

# RL-QL exam 2

密室摸金

LV Pushing Group

共有16个密室组成一个迷宫，密室两两相连，相邻密室之间有可进出的门，每进入一个密室都会获得相应的宝藏奖励（或暗器惩罚），这16个密室只有第15个密室有真正逃离迷宫的门（ $S_{win}$ ，记为 $S_w$ ）其余的密室都没有可以逃离迷宫的门。最外围的密室都有假的“逃离门”（ $S_{failure}$ ，记为 $S_f$ ），一旦进入就掉如悬崖了。

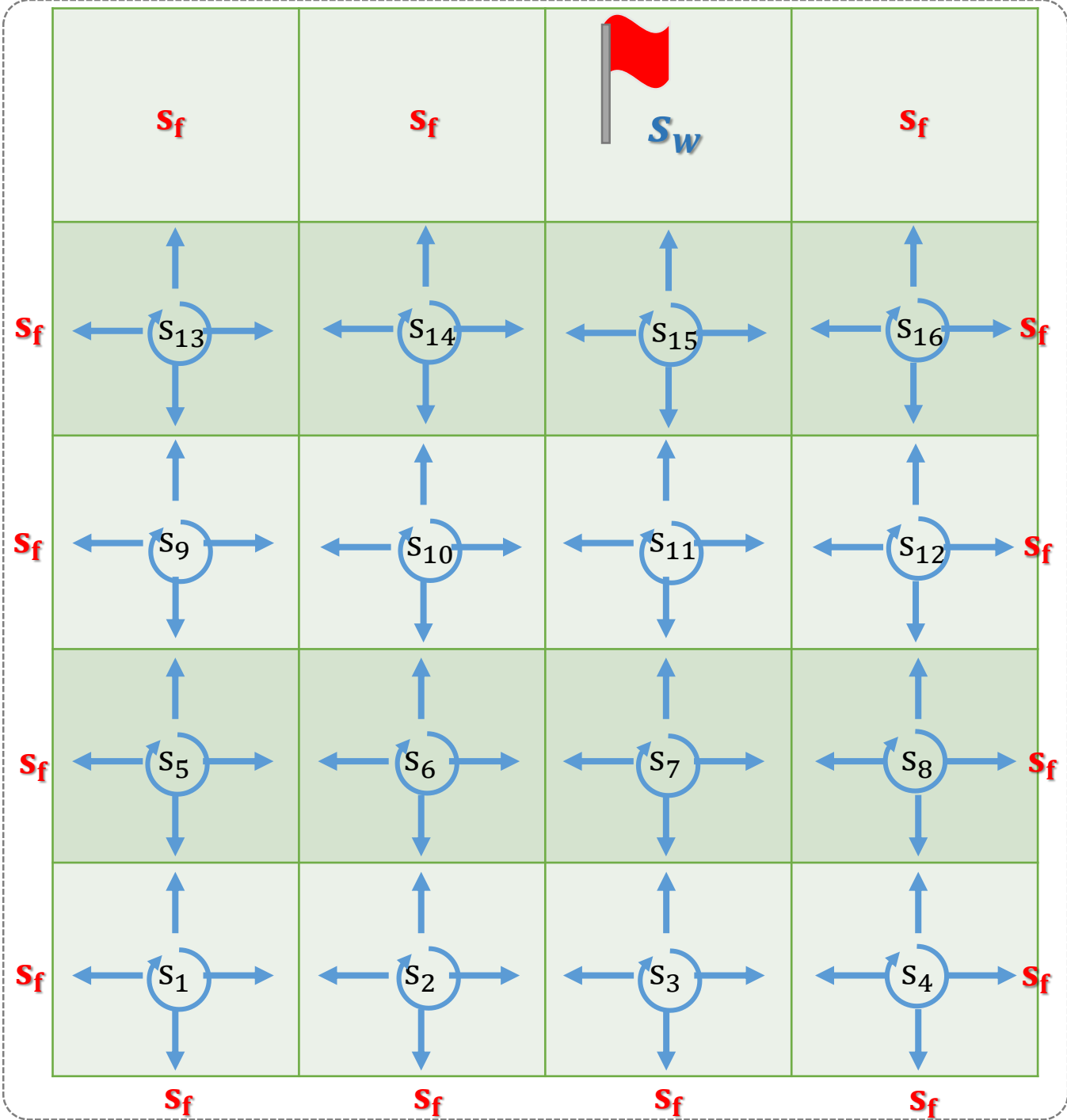
问题：将玩家随机放入一个密室，让玩家通过Q learning学习如何选择actions，使得拿到更多的奖励并获得游戏的胜利（离开迷宫）。






$$S = \{s_i\}_{i=1}^{16}$$






$$a = \{'up', 'down', 'left', 'right', 'stay'\}$$

$R$ 表和 $T$ 表已经设计好，随机初始化一个 $Q$ 表，并进行学习。

结果呈现:  
随机给定起点, 输出 最优的选择路径;



<i><b>R</b></i>	 <i><b>a</b><sub>1</sub></i>	 <i><b>a</b><sub>2</sub></i>	 <i><b>a</b><sub>3</sub></i>	 <i><b>a</b><sub>4</sub></i>	 <i><b>a</b><sub>5</sub></i>
<i>s</i> <sub>1</sub>	0.1	-0.5	-0.5	0.1	-0.2
<i>s</i> <sub>2</sub>	0.1	-0.5	-0.1	0.1	-0.2
<i>s</i> <sub>3</sub>	0.1	-0.5	-0.1	-0.1	-0.2
<i>s</i> <sub>4</sub>	0.1	-0.5	0.1	-0.5	-0.2
<i>s</i> <sub>5</sub>	0.1	-0.1	-0.5	0.1	-0.2
<i>s</i> <sub>6</sub>	0.1	-0.1	-0.1	0.1	-0.2
<i>s</i> <sub>7</sub>	0.1	-0.1	-0.1	-0.1	-0.2
<i>s</i> <sub>8</sub>	0.1	-0.1	0.1	-0.5	-0.2
<i>s</i> <sub>9</sub>	0.1	-0.1	-0.5	0.1	-0.2
<i>s</i> <sub>10</sub>	0.1	-0.1	-0.1	0.1	-0.2
<i>s</i> <sub>11</sub>	0.1	-0.1	-0.1	-0.1	-0.2
<i>s</i> <sub>12</sub>	0.1	-0.1	0.1	-0.5	-0.2
<i>s</i> <sub>13</sub>	-0.5	-0.1	-0.5	0.1	-0.2
<i>s</i> <sub>14</sub>	-0.5	-0.1	-0.1	0.1	-0.2
<i>s</i> <sub>15</sub>	0.5	-0.1	-0.1	-0.1	-0.2
<i>s</i> <sub>16</sub>	-0.5	-0.1	0.1	-0.5	-0.2

<i><b>T</b></i>	 <i><b>a</b><sub>1</sub></i>	 <i><b>a</b><sub>2</sub></i>	 <i><b>a</b><sub>3</sub></i>	 <i><b>a</b><sub>4</sub></i>	 <i><b>a</b><sub>5</sub></i>
<i>s</i> <sub>1</sub>	<i>s</i> <sub>5</sub>	<i>s<sub>f</sub></i>	<i>s<sub>f</sub></i>	<i>s</i> <sub>2</sub>	<i>s</i> <sub>1</sub>
<i>s</i> <sub>2</sub>	<i>s</i> <sub>6</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>1</sub>	<i>s</i> <sub>3</sub>	<i>s</i> <sub>2</sub>
<i>s</i> <sub>3</sub>	<i>s</i> <sub>7</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>2</sub>	<i>s</i> <sub>4</sub>	<i>s</i> <sub>3</sub>
<i>s</i> <sub>4</sub>	<i>s</i> <sub>8</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>3</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>4</sub>
<i>s</i> <sub>5</sub>	<i>s</i> <sub>9</sub>	<i>s</i> <sub>1</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>6</sub>	<i>s</i> <sub>5</sub>
<i>s</i> <sub>6</sub>	<i>s</i> <sub>10</sub>	<i>s</i> <sub>2</sub>	<i>s</i> <sub>5</sub>	<i>s</i> <sub>7</sub>	<i>s</i> <sub>6</sub>
<i>s</i> <sub>7</sub>	<i>s</i> <sub>11</sub>	<i>s</i> <sub>3</sub>	<i>s</i> <sub>6</sub>	<i>s</i> <sub>8</sub>	<i>s</i> <sub>7</sub>
<i>s</i> <sub>8</sub>	<i>s</i> <sub>12</sub>	<i>s</i> <sub>4</sub>	<i>s</i> <sub>7</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>8</sub>
<i>s</i> <sub>9</sub>	<i>s</i> <sub>13</sub>	<i>s</i> <sub>5</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>10</sub>	<i>s</i> <sub>9</sub>
<i>s</i> <sub>10</sub>	<i>s</i> <sub>14</sub>	<i>s</i> <sub>6</sub>	<i>s</i> <sub>9</sub>	<i>s</i> <sub>11</sub>	<i>s</i> <sub>10</sub>
<i>s</i> <sub>11</sub>	<i>s</i> <sub>15</sub>	<i>s</i> <sub>7</sub>	<i>s</i> <sub>10</sub>	<i>s</i> <sub>12</sub>	<i>s</i> <sub>11</sub>
<i>s</i> <sub>12</sub>	<i>s</i> <sub>16</sub>	<i>s</i> <sub>8</sub>	<i>s</i> <sub>11</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>12</sub>
<i>s</i> <sub>13</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>9</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>14</sub>	<i>s</i> <sub>13</sub>
<i>s</i> <sub>14</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>10</sub>	<i>s</i> <sub>13</sub>	<i>s</i> <sub>15</sub>	<i>s</i> <sub>14</sub>
<i>s</i> <sub>15</sub>	<i>s<sub>w</sub></i>	<i>s</i> <sub>11</sub>	<i>s</i> <sub>14</sub>	<i>s</i> <sub>16</sub>	<i>s</i> <sub>15</sub>
<i>s</i> <sub>16</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>12</sub>	<i>s</i> <sub>15</sub>	<i>s<sub>f</sub></i>	<i>s</i> <sub>16</sub>