The Seventeenth International Symposium on Artificial Life and Robotics 2012 (AROB 17th '12),
B-Con Plaza, Beppu, Oita, Japan, January 19-21, 2012

Control of real world complex robots using a biologically inspired algorithm

Fabio DallaLibera[1,2], Shuhei Ikemoto[3], Hiroshi Ishiguro[2], and Koh Hosoda[3]

[1] JSPS Research Fellow / JST ERATO Asada Project / Padua University, Padua 35131, Italy
(fabio.dl@irl.sys.es.osaka-u.ac.jp)
[2] Osaka University, Toyonaka, Osaka 560-8531, Japan
[3] Osaka University, Suita, Osaka 565-0871, Japan

Abstract: Elementary living beings, like bacteria, are able to reach food sources using only limited and very noisy sensory information. In this paper we describe a very simple algorithm inspired from bacteria chemotaxis. We present a Markov chain model for studying the effect of noise on the the behavior of an agent that moves according to this algorithm, and we show that, counterintuitively, the application of noise can increase the expected average performance over a fixed available time. After this theoretical analysis, experiments on real world application of this algorithm are introduced. In particular, we show that the algorithm is able to control a complex robot arm, actuated by 17 McKibben pneumatic artificial muscles, without the need of any model of the robot or of its environment.

Keywords: Robotics, Control techniques, Artificial life

1 INTRODUCTION

Bacteria chemotaxis, the process by which bacteria move toward increasing concentrations of nutrients, is an important topic of the biological field. In particular, the movement of Escherichia Coli (E. Coli) was intensively studied [1]. This bacterium has only two ways of moving, rotating its flagella in clockwise or counterclockwise direction. When the flagella rotate counter-clockwisely they align into a single bundle and the bacterium swims in a straight line. Conversely clockwise rotations break the flagella apart and the bacterium tumbles in place. The bacterium keeps alternating clockwise and counterclockwise rotations. In absence of chemical gradients, the length of the straight line paths, i.e. the counterclockwise rotations, is independent of the direction.

In case of a positive gradient of nutrients, E. Coli reduces its tumbling frequency. In other terms, when the concentration of nutrients increases the bacterium proceeds in the same direction for a longer time. This simple strategy, usually modeled by a biased random walk, is able to drive bacteria to high concentrations of nutrients despite the difficulties in precisely sensing the gradient.

In [2] we proposed a simplified, more generic model for the movement of E. Coli, that we termed *Minimalistic Behavioral Rule* (MBR). Simulations with this model showed that while in most engineering approaches noise is considered as a nuance, that must be filtered to extract the underlying information, noise can actually help bacteria in reaching high concentrations of nutrients.

This paper provides a theoretical analysis of the phenomenon. A Markov chain model is introduced in section 3 to explain the beneficial effect of noise on the movement of an MBR-driven agent. The possibility of employing MBR for the control of complex systems is then shown experimentally by using a robot arm driven by pneumatic actuators (Section 4). Finally, Section 5 concludes the paper by discussing the results.

2 ALGORITHM

As briefly described in the previous section, E. Coli proceeds in the same direction for longer times when the conditions are improving, i.e. when the the concentration of nutrients increases, and proceeds by random walk otherwise. In [2] we defined a very general control algorithm that follows a similar principle. This algorithm, besides being able to explain phenomena regarding bacteria chemotaxis, can be applied for the control of robots, as will be shown in this paper. More formally, the algorithm can be described in the time-discrete domain as follows. Let us denote by E_t an evaluation function (e.g. the concentration of nutrient in the bacteria's case), and by ΔE_t its variation from time $t-1$ to time t. Let us then define a generic motor command $u_t \in \mathbb{R}^n$ that the agent sends to its actuators. The MBR sets u_{t+1} as

$$u_{t+1} = \begin{cases} u_t + \eta R & if\ \Delta E_t \geq 0 \\ random\ selection & otherwise \end{cases}. \tag{1}$$

where $R \in \mathbb{R}^n$ is a random variable, $\eta \in \mathbb{R}$ is a multiplying factor and "random selection" corresponds to randomly picking a motor command in the whole motor command space. Intuitively, MBR tells to change randomly the motor command when the conditions get worse, and keep the current motor command when conditions are improving. In more detail, when the conditions are improving, the motor command is not kept identical to the previous step, but it is slightly modified, with a random perturbation whose inten-

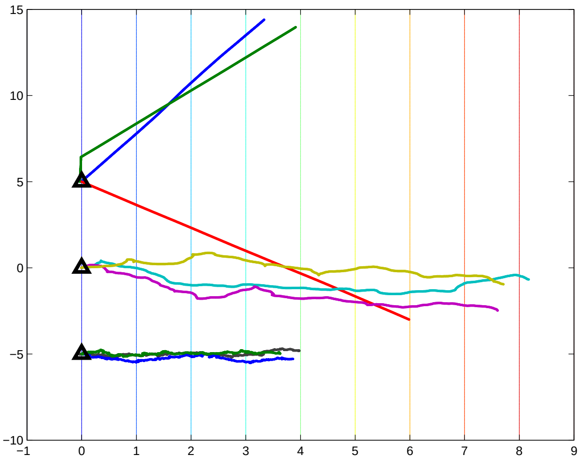

Fig. 1. Paths taken by agents with different values of η: agents starting from $[0,-5]^T$, $[0,0]^T$ and $[0,5]^T$ have, respectively low, medium and high values of η. Each trajectory corresponds to 100 time steps of the agent.

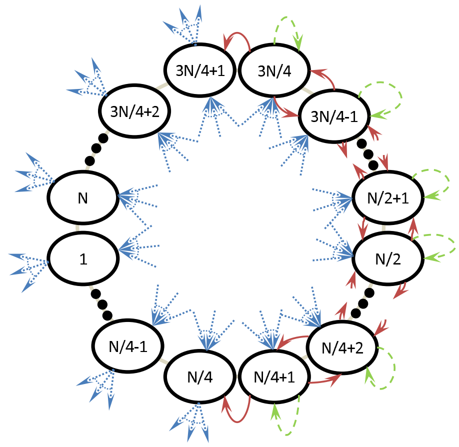

Fig. 2. Representation of the Markov chain. Blue dotted arrows represent the selection of a new node with uniform, $1/N$ probability. Red solid arrows represent transition to the neighboring states with probability v and green dashed arrows indicate loops of probability $1-2v$.

sity is modulated by η. Fig. 1 shows an example of the effect of this random perturbation on the movement of an MBR-controlled agent. Suppose to have an agent that moves on a 2D plane. Let $x \in \mathbb{R}^2$ be the agent position. Assume the agent to move by steps of fixed length s, along the angle indicated by $u_t \in \mathbb{R}$, i.e. $x_{t+1} = x_t + s \cdot [\cos(u_t), sin(u_t)]^T$ and to have an evaluation function that increases along the x axis, i.e. $E_t = [1,0]^T x_t$. If the perturbation η is too small, as for the agents starting from $[0,5]^T$ in Fig. 1, then every motor command u that that bring brings the agent closer to the target, even a very tangential and inefficient one, will be employed for a long time. If the perturbation is too big, as for the agents starting at the bottom of the figure, then the agent keeps changing direction, even when the movement is headed straight to the goal. If the perturbation amplitude η is appropriate, as for the agents starting in the middle of Fig. 1, then the agent will rapidly increase E_t with a good trajectory.

3 THEORETICAL ANALYSIS

In order to understand the reasons of the performance improvement given by an opportune value of $\eta \neq 0$, let us consider again example reported in the previous section. Let us discretize the direction taken by the agent into N possible states of a Markov chain, with the state i, $1 \leq i \leq N$ indicating the probability of taking the directions in the range $\left[-\pi + \frac{2\pi(i-1)}{N}, -\pi + \frac{2\pi+i}{N}\right)$. Let us restrict to the case where transitions due to the perturbation η are limited to neighboring states. In particular, let us assume to transit from a state i to one of its neighbors $i+1$ and $i-1$ with probability v, and remain in the same state with probability $1-2v$, as shown in Fig. 2. We note that only for the states corresponding to directions that lead to an increase of E_t, i.e. the states $\frac{N}{4}+1 \leq i \leq \frac{3N}{4}$, such outward transition to the neighboring states exist. The states corresponding to directions for which $\Delta E_t < 0$, i.e. the directions that imply a random selection in the MBR, have outward transitions that reach each state, including the current state, with probability $1/N$.

The evolution of the probability of the states is described by the following equations. Let us denote by $p_{i,t}$ the probability of being in state i at time t, and by $g = \frac{1-\sum_{i=N/4+1}^{3N/4} p_{i,t}}{N}$ the probability of being selected by random selection when $\Delta E_t < 0$. With these definitions

$$p_{i,t+1} =$$

$$\begin{cases} vp_{i+1,t} + g & if\ i = \frac{N}{4} \\ (1-2v)p_{i,t} + vp_{i+1,t} + g & if\ i = \frac{N}{4}+1 \\ (1-2v)p_{i,t} + v(p_{i+1,t} + p_{i-1,t}) + g & if\ \frac{N}{4}+2 \leq i \leq \frac{3N}{4}-1 \\ (1-2v)p_{i,t} + vp_{i-1,t} + g & if\ i = \frac{3N}{4} \\ vp_{i-1,t} + g & if\ i = \frac{3N}{4}+1 \\ g & otherwise \end{cases} \tag{2}$$

Intuitively, when $v \neq 0$ at every time step there is a "flow of probability" between the nodes. In detail, even starting from a uniform probability of the nodes $p_{i,0} = 1/N$, the nodes $1 \leq i \leq \frac{N}{4}$ and $\frac{3N}{4}+1 \leq i \leq N$ assume a lower probability in one time step. This causes a flow of probability from nodes $\frac{N}{4}+1$ to $\frac{N}{4}$, $\frac{N}{4}+2$ to $\frac{N}{4}+1$, etc. (and, symmetrically, from $\frac{3N}{4}-k$ to $\frac{3N}{4}-k+1$, for $0 \leq k \leq \frac{N}{4}-1$).

Fig. 3 shows the evolution of the probability for different values of v (corresponding, in the MBR, to different values of η) at different time instants t starting from a uniform distribution, i.e. starting from a random direction of the agent. We note that at the beginning a higher value of v is able to "speed up the flow", that makes the distribution peaky around states

$\frac{N}{2}$ and $\frac{N}{2}+1$, corresponding to angles close to 0. As times goes on, however, the distribution becomes more peaky for lower settings of v, as shown in the last panel of Fig. 3.

Let us analyze the chain for $t \to \infty$. The stationary distribution, computed from Eq. 2 and $p_{i,t+1} = p_{i_t}$, is

$$p_{i,t+1} = \begin{cases} \frac{24(n+4)v}{n(8+6n+n^2+96v)} & if i = \frac{N}{4} \wedge i = \frac{3N}{4}+1 \\ \frac{-48i^2+48i(n+1)-12n-9n^2}{n(8+6n+n^2+96v)} & if \frac{N}{4}+1 \le i \le \frac{3N}{4} \\ \frac{96v}{n(8+6n+n^2+96v)} & otherwise \end{cases} \tag{3}$$

If we take the derivative with respect to v we find that $\frac{dp_{i,\infty}}{dv}$ is always negative for $\frac{N}{4}+1 \le i \le \frac{3N}{4}$ (states that lead to positive ΔE_t) and always positive for the other states (that lead to a negative ΔE_t). In other terms, over short times high values of v are preferable, but as the time available increases, it is more efficient to have a lower value of v. In short, there is a trade-off between the speed with which good motor commands u are chosen, increased by higher values of η, and the precision obtained after a long period, improved by decreasing η. Section 5 will briefly discuss this issue.

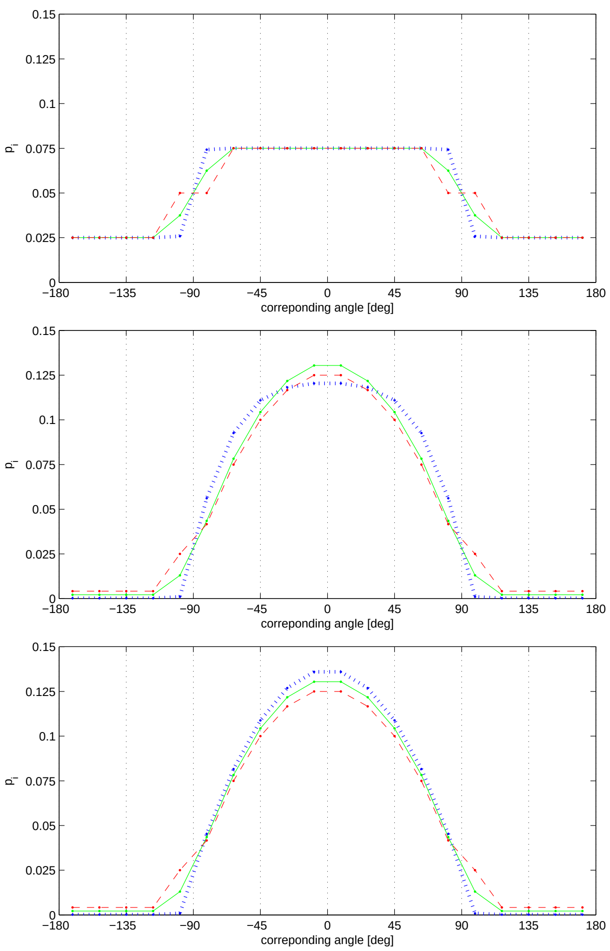

Fig. 3. Probability of a Markov chain of $N = 20$ states starting from a uniform distribution at time $t = 0$. Each panel corresponds to a time instant: $t = 1$ (top panel), $t = 10$ (central panel) and $t = 10^{10}$ (bottom panel). Each panel reports the distribution obtained 3 different values of v: $1/64$ (blue dotted line), $1/4$ (green solid line) and $1/2$ (red dashed line).

These relationships between the noise intensity η, the distribution of the directions taken by the agent and time can be confirmed by numerical simulations of the MBR, with results that resemble the Markov chain approximation here introduced. Actually, the probability density function $p(\theta)$ of taking direction θ is the solution of the following Fredholm integral equation of the second type:

$$\begin{aligned} p(\theta) &= \int_{-\pi/2}^{\pi/2} K(\theta,\phi)p(\phi)d\phi + \frac{1-\int_{-\pi/2}^{\pi/2} p(\phi)d\phi}{2\pi} \\ &= \frac{1}{2\pi} + \int_{-\pi/2}^{\pi/2} (K(\theta,\phi) - \frac{1}{2\pi})p(\phi)d\phi \end{aligned} \tag{4}$$

where K is the probability of taking the direction ϕ when the current direction is θ due to the random perturbation R. For instance, if R is Gaussian, then K is a Von Mises distribution centered in θ, i.e. $K(\theta,\phi) = \frac{e^{cos(\phi-\theta)}}{2\pi I_0(1)}$ with $I_0(x)$ the modified Bessel function of order 0.

However, no closed form can be obtained for $p(\theta)$. The main advantage of the Markov model proposed in this section is thus to allow making predictions on the behavior of an MBR controlled agent, using simple in analytical formulae, without lengthy Monte Carlo simulations.

4 PRACTICAL EXPERIMENTS

Previously reported simulation experiments [3] show that the application of MBR is not restricted to the simple 2D case described beforehand. In particular, it was shown that MBR is able to increase the evaluation function E_t of systems with dynamics

$$x_{t+1} = f(x_t, u_t, \ldots, u_{t-d}) \tag{5}$$

even when $x_t \in \mathbb{R}^n$ $u_t \in \mathbb{R}^m$ with n and m in the order of 40 or more dimensions. It was also shown that the function f can be a nonlinear function of the previous inputs $u_{t-1}, \ldots, u_{t-d}$, and in particular, even when a dead time is present in the system, i.e. $x_{t+1} = f(x_t, u_{t-d})$.

To test the MBR approach in a real world setting, we employed the biologically inspired robot arm shown in Fig. 4. This arm has 7 degrees of freedom driven by 17 McKibben pneumatic muscles. Each muscle is equipped with a pressure sensor, used for closed loop pressure control.

The robot is controlled by setting the variation of the pressure in each of the 17 pneumatic actuators, i.e. $u_t \in \mathbb{R}^{17}$. The task chosen consists in reaching repeatedly three targets in the robot reachable space. These targets are located at the robot's right, left and bottom part of the reachable space.

To verify the robustness of the approach to sensory input noise, we tested three ways of measuring the end effector position, from which the function E_t is computed:

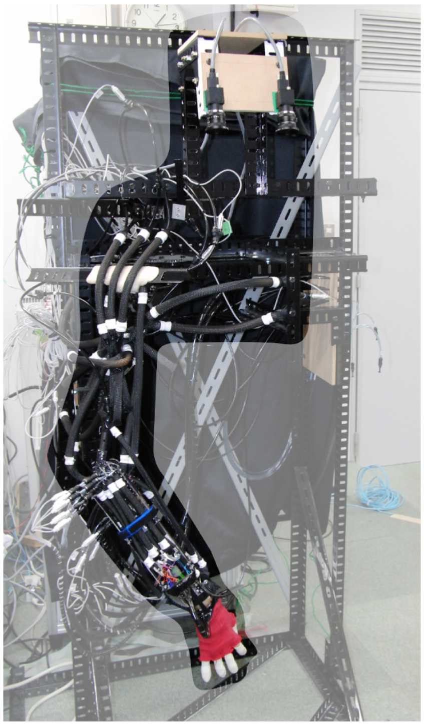

Fig. 4. The robot arm used in the experiment.

1. A four dimensional vector composed by the end-effector centroid in the images of the two cameras mounted on the robot's head.
2. The three dimensional position of the end effector, obtained using stereo computation.
3. A three dimensional position of the end-effector observed by a motion capture system (EvaRT, Motion Analysis Ltd.)

The task could be achieved with all the three types of the information, showing the adaptability and the robustness of the approach. We note that the value of η was chosen empirically, without fine tuning, and kept constant in all the settings. In fact, although not optimal, any $\eta \neq 0$ can be employed for achieving the task.

The real world experiment is different for many points compared with the system analyzed in Section 3: the dimensions of the input command u, of the space x, the mapping between u and x, the noise in the sensory system, etc. Nonetheless, as Fig. 5 shows, the distribution of the cosine between the optimal direction (straight to the target) and the direction taken by the robot during the whole experiment is a decreasing function, as predicted by our simplified model.

5 CONCLUSIONS

In this paper we proposed a Markov chain model to study MBR, a simple, biologically inspired algorithm for robot control. We showed that the algorithm is able to drive a pneumatic robot arm's end-effector to desired targets without any knowledge on the mapping between the control signal u and the resulting evaluation function variation ΔE_t or on the directions taken by the links when each muscle is contracted. We analyzed the distribution of the cosine between the optimal movement and the direction actually taken by the robot, and showed that the probability is a decreasing function of the angle, as predicted by our model, despite the numerous differences between the model and the real world setup.

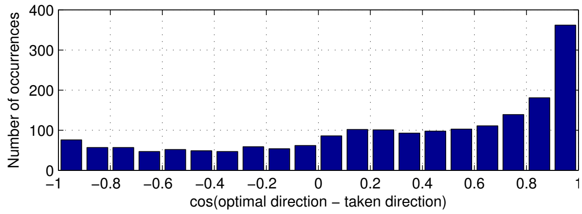

Fig. 5. Frequency of the cosine between the direction taken by the MBR-controlled robot arm and the direction straight to the target location.

Future works will need to define policies for changing η online and maximize the performances. We note that starting with a high η and decrease it over time, as in Simulated Annealing, may allow obtaining solutions with high precision (due to a low final η) in a short time (due to an initially high η). However, estimating an appropriate time constant for the adaptation of η is very complex. In fact, this depends on how fast the optimal u changes, which, in turn, depends both on the evaluation function E and on the system dynamics f.

We note that, despite the similarities, MBR presents advantages over classical Simulated Annealing. MBR does not require a good estimation of the evaluation function, but only a binary value specifying if it is increasing or not. Furthermore, it does not need to return to the one state x_t after noticing that x_{t+1} is a less favorable state. This is fundamental when dealing with real word robots of unknown dynamics. Comparison of the two algorithms in real world setups is, however, an important point to be focused in the future.

REFERENCES

[1] J. Adler. The sensing of chemicals by bacteria. *Scientific American*, 234:40–47, 1976.

[2] S. Ikemoto, F. DallaLibera, and H. Ishiguro. Stochastic resonance emergence from a minimalistic behavioral rule. *Journal of Theoretical Biology*, 273(1):179–187, 2011.

[3] F. DallaLibera, S. Ikemoto, T. Minato, H. Ishiguro, E. Menegatti, and E. Pagello. A parameterless biologically inspired control algorithm robust to nonlinearities, dead-times and low-pass filtering effects. In *Proceedings of SIMPAR2010*, pages 362–373.