



Скрытая марковская модель (Hidden Markov Model)

Наталья Корепанова, korepanova@bigdatateam.org

Инструктор, [BigDataTeam](#)

Data Science Tech Lead, [IPONWEB](#)

Профиль на [LinkedIn](#)

9 декабря 2020, Москва



MEGAFON



- ▶ Марковская цепь
- ▶ Скрытая марковская модель (СММ)
- ▶ Три задачи СММ



Марковская цепь



Марковская цепь (1906)

Рассмотрим систему, которую в любой момент времени можно описать одним из N состояний S_1, \dots, S_N .

В каждый момент времени система может изменить свое состояние или остаться в том же состоянии. Моменты времени обозначим $t=1, 2, \dots$, а состояние в момент времени t обозначим q_t .

Марковское свойство: $P[q_t = S_j \mid q_{t-1} = S_i, q_{t-2} = S_{k'} \dots] = P[q_t = S_j \mid q_{t-1} = S_i]$ для любого t .

Тогда систему можно описать множеством вероятностей $a_{ij} = P[q_t = S_j \mid q_{t-1} = S_i]$, где $1 \leq i, j \leq N$, $a_{ij} \geq 0$ и $\sum_j a_{ij} = 1$.

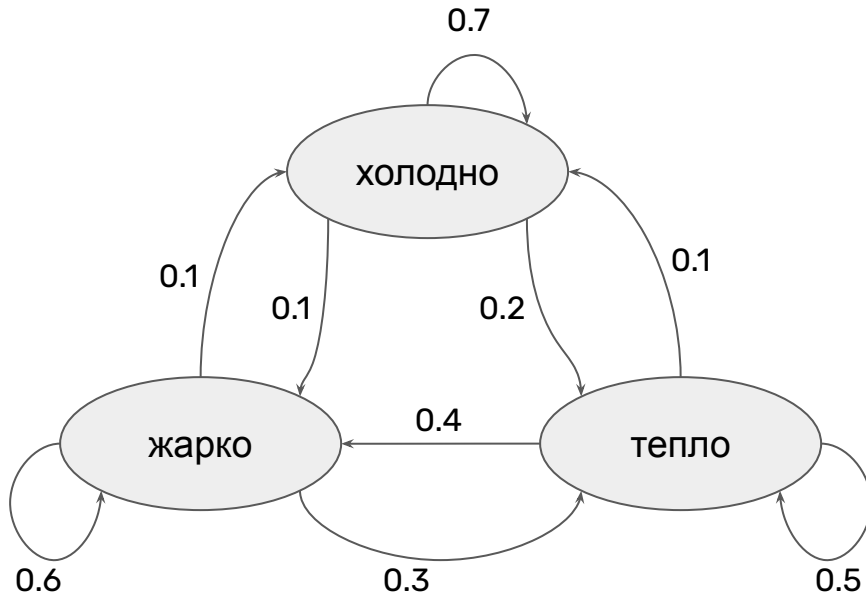
И вероятностями начального состояния системы: $\pi_i = P[q_1 = S_i]$, где $1 \leq i \leq N$, $\pi_i \geq 0$ и $\sum_i \pi_i = 1$.



Марковская цепь

Лето в Москве

Состояния: S_1 = Холодно, S_2 = Тепло, S_3 = Жарко





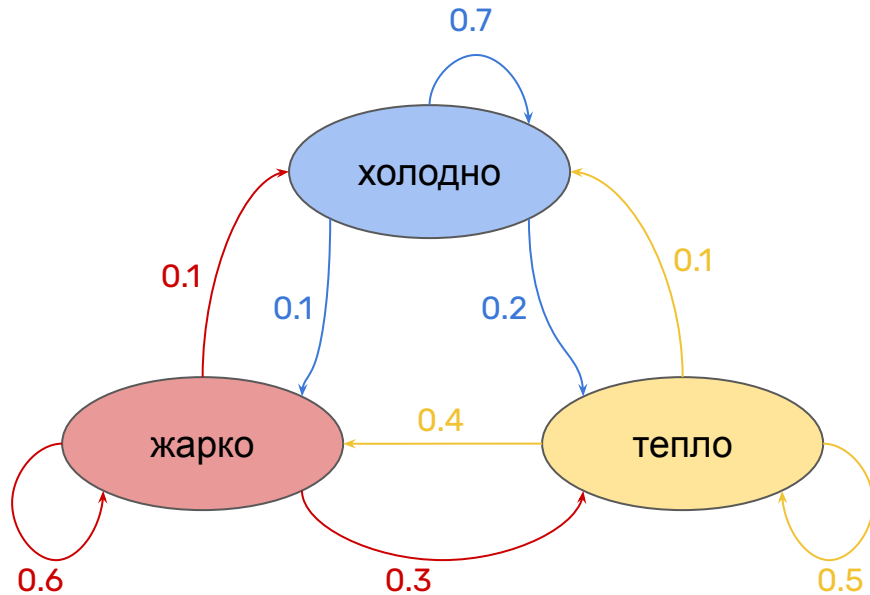
Марковская цепь

Лето в Москве

Состояния: S_1 = Холодно, S_2 = Тепло, S_3 = Жарко

Матрица переходов $\{a_{ij}\}$:

0.7	0.2	0.1
0.1	0.5	0.4
0.1	0.3	0.6





Марковская цепь

Лето в Москве

Состояния: S_1 = Холодно, S_2 = Тепло, S_3 = Жарко

Матрица переходов $\{a_{ij}\}$:

0.7	0.2	0.1
0.1	0.5	0.4
0.1	0.3	0.6

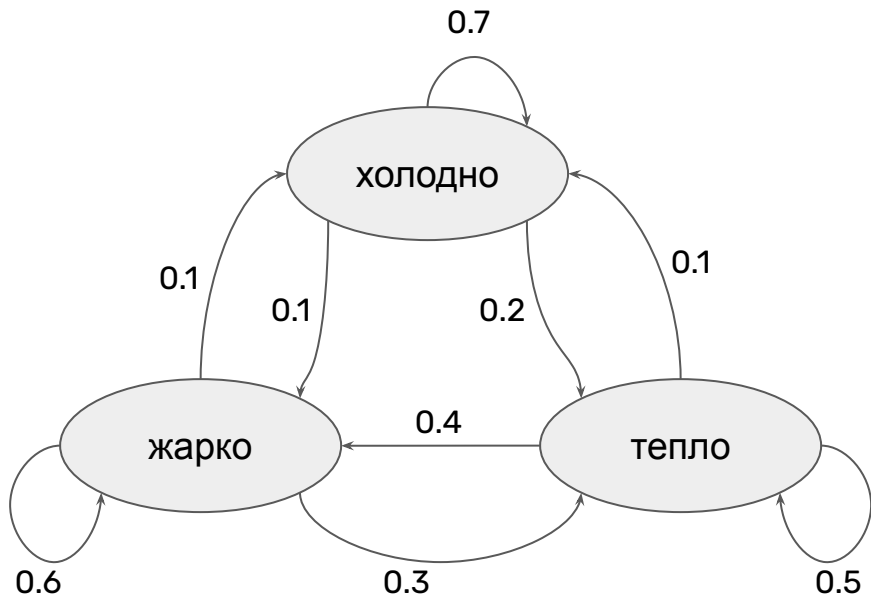
Пример последовательности:

1 июня - холодно

2 июня - тепло

3 июня - тепло

4 июня - жарко





Марковская цепь

Лето в Москве

Состояния: S_1 = холодно, S_2 = тепло, S_3 = жарко

Матрица переходов $\{a_{ij}\}$:

0.7	0.2	0.1
0.1	0.5	0.4
0.1	0.3	0.6

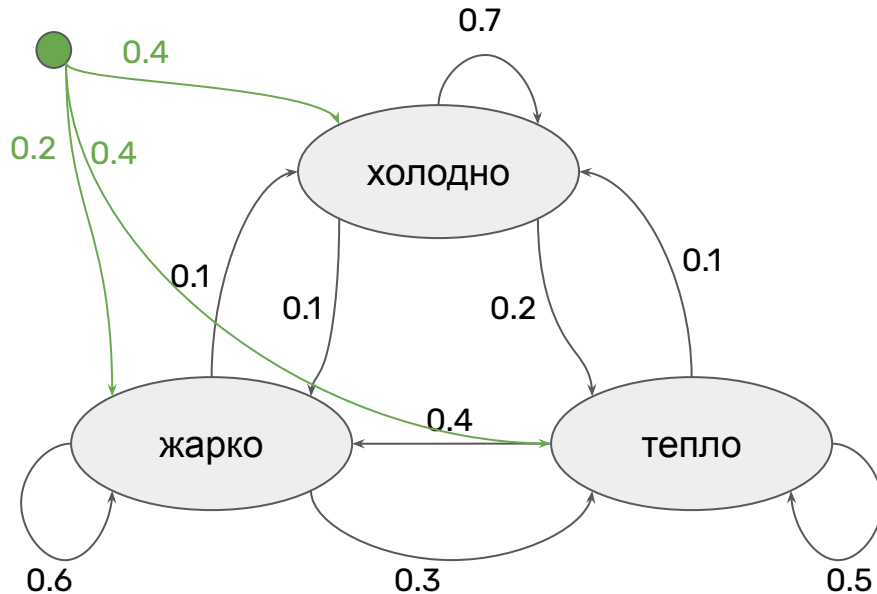
Пример последовательности:

1 июня - холодно

2 июня - тепло

3 июня - тепло

4 июня - жарко



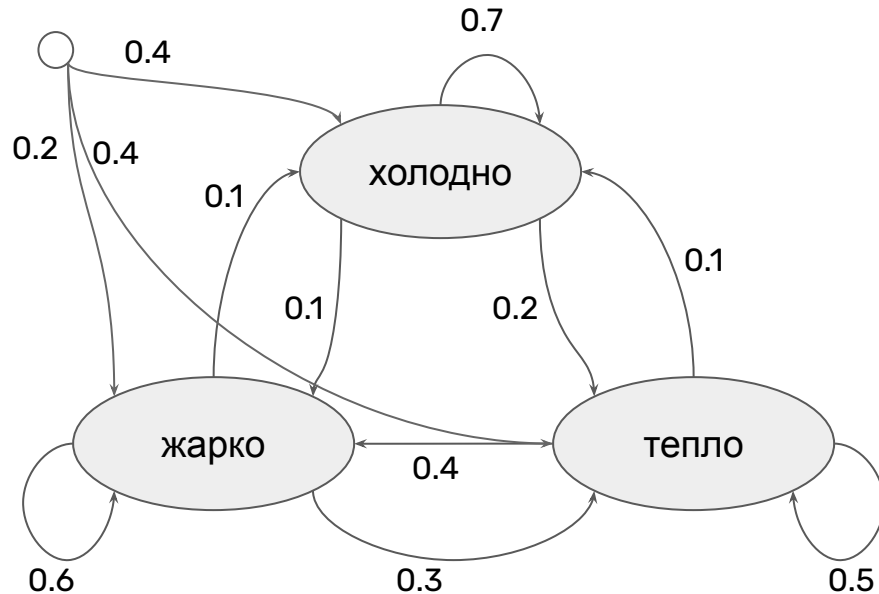
$$\pi_1 = P[1 \text{ июня холодно}] = 0.4$$

$$\pi_2 = P[1 \text{ июня тепло}] = 0.4$$

$$\pi_3 = P[1 \text{ июня жарко}] = 0.2$$



Марковская цепь



Лето в Москве

Пример последовательности:

1 июня - холодно

2 июня - тепло

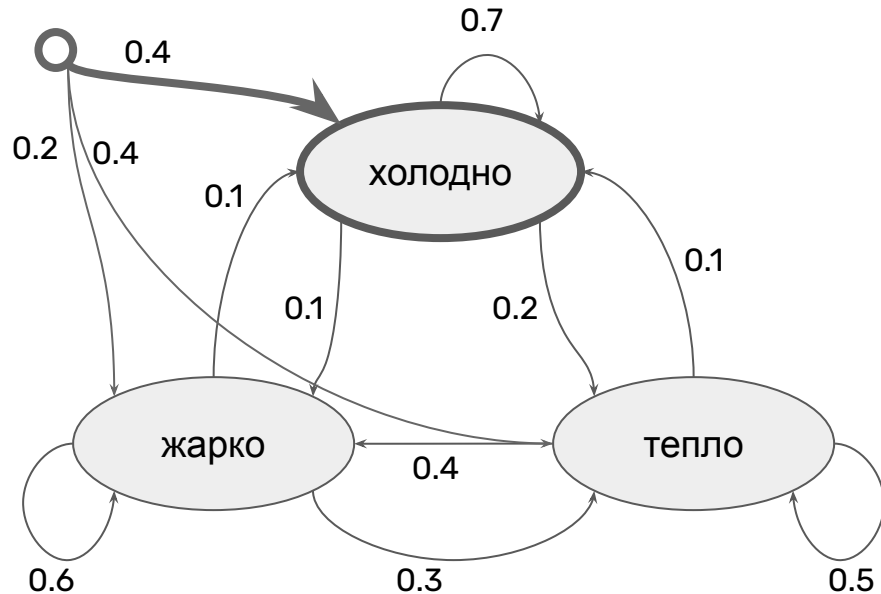
3 июня - тепло

4 июня - жарко

Какова вероятность такой последовательности?



Марковская цепь



Лето в Москве

Пример последовательности:

1 июня - холодно

2 июня - тепло

3 июня - тепло

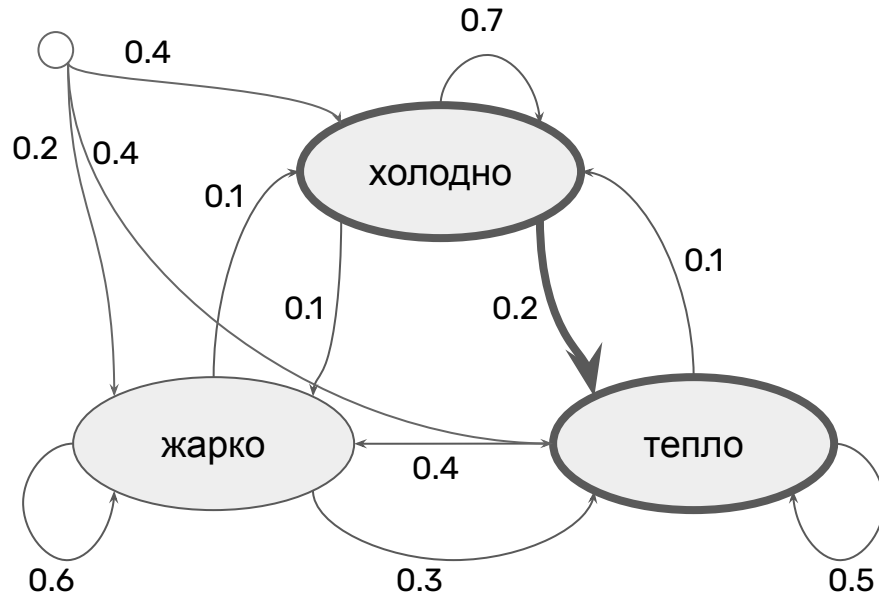
4 июня - жарко

Какова вероятность такой последовательности?

0.4



Марковская цепь



Лето в Москве

Пример последовательности:

1 июня - холодно

2 июня - тепло

3 июня - тепло

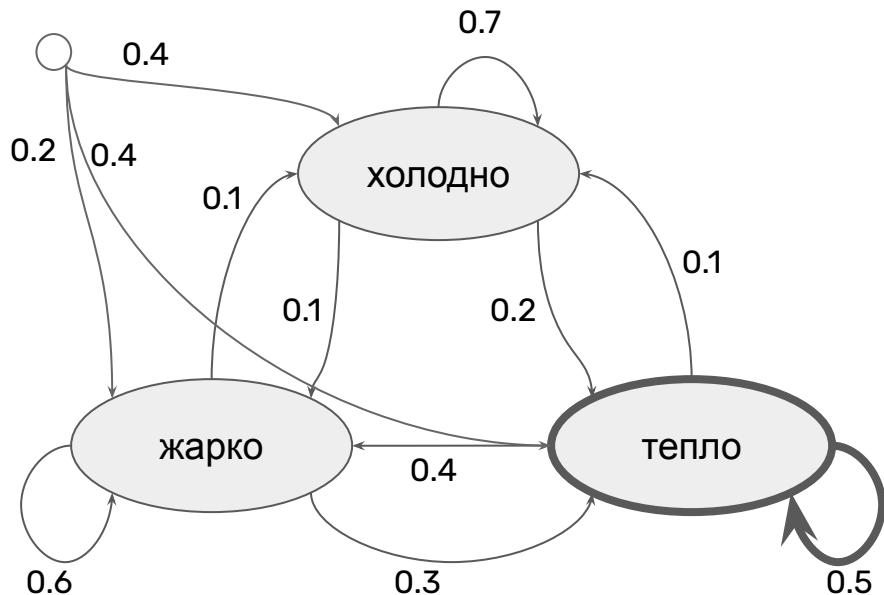
4 июня - жарко

Какова вероятность такой последовательности?

$$0.4 * 0.2$$



Марковская цепь



Лето в Москве

Пример последовательности:

1 июня - холодно

2 июня - тепло

3 июня - тепло

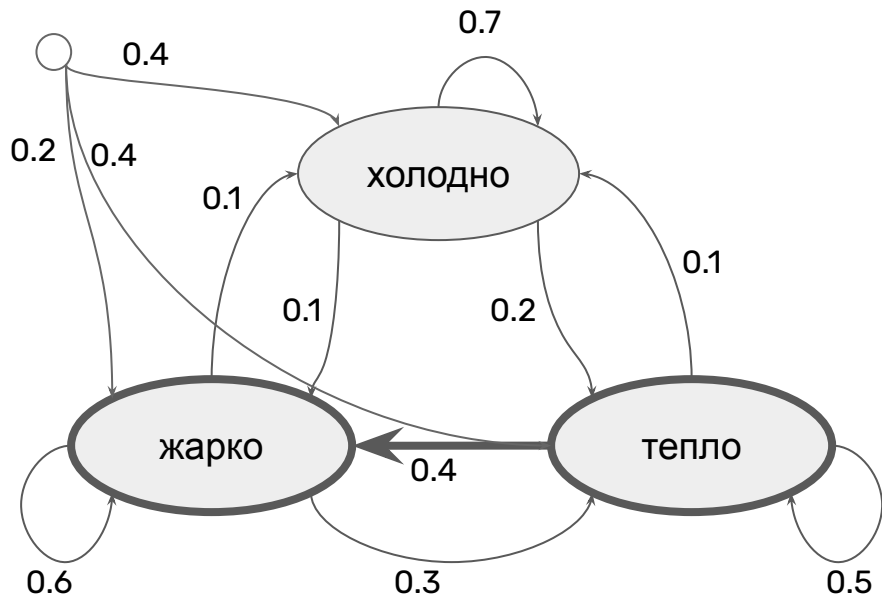
4 июня - жарко

Какова вероятность такой последовательности?

$$0.4 * 0.2 * 0.5$$



Марковская цепь



Лето в Москве

Пример последовательности:

1 июня - холодно

2 июня - тепло

3 июня - тепло

4 июня - жарко

Какова вероятность такой последовательности?

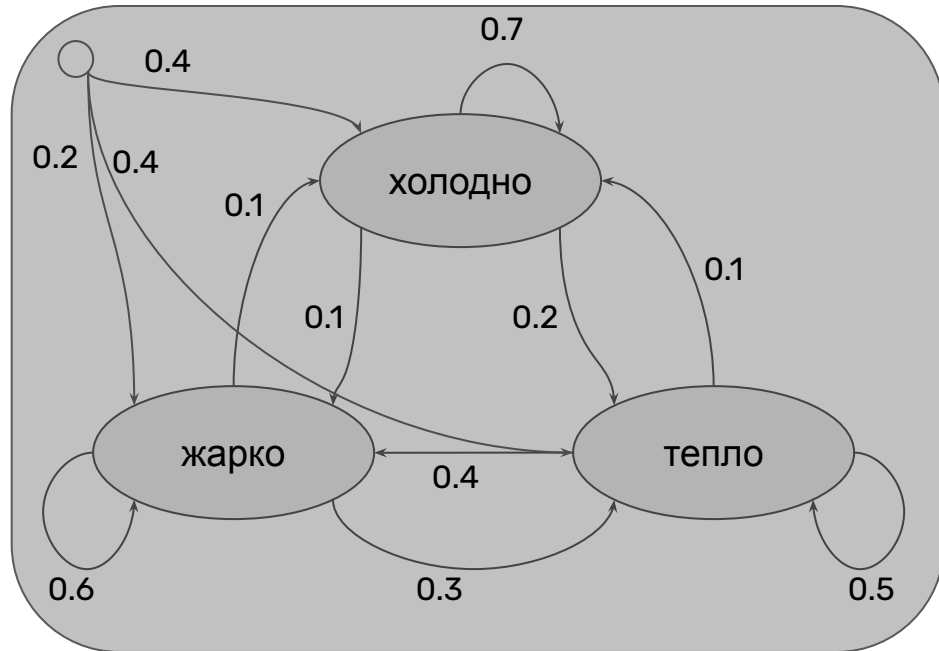
$$0.4 * 0.2 * 0.5 * 0.4 = 0.016$$



СММ - Скрытая марковская модель (HMM - Hidden Markov Model)



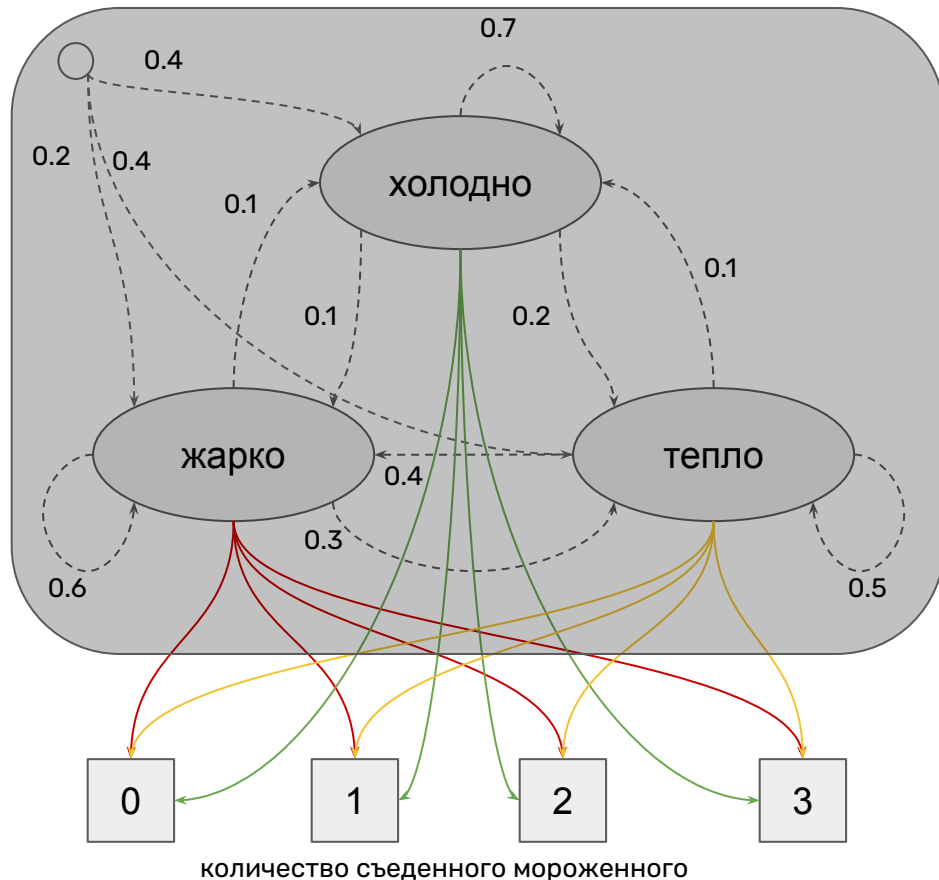
Скрытая марковская модель



Представим, что в далеком 2200 не сохранилось климатических данных о погоде в Москве в 2020 году. Сохранился только дневник Пети, который каждый день лета записывал количество съеденного мороженого (больше 3 не съедал).



Скрытая марковская модель



Представим, что в далеком 2200 не сохранилось климатических данных о погоде в Москве в 2020 году. Сохранился только дневник Пети, который каждый день лета записывал количество съеденного мороженого (больше 3 не съедал).

холодно, тепло, жарко - скрытые состояния

количество съеденного мороженого - наблюдения



Скрытая марковская модель

Марковская цепь, состояния которой скрыты от наблюдателя

+

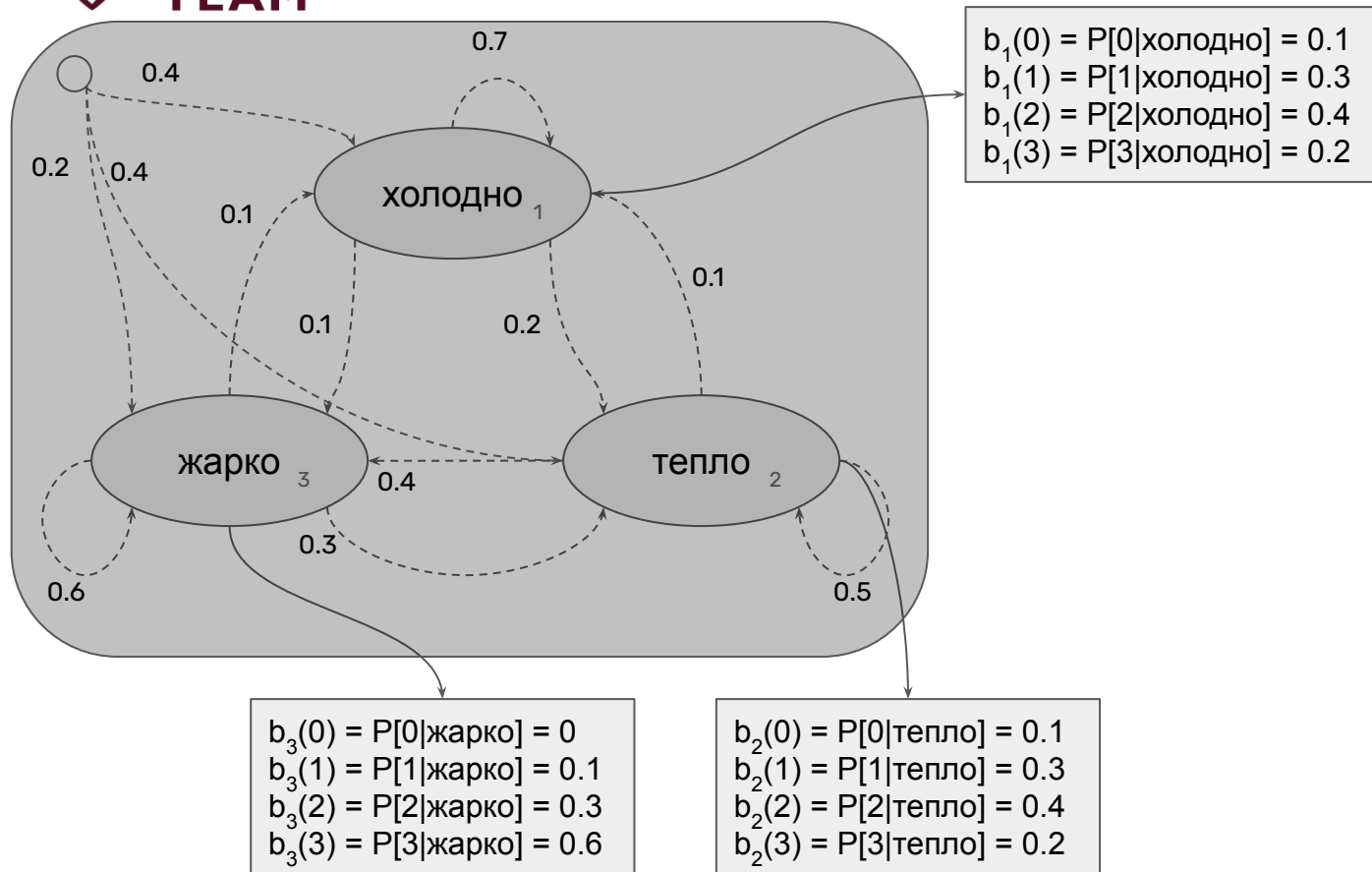
В каждый момент времени $t = 1, 2, \dots$ наблюдаем одно из M событий из множества v_1, \dots, v_M . Событие, наблюдаемое в момент времени t будем обозначать o_t .

$b_j(k) = P[o_t = v_k \mid q_t = S_j]$ - вероятность в момент времени t наблюдать v_k при условии, что система находится в состоянии S_j

$1 \leq j \leq N, 1 \leq k \leq M$



Скрытая марковская модель





Три задачи СММ



- ▶ **Правдоподобие**
- ▶ **Декодирование**
- ▶ **Обучение**



Три задачи СММ: правдоподобие

Дано:

модель $\lambda = (\{a_{ij}\}, \{b_j(k)\}, \{\pi_i\})$;

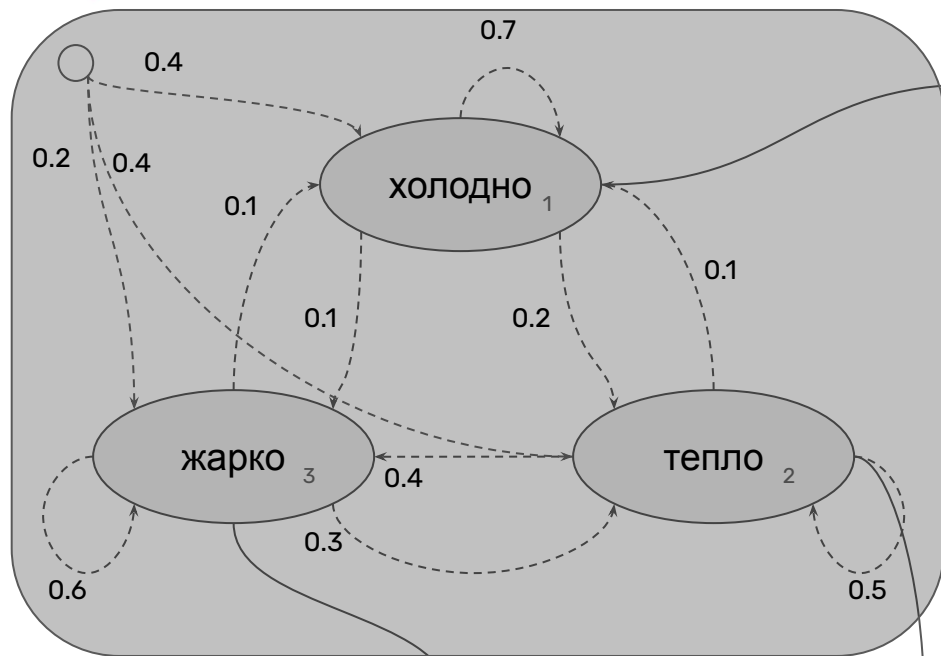
наблюдаемая последовательность $o_1 o_2 \dots o_T$

Необходимо вычислить вероятность наблюдаемой последовательности при заданной модели

$$P[o_1 o_2 \dots o_T \mid \lambda] = ?$$



Три задачи СММ: правдоподобие



$$\begin{aligned}b_1(0) &= P[0|\text{холодно}] = 0.1 \\b_1(1) &= P[1|\text{холодно}] = 0.3 \\b_1(2) &= P[2|\text{холодно}] = 0.4 \\b_1(3) &= P[3|\text{холодно}] = 0.2\end{aligned}$$

Какова вероятность того, что Петя съел
1 июня - 2 мороженых
2 июня - 1 мороженое
3 июня - 0 мороженых ?

$$\begin{aligned}b_3(0) &= P[0|\text{жарко}] = 0 \\b_3(1) &= P[1|\text{жарко}] = 0.1 \\b_3(2) &= P[2|\text{жарко}] = 0.3 \\b_3(3) &= P[3|\text{жарко}] = 0.6\end{aligned}$$

$$\begin{aligned}b_2(0) &= P[0|\text{тепло}] = 0.1 \\b_2(1) &= P[1|\text{тепло}] = 0.3 \\b_2(2) &= P[2|\text{тепло}] = 0.4 \\b_2(3) &= P[3|\text{тепло}] = 0.2\end{aligned}$$



Три задачи СММ: правдоподобие

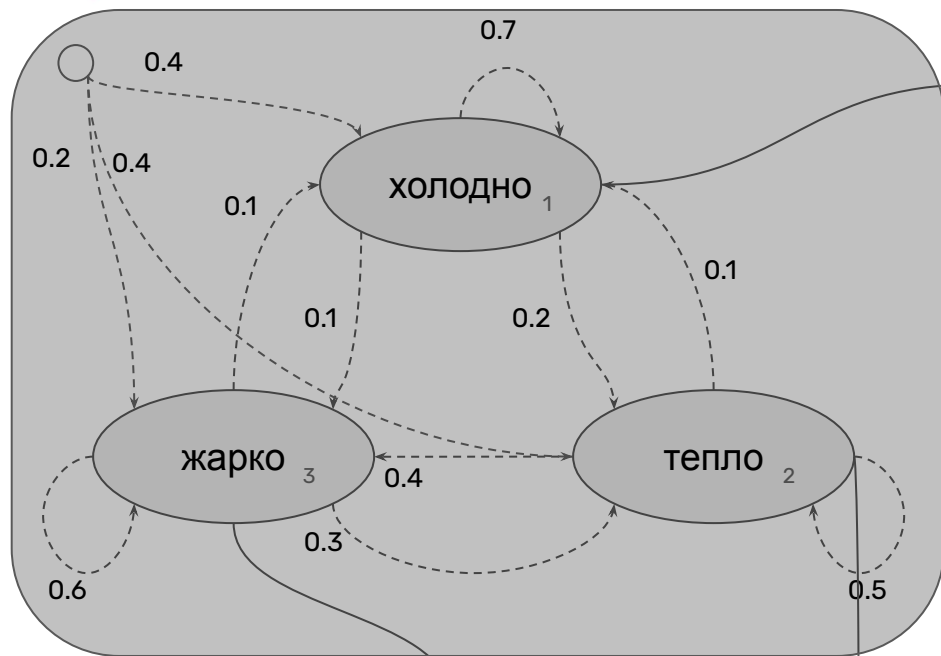
$O = o_1 o_2 \dots o_T$ - последовательность наблюдаемых событий

$Q = q_1 q_2 \dots q_T$ - последовательность скрытых состояний

$$P[O] = \sum_Q P[O, Q] = \sum_Q P[O | Q] P[Q]$$



Три задачи СММ: правдоподобие



$$\begin{aligned}b_1(0) &= P[0|\text{холодно}] = 0.1 \\b_1(1) &= P[1|\text{холодно}] = 0.3 \\b_1(2) &= P[2|\text{холодно}] = 0.4 \\b_1(3) &= P[3|\text{холодно}] = 0.2\end{aligned}$$

Какова вероятность того, что Петя съел
1 июня - 2 мороженых
2 июня - 1 мороженое
3 июня - 0 мороженых ?

O = 2,1,0

Q = холодно,холодно,холодно
холодно,холодно,тепло
холодно,холодно,жарко
холодно,тепло,холодно

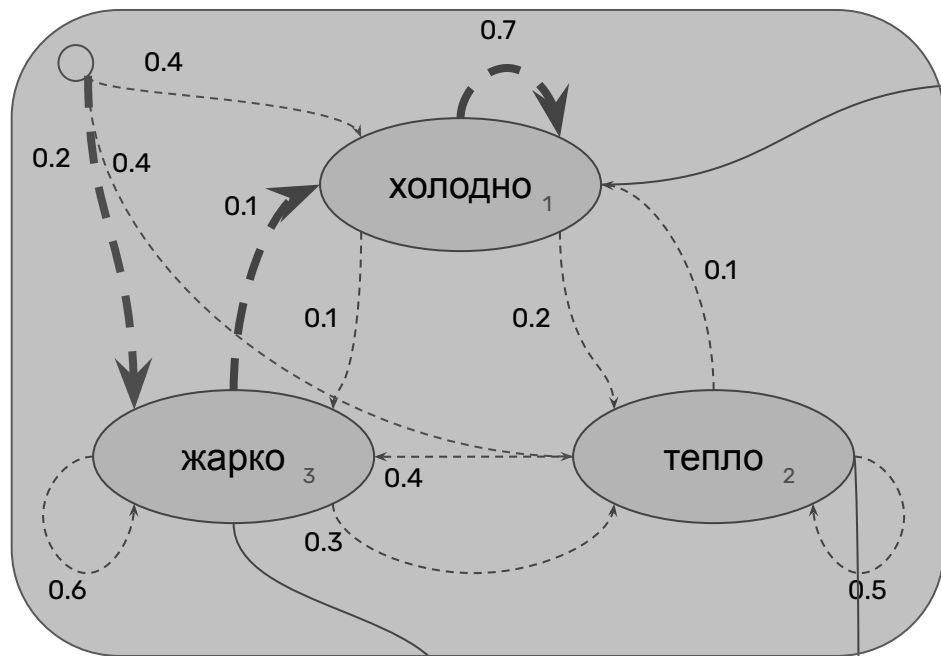
...

$$\begin{aligned}b_3(0) &= P[0|\text{жарко}] = 0 \\b_3(1) &= P[1|\text{жарко}] = 0.1 \\b_3(2) &= P[2|\text{жарко}] = 0.3 \\b_3(3) &= P[3|\text{жарко}] = 0.6\end{aligned}$$

$$\begin{aligned}b_2(0) &= P[0|\text{тепло}] = 0.1 \\b_2(1) &= P[1|\text{тепло}] = 0.3 \\b_2(2) &= P[2|\text{тепло}] = 0.4 \\b_2(3) &= P[3|\text{тепло}] = 0.2\end{aligned}$$



Три задачи СММ: правдоподобие



$$\begin{aligned} b_1(0) &= P[0|\text{холодно}] = 0.1 \\ b_1(1) &= P[1|\text{холодно}] = 0.3 \\ b_1(2) &= P[2|\text{холодно}] = 0.4 \\ b_1(3) &= P[3|\text{холодно}] = 0.2 \end{aligned}$$

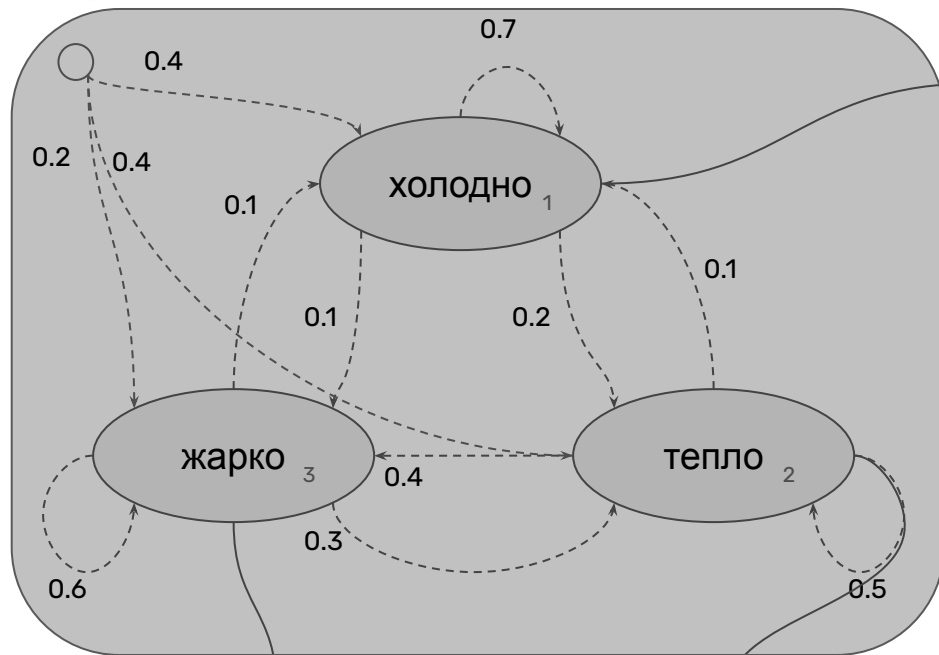
$$P[\text{жарко, холодно, холодно}] = 0.2 * 0.1 * 0.7$$

$$\begin{aligned} b_3(0) &= P[0|\text{жарко}] = 0 \\ b_3(1) &= P[1|\text{жарко}] = 0.1 \\ b_3(2) &= P[2|\text{жарко}] = 0.3 \\ b_3(3) &= P[3|\text{жарко}] = 0.6 \end{aligned}$$

$$\begin{aligned} b_2(0) &= P[0|\text{тепло}] = 0.1 \\ b_2(1) &= P[1|\text{тепло}] = 0.3 \\ b_2(2) &= P[2|\text{тепло}] = 0.4 \\ b_2(3) &= P[3|\text{тепло}] = 0.2 \end{aligned}$$



Три задачи СММ: правдоподобие



$$\begin{aligned}b_1(0) &= P[0|\text{холодно}] = 0.1 \\b_1(1) &= P[1|\text{холодно}] = 0.3 \\b_1(2) &= P[2|\text{холодно}] = 0.4 \\b_1(3) &= P[3|\text{холодно}] = 0.2\end{aligned}$$

$$P[\text{жарко, холодно, холодно}] = 0.2 * 0.1 * 0.7$$

$$P[2,1,0 \mid \text{жарко, холодно, холодно}] = 0.3 * 0.3 * 0.1$$

$$\begin{aligned}b_3(0) &= P[0|\text{жарко}] = 0 \\b_3(1) &= P[1|\text{жарко}] = 0.1 \\b_3(2) &= P[2|\text{жарко}] = \mathbf{0.3} \\b_3(3) &= P[3|\text{жарко}] = 0.6\end{aligned}$$

$$\begin{aligned}b_2(0) &= P[0|\text{тепло}] = 0.1 \\b_2(1) &= P[1|\text{тепло}] = 0.3 \\b_2(2) &= P[2|\text{тепло}] = 0.4 \\b_2(3) &= P[3|\text{тепло}] = 0.2\end{aligned}$$

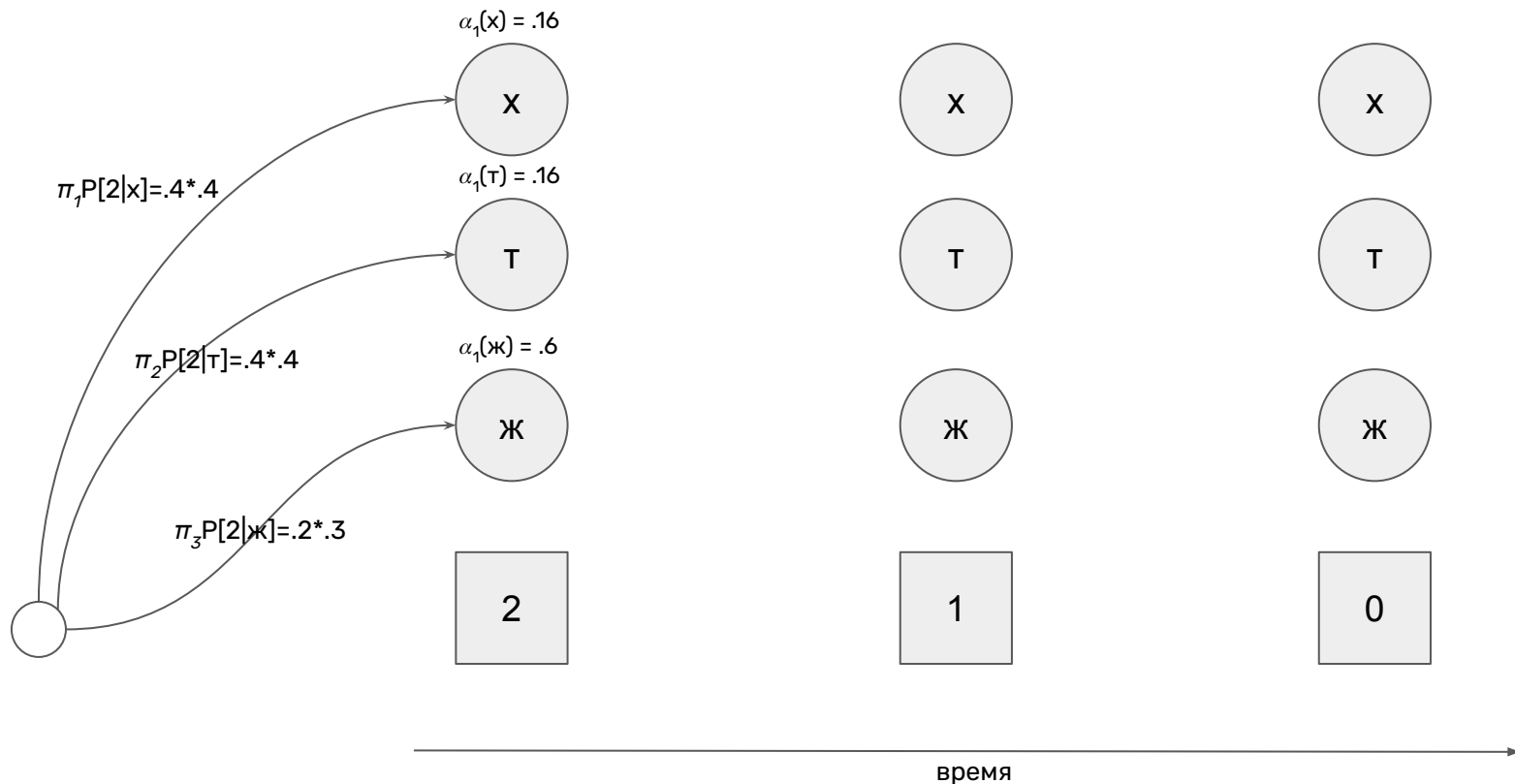
Наивное решение - оценить такие вероятности для каждой возможной комбинации скрытых состояний - высокая сложность вычислений $O(TN^T)$

Алгоритм прямого хода (динамическое программирование) - сложность вычислений $O(TN^2)$



Три задачи СММ: правдоподобие

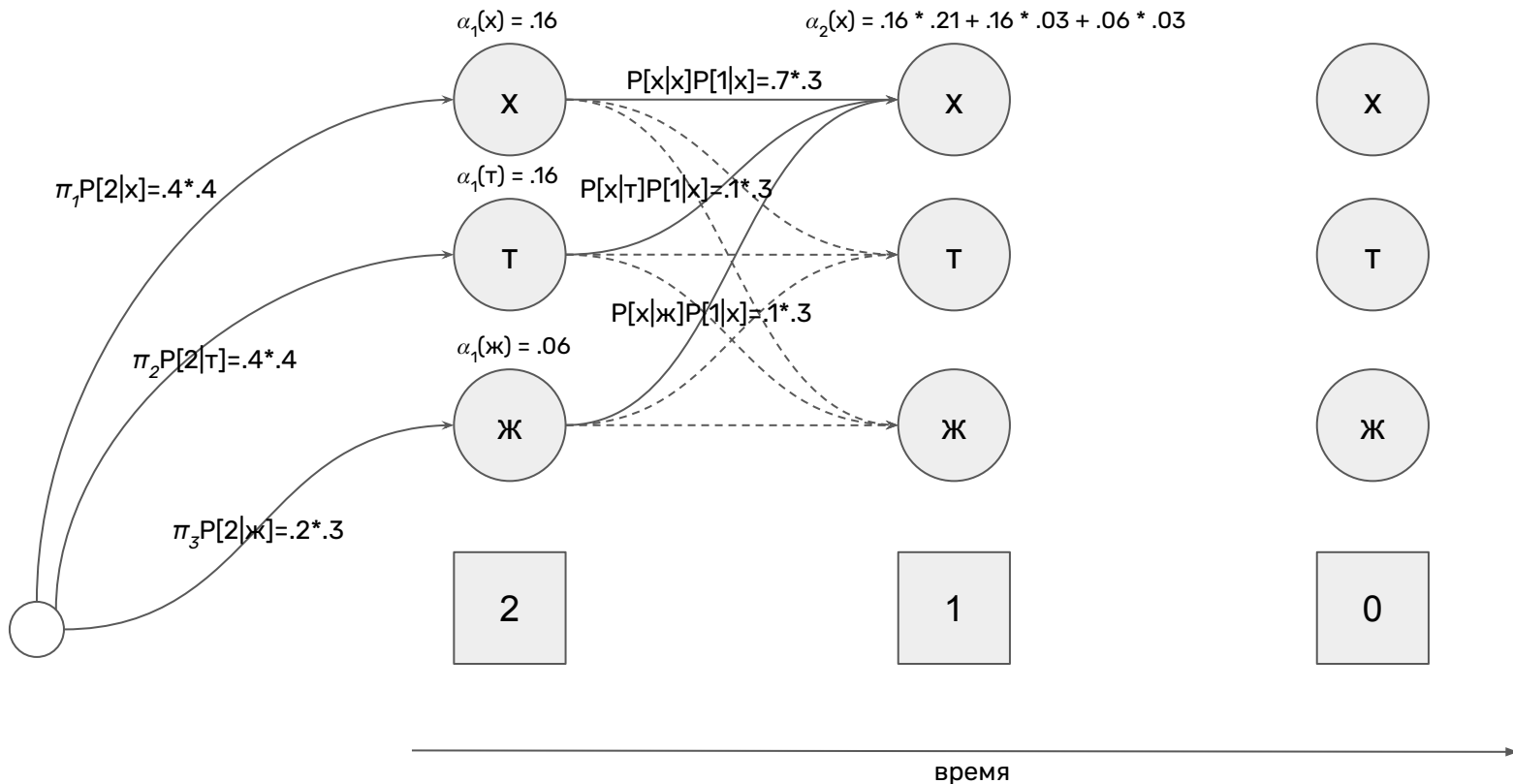
Алгоритм прямого хода





Три задачи СММ: правдоподобие

Алгоритм прямого хода





Три задачи СММ: правдоподобие

Алгоритм прямого хода

1. Initialization:

$$\alpha_1(j) = \pi_j b_j(o_1) \quad 1 \leq j \leq N$$

2. Recursion:

$$\alpha_t(j) = \sum_{i=1}^N \alpha_{t-1}(i) a_{ij} b_j(o_t); \quad 1 \leq j \leq N, 1 < t \leq T$$

3. Termination:

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i)$$

```
function FORWARD(observations of len  $T$ , state-graph of len  $N$ ) returns forward-prob

  create a probability matrix forward[ $N, T$ ]
  for each state  $s$  from 1 to  $N$  do                                ; initialization step
    forward[ $s, 1$ ]  $\leftarrow \pi_s * b_s(o_1)$ 
  for each time step  $t$  from 2 to  $T$  do                                ; recursion step
    for each state  $s$  from 1 to  $N$  do
      forward[ $s, t$ ]  $\leftarrow \sum_{s'=1}^N \text{forward}[s', t-1] * a_{s',s} * b_s(o_t)$ 

  forwardprob  $\leftarrow \sum_{s=1}^N \text{forward}[s, T]$                                 ; termination step
  return forwardprob
```



Три задачи СММ: декодирование

Дано:

модель $\lambda = (\{a_{ij}\}, \{b_j(k)\}, \{\pi_i\})$;

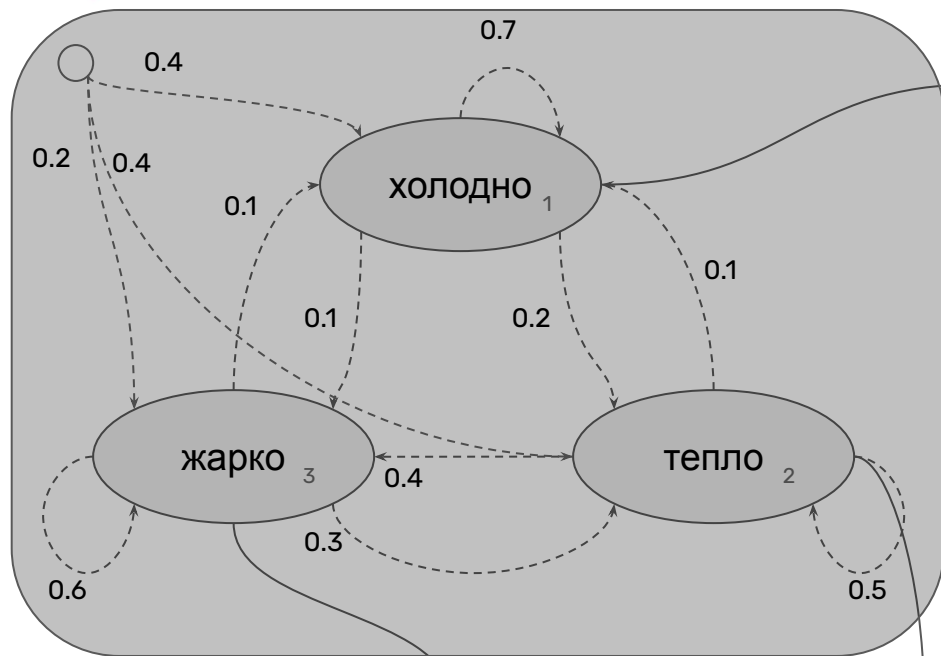
наблюдаемая последовательность $o_1 o_2 \dots o_T$

Необходимо найти наиболее вероятную последовательность
скрытых состояний

$$\operatorname{argmax} P[q_1 q_2 \dots, q_T \mid \lambda, o_1 o_2 \dots o_T]$$



Три задачи СММ: декодирование



$$\begin{aligned} b_1(0) &= P[0|\text{холодно}] = 0.1 \\ b_1(1) &= P[1|\text{холодно}] = 0.3 \\ b_1(2) &= P[2|\text{холодно}] = 0.4 \\ b_1(3) &= P[3|\text{холодно}] = 0.2 \end{aligned}$$

Какая последовательность погодных условий наиболее вероятна, при условии, что Петя съел

1 июня - 2 мороженных

2 июня - 1 мороженое

3 июня - 0 мороженных?

Перебирать все варианты вычислительно долго!

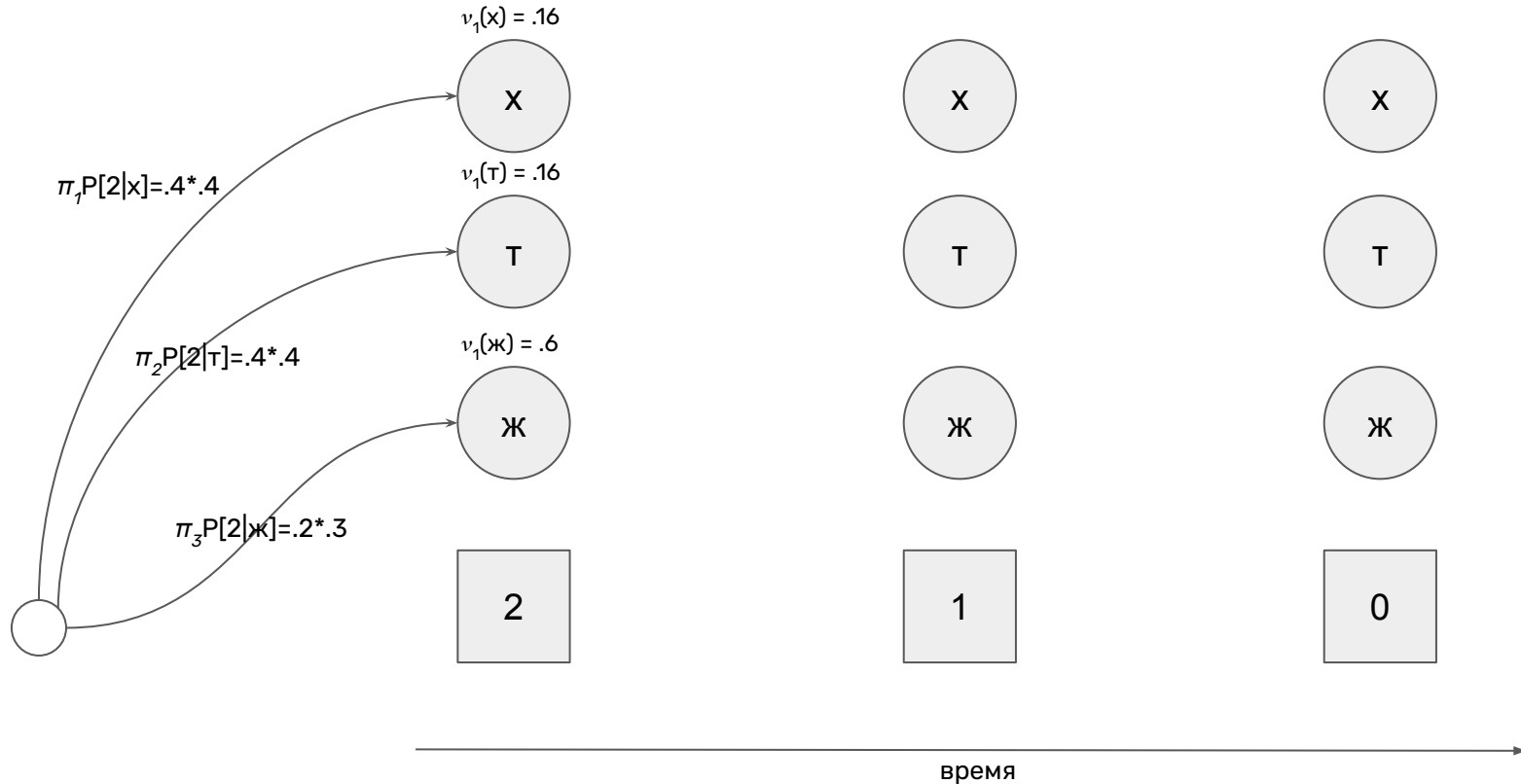
$$\begin{aligned} b_3(0) &= P[0|\text{жарко}] = 0 \\ b_3(1) &= P[1|\text{жарко}] = 0.1 \\ b_3(2) &= P[2|\text{жарко}] = 0.3 \\ b_3(3) &= P[3|\text{жарко}] = 0.6 \end{aligned}$$

$$\begin{aligned} b_2(0) &= P[0|\text{тепло}] = 0.1 \\ b_2(1) &= P[1|\text{тепло}] = 0.3 \\ b_2(2) &= P[2|\text{тепло}] = 0.4 \\ b_2(3) &= P[3|\text{тепло}] = 0.2 \end{aligned}$$



Три задачи СММ: декодирование

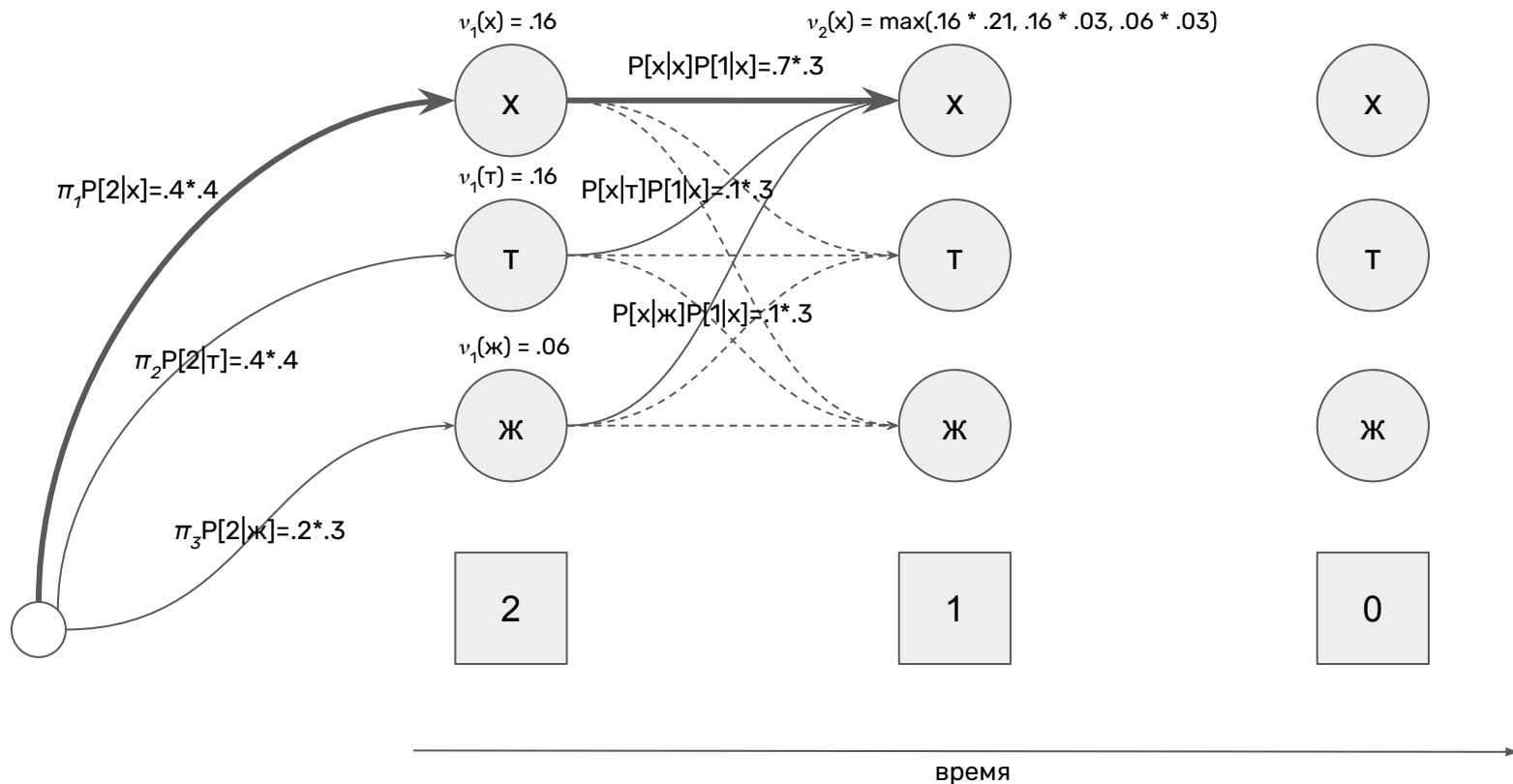
Алгоритм Витерби





Три задачи СММ: декодирование

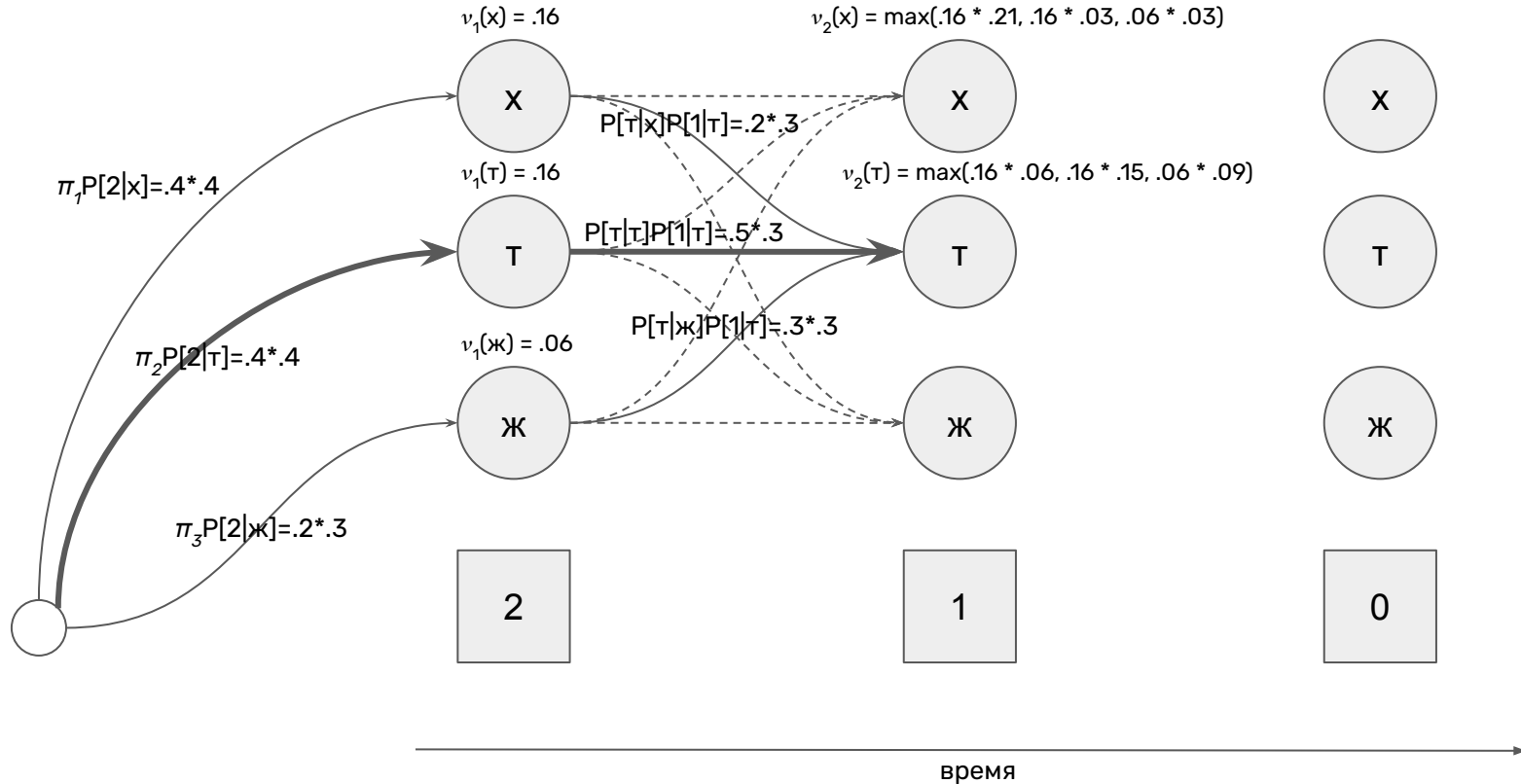
Алгоритм Витерби





Три задачи СММ: декодирование

Алгоритм Витерби





Три задачи СММ: декодирование

Алгоритм Витерби

1. Initialization:

$$\begin{aligned}v_1(j) &= \pi_j b_j(o_1) & 1 \leq j \leq N \\bt_1(j) &= 0 & 1 \leq j \leq N\end{aligned}$$

2. Recursion

$$\begin{aligned}v_t(j) &= \max_{i=1}^N v_{t-1}(i) a_{ij} b_j(o_t); & 1 \leq j \leq N, 1 < t \leq T \\bt_t(j) &= \operatorname{argmax}_{i=1}^N v_{t-1}(i) a_{ij} b_j(o_t); & 1 \leq j \leq N, 1 < t \leq T\end{aligned}$$

3. Termination:

$$\text{The best score: } P^* = \max_{i=1}^N v_T(i)$$

$$\text{The start of backtrace: } q_T^* = \operatorname{argmax}_{i=1}^N v_T(i)$$

function VITERBI(*observations* of len T , *state-graph* of len N) **returns** *best-path*, *path-prob*

create a path probability matrix *viterbi*[N, T]

for each state s **from** 1 **to** N **do** ; initialization step

$viterbi[s, 1] \leftarrow \pi_s * b_s(o_1)$

$backpointer[s, 1] \leftarrow 0$

for each time step t **from** 2 **to** T **do** ; recursion step

for each state s **from** 1 **to** N **do**

$viterbi[s, t] \leftarrow \max_{s'=1}^N viterbi[s', t-1] * a_{s', s} * b_s(o_t)$

$backpointer[s, t] \leftarrow \operatorname{argmax}_{s'=1}^N viterbi[s', t-1] * a_{s', s} * b_s(o_t)$

$bestpathprob \leftarrow \max_{s=1}^N viterbi[s, T]$; termination step

$bestpathpointer \leftarrow \operatorname{argmax}_{s=1}^N viterbi[s, T]$; termination step

$bestpath \leftarrow$ the path starting at state $bestpathpointer$, that follows $backpointer[]$ to states back in time

return *bestpath*, *bestpathprob*

<https://web.stanford.edu/~jurafsky/slp3/A.pdf>



Три задачи СММ: обучение

Дано:

множество скрытых состояний S_1, \dots, S_N

наблюдаемая последовательность $o_1 o_2 \dots o_T$

Необходимо подобрать $\lambda = (\{a_{ij}\}, \{b_j(k)\}, \{\pi_i\})$ так, чтобы максимизировать $P[o_1 o_2 \dots o_T \mid \lambda]$

Решение: Алгоритм Баума-Велша



- ▶ Определение частей речи (POS-tagging)
- ▶ Распознавание речи
- ▶ Синтез речи
- ▶ Машинный перевод
- ▶ Распознавание рукописного текста
- ▶ Анализ временных рядов



Вы узнали:

- ▶ Как устроена марковская цепь
- ▶ Что представляет собой скрытая марковская модель
- ▶ Какие задачи можно ставить и как их решать



<https://web.stanford.edu/~jurafsky/slp3/A.pdf>

<https://web.stanford.edu/~jurafsky/slp3/8.pdf>

http://www.cs.cornell.edu/courses/cs4758/2012sp/materials/hmm_paper_rabiner.pdf



Спасибо за внимание! Вопросы?

Feedback: [день 1](#)

Наталья Корепанова, korepanova@bigdatateam.org

Инструктор, [BigDataTeam](#)

Data Science Tech Lead, [IPONWEB](#)

Профиль на [LinkedIn](#)