A Reinforcement Learning Approach for Goal-directed Locomotion of a Complex Snake-Like Robot

Sromona Chatterjee[1], Florentin Wörgötter[1], Minija Tamosiunaite[1], Poramate Manoonpong[1,2]
Yoshihide Enomoto[3], Ryo Ariizumi[3], Fumitoshi Matsuno[3]

1. Bernstein Center for Computational Neuroscience, Georg-August-Universität Göttingen, Germany.

2. Marsk Mc-Kinney Moller Institute, University of Southern Denmark, Denmark.

3. Department of Mechanical Engineering and Science, Kyoto University, Japan.

Snake-like robots conventionally use undulation movements for locomotion [1], [2]. In contrast, a nonstandard snake-like robot having four screw-drive units which are connected serially by three active joints [3] uses propulsion by rotating its screw-drive units to move. Employing this screw-drive mechanism, undulation is not required for its locomotion. This helps it to move in narrow spaces. However, generating its goal-directed locomotion is a challenging control problem. In this work, we tackle the problem by using a reinforcement learning approach, called Policy Improvement using Path Integrals (PI^2) [4]. PI^2 is numerically simple and has an ability to deal with high dimensional systems. Here, it is used as a model-free learning mechanism to find a proper combination of seven locomotion control parameters (three yaw joint angles and four screw unit angular velocities) of the robot for moving towards a given goal. The learning process is achieved using simulation and the learned parameters are successfully transferred to the real robot.

Experiments show that the robot with different body configurations (like, straight-line, zigzag, arc, etc.) can effectively move toward any given goal (see supplementary video at http://www.manoonpong.com/BCCN2014/SnakeRobot.wmv). In this way, a large repertoire of robot behaviors are obtained and used as motor primitives for generating other new behaviors online. By selecting different primitives and properly chaining them along with parameter interpolation and/or sensory feedback techniques, the robot can successfully handle complex tasks like, reaching a single goal or multiple goals while avoiding obstacles and compensating to a change of its body shape.

Taken together this study suggests how the PI^2 reinforcement learning approach can be used to solve coordination problem of many degrees-of-freedom systems, like the nonstandard snake-like robot, and to generate goal-directed locomotion in a complex environment.

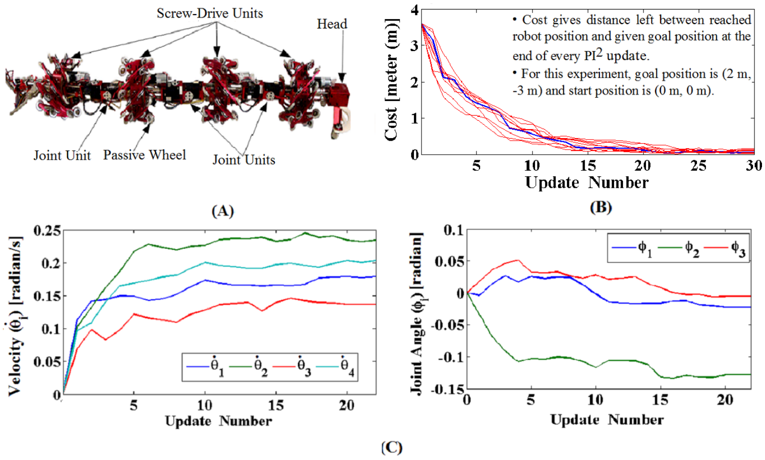

Figure 1: (A) Snake-like robot with screw-drive units. (B) Learning converges to lowest cost for all 10 runs, taking 15 updates for the average run in blue. (C) Learning of 4 screw velocities and 3 joint angles to give final parameters for reaching the goal.

Acknowledgments

We acknowledge Emmy Noether grant MA4464/3-1 of the Deutsche Forschungsgemeinschaft, Bernstein Center for Computational Neuroscience II Göttingen (BCCN grant 01GQ1005A, project D1), HeKKSaGOn network.

References

[1] S. Hirose, ”Biologically Inspired Robots: Snake-Like Locomotors and Manipulators”, Oxford University Press, Oxford, 1993
[2] P. Liljebck, K.Y. Pettersen, . Stavdahl and J.T. Gravdahl, ”A Review on Modeling, Implementation, and Control of Snake Robots”, Robotics and Autonomous Systems, vol. 60 (1), pp. 29–40, 2012
[3] H. Fukushima, S. Satomura, T. Kawai, M. Tanaka, T. Kamegawa, and F. Matsuno, ”Modeling and Control of a Snake-Like Robot Using the Screw-Drive Mechanism”, IEEE Trans. Robot., vol. 28 (3), pp. 541–555, 2012
[4] E. Theodorou, J. Buchli, and S. Schaal, ”A Generalized Path Integral Control Approach to Reinforcement Learning”, Journal of Machine Learning Research, vol. 11, pp. 3137–3181, 2010