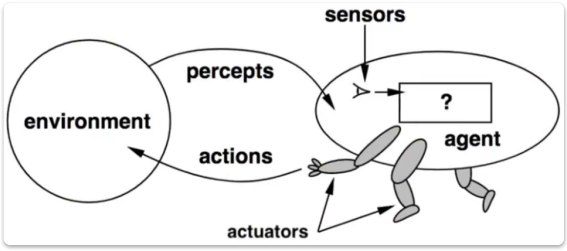
Lezione06

Table of contents

Agents



Nell'intelligenza artificiale, un agente viene inteso come un qualcosa che:

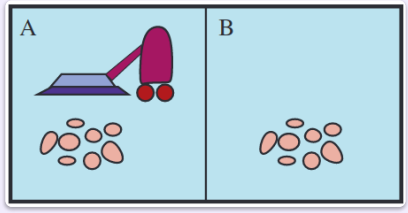
- cattura **percetti** da un ambiente e li legge, creando prodotti detti **azioni**;
- prende azioni in completa autonomia tramite **attuatori** per reagire ai cambiamenti, dipesi dai percetti, per raggiungere il proprio obbiettivo;
- potrebbe migliorare le proprie prestazioni imparando, anche se non è un algoritmo (non termina);
- potrebbe usare la conoscenza per raggiungere la *razionalità*.

Entrato in uso nella ricerca, il termine ha avuto variazioni del significato. Quello che vedremo sarà a ogni modo legato o alla definizione di un testo, o alle definizioni formalizzate per motivi di tempo.

Agents & Environments

Per motivi di semplicità, qual'ora il nostro problema sia molto difficile da descrivere rispetto al metodo da utilizzare per raggiungere l'obbiettivo, vale la pena non farlo.

Nel libro "Artificial Intelligence: A Modern Approach" degli autori S.J. Russel e P. Norvig, viene fatta una caratterizzazione degli agenti, in base all'ambiente in qui agiscono.

> ⁝☰ **The vacuum-cleaner's world**
>
> 
>

Nel mondo dell'aspirapolvere, esistono celle che possono essere occupate per essere pulite, che quindi sono o sporche, o pulite. Versioni diverse del mondo dell'aspirapolvere, permettono diverse regole riguardo "cosa" l'agente può percepire.

Fully observable vs partially observable

Un ambiente può essere **accessibile** o **non accessibile** (parzialmente accessibile).
Ci chiediamo se il nostro sistema ha *sensori* sufficientemente sofisticati, per arrivare agli obbiettivi. Sulla base di questi obbiettivi, non ci interessano altre informazioni che contornano la soluzione, ma soltanto un punto specifico (esempio: catturare pacchetti Wi-Fi).

- Dire che un ambiente è **completamente osservabile** vuole dire che abbiamo tutte le informazioni che ci servono per completare un ragionamento razionale: fare la scelta ottima; come negli scacchi.
- Dire che un ambiente è **parzialmente osservabile** vuole dire che non conosciamo tutto l'ambiente; proprio come un gioco di carte.

Deterministic vs non-deterministic

Un ambiente anche se completamente osservabile, potrebbe essere **stocastico**.
Usare questo termine significa che stiamo introducendo una variabile di dubbio, o casualità (esempio: pescare una carta da un mazzo): una certa probabilità viene associata a ogni evento e viene usata dal nostro agente per compiere azioni e raggiungere obbiettivi.

- Gli eventi sono **deterministici**; avvengono con probabilità 1.
- Chiesta un'azione, questa verrà compiuta per come ce l'aspettiamo.

> ⓘ **"Stocastico": che ha distribuzione probabilistica o pattern analizzabili ma non deducibili con precisione**

Static vs dynamic

- Ambiente **dinamico**
  Se l'unica cosa che cambia nell'ambiente è data dalle azioni dell'agente, allora questo è capace di controllarlo (pianificazione). L'unico problema diventa la sequenza di azioni per arrivare al risultato; possiamo interessarci quindi del potere computazionale necessario per farlo.
- Ambiente **statico**
  L'agente non necessita di percepire l'ambiente, il che rende la scelta dell'azione successiva, più facile da eseguire.

Discrete vs continuous

Se il numero di percetti che è possibile catturare dall'ambiente è **discreto**, allora sarà possibile associare matrici per ognuno che identifica se compiere o meno un'azione.
La grandezza delle matrici fa diventare la matematica dietro il metodo, molto raffinata, perché se il numero diventa continuo, avere matrici arbitrariamente grandi diventa un ostacolo piuttosto che un pregio.

Maggior parte dei casi, include ambienti *finiti*, un sotto-ambiente degli ambienti discreti. Per casi particolari, ambienti **continui** sono interessanti da prendere.

Episodic vs sequential

- In un ambiente a **episodi**, solamente il percetto che arriva in ingresso influenza la prossima azione da prendere. Se indipendentemente da tutti i ragionamenti, possiamo sostituire con una mappa che associa percetto/azione, allora l'agente intende l'ambiente è episodico.

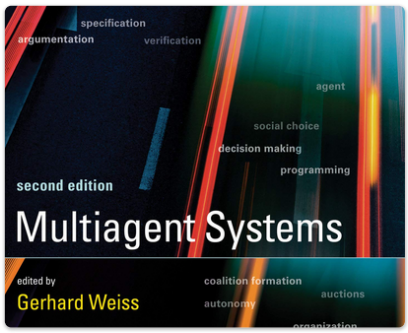
- In una ambiente **sequenziale**, la memoria necessaria per compiere l'azione non è finita, come in una macchina di Turing. Possono essere utilizzati percetti passati, complicando la soluzione dell'agente.

Single-agent vs multi-agent

Delle volte, il mondo contiene *altri* agenti.
Stiamo ipotizzando che questi agenti raggiungano degli obbiettivi, stiamo dicendo che molto probabilmente alcuni di questi avranno lo stesso obbiettivo.
Sapere che parti del mondo si muovono non secondo regole naturali ma strategicamente, farà sì che l'agente le sfrutti per lavorare meglio all'interno dell'ambiente.

- Agenti ragionano in modo **cooperativo** per massimizzare l'utilità complessiva.
- Agenti sono **competitivi**, costruendo un modello che descrive l'avversario, possiamo giocare su questo per vincere (teoria dei giochi).
- Agenti sono **strategici** se l'unica cosa che cambia sono le azioni degli alri.

Weiss Agents

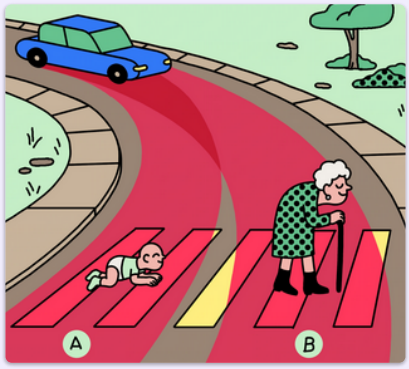


La copertina del libro "Multiagent Systems"

Dal libro "Multiagent systems" dell'autore G. Weiss, viene fatta classificazione degli agenti non più tanto in termini dell'ambiente in cui interagisce, ma come vengono realizzati.

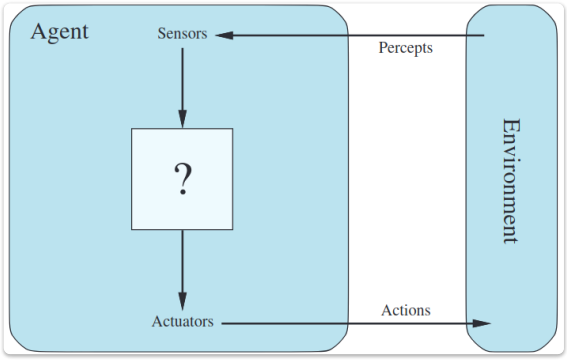
- **Reactive agents**
  Dato un percetto, l'agente ha mappa interna che gli permette di decidere l'azione da intraprendere. Se non memorizzato uno stato di computazione ma usata una funzione pura, quello che andiamo a fare è una procedura (come la porta vista nella lezione scorsa).
- **Layered agents**
  Mediante un'architettura di reazione, diciamo quali azioni compiere dopo avere ricevuto un percetto. Tipicamente i livelli sono 3: alcuni percetti sono associati a reazioni immediate, mentre altri sono più complessi ed elaborati.
- **Logic-based agents**
  Se vogliamo permettere all'agente di avere obbiettivi propri, non tanto soltanto perché reagisce ad eventi ma piuttosto perché ragiona, allora siamo nella condizione logica. L'agente pensa a cosa deve fare per raggiungere il "goal", usando deduzione logica.
- **Belief-Desire-Intention** (BDI) **agents**
  Lo scopo di questa architettura è quello di descrivere gli obbiettivi da perseguire: l'agente ha degli obbiettivi e questi prendono nome d'*intenzioni*, che devono essere consistenti e precise in un istante di tempo. Con *desideri* s'intende obbiettivi che non sono ancora diventati piano d'azione, ma che vogliono essere perseguiti. Un *pensiero* (o credenza) sarebbe una convinzione che non ha ricevuto conferma, finché non studiata.

> **The autonomous car: a philosophical dilemma**



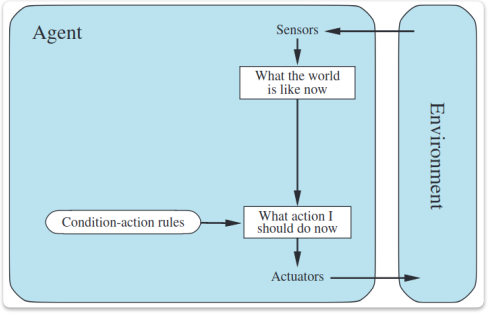
Se una macchina, un treno, o un tram, fossero macabramente costretti a decidere se colpire una persona piuttosto che un'altra, per come sono fatti gli agenti, questi farebbero sempre la decisione più razionale.

Russel & Norvig Agents



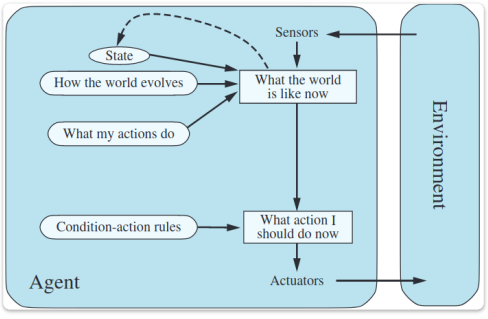
Sempre nel testo visto in precedenza, viene fatta una classificazione degli agenti.

- Simple (reflex) agents



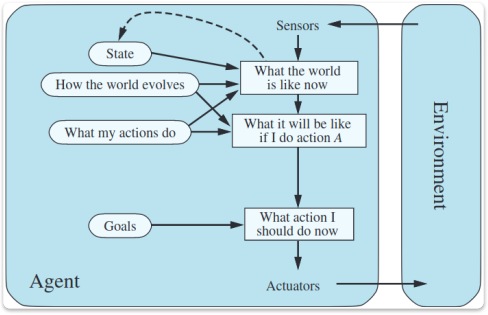
  Una mappa condizione/azione, prende un percetto e decide la reazione. È un modo molto semplice.

- Model-based (reflex) agents

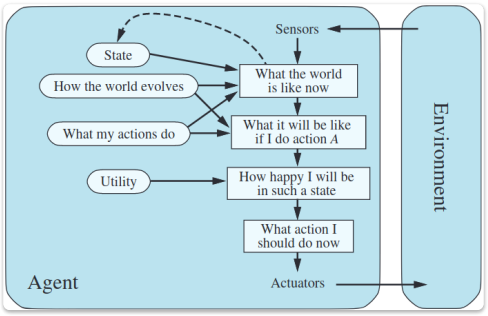


  Il percetto viene acquisito dal sensore e la scelta presa in modo più raffinato.

- Goal-based agents



  Non basta sapere come il mondo evolve, dobbiamo mettere dentro anche cosa vogliamo fare. Nel momento in cui consideriamo "goal" espliciti (potenzialmente in conflitto tra di loro), l'agente fa scelte per arrivare all'obbiettivo.
- Utility-based agents



  Con utlità quantificata come numero, l'agente sceglierà cosa compiere nel tentativo di massimizzare questa misura. L'agente decide tenendo conto che a lungo termine questa debba essere al massimo possibile.

Agenti software

- Software agents
- Distributed agents
- Mobile agents
- Autonomous agents
- Intelligent agents

14/03/2023