→ Maximize PROFIT
↳ Revenue − Cost.

↱ 100 Barrels of Oil.

At yr 1, 2 . . . will go till 10.

ACTIONS.
~~CHOICES~~ and associated aspects.

↳ 1. x_t → Barrel extracted + processed
↳ COST → x_t^2

2. Sell x^t at price p_t.

3. Exploration → will lead to discovery b^t.
↳ known value.

Discount factor = d.

* STATE vars:–

1. Barrel of Oil → 'oil_qt'
2. Year of exploration → y_i

→ y_i will go from ~~0~~ 1 to 10 both inclusive.

→ Initial oil_qt is 100 barrels.
↳ And should not go below 0.

* STOP CRITERION
No matter how much we extract, we cannot perform actions.
∴ Profit ('Yr=11') = 0

→ We cannot perform Extracting + selling actions if available oil_qt is 0

⇒ But we dont need to phrase this as a STOP CRITERION

⇒ This 0 can be included in the choices ~~tco~~ at each state from the combination map.

⇒ ⇒ At all combinations of the state variable, we need to EVALUATE profit (+ change in state variables) generated ~~for~~ from all possible choices.

↳ PROFIT is generated by performing sequence of actions.

** Input state = oil_qt, T

↳ Choice → Extract x_t quantity from the oil_qt

↳ x_t → 20

↳ $\leq$ oil_qt

* DYNAMICS.

⇒⇒ Cost associated with extraction of $x_t \rightarrow x_t^2$

Revenue → $x_t \times p_t$

PROFIT → $-x_t^2 + x_t \times p_t$

Discounted Profit → $\delta^t \times$ Profit

↳ $\frac{1}{d}$.

Output →

⇒⇒ barrel left → $oil\_q_t - x_t + b_t$

Explored.

time → T+1

⇒⇒ From stop criterion, we know x_t is 0 for $Yr \geq 11$.

→ For yr = 10 let oil qt = oil_qt.
Then x_t will be in range 0 to oil qt (both inclusive)

CHOICES

for all possible values of x_t

↳ At any x_t.

$Profit = -x_t^2 + x_t \times p_t + \delta \times \overline{Profit}$ [oil_qt − x_t, t+1]

$$\underset{x_t}{\text{Profit}} = -x_t^2 + x_t p_t + \frac{1}{d}\left(\text{Profit}\langle \text{oil\_q}_t + b_t - x_t, t+1 \rangle\right)$$

↳ We will have series of x_t

↳ We need to compute max (statistic) of this series, as well as state_value (x_t) that is genera-ting this max value

↳ And cache them for future (back-time) calculation.

Dynamic (Pseudo Bellman) Equation

$$\text{Profit}(O_t, t) = \max_{0 < x_t < O_t} \left[x_t p_t - x_t^2 + \frac{1}{d}\left[\text{Profit}(O_t - x_t + b_t, t+1)\right]\right]$$

→ We know $\text{Proift}(O_t, t = 11) = 0$

↳ This can help compute $\text{Profit}(O_t, t = 10)$ by Pivoting

Iterate till $\gamma_\theta = 1$