Advanced Behavior-based Control of Bipedal Locomotion

Karsten Berns, Qi Liu
Robotics Research Lab
Department of Computer Science
University of Kaiserslautern, Germany
Research at Robotics Research Lab (RRLab)

AUTONOMOUS MOBILE ROBOTS

HUMANOID ROBOTS

OFF-ROAD ROBOTS

SERVICE ROBOTS

METHODOLOGY AND ALGORITHMS
Human Robot Interaction
Behavior-based Control (iB2C)

- Fundamental unit in iB2C: Control/Perception module
  \[ B = (f_r, F) \]
- \( f_r \): Target rating function
- \( F \): Transfer function determines output vector

- \( \vec{e} = (d, \sigma) \in \mathbb{R}^{2m} \): Input vector
- \( \vec{u} = (d, \sigma) \in \mathbb{R}^{2n} \): Output vector
- \( d \in \mathbb{R} \): Data value
- \( \sigma \in \mathbb{R}^+ \): Standard deviation of \( d \)
  - Percept: Uncertainty
  - Behavior: Error margin

- \( s \in [0,1] \): Stimulation
- \( i \in [0,1] \): Inhibition
- \( a \in [0,1] \): Activity
- \( r \in [0,1] \): Target rating
Fusion Behavior Module

- Common interface
- Coordinate $p$ competing behaviors $B_c$
- $F$ is the fusion function processing input values to a merged output control vector $\vec{u}$

- Maximum fusion: $\vec{u} = \vec{u}_s$, $a = \max_c(a_c)$, $r = r_s$ where $s = \arg\max_c(a_c)$

- Weighted average fusion: if $(1 = \max_c(r_c))$ then maximum fusion else

\[
\vec{u} = \frac{\sum_{j=0}^{p-1} a_j \cdot \vec{u}_j}{\sum_{k=0}^{p-1} a_k} \quad a = \frac{\sum_{j=0}^{p-1} a_j^2}{\sum_{k=0}^{p-1} a_k} \quad r = \frac{\sum_{j=0}^{n-1} a_j \cdot r_j}{\sum_{k=0}^{n-1} a_k}
\]
Behavior Networks
Background Biological inspired Robots (FZI)
Challenges in Bipedal Locomotion Control

- Low energy consumption
- Control of stiffness, complaint
- Locomotion capability in rough environments
- Sensitive to any external disturbances

=> What principles can be transferred from nature
Elastic operation of the leg can passively stabilize running in the presence of external disturbances without changing the angle of attack or the stiffness [Blickhan 07].

Functional morphology
- Mass distribution
- Geometry of trunk and extremities
- Low resistant elastic actuator

Exploitation of inherent dynamics
Mechanical Model of the Bipedal Robot

- Anthropomorphic biped
  - mass distribution
  - segment length
- Weight 76kg
- Height 1.8m
- DOF 21
  - 6 DoF per leg
  - 3 DoF spine
  - 3 DoF per arm
- Physics simulation environment - Newton
# Mechanical Parameters of the Simulated Biped

<table>
<thead>
<tr>
<th>Joint</th>
<th>min. angle [rad]</th>
<th>max. angle [rad]</th>
<th>max. torque [Nm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spine X</td>
<td>-1.0</td>
<td>1.0</td>
<td>180</td>
</tr>
<tr>
<td>Spine Y</td>
<td>-1.0</td>
<td>1.0</td>
<td>180</td>
</tr>
<tr>
<td>Spine Z</td>
<td>-1.0</td>
<td>1.0</td>
<td>100</td>
</tr>
<tr>
<td>Shoulder X left</td>
<td>0.0</td>
<td>2.0</td>
<td>80</td>
</tr>
<tr>
<td>Shoulder X right</td>
<td>-2.0</td>
<td>0.0</td>
<td>80</td>
</tr>
<tr>
<td>Shoulder Y</td>
<td>-1.0</td>
<td>1.0</td>
<td>80</td>
</tr>
<tr>
<td>Elbow Y</td>
<td>-2.0</td>
<td>0.0</td>
<td>30</td>
</tr>
<tr>
<td>Hip X</td>
<td>-1.0</td>
<td>1.0</td>
<td>220</td>
</tr>
<tr>
<td>Hip Y</td>
<td>-0.85</td>
<td>0.15</td>
<td>220</td>
</tr>
<tr>
<td>Hip Z</td>
<td>-1.0</td>
<td>1.0</td>
<td>150</td>
</tr>
<tr>
<td>Knee Y</td>
<td>0.0</td>
<td>2.0</td>
<td>150</td>
</tr>
<tr>
<td>Ankle X</td>
<td>-0.5</td>
<td>0.5</td>
<td>80</td>
</tr>
<tr>
<td>Ankle Y</td>
<td>-0.5</td>
<td>0.5</td>
<td>150</td>
</tr>
</tbody>
</table>
Hierarchical Layout of Motion Control

• Hierarchical control concept
• Movement based on synergies and
• Reflexes (modulated by intensity of stimulus)
Phases of Walking

(1) weight acceptance
(2) loading or propulsion
(3) trunk stabilization during double support
(4) toe lift-off
(5) heel strike

- Fixed sequence of walking phases
- Weighting and scaling dependent on walking speed
Reflex Function during Walking

Function of stretch reflex, load receptor reflexes and cutaneous reflexes during locomotion [Zehr 99].

Stabilisation of Walking based on reflexes in different phases
Behavior-Based Bio-inspired Bipedal Locomotion
Human-like Walking Phases

- Five walking phases (see Lacquaniti et al.)
- Transitions triggered by sensor events
- Bilateral synchronization → robust state switches
Example: Initiation of Walking

- How to take weight from swing leg?
- Analyse EMG from walking initiation
- Introduce new motor pattern and trigger it during initiation phase
- Let passive dynamics do the rest
- Ground reaction forces and angles comparable to human data

[S.G.B. Kirker et al., Stepping before standing, J. Neuropsychol. 2000]
Design of Motor Patterns

- Motor patterns shape passive dynamics (feed-forward control)
- Analysis of muscle activities, kinematical, and kinetic data: derive motor patterns (and other control units)
- Same parametrized equation for all motor patterns:

\[
\tau = \begin{cases} 
\frac{1}{2} + \frac{1}{2} \sin \left( \pi \left( \frac{t}{T_1} - \frac{1}{2} \right) \right) & 0 \leq t < T_1 \\
1 & T_1 \leq t < T_2 \\
\frac{1}{2} - \frac{1}{2} \sin \left( \pi \left( \frac{t-T_2}{T_3-T_2} - \frac{1}{2} \right) \right) & T_2 \leq t \leq T_3
\end{cases}
\]
Feedback Local Reflexes

- Tight coupling between sensor information and motor action
- Linear/nonlinear relation between sensor data and control output
- Event-based control
- Example: Lock Hip
  - Active when hip angle approaches target position, generate braking torque
    \[ \hat{\tau}_{lp} = -\omega_{leg} \cdot K_{torque} \]
  - Until rotation velocity is 0, hold leg in target position \( \hat{a}_{lp} \) for preparation of heel strike
Implementation of lowest Level of Control
Control Block Diagram

\[ \tau_{\text{control}} = \frac{s_\tau (\tau_{\text{target}} s_\tau) + s_\alpha (\tau_{\text{pos}} s_\alpha)}{s_\tau + s_\alpha} \]

\[ \tau_{\text{spring}} = \text{sgn}(\alpha_0 - \alpha) \cdot K_{\text{spring}} \cdot (\alpha_0 - \alpha)^2. \]
Postural Control is Supraspinal

- [Hof 08]: postural control based on extended center of mass ($X_{\text{coM}}$)
  \[
  X_{\text{com}} = d + \frac{\dot{d}}{\omega_0}
  \]

- For each leg and each direction (frontal and sagittal plane)
- Postural reflexes: Upright Trunk, Forward Velocity, Lateral Balance Ankle, Lateral Foot Placement
- Approximation of $X_{\text{coM}}$ trajectory
- Derivation results in reflex action
Example: Network during Walking Phases 4/1

Walking Initiation → Walking SPG → Walking → Cyclic Walking → Stable Standing

Walking SPG:
1. Weight Accept.
2. Propulsion
3. Trunk Stabiliz.
4. Leg Swing
5. Heel Strike


Stabilize Pelvis R → Lock Hip L → Initialize Swing L → Arm Swing R

Optimization → Stabilization → Relax Joints
First Results
Problems of B4LC System

- Not optimized in current locomotion skills
  - Energy consumption
  - Velocity control
  - Stability

- Not adaptive to more challenging disturbances
  - Rough terrains
  - Large obstacles
  - Stairs
  - Slopes
Optimization of Motor Patterns

\[ T_1 \] the starting time of maximum torque
\[ T_2 \] the ending time of maximum torque
\[ T_3 \] the total time of torque command

\[
\hat{\tau} = A \times \begin{cases} 
\frac{1}{2} + \frac{1}{2} \sin \left( \pi \left( \frac{t}{T_1} + \frac{1}{2} \right) \right) & 0 \leq t \leq T_2 \\
1 & T_1 \leq t \leq T_2 \\
\frac{1}{2} - \frac{1}{2} \sin \left( \pi \left( \frac{t-T_2}{T_3-T_2} - \frac{1}{2} \right) \right) & T_2 \leq t \leq T_3
\end{cases}
\]
Optimization Methodology

- Optimization module using Particle Swarm Optimization
- Parameter values are calculated until the desired fitness functions are obtained
Example: Optimization on Even Ground

- Optimization for **energy consumption, locomotion stability, and walking speed**
Example: Optimization on Even Ground

- Fitness functions:

\[ f_1 = s_w \]
\[ f_2 = \sum_{i=0}^{T} \frac{\Delta X_{com_i}}{s_w} \]
\[ f_3 = \sum_{i=0}^{T} \frac{|V_{ref} - V_i|^2}{T} \]
\[ f_4 = \sum_{i=0}^{T} \frac{\tau_i^h}{T} \]

- Robustness
- Stability
- walking speed control
- energy consumption
Example: Optimization on Even Ground

Optimization of the motor patterns LP and AHS

- Iterations 0 - 25
- Iterations 25 - 50
- Iteration 50 - 175
- Iterations 175 - 215

Speed x 4
Example: Optimization on Even Ground

![Graph showing the comparison of Lock Hip torque before and after optimization.]

<table>
<thead>
<tr>
<th></th>
<th>Accumulated joint torques during one cycle</th>
<th>Velocity deviation during 10 seconds (%)</th>
<th>Accumulated $\Delta X_{com}$ during one cycle (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ankle (Nm)</td>
<td>hip (Nm)</td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>3842.2</td>
<td>746.3</td>
<td>8.3</td>
</tr>
<tr>
<td>After</td>
<td>3142.1</td>
<td>657.9</td>
<td>0.04</td>
</tr>
<tr>
<td>Improvement</td>
<td>18.22%</td>
<td>11.83%</td>
<td>8.26</td>
</tr>
<tr>
<td></td>
<td>2.721</td>
<td>1.232</td>
<td>1.489</td>
</tr>
</tbody>
</table>
RL for Reflexes - Locomotion on Uneven Ground

- Reflexes (R:Control Forward Velocity) produce compensating torques at ankle joints to reject disturbances.
Example: Locomotion on Uneven Ground

Locomotion on even terrain with the optimized motor patterns

speed x 2
Example: Optimization on Uneven Ground

<table>
<thead>
<tr>
<th></th>
<th>Accumulated $\Delta X_{com}$ during one cycle (m)</th>
<th>Accumulated compensation torque (Nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>3.586</td>
<td>940.3</td>
</tr>
<tr>
<td>after</td>
<td>3.198</td>
<td>855.4</td>
</tr>
<tr>
<td>improvement</td>
<td>0.388</td>
<td>9.02%</td>
</tr>
</tbody>
</table>
B4LC based Advanced Walking Skills

- Curve walking
- Upslope walking
- Speed control
- Push recovery
- Push recovery standing
- Walking Termination
Various Speed Locomotion
Push Recovery Locomotion

Locomotion without push recovery and with push recovery
F = 300 N
Stepping over Obstacle Locomotion

- The control units at the hip and knee joints are refined
  - Hip joint movement, e.g. motor pattern Active leg swing, reflexes Lock hip
  - Knee joint movement, e.g. reflexes Knee flexion
Stepping over Obstacle Locomotion

- Using PSO to search the parameters of 8 parameters
- 20cm height and 15cm width
Experimental Validation
SEA – Theoretical Characteristics

- **Force**
  - Continuous: 800 N
  - Maximal: 2800 N
- **Max speed:** 400 mm/s
- **Max travel:** 145 mm
- **Weight:** ca 1,3 kg
- **Dimensions**
  - **Length:** 225 mm
  - **Diameter:** 84.5 mm
Open-loop actuation

Open-loop Sine 2Hz 5A
Robotic Leg with biarticular elements

- Leg with mono- and biarticular actuation
  - 3 mono-articular SEAs at each joint
  - 2 bi-articular SEAs spanning
    - Hip and knee
    - Knee and ankle
- Prosthetic foot
  - Simulates human foot arch
  - Introduces additional compliance
Thank you for your attention!
# Reflex Controllers for Cyclic Walking I

<table>
<thead>
<tr>
<th>Phase</th>
<th>Reflexes</th>
<th>Sensor Event</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 2</td>
<td><em>Lateral Balance Ankle</em></td>
<td><em>Always on during these phases</em></td>
<td>Generating torques to ankle x</td>
</tr>
<tr>
<td>1</td>
<td><em>Control Forward Velocity Ankle</em></td>
<td><em>Always on during this phase</em></td>
<td>Correcting torques to ankle y and knee</td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td><em>Stabilize Pelvis</em></td>
<td><em>Always on during these phases</em></td>
<td>Correcting angle to hip x</td>
</tr>
<tr>
<td>2</td>
<td><em>Keep Knee Angle</em></td>
<td><em>Always on during this phase</em></td>
<td>Keeping the knee neutral</td>
</tr>
<tr>
<td>4</td>
<td><em>Cutaneous reflex</em></td>
<td><em>Only on during the beginning of the swing phase (when small ground contact still existing)</em></td>
<td>Generating torques to ankle y and knee</td>
</tr>
<tr>
<td>4 &amp; 5</td>
<td><em>Lateral Foot Placement</em></td>
<td><em>Always on during these phases</em></td>
<td>Correcting the desired angle of hip x</td>
</tr>
</tbody>
</table>
## Reflex Controllers for Cyclic Walking II

<table>
<thead>
<tr>
<th>Phase</th>
<th>Reflexes</th>
<th>Sensor Event</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 &amp; 5</td>
<td><em>Lock Hip</em></td>
<td>Only <em>on</em> when hip x in activation zone</td>
<td>Keeping hip x stiff</td>
</tr>
<tr>
<td>4</td>
<td><em>Lock Knee</em></td>
<td>Only <em>on</em> when angle of knee in activation zone</td>
<td>Keeping knee stiff</td>
</tr>
<tr>
<td>4 &amp; 5</td>
<td><em>Keep Ankle Y Angle</em></td>
<td>Always <em>on</em> during these phases</td>
<td>Correcting angle of ankle y</td>
</tr>
<tr>
<td>5</td>
<td><em>Heel Strike</em></td>
<td><em>On</em> after heel contact measured</td>
<td>Generating torques to ankle y</td>
</tr>
<tr>
<td>1</td>
<td><em>Upright Trunk</em></td>
<td>Always <em>on</em> in all phases</td>
<td>Generating torques to ankle y and knee</td>
</tr>
<tr>
<td>4 &amp; 5</td>
<td><em>Lateral Foot Placement</em></td>
<td>Always <em>on</em> during these phases</td>
<td>Generating torques to hip x</td>
</tr>
</tbody>
</table>
# Motor Patterns for Cyclic Walking

<table>
<thead>
<tr>
<th>Phase</th>
<th>Motor Patterns</th>
<th>Sensor Event</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 &amp; 3</td>
<td><em>Leg Propel</em></td>
<td>Triggered since phase 2</td>
<td>Generating torques to ankle y and knee</td>
</tr>
<tr>
<td>3</td>
<td><em>Hip Swing</em></td>
<td>Stimulated from phase 3 onwards</td>
<td>Generating torques for hip x,y,z</td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td><em>Arm Swing</em></td>
<td>Always <em>on</em> during these phases</td>
<td>Generating torques for ipsilateral and contralateral arms</td>
</tr>
<tr>
<td>5</td>
<td><em>Heel Strike</em></td>
<td><em>On</em> after heel contact measured</td>
<td>Generating torques to ankle y</td>
</tr>
<tr>
<td>5 &amp; 1</td>
<td><em>Weight Acceptance</em></td>
<td>Activated during these phases</td>
<td>Generating torques to ankle y and knee</td>
</tr>
</tbody>
</table>
# Control Units for Advanced Locomotion I

<table>
<thead>
<tr>
<th>Phase</th>
<th>Motor Pattern / Reflex</th>
<th>Locomotion Skills</th>
<th>Actions</th>
<th>New unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 &amp; 3</td>
<td><em>Leg Propel</em></td>
<td>Various speed walking</td>
<td>Adapt ankle torques</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td><em>Hip Swing</em></td>
<td>Various speed walking</td>
<td>Adapt hip torques for different step length</td>
<td>No</td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td><em>Arm Swing</em></td>
<td>Various speed walking</td>
<td>Adapt arm swing frequency for various speed</td>
<td>No</td>
</tr>
<tr>
<td>4 &amp; 5</td>
<td><em>Lock Hip</em></td>
<td>Various speed walking</td>
<td>Adapt step length for various speed</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td><em>Lock Knee</em></td>
<td>Various speed walking</td>
<td>Enable ground clearance</td>
<td>No</td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td><em>Lateral Balance Ankle</em></td>
<td>Various speed walking</td>
<td>Adapt lateral ankle torques</td>
<td>No</td>
</tr>
</tbody>
</table>
## Control Units for Advanced Locomotion II

<table>
<thead>
<tr>
<th>Phase</th>
<th>Motor Pattern / Reflex</th>
<th>Locomotion Skills</th>
<th>Actions</th>
<th>New unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 &amp; 3</td>
<td><em>Leg Propel</em></td>
<td>Push recovery</td>
<td>Adapt ankle torques for instability</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td><em>Hip Swing</em></td>
<td>Push recovery</td>
<td>Adapt hip torques for larger step length</td>
<td>No</td>
</tr>
<tr>
<td>4 &amp; 5</td>
<td><em>Lock Hip</em></td>
<td>Push recovery</td>
<td>Adapt hip angle for smooth heel strike</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td><em>Hip extension</em></td>
<td>Push recovery</td>
<td>Enable hip extension for stance leg when push in early swing phase</td>
<td>Yes</td>
</tr>
<tr>
<td>1</td>
<td><em>Knee Flexion</em></td>
<td>Push recovery</td>
<td>Enable knee flexion for stance leg when push in early swing phase</td>
<td>Yes</td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td><em>Hold Knee</em></td>
<td>Push recovery</td>
<td>Increase knee stiffness in recovery phase</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Control Units for Advanced Locomotion III

<table>
<thead>
<tr>
<th>Phase</th>
<th>Motor Pattern / Reflex</th>
<th>Locomotion Skills</th>
<th>Actions</th>
<th>New unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Hip Swing</td>
<td>Stepping over obstacle</td>
<td>Adapt hip torques for larger swing angle</td>
<td>No</td>
</tr>
<tr>
<td>4 &amp; 5</td>
<td>Lock Hip</td>
<td>Stepping over obstacle</td>
<td>Adapt hip angle for larger swing angle</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>Knee Flexion</td>
<td>Stepping over obstacle</td>
<td>Actively control knee joint until knee is stretched</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Example: Network during Obstacle Walking
Phases 4/1

- The control units at the hip and knee joints are refined
  - Hip joint: Leg swing & Lock hip
  - Knee joint: Knee flexion
Joint Control

\[ \tau_{\text{control}} = \frac{s_\tau (\tau_{\text{target}} s_\tau) + s_\alpha (\tau_{\text{pos}} s_\alpha)}{s_\tau + s_\alpha} \]

\[ \tau_{\text{spring}} = \text{sgn}(\alpha_0 - \alpha) \cdot K_{\text{spring}} \cdot (\alpha_0 - \alpha)^2. \]
Key Design Factors

- **Compliance**
  - Impact tolerance
    - Deal with uncertainties
  - Energy storage
    - Reduced metabolic cost
    - Increased mechanical output

- **Modulation of stiffness/damping**

- **Passivity and torque(force control)**
  - Ballistic movements

- **Hierarchical feedback mechanisms**
  - Physically distributed
RRLAB SEA Implementation

- Drivetrain
  - Robodrive 70x18
  - Ball screw pitch of 8mm
    - Torque to force efficiency of 97%
    - Force to torque efficiency of 97%
    → Increased energy recuperation
  - Ball nut integrated in rotor shaft
- Spring system
  - Proximal placement [1]
  - Standard die springs
  - Linear guides
  - Rotational spring bushings
Bi-articular muscles

- Muscles acting across two joints
- Redirect muscle action
  - e.g. vertical jumping
- Energy transfer between proximal and distal joints
  - Implication can be found in torque profiles obtained from Luksch[10]

Reflexes and Motor Patterns

Reflex action provides the basic feedback mechanism in human motion control. Local reflexes show a tight coupling between sensor information and motor action. Reflexes show a feedback controller-like behavior with a linear or nonlinear relation between sensor data and control output. The other type of reflexes works event-based: as soon as a certain sensor event occurs, the output state is changed.

The reflex action at one place of the robot can be the result of a sensor event at the opposite end of the machine, possibly supported by a simplified dynamical model. The state variables most frequently used or calculated by postural reflexes are estimations of the upper body orientation, the rough position or velocity of the center of mass, or load distribution in the feet. In collaboration with the local reflexes, the postural reflexes enhance the global stability of the biped.

The stimulation of certain regions in the spinal cord result in muscle action producing coordinated joint or limb motions. These components or motion primitives seem also to be recruited in phases of locomotion. The corresponding control unit of this concept is called motor pattern. They produce uniform patterns of torques for one or more joints in a feed-forward manner. They always work locally.
Optimization Methodology

- It searches a space by adjusting the trajectories of individual vectors, called ‘particles’, as they are conceptualized as moving as points in multidimensional space.
- The velocity and position of particle are updated as:
  \[ v_{i}^{k+1} = \omega \cdot v_{i}^{k} + c_1 \cdot Rand \cdot (Pbest_{i}^{k} - x_{i}^{k}) \]
  \[ + c_2 \cdot Rand \cdot (Gbest^{k} - x_{i}^{k}) \]

  \( \omega \): inertia weight  
  \( c_1 \) and \( c_2 \): acceleration constant  
  \( rand \): random number between 0 and 1  
  \( Gbest \): best position of group in iteration \( k \)  
  \( Pbest_{i}^{k} \): best position of the particle \( i \) in iteration \( k \)  
  \( x_{i}^{k+1} = x_{i}^{k} + v_{i}^{k+1} \)

\( x_{i}^{k} \): position of particle \( i \) in iteration \( k \)  
\( v_{i}^{k} \): velocity of particle \( i \) in iteration \( k \)
Expectation-Maximization based Reinforcement Learning I

- Action $a$ is calculated by combining the weight parameters $\theta$ of the radial basis function network and the basis function $\Phi(s)$
  
  $$a = \theta^T \Phi(s)$$

  where $\Phi_j(s) = \exp\left(-\frac{(s-\mu_j)^2}{2\sigma^2}\right)$ for $j = 1, 2, \ldots, N$

- Undiscounted accumulated reward in the iteration $n$
  
  $$R(n) = \sum_{0}^{T} r(t)$$

- The parameter $\theta_n$ in the iteration $n$ is updated to $\theta_{n+1}$ in the iteration $n + 1$ with previous $K$ best iterations
  
  $$\theta_{n+1} = \theta_n + \frac{\sum_{k=1}^{K}(\theta_k - \theta_n)R(k)}{\sum_{k=1}^{K}R(k)}$$

- Stop learning until $\theta_{n+1} = \theta_n$
Expectation-Maximization based Reinforcement Learning II

- Learning module is between Motion Phase and Reflexes.
- The sensory information is feedback as state inputs to RL module.
- RL module calculates action values based on radial basis function network.
- The action is exerted at the reflexes.
- Rewards to update network until expected returns obtained.
Behavior-Based Bio-inspired Bipedal Locomotion

B4LC system

- Even ground locomotion
- Curve locomotion
- Upslope locomotion
- Speed control
- Push recovery
- Balancing standing
State-of-the-art in bipedal robots

- Lola (TUM), ASIMO (HONDA), HRP-4C (JAIST), Atlas (Google).
Postural Control

Schematic diagram of the inverted pendulum model [Hof08].

**Implications for Robotics:**

Postural control includes maintaining body stability in the sagittal and front plane and controlling forward velocity. This is a high-level skill requiring an estimation of the robot’s pose using information on joint angles, acceleration, and velocity information from an inertial measurement unit, and optical flow from vision, if available. **Adjusting the foot placement** seems to have the major influence on whole body balancing. Anticipatory torque patterns seem to be necessary to compensate segment movements resulting from mass inertia during normal walking, especially in the hip and trunk joints.
Various Speed Locomotion

$T_1$ $T_2$ $T_3$ are searched with respect to different velocities
Various Speed Locomotion

- Searching the 18 parameters using PSO
- Using linear least square to define the functions that parameters with respect to walking velocities
- Good performance from 0.625 m/s to 1.625 m/s
Push Recovery Locomotion

Using RL module to control the hip, ankle and knee joints movement during pushes

Reward functions considering stability in sagittal and frontal plane

Reward functions:

$$r(t) = e^{(-k_s|\Delta X_{com,t}| - k_t|\Delta X_{com,t}|)}$$

$$R_s(e) = k_{su}e^{-(\text{step}_{\text{max}} - \text{step}_{\text{ach}}(e))/\text{step}_{\text{max}}}$$

$$R(e) = \sum_{t=1}^{t_e} r(t) + R_s(e)$$