INF623

2024/1

UFV

Inteligência Artificial

A16: Raciocínio Probabilístico III

Plano de aula

- Processos (ou cadeias) de Markov
- Modelo de transição
- Inferência
- Amostragem
- Distribuição estacionária

Exemplo 1: previsão do tempo

Considere o problema de previsão do tempo (ensolarado, chuvoso) do dia seguinte dado o panorama dos $k = 4$ dias passados:

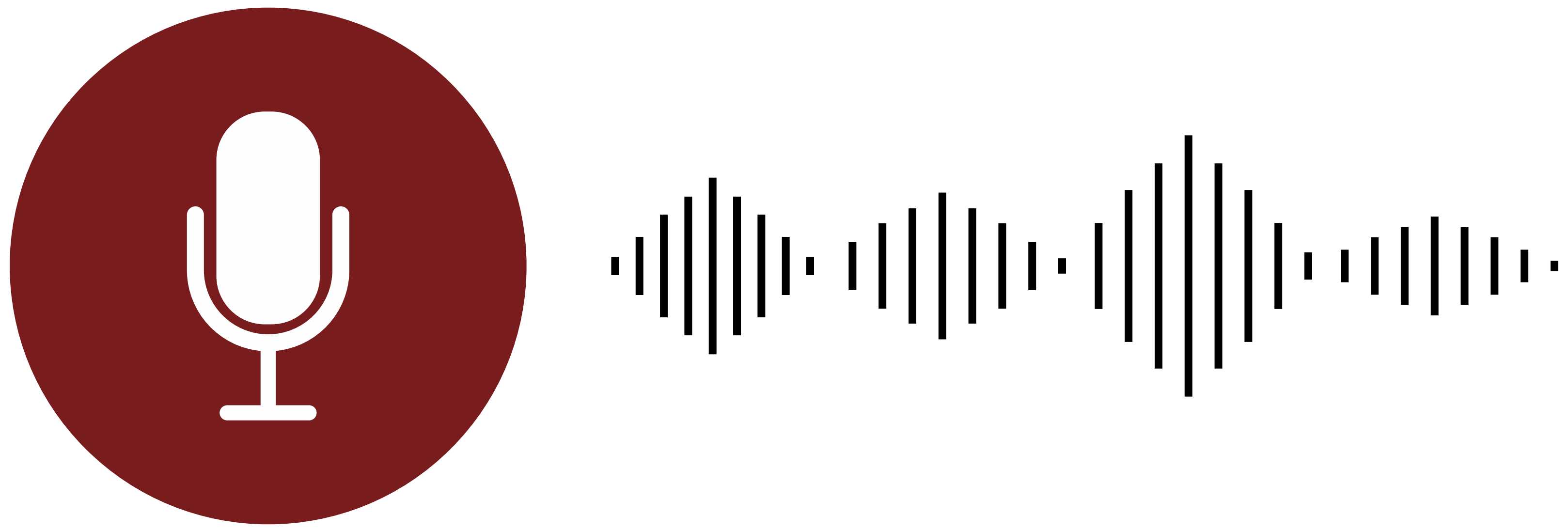
Inferência probabilística no tempo

Em Inteligência Artificial, frequentemente precisamos fazer inferência no tempo (ou espaço):

- Reconhecimento de voz
- Localização de robôs
- Atenção do usuário
- Monitoramento médico

- Para isso, é necessário introduzir tempo (ou espaço) nos modelos probabilísticos.

Reconhecimento de voz



P("A palavra rubrica tem **acento**?") = 0.23

P("A palavra rubrica tem **assento**?") = 0.1

Inferência probabilística no tempo

Até agora, vimos como fazer inferência $P(Q \mid E)$ dado variáveis Q e E de consulta e evidência, respectivamente.

$$P(Q \mid E) = \alpha \sum_h P(Q, h, e)$$

- Essa formalização não representa a relação temporal entre as variáveis.
- **Processos de Markov nos permitem estabelecer essas relações!**

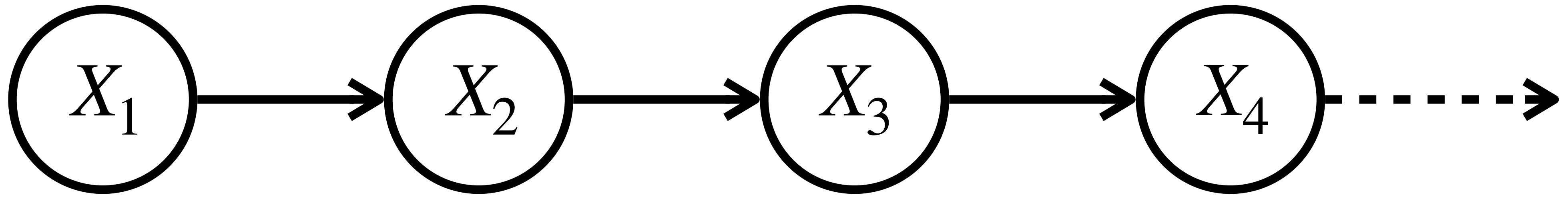
Processos (ou cadeias) de Markov

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

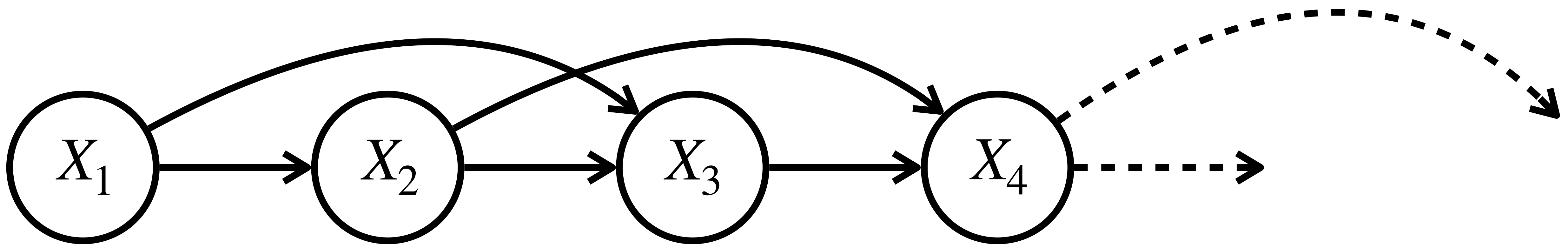
Para $k = 1$, dizemos que o modelo é de primeira ordem:

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



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 $k = 3,4,\ldots$

UFV

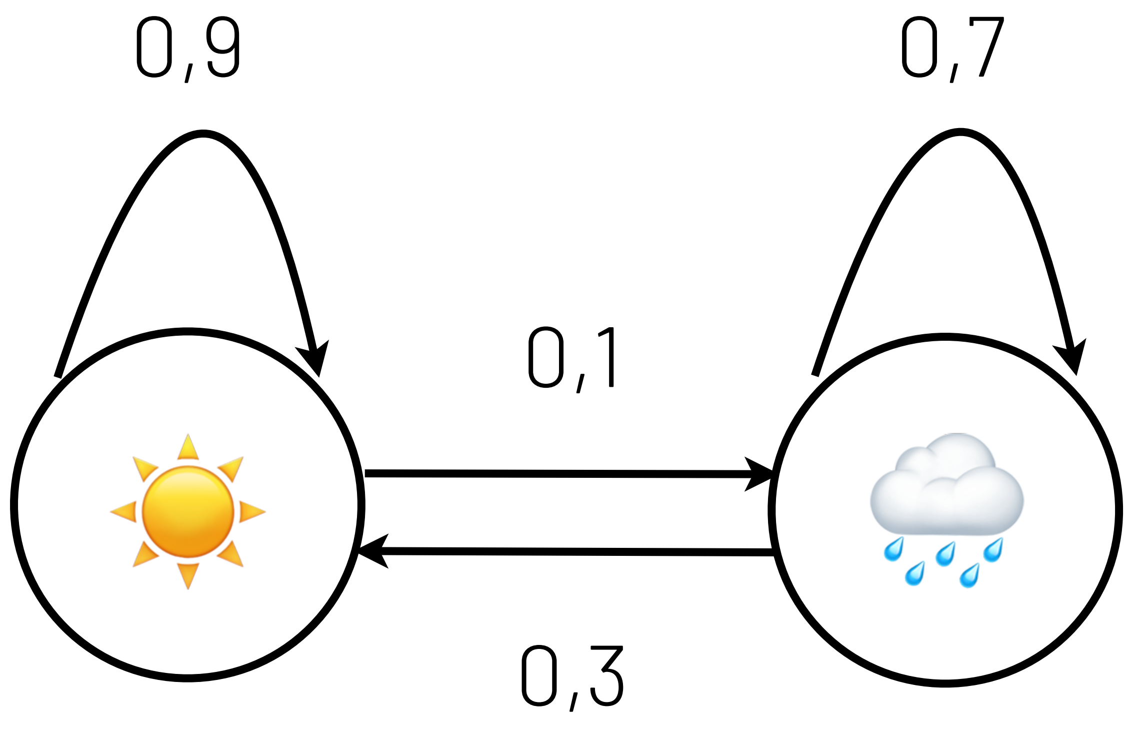
Modelo de Transição

Um processo de Markov é definido por uma distribuição incial $P(X_1)$ e um **modelo de transição** $P(X_t \mid X_{t-1})$ que especifique a distribuição de probabilidade do próximo evento com base nos valores possíveis do evento atual.

Modelo de Transição [para previsão do tempo]

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

(a) Tabela



(b) Grafo

Distribuição conjunta de processos de Markov

$$P(X_t | X_{t-1}) \qquad X_1 \to X_2 \to X_3 \to X_4 \dashrightarrow$$

- Pela regra da cadeia:

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2 | X_1)P(X_3 | X_1, X_2)P(X_4 | X_1, X_2, X_3)$$

- Pela suposição de Markov para $k = 1$:

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2 | X_1)P(X_3 | X_2)P(X_4 | X_3)$$

- De uma maneira geral, para T estados:

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

Inferência

[Lembrete]
$$P(X = x) = \sum_y P(X = x, Y = y)$$
$$P(A, B) = P(A \mid B)P(B)$$

Qual a distribuição de probabilidades $P(X_t)$ no tempo t?

Simulação para frente:

Samemos $P(X_1, X_2, \ldots, X_T) = P(X_1)\prod_{t=2}^{T} P(X_t \mid X_{t-1})$

Marginalização

$$P(X_t) = \sum_{x_{t-1}} P(X_t, x_{t-1})$$

Regra do produto

$$= \sum_{x_{t-1}} P(X_t \mid x_{t-1})P(x_{t-1})$$

Como $P(X_1)$ é dado, podemos calcular $P(X_t)$ iterativamente, começando por $P(X_2)$ a partir de $P(X_1)$, depois $P(X_3)$ a partir de $P(X_2)$ e assim sucessivamente até $P(X_t)$ a partir de $P(X_{t-1})$

Exemplo

[Lembrete]

$$P(X_t) = \sum_{x_{t-1}} P(X_t | x_{t-1}) P(x_{t-1})$$

Qual a probabilidade de fazer sol no segundo dia $P(X_2 = sol)$?

Assuma que $P(X_1 = sol) = 1.0$

$$P(X_t = x_t) = \sum_{x_{t-1}} P(x_t | x_{t-1}) P(x_{t-1})$$

$$P(X_2 = sol) = P(X_2 = sol | X_1 = sol) P(X_1 = sol) + P(X_2 = sol | X_1 = chuva) P(X_1 = chuva)$$

$$P(X_2 = sol) = 0{,}9 \times 1{,}0 + 0{,}3 \times 0{,}0 = 0{,}9$$

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

Exemplo

[Lembrete]

$$P(x_t) = \sum_{x_{t-1}} P(x_t \mid x_{t-1}) P(x_{t-1})$$

Qual a distribuição de probabilidade $P(X_4)$ no quarto dia?

Assuma que $P(X_1) = \begin{bmatrix} 1.0 \\ 0.0 \end{bmatrix}$

$$\begin{aligned} P(X_2 = sol) = \; & P(X_2 = sol \mid X_1 = sol) P(X_1 = sol) + \\ & P(X_2 = sol \mid X_1 = chuva) P(X_1 = chuva) \end{aligned}$$

$$\begin{aligned} P(X_2 = chuva) = \; & P(X_2 = chuva \mid X_1 = sol) P(X_1 = sol) + \\ & P(X_2 = chuva \mid X_1 = chuva) P(X_1 = chuva) \end{aligned}$$

$$P(X_2) = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix} \longrightarrow P(X_3) = \begin{bmatrix} 0.84 \\ 0.16 \end{bmatrix} \longrightarrow P(X_4) = \begin{bmatrix} 0.804 \\ 0.196 \end{bmatrix}$$

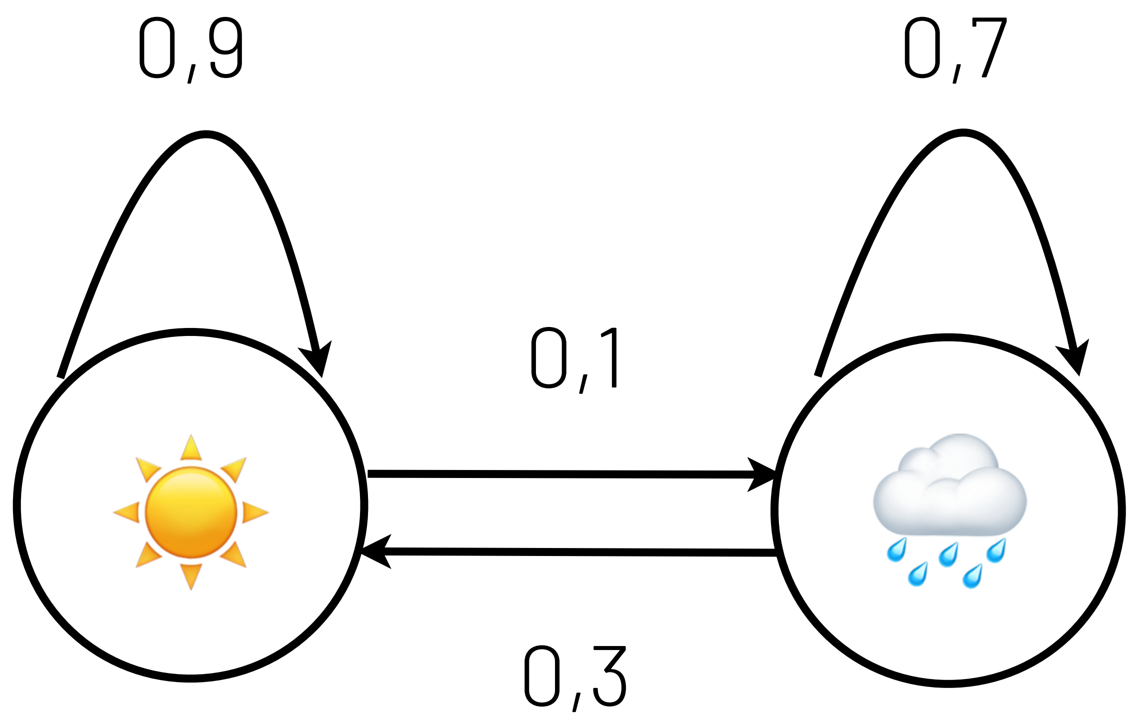
X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
☀️	☀️	0,9
☀️	🌧️	0,1
🌧️	☀️	0,3
🌧️	🌧️	0,7

Amostragem

Podemos gerar sequências com processos de Markov, começando com um estado inicial e amostrando novos estados de acordo com o modelo de transição

$t = 0$ $\quad t = 1$ $\quad t = 2$ $\quad \ldots$

Modelo de Transição



Distribuição estacionária

Qual a distribuição de probabilidade $P(X_\infty)$ no tempo $t = \infty$?

- Considerando estado inicial igual a sol

$$\underset{P(X_1)}{\begin{bmatrix}1.0\\0.0\end{bmatrix}} \quad \underset{P(X_2)}{\begin{bmatrix}0.9\\0.1\end{bmatrix}} \quad \underset{P(X_3)}{\begin{bmatrix}0.84\\0.16\end{bmatrix}} \quad \underset{P(X_4)}{\begin{bmatrix}0.804\\0.196\end{bmatrix}} \quad \dashrightarrow \quad \underset{P(X_\infty)}{\begin{bmatrix}0.75\\0.25\end{bmatrix}}$$

- Considerando estado inicial igual a chuva

$$\underset{P(X_1)}{\begin{bmatrix}0.0\\1.0\end{bmatrix}} \quad \underset{P(X_2)}{\begin{bmatrix}0.3\\0.7\end{bmatrix}} \quad \underset{P(X_3)}{\begin{bmatrix}0.48\\0.52\end{bmatrix}} \quad \underset{P(X_4)}{\begin{bmatrix}0.588\\0.412\end{bmatrix}} \quad \dashrightarrow \quad \underset{P(X_\infty)}{\begin{bmatrix}0.75\\0.25\end{bmatrix}}$$

- Considerando qualquer distribuição inicial $P(X_1)$

$$\underset{P(X_1)}{\begin{bmatrix}p\\1-p\end{bmatrix}} \quad \dashrightarrow \quad \underset{P(X_\infty)}{\begin{bmatrix}0.75\\0.25\end{bmatrix}}$$

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

UFV

Distribuição estacionária

Qual a distribuição de probabilidade $P(X_\infty)$ no tempo $t = \infty$?

$$P_\infty(sun) = P(sun \mid sun)P_\infty(sun) + P(sun \mid rain)P_\infty(rain)$$
$$P_\infty(rain) = P(rain \mid sun)P_\infty(sun) + P(rain \mid rain)P_\infty(rain)$$

$$P_\infty(sun) = 0{,}9P_\infty(sun) + 0{,}3P_\infty(rain)$$
$$P_\infty(rain) = 0{,}1P_\infty(sun) + 0{,}7P_\infty(rain)$$

$$P_\infty(sun) = 3P_\infty(rain)$$
$$P_\infty(rain) = 1/3P_\infty(rain)$$

Além disso:

$$P_\infty(sun) + P_\infty(rain) = 1$$

$$P_\infty(sun) = \frac{3}{4}$$

$$P_\infty(sun) = \frac{1}{4}$$

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

Distribuição estacionária

A distribuição de probabilidade $P(X_\infty)$ no tempo $t = \infty$ é chamada de **distribuição estacionária** e possui a seguinte propriedade:

$$P_\infty(X) = P_{\infty+1}(X) = \sum_x P(X|x)P_\infty(x)$$

- **Para a maioria dos processos de Markov:**
  - A influência da distribuição inicial diminui mais e mais ao longo tempo
  - A distribuição obtida ao final é independente da distribuição inicial

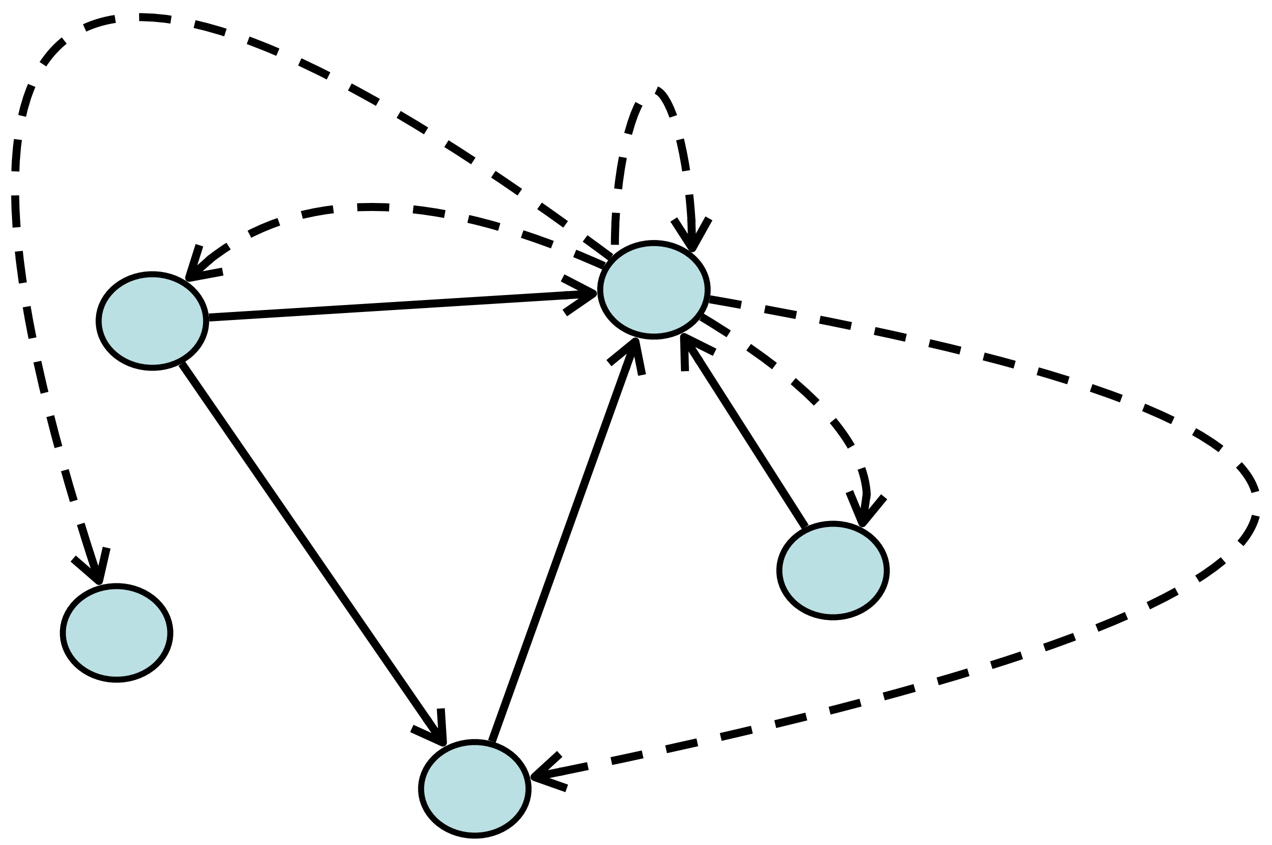
UFV

Aplicação de distribuições estacionárias

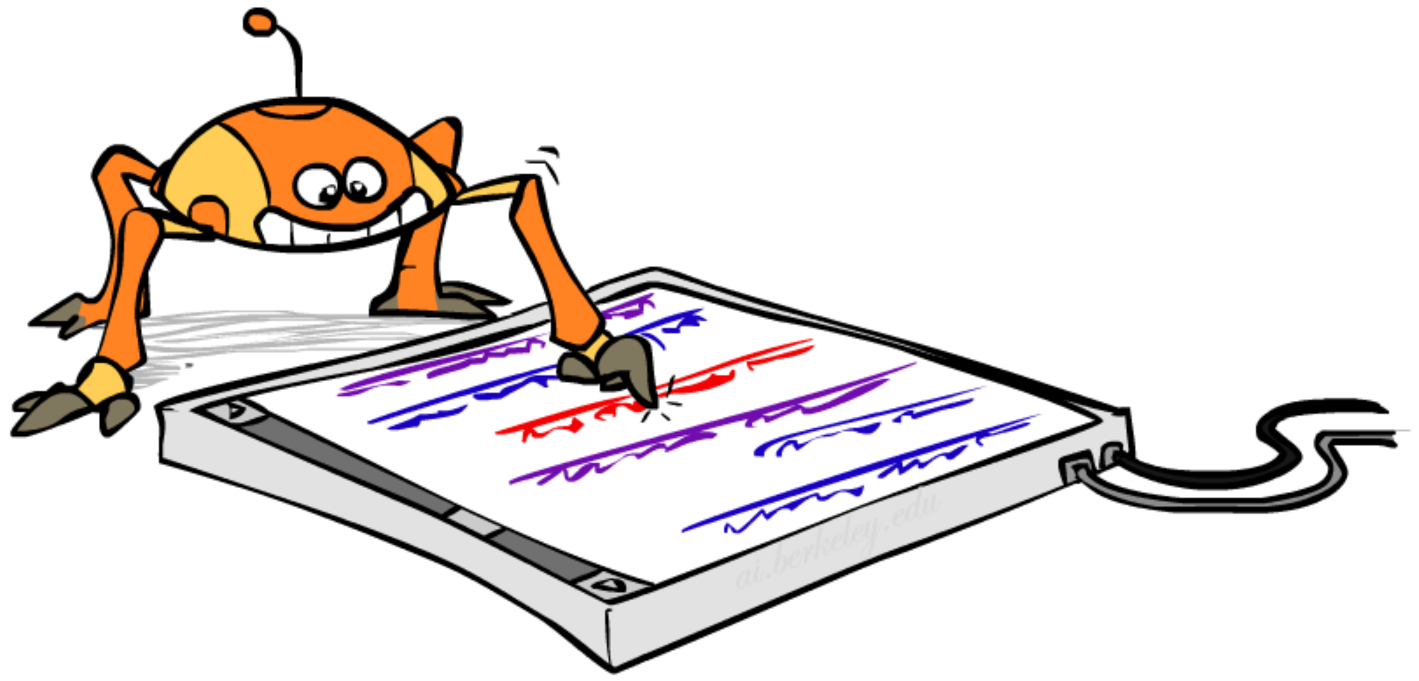
- **Algoritmo PageRank para ranquear páginas web**

  Modela um usuário navegando na internet, seguindo links aleatoriamente

  - Cada página web é um estado
  - Distribuição inicial: uniforme
  - Transições:
    - Com probabilidade p, escolha uma página web aleatoriamente (linhas tracejadas, nem todas estão mostradas)
    - Com probabilidade $1 - p$, siga um link de saída aleatório (linhas sólidas)



- **Distribuição estacionária**
  - Terá a probabilidade do usuário terminar a navegação em cada página web
    - E.g. probabilidade de terminar na página da Wikipedia
  - Abordagem consideravelmente robusta a link spams



Próxima aula

A17: Raciocínio Probabilístico IV

Estados ocultos, modelos ocultos de Markov