Design and Optimization of an Omnidirectional Humanoid Walk: A Winning Approach at the RoboCup 2011 3D Simulation Competition

Patrick MacAlpine, Samuel Barrett, Daniel Urieli, Victor Vu, and Peter Stone

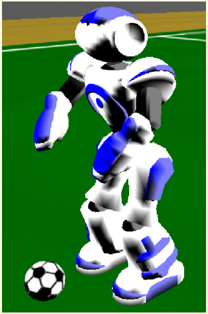
Department of Computer Science, The University of Texas at Austin

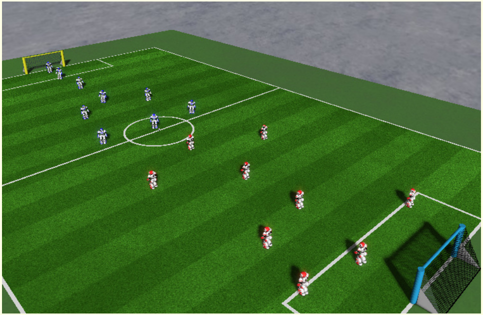
July 26, 2012

LARG
Learning Agents Research Group
The University of Texas at Austin

2011 RoboCup 3D Simulation Domain

- Teams of 9 vs 9 autonomous robots play soccer
- Realistic physics using Open Dynamics Engine (ODE)
- Simulated robots modeled after Aldebaron Nao robot
- Robot receives noisy visual information about environment
- Robots can communicate with each other over limited bandwidth channel





Competition Results

RoboCup	2010	2011
Goals For:	11	
Goals Against:	17	
Record (W-L-T):	4-5-1	
Place:	Outside Top-8	

Competition Results

RoboCup	2010	2011
Goals For:	11	136
Goals Against:	17	
Record (W-L-T):	4-5-1	
Place:	Outside Top-8	

Competition Results

RoboCup	2010	2011
Goals For:	11	136
Goals Against:	17	0
Record (W-L-T):	4-5-1	
Place:	Outside Top-8	

Competition Results

RoboCup	2010	2011
Goals For:	11	136
Goals Against:	17	0
Record (W-L-T):	4-5-1	24-0-0
Place:	Outside Top-8	

RoboCup	2010	2011
Goals For:	11	136
Goals Against:	17	0
Record (W-L-T):	4-5-1	24-0-0
Place:	Outside Top-8	1st

Competition Results

RoboCup	2010	2011
Goals For:	11	136
Goals Against:	17	0
Record (W-L-T):	4-5-1	24-0-0
Place:	Outside Top-8	1st

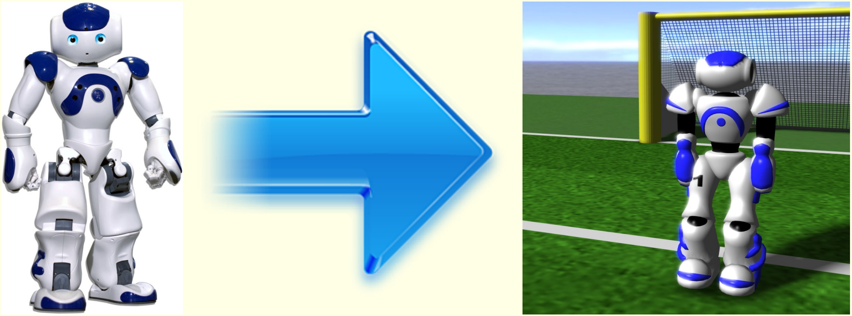
BIG IMPROVEMENT!

Motivation (2010 Walk)

- Consists of many fixed frame based skills
- Unable to quickly react (not omnidirectional)
- Not as stable as desired (completely open loop)

Omnidirectional Walk Engine

- Double linear inverted pendulum model
- Based closely on that of walk engine by Graf et al
- Mostly open loop but not entirely
- Designed on actual Nao robot



Walk Engine Parameters

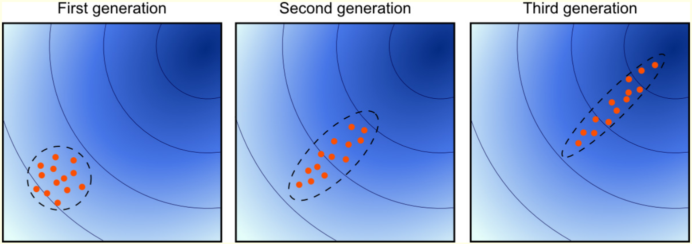
Notation	Description
$\mathbf{maxStep_i}$	Maximum step sizes allowed for x, y, and θ
$\mathbf{y_{shift}}$	Side to side shift amount with no side velocity
$\mathbf{z_{torso}}$	Height of the torso from the ground
$\mathbf{z_{step}}$	Maximum height of the foot from the ground
$\mathbf{f_g}$	Fraction of a phase that the swing foot spends on the ground before lifting
f_a	Fraction that the swing foot spends in the air
$\mathbf{f_s}$	Fraction before the swing foot starts moving
f_m	Fraction that the swing foot spends moving
$\boldsymbol{\phi_{length}}$	Duration of a single step
$\boldsymbol{\delta}$	Factors of how fast the step sizes change
y_{sep}	Separation between the feet
$\mathbf{x_{offset}}$	Constant offset between the torso and feet
$\mathbf{x_{factor}}$	Factor of the step size applied to the forwards position of the torso
$\mathbf{err_{norm}}$	Maximum COM error before the steps are slowed
$\mathbf{err_{max}}$	Maximum COM error before all velocity reach 0

Parameters of the walk engine with the optimized parameters shown in bold

Initial Walk Parameters

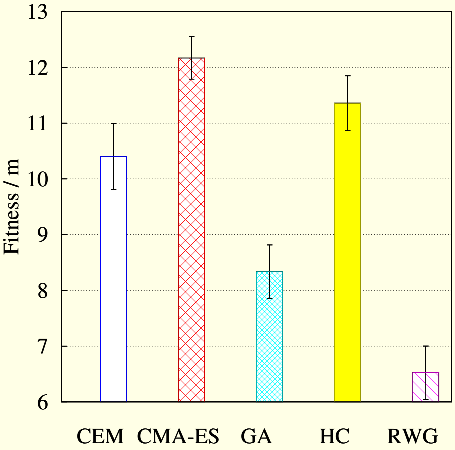
- Designed and hand-tuned to work on the actual Nao robot
- Provides a slow and stable walk

CMA-ES (Covariance Matrix Adaptation Evolutionary Strategy)



- Evolutionary numerical optimization method
- Candidates sampled from multidimensional Gaussian and evaluated for their fitness
- Weighted average of members with highest fitness used to update mean of distribution
- Covariance update using evolution paths controls search step sizes

Learning Algorithms Evaluation



CEM	Cross Entropy Method
CMA-ES	Covariance Matrix Strategy Evolutionary Strategy
GA	Genetic Algorithm
HC	Hill Climbing
RWG	Random Weight Guessing

Drive Ball to Goal Optimization

- Parameters (14) optimized through CMA-ES across a cluster
  - Population of 150 across 200 generations = 210,000 evaluations in less than a day
- Reward = distance robot dribbles ball toward goal in 30 seconds
- Wins by an average goal difference of 5.54 against *Initial* agent and 2.99 against *2010 walk* agent

Problems with Drive Ball to Goal Optimization

- Agent not that fast
  - .43 m/s compared to .6 m/s speed of 2010 walk
- Agent unstable when stopping
- Agent overfits to when dribbling is going well

Walk Forward Agent

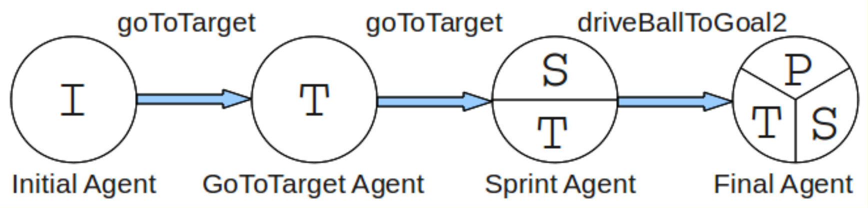
- Robot walks forward for 10 seconds from a complete stop
- Reward = distance robot walks forward
- Faster speed of .78 m/s up from .43 m/s

Attempts to transition between *driveBallToGoal* walk parameters (red 'D') and new *walkForward* parameters (yellow 'F')

Multiple Walk Parameter Set Optimization

- Learn three different parameter sets for three different subtasks
  - Going to a target
  - Sprinting forward
  - Positioning around the ball when dribbling

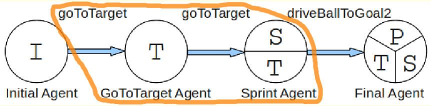
- Parameters learned through a layered learning approach
  - Parameter sets learned sequentially
  - Each parameter set learned *in conjuction* with each other
  - Robot able to seamlessly transition between parameter sets



I = *initial*, T = *goToTarget*, S = *sprint*, P = *positioning*

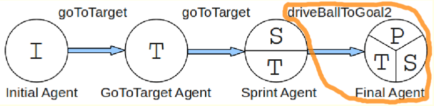
Go to Target Optimization

- *GoToTarget* agent wins on average by 2.04 goals against *DriveBallToGoal* agent
- *GoToTarget* agent speed at .64 m/s and with *Sprint* agent increased to .71 m/s



Drive Ball To Goal 2 Optimization

- Dribble ball toward goal for 15 seconds from multiple starting points around ball
- *Final* agent wins by an average goal difference of .15 against *Sprint* agent



Final Agent Video

Red 'T' = *gotoTarget* parameters, yellow 'S' = *sprint* parameters, cyan 'P' = *positioning* parameters

Walk Agent Performance

Game results of agents with different walk parameter sets. Entries show the average goal difference (row − column) from 100 ten minute games. Values in parentheses are the standard error.

	Initial	DriveBallToGoal	GoToTarget
Final	8.84(.12)	2.21(.12)	.24(.08)
GoToTarget	8.82(.11)	2.04(.11)	
DriveBallToGoal	5.54(.14)		

Walk Agent Performance

Game results of agents with different walk parameter sets. Entries show the average goal difference (row − column) from 100 ten minute games. Values in parentheses are the standard error.

	Initial	DriveBallToGoal	GoToTarget
Final	8.84(.12)	2.21(.12)	.24(.08)
GoToTarget	8.82(.11)	2.04(.11)	
DriveBallToGoal	5.54(.14)		

- Final competition agent beat agent with 2010 walk by an average goal difference of 6.32 goals across 100 games

Competition Analysis

Average goal difference across 100 games against other agents in the competition

Rank	Team	Goal Difference
3	apollo3d	1.45 (.11)
5-8	boldhearts	2.00 (0.11)
5-8	robocanes	2.40 (0.10)
2	cit3d	3.33 (0.12)
5-8	fcportugal3d	3.75 (0.11)
9-12	magmaoffenburg	4.77 (0.12)
9-12	oxblue	4.83 (0.10)
4	kylinsky	5.52 (0.14)
9-12	dreamwing3d	6.22 (0.13)
5-8	seuredsun	6.79 (0.13)
13-18	karachikoalas	6.79 (0.09)
9-12	beestanbul	7.12 (0.11)
13-18	nexus3d	7.35 (0.13)
13-18	hfutengine3d	7.37 (0.13)
13-18	futk3d	7.90 (0.10)
13-18	naoteamhumboldt	8.13 (0.12)
19-22	nomofc	10.14 (0.09)
13-18	kaveh/rail	10.25 (0.10)
19-22	bahia3d	11.01 (0.11)
19-22	l3msim	11.16 (0.11)
19-22	farzanegan	11.23 (0.12)

- Across 2100 games played won all but 21 games which ended in ties (no losses)

Competition Analysis

Average goal difference across 100 games against other agents in the competition

Rank	Team	Goal Difference
3	apollo3d	1.45 (.11)
5-8	boldhearts	2.00 (0.11)
5-8	robocanes	2.40 (0.10)
2	cit3d	3.33 (0.12)
5-8	fcportugal3d	3.75 (0.11)
9-12	magmaoffenburg	4.77 (0.12)
9-12	oxblue	4.83 (0.10)
4	kylinsky	5.52 (0.14)
9-12	dreamwing3d	6.22 (0.13)
5-8	seuredsun	6.79 (0.13)
13-18	karachikoalas	6.79 (0.09)
9-12	beestanbul	7.12 (0.11)
13-18	nexus3d	7.35 (0.13)
13-18	hfutengine3d	7.37 (0.13)
13-18	futk3d	7.90 (0.10)
13-18	naoteamhumboldt	8.13 (0.12)
19-22	nomofc	10.14 (0.09)
13-18	kaveh/rail	10.25 (0.10)
19-22	bahia3d	11.01 (0.11)
19-22	l3msim	11.16 (0.11)
19-22	farzanegan	11.23 (0.12)

- Across 2100 games played won all but 21 games which ended in ties (no losses)
- Agent with 2010 walk would have finished in tenth place

Walk in Competition

Action from the second half of the 2011 RoboCup 3D Simulation Final

Omnidirectional Walk Optimization Summary

Omnidirectional Walk Optimization Summary

- An optimization task should be representative of the overall task

Omnidirectional Walk Optimization Summary

- An optimization task should be representative of the overall task
- Parameter sets can be combined but must be learned *in conjunction* with each other

Omnidirectional Walk Optimization Summary

- An optimization task should be representative of the overall task
- Parameter sets can be combined but must be learned *in conjunction* with each other
- Machine learning using CMA-ES is very effective for parameter optimization

Related Work

- N. Hansen. The CMA Evolution Strategy: A Tutorial, January 2009.
- C. Graf, A. Härtl, T. Röefer, and T. Laue. A robust closed-loop gait for the standard platform league humanoid.
- N. Shafii, L. P. Reis, and N. Lao. Biped walking using coronal and sagittal movements based on truncated Fourier series, January 2010.
- J. E. Pratt. Exploiting Inherent Robustness and Natural Dynamics in the Control of Bipedal Walking Robots. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, June 2000.
- N. Kohl and P. Stone. Machine learning for fast quadrupedal locomotion, 2004.
- D. Urieli, P. MacAlpine, S. Kalyanakrishnan, Y. Bentor, and P. Stone. On optimizing interdependent skills: A case study in simulated 3d humanoid robot soccer, 2011.
- P. MacAlpine, D. Urieli, S. Barrett, S. Kalyanakrishnan, F. Barrera, A. Lopez-Mobilia, N. Stiurca, V. Vu, and P. Stone. UT Austin Villa 2011: A Winning Approach to the RoboCup 3D Soccer Simulation Competition, 2012.
- P. MacAlpine and P. Stone. Using Dynamic Rewards to Learn a Fully Holonomic Bipedal Walk, 2012.

Related Work

- N. Hansen. The CMA Evolution Strategy: A Tutorial, January 2009.
- C. Graf, A. Härtl, T. Röefer, and T. Laue. A robust closed-loop gait for the standard platform league humanoid.
- N. Shafii, L. P. Reis, and N. Lao. Biped walking using coronal and sagittal movements based on truncated Fourier series, January 2010.
- J. E. Pratt. Exploiting Inherent Robustness and Natural Dynamics in the Control of Bipedal Walking Robots. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, June 2000.
- N. Kohl and P. Stone. Machine learning for fast quadrupedal locomotion, 2004.
- D. Urieli, P. MacAlpine, S. Kalyanakrishnan, Y. Bentor, and P. Stone. On optimizing interdependent skills: A case study in simulated 3d humanoid robot soccer, 2011.
- P. MacAlpine, D. Urieli, S. Barrett, S. Kalyanakrishnan, F. Barrera, A. Lopez-Mobilia, N. Stiurca, V. Vu, and P. Stone. UT Austin Villa 2011: A Winning Approach to the RoboCup 3D Soccer Simulation Competition, 2012.
- P. MacAlpine and P. Stone. Using Dynamic Rewards to Learn a Fully Holonomic Bipedal Walk, 2012.

Related Work

- N. Hansen. The CMA Evolution Strategy: A Tutorial, January 2009.
- C. Graf, A. Härtl, T. Röefer, and T. Laue. A robust closed-loop gait for the standard platform league humanoid.
- N. Shafii, L. P. Reis, and N. Lao. Biped walking using coronal and sagittal movements based on truncated Fourier series, January 2010.
- J. E. Pratt. Exploiting Inherent Robustness and Natural Dynamics in the Control of Bipedal Walking Robots. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, June 2000.
- N. Kohl and P. Stone. Machine learning for fast quadrupedal locomotion, 2004.
- D. Urieli, P. MacAlpine, S. Kalyanakrishnan, Y. Bentor, and P. Stone. On optimizing interdependent skills: A case study in simulated 3d humanoid robot soccer, 2011.
- P. MacAlpine, D. Urieli, S. Barrett, S. Kalyanakrishnan, F. Barrera, A. Lopez-Mobilia, N. Stiurca, V. Vu, and P. Stone. UT Austin Villa 2011: A Winning Approach to the RoboCup 3D Soccer Simulation Competition, 2012.
- P. MacAlpine and P. Stone. Using Dynamic Rewards to Learn a Fully Holonomic Bipedal Walk, 2012.

Future Work

- Attempt to apply learned walks in simulation to actual Nao robots
- Experiment with other walk engines
- Model walk trajectories after those taken by human infants learning to walk
- Experiment with heterogenous robot models

More Information

UT Austin Villa 3D Simulation Team homepage:
www.cs.utexas.edu/~AustinVilla/sim/3dsimulation/

Email: patmac@cs.utexas.edu

This work has taken place in the Learning Agents Research Group (LARG) at UT Austin. LARG research is supported in part by NSF (IIS-0917122), ONR (N00014-09-1-0658), and the FHWA (DTFH61-07-H-00030).