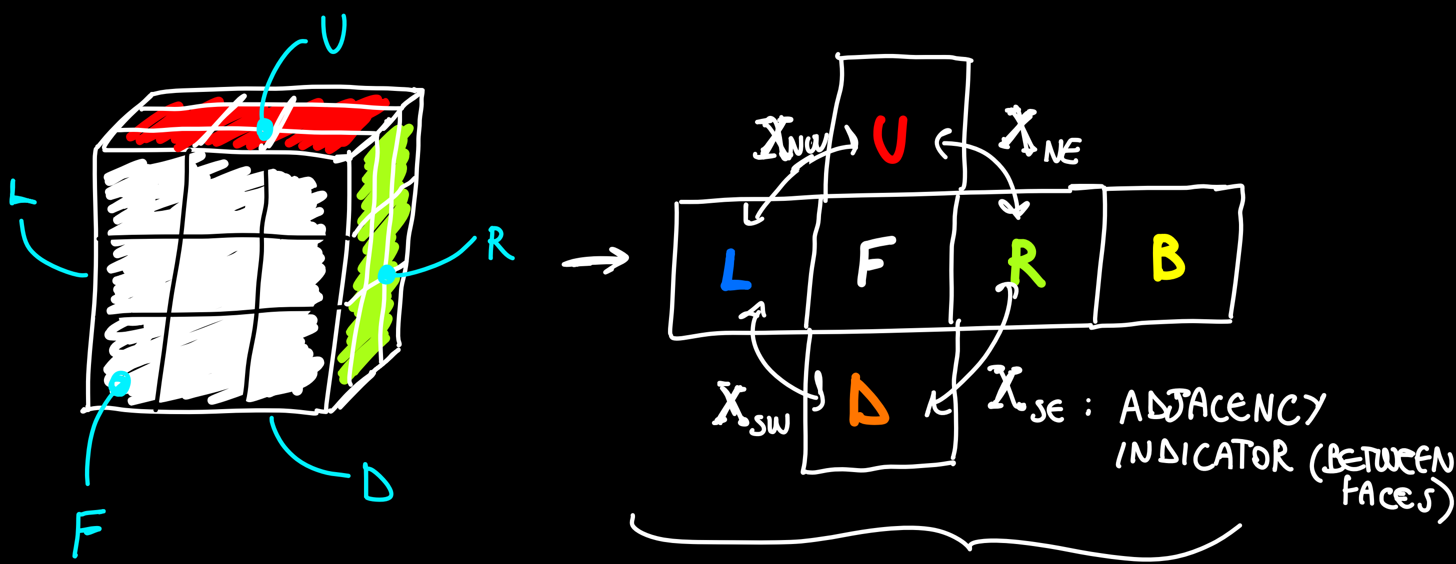
RUBIK'S MDP

"OPEN BOX" MODEL

FACES $\mathcal{Y} = \{F, B, U, \overset{f_1}{D}, L, R\} \rightarrow$ FIXED

COLORS $\mathcal{X} = \{W, Y, R, O, B, G\} \rightarrow$ ASSIGNED TO $\mathcal{Y}$



Send

STATE SPACE $S := \mathcal{Y}_t$, EACH $f_i^{(t)} \in \mathcal{Y}_t$ IS A 3x3 MAT

ACTIONS
- → SWITCH( ) Λ : MODIFIES POV (HORIZ., VERT.)
- → ROTATE( ) Γ : PERFORMS ROTATION ONLY FOR $f_1^t = F^t$ (CLOCKWISE, COUNTERCLOCKWISE)

$$\Lambda, \Gamma : \mathcal{Y}^{(t)} \mapsto \mathcal{Y}^{(t+1)}$$

Λ SWITCH POV. LET Λ_h HORIZ. SWITCH, Λ_v VERT. SWITCH DOESN'T MODIFY THE SINGLE MATRICES $f_i^{(t)}$ BUT REASSIGNES THEM TO DIFFERENT KEYS

EXAMPLE (f_i's NOT SHUFFLED YET)

$$\begin{pmatrix} & R & & \\ B & W & G & Y \\ & O & & \end{pmatrix}^{(t)} \xrightarrow{\Lambda_h} \begin{pmatrix} & R' & & \\ W & G & Y & B \\ & O' & & \end{pmatrix}^{(t+1)}$$

WE NEED TO TAKE PROSPECTIVE INTO ACCOUNT, HENCE

$$R' = \underbrace{T_{cw}\left(\frac{\pi}{2}\right)}_{\text{CLOCKWISE 3D ROTATION MATRIX}} R$$

$$O' = T_{ccw}\left(\frac{\pi}{2}\right) O$$

IN GENERAL

$$\Lambda_h\left[\begin{pmatrix} & f_5 & & \\ f_1 & f_2 & f_3 & f_4 \\ & f_6 & & \end{pmatrix}^{(t)}\right] = \begin{pmatrix} & f_5' & & \\ f_2 & f_3 & f_4 & f_1 \\ & f_6' & & \end{pmatrix}^{(t+1)}$$

WHERE $f_5' = T_{cw}\left(\frac{\pi}{2}\right) f_5$

$f_6' = T_{ccw}\left(\frac{\pi}{2}\right) f_6$

$$\Lambda_v\left[y^{(t)}\right] = \begin{pmatrix} & f_4'' & & \\ f_1' & f_5 & f_3' & f_6'' \\ & f_2 & & \end{pmatrix}$$

WHERE $f_1' = T_{cw}\left(\frac{\pi}{2}\right) f_1$ $\quad f_4'' = T(\pi) f_4$

$f_3' = T_{ccw}\left(\frac{\pi}{2}\right) f_3$ $\quad f_6'' = T(\pi) f_6$

T

MODIFIES THE CURRENT f_i IN F FRONT KEY

$$T_{cw}[y^{(t)}] = \begin{pmatrix} & f_5^* & & \\ f_1^* & f_2' & f_3^* & f_4 \\ & f_6^* & & \end{pmatrix}$$

WHERE

$$f_2' = T_{cw}\left(\frac{\pi}{2}\right) f_2$$

$$f_3^* = \underbrace{\overbrace{X_{NE}\,\mathcal{R}_{III}}^{\mathcal{C}_I}}_{\text{CHANGED COLUMN I WHICH TOOK VALUES OF THE ROW III OF THE ADJACENT } (X_f)\ f} \oplus \underbrace{\mathcal{C}_{II} \oplus \mathcal{C}_{III}}_{\text{SAME COLUMNS II AN III}}$$

WE REASON WITH COLUMNS FOR PLANES AND ROWS FOR PLANES

$$f_6^* = \underbrace{X_{SE}\,\mathcal{C}_I}_{\mathcal{R}_I} \oplus \mathcal{R}_{II} \oplus \mathcal{R}_{III}$$

$$f_1^* = \mathcal{C}_I \oplus \mathcal{C}_{II} \oplus \overbrace{X_{SW}\,\mathcal{R}_I}^{\mathcal{C}_{III}}$$

$$f_5^* = \mathcal{R}_I \oplus \mathcal{R}_{II} \oplus \overbrace{X_{NW}\,\mathcal{C}_{III}}^{\mathcal{R}_{III}}$$

IN GENERAL

$$T_d[y^{(t)}] = \{f_2', f_4, f_5^*, f_6^*, f_1^*, f_3^*\}$$

$$f_2' = T_{cw}\left(\frac{\pi}{2}\right)$$

$$f_i^* = \bigoplus_{*j=1}^{2} \mathcal{H}_{j*} \oplus X_* \mathcal{H}_* \qquad (\mathcal{H} = \mathcal{C} \wedge \mathcal{R})$$

SAME REASONING APPLIES TO Γ'_{ccw}

HENCE

$$S := \mathcal{Y}^{(t)} = \{ f_i^{(t)} \}_{i=1}^{6}$$

$$A := \{ \underbrace{\Lambda_h}_{a}, \Lambda_v, \Gamma_{cw}, \Gamma_{ccw} \}$$

$$P : S \times A \times S' \mapsto [0,1]$$ IS DETERMINISTIC

$$P(s' | s, a) = 1 \quad \forall s, a, s'$$

$$\gamma = [0,1]$$

$$R : S \times A \longrightarrow \mathbb{R}$$

R REWARD FUNCTION IS YET TO BE STUDIED, BUT COULD TAKE SOME APPROACH BASED ON DISTANCE FROM S_{end} SOLUTION STATE. FOR NOW LET'S DENOTE AS $\| \cdot \|_R$ A DISTANCE BASED ON SOME RUBIK'S CUBE METRIC

$$R(s,a) = \| s' \|_R$$

AND WE HAVE OUR MDP

$$\mathcal{M}_R = \langle S, A, R, P, \gamma \rangle$$