1. Optimal policy
   1. Move right until:
      * Exit reached, or wall bounced. If bounce off wall, move up
      * Value of r being 0 leads to this instead of going right then up because benefit of hitting r is not meaningful, so best to minimize chance of -1
   2. Move up until:
      * Exit reached, or wall bounced. If bounce off wall, move right
      * Value of r being 0 leads to this instead of going right then up because benefit of hitting r makes increased chance of hitting -1 worth it
2. Value iteration

|  |  |
| --- | --- |
| 0 | 0 |
| + | 9.8 |
|  | .80 |
|  |  |