

## Un'introduzione all'intelligenza artificiale

Cristina Baroglio

1

- Docente:
  - Cristina Baroglio
  - Ufficio 4
  - 011-6706703
  - [baroglio@di.unito.it](mailto:baroglio@di.unito.it)
  - [www.di.unito.it/~baroglio](http://www.di.unito.it/~baroglio)
- Corso:
  - 60 ore
  - Materiale delle lezioni:
    - <http://informatica.i-learn.unito.it>
  - Dopo le lezioni pubblico una falsariga di quanto spiegato insieme alla registrazione della lezione
  - Esame: orale
  - Libro di testo: *Intelligenza Artificiale, un approccio moderno* (S. Russell, P. Norvig) Prentice Hall
  - <http://aima.cs.berkeley.edu/>

Cristina Baroglio

2

## Sfatiamo qualche leggenda

## Al produce robot che si dividono in 2 categorie



- Robot cattivi e pericolosi



STAR WARS

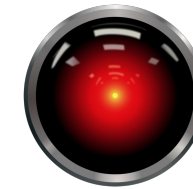
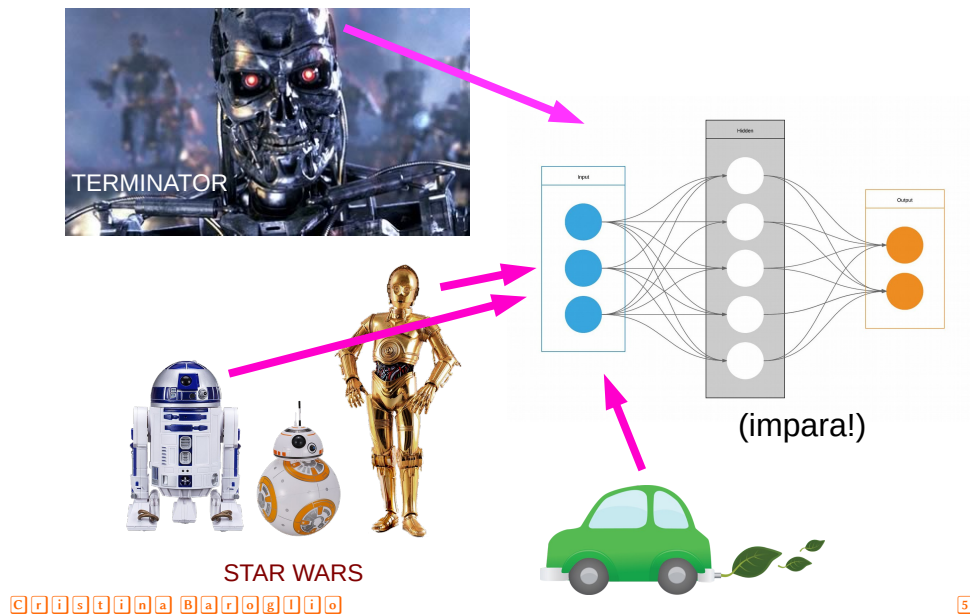
- Robot buoni, innocui e pure un po' buffi

Cristina Baroglio

3

Cristina Baroglio

4



- **Percepisce**
- **Controlla**
- **Ha propositi**
- **È più intelligente degli esseri umani**
- **È incontrollabile**
- **È pericolosa**

Cristina Baroglio

6

## Leggende!

- AI è robotica sono discipline distinte, con qualche intersezione
- I robot non sono più intelligenti di una lavatrice
- Siamo ad anni luce di distanza dalla realizzazione di agenti senzienti
- Il software non è né buono né cattivo (ma può essere scritto male)
- Le reti neurali apprendono ma solo la funzione obiettivo (e pure in maniera approssimativa)

Cristina Baroglio

7

## Leggende!

- AI è robotica sono discipline distinte, con qualche intersezione
- I robot non sono più intelligenti di una lavatrice
- Siamo ad anni luce di distanza dalla realizzazione di agenti senzienti
- Il software non è né buono né cattivo (ma può essere scritto male)
- Le reti neurali apprendono ma solo la funzione obiettivo (e pure in maniera approssimativa)
- ... eppure al TG hanno detto che due chatbot hanno inventato un linguaggio nuovo incomprensibile ai loro programmatori e sono state spente perché non si sapeva cosa stessero dicendo ...

Cristina Baroglio

8

- **Data e luogo di nascita dell' AI:**
  - Dartmouth Conference (USA), 1956
  - Nome scelto da John McCarthy
- **In precedenza (primi anni 50):**
  - Una macchina può “pensare” ed essere considerata intelligente?
  - Vari approcci: cybernetics, automata theory, complex information processing
  - Turing Test
- **Primo anni 60:**
  - Primi tentativi (scacchi, giochi, dimostrazioni automatiche)

Cristina Baroglio

9



Computer nel 1956

John McCarthy

Cristina Baroglio

10

## Più precisamente

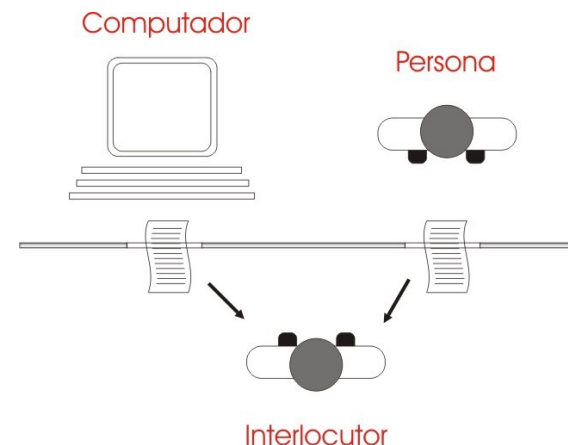
## Turing Test: una macchina è intelligente?

- I computer erano un po' diversi da come li conosciamo:
  - Niente tastiere, dischi magnetici, monitor
  - Al MIT: primi esperimenti con computer ad input diretto da tastiera (e non da schede perforate)
  - All' IBM: creazione dei precursori dei supporti di memoria a disco magnetico (invece dei nastri)

<http://www.computerhistory.org/timeline/1956/>

Cristina Baroglio

11



Interlocutor non vede chi c'è dall'altra parte del muro

Interazione:

- Interlocutor scrive delle domande
- Chi è dall'altra parte fornisce risposte scritte
- L'interlocutore capirà se si tratta di una persona o di un computer?

(immagine di Juan Alberto Sánchez Margallo)

Cristina Baroglio

12

Q: Please write me a sonnet on the subject of the Forth Bridge.

A: Count me out on this one. I never could write poetry.

Q: Add 34957 to 70764.

A: (Pause about 30 seconds and then give as answer) 105621.

Q: Do you play chess?

A: Yes.

XYZ ?



ABC !

XYZ ?



ABC !

Basta produrre gli *output attesi* per dire che vi è comprensione?

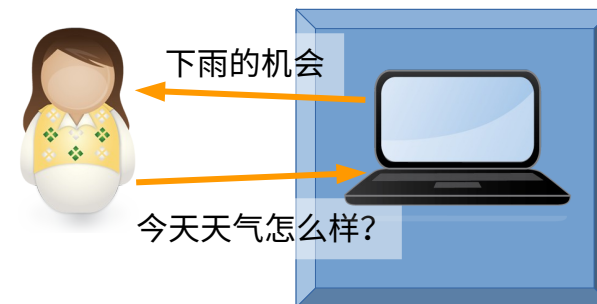
## Test di Turing: intelligenza e comprensione

Il fatto che di fronte agli stessi input un uomo e un computer producano gli stessi output significa che entrambi capiscono cosa stanno facendo?

Searle, John. R. (1980) *Minds, brains, and programs*. *Behavioral and Brain Sciences* 3 (3): 417-457

## Esperimento della stanza cinese (John Searle)

L'investigatore interagisce con un computer, programmato per rispondere con certi ideogrammi cinesi ad altri ideogrammi cinesi ricevuti in input. Il programma è molto ben scritto. La persona non vede il computer che è chiuso in una stanza

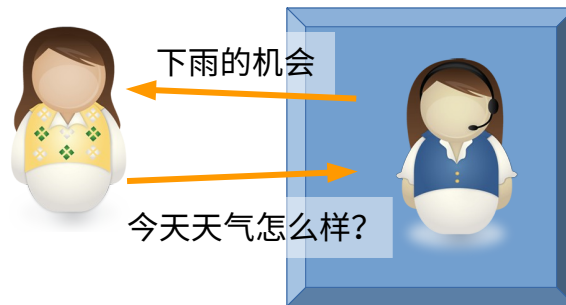


Parla cinese?

Lo capisce?

Una **persona** chiusa in una stanza ha istruzioni per rispondere con certi ideogrammi cinesi in risposta ad altri ideogrammi cinesi

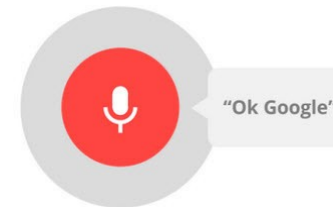
Parla cinese?  
Lo capisce?



*“Instantiating a computer program is never by itself a sufficient condition of intentionality”*

Cristina Baroglio

17



Impressionante

Capisce quel che dite ... *oppure no?*

Cristina Baroglio

18

## Strong e weak AI

### 1) Strong AI:

è possibile riprodurre l' intelligenza umana?

- Studio del pensiero e del comportamento umano (scienze cognitive)

### 2) Weak AI:

è possibile trovare dei modi per risolvere problemi che, se risolti dagli esseri umani richiederebbero intelligenza?

- task-oriented, studio del pensiero e del comportamento razionale

Cristina Baroglio

19

## Quanto è difficile attraversare una strada?



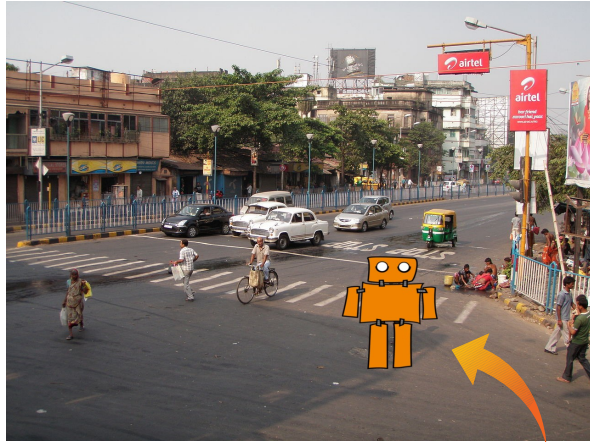
Un uomo  
attraversa la  
strada

Cristina Baroglio

20



## Quanto è difficile attraversare una strada?

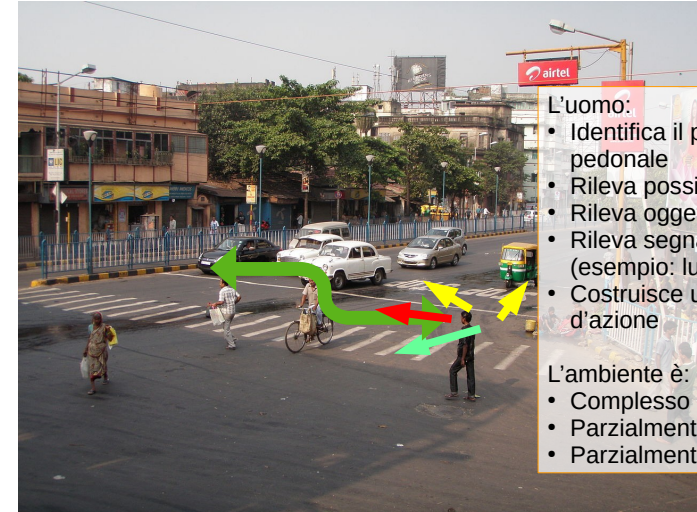


Un agente  
artificiale  
attraversa la  
strada

Cristina Baroglio

21

## Quanto è difficile attraversare una strada?



L'uomo:

- Identifica il passaggio pedonale
- Rileva possibili ostacoli
- Rileva oggetti in movimento
- Rileva segnali significativi (esempio: luce semaforica)
- Costruisce un piano d'azione

L'ambiente è:

- Complesso
- Parzialmente prevedibile
- Parzialmente collaborativo

Cristina Baroglio

22

?

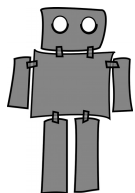
## Emergono le prime astrazioni

Come programmo un agente artificiale in grado di:

- Identificare un passaggio pedonale
- Rilevare possibili ostacoli
- Rilevare oggetti in movimento
- Rilevare segnali significativi (esempio: luce semaforica)
- Costruire un piano d'azione

In un ambiente che è:

- Complesso
- Parzialmente prevedibile
- Parzialmente collaborativo



Cristina Baroglio

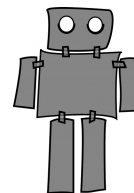
23

Come programmo un **agente artificiale** in grado di:

- Identificare un passaggio pedonale
- Rilevare possibili ostacoli
- Rilevare oggetti in movimento
- Rilevare segnali significativi (esempio: luce semaforica)
- Costruire un piano d'azione

In un **ambiente** che è:

- Complesso
- Parzialmente prevedibile
- Parzialmente collaborativo

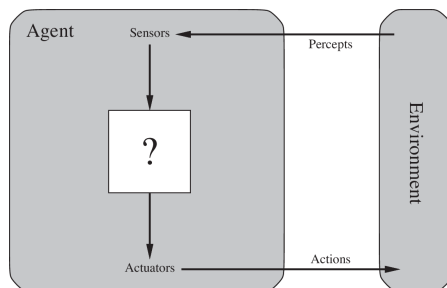


**Binomio <agente, ambiente>**

Cristina Baroglio

24

- Agente:** astrazione che rappresenta un qualsiasi sistema che percepisce il proprio ambiente tramite dei sensori ed agisce su di esso tramite degli attuatori



Non esistono agenti che non siano situati in un ambiente

Agente e ambiente costituiscono un binomio inscindibile

Il '?' rappresenta la funzione deliberativa esercitata dall'agente, che permette di determinare l'azione da eseguire

## Completamente osservabile

in ogni istante i sensori danno accesso a tutti gli aspetti dell'ambiente rilevanti per la scelta dell'azione

## Parzialmente osservabile

i sensori danno accesso solo a parte dell'informazione rilevante (cause: sensori imprecisi oppure non in grado di rilevare alcuni dati)

## Deterministico

lo stato successivo è determinato dallo stato corrente e dall'azione applicata

## Stocastico

applicando più volte una stessa azione in uno stesso stato si possono raggiungere stati diversi. Si dice **strategico** quando è stocastico solo per quanto riguarda le azioni degli altri agenti.

## Episodico

l'esperienza degli agenti è divisa in episodi atomici: un episodio è dato da una percezione seguita da una singola azione (esempio: classificazione)

## Sequenziale

Attività composta da più passi ognuno dei quali in generale influenzerà i successivi



## Statico

l'ambiente non cambia mentre l'agente "pensa" (cioè mentre decide quale azione eseguire)

## Dinamico

l'ambiente può cambiare mentre l'agente "pensa"

## Discreto

possono essere discreti stato, tempo, percezioni, azioni (esempio: gli scacchi hanno stati, percezioni, azioni discreti)

## Continuo

possono essere continui stato, tempo, percezioni, azioni (esempio: gli scacchi hanno tempo continuo)

## Singolo Agente

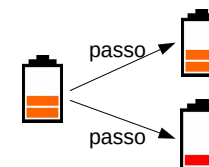
Viene modellata come agente una sola entità

## Multiagente

Vengono modellate come agenti più entità

## Parzialmente osservabile / stocastico:

vi è una connessione. Spesso viene visto come stocastico un ambiente che è parzialmente osservabile perché non si ha la percezione di quegli aspetti che renderebbero deterministico il mondo. Esempio: batteria con carica bassa, media alta, si consuma in modo continuo ma l' agente percepisce solo i tre stati citati. La stessa azione passo eseguita con carica media potrebbe portare sia a un carica bassa sia a rimanere nello stato carica media.



## Singolo agente / multiagente:

talvolta il programmatore deve decidere se modellare alcune entità del mondo come agenti o come elementi dell' ambiente. Quali devono essere rappresentate come agenti? Quelle il cui comportamento tenta di massimizzare una misura di prestazione che dipende anche dal comportamento di altri agenti.

## Caso più complesso:

ambiente parzialmente osservabile, stocastico, sequenziale, dinamico, continuo e multiagente.

## Esempio di ambiente: parole crociate

- Osservabilità: totale
- Agenti: singolo
- Deterministico: deterministico
- Episodico: sequenziale
- Statico: statico
- Discreto: discreto

Cristina Baroglio

29

## Esempio di ambiente: guidare un taxi

- Osservabilità: parziale
- Agenti: multiagente
- Deterministico: stocastico
- Episodico: sequenziale
- Statico: dinamico
- Discreto: continuo

Cristina Baroglio

30

## Esempio di agente: termostato



Cristina Baroglio

31

## Esempio di agente: roomba & simili



Cristina Baroglio

32



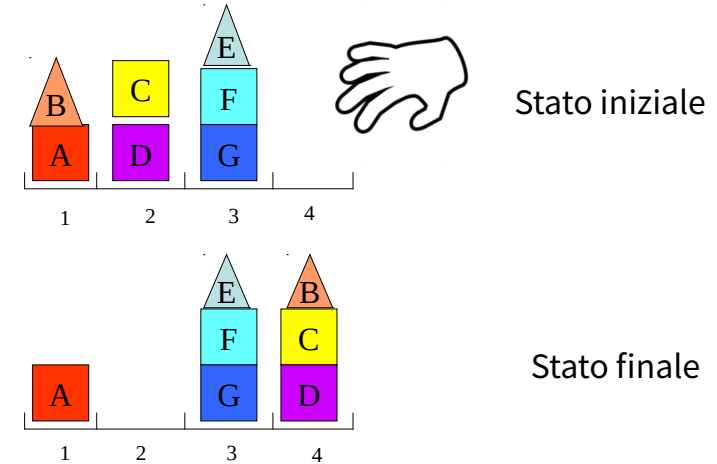
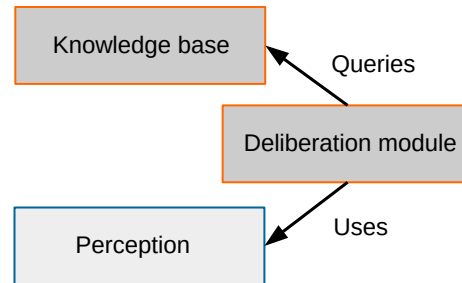
## Non-AI software

- Risolve un singolo compito
- Tipicamente strutturato come sequenza di passi (**how**)

```
list ins_ord(list L, int dato) {
    if (!L) return(crea_nodo(dato));
    else if (L->d < dato) {
        L->next =
            ins_ord(L->next, dato);
        return L;
    }
    else {
        list tmp = crea_nodo(dato);
        tmp->next = L;
        return(tmp);
    }
}
```

## AI software

- Separa una descrizione dichiarativa (**what**) da un programma generale
- Lo stesso programma è applicato a diverse descrizioni per risolvere problemi diversi



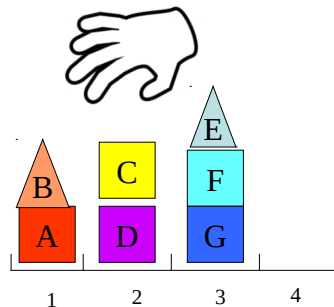
AGENTE AUTONOMO NEL MONDO DEI BLOCCHI

## Soluzione trovata dall'agente

Prendi(C), Metti(C,4), Prendi(B), Impila(B,D), Prendi(C), Impila(C,A), Prendi(B), Impila(B,C), Prendi(D), Metti(D,4), Prendi(B), Metti(B,2), Prendi(C), Impila(C,D), Prendi(B), Impila(B,C)

Localmente molte mosse possibili, nello stato iniziale:

- 1) Prendi(B),
- 2) Prendi(C),
- 3) Prendi(E)



## Soluzione trovata dall'agente

Prendi(C), Metti(C,4), Prendi(B), Impila(B,D), Prendi(C), Impila(C,A), Prendi(B), Impila(B,C), Prendi(D), Metti(D,4), Prendi(B), Metti(B,2), Prendi(C), Impila(C,D), Prendi(B), Impila(B,C)

Localmente **molte mosse possibili**, nello stato iniziale:

- 1) Prendi(B),
- 2) Prendi(C),
- 3) Prendi(E)

Solo **qualcuna utile** per raggiungere stato finale, **come scegliere?**

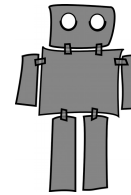
Tra molte soluzioni, **qual è la migliore?**



Da toy problem a real-world problem

Cristina Baroglio

37



Come programmo un agente artificiale in grado di:

- Identificare un passaggio pedonale
- Rilevare possibili ostacoli
- Rilevare oggetti in movimento
- Rilevare segnali significativi
- Costruire un piano d'azione

**PERCEPIRE**

In un ambiente:

- Complesso
- Parzialmente prevedibile
- Parzialmente collaborativo

**COSA?**

**COME?**

**DATI O INFORMAZIONI?**

Cristina Baroglio

38

## Un esempio: identificare un passaggio pedonale



Prospettiva  
Colore  
Qualità dell'immagine  
Estrazione dal contesto  
Modi alternativi di rappresentare la stessa cosa

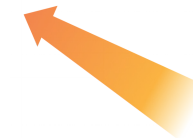
Cristina Baroglio

39

## Un esempio: identificare un passaggio pedonale



Supponiamo di avere catturato questa immagine con una telecamera. Noi vediamo delle strisce bianche sul manto stradale, ma cosa "vede" la telecamera?



Pixel, codifiche numeriche di colori con i quali l'immagine è approssimata

Cristina Baroglio

40



“vedere” non è incamerare dati ma elaborare quei dati in informazioni secondo modelli che noi umani (e molti animali) sappiamo costruire autonomamente

La telecamera cattura approssimazioni digitali **che noi interpretiamo** come immagini



**NB: “Vediamo” le strisce solo quando ci interessa attraversare**

Cristina Baroglio

41

- **Automazione del calcolo** (da metà anni 50, pochi dati, molti calcoli)
- **Automazione di procedure amministrative e contabili** (da metà anni 60, pochi calcoli, grandi moli di dati alfanumerici)
- **Automazione di fabbrica** (da metà anni 70, primi robot industriali, ambiente predeterminato, real-time)
- **Automazione di ufficio** (da metà anni 80, primi PC, primi strumenti per utenti non esperti)
- **Automazione della ricerca delle informazioni** (da fine anni 90, Internet, WEB, motori di ricerca)

Cristina Baroglio

42

## Da automazione ad autonomia

- **Automazione:** ormai standard in molte attività diverse
  - Si deve programmare il computer (il device) a fare ogni passo
  - Applicabile in domini fortemente ripetitivi (robot di saldatura in ambiente industriale)
- **Autonomia:** dare ad un agente artificiale un compito ad alto livello e deve lui risolverlo

Cristina Baroglio

43

## Agente autonomo

- Deve avere capacità di azione
- Se è autonomo:
  - Riceve solo compiti ad alto livello
  - Ragiona ed esplora alternative (molte mosse possibili ad ogni istante, numero esponenziale di possibilità da esplorare)
  - Riconosce quando non si può andare avanti su una strada
  - Riconosce che si è già stati in quella situazione
  - Semplificando, prima ragiona e poi agisce (esegue il programma che si è inventato)

Cristina Baroglio

44

### Gli agenti autonomi come percepiti al di fuori della comunità di AI:

a **self-conscious, uncontrollable** entity whose *autonomy emerges as a property "extra-program"*.

An agent can take initiatives that were not previously encoded in its process



### Gli agenti autonomi come percepiti al di fuori della comunità di AI:

a **self-conscious, uncontrollable** entity whose *autonomy emerges as a property "extra-program"*.

An agent can take initiatives that were not previously encoded in its process

In AI, gli agenti autonomi sono un modo di concepire i programmi, in cui controllo e logica (o modello) sono chiaramente separati

Un agente fa **sempre** ciò che è programmato a fare



## Da automazione ad autonomia



Metro, rover, auto senza guidatore sono esempi di agenti autonomi che eseguono **compiti simili** in **ambienti differenti**, richiedendo così diversi gradi di autonomia

**Cogliamo un'intuizione!**



## Metropolitana automatica



- Ambiente "semplice"
- Azione successiva computata in modo deterministico

### Ambiente:

una rotaia, un punto di partenza, uno di arrivo

### Sensori:

velocità, accelerazione, posizione, porte

### Azioni:

Accelera (decelera)

### Agent loop:

- 1) Leggi i sensori
- 2) Applica la funzione di controllo a letture + posizione della prossima fermata e calcola la prossima accelerazione (decelerazione)
- 3) Applica l'azione calcolata

## Rover su Marte: parzialmente autonomo



- Ambiente ancora “semplice”: no altri agenti, no fenomeni atmosferici strani
- Nozione di “goal programmabile”: operatore umano assegna goal (destinazione)
- **radio signal delay avg ~14 minute** ⇒ al rover è necessaria autonomia per condurre l'esplorazione

### Ambiente:

una piana (più o meno), qualche roccia intorno

### Sensori:

distanza, contatto, telecamera, velocità, accelerazione, posizione, ...

### Azioni:

Accelera (decelera), gira

### Agent loop:

- 1) Leggi i sensori
- 2) Applica la funzione di controllo a letture e calcola la prossima accelerazione (decelerazione)
- 3) Applica l'azione calcolata

Cristina Baroglio

49

## Auto senza guidatore



- Ambiente complesso: molti altri agenti, imprevedibile
- goal (destinazione) fornito dall'operatore

### Ambiente:

strada affollata di auto, pedoni, cani, ...  
condizioni atmosferiche (pioggia), giorno/notte  
...

### Sensori:

distanza, contatto, telecamera, velocità, accelerazione, posizione, ...

### Azioni:

Accelera (decelera), gira

### Agent loop:

- 1) Leggi i sensori
- 2) Applica la funzione di controllo a letture e calcola la prossima accelerazione (decelerazione)
- 3) Applica l'azione calcolata

Cristina Baroglio

50

## Fare la cosa giusta: cosa vuol dire?

- La **funzione deliberativa** di un agente determina le azioni che saranno eseguite
- In termini informali un agente è **razionale** quando “*fa la cosa giusta*”, cioè opera per conseguire il “successo”
- Occorre una guida: una **misura di prestazione**
- Tale misura dipende *dall' effetto* che si desidera conseguire

Cristina Baroglio

51

## Comportamento razionale

Il **comportamento razionale** di un agente dipende da **4 fattori**:

1. Azioni nelle facoltà dell'agente
2. Misura della prestazione
3. Conoscenza dell'ambiente
4. Percezione

Un **agente razionale** dovrebbe scegliere sempre un'azione che massimizza la misura di prestazione attesa, data la particolare sequenza percettiva in oggetto e le informazioni derivabili dalla conoscenza dell'ambiente

Cristina Baroglio

52



## RAZIONALITÀ

### Ottimizza il risultato atteso

Possono intercorrere fattori ignoti o imprevedibili che impediscono di conseguire il risultato atteso

Voglio che al compleanno di mio figlio vi sia la torta più bella e meno costosa. Conosco le offerte dei pasticceri della zona e i gusti di mio figlio. Ho sufficiente denaro in tasca. So fare acquisti.

*Cosa non posso prevedere:*  
la nonna porterà una torta identica (regalandola)

*Cosa non so*  
Offerte di una nuova pasticceria

## ONNISCENZA/CHIAROVEGGENZA

### Ottimizza il risultato reale

Non possono intercorrere fattori ignoti o imprevedibili

Es. non compero nessuna torta perché so che la nonna regalerà la torta perfetta. Ottengo il prezzo migliore perché so di una nuova pasticceria che fa delle offerte.