

Reward

- eat Food	+10
- Game over	-10
- else	0

Action

$[1, 0, 0] \rightarrow$ Straight
 $[0, 1, 0] \rightarrow$ Right Turn
 $[0, 0, 1] \rightarrow$ Left Turn

State

Danger [S,R,L]
Direction [L,R,U,D]
Food [L,R,U,D]

} Boolean values

Learning model

Q value = Quality of action

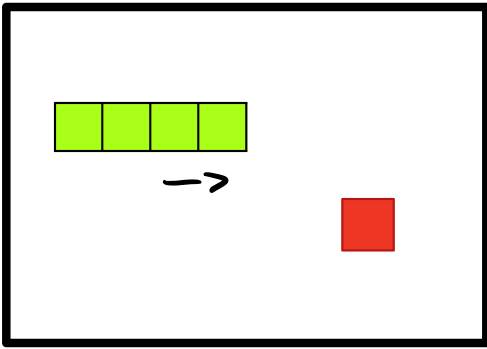
C. Init Q value

1. Choose Action (Model.Predict(State))

2. Perform action

3. Measure Reward

4. Update Q value (train model)



$\Rightarrow \text{get}(\text{state}) \text{ game} \Rightarrow \begin{bmatrix} 0, 0, 0, \\ 0, 1, 0, 0, \\ 0, 1, 0, 1 \end{bmatrix} \Rightarrow \text{MODEL} \Rightarrow \text{Action} \Rightarrow \underline{\underline{\text{Repeat}}}$