



Agentes Baseados em Utilidade - Decisões Sequenciais

Exemplo: Ambiente 4x3

■ Recompensas $R(s)$:

- Dois estados finais $R(s) = +1$ ou $R(s) = -1$
- $R(s) = -0.5$ na posição (2,2)
- $R(s) = +0.2$ na posição (4,1)
- $R(s) = r$ (valor fixo) para todos estados não terminais

| | | | | |
|---|--------------|-----------------|--------------|-----------------|
| 3 | <div>r</div> | <div>r</div> | <div>r</div> | <div>+1</div> |
| 2 | <div>r</div> | <div>-0.5</div> | <div>r</div> | <div>-1</div> |
| 1 | <div>r</div> | <div>r</div> | <div>r</div> | <div>+0.2</div> |
| | 1 | 2 | 3 | 4 |

■ Implemente o algoritmo de MDP com as equações de Bellman e retorne a política retorna para as seguintes situações:

- $r = -0.4$
- $r = -0.04$
- $r = -0.0004$
- Comente as diferenças entre as políticas retornadas

