

**ROBOSOCCER  
&  
REINFORCEMENT LEARNING**

Jiří Iša, 4.12.2003



By the year 2050, a team of fully autonomous humanoid robot soccer players shall win the soccer game, comply with the official rule of the FIFA, against the winner of the most recent World Cup.

*[www.robocup.org](http://www.robocup.org)*

# Co je RoboCup?

- Technické konference
- Mistrovství světa v robotickém fotbalu a přidružené konference
- Mezinárodní výzkum
- Vzdělávací iniciativa
- Výzva

# Co je na fotbalu tak vyzývavého?

- Autonomní agenti
- Týmová spolupráce
- Volba strategie
- Rozhodování v reálném čase
- Robotické (technické) problémy
- Dynamické prostředí
- Neúplná informace

# Porovnání šachu a fotbalu

	<b>Šachy</b>	<b>RoboFotbal</b>
<i>Prostředí</i>	Statické	Dynamické
<i>Změna stavu</i>	Tahy	Průběžně
<i>Informace o prostředí</i>	Úplná	Neúplná
<i>Vstupy</i>	Symbolické	Reálné
<i>Řízení</i>	Centralizované	Distribuované

# A k čemu je to dobré?

- Samo o sobě k ničemu
- Milník ve vývoji
- Program Apollo: Dopravit člověka na Měsíc a zpět
- Spousta objevů jako vedlejší efekt
- Investorsky poutavé :-)

# Historie

- 1992 prof. Alan Mackworth: "On seeing robots"
- 1993 Minoru Asada, Yasuo Kuniyoshi, Hiroaki Kitano; Itsoku Noda -> J-League
- 1996 Pre-RoboCup: 8 družstev, střední liga
- 1997 Nagoya: 1. mistrovství
  - Přes čtyřicet účastníků
  - Více jak 5000 diváků

# Tři základní pilíře RoboCupu

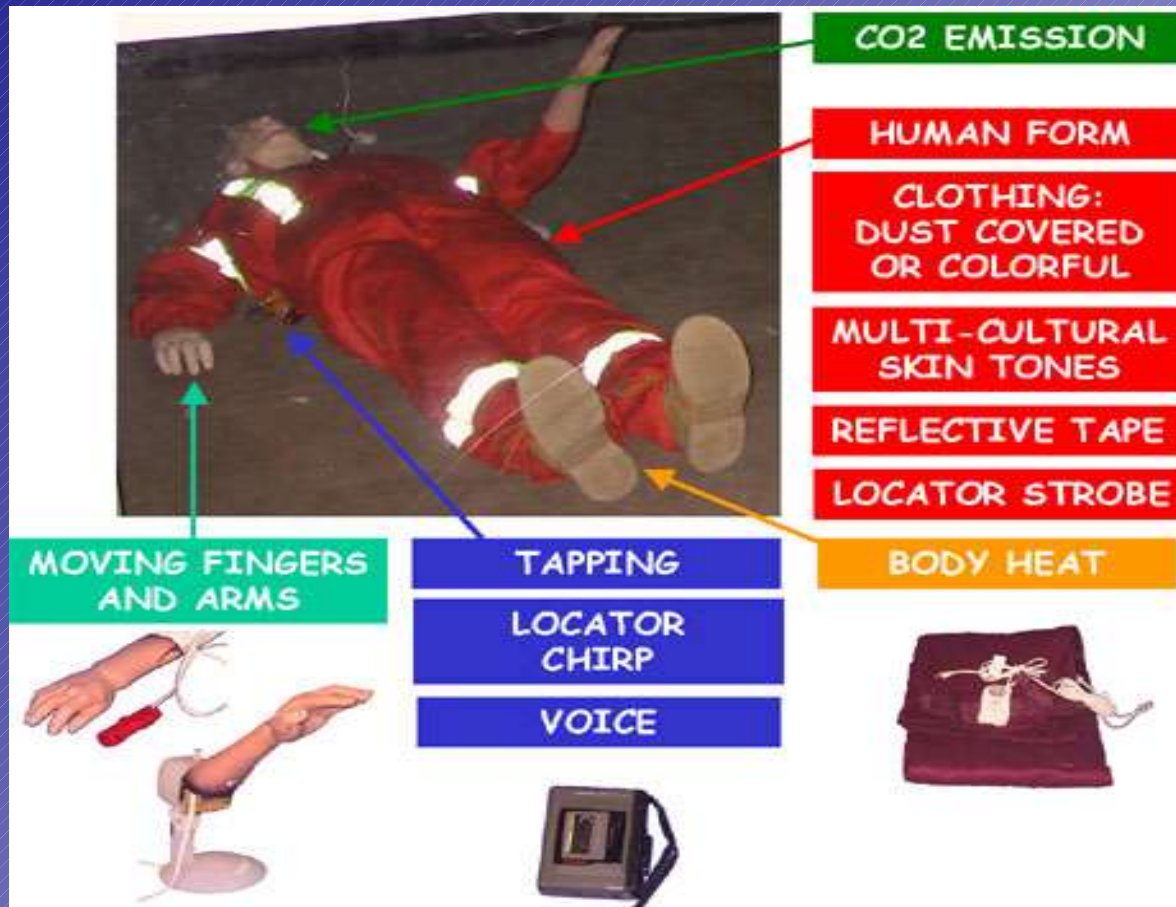
- RoboCup Soccer
- RoboCup Rescue
- RoboCup Junior

# RoboCup Rescue

- Simulované havárie

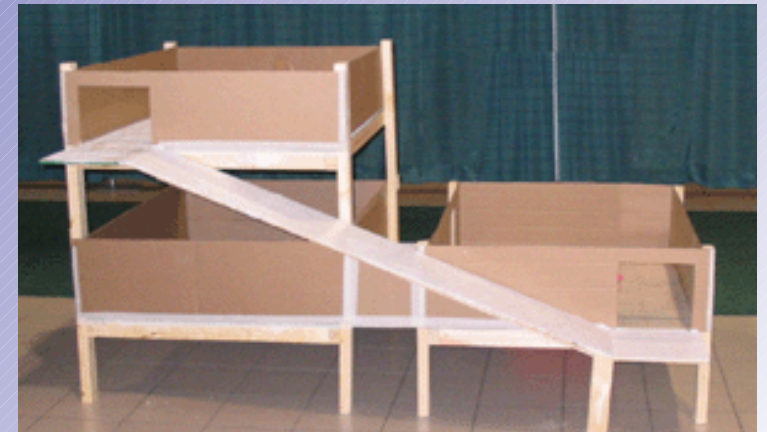


# RoboCup Rescue - Oběť



# RoboCup Junior

- RoboSoccer
- Rescue
- Dance



# RoboSoccer

- Liga malých robotů – F180
- Střední liga – F2000
- Liga humanoidních robotů
- Sony čtyřnozí roboti
- Simulovaný fotbal

# F180 - Liga malých robotů

- Hráč se musí v každém okamžiku (i při použití kopacího zařízení) vejít do válce o průměru 18 cm.
- Hráč smí mít na výšku 22.5 cm, používá-li vlastní optické zařízení. Jinak nesmí přesáhnout 15 cm (např. při společné kameře umístěné nad hřištěm)
- V každém okamžiku musí být 80% míče mimo půdorys hráče

# F180 - Herní plocha

- Hřiště: 2,9m x 2,4m
- Povrch: matně zelený, dřevěný
- Ohraničení hřiště: bílá šikmina o výšce 5cm a sklonu  $45^\circ$  ( $\pi/4$ ;-)

# F180 - Pravidla

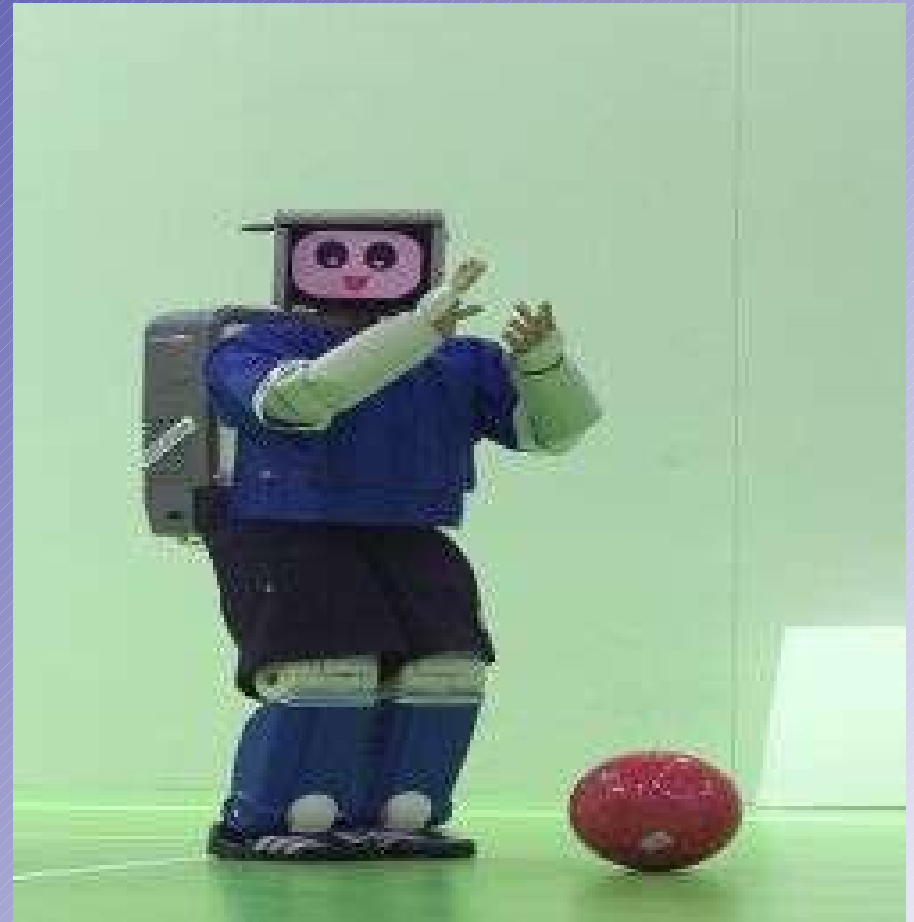
- Bezdrátová komunikace s počítačem
- 2x10min
- FIFA pravidla (s výjimkou ofsidu)
- Nejvýše pět hráčů, z toho jeden předem určený brankář

# F2000 - Liga středních robotů

- 50cm x 50cm (po méně jak třetinu hrací doby povoleno 60cm x 60cm – kvůli kopacím zařízením)
- Výška: 30 – 85 cm
- Maximální hmotnost: 80kg
- Pravidlo jedné třetiny míče
- Vlastní vidění
- Bezdrátová komunikace s počítačem

# Liga humanoidních robotů

- Od roku 2002
- Stoj na jedné noze (Tokyo)
- Chůze (penalizace za dopomoc)
- Penalty
- Penalizační bodování



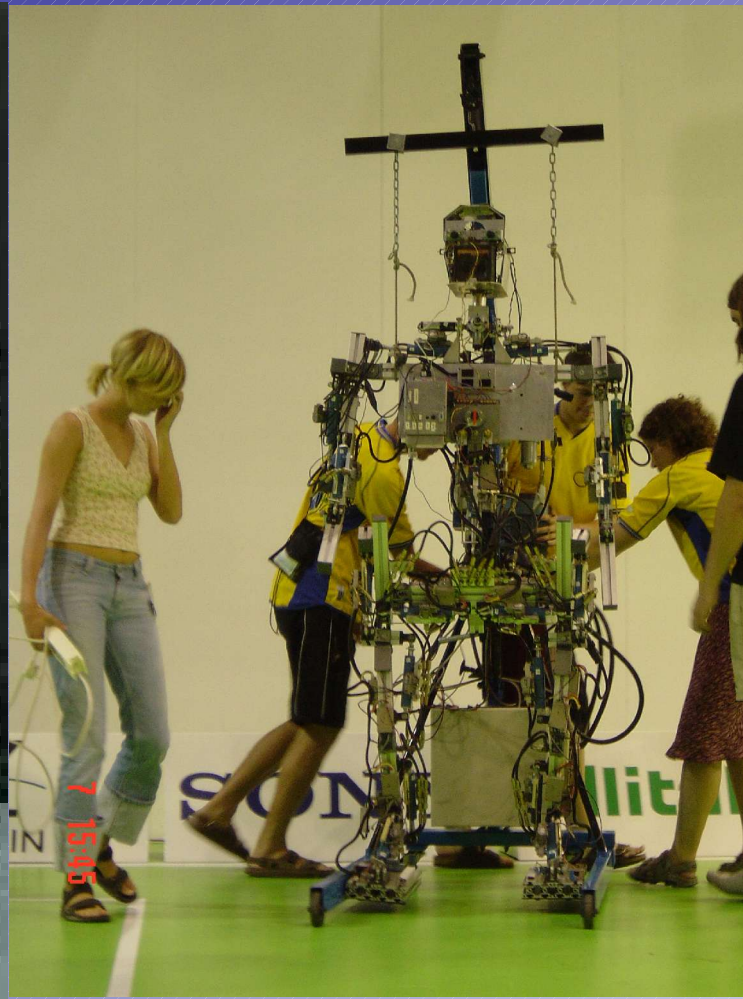
# Humanoidní - do 40cm



# Humanoidní - do 80cm



# Humanoidní - nad 80cm



# Sony čtyřnozí

- Stejná hardwarová platforma
- Zákaz jakýchkoli hardwarových modifikací



# REINFORCEMENT LEARNING

# Motivace

- Změna prostředí -> schopnosti nízké úrovňe
- Neznámý soupeř -> schopnosti vyšší úrovňe

# Potřeby

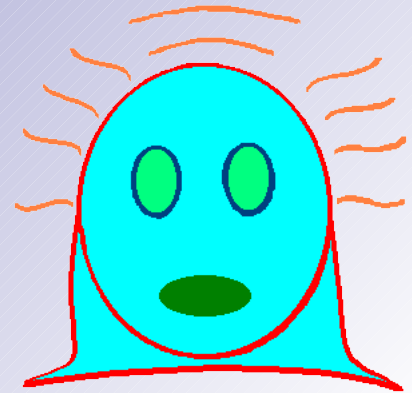
- Změna v reálném čase
- Stabilita

# Postup

- $Q_{k+1}(s_t, a_t) := (1-\alpha)Q_k(s_t, a_t) + \alpha[R(s_t, a_t) + \gamma \max_{a \in A} Q_k(s_{t+1}, a_{t+1})]$ 
  - $s_t$  ... stav v čase  $t$
  - $a_t$  ... akce vykonaná v čase  $t$
  - $Q_k(s_t, a_t)$  ...  $k$ -tá verze ohodnocení volby akce  $a$  za stavu  $s$
  - $\alpha$  ... rychlost učení
  - $\gamma$  ... faktor pesimismu (*discount factor*)

# Wumpus

- *Wumpus*:
  - Smrad'och -> Na polích okolo *Wumpa* je smrad
  - Tři *Wumpové* střeží poklad v jeskyni o rozměrech 6x6 polí
  - Lovec pokladů (Tomb Raider) hledá zlato a snaží se nešlápnout na *Wumpa* (=smrt usmrděním)



# Wumpus - zjednodušení

- Zlato je dosažitelné po nesmradlavých polích
- Výchozí pole je nesmradlavé
- Zlato je v nejvzdálenějším kvadrantu od výchozího pole *Lovce pokladů*

# Wumpus - vjemy

- Smrad na aktuální pozici
- Ví, které pole už navštívil (každé jednotlivě)
- Ví, odkud přišel na aktuální pozici

# Předprogramovaný *Lovec pokladů*

- Algoritmus:
  - Je-li vedle tebe neprozkoumané pole a tady to nesmrdí, jdi tam
  - Smrdí-li to tu, vrať se na předchozí pole a zkus to jinudy
- Průchod do hloubky
- Pro každou možnou kombinaci *situace x akce* je dáno ohodnocení, volí se nejvhodnější akce

# Posilovaný Lovec pokladů

- Trest za smrt (-1000 bodů)
- Bonus za zisk zlata (1000 bodů)
- Pomocí reinforcement learning se snaží najít nejvhodnější ohodnocení dvojic *situace x akce*

# Porovnání

- Učící se *Lovec* dosáhne do dvaceti kol výkonů podávaných předprogramovaným *Lovcem*
- Po čtyřiceti kolech je učící se *Lovec* v průměru o dva kroky rychlejší
- Po zhruba dvoustech kolech dojde k přeučení a oba *Lovci dosahují* stejných výsledků

# RoboSoccer - Layered Learning

- Layered Learning:
  - CMUUnited 97, CMU98:
    - Manuela Veloso, Peter Stone – *A layered approach to learning client behaviors in the RoboCup soccer server; Applied Artificial Intelligence, 12:165-188, 1998*
    - 110:0; 66:0

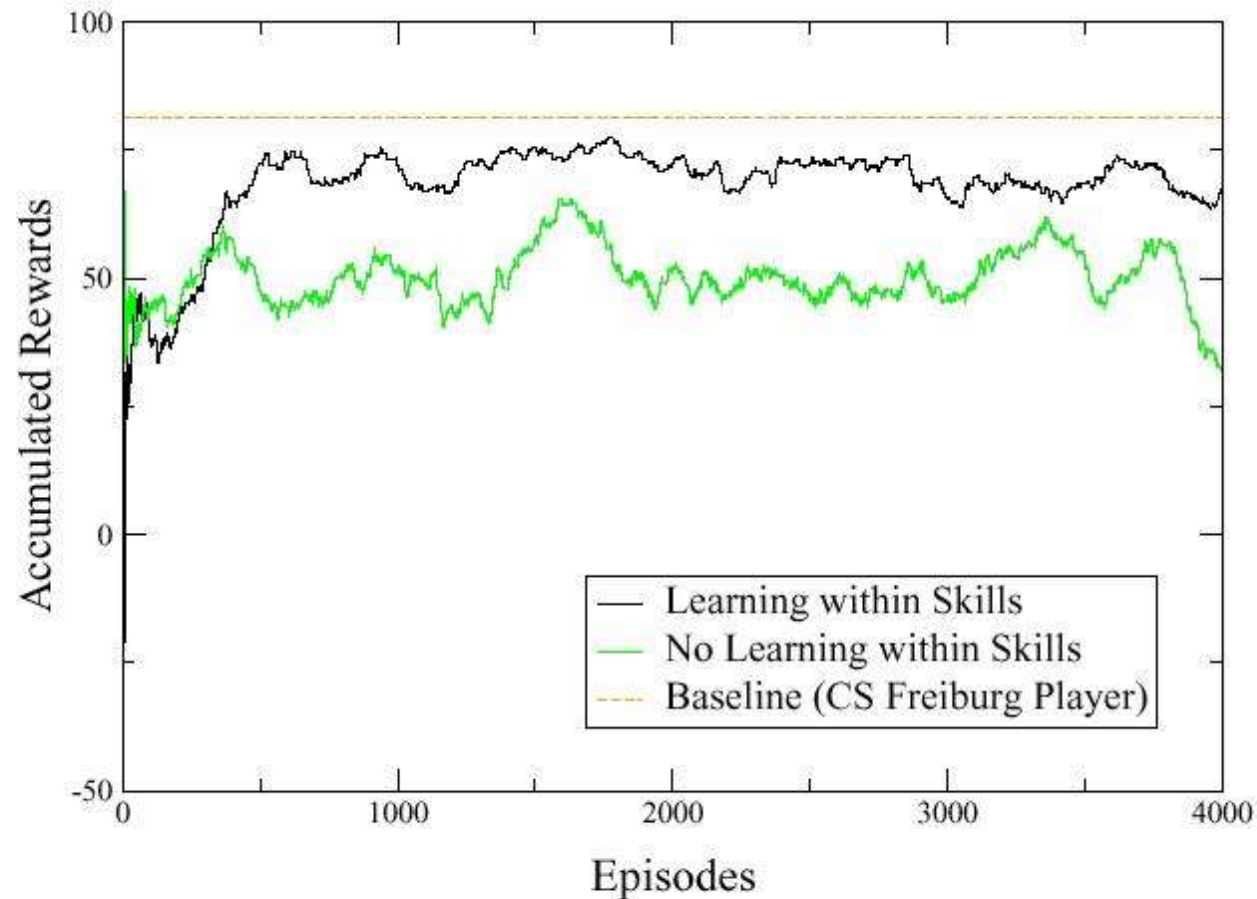
# Freiburg

- Dvě úrovně učení:
  - Nižší úroveň:
    - Najít míč
    - Běžet s míčem
    - Vystřelit na bránu
  - Vyšší úroveň:
    - Kterou akci zvolit

# Freiburg - cíl pokusu

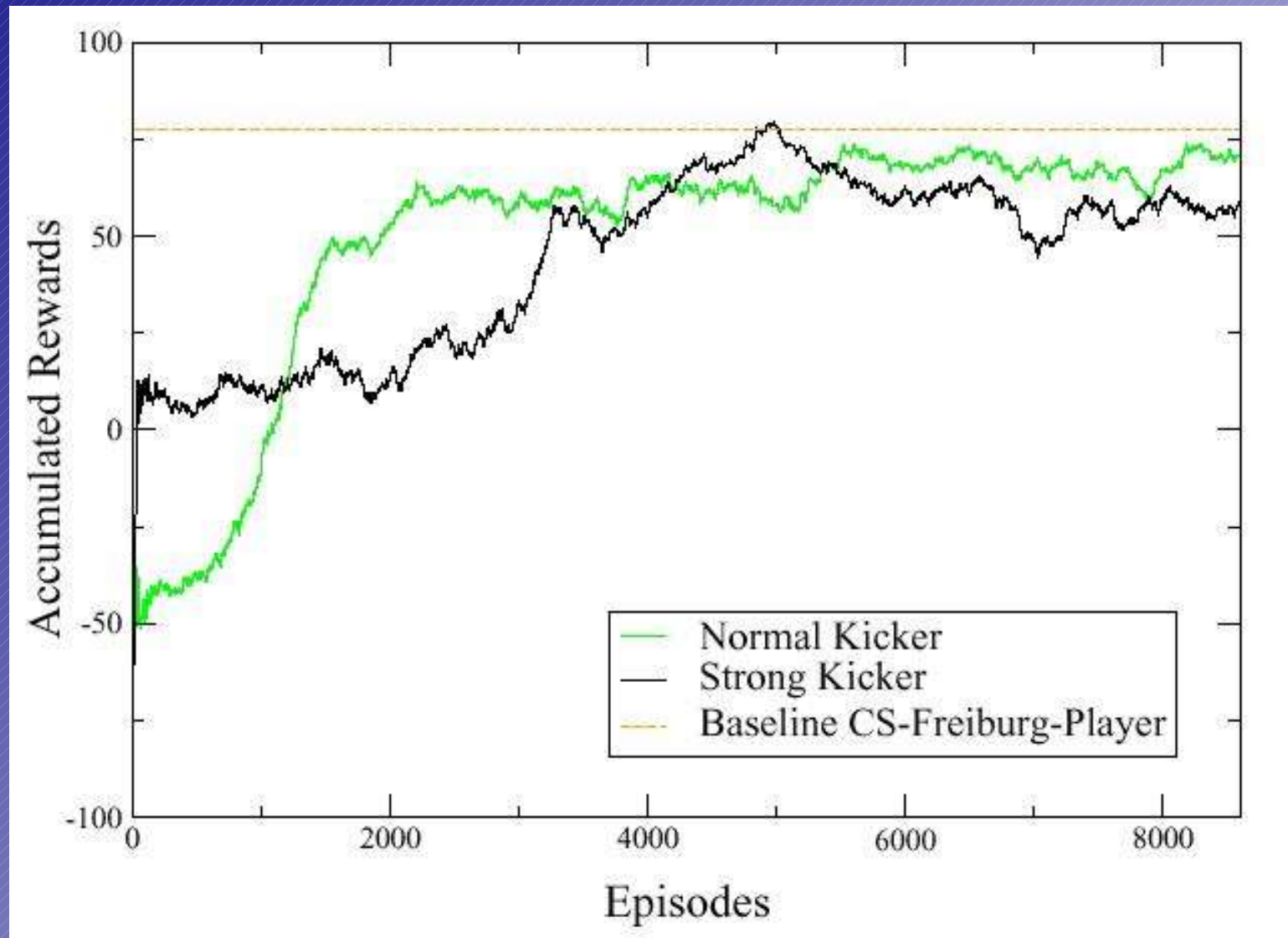
- Reakce na změny prostředí
- Rozdílnost hardware
- Porovnat jedno- a dvouúrovňové učení

# Jedno- vers. dvouúrovňové učení



**Fig. 1.** Learner versus static goalkeeper with and without learning within skills

# David a Goliáš



# Zdroje

- *Towards a Life-Long Learning Soccer Agent:* A. Kleiner, Markus Dietel, Bernhard Nebel
- [www.robocup.org](http://www.robocup.org)
- *RoboCup Soccer Server – Users Manual:* Itsuki Noda, Oliver Obst, ...
- *Model Based Reinforcement Learning:* Richard S. Sutton, Leonid Kuvyaev
- *The CMUnited-98 Champion Simulator Team:* Peter Stone, Manuela Veloso, Patrick Riley
- *Reinforcement Learning in The Multirobot domain:* Maja J. Matarić
- <http://citeseer.nj.nec.com>

?