Lecture 11b: Decision Theory

CSCI 360

Introduction to Artificial Intelligence

USC

Here is where we are…

	3/1		Project 2 Out	
9	3/4 3/6	3/5 3/7	Quantifying Uncertainty Bayesian Networks	[Ch 13.1-13.6] [Ch 14.1-14.2]
10	3/11 3/13	3/12 3/14	*(spring break, no class)* *(spring break, no class)*	
11	3/18 3/20	3/19 3/21	Inference in Bayesian Networks Decision Theory	[Ch 14.3-14.4] [Ch 16.1-16.3 and 16.5]
	3/23		Project 2 Due	
12	3/25 3/27	3/26 3/28	***Advanced topics*** *(Chao traveling to National Science Foundation)* ***Advanced topics*** *(Chao traveling to National Science Foundation)*	
	3/29		Homework 2 Out	
13	4/1 4/3	4/2 4/4	Markov Decision Processes Decision Tree Learning	[Ch 17.1-17.2] [Ch 18.1-18.3]
	4/5 4/5		Homework 2 Due Project 3 Out	
14	4/8 4/10	4/9 4/11	Perceptron Learning Neural Network Learning	[Ch 18.7.1-18.7.2] [Ch 18.7.3-18.7.4]
15	4/15 4/17	4/16 4/18	Statistical Learning Reinforcement Learning	[Ch 20.2.1-20.2.2] [Ch 21.1-21.2]
16	4/22 4/24	4/23 4/25	Artificial Intelligence Ethics Wrap-Up and Final Review	
	4/26		Project 3 Due	
	5/3	5/2	**Final Exam** (2pm-4pm)	

Outline

- What is AI?
- Part I: Search
- Part II: Logical reasoning
- **Part III: Probabilistic reasoning**
  - Quantifying Uncertainty
  - Bayesian Networks
  - Inference in Bayesian Networks
  - **Decision Theory**
  - Markov Decision Processes
- Part IV: Machine learning

Recap: *Making rational decisions*

Rational decision depends on

(1) The relative importance of various goals and

(2) likelihood that (and degree to which) they will be reached

Decision theory = Utility theory + Probability theory

Choose an action that yields the ***maximum expected utility (MEU)***, averaged over all the possible outcomes of the action, weighted by the probability

Recap: *Deterministic environment*

- In a **deterministic** environment, given the current state ***s***, compute next state ***s'***
  - ***$s' = RESULT(s_0, a)$***

- What about a **non-deterministic** environment where the agent does not even know the current state?
  - ***RESULT(a)*** is random, meaning that the outcome is a <u>lottery</u>

Recap: *Deterministic environment*

- In a **deterministic** environment, given the current state ***s***, compute next state ***s'***
  - ***$s' = RESULT(s_0, a)$***

- What about a **non-deterministic** environment where the agent does not even know the current state?
  - ***RESULT(a)*** is random, meaning the outcome is a lottery

The probability of each outcome (***s'***), given evidence (***e***) and action (***a***), is

$$\boxed{P(\textsc{Result}(a) = s' \mid a, \mathbf{e})}$$

Example

- *Taking a course with* ***Professor X***

getting B

deterministic

- *Taking a course with* ***Professor Y***

getting A

getting C

non-deterministic

Example

- *Taking a course with* ***Professor X***

getting B(3)

- *Taking a course with* ***Professor Y***

getting A(4) getting C(2)

✓ *Professor Y (seeking the best case – 4.0)*
✓ *Professor X (avoiding the worst case – 2.0)*

Which one do **you** choose?

Example

- *Taking a course with* ***Professor X***

getting B(3)

(100%)

3*100% = 3

- *Taking a course with* ***Professor Y***

getting A(4) getting C(2)

(50%) (50%)

4*50% + 2*50% = 3

SHOULD

Which one ~~do~~ you choose?

Example

- *Taking a course with* ***Professor X***

getting B(3)

(100%)

3*100% = 3

- *Taking a course with* ***Professor Y***

getting A(4) getting C(2)

(75%) (25%)

4*75% + 2*25% = 3.5

SHOULD

Which one ~~do~~ you choose?

Example

- *Taking a course with* ***Professor X***

getting B(3)

(100%)

3*100% = 3

- *Taking a course with* ***Professor Y***

getting A(4) getting C(2)

(25%) (75%)

4*25% + 2*75% = 2.5

SHOULD

Which one ~~do~~ you choose?

Maximum expected utility (MEU)

- In the previous example, we have applied the principle of maximum expected utility (MEU)
- Utility function, ***U(s'),*** expresses the desirability of state ***s'***
- Example:

$$\mathbf{s'} = \{\ getting\ A,\ getting\ C\ \}$$

$\mathbf{U}(\ getting\ A\ ) = 4.0$

$\mathbf{U}(\ getting\ C\ ) = 2.0$

Maximum expected utility (MEU)

- Utility function, ***U(s')***, expresses the desirability of state ***s'***

- Expected utility, ***EU(a|e)***, is the weighted average

$$EU(a|\mathbf{e}) = \sum_{s'} P(\textsc{Result}(a) = s' \mid a, \mathbf{e})\, U(s')$$

*a: with professor **Y***   *e: taking the course*   ***s'*** = { *getting A*, *getting C* }

Example:

*P(**RESULT**(a) = getting A | a, e) * U(getting A) = 75% * 4.0*

*P(**RESULT**(a) = getting C | a, e) * U(getting C) = 25% * 2.0*

EU(a|e) = 75%*4.0 + 25%*2.0

Maximum expected utility (MEU)

- Choosing an action that maximizes the expected utility

$$\boxed{action = \underset{a}{\operatorname{argmax}}\, EU(a|\mathbf{e})}$$

- This principle defines all of AI

Maximum expected utility (MEU)

- Choosing an action that maximizes the expected utility

$$\boxed{action = \operatorname*{argmax}_a EU(a|\mathbf{e})}$$

This principle defines all of AI

$$\boxed{EU(a|\mathbf{e}) = \sum_{s'} P(\textsc{Result}(a) = s' \mid a, \mathbf{e})\, U(s')}$$

Estimating the state s′ requires perception, learning, knowledge representation, and inference…

Computing the probability requires Bayesian networks…

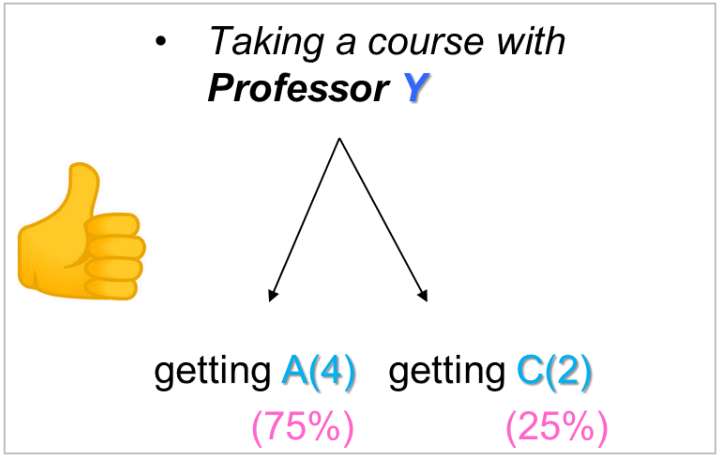
Computing the utility requires search or planning (to estimate how good s′ is in terms of reaching the goal)

Outline of today's lecture

- **Maximum expected utility (MEU)**
- **Axioms of Utility Theory**
  - E.g., why is the principle of MEU rational?
  - Characteristics of rational preferences
  - Human irrationality
- **Decision making**

Axioms of utility theory

- **Lottery**: the set of *probability-weighted* outcomes for each action $L = [\, p_1, S_1;\ \ p_2, S_2;\ \ \ldots ;\ p_n, S_n \,]$
  - S_i denotes an outcome
  - p_i denotes the probability of S_i



Example

p_1 = 75%
S_1 = getting A

p_2 = 25%
S_2 = getting C

Axioms of utility theory

- **Lottery**: the set of *probability-weighted* outcomes for each action $L = [p_1, S_1;\ p_2, S_2;\ \ldots;\ p_n, S_n]$

- Constraints that any **rational preference** must obey:
  - Orderability

    Exactly one of $(A \succ B)$, $(B \succ A)$, or $(A \sim B)$ holds.
  - Transitivity

    $$(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)\ .$$
  - Continuity

    $$A \succ B \succ C \Rightarrow \exists p\ [p, A;\ 1-p, C] \sim B\ .$$
  - Substitutability

    $$A \sim B \Rightarrow [p, A;\ 1-p, C] \sim [p, B; 1-p, C]\ .$$
  - Monotonicity

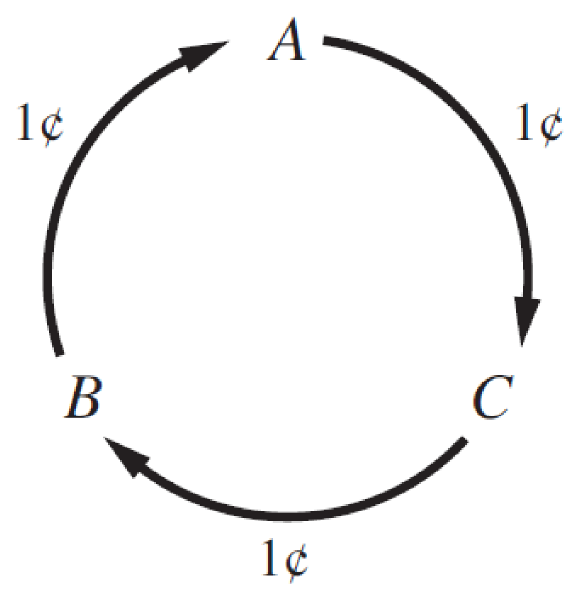
    $$A \succ B \Rightarrow (p > q \Leftrightarrow [p, A;\ 1-p, B] \succ [q, A;\ 1-q, B])\ .$$
  - Decomposability

    $$[p, A;\ 1-p, [q, B;\ 1-q, C]] \sim [p, A;\ (1-p)q, B;\ (1-p)(1-q), C]\ .$$

Self-evident irrationality

- For example, an agent with “intransitivity” preference can be induced to give away all its money

- If *B>C* , then an agent who has *C* would pay 1 cent to get *B*
- If *A>B* , then an agent who has *B* would pay 1 cent to get *A*
- If *C>A* , then an agent who has *A* would pay 1 cent to get *C*



Theorem *[von Neumann and Morgenstern, 1944]*

- If an agent's preferences obey the axioms of utility, then
  - there exists a function U(s') that captures the preferences; and

$$
\begin{aligned}
U(A) > U(B) &\Leftrightarrow A \succ B \\
U(A) = U(B) &\Leftrightarrow A \sim B
\end{aligned}
$$

  - the utility of a lottery is the sum of the probability of each outcome times the utility of that outcome

$$U([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i U(S_i)$$

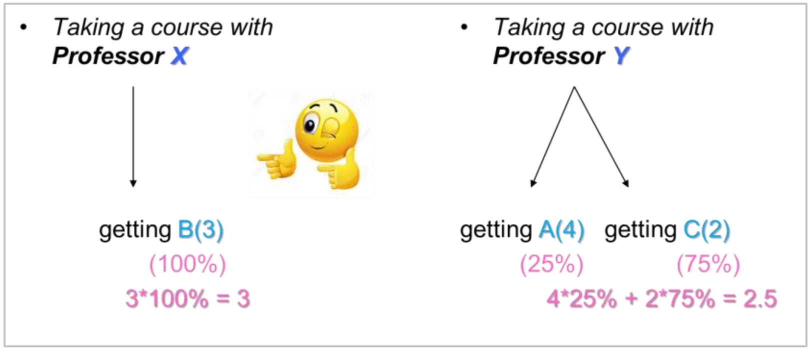
Theorem *[von Neumann and Morgenstern, 1944]*

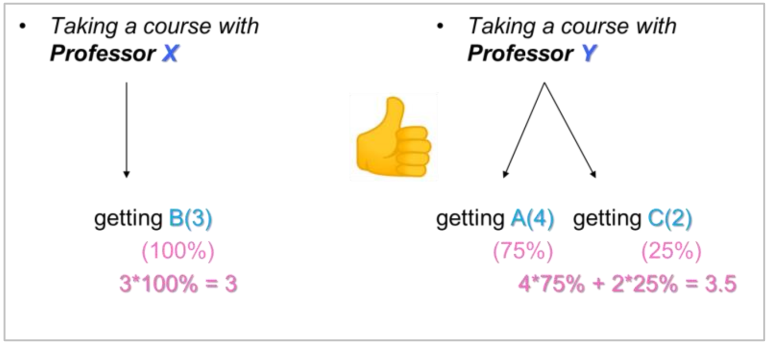
- If an agent's preferences obey the axioms of utility, then
  - there exists a function U(s') that captures the preferences; and

$$U(A) > U(B) \Leftrightarrow A \succ B$$
$$U(A) = U(B) \Leftrightarrow A \sim B$$

  - the utility of a lottery is the sum of the probability of each outcome times the utility of that outcome

$$U([p_1, S_1; \ldots ; p_n, S_n]) = \sum_i p_i U(S_i)$$





Utility function is not unique

- For example, an agent's behavior would not change if its utility function ***U(s)*** were transformed "linearly"

$$\boxed{U'(S) = aU(S) + b}$$

where a and b are constants and $a > 0$.

This is called "affine" transformation.

Question for you (1)

- Given two options
  - (A). You will get $240 for sure
  - (B). You will get $1000 with 25% chance, and nothing with 75%

You might want to choose (A)

U(B) = U(1000*25% + 0*75%) = U($250)

U(A) = U($240)

I get it – you **don't** care about money

Question for you (2)

- Given two options
  - (C). You will lose $760 for sure
  - (D). You will lose $1000 with 75% chance, and nothing with 25%

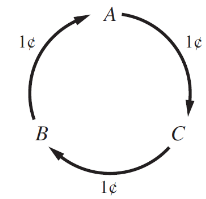
You might want to choose (D)

U(D) = U(-1000*75% + 0*25%) = U(-$750)

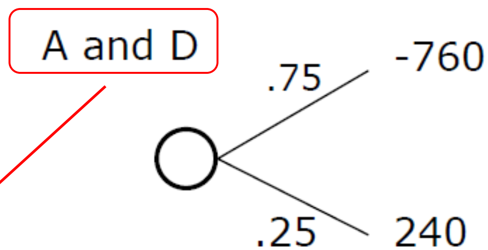
U(C) = U(-$760)

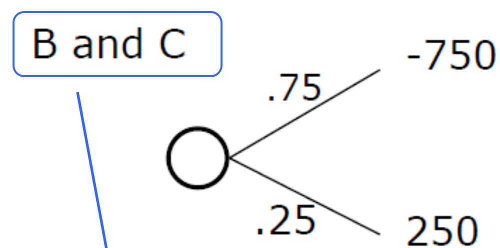
Wait a second – you **do** care about money

Human irrationality



- Most people prefer ***A in Question 1*** and ***D in Question 2***





Net loss, given both questions:

240*0.25 – 760*0.75 = -$510

Net loss, given both questions:

250*0.25 – 750*0.75 = - $500

Outline of today's lecture

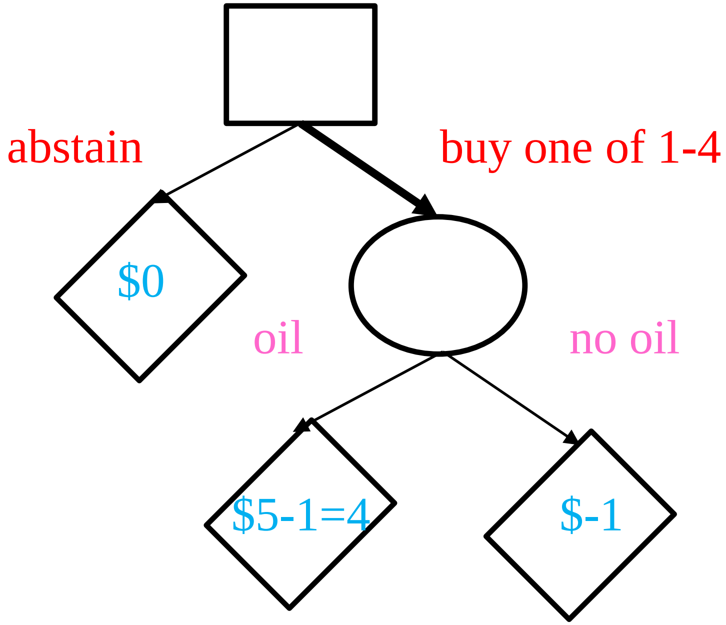
- **Maximum expected utility (MEU)**
- **Axioms of Utility Theory**
  - E.g., why is the principle of MEU rational?
  - Characteristics of rational preferences
  - Human irrationality
- **Decision making**

Example: oil drilling

- Situation:
  - You can buy one of 4 indistinguishable blocks of ocean drilling rights (numbered 1-4), exactly one of which contains oil worth 5 (million) dollars.
  - The price of each block is 1 (million) dollars.
- Question:
  - Should you buy one or should you abstain from the purchase?

Example: oil drilling

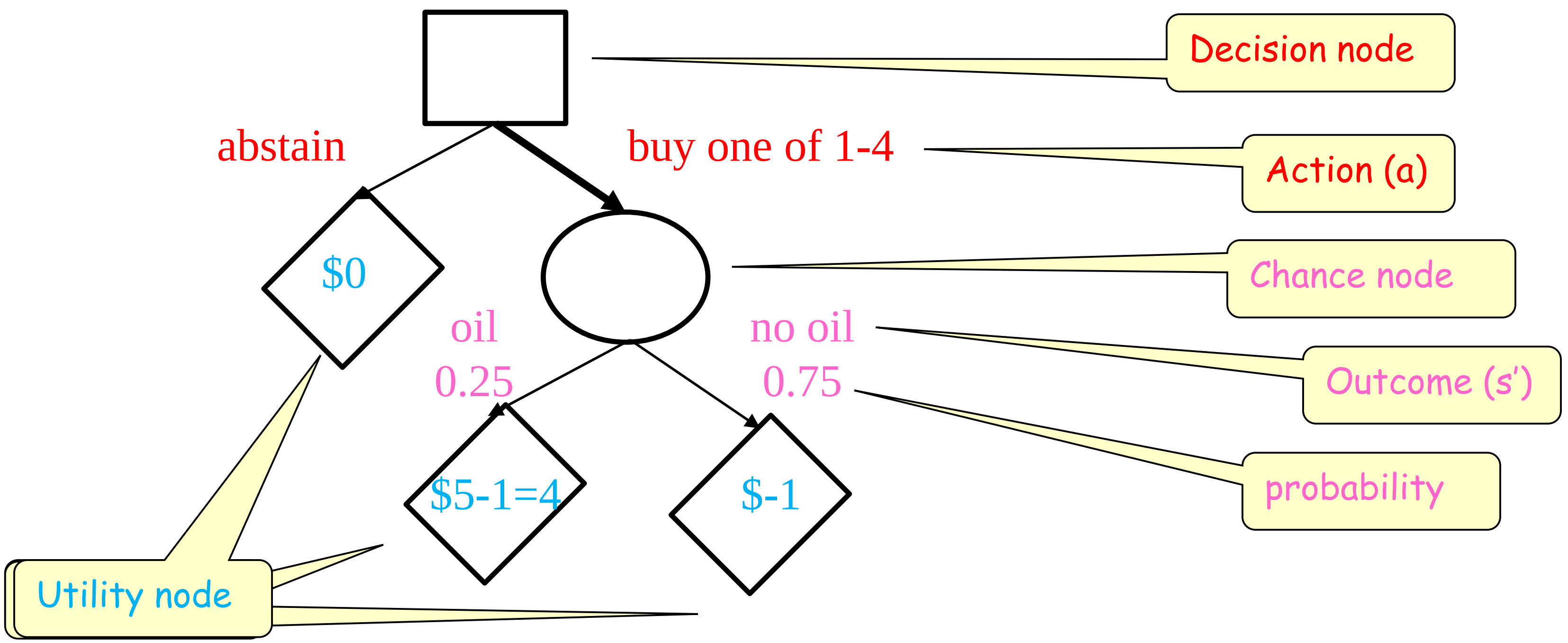
- Situation:
  - You can buy one of 4 indistinguishable blocks of ocean drilling rights (numbered 1-4), exactly one of which contains oil worth 5 (million) dollars. The price of each block is 1 (million) dollars.
- Question:
  - Should you buy one or should you abstain from the purchase?



Human decision makers are often **risk-averse** in high-stake situations that do not repeat; they consider the worst-case more than average outcome.

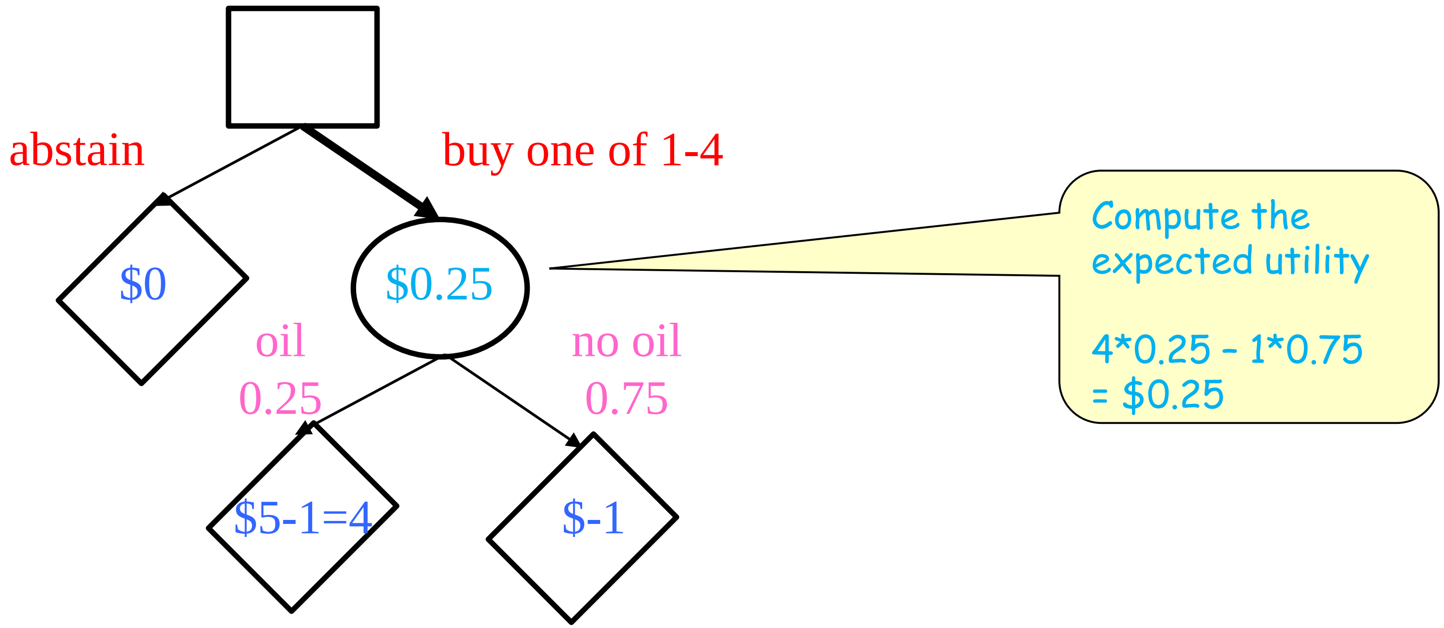
Example: oil drilling

- Situation:
  – You can buy one of 4 indistinguishable blocks of ocean drilling rights (numbered 1-4), exactly one of which contains oil worth 5 (million) dollars. The price of each block is 1 (million) dollars.
- Question:
  – Should you buy one or should you abstain from the purchase?



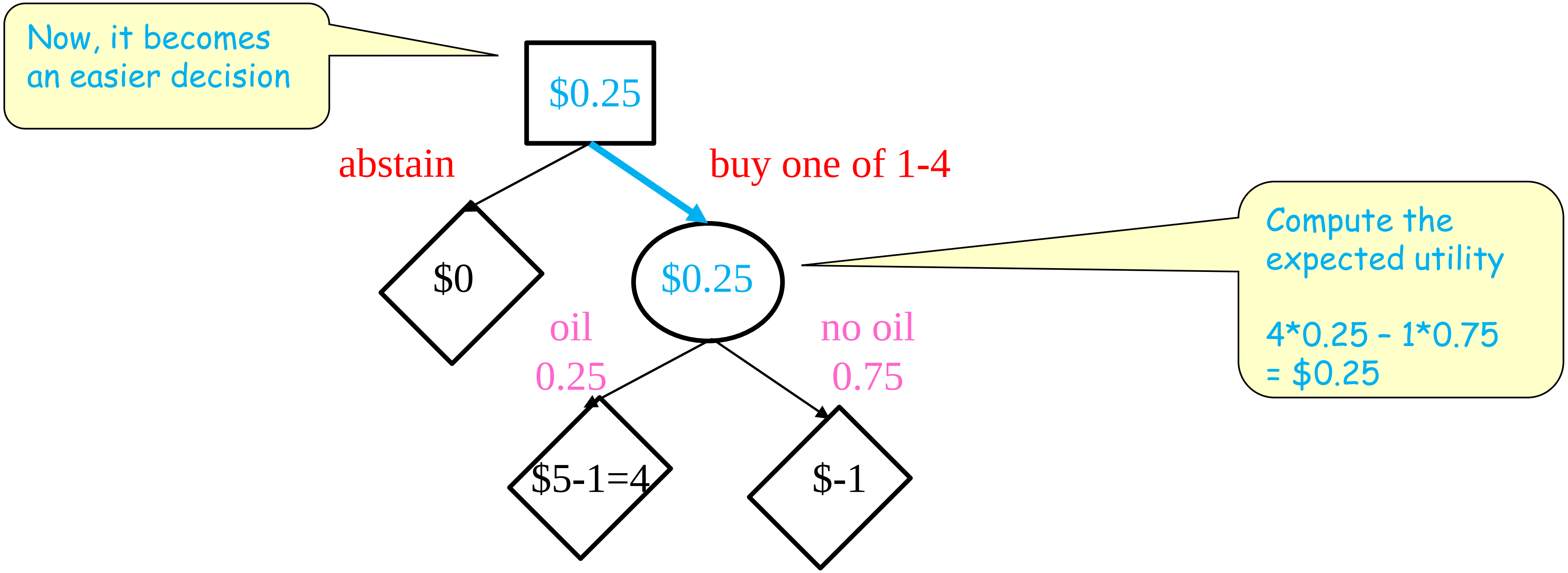
Example: oil drilling

- Situation:
  - You can buy one of 4 indistinguishable blocks of ocean drilling rights (numbered 1-4), exactly one of which contains oil worth 5 (million) dollars. The price of each block is 1 (million) dollars.
- Question:
  - Should you buy one or should you abstain from the purchase?



Example: oil drilling

- Situation:
  - You can buy one of 4 indistinguishable blocks of ocean drilling rights (numbered 1-4), exactly one of which contains oil worth 5 (million) dollars. The price of each block is 1 (million) dollars.
- Question:
  - Should you buy one or should you abstain from the purchase?

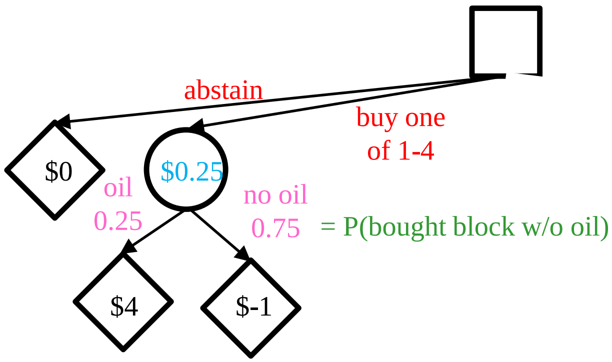


Example: oil drilling (more info)

- Situation:
  - A seismologist can survey Block 1 for you and find out definitively "whether it contains oil", for 1 (million) dollars.
- Question:
  - Should you buy the information?

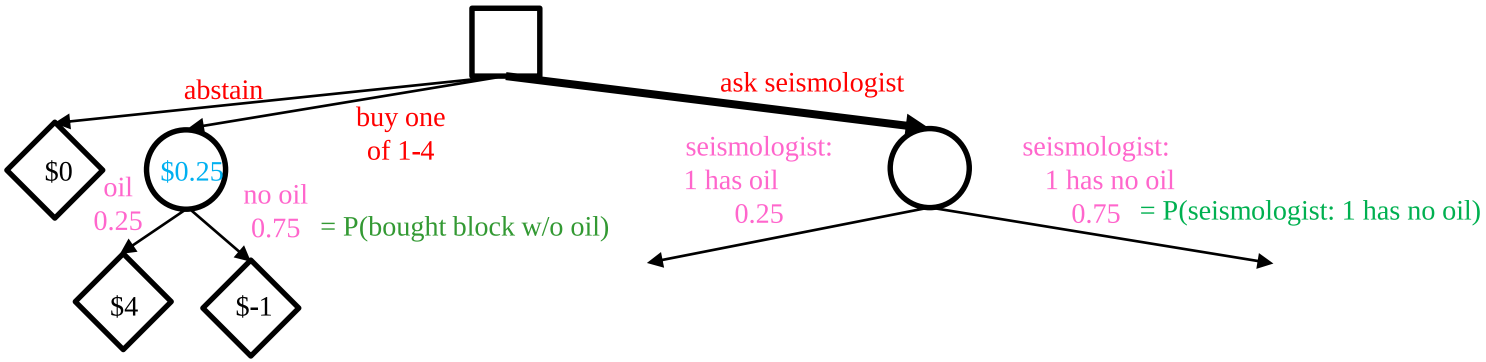
Example: oil drilling (more info)

- Situation:
  - A seismologist can survey Block 1 for you and find out definitively “whether it contains oil”, for 1 (million) dollars.
- Question:
  - Should you buy the information?



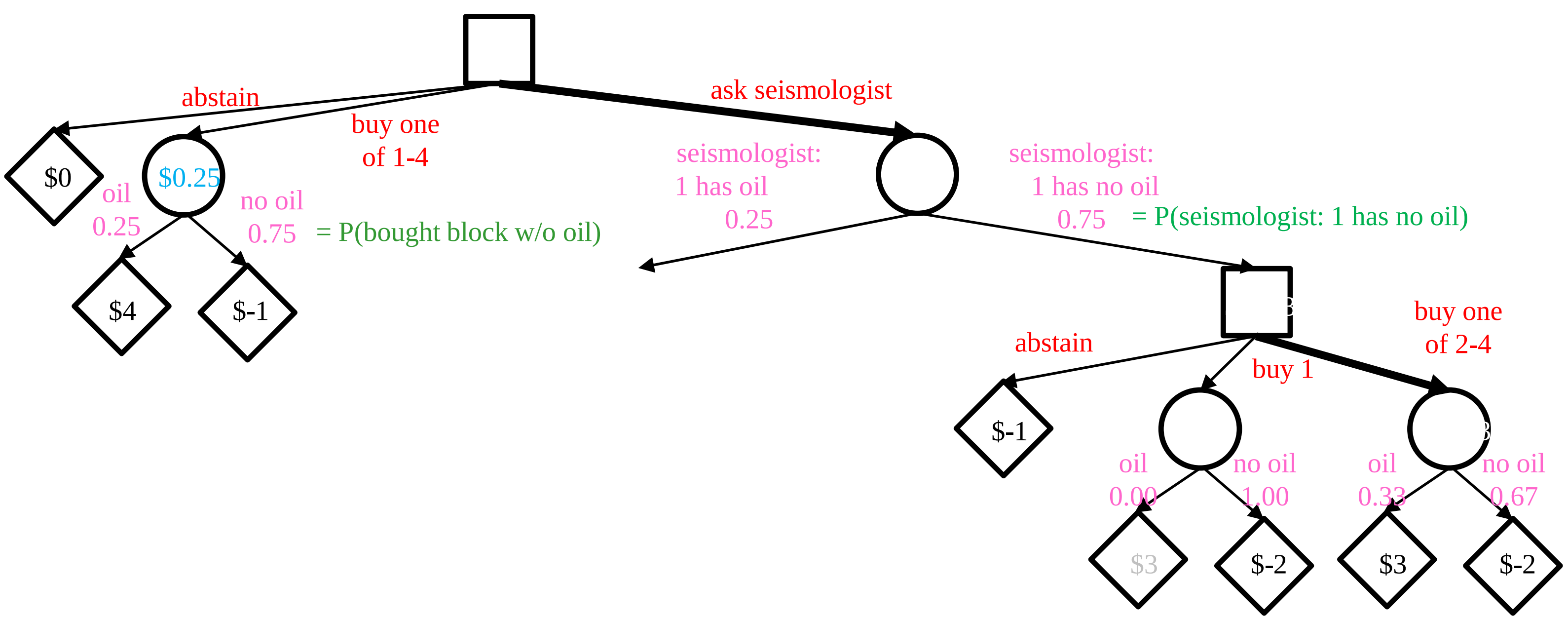
Example: oil drilling (more info)

- Situation:
  - A seismologist can survey Block 1 for you and find out definitively “whether it contains oil”, for 1 (million) dollars.
- Question:
  - Should you buy the information?



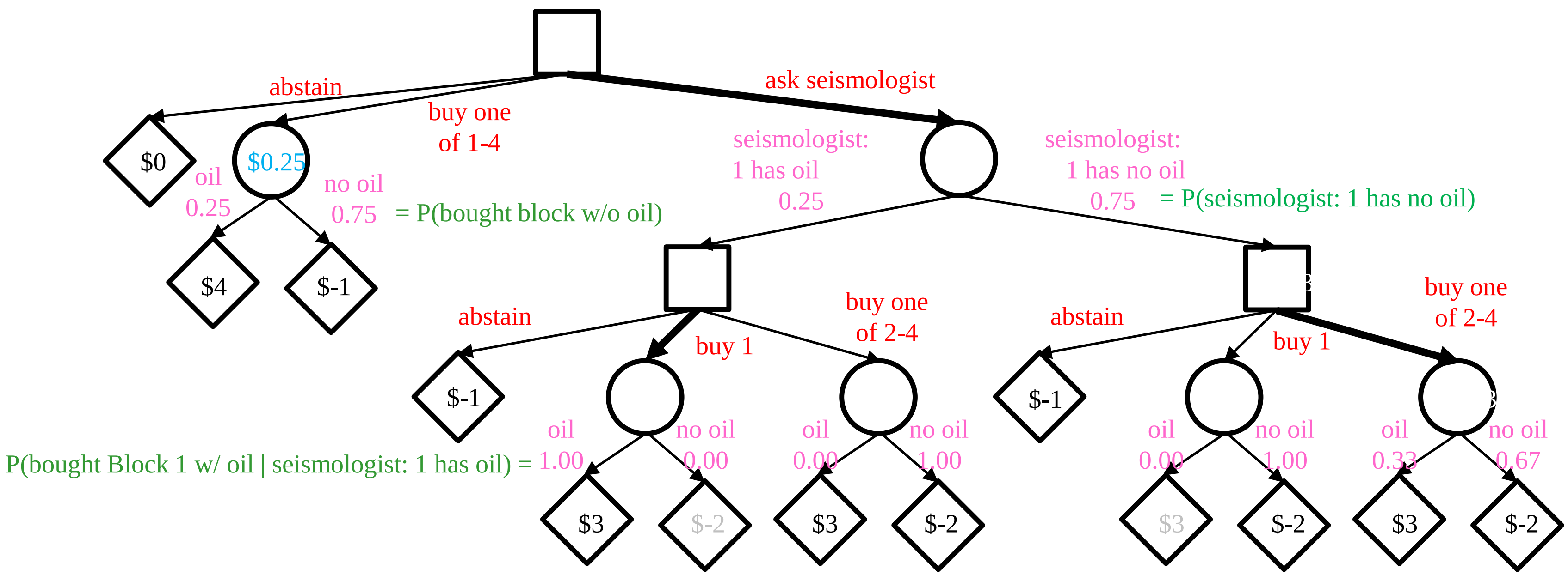
Example: oil drilling (more info)

- Situation:
  - A seismologist can survey Block 1 for you and find out definitively “whether it contains oil”, for 1 (million) dollars.
- Question:
  - Should you buy the information?



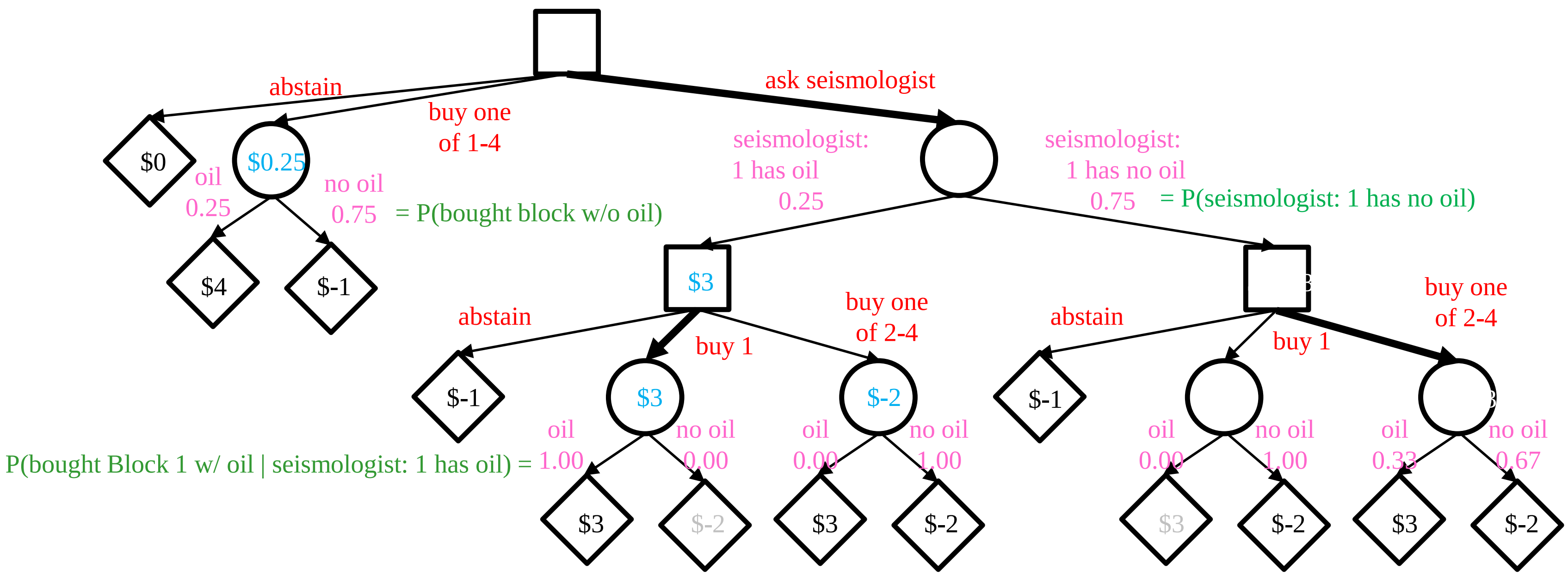
Example: oil drilling (more info)

- Situation:
  - A seismologist can survey Block 1 for you and find out definitively "whether it contains oil", for 1 (million) dollars.
- Question:
  - Should you buy the information?



Example: oil drilling (more info)

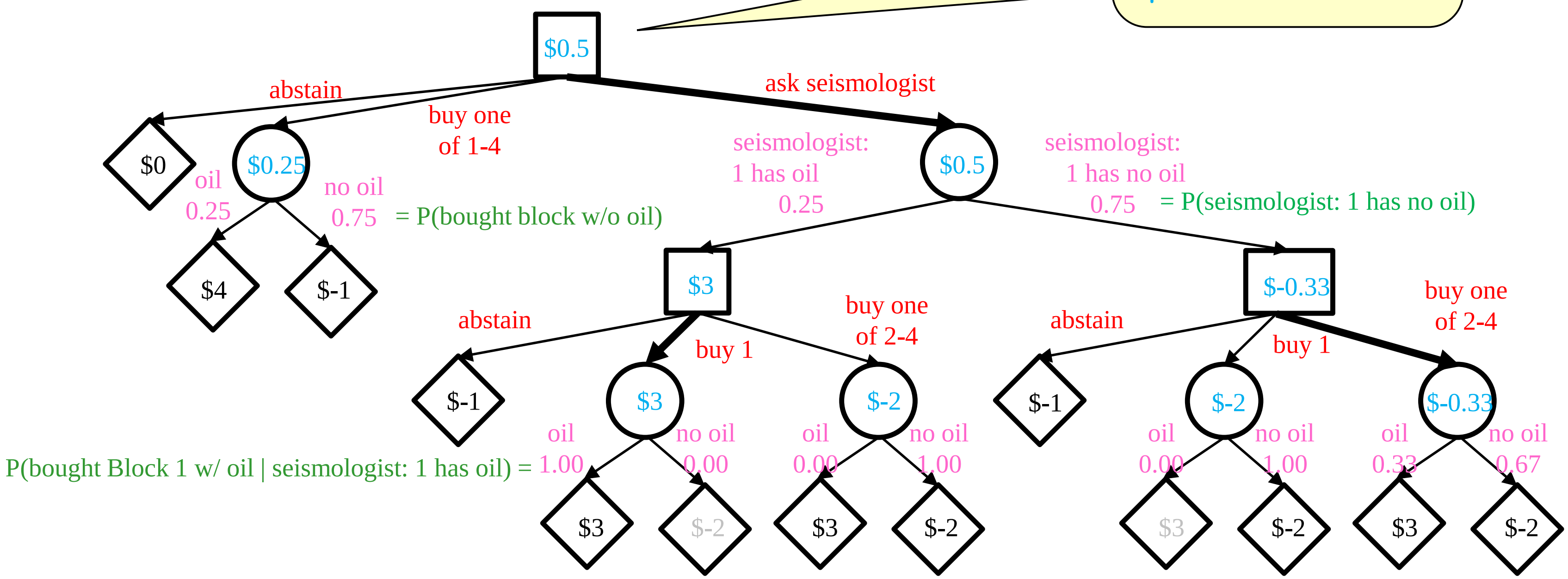
- Situation:
  - A seismologist can survey Block 1 for you and find out definitively “whether it contains oil”, for 1 (million) dollars.
- Question:
  - Should you buy the information?



Example: oil drilling (more info)

- Situation:
  - A seismologist can survey Block 1 for you and find out definitively "whether it contains oil", for 1 (million) dollars.
- Question:
  - Should you buy the information?

Expected utility has improved from $0.25 to $0.5



Example: oil drilling (more info)

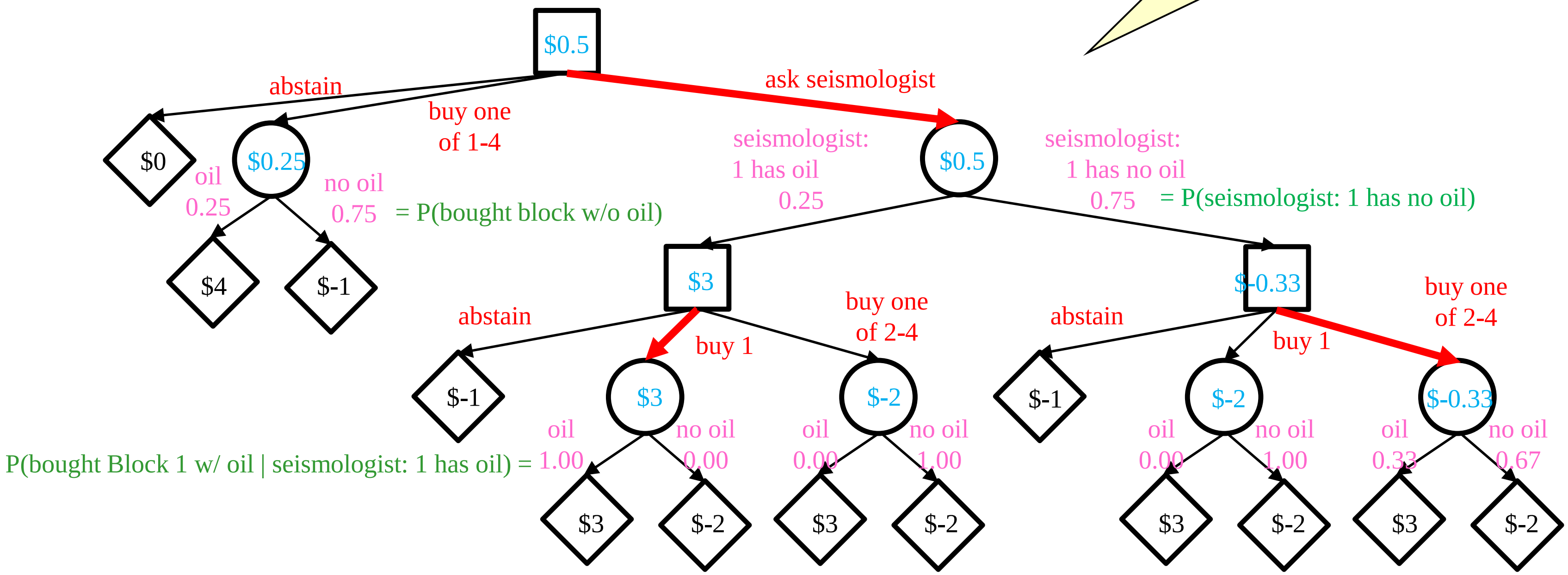
- **Answer:**
  - Yes, for an expected profit of 0.5 (million) dollars.
- **The plan:**
  - Ask the seismologist.
  - If the seismologist says that Block 1 contains oil, buy it.
  - If the seismologist says that Block 1 does not contain oil, buy one of the other blocks.

- this is a **conditional plan**, not a sequence of actions. In general, plans are no longer sequences of actions for probabilistic environment.

Example: oil drilling (more info)

- Situation:
  - A seismologist can survey Block 1 for your and find out definitively "whether it contains oil", for 1 (million) dollars.
- Question:
  - Should you buy the information?

The probabilistic plan



Example: oil drilling *(yet another question)*

- Situation:
  - The seismologist now wants <u>more money</u> for the information.
- Question:
  - <u>How much</u> should you pay at most for the information?

Information is money…

- Answer:
  - **1.25 (million) dollars**, which makes you indifferent between

  (a) buying one of the blocks unseen or

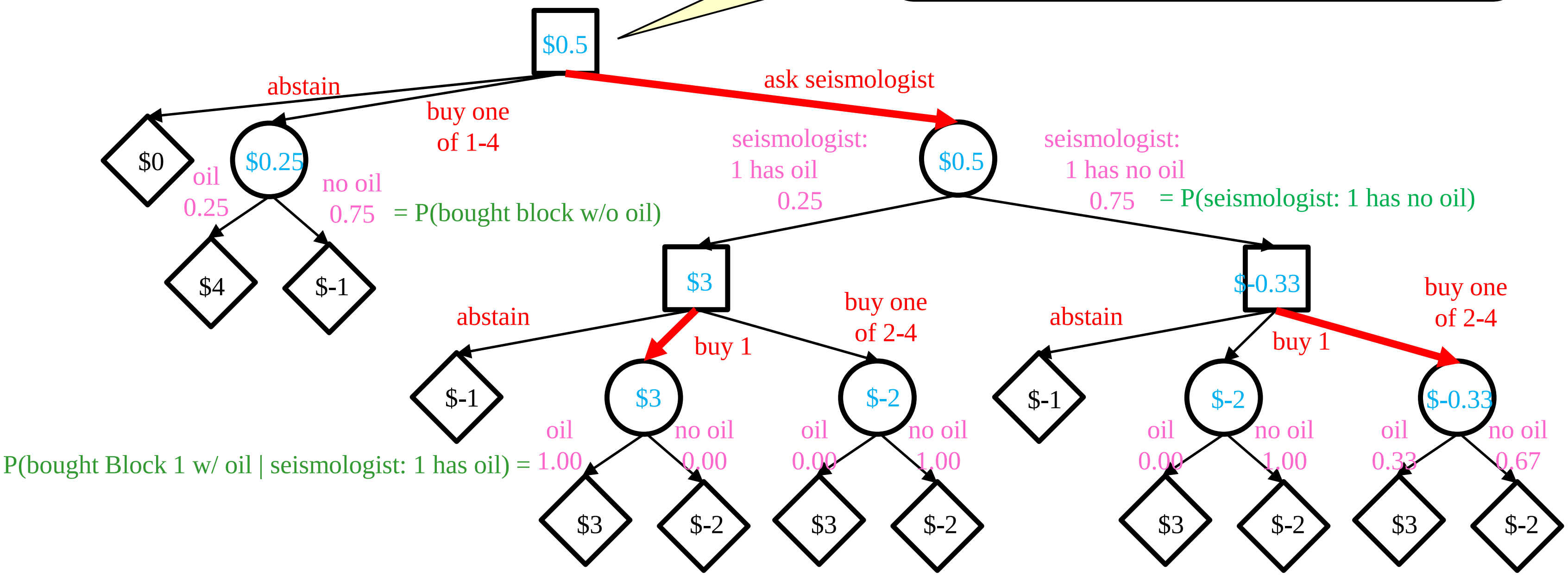
  (b) paying the seismologist for the information, for an expected profit of 0.25 (million) dollars in both cases.

Example: oil drilling *(yet another question)*

- Situation:
  - A seismologist can survey Block 1 for your and find out definitively "whether it contains oil", for 1 (million) dollars.
- Question:
  - Should you buy the information?

How much is the information worth?

$0.5 - $0.25
= $0.25



Summary

- **Basis of Utility Theory**
  - E.g., why is the principle of MEU rational?
  - Characteristics of rational preferences
  - Human irrationality
- **Decision making**