Adaptive Behavioural Modulation and Hysteresis in an Analogue of a Kite Control Task

Allister Furey, Inman Harvey

Centre for Computational Neuroscience and Robotics
University of Sussex, Brighton, United Kingdom
a.d.j.furey@sussex.ac.uk, inmanh@sussex.ac.uk

Abstract. We define a simplified analogue of a kite control task that requires, in its simplest form, a situated artificial agent to switch between two mutually exclusive behaviours. In more complex versions of the task, the agent is required to adapt to changes within its environment that occur on different temporal scales. We describe the failure to evolve successful agents when a decision threshold is defined artificially and conversely the evolution of successful agents when they themselves are allowed to determine their own threshold through interaction with the environment. Agents are demonstrated capable of adapting both their switching behaviour and spatial domain according to environmental changes on three temporal scales, on the fastest of which, the agents behave in an opportunistic manner.

1. Introduction

There are novel wind energy systems in development using tethered airfoils – kites, as the wind intercepting element. The kites require active control that ideally modulates the flight behaviour according to environmental conditions, often to maintain the operation of the system within certain hard limits of viability. Additionally, the most efficient currently posited system configuration requires alternation between two discrete modes of operation. The point of switching between these two, mutually exclusive actions or 'behaviours' could be determined by a simple but rigid threshold or heuristic that is predetermined. However, it is likely to be beneficial to have an adaptive strategy that modulates behavioural hysteresis actively through some mechanism that takes into account the current state of the agent and its environment.

Here we present a simplified analogue or 'toy version' of the kite control task where an environmental resource with spatially heterogeneous distribution is exploited by an evolved agent. The task is designed so as to mirror key aspects of the kite control problem at a 'higher' level i.e. that of behavioural modulation and/or hysteresis without the complication and heavy computational burden of simulating the dynamics of kite flight, treated elsewhere [2,3]. The task is designed to be simple enough to allow evolution to proceed rapidly whilst still retaining a dynamical embodied interaction between the agent and its environment. Assessing the adaptive abilities of the evolved agents may indicate the potential of evolutionary robotics for generating kite control behaviours. Additionally, determining the best way in which to evolve adaptive behaviour in this simple analogue task may inform future work in generating similar behaviours in real kite control.

2. Background and Motivation

The use of power kites for production of renewable energy from the wind is highly attractive due to the prospect of accessing the stronger and more consistent winds at altitudes above that of conventional wind turbines. This is especially so given that power generated increases with the cube of the wind velocity [4,6], meaning large gains in power output per unit area of the wind intercepting element can be realised. The lack of requirement for expensive and visually intrusive towers and much reduced territorial occupation [4,5] compounds the potential socio-economic advantages of such systems.

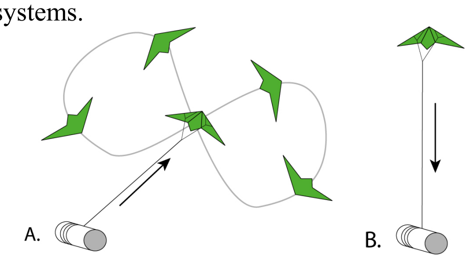

Fig. 1. Generation (reel out) phase A and retraction (reel in) phase B.

Most current concepts (for review see [5]) for implementing kite energy operate with two functional phases as per Fig. 1. In the first (A), the kite is steered actively through the airspace in order to increase the apparent wind speed across the kite. The kite therefore pulls its lines out from a reel at ground level which is coupled to a dynamo, thus generating electricity. We have demonstrated previously in simulation [2,3] that an evolutionary robotics approach can be used to produce neural network controllers that both find an optimal flight trajectory for phase A and maintain it in the face of significant environmental perturbations. (see Fig. 2, right). In the second phase (B), the lines are retracted with the kite controlled so as to reduce resistance, the process is therefore cyclical and sustainable indefinitely whilst the wind allows. As long as A produces more energy than consumed in B then there is a net gain of energy. The kite control problem has some further key characteristics, firstly, the environment is highly dynamic, both the wind speed and direction can change rapidly. Secondly there are some hard constraints within which the agent must operate in order to keep the kite aloft and not overpower the hardware, as discussed below. Given these characteristics a high quality controller will have to perform a) in-behaviour adaptive modification in order to maximise productivity whilst b) meeting the viability limits imposed by the operational constraints.

Our exercise here of building a task analogous to the kite control problem is intended primarily to be informative as to the validity of an evolutionary robotics (ER) approach to tackling its behavioural aspects and to gain insight into how it might best be implemented. With this analogous task, without necessarily expecting direct crossapplicability, we aim to address two components of the kite control problem that may benefit from an adaptive controller. The first issue is one of switching between the two separate behaviours, that of generation and retraction. There is some time or performance cost to the switching from generation to retraction, given that the kite should be ideally be steered to zenith or the side of the wind window in order to minimise both the force and the time required to retract the lines. A heuristic of switching at predetermined points at some sensible intermediate values has been employed previously in the literature [4] constituting one potential strategy. Another possibility is much more frequent alternation, whereby retraction occurs within a given portion of every figure eight trajectory as posited by Houska [6]. The latter strategy avoids some of the problems of the former one, namely a reduction in yield at lower altitudes, and that of potentially reaching a hard line length limit and resulting wasted generation time or hardware damage. However, by not placing the kite at zenith or windows edge, greater energy needs to be expended to retract the kite and the momentum of the kite is reduced. We intend to approach this trade-off with our toy problem using standard ER techniques, and the as yet unaddressed question of whether to modify the switching points according to environmental conditions, a question closely coupled to the second issue.

This second issue is specifically related to the spatial distribution of the wind resource in relation to the kite tether point. As mentioned above, in normal wind conditions, the kite is steered in a looping pattern (see Fig. 2, right) directly downwind of the tether point (the ++ and + area in Fig. 2, left), in order to maximise the aerodynamic forces away from the reel and therefore the energy that is generated. However, in strong wind conditions, this strategy will generate forces that may damage the generating hardware, the kite, or its lines.

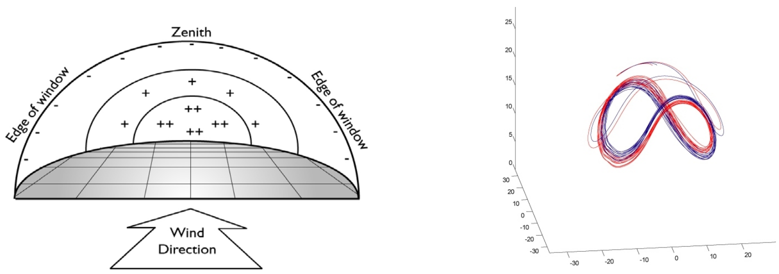

Fig. 2. (Left) The wind window concept. The most powerful portion of the window is signified by ++, less powerful by + and the least powerful by -. (Right) A plot of the trajectory of the kite wingtips in a 42 second trial as controlled an evolved neural network, starting from zenith.

This overpowered situation can be prevented by simply avoiding the most powerful region of flight space. As the wind speed increases, the extent of the avoided region should increase. Conversely with a weak wind resource the agent must confine itself to the most powerful region of flight space. These changes constitute short term, reactive modifications to the kite's spatial domain of activity and are the second level of adaptation that we shall address in the model. As we detail below, we intend to explore how changes in the intensity of the light source on different temporal scales affect the agent and its behaviour.

3. Methodology

3.1 Agent and task design

We have designed a task that captures some of the key properties of the kite generation strategy problem without the computational burden of modelling the aerodynamics of the kite. Our task is based around a simple 2 wheeled, 2 sensor agent (see Fig. 3, left) situated within an unbounded 2d area. The analogy for the flying lines is a battery, charged by an orientation independent 'solar panel' on the agent. As with the lines, there are upper and lower hard limits to the battery level. As success is judged by quantity of energy passed to the battery, the agent can score more highly by repetitively switching between charging and discharging the battery, corresponding to the reeling out and in of the lines respectively.
The agent's sensors are mounted 90 degrees apart on its perimeter, each has a 180 degree range and provides a signal describing its local light intensity. Within the simulated environment there is a point source light which constitutes an energy resource which the agent must exploit, the equivalent of the wind in the kite control problem. As the intensity of the light drops away exponentially with distance from the source, there is light intensity gradient which mirrors the wind window of the kite problem in that the energy resource is distributed spatially. This radial distribution regionalises the arena into 3 key areas as per Fig. 3 (right).

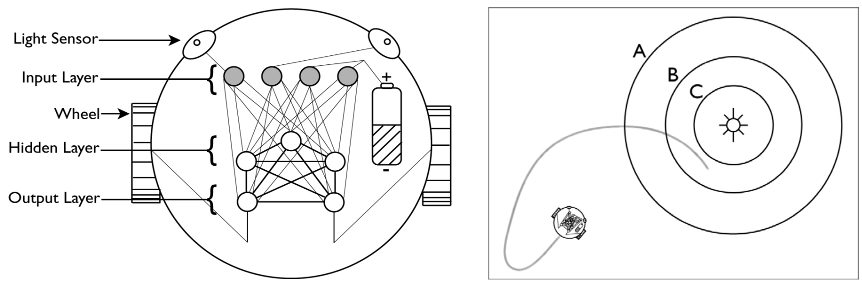

Fig. 3. (Left) A diagram of the agent and its neurocontroller architecture, light grey neural connections are feed forward, bold black neural connections are reciprocal with asymmetric weights. The battery connects to two input neurons, one connection carrying battery level data and the other, rate of change of battery level. (Right) Diagram of a portion of the experimental area, light source marked by the sun icon. Between line C and the sun marks the overpowered region, between B and C, the region where the charging rate plateaus. A marks the line at which the switching threshold is met.

As would be expected in a kite hardware implementation, and in common with conventional wind turbines, we implement both a plateau of the generation/charge rate at some moderate to high light level (line B in Fig. 3, right), and a 'cut out' at a higher level still, where the environmental power source, if it were coupled to the generator, would overload and damage it. To mimic the reduction of line tension by spooling out the lines with the generator uncoupled, when a threshold point is reached corresponding to entering the area between line C in Fig. 3 (right), and the light, the

battery continues to increase but the agent does not accumulate fitness until it leaves the region. If the threshold is exceeded when the battery is full, the agent is considered to be broken and the trial is ended prematurely, implications for the agents fitness as detailed below. Similarly, if the agents battery reaches zero, the agent has failed to collect sufficient energy from the environment and the trial is also ended.
In order to restrict the scope of the problem for ease of analysis and potentially increased evolution speed, we first simplify the problem of making some trade off between time cost of switching and energy required to switch by defining a lower force threshold at which switching automatically occurs. Once the agent crosses line A, at which point the battery charge rate input to the neural network will become negative, it commits to switch behaviour although the 'decision' to do so may have been made at some earlier point. As the boundary is determined by an environmental threshold independent of the agents activity, we term this regime 'threshold switching'. We subsequently assess a variation in which the rules are unchanged, except that the agent has some control over its own switching point. This is implemented implicitly by introducing a battery cost to movement, if the agent stays still, it will charge until full at any point in the environment at a rate determined by distance to the light, it is movement at a speed which consumes more energy than is recouped from the environment that will drive down the battery. The agent in this scenario is forced to make a more complex, but more realistic trade off in that moving far from the high-quality resource will speed up discharging, but mean that for portions of the return journey, time is spent near energy neutral and at suboptimal charging speeds. We term this regime 'enactive switching'.

3.2 Neural network and agent dynamics

The neural network that drives the agents behaviour is a small continuous time recurrent neural network (CTRNN) of 4 input neurons, 3 hidden layer neurons and 2 output layer neurons, with both inhibitory and excitatory connections permissible. Each of the four input layer neurons receives input constituting one of the following four data values; left light sensor value, right light sensor value, proportion of battery capacity remaining, battery charge rate. Each input neuron connects to every hidden and output layer neuron, but receives no other input (see Fig. 3, left). The hidden and output layer neurons are fully interconnected, connection weights are asymmetric. A single neuron derives its dynamics from Eq.1:

$$\tau_j \dot{a}_j = -a_j + \sum w_{ij}\sigma(o_i - \theta_i) \qquad \textbf{(1)}$$

θ is the bias term, w the weight and σ is the sigmoid function, τ_j is the neurons time constant, and o_i the old activation values from the previous timestep. In the simulation neuron activation values are integrated using Euler integration with a timestep of 0.1. The sigmoided values of the two output neurons form the motor activation values for the left and right wheels respectively. Average wheel speed determines the agent's velocity and difference in wheel speed determines the agents angle, dependent on the agents diameter. The agents position is also determined by Euler integration with a step size of 0.1. As the motivation of this work is to answer the questions whether ER can as tool generate the kind of switching behaviour required for kite control devices and how best to implement the ER strategy, we do not intend to analyse the network dynamics that result from the evolved weights.

3.3 Genetic Algorithm and trial configuration

We use a simple microbial genetic algorithm [1], with a population of genotypes or artificial 'DNA', randomly initialised, specifying possible parameter values. Pairs of neurocontroller parameter sets taken from the population of 20 have their performance evaluated. The winner overwrites the loser with small additive random mutations ±0.05 added to a random 50% subset of its parameter values. This process continues for several hundred iterations, each consisting of 10 competitive trials. For all tasks, fitness is judged by the amount that the battery is charged, over the time course of the trial. As the battery cannot be further charged once a hard upper limit is reached, and given that the trial length is set between 30 and 300 seconds such that the battery could be charged and discharged fully at least twice, the agent is effectively rewarded for switching between charging and discharging behaviour.

We perform three sets of experiments, the common elements to which are as follows: for each of the five trials that constitute one evaluation, the agent is reinitialised at a new random orientation, in a random location within a square of side 30 units. Its battery level is randomised between a minimum of 20 units and its maximum value of 60 units. Competing pairs share the same random initialisations for fair comparison.

In the first set of experiments, we evolve the agent (1000 generations) to maximise charging when the light intensity value of 5 remains virtually constant between experiments (+- noise of 0.2). In the second we use a successful population from the first experiment as a seed for evolution (500 generations) during which we vary the intensity of the light from evaluation to evaluation within a range of 1 to 10. In the third, (500 generations) the intensity of the light is varied within each experiment, with either continual, linear slow increases/decreases that run the whole length of the trial, constituting a gradual trend, or immediate discontinuous jumps between the original and higher or lower intensity values, constituting 'gusts' and 'lulls'. In this final experiment, we only judge the agent on its fitness in the second half of the experiment, the agent is therefore given an opportunity to have adapted to the changes in the environment after a period of unassessed interaction.

4. Results

The first and most salient result was the poor performance of the agents evolved under the threshold switching regime. Although some dozens of evolutionary runs were completed, most resulted in the agent approaching the light and either stopping in that position or moving in a small stereotypic trajectory near the light. This was also true when the population was seeded with successful phototactic agents. In doing so they successfully completed one behaviour but were unable to switch to another. This incomplete strategy consisted of approaching the light and then successfully stopping in the plateau zone of the resource, but never moving from that position. Due to the failure on this simple task, threshold switching was not carried on to either of the other two tasks.

In contrast, the controllers evolved with ‘enactive’ switching evolved a variety of switching strategies all of which involved repetitively moving closer and further from the light. Interestingly, even though these agents were not evolved under the rules of threshold selection, they performed equally as well and with the same apparent strategy when operating under those rules. They seemed to be capable of generalizing from the smooth gradient of switching that they were evolved under to a sharp threshold, and are therefore surprisingly independent of the need to control their rate of discharge, once the behaviour is fully evolved. All the subsequent plots are of the ‘enactive switching’ evolved agents, but operating under threshold selection rules for clarity of interpretation. With the exception of one of several alternative suboptimal solutions shown in Fig. 4 (left), all plots are of the best performing agent from the best evolutionary run. After evolution, populations tended to be highly converged, with each agent from a given population pursuing near-identical strategies that were highly robust to changes in agent starting location or orientation and light dynamics, at least within ranges experienced in evolution.

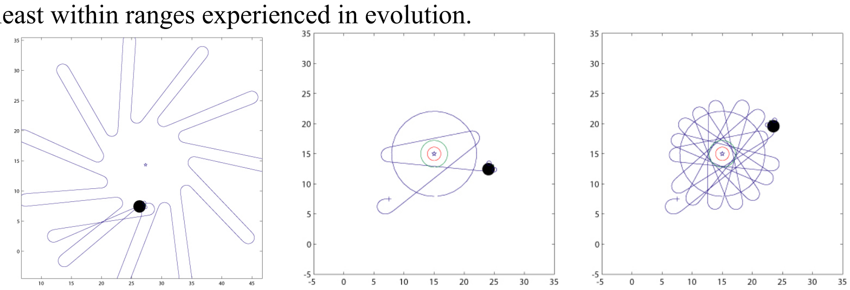

Fig. 4. An interesting but sub-optimal strategy is shown (left). The first few passes of the trajectory of the best performing agent with light source intensity of 5 are shown (centre) as well as the trajectory at the end of the experiment (right).

As the centre and right plots of Fig. 4 demonstrate, the agents skim the edge or just enter the middle ring that marks the region where charging rate plateaus. They then continue the same trajectory until they make an abrupt turn once in the discharging portion of the arena.

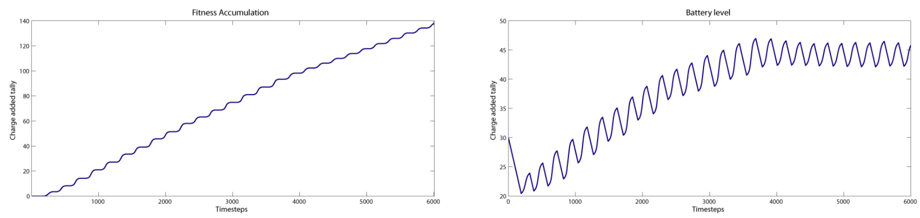

Fig. 5. Plots demonstrating the accumulation of fitness over the experiment shown in Fig. 4 (left) and the cycling of battery level in the same experiment (right).

Although it is not immediately apparent from Fig. 4, Fig. 5 (right) shows how this cycle initially is skewed in favour of charging, resulting in a gradual accumulation of battery which then settles out and oscillates round some intermediate capacity value (c. 75% of capacity). After a further period of evolution with the light intensity varying between experiments by up to a factor of 10, the controllers were able to adapt well to changes in the light intensity. This seemed to be achieved by modulation of the turning angle on exit from the charging area, however as Fig. 6 (left) shows, the agent enters the charging zone at the same angle. Changing the discharge rate had a similar effect with wide turns keeping the agent in the discharging area longer when the discharge rate was lowered (Fig. 6, centre) and tight turns occurring when the discharge rate is elevated (Fig. 6, right)

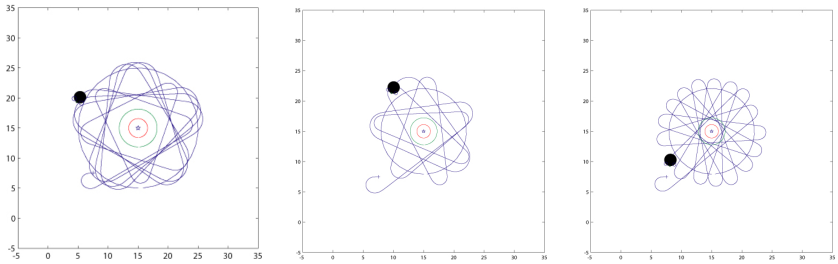

Fig. 6. Adaptation to an increased light intensity of 10 (left), lowered (centre) and elevated (right) discharge rates.

Finally, we compared the performance of the intensity-adaptive agents from the prior experiment with that of agents subject to further evolution in which the light intensity changed within the experiment on one of two temporal scales. When slow linear changes of intensity occurred, the performance of both sets of agents was essentially equal, all agents continuously modulating their trajectories as per Fig. 7 (left).

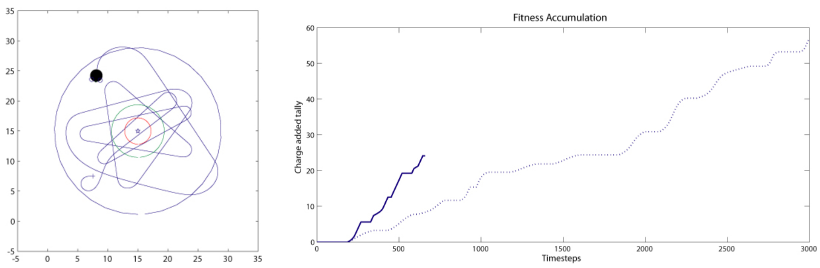

Fig. 7. (left) The modulation of trajectories with continuous and gradual increase in light intensity (1-20). The right plot compares the fitness of the agents with (dotted line) and without (solid line) additional evolution with dynamic light strength when exposed to rapid perturbations in light intensity.

However when the intensity changes were rapid, gust-like perturbations, there was a marked improvement in the population that was further evolved (see Fig. 7 right), which accumulates fitness more slowly than the agents naïve to intraexperimental changes in light intensity, but this trade-off allows it to survive indefinitely in the highly dynamic environment. Fig. 8 illustrates that the strategy of the non-naïve agents in response to gust-like perturbations in light intensity is an opportunistic one, in which during a gust the agent is always charging and immediately after a gust, always discharging. Additionally Fig. 8 seems to imply that the agent takes the strategy of charging at an appropriate rate for its battery level, presumably by regulating its turns to lead it closer or further from the source according to its battery state as it approaches the turn.

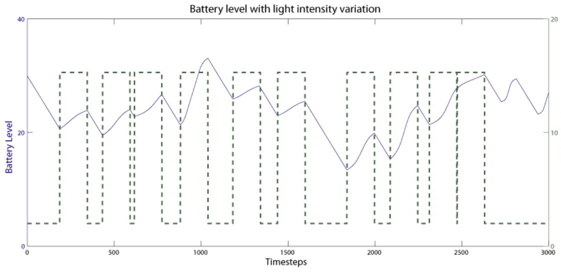

Fig. 8. A dual plot showing the change in battery level over time (solid line) concurrently with the sudden changes in the light intensity (dotted line)

At the start, when its battery is relatively high, it takes relatively small quantities of the available increased resource within gusts, then after much charge is lost during the lull that begins at c.1600 timesteps, much more energy is extracted from the subsequent three gusts. When the variation stops at the end of the trial, the agent reverts to it's default behaviour. The fact that neither the charge speed nor the reversion to oscillatory behaviour seems consistent is likely due to the agents natural in/out oscillations and the delay inherent in turning where the agent is slowed due to the forward operation of only one wheel.

5. Discussion

Perhaps the most pertinent lesson of this exercise for future applications of evolutionary robotics for actual adaptive kite control was the relative success of evolution using the threshold and enactive rules sets for switching between charging and discharging behaviours. Imposing a sharp, externally-defined boundary between two behaviours prevented evolution of successful agents, contrary to their enactive peers which, through their own interactions with the environment determined their own behavioural boundary or lack of it. Given this experience, it would certainly seem prudent to assess an similar strategy for evolving reel-in and reel-out behaviours for kite control, especially given that a sharp boundary could be subsequently imposed on the enactive switching agents in the toy case with little impact on performance, which may be a necessity due to hardware constraints in kite energy systems.

The strategy of the most successful neurocontrollers as presented in Figures 4 through 8 was an elegant one that defied our prior expectations of how the problem would be solved. Instead of heading directly for the energy source and then doubling back for discharging, the supposedly discrete behaviours were subsumed into one action constituting a single pass. The ‘decision’ is then reduced to when, and at what angle to turn in order to start the next charge/discharge pair whose duration and outcome, in a static environment at least, is essentially predetermined. The ability of the evolved neurocontrollers to adapt to changes in light intensity and indeed discharging rates (Fig. 6) both between and within (Figs. 7,8) experiments dispelled our early concerns as to the fragility of this strategy. Given that here, supposed problems such as a trade off between persistence and dithering almost appear to be resolved almost by the environment itself, this type of behaviour seems best described as an ongoing interaction between the agent and its environment and less well described as one of discrete decisions at the top of a hierarchy being passed down to subordinate effector systems. This may be a product of the relative simplicity of the task and the lack of constraints on the network connectivity, indeed the imposition of symmetry or other neuroanatomically inspired constraints would be an interesting comparison, and could potentially generate more transparent internal dynamics.

The final experiment in which neurocontrollers were further evolved with exposure to dynamical environments produced the unexpected result of two co-existing strategies. The standard light passing behaviour switched seamlessly into one in which the environment was permitted to almost completely dominate whether the agent charged or discharged. This apparently submissive behaviour was much more robust than that of the naïve agents, with the agent exerting a more subtle control by increasing or decreasing its exposure to resource variations in order to regulate its state to a level that was notably more conservative than in the prior experiments (50% vs 75%). A thorough analysis of this final behaviour in response to a range of gust intensities and durations and indeed more realistic variations in intensity would be enlightening and is the subject of ongoing work.

In summary, this work has succeeded in its brief of informing future work in the control of kite energy systems at a behavioural level, specifically by suggesting that highly adaptive spatially embedded, behaviour-switching agents can be evolved if decision boundaries are determined by the agents themselves. The opportunistic behaviour seen in the most complex task bodes well for the application of ER to the development of kite control behaviours.

6. References

1. Harvey, I. (2001). Artificial Evolution: a Continuing SAGA, In *Evolutionary Robotics: From Intelligent Robots to Artificial Life*, T. Gomi (ed.) Springer-Verlag.
2. Furey, A.D.J. and Harvey, I. (2007). Evolution of Neural Networks for Active Control of Tethered Airfoils. *Proc. 9th European Conference on Artificial Life*, pages 746-756.
3. Furey, A.D.J. and Harvey, I. (2008). Robust adaptive control for kite wind energy using evolutionary robotics. *Proc. Biological Approaches for Engineering*, (in Press).
4. Canale M., Fagiano L., Ippolito M., Milanese M. (2006). Control of tethered airfoils for a new class of wind energy generator. *Proc. 45th IEEE Conference on Decision and Control*, 4020-2046.
5. Lansdorp B. and Ockels W.J. (2005). Comparison of concepts for high-altitude wind energy generation with ground based generator. *Proc. 2nd China IREETEC*, 409-417.
6. Houska B. (2007). Robustness and stability optimisation of open-loop controlled power generating kites. *PhD Thesis*, University of Heidelberg.