



THE UNIVERSITY OF TEXAS AT AUSTIN  
McCOMBS SCHOOL OF BUSINESS

# Decision Trees 2

---

## Lecture 7

STA 371G

## Announcement: Midterm 1 signup

On Thursday at 7 PM, you'll get a link to sign up for a time to take Midterm 1 in the ModLab. You can pick any of the following times, but sign up ASAP as space is limited!

- Thursday, February 22 @ 5:00 PM
- Thursday, February 22 @ 6:45 PM
- Friday, February 23 @ 8:00 AM
- Friday, February 23 @ 9:45 AM
- Friday, February 23 @ 11:30 AM
- Friday, February 23 @ 1:15 PM

## Oil drilling

- Suppose you are planning to drill for oil in a newly-discovered field.



## Oil drilling

- Suppose you are planning to drill for oil in a newly-discovered field.
- Setting up the drilling equipment costs \$1M.



## Oil drilling

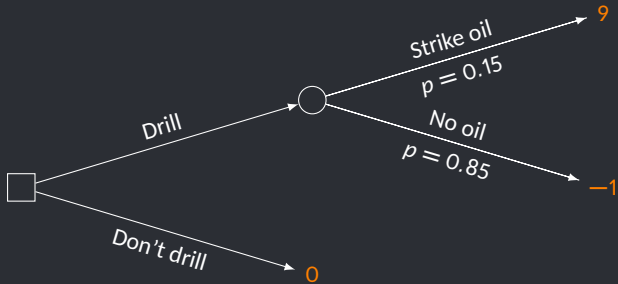
- Suppose you are planning to drill for oil in a newly-discovered field.
- Setting up the drilling equipment costs \$1M.
- If you strike oil, you will generate \$10M.

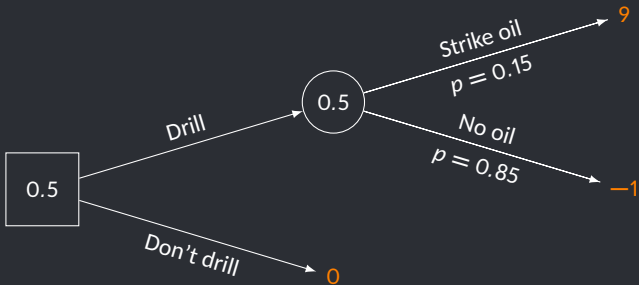


## Oil drilling

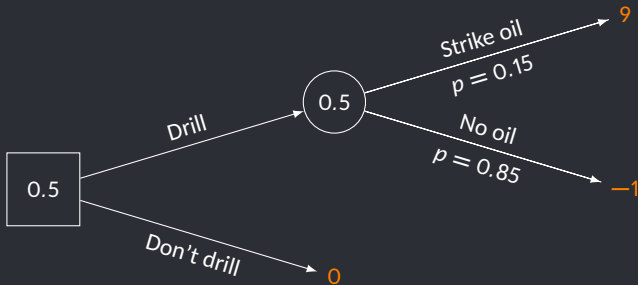
- Suppose you are planning to drill for oil in a newly-discovered field.
- Setting up the drilling equipment costs \$1M.
- If you strike oil, you will generate \$10M.
- But there's only a 15% chance that a random oil field has oil.











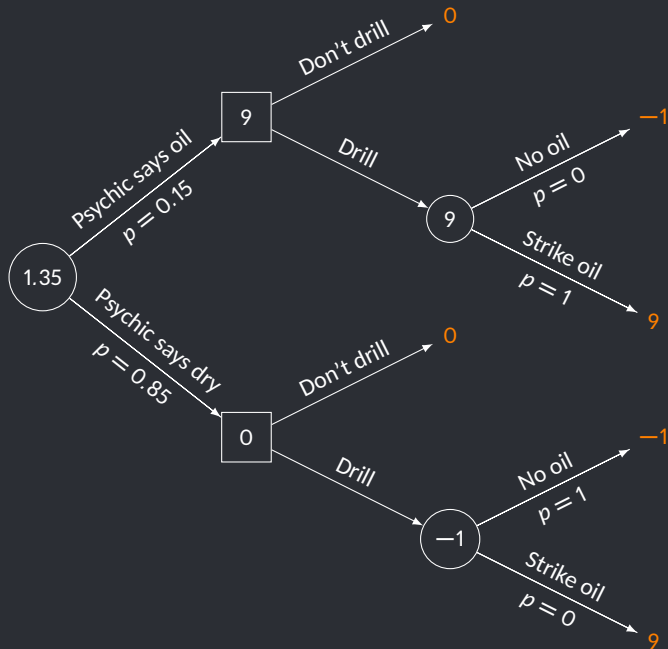
Without any information, we expect to make \$0.5M by deciding to drill.

## How much is perfect information worth?

- Suppose we have a psychic that gives us *perfect information* about whether the field has oil or not.

## How much is perfect information worth?

- Suppose we have a psychic that gives us *perfect information* about whether the field has oil or not.
- The psychic is always right!



- Suppose we commission a geological survey of the field; this is *imperfect information*.



- Suppose we commission a geological survey of the field; this is *imperfect information*.
- When a field is oil-rich, the survey will indicate this 60% of the time.



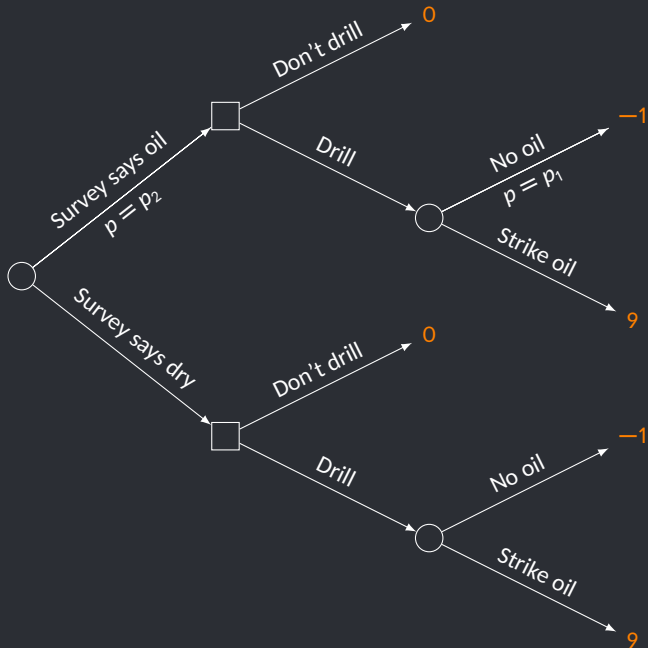
- Suppose we commission a geological survey of the field; this is *imperfect information*.
- When a field is oil-rich, the survey will indicate this 60% of the time.
- When a field is dry, the survey will indicate this 80% of the time.



- Suppose we commission a geological survey of the field; this is *imperfect information*.
- When a field is oil-rich, the survey will indicate this 60% of the time.
- When a field is dry, the survey will indicate this 80% of the time.
- How much should we be willing to pay for the survey?







Let's say  $O$  is the event that you would strike oil in the oilfield, and  $S$  is the event that the survey indicates there is oil present.

Let's say  $O$  is the event that you would strike oil in the oilfield, and  $S$  is the event that the survey indicates there is oil present.

There's only a 15% chance that a random oil field has oil.

$$P(O) = 0.15, \quad P(\overline{O}) = 0.85$$

Let's say  $O$  is the event that you would strike oil in the oilfield, and  $S$  is the event that the survey indicates there is oil present.

There's only a 15% chance that a random oil field has oil.

$$P(O) = 0.15, \quad P(\overline{O}) = 0.85$$

When a field is oil-rich, the survey will indicate this 60% of the time.

Let's say  $O$  is the event that you would strike oil in the oilfield, and  $S$  is the event that the survey indicates there is oil present.

There's only a 15% chance that a random oil field has oil.

$$P(O) = 0.15, \quad P(\overline{O}) = 0.85$$

When a field is oil-rich, the survey will indicate this 60% of the time.

$$P(S|O) = 0.6,$$

Let's say  $O$  is the event that you would strike oil in the oilfield, and  $S$  is the event that the survey indicates there is oil present.

There's only a 15% chance that a random oil field has oil.

$$P(O) = 0.15, \quad P(\overline{O}) = 0.85$$

When a field is oil-rich, the survey will indicate this 60% of the time.

$$P(S|O) = 0.6, \quad P(\overline{S}|O) = 0.4$$

Let's say  $O$  is the event that you would strike oil in the oilfield, and  $S$  is the event that the survey indicates there is oil present.

There's only a 15% chance that a random oil field has oil.

$$P(O) = 0.15, \quad P(\overline{O}) = 0.85$$

When a field is oil-rich, the survey will indicate this 60% of the time.

$$P(S|O) = 0.6, \quad P(\overline{S}|O) = 0.4$$

When a field is dry, the survey will indicate this 80% of the time.

Let's say  $O$  is the event that you would strike oil in the oilfield, and  $S$  is the event that the survey indicates there is oil present.

There's only a 15% chance that a random oil field has oil.

$$P(O) = 0.15, \quad P(\overline{O}) = 0.85$$

When a field is oil-rich, the survey will indicate this 60% of the time.

$$P(S|O) = 0.6, \quad P(\overline{S}|O) = 0.4$$

When a field is dry, the survey will indicate this 80% of the time.

$$P(S|\overline{O}) = 0.2,$$



Let's say  $O$  is the event that you would strike oil in the oilfield, and  $S$  is the event that the survey indicates there is oil present.

There's only a 15% chance that a random oil field has oil.

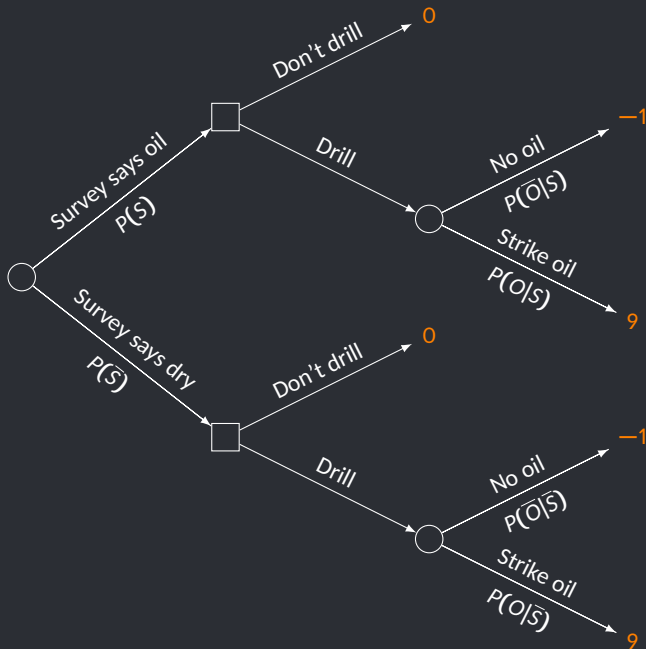
$$P(O) = 0.15, \quad P(\overline{O}) = 0.85$$

When a field is oil-rich, the survey will indicate this 60% of the time.

$$P(S|O) = 0.6, \quad P(\overline{S}|O) = 0.4$$

When a field is dry, the survey will indicate this 80% of the time.

$$P(S|\overline{O}) = 0.2, \quad P(\overline{S}|\overline{O}) = 0.8$$



Let's calculate the required probabilities using Bayes' rule:

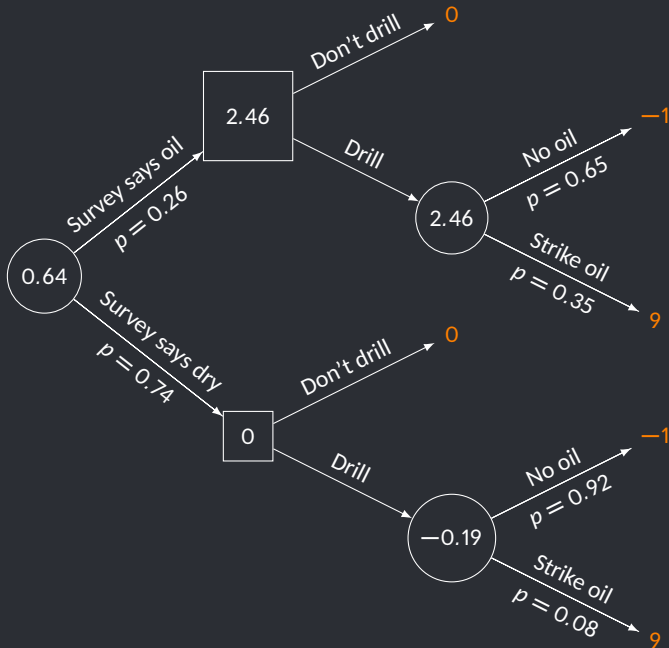
$$\begin{aligned}P(O|S) &= \frac{P(S|O)P(O)}{P(S|O)P(O) + P(S|\bar{O})P(\bar{O})} \\&= \frac{(0.6)(0.15)}{(0.6)(0.15) + (0.2)(0.85)} \\&= 0.35\end{aligned}$$

$$\begin{aligned}P(\bar{O}|S) &= \frac{P(S|\bar{O})P(\bar{O})}{P(S|\bar{O})P(\bar{O}) + P(S|O)P(O)} \\&= \frac{(0.2)(0.85)}{(0.2)(0.85) + (0.6)(0.15)} \\&= 0.65\end{aligned}$$



We also need to know the probability that the survey will indicate oil:

$$\begin{aligned}P(S) &= P(S \text{ and } O) + P(S \text{ and } \bar{O}) \\&= P(S|O)P(O) + P(S|\bar{O})P(\bar{O}) \\&= (0.6)(0.15) + (0.2)(0.85) \\&= 0.26\end{aligned}$$

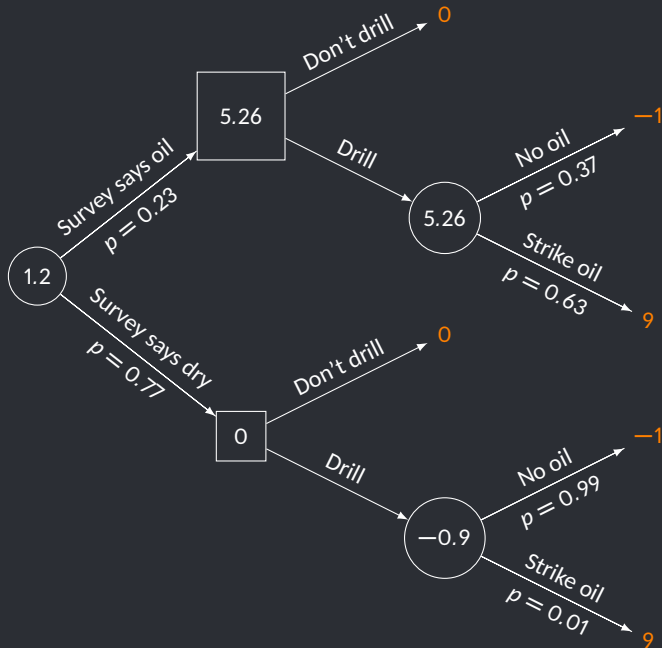


- The information from the survey is worth the difference between these two trees ( $0.64 - 0.5 = \$0.14\text{M}$ ); this is EVSI (Expected Value of Sample Information)



- The information from the survey is worth the difference between these two trees ( $0.64 - 0.5 = \$0.14\text{M}$ ); this is EVSI (Expected Value of Sample Information)
- Suppose we had a better survey, which correctly identifies oil-rich fields 95% of the time, and correctly identifies dry fields 90% of the time.







## Summary

- With no information, the expected value of the tree was \$0.5M.

## Summary

- With no information, the expected value of the tree was \$0.5M.
- With perfect information (the psychic), the expected value of the tree was \$1.35M. The EVPI (expected value of perfect information) is  $1.35 - 0.5 = \$0.85\text{M}$ .

## Summary

- With no information, the expected value of the tree was \$0.5M.
- With perfect information (the psychic), the expected value of the tree was \$1.35M. The EVPI (expected value of perfect information) is  $1.35 - 0.5 = \$0.85\text{M}$ .
  - No survey (or any other kind of information about this oil field) could ever be worth more than \$0.85M.

## Summary

- With no information, the expected value of the tree was \$0.5M.
- With perfect information (the psychic), the expected value of the tree was \$1.35M. The EVPI (expected value of perfect information) is  $1.35 - 0.5 = \$0.85\text{M}$ .
  - No survey (or any other kind of information about this oil field) could ever be worth more than \$0.85M.
- With imperfect information (the better survey), the expected value of the tree was \$1.2M. The EVSI (expected value of sample information) is  $1.2 - 0.5 = \$0.7\text{M}$ .

## Summary

- With no information, the expected value of the tree was \$0.5M.
- With perfect information (the psychic), the expected value of the tree was \$1.35M. The EVPI (expected value of perfect information) is  $1.35 - 0.5 = \$0.85\text{M}$ .
  - No survey (or any other kind of information about this oil field) could ever be worth more than \$0.85M.
- With imperfect information (the better survey), the expected value of the tree was \$1.2M. The EVSI (expected value of sample information) is  $1.2 - 0.5 = \$0.7\text{M}$ .
  - It's worth paying up to \$0.7M for this particular survey.