Decision-Theoretic Agents

- **Quantitative management of uncertainty**

  Keep track of the probability of each possible state of the world.

- **Quantitative management of goals**

  Replace goal predicate $goal(s)$ with utility function $U(s)$.

- **Choose actions to maximize expected utility**

  Allows agents to make tradeoffs.

```
function DT-AGENT(percept) returns an action
   static: a set probabilistic beliefs about the state of the world

   calculate updated probabilities for current state based on
        available evidence including current percept and previous action
   calculate outcome probabilities for actions,
        given action descriptions and probabilities of current states
   select action with highest expected utility
        given probabilities of outcomes and utility information
   return action
```

Introduction to Probability

A **random variable** is a variable (e.g., heading of robot) whose value is not known with certainty.

Every random variable has a **type**, which specifies the set of possible values that it can have. (e.g., $\{north, south, east, west\}$).

Every random variable has a **probability distribution** which represents the agent's beliefs about the value of the variable.

$$P(heading = north) = 0.1$$
$$P(heading = south) = 0.1$$
$$P(heading = east) = 0.7$$
$$P(heading = west) = 0.1$$

The probabilities in a distribution must lie between 0 and 1 (inclusive) and sum to 1.

$$P(heading = north) = 1 \text{ implies}$$
$$P(heading = south) = 0$$
$$P(heading = east) = 0$$
$$P(heading = west) = 0$$

Probability and Propositional Logic

We can also write probabilities for sentences in propositional logic (i.e., logic without logical variables or quantifiers):

$$P(heading = north \ \vee \ heading = south) = 0.2$$
$$P(\neg(heading = north)) = 0.9$$

Probability law of inclusion and exclusion:

$$P(A \ \vee \ B) = P(A) + P(B) - P(A \ \wedge \ B)$$

Writing a Distribution as a Table

We often write a distribution as a probability table instead of as a set of statements:

$heading$	$P(heading)$
$north$	0.1
$south$	0.1
$east$	0.7
$west$	0.1

In the table, every logical sentence corresponds to a region, and its probability can be computed by summing the cells in that region.

We will use the special notation $\mathbf{P}(heading)$ to refer to the **entire** table (i.e., the entire distribution).

Boolean Random Variables

We will write Boolean random variables using a slightly different notation. Instead of writing:

$$P(\mathsf{PowerOn} = 1) = 0.95$$
$$P(\mathsf{PowerOn} = 0) = 0.05$$

we will write

$$P(\mathsf{PowerOn}) = 0.95$$
$$P(\neg\mathsf{PowerOn}) = 0.05$$

Multiple Random Variables: The Joint Distribution

When there are multiple random variables, the most important thing to consider is the joint distribution:

	Starts		¬Starts	
	Gas	¬Gas	Gas	¬Gas
GasGauge	.5040	.0006	.2160	.0594
¬GasGauge	.0560	.0014	.0240	.1386

We will use three notations:

- **Single Row:** $P(\mathsf{Gas}, -\mathsf{GasGauge}, \mathsf{Starts}) = P(\mathsf{Gas} \wedge \neg\mathsf{GasGauge} \wedge \mathsf{Starts}) = .0560.$
- **Whole Table:** $\mathbf{P}(\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts})$ stands for the entire joint distribution.
- **Summation Shorthand:** $\sum_{\mathsf{Gas}} P(\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts}) = P(\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts}) + P(-\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts})$

  For the distribution, we write this as $\sum_{\mathsf{Gas}} \mathbf{P}(\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts}) = \mathbf{P}(\mathsf{GasGauge}, \mathsf{Starts})$

Marginal Probabilities

Suppose we want to know the probability that $\mathsf{GasGauge}$ is true. We obtain this by summing over the entire top row of the table.

$$P(\mathsf{GasGauge}) = .5040 + .0006 + .2160 + .0594 = 0.78$$

Another example: the probability that Gas is false:

$$P(-\mathsf{Gas}) = .0006 + .0014 + .0594 + .1386 = 0.2$$

This is called *marginalizing*. We write it as

$$P(-\mathsf{Gas}) = \sum_{GasGauge, Starts} P(-\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts})$$

$\mathbf{P}(\mathsf{Gas})$ is called the *marginal distribution* of Gas.

Conditional Probability: Belief Updating

Suppose we want to know the probability that the car will start given that we have observed GasGauge and found it to be true.

$$P(\mathsf{Starts}|\mathsf{GasGauge}) = \frac{P(\mathsf{Starts}, \mathsf{GasGauge})}{P(\mathsf{GasGauge})} = \frac{.5040 + .0006}{.5040 + .0006 + .2160 + .0594} = .6469$$

	Starts		–Starts	
	Gas	–Gas	Gas	–Gas
GasGauge	.5040	.0006	.2160	.0594
–GasGauge	.0560	.0014	.0240	.1386

Since GasGauge is true, one of the four situations from the top row of the table must hold. We *renormalize* the probabilities and then marginalize to compute $\mathbf{P}(\mathsf{Starts})$.

Every time the agent learns the value of some random variable, it can update the joint distribution to include only the remaining relevant variables.

Conditional Probability Distribution

The conditional probability distribution $\mathbf{P}(\mathsf{GasGauge}|\mathsf{Gas})$:

	GasGauge	¬GasGauge
Gas	.9	.1
¬Gas	.3	.7

Each row must sum to 1.

Chain Rule of Probability

$$P(A,B,C) = P(A|B,C) \cdot P(B|C) \cdot P(C)$$

for any number of variables in any order.

Proof:

$$\begin{aligned}
P(A,B,C) &= \frac{P(A,B,C)}{P(B,C)} \cdot P(B,C) \\
&= P(A|B,C) \cdot P(B,C) \\
&= P(A|B,C) \cdot \frac{P(B,C)}{P(C)} \cdot P(C) \\
&= P(A|B,C) \cdot P(B|C) \cdot P(C)
\end{aligned}$$

This holds for distributions as well:

$$\mathbf{P}(A,B,C) = \mathbf{P}(A|B,C) \cdot \mathbf{P}(B|C) \cdot \mathbf{P}(C)$$

Expected Values

Suppose that our vacuum agent is in a square where there is dirt, but suppose that its suck action only succeeds 60% of the time. Then if it performs a suck in this state, there are two possible resulting states: A and B. In state A, the dirt is gone, and in state B, the dirt is still there. Suppose that the utility of state A is $+10$ and the utility of state B is 0. What is the expected value of performing the suck action?

result state S_{t+1}	$P(S_{t+1})$	$U(S_{t+1})$
A	0.60	+10
B	0.40	0

$$EU(S_{t+1}) = \sum_x P(S_{t+1} = x) \cdot U(S_{t+1} = x)$$

In this case: $EU(S_{t+1}) = 0.60 \cdot 10 + 0.40 \cdot 0 = 6$

Implementing a Decision-Theoretic Agent: Version 1

Suppose we have a conditional joint distribution $\mathbf{P}(S_t, O_t, S_{t+1}|A_t)$.

At time t, the agent performs the following:

- **Observe** O_t. After observing O_t, we can remove the cells for the other possible observations, and renormalize the table to get $\mathbf{P}(S_t, S_{t+1}|O_t, A_t)$.
- **Choose** A_t. For each possible value of A_t, compute the following:

$$\mathbf{P}(S_{t+1}|O_t, A_t) = \sum_{S_t} \mathbf{P}(S_{t+1}, S_t|O_t, A_t)$$

  Then, compute

$$EU(S_{t+1}|O_t, A_t) = \sum_{S_{t+1}} P(S_{t+1}|O_t, A_t)U(S_{t+1})$$

  Choose the action A_t that has the highest value of $EU(S_{t+1}|O_t, A_t)$.

Problem: The Joint Distribution is Too Big

To perform a sequence of k actions, we need the joint distribution

$$\mathbf{P}(S_0, O_0, A_0, S_1, O_1, A_1, \ldots, S_k).$$

This would be enormous!

$\boxed{\text{Key Observation: This joint distribution contains structure that can be exploited.}}$

The structure is known as **conditional independence**.

It can be exploited by a method known as **belief networks**.

Belief Networks (1): Independence

Two random variables X and Y are *independent* if $P(X, Y) = P(X) \cdot P(Y)$ for all values of X and Y. This can be written as $\mathbf{P}(X, Y) = \mathbf{P}(X) \cdot \mathbf{P}(Y)$. Intuitively, there is no causal connection betwen X and Y.

Example:

- **X**: coin with probability $X = true$ of .4, and $X = false$ of .6.
- **Y**: coin with probability $Y = true$ of .8, and $Y = false$ of .2.

Joint distribution:

	$X = true$	$X = false$
$Y = true$	.32	.48
$Y = false$	.08	.12

Belief Networks (2): Conditional Independence

Two random variables X and Y are *conditionally independent given* Z if $P(X, Y|Z) = P(X|Z) \cdot P(Y|Z)$ for all values of X, Y, and Z.

Example:

$$\mathbf{P}(\mathsf{Starts}, \mathsf{GasGauge}|\mathsf{Gas}) = \mathbf{P}(\mathsf{Starts}|\mathsf{Gas}) \cdot \mathbf{P}(\mathsf{GasGauge}|\mathsf{Gas})$$

	Starts		¬Starts	
	GasGauge	¬GasGauge	GasGauge	¬GasGauge
Gas	.630	.070	.270	.030
¬Gas	.003	.007	.297	.693

	GasGauge	¬GasGauge
Gas	.9	.1
¬Gas	.3	.7

	Starts	¬Starts
Gas	.7	.3
¬Gas	.01	.99

Simplifying the Joint Distribution by Conditional Independence

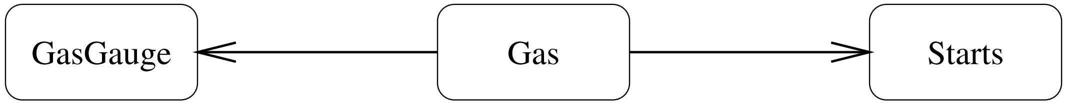
By the definition of conditional probability, we have

$$\mathbf{P}(\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts}) = \mathbf{P}(\mathsf{GasGauge}, \mathsf{Starts}|\mathsf{Gas})\mathbf{P}(\mathsf{Gas})$$

But by conditional independence, the first factor can be rewritten as

$$\mathbf{P}(\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts}) = \mathbf{P}(\mathsf{GasGauge}|\mathsf{Gas})\mathbf{P}(\mathsf{Starts}|\mathsf{Gas})\mathbf{P}(\mathsf{Gas})$$

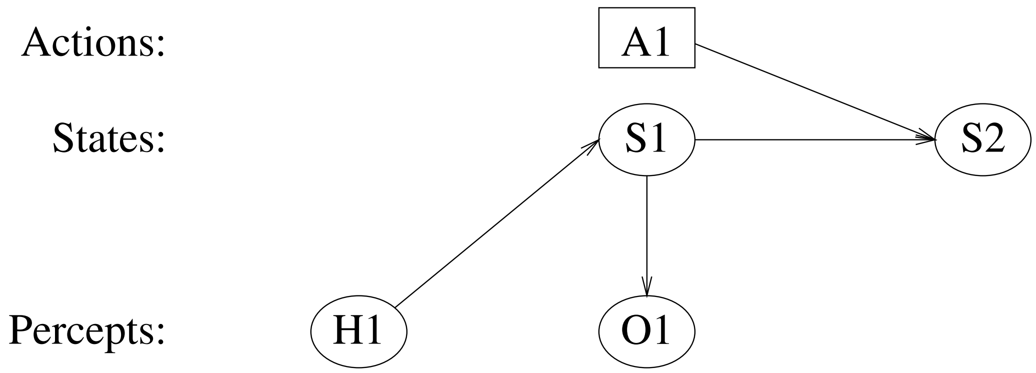
Bayes Network Diagram



- **One node for each random variable**
- **Nodes with only outgoing arrows store marginal distributions**
  The Gas node stores $\mathbf{P}(\mathsf{Gas})$.
- **Nodes with incoming arrows store conditional probability tables**
  The Starts node stores $\mathbf{P}(\mathsf{Starts}|\mathsf{Gas})$.
- **Only direct dependencies are shown.** If two nodes are not directly connected, then they are conditionally independent on any set of nodes that disconnects them.
  Starts and $\mathsf{GasGauge}$ are conditionally independent given Gas.
- **Joint distribution is product of the distributions of the nodes.**

$$\mathbf{P}(\mathsf{Gas}, \mathsf{GasGauge}, \mathsf{Starts}) = \mathbf{P}(\mathsf{GasGauge}|\mathsf{Gas})\mathbf{P}(\mathsf{Starts}|\mathsf{Gas})\mathbf{P}(\mathsf{Gas})$$

Implementing a Decision-Theoretic Agent: Version 2



- $\mathbf{P}(S_t|H_t)$: Probability of current state given our percept/action history H_t.
- $\mathbf{P}(S_{t+1}|S_t, A_t)$: How the world changes when we act.
- $\mathbf{P}(O_t|S_t)$: How our sensors work.
- A_t is not a random variable, because we can choose its value. (This is why it appears in a box instead of an oval.)

We can compute the joint probability distribution as the product of the individual distributions

$$\mathbf{P}(S_t, O_t, S_{t+1}|A_t, H_t) = \mathbf{P}(S_t|H_t) \cdot \mathbf{P}(O_t|S_t) \cdot \mathbf{P}(S_{t+1}|S_t, A_t)$$

Implementing a Decision-Theoretic Agent: Version 2 (2)

At each time step, the agent does the following:

- **Receive current percept:** O_t
- **Update beliefs about S_t:** Compute

$$P(S_t|H_t, O_t) = \alpha P(S_t|H_t) \cdot P(O_t|S_t)$$

  where $\alpha = 1/\sum_{S_t} P(S_t|H_t) \cdot P(O_t|S_t)$.

- **Choose a new action A_t**

$$\begin{aligned} \operatorname*{argmax}_{A_t} EU(S_{t+1}|H_t, O_t, A_t) &= \sum_{S_{t+1}} P(S_{t+1}|H_t, O_t, A_t) U(S_{t+1}) \\ &= \sum_{S_{t+1}} \sum_{S_t} P(S_{t+1}|S_t, A_t) \cdot P(S_t|H_t, O_t) \cdot U(S_{t+1}) \end{aligned}$$

- **Compute beliefs about the next state:**
  Compute $\mathbf{P}(S_{t+1}|H_{t+1}) = \mathbf{P}(S_{t+1}|H_t, O_t, A_t)$

Over time, H_{t+1} contains all of the history of the computation. But we don't need to represent this explicitly if we just represent the probability distribution $\mathbf{P}(S_{t+1}|H_{t+1})$.