

# **OSNOVE UMETNE INTELIGENCE**

2018/19

*uvod*

*intelligentni agenti*

# Osnove umetne intelligence

## Izvajalci

- nosilec:  
Zoran Bosnić  
kabinet: 2. nadstropje, desno, R2.17
- asistenti:  
Jure Žabkar, 3. nadstropje, levo, R3.54 (LUI)  
Martin Jakomin, 2. nadstropje, desno, R2.37 (LKM)  
Anita Valmarska  
Jan Kralj

## Predmet

- 3 ure predavanj, 2 uri vaj
- učilnica
- forum



# Cilji predmeta

- seznanitev z osnovnimi **idejami, koncepti in metodami** umetne inteligence
- pridobiti sposobnost **reševanja problemov** z metodami umetne inteligence
- pridobiti zmožnost **razumevanja literature** s področja umetne inteligence
- razumevanje **dosežkov** umetne inteligence glede na njihove **implikacije v filozofiji in psihologiji**



# Obveznosti predmeta



## 1. domače naloge (kvizi)

- 6 nalog
- 3 poskusi za vsako nalogo
- pogoj: potrebno pozitivno ( $\geq 50\%$ ) rešiti vsaj 5 od 6 domačih nalog

## 2. seminarska naloga

- neobvezna, vredna 20% končne ocene
- teme bodo objavljene na učilnici, predlagate lahko tudi lastne
- ocena se prišteje k rezultatu izpita, če je ta pozitiven
  - prijava teme seminarske naloge: do 30. 10. 2018
  - oddaja in zagovor seminarske naloge: do 18. 1. 2019

## 3. končni izpit

- pisni in/ali ustni izpit
- vreden 80% končne ocene
- pogoj za pristop k izpitu: opravljene DN

**KONČNA OCENA** je utežena vsota izpita (80%) in seminarja (20%).  
Končna ocena je pozitivna, če je končni izpit pozitiven in utežena vsota obveznosti znaša  $\geq 50\%$ .

# Literatura

- I. Bratko, Prolog Programming for Artificial Intelligence, 4th edition, Pearson Education, Addison-Wesley 2011, ISBN: 0201403757.
- S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Third edition, Pearson Education, Prentice-Hall 2010, ISBN: 0136042597.
- I. Bratko, Prolog in umetna inteligenca, Založba FE in FRI, ponatis 2011.
- I. Kononenko, Strojno učenje, Založba FE in FRI, 2005.
- I. Kononenko, M. Robnik Šikonja, Inteligentni sistemi. Založba FE in FRI, 2010.

literatura... ?



# Vsebina predmeta

## Osnove umetne inteligence

- inteligentni agenti
- strojno učenje:
  - problemski prostor, hipoteze, ocenjevanje učenja
  - gradnja odločitvenih dreves
  - učenje iz šumnih podatkov, rezanje
  - manjkajoči atributi, regresija, naivni Bayes
- reševanje problemov kot preiskovanje grafov
  - neinformirani preiskovalni algoritmi
  - informirani preiskovalni algoritmi
  - lokalno preiskovanje
  - grafi AND/OR, nedeterministično okolje
- igranje iger
- planiranje, razporejanje opravil
- predstavitev negotovega znanja, Bayesovske mreže
- avtomatsko sklepanje



# **Uvod v umetno inteligenco**

# Kaj je umetna inteligenca?

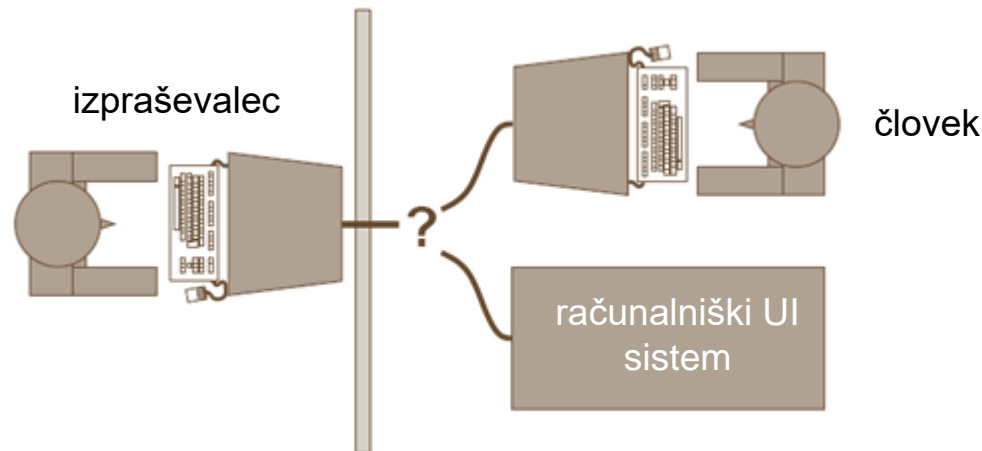


- cilj umetne intelligence: razumeti in zgraditi *intelligentne* sisteme, na osnovi razumevanja človeškega razmišljanja, sklepanja, učenja in komuniciranja (izvajanja dejanj, pojasnjevanja, komentiranja)
- različne definicije, ki izvirajo iz različnega pojmovanja človeške intelligence v in cilja, ki ga pričakujemo od računalniškega sistema
- težava: definicija *intelligence*?
  - ali je vse, kar počne človek, inteligentno (refleksi, nekonsistentnost, napake, učinkovitost)?
  - ali zasledovati modeliranje človeka ali modeliranje ideala (princip ***racionalnosti*** – optimalnosti – "*doing the right thing*")
- umetna inteligenca:
  - sistemi, ki se vedejo/razmišljajo kot človek/racionalno
  - kako je s kreativnostjo, čustvi, zavestjo?
- velika medijska pokritost ("*AI will conquer the world*" (!@#!))



# Turingov test

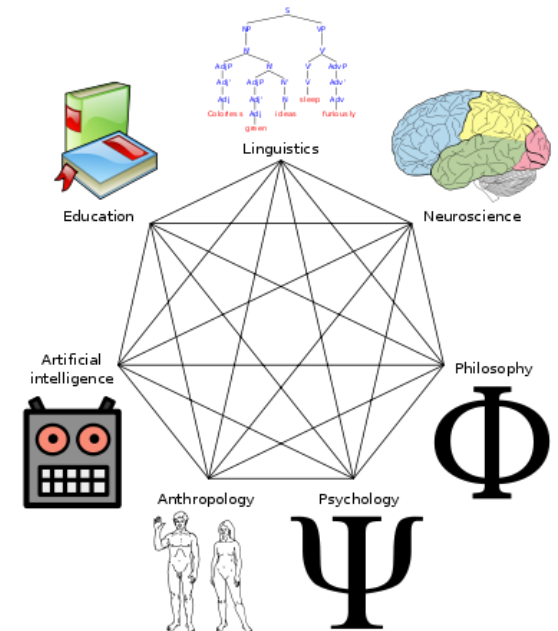
- sistem, ki se "vede kot človek"
- praktični preizkus (the imitation game), ki ga predlaga Turing (1950) za testiranje, ali je sistem dosegel stopnjo inteligence, primerljivo s človekom
- računalnik "opravi" preizkus, če človeški izpraševalec po računalnikovih odgovorih na zastavljena vprašanja ne more ugotoviti, ali odgovore podaja človek ali računalnik



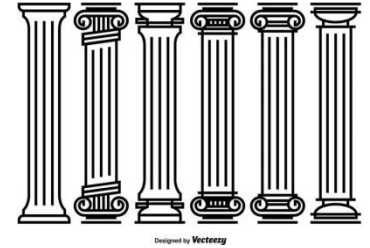
- ideja napeljuje na sposobnosti inteligentnih sistemov: obdelava naravnega jezika, predstavitev znanja, avtomatsko sklepanje, strojno učenje, računalniški vid, robotika
- problem: testa ni možno reproducirati ali podvreči matematični analizi

# Kognitivna znanost

- sistem, ki "razmišlja kot človek"
- področje, komplementarno UI
- združuje področje UI s področji psihologije, filozofije, antropologije, izobraževanja, lingvistike, nevroznanosti
- simuliranje človekovega razmišljanja na osnovi:
  - opazovanja lastnih (človeških) misli
  - psihološkimi poskusi (opazovanje vedenja oseb)
  - posnetki možganov (opazovanje med delovanjem)



# Temelji umetne inteligence



## Filozofija

- logika, metode sklepanja
- relacija med mišljenjem in fizičnimi možgani
- vir znanja, povezava z akcijami, racionalnost (UI potrebuje vse tri komponente)

## Matematika

- formalna predstavitev problemov, dokazovanje pravilnosti
- algoritmi, izračunljivost, odločljivost (decidability) in obvladljivost (tractability) problemov
- verjetnostni račun

## Psihologija

- način človeškega mišljenja in izvajanja dejanj
- prilagajanje, zaznavanje, motorična kontrola

## Jezikovna znanost

- relacija med jezikom in razmišljanjem (predstavitev znanja, slovnica)

## Nevroznanost

- preučevanje živčnega sistema (možgani), ki je osnova za mentalno aktivnost (nevroni, EEG)

## Ekonomija

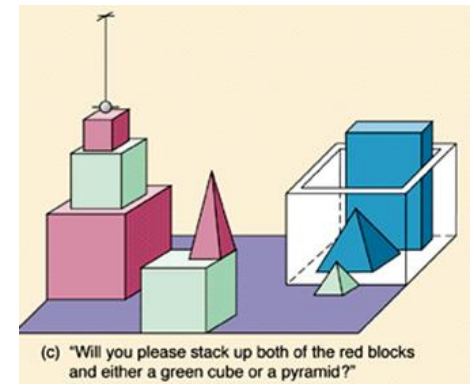
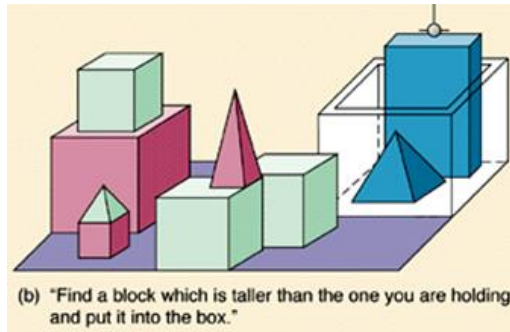
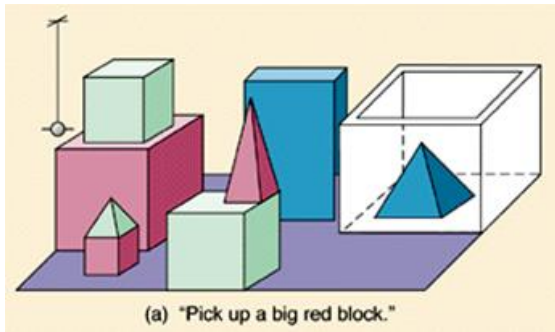
- maksimizacija izkupička, načrtovanje rezultatov po zaporedju korakov (operacijske raziskave)

## Teorija nadzora (control theory)

- avtonomni sistemi, ki nadzirajo lastno delovanje

# Zgodovina UI

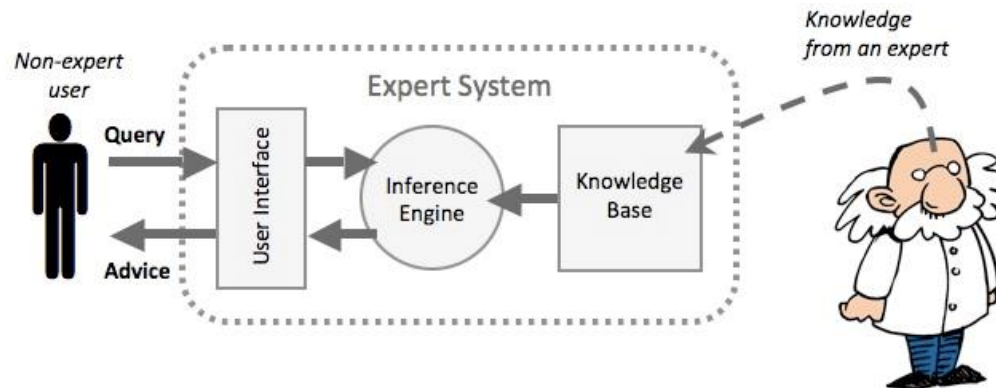
- **1943:** McCulloch & Pitts: model možganov (nevronska mreža)
- **1950:** Turingovo delo "Computing Machinery and Intelligence"
- **1952-69:**
  - General Problem Solver
  - Lisp kot prevladujoči jezik za programiranje na področju AI
  - zbirke problemov Microworlds (npr. *blocks*)



- **1966-80:**
  - problemi: računska zahtevnost, zastoj pri razvoju nevronske mreže
  - genetski algoritmi
  - sistemi znanja (angl. knowledge-based systems) – izpeljava novih ugotovitev iz obstoječih baz znanja

# Zgodovina UI (nad.)

- **1980-95:**
  - pojav ekspertnih sistemov (medicina)
  - 1987: UI uradno postane "znanost" z uporabo uveljavljenih znanstvenih pristopov in metod
  - zametki odkrivanja zakonitosti iz podatkov (angl. *data mining*): osredotočenost na vzorce v podatkih in ne primarno na algoritme



- **1995-**
  - porast števila avtomatski agentov (Internet, semantični splet, iskalniki, priporočilni sistemi, povzetki spletnih strani)

# Sodobne aplikacije UI

- avtonomna (robotska) vozila
- vožnja robota po kompleksni poti
- robotska opravila (pospravljanje pomivalnega stroja?)
- razpoznavanje govora (klicni centri, pametni telefoni, ...)
- avtonomno načrtovanje (avtomatski piloti, NASA)
- igranje iger (1997: šahovska igra med Deep Blue in Kasparovom, namizni tenis, karte)
- načrtovanje logistike (optimizacija porabe sredstev pri operacijah)
- detekcija neželene elektronske pošte
- robotika (Roomba, raziskovanje vesolje, odstranjevanje min)
- strojno prevajanje
- nakupovanje živil za naslednji teden
- iskanje in izpeljava novih matematičnih dokazov
- podajanje ustreznih nasvetov na pravnem področju
- prevajanje med dvema jezikoma v realnem času
- izvajanje pogovora z osebo
- izvedba kompleksne kirurške operacije
- analiza in sinteza jezika
- ...

# Sodobne aplikacije UI

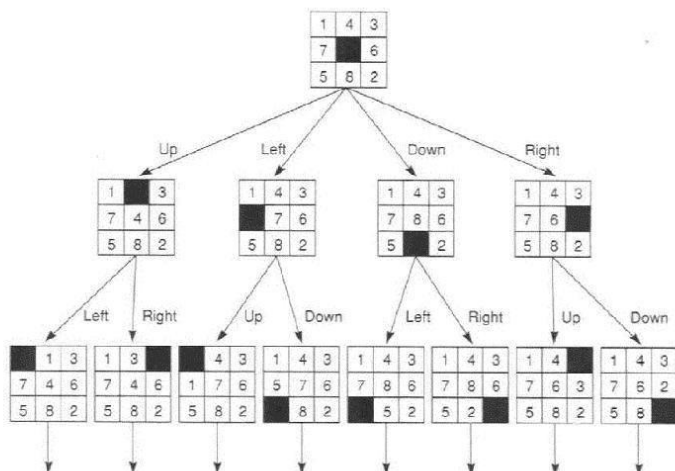


*Henry Squirrel was thirsty. He walked over to the river bank where his good friend Bill Bird was sitting. Henry slipped and fell in the river. Gravity drowned. The End.*

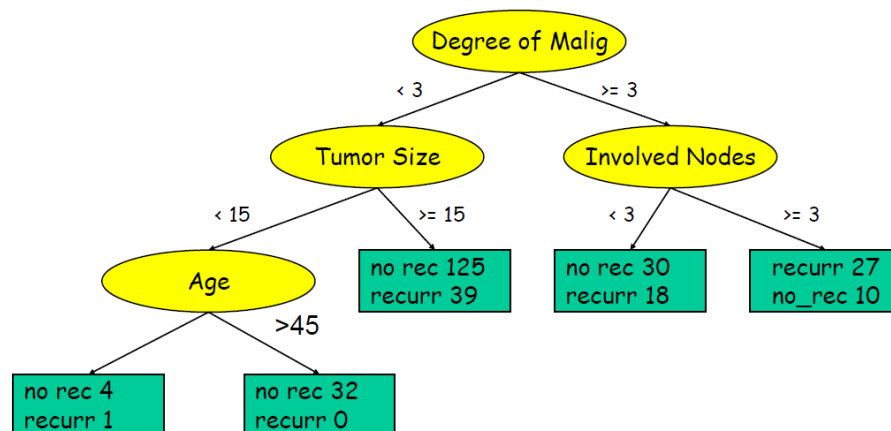
*Once upon a time there was a dishonest fox and a vain crow. One day the crow was sitting in his tree, holding a piece of cheese in his mouth. He noticed that he was holding the piece of cheese. He became hungry, and swallowed the cheese. The fox walked over to the crow. The End.*

# Področja UI

- reševanje problemov s preiskovanjem grafov
- planiranje po principu sredstev in ciljev (*means-end planning*)
- predstavitev znanja, ekspertni sistemi, tehnologija znanja, ontologije
- avtomatsko sklepanje (*reasoning, inference*)
- strojno učenje (*machine learning*) in odkrivanje zakonitosti v podatkih (*podatkovno rudarjenje, data mining, knowledge discovery in databases*)



preiskovanja prostora stanj igre s ploščicami

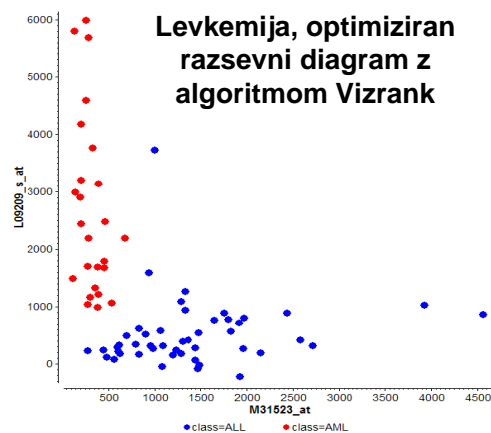
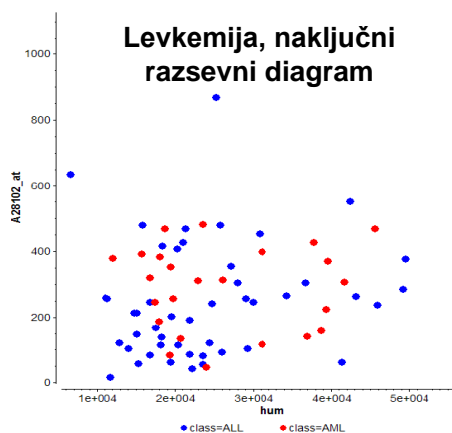


strojno učenje – odločitveno drevo za napovedovanje ponovitve raka na dojki



# Področja UI (nad.)

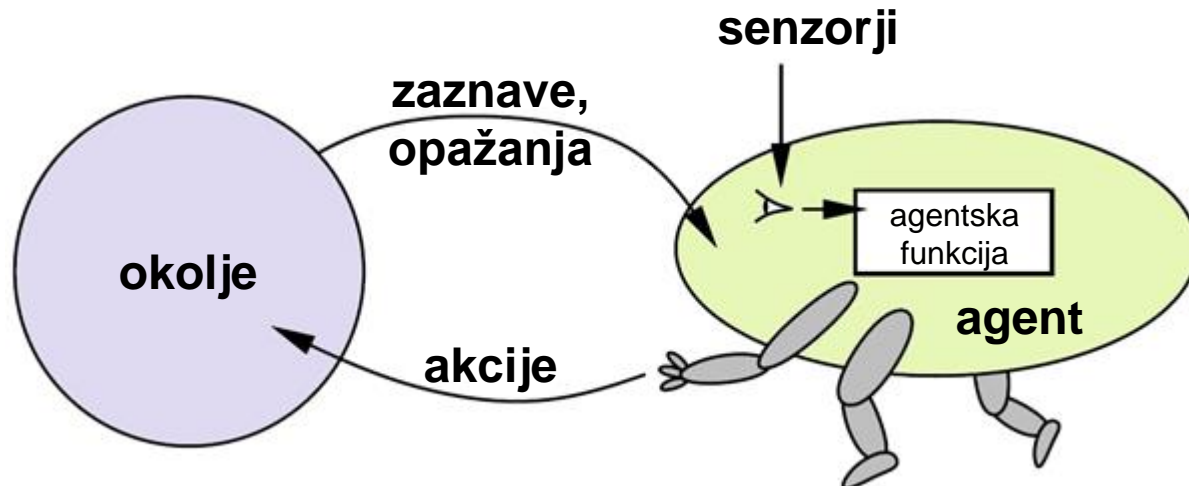
- obdelava naravnega jezika (*natural language processing*)
- robotika
- evolucijsko računanje (genetski algoritmi/programiranje)
- računalniški vid, zaznavanje (*computer vision*)



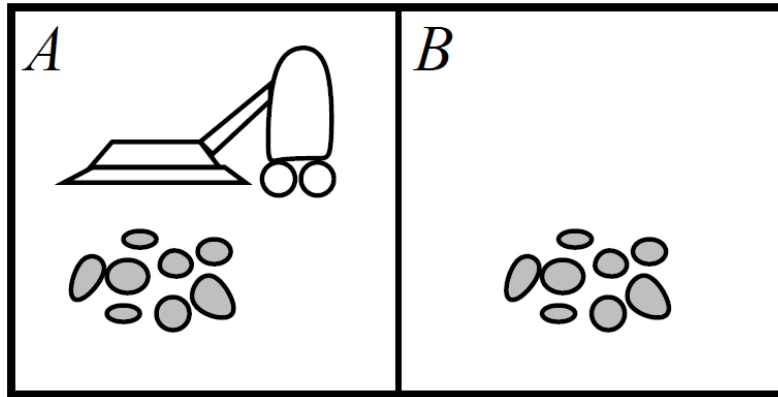
# **Intelligentni agenti**

# Inteligentni agent

- **inteligentni agent** je sistem, ki implementira *inteligenco* po principu *racionalnosti* (*optimalnosti*) – v vsaki situaciji izbere najboljšo možno akcijo
- **agenti** zaznavajo **okolje** skozi **senzorje** in na podlagi **zaznav** izvajajo **akcije** (primeri: robotska kamera, avtonomno vozilo, delniški program, termosta, človek?)
- **agentska funkcija** definira akcijo glede na zgodovino zaznav

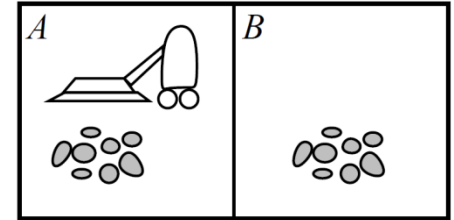


# Primer agenta



- **zaznave:** lokacija (*A* ali *B*) in čistoča (*čisto* ali *umazano*)
- **akcije:** *levo*, *desno*, *sesaj*, *čakaj*
- Kaj je idealna agentska funkcija?
- Kako predstaviti agentsko funkcijo? Ali je možno implementirati s kratkim programom?
- Kako oceniti uspešnost delovanja našega agenta?

# Primer agenta



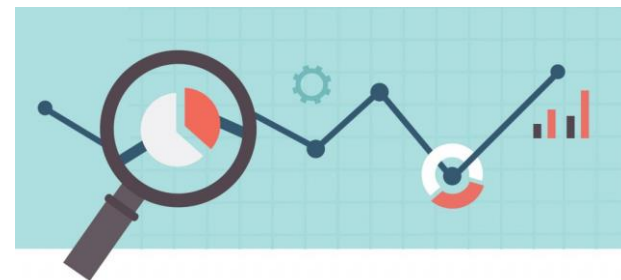
- možna definicija agentske funkcije:

Zgodovina zaznav	Akcija
[A, čisto]	desno
[A, umazano]	sesaj
[B, čisto]	levo
[B, umazano]	sesaj
[A, čisto], [A, čisto]	desno
[A, čisto], [A, umazano]	sesaj
...	...

- možna implementacija agentskega programa:

```
function SESALEC_AGENT([lokacija, čistoča]) returns akcija
  if čistoča = umazano then return sesaj
  else if lokacija = A then return desno
  else if lokacija = B then return levo
```

# Mera uspešnosti



- agent spreminja okolje z zaporedjem svojih akcij
- **mera uspešnosti** oceni način **sprememb okolja** glede na akcije agente
  - zakaj ne "sprememb stanja agenta" namesto sprememb okolja?
- primer mer za sesalec:
  - količina posesanega prahu? (zakaj ne?)
  - čas, da sta oba prostora čista?
  - poraba elektrike, povzročen hrup?
- racionalen agent **maksimira pričakovano mero uspešnosti** na podlagi zaporedja zaznav in vgrajenega znanja
  - *racionalen* ne pomeni *optimalen* ali *vseveden* (zaznave lahko ne podajo vseh informacij, izidi so lahko nepričakovani)
  - okolje je popolnoma znano le v redkih primerih (primer hrošča rogača in talne ose)
- racionalen agent zato: **raziskuje**, se **uči**, je **avtonomen**
  - primeri?

# Opredelitev naloge agenta

Za vsakega agenta je potrebno določiti:

- mero uspešnosti (Performance measure)
- pričakovano okolje (Environment)
- možne akcije (Actuators)
- senzorje. (Sensors)



Primer 1: agent za avtomatsko internetno nakupovanje:

- mera uspešnosti: cena, kvaliteta, primernost, učinkovitost
- okolje: trenutna in bodoča spletna mesta, proizvajalci, dobavitelji
- akcije: prikaz uporabniku, oddaja zahtevkov HTTP, izpolnjevanje obrazcev
- senzorji: strani HTML (besedila, slike, animacije, skripte)



Primer 2: avtonomno vozilo:

- mera uspešnosti: varnost, cilj, zaslužek, legalnost, udobje
- okolje: cesta, promet, pešci, vreme
- akcije: navigacija, pospeševanje, zaviranje, hupanje, displej/govor
- senzorji: zajem slike, pospeškometri, senzorji v motorju, GPS, tipkovnica



# Vrste okolij

Okolja se ločijo glede na:

1. **s polno informacijo** (observability): senzorji podajajo vse informacije, ki so potrebne za odločanje (npr. sesalec bi lahko imel senzor za stanje čistoče v obeh sobah)
2. **eno/več-agentnost** (single/multi agent): glede na to, ali obravnavamo druge entitete v okolju tudi za agente in optimiziramo mero uspešnosti glede na njihovo pričakovano obnašanje (npr. v šahu maksimiziramo svojo uspešnost tudi skozi minimizacijo uspešnosti nasprotnega agenta)
  - razlika: kompetitivna in kooperativna okolja (npr. avtonomna vozila)
3. **determinističnost/stohastičnost** glede na to, ali je naslednje stanje v okolju popolnoma odvisno le od trenutnega stanja in akcije agenta (npr. vožnja je lahko stohastična zaradi nepredvidenih okvar avtomobila):
  - kakšna je razlika med netransparentnim determinističnim in stohastičnim okoljem?





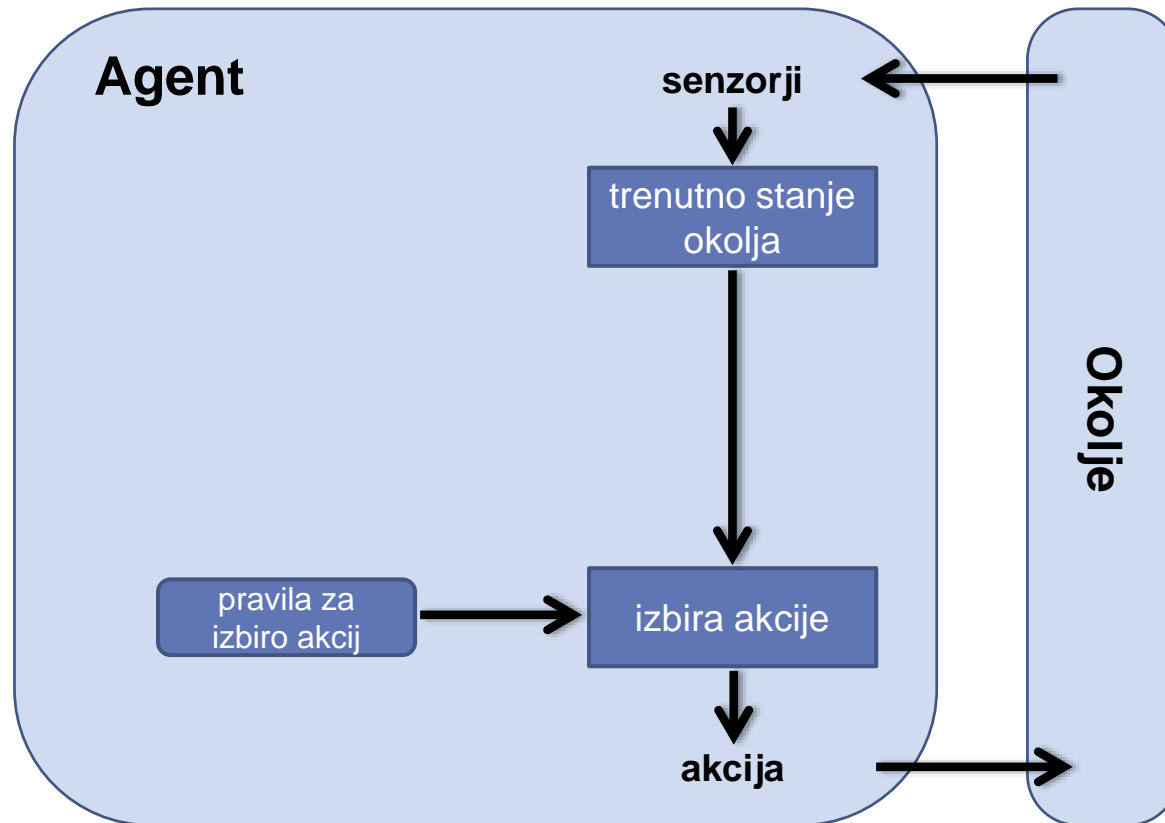
# Vrste okolij (nad.)

4. **epizodičnost/zaporednost**: glede na to, ali je naslednje stanje okolja odvisno samo od akcije, izvedene v prejšnjem okolju (npr. epizodično: zaznavanje okvarjenih izdelkov; zaporedno: vožnja z avtonomnim vozilom)
5. **statičnost/dinamičnost**: glede na to, ali se okolje lahko spreminja medtem, ko se agent odloča (npr. statično: reševanje križank; dinamično: vožnja avtonomnega vozila)
6. **diskretna/zvezna**: glede na to, kako agent obravnava čas (kot diskretne ali zvezne korake (npr. diskretno: igra šaha, zvezno: vožnja)

	Pasijansa	Človek ne jezi se	Internetno nakupovanje	Avtonomno vozilo
transparentno?	DA	DA	NE	NE
eno-agentno?	DA	NE	DA ?	NE
deterministično?	DA	NE	DA ?	NE
epizodično?	NE	NE	NE	NE
statično?	DA	DA	NE	NE
diskretno?	DA	DA	DA	NE

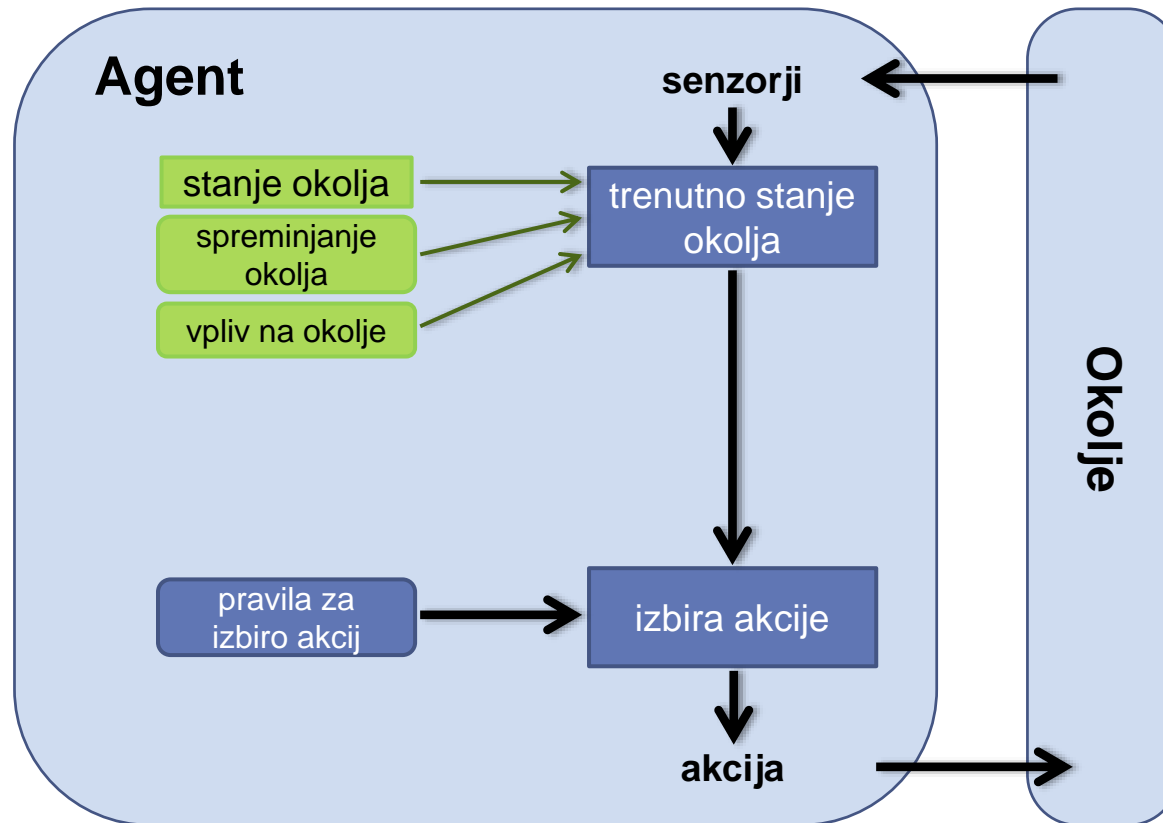
# Vrste agentov

1. **refleksni agent:** izbira akcijo glede na trenutne zaznave (zgodovina ni pomembna)
  - primer: avtomatski sesalec iz prejšnjega primera



# Vrste agentov (nad.)

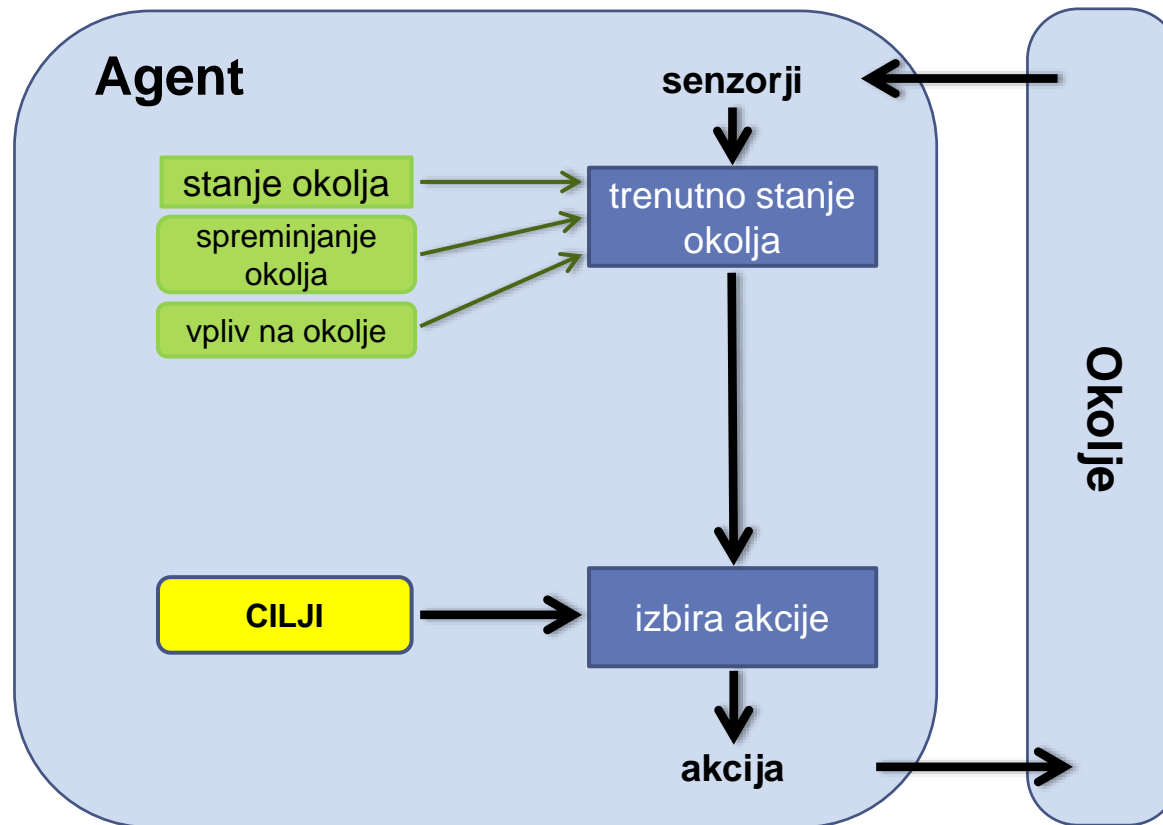
2. **refleksni agent z modelom okolja:** izbira akcijo glede na trenutne zaznave, pri čemer upošteva potekajoče spremembe v okolju in svoje vplive na okolje
- primer: avtomatsko vozilo, zaviranje vozila pred njim lahko pomeni zastoje



# Vrste agentov (nad.)

## 3. ciljno-usmerjen agent: bolj prilagodljiv, prilagaja svoje akcije glede na doseganje cilja

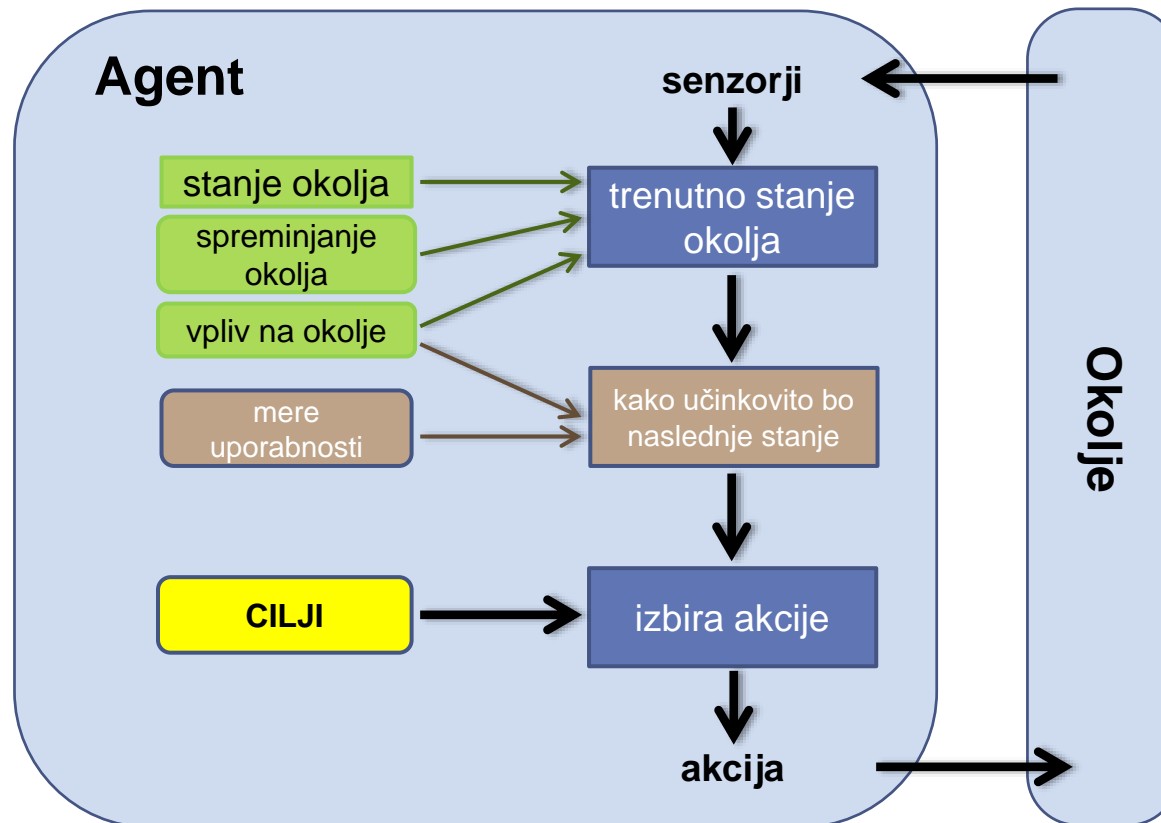
- primer: avtomatsko vozilo, cilj je končna destinacija,
- UI: področje preiskovanja in planiranja



# Vrste agentov (nad.)

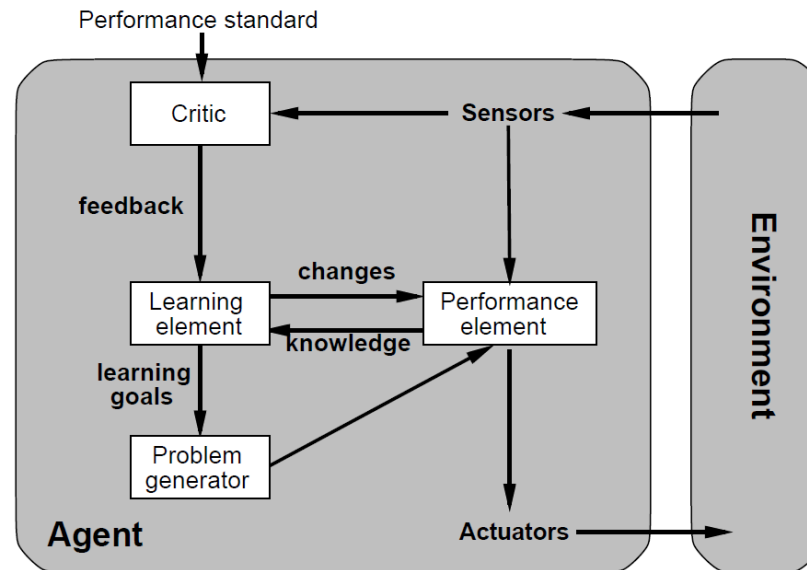
## 4. uporabno-usmerjen agent: iščejo tak način realizacije cilja, ki je tudi najbolj učinkovit (uporaben)

- primer: avtomatsko vozilo, cilj je končna destinacija,
- UI: področje preiskovanja in planiranja



# Samodejno učenje

- Vse agente lahko dopolnimo tako, da se samodejno učijo in izboljšujejo





**Strojno učenje**