How Co-Evolution can Enhance the Adaptive Power of Artificial Evolution: Implications for Evolutionary Robotics

Stefano Nolfi[1] Dario Floreano[2]

[1]Institute of Psychology, National Research Council
Viale Marx 15, Roma, Italy
stefano@kant.irmkant.rm.cnr.it
[2]LAMI - Laboratory of Microcomputing
Swiss Federal Institute of Technology
EPFL, Lausanne, Switzerland
Dario.Floreano@epfl.ch

Abstract. Co-evolution (i.e. the evolution of two or more competing populations with coupled fitness) has several interesting features that may potentially enhance the power of adaptation of artificial evolution. In particular, as discussed by Dawkins and Krebs [2], competing populations may reciprocally drive one another to increasing levels of complexity by producing an evolutionary "arms race". In this paper we will investigate the role of co-evolution in the context of evolutionary robotics. In particular, we will try to understand in what conditions co-evolution can lead to "arms races" in which two populations reciprocally drive one another to increasing levels of complexity.

1. Introduction

Co-evolution (i.e. the evolution of two or more competing populations with coupled fitness) has several interesting features that may potentially enhance the adaptation power of artificial evolution. First, because the performance of the individual in a population depends also on the individual strategies of the other population which vary during the evolutionary process, the ability for which individuals are selected is more general (i.e. it has to cope with a variety of different cases) than in the case of an evolutionary process in which co-evolution is not involved. The generality of the selection criterion is a very important property because the more general the criterion, the larger the number of ways of satisfying it (at least partially) and the greater the probability that better and better solutions will be found by the evolutionary process.

Consider for example the well-studied case of two co-evolving populations of predators and prey [1]. If we ask the evolutionary process to catch one individual prey we may easily fail. In fact, if the prey is very efficient, the probability that an individual with a randomly generated genotype may be able to catch it is very low. As a consequence, all individuals will be scored with the same null value and the selective process cannot operate. On the contrary, if we ask the evolutionary process to find a predator able to catch a variety of different preys, it is much more probable that it will find an individual in the initial generations able to catch at least one of them and then select better and better individuals until one predator able to catch the original individual prey is selected.

Secondly, competing co-evolutionary systems are appealing because the ever-changing fitness landscape, due to changes in the co-evolving species is potentially useful in preventing stagnation in local minima. From this point of view, co-evolution may have consequences similar to evolving a single population in an ever-changing environment. Indeed the environment changes continuously given the fact that the co-evolving species is part of the environment of each evolving population.

Finally, the co-evolution of competing populations may produce increasingly complex evolving challenges. As discussed by Dawkins and Krebs [2] competing populations may reciprocally drive one another to increasing levels of complexity by producing an evolutionary "arms race". Let us again consider the predator and prey case: the success of predators implies a failure of the prey and conversely, when preys evolve to overcome the predators they also create a new challenge for them. Similarly, when the predators overcome the new preys by adapting to them, they create a new challenge for the preys. Clearly the continuation of this process may produce an ever-greater level of complexity. As Rosin and Belew [3] point out, it is like producing a *pedagogical* series of challenges that gradually increase the complexity of the corresponding solutions.

This nice property overcomes the problem that if we ask evolution to find a solution to a complex task we have a high probability of failure while if we ask evolution to find a solution first to a simple task and then for progressively more complex cases, we are more likely to succeed. Consider the predators and preys case again. At the beginning of the evolutionary process, the predator should be able to catch its prey which have a very simple behavior and are therefore easy to catch, likewise, prey should be able to escape simple predators. However, later on, both populations and their evolving challenges will become progressively more and more complex. Therefore, even if the selection criterion remains the same, the adaptation task will become progressively more complex and more general.

Unfortunately however a continuous increase in complexity is not guaranteed. In fact, co-evolving populations may cycle between alternative class of strategies that although they do not produce advantages in the long run may produce a temporary improvement over the co-evolving population. Imagine, for example, that in a particular moment population A adopts the strategy A_1 which gives population A an advantage over population B which adopts strategy B_1. Imagine now that there is a strategy B_2 (similar to B_1) that gives population B an advantage over strategy A_1. Population B will easily find and adopt strategy B_2. Imagine now that there is a strategy A_2 (similar to A_1) that provides an adaptive advantage over strategy B_2. Population A will easily find and adopt strategy A_2. Finally imagine that previously discovered strategy B_1 provides an advantage over strategy A_2. Population B will come back to strategy B_1. At this point also population A will come back to strategy A_1 (because, as explained above, it is effective against strategy B_1) and the cycle of the same strategies will be repeated over and over again (Fig. 1).

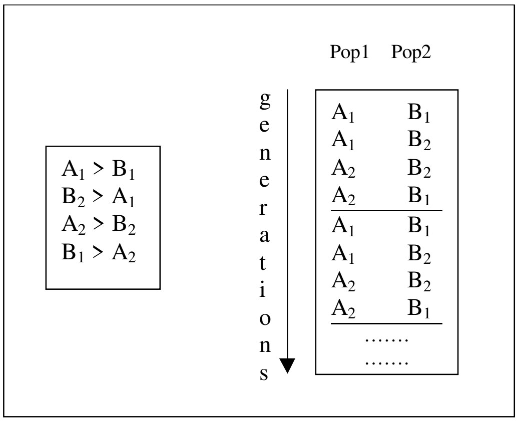


Fig. 1. The same strategies (A_1 and A_2 in population A) and (B_1 and B_2 in population B) may be selected over and over again throughout generations as is shown in the right hand side of the figure if the interaction between them looks like what is represented on the left side of the Figure. In this case the repeated cycle corresponds to 4 different combinations of strategies

Notice how the cycling may involve two or more different strategies for each population but also two or more different groups of strategies.

Of course this type of phenomena may cancel out all the previously described advantages because, although evolution may never get stuck in a particular solution, the number of different solutions discovered might be quite limited. Moreover there is no need to discover progressively more complex strategies. It is sufficient to re-discover previously selected strategies that can be obtained with a limited number of changes.

In this paper we will investigate the role of co-evolution in the context of evolutionary robotics. In particular, we will try to understand in what conditions co-evolution can lead to "arm races" in which two populations reciprocally drive one another to increasing levels of complexity.

2. Co-Evolving Predator and Prey Robots

Several researchers have investigated co-evolution in the context of predators and prey in simulation [1, 4, 5, 6]. More recently, we have tried to investigate this framework first by using realistic simulations based on the Khepera [7, 8] and subsequently the real robots [9].

In this section, we will first describe our experimental framework and the results obtained in a simple case. Then, we will describe other two experimental conditions more suitable to the emergence of 'arm races' between the two competing populations.

2.1 The Experimental Framework

As often happens, predators and prey belong to different species with different sensory and motor characteristics. Thus, we employed two Khepera robots, one of which (the Predator) was equipped with a vision module while the other (the Prey) had a maximum available speed set to twice that of the predator. The prey has a black protuberance, which can be detected by the predator everywhere in the environment. The two species could evolve in a square arena 47 x 47 cm in size with high white walls so that predator could always see the prey (within the visual angle) as a black spot on a white background (see Fig. 2).

Both individuals were provided with eight infrared proximity sensors (six on the front side and two on the back) which had a maximum detection range of 3-4 cm in our environment. For the predator we considered the K213 module of Khepera which is an additional turret that can be plugged in directly on top of the basic platform. It consists of a 1D-array of 64 photoreceptors which provide a linear image composed of 64 pixels of 256 gray-levels each, subtending a view-angle of 36°. However the K231 module also allows detection of the position in the image corresponding to the pixel with minimal intensity. We exploited this facility by dividing the visual field into five sectors of about 5° each corresponding to five simulated photoreceptors. If the pixel with minimal intensity lay inside the first sector, then the first simulated photoreceptor would become active; if the pixel lay inside the second sector, then the second photoreceptor would become active, etc. From the motor point of view, we set the maximum wheel speed in each direction to 80mm/s for the predator and 160mm/s for the prey.

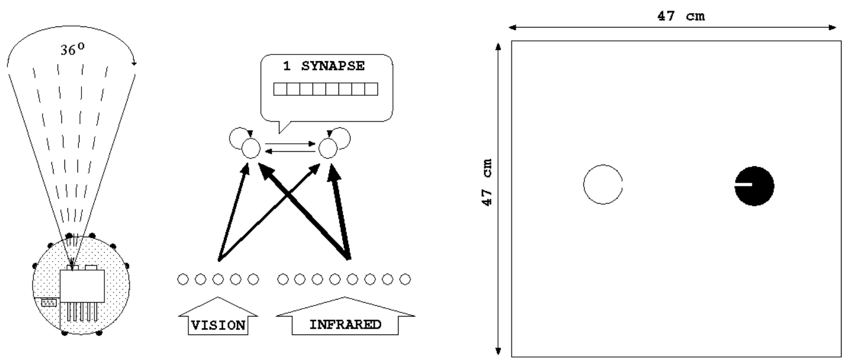


Fig. 2. Left and center: details of simulation of vision, of neural network architecture, and of genetic encoding. The prey differs from the predator in that it does not have 5 input units for vision. Eight bits code each synapse in the network. Right: Initial starting position for prey (left, empty disk with small opening corresponding to frontal direction) and predator (right, back disk with line corresponding to frontal direction) in the arena. For each competition, the initial orientation is random

In line with some of our previous work [10], the robot controller was a simple perceptron comprising two sigmoid units with recurrent connection at the output layer. The activation of each output unit was used to update the speed value of the corresponding wheel every 100ms. In the case of the predator, each output unit received connections from five photoreceptors and from eight infrared proximity sensors. In the case of the prey, each output unit received input only from 8 infrared proximity sensors, but its activation value was multiplied by 2 before setting the wheel speed. This structure, which is well-suited for the evolution of Braitenberg-like obstacle avoidance, was chosen as being a minimally sufficient architecture to evolve something interesting while maintaining system complexity at a manageable level; for the same reason, the architecture was kept fixed, and only synaptic strengths and output units threshold values were evolved.

In order to keep things as simple as possible and given the small size of the parameter set, we used direct genetic encoding [11]: each parameter (including recurrent connections and threshold values of output units) was encoded using 8 bits. Therefore, the genotype of the predator was 8 x (30 synapses + 2 thresholds) bits long while that of prey was 8 x (20 synapses + 2 thresholds) bits long.

Two populations of 100 individuals were each co-evolved for 100 generations. Each individual was tested against the best competitors of the previous generations (a similar procedure was used in [6, 12]). In order to improve co-evolutionary stability, each individual was tested against the best competitors of the ten previous generations (on this point see also below). At generation 0, competitors were randomly chosen within the same generation, whereas in the other 9 initial generations they were randomly chosen from the pool of available best individuals of previous generations.

For each competition, the prey and the predator were always positioned on a horizontal line in the middle of the environment at a distance corresponding to half the environment width, but always at a new random orientation. The competition ended either when the predator touched the prey or after 500 motor updates (corresponding to 50 seconds at maximum on the physical robot). The fitness function for each competition was simply 1 for the predator and 0 for the prey if the predator was able to catch the prey and, conversely 0 for the predator and 1 for the prey if the latter was able to escape the predator. Individuals were ranked after fitness performance in descending order and the best 20 were allowed to reproduce by generating 5 offspring each. Random mutation (bit substitution) was applied to each bit with a constant of probability pm=0.02[1].

For each set of experiments we ran 10 replications starting with different randomly assigned genotypes.

In this paper we will refer to data obtained in simulation. A simulator developed and extensively tested on Khepera by some of us [13] was used. However some of the experiments described have also been successfully replicated on real [9].

---

[1] The parameters used in the simulations described in this paper are mostly the same as in the simulation described in [7]. However, in these experiments we used a simpler fitness formula (a binary value instead of a continuous value proportional to the time necessary for the predator to catch the prey). Moreover, to keep the number of parameters as small as possible, we did not use crossover. In the previous experiments, in fact, we did not notice any significant difference in experiments conducted with different crossover rates.

2.2 Measuring Adaptive Progress in Co-Evolving Populations

In the co-evolutionary case, the Red Queen effect [14] makes it hard to monitor progress by taking measures of the fitness throughout generations. In fact, because fitnesses are defined relative to a co-evolving set of traits in the other individuals, the fitness landscapes for the co-evolving individuals vary. As a consequence, for instance, periods of stasis in the fitness value of the two populations may correspond to a period of tightly-coupled co-evolution.

To avoid this problem, different measure techniques have been proposed. Cliff and Miller [15] have devised a way of monitoring fitness performance by testing the performance of the best individual in each generation against all the best competing ancestors which they call CIAO data (Current Individual vs. Ancestral Opponents).

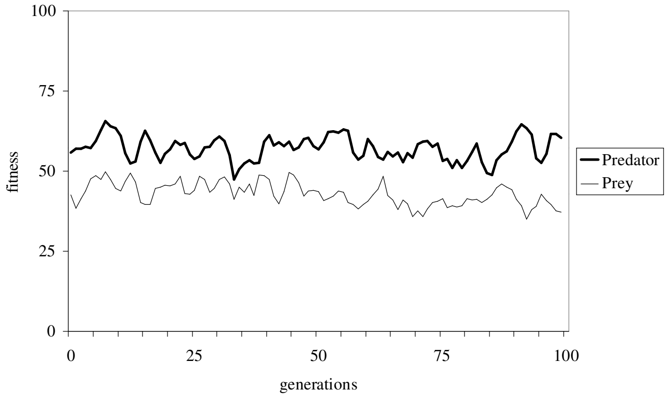
A variant of this measure technique has been proposed by some of us and has been called Master Tournament [7]. It consists in testing the performance of the best individual of each generation against each best competitor of all generations. This latter technique may be used to select the best solution from an optimization point of view [7]. Both techniques may be used to measure co-evolutionary progress (i.e. the discovery of more general and effective solutions).

2.3 Co-Evolution of Predator and Prey Robots: A Simple Case

The results obtained by running a set of experiments with the parameter described in Section 2.1 are shown below. Fig. 3 represents the results of the Master Tournament, i.e the performance of the best individual of each generation tested against all best competitors. The top graph represents the average result of 10 simulations. The bottom graph represents the result of the best run.

These results show that, at least in this case, phases in which both predators and preys produce increasingly better results are followed by sudden drops in performance. As a consequence, if we look at the average result of different replications in which increase and drop phases occur in different generations, we observe that performance does not increase at all throughout the generations. In other words the efficacy and generality of the different selected strategies does not increase evolutionarily. In fact, individuals of later generations do not necessarily score well against competitors of much earlier generations (see Fig. 4, right side). Similar cases have been described [3, 6].

The 'arm races' hypothesis would be verified if, by measuring the performance of each best individual against each best competitor, a picture approximating that shown on the left side of Fig. 4 could be obtained. In this ideal situation, the bottom-left part of the square, which corresponds to the cases in which predators belong to more recent generations than the prey, is black (i.e. the predator wins). Conversely, the top right part of the square, which corresponds to the cases in which the prey belong to more recent generations than the predators, is white (i.e. the prey wins). Unfortunately, what actually happens in a typical run is quite different (see right part of Fig. 4). The distribution of black and white spots does not differ significantly in the two sub-parts of the square.



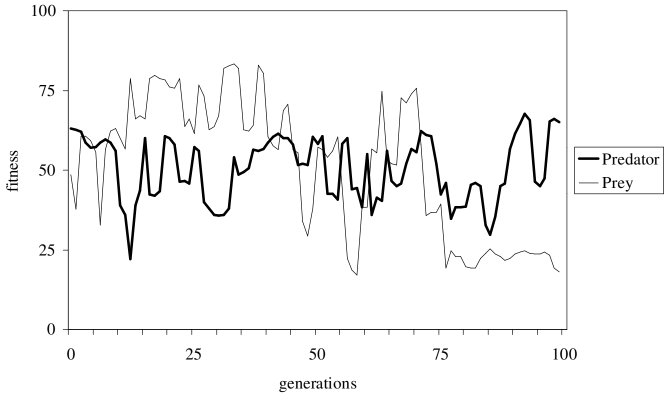


Fig. 3. Performance of the best individuals of each generation tested against all the best opponents of each generation (Master Tournament). Performance may range from 0 to 100 because each individual is tested once against each best competitor of 100 generations. The top graph shows the average result of 10 different replications. The bottom graph shows the result in the best replication (i.e. the simulation in which predators and preys of a given generations attain their best performance). Data were smoothed using rolling average over three data points

This does not imply that the co-evolutionary process is unable to find interesting solutions [7]. This merely means that effective strategies may be lost instead of being retained and refined. Such good strategies, in fact, are often replaced by other strategies that, although providing an advantage over the current opponents, are much less general and effective in the long run. In particular, this type of process may lead to the cycling process described in Section 1.2 in which the same strategies are lost and re-discovered over and over again.

The cycling between the same class of strategy is actually what happens in these experiments. If we analyze the behaviors of the best individuals of successive generations we see that in all replications, evolving predators discover and rediscover two different classes of strategies: (A_1) track the prey and try to catch it by approaching it; (A_2) track the prey while remaining more or less in the same area and attacking the prey only on very special occasions (when the prey is in a particular position relative to the predator). Similarly the prey cycles between two class of strategies: (B_1) stay still or hidden close to a wall waiting for the predator and eventually try to escape when the IR sensors detect the predator; (B_2) move fast in the environment, avoiding walls.

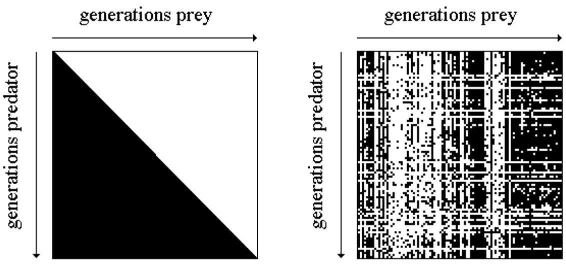


Fig. 4. Performance of the best individuals of each generation tested against all the best opponents of each generation. The black dots represent individual tournaments in which the predators win while the white dots represent tournaments in which the prey wins. The picture on the left represents an ideal situation in which predators are able to catch all prey of previous generations and the prey are able to escape all predators of previous generations. The picture on the right represents the result for the best simulation (the same shown in Fig. 3).

Now, as in Fig. 1, the strategy A_1 is generally effective against B_1, in fact the predator will reach the prey if the prey does not move too much and has a good chance to succeeding given that the prey can only detect predators approaching from certain directions. Strategy B_2 is effective against strategy A_1 because the prey is faster than the predator and so, if the predator tries to approach a moving fast prey, it has little chance of catching it. Strategy A_2 is effective against strategy B_2 because, if the prey moves fast in the environment, the predator may be able to catch it easily by waiting for the prey itself to come close to the predator. Finally, strategy B_1 is very effective against strategy A_2. In fact if the predator does not approach the prey and the prey stays still, the prey will never risk being caught. This type of relation between different strategies produces a cycling process similar to that described in Fig. 1.

What actually happens in the experiments is not so simple as in the description we have just given because of several factors: (1) the strategies described are not single strategies but classes of similar strategies. So for example there are plenty of different ways for the predator to approach the prey and different ways may have different probabilities of being successful against the same opposing strategies; (2) the advantage or disadvantage of each strategy against another strategy varies quantitatively and is probabilistic (each strategy has a given probability of beating a competing strategy); (3) populations at a particular generation do not include only one strategy but a certain number of different strategies although they tend to converge toward a single one; (4) different strategies may be easier to discover or re-discover than others.

However the cycling process between the different class of strategies described above can be clearly identified. By analyzing the behavior of the best individuals of the best simulation (the same as that described in Fig. 3 and 4), for example, we can see that the strategy B_2 discovered and adopted by preys at generation 21 and then abandoned after 15 generations is rediscovered and re-adopted at generation 58 and then at generation 98. Similarly the strategy A_2, first discovered and adopted by the predator at generation 10 and then abandoned after 28 generations for strategy A_1, is then rediscovered at generation 57. Interestingly, however, preys also discover a variation of strategy B_1 that includes also some of the characteristics of strategy B_2. In this case, preys move in circles waiting for the predator as in strategy B_1. However, as soon as they detect the predator with their IR sensors, they start to move quickly exploring the environment as in strategy B_2. This type of strategy may in principle be effective against both strategies A1 and A2. However sometimes preys detect the predator too late, especially when the predator approaches the prey from its left or right rear side which is not provided with IR sensors. Also, it might be that this hybrid strategy which is effective against both predator-strategies, it is not as effective against either predator strategy as the appropriate `pure' escape strategies. Therefore the hybrid strategy, despite its generalized effectiveness, is eventually turned into one of the pure strategies, namely the one that is more effective against whatever strategy is, at that time, being adopted by the predator.

2.4 Testing Individuals against All Discovered Solutions

In a recent article, Rosin and Belew [3], in order to encourage the emergence of 'arms races' in a co-evolutionary framework suggested saving and using as competitors all the best individuals of previous generations:

> *So, in competitive coevolution, we have two distinct reasons to save individuals. One reason is to contribute genetic material to future generations; this is important in any evolutionary algorithm. Selection serves this purpose. Elitism serves this purpose directly by making complete copies of top individuals.*
> *The second reason to save individuals is for purposes of testing. To ensure progress, we may want to save individuals for an arbitrarily long time and continue testing against them. To this end, we introduce the 'hall of fame', which extends elitism in time for purposes of testing. The best individual from every generation, is retained for future testing.*
>
> From Rosin and Belew [3], pp. 8.

This type of solution is of course implausible from a biological point of view. Moreover, we may expect that, by adopting this technique, the effect of the co-evolutionary dynamic will be progressively reduced throughout generations with the increase in number of previous opponents. In fact, as the process goes on, there is less and less pressure to discover strategies that are effective against the opponent of the current generation and greater and greater pressure to develop solutions capable of improving performance against opponents of previous generations.

However, as the authors show, this method may be much more effective than a co-evolutionary framework in which individuals compete only with opponents of the same generation. More specifically, we think, it may be a way to overcome the problem of the cycling of the same strategies. In this framework in fact, ad hoc solutions that compete successfully against the opponent of the current generation but do not generalize to opponents of previous generations cannot spread in evolving populations.

We applied the *hall of fame* selection regime to our predator and prey framework and measured the performance of each best individual against each best competitor (Master Tournament). As shown in Fig. 5 and 6, in this case, we obtain a progressive increase in performance. Results are obtained by running a new set of 10 simulations in which each individual is tested against 10 opponents randomly selected from all previous generations. All the other parameters remain the same.

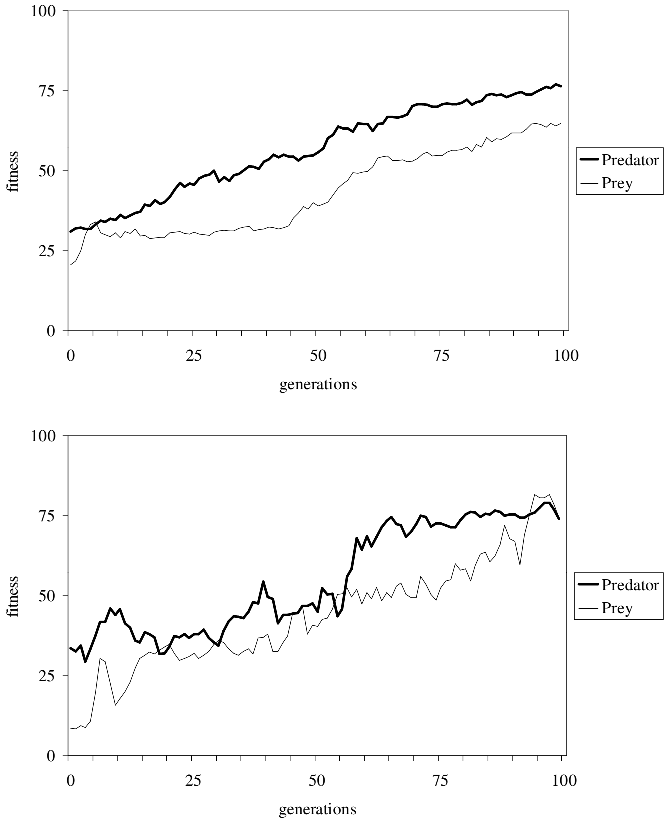


Fig. 5. Performance of the best individuals of each generation tested against all the best opponents of each generation (Master Tournament). The top graph shows the average result of 10 different replications. The bottom graph shows the result in the best replication (i.e. the simulation in which predators and prey of a given generation attain the best performance). Data were smoothed using a rolling average over three data points

Fig. 5 shows how in this case the average fitness of the best individuals tested against all best competitors progressively increases throughout generations, ultimately attaining near to optimal performances. Fig. 6 shows how this is accomplished by being able to beat most of the opponents of previous generations. The results do not exactly match the ideal situation described in Fig. 4 (left side) in which predators and prey are able to beat all individuals of previous generations. In the best simulation described in Fig. 5 (bottom graph) and Fig. 6, for example, there are two phases in which preys are unable to beat most of the predators of few generations before. The general picture, however, approximates the ideal one.

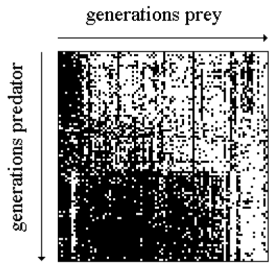


Fig. 6. Performance of the best individuals of each generation tested against all the best opponents of each generation. Black dots represent individual tournaments in which the predators win while white dots represent tournaments in which the prey wins. Result for the best simulation (the same shown in Fig. 5)

If we look at the strategies selected in this set of experiments we see that they are of the same class as those described in the previous Section. However, in this case the strategies are more stable (i.e. in general they are not suddenly replaced by another strategy of a different class). This enables the co-evolutionary process to progressively refine the current strategies instead of cycling between different classes of strategies restarting each time from about the same point.

2.5 How 'Arms Races' can Continue Progressively to Produce More General Solutions in Certain Conditions

In section 2.3 we showed how 'arms races' spontaneously emerge in a co-evolutionary framework. However, we also showed how the innovations produced by such a process may be easily be lost because the evolutionary process tends to fail in a dynamic attractor in which the same type of solutions are adopted over and over by the two co-evolving populations. In section 2.4 we showed how the tendency to cycle between the same type of strategies may be reduced by preserving all previously discovered strategies and by using all of them to test the individual of the current population. However we also pointed out that this techniques which is biologically implausible, has its own drawbacks which may prevent it from scaling up.

In doing so, however, we also learned what characteristics may cause the sudden loss of the acquired abilities which often have to be rediscovered later on. As we showed in Section 2.3, evolution tends to produce the alternation of the same solutions over and over when there are two or more different classes of solutions that interact in a certain way among themselves. This implies that, if such conditions are not verified, 'arms races' should in principle be able to produce better and better solutions without falling into cycling periods.

Of course, it is not easy to predict the cases in which the conditions that produce cycling between the same strategies are absent. However, by analyzing the type of solutions selected by evolution in the experiments described above, we can try to make some predictions. One thing to consider, for example, is that the prey has a limited sensory system that enables it to perceive predators only at a very limited distance and not from all relative directions (there are no IR sensors able to detect predators approaching from the rear-left and rear-right side). Given this limitation, the prey cannot improve its strategy above a certain level. It can compete with co-evolving predators only by suddenly changing strategy as soon as predators select an effective strategy against them. However, if we increase the richness of the prey's sensory system we may expect that the prey will be able to overcome well adapted predators by refining its strategy instead of radically changing its behavior.

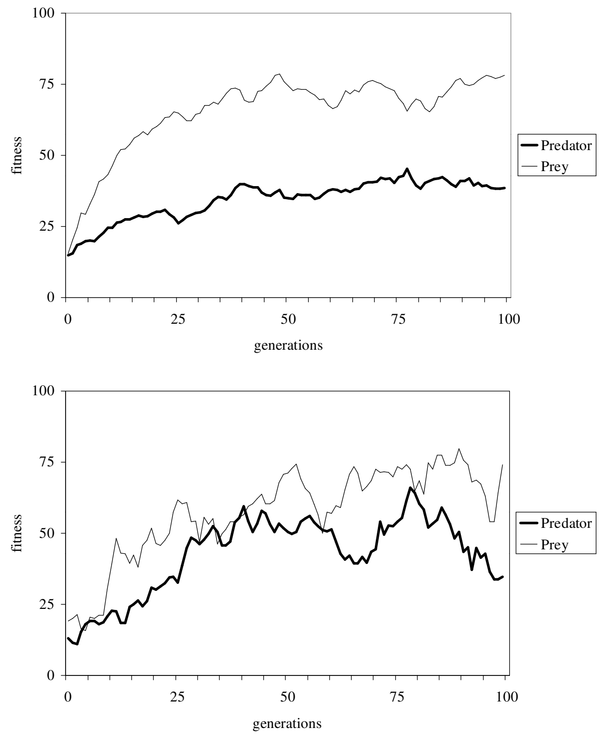


Fig. 7. Performance of the best individuals of each generation tested against all the best opponents of each generation (Master Tournament). The top graph shows the average result of 10 different replications. The bottom graph shows the result in the best replication (i.e. the simulation in which predators and prey of a given generation attain the best performance). Data were smoothed using a rolling average over three data points

To investigate this hypothesis we ran a new set of simulations in which the prey also was provided with a camera able to detect the predators' relative position. For the prey we considered another turret under development at LAMI, which consists of an 1D-array of 150 photoreceptors which provide a linear image composed of 150 pixels of 256 gray levels each [16]. We chose this wider camera because the prey, by escaping the predators, will only occasionally perceive opponents in their frontal direction. As, in the case of predators, the visual field was divided into five sectors of 44° corresponding to five simulated photoreceptors subtending a view-angle of 220°. As a consequence, in this experiment, both predator and prey are controlled by a neural network with 13 sensory neurons. Moreover, in this case, both predator and prey could see their competitors as a black spot against a white background. As in the experiments described in Section 2.3, individuals were tested against the best competitors of the 10 previous generations (not against competitors selected from all previous generations as in the experiments described in Section 2.4). All other parameters remained the same.

If we measure the average performance of the best predators and prey of each generation tested against all the best opponents of each generation (Master Tournament) a significant increase in performance throughout generations is observed (Fig. 7). In the case of the best replication, in particular, although predators' performance decrease in the last 20 generations, the best individuals up to generation 80 are able to overcome most of their opponents of previous generations (Fig. 8).

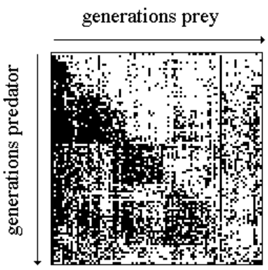


Fig. 8. Performance of the best individuals of each generation tested against all the best opponents of each generation. Black dots represent individual tournaments in which the predators win while white dots represent tournaments in which the prey wins. Result for the best simulation (the same as that shown in Fig. 7)

These results show how by changing the initial conditions (in this case by changing the sensory system of one population) 'arms races' can continue to produce better and better solutions in both populations without falling into cycles.

Interestingly, in their simulations in which also the sensory system of the two co-evolving populations was under evolution, Cliff and Miller observed that ".. pursuers usually evolved eyes on the front of their bodies (like cheetahs), while evaders usually evolved eyes pointing sideways or even backwards (like gazelles)." [16, pp.506]. The authors did not provide enough data in their paper to understand whether their simulations fell into solution cycles. However, even though both the nervous system and the sensory system were under co-evolution in their case, it seems that Cliff and

Miller did not observe any co-evolutionary progress toward increasingly general solutions. In fact, they report that ‘co-evolution works to produce good pursuers and good evaders through a pure bootstrapping process, but both types are rather specially adapted to their opponents’ current counter-strategies.” [16, pp. 506]. However, it should be noted that there are several differences between Cliff and Miller experiments and ours. The fitness function used in their experiments, in fact, is more complex and includes additional constraints that try to force evolution in a certain direction (e.g. predators are scored for their ability to approach the prey and not only for their ability to catch it). Moreover, the genotype-phenotype mapping is much more complex in their cases and includes several additional parameters that may effect the results obtained.

3. Discussion

Evolutionary Robotics is a promising new approach to the development of mobile robots able to act quickly and robustly in real environments. One of the most interesting features of this approach is that it is a completely automatic process in which the intervention of the experimenter is practically limited to the specification of a criterion for evaluating the extent to which evolving individuals accomplish the desired task. However, it is still not clear how far this approach can scale up.

From this point of view, one difficult problem is constituted by the fact that the probability that one individual within the initial generations is able to accomplish the desired task, at least in part, is inversely proportional to the complexity of the task itself. As a consequence, if we apply this methodology to solving a complex task we are likely fail because all individuals of the initial generations are scored with the same zero values and as a consequence the selection mechanism cannot operate. We will refer to this problem as the *bootstrap problem*.

This problem arises from the fact that in artificial evolution people usually start from scratch (i.e. from individuals obtained with randomly generated genotypes). In fact, one possible solution to this problem is the use of ‘incremental evolution’. In this case, we start with a simplified version of the task and, after we get individuals able to solve such a simple case, we progressively move to more and more complex cases [17, 18, 19]. This type of approach can overcome the bootstrap problem, although it also has the negative consequence of increasing the amount of supervision required and the risk of introducing inappropriate constraints. In the case of incremental evolution in fact, the experimenter should determine not only an evaluation criterion but also a ‘pedagogical’ list of simplified criteria. In addition the experimenter should decide when to change the selection criterion during the evolutionary process. Some of these problems may arise also when, although the selection criterion is left unchanged throughout the evolutionary process, it is designed to include rewards also for sub-components of the desired behavior [20].

Another possible solution of the bootstrap problem is the use of co-evolution. Co-evolution of competing populations, in fact, may produce increasingly complex evolving challenges spontaneously without any additional human intervention. Unfortunately however, no continuous increase in complexity is guarantied. In fact, the co-evolutionary process tends to fail into dynamical attractors in which the same solutions are adopted by both populations over and over (we will refer to this problem as the *cycling problem*). What happens is that at a certain point one population, in order to overcome the other population, finds it more useful to suddenly change its strategy instead of continuing to refine it. This is usually followed by a similar rapid change of strategy in the other population. The overall results of this process is that most of the characters previously acquired are not appropriate in the new context and therefore are lost. However, later on, a similar sudden change may bring the two populations back to the original type of strategy so that the lost characters are probably rediscovered again and again.

The effect of the cycling problem may be reduced by preserving all the solutions previously discovered in order to test the individuals of the current generations [3]. However, this method has drawbacks that may affect some of the advantages of co-evolution. In fact, as the process goes on there is less and less pressure to discover strategies that are effective against the opponent of the current generation and increasing pressure to develop solutions able to improve performance against opponents of previous generations which are no longer under co-evolution.

We believe that the cycling problem, as the local minima problem in gradient-descent methods (i.e. the risk of getting trapped in a sub-optimal solution when all similar solutions produce a decrease in performance), is an intrinsic problem of co-evolution that cannot be eliminated completely. However, we also believe that the negative effects of such a problem do not apply to all cases and so completely cancel out the advantages of co-evolution. There may be cases, such as that described in Section 2.5, in which co-evolution may progressively produce more complex solutions for a large number of generations without losing the acquired characters by cycling between different types of solutions.

Moreover, different mechanisms may be able to limit the problems caused by the tendency to cycle between the same types of solutions.

Ontogenetic plasticity, for example, may allow individuals of one population to cope with different classes of strategies adopted by the second population, thus reducing the adaptive advantage of a sudden shift in the behavior causing the cycling problem (on the effects of some forms of ontogenetic plasticity in this framework see [8]). Interestingly, one can argue that co-evolution not only creates the adaptive pressure for ontogenetic adaptation (i.e. the ability to adapt during one's lifetime to different types of opponents' strategies produce a significant increase in the adaptation level of one individual) but also create the conditions in which ontogenetic adaptations can easily arise. In fact, coevolution, by falling into cycles of different classes of strategies, tends to select individuals which can shift from one class of strategy to another with only a few changes at the genotype level. It is plausible to argue that, for such individuals, a limited number of changes during ontogeny will be able to produce the required behavioral shift. In other words, it will be easier for co-evolving individuals to change their behavior during their lifetime to adopt strategies already adopted by their close ancestors thanks to the cycles occurring in previous generations.

Another factor that may limit the cycling problem is the richness of the environment. In the case of co-evolution, competing individuals are part of the environment. This means that part, but not all of the environment is undergoing co-evolution. Now the probability that a sudden shift in behavior will produce viable individuals is inversely proportional to the richness of the environment that is not under co-evolution. Imagine, for example, that an ability acquired under co-evolution, such as the ability to avoid inanimate obstacles, involves a characteristic of the environment which is not under co-evolution. In this case it will be less probable than a sudden shift in strategy involving the lost of such ability will be retained. In fact the acquired character will always have an adaptive value independently of the current strategies adopted by the co-evolving population. The same argument applies to the cases in which one population is co-evolving against more than one other population. The probability of retaining changes that involve a sudden shift in behavior will decrease because, in order to be retained, such changes would have to provide an advantage over both co-evolving populations.

References

1. Miller, G. F., Cliff, D.: Co-Evolution of Pursuit and Evasion I: Biological and Game-Theoretic Foundations, Technical Report CSRP311, School of Cognitive and Computing Sciences, University of Sussex (1994)
2. Dawkins, R., Krebs, J.R.: Arms races between and within species. Proceedings of the Royal Society of London B, **205** (1979) 489:511
3. Rosin, C.D., Belew, R.D.: New methods for competitive coevolution, Evolutionary Computation, **5** (1997) 1-29
4. Koza, J.R.: Evolution and co-evolution of computer programs to control independently-acting agents. In Meyer, J.A., Wilson, S. (eds.): From Animals to Animats: Proceeding of the First International Conference on Simulation of Adaptive Behavior, Cambridge, MA:MIT Press (1991)
5. Koza, J.R.: Genetic programming: On the programming of computers by means of natural selection, Cambridge, MA: MIT Press (1992)
6. Cliff, D., Miller, G.F.: Co-evolution of pursuit and evasion II: Simulation Methods and results. In Maes, P., Mataric, M., Meyer, J-A., Pollack, J., Roitblat, H., Wilson, S. (eds.): From Animals to Animats IV: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior, Cambridge, MA: MIT Press-Bradford Books (1996)
7. Floreano, D., Nolfi, S.: God Save the Red Queen! Competition in Co-Evolutionary Robotics. In Koza, J-R., Kalyanmoy, D., Dorigo, M., Fogel, D.B., Garzon, M., Iba, H., Riolo, R.L. (eds): Genetic Programming 1997: Proceedings of the Second Annual Conference, San Francisco, CA: Morgan Kaufmann (1997)
8. Floreano, D., Nolfi, S.: Adaptive behavior in competing co-evolving species. In Husband, P., Harvey, I. (eds): Proceedings of the Fourth European Conference on Artificial Life, Cambridge, MA: MIT Press (1997)
9. Floreano, D., Nolfi, S., Mondada, F.: Co-evolutionary Robotics: From Theory to Practice, Technical Report, EPFL, Swiss Federal Institute of Technology (1998)
10. Floreano, D., Mondada, F.: Automatic creation of an autonomous agent: Genetic evolution of a neural-network driven robot. In Cliff, D., Husband, F., Meyer J-A., Wilson, S. (eds.): From Animals to Animats III: Proceedings of the Third International Conference on Simulation of Adaptive Behavior, Cambridge, MA: MIT Press-Bradford Books (1994)
11. Yao, X.: A review of evolutionary artificial neural networks, International Journal of Intelligent Systems, **4** (1993) 203-222
12. Sims, K.: Evolving 3D morphology and behavior by competition. Artificial Life, **1** (1995) 353-372

13. Miglino, O., Lund, H. H., Nolfi, S.: Evolving Mobile Robots in Simulated and Real Environments. Artificial Life, **4** (1995) 417-434
14. Ridley, M.: The Red Queen: Sex and the evolution of human nature. Viking: London (1993)
15. Cliff, D., Miller, G.F.: Tracking the read queen: Measurement of adaptive progress in co-evolutionary simulations. In Moran, F., Moreno, A., Merelo, J.J., Chacon, P. (eds.): Advances in Artificial Life: Proceedings of the Third European Conference on Artificial Life, Berlin:Springer Verlag (1995)
16. Landolt, O.: Description et mise en oeuvre du chip ED084V2A, Technical Report 16-11-95, Centre Suisse d'Electronique et Microtechnique, Switzerland (1996)
17. Floreano, D.: Emergence of Home-Based Foraging Strategies in Ecosystems of Neural Networks. In: Meyer, J-A., Roitblat, H.L., Wilson, S. (eds.): From Animals to Animats II: Proceedings of the Second International Conference on Simulation of Adaptive Behavior, Cambridge, MA:MIT Press-Bradford Books (1993)
18. Harvey, I., Husbands, P., Cliff, D.: Seeing The Light: Artificial Evolution, Real Vision. In Cliff, D., Husbands, P., Meyer, J-A., Wilson, S. (eds.) From Animals to Animats III: Proceedings of the Third International Conference on Simulation of Adaptive Behavior, Cambridge, MA: MIT Press-Bradford Books (1994)
19. Gomez, F., Miikkulainem, R.: Incremental Evolution of Complex General Behavior, Adaptive Behavior, **5** (1997) 317-342
20. Nolfi, S. Evolving non-trivial behaviors on real robots: a garbage collecting robot, Robotics and Autonomous System 646, Special issue on "Robot learning: The new wave" (1997) 1-12