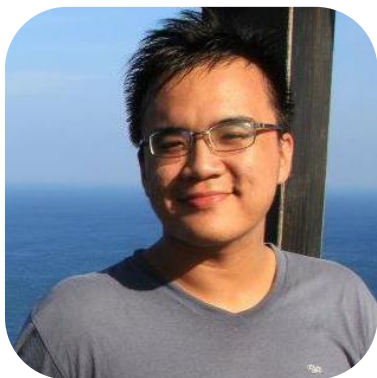


Asking for Knowledge : Training RL Agents to Query External Knowledge Using Language



Iou-Jen Liu*



Xingdi Yuan*



Marc-Alexandre Côté*



Pierre-Yves Oudeyer



Alex Schwing



*equal contribution

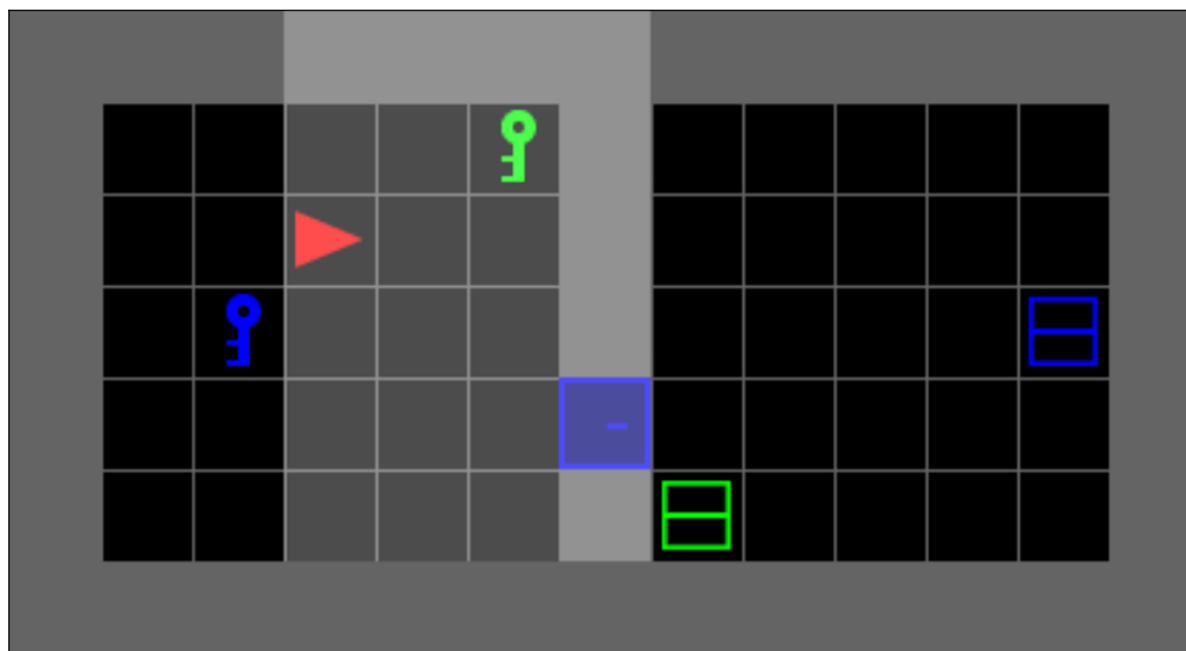
To solve difficult tasks, humans ask questions to acquire knowledge from external sources

To solve difficult tasks, humans ask questions to acquire knowledge from external sources

Classical RL agents lack such an ability, i.e. asking for knowledge in **natural language** [Mnih et al. 2015; Hessel et al. 2017; Haarnoja et al. 2018; Yu et al. 2020]

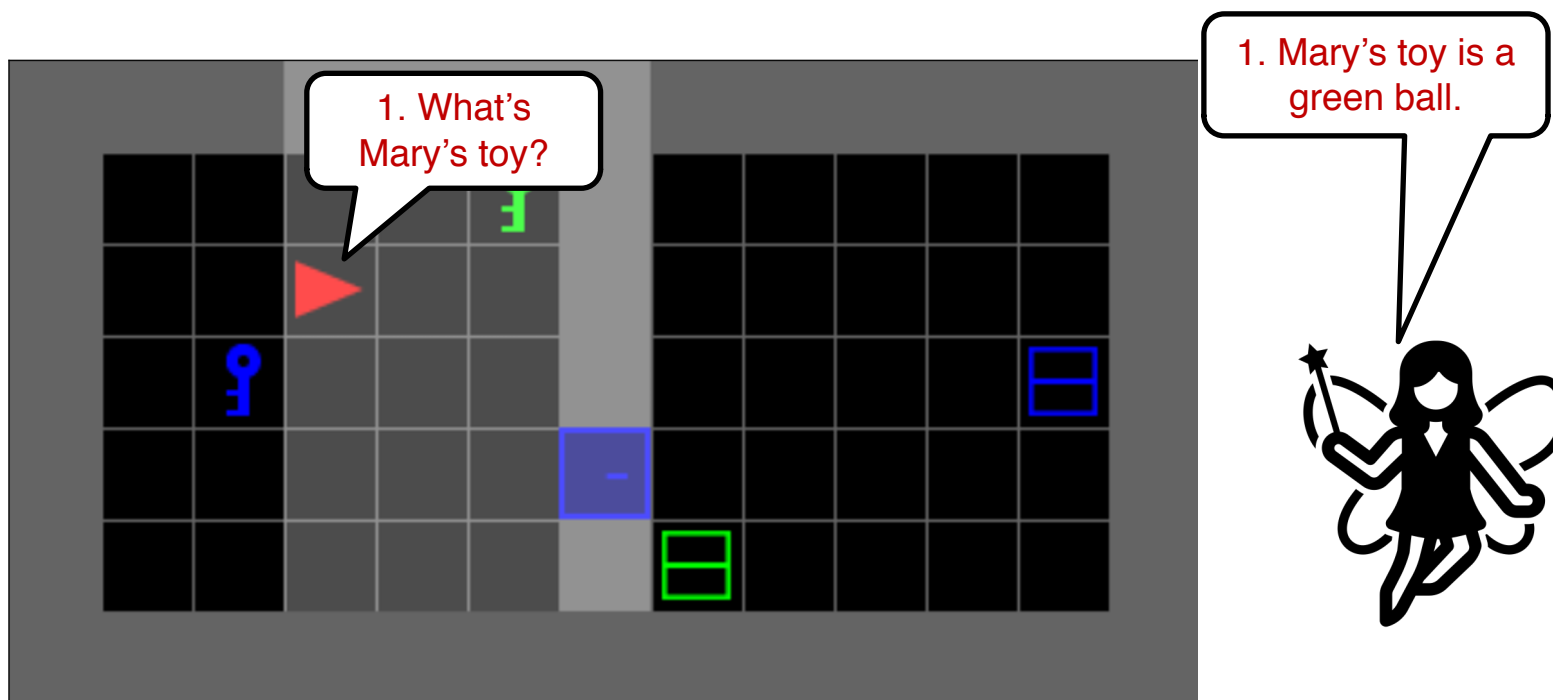
To evaluate, we create *Q-BabyAI* and *Q-TextWorld* environments

To evaluate, we create *Q-BabyAI* and *Q-TextWorld* environments



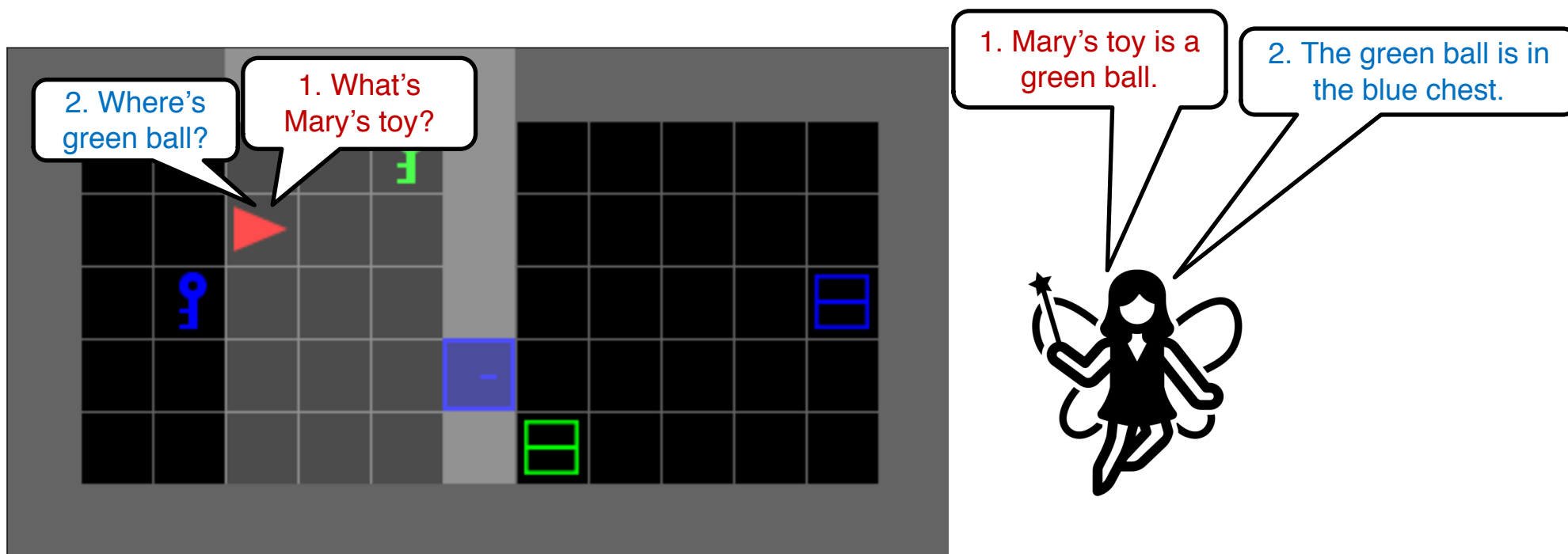
Task: Find the key to the door and find Mary's toy.

To evaluate, we create *Q-BabyAI* and *Q-TextWorld* environments



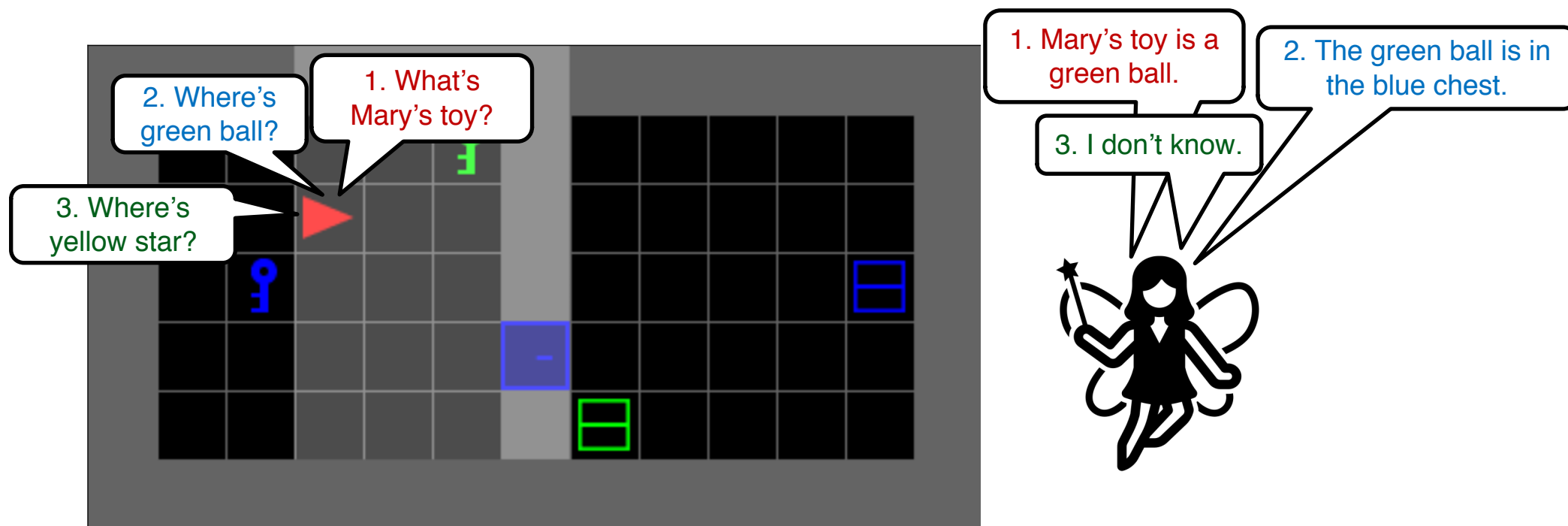
Task: Find the key to the door and find Mary's toy.

To evaluate, we create *Q-BabyAI* and *Q-TextWorld* environments



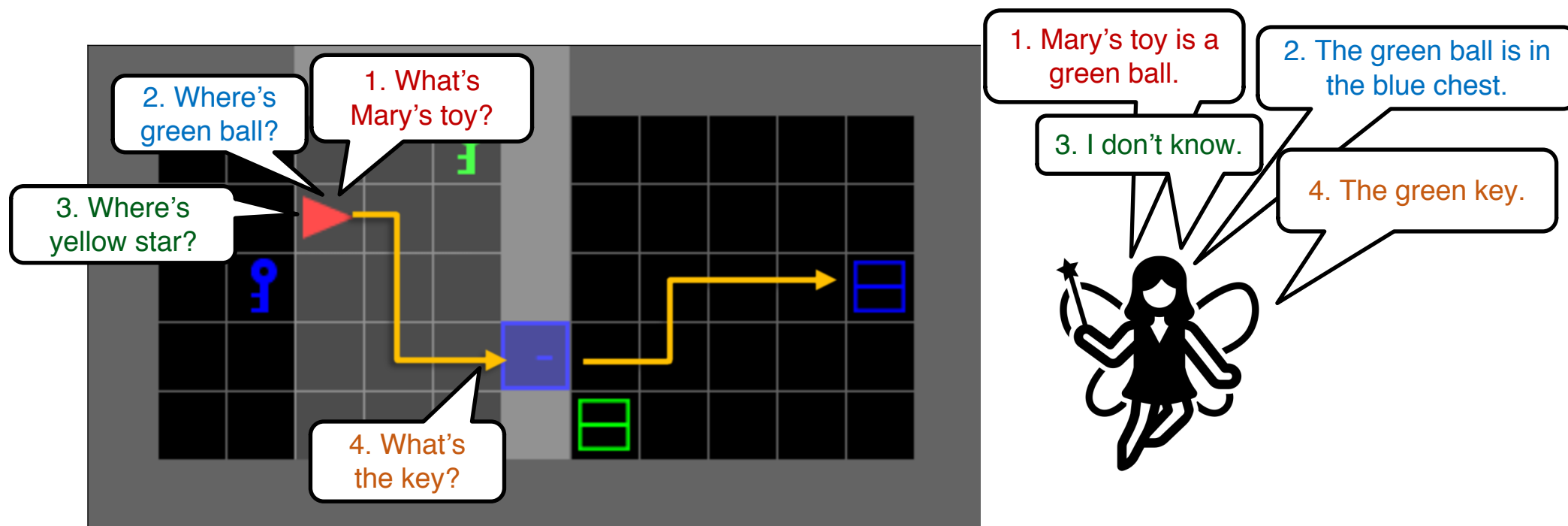
Task: Find the key to the door and find Mary's toy.

To evaluate, we create *Q-BabyAI* and *Q-TextWorld* environments



Task: Find the key to the door and find Mary's toy.

To evaluate, we create *Q-BabyAI* and *Q-TextWorld* environments



Task: Find the key to the door and find Mary's toy.

Challenges

Challenges

- Irrelevant information

Challenges

- Irrelevant information
- Large language generation space

Challenges

- Irrelevant information
- Large language generation space
- Delayed reward for making meaningful queries

Challenges

- Irrelevant information
- Large language generation space
- Delayed reward for making meaningful queries

Asking for knowledge (AFK) agents

Challenges

- Irrelevant information
- Large language generation space
- Delayed reward for making meaningful queries

Asking for knowledge (AFK) agents

- Notebook
 - Non-parametric memory

Challenges

- Irrelevant information
- Large language generation space
- Delayed reward for making meaningful queries

Asking for knowledge (AFK) agents

- Notebook
 - Non-parametric memory
- Pointer mechanism for generating template language query

Challenges

- Irrelevant information
- Large language generation space
- Delayed reward for making meaningful queries

Asking for knowledge (AFK) agents

- Notebook
 - Non-parametric memory
- Pointer mechanism for generating template language query
- Episodic exploration strategy

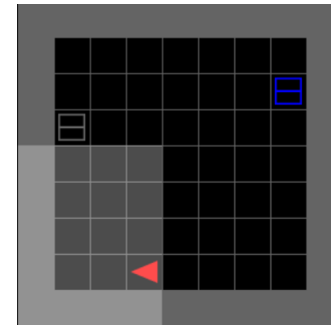
Results

	Tasks	No Query	Query Baseline	AFK (Ours)
Lv. 1	♣	50.5±2.0	49.8±1.2	100.0±0.0
	♠	68.3±2.4	73.8±1.2	100.0±0.0
	♦	98.9±0.8	99.3±0.3	100.0±0.0
	♥	99.7±0.3	85.3±22.3	100.0±0.0
Lv. 2	♣♠	0.0±0.0	0.0±0.0	90.3±1.8
	♣♦	0.1±0.1	0.6±0.5	94.3±2.3
	♣♥	0.0±0.0	0.0±0.0	96.0±2.6
	♠♦	0.4±0.1	0.2±0.2	100.0±0.0
	♠♥	0.0±0.0	0.0±0.0	19.6±39.4
	♦♥	10.8±3.1	10.2±2.6	89.8±3.8
Lv. 3	♣♠♦	0.0±0.0	0.0±0.0	0.15±0.2
	♣♠♥	0.0±0.0	0.0±0.0	0.0±0.0
	♣♦♥	0.0±0.0	0.0±0.0	1.6±0.6
	♠♦♥	4.3±1.0	4.4±0.8	5.8±1.1
Lv. 4	♣♠♦♥	0.0±0.0	0.0±0.0	0.0±0.1

Table 1. Success rate (%) on *Q-BabyAI*. ♣: **Object in Box**, ♠: **Danger**, ♦: **Go to Favorite**, ♥: **Open Door**.

Results

	Tasks	No Query	Query Baseline	AFK (Ours)
Lv. 1	♣	50.5±2.0	49.8±1.2	100.0±0.0
	♠	68.3±2.4	73.8±1.2	100.0±0.0
	♦	98.9±0.8	99.3±0.3	100.0±0.0
	♥	99.7±0.3	85.3±22.3	100.0±0.0
Lv. 2	♣♠	0.0±0.0	0.0±0.0	90.3±1.8
	♣♦	0.1±0.1	0.6±0.5	94.3±2.3
	♣♥	0.0±0.0	0.0±0.0	96.0±2.6
	♠♦	0.4±0.1	0.2±0.2	100.0±0.0
	♠♥	0.0±0.0	0.0±0.0	19.6±39.4
	♦♥	10.8±3.1	10.2±2.6	89.8±3.8
Lv. 3	♣♠♦	0.0±0.0	0.0±0.0	0.15±0.2
	♣♠♥	0.0±0.0	0.0±0.0	0.0±0.0
	♣♦♥	0.0±0.0	0.0±0.0	1.6±0.6
	♠♦♥	4.3±1.0	4.4±0.8	5.8±1.1
Lv. 4	♣♠♦♥	0.0±0.0	0.0±0.0	0.0±0.1



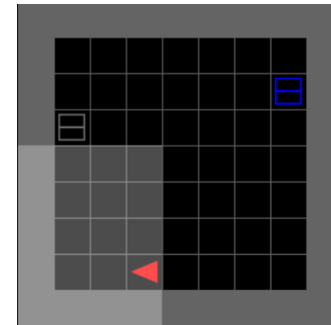
Object in Box (♣)

Table 1. Success rate (%) on *Q-BabyAI*. ♣: **Object in Box**, ♠: **Danger**, ♦: **Go to Favorite**, ♥: **Open Door**.

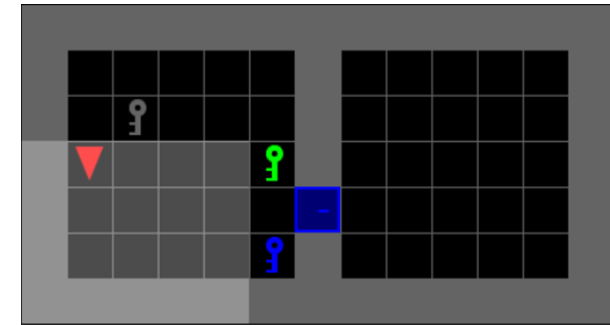
Results

	Tasks	No Query	Query Baseline	AFK (Ours)
Lv. 1	♣	50.5±2.0	49.8±1.2	100.0±0.0
	♠	68.3±2.4	73.8±1.2	100.0±0.0
	♦	98.9±0.8	99.3±0.3	100.0±0.0
	♥	99.7±0.3	85.3±22.3	100.0±0.0
Lv. 2	♣♠	0.0±0.0	0.0±0.0	90.3±1.8
	♣♦	0.1±0.1	0.6±0.5	94.3±2.3
	♣♥	0.0±0.0	0.0±0.0	96.0±2.6
	♠♦	0.4±0.1	0.2±0.2	100.0±0.0
	♠♥	0.0±0.0	0.0±0.0	19.6±39.4
	♦♥	10.8±3.1	10.2±2.6	89.8±3.8
Lv. 3	♣♠♦	0.0±0.0	0.0±0.0	0.15±0.2
	♣♠♥	0.0±0.0	0.0±0.0	0.0±0.0
	♣♦♥	0.0±0.0	0.0±0.0	1.6±0.6
	♠♦♥	4.3±1.0	4.4±0.8	5.8±1.1
Lv. 4	♣♠♦♥	0.0±0.0	0.0±0.0	0.0±0.1

Table 1. Success rate (%) on *Q-BabyAI*. ♣: **Object in Box**, ♠: **Danger**, ♦: **Go to Favorite**, ♥: **Open Door**.



Object in Box (♣)

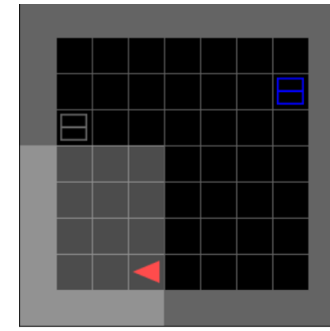


Open Door (♥)

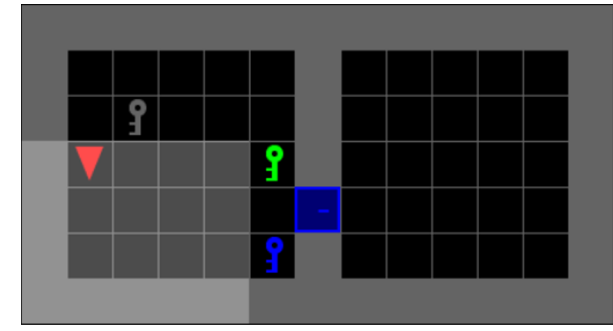
Results

	Tasks	No Query	Query Baseline	AFK (Ours)
Lv. 1	♣	50.5±2.0	49.8±1.2	100.0±0.0
	♠	68.3±2.4	73.8±1.2	100.0±0.0
	♦	98.9±0.8	99.3±0.3	100.0±0.0
	♥	99.7±0.3	85.3±22.3	100.0±0.0
Lv. 2	♣♠	0.0±0.0	0.0±0.0	90.3±1.8
	♣♦	0.1±0.1	0.6±0.5	94.3±2.3
	♣♥	0.0±0.0	0.0±0.0	96.0±2.6
	♠♦	0.4±0.1	0.2±0.2	100.0±0.0
	♠♥	0.0±0.0	0.0±0.0	19.6±39.4
	♦♥	10.8±3.1	10.2±2.6	89.8±3.8
Lv. 3	♣♠♦	0.0±0.0	0.0±0.0	0.15±0.2
	♣♠♥	0.0±0.0	0.0±0.0	0.0±0.0
	♣♦♥	0.0±0.0	0.0±0.0	1.6±0.6
	♠♦♥	4.3±1.0	4.4±0.8	5.8±1.1
Lv. 4	♣♠♦♥	0.0±0.0	0.0±0.0	0.0±0.1

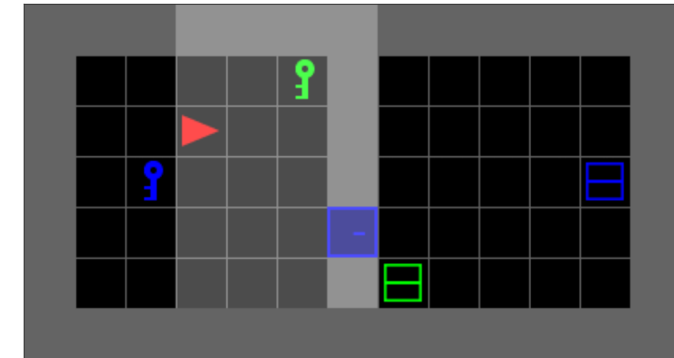
Table 1. Success rate (%) on *Q-BabyAI*. ♣: **Object in Box**, ♠: **Danger**, ♦: **Go to Favorite**, ♥: **Open Door**.



Object in Box (♣)

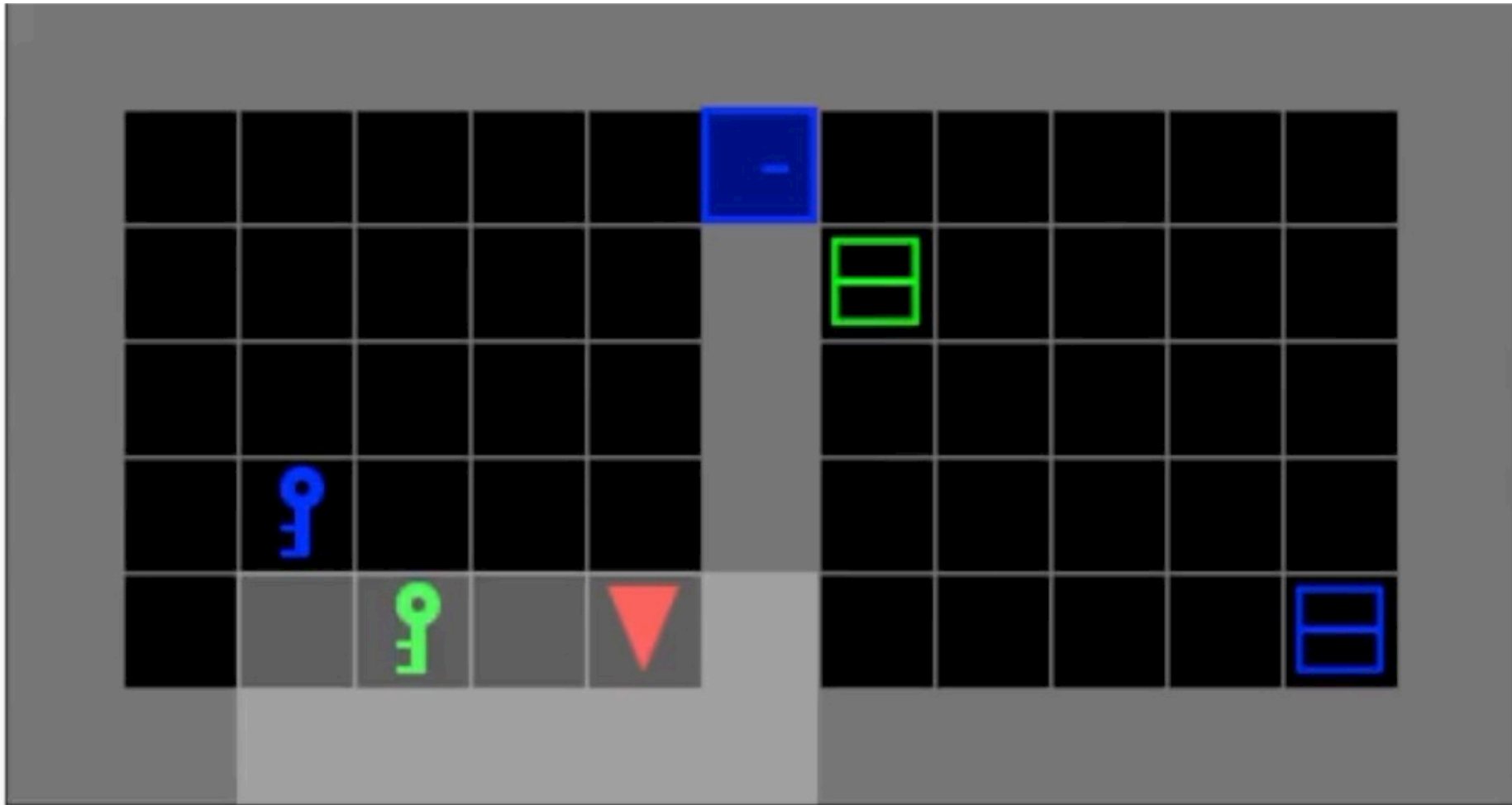


Open Door (♥)



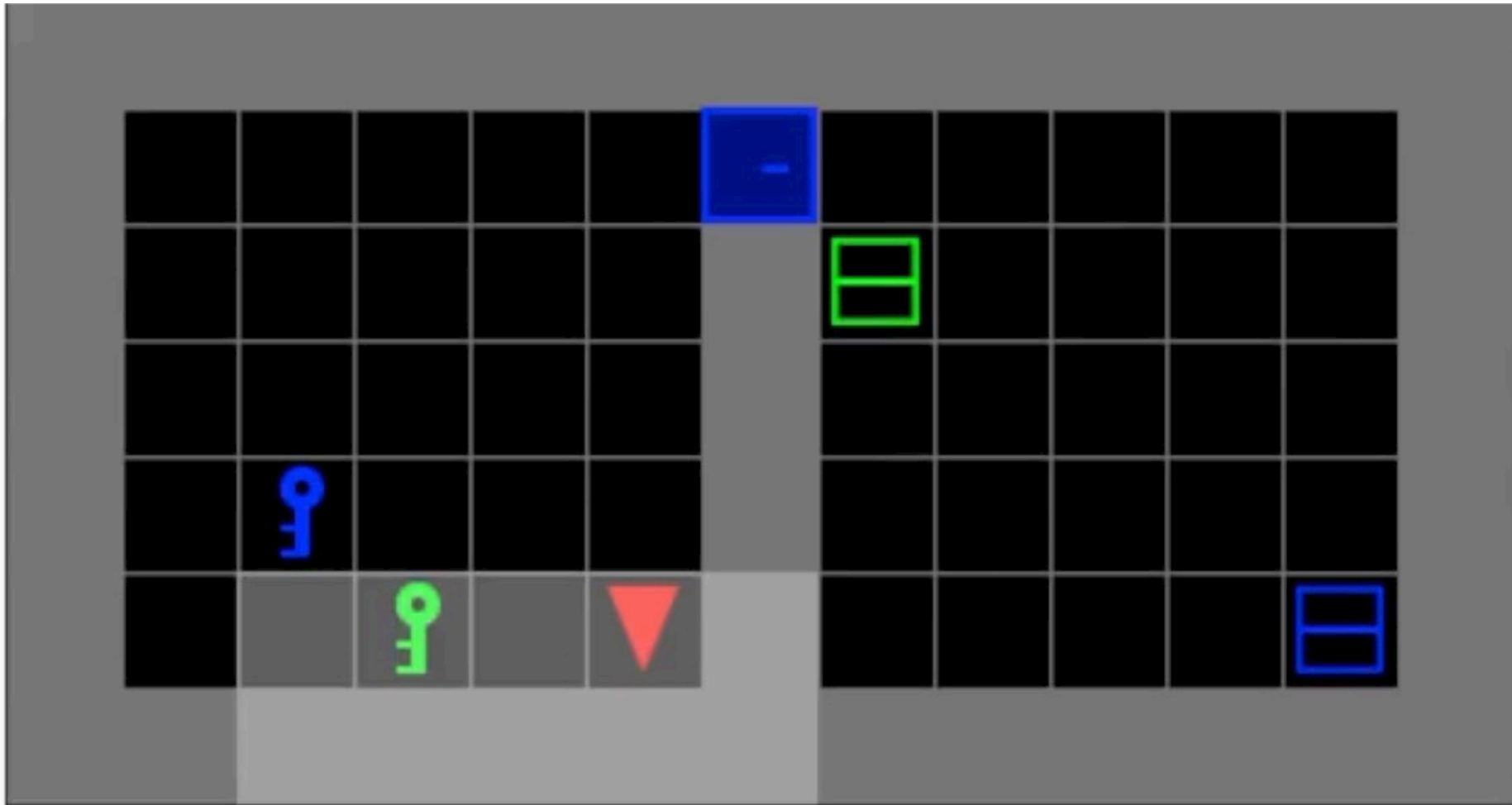
Object in Box + Open Door (♣♥)

AFK Agents in Q-BabyAI



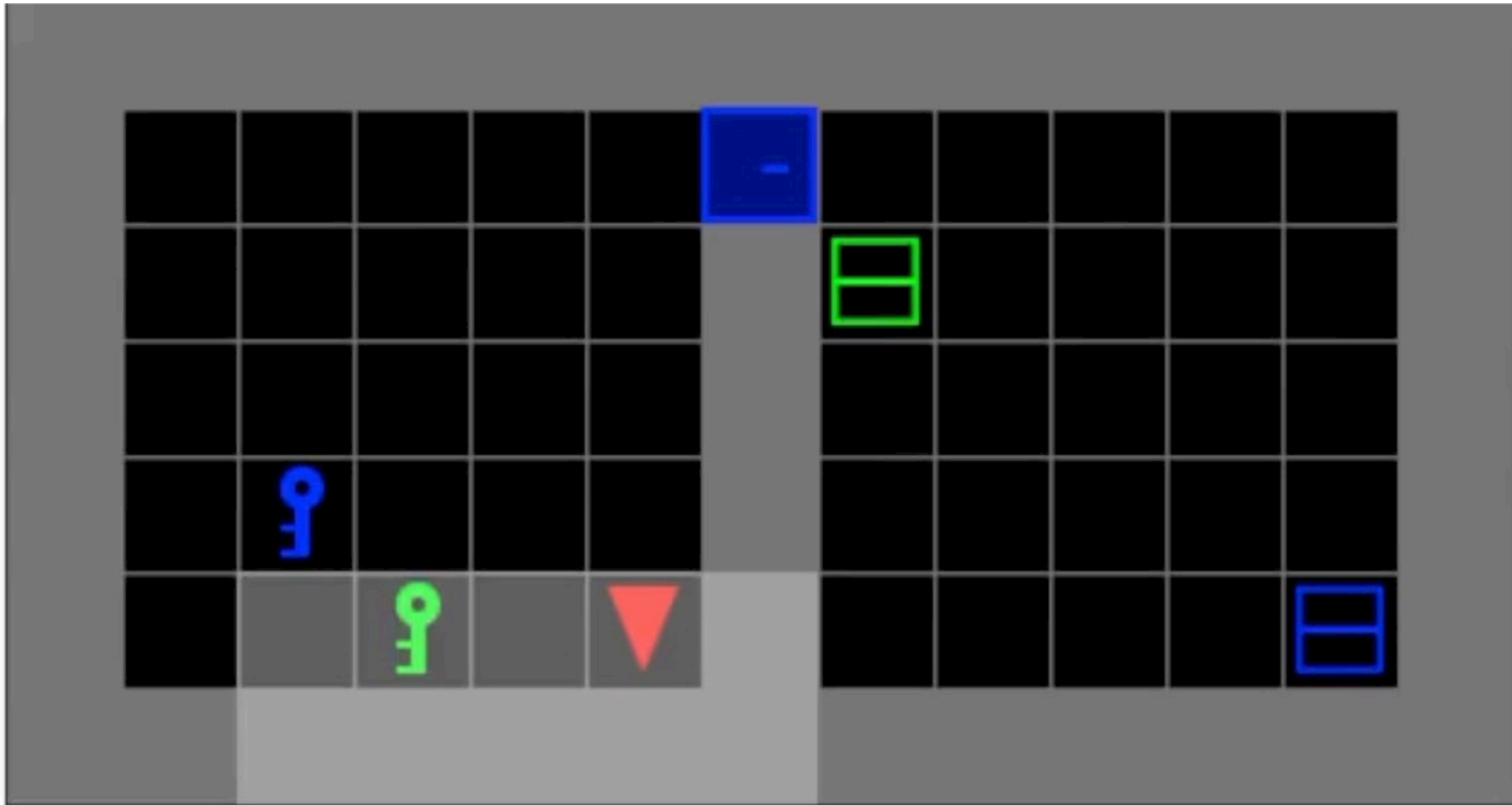
Instr: find the key to the door, and find mary toy step: 0 action: None
Q: None A: None

AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 0 action: None
Q: None A: None

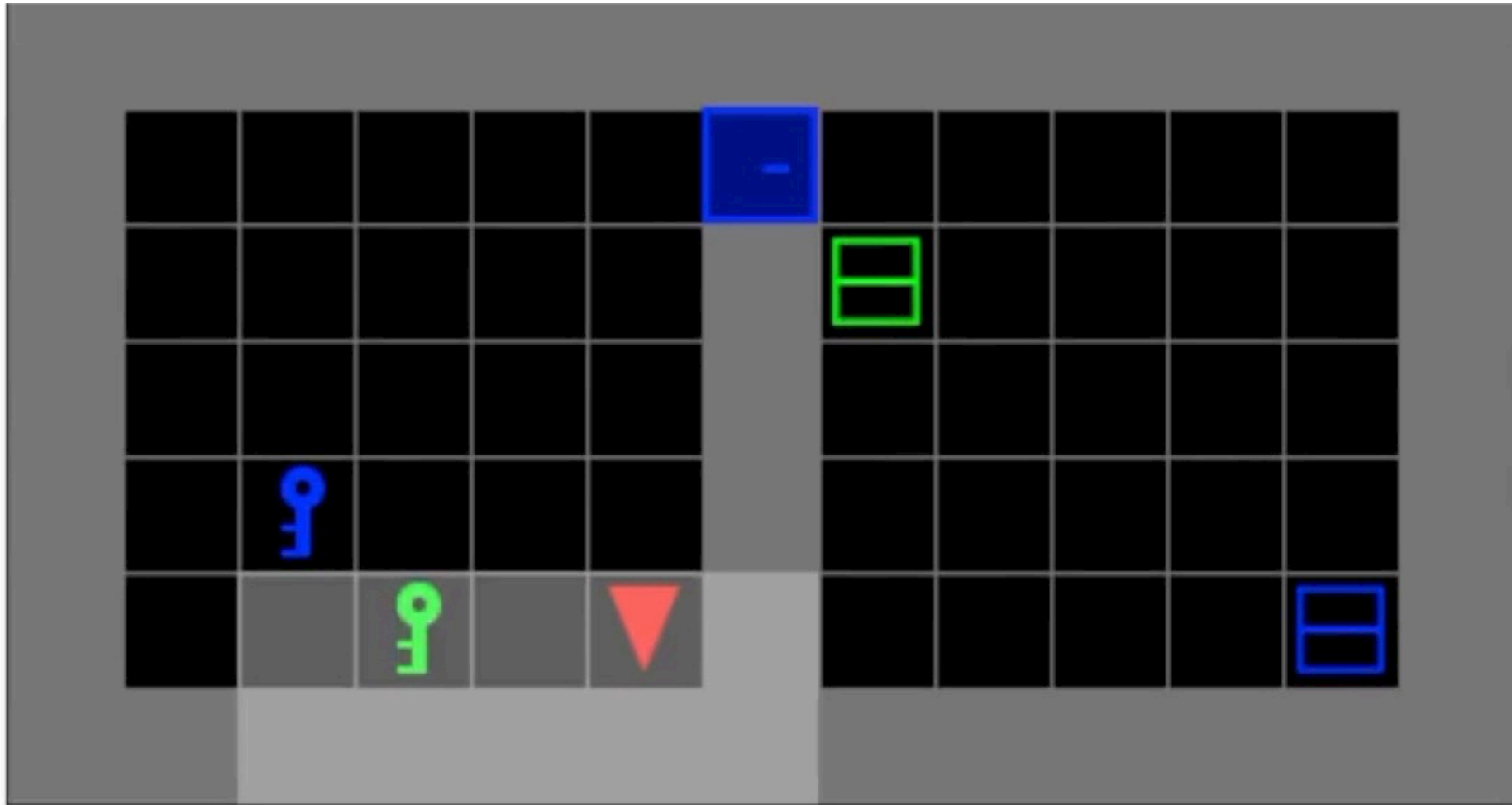
AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 0 action: None

Q: None A: None

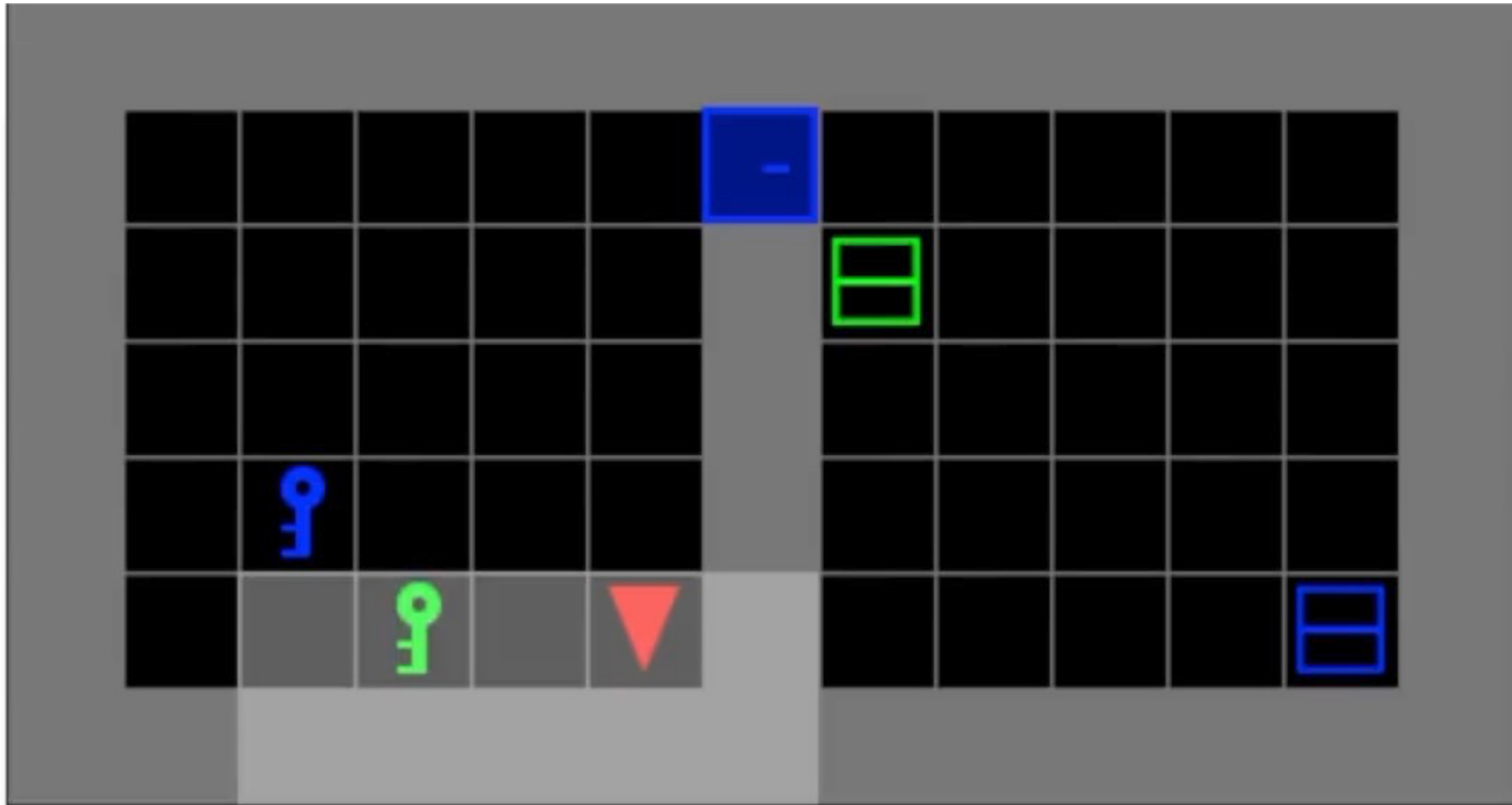
AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 1 action: ['what', 'is', 'mary', 'toy']

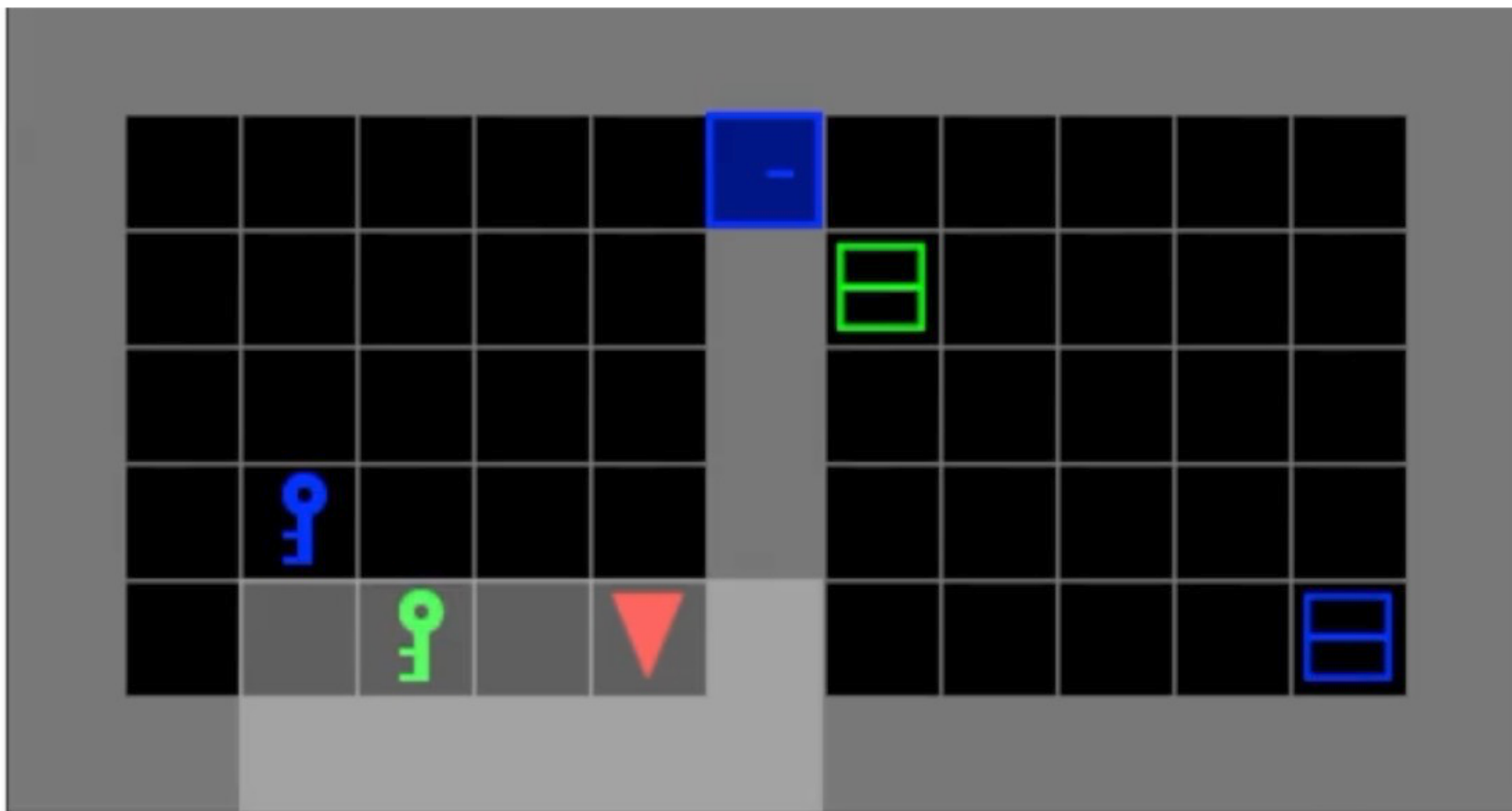
Q: what is mary toy A: mary toy is blue ball

AFK Agents in Q-BabyAI



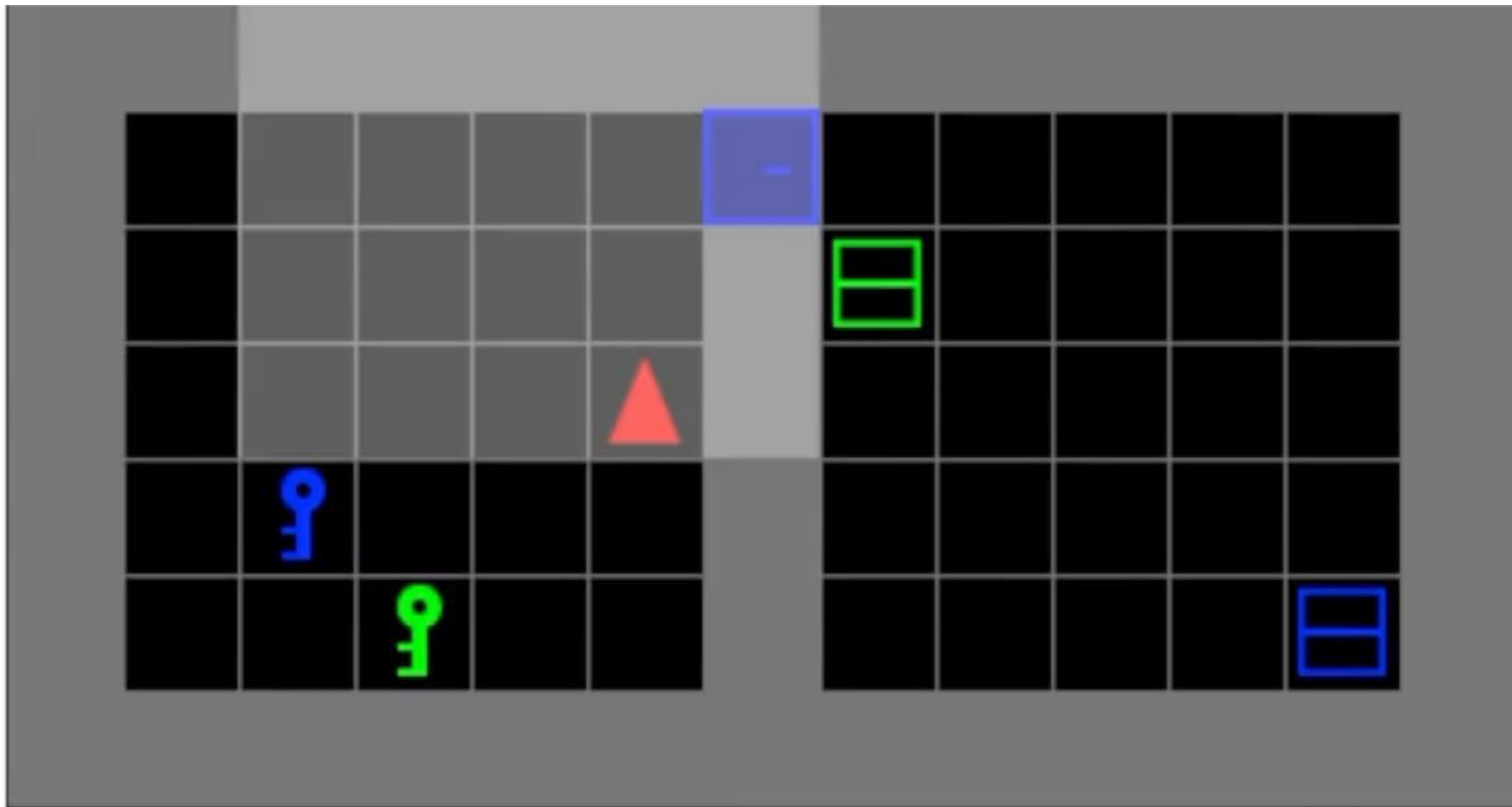
Instr: find the key to the door, and find mary toy step: 2 action: ['where', 'is', 'blue', 'ball']
Q: where is blue ball A: while red ball in the restroom, blue ball in jack suitcase

AFK Agents in Q-BabyAI



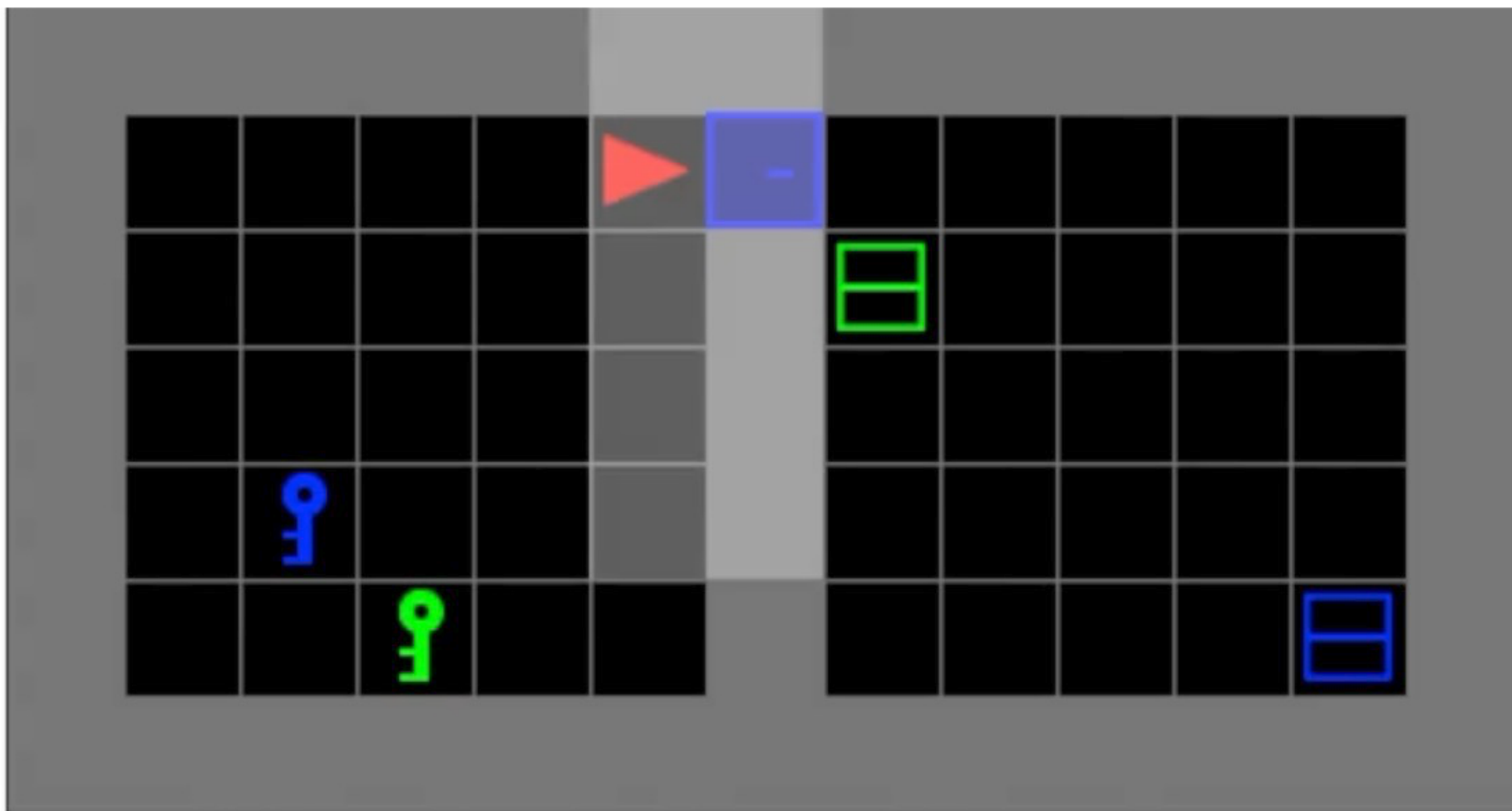
Instr: find the key to the door, and find mary toy step: 3 action: ['what', 'is', 'jack', 'suitcase']
Q: what is jack suitcase A: different from tim, jack suitcase green box

AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 8 action: ['forward']
Q: where is green box A: green box in room1

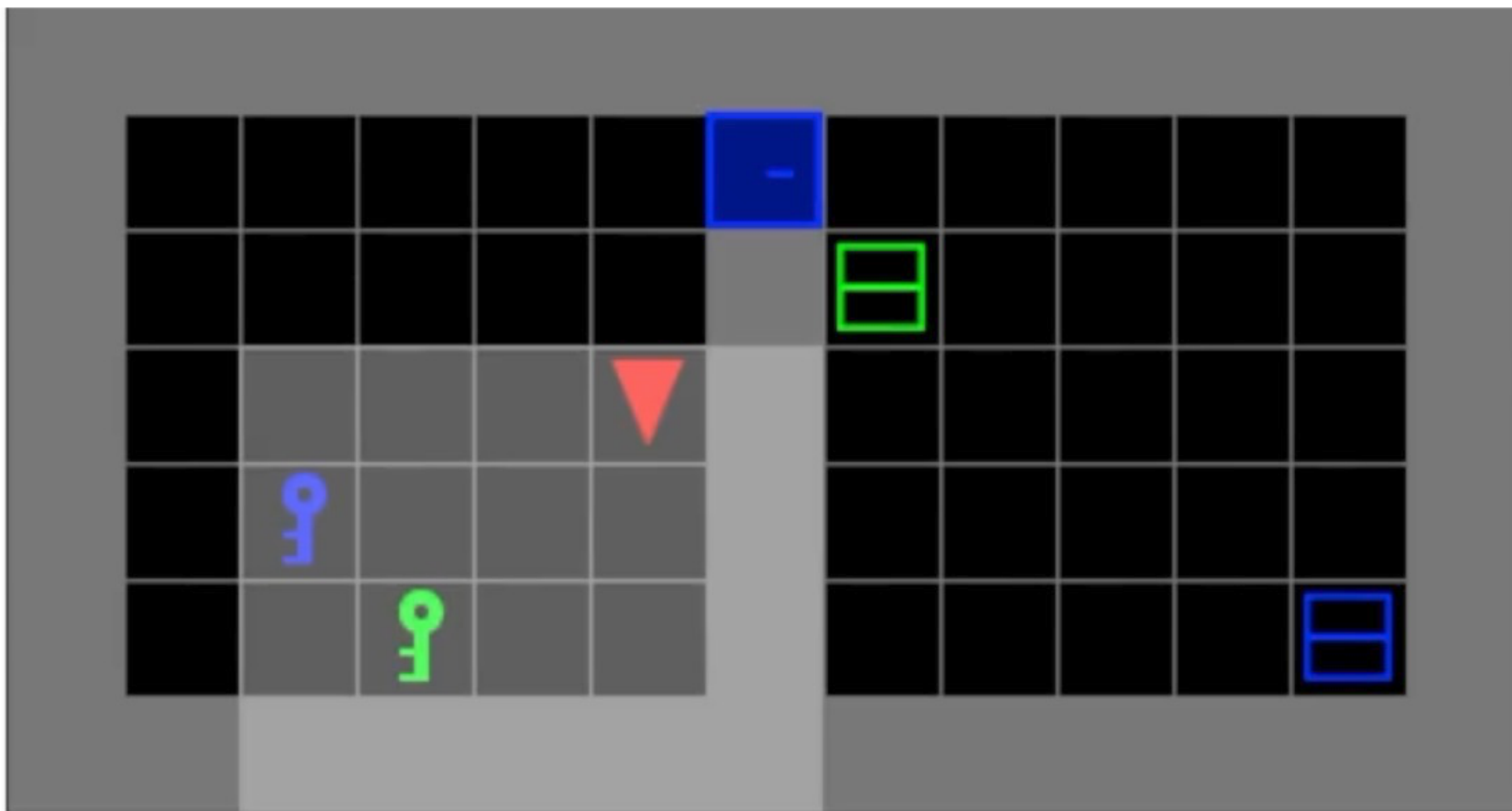
AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 12 action: ['what', 'is', 'key']

Q: what is key A: blue key to the door

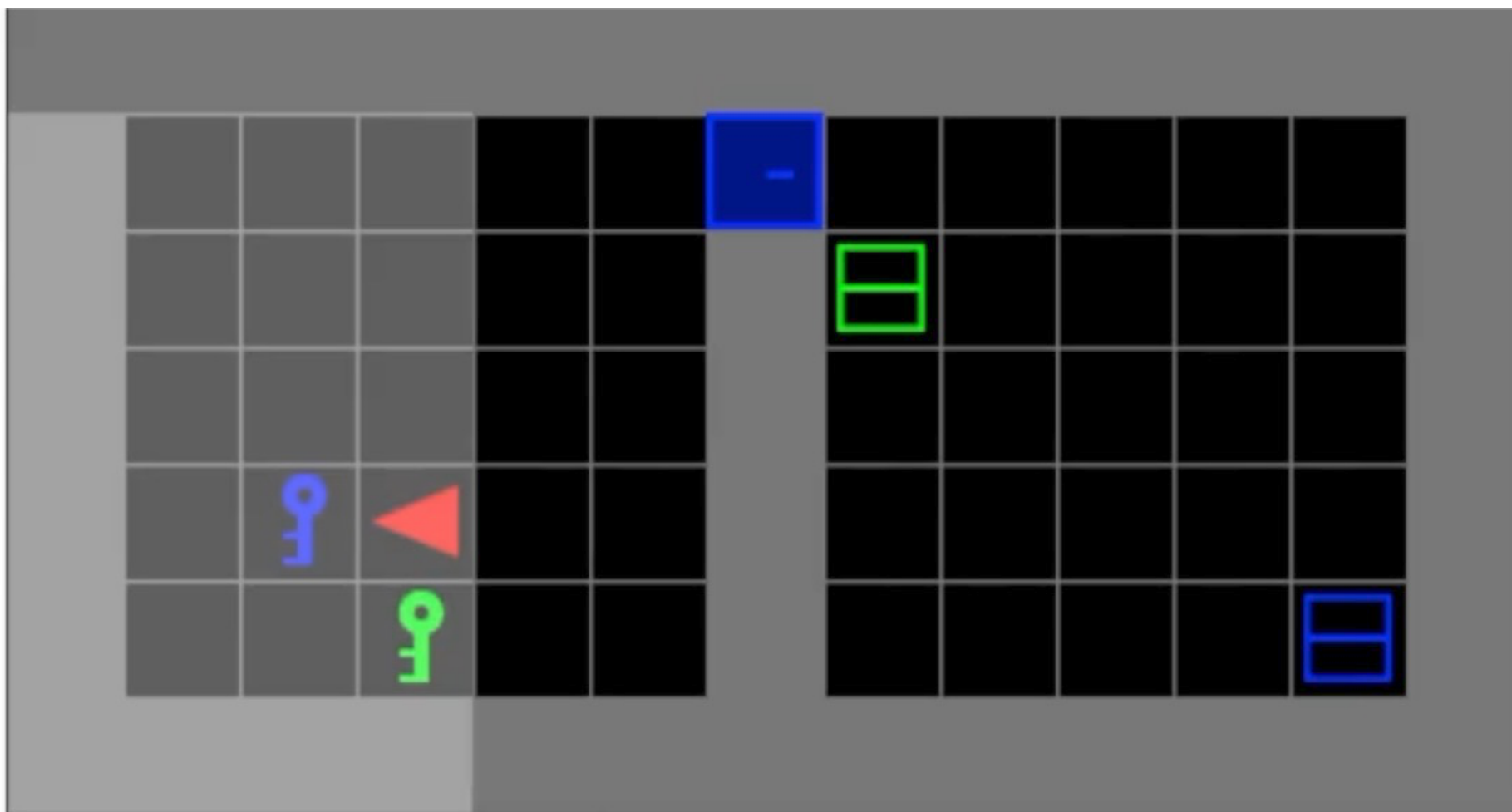
AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 16 action: ['forward']

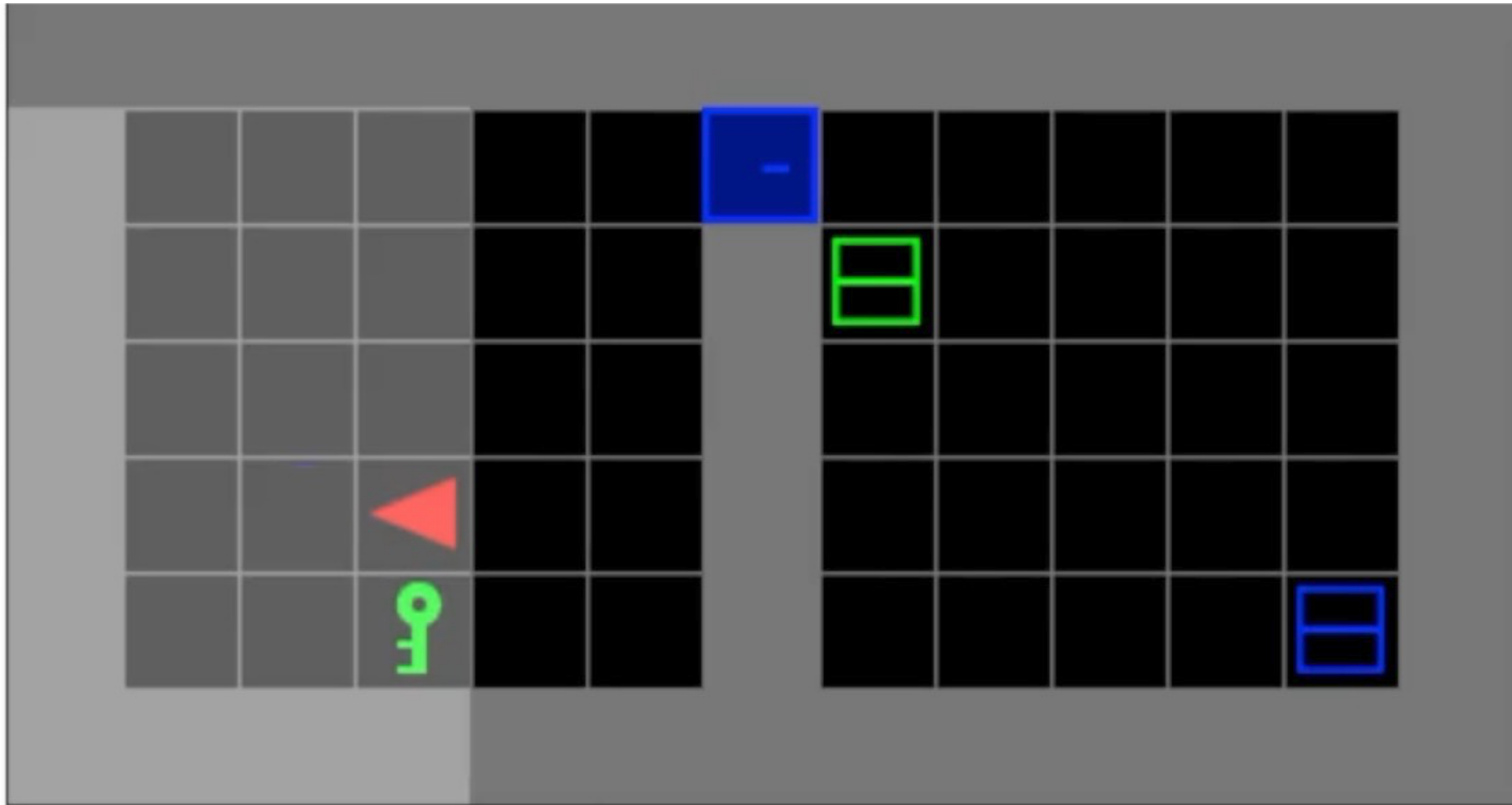
Q: where is blue key A: blue key in room0

AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 20 action: ['forward']
Q: where is blue key A: blue key in room0

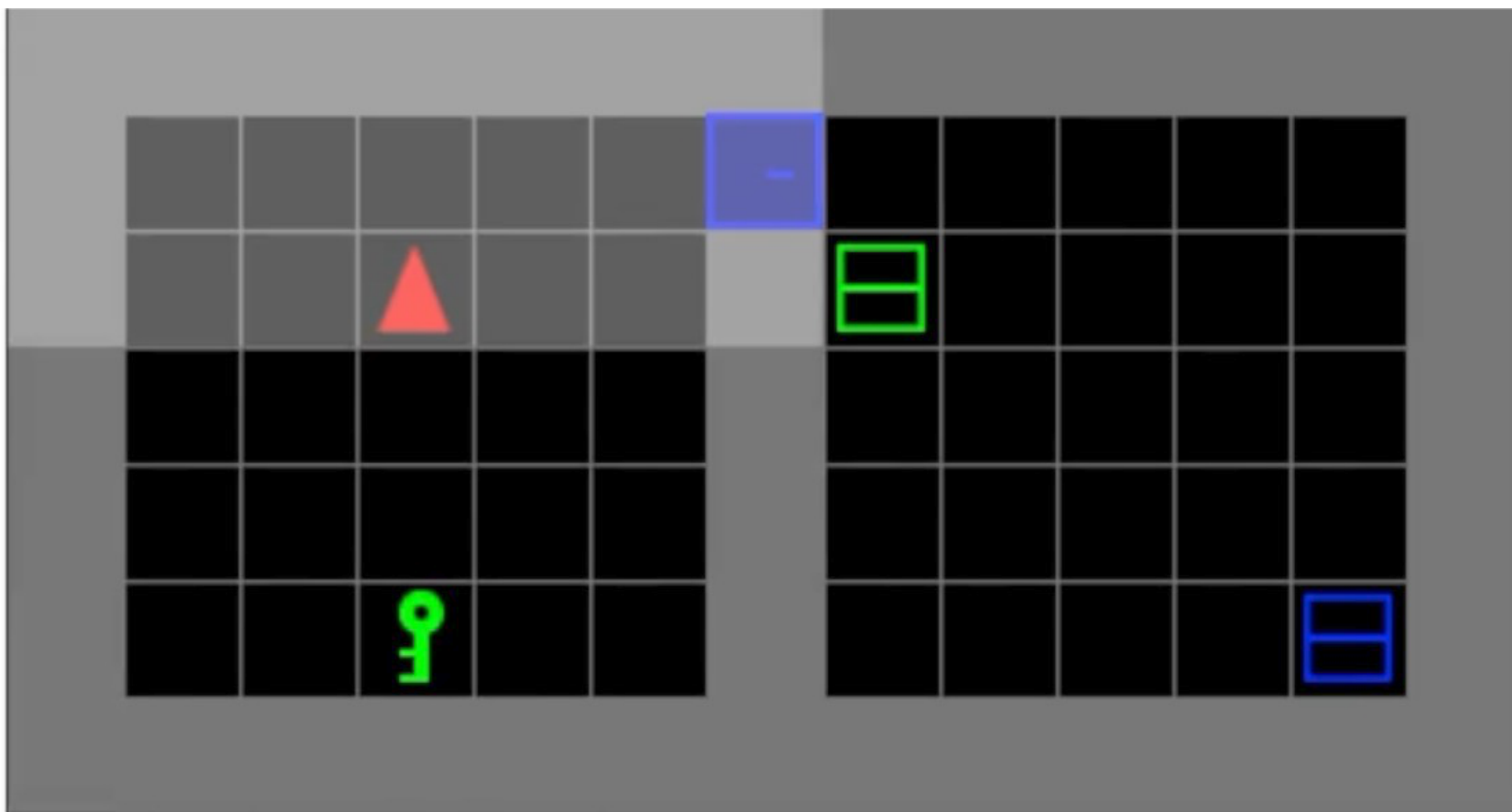
AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 21 action: ['pickup']

Q: where is blue key A: blue key in room0

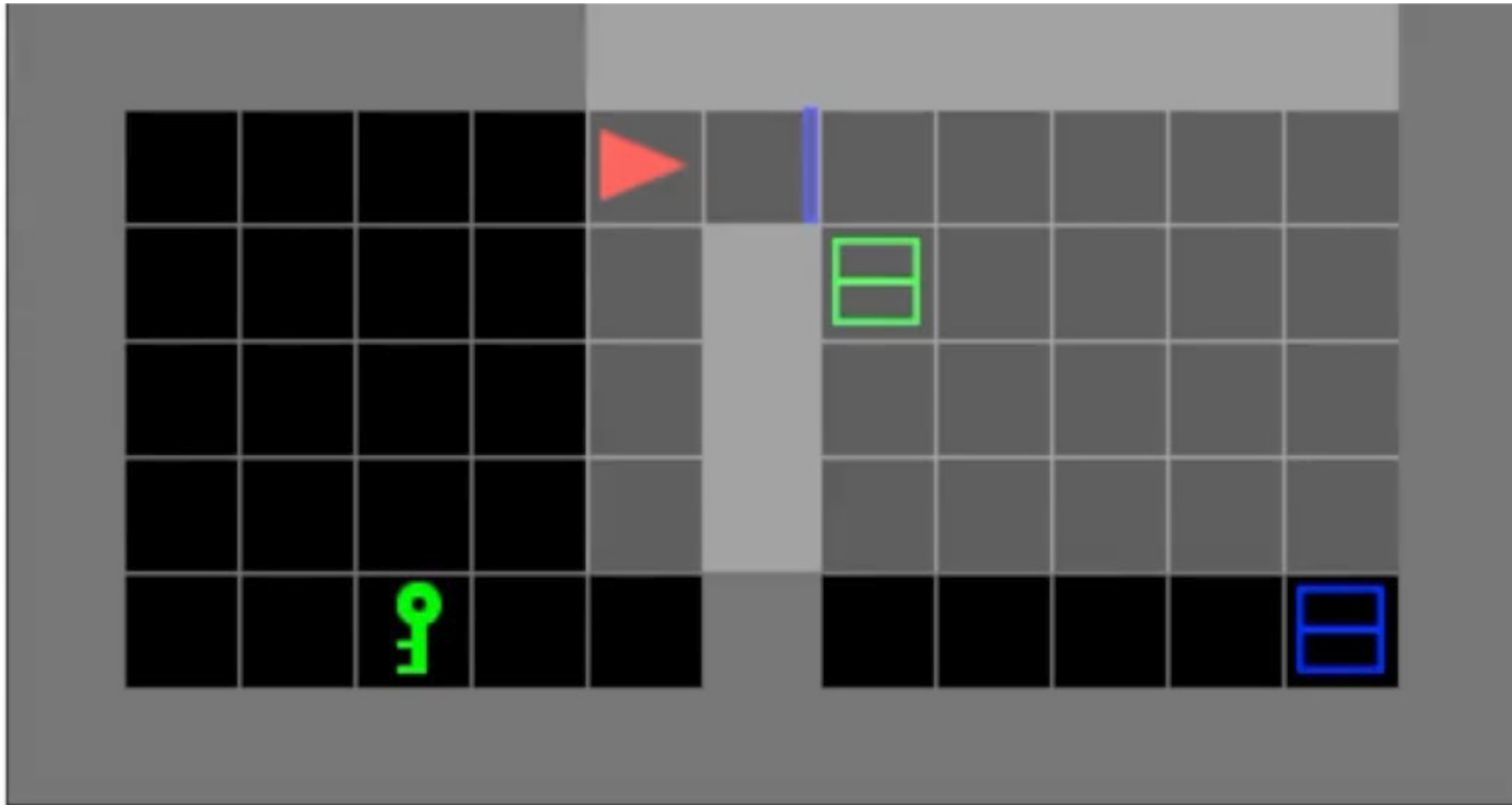
AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 24 action: ['forward']

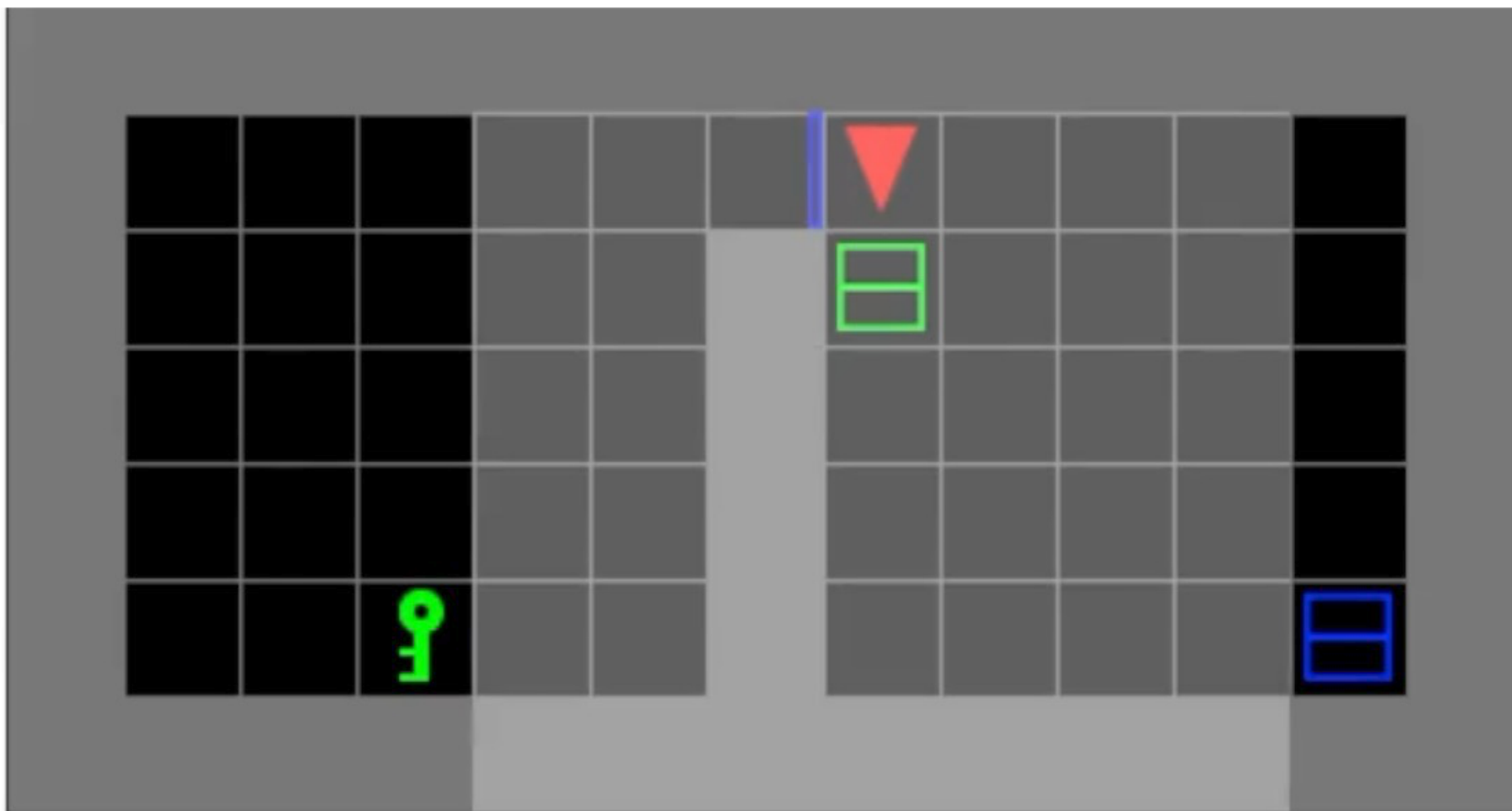
Q: where is blue key A: blue key in room0

AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 29 action: ['toggle']
Q: where is blue key A: blue key in room0

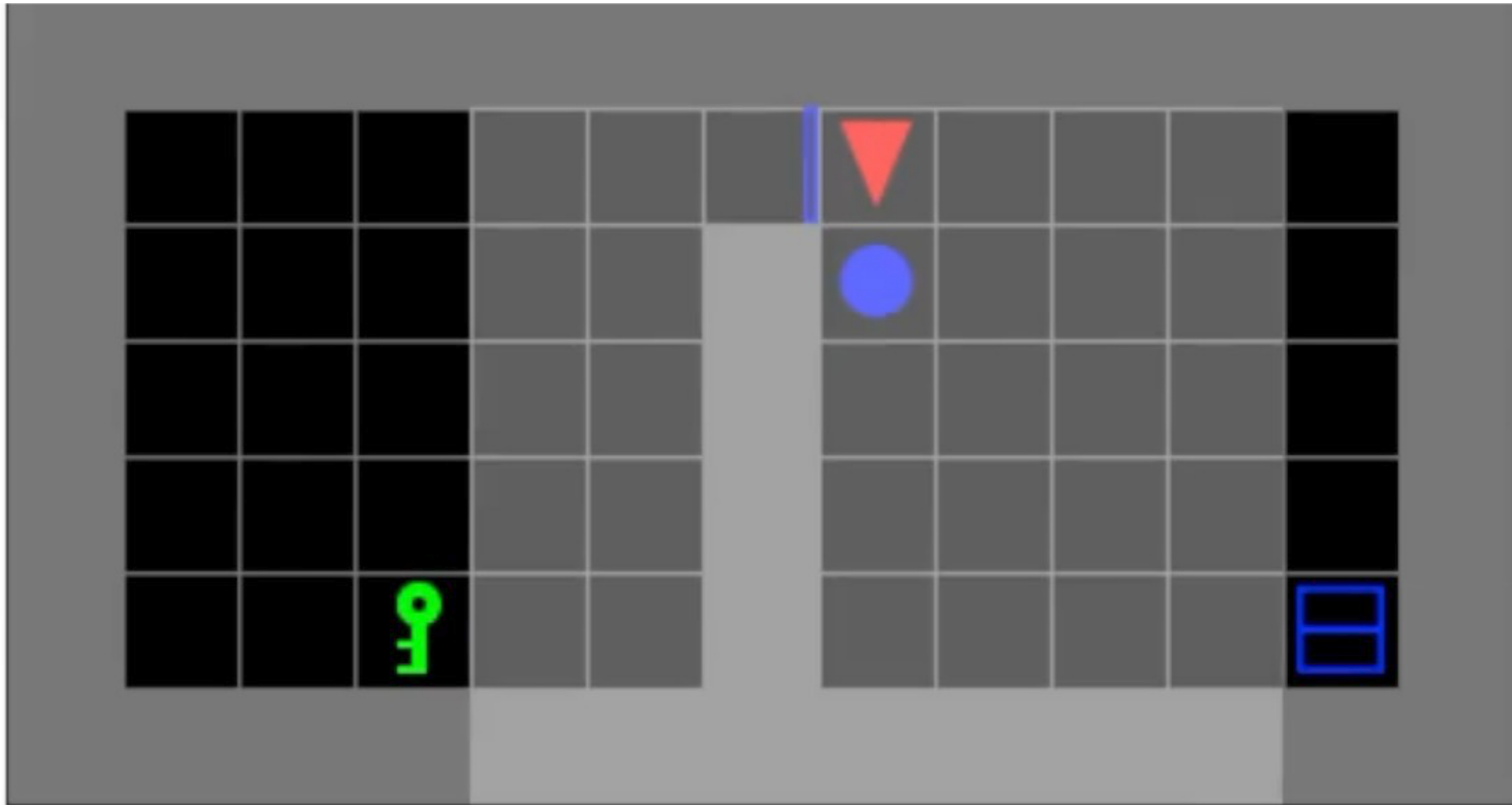
AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 32 action: ['right']

Q: where is blue key A: blue key in room0

AFK Agents in Q-BabyAI



Instr: find the key to the door, and find mary toy step: 33 action: ['toggle']

Q: where is blue key A: blue key in room0

Asking for Knowledge Agent (AFK)

- An initial step to study agents that learn to query knowledge sources with language
- Outperform baselines on *Q-BabyAI* and *Q-TextWorld*

Please see us at the poster session for more details!



<https://ioujenliu.github.io/AFK>