Porazdeljene inteligentne programske tehnologije

Učenje v večagentnih sistemih

Danijel Skočaj
Univerza v Ljubljani
Fakulteta za računalništvo in informatiko

Literatura: MAS, G. Weiss, Poglavje 6

Učenje

- Kaj je učenje?
  - inteligentni agenti so sposobni učenja
  - učeči se agenti so inteligentni
  - nadgrajevanje znanja
  - izboljševanje sposobnosti
  - povečevanje verjetnosti, da se bo v prihodnosti boljše odrezal
  - učenje na izkušnjah
  - prilagajanje na novo zaznane vhodne signale
  - posplošitev znanja

Strojno učenje in MAS

- Strojno učenje (machine learning, ML)
  - uveljavljena disciplina
  - zelo veliko predlaganih metod za učenje
  - aplicirana na zelo veliko domenah
  - ponavadi centralizirano učenje
  - izoliran, samostojen proces
- Večagentni sistemi (MAS)
  - porazdeljeni sistemi
  - prepleteni procesi
  - ponavadi strukturna organizacija in funkcionalno obnašanje predvideni v naprej in fiksni
- Presek ML in MAS -> učenje v večagentnih sistemih!
  - nadgraditi MAS s sposobnostjo učenja
  - razširiti pogled na strojno učenje z večagentnim učenjem (lahko izboljša razumevanje splošnih principov učenja)

Zakaj učenje v MAS

- Razlogi za učenje
  - večagentni sistemi naj bi delovali v kompleksnem okolju
    - veliko, odprto, dinamično, nepredvidljivo okolje
  - nemogoče je v naprej predvideti vse mogoče situacije
    - kakšno bo okolje
    - koliko agentov bo prisotnih
    - kako se bodo agenti odzivali
    - kako se bo spreminjalo okolje…
  - nemogoče je v naprej specificirati sistem in programirati vsega potrebnega znanja

  => potrebno je omogočiti učenje – da agenti izboljšujejo svoje in skupne sposobnosti
- učenje v MAS je kvalitativno drugačno kot izolirano
  - vpliv drugih agentov na učenje (pomagajo, zavirajo, omogočajo učenje…)
  - sodelovanje, posredovanje znanja, komunikacija…

Karakterizacija učenja

- Učenje v MAS je večstranski problem
- različni pristopi, različna terminologija:
  - mutual learning
  - cooperative learning
  - collaborative learning
  - co-learning
  - team learning
  - social learning
  - shared learning
  - plularistic learning
  - organizational learning
  - interactive learning
  - active learning

Glavni kategoriji

- Dve glavni kategoriji
  - centralizirano (izolirano) učenje
  - decentralizirano (interaktivno) učenje
- Učni proces: vse aktivnosti (načrtovanje, sklepanje, odločanje, popravljanje predstavitev, ipd.), ki so izvedene s ciljem doseči določeni učni cilj (se nečesa naučiti).
- Centralizirano učenje:
  - celoten proces učenja je izveden s strani enega agenta
  - agent je pri tem popolnoma neodvisen od ostalih agentov
  - obnaša se, kot bi bil sam
- Decentralizirano učenje
  - več agentov je vpletenih v en učni proces
  - učne aktivnosti se izvajajo na več agentih
  - potrebnih je več agentov

Različni tipi učenja

- V večagentnem sistemu je lahko hkrati aktivnih več centraliziranih učečih se agentov
- lahko imajo različne ali isti cilj
- lahko je več agentov ali skupin agentov, ki so lahko vpleteni v decentralizirane učne procese
- lahko imajo različne ali isti cilj
- vsak agent je lahko vključen v več centraliziranih ali decentraliziranih učnih procesov

Značilnosti različnih tipov učenja

- Stopnja decentralizacije
  - porazdeljenost
    - en agenta
    - veliko agentov v učnem procesu
  - paralelizem
    - zaporedno izvajanje
    - popolnoma paralelizirano izvajanje

Značilnosti različnih tipov učenja

- Interakcija
  - stopnja interakcije
    - samo opazovanje
    - enostavno pošiljanje signalov
    - izmenjava informacij
    - kompleksni dialogi in pogajanja
  - vztrajnost interakcije
    - kratkoročna
    - dolgoročna
  - pogostost interakcije
    - redka
    - zelo pogosta
  - urejenost vzorca interakcije
    - popolnoma nestrukturiran
    - popolnoma hierarhičen
  - spremenljivost interakcije
    - fiksna
    - spremenljiva

Značilnosti različnih tipov učenja

- Vpletenost agenta
  - relevanca vpletenosti
    - vpletenost nekega agenta ni nujno potrebna
    - cilj se ne da doseči brez sodelovanja agenta
  - vloga igrana med vpletenostjo
    - splošen agent (agent lahko izvede vse učne aktivnosti)
    - specialist (agent je specializiran za določene učne aktivnosti)
- Cilji
  - tip izboljšave, ki se jo hoče doseči z učenjem
    - izboljšanje posameznega agenta
    - izboljšanje skupine agentov
  - kompatibilnost učnih ciljev, ki jih zasledujejo agenti
    - konfliktni
    - komplementarni učni cilji

Značilnosti različnih tipov učenja

- Učna metoda
  - rutinsko učenje
    - vgrajeno znanje, ni potrebno sklepanje pri učenju
  - učenje z navodili in nasveti
    - operacionalizacija – pretvorba podane informacije v obliko, ki jo uporablja učenec in integracija s predhodnim znanjem
  - učenje iz primerov in izkušenj
    - izločanje in izboljševanje znanja ter generalizacija le tega iz pozitivnih in negativnih primerov in izkušenj
  - učenje z analogijo
    - prenos znanja z rešenega na nov problem
  - učenje z raziskovanjem
    - zajemanje novega znanja z opazovanjem, izvajanjem eksperimentov, generiranjem in testiranjem hipotez)

Značilnosti različnih tipov učenja

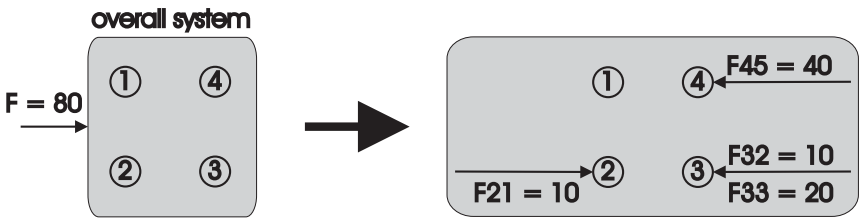
- Odziv pri učenju
  - nadzorovano učenje (supervised learning)
    - odziv je v naprej podan, učenje mora nadgraditi model tako, da se bo rezultat ujema z odzivom
    - okolje (ali drugi agent, človek) deluje kot učitelj
  - vzpodbujevalno učenje (reinforcement learning)
    - (spodbujevano učenje, učenje z ojačitvijo, ojačano učenje, vsiljeno učenje, učenje z vzpodbujanjem, vzpodbujevano učenje, vzpodbujevalno učenje)
    - za odziv učenec dobi podatek o koristi učne akcije
    - cilj je dolgoročno maksimizirati koristi
    - okolje deluje kot kritik
  - nenadzorovano učenje
    - učenec ne dobi nobenega odziva
    - sam mora ugotoviti koristi učnih akcij s poskušanjem ali samoorganizacijo
    - okolje deluje zgolj kot opazovalec

Problem dodelitve zaslug (CAP)

- Credit-Assignment Problem (CAP)
- Kako pravilno upoštevati odziv – dodeliti zasluge (pozitivni odziv) ali penale (negativni odziv) za skupno spremembo v učinkovitosti (pozitivno ali negativno)
- Kako vedeti katera aktivnost v učnem procesu je bolj prispevala k spremembi?
- V večagentnem sistemu je ta problem še bolj izrazit
- Dva podproblema:
  - med-agentni problem (inter-agent CAP)
    - določitev eksternih akcij agentov, ki jim gredo zasluge za celotno spremembo učinkovitosti
    - katera zunanja akcija katerega agenta je koliko prispevala k spremembi
  - znotraj-agentni problem (intra-agent CAP)
    - določitev internih akcij (sklepanj, odločitev) agenta, ki jim gredo zasluge za določeno eksterno akcijo
    - katero znanje, sklepanja in odločitve (notranje akcije) so vodile k neki zunanji akciji
- V praksi je zelo težko ločiti ta dva podproblema!

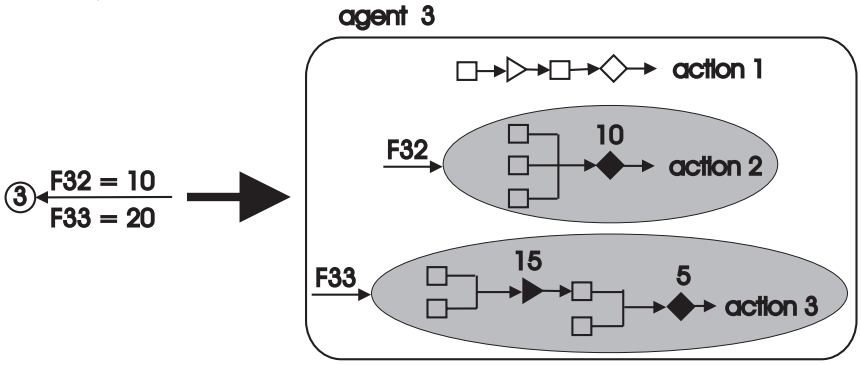
Med-agentni CAP

- Primer
  - štiriagentni sistem
  - izvedenih je bilo več akcij, s končnim odzivom F=80
  - Dekompozicija na osnovi akcij posameznih agentov:
    $F_{i,j}$ - i–ti agent in j–ta eksterna akcija



Znoraj-agentni CAP

- Primer za Agenta 3, ki je izvedel tri interne akcije
- Notranje akcije temeljijo na:
  - □ notranje znanje
  - ◇ odločitve
  - ▷ sklepanja



Učenje in koordinacija aktivnosti

- Kako naj se agenti naučijo primerno koordinirati njihove aktivnosti
- Agenti se lahko koordinirajo na v naprej določen način, s pravili, pogajanji, ipd.
- Lahko pa se tudi naučijo in prilagajajo novim situacijam
- Agenti se morajo naučiti prilagajati e ostalim agentom in koordinirati aktivnosti

- Dva primera:
  - Vzpodbujevalno učenje za koordinacijo v večagentnih sistemih, kjer se agenti ne zavedajo drug drugega
  - skupina sodelujočih učencev

Vzpodbujevalno učenje

- Reinforcement learning

- Agenti imajo opis trenutnega stanja in morajo izbrati naslednjo akcijo, ki naj bi maksimizirala odziv
- Okolje učenca se lahko modelira z Markovim procesom z diskretnim časom in končnim stanjem
- Okolje v MDP je predstavljeno s četverčkom <*S*,*A*,*P*,*r* >
  - S je množica stanj
  - A je množica akcij
  - $P : S \times S \times A \mapsto [0,1]$
  - $r : S \times A \mapsto \Re$ skalarna funkcija nagrade (odziva)
- Vsak agent ohranja strategijo π, ki preslika trenutno stanje v želeno akcijo, ki naj bi bila izvedena nad tem stanjem

- Različni pristopi k vzpodbujevalnemu učenju

Q-učenje

- Poišči strategijo π^*, ki maksimizira vsoto vseh prihodnjih nagrad za vsa možna stanja s∈S
- Odločitvena funkcija je predstavljena s funkcijo (tabelo)
  $Q : S \times A \mapsto \Re$
  = dolgoročna nagrada za vsak par stanje-akcija

- Q-vrednosti se sproti spreminjajo

Učenje klasifikatorskih sistemov

- Klasifikatorski sistemi
  - temeljijo na pravilih
  - se učijo z uravnavanjem moči posameznih pravil
  - ob upoštevanju odziva okolja
  - in z odkrivanjem novih pravil z genetskimi algoritmi

- En klasifikator ima en pogoj in eno akcijo:
  - klasifikator i: (c_i, a_i)
  - moč klasifikatorja i ob času t: $S_t(c_i,a_i)$

BBA algoritem

- Cikel:
  - Vsem klasifikatorjem se na začetku priredijo naključne moči
  - na vsakem koraku se izbere eden izmed klasifikatorjev, ki vžge
  - na osnovi njegve akcije se dobi odziv iz okolja
  - moči klasifikatorjev se ustrezno popravijo
- Cikel se ponovi, niz takih ciklov tvori poizkus

- BBA agoritem (Bucket Brigade Algorithm) za dodeljevanje zaslug:
  - klasifikatorju, ki vžge se pripišejo skoraj vse zasluge
  - del zaslug se pripiše klasifikatorju, ki je vžgal pred njim

- Shema učenja zelo podobna Q-učenju

Izolirani, sočasni, vzpodbujevalni učenci

- Isolated, Concurrent Reinforcement Learners

- Pri vzpodbujevalnem učenju agenti tvorijo preslikave med zaznavami in akcijami in s primerno izbiro akcij skušajo maksimizirati odziv okolja
- Uporabijo se lahko tudi, ko imajo agentje malo predhodnega znanja in poznavanja sposobnosti in ciljev drugih agentov

- Naj agentje izberejo komunikacijo pri učenju koordinacije?
  - komunikacija pogosto olajša skupinske procese
  - ne garantira koordiniranega obnašanja
  - je časovno potratna
  - lahko odvrača od drugih problemov
  - naj bi se ne prenašale velike količine informacij
  - naj bi se agentje ne preveč zanašali na komunikacijo

Izolirani, sočasni, vzpodbujevalni učenci

- Izolirani, sočasni, vzpodbujevalni učenci
- vsak agent se uči zase, brez komunikacije z ostalimi
- upošteva samo odziv okolja
- ne modelira drugih agentov

- Ali je možno uporabiti enak učni mehanizem za kooperativna in ne-kooperativna okolja?

Karakteristike CIRL

- Običajna predpostavka, da dinamika okolja ni odvisna d drugih agentov, ne drži
- Karakteristike CIRL (concurrent, isolated, reinforcement learning):
  - povezovanje agentov (agent coupling)
    - tesno povezan sistem: akcije enega agenta pogosto in močno vplivajo na načrte drugih agentov
    - šibko povezan sistem: akcije enega agenta redko in šibko vplivajo a načrte drugih agentov
  - relacije med agenti
    - kooperativni agenti: agenti sodelujejo v skupini pri reševanju skupnega problema
    - nevtralni agenti: agenti komunicirajo, ker imajo skupne vire, nimajo pa posebnih odnosov
    - agenti tekmeci: agenti imajo nasprotujoče interese
  - čas odziva
    - odziv je na voljo takoj ali čez nekaj časa
  - kombinacije optimalnega obnašanja
    - koliko vrst obnašanja optimalno reši problem (od ena do neskončno pri različnih problemskih domenah)

CIRL eksperimenti

- Potiskanje kocke:
  - dva agenta se individualno naučita potiskati kocko po začrtani poti od starta do cilja
  - kooperatvna situacija (oba agenta mata isti cilj)
  - kompetitivna situacija (vsak agent ima svoj cilj)
  - odziv temelji na razdalji kocke od začrtane poti
  - karakteristike:
    - sočasno delovanje dveh agentov
    - tesno povezan sistem
    - takojšen odziv
    - več optimalnih obnašanj

CIRL eksperimenti

- Souporaba virov:
  - nalogi sta podani
  - dva agenta se morata naučiti deliti vire skozi čas
  - karakteristike:
    - tesno povezan sistem
    - zapoznel odziv
    - eno samo optimalno obnašanje
- Navigacija robota
  - dva robota se naučita navigirati po sekajočih se poteh na mreži brez trčenja
  - karakteristike:
    - spremenljivo močno povezovanje
    - takojšen odziv
    - več optimalnih obnašanj

Zaključki CIRL eksperimentov

- Tako prijatelji kot nasprotniki se lahko naučijo koordinacije s CIRL
- Ni potrebno ne predhodno znanje, ne ekspliciten model drugih agentov
- Slabše se CIRL izkaže, ko
  - je sistem tesno povezan
  - odziv je zapoznel in
  - obstaja samo eno ali malo optimalnih obnašanj
- Ta problem bi lahko rešili z učenjem po fazah
  - nekaj časa bi se agent učil
  - potem bi izvrševal trenutno strategijo
  - potem bi se spet učil, itn.
  - dva agenta bi e lahko koordinirala, da se eden uči, ko se drugi ne
- V kooperativnih situacijah se agenti naučijo komplementarne strategije (se specializirajo)
- Agenti lahko prenesejo znanje na podobne situacije

Interaktivno vzpodbujevalno učenje koordinacije

- Agenti med seboj eksplicitno komunicirajo ob odločitvah o individualnih ali skupinskih akcijah
- Prilagojen BBA algoritem
  - agent lahko spremlja množico akcij, o katerih premišljajo drugi agenti
  - temu ustrezno lahko nekompatibilne akcije izloči iz svoje množice akcij
- Dva algoritma (ki izvirata iz BBA)
  - ACE (Action Estimation)
  - AGE (Action Group Esimation)
  - zahtevata različen nivo vpletenosti in koordinacije med vpletenimi agenti
- Agenti imajo isti cilj – sodelujejo, da bi optimizirali skupinski cilj

ACE – Action Estimation Algorithm

- Podana je zaznava S_i trenutnega stanja S (agenta a_i)
- vsak agent najprej izračuna množico akcij $A_i(S)$, ki jih lahko izvede v stanju S
- za vsako akcijo A_i^j izračuna relevanco $E_i^j(S)$ glede na cilj
- za vse akcije nad pragom agent izračuna in objavi ponudbo
- za izvajanje se izbere akcija z najvišjo ponudbo
- nekompatibilne akcije so izločene
- ta proces se ponavlja dokler niso vse akcije, za katere so bile poslane ponudbe, bodisi izločene bodisi izbrane
- izbrane akcije tvorijo kontekst aktivnosti (activity context)
- BBA mehanizem nato zmanjša oceno izbrane akcije za vrednost ponudbe, ki se razporedi med akcije v prejšnjem kontekstu akcij
- če se po izvedbi akcij v trenutnem kontekstu aktivnosti dobi odziv, se le-ta enakomerno razporedi med vse akcije

AGE – Action Group Estimation Algorithm

- naprej se zberejo vse izvršljive akcije vseh agentv v danem stanju okolja
- nato se izračunajo množice kontekstov akcij, pri čemer je vak kontekst sestavljen iz množice med seboj kompatibilnih akcij
- za vsak kontekst aktivnosti se nato zbere ponudbe od vsakega agenta za vse njegove akcije
- izbere se kontekst aktivnosti, ki ima najvišjo vsoto ponudb
- izbrane akcije se izvedejo na dotičnih agentih
- spremenijo se ocene relevanc za akcije $E_i^j(S,A)$
- kupna vrednost ponudb se prerazporedi med akcije v prejšnjem kontekstu aktivnosti

Primerjava ICE in AGE algoritmov

- Oba algoritma omogočata učenje koordiniranega obnašanja
- agenti se naučijo precej bolj učinkovito izvrševati naloge kot pri naključni izbiri akcij
- AGE algoritem se nauči bolj učinkovite koordinacije
- AGE algoritem je računsko prostorsko in časovno bolj kompleksen
- algoritma nista dosegla globalno optimalnega obnašanja
  - zaradi omejenega zaznavanja
  - ker nista mogla razločiti nekaterih različnih globalnih stanj

Učenje o in od drugih agentov

- ne samo učenje koordinacije
- agenti se učijo izboljšati njihove individualne sposobnosti
- včasih gre to izboljšanje na račun drugih

- agenti naj se učijo o drugih, da bi bolje izkoristili priložnosti
- agenti se učijo, kako lahko na njihovo učenje vplivajo drugi agenti

- predvidevanje, napoved obnašanja drugih agentov
- interakcija z drugimi agenti
- skupni pogled na svet

Učenje o in od drugih agentov

- Tri pristopi:

- Učenje organizacijskih vlog
  - agentje v skupini se naučijo različnih komplementarnih vlog
  - na ta način izboljšajo skupno učinkovitost

- Učenje v tržnih pogojih
  - informacijski agenti, ki prodajajo in kupujejo na trgu
  - se prilagajajo trenutnim razmeram na trgu
  - gradijo modele drugih agentov

- Učenje boljše strategije za igro proti nasprotniku
  - namizne in podobne igre (med dvema igralcema)
  - zgrajen model nasprotnikove strategije pomaga izboljšati lastno strategijo

Učenje organizacijskih vlog

- Agenti se učijo prilagajanja na vloge, specifične za dano situacijo, v kooperativnih problemih
- vsak agent lahko igra eno ali več vlog v vsaki situaciji
- naloga je izbrati najboljšo vlogo (ki bo bolj pomagala pri reševanju problema)

- Uporaba ocen UPC (Utility, Probabilty and Cost) posamezne vloge v dani situaciji
  - korist, verjetnost, cena
  - vsa stanja sveta S so preslikana v manjšo množico situacij
  - korist (utility): ocena končne koristi v končnem stanju, če agent sprejme dano vlogo v dani situaciji
  - verjetnost (probabilty): verjetnost, da bo agent dosegel uspešno zadnje stanje, če bo sprejel dano vlogo
  - cena (cost): računski stroški, ki pri tem nastanejo
  - potencial (potencial) vloge: uporabnost vloge pri odkrivanju primerne globalne informacije in omejitev

Učenje organizacijskih vlog

- Naj bodo S_k, R_k množice vektorjev situacij in vlog agenta k
- Agent hrani do $|S_k| \cdot |R_k|$ vektorjev vrednosti UPC in potencialov za različne vloge v različnih situacijah

- Med učno fazo je verjetnost za izbiro vloge r v situaciji s

$$\Pr(r) = \frac{f(U_{rs}, P_{rs}, C_{rs}, Potential_{rs})}{\sum_{j \in R_k} f(U_{js}, P_{js}, C_{js}, Potential_{js})}$$

kjer je f funkcija, ki se uporablja za ocenitev vloge

- Ko je učna faza mimo, se vloga izbere najbolje ocenjeno vlogo: $r = \arg\max_{j \in R_k} f(U_{js}, P_{js}, C_{js}, Potential_{js})$

Učenje organizacijskih vlog

- UPC vrednosti se naučijo z vzpodbujevalnim učenjem
- Naj bodo $\hat{U}_{rs}^{n}, \hat{P}_{rs}^{n}, \hat{P}otential_{rs}^{n}$ vrednosti po n korakih za vlogo r v situaciji s
- Naj bodo S situacije srečane med začetkom uporabe vloge r v situaciji s in končnim stanjem F (s končno koristjo U_F)
- Vrednosti za koristi (utility) se osvežujejo tako:

$$\hat{U}_{rs}^{n+1} \leftarrow (1-\alpha)\cdot \hat{U}_{rs}^{n} + \alpha \cdot U_F$$

- Osvežijo se vse vloge v vseh situacijah v S (na poti do F)
- Naj $O: S \rightarrow [0,1]$, pove ali je dano stanje uspešno ali ne
- Osvežitveno pravilo za verjetnost: $\hat{P}_{rs}^{n+1} \leftarrow (1-\alpha)\cdot \hat{P}_{rs}^{n} + \alpha \cdot O(F)$
- Naj $Conf(S)$ vrne 1, če obstajajo na poti do F konfliktne poti, ki se razrešijo
- Osvežitveno pravilo za potencial:

$$\hat{P}otential_{rs}^{n+1} \leftarrow (1-\alpha)\cdot \hat{P}otential_{rs}^{n} + \alpha \cdot Conf(S)$$

Učenje organizacijskih vlog

- Osvežitvena pravila za cene so odvisna od domene/problema
- Najtežji problem je kako oceniti funkcijo f
- Primer:

$$f(U, P, C\ , Potential\ ) = U * P + Potencial$$

Učenje v tržnih pogojih

- Agenti lahko kupujejo in prodajajo dobrine na elektronskem trgu
- trg je odprt za nove agente
- agenti so sebični in hočejo maksimizirati svoj dobiček, vsak igra zase
- tržni mehanizem uravnava trgovanje
- agenti lahko dostopajo enako do vseh ostalih agentov

- kupec objavi povpraševanje
- prodajalci objavijo ponudbe
- kupec izbere najbolj ugodno ponudbo in plača
- prodajalec izroči dobrino kupcu

- kupec hoče maksimizirati vrednost (odvisna od cene in kvalitete)
- prodajalec hoče maksimizirati dobiček (odvisen d cene in stroškov)

Tri tipi agentov

- Agenti 0. stopnje
  - ne modelirajo obnašanja drugih agentov
  - cene postavljajo na osnovi akumuliranih izkušenj
- Agenti 1. stopnje
  - Agenti analizirajo preteklo obnašanje ostalih agentov in skušajo predvideti njihove preference glede nakupov in prodaj
  - ostale agente modelirajo le kot agente 0. stopnje (ne upoštevajo, da lahko tudi ti agenti modelirajo druge agente)
- Agenti 2. stopnje
  - Agenti modelirajo druge agente kot agente 1. stopnje

Učenje za zmago nad nasprotnikom

- Namizne igre med dvema igralcema

- Minimax algoritem – minimiziraj maksimalno možno izgubo
- Predvidi, da bo nasprotnik naredil zate najslabšo potezo
- če nasprotnik lahko zmaga v naslednjem koraku, mu to prepreči

- M* algoritem je posplošitev Minimax algoritma
  - uporabi model nasprotnika za izbiro naslednje poteze
  - M^0 ustreza minimax algoritmu (če je evaluacijska funkcija f_0 znana)
  - višji redi algoritma M iščejo rekurzivno bolj v globino
  - igralec M^n predpostavlja, da je nasprotnik igralec M^{n-1} itn.
  - igralec M^1 igra boljše kot igralec M^0 , če je model nasprotnika poznan
  - Vprašanje pa je kako dobiti natančen model nasprotnika!
  - Nauči se model nasprotnika!

Učenje in komunikacija

- Kako sta učenje in komunikacija povezana?
- zahteve za učinkovito izmenjavo informacij med agenti

- Dve glavni relaciji:
  - Učenje za komunikacijo
    - z učenjem se zmanjša količina informacije med agenti
    - komunikacija je počasna in draga, učenjem se jo lahko zmanjša
  - Komunikacija za učenje
    - s komunikacijo se omogoči izmenjava informacij, ki jih agenti potrebujejo pri učenju
    - učenje potrebuje informacije, komunikacija jih lahko priskrbi

Učenje in komunikacija

- Relevantna vprašanja:
  - Kaj komunicirati?
    - kakšna informacija je zanimiva za ostale
  - Kdaj komunicirati?
    - do kolikšne mere naj agent sam poskuša rešiti problem predno vpraša druge za pomoč
  - S kom komunicirati?
    - katere agente zanima naša komunikacija, kateri nam lahko pomagajo
  - Kako komunicirati?
    - na katerem nivoju
    - kateri jezik in protokol naj uporabljajo
    - neposredno ali preko table

Skupna ontologija

- Agentje morajo vedeti o čem se pogovarjajo kaj se na kaj nanaša
- Glavna naloga učenja v porazdeljenem sistemu je zagotoviti enak/skupni pomen simbolov
- Prizemljitev simbolov (symbol grounding problem)
  - prizemljitev (povezava) simbolov v realen svet
  - asociacije med visokonivojskimi simboli in nizkonivojskimi (realnimi) signali (ali entitetami izpeljanimi iz njih)
  - pogoj za realen inteligentni sistem?
- Problem skupnega pomena (shared meaning problem)
  - prizemljitev simbolov v MAS

Učenje za komunikacijo

- Zmanjševanje količine komunikacije z učenjem
- Pogodbene mreže
  - Manager objavi naloge
  - potencialni pogodbeniki pošljejo ponudbe
  - manager izbere najboljšo ponudbo, sklenitev pogodbe
- Pošiljanje ponudb vsem je lahko zelo potratno v velikih sistemih z veliko agenti in/ali veliko nalogami
- Rešitve:
  - usmerjeno naslavljanje
  - neposredno komuniciranje
  - včasih težko izvedljivo, naslovniki niso vedno v naprej poznani
- Elegantnejša rešitev:
  - učenje naslovnikov

Učenje naslovnikov

- Glavni cilj: zmanjšanje količine komunikacije z zajemanjem znanja o drugih agentih (oz. njihovih sposobnostih)
- Z naučenim znanjem je lahko naslavljanje bolj usmerjeno ali neposredno
- Sklepanje na osnovi primerov (Case-based reasoning):
  - problem se reši na osnovi rešitve podobnega primera v preteklosti
  - tvorjenje parov problem-rešitev
  - ko pride nov problem, se najprej pogleda, če že obstaja podoben rešen problem (iskanje problema)
    - če je problem popolnoma nov, se reši od začetka
    - če obstaja podoben problem, se njegova rešitev uporabi kot začetni približek in se ga nato reši
  - Vsi problemi se skupaj z rešitvami shranijo v bazi primerov (case storage)

Učenje naslovnikov

- Vsak agent hrani svojo bazo primerov
- Primer je sestavljen iz
  - specifikacije naloge
  - informacijo o tem kateri agent je že rešil to nalogo in kako dober je pri tem bil
- Specifikacija naloge:
  - atributi naloge
  - vrednosti atributov
- Mera podobnosti med nalogami
  - mera podobnosti med atributi in njihovimi vrednostmi
- Meri podobnosti med atributi in vrednostmi so odvisni od domene, recimo: 1, če sta enaki ali 0 sicer (identiteta)
- Podobnosti upragovimo, da dobimo množico podobnih nalog
- Odločitev komu poslati povpraševanje:
  - za vsakega agenta agent izračuna množico podobnih nalog
  - izračuna primernost agenta z obtežitvijo s preteklimi uspehi
- Agent nato obvesti samo najbolj primerne agente!

Komunikacija za izboljšanje učenja

- Agent ni vsemogočen niti vseveden (razen v zelo omejenih pogojih)
- Pomanjkanje informacije se lahko nanaša na
  - okolje, v katerem je umeščen in problem, ki naj ga reši
  - ostale agente
  - odvisnosti med različnimi aktivnostmi in učinek nekaterih agentovih aktivnosti na okolje in na aktivnosti drugih agentov
- Pomanjkanje informacij se lahko zmanjša s komunikacijo med agenti in se izboljša učenje
  - učenje, ki temelji na nizkonivojski komunikaciji (izmenjava manjkajočih delov informacije)
    - informacija v skupni rabi
  - učenje, ki temelji na visokonivojski komunikaciji (bolj kompleksna interakcija, pogajanja, medsebojno razlaganje in kombiniranje znanja)
    - skupno razumevanje

Komunikacija za izboljšanje učenja

- Komunikacija kot orodje za izboljševanje učenja
- Komunikacija je del učnega procesa (zajemanja znanja)
- Naj bo učenje obogateno s komunikacijo?
  - Kako hitro se učni rezultati dosežejo z ali brez komunikacije?
  - kakšna je kvaliteta rezultatov doseženih z ali brez komunikacije?
  - kako kompleksen je celotni učni proces z ali brez komunikacije?

- Komunikacija lahko zelo poenostavi učenje (zagotovi nove informacije)
- Komunikacija lahko tudi upočasni učenje (če pride do napake in se te napake uporabijo v učnem procesu/naučenih modelih)

Primer 1

- Lovimo skupaj!

- Domena lovec-plen
- kvadratna mreža, 4 možne poteze
- žrtve se gibljejo naključno
- lovci želijo ujeti čim več plena
- lovci imajo omejen vidi doseg

- strategija lovcev temelji na Q-učenju
- uspeh lahko izboljšajo s komunikacijo:
  - Komuniciranje senzorskih podatkov
    - podatki o opaženem plenu, o medsebojnih premikih
    - centraliziran nadzor porazdeljenih senzorjev
  - Komuniciranje odločitvenih strategij in odločitev o akcijah
    - sporočajo si odločitve/aktivnosti (recimo vrednosti Q(ws,a))
    - centraliziran nadzor porazdeljenih efektorjev

Primer 2 – Učenje s konsenzom

- Učenje z visokonivojsko komunikacijo
- Učenje s konsenzom
- Skupina agentov komunicira z uporabo table
- Uporabljajo enostaven jezik, ki vsebuje naslednjih 9 operatorjev za hipoteze:
  - dodajanje in odstranjevanje hipotez s table:
    - ASSERT(H)
    - PROPOSE(H,C)
    - WITHDRAW(H)
  - evaluacija hipotez:
    - CONFIRM(H,C)
    - DISAGREE(H,C)
    - NOOPINION(H)
    - MODIFY(H,G,C)
  - sprememba stanja hipotez in sprejem hipotez:
    - AGREED(H,T)
    - ACCEPT(H)

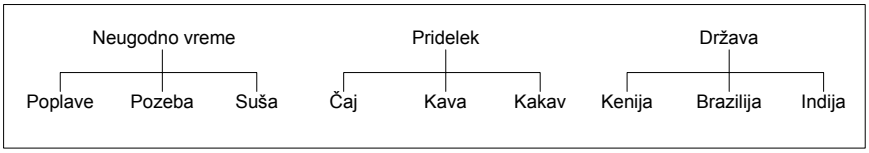
Učenje s konsenzom

- Agent predlaga hipotezo H (s PROPOSE)
- ostali agenti odgovorijo (s CONFIRM, DISAGREE, NOOPINION ali MODIFY)
- Prvi agent nato ugotovi skupno stopnjo zaupanja T v hipotezo H.
  - upošteva tako pozitivne odgovore (CONFIRM in MODIFY)
  - kot tudi negativne odgovore (DISAGREE)

Učenje s konsenzom - primer

- Koliko bo stala skodelica kave?

- Tri agenti, ki bi radi vedeli kako se bodo spreminjale cene kave, čaja in kakava
- Taksonomije znanja agentov:



Učenje s konsenzom - primer

- Lokalno znanje:
  - *Agent1: Glavni-Proizvajalec(Kenija,Kava)*
    Glavni-Proizvajalec(Kenija,Čaj)
  - *Agent2: Glavni-Proizvajalec(Brazilija,Kava)*
    Glavni-Proizvajalec(Brazilija,Kakav)
  - *Agent3: Glavni-Proizvajalec(Indija,Čaj)*

- Zaznani podatki in generalizacije:
  - *Agent1: Vreme(Kenija,Suša),Cena(Čaj,Raste)*
    Vreme(Kenija,Suša),Cena(Kakav,Stabilna)
    Vreme(Kenija,Pozeba),Cena(Kava,Raste)
    GEN: Vreme(Kenija,Neugodno) in
    Glavni-Proizvajalec(Kenija,Pridelek)->
    Cena(Pridelek,Raste)
  - *Agent2: Vreme(Brazilija,Pozeba),Cena(Kava,Raste)*
    Vreme(Brazilija,Poplave),Cena(Kakav,Raste)
    GEN: Vreme(Brazilija,Neugodno) -> Cena(Pridelek,Raste)
  - *Agent3: Vreme(Indija,Poplave),Cena(Čaj,Raste)*
    GEN: Vreme(Indija,Poplave) -> Cena(Čaj,Raste)

Učenje s konsenzom - primer

- Hipoteze:
  - H1= *Vreme(Indija,Poplave) -> Cena(Čaj,Raste)*
  - *H2= Vreme(Država,Neugodno) -> Cena(Pridelek,Raste)*
  - *H3=Vreme(Država,Neugodno) in Glavni-Proizvajalec(Država,Pridelek) -> Cena(Pridelek,Raste)*

- *Potek interakcije med agenti:*

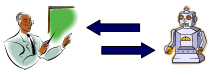
Agent3: PROPOSE(H1,0.6)
Agent2: MODIFY(H1,H2,0.5)
Agent1: MODIFY(H1,H3,0.55)
H1: 0.775
Agent1: MODIFY(H2,H3,0.45)
Agent3: CONFIRM (H2,0.6)
H2: 0.78
Agent3: CONFIRM (H3,0.6)
Agent2: CONFIRM (H3,0.5)
H3: 0.8

Agent3: WITHDRAW(H1)
Agent2: WITHDRAW(H2)
Agent1: AGREED(H3,0.8)
Agent3: ACCEPT(H3)
Agent2: ACCEPT(H3)

Primer - Različni načini učenja

Popolnoma usmerjano

Deloma nadzorovano

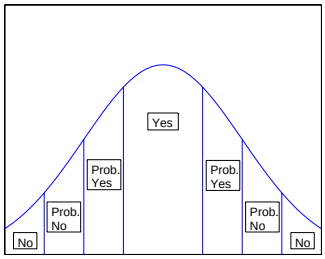


Popolnoma samostojno

- Popolnoma usmerjano — **PU** – Za pravilno interpretacijo vedno poskrbi človek.
- Deloma nadzor. (konzervativno) — **DNk** – Vpraša uporabnika vedno, ko ni popolnoma prepričan.
- Deloma nadzor. (liberalno) — **DNl** – Vpraša uporabnika samo, ko je zelo nezanesljiv.
- Pop. samostojno (konzervativno) — **PSk** – Osveži znanje samo, ko misli, da ima gotovo prav.
- Pop. samostojno (liberalno) — **PSl** – Osveži tudi, ko misli, da ima najbrž prav.

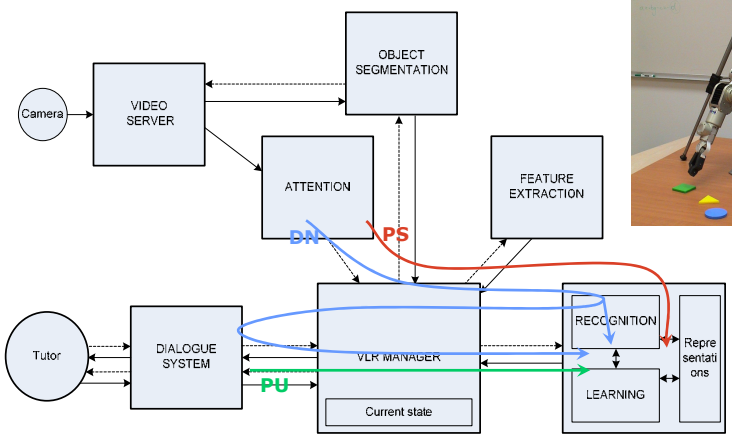
Pravila za osveževanje znanja

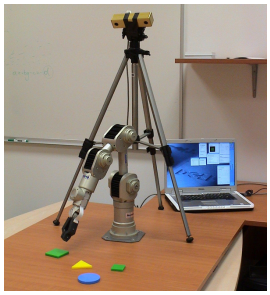
- Učni proces:
  - zaznaj predmet
  - segmentiraj predmet
  - poizkušaj razpoznati predmet
  - odloči se za naslednji korak
    - osveži
    - vprašaj
  - osveži predstavitev
- Tabela za osveževanje:



	Da	Najbrž da	Najbrž ne	Ne	Ne vem
PU	vprašaj	vprašaj	vprašaj	vprašaj	vprašaj
DNk	osveži	vprašaj	vprašaj	/	vprašaj
DNl	osveži	osveži	osveži	/	vprašaj
PSk	osveži	/	/	/	/
PSl	osveži	osveži	/	/	/

Implementiran sistem





Primer dialoga

'H: This is a red, small, and square object.'
'R: OK.'
'H: What do you see?'
'R: I don't know this object.'
'H: It is green, small, and triangular.'
'R: OK.'
'H: What do you see?'
'R: I see a small and square object.'
'H: Yes, that's correct.'
'R: OK.'
'H: It is also blue.'
'R: OK.'
'H: What do you see?'
'R: I see a green object. I think it is also large.'
'H: Yes, that's correct.'
'R: OK.'
'H: It is also rectangular.'
'R: OK.'
'R: I don't know this object. What is it like?'
'H: It is blue, small, and rectangular.'
'R: OK.'

Popolnoma usmerjano

'R: Is this object large?'
'H: Yes, that's correct.'
'R: OK.'
'H: It is also yellow and circular.'
'R: OK.'
'R: Is this object red?'
'H: No, it is not red.'
'R: OK.'
'H: It is yellow.'
'R: OK.'
'H: What do you see?'
'R: I see a red, small, and triangular object.'

Deloma nadzorovano

Eksperimenti

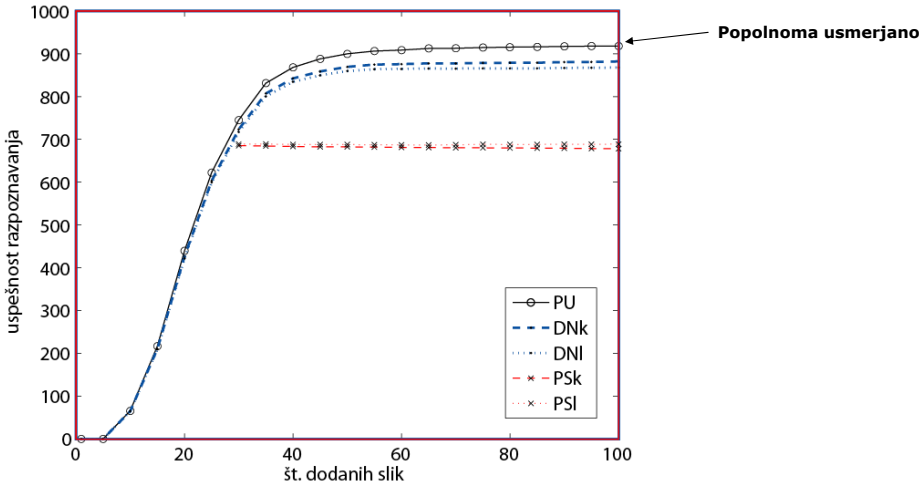
- Učenje in razpoznavanje vizualnih lastnosti
  - 6 vizualnih značilnic
    - 3 značilnice videza (Hu,Sa,In)
    - 3 značilnice oblike (Ar,Pr,Cm)
  - 10 vizualnih lastnosti
    - 4 barve (Rd,Gr,Bl,Yl)
    - 2 velikosti (Sm,Lr)
    - 4 oblike (Sq,Cr,Tr,Rc)
  - Pravilni opisi poznani – simulirani odgovori uporabnika
  - Evolucija uspešnosti razpoznavanja skozi čas

Tabela uspešnosti:

	YES	PY	PN	NO	DK
YES	1	0.5	-0.5	-1	0
NO	-1	-0.5	0.5	1	0

Eksp. rezultati – vizualne lastnosti

- Uspešnost razpoznavanja

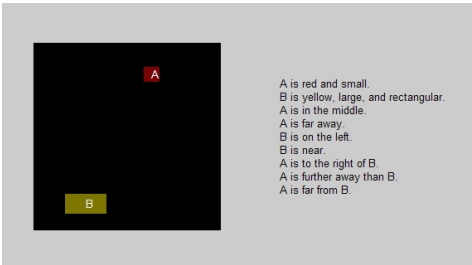


Učenje prostorskih relacij

- 2 objekta (A and B)
- 5 značilnic (x,y,dx,dy,d)
- 11 prostorskih relacij:
  - levo od (TL: to the left of: "A is to the left of B")
  - desno od (TR: to the right of: "A is to the right of B")
  - bližje kot (CT: closer than: "A is closer to me than B")
  - dlje kot (FT: further away than: "A is further away from me than B")
  - blizu do (NT: near to: "A is near to B")
  - daleč od (FF: far from: "A is far from B")
  - na levi (OL: on the left: "A is on the left")
  - na sredini (OR: on the right: "A is on the right")
  - na desni (IM: in the middle: "A is in the middle")
  - blizu (NR: near: "A is near")
  - daleč (FA: far away: "A is far away")

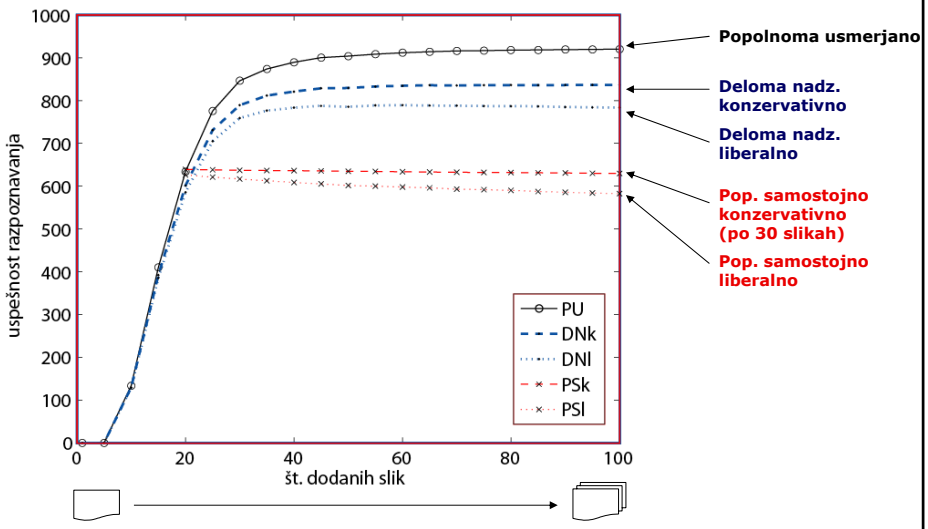
Dva primera

- Avtomatski opis prizora



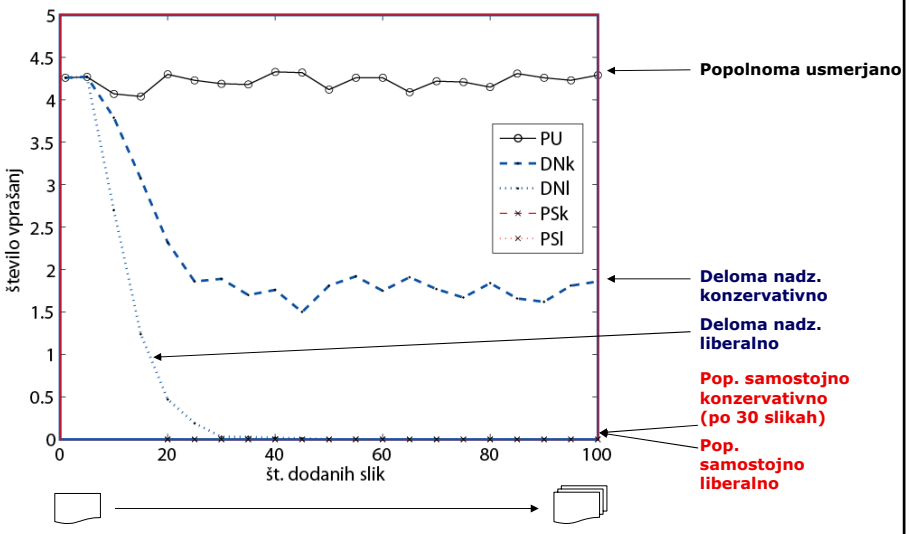
Eksp. rezultati – prostorske relacije

- Uspešnost razpoznavanja



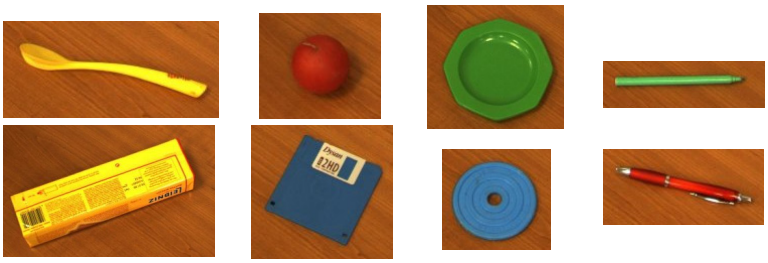
Eksp. rezultati – prostorske relacije

- Število vprašanj

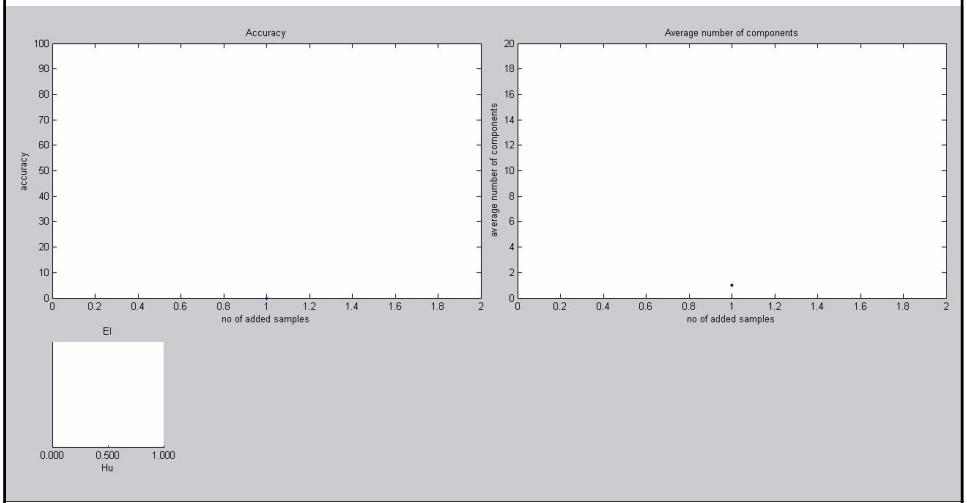


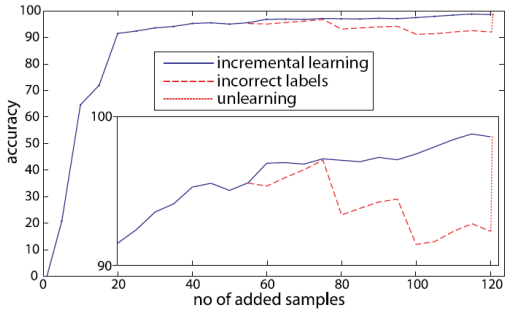
Eksperimentalni rezultati

- Učenje in razpoznavanje vizualnih lastnosti vsakodnevnih predmetov
  - 6 vizualnih značilnic
    - 3 značilnice videza (Hu,Sa,In)
    - 3 značilnice oblike (Ar,Co,Ec)
  - 6 vizualnih lastnosti
    - 4 barve (Rd,Gr,Bl,Yl)
    - 2 obliki (Cm,El)



Eksperimentalni rezultati



Eksperimentalni rezultati

- Učenje z napačnimi podatki zelo poslabša rezultate učenja
  - slaba komunikacija
  - nezanesljiv vir informacije
- Rešitev: popravljanje naučenih predstavitev (unlearning)