Теория автоматов

Лекция 1: Дискретная математика

Дьулустан Никифоров

Кафедра ИТ
Северо-Восточный Федеральный Университет

Осень 2024

Введение

- Как и ОВС, этот предмет является фундаментальным для специалистов в области Computer Science.
- ОВС — КАК работает компьютер?
  Теория автоматов и формальных языков — ЧТО делает компьютер?
- По-хорошему, наш предмет называется — Theory of Computation.
  В более широком контексте: Theoretical Computer Science.
- В этом курсе: компьютер — это абстрактная “машина”. Мы будем изучать, что умеют делать такие абстрактные машины.

Введение

- Theory of computation — мы математически точно определим, что такое computation, что такое алгоритм.

Темы

- Regular expressions (регулярные выражения).
- Deterministic finite automata (детерминированные конечные автоматы)
- Nondeterministic finite automata (недетерминированные конечные автоматы)
- Regular grammars and languages (регулярные грамматики и языки)
- Context-free grammars (контекстно-свободные грамматики)
- Pushdown automata (магазинные автоматы)
- Turing machine (машина Тьюринга)

Рекомендованная литература

- Michael Sipser. Introduction to the theory of computation (1997) $\leftarrow$ **главная книга курса**.
- Hopcroft, Motwani, Ullman. Introduction to Automata Theory, Languages, and Computation (2006).
- Пентус, Пентус. Теория формальных языков (2003).

Дискретная математика: Множества

Перед тем, как начнем с теорией формальных языков, убедимся, что у нас есть необходимые инструменты:

- **множество (set)** — набор элементов, элементы уникальны и не упорядочены.
  Примеры: $A = \{2, -\frac{\pi}{2}, 10101010, (1,2,3)\}, \emptyset = \{\}, \mathbb{N} = \{1, 2, 3, \dots\}, \mathbb{Z} = \{dots, -2, -1, 0, 1, 2, \dots\}, \mathbb{Q}$(множество рациональных чисел), $\mathbb{R}$ (множество действительных чисел).
- $x \in A$ — элемент x **принадлежит** множеству A (x belongs to A), A **содержит** x (A contains x).
  $y \notin A$ — элемент y **не принадлежит** множеству A (y does not belong to A), A **не содержит** x (A does not contain x).
- $A \subseteq B$ — A **подмножество (subset of)** B,
  $A \subsetneq B$ — A **строгое подмножество (proper subset of)** B.

Дискретная математика: Множества

Вот некоторые популярные множества, которые следует знать:

- $\mathbb{N} = \{1, 2, 3, ...\}$ — множество **натуральных чисел (natural numbers)**.
- $\mathbb{Z} = \{... -3, -2, -1, 0, 1, 2, 3, ...\}$ — множество **целых чисел (integers)**.
- $\mathbb{Q} = \{\frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}\}$ — множество **рациональных чисел (rational numbers)**.
- $\mathbb{R}$ — множество **действительных (вещественных) чисел (real numbers)**.
- $\mathbb{P} = \{2, 3, 5, 7, 11, ...\}$ — множество **простых чисел (prime numbers)**.

Дискретная математика: Множества

Перед тем, как начнем с теорией формальных языков, убедимся, что у нас есть необходимые инструменты:

- $A \cup B$ — **объединение (union)** ,
  $A \cap B$ — **пересечение (intersection)** ,
  $A \setminus B$ — **разность (difference)** .
- Определим математически строго:
  $A \cup B = \{x \mid x \in A$ **или** $x \in B\}$,
  $A \cap B = \{x \mid x \in A$ **и** $x \in B\}$,
  $A \setminus B = \{x \mid x \in A$ **и** $x \notin B\}$,
- Примеры:
  $A = \{2, 3, 5, 7, 11\}, B = \{2, 4, 8\}$

Дискретная математика: Множества

- $\{1,5,7\} = \{5,1,7\} = \{7,1,5\}$ — множества *неупорядочены*.
- $\{2,4,14,2,2\} = \{2,4,14\}$ — элементы во множестве *уникальны*.
- $A = \{0,2,4,6,8,10\}$, $B = \{2,4,8,16,32\}$, $C = \{2,3,5,7\}$. Тогда:

Дискретная математика: Множества

- $\{1,5,7\} = \{5,1,7\} = \{7,1,5\}$ — множества *неупорядочены*.
- $\{2,4,14,2,2\} = \{2,4,14\}$ — элементы во множестве *уникальны*.
- $A = \{0,2,4,6,8,10\}$, $B = \{2,4,8,16,32\}$, $C = \{2,3,5,7\}$.
  Тогда:
  $A \cap B = \{2,4,8\}$
  $A \cup B = \{0,2,4,6,7,10,16,32\}$
  $B \cap C = \{2\}$
  $B \cup C = \{2,3,4,5,7,8,16,32\}$
  $A \setminus B = \{0,6,10\}$
  $B \setminus A = \{16,32\}$
  $A \cap B \cap C = \{2\}$
  $A \cup B \cup C = \{0,2,3,4,5,6,7,8,10,16,32\}$

Пусть A — множество всех русских имен, начинающихся на ‘К’. Тогда можем сделать такие утверждения:

Пусть A — множество всех русских имен, начинающихся на 'К'. Тогда можем сделать такие утверждения:

- 'Коля' $\in A$;
- 'Карл', 'Кира' $\in A$;
- 'Дима' $\notin A$;
- { 'Коля', 'Кира', 'Карл' } $\subseteq A$;
- { 'Коля', 'Кира', 'Карл' } $\subsetneq A$;
- Пусть B — множество всех русских имен, заканчивающихся на 'а'.

Дискретная математика: Множества

Пусть A — множество всех русских имен, начинающихся на ‘К’.
Тогда можем сделать такие утверждения:

- ‘Коля’ $\in A$;
- ‘Карл’, ‘Кира’ $\in A$;
- ‘Дима’ $\notin A$;
- { ‘Коля’, ‘Кира’, ‘Карл’ } $\subseteq A$;
- { ‘Коля’, ‘Кира’, ‘Карл’ } $\subsetneq A$;
- Пусть B — множество всех русских имен, заканчивающихся на ‘а’.
  Тогда: ‘Коля’ $\notin B$, ‘Коля’ $\notin A \cap B$, ‘Коля’ $\in A \cup B$.
  ‘Кира’ $\in A \cap B$.

Дискретная математика: Множества

Definition

- $2^A = \bigcup\limits_{B \subseteq A}$ — **power set** of A.
  Power set of A — это множество всех подмножеств A.
- **sequence (последовательность)** — упорядоченный набор элементов: $(a_1, a_2, \dots, a_k, \dots)$.
- **tuple (кортеж)** — конечная последовательность (finite sequence).
  $(a_1, a_2, \dots, a_k)$ — **k-tuple**.
- **pair (пара)** — 2-tuple.

Definition

Декартово произведение (Cartesian product) двух множеств A и B:

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}.$$

Пусть $A = \{0, 5, 9\}, \quad B = \{1, 2, 6, 9\}$.

Тогда:

- $2^A = ?$
- $2^B = ?$
- $A \times B = ?$
- $B \times B = ?$

Пусть $A = \{0, 5, 9\}, \quad B = \{1, 2, 6, 9\}$.

Тогда:

- $2^A = \{\emptyset, \{0\}, \{5\}, \{9\}, \{0, 5\}, \{0, 9\}, \{5, 9\}, \{0, 5, 9\}\}$
- $2^B =$ $\{\emptyset, \{1\}, \{2\}, \{6\}, \{9\}, \{1, 2\}, \{1, 6\}, \{1, 9\}, \{2, 6\}, \{2, 9\}, \{6, 9\},$ $\{2, 6, 9\}, \{1, 6, 9\}, \{1, 2, 9\}, \{1, 2, 6\}, \{1, 2, 6, 9\}\}$
- $A \times B = \{(0, 1), (0, 2), (0, 6), (0, 9), (5, 1), (5, 2), (5, 6), (5, 9),$ $(9, 1), (9, 2), (9, 6), (9, 9)\}$
- $B \times B = \{(1, 1), (1, 2), (1, 6), (1, 9), (2, 1), (2, 2), (2, 6), (2, 9),$ $(6, 1), (6, 2), (6, 6), (6, 9), (9, 1), (9, 2), (9, 6), (9, 9)\}$

$A = \{1, 2\}, \quad B = \{3, 4\}$

Можно Декартово умножать несколько раз:
$A \times A \times B =$
$\{(1,1,3),(1,2,3),(2,1,3),(2,2,3),(1,1,4),(1,2,4),(2,1,4),(2,2,4)\}$
Заметьте, что при этом принято делать не пары пар (типа $((1,1),3))$ как можно было бы ожидать, а сразу составлять тройки.

Также, **возведение множества в степень** означает, что она Декартово умножается на себя:

$$A^n = \underbrace{A \times A \times \ldots \times A}_{n \text{ раз}}$$

Дискретная математика: Бинарное отношение

Бинарное отношение (binary relation) R на множестве A — это подмножество $A \times A$.

- R — бинарное отношение на A.
  Мы пишем $aRb = true$, чтобы обозначить, что $(a, b) \in R$.
  Обычно просто пишем aRb.
- Какие бинарные отношения мы часто видим в жизни?

Дискретная математика: Бинарное отношение

Бинарное отношение (binary relation) R на множестве A — это подмножество $A \times A$.

- R — бинарное отношение на A.
  Мы пишем $aRb = true$, чтобы обозначить, что $(a, b) \in R$.
  Обычно просто пишем aRb.
- Какие бинарные отношения мы часто видим в жизни?
  Например, $<, >, \leq, \approx, \subset, \rightarrow$.
- Бинарное отношение R:
  - **рефлексивное (reflexive)** если для всех x: xRx.
  - **симметричное (symmetric)** если для всех x, y: $xRy \Leftrightarrow yRx$.
  - **транзитивное (transitive)** если для всех x, y, z: xRy and $yRz \Rightarrow xRz$.
- Если все три свойства выполняются, то R называется **отношение эквивалентности (equivalence relation)**.

Definition

Бинарное отношение R, определенное на множестве S называется:

- **рефлексивным (reflexive)**, если для всех $x \in S$ выполняется xRx.
- **симметричным (symmetric)**, если для всех $x, y \in S$ выполняется $xRy \Leftrightarrow yRx$ (т.е. xRy тогда и только тогда, когда yRx).
- **транзитивным (transitive)**, если для всех x, y, z выполняется $xRy, yRz \Rightarrow xRz$ (т.е. если xRy и yRz, то должно быть xRz).

Рефлексивное, симметричное и транзитивное бинарное отношение назвается **отношением эквивалентности (equivalence relation)**.

Пример 1: Определим бинарное отношение R на множестве $S = \{-3, -2, -1, 0, 1, 2, 3\}$ — xRy, если $|x| > |y|$ (где $|x|$ означает модуль числа x).

Можно полностью расписать это бинарное отношение:

$$\begin{aligned} R = \{ & (-3,-2), (-3,-1), (-3,0), (-3,1), (-3,2) \\ & (-2,-1), (-2,0), (-2,1), (-1,0), \\ & (3,-2), (3,-1), (3,0), (3,1), (3,2), \\ & (2,-1), (2,0), (2,1), (1,0)\} \end{aligned}$$

Можно показать, что R *нерефлексивное, несимметричное, транзитивное*.

Пример 2: Определим бинарное отношение $\equiv_m$ на множестве целых чисел $\mathbb{N}$ — отношение сравнимости по модулю m.
$a \equiv_m b$ означает что целое число a имеет тот же остаток при делении на m, что и b (на прогерском: $a\%m == b\%m$).

Можно показать, что $\equiv_m$ *рефлексивное, симметричное и транзитивное* $\Rightarrow$ $\equiv_m$ является *отношением эквивалентности*.

- **мощность (cardinality)** множества A — типа, "количество" элементов в массиве. Обозначается $|A|$.
  $A = \{1, 2, 3\} \Rightarrow |A| = 3$,
  $|\emptyset| = 0$.
  Бесконечное множество $\Rightarrow$ бесконечная мощность, но бесконечность бесконечности рознь.
- Множества, которые имеют одинаковую мощность со множеством $\mathbb{N}$, называются **счётно бесконечными (countably infinite)**.
- Что насчет множества всех чётных натуральных чисел? множества всех целых чисел? множества всех простых чисел?

Дискретная математика: Бесконечность

- **мощность (cardinality)** множества A — типа, "количество" элементов в массиве. Обозначается $|A|$.
  $A = \{1, 2, 3\} \Rightarrow |A| = 3$,
  $|\emptyset| = 0$.
  Бесконечное множество $\Rightarrow$ бесконечная мощность, но бесконечность бесконечности рознь.
- Множества, которые имеют одинаковую мощность со множеством $\mathbb{N}$, называются **счётно бесконечными (countably infinite)**.
- Что насчет множества всех чётных натуральных чисел? множества всех целых чисел? множества всех простых чисел?
  они все счётно бесконечные.
- Важный факт жизни: $\mathbb{Q}$ — бесконечно счётное.

Пример: докажем, что множество целых чисел $\mathbb{Z}$ — счётно бесконечное.

Для этого надо провести взаимно-однозначное соответствие (биекцию) между $\mathbb{Z}$ и $\mathbb{N}$. Простыми словами, надо занумеровать $1, 2, 3, ...,$ все целые числа!

Один способ сделать это:
$1 \leftrightarrow 0$
$2 \leftrightarrow -1$
$3 \leftrightarrow 1$
$4 \leftrightarrow -2$
$5 \leftrightarrow 2$
$\ldots$

Дискретная математика: Бесконечность

Theorem

Множество действительных чисел $\mathbb{R}$ несчётно бесконечно (uncountably infinite).

Идея доказательства:

- Допустим от противного, что мы смогли пронумеровать все действительные числа.

$$
\begin{array}{l}
0 \mapsto 0.\ a_{0,0}\ a_{0,1}\ a_{0,2}\ a_{0,3}\ a_{0,4} \dots \\
1 \mapsto 0.\ a_{1,0}\ a_{1,1}\ a_{1,2}\ a_{1,3}\ a_{1,4} \dots \\
2 \mapsto 0.\ a_{2,0}\ a_{2,1}\ a_{2,2}\ a_{2,3}\ a_{2,4} \dots \\
3 \mapsto 0.\ a_{3,0}\ a_{3,1}\ a_{3,2}\ a_{3,3}\ a_{3,4} \dots \\
4 \mapsto 0.\ a_{4,0}\ a_{4,1}\ a_{4,2}\ a_{4,3}\ a_{4,4} \dots \\
\vdots \qquad\quad \vdots \qquad\qquad\qquad\quad \vdots
\end{array}
$$

- Возьмем число, собранное из диагональных элементов этой таблицы, и изменим каждый диагональный элемент, назовем это число x.

Идея доказательства:

- Тогда x не может совпадать ни с каким числом в таблице: с любым числом у x точно есть несовпадающая цифра (тот самый диагональный элемент).
- Значит числа x нету в таблице $\Rightarrow$ противоречие с тем, что все действительные числа были пронумерованы!
- Теорема доказана.

Это знаменитый **метод диагонализации** Кантора.

Theorem (Теорема Кантора)

Если множество A имеет бесконечную мощность, то 2^A имеет большую мощность, чем A.

Идея доказательства:

- Допустим от противного, что A имеет такую же мощность, что и 2^A.
- Тогда можно сделать полное соответствие между A и 2^A: для каждого $a \in A$ будет свой $S_a \in 2^A$.
- Построим такое множество X: для каждого $a \in S$,
  - если $a \in S_a$, то не включаем a в X;
  - если $a \notin S_a$, то включаем a в X;
- Тогда полученное множество X не совпадает ни с каким множеством из 2^A (это следует из нашего построения) $\Rightarrow$ противоречие
- Теорема доказана.

Дискретная математика: Бесконечность

- Возьмем $\mathbb{N}$ — самая “маленькая” бесконечность (счётная). Обозначим её $\mathcal{N}_0$.
- Обозначим $\mathcal{N}_1 = 2^N$. Тогда по теореме Кантора, $|\mathcal{N}_1| > |\mathcal{N}_0|$.
- Можно доказать, что $|\mathbb{R}| = |\mathcal{N}_1|$.
- Можем построить бесконечную цепочку:
  $|\mathcal{N}_0| < |\mathcal{N}_1| < |\mathcal{N}_2| < |\mathcal{N}_3| < \dots$,
  где $\mathcal{N}_i = 2^{\mathcal{N}_{i-1}}$.
- *Важная мысль*:
  Одна бесконечность может быть бесконечнее другой бесконечности, однако нет предела как бесконечна может быть бесконечность.

- Мы поняли, что множества $\mathbb{R}$ и $2^{\mathbb{N}}$ более мощные, чем $\mathbb{N}$: $|\mathbb{N}| = |\mathbb{Q}| < |\mathbb{R}| = |2^{\mathbb{N}}|$.
- Какой вопрос возникает?

Дискретная математика: Бесконечность

- Мы поняли, что множества $\mathbb{R}$ и $2^{\mathbb{N}}$ более мощные, чем $\mathbb{N}$: $|\mathbb{N}| = |\mathbb{Q}| < |\mathbb{R}| = |2^{\mathbb{N}}|$.
- Какой вопрос возникает?
  А есть ли бесконечное множество лежащее между $\mathbb{N}$ и $\mathbb{R}$ (т.е. множество A: $|\mathbb{N}| < |A| < |\mathbb{R}|$)?

Дискретная математика: Бесконечность

- Мы поняли, что множества $\mathbb{R}$ и $2^{\mathbb{N}}$ более мощные, чем $\mathbb{N}$: $|\mathbb{N}| = |\mathbb{Q}| < |\mathbb{R}| = |2^{\mathbb{N}}|$.
- Какой вопрос возникает?
  А есть ли бесконечное множество лежащее между $\mathbb{N}$ и $\mathbb{R}$ (т.е. множество A: $|\mathbb{N}| < |A| < |\mathbb{R}|$)?
- Ответ: было доказано, что нельзя ни найти такое множество, ни доказать, что её не существует!
- *Важная мысль*: WTF?