, J. M. Carmena, and M. M. Maharbiz, "Wireless recording in the peripheral nervous system with ultrasonic neural dust," *Neuron*, vol. 91, no. 3, pp. 529–539, 2016.
- [5] M. Sitti, H. Ceylan, W. Hu, J. Giltinan, M. Turan, S. Yim, and E. Diller, "Biomedical applications of untethered mobile milli/microrobots," *Proceedings of the IEEE*, vol. 103, no. 2, pp. 205–224, Feb. 2015.
- [6] D. D. Chaplin, "Overview of the immune response," *The Journal of Allergy and Clinical Immunology*, vol. 125, pp. S3–23, Feb. 2010.
- [7] T. Newman, "How the immune system works," *Medical News Today*, Jan. 2018. [Online]. Available: <https://www.medicalnewstoday.com/articles/320101.php>
- [8] P. Liu, A. T. Liu, D. Kozawa, J. Dong, J. F. Yang, V. B. Koman, M. Saccone, S. Wang, Y. Son, M. H. Wong, and M. S. Strano, "Autoperforation of 2d materials for generating two-terminal memristive janus particles," *Nature Materials*, vol. 17, pp. 1005–1012, Oct. 2018.
- [9] W. F. Paxton, K. C. Kistler, C. C. Olmeda, A. Sen, S. K. St. Angelo, Y. Cao, T. E. Mallouk, P. Lammert, and V. H. Crespi, "Catalytic nanomotors: autonomous movement of striped nanorods," *Journal of the American Chemical Society*, vol. 126, pp. 13 424–13 431, 2004.
- [10] H. H. Hardy, R. E. Collins, and R. E. Calvert, "A digital computer model of the human circulatory system," *Medical and Biological Engineering and Computing*, vol. 20, no. 5, pp. 550–564, Sep. 1982.
- [11] A. L. Strehl, L. Li, and M. L. Littman, "Reinforcement learning in finite MDPs: PAC analysis," *Journal of Machine Learning Research*, vol. 10, pp. 2413–2444, Dec. 2009.
- [12] G. Mathew and I. Mezić, "Metrics for ergodicity and design of ergodic dynamics for multi-agent systems," *Physica D: Nonlinear Phenomena*, vol. 240, no. 4, pp. 432–442, 2011.
- [13] I. Abraham, A. Prabhakar, and T. D. Murphey, "Active area coverage from equilibrium," *Workshop on the Algorithmic Foundations of Robotics (WAFR)*, 2018.
- [14] A. Pervan and T. D. Murphey, "Low complexity control policy synthesis for embodied computation in synthetic cells," *The 13th International Workshop on the Algorithmic Foundations of Robotics*, Dec. 2018.
- [15] G. Tesauro, "Extending Q-learning to general adaptive multi-agent systems," *Advances in Neural Information Processing Systems*, pp. 871–878, 2004.