Конспект по дискретной математике
II семестр

Коченюк Анатолий

13 апреля 2021 г.

Глава 1

Дискретная теория вероятностей

1.1 Введение

Определение 1 (Вероятностное пространство)**.**
Ω – элементарные исходы, неделимые дальше.

p – дискретная плотность вероятности.

$p : \Omega \to [0, 1] \quad \sum\limits_{q \in \Omega} p(\omega) = 1$

Замечание. В случае дискретного вероятностного пространства $|\Omega|$ – не более, чем счётное.

Пример (Честная монета)**.** $\Omega = \{0, 1\} \quad p(0) = p(1) = \frac{1}{2}$

Пример (Нечестная монета)**.** $\Omega = \{0, 1\} \quad p(1) = p, p(0) = q$ – различные числа. $p + q = 1$

Ещё одно название – распределение Бернулли

Пример (Честная игральная кость)**.** $\Omega = \{1, 2, 3, 4, 5, 6\} \quad p(\omega) = \frac{1}{6}$

Определение 2. Событие, случайное событие – $A \subseteq \Omega$

Замечание. Неправильное определение – то, что может произойти, а может не произойти.

$\emptyset \subseteq \Omega \quad \Omega \subseteq \Omega$ – примеры, когда никогда не происходит и всегда происходит

Замечание. Для недискретного случая неверно, что <u>любое</u> подмножество Ω это событие

> **Определение 3.** Вероятность события $P(A) = \sum\limits_{\omega \in A} p(\omega)$
>
> p берёт элементарные исходы. $P, \mathbb{P}$ – вероятность события

Пример. Событие $E = \{2, 4, 6\} \quad P(E) = p(2) + p(4) + p(6) = \frac{3}{6} = \frac{1}{2}$

$O = \{1, 3, 5\}$

Замечание. Не существует вероятностного пространства с бесконечным числом равновероятных исходов

$p(\omega) = 0 \quad \sum = 0$

$p(\omega) = a > 0 \quad \sum = a \cdot (+\infty) = +\infty$

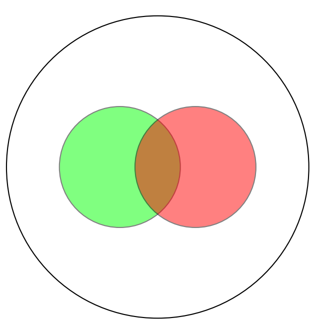
Пример. Событие $B(IG) = \{4, 5, 6\} \quad P(B) = \frac{1}{2}$

> **Определение 4** (Независимое событие)**.** События A, B независимы, если $P(A \cap B) = P(A) \cdot P(B)$

Пример. $E \cap O = \emptyset \quad B \cap E = \{4, 6\}$

$P(E \cap O) = \emptyset \quad P(O) \cdot P(B) = \frac{1}{4} \neq 0$

$P(B) \cdot P(E) = \frac{1}{4} \neq \frac{1}{3} = P(B \cap E)$



$P(A \cap B) = P(A) \cdot P(B)$

$\frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(\Omega)}$

Определение 5 (Условная вероятность)**.** $P(A|B) = \frac{P(A\cap B)}{P(B)}$

Замечание. Альтернативное определение независимости, не поддерживающее 0: $P(A|B) = P(A)$

$V = \{5, 6\}$

$P(V \cap E) = \frac{1}{6}$

$P(V) = \frac{1}{3} \quad P(E) = \frac{1}{2} \quad P(V) \cdot P(E) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6} = P(V \cap E)$

Определение 6 (Произведение вероятностных пространств)**.**
$\Omega_1, p_1 \qquad \Omega_2, p_2$

$\Omega = \Omega_1 \times \Omega_2$

$p\left(\langle \omega_1, \omega_2 \rangle\right) = p_1\left(\omega_1\right) \cdot p_2(\omega_2)$

Теорема 1. $\forall A_1 \subseteq \Omega_1$ и $\forall A_2 \subseteq \Omega_2$

$A_1 \times \Omega_2$ и $\Omega_1 \times A_2$ – независимы

Доказательство. $P\left(A_1 \times \Omega_2 \cap \Omega_1 \times A_2\right) = P\left(A_1 \times A_2\right) = \sum\limits_{\substack{a \in A_1 \\ b \in A_2}} p\left(\langle a, b \rangle\right) =$

$$= \sum_{a \in A_1} \sum_{b \in A_2} p_1(a) \cdot p_2(b) = \sum_{a \in A_1} p_1(a) \left(\sum_{b \in A_2} p_2(b) \right) = P_1(A_1) \cdot P_2(A_2) \qquad \blacksquare$$

Определение 7. $A_1, A_2, \ldots, A_n$

1. Попарно независимые $\quad A_i$ и A_j независимы
2. Независимы в совокупности $\forall I \subseteq \{1, 2, \ldots, n\}$ $P\left(\bigcap\limits_{i \in I} A_i\right) = \prod\limits_{i \in I} P(A_i)$

   $P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$

Пример. Кидаем две монеты $\Omega = \{00, 01, 10, 11\}$

$A_1 = \{10, 11\} \quad A_2 = \{01, 11\} \quad A_3 = \{01, 10\}$ – независимы попарно, но не в совокупности

Определение 8 (Формула полной вероятности)**.**
$\Omega = A_1 \cup A_2 \cup \ldots \cup A_n \quad i \neq j \implies A_i \cap A_j = \emptyset$

Совокупность таких А-шек называется полной системой событий.

Дано: вероятности $P(A_i) \quad P(B|A_i)$ Найти: $P(B)$

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

– формула полной вероятности

Найти: $P\left(A_j|B\right)$

A_1 – болен, A_2 – здоров, B – положительный результат теста

$P(A_2|B)$

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j) \cdot P(A_j)}{\sum\limits_{i=1}^{n} P\left(B|A_i\right) \cdot P(A_i)}$$

– формула Байеса

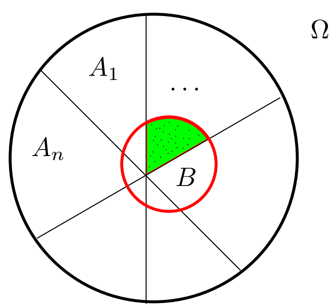


Рис. 1.1: B

1.2 Случайные величины

Замечание. Неправильное (наивное) определение – величина, принемающая слуйное значение.

Она может быть константой. Что такое величина?

Определение 9 (Случайная величина)**.** $\xi : \Omega \to \mathbb{R}$ – $\mathbb{R}$ -значная функция

Ω, p – вероятностное пространство.

Пример. Если взять случайные текст длинной 1Кб. Вариантов текста очень много и бессмысленно их рассматривать отдельно, интересует какое-то свойство, величина.

Графы, $2^{\binom{n}{2}}$ штук. Но нас интересует какая-то (численная) характеристика элементарного исхода.

Пример. $D(ice) = \{1, 2, 3, 4, 5, 6\}$

$\Omega = D^2 \quad p\left(\langle i, j\rangle\right) = \frac{1}{36}$

$\xi : \Omega \to \mathbb{R} \qquad \xi(\langle i, j\rangle) = i + j$

Пример (Случайные графы)**.** $G(4, \frac{1}{2})$ – случайный граф, 4 вершины, каждое ребро существует с вероятностью $\frac{1}{2}$

$\Omega = \mathbb{B}^6 \quad p(G) = \frac{1}{64}$

$\xi(G) =$ количество компонент связности

Пример. $\Omega = \{1, 2, 3, 4, 5, 6\}$ $\xi(w) = w$

Пример. $\Omega = \{1, 2, 3, 4, 5, 6\} \quad E = \{2, 4, 6\}$

$\chi_E(\omega) = \begin{cases} 1, & \omega \in E \\ 0, & \omega \notin E \end{cases}$ – индикаторная случайная величина

Определение 10. $\Omega, p \quad \xi$

$[\xi = i] = \{\omega | \xi(\omega) = i\} \subseteq \Omega$

$P\left([\xi = i]\right) = P\left(\xi = i\right) = f_\xi(i) \quad f : \mathbb{R} \to \mathbb{R}$

$f_\xi(i) = P\left(\xi = i\right)$ – дискретная плотность вероятности случайной величины ξ

$F_\xi(i) : \mathbb{R} \to \mathbb{R} = P\left(\xi \leqslant i\right)$ – функция распределения

Замечание. Непрерывная vs Дискретная вероятность

Пример.

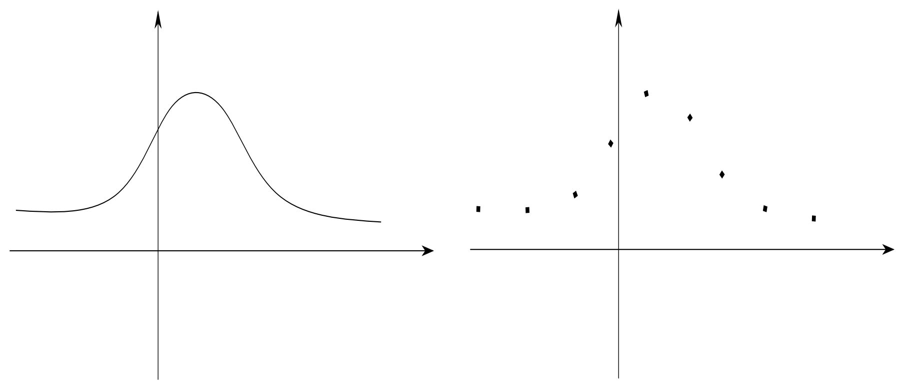


Рис. 1.2: непрерывная и дискретная вероятность

Замечание. $\delta(x) = \begin{cases} 0, & x \neq 0 \\ +\infty, & x = 0 \end{cases}$

$\int\limits_{-\infty}^{+\infty} \delta(x) dx = 1$

$f_\xi(i) = P\left(\xi = i\right) \qquad f_\xi = \int\limits_{-\infty}^{+\infty} f_{xi}(x) = F_\xi(i)$

$f_\xi(x) = \sum\limits_i P\left(\xi = i\right) f\left(x - i\right)$

Пример. $\Omega = \mathbb{B}^{1000} \quad p(\omega) = \frac{1}{2^{1000}}$

$\xi(w) =$ число 1 в ω

$|$множество значений $\xi| = 1001 \qquad p\left(\xi = i\right) = \frac{\binom{1000}{i}}{2^{1000}}$

Замечание. Случайные числа обозначаются строчными греческими или заглавными латинскими из конца алфавита (X, Z)

Замечание (Что можно делать со случайными величинами)**.** ξ, η – функции

$\xi^2 \quad 2\xi \quad \xi + \eta \quad \xi \cdot \eta \quad \xi^\eta \quad \sin\xi \quad e^\eta \quad \frac{1+\xi}{\eta}$ (всё то же, что мы можем делать с функциями.

Пример. $\Omega = D^2 \quad \xi_1\left(\langle i, j\rangle\right) = i \qquad \xi_2\left(\langle i, j\rangle\right) = j$ – одинаково распределённые случайные величины

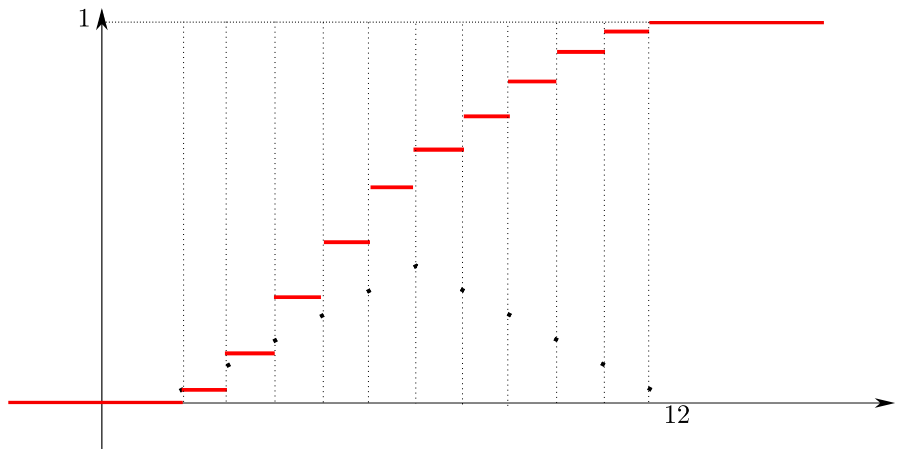


Рис. 1.3: Два-кубика (функция распределения)

Пример. $\Omega = F \quad id(\omega) = \omega$

$1, 2, \ldots, 6$ – каждый с вероятностью $\frac{1}{6}$. Другое вероятностное пространство относительно предыдущего примера, но всё равно одинаковое распределение.

$\xi = (i + j) = \xi_1 + \xi_2 \qquad \xi = (i + j)\%6 + 1$ – у второй то же распределение, что у верхних, но она уже совсем другая.

Определение 11. Математическое ожидание

$$E_\xi = \sum_{\omega} p\,(\omega)\,\xi(\omega).$$

Утверждение 1. $E_\xi = \sum_i ip(\xi = i)$

Доказательство.

$$\begin{aligned} E_\xi &= \sum_\omega p(\omega)\,\xi(\omega) \\ &= \sum_i \sum_{w:\xi(\omega)=i} p(\omega)\cdot\xi(\omega) \\ &= \sum_i \sum_{\omega:\xi(\omega)=i} p(\omega)\cdot i \\ &= \sum_i i \sum_{w:\xi(\omega)=i} p(\omega) \\ &= \sum_i iP(\xi=i) \end{aligned}$$

.

■

Пример. $\Omega = D \quad \xi = id$

$E_\xi = \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 2 + \frac{1}{6}\cdot 3 + \frac{1}{6}\cdot 4 + \frac{1}{6}\cdot 5 + \frac{1}{6}\cdot 6 = \frac{21}{6} = \frac{7}{2} = 3,5$

Пример. $\Omega = D^2 \quad \xi(\langle i,j\rangle) = i+j$

$E_\xi = \frac{1}{36}\cdot(2\cdot 1 + 3\cdot 2 + 4\cdot 3 + 5\cdot 4 + 6\cdot 5 + 7\cdot 6 + 8\cdot 5 + 9\cdot 4 + 10\cdot 3 + 11\cdot 2 + 12\cdot 1)$ – здесь среднее значение оказалось наиболее частым, но так оно не всегда (пример с 3,5)

Теорема 2. $E_{\lambda\xi} = \lambda E_\xi$

$E(\xi+\eta) = E_\xi + E_\eta$

Доказательство. $E_{\lambda\xi} = \sum\limits_\omega p(\omega)\,\lambda\xi(\omega) = \lambda E_\xi$

$E(\xi+\eta) = \sum\limits_\omega p(\omega)\,(\xi(\omega)+\eta(\omega)) = E_\xi + E_\eta$ ■

Утверждение 2. Если ξ и η одинаково распределены, то $E_\xi = E_\eta$

Пример. Бросим кубик один раз, ξ_1 – что выпало сверху, ξ_2 – что выпало снизу

$E(\xi_1+\xi_2) = 7.$ – не играет роли как числа друг относительно друга расположены.

МАТОЖИДАНИЕ ЛИНЕЙНО ВСЕГДА

Пример. $\Omega = S_n \quad p(\omega) = \frac{1}{n!}$

$\xi(\pi) = |\{i|\pi[i] = i\}| \quad 0 \ldots n$, кроме $n-1$

$E_\xi = \sum\limits_{j=1}^{n} \xi_i = 1$

$\xi_i(\pi) = \begin{cases} 1, & \pi[i] = i \\ 0, & \text{иначе} \end{cases}$

$E_{\xi_i} = \frac{1}{n}$

$\xi = \sum\limits_{j=1}^{n} \xi_i$

1.3 Независимые случайные величины

Определение 12 (удобное)**.** Случайные величины ξ и η независимы, если события $[\xi = \alpha]$ и $[\eta = \beta]$ – независимы $\forall \alpha, \beta$

Определение 13 (нормальное)**.** $[\xi \leqslant \alpha]$ и $[\eta \leqslant \beta]$ – независимы для $\forall \alpha,]beta$

Пример. $\Omega = \Omega_1 \times \Omega_2$

$\xi_1(\langle \omega_1, \omega_2 \rangle) = f(\omega_1)$

$\xi_2(\langle \omega_1, \omega_2 \rangle) = g(\omega_2)$

A и B независимы, χ_A, χ_B – независимы

Теорема 3. ξ, η – независимы $\implies E(\xi \cdot \eta) = E_\xi \cdot E_\eta$

Доказательство. $E\xi \cdot \eta = \sum\limits_{\alpha} \alpha \cdot P(\xi \cdot \eta = \alpha) = \sum\limits_{i,j} \alpha P([\xi = i] \cap [\eta = j]) = \sum\limits_{i} \sum\limits_{j} ijP(\xi = i) P(\eta = j) = E_\xi E_\eta$

$i \cdot j = \alpha \quad i \in R_\xi \quad j \in R_\eta$ ∎

Пример. $\Omega = \{0, 1\} \quad p = \frac{1}{2} \quad \xi(i) = 2i \quad E_\xi = 1$

$\Omega = S_n \quad p = \frac{1}{n!} \quad \xi =$ число неподвижных точек $\quad E_\xi = 1$

Матожидание одно, но ведут себя совершенно по разному.

Определение 14 (Дисперсия)**.** $D_\xi = Var\left(\xi\right)$

$D_{xi} = E\left(\xi - E_\xi\right)^2 = E\left(\xi^2 - 2\xi E_\xi + \left(E_\xi\right)^2\right) = E_{xi}^2 - 2E_\xi E_\xi + \left(E_\xi\right)^2 = E_{\xi^2} - \left(E_\xi\right)^2$

Теорема 4. $D_{c\xi} = c^2 D_\xi$

Если ξ и η независимы, то $D_{\xi+\eta} = D_\xi + D_\eta$

Доказательство. Упражнение ∎

Вспомним. $\xi, \eta : \Omega \to \mathbb{R}$

$F_\xi(a) = P\left(\xi \leqslant a\right)$

$f_\xi(a) = P\left(\xi = a\right)$

$F_\xi(a) = \sum\limits_{b \leqslant a} f_\xi(b)$

$E_\xi = \sum\limits_{\omega \in \Omega} p(\omega)\xi(\omega) = \sum\limits_a a \cdot P\left(\xi = a\right)$

$E(\xi + \eta) = E_\xi + E_\eta$ ∎

$E\left(\xi - E\xi\right) = E_\xi - EE_\xi = 0$ матожидане отклонения от матожидания равно нулю..

Хочется смотреть насколько величина отклоняется от своего матожиданя. Для этого используется понятие дисперсии:

$D_\xi = E\left(\xi - E\xi\right)^2 = E\xi^2 - \left(E\xi\right)^2$

$D\left(\xi + \eta\right) = E\xi^2 + E\eta^2 + 2E\xi\eta - \left(E\xi\right)^2 - \left(E\eta\right)^2 - 2E\xi E\eta = D_\xi + D_\eta + 2\left(E_{\xi\eta} - E_\xi E_\eta\right)$.

В случае независимых случайных величин дисперсия линейна. Иначе она отличается на ковариацию:

Определение 15 (Ковариация)**.**

$$Cov\left(\xi, \eta\right) = E_{\xi\eta} - E_\xi E_\eta.$$

$D_\xi = Cov\left(\xi, \xi\right)$

Определение 16 (Корреляция).

$$Corr\,(\xi,\eta) = \frac{Cov\,(\xi,\eta)}{\sqrt{D_\xi D_\eta}}.$$

Теорема 5. Корреляция двух случайных величин лежит между -1 и 1.

$$-1 \leqslant Corr(\xi,\eta) \leqslant 1.$$

Доказательство. $\alpha = \xi - \lambda\eta$

$D_\alpha = E_{\xi^2} - 2\lambda E_{\xi\eta} + \lambda^2 E\left(\eta^2\right) - (E(\xi))^2 + 2\lambda E_\xi E_\eta - \lambda^2\left(E_{\eta^2}\right) \geqslant 0$

$D_\xi + 2\lambda Cov\,(\eta,\eta) + \lambda^2 D_\eta$

$4Cov\,(\xi,\eta)^2 - 4D_\xi D_\eta \leqslant 0$ ■

1.4 Хвостовые неравенства

$\xi \quad E\xi = 10 \quad \xi \geqslant 0$

$P\,(\xi \geqslant 100) < \frac{1}{10}$

Теорема 6 (Неравенство Маркова). $\xi \not\equiv 0 \quad \xi \geqslant 0 \quad P\,(\xi \geqslant a \cdot E_\xi) \leqslant \frac{1}{a}$

Доказательство. $E_\xi = \sum\limits_v v \cdot P\,(\xi = v) = \sum\limits_{v < a \cdot E_\xi} v \cdot P\,(\xi = v) + \sum\limits_{v \geqslant a \cdot E_\xi} v \cdot P\,(\xi = v) \geqslant 0 + a \cdot E_\xi \sum\limits_{v \geqslant a \cdot E_\xi} P\,(\xi = v) = a \cdot E_\xi \cdot P\,(\xi \geqslant a \cdot E_\xi)$ ■

Пример. $a = \frac{c}{E_\xi} \quad P\,(\xi \geqslant c) \leqslant \frac{E_\xi}{c}$

$D_\xi = E\,(\xi - E\xi)^2$

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

$P((\xi - E_{xi})^2 \geqslant a^2 \cdot D_\xi) \leqslant \frac{1}{a^2}$

$\sigma = \sqrt{D_\xi}$ – среднеквадратичное отклонение

Теорема 7 (Неравенство Чебышева). $P\,(|\xi - E_\xi| \geqslant a\sigma) \leqslant \frac{1}{a^2}$

$P\,(|\xi - E_\xi| \geqslant c) \leqslant \frac{D_\xi}{c^2}$

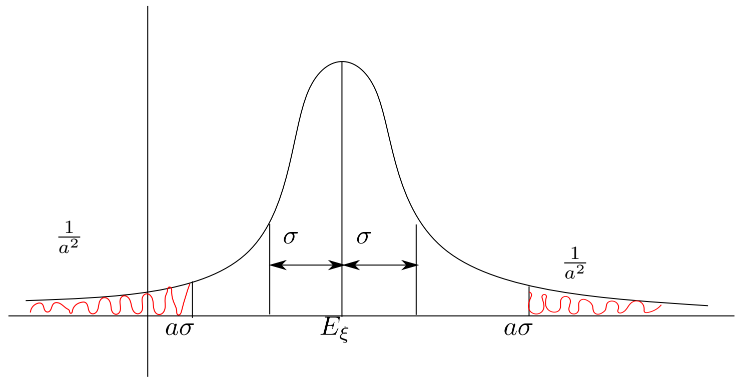


Рис. 1.4: drawing

Задача 1 (10 монет, найти количество “1”)**.**

$$E\xi = 5 \quad D\xi = 2{,}5$$

$$P(\xi \leqslant 0) \leqslant P(|\xi - E\xi| \geqslant 5) \leqslant \frac{2{,}5}{25} = \frac{1}{10}$$

Замечание. С одной стороны, у неравенства есть плюс: оно **всегда** работает; всегда (!)

С другой — иногда оценки получаются, очень грубыми. В нашем примере ответ $\leqslant \dfrac{1}{10}$, а в жизни — $\dfrac{1}{1024}$.

Пример. Нечестная монета $p \neq \frac{1}{2}$. Хотим выяснить чем чаще выпадает.

Бросили: c единиц, $n - c$ нулей. Предположим, что $c < \frac{n}{2} \quad p > \frac{1}{2} \quad pn > \frac{n}{2}$

$P(\xi = c) \leqslant P(\xi \leqslant c) \leqslant P(|\xi - pn| \leqslant pn - c) \leqslant P\left(|\xi - pn| \leqslant \frac{n}{2} - c\right) \leqslant \frac{n}{4 \cdot \left(\frac{n}{2} - c\right)^2}$

Теорема 8 (Граница Чернова (без доказательства))**.** $\xi_i \quad P(\xi_i = 1) \quad P(\xi_i = 0) = 1$

$\xi = \sum\limits_{i=1}^{n} \xi_i \qquad E_\xi = np = \mu$

$P(|\xi - \mu| \geqslant \delta\mu) < e^{-\mu\frac{\delta^2}{3}}$

Пример. Случайная величина ξ. Хотим узнать матожидание. Проведём эксперимент n раз: $\xi_1, \xi_2, \ldots, \xi_n$

$$P\left(\left|\frac{\sum \xi_i}{n} - E_\xi\right| > c\right) \leqslant \frac{D_\xi}{n\varepsilon^2}.$$

$\xi : \Omega \to \mathbb{Z}^+$

$E_\xi = \sum\limits_{i=0}^{n} i \cdot P\left(\xi = i\right) = \sum\limits_{i=0}^{n} \left(P\left(\xi \geqslant i\right) - P\left(\xi \geqslant i+1\right)\right) = \sum\limits_{i=1}^{n} P\left(\xi \geqslant i\right)$

1.5 Теория информации

Определение 17 (Что такое информации)**.**
Информация = − неопределённость

непределённость H1. Что-то узнали, стала неопределённость H2. полученная информация I =H1-H2 $= -\Delta$ H

Хочется убрать наблюдателя, нас, из определения, чтобы не было кого-то, кто узнаёт и меняет неопределённость. Надо ввести объективную модель:

Определение 18 (Случайный источник)**.** Ω – вероятностное пространство. Есть исходы $p_1, p_2, \ldots, p_n$

Чёрный ящик с красной кнопкой и дисплеем. Основан на вероятностном пространстве

$\xi_1, \xi_2, \ldots, \xi_m \cdots$

$P(\xi_i = a) = p_a \quad a = 1 \ldots n$

Случайный источник $p_1, p_1, \ldots, p_n$. Хотим померять сколько информации содержится в одном результате эксперимента.

$H\left(p_1, p_2, \ldots, p_n\right) : RS(randomsources) \to \mathbb{R}^+$

Частный случай $p_i = \frac{1}{n}$

$h(n) = H\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$

1. $h(n+1) > h(n)$
2.

   Пример. $\Omega = \{(1,1), (1,2), \ldots, (1, m_1)(2,1), \ldots, (2, m_2), \ldots, (k,1), \ldots, (k, m_k)\}$

   $n = m_1 + m_2 + \ldots m_n$

   $p(i,j) = 1_{ij} \quad p_i = \sum\limits_{j=1}^{m_k} q_{ij}$

Первый ряд $(1, *)$ – p_1. Второй p_2 ... Последний p_k

Если случайный источник показывает только первое число это эквивалентно $H(p_1, p_2, \ldots, p_n)$

Теперь представим, что мы сначала узнаём первую компоненту, а потом открываем вторую

$\sum\limits_{i=1}^{k} p_i H\left(\frac{q_{i1}}{p_i}, \ldots, \frac{q_{im_i}}{p_i}\right)$

Если провести эксперимент сразу, получим $q_{11}, \ldots, q_{im_i}$

$q_{ij} = p_i q_{ij}$

$$H\left(p_1 r_{11}, p_1 r_{12}, \ldots, p_1 r_{1k_1}, p_2 r_{21}, \ldots, p_k r_{km_k}\right) = H\left(p_1, p_2, \ldots, p_k\right) + \sum_{i=1}^{k} p_i H\left(r_{i1}, \ldots, r_{im_i}\right).$$

3. Для фиксированного n H непрерывная как функция $\mathbb{R}^n \to \mathbb{R}$

Теорема 9. $H\left(p_1, p_2, .., p_n\right) = -\sum\limits_{i=1}^{n} p_i \log p_i$

Лемма 1. $h\left(nm\right) = H_(n) + h(m)$ Следует из второго свойства

Доказательство. $k = n \quad m_i = m \quad p_i = \frac{1}{n} \quad q_{ij} = \frac{1}{nm} \quad r_{ij} = \frac{1}{m}$

$h(nm) = H\left(q_{11}, q_{12}, \ldots, q_{nm}\right) = H\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) + \sum\limits_{i=1}^{n} \frac{1}{n} H\left(\frac{1}{m}, \ldots, \frac{1}{m}\right) = h(n) + h(m)$ ■

Определение 19. $h(2) = \alpha$ (может с точностью до мультипликативной константы задать)

Лемма 2. $h(2^k) = k\alpha$

Лемма 3. $h(n) = \alpha \log_2 n$

Доказательство. $2^i \leqslant n^r < 2^{i+1} \quad r \in \mathbb{N}$

$h(i) \leqslant h\left(n^r\right) < \alpha(i+1)$

$\alpha \cdot i \leqslant r \cdot h(n) < \alpha(i+1)$

$\alpha \cdot \frac{i}{r} \leqslant h(n), \alpha\frac{i+1}{r}$

$i \leqslant r \log_2 n < i+1$

$\alpha\frac{i}{r} \leqslant \alpha \log_2 n < \alpha\frac{i+1}{r}$

$\forall r \quad |h(n) - \alpha \log_2 n| \leqslant \frac{\alpha}{r}$

$\implies h(n) = \alpha \log_2 n$ ■

Доказательство теоремы. Рациональный $p_i = \frac{a_i}{b} \quad m_i = a_i \quad r_{ij} = \frac{1}{a_i} \quad q_{ij} = \frac{1}{b} \quad q_{ij} = p_i r_{ij}$

$$H\left(\underbrace{\frac{1}{b}, \ldots, \frac{1}{b}}_{b}\right) = H(p_1, \ldots, p_k) + \sum_{i=1}^{k} p_i H\left(\underbrace{\frac{1}{a_i}, \ldots, \frac{1}{a_i}}_{a_i}\right)$$

$$\left(\sum_{i=1}^{k} p_i\right) h(b) = H(p_1, \ldots, p_k) + \sum_{i=1}^{k} p_i h(a_i)$$

$$H(p_1, \ldots, p_k) = \sum_{i=1}^{k} p_i \left(\alpha \log_2 b - \alpha \log_2 a_i\right)$$

Функция непрерывна и она верна для рациональных, следовательно она верна для всех ■

Замечание. $h(2) = \alpha$ – бит

А теперь мы хотим перевести определение информации на неслучайный источник

Ответ 1. Это тогда будет не совсем корректно с математической точки зрения. Когда смотришь на конкретные детерминированные данные. ■

Ответ 2. Изучение среднего не совсем антинаучное занятие. Внешне оно ведёт себя как случайные величины. ■

Пример. Есть строка s, в которой мы хотим померить информацию.

$s \in \Sigma^* \quad n = |\Sigma|$

$|s| = l \quad f_i$ – количество символов c_i в строке s

$p_i = \frac{f_i}{L}$

Допустим, что символы выдаёт случайный источник, который выдал символы $s_1, s_2, \ldots, s_L$. Статистически эта строка похожа на s. *натягивание на глобус* Допустим, что количество информации в строке s равно количеству в строке $\widetilde{s}$

$I(\widetilde{s}) = J \cdot H\left(p_1, p_2, \ldots, p_n\right) = -L \sum\limits_{i=1}^{n} p_i \log_2 p_i$

Вспомним арифметическое кодирование

$q = A(s)$ – длина арифместического кодирования

$A(s) \leqslant -\log_2(b_L - a_L) = -\log_2\left(p_{s_1} \cdot \ldots \cdot p_{s_L}\right) = -\log_2\left(\prod_{i=1}^{n} p_i^{f_i}\right) = -\sum\limits_{i=1}^{n} \underbrace{\frac{L}{p_i}}_{p_i} \log_2 p_i =$

$= I\left(\widetilde{s}\right) = L \cdot H\left(p_1, p_2, \ldots, p_n\right)$

Теорема 10. Длина кода, после арифметического кодирования не превышает энтропию Шеннона

Замечание. Арифметическое кодирование асимптотически оптимально среди тех, которые не учитывают взаимное расположение символов.

Пример (Нижняя оценка для сортировки)**.** Пусть $a_1, \ldots, a_n$ – перестановка и мы хотим её отсортировать

Утверждение 3. От одного сравнения мы получаем не больше 1 бита информации

Рассмотрим все перествки. В каждой содержится

$\log_2 n! = \sum\limits_{i=1}^{n} \log_2 i \geqslant \sum\limits_{i=\frac{n}{2}}^{n} \log_2 \frac{n}{2} = \Omega\left(n \log n\right)$

1.6 Цепи Маркова

$b = (b_1, b_2, \ldots, b_n)$ – b_i вероятность находиться в состоянии i

$C \quad c = (c_1, c_2, \ldots, c_n)$ – случайная величина после одного перехода

Матрица перехода p_{ij} – вероятность перейти из i в j

$$P = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$c_i = P(C = i) = \sum\limits_{j=1}^{n} P\left(x = i | B = j\right) P(B = j) = \sum\limits_{j=1}^{n} p_{ji} \cdot b_j$

$b^0 = (1, 0, 0, 0)$ – нулевой шаг

$b^1 = (0, \frac{1}{2}, \frac{1}{2}, 0)$

Рассмотрим судьбу м.ц. после поглощения. Жизнь происходит внутр одной сильно связанной компоненты – эргадического класса.

1. $d > 1$ длинна любого цикла кратна d. Циклический класс
2. НОД(длин всех циклов) = 1.

Теорема 11 (Эргадическая для регулярных цепей)**.** М.ц. такова, что $p_{ij} > 0 \forall i, j$

Тогда $\exists\, b \quad \forall\, b^0 \quad b^0 P^n \to b$

(b удолветворяет равенству $b = bP$

Доказательство. $(b^0 A)_i = \sum\limits_{j=1}^{n} b_j^0 \cdot A_{ji} = \left(\sum\limits_{j=1}^{n} b_j^0 \right) \widetilde{a}_i = \widetilde{a}_i$

$\sqsupset \forall j \quad a_{ji} = \widetilde{a}_i$

$P^n \to A$, которая удовлетворяет условию выше.

$m_i^t = \min_j (P^t)_{ji} \quad M_i^t = \min_j (P^t)_{ji}$

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

$\delta = \min_{i,j} \quad \delta > 0$

$$
\begin{aligned}
P_{ji}^{t+1} &= \sum_{k=1}^{n} P_{jk}^t P_{ki} \\
&\leqslant \overbrace{\sum_{\substack{k=1 \\ k \neq posMin}}^{n} P_{jk}}^{1} M_i^t + P_{j\ posMin}(m_i^t - M_i^t) \\
&\leqslant M_i^t + \delta \left(m_i^t - M_i^t \right).
\end{aligned}
$$

Аналогично с максимумом, оцениваем всё снизу минимумов, кроме максимума

$$M_i^{t+1} \leqslant M_i^t + \delta \left(m_i^t - M_i^t \right).$$

$$-m_i^{t+1} \leqslant -m_i^t + \delta \left(m_i^t - M_i^t \right).$$

$$M_i^{t+1} - m_i^{t+1} \leqslant \left(M_i^t - m_i^t \right) (1 - 2\delta) \leqslant (1 - 2\delta)^{t+1} \to 0.$$

Теперь у $b = bP$

$(I - P)b = 0$

$Rg\,(I - P) = n - 1$

$\sum b_i = 1$

P^{2^c}

$bP^n 0 > b \quad bP^{n+1} \to bP$

$b = bP$ ■

Вернёмся к вопросу что происходит после поглощения.

$\lessdot$ эргадический класс $A \quad \widetilde{p} = \sum\limits_{a \in A} (b^o NR)_a$

$\widetilde{b}^0 = \left(b^0 NR\right)_{A - \frac{1}{p}}$

$\exists$ предельное $b: \quad \widetilde{b}^0 A^n \to b$

Конечное распределение $b\widetilde{p}$

Скрытые Марковские модели. Мы решали до этого прямую задачу – брали м.ц. с извстными матрицами перехода и смотрели на их характеристики.

Есть обратная: Есть состояние и мы хотим узнать матрицу перехода.

Ещё задача: Есть немарковский процесс и мы хотим аппроксимировать его марковским.

1.7 Формальные языки

Алфавит – Σ, конечное непустое множество

слово, цепочка, строка $\Sigma^* = \bigcup\limits_{k=0}^{\infty} \Sigma^k$

Формальный язык $L \subseteq \Sigma^*$

<TODO>

Определение 20. Описание языка – слово конечной длины

Всего “описаний” счётное множество

- Распознавание – по слову возвращаем булевский флаг – есть слово в нашем языке или нет
- Порождение – описывает как породить возможно бесконечное количество слов

Определение 21. Перечиление слов. $\{01, 011, 10, 1010\}$. Но так можно описать только конечные языки (содержащие конечное количество слов

Пример. Правильные скобочные последовательности ε – псп

A, B – псп $\implies AB$ – псп

A – псп $\implies (A)$ – псп

Это порождение. Можно описать распознованием: баланс в любой момент неотрицательный, баланс в конце $= 0$

Между этими способми есть логический переход:

- Распозновать $\to$ Порождать. порождаем всё, что можем распознать
- Обратно: распознаём всё, что в какой-то момент порождаем

Пример. C++ без ограничений по памяти. Пограмма P

$L = \{\omega \mid p(\omega) = 1\}$

1.7.1 Регулярные = Автоматные языки

Определение 22. Конкатенация: $\alpha \in \Sigma^k, \beta \in \Sigma^l \quad \alpha\beta \in \Sigma^{k+l}$

$\gamma = \alpha\beta \quad \gamma_i = \begin{cases} i \leqslant k & \implies \alpha_i \\ i > k & \implies \beta_{i-k} \end{cases}$

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

Пример. $AB = \{x \mid x = yz, y \in A, z \in B\}$

$A = \{0, 01\} \quad B = \{0, 10\}$

$AB = \{00, 010, 0110\}$

Базовые операции:

1. Объединение $A \cup B$
2. Конкатенация AB

   Возведение в степень $A^k = \underbrace{AA\dots A}_{k} \quad A^0 = \{\varepsilon\}$
3. Замыкание Клини $A^* = \bigcup\limits_{k=0}^{\infty} A^k$

Определение 23. $Reg_0 = \{\emptyset, \{\varepsilon\}, \{c\} \, \forall c \; \in \; C\}$

$Reg_{i+1} = Reg_i \cup \{A \cup B, AB, A^* \mid A, B \in Reg_i\}$

$Reg_1 = \{\emptyset, \varepsilon, a, b, \dots, \{a, b\}, \{a, \varepsilon\}, \dots, ab, aa, \dots, \{\varepsilon, a, aa, aaa, \dots\}, \dots, \{\varepsilon, b, bb, bbb, \dots\}\}$

$Reg = \bigcup\limits_{k=0}^{\infty} Reg_k$

Лемма 4. $A, B \in Reg$:

1. $A \cup B \in Reg$
2. $AB \in Reg$
3. $A^* \in Reg$

$A \in Reg_i, B \in Reg_j$

Они все принадлежат $Reg_{\max\{i,j\}+1}$

Определение 24. Назовём семейство языков $X \in Good \quad X \subseteq 2^{\Sigma^*}$

$X = set \langle lang \rangle$

Good: set<set<lang> >

1. $Reg_0 \in X$
2. X замкнуто относительно $A \cup B, AB, A^*$, а.и.

$$A, B \in X \implies AB, A \cup B, A^* \in X \qquad A, B : lang.$$

Теорема 12. $Reg = \bigcap\limits_{u \in Good} U$

Доказательство. to be written ■

Определение 25 (Описание).

- $\emptyset$ ε c
- A α B β:
  - AB $\alpha\beta$ – средний приоритет
  - $A \cup B$ $\alpha|\beta$ – минимальный приоритет
  - A^* α^* – максимальный приоритет

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

$(0 \mid 11)^*$ – язык в кортором единицы идут парами

Такие описания называют академическими регулярными выражениями

$\alpha^+ = \alpha\alpha^*$

$\alpha^k = \underbrace{\alpha\alpha\ldots\alpha}_{k}$

Пример. $0^*| \left(0^*10^*10^*\right)^*$ – язык, содержащий чётное число единиц

Пример. Проверка чётное ли число единиц

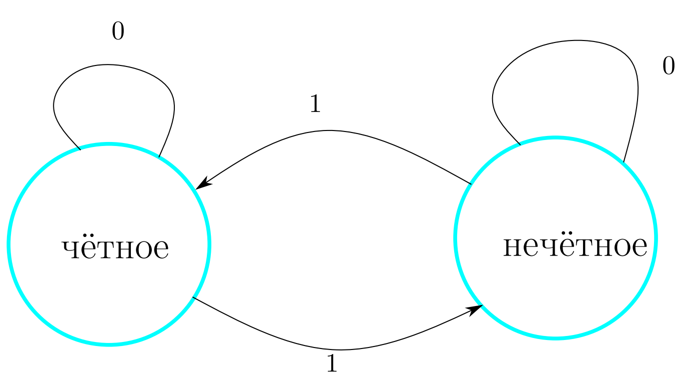


Рис. 1.5: check-one

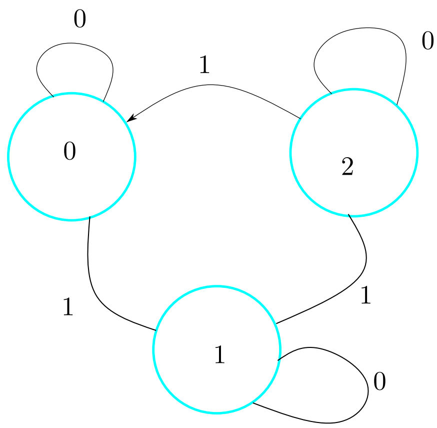


Рис. 1.6: check-div3

Определение 26. Детерминированный Конечный Автомат ДКА DFA

$$A = \langle \Sigma, Q, S \in \Sigma, T \subseteq Q, \delta : Q \times \Sigma \to Q \rangle .$$

- Σ – алфавит
- Q – конечное множество состояний
- S – начальное состояние
- T – допускающие состояния
- δ – функция переходов

$Snap = Q \times \Sigma^*$

Пеехoд:

1. $\alpha = c\beta \quad c \in \Sigma$
2. $r = \delta(q, c)$

Пример. $\langle e, 0101 \rangle \vdash \langle e, 101 \rangle \vdash \langle o, 01 \rangle \vdash \langle o, 1 \rangle \vdash \langle e, \varepsilon \rangle$

$\mathscr{L}(A) = \{\omega \mid \langle s, \omega \rangle \vdash^* \langle t, \varepsilon \rangle , t \in T\}$

Теорема 13 (Клини)**.** $Reg = Aut$

$Aut = \{X \mid \exists \text{ ДКА } A : \quad X = L(A)\}$

1.8 Недетерминированный конечный автомат

x – допускаюется Недетерминированным Конечным Автоматом $\iff$ $\exists$ последовательность переходов по символам x, заканчиващееся в допускающем состоянии

L – формальный язык. $L \times \Sigma^*$

Артур: $x \mapsto x \in L$?

Мерлин: Убедить Арутра, что $x \in L$

Замечание. Артуру в случае неопределённости выгодно слушать Мерлина.

Если $x \notin L$, то Мерлин не сможет испортить своими советами, потому что в автомате просто нет такой последовательности, на которую можно направить, чтобы попасть в допускающее

Если $x \in L$, то, внезпано, интересы Артура и Мерлина совпадают

Замечание (интерпретация через миры)**.** На каждом шаге, где недетерминирован следующий шаг, создаётся два мира, на каждый из шагов. Если хотя бы в одном дошли до допускающего, то слово принадлежит.

Определение 27 (НКА)**.** $\left(\Sigma, Q, S \subseteq Q, T \subseteq Q, \delta : Q \times E \to 2^Q\right)$

Стартовых состояний может быть несколько, хотя почти никогда не нужно

Состояние – $\langle q, x\rangle \quad q \in Q, \quad x \in X^*$

$\langle q, x\rangle \vdash \langle r, y\rangle$

1. $x = cy, \quad c \in \Sigma$
2. $r \in \delta\left(q, c\right)$

x – допускается A, если $\langle s, x\rangle \vdash^* \langle t, \varepsilon\rangle\,, \quad t \in T$

Пример.

```
class DFA {
    // 0 .. n-1 -- Q; 0 .. c-1 -- Sigma
    s : int
    t :  vector<bool>(n)
    delta: vector<vector<int>> (n,c)

    bool accept(x) {
        cur = s
        for (i = 0 .. len(x) - 1)
```

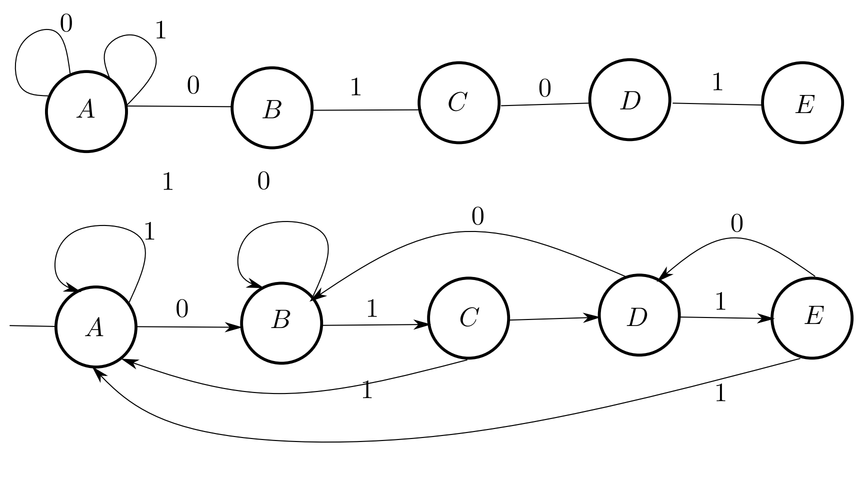


Рис. 1.7: auto

```
            cur = delta[cur][x[i]
        return t[cur]
    }
}
```

$O(len(x))$

1.9 Динамическое программирование

```
class NFA {
    // 0 ... n-1 -- Q; 0 .. c-1 -- Sigma
    s : int
    t : vector<bool>(n)
    delta : vector<vector<set<int>>>(int)

    can[i][q] -- можно ли прочитав i символов x'а оказаться в состоянии q

    bool accept(x):
        can[0][s] = true
        for (i = 0 .. len(x) - 1)
            for (q = 0 .. n-1)
                if (can[i][q])
                   for (r : delta[q][x[i]])
                    can[i+1][r] = true
        for (q = 0 .. n-1)
            if (can[len(x)][q] &&  t[q])
                return true
```

```
}
```

$O\left(len(x)\cdot(n^2+m)\right)$

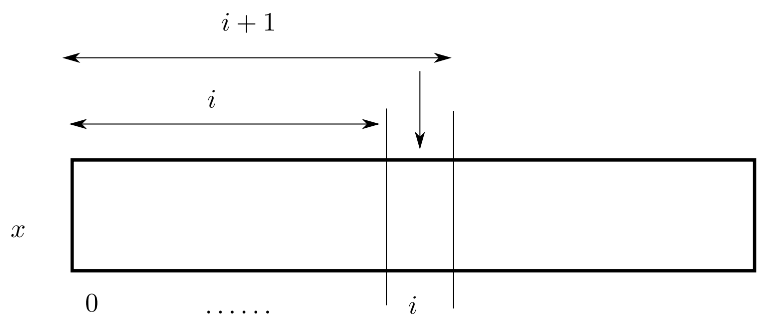


Рис. 1.8: nfa-step

Утверждение 4. Для L $\exists$ НКА A_n $\iff$ для L $\exists$ ДКА A_D

Доказательство. $\forall$ ДКА является частным случаем НКА

$\Longleftarrow$ очевидно

$\Longrightarrow$ Алгоритм Томпсона

```
next(a : vector<bool>, c) vector<bool>
    res = vector<bool>(n)
    for (q = 0 .. n-1)
        if a[q]
            for r : delta[q][c]
                res[r] = true
    return res

bool accept(x)
    can[0][s] = true
    for (i = 0 .. len(x) - 1)
        can[i+1] = next(can[i], x[i])
```

$(\Sigma, Q_D = 2^{Q_N}, \{s\}, T_D = \{A \mid A\cap T_n \neq \emptyset\}, \delta_D(A,c) = \{r \mid \exists q \in A,\ r \in \delta_N(q,c)\}$

∎

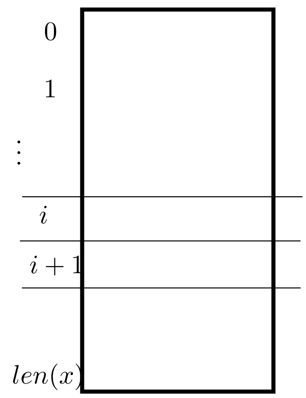


Рис. 1.9: tomps

Замечание. Казалось бы, вот Детерминированный автоматы такие хорошие, зачем нужны другие? Но он имеет экспоненциальное количество состояний (верхняя оценка) по сравнению с Недерминированным

Но в реальности можно улучшать, убирая состояния, которые недостижимы (например $\{B, C\}$ может не встречаться одновременно никогда) (в недерминированном)

Иначе можно начать со стартовых состояний и делать очередь всех состояний, в которых мы можем быть. Именно такую конструкцию обычно и называют Алгоритмом Томпсона.

Конструкция описанная выше называется Конструкцией подмонжеств.

1.10 ε-НКА

Разрешим на переходе писать не символ, а ε. Переходя по нему строка на входе не меняется

Пример. $\left((0|1)^* 00| (0|1)^* 11\right) (0|1)^*$

$0^*1^*2^*$

Утверждение 5. $\forall \varepsilon$-НКА $\exists$ эквивалентный НКА без ε переходов

Доказательство.

1. Рассмотрим граф ε-переходов. Этому графу мы сделаем транзитивное замыкание. И добавим новые рёбра как ε-переходы в наш граф

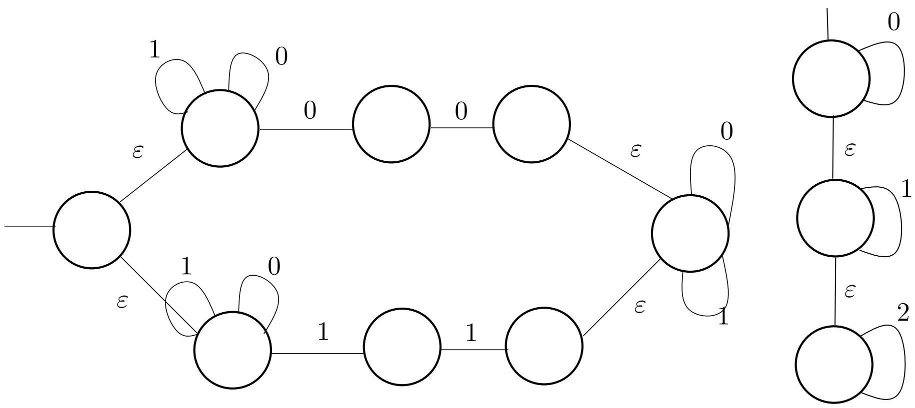


Рис. 1.10: epsauto

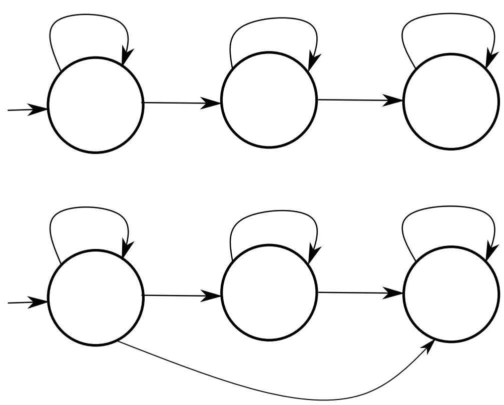


Рис. 1.11: epsgraph

Язык не поменялся. Вместо прохождения по новому переходу, можно делать n эпсилон-переходов в старом графе

2. Для каждой конструкции: Из p есть ε-переход в q терминальный, сделаем p тоже терминальным

> **Утверждение 6.** Если x Допускалось раньше, $\iff$ допускается и сейчас. Последний переход не ε

3. Рассмотрим тройки вершин, что Из p есть ε переход в q откуда переход по c в r. Тогда добавим ребро из p в r по c

Утверждение 7. Если слово можно допустить, то его можно допустить вообще не делая ε-переход

4. удалим все ε-переходы

■

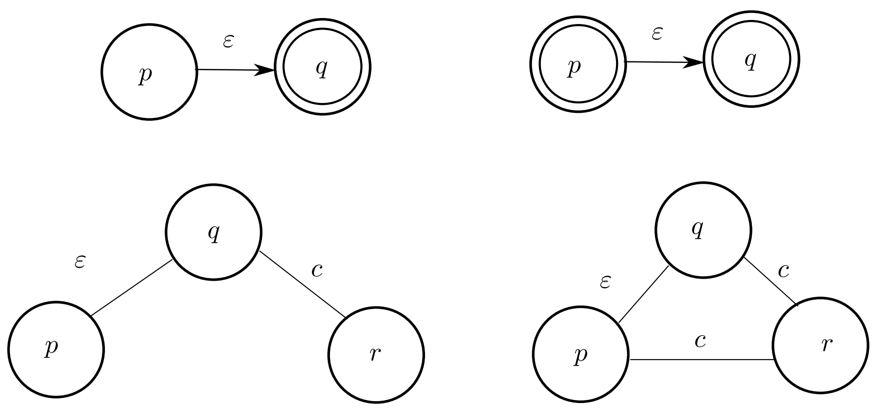


Рис. 1.12: proof-pic

Утверждение 8. Для любого Языка следующие три утверждения эквивалентны:

1. Можно построить ДКА
2. Можно построить НКА
3. Можно построить ε-НКА

$2 \implies 1$ – Томпсон

$3 \implies 1$ – ε-замыкание

Теорема 14 (Клини)**.** $\mathrm{Reg} = \mathrm{Aut}$

Доказательство.

$\mathrm{Reg} \subseteq \mathrm{Aut}$ Докажем по индекции, что $\forall i \quad \mathrm{Reg}_i \subseteq Aut$

Будем строить ε-НКА с одним терминальным состоянием

База: $\emptyset$. Стратовое и терминальное состояние и никаких переходов

ε – одна ε-стрелка

c – одна c-стрелка

Переход: $\text{Reg}_i \subseteq Aut \implies \text{Reg}_{i+1} \subseteq Aut$

- $L = A \cup B$
- $L = AB$
- $L = A^*$

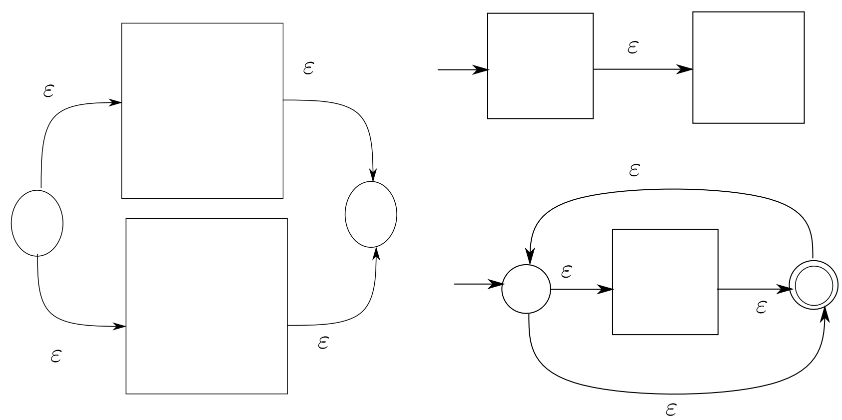


Рис. 1.13: move-klini

■