Bridging the Reality Gap — A Dual Simulator Approach to the Evolution of Whole-Body Motion for the Nao Humanoid Robot

Malachy Eaton
Dept. Computer Science and Information Systems, University of Limerick, Limerick, Ireland

Keywords: Evolutionary Algorithms, Humanoid Robotics, Ball Kicking, Evolutionary Robotics, Evolutionary Humanoid Robotics, Whole-Body-Motion.

Abstract: We describe a novel approach to the evolution of whole-body behaviours in the Nao humanoid robot using a multi-simulator approach to the alleviation of the reality gap issue. The initial evolutionary process takes place in the V-REP simulator. Once a viable whole-body motion has been evolved, this evolved motion is subsequently transferred for testing onto another simulation platform – Webots. Only when the evolved kicking behaviour has been demonstrated to also be viable on the Webots platform is this behaviour then transferred onto the real Nao robot for testing. This eliminates the time-consuming process of transferring behaviours onto the real robot which have little chance of successfully crossing the reality gap, and also minimises the potential for damage to the real Nao robot and/or it's environment. By using this novel approach of employing two different simulators, each with its own individual strengths and weaknesses, we reduce the likelihood that any individual behaviour will be able to exploit individual simulators' weaknesses, as the other simulator should pick up on this weak point. Using this procedure we have successfully evolved ball kicking behaviour in simulation, which has transferred with reasonable fidelity onto to the real Nao humanoid.

1 INTRODUCTION

The field of humanoid robotics addresses the creation of mobile robots that are broadly humanlike in their gross anatomy and/or aspects of their behaviour. Humanoid robots have several advantages, not least of which is their potential ability to operate in environments designed for humans, thus potentially having the ability to handle tasks that may be time-consuming, distasteful, or even dangerous for humans to perform (Eaton, 2015).

A robot in common use today by researchers into human-like behaviours is the Nao humanoid robot from Aldebaran Robotics (Gouaillier et al., 2009). This robot has up to 25 degrees of freedom and stands 58cm tall. A version of this robot is used in the RoboCup Standard Platform League (SPL), with the eventual avowed aim of producing, by the year 2050 a humanoid robot team that will be able to take on (and beat) the current human World Cup champions (Kitano and Asada, 1998) (Kitano et al., 1998). Central, of course, to being able to take up this challenge is the development of an effective kicking action, which is the area we address in this paper. While some work has been done to date on the automatic generation of kicking motions through parameter optimisation or other means (e.g. (Jouandeau and Hugel, 2014) ,(Li et al., 2015)), little work has been done on the direct evolution of individual joint motions for the robot, which is the approach we take. In general the field of evolutionary robotics seeks to evolve some, or all aspects of a robots controller and/or morphology (Nolfi and Floreano, 2000), (Bongard, 2013).

Much of the work in the area of generating soccer-centric skills has involved simulated robots for the RoboCup3D simulation league (Depinet et al., 2014), however our emphasis is on the evolution of behaviours which can be transferred effectively onto the real robot.

1.1 The "Reality Gap" Issue

A major issue that arises in this regard is the so-called "reality gap"; that is the potential disparity between evolved (or otherwise generated) behaviours in simulation, and their actual implementation on the real robot. This can be of particular importance in the

Eaton, M.
Bridging the Reality Gap — A Dual Simulator Approach to the Evolution of Whole-Body Motion for the Nao Humanoid Robot.
DOI: 10.5220/0006052301860192
In *Proceedings of the 8th International Joint Conference on Computational Intelligence (IJCCI 2016) - Volume 1: ECTA*, pages 186-192
ISBN: 978-989-758-201-1

evolution of behaviours for multi-jointed robots with many degrees of freedom, as in the case discussed in this paper. Various approaches have been taken to alleviate this issue, including the transferability approach (Koos et al., 2013), the grounded simulated learning approach (Farchy et al., 2013), the leveraging multiple simulators approach (Boeing and Bräunl, 2012), combining evolution in simulation with pre-programmed behaviours (Duarte et al., 2012), using an EA to tune the parameters of a simulator (Laue and Hebbel, 2009), fitness function correction interleaving simulated and real data (Iocchi et al., 2007), coevolution of controller and simulator (Lipson et al., 2006), (Bongard and Lipson, 2004), the “back to reality approach” (Zagal and Ruiz-Del-Solar, 2007), (Zagal et al., 2004), the online adaptation approach (Floreano and Urzelai, 2001), the envelope of noise approach (Jakobi, 1997a), (Jakobi, 1997b), and scaled experimentation (Eaton, 2015).

Although there has been work done to date on leveraging the effects of multiple physics simulators (Boeing and Bräunl, 2012), (Boeing, 2009) for evolutionary robotics experiments, to our knowledge this is one of the few, if any, which utilises the advantages of using multiple simulation packages, rather than just the core physics engines.

1.2 Simulators used — Webots and V-REP

The two simulation packages we use are Webots (Michel, 2004) and the Virtual Robot Experimentation Platform (V-REP) (Freese, 2010). Both of these packages have been used extensively in the simulation of a wide variety of robots including wheeled and legged robots of a variety of types, and also for the simulation of humanoid robots. Webots, which in original form dates from 1996, is one of the longest running simulators in continuous development suited for the detailed simulation of complex robotic environments. V-REP is a more recent arrival dating from around the start of this decade, and which describes itself as the Swiss army knife among robot simulators; an example scene from the V-REP simulator is given in Fig. 1. Regarding physics engines Webots relies on the Open Dynamics Engine (ODE), while V-REP provides a choice of 4 engines, ODE, the Bullet physics library, the Vortex Dynamics Engine and the Newton Dynamics engine. For the work described here we utilise the Bullet physics engine.

1.3 Overall Approach

Our approach, then, is to run the evolutionary experiments on a simulated Nao robot in the V-REP package, and then to transfer successfully evolved controllers into the Webots environment for further testing and validation of their overall performance.

One advantage of our approach is that it is unlikely that simulation weaknesses that would manifest themselves on one simulator would occur on the other , and vice versa. Another advantage of our approach is that while Webots is a proprietary simulation package, a fully functional version of V-REP is freely available for non-commercial use. We have observed through experimentation that a significant proportion of behaviours evolved using the V-REP platform do not transfer successfully to the real Nao robot. As the process of transferral to the real robot can be quite time-consuming , and as the potential for damage to the robot and/or its environment on execution of an incorrectly evolved motion involving quite rapid whole-body motion such as kicking is nontrivial, it is highly desirable only to transfer motions to the real Nao robot which have a high probability of success.

We have observed that it is very unlikely that a behaviour evolved in V-REP, but that fails to operate successfully in Webots will transfer onto the real robot with any degree of fidelity, however if validated in the Webots simulator a high percentage should transfer with reasonable accuracy. Preliminary experimental verification of this observation is discussed in section 3.

Another advantage is that while similar models of the Nao robot are used in each simulator, there are certain differences. For example, it is known that certain problems exist in the precise positioning of the centre of mass (COM) of some parts of the simulated Nao in V-REP. Again, by using multiple simulators the expectation is that exactly similar problems will not exist in all simulators.

While the work in (Boeing and Bräunl, 2012), (Boeing, 2009) involved a parallel evolutionary process, with each individual being tested in parallel on several physics simulators, and the results obtained from the evaluations being combined to generate an overall fitness for the individual, we employ a serial evolutionary process, with each individual in a generation being evaluated initially on the V-REP simulator as part of the evolutionary process. Only successful individuals are then transferred to the Webots simulator for validation of their performance before being then transferred to the real robot.

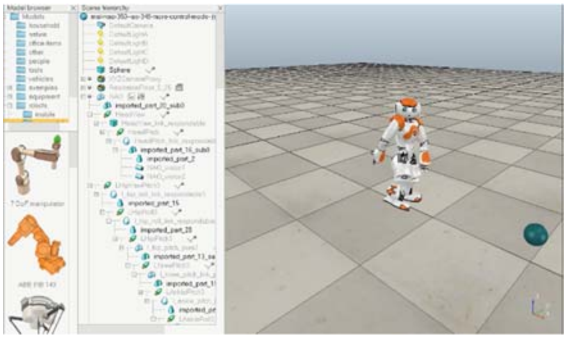

Figure 1: An example scene from the V-REP simulator. On the far left are examples of some of the robots that can be simulated, next to this is the scene hierarchy for the current scene, on the right is an example of an evolved kick.

2 FITNESS FUNCTION

For the evolution of ball-kicking behaviour we base our fitness function *f* on the distance travelled by the ball in the forward direction in the time allowed for each evaluation cycle. If the robot falls over mid-cycle we base the fitness on the distance travelled by the ball until the robot falls. This is to encourage stable and replicable kicking motions which should not cause undue strain to the real robot when transferred from simulation. So, if the robot fails to move the ball in the forward direction in the period in which it remains upright, or the time limit (T) expires, the fitness function is simply

$$f=100*t \tag{1}$$

where t is the time that the robot remains upright (t=T if the robot does not fall in the experimentation period). T is set at 5 seconds for all the experiments described here. If the robot does manage to move the ball some distance in the forward direction the fitness is then given by

$$f=(1+d)*100*t \tag{2}$$

where d is the distance travelled by the ball. This fitness function is designed in order to reward both the robot remaining upright, and the ball being moved in the forward direction. We note, of course, that no constraint of remaining upright is placed on human soccer players, it may indeed even be advantageous to a player to conclude a kicking motion on the ground in certain circumstances. The ball used in these experiments was of roughly similar diameter to that used in the RoboCup Standard Platform League (SPL).

3 EXPERIMENTAL DETAILS

3.1 Genome Composition

The genome length is 416 bits in total. This comprises 4 bits per joint angle (allowing for a total of 16 different angle positions per joint) for each of the 24 modelled joints of the robot, for each of 4 keyframe values. 4 keyframes were chosen as it was considered that this would be a sufficient number to characterise a complete kicking motion. While a certain amount of a-priori knowledge was involved in this decision, very little was specified about the joint values associated with each keyframe, apart from the fact that each joint has to keep within the maximum and minimum ranges as given by the specifications for the physical Nao robot. These maximum and minimum ranges are then modified by the joint restriction values evolved in each individual robots' genome as discussed below. Lower values of this 16-bit parameter correspond to higher joint ranges.

Typically this parameter starts at quite a low value early in the evolutionary process as the robot tends to move in a "thrashing" fashion, which may, or may not, cause ball movement. This value then increases as the robot restricts its joint movement range in order to increase the probability of not falling over due to "thrashing" motions. As the evolution progresses this value typically decreases gradually as the robot "frees up" its joints in order to more effectively perform the actions required (Eaton, 2007), (Eaton, 2013). As an example of this progression, for the experiment detailed in the next section the average value of this parameter for the best genome of generation 1 was 2.41, rising to a maximum value of 3.91 in generation 124, and reducing gradually to a value of 3 in generation 500.

A final 16 bits encodes movement durations for each of the 4 keyframes.

3.2 Keyframe Interpolation

The interpolation between the keyframe values is carried out by the V-REP and the Webots simulators themselves using inbuilt functions within each simulator. Once a sequence of 4 keyframes is completed the process cycles over until the time limit is exceeded or the run terminates for some other reason (the robot falling over).

3.3 Evolutionary Algorithm and Robot Control

The code for the evolutionary algorithm and robot control software in the V-REP simulator is written in the Lua programming language, while the corresponding controller for the Webots simulator and subsequent transfer to the real robot is written in Python. This transferral is semi-automated at present, however it is planned to fully automate this process in the future.

A population size of 120 was used using a mutation rate of 0.01 and a crossover probability of 0.2. These values were arrived at after some experimentation. The genetic algorithm employs tournament selection, and single-point crossover. It also employs elitism, where the best individual of each generation is guaranteed safe passage to the next generation.

4 EXPERIMENTAL RESULTS

4.1 Evolution of Kicking Behaviour

Three runs over 500 generations were performed, each taking about a day to complete on a 64-bit Dell 2.3GHz XPS 15 computer with an Intel i7 quad-core CPU and 16GB of RAM. The results obtained were then averaged to produce the fitness graph as shown in Fig.2.

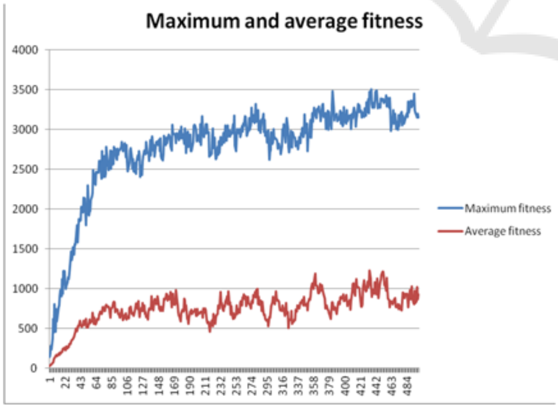

Figure 2: Maximum and average fitness, averaged over three runs for 500 generations for the evolution of ball kicking behaviour.

An effective kicking behaviour involves learning to stand on one foot, and the maintenance of balance on this foot while delivering a substantial blow to the ball with the other foot. For our work we also wish to maintain this balance (i.e. the robot does not fall over on completion of the kicking motion), if possible. Once the robot has learned to maintain its balance its kicking efficiency (as measured by the distance travelled by the ball) increases quite rapidly up to about generation 100, with more modest gains thereafter. Fig. 3 gives an example of an evolved kick as modelled in the Webots simulator.

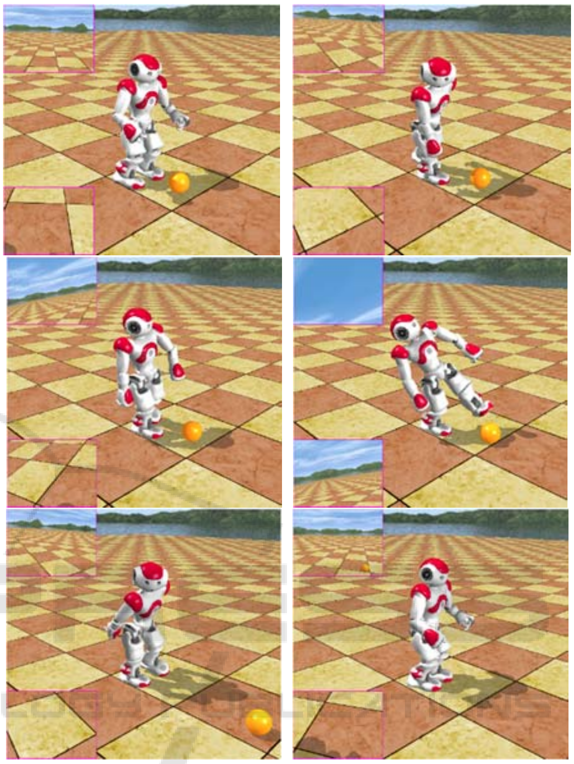

Figure 3: An example of an evolved kick, as transferred from V-REP to the Webots simulator; read from *top left* to *bottom right*.

Fig. 4 then shows this kick as transferred to the Nao humanoid robot using the procedure outlined earlier. The main portion of this kicking motion transfers directly onto the robot without need for human intervention. The only minor point of instability occurs in the final steadying motion before the robot comes to rest on completion of the kick, we conjecture that this is due to a friction mismatch between the surface the real robot rests on, and the values used in the V-REP and Webots simulators. However the majority of the kicking behaviour transferred directly to the robot resulting in an effective and quite human-like striking action. The entire behaviour sequence depicted in Fig.4 took place without the need for human intervention.

Effective kicking behaviour evolved in all three runs. In one of the runs predominately right-footed kicks were evolved, whereas a left-footed kick, as demonstrated in Fig. 3 and Fig. 4, was evolved in the other two runs, thus demonstrating the robustness and flexibility of our approach.

It should be noted that tests conducted on the real Nao robot were conducted at a reduced speed than the V-REP and Webots environments to reduce the likelihood of damage to the robot; it was also found that a behaviour was more likely to transfer successfully from simulated to real robot if conducted at a lower speed. However this reduction of speed was not, in general, found to cause a major diminution in the effectiveness of the behaviours evolved.

4.2 Validation of Our Approach

As an additional preliminary test of the effectiveness of our approach we chose 10 genomes at random from the first of the 3 runs. All of the fitness's of the evolved behaviours were the best of their generation and were around the 3000 mark, corresponding to an effective kick as evolved in the V-REP simulator.

Of these 10 behaviours two resulted in consistently unstable behaviour over several evaluations in the Webots environment. When transferred to the real Nao humanoid unstable motions also resulted, with the robot falling over and having to be manually restrained to avoid damage to the robot.

Of the remaining 8 motions three resulted in motions in which the either the robot either fell over in one of the Webots test evaluations, or, while not falling over exhibited significant instability at some point in the sequence. Of these three test runs, when transferred to the real Nao robot, two resulted instability (robot falling over), and one corresponded to an effective kicking motion, however exhibiting significant instability towards the end of the motion sequence.

Five of the 10 motions tested in the Webots simulator resulted in effective stable kicking motions. Of these 5 motions when transferred to the real Nao, all but one resulted in successful stable kicks.

Based on the results of these initial experiments we would not now generally consider testing any evolved motion on the real Nao robot that had not been successful in both the V-REP and Webots environments, to avoid potential strain on the Nao robots' actuators and/or actual damage to the robot or its environs.

5 CONCLUSIONS

In this paper we have demonstrated the evolution of kicking behaviour in the Nao humanoid robot. Using a novel dual-simulator approach with a fitness function based solely on the stability of the robot and the distance travelled by the ball, effective kicking behaviours were developed which were demonstrated to transfer, with reasonable fidelity, to the real Nao robot.

To our knowledge this is the first time a multi-simulator approach to the evolution of robot behaviours in the manner described in this paper. Also to our knowledge this is one of the few, if not the only, work which involves the evolution of kicking behaviours for direct transferral to the real Nao humanoid, rather than for use in the RoboCup simulation environment.

Figure 4: The evolved kick from Fig.3 as transferred to the real Nao robot. Compare these whole-body motions with the first four frames of Fig. 3.

We intend to extend the approach presented in this paper to the evolution of further behaviours of a more complex nature, including involving multiple robots.

ACKNOWLEDGEMENTS

My thanks to Jason Brownlee for his Lua GA implementation which provided the inspiration for our GA coding. My appreciation also to Norah Power for her assistance in the experimental phase of this work. Finally, also thanks to the reviewers of this paper for their helpful and constructive comments.

REFERENCES

Boeing, A., 2009. *Design of a Physics Abstraction Layer for Improving the Validity of Evolved Robot Control Simulations.* Ph.D Dissertation, School of Electrical, Electronic and Computer Engineering, The University of Western Australia, WA, 2009.

Boeing, A., and Bräunl, T., 2012. Leveraging multiple simulators for crossing the reality gap. In: *Control Automation Robotics and Vision (ICARCV), 2012 12th International Conference on* (pp. 1113–1119). IEEE.

Bongard, J. C. , 2013. Evolutionary robotics. *Communications of the ACM*, 56(8), 74–83.

Bongard, J. C., and Lipson, H., 2004. Once more unto the breach: Co-evolving a robot and its simulator. In: *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems* (ALIFE9) (pp. 57–62).

Depinet, M., MacAlpine, P., & Stone, P. ,2014. Keyframe sampling, optimization, and behavior integration: Towards long-distance kicking in the robocup 3d simulation league. In *RoboCup 2014: Robot World Cup XVIII* (pp. 571-582). Springer International Publishing.

Duarte, M., Oliveira, S., and Christensen, A. L., 2012. Automatic synthesis of controllers for real robots based on preprogrammed behaviors. In: *From Animals to Animats 12* (pp. 249–258). Springer Berlin Heidelberg.

Eaton, M., 2007. Evolutionary humanoid robotics: past, present and future, In: *50 Years of Artificial Intelligence: Essays Dedicated to the 50th Anniversary of Artificial Intelligence LNAI 4850*, Springer, pp. 42–53.

Eaton, M., 2013. An Approach to the Synthesis of Humanoid Robot Dance Using Non-interactive Evolutionary Techniques. In: *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on* (pp. 3305–3309). IEEE.

Eaton, M., 2015. *Evolutionary Humanoid Robotics*. Springer Berlin Heidelberg.

Farchy, A., Barrett, S., MacAlpine, P., and Stone, P., 2013. Humanoid robots learning to walk faster: From the real world to simulation and back. In: *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems* (pp. 39–46). International Foundation for Autonomous Agents and Multiagent Systems.

Floreano, D., and Urzelai, J., 2001. Evolution of plastic control networks. *Autonomous robots, 11(3),* 311–317.

Freese, M.S., Singh, S, Ozaki, F, and Matsuhira N., 2010. Virtual robot experimentation platform v-rep: a versatile 3d robot simulator. In *Simulation, Modeling, and Programming for Autonomous Robots,* pages 51–62. Springer, 2010.

Gouaillier, D., Hugel, V., Blazevic, P., Kilner, C., Monceaux, J., Lafourcade, P., Mariner, B., Serre, J., and Maisonnier, B.,2009 . Mechatronic design of NAO humanoid. In: *Robotics and Automation, 2009. ICR'09. IEEE International Conference on* (pp. 769–774). IEEE.

Iocchi, L., Libera, F. D., and Menegatti, E. , 2007. Learning Humanoid soccer actions interleaving simul ated and real data, in: *Proc. of The Second Workshop on Humanoid Soccer Robots, IEEE-RAS 7th International Conference on Humanoid Robots,* Pittsburgh, 2007

Jakobi, N. , 1997b. Half-baked, ad-hoc and noisy: minimal simulations for evolutionary robotics. In: P. Husbands and I. Harvey, *Proceedings of the Fourth European Conference on Artificial Life*. Cambridge, MA: MIT Press.

Jakobi, N., 1997a. Evolutionary robotics and the radical envelope-of-noise hypothesis. *Adaptive Behavior, 6(2),* 325–368.

Jouandeau, N., & Hugel, V. ,2014. Optimization of parametrised kicking motion for humanoid soccer player. In *Autonomous Robot Systems and Competitions (ICARSC), 2014 IEEE International Conference on* (pp. 241-246). IEEE.

Kitano, H., and Asada, M.,1998. RoboCup humanoid challenge: That's one small step for a robot, one giant leap for mankind. In: *Intelligent Robots and Systems, 1998. Proceedings, 1998 IEEE/RSJ International Conference on* (Vol. 1, pp. 419–424). IEEE.

Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I., Osawai, E., and Matsubara, H., 1998. Robocup: A challenge problem for AI and robotics. In: *RoboCup-97: Robot Soccer World Cup I* (pp. 1–19). Springer Berlin Heidelberg.

Koos, S., Mouret, J. B., and Doncieux, S..,2013. The transferability approach: Crossing the reality gap in evolutionary robotics. *Evolutionary Computation, IEEE Transactions on, 17(1),* 122–145.

Laue, T., and Hebbel, M., 2009. Automatic parameter optimization for a dynamic robot simulation. In: *RoboCup 2008: Robot Soccer World Cup XII* (pp. 121–132). Springer Berlin Heidelberg.

Li, X., Liang, Z., & Feng, H. ,2015. Kicking motion planning of Nao robots based on CMA-ES. *In Control and Decision Conference (CCDC), 2015* 27th Chinese (pp. 6158-6161). IEEE.

Lipson, H., Bongard, J. C., Zykov, V., and Malone, E., 2006. Evolutionary Robotics for Legged Machines: From Simulation to Physical Reality. In: Arai, T. et al (eds.*), Intelligent Autonomous Systems 9(IAS-9)*, pp.11–18.

Michel, O., 2004. Webots: Professional Mobile Robot Simulation. *International Journal of Advanced Robotic Systems, Vol. 1, No. 1*, 39–42.

Nolfi, S., and Floreano, D.,2000. *Evolutionary robotics. The biology, intelligence, and technology of self-organizing machines*. MIT Press.

Zagal, J. C., and Ruiz-Del-Solar, J., 2007. Combining simulation and reality in evolutionary robotics. *Journal of Intelligent and Robotic Systems, 50(1)*, 19–39.

Zagal, J. C., Ruiz-del-Solar, J., and Vallejos, P., 2004. Back to reality: Crossing the reality gap in evolutionary robotics. In: *IAV 2004 the 5th IFAC Symposium on Intelligent Autonomous Vehicles*, Lisbon, Portugal.