INF623

2024/1

UFV

Inteligência Artificial

A17: Raciocínio Probabilístico IV

Plano de aula

- Modelos Ocultos de Markov
- Inferência
  - Filtragem
  - Previsão
  - Suavização
  - Explicação mais provável

UFV

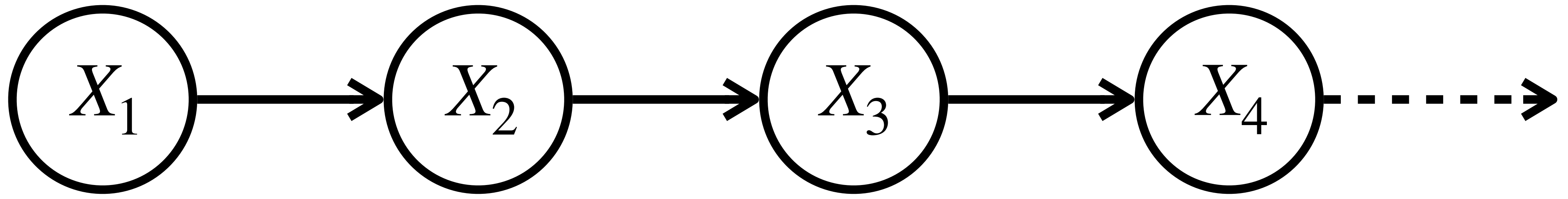
Processos (ou cadeias) de Markov

Um **processo de Markov** é uma sequência de variáveis aleatórias onde a distribuição de cada variável depende apenas de um número fixo k de variáveis anteriores $X_{t-1}, X_{t-2}, \ldots, X_{t-k}$:

Suposição de Markov

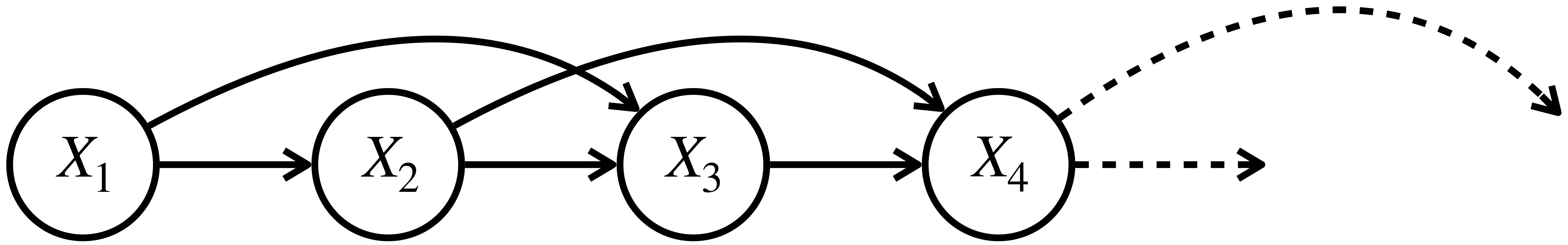
Para $k = 1$, dizemos que o modelo é de primeira ordem:

$P(X_t \mid X_{t-1})$



Para $k = 2$, dizemos que o modelo é de segunda ordem:

$P(X_t \mid X_{t-1}, X_{t-2})$



e assim por diante para $k = 3,4,...$

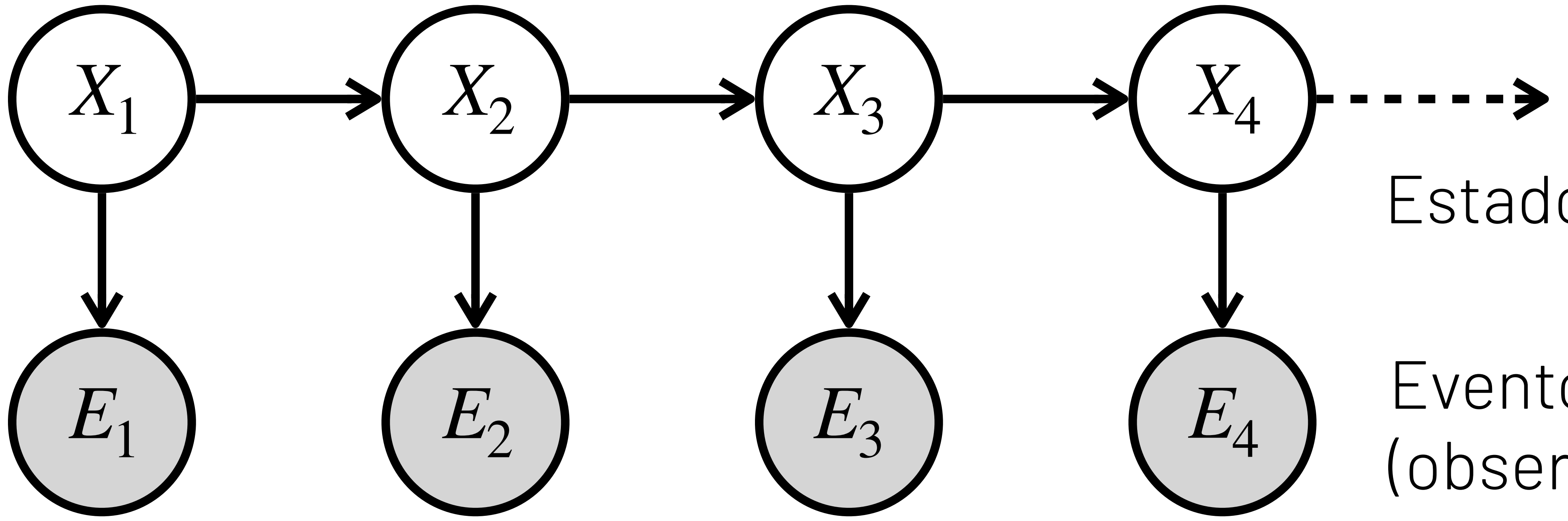
UFV

Estados ocultos

- Processos de Markov são úteis em situações onde queremos prever estados futuros a partir de um estado inicial:
  - $P(X_2 = 🌧️ \mid X_1 = ☀️)$
- No entanto, na prática, frequentemente não observamos os estados $X_1 = ☀️$ diretamente, mas temos um "*sensor*" que nos dá informações sobre eles:
  - Reconhecimento de voz
  - Rastreamento de robôs
  - Atenção do usuário

Modelos ocultos de Markov (HMM)

Modelos ocultos de Markov são processos de Markov com estados ocultos X_t, onde cada estado produz uma observação E_t.

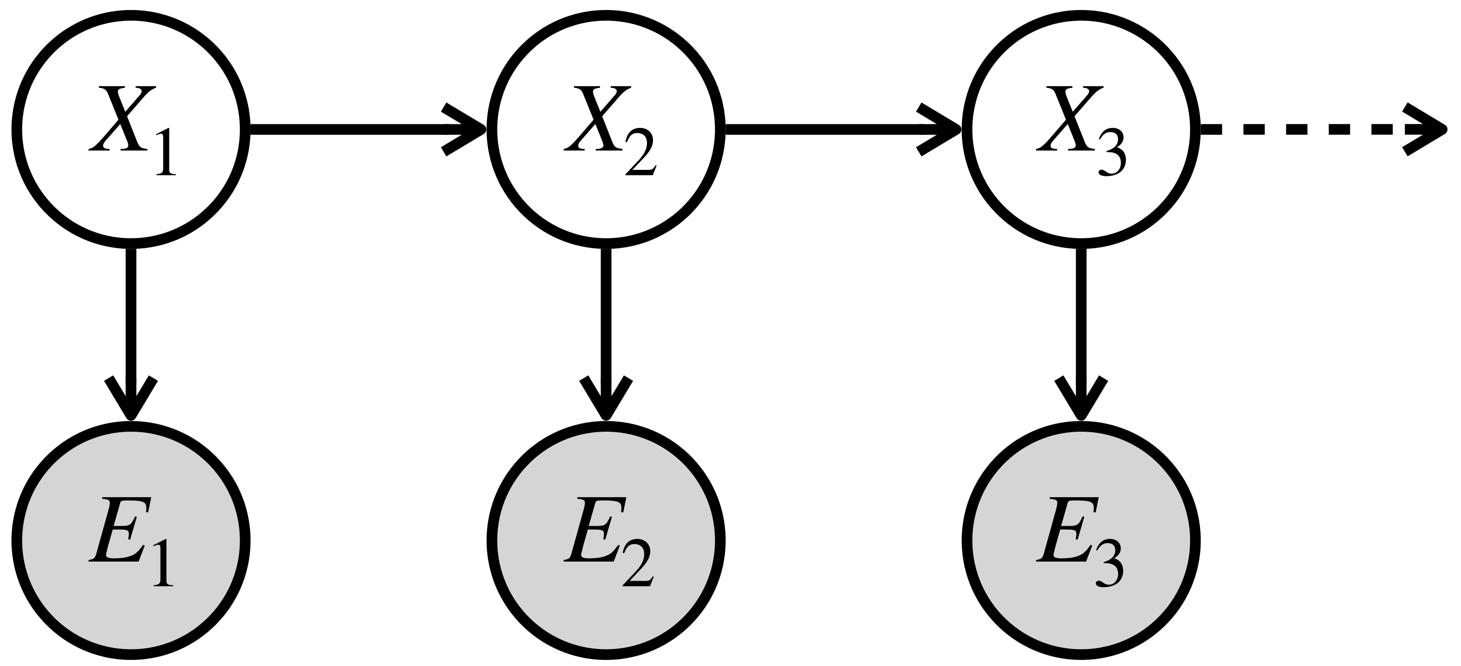


Um modelo oculto de Markov é definido por:

- Distribuição inicial – $P(X_1)$
- Modelo de transições – $P(X_t \mid X_{t-1})$
- Modelo de sensor – $P(E_t \mid X_t)$

UFV

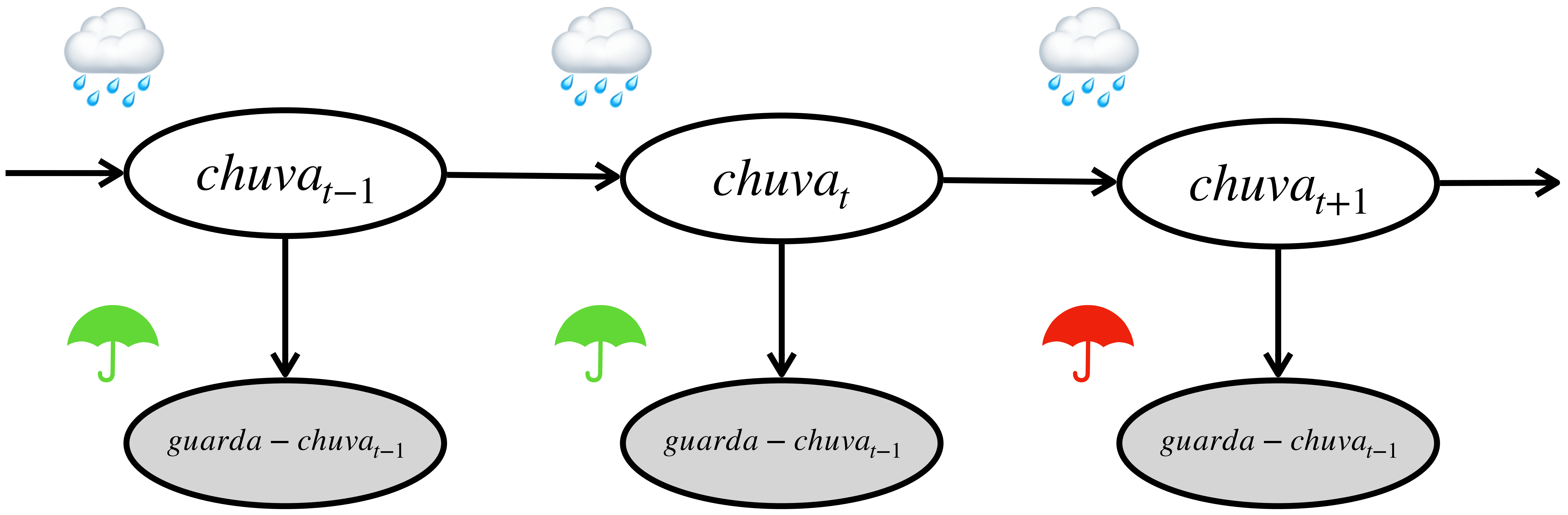
Exemplos reais de HMMs



- **Reconhecimento de voz**
  - Observações são sinais de aúdio
  - Estados são as palavras ditas em cada posição da frase
- **Traduação de texto**
  - Observações são palavras
  - Estados são opções de tradução
- **Rastreamento de robôs**
  - Observações são leituras de sensores de navegação
  - Estados são as posições no mapa

Modelo de sensor

Além do estado inicial e do modelo de transição, um HMM possui um **modelo de sensor** $P(E_t \mid X_t)$ que especifica a distribuição de probabilidade do estado X_t produzir uma observação E_t.



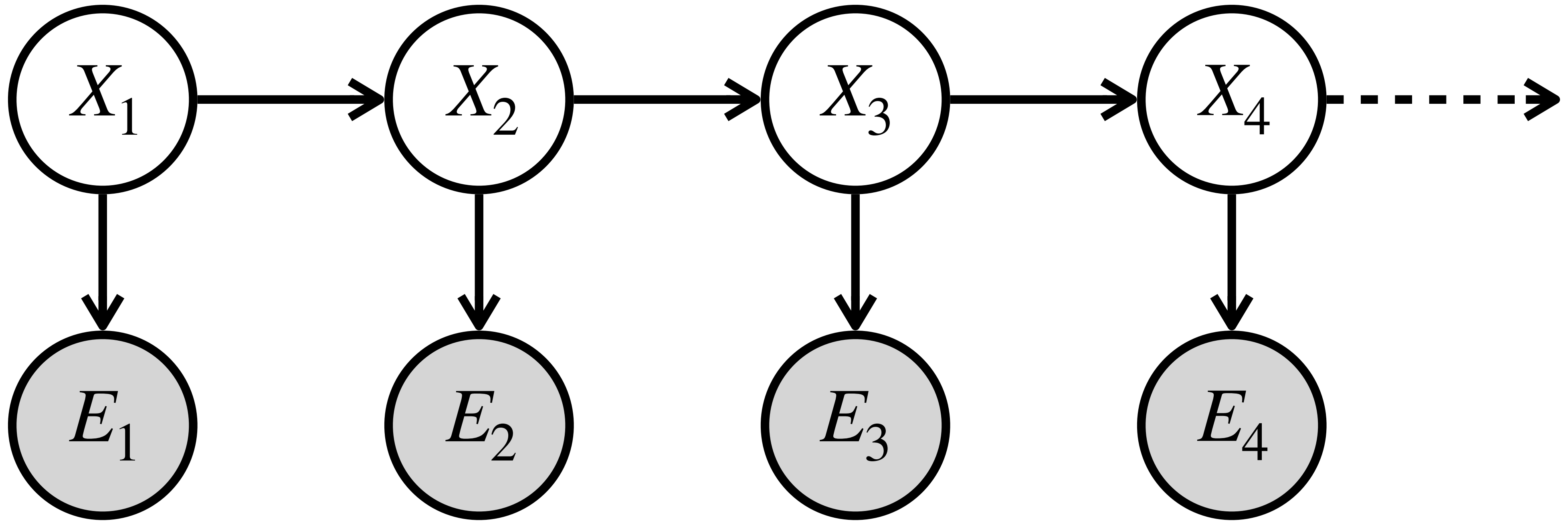
X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
☀️	☀️	0,9
☀️	🌧️	0,1
🌧️	☀️	0,3
🌧️	🌧️	0,7

X_t	E_t	$P(E_t \mid X_t)$
☀️	☂️	0,2
☀️	☂️	0,8
🌧️	☂️	0,9
🌧️	☂️	0,1

Suposição de sensores de Markov

A probabilidade do evento E_t depende apenas do estado X_t

Distribuição conjunta de HMMs



- Para 4 estados:

$$P(X_1, E_1, X_2, E_2, X_3, E_3, X_4, E_4) = P(X_1)P(E_1 | X_1)P(X_2 | X_1)P(E_2 | X_2)P(X_3 | X_2)P(E_3 | X_3)P(X_4 | X_3)P(E_4 | X_4)$$

- De uma maneira geral, para T estados:

$$P(X_1, E_1, X_2, E_2, \ldots, X_T, E_T) = P(X_1)P(E_1 | X_1)\prod_{t=2}^{T} P(X_t | X_{t-1})P(E_t | X_t)$$

Inferência

Um HMM pode ser utilizado para diferentes tarefas de inferência:

- **Filtragem**: $P(X_t \mid e_1, ..., e_t)$?

  A probabilidade de chover hoje, dado todas as observações de guarda-chuva até o momento

- **Previsão**: $P(X_{t+k} \mid e_1, ..., e_t)$?

  A probabilidade de chover depois de amanhã, dado todas as observações de guarda-chuva até o momento

- **Suavização**: $P(X_k \mid e_1, ..., e_t)$?

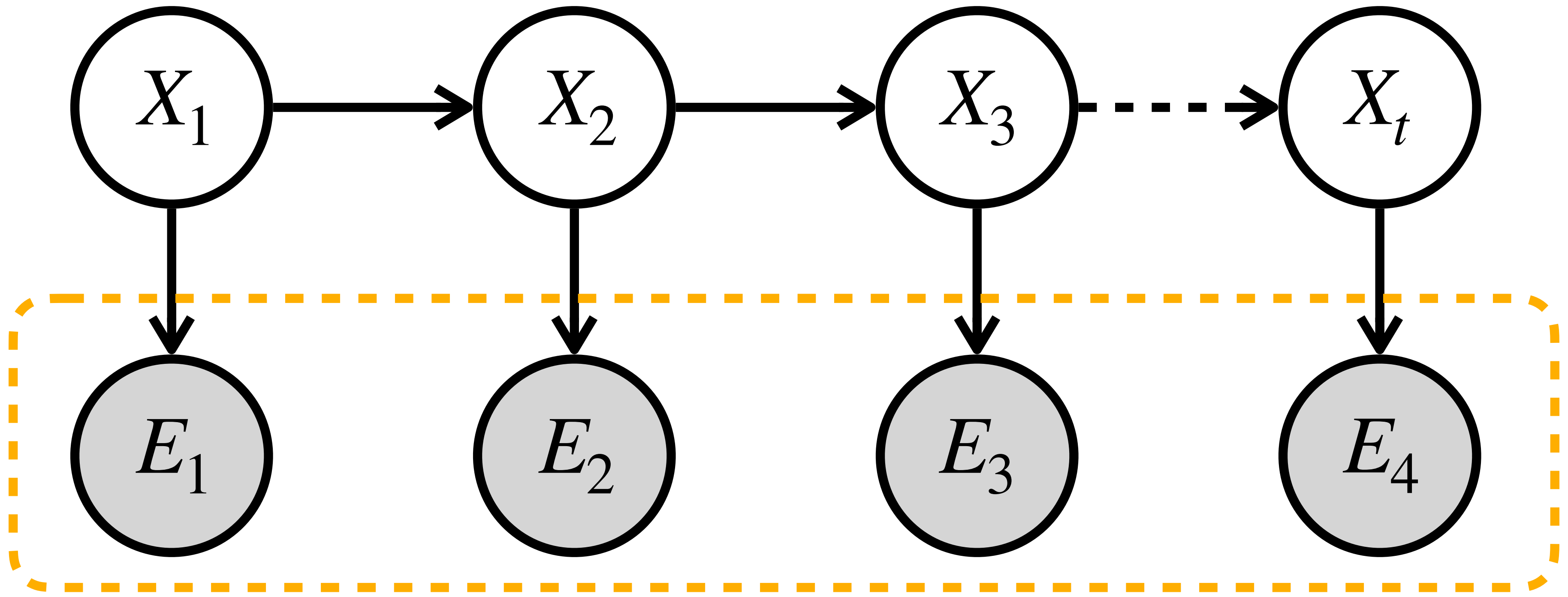
  A probabilidade de que choveu na última quarta-feira, dado todas as observações de guarda-chuva até o momento

- **Explicação mais provável**: $argmax_{x_{1:t}} \; P(x_{1:t} \mid e_1, ..., e_t)$

  A sequência de estados mais provável de ter gerado as observações obtidas.

Filtragem

Qual a probabilidade $B_t(X) = P(X_t \mid e_1, ..., e_t)$, dado todas as observações até o momento?

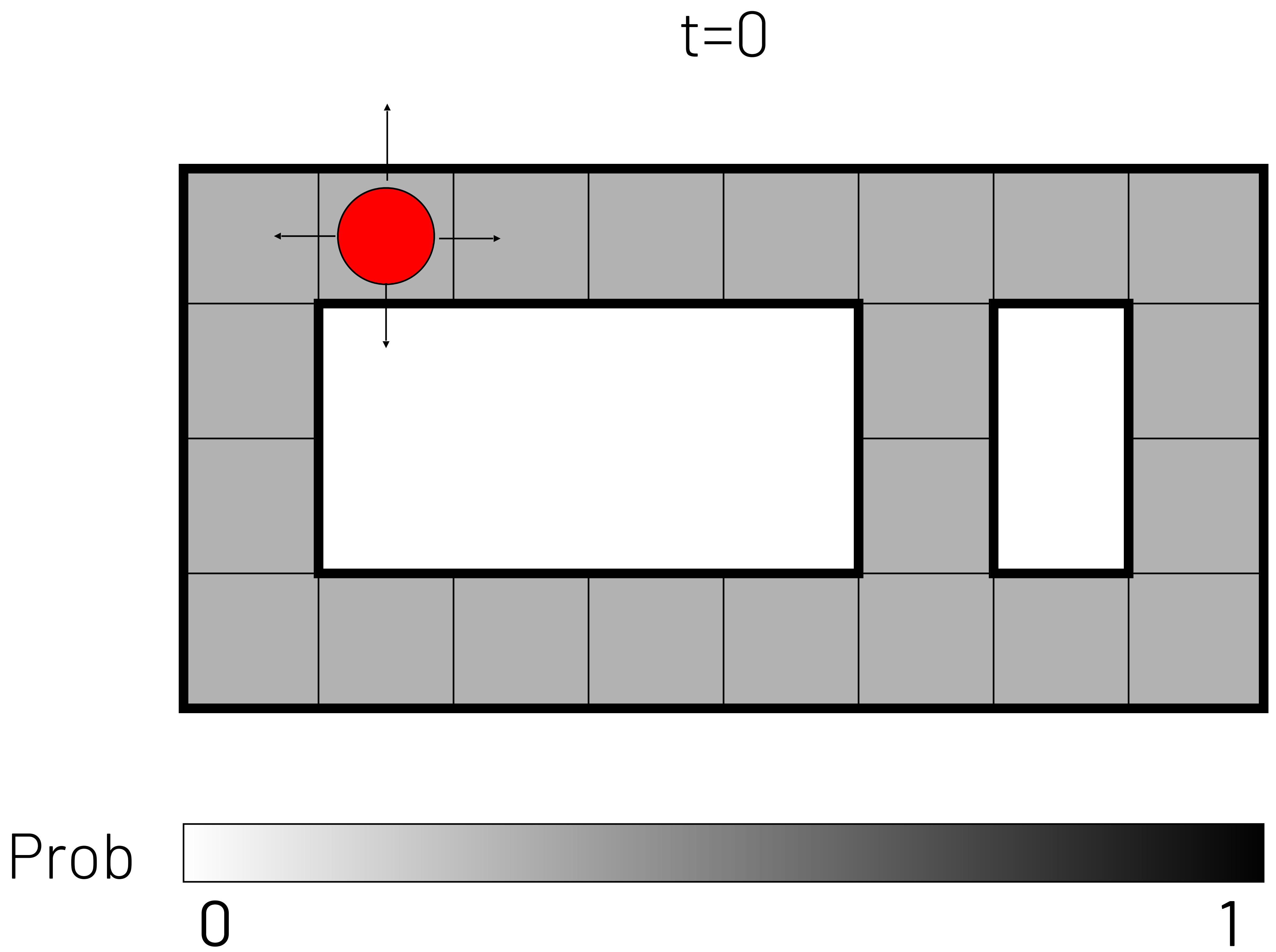


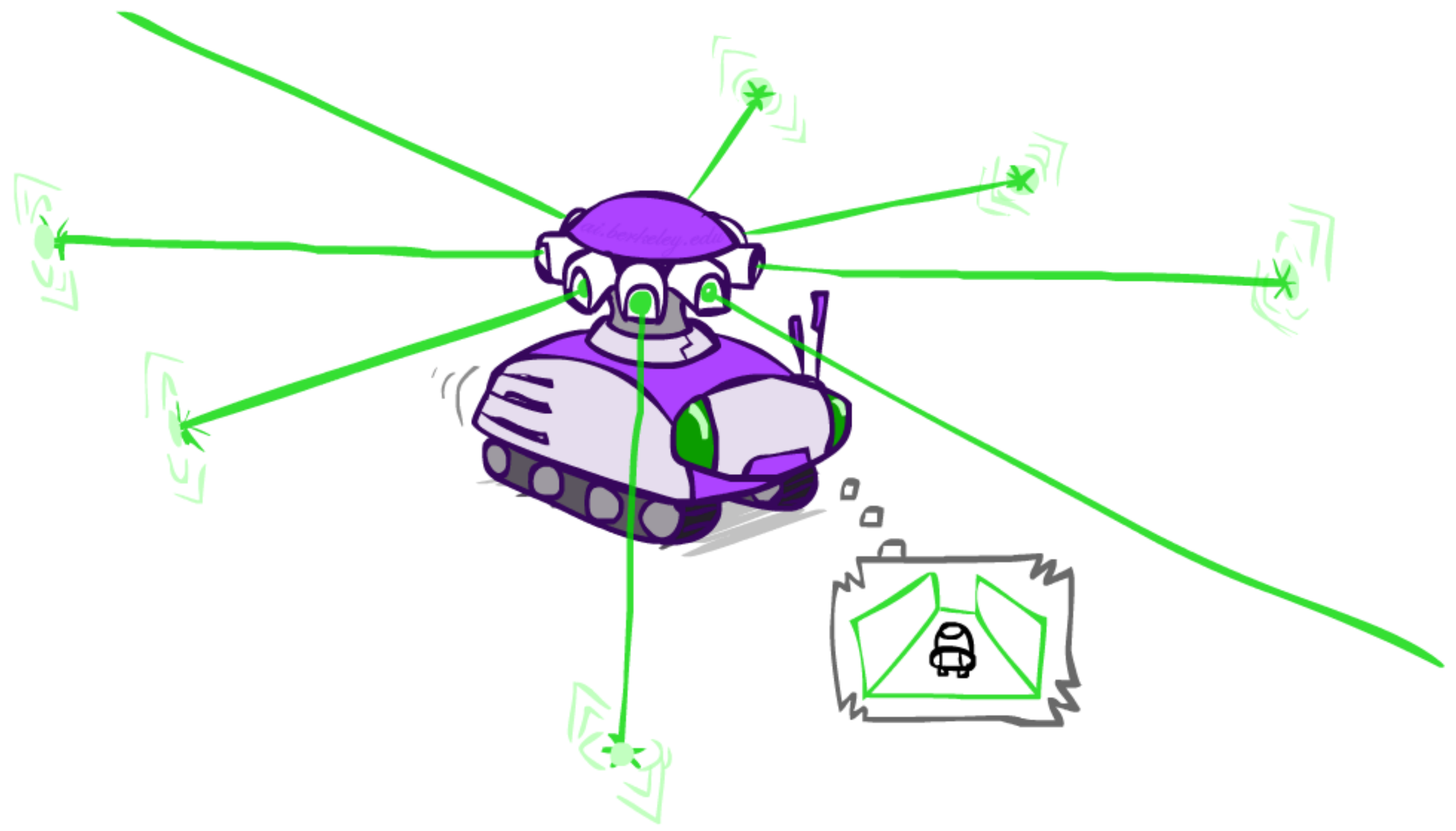
Queremos: $P(X_t \mid e_1, ..., e_t) = B_t(X)$

Observamos

- Começar $B_1(X) = P(X_1 \mid e_1)$, normalmente uniforme
- Atualizamos $B_t(X)$ a cada nova evidência e_t
- O **filtro de Kalman** foi inventado na década de 60 e implementado pela primeira vez como método de estimativa de trajetória para o programa Apollo.

Exemplo: Localização de robôs





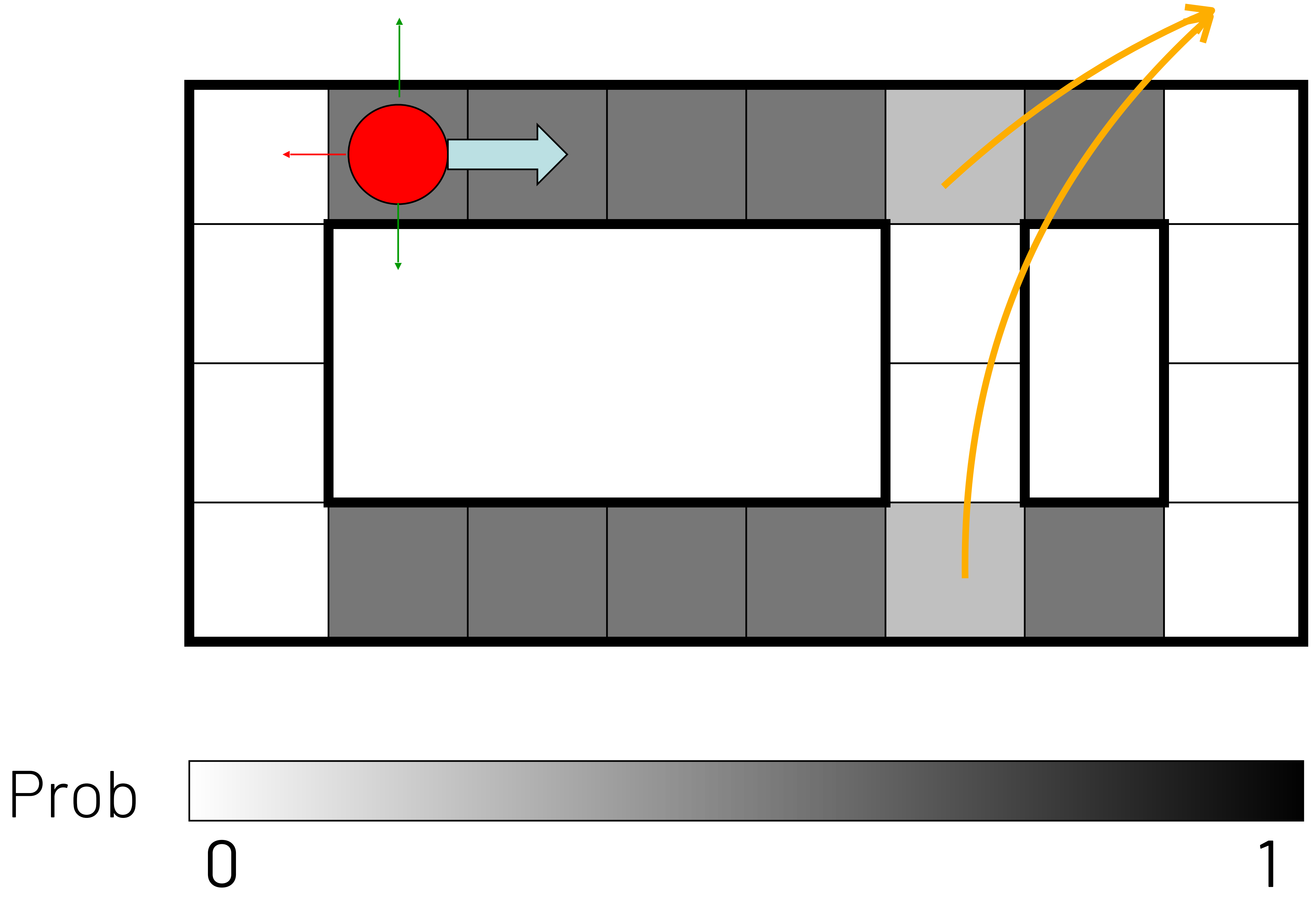
Modelo de sensor: consegue ler em quais direções há uma parede, nunca mais do que 1 erro

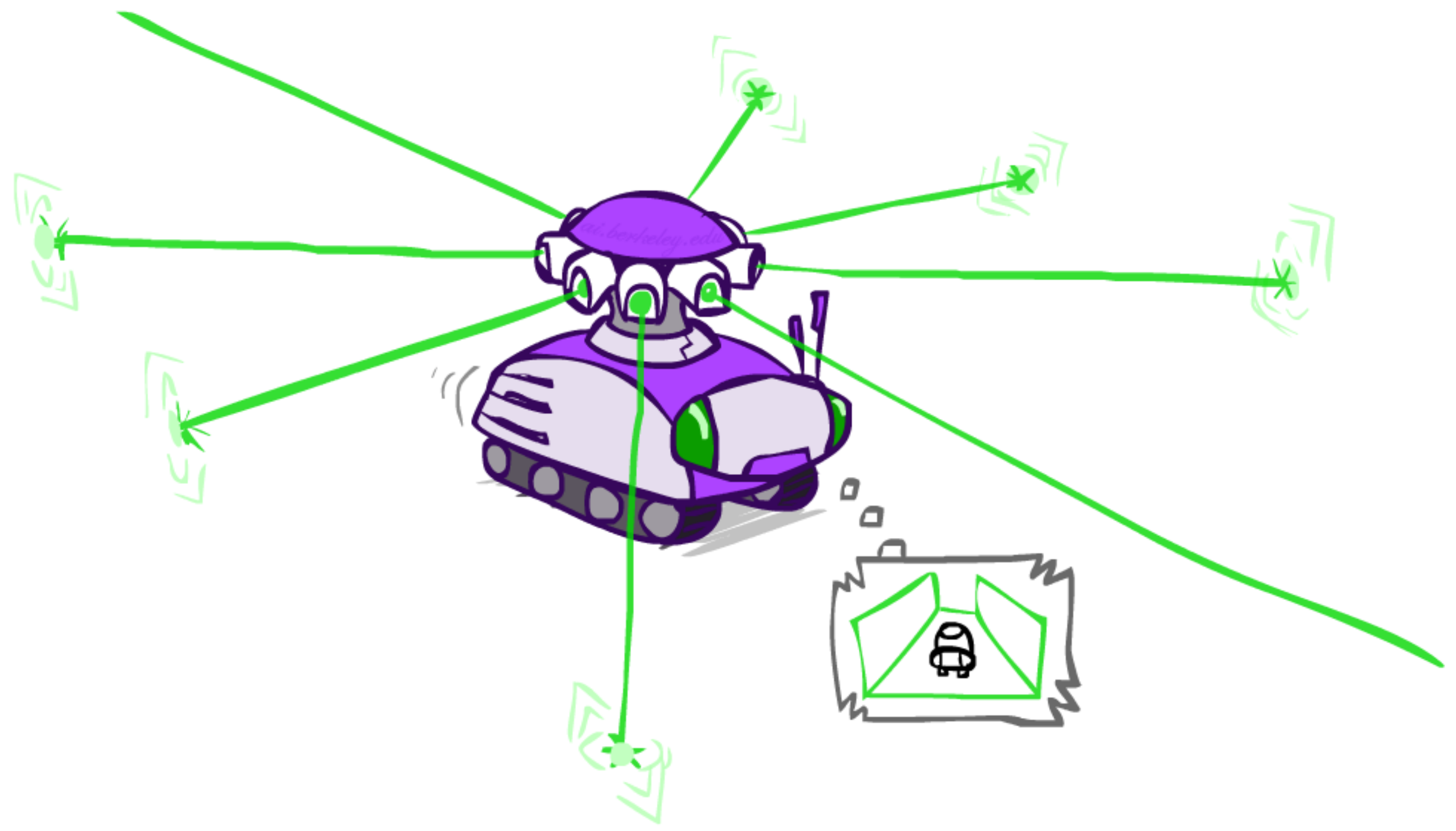
Modelo de movimento: tem uma baixa probabilidade de falha.

Exemplo: Localização de robôs

t=1

Cinza mais claro: consistente com as medidas do sensor, mas pode ter sido erro



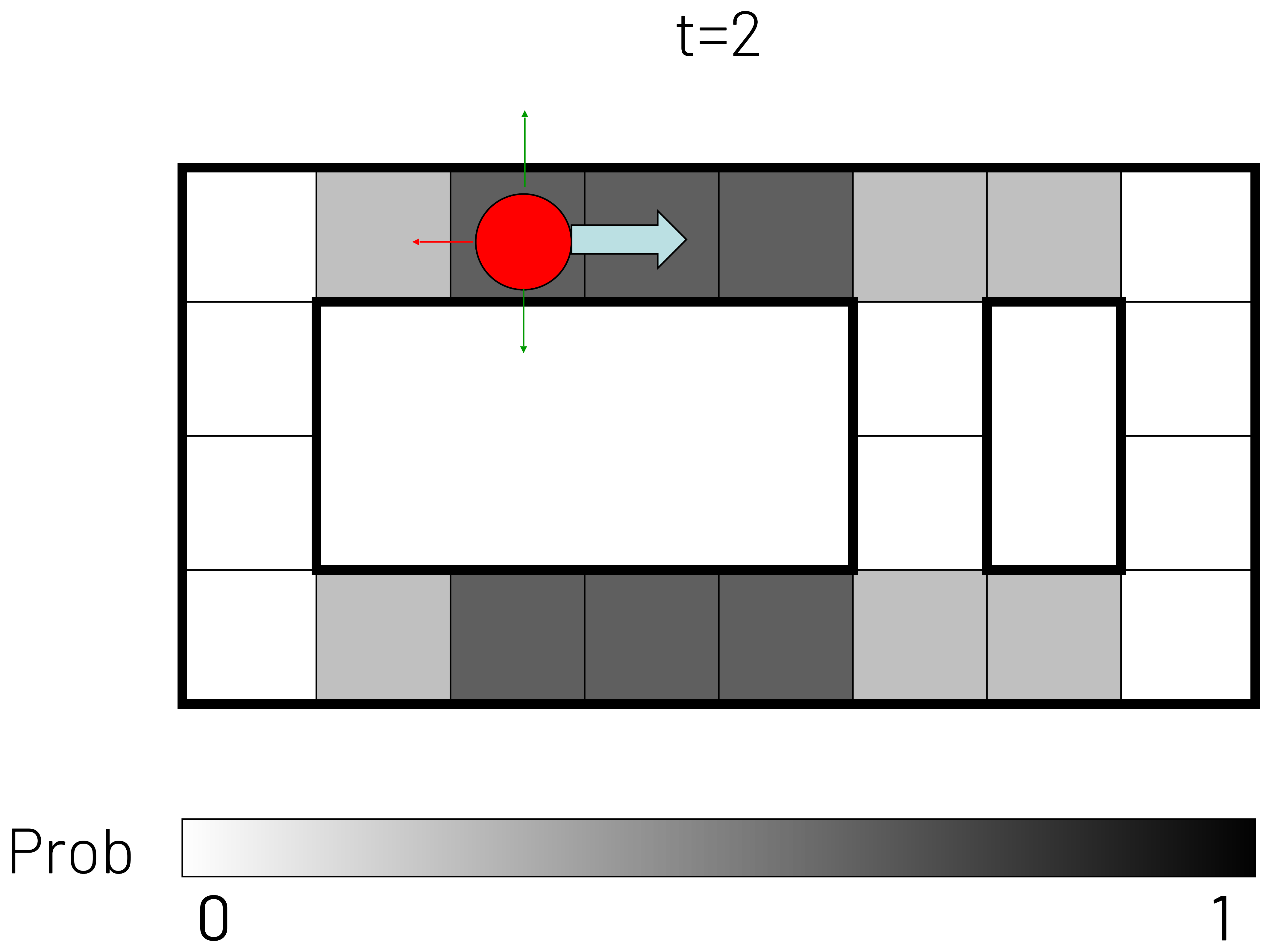


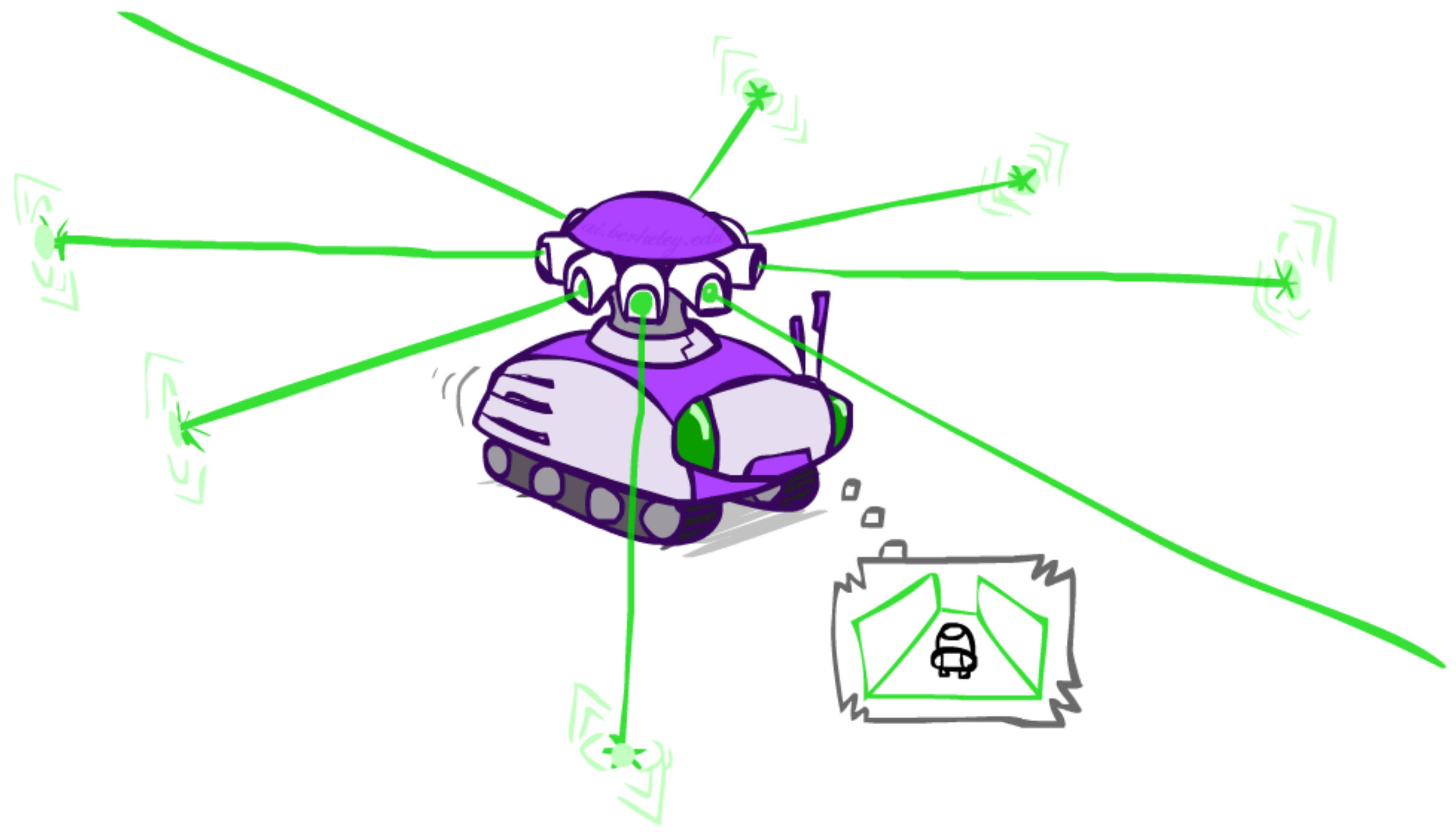
Modelo de sensor: consegue ler em quais direções há uma parede, nunca mais do que 1 erro

Modelo de movimento: tem uma baixa probabilidade de falha.

UFV

Exemplo: Localização de robôs

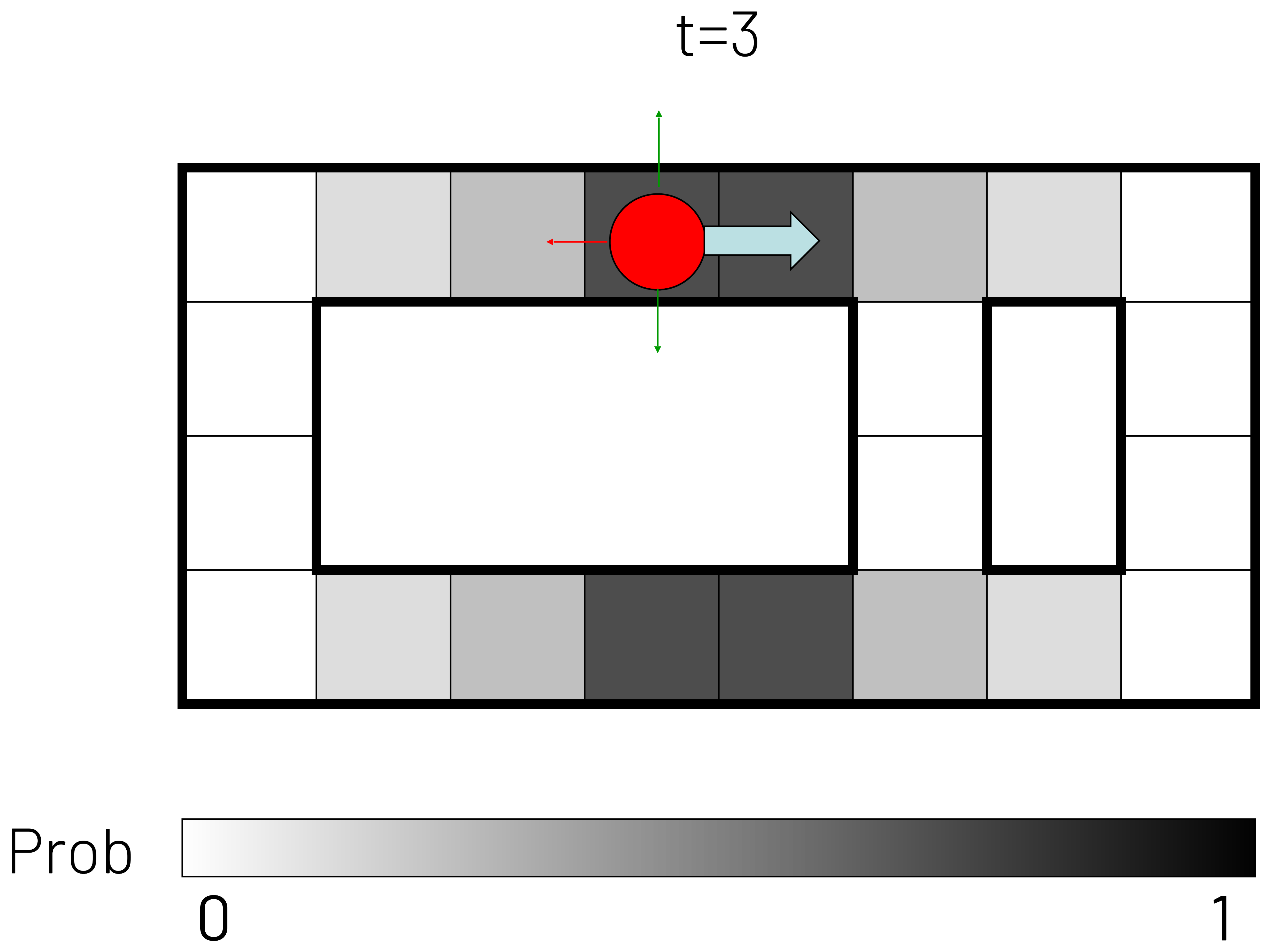


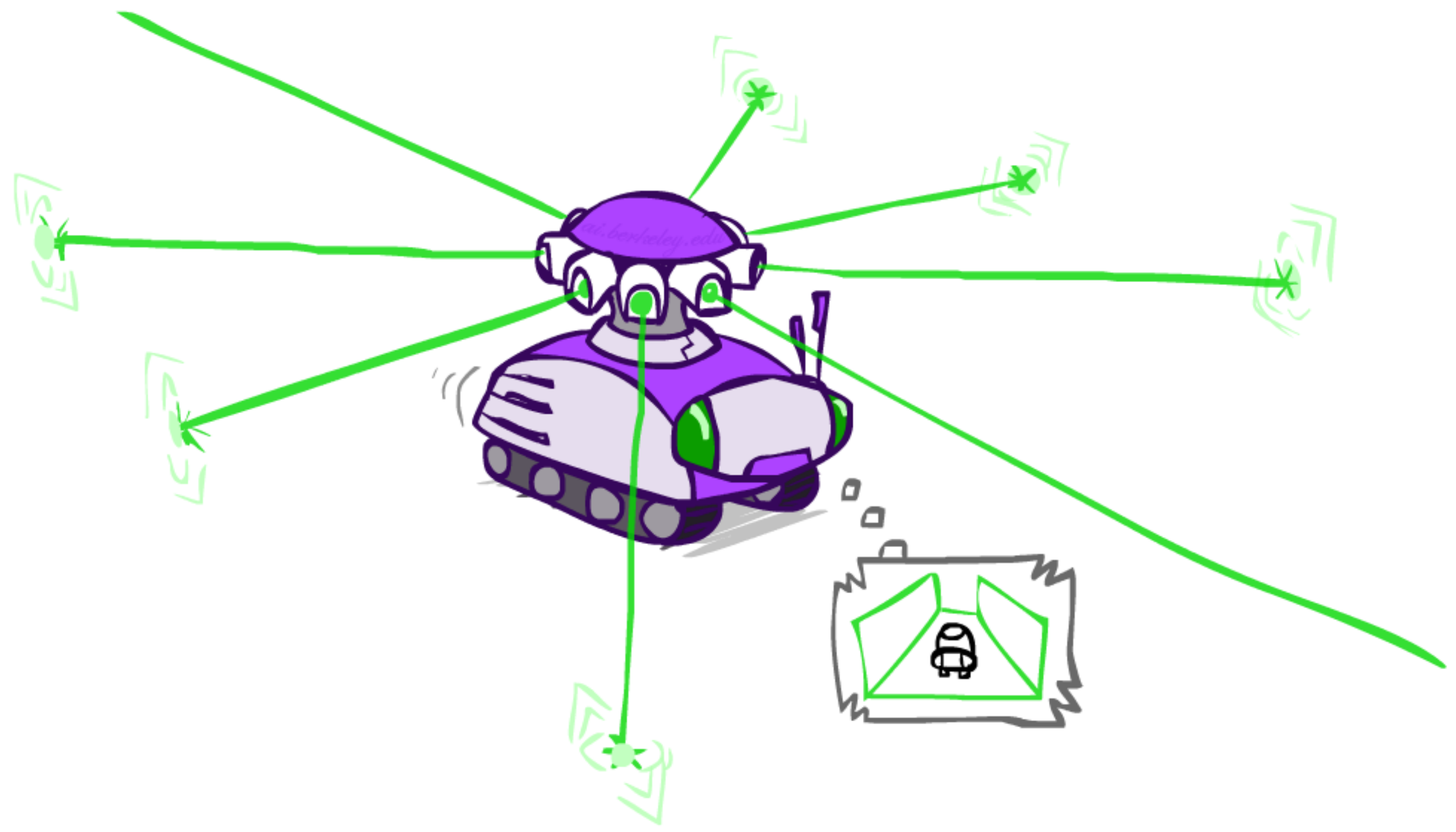


Modelo de sensor: consegue ler em quais direções há uma parede, nunca mais do que 1 erro

Modelo de movimento: tem uma baixa probabilidade de falha.

Exemplo: Localização de robôs

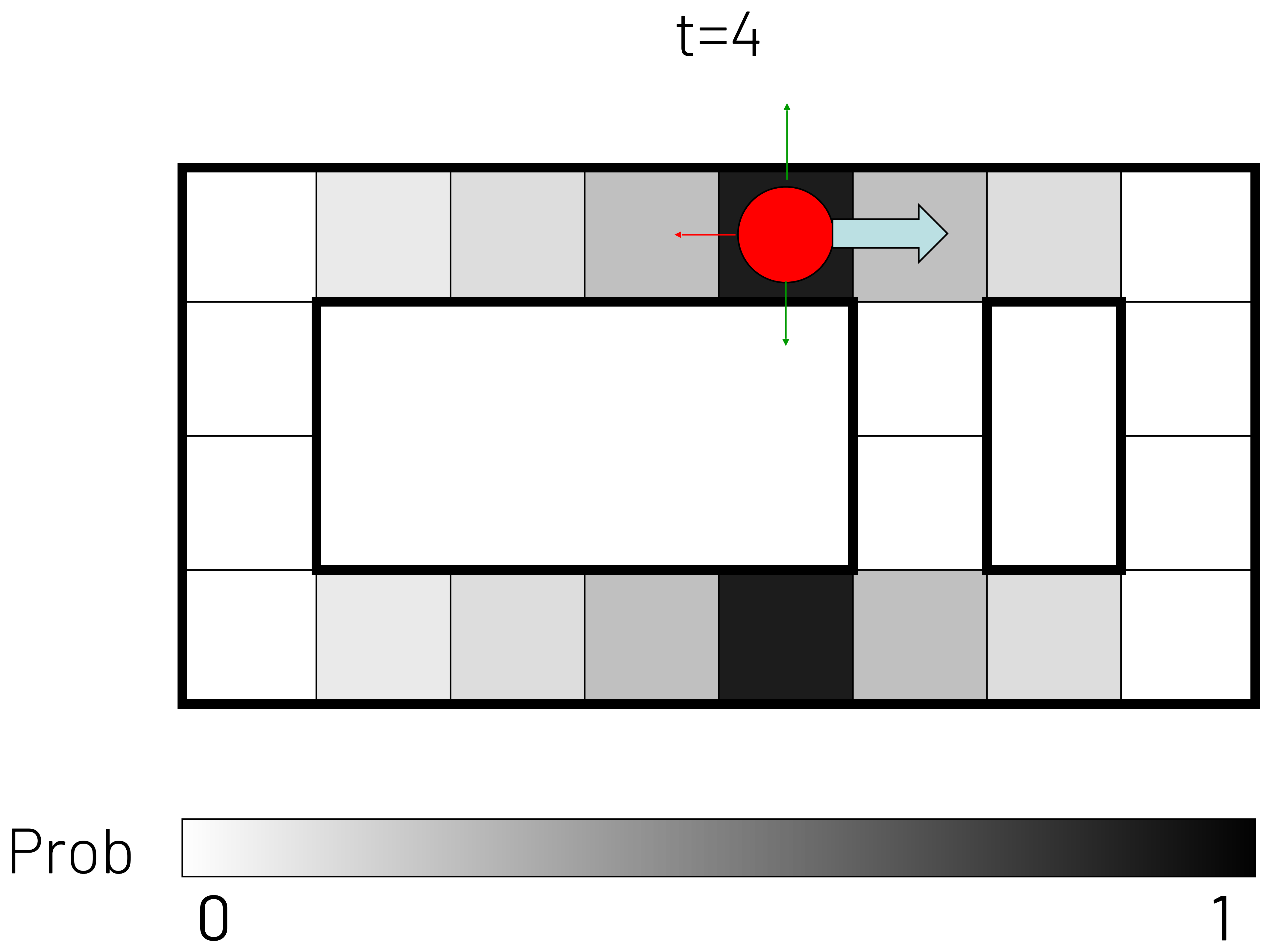


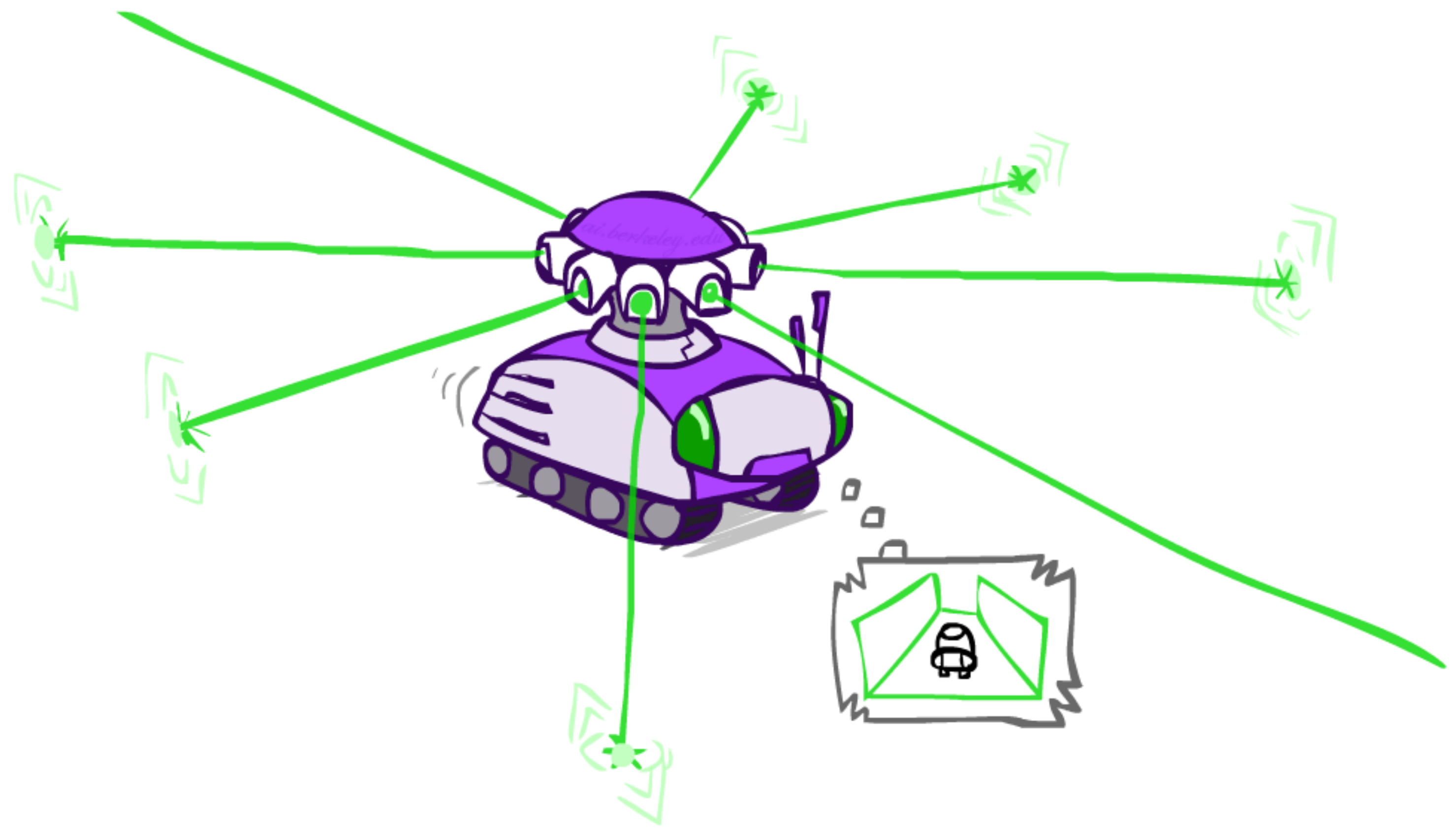


Modelo de sensor: consegue ler em quais direções há uma parede, nunca mais do que 1 erro

Modelo de movimento: tem uma baixa probabilidade de falha.

Exemplo: Localização de robôs

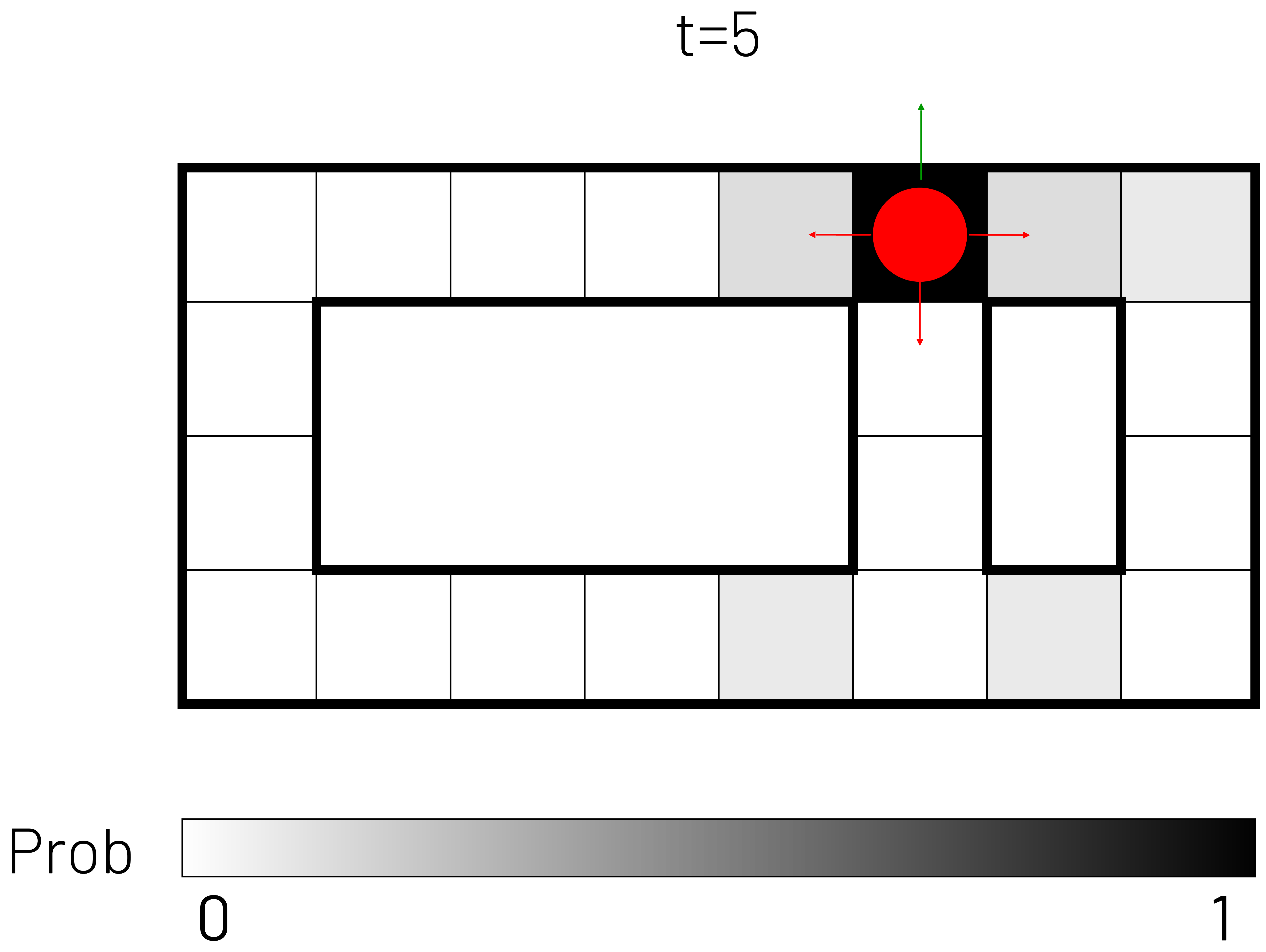


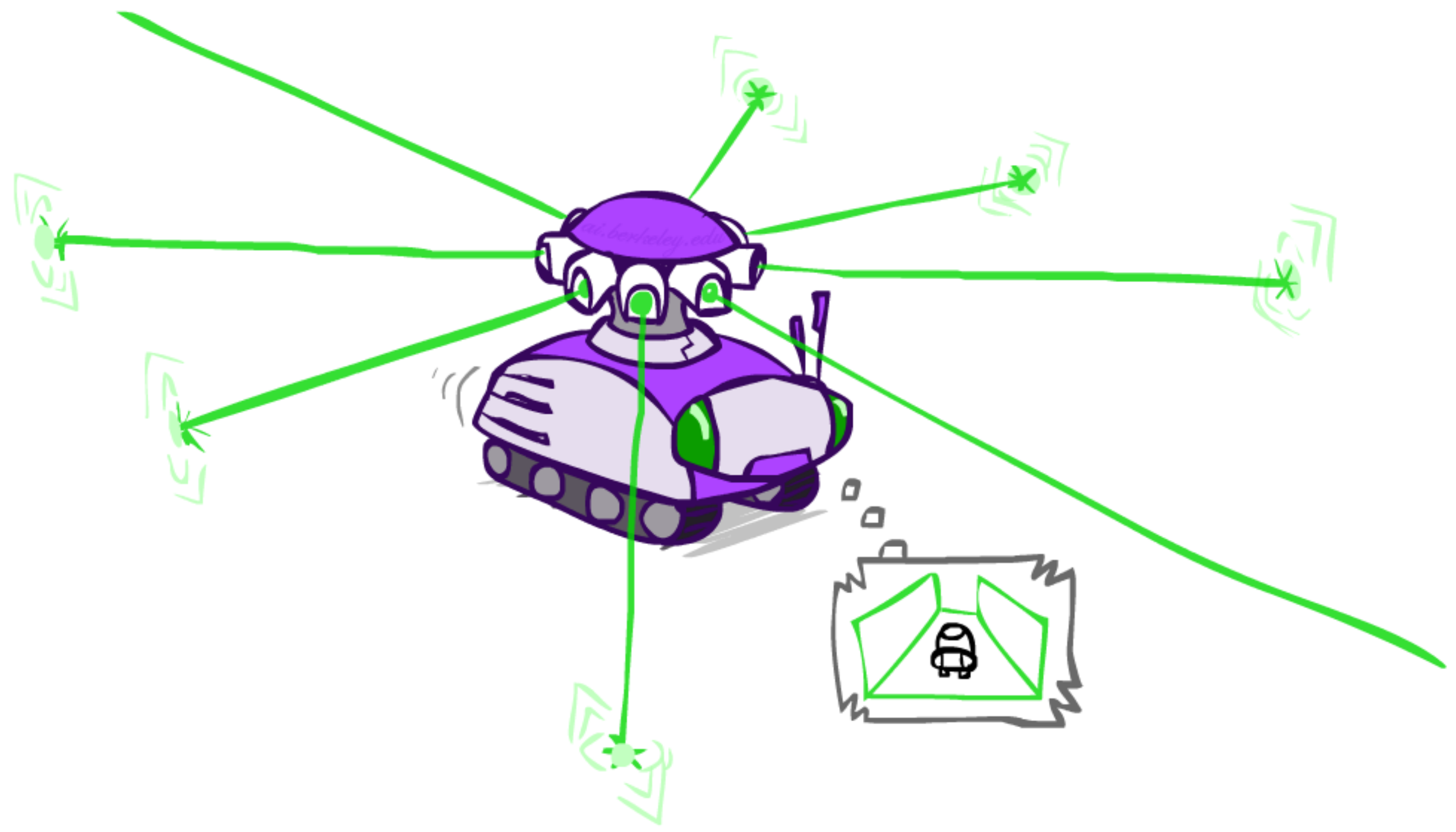


Modelo de sensor: consegue ler em quais direções há uma parede, nunca mais do que 1 erro

Modelo de movimento: tem uma baixa probabilidade de falha.

Exemplo: Localização de robôs



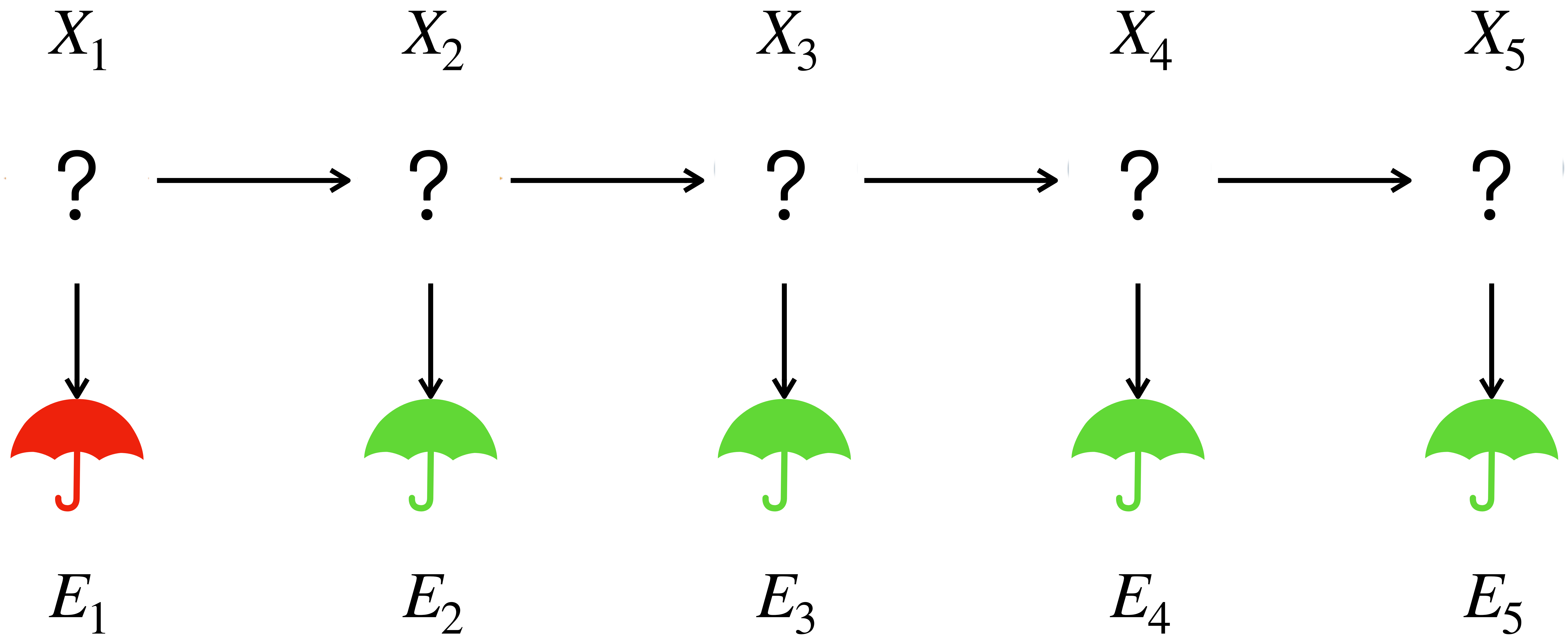


Modelo de sensor: consegue ler em quais direções há uma parede, nunca mais do que 1 erro

Modelo de movimento: tem uma baixa probabilidade de falha.

Explicação mais provável

Qual é a sequência de estados mais provável de ter gerado as observações obtidas?



Próxima aula

A18: Processos de decisão de Markov I

Formalização matemática, exemplos, política, utilidade, descontos