



LATVIJAS UNIVERSITĀTE  
DATORIKAS  
FAKULTĀTE

# ATDARINOŠĀS MAŠĪNMĀCĪŠANĀS PIELIETOJUMS ROBOTIKĀ

## MAGISTRA KURSA DARBS

Darba autors: **Pēteris Račinskis**  
Darba vadītājs: Dr. Sc. comp. Modris Greitāns

### DARBA MĒRĶIS

**Struktūra:** teorētisks nozares pārskats

**Motivējošā problēma:** atkritumu šķirošanas līnijās vai citur. Robots no konveijera ņem neregulāras formas objektus, šķiro tvertnēs ar metienu palīdzību.

**Tematika:** atdarināšanu kā piemērotu risinājumu apsver Elektronikas un datorzinātņu institūts, mazāki šķēršļi nekā *reinforcement learning*.

**Jānoskaidro:** nozares apakšvirzieni; kādi pētījumi jau veikti; kādas metodes pastāv; kādi ir to sasniegtie rezultāti; vai var piemērot tos motivējošam uzdevumam.

**Jāizstrādā:** provizorisks risinājums, plāns tālākai pētnieciskai darbībai

### Pētniecības virziens: TRAJEKTORIJU KOPĒŠANA

**Uzdevums:** dotas trajektorijas ar tiešiem novērojumiem  $s_t$  un darbībām  $a_t$ . Kā precīzāk, efektīvāk atdarināt?

**Metodes:** uzvedības klonēšana - modelis trenēts aproksimēt demonstrāciju,  $\pi(s_t) \cong \pi_{\text{dem}}(s_t)$ . Modeļu tipi - neironu tīkli, SVM, citi klasifikatori vai regresori.

**Izaicinājumi:** stratēģiju inducēto stāvokļu sadalījumu diverģence - situācijas, kas demonstrāciju kopā nav redzētas. Pielieto statistiskas korekcijas, uzdevumu dekompozīciju, *inverse reinforcement learning*.

### Pētniecības virziens: VISPĀRINĀŠANA, ADAPTĀCIJA

**Uzdevums:** demonstrācijas nav optimālas - uzlabot; atdarināt jaunus paraugus ar minimālu papildus apmācību.

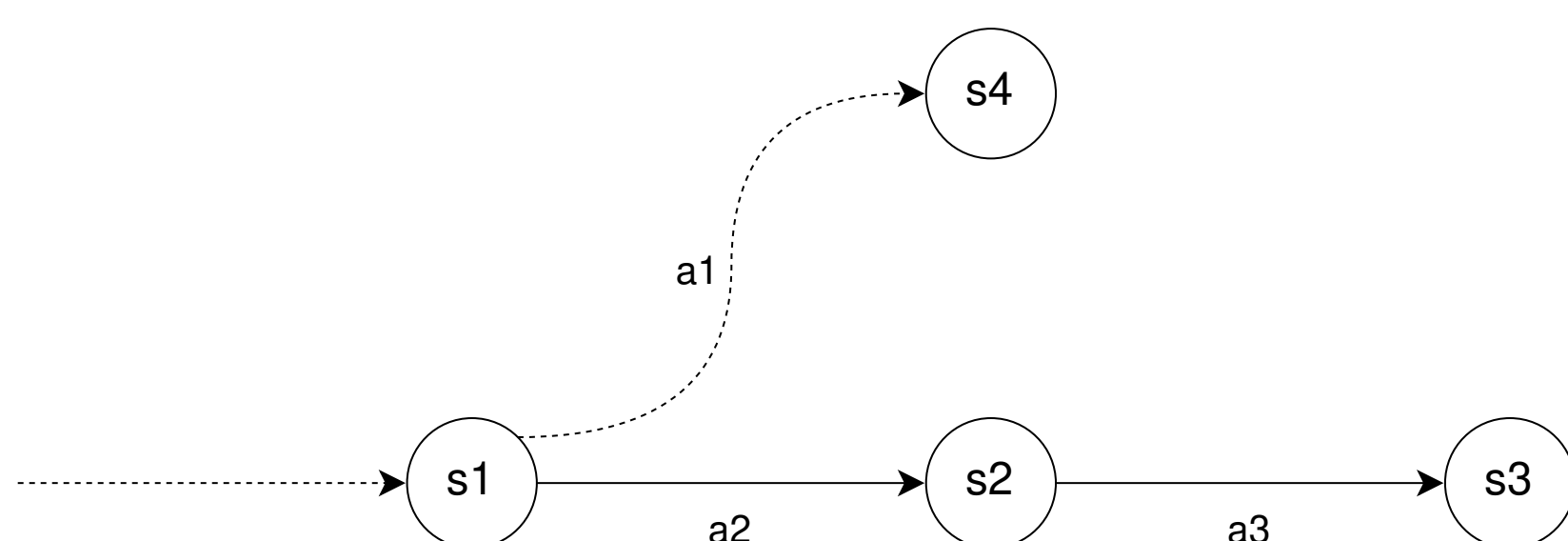
**Metodes:** *reinforcement learning* atalgojuma funkciju atjaunošana un ekstrapolācija; tūlītēja (*one-shot*) atdarināšana, par modeļa argumentu ņemot veselu demonstrāciju; demonstrāciju kopas kā inicializācija *reinforcement learning* procesiem; demonstrāciju iegūšana no nestrukturētām datu kopām (*learning from play*), pārejas stratēģiju iegūšana starp patvaļīgiem stāvokļiem.

### ATDARINOŠĀ MAŠĪNMĀCĪŠANĀS IMITATION LEARNING

**Uzdevums:** Markova lēmumu process (MDP)

**Zināmas:** demonstrāciju trajektorijas ( $s_t, a_t$ )

**Iegūstama:** stratēģija  $\pi(s_t) = a$



### Pētniecības virziens: NOVĒROJUMU IEGŪŠANA, PAPILDINĀŠANA

**Uzdevums:** dotas trajektorijas ar netiešiem novērojumiem  $o_t = f(s_t)$  (piemēram, attēliem), iztrūkstošām darbībām  $a_t$ . Nepietiekami treniņa datu apjomi. Mainīgi darba apstākļi.

**Metodes:** aizvieto trūkstošās darbības ar sistēmas dinamikas iemācīšanos, demonstrācijās novēroto pāreju klasifikāciju. Mainīgas slodzes var kompensēt, ja eksperimentāli nosaka nominālās slodzes. Papildināšanai - datus sintezē simulatorā. Attēliem - konvolūciju neironu tīkli, perspektīvu pārnese ar enkoderiem.

### RISINĀJUMA PLĀNS

**Izstrādes vieta:** Elektronikas un datorzinātņu institūts, robotikas laboratorija

**Pieņēmumi:** objekta satveršanas un klasifikācijas uzdevumi ir atrisināti.

**Demonstrāciju ģenerēšana:** izmantojot kinemātikas ieraksta aprīkojumu (kmeras) vai VR saskarni. Iespējams - kompaktas metiena reprezentācijas aprēķins.

**Būvējamais modelis:** parametrizēts pēc metiena galamērķa, jaunu metienu programmēšana ar koordinātēm. Vispārināma apmācība no demonstrācijām, papildināta ar *reinforcement learning* optimizācijai.