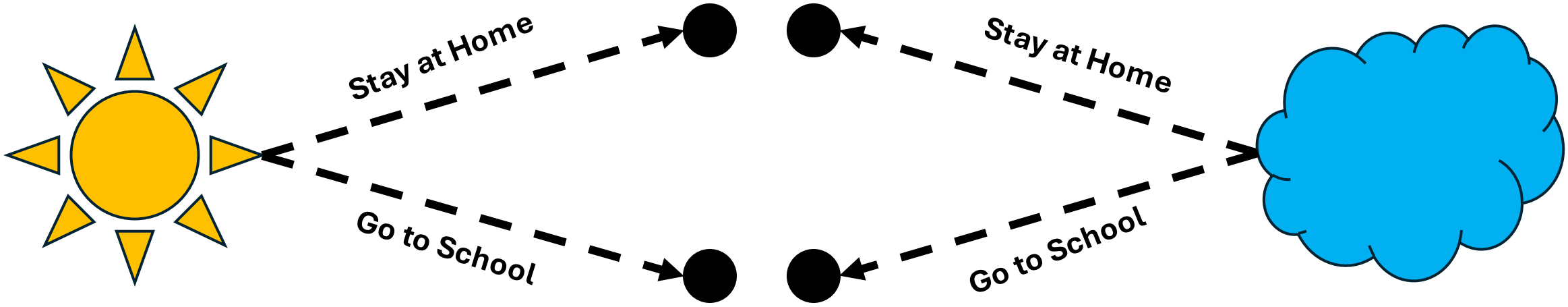
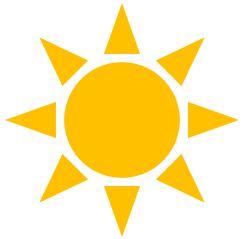
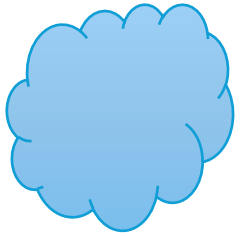
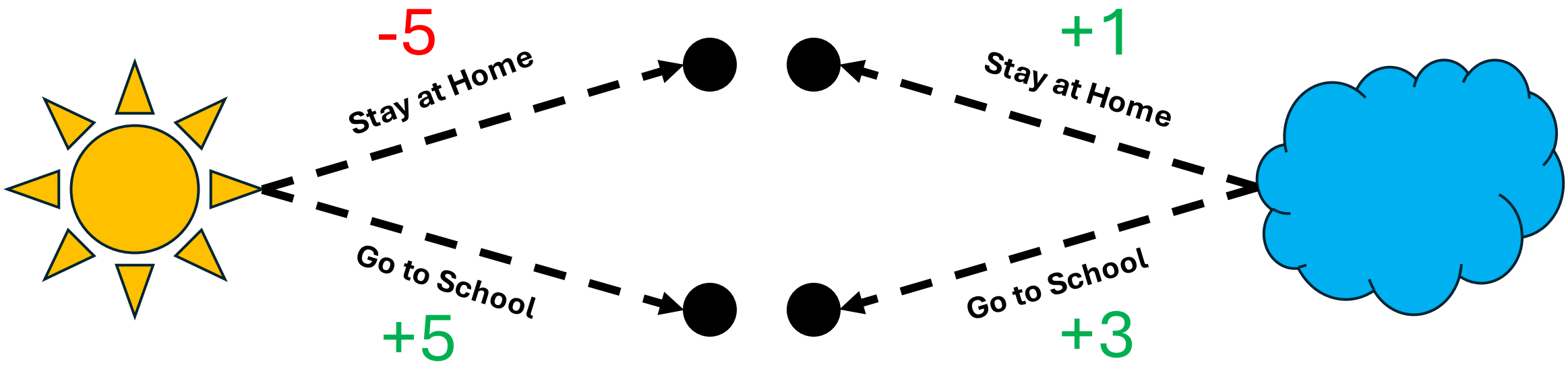


States s: {Sunny, Cloudy}

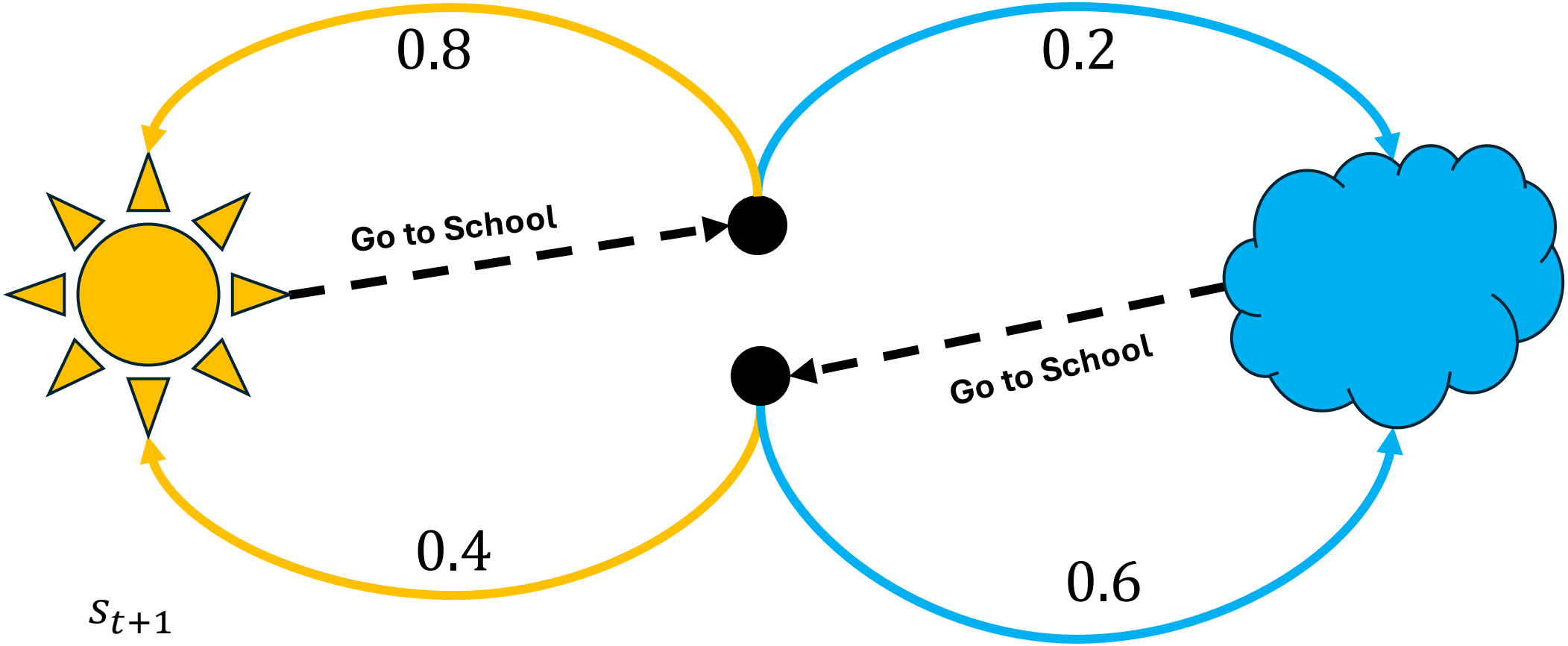
Actions a: {Go to School, Stay at Home}

Discount $\gamma = 0.9$

	Go to School	Stay at Home
	+5	-5
	+3	+1



$R_{school} = ?$ $R_{stay} = ?$

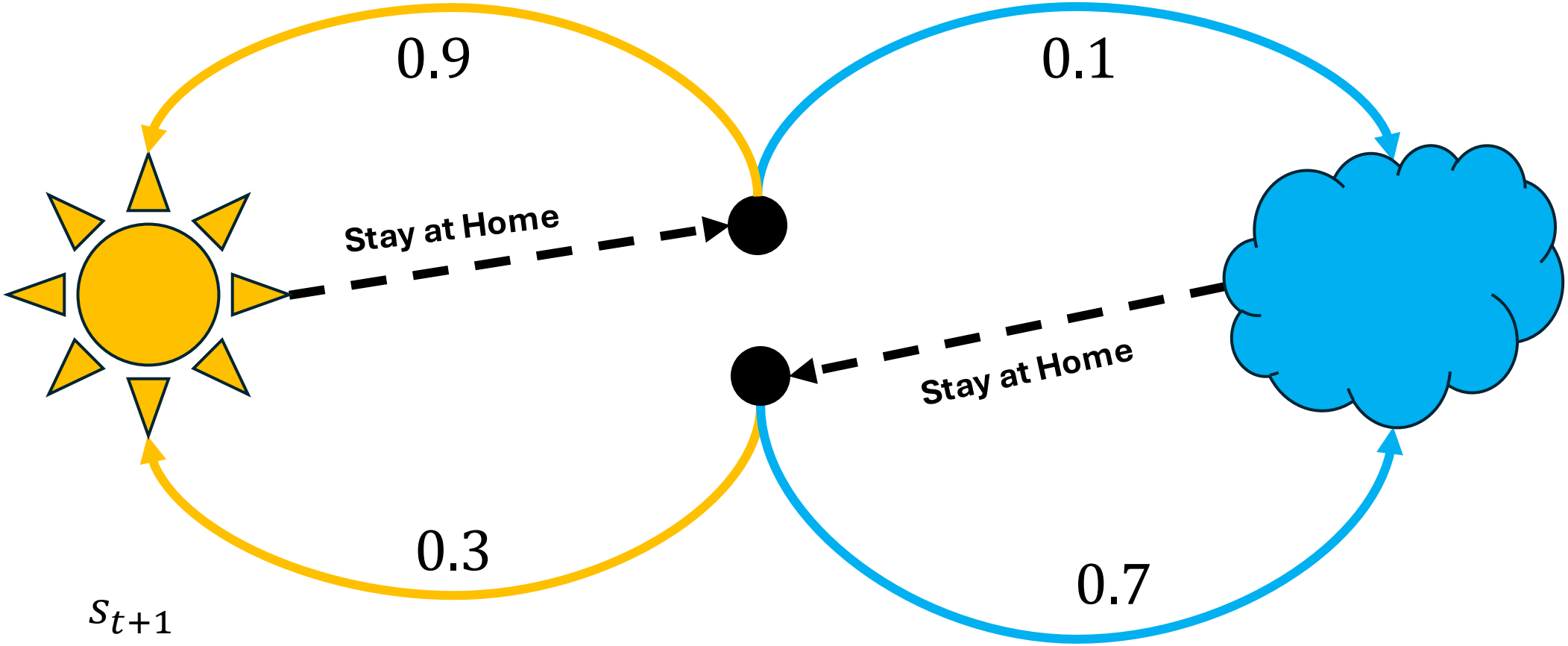


s_{t+1}

s_t		
	0.8	0.2
	0.4	0.6

State Transition Matrix

$$P_{school} = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}$$



s_{t+1}

s_t		
	0.9	0.1
	0.3	0.7

State Transition Matrix

$$P_{stay} = \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix}$$

Step 1: Compute state-wise average reward under the policy π

1. $Find\ r_\pi\ for\ sunny = ?$

2. $Find\ r_\pi\ for\ cloudy = ?$

3. $Find\ r_\pi\ matrix = ?$

Step 2: Compute the policy transition matrix

Row 1 (Sunny):

- $P\pi(1,1) = ??$
- $P\pi(1,2) = ??$

Row 2 (Cloudy):

- $P\pi(2,1) = ??$
- $P\pi(2,2) = ??$

$$4.\, Find\ P_\pi = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$$

Step 3: Write the Bellman expectation equations $v_\pi(sunny)$

General Form:

$$v_\pi(s) = r_\pi(s) + \gamma \sum P_\pi(s, s')\, v_\pi(s')$$

5. $find\ v_1 = ??$

Step 3: Write the Bellman expectation equations $v_\pi(cloudy)$

General Form:

$$v_\pi(s) = r_\pi(s) + \gamma \sum P_\pi(s, s')\, v_\pi(s')$$

6. $Find\ v_2 = ?$

Step 4: Solve for $v_\pi(cloudy)$

7. $v_\pi(cloudy) = ??$

Step 4: Solve for $v_\pi(sunny)$

$8. v_\pi(sunny) = ???$

Step 5: Write the Bellman optimality equations

General Form:

$$v_*(s) = \max_a \{R(s,a) + \gamma s' \sum P(s' \mid s,a) v*(s')\}$$

Find Sunny (v_1) using Go to School:

$$9.\, v_*(sunny) = ???$$

Find Cloudy (v_2) using Go to School:

$$10.\, v_*(cloudy) = ???$$

Step 6: Solve for v_*(cloudy)

11. v_*(cloudy) $= ???$

Step 6: Solve for v_*(sunny)

$$12.\ v_*(\text{cloudy}) = 21.538$$

Step 7: Solve for q_*

13. $\mathrm{q}(1, School) = ??$

14. $\mathrm{q}(1, Home) = ??$

15. $\mathrm{q}(2, School) = ??$

16. $\mathrm{q}(2, Home) = ??$