

2D grid-world: Starting from blue reach green

1 - Data collection: Use random policy to generate data


cell numbers: agent visitation

0	0	0	0	0
0	0	0	0	0
0	0	X	0	0
0	0	0	0	0
1	0	0	0	0

2D grid-world: Starting from blue reach green

1 - Data collection: Use random policy to generate data

cell numbers: agent visitation

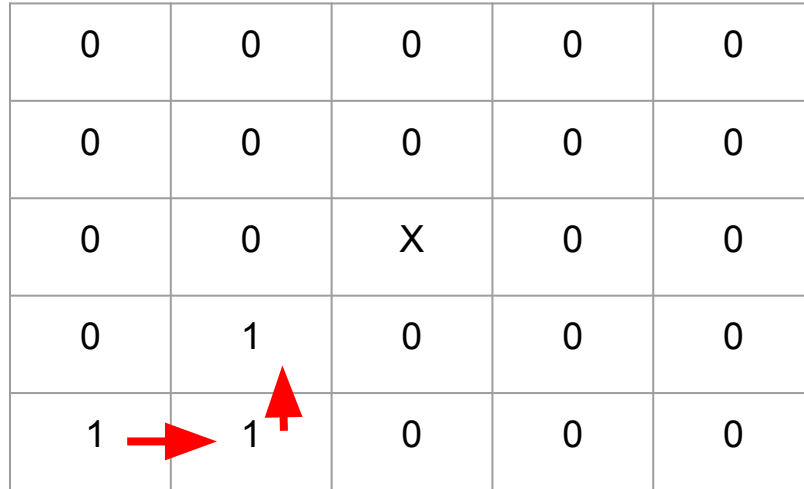
0	0	0	0	0
0	0	0	0	0
0	0	X	0	0
0	0	0	0	0
1	 1	0	0	0

2D grid-world: Starting from blue reach green

1 - Data collection

cell numbers: agent visitation

0	0	0	0	0
0	0	0	0	0
0	0	X	0	0
0	1	0	0	0
1	1	0	0	0



The diagram illustrates a 5x5 grid world. The cells contain the following values:

- Row 1: 0, 0, 0, 0, 0
- Row 2: 0, 0, 0, 0, 0
- Row 3: 0, 0, X, 0, 0
- Row 4: 0, 1, 0, 0, 0
- Row 5: 1, 1, 0, 0, 0

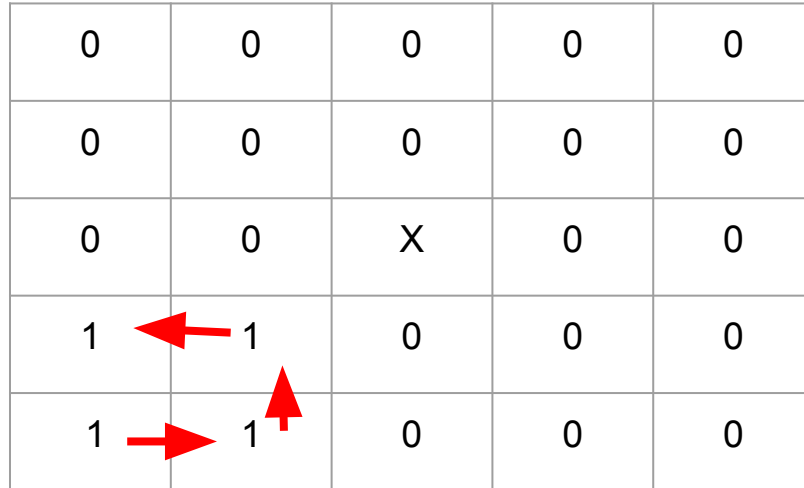
Red arrows indicate the agent's path: starting from the bottom-left cell (row 5, column 1), moving right to (row 5, column 2), and then up to (row 4, column 2).

2D grid-world: Starting from blue reach green

1 - Data collection: Use random policy to generate data

cell numbers: agent visitation

0	0	0	0	0
0	0	0	0	0
0	0	X	0	0
1	← 1	0	0	0
1	→ 1	0	0	0



2D grid-world: Starting from blue reach green

1 - Data collection: Use random policy to generate data

cell numbers: agent visitation

0	0	0	0	0
0	0	0	0	0
0	0	X	0	0
1	2	0	0	0
1	1	0	0	0

The diagram illustrates a 5x5 grid world. The cells contain visitation counts. A red 'X' is located in the third row, third column. Red arrows indicate the agent's path: starting at (4,1), moving right to (4,2), then up to (3,2), then left to (3,1), then up to (2,1), and finally right to (2,2).

2D grid-world: Starting from blue reach green

After many iterationsAll cells
should be visited uniformly.

2D grid-world: Starting from blue reach green

1 - Data collection: Use random policy to generate data

cell numbers: agent visitation

N	N	N	N	N
N	N	N	N	N
N	N	X	N	N
N	N	N	N	N
N	N	N	N	N

Show here environment and distribution data!!!

2D grid-world: Starting from blue reach green

Collected state-action data distribution

2D grid-world: Starting from blue reach green

2 - Optimal policy

cell numbers: agent visitation

0	0	0	1	→ 1
0	1	→ 1	→ 1	0
0	1	X	0	0
0	1	0	0	0
1	0	0	0	0

2D grid-world: Starting from blue reach green

Optimal policy state-action distribution:

