

Novelty Search in Competitive Coevolution

Jorge Gomes^{1,2}, Pedro Mariano², and Anders Lyhne Christensen^{1,3}

¹Instituto de Telecomunicações, Lisbon, Portugal

²LabMAG – Faculdade de Ciências da Universidade de Lisboa, Portugal

³Instituto Universitário de Lisboa (ISCTE-IUL), Lisbon, Portugal

`jgomes@di.fc.ul.pt, plmariano@fc.ul.pt, anders.christensen@iscte.pt`

Abstract. One of the main motivations for the use of competitive coevolution systems is their ability to capitalise on arms races between competing species to evolve increasingly sophisticated solutions. Such arms races can, however, be hard to sustain, and it has been shown that the competing species often converge prematurely to certain classes of behaviours. In this paper, we investigate if and how novelty search, an evolutionary technique driven by behavioural novelty, can overcome convergence in coevolution. We propose three methods for applying novelty search to coevolutionary systems with two species: (i) score both populations according to behavioural novelty; (ii) score one population according to novelty, and the other according to fitness; and (iii) score both populations with a combination of novelty and fitness. We evaluate the methods in a predator-prey pursuit task. Our results show that novelty-based approaches can evolve a significantly more diverse set of solutions, when compared to traditional fitness-based coevolution.

Keywords: Competitive coevolution, behavioural diversity, novelty search, convergence, evolutionary robotics

1 Introduction

In a coevolutionary system, two or more populations simultaneously evolve, and evaluations are conducted with an individual from each population. In the case of competitive coevolution systems (CCESs), the populations represent competing species, where each succeed at the expense of the opposing species. CCESs are appealing for a number of reasons [18, 14]: (i) they are suited to domains where a notion of absolute fitness might not exist; (ii) each population acts as a progressively more challenging opponent to the other population, and (iii) a CCES may be less prone to stagnation than non-coevolutionary methods, because of the ever-changing fitness landscape.

In practice, however, many of these advantages are frequently not observed [15]. The key to successful coevolutionary learning is a continuous arms race between the opposing species [7]. However, such an arms race is not easy to establish [6], and even when an arms race does occur, there is no guarantee that it will lead to good solutions. A number of techniques to help sustain an arms

race have been proposed [17], but most of them are focused on maintaining performance against opponents from earlier generations. While such techniques can contribute to a continuous and objective improvement of the solutions, they do not guarantee that a wide range of solutions are discovered, which is one of the most appealing arguments for the use of competitive coevolution [14]. The goal of a CCES is not necessarily to find the (near-)optimal solution for a problem [15]. The discovery of diverse solutions might be equally valuable, since there is no absolute measure of quality. Consider, for instance, the evolution of agents to play against human players: having a good diversity of playing strategies might be more valuable than having a single near-optimal player [19]. Moreover, it has been shown that promoting diversity in a population can potentially lead to improvements in the population’s generalisation capacity [3].

Evolutionary techniques based on behavioural diversity, such as novelty search (NS) [11], have shown effective in overcoming behavioural convergence. These techniques drive evolution towards behavioural innovation, often resulting in a more effective evolutionary process that produces a greater diversity of solutions compared with traditional fitness-based evolution. Although most studies focus on non-coevolutionary domains [12, 10], we recently showed that NS can be successfully used to overcome stable states in cooperative coevolution [8].

In this paper, we study the adaptation of behavioural diversity techniques to competitive coevolution. This combination of techniques is especially challenging because in a CCES, all populations are required to be fairly competitive at all times, which may conflict with the diversity objective [1]. We propose three different ways to apply novelty search to a CCES, and compare them with traditional fitness-based coevolution in a predator-prey pursuit task. We assess the proposed techniques along three dimensions: (i) the quality of the best solutions achieved; (ii) the exploration of the behaviour space; and (iii) the diversity of high-quality solutions evolved.

2 Related Work

2.1 Premature Convergence in Competitive Coevolution

Many reports of convergence of coevolving populations to undesirable regions of the solution space can be found in the literature (for examples, see [5, 2, 16]). One of the main causes for failure is that the coevolutionary process easily gets trapped in a *mediocre stable state* [7]: a cycle where a limited set of solutions is adopted by the populations over and over again. This cycle can give the impression of competition, without actually causing the evolutionary process to explore new solutions or to improve the quality of the individuals.

Another reason for the loss of diversity in the populations can be the lack of a fitness gradient [18]. It is possible that one population becomes so dominant that it turns into an *unhittable target* for its competitors. As such, the selection pressure disappears, and the populations stop improving. Ashlock et al. [1] tried to overcome the lack of a fitness gradient by rewarding individuals that generate

a high variability of fitness scores in the opposing population. The approach was, however, unsuccessful, since individuals that allow for a high variability of fitness typically do not pose a significant challenge.

A number of strategies for cultivating fruitful arms races have been proposed. Rosin & Belew [17] propose three techniques that aim to select a diverse and challenging set of competitors to test the individuals: (i) competitive fitness sharing, (ii) shared sampling, and (iii) the hall of fame. Other works have focused on the characteristics of the task setup that are favourable to the emergence and sustainability of arms races [13]. It has also been shown that shaping the environment or the fitness function throughout evolution can help avoid convergence to uninteresting and non-diverse solutions [5].

2.2 Novelty Search

Novelty search (NS) [11] is an evolutionary technique in which the population is driven towards behavioural novelty. NS has the potential to avoid premature convergence, and evolve a wide diversity of solutions in a single evolutionary run, as opposed to fitness-driven evolution that typically converges to a certain class of solutions [10]. The approach has been applied in the domain of evolutionary robotics and evolution of agent controllers with considerable success [12, 10, 8].

Novelty search algorithm The distinctive aspect of NS is how the individuals of the population are scored. Instead of scoring individuals according to how well they perform a given task, which is typically measured by a fitness function, individuals are scored based on their behavioural novelty. Behavioural novelty is measured with a *novelty metric* that quantifies how different an individual is from other, previously evaluated individuals. The distance between two individuals is typically given by the distance between their *behaviour characterisation vectors*. These vectors are commonly composed of behavioural traits that the experimenter considers relevant for the task.

To measure how far an individual is from other individuals in behaviour space, the novelty metric relies on the average behaviour distance of that individual to the k nearest neighbours. Potential neighbours include the other individuals of the current generation and a random sample of individuals from previous generations (stored in an archive). Candidates from sparse regions of the behaviour space therefore receive higher novelty scores, which results in a constant evolutionary pressure towards behavioural innovation.

Combining novelty and fitness As NS is guided by behavioural innovation alone, its performance can be greatly affected by the size and shape of the behaviour space [10]. Since NS is essentially an exploratory technique, most of the effort may be spent in behaviour regions that are irrelevant for the fulfilment of the task. Therefore, NS is often combined with fitness-based evolution to promote exploration of high-fitness behaviour regions.

In our experiments, we use a variant of *progressive minimal criteria novelty search* (PMCNS) [9] to combine novelty and fitness. PMCNS restricts exploration to behaviour regions associated with relatively high fitness scores. At

each generation g , the individuals’ selection scores are assigned in the following way: if the fitness score of an individual is above the minimal criterion mc_g , its novelty score alone is used for selection, otherwise it receives a score of zero. The mc_g criterion corresponds to the P -th percentile value of the fitness scores of all individuals of generation g .

3 Approach

The application of behavioural diversity techniques to a CCES poses a number of challenges. In order to evolve effective solutions, coevolution requires an arms race between the coevolving species. This implies that all species are simultaneously improving to defeat one another: they learn to exploit flaws in the opponents’ strategy. Behavioural diversity techniques are, however, essentially exploratory: evolution is driven towards behavioural novelty, not necessarily better solutions (i.e, solutions that are able to defeat the opponents). Prioritising diversity instead of competitiveness can compromise the effectiveness of coevolution, as shown in [1]. On the other hand, it has also been shown that when a population has converged to a strategy, it can be beneficial to reward alternate strategies, in order to foster the evolution of new behavioural traits [5].

We study the application of NS to a CCES with two competing species. The individuals’ fitness score is measured in the traditional way: by competing against a representative sample of the individuals from the opposing species. In these competitions, we record the individual’s behaviour as a task-specific behaviour characterisation vector. The novelty score of an individual is then calculated by computing the distance between its characterisation vector and the vectors of current and past individuals from the same species. The following methods are evaluated in this paper:

- Fit** Selection based exclusively on the fitness score, in both populations. This method is used as a baseline.
- NS-Both** Selection based exclusively on the novelty score in both populations.
- NS- p** Selection based on the novelty score in population p (in our experiments, either *Pred* for predator, or *Prey*), and based on the fitness score in the other population.
- PMCNS** Both populations are scored with PMCNS based on both novelty and fitness scores. The individuals are rewarded for displaying behavioural novelty while still meeting the minimal fitness criterion.

4 Experimental Setup

4.1 Predator-Prey Pursuit

The predator-prey pursuit task is a common testbed for CCESs [13, 4]. In this task, two agent controllers are coevolved: one for a predator and one for a prey. The predator’s objective is to capture (touch) the prey, and the prey’s objective is

to escape the predator. The agents operate in a closed square arena of $75 \times 75 \text{ cm}^2$. A simulation trial ends if the predator captures the prey or if 100 seconds elapse. The initial conditions and setup of the task are depicted in Figure 1.

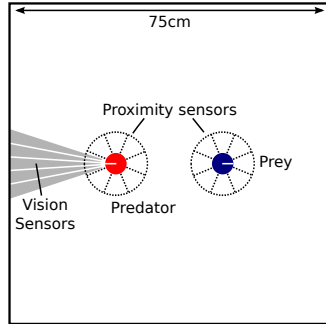


Fig. 1. Setup of the predator-prey task. The agents are initially placed in fixed positions, facing opposing directions.

The fitness of the prey is given by the fraction of simulations it avoided capture. The fitness of the predator is the opposite: the fraction of simulations it was able to capture the prey. The behaviour characterisation used in the novelty-based setups is the same for both the predator and the prey. It captures behavioural traits that are intuitively relevant in the context of the task. The characterisation is a vector of four elements: (i) the simulation length; (ii) the mean distance to the other agent throughout the simulation; (iii) the mean agent movement speed; and (iv) the mean distance between the agent and the closest wall. All elements are normalised to $[0, 1]$.

4.2 Coevolutionary Algorithm

The predator and prey controllers are coevolved in two separate populations. The weights of the neural controllers are directly encoded in the chromosomes. Each population is evolved with a simple evolutionary strategy with the following parameters: no crossover, a gene mutation rate of 5%, population size of 200, and tournament selection of size 5. At the end of each generation, the individual with the highest fitness score of each population is added to the respective *hall of fame* [17]. The individuals of each population are evaluated against a set of ten competitors randomly drawn from the opposing species' hall of fame.

NS is implemented as described in Section 2.2. Each population has its own novelty archive. Individuals are added to the archive with a probability of 3%. The archive can hold at most 1000 individuals: after reaching the limit, random individuals are removed to allow space for new ones. PMCNS was configured with a percentile value P of 0.5, i.e., the median fitness of the population. Values of 0.25 and 0.75 were also tested, but yielded inferior performance.

The experimental setup is based on [14]. The prey and the predator move at a maximum speed of 5 cm/s. Both the prey and predator have eight proximity sensors, evenly distributed on their bodies, with a maximum range of 5 cm. The proximity sensors can detect both the walls and the other agent. The predator is additionally equipped with five vision sensors to detect the prey. The vision sensors have unlimited range and when combined provide a view angle of 40° . All sensors are binary: they return 1 if something is detected, and 0 otherwise. Each agent is controlled by a feed-forward neural network. Sensor readings are fed to the network, and two outputs control the speed of the two wheels. The networks of the predator and the prey have respectively 7 and 5 hidden neurons.

5 Results

5.1 Quality of the Solutions

In all evolutionary runs, the individuals with the highest fitness score from both populations in every generation were recorded. To obtain an objective measure of the quality of these individuals, we performed a two-step master tournament [14], obtaining a *master fitness* and a *master behaviour characterisation*. First, the best individuals of each evolutionary run were identified by evaluating them against the individuals of the opposing species from the same run. Second, the individuals were evaluated against all the individuals identified in the first step, from all evolutionary treatments. The analysis of the individuals' master fitness can be seen in Figure 2 and Table 1 (Best fitness column).

These results show that all evolutionary treatments were able to achieve high-quality predator controllers. The differences between the treatments are relatively small. *Fit* is only significantly different from *NS-Both*, which is significantly inferior to all other treatments (Mann-Whitney U test, p -value < 0.01). The evolution of prey controllers, on the other hand, revealed larger differences between the evolutionary treatments. The treatment in which novelty had the biggest influence (*NS-Both*) was the lowest performing treatment with respect to the quality of the solutions (p -value < 0.01). Conversely, the treatments with highest fitness pressure yielded the highest scoring solutions, with none of the novelty-based treatments outperforming *Fit* (p -value < 0.01). This result confirms that in order to succeed, the populations need to have a strong selection pressure towards defeating the opposing population. In the predator-prey task, using NS to promote exploration of the behaviour space did not yield significant advantages regarding the achieved fitness scores.

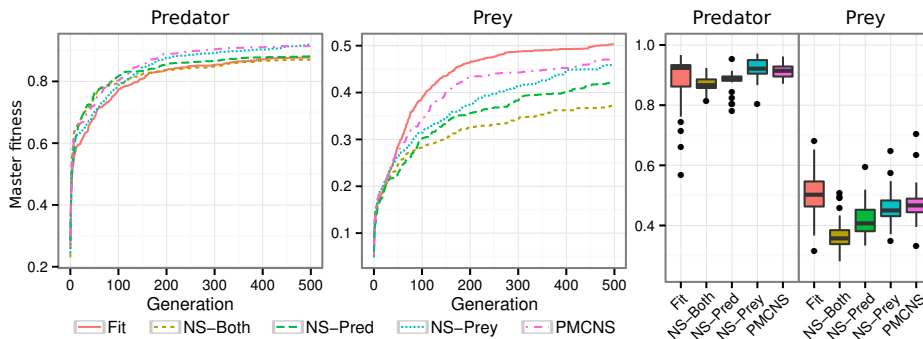


Fig. 2. Left and middle: best fitness found so far at each generation, for each population and each method. The fitness values correspond to the *master fitness*, and are averaged over 30 evolutionary runs. Right: boxplots of the highest fitness scores found in each evolutionary run.

Table 1. Best fitness found in each evolutionary run; Average coverage of the behaviour space in a whole evolutionary run; Average coverage of the high-fitness behaviour regions. Values are averages of 30 evolutionary runs for each method. Standard deviations are in parentheses.

| | Best fitness | | Global exploration | | Elite exploration | |
|----------------|--------------|-------------|--------------------|-------------|-------------------|-------------|
| | Predator | Prey | Predator | Prey | Predator | Prey |
| <i>Fit</i> | 0.88 (0.10) | 0.50 (0.09) | 0.42 (0.05) | 0.37 (0.05) | 0.49 (0.11) | 0.46 (0.06) |
| <i>NS-Both</i> | 0.87 (0.03) | 0.37 (0.06) | 0.64 (0.03) | 0.60 (0.03) | 0.62 (0.04) | 0.60 (0.05) |
| <i>NS-Pred</i> | 0.88 (0.04) | 0.42 (0.06) | 0.64 (0.02) | 0.40 (0.04) | 0.67 (0.06) | 0.43 (0.08) |
| <i>NS-Prey</i> | 0.92 (0.04) | 0.46 (0.06) | 0.43 (0.03) | 0.59 (0.03) | 0.50 (0.04) | 0.70 (0.04) |
| <i>PMCNS</i> | 0.91 (0.03) | 0.48 (0.07) | 0.59 (0.03) | 0.53 (0.04) | 0.70 (0.03) | 0.59 (0.06) |

5.2 Behaviour Space Exploration

One of the main advantages of novelty-based evolutionary techniques is the diversity of evolved solutions. To make a quantitative analysis of the behaviour space exploration, the space was first divided in regions of equal size: each dimension of the behaviour characterisation was discretised into 5 levels, resulting in a total of 625 regions. We then calculated how many times each region was visited throughout each evolutionary run (based on the *master characterisations*). This distribution was compared with the uniform distribution to obtain a measure of behaviour space coverage. The non-visited regions of the behaviour space (not reached by any method) were excluded, see Table 1 – Global exploration. To measure the exploration of the high-fitness behaviour regions, we followed the same procedure but excluded the regions where no reasonably good solutions were found (master fitness below 0.8 for the predator population, and 0.3 for the prey population), see Table 1 – Elite exploration.

Fit has the lowest degree of exploration, which suggests that it typically converges to a single class of solutions. *NS-Both*, on the other hand, has the highest degree of global exploration, but the same does not hold for elite exploration. The lack of elite exploration can explain the inferior performance of *NS-Both* regarding the achieved fitness scores: too much effort is spent on the exploration of low-fitness behaviour regions. The results show that *PMCNS* is able to avoid spending too much effort on exploration of low-fitness behaviours. The global exploration in *PMCNS* is still significantly higher than *Fit*, and it has one of the highest degrees of elite exploration.

Comparing *NS-p* to *Fit*, it is possible to observe that the fitness-driven population in *NS-p* has similar exploration and fitness scores (see in Table 1 – Prey: *NS-Pred*, *Fit*, and Predator: *NS-Prey*, *Fit*). The fitness-driven population is not significantly affected by the greater behavioural diversity in the opposing, novelty-driven population. On the other hand, if we compare the novelty-driven population with *NS-Both*, we can see that it is favourable to have one of the populations driven by fitness (see in Table 1 – Predator: *NS-Pred*, *NS-Both*, and Prey: *NS-Prey*, *NS-Both*). When a novelty-driven population competes against a fitness-driven population, it is able to achieve significantly higher elite explo-

ration scores, which also translates into higher fitness scores reached. By co-evolving with a highly competitive, fitness-driven population, the novelty-driven population tends to move towards better solutions – even though their individuals are selected based on their novelty scores alone.

5.3 Diversity of Effective Solutions

The behaviour of the most effective preys is always very similar: they move at full speed in circular trajectories around the arena, using the proximity sensors to avoid the walls and to escape the predators when being chased. This behaviour typically fails when the predator quickly approaches the prey head on or from the sides, or when the prey is being chased and gets trapped between a wall and the predator. The set of successful predator strategies evolved was more diverse, especially in novelty-based treatments. To visualise the diversity of predator behaviours, we reduced the four dimensions of the characterisation using a Kohonen self-organising map. The Kohonen map was trained using a sample of the predator behaviours (*master characterisations*) found in all evolutionary runs. The solutions evolved in each run were then mapped individually. Figure 3 shows one typical evolutionary run of each treatment. *NS-Prey* is omitted since the exploration in the predator population is similar to *Fit*.

The results showed by the exploration maps are in accordance with the exploration scores (Table 1). The evolutionary runs of *Fit* typically converge to region B (23 out of 30 runs). In seven of those runs, solutions in region A were also discovered. Only one run of *Fit* explored regions C and D, although these regions are associated with solutions of similar quality to the solutions of regions A and B. *NS-Both* explored the behaviour space relatively uniformly, without any noticeable bias towards specific regions. *NS-Pred* is similar to *NS-Both*, but clearly spends more effort in high-fitness regions. The exploration pattern of the *PMCNS* treatment is in between that of *NS-Both* and *Fit*: a single evolutionary run typically converges to multiple classes of high-quality solutions.

Observing the predator solutions in action confirms that the novelty-based treatments can evolve an interesting and diverse set of behaviours. The follow-

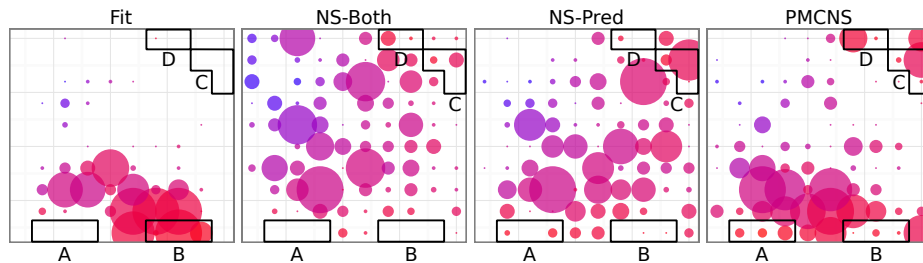


Fig. 3. Exploration of the predator behaviour space in typical evolutionary runs. The diameter of each circle is proportional to the number of individuals found belonging to that behaviour region. The highlighted regions A, B, C and D are associated with the highest fitness scores.

ing behaviour descriptions correspond to the highest scoring predator solutions found in each of the behaviour regions highlighted in Figure 3.

- (A) Moves backwards and in a curved trajectory until the prey is detected with the vision sensors. Afterwards, the predator stops and rotates to face the prey. When the prey gets close, the predator rushes towards it.
- (B) Moves forwards at full speed, avoiding the walls, until the vision sensors detect the prey. Afterwards, the predator chases the prey at full speed.
- (C) Moves slowly and directly towards a wall, and then stays there, facing the centre of the arena. When a prey passes nearby, the predator quickly rushes forward to catch the prey.
- (D) Does not initially move, only rotates to find the prey. When the prey is located, the predator keeps rotating to face the prey, and moves backwards very slowly. When the prey is near, the predator rushes towards it.

6 Conclusions

We proposed three methods to promote behavioural diversity in competitive coevolution: selection based on novelty score in both populations (*NS-Both*), novelty score only in one of the populations (*NS-p*), and a combination of novelty and fitness scores in both populations (*PMCNS*). These methods were compared with traditional fitness-based coevolution (*Fit*) in a predator-prey task.

With respect to the highest fitness scores achieved in the evolutionary runs, there was no significant advantage of novelty-based methods, when compared to fitness-based evolution. The novelty-based methods did, however, display significantly higher degrees of behaviour space exploration. The *NS-Both* method had the lowest performance among the novelty-based methods, since most of the exploration effort was spent in behaviour regions associated with relatively low fitness scores. However, when a novelty-driven population (*NS-p*) competed against a fitness-driven population, results showed that the novelty-driven population had a significantly stronger tendency to explore high-fitness behaviour regions. *PMCNS* was effective all-around: it could achieve fitness scores comparable to *Fit*, and could discover a wide range of high-quality solutions in a single evolutionary run, for both predator and prey controllers.

The novelty-based methods could consistently explore new behaviour regions and find classes of solutions that *Fit* rarely or never reached. In particular, a diverse set of interesting high-quality predator solutions was identified. Our experiments showed that novelty-based techniques can be used to avoid behavioural convergence and discover a broad diversity of solutions in competitive coevolution systems. The effectiveness of the novelty-based techniques depends on the balance between diversity and competitiveness, since the populations require challenging opponents in order to achieve solutions of high objective quality.

Acknowledgements This research is supported by Fundação para a Ciência e Tecnologia (FCT) grants PEst-OE/EEI/LA0008/2013, PEst-OE/EEI/UI0434/2014, SFRH/BD/89095/2012 and EXPL/EEI-AUT/0329/2013.

References

1. Ashlock, D., Willson, S., Leahy, N.: Coevolution and tartarus. In: Congress on Evolutionary Computation (CEC). vol. 2, pp. 1618–1624. IEEE Press (2004)
2. Avery, P., Louis, S.: Coevolving team tactics for a real-time strategy game. In: Congress on Evolutionary Computation (CEC). pp. 1–8. IEEE Press (2010)
3. Chong, S.Y., Tino, P., Yao, X.: Relationship between generalization and diversity in coevolutionary learning. *IEEE Transactions on Computational Intelligence and AI in Games* 1(3), 214–232 (2009)
4. Cliff, D., Miller, G.F.: Tracking the red queen: Measurements of adaptive progress in co-evolutionary simulations. In: *Advances In Artificial Life*, pp. 200–218. Springer (1995)
5. Dziuk, A., Miikkulainen, R.: Creating intelligent agents through shaping of coevolution. In: Congress on Evolutionary Computation (CEC). pp. 1077–1083. IEEE Press (2011)
6. Ebner, M., Watson, R.A., Alexander, J.: Coevolutionary dynamics of interacting species. In: *European Conference on the Applications of Evolutionary Computation (EvoApplications)*, Lecture Notes in Computer Science, vol. 6024, pp. 1–10. Springer (2010)
7. Ficici, S.G., Pollack, J.B.: Challenges in coevolutionary learning: Arms-race dynamics, open-endedness, and mediocre stable states. In: *Artificial Life*. pp. 238–247. MIT Press (1998)
8. Gomes, J., Christensen, A.L., Mariano, P.: Avoiding convergence in cooperative coevolution with novelty search. In: *International Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*. pp. 1149–1156. IFAAMAS (2014)
9. Gomes, J., Urbano, P., Christensen, A.L.: Progressive minimal criteria novelty search. In: *Ibero-American Conference on Artificial Intelligence (IBERAMIA)*, Lecture Notes in Computer Science, vol. 7637, pp. 281–290. Springer (2012)
10. Gomes, J., Urbano, P., Christensen, A.: Evolution of swarm robotics systems with novelty search. *Swarm Intelligence* 7(2–3), 115–144 (2013)
11. Lehman, J., Stanley, K.O.: Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary Computation* 19(2), 189–223 (2011)
12. Mouret, J.B., Doncieux, S.: Encouraging behavioral diversity in evolutionary robotics: an empirical study. *Evolutionary Computation* 20(1), 91–133 (2012)
13. Nolfi, S.: Co-evolving predator and prey robots. *Adaptive Behavior* 20(1), 10–15 (2012)
14. Nolfi, S., Floreano, D.: Coevolving predator and prey robots: Do arms races arise in artificial evolution? *Artificial Life* 4(4), 311–335 (1998)
15. Popovici, E., Bucci, A., Wiegand, R.P., De Jong, E.D.: Coevolutionary principles. In: *Handbook of Natural Computing*, pp. 987–1033. Springer (2012)
16. Reisinger, J., Bahçeci, E., Karpov, I., Miikkulainen, R.: Coevolving strategies for general game playing. In: *Computational Intelligence and Games*. pp. 320–327. IEEE Press (2007)
17. Rosin, C.D., Belew, R.K.: New methods for competitive coevolution. *Evolutionary Computation* 5(1), 1–29 (1997)
18. Watson, R.A., Pollack, J.B.: Coevolutionary dynamics in a minimal substrate. In: *Genetic and Evolutionary Computation Conference (GECCO)*. pp. 702–709. Morgan Kaufmann (2001)
19. Yannakakis, G.N., Hallam, J.: Modeling and augmenting game entertainment through challenge and curiosity. *International Journal on Artificial Intelligence Tools* 16(6), 981–999 (2007)