Toward a Gene Regulatory Network Model for Evolving Chemotaxis Behavior

Neale Samways, Yaochu Jin, Xin Yao, Bernhard Sendhoff

2008

Toward a Gene Regulatory Network Model
for Evolving Chemotaxis Behavior

Neale Samways, Yaochu Jin, Xin Yao, and Bernhard Sendhoff

Abstract— Inspired from bacteria, a gene regulatory network model for signal transduction is presented in this paper. After describing experiments on stabilizing the population size for sustained open-end evolution, we examine the ability of the model to evolve gradient-following behavior resembling bacterial chemotaxis. Under the conditions defined in this paper, an overwhelming chemotaxis behavior does not seem to emerge. Further experimentation suggests that chemotaxis is selectively favored, however, it is shown that the gradient information, which is critical for evolving chemotaxis, is heavily degraded under the current regime. It is hypothesized that lack of consistent gradient information results in the selection of non-chemotaxis behavior. Future work on revising the model as well as the environmental setups is discussed.

I. INTRODUCTION

One challenging target in computational systems biology is to understand the dynamics of gene regulatory networks, which plays a central role in understanding natural evolution and development [1]. To this end, various models of gene regulatory networks for signal transduction have been suggested. These models attempt to investigate, mimic and utilize genetic regulation in a computational framework inspired from many classes of biological systems, particularly the most comprehensively understood bacteria such as *escherichia coli*. Simpler than multicellular and prokaryotic organisms, bacteria surprisingly exhibit a range of complex behaviors, one of which is chemotaxis. This behavior involves directed movement up gradients of attractant (such as a nutrient) or down gradients of repellent.

As bacterial behavior is heavily influenced by external environmental conditions (including other bacteria), one approach for investigating genetic regulatory systems involves the modeling of populations or complete ecologies within a dynamic environment. Such models are often termed individual based models (IbMs) [2], [3], where every individual possesses its own set of state variables and parameters allowing for dynamic behaviors [4]. In the following, we discuss briefly a few biologically motivated IbM systems that closely relate to our work.

COSMIC (COmputing System of Microbial InteraCtions) [2], [5], [6] is a computational model of bacterial growth and evolution, predicated on phenomena known to occur in bacterial cells. The goal of this system was to evolve a controller capable of directing chemotaxis through a genetic regulatory network. The gene regulatory network contains genes that can produce proteins (receptors) for the detection of nutrient concentration therefore presenting an input to the genetic mechanisms. The environment of COSMIC permits cell movement as well as competition for resources, enforcing an implicit fitness function across the population. A uniform level of nutrient is initially present in the environment, and upon ingestion and metabolism leads to an increase in cell size and subsequent reproduction. During the evolution, mutation in the form of sequence insertion and deletion is performed. Experiment results show that COSMIC can generate various behaviors, though it is unclear whether chemotaxis has evolved [5].

Penner et al. [7], [8], [9] developed an IbM bacterial ecosystem that comprises of individual bacterium encoded on a DNA–like genome, inhabiting a simulated agar plate. The agents produce proteins according to the genome, whilst interactions between proteins and DNA give rise to a regulatory network. An artificial chemistry consisting of a rule-base mediates genetic regulation and is central to the system. This rule base is user–defined based on known reactions to occur in *e.coli* and not evolved. Cellular behaviors are expressed upon the creation of proteins, whilst interaction (subject to the rules of the artificial chemistry) occurs between modules of the bacteria and local environment. Contents of the environment change as a result of interaction with the bacteria and through decay and diffusion. Although lacking in depth analysis, the results indicate the behavior of the model is congruent with *in vivo* experimentation and an increased tumbling frequency upon depleted levels of nutrient within the environment is observed. In a later stage of the experiment, evolutionary mechanisms have been added to a highly simplified model through the addition of mutation and recombination operations acting on the artificial chemistry. Bacteria circle and color patterns are evolved, though no analysis has been done on whether chemotaxis evolves.

There are also other IbMs developed incorporating less 'low–level' biological details of regulation, such as learning classifier systems. Both HERBY [10] and RUBAM [4] involve populations of organisms and their adaptation to environmental factors during evolutionary and organismal lifetime. HERBY simulates an ecology in which agents move and eat within a discretized grid world, whilst RUBAM models bacteria using a fuzzy classifier system, and aims at evolving chemotaxis. This system represents a much more 'coarse-grained' model than those presented above, and has a focus on application to real world problems [2], [6].

The BACSIM model, as outlined in [11] is a simulation of the growth and behavior of bacteria associated with

Neale Samways and Xin Yao are with the School of Computer Science, University of Birmingham, UK, B15 2TT. Yaochu Jin and Bernhard Sendhoff are with Honda Research Institute Europe GmbH, Offenbach am Main, Germany. Email: {yaochu.jin;bernhard.sendhoff}@honda-ri.de.

properties including substrate uptake, metabolism and reproduction. Although this model is capable of producing qualitative population growth data consistent with that from *in vitro* experiments, individual cells are not associated with genetic material, or directly amenable to evolution. A more biologically motivated systems for the evolution of chemotaxis was developed by Soyer et al. [12]. Although the model shows impressive results, its evolutionary framework is notably artificial in that it uses an explicit fitness function that captures essential features of chemotaxis.

As a whole, although various systems have been developed offering some interesting results, no IbM incorporating a biologically–inspired genetic regulation mechanism has sufficiently shown success in evolving chemotaxis behavior.

The model outlined in this paper was developed with the intention of investigating the way in which genetic control develops through evolution with inspiration from natural systems. The long-term goal of this research concerns the testing of a hypothesis that genetic control networks become increasingly influenced by internal conditions as evolution proceeds.

The remainder of the paper is organized as follows. Section II describes the main components and operations of the proposed model. In Section III, experimental setups are given. Section IV presents and discusses the results of two experiments with the model, concentrating on managing a stable population size and thus evolving chemotaxis behavior. Section V summarizes the paper and suggests the future work.

II. Model Description

Like existing IbMs, the model suggested in this paper is an abstraction of *in vivo* bacteria such as *e.coli*, including the gene regulation for signal transduction and behavioral functions. At the core of the system, a population of individuals acts in a simple virtual environment, with particular behaviors governed by an artificial genetic regulatory network (GRN). In the following, a more detailed description of the model is given.

A. The Environment

The environment in which the population exists is considered an abstract representation of an agar plate. For simplicity, the entire space is divided into a grid, with each location (site hereafter) maintaining an independent concentration of nutrient. Although the world is discrete with respect to nutrient, individuals are associated with a location in two dimensional real space. The immediate environment of an individual corresponds to the site to which its location translates; a simple function maps a two-dimensional real location to a site within the grid by rounding co–ordinates down to an integer value. Boundary conditions for the world are enforced through simple clamping of any individuals attempting to cross a boundary, until directed away by appropriate movement behavior.

Nutrient is added to the environment at regular time intervals. During nutrient deployment, a randomly chosen site is selected as the centroid and receives a maximum increment in nutrient concentration. Details for nutrient deployment will be given in Section III-A.

B. Genetic Representation

All individuals possess a genome comprising of virtual DNA. The genome consists of a variable number of operons, each of which is composed of one promoter denoted by $\mathbf{P} = <0, 1, 0, 1>$ and three genes. Each gene has two codons encoding a protein, with each codon, i.e., a triplet of bases $b \in \{0, 1, 2, 3\}$ representing the four nucleotides, A, C G, and T, coding an amino acid. In other words, six bases together will code one protein.

Thus, the compliment of proteins an individual is capable of producing is exclusive and determined by the gene content of the genome. The mapping from gene sequence to protein is fixed and redundant, such that the total compliment of distinct proteins is 16. Among these 16 proteins, three are associated with behavioral functions, while the remaining play the role of transcription factors. Details of the encoded functions will be discussed in Section II-E.

C. Gene Regulation through Bio-chemical Interactions

In this model, gene regulation is modeled through the gene-protein and protein-protein interactions defined by an interaction matrix.

Based on the expression level of the genes and influenced by interactions between genes and proteins, each individual has concentrations of chemicals that change over time. At any given instant the internal states, i.e., the concentration of the chemicals, can be represented with a vector as follows:

$$\vec{P}_t = \left(\{[\phi_1], [\phi_2], \ldots, [\phi_n]\}_t\right), \tag{1}$$

where $[\phi_x]$ denotes the concentration of chemical ϕ_x, n is the total number of internal chemicals. In this work, internal chemicals include the 16 encoded proteins plus one additional protein indicating the stored energy.

Changes in internal state result from the modification of chemical concentrations, occurring synchronously in a process illustrated by Equations (2) and (3). In Equation (2) the state at time $t + 1$ is calculated as the addition of a vector of changes $\vec{D}_t$ to the state vector at the previous time step. In Equation (3), values $Q_1 ..., Q_n$ represent the changes for individual chemicals, determined by the mechanisms of the GRN.

$$\vec{P}_{t+1} = \vec{P}_t + \vec{D}_t, \tag{2}$$
$$\vec{D}_t = \langle Q_1, Q_2, \ldots, Q_n \rangle. \tag{3}$$

The GRN of an individual results from the genes and artificial chemistry. The chemistry describes possible binary interactions between all chemicals, and is represented by a two–dimensional matrix $\mathbf{C}$. Each matrix element $c_{p,q}$ describes the effect of chemical q on chemical p. As the effect of every chemical pair must be specified, the matrix is necessarily square and of size $n \times n$. It is evident that

a row represents the effect of all chemicals on a single chemical associated with the row. The set of permissable interactions between any two given chemicals, q and p, is given in Equation (4).

$$c_{p,q} \quad \in \quad \left\{ 0, \quad \frac{1}{1+e^{-[\phi_q]}}, \frac{1}{1+e^{[\phi_q]}} \right\}, \quad (4)$$

where: $\quad (p,q) \quad \in \quad \mathbb{Z}, \qquad 1 \leq (p,q) \leq n.$

The element 0 indicates a null interaction, whilst the other expressions describe a sigmoidal relation between reactant concentration and change in product concentration. The change in concentration of a single chemical a can be determined by:

$$\forall Q_a \in \vec{D} : Q_a \quad = \quad \left[\tanh \left(\sum_{i=1}^{n} c_{a,i} \right) \cdot \alpha \right] + \delta, \quad (5)$$

where, δ is the production rate without any interactions, α represents the basal production level of a gene, ensuring a fixed level of protein production in the absence of any non–null interaction. Both increment and decrement in chemical concentration are permitted, which simulates the activation and repression of gene expression.

D. Genetic Variations

When the model is evolved, variations to the genome will be performed. Three genetic operations are adopted in this work, including single nucleotide polymorphism (SNP), gene duplication, and gene deletion.

SNP involves a simple, random re–assignment of a base at a given locus of the DNA. Where this mutation is determined to occur, the new value of the mutating base is selected with equal probability from the four bases. Genetic duplication can occur once during reproduction, with a fixed pre–set probability. When duplication occurs, a 'duplication length' is selected randomly between 1 and the length of the parental DNA. Subsequently, a suitable splice locus is determined at random such that the remainder of bases beyond this locus is equal to or greater than the duplication length. A string of bases of duplication length are spliced into the offspring DNA at the splice locus. The spliced DNA is a duplicate of the parental DNA beginning from the splice locus. Effectively a chunk of DNA is repeated contiguously within the offspring genome thus increasing its length by the duplication length. Genetic deletion is notionally similar to duplication, however a length of DNA of 'deletion length' is expunged from the offspring DNA at a suitable splice locus. The resulting offspring genome is therefore shorter than that of the parent by a length equal to the deletion length. Owing to the redundant nature of the DNA, it is possible for all types of genetic variations to affect only 'junk DNA', therefore having no effect on the phenotype.

As with the genome, the interaction matrix of a newly generated individual is subject to probabilistic mutation too. Each matrix element is mutated with a pre–defined, fixed probability. When mutation occurs, the value of the element in question is reassigned to any permitted value, with equal probability. As a given individual may or may not possess genes to produce the full compliment of proteins, it is possible that certain mutations do not affect the behavior of the individual. However, as a ramification of the role of the interaction matrix, a single mutation can modify the connectivity of the GRN, drastically modifying behavior.

E. Behavioral Actions and Action Selection

Each individual has a number of behavioral actions, including sensing the gradient of nutrient, running, tumbling, reproducing, and dying. Among them, sensing the nutrient gradient is realized by modulating the concentration of a "sensing" protein representing the signaling pathway between the environment and receptor, thus presenting as an input to the gene regulatory network. The action of reproducing and dying is determined by the internal energy level of the individual, represented by the concentration of an "energy indicating" protein. When the energy level is higher than a prescribed threshold, a bacterium will reproduce an offspring. The concentration of all chemicals, as well as the internal energy, is divided equally between the parent and its offspring. When the energy level goes to zero, the individual dies. Note that the concentration of the energy indicating protein is not directly influenced by the internal states.

The two motile behaviors are determined by the concentration of two other proteins, one "tumbling" protein, and one "running" protein. The concentration of these two proteins are directly determined by the expression level of the corresponding genes. When the concentration level of the running protein is larger than the threshold, the individual performs the running behavior. Similarly, when the concentration level of tumbling protein exceeds the predefined threshold, the individual tumbles. If the concentration of both proteins is above the threshold, the winner-take-all strategy is adopted. On the other hand, no motile actions will be invoked if the concentration of the motile proteins is below the threshold.

Each individual is associated with a heading value in the range $[0-2\pi)$, which is randomly assigned when the individual is generated, describing its current orientation. This value remains fixed unless a tumble behavior is invoked, upon which a random re–assignment is made. When an individual runs, it moves in the environment along the current orientation for a fixed number of sites.

Consumption of a fixed amount of nutrient (environment permitting) occurs automatically once per epoch for all individuals. Where nutrient is consumed, a decrement is made to the concentration of nutrient within the environment, whilst an increment made to the concentration of the energy indicating protein of the corresponding individual.

At the end of each epoch, the concentration of all proteins will be updated depending on the existing concentration and the interaction matrix. Based on the updated protein concentration, corresponding behaviors will be performed.

TABLE I

EXPERIMENTAL PARAMETERS

Parameter	Value
Maximum food intake (per epoch)	1.0
Nutrient to energy conversion rate	0.1
Life energy cost (per epoch)	0.0001
Run / tumble energy cost	0.003
Behavioral protein trigger threshold	25.0
Reproduction trigger threshold	25.0
Maximum movement range	1.5

III. EXPERIMENTAL SETUPS

To verify the functioning of the system, two sets of experiments were undertaken. The goal of Experiment 1 (Exp. 1) is to elucidate appropriate parameter ranges regarding food deployment for attaining a stable population size. Experiment 2 (Exp. 2) was conducted in an evolutionary context, with the aim of investigating necessary factors for the evolution of chemotaxis behavior.

In both experiments, the environment is defined by a 500 by 500 grid, consisting of 250000 sites. The initial population size is 50, and simulation terminates either when the population size reaches zero, or when the number of completed epochs reaches 10^8. In Exp. 2, the probability for all three genetic operations is 0.03, and the probability mutating the interaction matrix is 0.004. For each of the experimental conditions, 10 independent runs were undertaken. All other parameter setups used in the experiments, unless otherwise stated, are listed in Table I. While most parameters are self-explaining, the maximum movement range means the number of sites individuals move during the running behavior.

A. Exp. 1: Fixed, Hand–Coded Genomes

To evolve chemotaxis behavior, a stable population size needs to be maintained. For this purpose, nutrient is deployed according to a cone-like distribution, with N_c the units of food to be deployed on the centroid and N_w (number of sites) the radium of the food distribution. The location of the centroid was chosen with equal probability within a square of 200×200 sites concentric with the environment.

Experiments were conducted for obtaining parameter configurations, including N_c, N_w, and the period of food deployment, resulting in a stable and computationally tractable population size. In this experiment, the initial population was seeded with a hand–coded genome and zero interaction matrix, thereby enforcing running and tumbling with equal frequency.

B. Exp. 2: Evolution of Genome and Interaction Matrix

To investigate the possible emergence of chemotaxis, further experiments were performed in which evolution was permitted via mutation of the genome and interaction matrix. Although evolutionary fitness is implicit, it was hypothesized that individuals capable of following an increasing gradient of nutrient would survive and reproduce at a greater rate

than those that could not. In this circumstance, chemotaxing individuals would proliferate in the population.

The initial population comprised of individuals with a randomly generated genome and interaction matrix. Genomes of the initial individuals were created by generating random strings from the elements of the DNA set with equal probability. Interaction matrices were synthesized in a similar manner, however the probability of an element being non–zero was set to 0.3 to prevent massively interconnected networks within the initial population.

IV. SIMULATION RESULTS

A. Ex 1: Stabilizing the Population Size

A variety of ad–hoc experiments were conducted, in which the parameters pertaining to nutrient deployment were varied. Runs illustrating an unstable population size are shown in Fig. 1, which results from an infrequent nutrient deployments coupled with higher energy costs for running and tumbling, Furthermore, an energy cost 0.1 units for reproduction is employed.
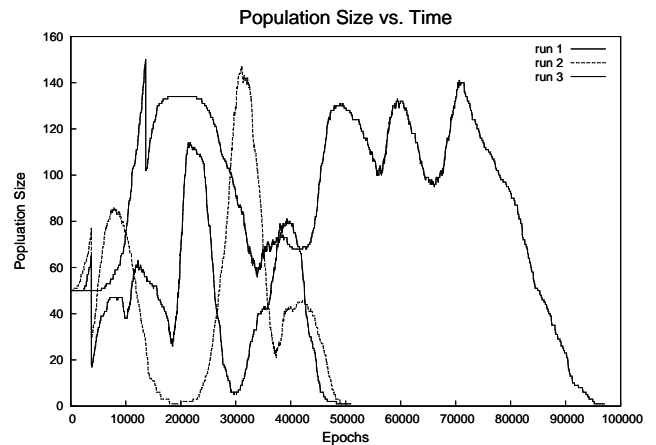

Fig. 1. Unstable populations arising from parameter settings: Energy for running / tumble 0.03, maximum movement range = 1.0, food deployment period = 2000, maximum volume of food deployments = 20.

Experimentation has shown that stability in population size is mainly dependent on the energy cost for running and tumbling, the energy threshold for reproduction, and the volume and period of nutrient deployment. Our results showed that the resting population size is proportional to the product of the deployment period and volume. Where this product is fixed, but the period varies, the amplitude in oscillation is observed as being proportional to the period.

Fig. 2 shows the population size averaged over 10 runs for seven different nutrient deployment periods, where N_c is set to 0.4 units and N_w is fixed to 25 sites. We see that a computationally tractable and stable population size can be obtained when the period of food deployment is roughly between 500 and 2000. Note that when the food deployment period is larger than 50, most of the 50 initial individuals live on their initial energy and die after about 40,000 epoches. because the initial energy is not sufficient for reproduction, and the probability of finding nutrient with random work is
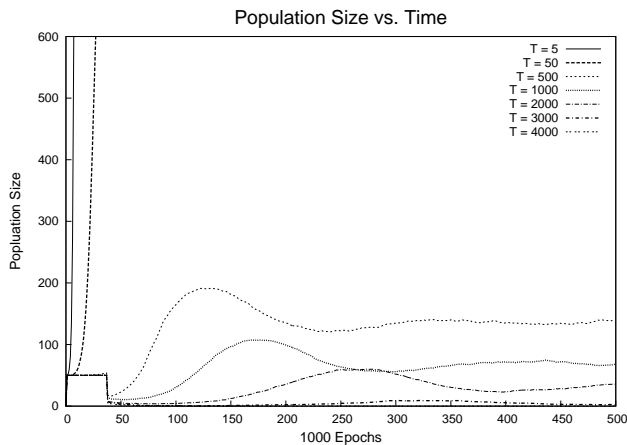
Fig. 2.   Changes of population size with the food deployment period.

low. Fig. 3 shows a stable population size when the food deployment period is set to 500, i.e., food will be deployed every 500 epoches. Although oscillation is apparent, a mean population size of approximately 135 is observed.
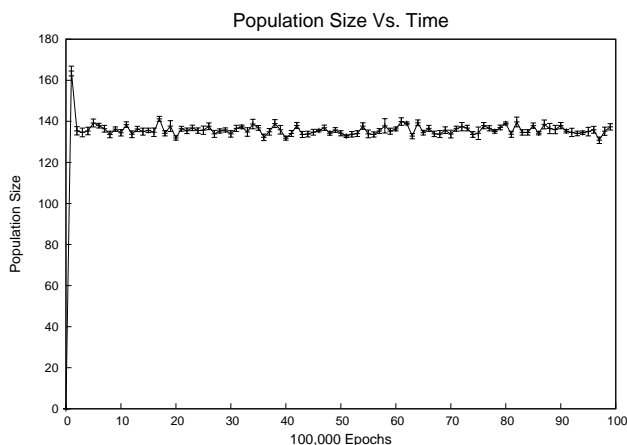

Fig. 3.   Population Size across time for a population of hand–coded, random walk individuals

In Fig. 4 the aggregate behaviors for the population are given. This shows the number of individuals performing each of the three possible behaviors at different epochs. Idling, in which neither running or tumbling is performed, is the most frequently exhibited behavior whilst running and tumbling occur with an approximately equal frequency. For clarity, Fig. 5 shows the instance of running and tumbling only.

The results of Exp. 1 (see Fig. 3) illustrate that a relatively stable population size is obtained with the reported parameters. This shows the emergence of a population capable of being sustained by the amount and period of nutrient deployed under these conditions. The unguided movement in the form of random walk results in a net gain in energy capable of allowing the perpetuation of the population.

Individuals whose path takes them toward a high concentration of nutrient reproduce prolifically as they approach the peak, as do their offspring. This results in a higher density of individuals around the peaks of nutrient deployment. Those
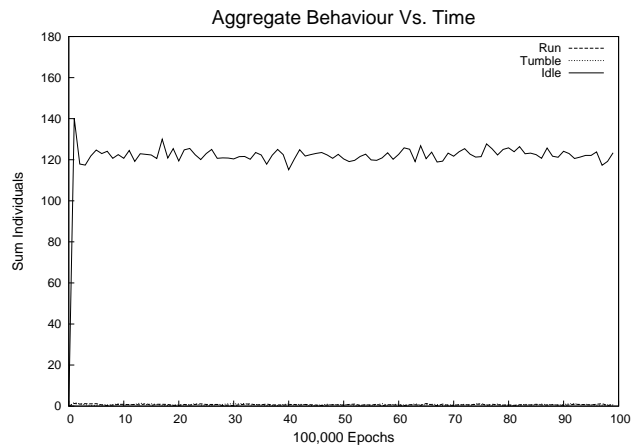

Fig. 4.   Aggregate behaviors across time for a population of hand–coded, random walk individuals
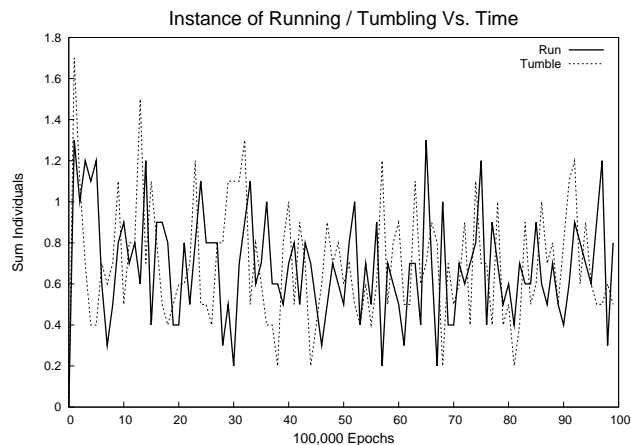

Fig. 5.   Instance of running and tumbling across time for a population of hand–coded, random walk individuals

individuals moving away from the peak simply run out of energy and die.

The population size remains relatively constant over time as a kind of homeostasis emerges. If the population size decreases, the amount of nutrient in the environment increases. With the greater amount of nutrient distributed over a smaller population, more reproduction can occur. The resultant larger population cannot then be sustained by the available nutrient, causing a corresponding decrease in the population size as individuals run out of energy and die.

Fig. 4 shows idling to be the most commonly expressed behavior in Exp. 1. During periods of idling, update of the GRN occurs, however the concentrations of the functional proteins are not sufficient to trigger a behavior. As evolution is switched off in Exp. 1, the ratios of the behaviors remain largely unchanged.

B. Exp. 2: Evolutionary Experiments

According to the experimental results in Exp. 1, a period of 500 is used for Exp. 2. Note that in Exp. 1, all elements of the interaction matrix are set to zero, whereas in this experiment, the interaction matrix is initialized randomly with a probability of 0.3. Owing to this stochastic nature of

initialization, it may happen that the initial population does not contain an adequate number of viable individuals, and therefore terminated prematurely. If this was the case, the run was repeated until a duration of $10,000,000$ epochs was obtained.

The profile of population size when evolution is switched on is given in Fig. 6. Toward the end of the run, a stable population size of approximately 50 is observed. A comparison between the populations of Exp. 1 and Exp. 2 during the initial stages of each experiment is shown in Fig. 9. The increase in size of the evolving population can be seen to occur prior to corresponding increase in the non–evolving population. The maximum and average population size attained by the non–evolving population is notably larger than that of the evolving population. This is somewhat counter–intuitive, as it might be expected that an evolved population would both acquire and expend energy more efficiently.

Fig. 8. Instance of running / tumbling across time for an evolving population of randomly initialized individuals

Fig. 9. Comparison of population size for Exp. 1 and Exp. 2 for first part of the runs
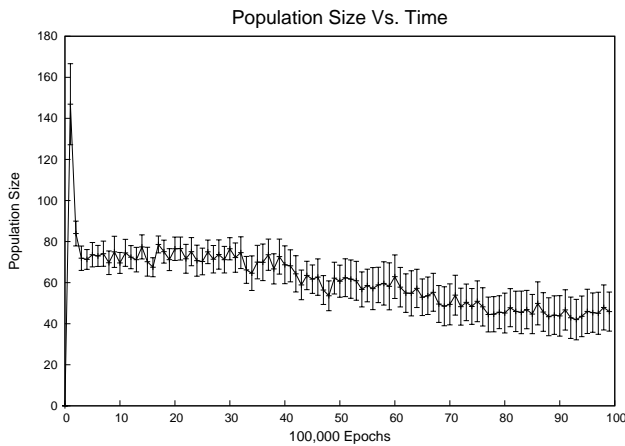

Fig. 6. Population Size across time for an evolving population of randomly initialized individuals
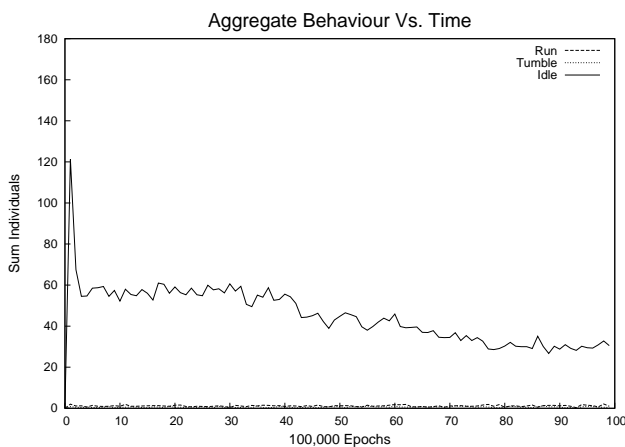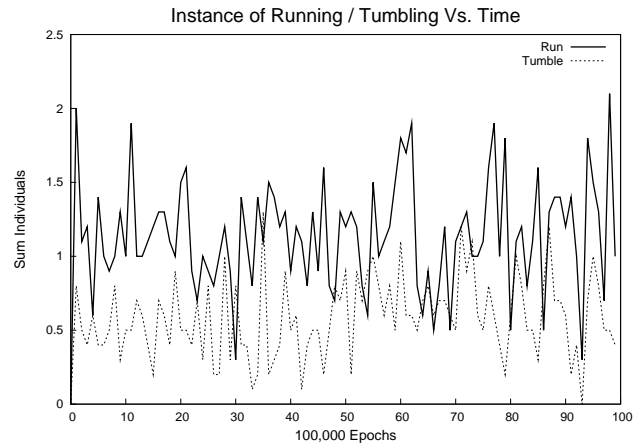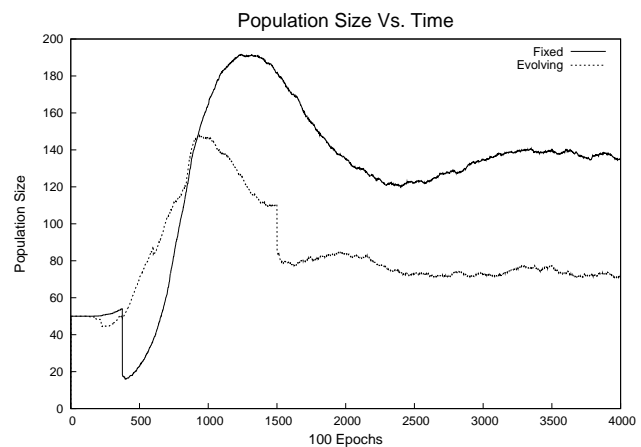

Fig. 7. Aggregate behaviors across time for an evolving population of randomly initialized individuals

To investigate whether chemotaxis evolves, we compare the average number of individuals in the population performing running, tumbling and idling. Fig. 7 shows the aggregate behaviors for the population of Exp. 2. As the run progresses, the relative number of idling individuals decreases. The instance of running and tumbling individuals remains comparatively constant throughout, with running being slightly more frequent than tumbling, as shown in Fig. 8.

In Exp. 2, the ratio of exhibited behaviors can be seen to change over time. The most obvious difference is the decrease in instances of idling. Whilst this is seen to decrease, the absolute number of running and tumbling individuals remains comparatively constant, suggesting the population to evolve toward more frequent movement. Fig. 8 indicates tumbling to occur less frequently than running.

The change in behavior results in extended runs, with two important ramifications. Firstly, the population becomes dispersed over a greater area of the environment. This allows for a comparatively quicker discovery of nutrient, as any given deployment is likely to be in the proximity an individual. Similarly, the greater range of movement of an individual is more likely to bring it in contact with a nutrient deployment – the individual will be 'covering more ground'.

Secondly, the increased movement will lead individuals

away from nutrient peaks. Although offspring are likely to be born close to a nutrient deployment (as a result of the reproduction mechanism), those configured to run disproportionally more than tumble will move away from their locus of creation comparatively more quickly. The effect of this rapid displacement is that individuals will move away from the nutrient concentration, therefore accruing less energy. Although nutrient will be consumed, a greater amount of energy is expended on movement as the individuals travel a greater distance within the environment.

Fig. 9 is helpful in indicating why this situation might arise. At the start of any run, the initial population size is small and therefore unable to cover the environment efficiently. However, during this time, nutrient is deployed regularly and accrues within the environment. Randomly initialized and evolved individuals exhibiting comparatively frequent movement will travel a greater distance, increasing the likelihood of encountering the previously deployed nutrient. Therefore, in the evolving population, the first generations are likely to comprise of comparatively frequently moving individuals.

Based on the results of Exp. 1 and Exp. 2, it is inferred that an overwhelming chemotaxis behavior has not emerged. If individuals were able to follow the increasing nutrient gradient (thus a higher percentage of individuals show a running behavior), more nutrient would be ingested whilst less energy would be expended on traversing areas of lower nutrient concentration. This positive energy deficit would theoretically sustain a larger population for any constant amount of nutrient. In contrast, the population appears to evolve toward a strategy of increased coverage of the environment.

C. Possible Reasons for the Non–Evolution of Chemotaxis

There are several plausible reasons why chemotaxis does not appear to evolve. Firstly, evolution might be unable to 'find' the genetic configuration for chemotaxis. Should this be the case, it is obvious that the evolving system will never produce chemotaxing individuals.

Secondly, it is possible that the selection probability on chemotaxing individuals is not large enough for proliferation. Were this the case, although chemotaxing individuals could arise they might die out before reproducing in sufficient number to infiltrate the population.

It is unlikely that the apparent lack of chemotaxis results from the first scenario. Our experimentation has shown that a single interaction within the interaction matrix is sufficient for chemotaxis (see below). Given the mutation rates, population turnover and initial randomization of the chemical interaction matrix, it is likely that this mutation arises within the course of the runs.

To investigate the second possible reason, experiments were conducted to assess whether chemotaxis would be selected for above simple random walk. An initial population of 50 individuals was created, comprising of 25 random walk individuals and 25 chemotaxing individuals. The chemotaxing individuals were configured such that a single inhibitory connection was present between the external

nutrient signaling chemical and the tumble inducing protein. All forms of mutation were disabled, and the conditions were otherwise identical to that of the previously described experiments.

The composition of the obtaining population over time is shown in Fig. 10. Although the results represent a single simulation, further runs provided qualitatively identical results; the random walk population invariably became extinct. This result suggests that chemotaxis have a selective advantage.
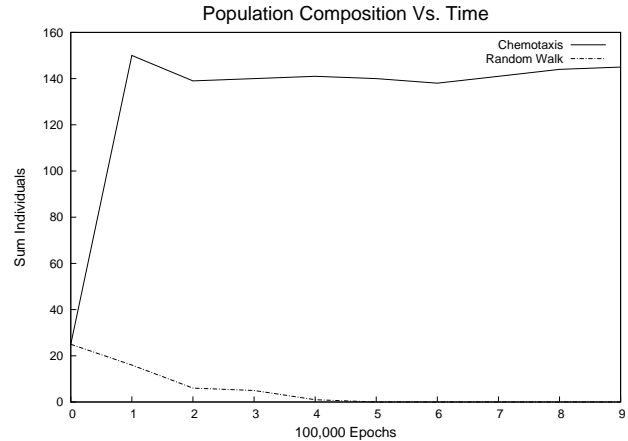

Fig. 10. Selection Pressure

Thirdly, chemotaxis could be a sub–optimal strategy for survival in certain conditions, in favor of some other pattern of movement. For chemotaxis to be advantageous, gradient information is necessary, yet it must be remembered that this information is destroyed as individuals consume nutrient. To investigate the effect of the activity of the population on the nutrient environment, the spatial concentration of nutrient for an experimental run was examined across time.

Figs. 11 and 12 illustrate the nutrient landscape at different epochs. From these illustrations, the ruggedness of the nutrient landscape is revealed. In consideration that the external nutrient sensing system only detects whether an increase or decrease in nutrient is encountered, it is obvious that the gradient information is extremely noisy. This potentially prevents chemotaxing individuals from achieving a significantly greater selection probability than those performing random walk.

Considering the previously discussed situation in which frequently moving individuals rapidly proliferate, it is understandable why chemotaxis would not evolve under these conditions. A population of frequently moving, random walk individuals will cover a comparatively large amount of the environment. The ingestion by this population of any encountered nutrient will rapidly obfuscate any meaningful gradient information. In the lack of gradient information, the ability to chemotax becomes decreasingly beneficial.

V. Conclusions and Future Work

The experimental results show the system to be functioning on many levels. Under the configurations outlined, open ended runs can be sustained, potentially allowing for ongoing
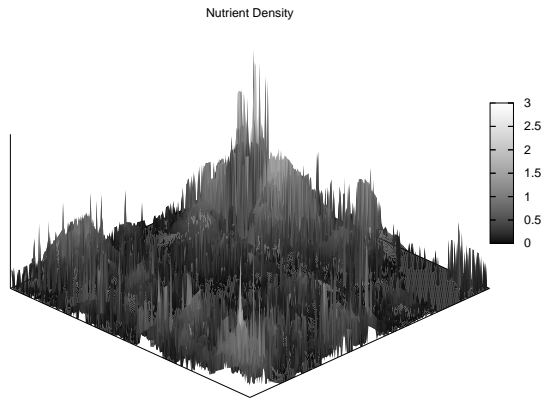
Fig. 11. Nutrient density at epoch 1000. The x,y plane correspond to the virtual world, whilst the z axis indicates the associated nutrient concentration. Only the area in which nutrient deployments are made is shown.
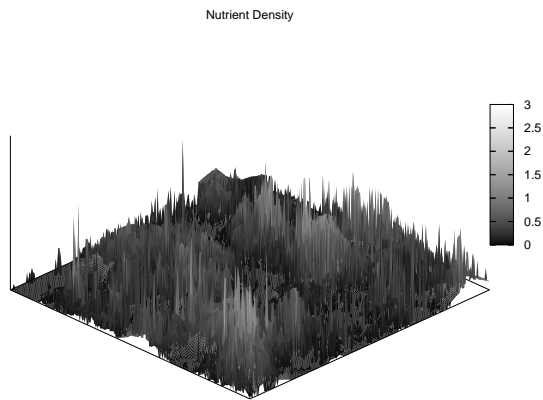

Fig. 12. Nutrient density at epoch 1499. As with Fig. 11, an area of the virtual world is shown.

evolution in future experiments. Unfortunately, evolution does not appear to lead to the emergence of chemotaxis in the present setup. Our further experiment results imply that this may be attributed to the destruction of gradient information by randomly moving individuals. This mismatch between the expected results (the emergence of chemotaxis) and those obtained indicate that developing an evolutionary system in which the fitness function is implicit is nontrivial.

Having considered why chemotaxis does not appear to evolve, the aim of immediate future work is to establish whether the given explanations are accurate. To determine the effect of the destruction of gradient information, the system will be modified to incorporate a spatio–temporal diffusion of nutrient. In addition to providing a scenario more congruent with natural systems, it is anticipated that this will prevent the observed ruggedness, giving individuals with the ability to chemotax a significant advantage.

Secondly, a more rigorous metric for the determination of whether chemotaxis is emerging is necessary. As the single permissible method of gradient tracking in this scenario involves a large stochastic element, it is a nontrivial assess-

ment. One plausible method would involve repeated comparison between the behavior of an evolved (i.e. experimental) individual and a control individual in the form of pure random walk. In simple form, this would require measuring the mean distance between an individual and the peak of a nutrient gradient across time. Given suitably repeated measurement, where chemotaxis is apparent the experimental group should show a significantly lower mean distance from the nutrient peak. Under the settings used within the experiments outlined above, this task would become confounded as a result of the nutrient being consumed by the population. The distribution of nutrient can become rapidly multi–modal, with a large number of local minima. To counter this, it is possible to utilize a single deployment of nutrient, ignoring the effect of consumption on the environment. As the focus of the experiment concerns a single individual, reproduction can be deactivated therefore avoiding a perpetually growing population size.

REFERENCES

[1] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC, July 2006.
[2] C. Vlachos, R. Gregory, R. Paton, J. R. Saunders, and Q. H. Wu, "Individual-based modelling of bacterial ecologies and evolution," *Comparative and Functional Genomics*, vol. 5, pp. 100–104, 2004.
[3] W. Tang, Q. Wu, and J. Saunders, "Individual-based modeling of bacterial foraging with quorum sensing in a time-varying environment," in *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*, ser. LNCS 4447, 2007, pp. 280–290.
[4] R. Gregory, J. R. Saunders, and V. A. Saunders, "The paton individual–based model legacy," *Biosystems*, vol. 85, pp. 46–54, 2006.
[5] R. Gregory, "COSMIC: a model of cellular genetic interaction and evolution," Ph.D. dissertation, University of Liverpool, January 2004.
[6] C. Vlachos, R. C. Paton, J. R. Saunders, and Q. H. Wu, "A rule-based approach to the modelling of bacterial ecosystems." *Biosystems*, vol. 84, no. 1, pp. 49–72, April 2005.
[7] J. Penner, R. Hoar, and C. Jacob, "Bacterial chemotaxis in silico," in *The First Australian Conference on Artificial Life*, Camberra, ACT, Australia, 2003.
[8] ——, "Modelling bacterial signal transduction pathways through evolving artificial chemistries," in *The First Australian Conference on artificial Life*, Camberra, ACT, Australia, 2003.
[9] R. Hoar, J. Penner, and C. Jacob, "Transcription and evolution of a virtual bacteria culture," in *Congress on Evolutionary Computation*. Canberra: IEEE Press, 8–12 December 2003, pp. 54–61.
[10] P. P. Devine and R. C. Paton, "Herby: An evolutionary artificial ecology," in *International Conference on Evolutionary Computation*, 1997, pp. 309–314.
[11] J. U. Kreft, G. Booth, and J. W. Wimpenny, "Bacsim, a simulator for individual-based modelling of bacterial colony growth." *Microbiology*, vol. 144 (Pt 12), pp. 3275–3287, December 1998.
[12] O. Soyer, T. Pfeiffer, and S. Bonhoeffer, "Simulating the evolution of signal transduction pathways." *J Theor Biol*, vol. 241, pp. 223–232, 2006.