

# Дискретная математика

## II семестр

Лектор: Станкевич Андрей Сергеевич

зима/весна 2024

\_scarleteagle

imkochelorov

AberKadaber

## Оглавление

<b>1. Дискретная теория вероятностей .....</b>	<b>2</b>
1.1. Введение .....	2
1.2. Аксиоматическая теория вероятностей .....	2
1.3. Независимость событий .....	2
1.4. Прямое произведение вероятностных пространств .....	2
1.5. Условная вероятность .....	2
1.6. Случайная величина .....	3
1.7. Математическое ожидание .....	3
1.8. Дисперсия .....	3
1.9. Хвостовые неравенства .....	4
1.10. Энтропия .....	5
1.11. Марковские цепи .....	6
<b>2. Теория формальных языков .....</b>	<b>7</b>
2.1. Действия над языками .....	7
2.2. Детерминированные конечные автоматы:	7
2.2.1. Напоминание: .....	8
2.3. Недетерминированные конечные автоматы .....	8
2.4. $\varepsilon$ -НКА .....	8
2.5. Теорема Клини .....	8
2.6. Регулярные языки .....	9
2.6.1. Минимизация ДКА .....	9
2.7. Нерегулярные языки .....	9

# 1. Дискретная теория вероятностей

## 1.1. Введение

Случайности и вероятности давно интересовали людей, в основном в азартных играх. Первые теоремы дискретной теории вероятностей появились задолго до появления первого конечного автомата. Но философы всех времён задаются вопросом — “существует ли случайность?” Философская мысль делится на 2 направления — считающих, что случайности существуют, и не считающих. Это всё мы с вами изучать не будем

## 1.2. Аксиоматическая теория вероятностей

by Андрей Николаевич Колмогоров

Бытовое восприятие случайностей мешает в понимании формальной модели

**Определение:** множество элементарных исходов

$\Omega$  — множество всех возможных элементарных исходов случайного эксперимента, который может закончиться ровно одним из них

$\Omega$  может быть не более чем счётным, счётным, или не счётным. Но мы будем рассматривать дискретное множество элементарных исходов

$|\Omega|$  — конечно или счётно

**Определение:** элементарный исход

Элемент  $\omega \in \Omega$

**Определение:** дискретная плотность вероятности

$p : \Omega \rightarrow \mathbb{R} : p(\omega \geq 0), \sum_{\omega \in \Omega} p(\omega) = 1$

Если  $\Omega$  несчётна, то требуется другая теория

**Определение:** дискретное вероятностное пространство

Пара из множества элементарных исходов и дискретной плотности вероятностей  $(\Omega, p)$

**Примеры:**

1. Честная монета:

$$\Omega = \{0, 1\}, p(0) = p(1) = \frac{1}{2}$$

2. Нечестная монета (распределение Бернулли):

$$\Omega = \{0, 1\}, p(1) = p, p(0) = 1 - p$$

3. Честная игральная кость (1d6):

$$\Omega = \{1, 2, 3, 4, 5, 6\}, p(i) = \frac{1}{6}$$

1. “Честная” игральная кость (1d20):

$$\Omega = \{1, \dots, 20\}, p(20) = 1$$

4. Колода карт:

$$\Omega = \{\langle r, s \rangle \mid r = 1 \dots 13, s = 1 \dots 4\}, p(\langle r, s \rangle)$$

**Определение:** случайное событие

Подмножество элементарных исходов  $A \subset \Omega, P(A) = \sum_{a \in A} p(a)$

Дискретное множество элементарных исходов является случайным событием

**Примеры:**

1. Пустое событие  $P(\emptyset) = 0$

2. Достоверное событие (полное (?))  $P(\Omega) = 1$

3. Для честной монеты  $P(\{0\}) = \frac{1}{2}, P(\{1\}) = \frac{1}{2}$

4. Для честной 1d6  $P(\{1, 3, 5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}, P(\{5, 6\}) = \frac{1}{3}, P(\{1, 2, 3\}) = \frac{1}{2}$

## 1.3. Независимость событий

**Определение:** независимое случайное событие

$A$  и  $B$  независимы, если  $P(A \cap B) = P(A) \cdot P(B)$

**Примеры:**

Для честной игральной кости

$\text{Even} = \{2, 4, 6\}, \text{Big} = \{5, 6\}, \text{Small} = \{1, 2, 3\}$

$$\bullet P(\text{Even} \cap \text{Big}) = P(\{6\}) = \frac{1}{6} \quad P(\text{Even}) \cdot P(\text{Big}) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$

$$\bullet P(\text{Even} \cap \text{Small}) = P(\{2\}) = \frac{1}{6} \quad P(\text{Even}) \cdot P(\text{Small}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\bullet P(\text{Big} \cap \text{Small}) = P(\emptyset) = 0 \quad P(\text{Big}) \cdot P(\text{Small}) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

**Определение:** события, независимые в совокупности

$A_1, A_2, \dots, A_k$  — независимы в совокупности, если  $\forall I \subset \{1, 2, \dots, k\} \quad P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$

**Примеры:**

Для броска двух разных честных монет

$\Omega = \{00, 01, 10, 11\}, p(i \cdot j) = \frac{1}{4}$

$A = \{01, 00\}, B = \{10, 00\}, C = \{11, 00\}$  не независимы в совокупности

## 1.4. Прямое произведение вероятностных пространств

**Определение:** прямое произведение вероятностей пространств

$\langle \Omega_1, p_1 \rangle, \langle \Omega_2, p_2 \rangle$ , прямое произведение пространств  $\Omega = \Omega_1 \times \Omega_2 : p(\langle u_1, u_2 \rangle) = p_1(u_1) \cdot p_2(u_2)$

$$\sum_{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2} p(\langle \omega_1, \omega_2 \rangle) = \sum p_1(\omega_1) \cdot p_2(\omega_2) = \sum_{\omega_1} \left( p_1(\omega_1) \cdot \sum_{\omega_2} p_2(\omega_2) \right) = 1$$

**Пример:**

$A_1 \subset \Omega_1, A_2 \subset \Omega_2 \Rightarrow A_1 \times A_2 \subset \Omega_1 \times \Omega_2$  и  $\Omega_1 \times A_2$  — независимы

$$(A_1 \times \Omega_2) \cap (\Omega_1 \times A_2) = A_1 \times A_2$$

## 1.5. Условная вероятность

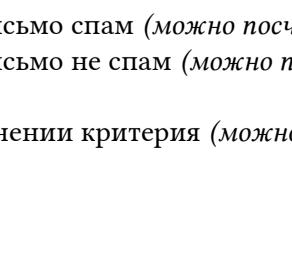
имеет смысл, если  $P(B) \neq 0$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) \neq 0$$

если  $A$  и  $B$  независимы, то

$$P(A|B) = \frac{P(A) \cdot P(B)}{P(B)} = P(A)$$

$$p_B(\omega) = \frac{p(\omega)}{P(B)}, P_B(A) = \frac{P(A \cap B)}{P(B)}$$



**Теорема:** (Формула полной вероятности)

$$\Omega = \underbrace{A_1 \cup A_2 \cup \dots \cup A_k}_{\text{полная система событий}}, A_i \cap A_j = \emptyset \text{ при } i \neq j$$

$B$

$$P(B) = \sum_{i=1}^k P(B|A_i)P(A_i)$$

$$\sum_{i=1}^k P(B|A_i)P(A_i) = \sum_{i=1}^k \frac{P(B \cap A_i)}{P(A_i)} \cdot P(A_i) = \sum_{i=1}^k P(B \cap A_i) = P(B)$$

**Задача:**

Две урны, в одной 3 черных и 2 белых, в другой 4 черных и 1 белый шар. С какой вероятностью можно вынуть черный шар?

$$A_1 \quad A_2 \quad \frac{1}{2}$$

$$P(B|A_1) = \frac{3}{5} \quad P(B|A_2) = \frac{4}{5} \quad P(B) = \frac{3}{5} \cdot \frac{1}{2} + \frac{4}{5} \cdot \frac{1}{2} = \frac{7}{10}$$

**Задача:**

$$P(B|A_i) \quad P(A_i) \quad \text{найти } P(A_i|B) = ?$$

Достоверность = 1 -  $P(B|A_2) = 99\%$

Надёжность =  $P(B|A_1) = 95\%$

$$A_1 \text{ — болен } \left(\frac{1}{100}\right)$$

$$A_2 \text{ — здоров } \left(\frac{99}{100}\right)$$

$$P(A_1|B) = \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.01 \cdot 0.99}$$

$$P(B|A_i) = \frac{P(A_i \cap B)}{P(A_i)}$$

**Определение:** формула Байеса

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^k P(B|A_j)P(A_j)}$$

**Определение:** Байесовский спам-фильтр

$A_1$  — спам

$A_2$  — не спам

$B$  — критерий

$P(B|A_1)$  — вероятность выполнения критерия, если письмо спам (можно посчитать)

$P(B|A_2)$  — вероятность выполнения критерия, если письмо не спам (можно посчитать)

Сам фильтр:  $P(A_1|B)$  — вероятность спама при выполнении критерия (можно вычислить, используя значения выше)

## 1.6. Случайная величина

**Определение:** случайная величина

$\langle \Omega, p \rangle$  – вероятностное пространство

$\xi : \Omega \rightarrow \mathbb{R}$  – случайная величина (численная характеристика вероятностного эксперимента)

**Примеры:**

Игральная кость

$\Omega = \{1, 2, 3, 4, 5, 6\}$

$\xi(\omega) = \omega$

$\eta$  – выигрыш Васи

**Пример 1:**

Вася получает 1 монету, если игральная кость падает нечётным числом и теряет 1 монету в ином случае:

$\omega$	1	2	3	4	5	6
$\eta$	1	-1	1	-1	1	-1

**Пример 2:**

Вася получает число монет, равное 2 в степени числа бросков кости, которые потребовались, чтобы выпало определённое число

$\omega$	1	01	001	0001	...
$\eta$	2	4	8	16	...

**Операции над случайными величинами:**

Произведение с числом:

$$\xi = c \cdot \eta \quad c \in \mathbb{R}$$

Сумма случайных величин:

$$\xi = \eta + \zeta$$

Произведение случайных величин:

$$\xi = \eta \cdot \zeta$$

Возведение в степень случайной величины:

$$\xi = \eta^\zeta$$

Можно даже рассмотреть синус случайной величины:

$$\xi = \sin \zeta$$

**Определение:** дискретная плотность распределения

$\xi$  – случайная величина

$f_\xi : \mathbb{R} \rightarrow \mathbb{R}$  – дискретная плотность распределения

$$f_\xi(a) = P(\xi = a) = P(\{\omega \mid \xi(\omega) = a\})$$

**Определение:** функция распределения

$\xi$  – случайная величина

$F_\xi : \mathbb{R} \rightarrow \mathbb{R}$  – функция распределения

$$F_\xi(a) = P(\xi \leq a)$$

**Пример:**

Подбросим 10 монет

$$\Omega = \mathbb{B}^{10}$$

$\xi(\omega)$  – число единиц

$$P(\xi = a) = \frac{\binom{10}{a}}{2^{10}}$$

$$f(a) = F(a) - F(a - \delta)$$

$$F(a) = \sum_{b \leq a} f(b)$$

## 1.7. Математическое ожидание

**Определение:** математическое ожидание

$\xi : \Omega \rightarrow \mathbb{R}$  – случайная величина

$$E\xi = \sum_{\omega \in \Omega} \xi(\omega) \cdot p(\omega) – математическое ожидание$$

**Пример:**

Игральная кость

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$\xi(\omega) = \omega$$

$$E\xi = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

Легко заметить, что 3.5 никогда не выпадает на игральной кости

Математическое ожидание не означает наиболее вероятные исход

**Пример 1:**

Вася получает 1 монету, если игральная кость падает нечётным числом и теряет 1 монету в ином случае:

$\omega$	1	2	3	4	5	6
$\eta$	1	-1	1	-1	1	-1

$$E\xi = 0$$

Математическое ожидание равно 0, но при этом, после 1 игры, Вася либо получит монету, либо потеряет

**Пример 2:**

Вася получает число монет, равное 2 в степени числа бросков кости, которые потребовались, чтобы выпало определённое число

$\omega$	1	01	001	0001	...
$\eta$	2	4	8	16	...

$$E\xi = +\infty$$

Математическое ожидание может равняться  $+\infty$

$$Ec = c$$

Математическое ожидание константы равняется константе

**Линейность математического ожидания:**

**Теорема:**

$$\xi = c \cdot \eta \quad E\xi = c \cdot E\eta$$

$$\xi = \eta + \zeta \quad E\xi = E\eta + E\zeta$$

$$E(\eta + \zeta) = \sum_{\omega \in \Omega} (\eta + \zeta)(\omega) = \sum_{\omega \in \Omega} (\eta(\omega) + \zeta(\omega)) = E\eta + E\zeta$$

**Примечание:**

$\xi$  и  $\eta$  – независимые, если

$\forall a, b \in \mathbb{R} \quad [\xi \leq a] \text{ и } [\eta \leq b]$  – независимые случайные события

$$P(\xi \leq a \wedge \eta \leq b) = P(\xi \leq a) \cdot P(\eta \leq b)$$

**Пример:**

$\Omega = S_n$  – перестановки  $n$  элементов

$$|\Omega| = n!$$

$$\xi(\sigma) = |\{i \mid \sigma_i = i\}| – количество неподвижных точек$$

$$n = 3 \quad \xi(\langle 1, 3, 2 \rangle) = 1$$

$$E\xi = 1$$

Мы можем посчитать математическое ожидание, не зная распределение

$$\xi_i = \begin{cases} 1, & \sigma_i = i \\ 0, & \text{иначе} \end{cases}$$

$$\xi_1(1 \ 3 \ 2) = 1$$

$$\xi_2(1 \ 3 \ 2) = 0$$

$$\xi_3(1 \ 3 \ 2) = 0$$

$$\xi = \sum_{i=1}^n \xi_i$$

$$P(\xi_i = 1) = \frac{1}{n}$$

$$E\xi_i = \frac{1}{n} \cdot P(\xi_i = 1) + 0 \cdot P(\xi_i = 0) = \frac{1}{n}$$

$$E\xi = \sum_{i=1}^n \frac{1}{n} = 1$$

Мы можем посчитать математическое ожидание, не зная распределение

**Пример:**

$\Omega = \mathbb{B}^n$  – перестановки  $n$  элементов

$$|\Omega| = 2^n$$

$$\xi(\sigma) = |\{i \mid \sigma_i = i\}| – количество неподвижных точек$$

$$n = 3 \quad \xi(\langle 1, 3, 2 \rangle) = 1$$

$$E\xi = 1$$

Мы можем посчитать математическое ожидание, не зная распределение

**Пример:**

$\Omega = \mathbb{B}^n$  – перестановки  $n$  элементов

$$|\Omega| = 2^n$$

$$\xi(\sigma) = |\{i \mid \sigma_i = i\}| – количество неподвижных точек$$

$$n = 3 \quad \xi(\langle 1, 3, 2 \rangle) = 1$$

$$E\xi = 1$$

Мы можем посчитать математическое ожидание, не зная распределение

**Пример:**

$\Omega = \mathbb{B}^n$  – перестановки  $n$  элементов

$$|\Omega| = 2^n$$

$$\xi(\sigma) = |\{i \mid \sigma_i = i\}| – количество неподвижных точек$$

$$n = 3 \quad \xi(\langle 1, 3, 2 \rangle) = 1$$

$$E\xi = 1$$

Мы можем посчитать математическое ожидание, не зная распределение

**Пример:**

$\Omega = \mathbb{B}^n$  – перестановки  $n$  элементов

$$|\Omega| = 2^n$$

$$\xi(\sigma) = |\{i \mid \sigma_i = i\}| – количество неподвижных точек$$

$$n = 3 \quad \xi(\langle 1, 3, 2 \rangle) = 1$$

$$E\xi = 1$$

Мы можем посчитать математическое ожидание, не зная распределение

**Пример:**

$\Omega = \mathbb{B}^n$  – перестановки  $n$  элементов

$$|\Omega| = 2^n$$

$$\xi(\sigma) = |\{i \mid \sigma_i = i\}| – количество неподвижных точек$$

$$n = 3 \quad \xi(\langle 1, 3, 2 \rangle) = 1$$

$$E\xi = 1$$

Мы можем посчитать математическое ожидание, не зная распределение

**Пример:**

$\Omega = \mathbb{B}^n$  – перестановки  $n$  элементов

$$|\Omega| = 2^n$$

$$\xi(\sigma) = |\{i \mid \sigma_i = i\}| – количество неподвижных точек$$

$$n = 3 \quad \xi(\langle 1, 3, 2 \rangle) = 1$$

$$E\xi = 1$$

Мы можем посчитать математическое ожидание, не зная распределение

</div

$\Omega$ ,  $p$  – вероятностное пространство

$\xi : \Omega \rightarrow \mathbb{R}$  – случайная величина (численная характеристика вероятностного эксперимента)

$E\xi = \sum_{\omega \in \Omega} \xi(\omega) \cdot p(\omega) = \sum_a a \cdot p(\xi = a)$  – мат ожидание

$E(\xi + \eta) = E\xi + E\eta$  – мат ожидание суммы равно сумме мат ожиданий

$F_\xi(a) = P(\xi \leq a)$  – функция распределения случайной величины

$f_\xi(a) = P(\xi = a)$  – плотность случайной величины

$\xi$  и  $\eta$  – независимые для  $\forall \alpha \beta$ , если  $[\xi = \alpha]$  и  $[\eta = \beta]$

$\xi$  и  $\eta$  – независимые  $\Rightarrow E(\xi, \eta) = E\xi \cdot E\eta$

$D\xi = E(\xi - E\xi)^2 = E\xi^2 - (E\xi)^2$  – дисперсия

$\xi$  и  $\eta$  независимые  $\Rightarrow D(\xi + \eta) = D\xi + D\eta$

$D(\xi + \eta) = D\xi + D\eta + 2(E\xi\eta - E\xi \cdot E\eta)$

$\text{Cov}(\xi, \eta) = E\xi\eta - E\xi \cdot E\eta$  – ковариация

$\xi$  и  $\eta$  независимые  $\Rightarrow \text{Cov}(\xi, \eta) = 0$

$\text{Cov}(\xi, \xi) = E(\xi \cdot \xi) - E\xi \cdot E\xi = D\xi$

$\text{Cov}(\xi, -\xi) = -E\xi^2 + (E\xi)^2 = -D\xi$  – ковариация может быть отрицательной

**Теорема:**

$$\text{Cov}(\xi, \eta)^2 \leq D\xi \cdot D\eta$$

**Определение:** корреляция

$$\text{Corr}(\xi, \eta) = \frac{\text{Cov}(\xi, \eta)}{\sqrt{D\xi \cdot D\eta}}$$
 – корреляция

$$-1 \leq \text{Corr}(\xi, \eta) \leq 1$$

$$D\xi = 0 \Leftrightarrow E(\xi - E\xi)^2 = 0 \Leftrightarrow p(\omega) > 0 \Rightarrow \xi(\omega) = E\xi$$

Корреляции с константой не бывает, иначе в формуле деление на ноль

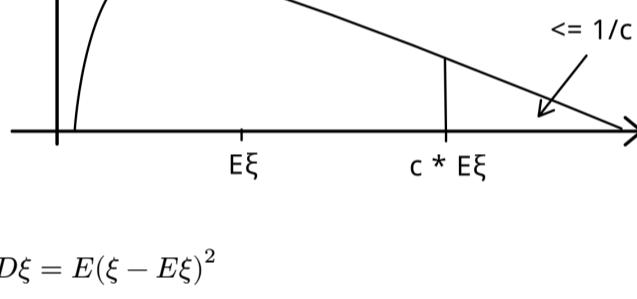
**Теорема:**

$$\text{Corr}(\xi, \eta) = 1 \Leftrightarrow \xi = c \cdot \eta, c > 0$$

$$= -1 \quad c < 0$$

Корреляция не означает причинно-следственной связи

## 1.9. Хвостовые неравенства



**Задача:**

Средняя зарплата 10 опрошенных человек – 50 тысяч. Сколько максимум человек может иметь зарплату больше или равную 250 тысяч рублей?

Максимум 2 человека, в случае, если 8 остальных имеют нулевую зарплату

**Неравенство Маркова** “очень богатых не может быть очень много”

$$\xi > 0 \quad E\xi$$

$$P(\xi \geq c \cdot E\xi) \leq \frac{1}{c}$$

**Доказательство:**

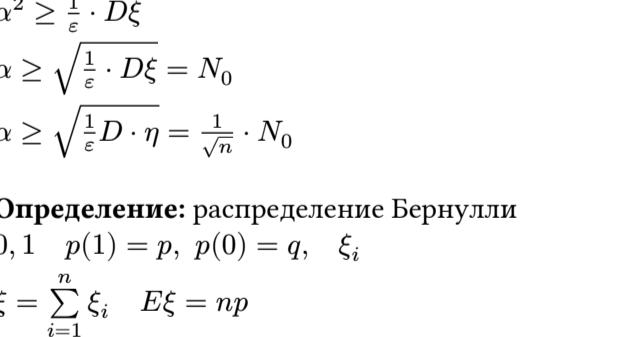
$$P(\xi \geq c \cdot E\xi) = \sum_{\omega: \xi(\omega) \geq c \cdot E\xi} p(\omega)$$

$$E\xi = \sum_{\omega} p(\omega) \cdot \xi(\omega) = \sum_{\omega: \xi(\omega) \geq c \cdot E\xi} p(\omega) \cdot \xi(\omega) + \sum_{\omega: \xi(\omega) < c \cdot E\xi} p(\omega) \cdot \xi(\omega) \geq$$

$$\geq \sum_{\omega: \xi(\omega) \geq c \cdot E\xi} p(\omega) \cdot c \cdot E\xi$$

$$E\xi \geq c \cdot E\xi \cdot P(\xi \geq c \cdot E\xi)$$

$$P(\xi \geq c \cdot E\xi) \leq \frac{1}{c}$$



$$D\xi = E(\xi - E\xi)^2$$

$$\eta = (\xi - E\xi)^2$$

$$P(\eta \geq c^2 E\eta) \leq \frac{1}{c^2}$$

$$P((\xi - E\xi)^2 \geq c^2 D\xi) \leq \frac{1}{c^2}$$

$$P(|\xi - E\xi| \geq c \cdot \sqrt{D\xi}) \leq \frac{1}{c^2}$$

**Определение:** среднеквадратическое отклонение

$$\sigma = \sqrt{D\xi}$$

**Теорема:** неравенство Чебышева

$$P(|\xi - E\xi| \geq c \cdot \sigma) \leq \frac{1}{c^2}$$

$$P(\eta \geq \alpha) \quad \alpha^2 = c^2 E\eta \quad c^2 = \frac{\alpha^2}{D\xi}$$

$$P(|\eta - E\eta| \leq \alpha) \leq \frac{D\xi}{\alpha^2}$$

$$\xi \quad n \text{ раз } \frac{1}{n} \sum_{i=1}^n \xi_i = \eta$$

$\xi_i$  – о. р. незав. сл. величины

$$E\eta = E\xi$$

$$D\eta = \frac{1}{n^2} \cdot n D\xi = \frac{1}{n} D\xi$$

нет. м. о.  $\mu$

$$P(|\xi - \mu| \geq \alpha) \leq \frac{D\xi}{\alpha^2} = \varepsilon$$

$$P(|\eta - \mu| \geq \alpha) \leq \frac{D\eta}{\alpha^2} = \frac{\varepsilon}{n}$$

Для какого  $\alpha$

$$P(|\xi - \mu| \geq \alpha) \leq \varepsilon$$

$$\frac{D\xi}{\alpha^2} \leq \varepsilon$$

$$\alpha^2 \geq \frac{1}{\varepsilon} \cdot D\xi$$

$$\alpha \geq \sqrt{\frac{1}{\varepsilon} \cdot D\xi} = N_0$$

$$\alpha \geq \sqrt{\frac{1}{\varepsilon} D \cdot \eta} = \frac{1}{\sqrt{n}} \cdot N_0$$

**Определение:** распределение Бернулли

$$0, 1 \quad p(1) = p, p(0) = q, \quad \xi_i$$

$$\xi = \sum_{i=1}^n \xi_i \quad E\xi = np$$

$$P(|\xi - np| \geq |\frac{1}{2}n - np|)$$



$$P_{\text{err}} \leq P(|\xi - np| \geq |\frac{1}{2}n - np|)$$

$$P(\eta \geq np - \frac{1}{2}n) \leq e^{-\frac{\delta^2}{2}np}$$

$$P(\eta \leq np + \frac{1}{2}n) \leq e^{-\frac{\delta^2}{2}np}$$

$$e^{-\frac{1}{54}n} \leq \varepsilon$$

$$-\frac{1}{54}n \leq \ln \varepsilon$$

$$n \geq 54 \ln \frac{1}{\varepsilon}$$

$$\xi \quad t \quad \eta = e^{t\xi}$$

$$P(\eta \geq e^{t\alpha}) \leq \frac{1}{e^{t\alpha}}$$

$$E\eta = Ee^{t\xi} - Ee^{\frac{t}{2}\sum_{i=1}^n \xi_i} = E \prod_{i=1}^n e^{t\xi_i} = \prod_{i=1}^n Ee^{t\xi_i} = \prod_{i=1}^n (p \cdot e^t + 1 - p)$$

Сегодня будет не так весело, как на вчерашней лекции парадигм

– Станкевич Андрей Сергеевич

Что такое информация?

## 1.10. Энтропия

информация = -неопределенность

**Определение:** энтропия случайнго источника

Количество информации, приходящейся на одно сообщение источника

$\Omega = \{1, 2, \dots, n\}$  – вероятностное пространство (случайный источник)

Вероятности  $p_1, p_2, \dots, p_n$

$H(p_1, p_2, \dots, p_n)$  – энтропия случайного источника

**Свойства:**

1.  $H$  непрерывна

2.  $H(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$  возрастает при росте  $n$

3. Закон аддитивности

$$\Omega = \{(i, j), i = 1 \dots n, j = 1 \dots m_i\}$$

$$p_i = P(w.\text{first} = i)$$

$$\sum_{i=1}^n p_i = 1$$

$$q_{ij} = p(w.\text{second} = j | w.\text{first} = i)$$

$$\sum_{j=1}^{m_i} q_{ij} = 1$$

$$p((i, j)) = p(w.\text{first} = i) \cdot p(w.\text{second} = j | w.\text{first} = i) = p_i \cdot q_{ij}$$

$$H(p_1 q_{11}, p_1 q_{12}, \dots, p_1 q_{1m_1}, p_2 q_{21}, \dots, p_n q_{nm}) = H(p_1, p_2, \dots, p_n) + \sum_{i=1}^n p_i \cdot H(q_{i1}, q_{i2}, \dots, q_{im_i})$$

$$h(n) = H(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}) \quad h : \mathbb{N} \rightarrow \mathbb{R}$$

$$h(n) \nearrow$$

$$p_i = \frac{1}{n}, \quad q_{ij} = \frac{1}{m}$$

$$h(m \cdot n) = h(n)$$

Такая функция единственна — логарифм. Докажем это

$$h(n) = ?$$

$$i \cdot h(n^i) = i \cdot h(n)$$

$$2^k \leq n^i \leq 2^{k+1} \quad k = \lfloor \log_2 n^i \rfloor = \lfloor i \cdot \log_2 n \rfloor$$

$$k \cdot h(2) \leq i \cdot h(n) \leq (k+1)h(2)$$

$$h(2) \frac{\lfloor i \cdot \log_2 n \rfloor}{i} \leq h(n) \leq h(2) \frac{\lfloor i \cdot \log_2 n \rfloor + 1}{i}$$

$$\forall i \left| h(n) - \frac{\lfloor i \cdot \log_2 n \rfloor}{i} \right| < \frac{1}{i}$$

$$\lfloor i \cdot \log_2 n \rfloor = i \cdot \log_2 n = \{i \cdot \log_2 n\}$$

$$\left| \frac{h(n)}{h(2)} - \log_2 n + \frac{\{i \cdot \log_2 n\}}{i} \right| < \frac{1}{i}$$

$$\frac{h(n)}{h(2)} = \log_2 n \quad h(2) = c$$

$$h(n) = c \cdot \log_2 n$$

Бит, если  $c = 1$

$$p_i = \frac{a_i}{b} \quad q_{ij} = \frac{1}{a_i} \quad m_i = a_i$$

$$H(p_i q_{ij}) = H(p_i) + \sum p_i H(q_{ij})$$

$$h(b) = H(p_1 \dots p_n) + \sum (p_i h(a_i))$$

$$\sum_{i=1}^n p_i \cdot \log_2 b = H(p_1 \dots p_n) + \sum_{i=1}^n p_i \cdot \log_2 a_i$$

$$- \sum_{i=1}^n (p_i (\log_2 a_i - \log_2 b)) = H(p_1 \dots p_n)$$

$$H(p_1 \dots p_n) = - \sum_{i=1}^n p_i \cdot \log p_i$$

**Пример 1:**

Возьмём TeX исходник учебника математического анализа, представимый в виде строки. Рассмотрим все символы в нём и вероятность получить конкретный символ, взяв случайный. Получим энтропию учебника математического анализа – количество информации. Но это не лучшая оценка информации учебника математического анализа, потому что после случайной перестановки символов энтропия (количество информации) не изменится, но количество информации по математическому анализу приравняется к 0

**Пример 2:**

Возьмём случайную строку длинной в миллион символов (число информации несколько миллионов бит). Лучший способ передать её (информацию о ней) – передать весь миллион символов (несколько миллионов бит возможно после небольшого сжатия)

Если мы захотим передать эту строку по телефонному звонку, нам придётся диктовать весь миллион символов. Но если мы просто захотим передать не фиксированную случайную строку в миллион символов, достаточно сказать “запиши случайную строку в миллион символов”. В обоих случаях количество переданной информации равно, но усилия для передачи не равны

**Определение:** сложность по Колмогорову

$$S \in \Sigma^*$$

$K(S)$  – наименьшая длина программы, выводящей  $S$  на пустом входе

$$|K_{\text{C++}}(S) - K_{\text{Java}}(S)| \leq C_{\text{C++, Java}}$$

Можно написать на языке программирования Java интерпретатор языка программирования C++ Его размер не зависит ни от размера программы на C++, ни от  $S$  – константа

$K(S)$  – случайная строка длины  $n$

$K(S) = \Omega(n)$  “почти наверное”

$$S = (01)^n \quad K(S) = O(\log_2 n)$$

C++ плохо выбирать, потому что он такой. Немного мерзкий язык. Выберем математическую модель C++

– Станкевич Андрей Сергеевич

**Пример 1:**

Колмогоровская сложность учебника по математическому анализу меньше, чем колмогоровская сложность случайной строки. Доказательства не будет

Кодирование Хаффмана – оптимальный префиксный код, независимо кодирует  $\forall$  символ

Арифметическое кодирование –

$$L \cdot f_i$$

$$\frac{p}{z^q} \cdot q \simeq -L \sum_{i=1}^n \frac{f_i}{L} \log_2 \frac{f_i}{L} = L \cdot H\left(\frac{f_1}{L}, \dots, \frac{f_n}{L}\right) - \text{энтропийный барьер}$$

Арифметическое кодирование – оптимальный алгоритм кодирования, не учитывающего взаимного расположения символов

Доказательство минимальной асимптотики алгоритма сортировки с помощью энтропийного барьера:

$$h(n!) = \log_2(n!) \simeq \log_2\left(\frac{n^n}{e^n} \sqrt{2\pi n}\right) = n \cdot \log_2 n - n \cdot \log_2 e + \log \sqrt{2\pi n} = \Sigma(n \cdot \log n)$$

## 1.11. Марковские цепи

**Определение:** Марковский процесс

Процесс, дальнейшее проходящее которого зависит от текущего состояния и не зависит от истории

**Пример:**

Не марковские процессы:

1. Ваше обучение в ВУЗе
2. обход графа в ширину и глубину

Если состояние Марковского процесса конечное число, а выбор нового состояния осуществляется случайно,

**Определение:** Марковская цепь

Математический объект состоящий из:

1. Множество состояний  $S = \{1, 2, \dots, n\}$  – конечное множество
2. Матрица переходов  $P_{n \times n} P_{ij} = P(v_{t+1} = j | v_t = i)$  – вероятность перейти из  $i$ -го состояния в  $j$ -ое
3. Начальное распределение вероятностей  $c^{(0)} = (c_1^{(0)} \dots c_n^{(0)})$

$$c^{(t)} = (c_1^{(t)}, c_2^{(t)}, \dots, c_n^{(t)}) \quad \sum_{i=1}^n c_i^{(t)} = 1$$

$$c_j^{(t+1)} = \sum_{j=1}^n P(v_{t+1} = j | v_t = i) = \sum_{i=1}^n p_{ij} c_i^{(t)}$$

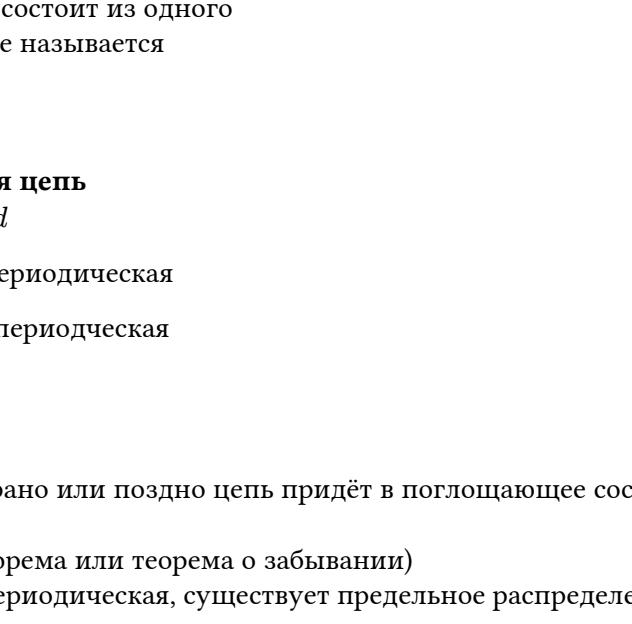
**Основное уравнение Марковских цепей:**

$$c^{(t+1)} = c^{(t)} \cdot P = c^{(t-1)} \cdot P \cdot P$$

В общем случае на  $t$ -ом шаге:  $c^{(t)} = c^{(0)} P^t$

**Пример:**

Студент решает, пойти ему в ВУЗ или нет. Рассмотрим его решения, как Марковскую цепь с 2 состояниями – учиться и играть. Из состояния учиться он с вероятностью 0.5 остается учиться и с вероятностью 0.5 уходит играть. Из состояния играть он с вероятностью 0.9 остается играть и с вероятностью 0.1 идет учиться



Марковский граф

Состояние 1 - учиться, состояние 2 - играть

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{pmatrix}$$

**Определение:** Компонента сильной связности (эргодический класс)

Компонента сильной связности – множество вершин, из каждой вершины которой можно попасть в любую другую вершину

**Определение:** Эргодическая Марковская цепь

Марковская цепь, состоящая из одного эргодического класса

**Определение:** Диаграмма порядка

Диаграмма порядка – граф, получаемый при “объединении” эргодических классов в одну вершину



**Определение:** Поглощающие (существенные) эргодические классы в диаграмме порядка

Поглощающие эргодические классы – классы из которых нельзя выйти

**Определение:** Поглощающее состояние

Если поглощающий класс состоит из одного состояния, то это состояние называется поглощающим

**Эргодическая Марковская цепь**

НОД длин всех циклов –  $d$

1.  $d > 1$  Марковская цепь периодическая

2.  $d = 1$  Марковская цепь апериодическая

**Теорема:** (о поглощении)

Любая цепь поглотится – “рано или поздно цепь придёт в поглощающее состояние”

**Теорема:** (Эргодическая теорема или теорема о забывании)

Если Марковская цепь не периодическая, существует предельное распределение (эргодическое), не зависящее от начального состояния

**Поглощающие Марковские цепи**

Жизнь Поглощающей Марковской цепи состоит из 2 этапов: до поглощения и после

1. Все поглощающие эргодические классы состоят из 1 состояния

2.  $\forall$  состояния  $\exists$  путь до поглощающего состояния

$1 \dots m, \underbrace{m+1 \dots n}_{\text{непогл}}$

1. почему обязательно произойдёт поглощение?

2.  $a_j, j = m+1 \dots n$  – вероятность поглощения в  $j$

3.  $g$  – время до поглощения

**Пример:**

Задача о пьянице (доказательство неизбежности поглощения)



Пьяница с 2 вершинами между домом и баром

$$P = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \Rightarrow Q = \begin{pmatrix} 0 & 0.5 \\ 0.5 & 0 \end{pmatrix} \Rightarrow I - Q = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} \Rightarrow N = (I - Q)^{-1} = \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{4}{3} \end{pmatrix}$$

$$c^{(0)} = (1 \ 0 \ 0 \ 0)$$

$$a = (1 \ 0) \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{4}{3} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \left(\frac{4}{3}, \frac{2}{3}\right) \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \left(\frac{2}{3}, \frac{1}{3}\right)$$

С вероятностью две трети мы закончим свой путь дома, с вероятностью одна третья мы закончим свой путь в баре, не так уж и плохо

– Станкевич Андрей Сергеевич

**Лектор:** Дмитрий Юрьевич

Вопросы, которые мы хотим изучить, изучая марковские цепи:

- С какой вероятностью марковская цепь окажется в том, или ином поглощающем эргодическом классе
- Какое среднее время жизни марковской цепи перед поглощением

$P_{ij}$  — вероятность  $i \mapsto j$

$$\begin{pmatrix} Q & R \\ O & I \end{pmatrix}$$

$P_{ij}^n$  — вероятность  $i \mapsto j$

$$b^0 = \frac{a^0}{m}$$

$$a^0 Q^n$$

$$N = (I - Q)^{-1}$$

$$a^0 N \vec{1}$$

**Эргодическая теорема:**

Не циклический эргодический класс  $\Rightarrow \exists b$  — стационарное распределение  $i$

$$\forall b^0 \quad b^0 P \rightarrow b$$

**Следствие:**

$$b = bP$$

Докажем для регулярных цепей: Регулярная цепь:  $P_{ij} > 0$

?  $\exists N : P^N$  — регулярная

**Доказательство:**

$$b = bP$$

$$b = (I - P) = 0$$

$$b > 0, \sum b_i = 1$$

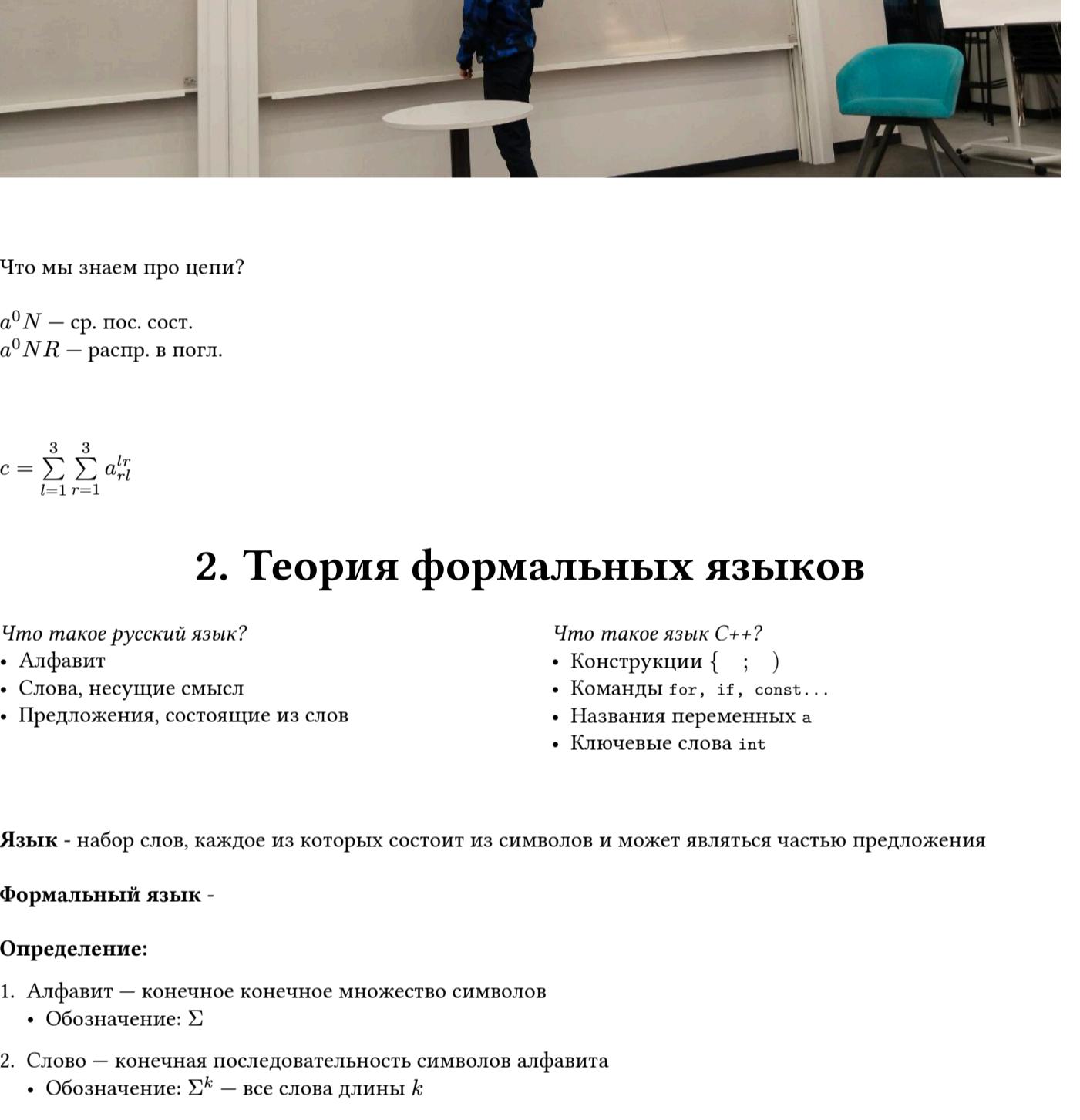
$$b = bA$$

$$bP = \sum_{k=1}^n b_k A_{ki}$$

$$A_{ki} = a_i \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ a_1 & a_2 & \dots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \dots & a_n \end{pmatrix}$$

$M_i^t$  — максимум в  $i$ -том столбце  $p^t$

$m_i^t$  — минимум в  $i$ -том столбце  $p^t$



$M_i^t - m_i^t \rightarrow 0$

$$\begin{aligned} p^{t+1} = p \cdot p^t : \sum_{k=1}^n p_{ik} M_j^t + p_{id} (M_j^t - m_j^t) &\leq M_j^t + \delta (M_j^t - m_j^t) \\ &\geq \sum_{k=1}^n p_{ik} m_j^t + p_{id} (M_j^t - m_j^t) \geq m_j^t + \delta (M_j^t - m_j^t) \end{aligned}$$

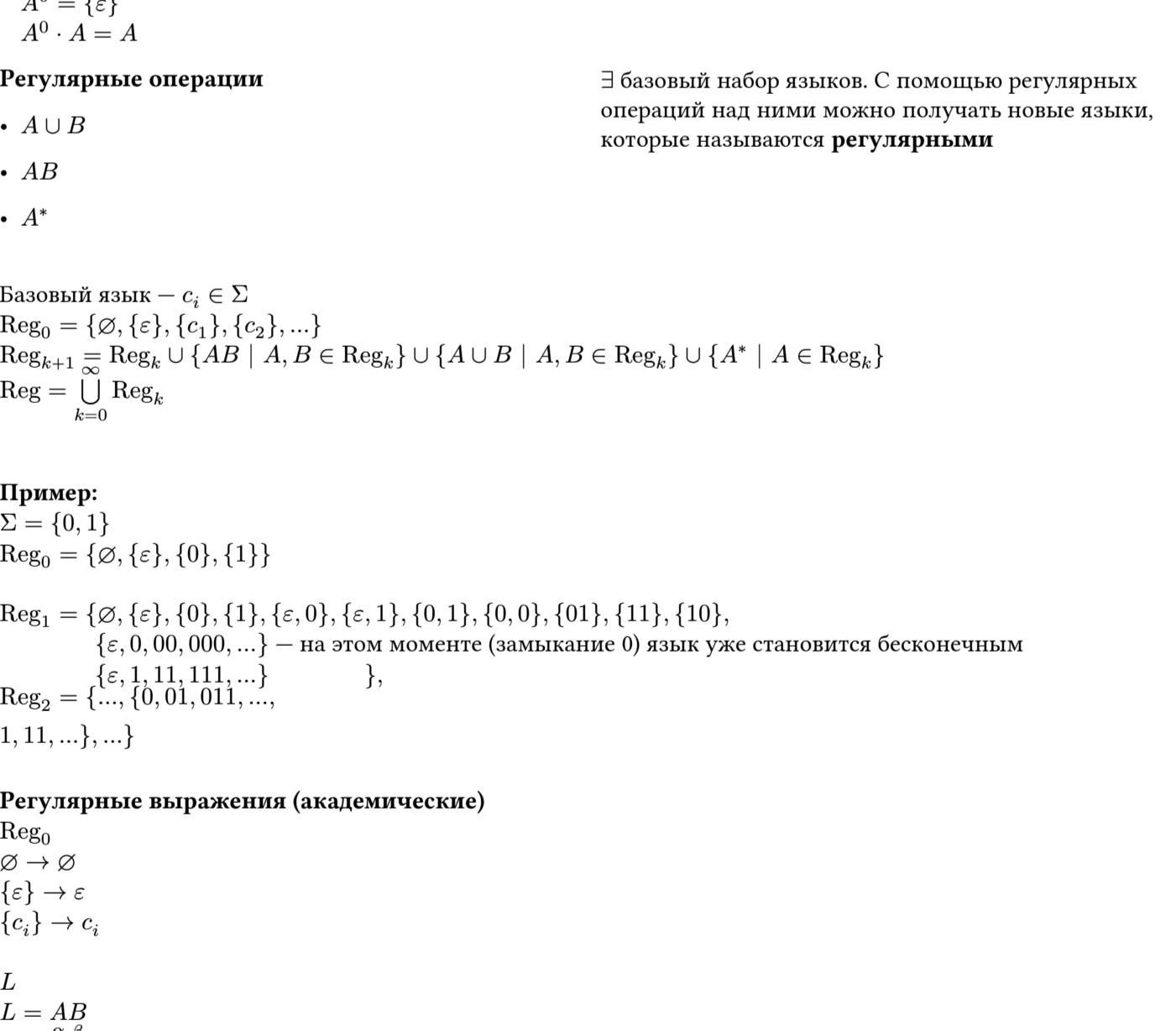
$d$  — номер строчки, в которой достигается минимум

$D$  — номер строчки, в которой достигается максимум

$$p_{ij}^{t+1} \leq M_j^t + \delta (m_j^t - M_j^t)$$

$$\downarrow \Downarrow$$

$$M_j^{t+1} \leq M_j^t + \delta (m_j^t - M_j^t)$$



Что мы знаем про цепи?

$$a^0 N$$
 — ср. пос. сост.

$$a^0 NR$$
 — распр. в погл.

$$c = \sum_{l=1}^3 \sum_{r=1}^3 a_{rl}^t$$

## 2. Теория формальных языков

Что такое русский язык?

- Алфавит
- Слова, несущие смысл
- Предложения, состоящие из слов

Что такое язык C++?

- Конструкции { ; }

- Команды for, if, const...

- Названия переменных a

- Ключевые слова int

Язык — набор слов, каждое из которых состоит из символов и может являться частью предложения

**Формальный язык** —

**Определение:**

1. Алфавит — конечное конечное множество символов

- Обозначение:  $\Sigma$

2. Слово — конечная последовательность символов алфавита

- Обозначение:  $\Sigma^*$  — все слова длины  $k$

3. Формальный язык — набор слов, каждое из которых состоит из конечного числа символов

- Обозначение:  $\bigcup_{k=0}^{\infty} \Sigma^k = \Sigma^*$

Описание языка — конечная строка

Число всех языков:  $2^{\Sigma^*}$  — несчетное

### 2.1. Действия над языками

- Теоретико множественные операции:  $\cup \cap / \bar{L} (U = \Sigma^*)$

Пусть  $\alpha, \beta$  — слова над  $\Sigma$ , тогда конкатенация слов  $\alpha\beta = \gamma \in \Sigma^{k+n}$

$(\alpha\beta)\gamma = \alpha(\beta\gamma)$

$\alpha\beta \neq \beta\alpha$

$\varepsilon$  — нейтральное слово:  $\varepsilon\alpha = \alpha\varepsilon = \alpha$

- Конкатенация языков

$A, B$  — формальные языки

$AB = \{\alpha\beta \mid \alpha \in A, \beta \in B\}$

$\gamma \in AB$

$\gamma = \alpha\beta, \alpha \in A, \beta \in B$

$|A|, |B|$

$|AB| \leq |A| \cdot |B|$

Конкатенация различных слов может давать одинаковый результат

**Пример:**

$A = \{0, 00\}$

$B = \{0, 00\}$

$AB = \{00, 000, 0000\}$

- $A^*$  — замыкание Клини  $A^* = \bigcup_{k=0}^{\infty} A^k$

$\alpha \in A^*$

$\alpha$  можно разложить на  $k$  подстрок, каждая из которых принадлежит  $A$

$\Sigma = \{0', 1', 2'\}$

$00\_1 \in \Sigma^5$

$A^0 = \{\varepsilon\}$

$A^0 \cdot A = A$

$|A| = 2$

$|A^*| = 3$

$\alpha = 0'1'2' \in A^*$

$\alpha = 0'1' + 1'2' \in A^0 \cup A^1$

$\alpha = 0'1' + 1'2' \in A^0 \cdot A^1$

$\alpha = 0'1' + 1'2' \in A^0 + A^1$

$\alpha = 0'1' + 1'2' \in A^*$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16} + A^{17}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16} + A^{17} + A^{18}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16} + A^{17} + A^{18} + A^{19}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16} + A^{17} + A^{18} + A^{19} + A^{20}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16} + A^{17} + A^{18} + A^{19} + A^{20} + A^{21}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16} + A^{17} + A^{18} + A^{19} + A^{20} + A^{21} + A^{22}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16} + A^{17} + A^{18} + A^{19} + A^{20} + A^{21} + A^{22} + A^{23}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A^{16} + A^{17} + A^{18} + A^{19} + A^{20} + A^{21} + A^{22} + A^{23} + A^{24}$

$\alpha = 0'1' + 1'2' \in A^0 + A^1 + A^2 + A^3 + A^4 + A^5 + A^6 + A^7 + A^8 + A^9 + A^{10} + A^{11} + A^{12} + A^{13} + A^{14} + A^{15} + A$

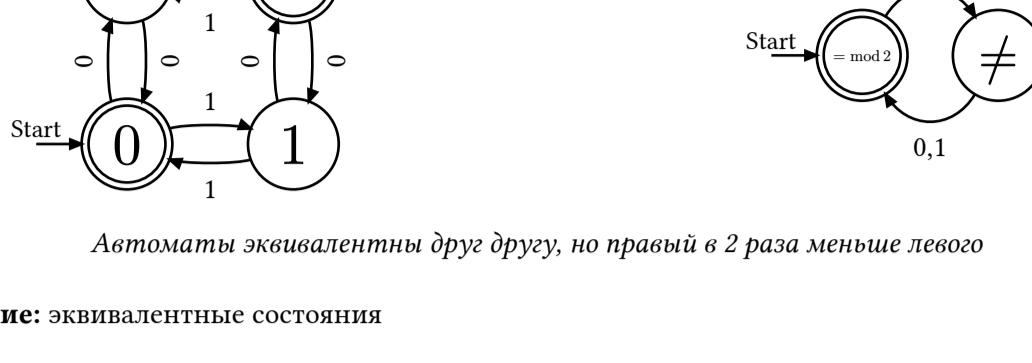


## 2.6. Регулярные языки

### 2.6.1. Минимизация ДКА

$L = \{\text{строки из } 0, 1: \#0 = \#1 \bmod 2\}$

Примеры автоматов для такого языка:



Автоматы эквивалентны друг другу, но правый в 2 раза меньше левого

**Определение:** эквивалентные состояния

$A$  - ДКА

$u, v \in Q$

$u \sim v$ , если  $\forall s \in \Sigma^*: [\delta(u, s) \in T] = [\delta(v, s) \in T]$

$\sim$  — отношение эквивалентности:

1. Рефлексивность:  $u \sim u$  — очевидно
2. Симметричность:  $u \sim v \Leftrightarrow v \sim u$  — очевидно.
3. Транзитивность:  $u \sim v \wedge v \sim w \Rightarrow u \sim w$

**Доказательство:** (от противного)

Из  $u$  приходим в терминальное состояние. Из  $w$  приходим в не терминальное состояние. Тогда, посмотрим на  $v$ . если  $v$  приходит в

терминальное состояние то  $v \not\sim w$ , в другом случае  $u \not\sim v$

**Занимательный факт:**

Для минимизации ДКА необходимо выделить все эквивалентные состояния и объединить их между собой

Так как  $\sim$  — отношение эквивалентности, мы можем выделять классы эквивалентности — вершины минимизированного автомата

Если до перестройки автомата между вершинами классов  $u$  и  $v$  было проведено ребро  $s_1$ , то оно должно быть проведено между классами  $K(u)$  и  $K(v)$

**Утверждение:**

$u \sim v, c \in \Sigma \Rightarrow \delta(u, c) \sim \delta(v, c)$

**Доказательство:**

Если  $\delta(u, c)$  и  $\delta(v, c)$  отличаются строкой  $s$ , то  $u$  и  $v$  отличаются строкой  $cs$

**Алгоритм выделения классов эквивалентности** ( $O(n^2 \cdot |\Sigma|)$ )

- $u \not\sim v \wedge \delta(a, c) = u \wedge \delta(b, c) = v \Rightarrow a \not\sim b$

used[n][n]

used[u][v] = True  $\Rightarrow u \not\sim v$

```
for x in T:
    for y !in T:
        used[x][y] = used[y][x] = True
        queue.add((x, y))
        queue.add((y, x))
while queue:
    u,v=queue.next()
    for c in Sigma:
        for x in Nodes.where(x=> u in x.at(c)):
            for y in Nodes.where(y=> v in y.at(c)):
                if not used[x][y]:
                    used[x][y]=True
                    queue.add((x,y))
```

- $u \not\sim v \wedge \delta(a, c) = u \wedge \delta(b, c) = v \Rightarrow a \not\sim b$

• Строим новый автомат  $A$

**Теорема:**  $A$  - ДКА,  $\forall u, v \in Q, u \not\sim v$ , все вершины  $A$  достижимы из  $S_A$ . тогда

1.  $\forall B$  - ДКА  $L(A) = L(B), |Q_B| \geq L(Q_A)$
2.  $\forall B$  - ДКА  $L(A) = L(B), |Q_B| = L(Q_A) \Rightarrow A$  и  $B$  изоморфны

**Доказательство:**

1. • Запишем автомат  $A$  и автомат  $B$  как один большой автомат с двумя стартовыми вершинами

- $S_A \sim S_B$
- $\forall v \in Q_A \exists u \in Q_B : v \sim u$

• Предположим, что  $|Q_B| < |Q_A| \Rightarrow \exists x, y \in Q_A : x \sim z, y \sim z \Rightarrow x \sim y$  — противоречие

2.  $|Q_A| = |Q_B|$  — биекция

$\Pi : Q_A \rightarrow Q_B$

$x \in \Pi(x) \Rightarrow \delta_A(x, c) \sim \delta_B(\Pi(X), c)$

## 2.7. Нерегулярные языки

**Классический пример нерегулярного языка:**

$L$  - язык ПСП.  $\{ "", "()", "(()", "())()", ... \}$

Для того, чтобы проверить, является ли строка ПСП, необходимо пройтись по строке, поддерживая баланс скобочек. Если во время прохода баланс окажется отрицательным или будет не равен 0 после прохода по строке, строка не ПСП

Чтобы создать автомат для проверки ПСП, необходимо создать по состоянию на каждое значение баланса скобок, что не получится сделать за конечное число вершин

**Доказательство:**

$L \notin \text{Reg}$ , от противного:

$\exists L \in \text{Reg} \Rightarrow \exists A$  - ДКА,  $L(A) = \text{ПСП}, |Q_A| = n$

Построим последовательность

$s_1 = "()", s_2 = "((", ..., s_{n+1} = s_n + ($

Тогда, по принципу Дирихле, будет хотя бы одно совпадение состояний.

**Лемма о разрастании (boost) (о накачке (pumping)):**

$L \in \text{Reg} \Rightarrow \exists n > 0 : \forall w \in L, |w| \geq n, \exists x, y, z : w = xyz, |xy| \leq n, y \neq \epsilon \Rightarrow \forall k \geq 0, xy^k z \in L$

**Отрицание леммы о разрастании:**

$\forall n > 0 \exists w \in L, |w| \geq n : \forall x, y, z : w = xyz, |xy| \leq n, y \neq \epsilon \exists k \geq 0 : xy^k z \notin L \Rightarrow L \notin \text{Reg}$

**Пример:**

$L = \{0^n 1^n\} = \{\epsilon, 01, 0011, ...\}$

$L \in \text{Reg}$

1.  $n > 0$

2. выбираем  $w = 0^n 1^n$

3.  $x = 0^a, y = 0^b, z = 0^{n-a-b} 1^n$

4.  $k = 0 \Rightarrow xk^k z = 0^a 0^{n-a-b} 1^n = 0^{n-b} 1^n$

$L \in \text{Reg} \Rightarrow \exists A$  - ДКА,  $n = |Q_A|$

