An Architecture for Action Selection in Robotic Soccer

Peter Stone

Joint work with
David McAllester

RoboCup

An international AI and Robotics research initiative

- Use soccer as a rich and realistic test-bed

Research challenges

- Multiple teammates with a common goal
- Multiple adversaries — not known in advance
- Real-time decision making necessary
- Noisy sensors and actuators
- Enormous state-space

CMUnited-99

- Stone, Riley, Veloso
- 1999 simulator league world champions
- 37-team field; Total score: 110–0 (8 games)

- Learned low-level behaviors
- Heuristic high-level action decision
  - Dribble; Shoot; Hold; Clear; Pass (10)

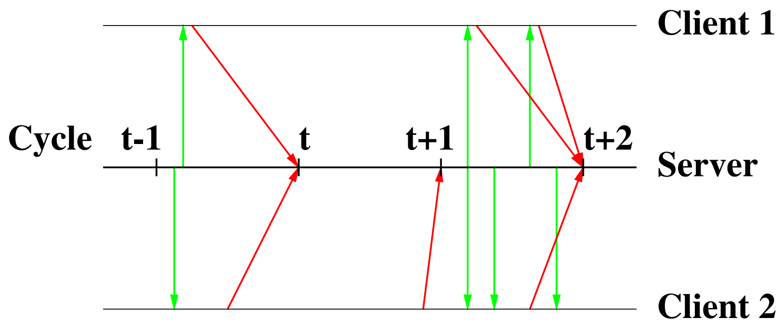
Here: Improvements over CMUnited-99

Outline

- **RoboCup simulator**
- Action Selection Architecture
- Leading Passes
- Force Field Control for Off-Ball Motion
- Results

RoboCup Simulator

- Distributed: each player a separate client
- Server models dynamics and kinematics
- Clients receive sensations, send actions



- Parametric actions: dash, turn, kick, say
- Abstract, noisy sensors, hidden state
  - Hear sounds from limited distance
  - See relative distance, angle to objects ahead
- $> 23^{10^9}$ states
- Limited resources: stamina
- Play occurs in real time ($\approx$ human parameters)

Outline

- RoboCup simulator
- **Action Selection Architecture**
- Leading Passes
- Force Field Control for Off-Ball Motion
- Results

Motivation

$\boxed{\text{Decisions based on a } \textbf{Value Function}}$

- $v(s) \equiv$ expected reward from state s (RL)
- $P(s'|s,a) \equiv$ probability of outcome s' when selecting option (action) a from s
- Select option with highest

$$\sum_{s'} P(s'|s,a)v(s')$$

Options

$\boxed{\text{An option can be scored and executed}}$

- Execute the option with the highest score
- Scoring:
  - $p_s \equiv$ probability of success
  - $v_s, v_f \equiv$ values of succeeding, failing
  - Score: $p_s v_s + (1 - p_s) v_f$
  - value function currently hand-written
  - Scoring across options must be comparable

Aside: Soft Boolean Expressions

$\boxed{\text{Avoid discontinuities}}$

- $x <^{\delta} y \in [0,\ 1]$ (continuous)

$$\begin{aligned} x = y &\Rightarrow x <^{\delta} y = 1/2 \\ x << 0 &\Rightarrow x <^{\delta} y \sim 0 \\ x >> 1 &\Rightarrow x <^{\delta} y \sim 1 \end{aligned}$$

- $\text{if}^*(p,\ x,\ y)$ assumes $p \in [0,\ 1]$

$$\text{if}^*(p, x, y) \equiv px + (1-p)y$$

  – Often write $\text{if}^*(x <^{\delta} y,\ z,\ w)$.

Pass Option

- Consider hundreds of passes:
  - angle increments of 4^o
  - speed increments of $0.2m/sec$
- I_t (I_o) $\equiv$ teammate (opponent) interception time
  - Approximate, **fast** computation
- Score: larger margin $\Rightarrow$ larger p_s

$$p_s = \text{if}^*(I_t <^5 I_o, .9, 0)$$

- v_s based on ball's predicted location after pass
- $v_f = 0$

Other Options

Shot Option: kick towards a point in the goal

- p_s related only to I_o
- $v_s >> 0$
- $v_f = 0$

Clear Option: kick the ball down the field

- p_s related only to I_o
- $v_s > 0$
- $v_f = 0$

Others: dribble, send, hold, cross, ...

- Difficult to calibrate many

Leading Passes

CMUnited-99: only direct passes
Now: hundreds considered

– Usually a pass option is selected
– **Many leading passes seen**

Movement without the ball is also crucial

CMUnited-99: SPAR

– Forces over limited regions
– Boundaries treated as hard constraints

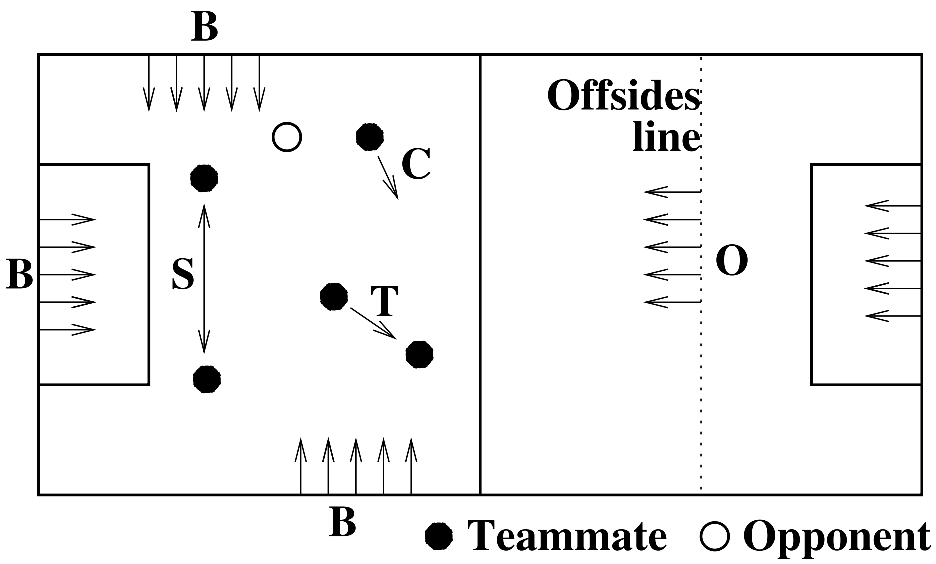
Outline

- RoboCup simulator
- Action Selection Architecture
- Leading Passes
- **Force Field Control for Off-Ball Motion**
- Results

Movement Off the Ball

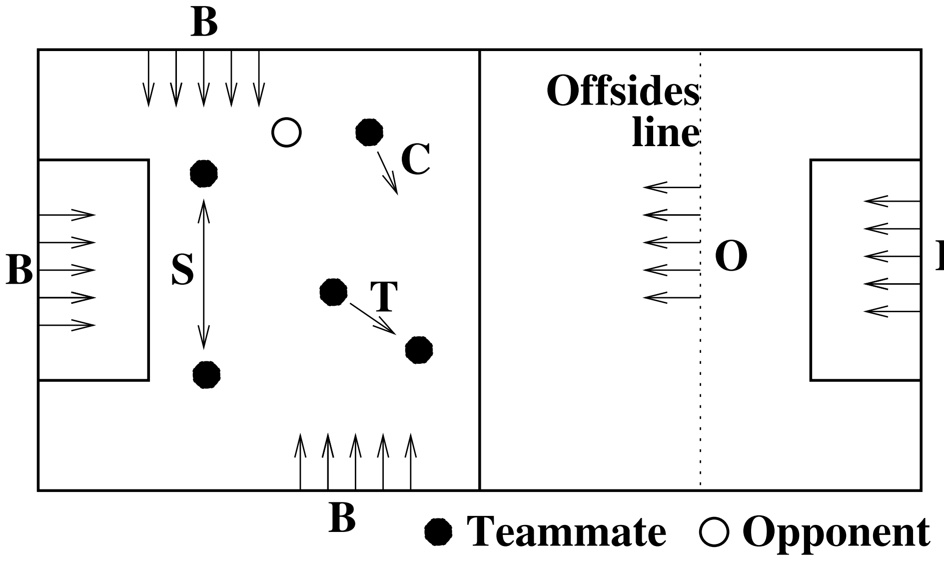
In principle: derivative of value function

Here: vector sum of force fields



$$d_b \equiv \text{distance of the player to the ball}$$
$$F \equiv B + O + \text{if}^*(d_b <^{10} 20,\ T + C,\ S)$$

Force Fields



Bounds-Repellent (B): Stay on the field
Offsides-Repellent (O): Stay on-sides
Strategic (S): Stay about 20m from teammates
Tactical (T): But not too close
Get-clear (C): Move away from “key” defender

Results

- Keepaway vs. CMUnited-99
  - Goal: maintain possession
  - No offensive or defensive reasoning

- Possession time in 95% confidence intervals

Program	Possession Time	Mean Ball x Position
CMUnited-99	5.7-6.6 sec	-19.5
New Team	16.9-18.7 sec	-33.6

Very insensitive to most parameters

Varying S

S^b: Force of unit magnitude towards the ball

S^d: Force downfield

S^*: S, $S + S^b$, $S + S^d$, or $S + S^b + S^d$

$$F \equiv B + O + \text{if}^*(d_b <^{10} 20,\ T + C,\ S^*)$$

Program	Possession Time	Mean Ball x Position
CMUnited	5.7-6.6	-19.5
S	16.9-18.7	-33.6
$S + S^b$	24.8-27.9	-35.9
$S + S^d$	22.2-25.2	25.7
$S + S^b + S^d$	23.7-26.8	26.6

Overall Results

- CMUnited-99 vs. CMUnited-99: **0.3 – 0.3**
- New Team vs. CMUnited-99: **2.5 – 0.3**

RoboCup-2000 Competition

- ATT-CMUnited-2000: 3rd place
  - Stone, Riley, McAllester, Veloso
  - Also included dynamic set plays

    [Riley & Veloso, 2001]
  - 35-team field; Total score: 26–11 (8 games)

Summary

- An option-based action-selection architecture
- Leading Passes in RoboCup soccer
- Force Field Control for Off-Ball Motion

Related Work

- Samba [Riekki & Roenig, '98]: force fields for action selection
- SPAR [Veloso et al., '99]: limited regions, hard constraints

Future Work

- Learn the option value functions using RL