Representation and Data-Driven Learning of Flexible Humanoid Robot Behaviours

Workshop on Speckled Computing – 2 Dec 2009

Subramanian Ramamoorthy

IPAB, School of Informatics

University of Edinburgh

Agenda

1. Introductory Remarks
2. Geometric Representations of Humanoid Behaviours
3. Skill Learning in a Non-stationary World
4. Future Directions

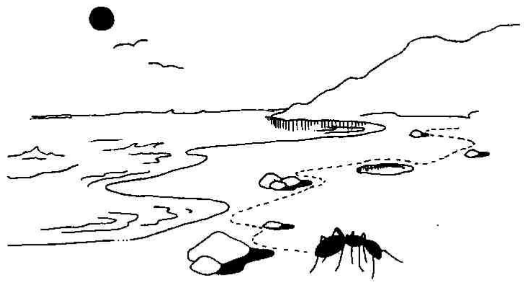
Towards Robustly Autonomous Agents

Research agenda:

Algorithms for decision making over time in autonomous agents (humanoid robots, trading agents, etc.)

- that can reliably achieve flexible behaviours (over large domains)
- in a constantly changing and uncertain/adversarial environment

- Simon's ant: complexity of behaviour is often related to interactions with external environment



- Conversely, if agent wants to achieve complex tasks under adversarial conditions, it needs suitably structured strategies

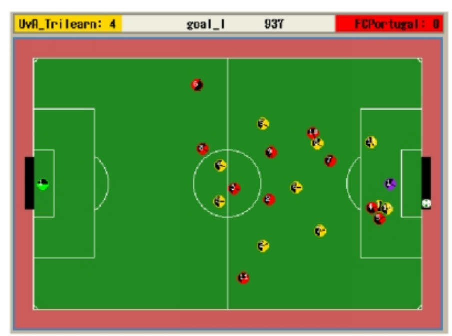
Autonomous Humanoid Robotics

We want humanoid robots to perform a number of different kinds of tasks:

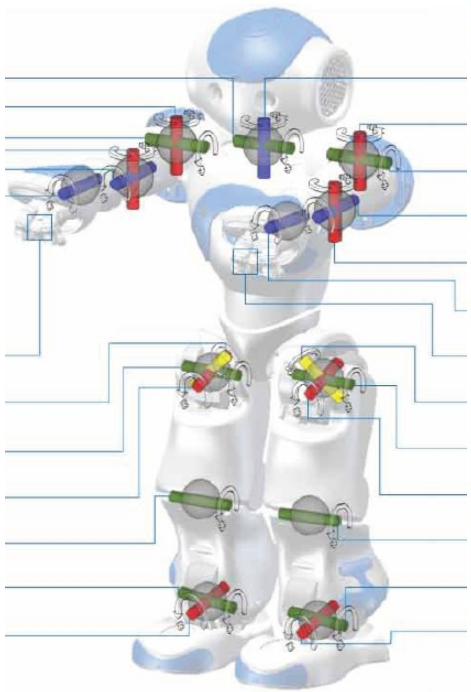
- Locomotion
- Manipulation
  - Full-body
  - Co-operative
- Higher level behaviours
  - Spatio-temporal movement
  - Information aggregation

Some representative domains we explore:

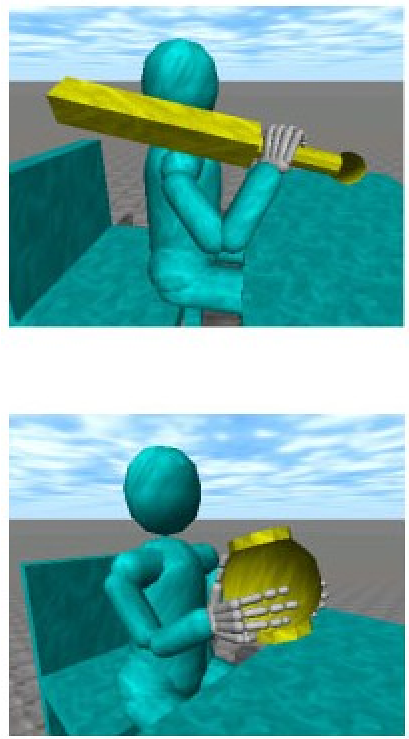
- Robotic soccer (RoboCup)
- Complex manipulation (e.g., flexible objects, cluttered environments)



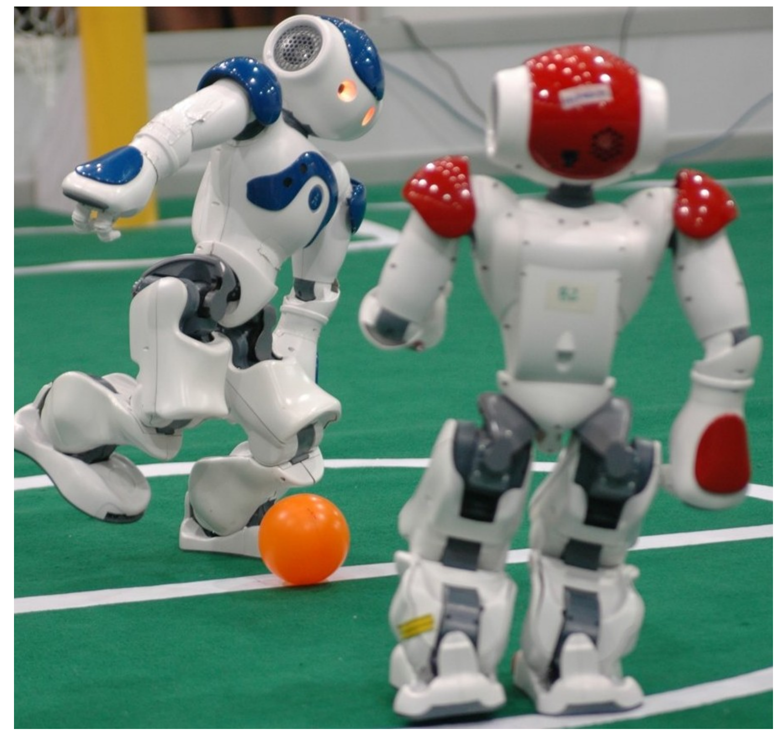
What Makes Humanoid Robotics Hard?



High-dim nonlinear dynamics (with imprecise models)



Complex constraints (again, with imprecise models) & Changing specifications



Strategic/adversarial Interactions with others

Two Major Questions

1. How to encode complex (dynamically dexterous) tasks, in the face of significant imprecision in models and specifications?
2. How to synthesize reactive strategies in the presence of strategic adversaries and in a non-stationary environment?

- There is a need for data-driven learning, e.g., from a combination of experience and expert demonstration
- Many direct approaches (ranging from RL in POMDPs to Bayesian statistical methods) can be intractable given combination of issues mentioned above

Our Approach

Two major techniques, at the high level:

1. Geometrical representations (low-dimensional, abstractions) of skills, e.g., manifolds

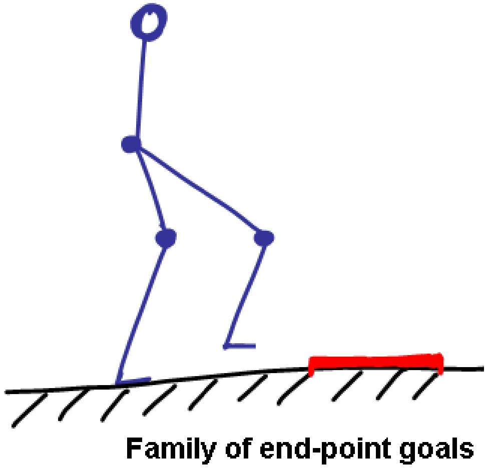
2. With the resulting abstractions, layered/hierarchical formulations of stochastic optimization/game problem:
   - trajectory optimization
   - strategic interactions with adversary

Agenda

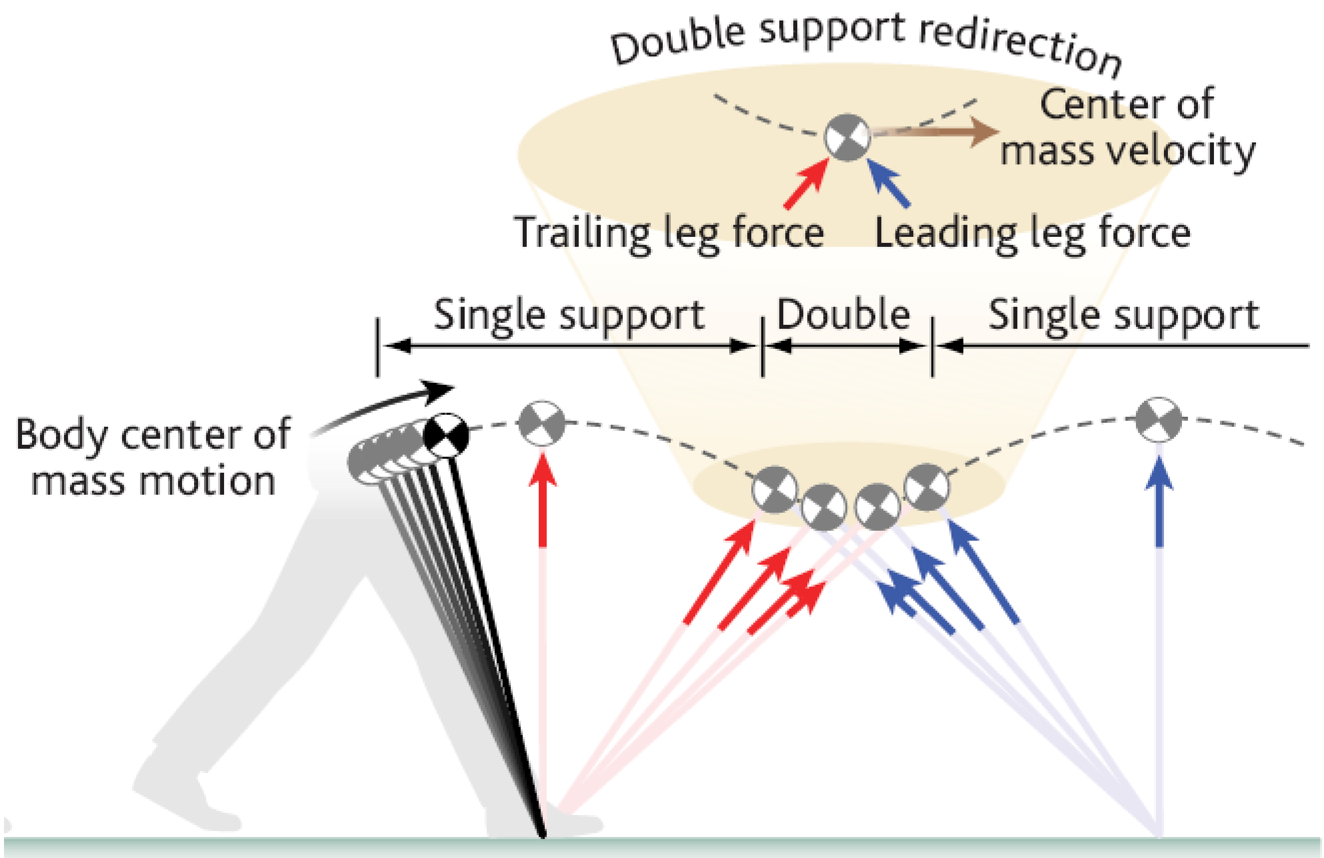
1. Introductory Remarks
2. Geometric Representations of Humanoid Behaviours
3. Skill Learning in a Non-stationary World
4. Future Directions

Canonical Task: Walking on Irregular Terrain

- No detailed models of dynamics
- Precisely specified footfalls
- Height/length variations
- Hard to **represent** & achieve with state of the art methods!



Compass Gait Walking: A Conceptual View



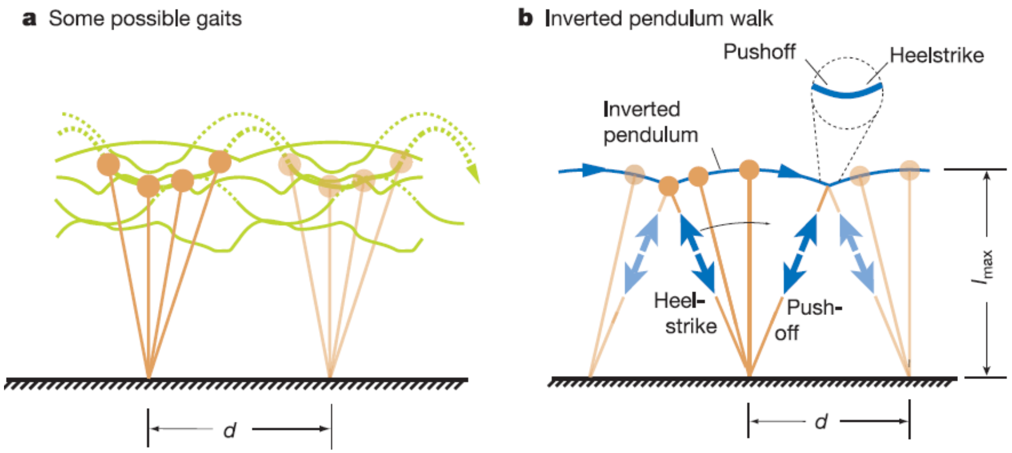
[Kuo, *Science* '05]

Why is the Compass Gait Model Meaningful?

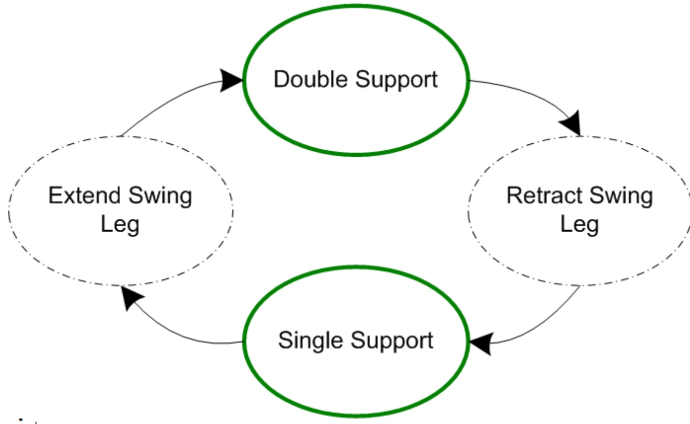
Biologically plausible template model

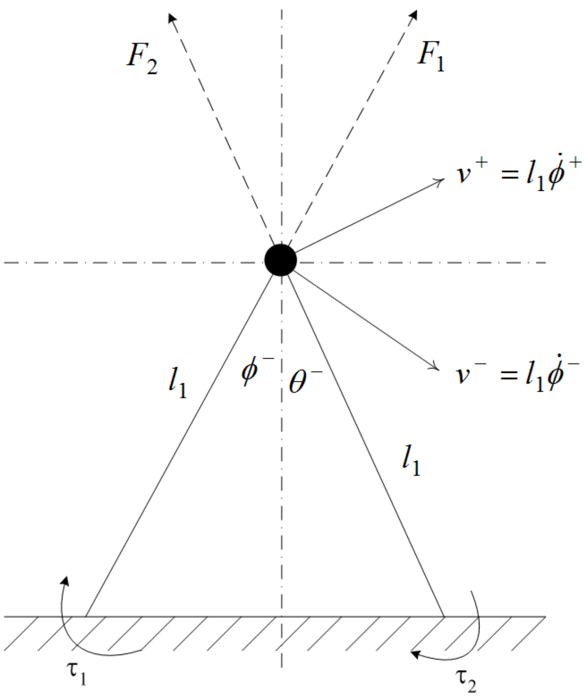
- Human energetics are pendulum-like (Cavagna, *JEB* '02)
- Infants learn this synergy before stable walking (Ivanenko, *JEB* '04)
- Neural representations for these variables (Poppele, *J. Neurophys.* '02)
- *Ab initio* optimization with this model recovers walking (Srinivasan & Ruina*, Nature '06)*

Question:
How does one use this model constructively, to plan trajectories on irregular terrain?

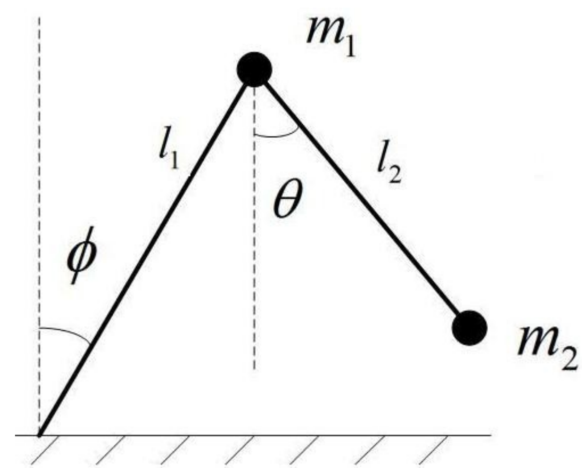


Compass Gait: Hybrid Dynamic Model



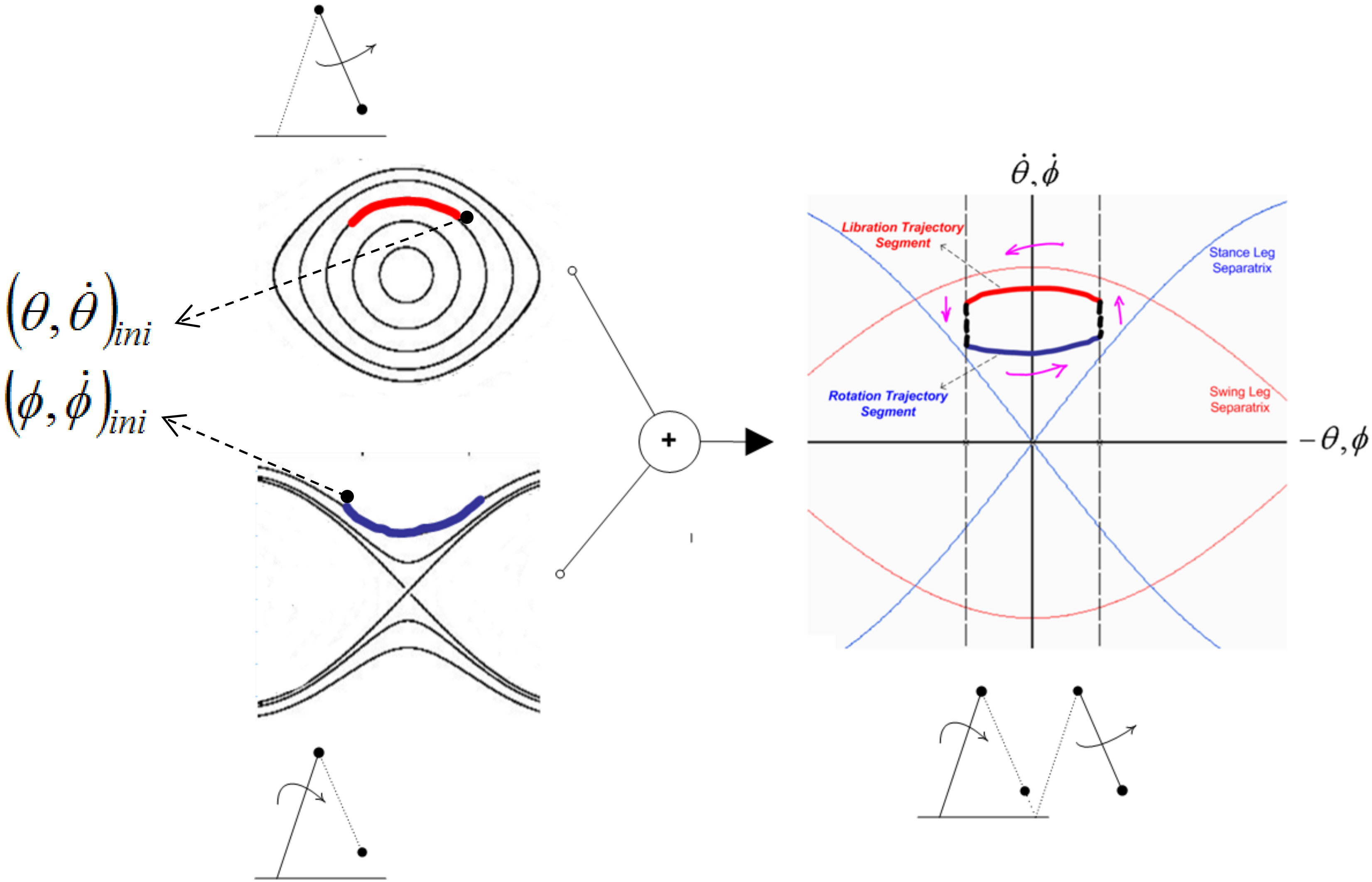


Double Support

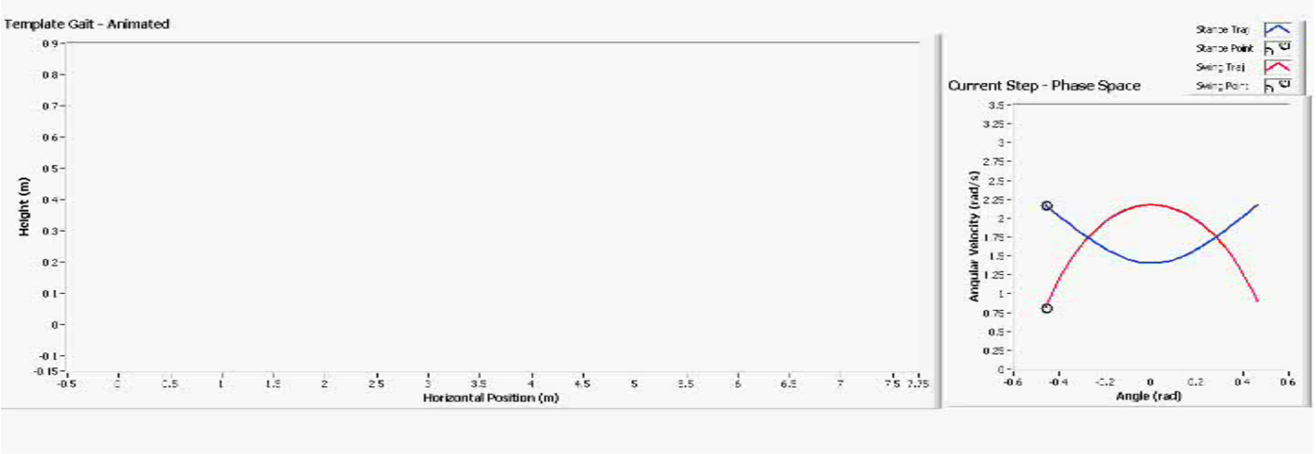


Single Support

Natural Parameterization of Gait



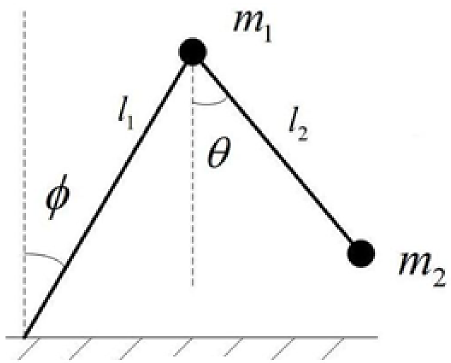
‘Ballistic’ Plan On Irregular Terrain



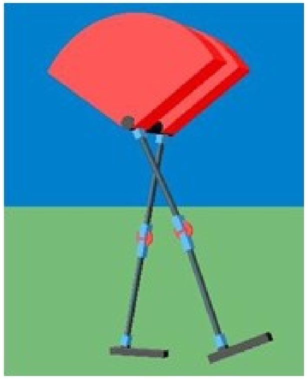
- Given terrain conditions, can reason about feasible trajectories
  - Can efficiently (dis)prove task achievement prior to execution
  - Velocity/force constraints are represented very naturally

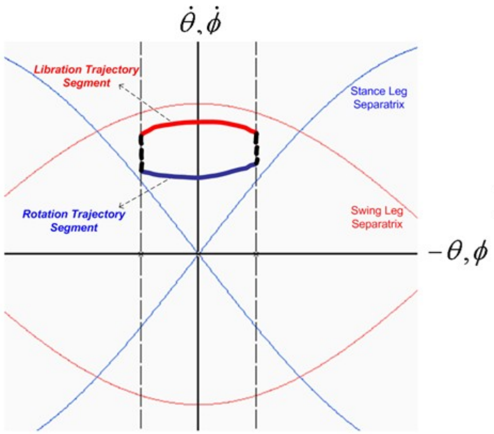
[Ramamoorthy+Kuipers, RSS ‘06]

Learning Dynamics of Multi-link Legged Robot

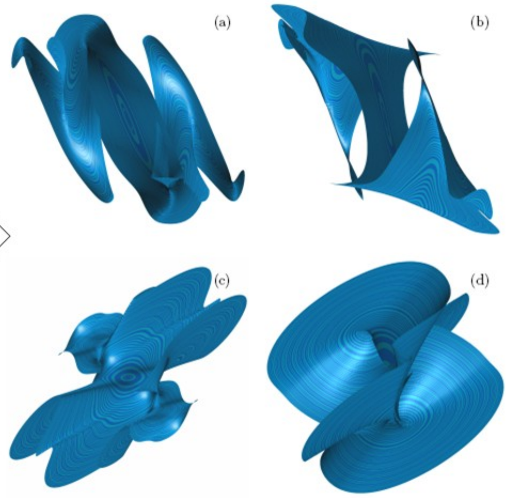


?

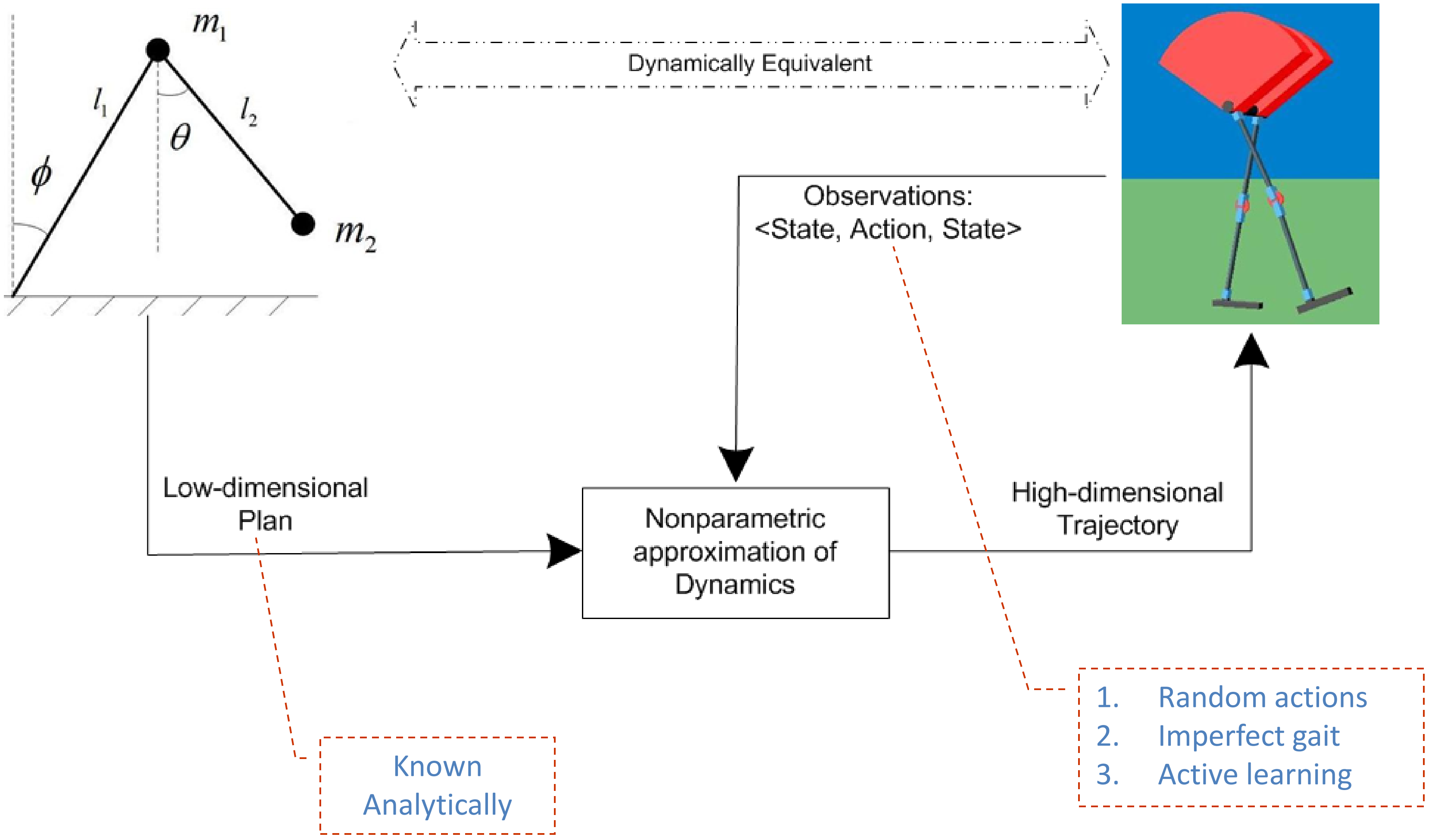




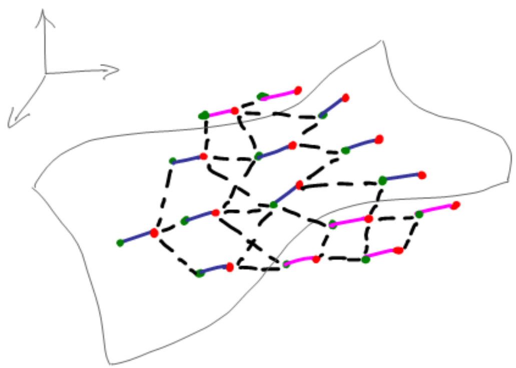
?



Trajectory Generation: Multi-link Legged Robot



Approximating Unknown Manifold from Data



$$
\begin{array}{ccc}
\mathcal{S}_H & \xrightarrow{\mathcal{M}_H} & \mathcal{S}_H \\
\downarrow \pi & & \downarrow \pi \\
\mathcal{S}_L & \xrightarrow{\mathcal{M}_L} & \mathcal{S}_L
\end{array}
$$

$$e_{comm} = f(\pi(x_j) - y_j),$$

for sequences $\{x_j\}, \{y_j\}, x_j \in \mathcal{S}_H, y_j \in \mathcal{S}_L$

Find sequence $\{x_j\}$, given template plan sequence $\{y_j\}$:

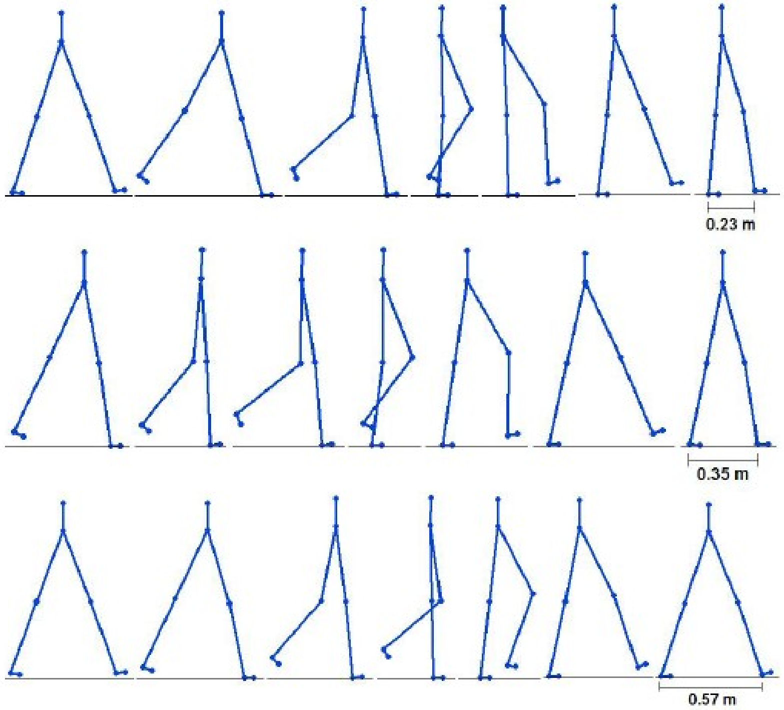
$$\{x_j\} = \text{argmin} \sum_{j=1}^{N} f(\pi(x_j) - y_j), x_j \in \mathcal{S}_H, y_j \in \mathcal{S}_L$$

Organize data in a k-NN graph

Where is *manifold* in the *graph*?

- Manifold ⇔ *Set of geodesic trajectories* restricted to it
- If the manifold encodes task – every geodesic must behave like template plan
- Diagram must commute!
  - Minimize commutativity error

Result: Controlled *Dynamic* Walking



	Manifold Approx.	Passive	Passive Period-2	Fast
$\mathcal{T}_{\text{int}_{rh}}$	16.7627	11.5357	13.3543	16.1479
$\mathcal{T}_{\text{int}_{rk}}$	7.3352	2.6817	6.19715	9.2198
$\mathcal{T}_{\text{int}_{ra}}$	18.1009	13.9009	15.0199	9.9699
$\mathcal{T}_{\text{int}_{lh}}$	14.9915	11.0399	17.0174	15.7704
$\mathcal{T}_{\text{int}_{lk}}$	6.00374	2.83644	5.9533	9.42043
$\mathcal{T}_{\text{int}_{la}}$	10.6606	13.6244	16.4178	15.8559
$\mathcal{T}_{intTotal}$	73.85464	55.61904	73.9598	76.3843

	Manifold Approx.	Passive	Passive Period-2	Fast
$cot = \left(\frac{\mathcal{E}}{mx}\right)$	1.63588	1.3672	1.68228	1.65216

[Ramamoorthy+Kuipers, ICRA ‘08]

Can We Proceed Without the *Low-dim* Model?

- Consider high-dim data drawn from an unknown low-dim manifold

$$x = \mathcal{M}(y)$$

- We can approximate the tangent space:

$$\mathcal{H}:\ x \in \mathbb{R}^D \mapsto \left[\frac{\partial}{\partial y_1}\mathcal{M}(y) \cdots \frac{\partial}{\partial y_d}\mathcal{M}(y)\right] \in \mathbb{R}^{D \times d}$$

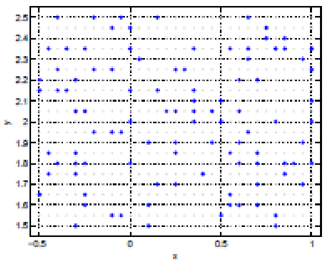
- This can be learned with a pair of optimization steps
- Simple example: 3-link arm
- The following error term defines the manifold:

$$err_{\mathcal{M}}(\mathbf{q}) = \min_{\{\epsilon^{ij}\}} \sum_{i,j \in N^i} \left\| \mathcal{H}_\theta(\bar{q}^{ij})\epsilon^{ij} - (q^i - q^j) \right\|_2^2$$

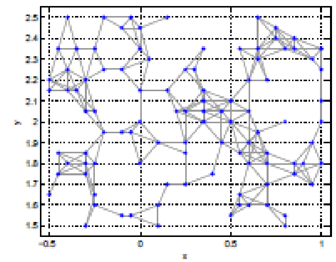
- Another error minimization defines geodesic paths:

$$err_{length}(\mathbf{q}) = \sum_{i=2}^{n} \left\| q^i - q^{i-1} \right\|_2^2$$

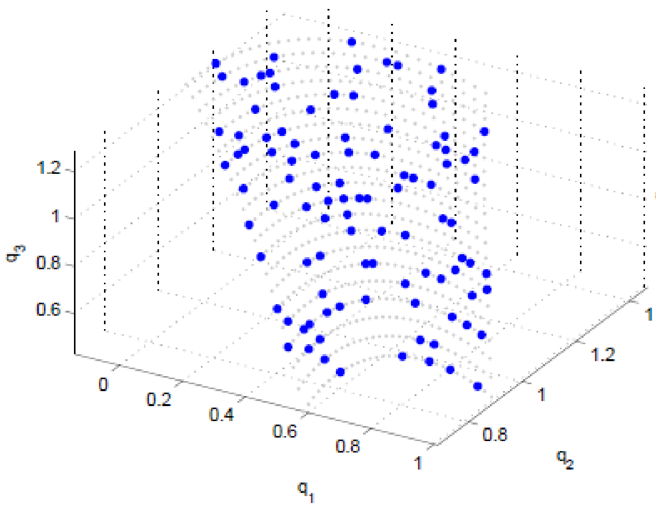
Learning Skill Manifold for 3-link Arm



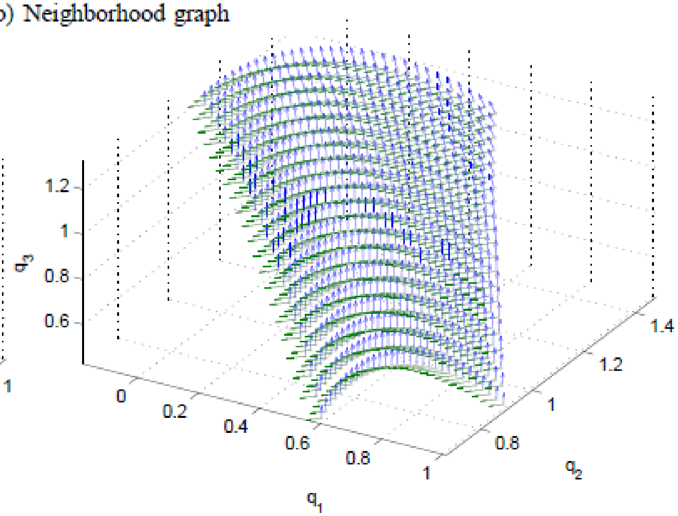
(a) Task space



(b) Neighborhood graph

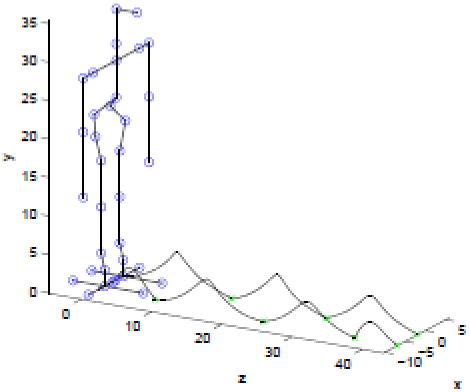


(c) Joint space

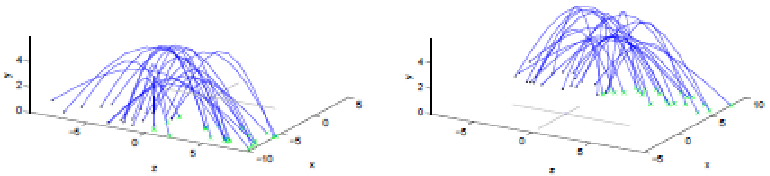


(d) Tangent space

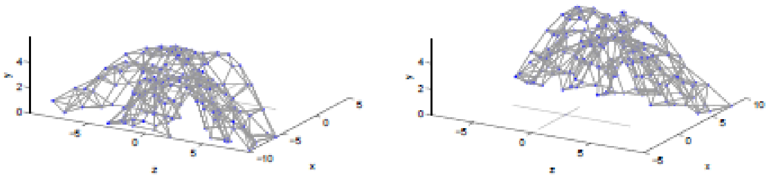
Application to Variable Step-length Walk



Approximated using:



(a) Right steps (b) Left steps



(c) Right neighborhood graph (d) Left neighborhood graph

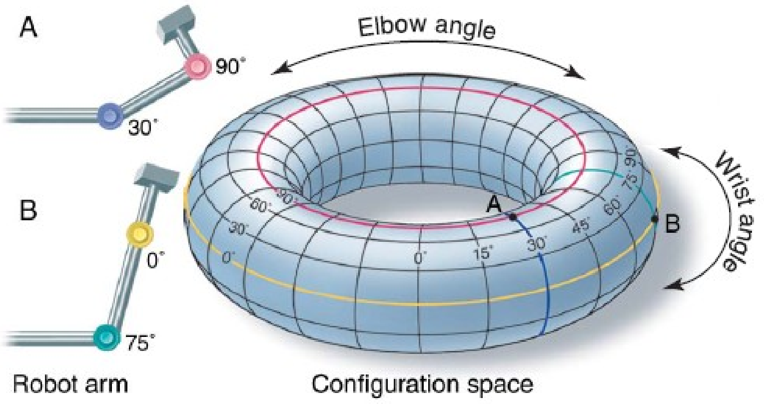
[Havoutis+Ramamoorthy, ICRA '10]

Agenda

1. Introductory Remarks
2. Geometric Representations of Humanoid Behaviours
3. Skill Learning in a Non-stationary World
4. Future Directions

Benefit of Geometric Representation

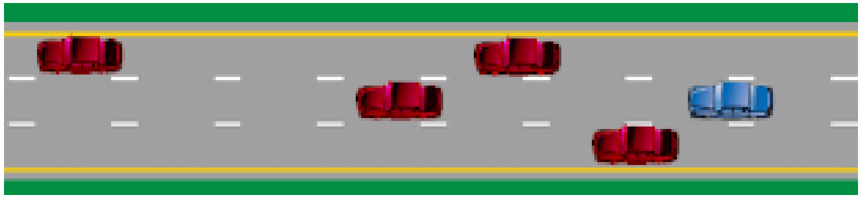
- Majority of dynamical behaviors are trajectories defined over abstract spaces (e.g., manifolds)
  - Traditional robotics problems:
    - Kinematics - manifold
    - Dynamics - metric on manifold
  - "Higher-level" considerations:
    - Shape
    - Topological restrictions



- **Skill learning in non-stationary world may be 'reduced' to adversarial navigation**

A Simple Formulation of Adversarial Navigation Problem

- Let us make the abstract spaces concrete
  - you are driving over a network of highways

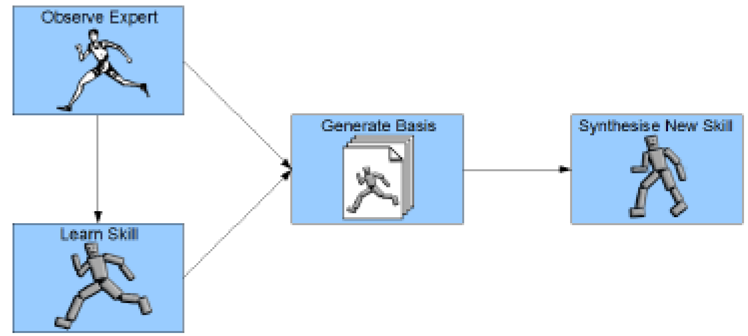


- Two sources of uncertainty:
  - Oncoming traffic (changing goals)
  - Changing navigability/costs

Imitation Learning of Basis Strategies

- We assume expert is solving an MDP with reward function $R^*(s) = w^* \cdot \phi(s)$

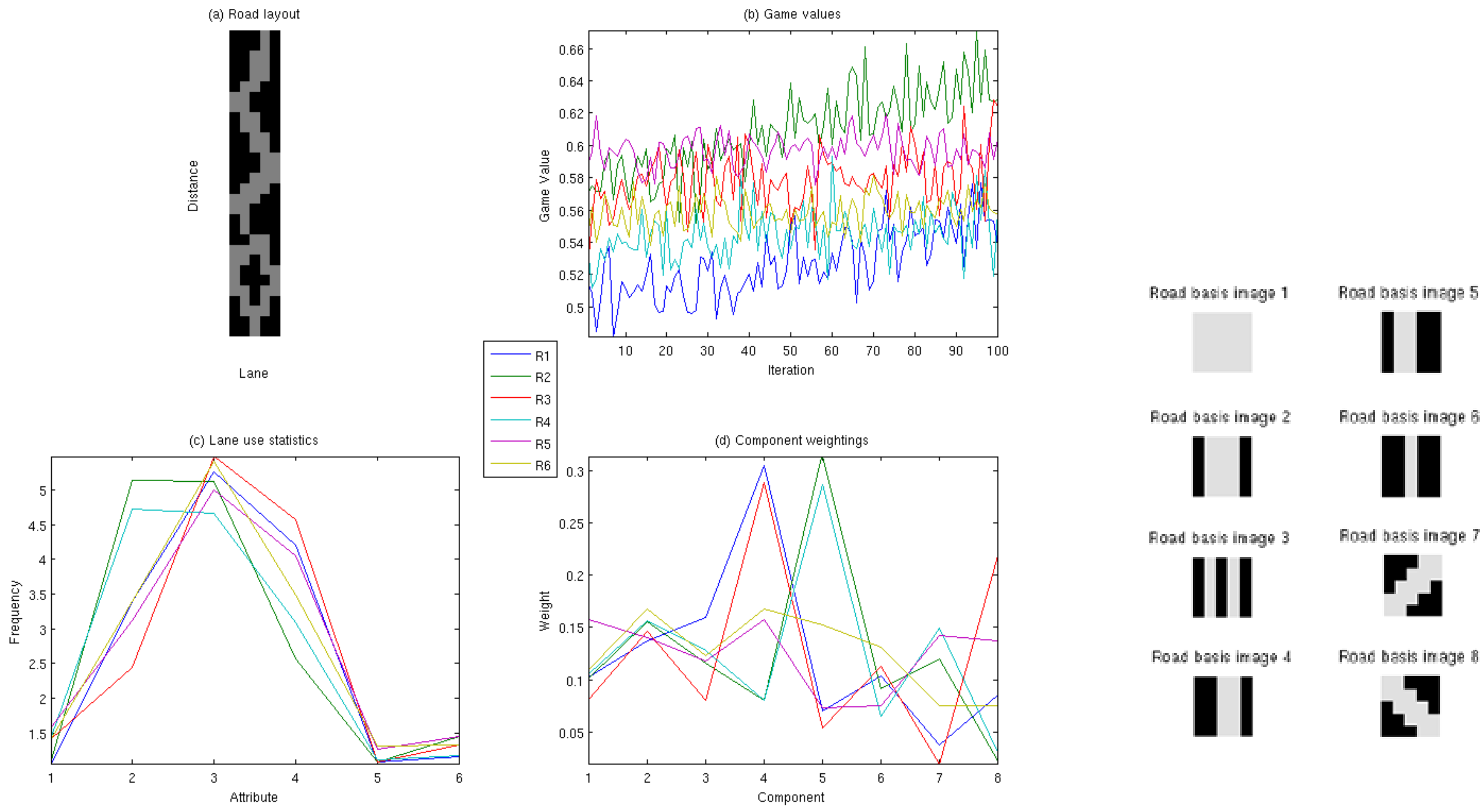
  where w^* is an unknown weight over features
- Goal is a mixed strategy maximizing $V(\psi) - V(\pi_E)$
- This can be computed as the Nash equilibrium of a game $v^* = \max_{\psi \in \Psi} \min_{w} w^T \mathbf{G} \psi$



Multiplicative Weight Updates:
[Syed+Schapire NIPS ‘08]

$$\mathbf{G}(\Psi, \tilde{\psi}, \mathbf{Q}) \longleftarrow$$
$$((E[R(\tilde{\psi}, \mathbf{Q})] - E[R(\Psi, \mathbf{Q})]) + 1)/2$$
$$\mathbf{W}(i) \longleftarrow \mathbf{w}^{(t)}(i) \beta^{\mathbf{G}(\psi_i, \tilde{\psi}, \mathbf{Q})} \text{ for all } i$$
$$\mathbf{w}^{(t+1)}(i) \longleftarrow \frac{\mathbf{W}_{(i)}}{\sum_i \mathbf{W}_{(j)}} \text{ for all } i = 1, \ldots, N$$

Learning to Drive in Novel Scenarios



[Rosman+Ramamoorthy, ICRA '10]

Agenda

1. Introductory Remarks
2. Geometric Representations of Humanoid Behaviours
3. Skill Learning in a Non-stationary World
4. Future Directions

Major Directions:
Learning from Experience + Demonstration

- RoboCup, etc.
  - Robust locomotion behaviours (walk, run, kick)
  - Composing skill manifolds
  - Adversarial navigation over these learnt manifolds (combine the two threads)
  - Multi-level (formation & player) skill learning and information aggregation in decentralized setting
- Topology-based motion synthesis (e.g., flexible object manipulation)
  - Qualitative task encoding
  - Shaped reinforcement learning

Acknowledgements

- Collaborators:
  - Sethu Vijayakumar, Taku Komura [Robot Learning, Motion Synthesis]
  - Rahul Savani [Computational Finance]

- Students:
  - Ioannis Havoutis, Aris Valtazanos, Benjamin Rosman, Majd Hawasly [Humanoid Robotics, RoboCup, Skill Learning]
  - Thomas Larkworthy [Reconfigurable Robotics]