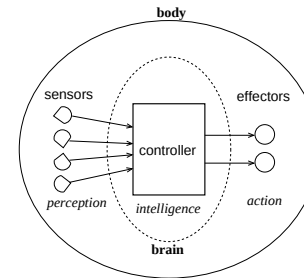


topics:

- intelligent agents: decision-making, rationality, environments
- markov decision processes
- probability theory

intelligent agents

- an *agent* is a system that is *situated* in an environment, and which is capable of *perceiving* its environment and *acting* in it to satisfy its design objectives.



different kinds of agents

- human "agent":

<i>environment:</i>	physical world
<i>sensors:</i>	eyes, ears, ...
<i>effectors:</i>	hands, legs, ...

- software agent:

<i>environment:</i>	e.g., UNIX operating system
<i>sensors:</i>	ls, ps, ...
<i>effectors:</i>	rm, chmod, ...

- internet agent:

<i>environment:</i>	the Internet
<i>sensors:</i>	http requests
<i>effectors:</i>	http commands

- embodied (robotic) agent:

<i>environment:</i>	physical world
<i>sensors:</i>	light meters, bumpers, thermometers, ...
<i>effectors:</i>	motors attached to wheels, treads, legs, grippers, ...

agent decision-making: what to do?

- need to know *what* to do in any given *state*
- what:
 - an *action* that the agent can take
- state:
 - a configuration of the agent and its environment, such as the position of all the pieces on a chess board, or the robots and the ball on a robot soccer field, or the position of a robot's gripper, or all the bids in an electronic market
- need to know *how* and *when* to evaluate success of any action
- how:
 - an objective performance measure
 - application specific
- when:
 - in discrete *episodes*, or over long periods?

rational agents

- rationality is concerned with *expected success* given the *information that is available*
- *ideal rational agent*:
for each percept sequence, an ideal rational agent will act to maximize its expected performance measure, on the basis of information provided by its percept sequence plus any information built in to the agent
- note that this does not preclude performing actions to *find things out*
- also note that real agents don't know enough to always make the best choices!
- more precisely, we can view an agent as a function:

$$f : P^* \rightarrow A$$

from sequences of percepts P to actions A .

agent: simple example

- for example, a “quadratic” agent:

percept	action
0	0
1	1
2	4
3	9
4	16
...	...

- this table can be viewed as a *specification* of the agent
- we don't have to *implement* agent as table lookup:

```
int agent( int n ) {  
    return n * n;  
}
```

design and structure of agents

- two components:
 - *program*: the thing which defines the mapping from percept sequences to actions
 - *architecture*: the “shell” into which the agent program fits
- i.e., agent = program + architecture
- an appropriate architecture can make the design of agent programs much easier
- *autonomy*
 - *autonomy* is a crucial concern for agents
 - run-time decisions are made by the agent *alone*—i.e., no human remote control
 - means behavior is based on *own* experience
 - implies *learning*, *adaptation*

agent environments

- the “PAGE” approach to defining agents:
 - percepts
 - actions
 - goals
 - environment
- example: wall-painting robot
 - percepts: color, obstacles, ...
 - actions: move forward, move backward, turn, paint, ...
 - goals: maximize number of repairs, safety, ...; minimize use of paint, ...
 - environment: classroom
- example: medical diagnosis system
 - percepts: symptoms, findings, patient answers, ...
 - actions: questions, tests, treatments, ...
 - goals: healthy patient, minimise costs, ...

- environment: patient, hospital, ...
- example: email manager
 - percepts: email arrived, headers, content of email, ...
 - actions: delete email, reorder email, obtain user attention, ...
 - goals: present important email first; hide junk email, ...
 - environment: mail reader, operating system, ...

environment characteristics

- accessible vs inaccessible
 - an accessible environment is one in which the agent can obtain complete, accurate, up-to-date information about the environment's state
 - most moderately complex environments (including, for example, the everyday physical world and the internet) are inaccessible
 - the more accessible an environment is, the simpler it is to build agents to operate in it
- deterministic vs non-deterministic
 - a deterministic environment is one in which any action has a single guaranteed effect—there is no uncertainty about the state that will result from performing an action
 - the physical world can to all intents and purposes be regarded as non-deterministic
 - non-deterministic environments present greater problems for the agent designer
- episodic vs non-episodic
 - in an episodic environment, the performance of an agent is dependent on a number of discrete episodes, with no link between the performance of an agent in different

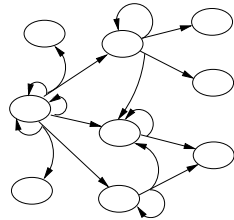
scenarios

- an example of an episodic environment would be a mail sorting system
- episodic environments are simpler from the agent developer's perspective because the agent can decide what action to perform based only on the current episode—it need not reason about the interactions between this and future episodes
- static vs dynamic
 - a static environment is one that can be assumed to remain unchanged except by the performance of actions by the agent
 - a dynamic environment is one that has other processes operating on it, and which hence changes in ways beyond the agent's control
 - The physical world is a highly dynamic environment
- discrete vs continuous
 - an environment is discrete if there are a fixed, finite number of actions and percepts in it
 - Russell and Norvig give a chess game as an example of a discrete environment, and taxi driving is an example of a continuous one

agent decision-making

- agent activity is frequently represented as a sequence of **states** and a set of **actions**
- in any state, s_i , the agent has a choice of which action, a_i , to take; the action will move them to another state, s_j
- a **transition function** defines a mapping from states to actions to states
- in a *non-deterministic environment*, there is not a simple transition function—i.e., an action can lead to multiple states
- in situations where we know (or can determine) which state we are in, then this is called **Markov decision process (MDP)**
- an MDP has the following formal model:
 - a state space S ;
 - a set of actions, $A(s) \subseteq A$, applicable in each state $s \in S$;
 - transition probabilities $\text{Pr}_a(s' | s)$ for $s, s' \in S$ and $a \in A$;
 - action costs $c(a, s) > 0$; and
 - a set of goal states $G \subseteq S$

- for each state we have a set of actions we can apply, and these take us to other states with some probability
- we don't know which state we will end up in, but we know which one we are in after the action (we have *full observability*)
- this gives us a problem space that could look something like this:



where each oval represents a state and each set of arrows represent actions; the arrows point to new states, given each action performed

- one can associate a number, a *transition probability*, with each arrow, indicating the likelihood that the action will result in the following state

policy

- the function that helps the agent decide what to do is called a **policy**, π , which maps states into applicable actions, $\pi(s_i) = a_i$
- a policy allows us to compute a *probability distribution* across all the trajectories (paths) from a given initial state (by following sequences of arrows from one oval to another, to another)
- this is the product of all the *transition probabilities*, $\Pr_{a_i}(s_{i+1} \mid s_i)$, along the trajectory
- goal states are taken to have no cost, no effects, so that if $s \in G$:
 - $c(a, s) = 0$
 - $\Pr(s \mid s) = 1$

- we can then calculate the expected cost of a policy starting in state s_0
- which is just the probability of the policy multiplied by the cost of traversing it:

$$\sum_{i=0}^{\infty} c(\pi(s_i), s_i)$$

- an *optimal policy* is then a π^* that has minimum expected cost for all states s

acting in an environment: issues

- when environments are non-deterministic, actions don't have predictable outcomes: actions can fail.
- this can partly be handled by trying to model the inherent *uncertainty* in actions
- when environments are inaccessible, we don't know everything about an environment
- an agent's knowledge is *incomplete* (i.e., not *fully observable*)
- we can make assumptions to "fill in" gaps in the knowledge-base... but these may turn out to be false
- when environments are dynamic, the world can change as a result of things other than the agent's actions
- one way of looking at this is that the world is non-deterministic and the world is inaccessible
- thus we have to model non-determinism/dynamism and handle assumptions

modeling uncertainty

- one thing we can do is to distinguish between things that we know will be true and things that may be true
- things that will always be true are *necessary truths*:
 - $2 + 2 = 4$
 - Simon is mortal
- things that may be true are *possible truths*:
 - Tomorrow will be sunny
 - A user will click on "check out" after placing something in her shopping cart
- one approach to dealing with these situations is to use **probability theory**
- thus we can say that 90% of the time the user will click "check out" after she places something in her shopping cart, but 10% of the time, she will click "log out" to disconnect from the shopping site

probability theory

- we start with a *sample space* Ω
- for instance, Ω for the action of rolling a die would be $\{1, 2, 3, 4, 5, 6\}$
- subsets of Ω then correspond to particular events, e.g., the set $\{2, 4, 6\}$ corresponds to the event of rolling an even number
- we use S to denote the set of all possible events:

$$S = 2^\Omega$$

- it is sometimes helpful to think of the sample space in terms of Venn diagrams—indeed all probability calculations can be carried out in this way
- a probability measure is a function:

$$\Pr : S \mapsto [0, 1]$$

such that:

$$\begin{aligned}\Pr(\emptyset) &= 0 \\ \Pr(\Omega) &= 1 \\ \Pr(E \cup F) &= \Pr(E) + \Pr(F), \text{ whenever } E \cap F = \emptyset\end{aligned}$$

- saying $E \cap F = \emptyset$ is the same as saying that E and F cannot occur together
- they are: *disjoint* or *exclusive*
- if the occurrence of an event E has no effect on the occurrence of an event F , then the two are said to be *independent*
- an example of two independent events are the throwing of a 2 on the first roll of a die, and a 3 on the second
- if E and F are independent, then:

$$\Pr(E \cap F) = \Pr(E) \cdot \Pr(F)$$

- when E and F are not independent, we need to use:

$$\Pr(E \cap F) = \Pr(E) \cdot \Pr(F|E)$$

where $\Pr(F|E)$ is the *conditional probability* of F given that E is known to have occurred.

- to see how $\Pr(F)$ and $\Pr(F|E)$ differ, consider F is the event “a 2 is thrown” and E is the event “the number is even”

conditional probability

- we can calculate conditional probabilities from:

$$\begin{aligned}\Pr(F|E) &= \frac{\Pr(E \cap F)}{\Pr(E)} \\ \Pr(E|F) &= \frac{\Pr(E \cap F)}{\Pr(F)}\end{aligned}$$

which, admittedly is rather circular, so...

- we can combine these two identities to obtain *Bayes' rule*:

$$\Pr(F|E) = \frac{\Pr(E|F) \Pr(F)}{\Pr(E)}$$

- another useful rule is *Jeffrey's rule*:

$$\Pr(F) = \Pr(F|E) \Pr(E) + \Pr(F|\neg E) \Pr(\neg E)$$

- more general versions are appropriate when considering events with several different possible outcomes

example

- (A) what is the probability of rolling a die twice and having the values rolled total 7?
(B) what is the probability of A, given C—that you roll a 4 first?
- to answer B, first consider the probability of C, rolling a 4: since the die has 6 sides, the probability is $\Pr(C) = 1/6$
- thus B is the probability of A given C: $\Pr(A|C)$
- to answer A, consider how many ways are there to make 7?
there are 6 ways: $1 + 6, 6 + 1, 2 + 5, 5 + 2, 3 + 4, 4 + 3$
- given that there are $6 \times 6 = 36$ possible ways of rolling two dice, then the probability of rolling a 7 after rolling a die twice is: $\Pr(A) = 6/36 = 1/6$
- $\Pr(C \cap A)$ is the probability of rolling a 4 and getting 7 after rolling two dice; since there is only one way to do this: $4 + 3$, then the probability is $1/36$
- finally, we use the equations on the previous slide to determine our conditional probability:

$$\Pr(A|C) = \frac{\Pr(C \cap A)}{\Pr(C)} = \frac{1/36}{1/6} = 1/6$$

second example

- what is the probability of picking the ace of spades from a deck of cards?
- let $Pr(E)$ be the probability of picking an ace: $Pr(E) = 4/52 = 1/13$
- let $Pr(F)$ be the probability of picking a spade: $Pr(F) = 13/52 = 1/4$
- then $Pr(E \cap F)$ is the probability of picking the ace of spades: $Pr(E \cap F) = 1/52$
This is the CONJUNCTION is E AND F — picking both an ace and a spade.
- let's determine the CONDITIONALS $Pr(E|F)$ and $Pr(F|E)$
- first, suppose I pick a card and tell you that it is a spade—what is the probability that it is the ace of spades?

$$Pr(E|F) = \frac{Pr(E \cap F)}{Pr(F)} = \frac{1/52}{1/4} = 4/52 = 1/13$$

- second, supposed I pick a card and tell you that it is an ace—what is the probability that it is the ace of spades?

$$Pr(F|E) = \frac{Pr(E \cap F)}{Pr(E)} = \frac{1/52}{1/13} = 13/52 = 1/4$$

a third, tricky, example

- what is the probability of picking two aces from a deck of cards, without replacement? i.e., first I pick an ace, then I pick again, without putting back the first ace
- let $Pr(E)$ be the probability of picking an ace the first time: $Pr(E) = 4/52 = 1/13$
- let $Pr(F)$ be the probability of picking an ace the second time: $Pr(F) = 3/51$
- I don't know what the CONJUNCTION is ($Pr(E \cap F)$) because it is a funny thing to express
- however, we do know the CONDITIONAL $Pr(F|E)$ because that is already inherent in the definition of $Pr(F)$, and in fact:

$$Pr(F|E) = Pr(F) = 3/51$$

- so now we can compute the CONJUNCTION:

$$Pr(F|E) = Pr(E \cap F) / Pr(E) = \frac{Pr(E \cap F)}{1/13} = 3/51$$

$$Pr(E \cap F) = 3/51 \cdot 1/13$$

expected value

- consider being offered a bet in which you pay \$2 if an odd number is rolled on a die, and win \$3 if an even number appears
- to analyse this prospect we introduce a *random variable* X , as the function:

$$X : \Omega \mapsto \mathfrak{R}$$

from the sample space to the values of the outcomes. Thus for $\omega \in \Omega$:

$$X(\omega) = \begin{cases} 3, & \text{if } \omega = 2, 4, 6 \\ -2, & \text{if } \omega = 1, 3, 5 \end{cases}$$

- the probability that X takes the value 3 is:

$$\begin{aligned} Pr(\{2, 4, 6\}) &= Pr(\{2\}) + Pr(\{4\}) + Pr(\{6\}) \\ &= 0.5 \end{aligned}$$

- how do we analyse how much this bet is worth to us?
- we need to calculate the *expected value* of X
- this is defined by:

$$E(X) = \sum_k k Pr(X = k)$$

where the summation is over all values of k for which $Pr(X = k) \neq 0$

- thus the expected value of X is \$1.50, and we take this to be the value of the bet
- and now we can make a first stab at defining how to decide what to do in an uncertain world.
- we choose the action which has *maximum expected value*