Robustness in artificial life

E. Silverman* and T. Ikegami

Department of General Systems Studies,
University of Tokyo,
Komaba, Tokyo 153-8902, Japan
Fax: +81-3-5454-4378
E-mail: erics@sacral.c.u-tokyo.ac.jp
E-mail: ikeg@sacral.c.u-tokyo.ac.jp
*Corresponding author

Abstract: Finding robust explanations of behaviours in Alife and related fields is made difficult by the lack of any formalised definition of robustness. A concerted effort to develop a framework which allows for robust explanations of those behaviours to be developed is needed, as well as a discussion of what constitutes a potentially useful definition for behavioural robustness. To this end, we describe two senses of robustness: robustness in systems; and robustness in explanation. We then propose a framework for developing robust explanations using linked sets of models, and describe a programme of research incorporating both robotics and chemical experiments which is designed to investigate robustness in systems.

Keywords: scientific explanation; robustness analysis; robotics experiments; self-organisation; chemical experiments; robustness; artificial life.

Reference to this paper should be made as follows: Silverman, E. and Ikegami, T. (xxxx) 'Robustness in artificial life', *Int. J. Bio-Inspired Computation*, Vol. X, No. Y, pp.000–000.

Biographical notes: Eric Silverman received his Doctorate in Computer Science from the University of Leeds in the UK in 2007. Following his studies, he was granted a JSPS Postdoctoral Research Fellowship and joined the Ikegami Laboratory at the University of Tokyo. There he worked on robotics experiments investigating behavioural robustness and new definitions of robustness in living systems. He has also performed research on the usage of computer simulations to study the evolution and development of social systems.

Takashi Ikegami received his Doctorate in Physics from the University of Tokyo in 1989. After being granted his doctorate, he began intensive study of self-reproduction, evolutionary theory and game theory. Currently, he is a Professor at the Department of General System Studies, also at the University of Tokyo. His research is centred on complex systems and artificial life, a field which aims to build a theory of life using dynamical systems perspectives. Recently, he has been working in a collaboration focused on chemical experiments involving self-moving oil droplets, and also on using autonomous robots to understand new concepts of biological robustness.

This paper is a revised and expanded version of a paper entitled [title] presented at [name, location and date of conference].

Comment [L1]: Author: Please complete where highlighted.

1 Introduction

Despite our frequent use of the term, the robustness concept as commonly defined in Alife and related disciplines is no longer adequate for producing real insight into the functions of biological life. Robustness in one methodology or virtual world does not imply robustness in another, and likewise does not imply that we can develop a robust explanation of the behaviour of interest. In order to move in this direction, we must develop a stronger understanding of the relationship between the differing senses of robustness – and through that understanding, formulate a new means of generating robust explanation using simulation and robotics.

Of course, within the field of artificial life, the potential background on which theories can be constructed is enormous. Alife began with the aim of investigating the larger questions surrounding life and its origins: how do populations develop and evolve? How might life have originally begun? Might we be able to investigate new forms of life through computer simulation, a sort of 'life-as-it-could-be' (Langton, 1992)?

Such an undertaking is far from straightforward. Beyond the fact that computer simulations are still a relatively recent development in science, our biological understanding of the origins and evolution of life is still quite tenuous. There is a pronounced lack of high-level overarching theories which can unify the drive in Alife to understand these larger questions about biological life. Without such a background, Alife researchers face the daunting prospect of developing a new method for examining the robustness of their simulations or robots.

2 The senses of robustness

Before we can propose a research programme designed to investigate robustness in Alife (Ikegami and Suzuki, 2008; Ikegami, 2009), we must develop a firmer definition of and relation between two senses of robustness: robustness in systems, and robustness in explanation. Each sense carries with it a set of foundational concepts which can guide us toward an understanding of how these senses can be related.

When discussing systems, robustness is often described as a property which gives the system a certain resilience against perturbation. A robust system is thus able to retain functionality despite variation. This concept is further elucidated in the following section.

In contrast, a robust explanation is a scientific explanation which can identify causal factors that underlie a phenomenon in a variety of circumstances. Thus, rather than an explanation which pertains to only once instance of behaviour in a system, a robust explanation will identify those elements which drive a system's behaviour as a whole. We examine this concept further in Section 4.

3 Robustness in systems

Jen (2005) presented a discussion of stability and robustness in her paper, 'Stable or robust? What's the difference?'. She argues that formalising the differences between these two properties is far from simple, given that robustness has never been explicitly defined, but rather is subject to multiple conflicting interpretations.

Jen attempts to address this problem by discussing the generally agreed-upon differences between stability and robustness, before progressing into a detailed discussion of the nature of robustness itself. She argues that robustness in general is broader in scope that stability, encompassing a larger range of systems and features of systems. She posits a definition of robustness as "a measure of feature persistence for systems or for features of systems, that are difficult to quantify or parameterise (i.e., to describe the dependence on quantitative variables), and with which it is therefore difficult to associate a metric or norm" (p.13).

Beyond this, however, she also characterises robustness as a measure of persistence against perturbation in systems in which those perturbations encompass not just fluctuations in inputs or parameters, but 'represent changes in system composition, system topology, or in the fundamental assumptions regarding the environment in which the system operates'. This differs substantially from stability analysis, in which one most often postulates a single perturbation. Robustness, in contrast, often requires the examination of many possible perturbations. Thus, robustness may exist or not exist on different levels of a system: the individuals in an Alife simulation may display robustness, for example, but that does not imply that the population as a whole is robust.

3.1 Extending Jen's definition

Kitano (2007) extends Jen's definition of robustness, noting that the particular separation of stability and robustness she proposes allows a system to display one property while lacking the other. In other words, a system could be both robust and unstable: elements of the system which is robust could be individually unstable, but still provide the system with robustness against perturbations. Similarly, systems could also be stable and not robust: single perturbations may not disturb the system, while multiple perturbations may overwhelm it.

Kitano's view leads us to the need for a strong distinction between stability and robustness within a system. If we imagine a system performing some behavioural task, we can imagine the elements of that system operating at a functional level to produce the desired behaviour. If we take Kitano's view into account, then we can argue that stability thus operates on this functional level – the sub-tasks and related elements of behaviour that lead to successful operation at the task level.

4 Robustness in explanations: Levins and robustness analysis

To address the second sense of robustness, robustness in explanations, we will examine the concept of robustness analysis and how it may apply to studies in Alife. Robustness analysis as a concept originated in the seminal paper 'The strategy of model-building in population biology' (Levins, 1966). This paper has been frequently cited and justifiably lauded for its insightful commentary on the pragmatic issues facing modellers in biological disciplines. However, his discussion of robustness analysis has been comparatively ignored, despite its apparent applicability to these difficulties in evaluating robustness in Alife.

Levins' conception of robustness analysis hinges upon the concept of studying multiple models of the same phenomenon. In his view it is important that each model is distinct from the other, containing different core assumptions or methodologies. As a consequence, he argues that "If these models, despite their different assumptions, lead to similar results, we have what we can call a robust theorem that is relatively free of the details of the model.... our truth is the intersection of independent lies" (p.20).

Thus, the modeller needs to be able to separate predictions which are relevant to the system of interest from those which are artefacts of the assumptions made in the construction of these distinct models. By assembling the predictions of these models together, and finding commonalities amongst them, the theorist attempts to find this 'intersection of independent lies' which is a robust theorem regarding the behaviour or system of interest.

4.1 Objections to robustness analysis, and Weisberg's clarification

Orzack and Sober (1993) launched a critique of Levins' paper, arguing that Levins' conception of robustness analysis is non-empirical and thus fundamentally flawed. The procedure outlined by Levins centres on the examination of models, not data, and this fact leads Orzack and Sober to conclude that the whole enterprise would be ineffective. In essence, they argue that such a procedure is only guaranteed to produce a correct robust theorem if the theorist knows ahead of time that one of the models in the set to be examined is true.

Their objections at first seem quite reasonable; after all, how could the modeller separate bonafide predictions from mere artefacts of their simplifying assumptions without this sort of foreknowledge? As Weisberg (2005) explains, this would further imply that if the models to be examined were indeed true, then robustness analysis becomes entirely unnecessary anyway; Levins proposes the method as a means for examining models which are highly idealised and thus not necessarily true or accurate.

Weisberg then provides an intensive examination of robustness analysis using predator-prey models and the Volterra principle as an example of a robust property, or a property common to multiple models which contain different idealising assumptions. This leads to a discussion of the need to find common structures between models: those structures found in different models which give rise to the robust property. Finally, he elucidates a four-step procedure for robustness analysis:

1 examine a group of similar but distinct models for a robust behaviour

2 find the core model structures which give rise to this robust property

3 interpret these common structures as descriptions of empirical phenomena

4 construct a robust theorem, and use stability analysis to examine the boundaries of that robust theorem and the behaviour it describes.

However, these four steps are not the end of the robustness analysis procedure. The theorist must also confirm that these common structures found in these disparate models are instantiated in the system of interest, and examine the effects of other possible causal factors on the function of that structure. Thus, we must discover if the common structure is present in the real system, and whether it is truly the primary causal element in the original behaviour of interest.

4.2 Problems with the Weisberg formulation

Thus far, Weisberg's reformulation of Levins' original idea seems quite fruitful. Weisberg expands the original simple and abstract concept into a pragmatic framework for the modeller and theorist, presenting a plausible means by which multiple models can contribute to an understanding of the real-world structures that produce a given behaviour. He further toughens his burgeoning framework against criticisms akin to those levelled at Levins by Orzack and Sober (1993).

However, the framework as described implies a higher level of general consensus amongst model-builders than seems to be present in Alife. While modellers in Alife may share certain simplifying assumptions when making models of similar phenomena, the methodologies used to create these models can vary enormously, ranging from cellular automata to evolutionary simulations to robotics.

Within his formulation, Weisberg specifically describes these sought-after common structures as mathematical constructs. While he makes allowances for common structures which may not fall into such clearly-defined territory, and thus must be divined through the theoretical expertise of the examiner of the collection of models, the advantage here clearly lies with the theorist that can construct a mathematical formulation of the common structures they see within that collection.

Say for example that our theorist is examining a collection of models which examine the evolution of parasitism. Models of such a phenomenon within Alife are incredibly diverse even when examining related phenomena, as we see in Tom Ray's Tierra and its parasites and hyper-parasites (Ray, 1994) as compared to Kaneko and Ikegami's homeochaotic examination of host-parasite dynamics (Kaneko and Ikegami, 1992). When examining models which contain not only different assumptions, but different means of constructing the various worlds these simulated systems inhabit, can we realistically follow Weisberg's plan for robustness analysis?

5 Weisbergian robustness analysis and Alife

In order to escape this conundrum, we need a unified framework under which to search for common structures in order to perform robustness analysis. Models in Alife can frequently share a conceptual relationship – they examine similar behaviours within biological systems, but using fundamentally different methods. The way forward, then, is to create experiments and simulations which share common grounding and related contexts.

Robust theories, as we have discussed here, require a framework of models to develop. The models within that framework must share a common background of assumptions and the same context. Our studies of robust systems reveal the importance of the relationship between context and environment. Further study of this relationship can allow us to develop a common background of assumptions on which to construct our framework of similar, but distinct models – each with differing simplifying assumptions, rather than different methods altogether.

6 Experimental insights on environment and robustness

Having discussed robustness in explanation, we return once more to robustness in systems – this time examining the results of a related experiment. The experiment detailed below has demonstrated the strong, highly complex links between environment and the generation of robust behaviour in a system. In particular, this novel chemical experiment (Hanczyc et al., 2007) provides an illustration of how real robust systems are a consequence of a balance between multiple causal factors – a balance that would not be revealed by a single perturbation of that system.

In this experiment, the following procedure was performed: adding oleic anhydride oil phase to highly alkaline water phase (in between pH 11 and 12) to see how the hydrolysis of the anhydride proceeds in a glass plate. Immediately the oil begins to react with the water, causing the oil phase to break up into smaller spherical droplets, ranging a few to several hundred microns in diameter. These droplets are like gliders in Conway's Game of Life, in that they move freely in space and interact with each other. In Figure 1, a snapshot of the initial reaction phase is presented. The wavy form shown is the interface which appears between the oil phase and the aqueous phase. After a few minutes, droplets begin to appear and start to move around. In some conditions, we see a population of droplets begin to cross the glass plate (see Figure 2).

We argue that the mechanism of the movement is caused by the coupling of the hydrolysis reaction at the interface with the fluid dynamics of the droplet. The surface tension of the droplet is determined by the amount of chemical product (oleic acid). Inhomogenity of chemical distribution on the surface causes the Marangoni force, which results in a convection flow (which goes from the leading to tailing portion along the inner surface of a droplet, then goes straight back to the leading portion along the axis through the centre). This flow transports fresh chemicals to the leading portion, thereby sustaining the reaction. Because of this coupling, the droplet can enhance its self-motility and prolong its non-equilibrium state, and thus its lifetime. This scenario has been largely confirmed with numerical simulations (Matsuno et al., 2007)

Figure 1 The hydrolysis of the oleic anhydride oil (top and centre) in the presence of alkaline water

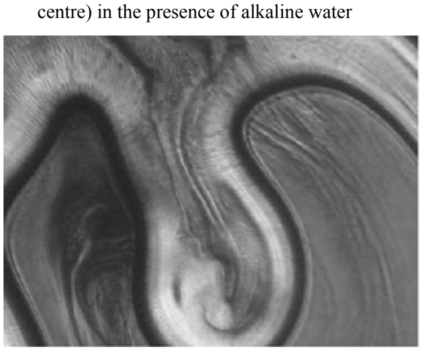
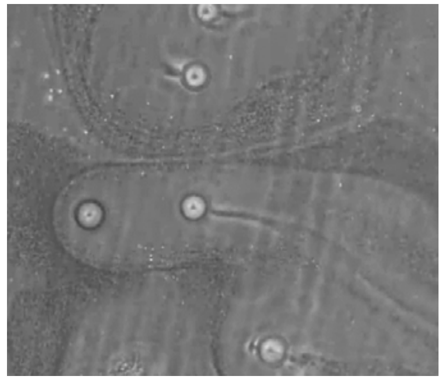


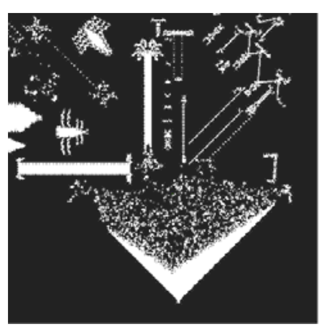
Figure 2 Population of droplets that self assembled after partial hydrolysis of the initial oil mass was added to the system, Figure 1



Note: Each droplet is several hundred μ metres in size.

The droplets observed here can change direction spontaneously, and when coming into contact they never fuse together. When gliders in the Game of Life collide with each other, however, they normally disappear entirely except for a few very special cases. The oil droplets are attracting or repulsing each other due to the flow patterns exhibited and product secreted around them. Therefore, the interaction range is quite wide in the case of these droplets. On the other hand, the interaction range of the Game of Life remains at one bit even if the organised pattern becomes huge. Also, a single bit flip in a Game of Life pattern usually causes fatal damage to those complex patterns (see Figure 3). From these facts, we observe that the oil droplets are far more robust than gliders in the Game of Life. However in terms of lifespan, gliders in the Game of Life are immortal, while the oil droplets have finite life spans of less than 10–15 minutes (though it is difficult to tell when the droplet dies) and are sensitive to factors in the external environment such as pH.

Figure 3 Break up of spaceship pattern in the Game of Life by flipping an arbitrary bit of the big triangular pattern in the figure



The lifespan of these droplets can be made much longer by using a mixture of reactive oleic anhydride and non-reactive nitrobenzene as a carrier oil. This results in droplets with controlled volume that are more amenable to analysis.

However, in the first experiment, the environmental conditions (such as pH, product concentration and Reynolds number) are self-organised by the system itself rather than being prepared by the experimenter. The self-moving behaviour of the self-organised droplets is different from that of the controlled droplets. In particular, the size of the self-organised droplets is suppressed under a few hundred micrometres, due to the moving mechanism being dominated by the convection flow. Further, the interaction between droplets is mediated by the chemical product around the droplets, but in the case of controlled droplets, these interactions appear more like ballistic collision. However this needs to be investigated further.

Comparing these droplet behaviours and the gliders of the Game of Life, we notice that both self-organisation and a rich complex initial state are required for biological robustness. Self-organisation tends to simplify the final outcome, limiting it to a low degree of complexity, while this low complexity assures the robustness of the outcome. The rich and complex initial state prevents the system from falling into a simple state which displays no interesting behaviours. Ikegami and Hanczyc (2009) has dubbed this balancing act between these two factors the maximalism design principle for biological robustness.

7 Quantifying robustness using robotics experiments

The above serves as an example of a novel study which provides some insight into factors that contribute to the robustness of a system. In particular, we notice the importance of complex interaction with the environment and the self-organisation of this chemical process, which can lead to the later development of robust behaviours. However, we still require a research framework which will allow us to produce more robust explanations of these sorts of behaviours. While our discussion of robustness analysis has illuminated some possibilities for developing studies which produce robust explanations, how might such a comprehensive approach to modelling assist our efforts to study robust systems?

Hubert et al. (2009) provides one example, using robots and simulation in a combined approach to study robustness in a simple system. While robotics has been used frequently to study elements of complex biological systems, these studies mention robustness only as a property of the system under study rather than as the central concern. This leads to experimenters simply perturbing their system and declaring it robust when the system remains functional. As we have seen, however, developing an understanding of how the system achieves that robustness is a highly complex task.

This study aims to examine these questions by studying a simple system which contains several interacting sub-systems. Hubert et al. use the Lego Mindstorms NXT (see Figure 4), an modular robotics system based upon the NXT microcontroller. In this case, the robot was given the task of navigating through a simple environment to reach a goal, and attempted to complete this task in several conditions using two different sensory modalities (visual and auditory). The goal is to develop a means to quantify robustness through studying the contribution of sub-systems (in this case, visual and auditory sensory systems) to the overall robustness of a system.

Figure 4 The Lego Mindstorms NXT

7.1 Robot experiment overview

Our experiments were performed in three stages. In the first stage, the robot used only the light source to reach the goal (L). In the second stage, it used only the sound source (S). The last stage allows the robot to use both sources to locate the goal (LS). Each of these stages required specific implementations of the robot's controller that are detailed below.

The performance of the robot is measured by its capacity to reach the goal in less than six minutes. For a single trial, the performance of a robot was one if it reached the goal and zero otherwise. The duration of the trial was not considered in the performance measure. We relate this measure of performance to the robustness of the robot and will use both terms to describe its capacity to reach the goal.

For the S and LS conditions, different levels of noise were added to the sound sensor in order to evaluate how the perturbation of one modality impacts the overall performance of the system. The noise was added by adding a value drawn from a uniform distribution to every measure reported by the sound sensor. The values for the noise presented in the results were always positive and represent the maximum MAX added. The range of those values is [–MAX; MAX].

7.2 Robot platform

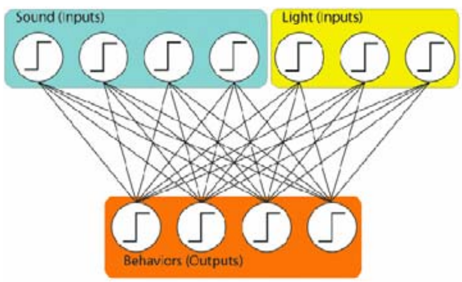
Our robotic platform is the Lego Mindstorms NXT (see Figure 4), which is a modular robot assembled from many elements such as motors, sensors and structural modules. The onboard processor is called the NXT and is a general purpose microprocessor whose function is to command the motors, retrieve information from the sensors and run custom software. The setup used in our experiments consisted of two motors, one light sensor and one sound sensor. The two latter sensors measure the intensity of light and the volume of sound respectively.

Our experiments were performed on the real platform using a speaker as a sound source and a fluorescent lamp as a light source. We also examined the robot performance in simulation, where we used real sensor readings from the robot to make it consistent with the real experiment.

7.3 The controller

The controller of the robot is a feedforward artificial neural network (NN) which possesses four or seven inputs based on the number of sources it must track and four outputs. No hidden neurons are present. The input and output neurons are fully interconnected, as seen in Figure 5, and the weights are tuned using Hebbian learning (Hebb, 1949).

Figure 5 The robot's NN controller (see online version for colours)



The inputs of the NN were pre-processed in order to obtain binary inputs. The pre-processing was necessary for the Hebbian learning to be stable and is different for each type of source in the environment. The outputs were squashed to a range of [0; 1] using a sigmoid function.

Before explaining the pre-processing, it should be noted that the robot was using two timescales to accomplish its task. The first one is the microprocessor timescale (MT) which corresponds to one step of the sensory-motor loop. One MT time-step involves one update of the sensors and of the motors. The second timescale is the neural timescale(NT) which corresponds to an update of the outputs of the NN.

In this NN, four inputs are for sound and three inputs are for light source processing. The current sensory inputs for sound and light are compared with the previous sensory inputs and the background value in order to let the robot achieve the goal reaching behaviour. The parameters used for this pre-processing were determined by hand. Further details of this pre-processing can be found in the Appendix.

The weights connecting the inputs to the outputs were tuned through Hebbian learning. Hebbian learning is an unsupervised learning algorithm which relies on correlations between inputs and outputs to decide if their connecting weights should be increased or not. There are many different implementations of Hebbian learning with different capabilities. The one we chose in our experiments is Oja's rule (Oja, 1982). This rule implements regular Hebbian learning while stabilising the growth of the weights.

Due to the unsupervised nature of Hebbian learning, we cannot expect the NN to converge to the desired behaviour without guidance. Hebbian learning will increase the value of a weight when the input and output are simultaneously activated, but in addition we must ensure that the network learns to demonstrate the appropriate behaviour for a given set of inputs. Given that the robot may use light, sound or both for navigation depending on the experimental condition, the appropriate behaviour for a given set of inputs can vary in each condition. Thus, prior to applying the Hebbian learning algorithm, we train the network to reflect the correct behaviour for the possible inputs in each experimental condition.

7.4 Robotics experiment results

After programming the robot, its behaviour was examined by alternating the environmental conditions as follows: an environment with only a light source (L condition); an environment with only a sound source (S condition); and an environment with both sources present (LS condition). The robot was then placed in different initial starting points, and its ability to reach the sound or light sources was evaluated over 1,000 trials while uniform sonic noise was added to its environment.

The results indicate that there is a certain amount of noise above which the LS condition always shows better performance than the S condition (but not the L condition). The performance shown in the LS condition is not a simple superposition of the L or S only condition, however. Above this noise threshold, a compensatory behaviour emerges: the robot utilises its ability to navigate in the visual modality to compensate for the added noise in the auditory environment.

We hypothesise from these observations that a certain amount of noise, whether it is from an internal or external source, encourages the robot to develop cross-modal performance as a means to navigate the environment (in this case light and sound modalities). This also implies that the concept of robustness is different from that of stability as we described in Section 3. The stability of a system will always degrade with the introduction of noise, but noise can prompt a robust system to become more robust (in this case, a robot can acquire cross-modal behaviours).

8 Building a research programme to study robustness

While this initial foray into studying robustness using robotics is a useful step forward on its own, the most important property of this study in the current analysis is the experimental design: a system with a complex initial state (a robot containing a NN controller and several sensors) which possesses several clearly-defined sub-systems, each of which can be studied both in a real environment and in simulation.

The modularity of this system also allows it to be expanded simply and quickly, and the resultant data can be shared between both the robot and the corresponding simulation. This research programme is moving toward the way forward we propose based upon Weisberg (2005): a framework of models with a shared background of assumptions. Upon this framework we can build a set of similar, but distinct models – perhaps utilising varied environments, subjecting the system to different forms of perturbation, or altering the construction of the robot or its controller – which will allow us to discern which common elements between those models contribute to the development of robust behaviours within the overall system.

Further, this study serves to link the insights gained through the oil droplet experiments to a new potential means to examine the components of a robust behaviour. The oil droplets provided an example of how the environment can allow a system to self-organise, which can allow the system to begin to develop complex and robust behaviours. The robotics study starts with an un-structured internal organisation (a robot with a NN controller with random parameters) then allows self-organisation to proceed; the robot is placed into the environment, and examining the resultant emergent behaviours allow us to probe how robust behaviours may develop and function. With this comprehensive, combined approach to the study of robustness, utilising several methodologies based upon linked base assumptions regarding robustness and self-organisation, we hope to further unravel the complexities of robustness.

9 Conclusions

Most studies of robustness to date have used the term only as a general descriptor for systems which are able to retain functionality despite perturbation. While Jen (2005) and Kitano (2007) have managed to develop more insightful *definitions* of robustness in complex systems, we still lack a robust *explanation* for how robust behaviours develop and function.

Our discussion of robustness analysis has illuminated the importance of developing a comprehensive research programme to develop such explanations. The need to combine similar but distinct models to discover the causal factors in systems that lead to robust behaviour has been discussed in great detail by Levins (1966), Orzack and Sober (1993) and Weisberg (2005). Our goal here has been to demonstrate that, even in the highly diverse field of artificial life, such an approach may still flourish.

In our case, a novel biochemical experiment together with simulation and robotics approaches are being used to develop an in-depth understanding of robustness and how we may quantify and examine its effects. The oil droplets have demonstrated the potential impact of the maximal design principle, which underlines the importance of 'half-way design' (of the initial states and architecture of a system) and letting a system self-organise in interacting with an environment, which can later lead to robust behaviours. Using a robotic platform, we are continuing to study further potential aspects of robustness by developing a series of robotic models which will allow us to probe the impact of various sub-systems on the robustness of the overall system.

In the case of the chemical experiment, the environmental conditions, pH and oleate concentration are controlled by the droplet motion. We define robustness of the droplets with respect to their ability to sustain self-moving behaviour. In contrast, gliders in the Game of Life appear to display self-moving behaviour but do not actually function in this way. This evolution of self-movement, autonomy and individuality appears to be a key prerequisite for developing robust behaviours.

When giving a robot the capacity for self-organisation, however, the robot cannot sustain its autonomous movement quite so easily. Developing robustness in this case appears to depend on the development of an appropriate use of time-scales for its behaviour; in particular, parameter settings for the robot's learning and forgetting during the process of Hebbian learning can affect the time-scales of the robot's behaviours. Finding a range of these parameters which allow proper functioning by utilising the background noise of the environment will allow the development of more robust behaviours.

As we move forward, our goal is to develop a strong definition of robustness: one which allows us to comprehend how it functions, and to quantify it in a variety of contexts. By increasing our understanding of how we can connect artificial systems with natural environments, we can further our development of a theoretical framework which provides a background of assumptions to inform our robotic and simulated models. Only via such a combined approach can we develop a robust explanation for robust behaviours in natural and artificial systems.

Acknowledgements

Eric Silverman wishes to thank the Japan Society for the Promotion of Science (JSPS) for supporting this work with a Postdoctoral Research Fellowship. We also thank our collaborators Martin Hanczyc, Naoto Horibe, Taro Toyota and Tadashi Sugawara for the chemical droplet experiment, and Julien Hubert and Eiko Matsuda for the robot experiments. This work has been partially supported by a Grant-in-Aid for Scientific Research on Priority Areas 'Emergence of adaptive motor function through interaction between body, brain and environment' from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

Hanczyc, M., Toyota, T., Ikegami, T., Packard, N. and Sugawara, T. (2007) 'Chemistry at the oil-water interface: self-propelled oil droplets', *Journal of the American Chemical Society*, Vol. 129, No. 30, pp.9386–9391.

Hebb, D. (1949) *The Organization of Behavior*, Wiley, New York.

Hubert, J., Matsuda, E., Silverman, E. and Ikegami, T. (2009) 'A robotic approach to understanding robustness', in *Proceedings of Mobiligence 2009*, Awaji, Japan.

Ikegami, T. (2009) 'Rehabilitating biology as a natural history', *Adaptive Behavior*, Vol. 17, pp.325–328.

Ikegami, T. and Hanczyc, M. (2009) 'The search for a first cell under the maximalism design principle', *Journal of Technoetic Arts*, Vol. 7, No. 2, pp.153–164.

Ikegami, T. and Suzuki, K. (2008) 'From homeostatic to homeodynamic self', *BioSystems*, Vol. 91, pp.388–400.

Jen, E. (2005) 'Stable or robust? What's the difference?', in *Robust Design: A Repertoire of Biological, Ecological and Engineering Case Studies*, Oxford University Press.

Kaneko, K. and Ikegami, T. (1992) 'Homeochaos: dynamical stability of symbiotic network with population dynamics and evolving mutation rates', *Physica D*, Vol. 56, pp.406–429.

Kitano, H. (2007) 'Towards a theory of biological robustness', *Molecular Systems Biology*, Vol. 3, p.137.

Langton, C. (1992) 'Life at the edge of chaos', in C.G. Langton, C. Taylor, J.D. Farmer and S. Rasmussen (Eds.): *Artificial Life II*, pp.41–91.

Levins, R. (1966) 'The strategy of model-building in population biology', *American Scientist*, Vol. 54, pp.421–431.

Matsuno, H., Hanczyc, M. and Ikegami, T. (2007) 'Self-maintained movements of droplets within convection flow', *LNAI*, Vol. 4828, pp.179–188.

Oja, E. (1982) 'Simplified neuron model as a principal component analyzer', *Journal of Mathematical Biology*, Vol. 15, No. 3, pp.267–273.

Orzack, S. and Sober, E. (1993) 'A critical assessment of Levins' 'The Strategy of Model Building (1966)', *Quarterly Review of Biology*, Vol. 68, pp.534–546.

Ray, T. (1994) 'An evolutionary approach to synthetic biology: Zen and the art of creating life', *Artificial Life*, Vol. 1, pp.179–209.

Weisberg, M. (2005) 'Robustness analysis', *Philosophy of Science*, Vol. 73, pp.730–742.

Appendix

The pre-processing applied on the inputs for the sound seeking task is as follows:

1 Input 0 is set to 1 if the current sound volume is higher than the volume measured 30 MT timesteps ago to which is added a small value of 0.03.

2 Input 1 is set to 1 if the current sound volume is lower than the volume measured 30 MT timesteps ago from which is subtracted a small value of 0.03.

3 Input 2 is set to 1 if the current sound volume is greater than or equal to the volume measured 30 MT timesteps ago.

4 Input 3 is set to 1 if none of the other inputs is activated. The controller goes through this list until one input is activated. The remaining ones are set to zero.

For the light seeking behaviour, the same system of rules is used but it is necessary to add an additional variable allowing the controller to distinguish between the ambient light in the room and the light coming from the goal. This is not necessary in the case of the sound as the room is quiet during the experiments. This memory, referred to as imprint, is updated every 120 MT timesteps and contains the intensity of the light at the time of its update. Every subsequent reading of the sensor is offset by this value. The following list details how the inputs are updated:

1 input 2 is set to 1 if the intensity of the light is lower than a threshold set to 0.01

2 if the current intensity is lower than the intensity 10 MT timesteps ago minus a small value of 0.01, then there are two choices:

   a input 1 is set to 1 if the robot goes backward

   b input 2 is set to 1 otherwise

3 if the current intensity fits in none of the above, input 1 is set to 1 if the robot goes backward and input 0 is set to 1 otherwise.

The need to test for the direction of the robot arises from the unidirectionality of the light sensor which only picks up light when facing the source directly. In that sense, going backward is not necessary but can nevertheless happen in the early stages of the learning process. Because of that possibility, it is necessary to allow the robot to reverse its direction. The memory of 10 MT timesteps used for the light differs from the 30 MT timesteps of the sound. Both values have been determined experimentally in order to improve the performance in a real world environment.

The sound being noisier, a value of 30 MT timesteps is necessary to ensure a correct evaluation of the tendency of the robot to approach it. The light shows less noise and only requires 10 MT timesteps.

When the task combines light and sound, the robot uses a controller with seven input neurons to allow it to combine both behaviours. In this condition the inputs are set according to the above algorithms. This means that at each NT timestep two inputs will be activated simultaneously: one for the sound and one for the light. The NN always has four outputs regardless of the task. These represent the four possible behaviours that can be activated by the robot. To determine which behaviour is activated, a winner-takes-all strategy is used and the output with the highest activation wins. The four behaviours are:

1 Output 0 maintains the current behaviour.

2 Output 1 inverts the current behaviour. If the robot is going forward, it will go backward at the next MT timestep and vice-versa.

3 Output 2 modifies the current behaviour to create a left turn while maintaining the same direction.

4 Output 3 modifies the current behaviour to create a right turn while inverting the current direction.