Learning Powerful Kicks on the Aibo ERS-7: The Quest for a Striker

Matthew Hausknecht and Peter Stone

{mhauskn,pstone}@cs.utexas.edu

Learning Agents Research Group

AUSTIN VILLA ROBOT SOCCER TEAM

THE UNIVERSITY OF TEXAS AT AUSTIN

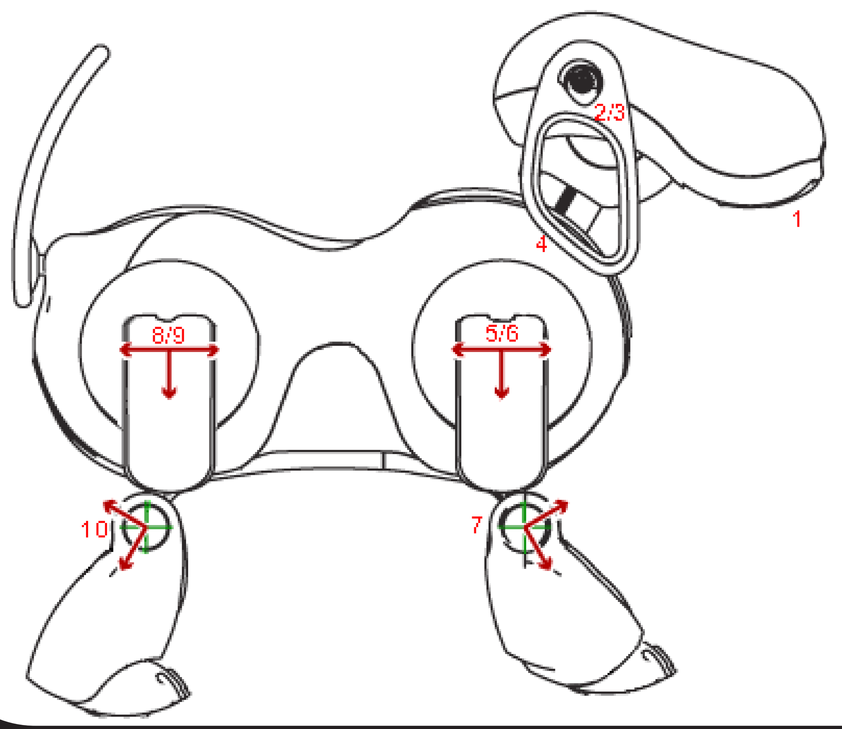
Abstract

- Machine learning applied to optimize kick power entirely on the real robot – Aibo ERS-7.
- Learned kick significantly more powerful than UT Austin Villa's best hand-coded kick.
- Model inversion to create a parameterized, variable distance kick.

Novelty

- First application of machine learning to the problem of kick learning entirely on the real robot.
- Learned with a larger parameter space than previous non-simulation work. [2] Very little domain engineering required.

Kick Parameterization



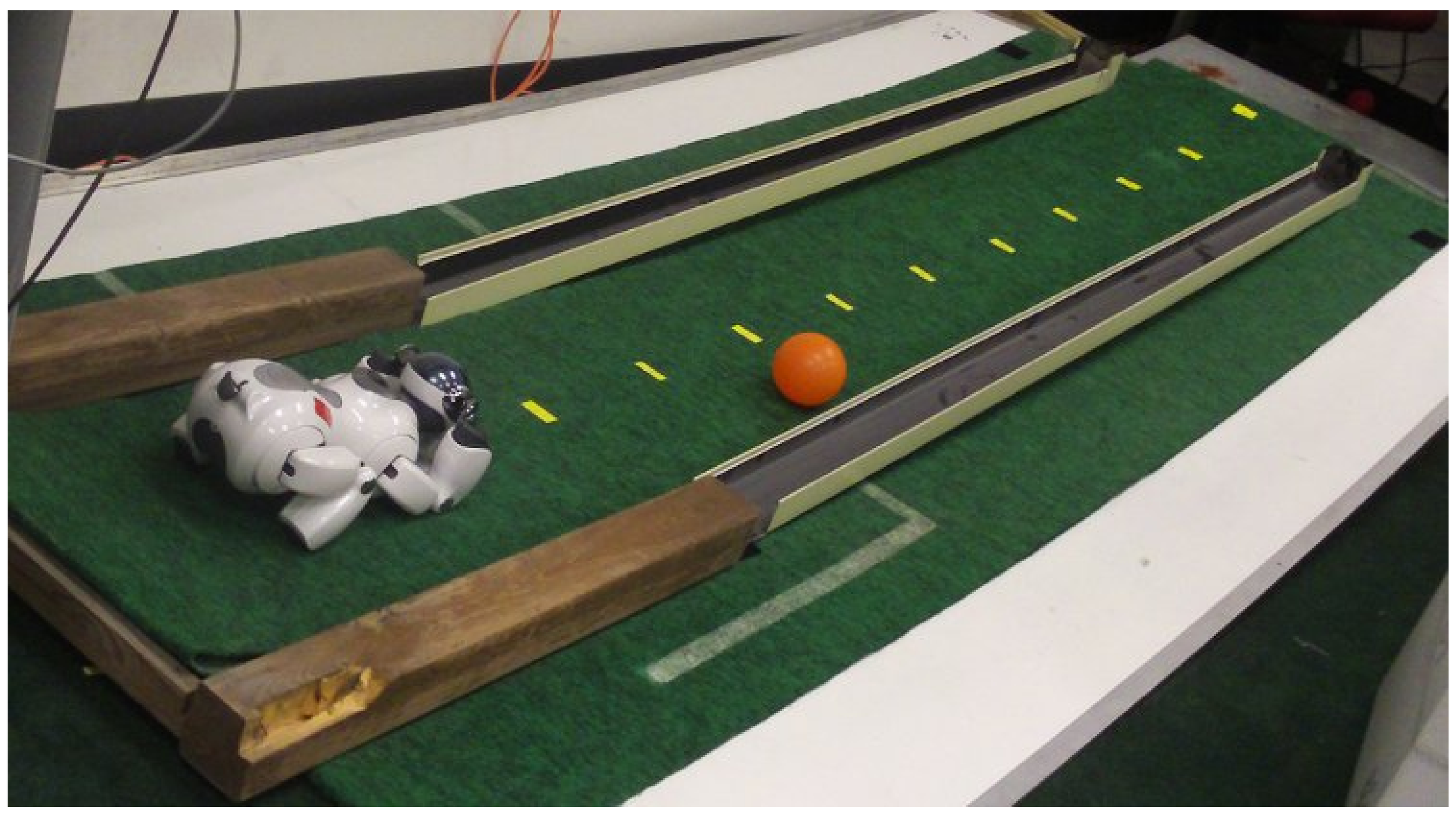
- 66 total parameters learned: 6 poses with 10 joints each + 6 pose timing parameters.

Parameter Space Reductions:

- Using bilateral symmetry, leg joints mirror each other.
- Eliminated both tail joints.
- Minimal domain specific knowledge required.

Learning Framework

- Adjustable incline ramp created for semi-autonomous learning.



Inclined ramp for optimizing kick distance

- Human required to reposition Aibo after each kick.
- Two metrics for kick power: time since kick until ball returns, kick distance up ramp.
- Speed: 7-8 seconds per trial.

Learning Algorithms

- Hill Climbing: 5 policies evaluated per iteration.
- Policy Gradient [1]: 10 policies per iteration, $\eta = 2.0$.
- Incremented random joint angle: $random(0, \frac{1}{10}) * jointRange$.

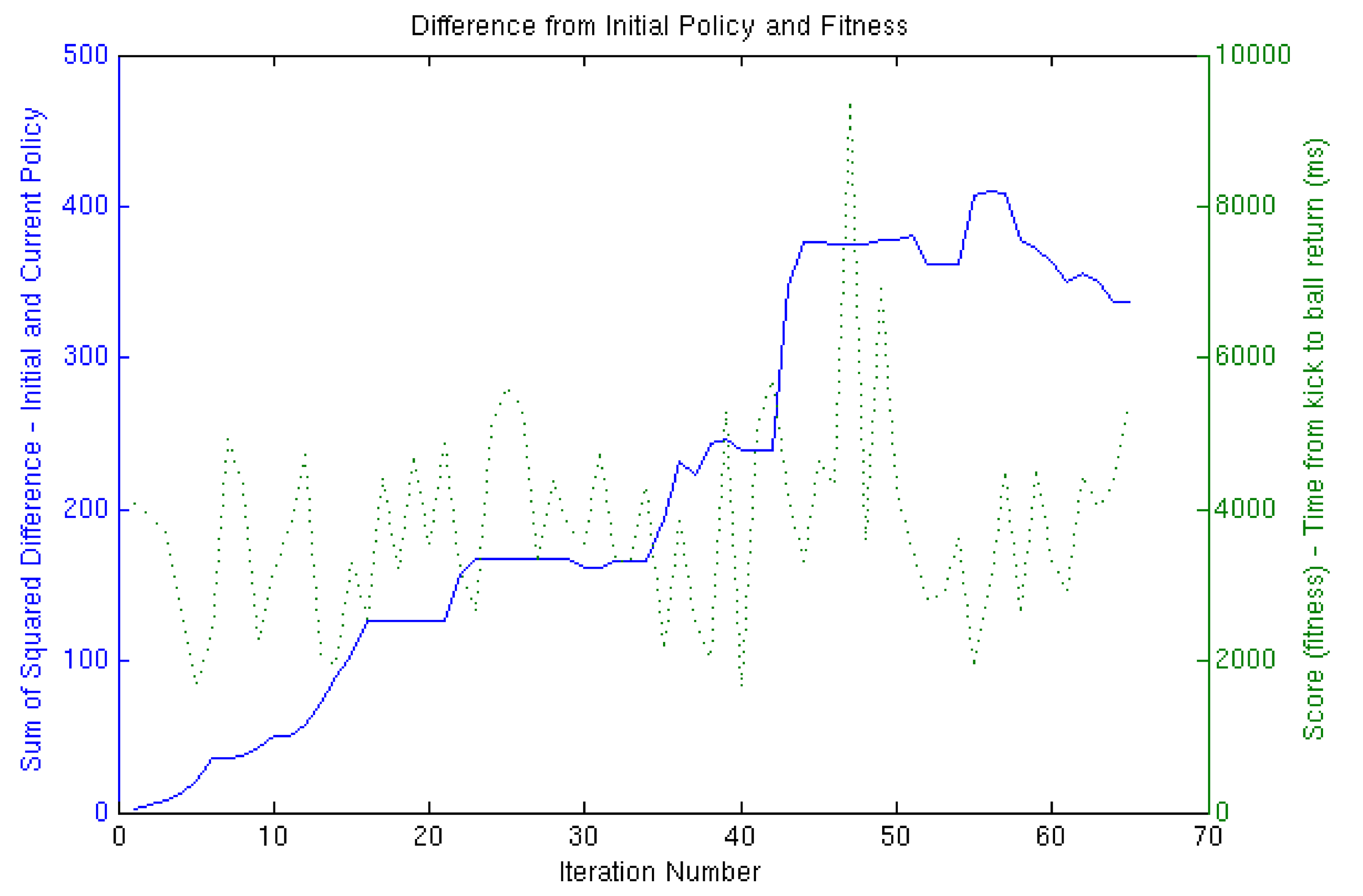
References

[1] N. Kohl, P. Stone. Machine learning for fast quadruped locomotion. In *Nineteenth National Conference on Artificial Intelligence* 2004

[2] S. Chernova, M. Veloso. An evolutionary approach to gait learning for four-legged robots. In *IROS'04* 2004

Learning

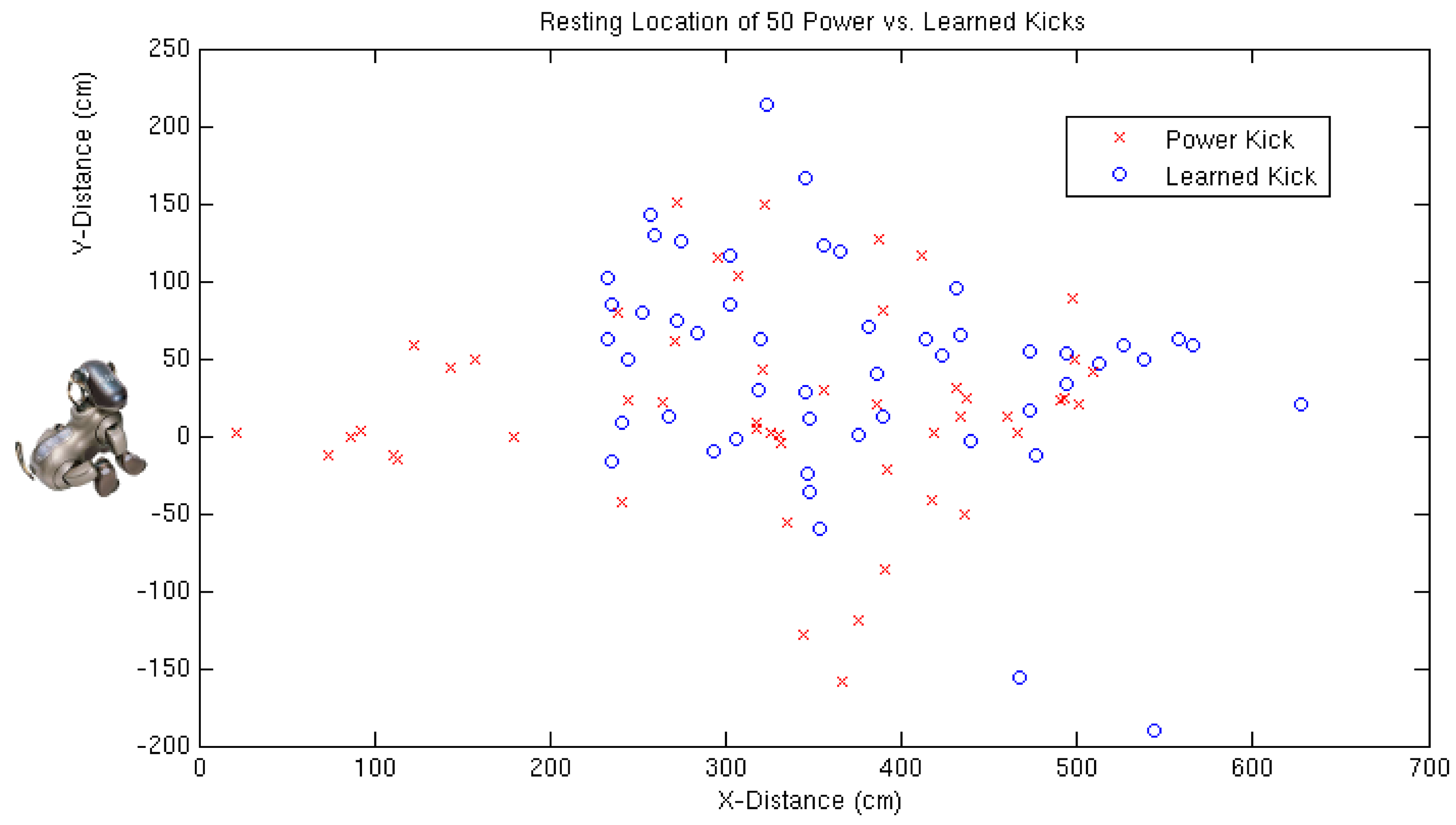
- Initial Policy: UT Austin Villa's Power Kick
- Policy Gradient run for 65 iterations (650 kicks)
- Subsequently, Hill Climbing used for 27 iterations on best policy.



Policy Gradient learning from power kick

Results

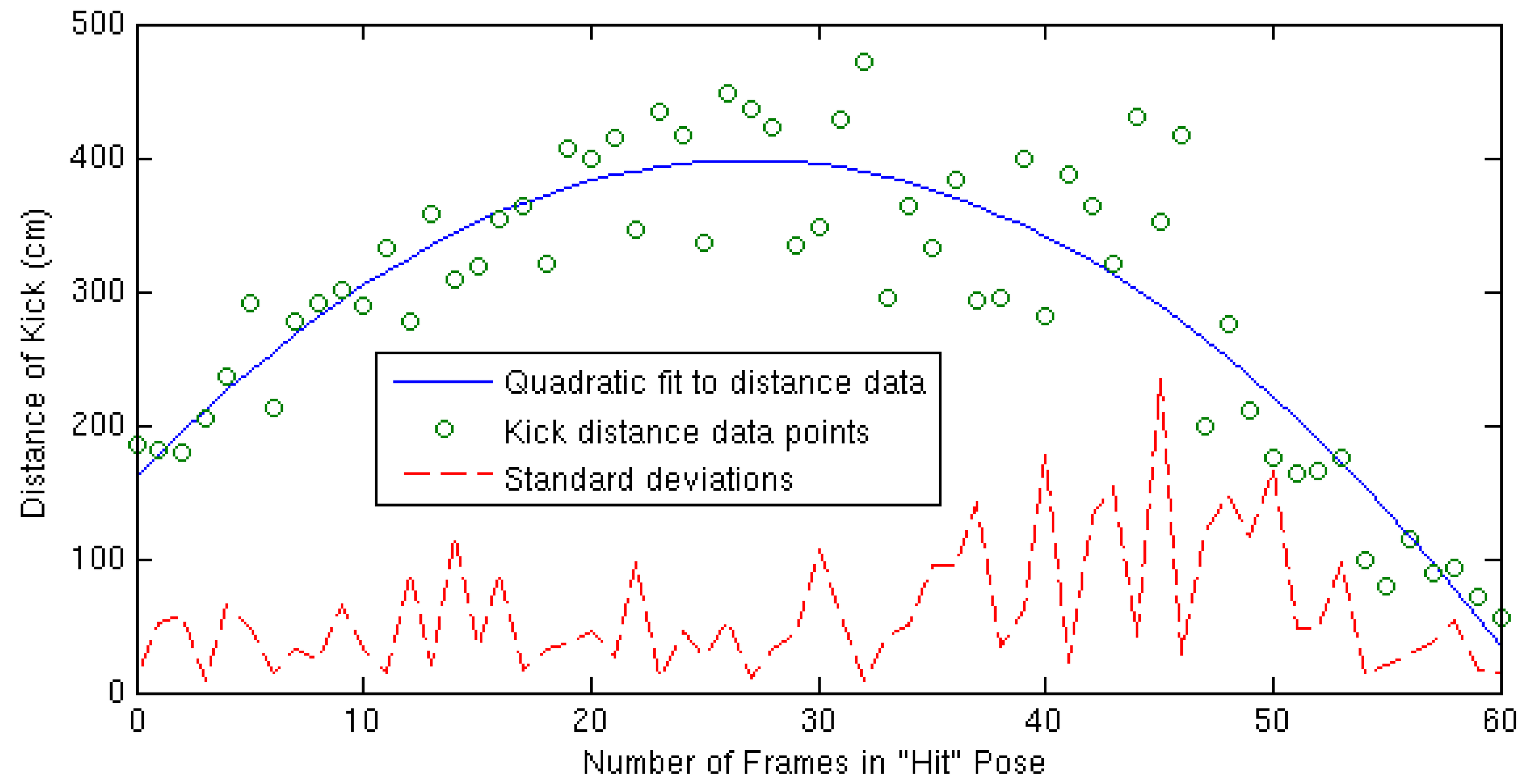
- Evaluated on 10 different Aibos with 5 kicks per Aibo.
- Learned Kick significantly more powerful than Power Kick: 373cm vs. 322cm on average.



Resting Location of 50 Power and Learned Kicks

Variable Distance Kick

- Model inversion applied to create variable distance kick.
- Identified "hit" pose in which Aibo makes contact with the ball.
- Varying the time in this pose was observed to change kick power.
- Quadratic curve fit to kick distance data points.
- Accurate to within 58cm of requested distance.



Quadratic function models the relationship between kick distance and pose time.

Acknowledgements

This work has taken place in the Learning Agents Research Group (LARG) at the Artificial Intelligence Laboratory, The University of Texas at Austin. LARG research is supported in part by grants from the National Science Foundation (CNS-0615104 and IIS-0917122), ONR (N00014-09-1-0658), DARPA (FA8650-08-C-7812), and the Federal Highway Administration (DTFH61-07-H-00030).