Howard Straubing «Finite Automata, Formal Logic, and Circuit Complexity»

Лекция 1(01.03.2012).

1. Логика для слов.

$\Sigma = \{a, b, ...\}$;
$|\Sigma| < +\infty$;
x, y, z позиции в слове;
$x < y$ позиция x стоит перед позицией y;
$Q_a x$ предикат в позиции x стоит буква a.

Пример. $\exists x(Q_a x \ \& \ \forall y \ (x \leq y))$ определяет язык $a\Sigma^*$.

$(a^*)^2$ записать нельзя;
X, Y, Z — множества позиций;

Пример. $\forall x \ Q_a x \ \& \ \exists X \ \forall x \ (x \in X \ \longleftrightarrow x + 1 \notin X) \ \& \ (x \notin X \leftrightarrow x + 1 \in X)$
$\& \ \forall x(\forall y \ (x \leq y) \rightarrow x \in X) \ \& \ \forall x(\forall y \ (y \leq x) \rightarrow x \notin X)$ описывает язык $(a^2)^*$.

Логика 1-го порядка. Переменные $x, y, z, ...$
числовые предикаты ($<$, следует за, делит и т.п.);
специальные предикаты для каждой буквы алфавита (Q_a, Q_b,...);
логические связки ($\&, \vee, \neg, \rightarrow, \leftrightarrow$);
кванторы($\forall, \exists$);
формулы, свободные и связанные переменные.

Если P n-местный предикат, то выражение $P(x_1, x_2, ..., x_n)$ — это формула (атомарная), в которой все переменные свободны.

Если Φ и Ψ формулы, в которых переменные $x_1, ..., x_n$ свободны, а переменные $y_1, ..., y_n$ связаны, то $\Phi\&\Psi, \Phi \vee \Psi, \Phi \rightarrow \Psi, \Phi \leftrightarrow \Psi, \neg\Phi$ — формулы в которых$x_1, ..., x_n$ — свободные, $y_1, ..., y_n$ — связанные.

Если Φ — формула, в которой x - свободная переменная, то $\forall x(\Phi)$ и $\exists x(\Phi)$ — формулы, в которых x — связанная, а остальные переменные свободны, если они были свободны вΦ и связаны, если они были связаны в Φ.

Логика 2-го порядка. Переменные $x, y, z, ..., X, Y, Z, \ldots$
числовые предикаты ($<$, следует за, делит и т.п.);
специальные предикаты для каждой буквы алфавита (Q_a, Q_b,...);
предикат ($\in$);
логические связки ($\&, \vee, \neg, \rightarrow, \leftrightarrow$);
кванторы($\forall, \exists$);
формулы, свободные и связанные переменные.

Если P n-местный предикат, то выражение $P(x_1, x_2, ..., x_n)$ —это формула (атомарная), в которой все переменные свободны, кроме того, атомарная формула может иметь вид $x \in Y$.

Если Φ и Ψ формулы, в которых переменные $x_1, ..., x_n, X_1, \ldots X_m$ свободны, а переменные $y_1, ..., y_n, Y_1, \ldots Y_m$ связаны, то $\Phi\&\Psi, \Phi \vee \Psi, \Phi \rightarrow \Psi, \Phi \leftrightarrow \Psi, \neg\Phi$ — формулы, в которых$x_1, ..., x_n, X_1, \ldots X_m$ — свободные, $y_1, ..., y_n, Y_1, \ldots Y_m$ — связанные.

Если Φ — формула, в которой X — свободная монадическая переменная, то $\forall X(\Phi)$ и $\exists X(\Phi)$ — формулы, в которых X — связанная, а остальные переменные свободны, если они были свободны в Φ и связаны, если они были связаны в Φ.

2. Теорема Бюхи.

Теорема (Бюхи(1960)). *Язык $L \subseteq \sum^*$ распознается конечным недетерминированным автоматом тогда и только тогда, когда он задается некоторой замкнутой формулой монадической теории 2-го порядка с отношением $=$ и единственным числовым предикатом «следует за».*

Доказательство. Необходимость. Пусть $\mathcal{A}$ — конечный автомат, который распознает L:

Q — $\{q_0, q_1, ..., q_{k-1}\}$;

q_0 — начальное состояние;

$F \subseteq Q$ — множество заключительных состояний;

$E \subseteq Q \times \Sigma \times Q$ — множество переходов (q, a, q').

$w \in L \iff$ существует путь из q_0 в F такой, что метки ребер из этого пути составляют слово w.

$w \in L \iff$ существуют множества $X_0, X_1, ..., X_{k-1} \subseteq \{1, 2, ..., |w|\}$, такие что выполняются условия:

(1) $\bigcup\limits_{i=0}^{k-1} X_i = \{1, 2, ..., |w|\}$;

(2) $i \neq j \Rightarrow X_i \cap X_j = \varnothing$;

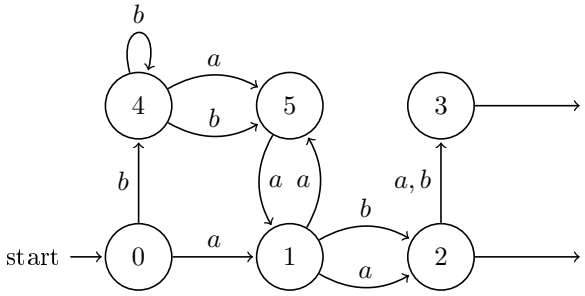
(3) $1 \in X_0$;

(4) $i \in X_j, i+1 \in X_l$ и a — i-ая буква в слове w, то $(q_j, a, q_l) \in E$;

(5) Если $|w| \in X_j$ и a — последняя буква в w, то есть такое $q \in F$, что $(q_j, a, q) \in E$.

Возьмем путь, помеченный w, и положим $i \in X_j \iff$ прочтя вдоль этого пути первые $i-1$ букву, мы окажемся в состоянии q_j. □

Пример. Рассмотрим язык, задаваемый автоматом:



и слово $w = b^3a^2b$. Тогда:

$1 \in X_0,\ 2 \in X_4,\ 3 \in X_4,\ 4 \in X_4,\ 5 \in X_5,\ 6 \in X_1$;

$X_0 = \{1\},\ X_2 = X_3 = \varnothing,\ X_1 = \{6\},\ X_4 = \{2, 3, 4\},\ X_5 = \{5\}$.

Лекция 2(08.03.2012).

Доказательство. Необходимость. Более подробно. Пусть L регулярен. Возьмем НКА, который распознает L (Q, Σ, q_0, F, E):

$E \subseteq Q \times \Sigma \times Q$;

$Q = \{q_0, q_1, ..., q_{k-1}\}$.

(*) $w \in L \iff$ существуют множества $X_0, X_1, ..., X_{k-1} \subseteq \{1, 2, ..., |w|\}$, такие что:

(1) $\bigcup\limits_{i=0}^{k-1} X_i = \{1, 2, ..., |w|\}$,

на логическом языке$\Phi_1 :\ (\forall x)(\bigvee\limits_{i=0}^{k-1}(x \in X_i))$;

(2)$i \neq j \Rightarrow X_i \cap X_j = \varnothing$,

$\Phi_2 : (\forall x)[\ \&\limits_{0 \leq i < j < k} \neg(x \in X_i \,\&\, x \in X_j)]$;

(3) $1 \in X_0$,

$\Phi_3 : (\forall x)[\forall y\,(x \neq y + 1) \to x \in X_0]$;

(4) $i \in X_j, i + 1 \in X_l$ и a — i-ая буква в слове w, то $(q_j, a, q_l) \in E$,

$\Phi_4 : (\forall x)[\forall y\,(y = x + 1 \to \&\limits_{0 \leqslant i,l < k}(x \in X_i \,\&\, y \in X_l \to \bigvee\limits_{S_{i,l}} Q_a x))]$,

где $S_{i,l} = \{a \mid (q_i, a, q_l) \in E\}$;

(5) Если $|w| \in X_j$ и a — последняя буква в w, то есть такое $q \in F$, что $(q_j, a, q) \in E$

$\Phi_5 : (\forall x)[\forall y\,(y \neq x + 1) \to \&\limits_{1 \leqslant j \leqslant k-1}(x \in X_j \to \bigvee\limits_{T_j} Q_a x)]$,

где $T_j = \{a \mid \exists a \in E\ (q_j, a, q) \in E\}$.

Если обосновать (*), то получится, что язык задаётся формулой

$$\exists X_0, \exists X_1, ..., \exists X_{k-1} \quad \Phi_1 \,\&\, \Phi_2 \,\&\, \Phi_3 \,\&\, \Phi_4 \,\&\, \Phi_5.$$

Если $w \in L$, то существует путь из q_0 в $q \in F$, метки которого образуют слово w.Тогда положим $i \in X_j$ тогда и только тогда, когда прочтя приставку слова w длины $i - 1$, мы оказываемся в состоянии q_j.

Обратно, пусть множества $X_0, X_1, ..., X_{k-1}$ со свойствами (1)-(5) существуют. По индукции по длине приставке w' слова w построим такой путь из q_0 в q_j, вдоль которого читается w'. Взяв приставку длины $|w| - 1$, получим путь в q_0, из которой есть стрелка j в состояние F, помеченное последней буквой w. □

Определение. Пусть Φ — формула первого порядка, в которой есть свободные переменные.

Пусть V — некоторое множество переменных, содержащее все свободные переменные из Φ.

Будем рассматривать «слова» над алфавитом $\Sigma \times 2^V$, т.е. «буквы» — это пары (a, U), где $U \subseteq V$.

V-слово — это последовательность букв из $\Sigma \times 2^V$:

$(a_1, U_1)(a_2, U_2)...(a_n, U_n)$ такая, что $\bigcup\limits_{i=1}^{n} U_i = V$ и $U_i \cap U_j = \varnothing$, при $i \neq j$.

Будем рассматривать такие формулы Φ, что каждая связанная переменная связывается ровно одним квантором.

Определение. Скажем, что формула Φ выполнена на V-слове

$w = (a_1, U_1)(a_2, U_2)...(a_n, U_n)$ (или V-слово w служит моделью для Φ), при некоторой интерпретации I. Обозначение: $w \models_I \Phi$. Если:

$w \models_I Q_a x \iff$ в w есть буква вида (a, S), где $x \in S$;

$w \models_I R(x_1, ..., x_k) \iff \bar{R}(j_1, ..., j_k)$, где $\bar{R}$ — это то отношение на множестве $1, ..., |w|$, которым интерпретируется R, а $j_1, ..., j_k$ — это позиции, которые занимают переменные $x_1, ..., x_k$;

$w \models_I \Phi\&\Psi \iff w \models_I \Phi$ и $w \models_I \Psi$;

$w \models_I \neg\Phi \iff w$ не служит моделью для Φ при I;
$w \models (\exists x)\Phi \iff$ существует такое $i \quad 1 \le i \le n$, что
$(a_1, U_1)...(a_{i-1}, U_{i-1})(a_i, U_i \cup \{x\})(a_{i+1}, U_{i+1})...(a_n, U_n) \models \Phi$.

$\forall x\, \Phi(x) \sim \neg\exists x\, \neg\Phi(x)$

Пример. $\Theta(x) \overset{I}{\rightleftharpoons}$ позиция x четна
$\exists x \forall y\, \neg(x < y)\,\&\,\Theta(x)$
Для формулы $\neg(x < y)\,\&\,\Theta(x)$
$V = \{x, y\},\ (a, \{y\})(b, \{x\})$ – пример V-слова.
Для формулы $\forall y\, \neg(x < y)\,\&\,\Theta(x)$ —множество четной длины, x в последней позиции
$V = \{x\},\ (a, \varnothing)(b, \{x\})$ — V-слово, удовлетворяющее этому множеству.

Лекция 3(15.03.2012).

Φ — формула монадической теории 2-го порядка (MSO), в которой есть свободные переменные как 1-го так и 2-го порядка.

Рассмотрим множества V_1, V_2.

V_i — множество переменных i-го порядка, содержащее все свободные переменные i-го порядка из Φ.

Будем рассматривать (V_1, V_2)-слова над алфавитом $\Sigma \times 2^{V_1} \times 2^{V_2}$;
типичная буква (a, S, T), где $a \in \Sigma,\ S \subseteq V_1,\ T \subseteq V_2$;
типичное (V_1, V_2)-слово $(a_1, S_1, T_1)(a_2, S_2, T_2)\dots(a_k, S_k, T_k)$;
$w \models_I \Phi$.
К атомарным формулам 1-го порядка относятся: $Q_a x,\ P(x_1, \dots, x_n)$,
к атомарным формулам 2-го порядка относится ещё $x \in X$.

$w \models_I (x \in X)$ — означает, что в w есть такая буква (a_i, S_i, T_i), что $x \in S_i,\ X \in T_i$.

Говорят, что $w \models_I (\exists X)\Psi$, если существует (возможно пустое) множество позиций $\mathcal{Y} \subseteq \{1, 2, \dots, k\}$, такое что слово w, полученное заменой каждой буквы (a_i, S_i, T_i), где $i \in \mathcal{Y}$, на букву $(a_i, S_i, T_i \cup \{x\})$, удовлетворяет Ψ.

Теорема (Бюхи(1960)). *Язык $L \subseteq \sum^*$ распознается конечным недетерминированным автоматом тогда и только тогда, когда он задается некоторой замкнутой формулой монадической теории 2-го порядка с отношением $=$ и единственным числовым предикатом «следует за».*

Доказательство. Достаточность. Пусть Φ — любая MSO формула. Докажем, что язык L_Φ всех (V_1, V_2)-слов, удовлетворяющих Φ, является регулярным (при любых V_1, V_2). Тогда теорема Бюхи — частный случай этого утверждения при $V_1 = V_2 = \varnothing$. Индукция по построению формулы.

Пусть L — множество всех (V_1, V_2)-слов. Легко проверить с помощью конечного автомата над алфавитом $\Sigma \times 2^{V_1} \times 2^{V_2}$, что каждая переменная первого порядка в V_1 встречается в точности один раз во входной строке, таким образом L – регулярный язык. Надо проверить, что если Φ – это атомарная формула, то то что ею задается – это регулярный язык.

$w \models Q_a x \iff$ в w есть буква (a, S, T), где $x \in S$. Легко проверить с помощью конечного автомата встречается ли конкретная переменная первого порядка в букве, чья первая компонента a. Пересечение множества всех таких слов с языком L – множество всех слов, удовлетворяющих $Q_a x$. Участвуют

только предикаты $x = y$, $y = x + 1$ ($w \models (y = x + 1) \iff$ в w есть фактор $(a_1, S_1, T_1)(a_2, S_2, T_2)$ такой, что $x \in S_1, y \in S_2$). Эти условия задаются конечными автоматами, упражнение построить автомат. С помощью конечного автомата можно проверить имеет ли любая буква x во второй компоненте и X в третьей, поэтому $x \in X$ – регулярный язык. Таким образом, утверждение доказано для атомарных формул.

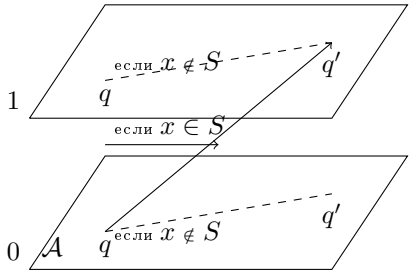
Если утверждение верно для формул Φ и Ψ, то оно верно и для $\Phi\&\Psi$ и$\neg\Phi$. В самом деле, т.к. регулярные языки замкнуты относительно пересечения и дополнения: $L_{\Phi\cap\Psi} = L_\Phi \cap L_\Psi \cap L$ и $L_{\neg\Phi} = L \setminus L_\Phi$ — регулярны.

Напомним, что $w \models (\exists x)\Phi \iff$ есть такое i, что
$(a_1, S_1, T_1)(a_2, S_2, T_2) \dots (a_i, S_i \cup \{x\}, T_i) \dots (a_k, S_k, T_k) \models \Phi$.

Пусть $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ — это тот автомат, который распознает язык L_Φ.

Построим новый автомат $\mathcal{M} = (Q \times \{0, 1\}, \Sigma, \delta', (q_0, 0), F \times \{1\})$.

Пусть $q \in Q, u \in \{0, 1\}$, переходы
$((q, u), (a, S, T), (q', u)) \in \delta'$, если $(q, (a, S, T), q') \in \delta$ и $x \notin S$;
$((q, 0), (a, S \setminus \{x\}, T), (q', 1)) \in \delta'$, если $(q', (a, S, T), q') \in \delta$ и $x \in S$.

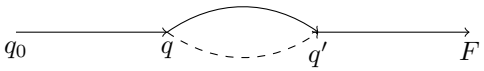


Легко видеть, что w принимается автоматом $\mathcal{M}$ тогда и только тогда, когда существует путь, соединяющий x со средней компонентой буквы слова w, так чтобы получилось слово, принимаемое автоматом $\mathcal{A}$. Таким образом, $\mathcal{M}$ распознает язык $L_{(\exists x)\Phi}$.

Напомним, что $w \models (\exists X)\Phi \iff$ есть такое $\mathcal{Y}$, что слово
$(a_1, S_1, T_1)(a_2, S_2, T_2) \dots (a_i, S_i, T_i \cup \{x\}) \dots (a_k, S_k, T_k) \models \Phi$, где $i \in \mathcal{Y}$.

Пусть $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ — это тот автомат, который распознает язык L_Φ.

Построим новый автомат $\mathcal{N} = (Q, \Sigma, \delta', q_0, F)$,
$\delta' = \{(q, (a, S, T \setminus \{x\}), q') |\ (q, (a, S, T), q') \in \delta\}$.



Автомат $\mathcal{N}$ распознает язык $L_{(\exists X)\Phi}$. □

3. Регулярные предикаты.

Пусть Φ —формула $\mathrm{MSO}(y = x+1)$, в которой есть предметные переменные.

Она определяет предикат на множестве позиций. А именно $(a_1, \dots, a_k) \in P_\Phi$ тогда и только тогда, когда $\Phi(a_1, \dots, a_k) = \text{И}$.

Пример. $x < y$ – регулярный предикат.

$x < y \leftrightharpoons y = x + 1 \ \vee \exists X\,(\forall z\,(z = x + 1 \to z \in X)\,\&$
$\forall t\,(y = t + 1 \to t \in X)\,\&\,\forall u\,(u \in X\,\&\,u + 1 \neq y \to \forall v\,(v = u + 1 \to v \in X)))$.

Пример. $x + y = z$ – нерегулярный предикат.

Докажем о/п. Пусть $\Phi(x, y, z) \leftrightharpoons x + y = z$. Тогда формула

$\Psi(x) \leftrightharpoons (\exists z)\ ((\forall y)\,(y \leqslant z)\,\&\,\Phi(x, x, z))$ определяет предикат «x – половина длины». Тогда $\exists x\ (\Psi(x)\,\&\,(\forall y)(y \leqslant x \to Q_a x)\,\&\,(y > x \to Q_b x))$ определяет нерегулярный язык $\{a^n b^n : n > 0\}$. Противоречие.

Домашнее задание. предикат - номер позиции делится на 3. Попробовать доказать (придумать формулу).

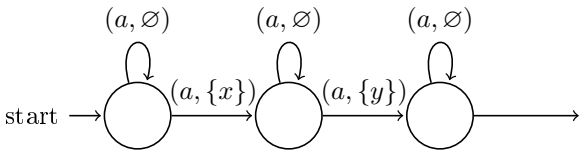
Лекция 4(22.03.2012).

Пусть $\Sigma = \{a\}$. Тогда предикат $Q_a x$ теряет свой смысл.

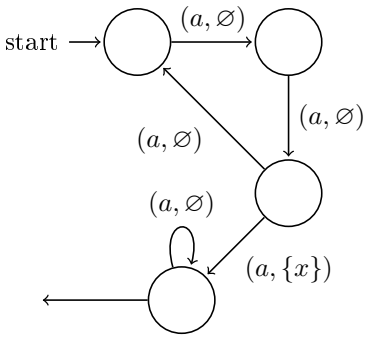
$(a, S, \varnothing) = (a, S)$.

Предикат регулярен, тогда и только тогда, когда он задается конечным автоматом.

Пример. Предикат $x < y$, тогда его распознает автомат:



Пример. Предикат «x делится на m», при $m = 3$ его распознает автомат:



Можно теперь применить алгоритм из доказательства теоремы Бюхи для нахождения формулы.

Определение. Порядковый тип слова.

Рассмотрим язык всех слов над алфавитом $\{(a, S) | S \subseteq \{x_1, \ldots, x_n\}\}$, задаваемый Ф. Обозначим через L_Φ. Берём слово w, берём все его буквы (a_i, S_i). Они образуют разбиение $\{x_1, \ldots, x_n\}$: $\{S_1, S_2, \ldots, S_h\}$ $(h \leqslant n)$. Записываем все переменные из S_i через знак $=$, а между ставим знак $<$. Такое выражение — порядковый тип w.

Пример. Пусть $n = 3$, переменные $\{x, y, z\}$.

Порядковый тип $x < y < z$, слово имеет вид

$$(a, \varnothing) \ldots (a, \varnothing)(a, \{x\})(a, \varnothing) \ldots (a, \varnothing)(a, \{y\})(a, \varnothing) \ldots (a, \varnothing)(a, \{z\})(a, \varnothing) \ldots (a, \varnothing).$$

Пример. Пусть $n = 3$, переменные $\{x, y, z\}$.

Порядковый тип $x = y < z$, слово имеет вид

$$(a, \varnothing) \ldots (a, \varnothing)(a, \{x, y\})(a, \varnothing) \ldots (a, \varnothing)(a, \{z\})(a, \varnothing) \ldots (a, \varnothing).$$

Теорема (Страубинг,1991)**.** *Предикат регулярен тогда и только тогда, когда он может быть задан формулой первого порядка, использующей предикаты «$x < y$» и «x делится на m».*

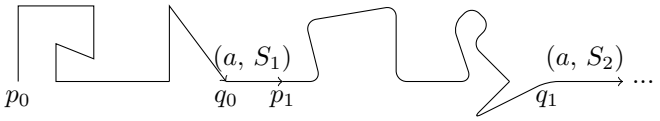
Доказательство. Достаточность очевидна. Подставляя формулы второго порядка в эти предикаты, получим, что задается с отношением «следует за», значит регулярен.

Необходимость. Пусть $P(x_1, \dots, x_n)$ – какой-то предикат и пусть он регулярен, т.е. задаётся какой-то формулой Φ $\mathrm{MSO}(y = x + 1)$. Пусть τ – порядковый тип. Обозначим через L_τ – множество всех слов типа τ. Достаточно доказать теорему для $L_\Phi \cap L_\tau$.

Итак, зафиксируем порядковый тип τ. Для простоты будем вместо $(a, \varnothing)$ писать a. Слово $w \in L_\Phi \cap L_\tau$ имеет вид $a^{m_0}(a, S_1)a^{m_1} \dots (a, S_j)a^{m_j}$, где $S_1, \dots, S_j$ однозначно определены типом τ. Пусть $\mathcal{A}$ – автомат, который распознает язык $L_\Phi \cap L_\tau$. Следом слова w в $\mathcal{A}$ назовем последовательность состояний

$$(p_0, q_0, p_1, q_1, \dots, p_j, q_j),$$

такую что a^{m_i} метит путь из p_i в q_i, а буква (a, S_i) помечает ребро из (q_{i-1}, p_i).



Через $L_{q,q'}$ обозначим множество всех слов вида a^m, которые метят путь из q в q'. Язык $L_{q,q'}$ – регулярный и является конечным объединением слов вида $(a^r) * a^s$, где $r, s \geqslant 0$.

$L_\Phi \cap L_\tau$ – это конечное объединение языков

$$L_{p_0,\, q_0}(a, S_1)L_{p_1,\, q_1}(a, S_2) \dots (a, S_j)L_{p_j,\, q_j} \ (*),$$

где объединение берётся по всевозможным следам.

Итак, достаточно написать формулу для языков вида $(*)$. Язык описывается следующими условиями:

(1) порядковый тип τ,
(2) $x_i = x_l + s + 1$,
(3) $x_i > x_l + s$,
(4) $x_i \equiv x_l + s (\mathrm{mod}\ r)$.

Расположение букв в слове выражается через булеву комбинаций условий (2),(3),(4).

Соответствующие формулы.

$x_i = x_l \rightleftharpoons \neg(x_i < x_l) \,\&\, \neg(x_l < x_i)$;

(1) конъюнкция выражений вида $x_i = x_l$ и $x_i < x_l$.

$x_i = x_l + 1 \rightleftharpoons (x_l < x_i) \,\&\, (\forall y)((x_l < y) \to (x_i = y) \vee (x_i < y))$;

(2) $(\exists y_1, \dots, y_s)((y_1 = x_i + 1) \,\&\, (x_l = y_s + 1) \,\&\, \underset{1 \leqslant m \leqslant s-1}{\&} (y_{m+1} = y_m + 1))$.

(3) $(\exists y_1, \dots, y_s)((y_1 = x_i + 1) \,\&\, (y_s < x_l) \,\&\, \underset{1 \leqslant m \leqslant s-1}{\&} (y_{m+1} = y_m + 1))$.

Если $0 < m < r$, то $x_i \equiv m (\mathrm{mod}\ r) \rightleftharpoons \exists z((z \equiv 0 (\mathrm{mod}\ r)) \wedge (x_i = z + m)$;

(4) $z \equiv x_i + s (\mathrm{mod}\ r) \rightleftharpoons \bigvee\limits_{m \in Z_r} ((x_i \equiv m (\mathrm{mod}\ r)) \wedge (z = m + s (\mathrm{mod}\ r)))$. $\square$

Пример. $(a^3)^*a^2(a,\{x_1\})(a^3)^*a(a,\{x_2\})(a^6)^*a^3$;
$x_2 > x_1, \quad x_2 = x_1 + 1(\mathrm{mod}\ 3)$.

Лекция 5(05.04.2012).

4. Языки бесконечных слов.

В логическом языке на самом деле не важно, чтобы слова были конечными. Всё сохраняет смысл для бесконечных слов. Идея Бюхи: использовать этот подход, чтобы определить регулярный язык бесконечных слов. Бесконечные слова имеют осязаемое практическое применение. Слова - протоколы некоторых процессов. Когда работает компьютер, то это потенциально бесконечная последовательность операций.

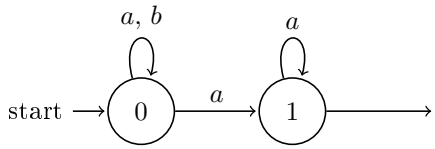
MSO(+1)
$\Sigma \quad Q_a$
Σ^ω — множество всех бесконечных вправо слов.

Определение. Язык $L \subseteq \Sigma^\omega$ назовем регулярным, если он состоит из всех слов, удовлетворяющих некой системе Φ формул MSO(+1).

Пусть $\mathcal{A}$ – конечный (недетерминированный) автомат.
$\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$;
$\delta : Q \times \Sigma \to 2^Q$.

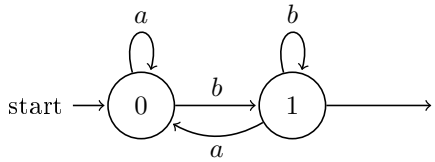
Определение. Скажем, что такой автомат $\mathcal{A}$ принимает слово $w \in \Sigma^\omega$ (в смысле Бюхи), если при чтении w автомат оказывается в состояниях из F бесконечное число раз.

Пример. Недетерминированный автомат:



распознает язык $L(\mathcal{A}) = \{a,b\}^*a^\omega$,
$(\exists y)(\forall x)(x > y \to Q_a x)$.

Пример. Детерминированный автомат:



распознает дополнение языка $L(\mathcal{A})$.
$L(\mathcal{B}) = (a^*ba^*)^\omega$,
$(\forall y)(\exists x)(x > y \ \&\ Q_b x)$.

Определение. Назовем таблицей некоторое множество множества состояний. Скажем, что автомат принимает слово в смысле Мюллера, если мы его читаем и выписываем те состояния, которые он посещает бесконечное число раз. Эти состояния должны быть в таблице.

Утверждение (Ландвебер, 1969). Язык $\{a,b\}^*a^\omega$ не распознается по Бюхи никаким детерминированным автоматом.

Доказательство. о/п. Пусть существует детерминированный автомат, который распознает язык $\{a,b\}^*a^\omega$. Он принимает слово ba^ω. Значит, это слово метит какой-то путь, который бесконечно много раз проходит через состояния из F. В силу детерминированности этот путь единственный. Пусть k_1таково, что ba^{k_1} метит путь из q_0 в состояния из F. Рассмотрим слово $ba^{k_1}ba^\omega \in L$. Значит, оно тоже метит какой-то путь, в котором бесконечно много раз встречаются состояния из F. Возьмем k_2 такое, что слово $ba^{k_1}ba^{k_2}$ заканчивается в состоянии из F. Рассмотрим $ba^{k_1}ba^{k_2}ba^\omega \in L$. Отсюда $ba^{k_1}ba^{k_2}\dots ba^{k_n}\dots$ бесконечно много раз посещает состояния из F. Но это слово содержит бесконечное число букв b. Противоречие. □

Теорема (Бюхи)**.** *Пусть $L \subseteq \Sigma^\omega$, L распознается (по Бюхи) конечным автоматом тогда и только тогда, когда L задается некоторым набором формул MSO(+1).*

Доказательство. Доказательство необходимости повторяет доказательство в случае конечных слов, за исключением одной модификации. А именно, вместо условия (5) нужно такое условие: одно из множеств X_q – бесконечно, где $q \in F$. $(\forall x)(\exists y)(y > x \ \&\ y \in X) \rightleftharpoons X$ бесконечно.

Доказательство достаточности повторяет доказательство в конечном случае для конечных слов, если принять на веру такой факт. Если $L_1, L_2 \subseteq \Sigma^\omega$ распознаются, то $L_1 \cup L_2$, $L_1 \cap L_2$, $\Sigma^\omega \setminus L_2$ распознаются. □

Утверждение. Если $L_1, L_2 \subseteq \Sigma^\omega$ – распознаваемы, то и $L_1 \cup L_2$ распознаваем.

Доказательство. Пусть $\mathcal{A}_i = (Q_i, \Sigma, q_0^{(i)}, \delta_i, F_i),\ i = 1,2$ – автомат, который распознает L_i. Будем считать, что $Q_1 \cap Q_2 = \varnothing$. Определим $\mathcal{A}$ так:
$Q = Q_1 \cup Q_2 \cup \{q_0\}$, q_0 – новое начальное состояние,
$F = F_1 \cup F_2$ – множество заключительных состояний,
$\delta = \delta_1 \cup \delta_2 \cup \{(q_0, a, q) | (\exists q \in Q_i) : (q_0^{(i)}, a, q) \in \delta_i\}$.
Тогда автомат $\mathcal{A}$ распознает язык $L_1 \cup L_2$. □

Предложение. *$L \in \Sigma^\omega$ распознается конечным автоматом (по Бюхи) тогда и только тогда, когда L можно представить в виде конечного объединения языков вида JK^ω, где J – регулярный язык в Σ^*, а K – регулярный язык в Σ^+.*

Доказательство. Необходимость. Пусть автомат $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ распознает язык L. Пусть J – это тот язык, который распознается этим автоматом в «конечном смысле», а K – это моножество тех непустых слов, которые в $\mathcal{A}$ читаются на пути из F в F. Через $L_{p,q}$ обозначим язык конечных слов, которые можно прочесть на пути из $p \in Q$ в $q \in Q$. Тогда $L = \bigcup\limits_{q \in F} L_{q_0,q} L_{q,q}^\omega$.

Достаточность. Чтобы доказать, достаточно (в силу предложения 1) проверить, что каждый язык вида JK^ω распознается по Бюхи конечным автоматом. Пусть $\mathcal{B} = (Q, q_0, F, \delta)$ – автомат, который распознает K. Возьмем новое состояние q_0' и для каждой стрелки $q_0 \xrightarrow{a} q$ нарисуем стрелку $q_0' \xrightarrow{a} q$, а для каждой стрелки $q \xrightarrow{a} f \in F$ нарисуем стрелку $q \xrightarrow{a} q_0'$. Обозначим через δ' расширенное множество переходов. Тогда автомат $\mathcal{B}' = (Q \cup \{q_0'\}, q_0', \{q_0'\}, \delta')$ распознает по Бюхи язык K^ω. Пусть автомат (P, p_0, T, Δ) – автомат, который распознает язык

J. Полагаем, что $P \cap (Q \cup \{q_0'\}) = \varnothing$. Пусть $\Delta' = \{(p, a, q_0') : \exists t \in T,\ (p, a, t) \in \Delta\}$. Тогда автомат

$$(P \cup Q \cup \{q_0'\}, p_0, \{q_0'\}, \Delta' \cup \Delta \cup \delta')$$

распознает язык JK^ω. □

Лекция 6 (12.04.2012)

$\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$;
$w \in \Sigma^+$;
$s(q, q', w) \rightleftharpoons$ есть путь из q в q', помеченный словом w;
$t(q, q', w) \rightleftharpoons -//-$ и проходящий через одно из состояний из F.
Введем отношение эквивалентности:
$u \sim v \rightleftharpoons$ для всех $q, q' \in Q,\ s(q, q', u) \longleftrightarrow s(q, q', v)$ & $t(q, q', u) \longleftrightarrow t(q, q', v)$.

Лемма 1. *У $\sim$ конечное число классов и каждый класс — регулярный язык.*

Доказательство. $\sim$-класс слова w определяется набором таких пар (q, q'), что $s(q, q', w)$ & $t(q, q', w)$. Поэтому число классов не больше, чем число множеств таких пар. Введем множества слов $S(q, q') = \{u|\ s(q, q', u)$ истинно $\}$ и $T(q, q') = \{u|\ t(q, q', u)$ истинно $\}$. Каждый класс – это булева комбинация таких множеств. Для слова w его $\sim$- класс:

$$\Big(\bigcup_{\substack{(q,q'),\\ \text{где } s(q,q',w)\\ \text{истинно}}} S(q,q') \cap \bigcup_{\substack{(q,q'),\\ \text{где } t(q,q',w)\\ \text{истинно}}} T(q,q')\Big) \setminus \Big(\bigcup_{\substack{(p,p'),\\ \text{где } s(p,p',w)\\ \text{ложно}}} S(p,p') \cap \bigcup_{\substack{(p,p'),\\ \text{где } t(p,p',w)\\ \text{ложно}}} T(p,p')\Big).$$

Поэтому осталось доказать, что регулярны $S(q, q')$ и $T(q, q')$.
$S(q, q') = L(q, q')$,
$T(q, q') = \bigcup\limits_{f \in F} L_{q,f} \cdot L_{f,q}$.
Так как множества вида $L(q, q')$ — регулярные языки, все доказано. □

Лемма 2. *Пусть J, K – $\sim$-классы, $L \subseteq \Sigma^\omega$ и $L = L(\mathcal{A})$. Если $JK^\omega \cap L \neq \varnothing$, то $JK^\omega \subseteq L$.*

Доказательство. Пусть $\alpha \in JK^\omega \cap L$. Тогда $\alpha = \alpha_0\alpha_1\alpha_2 \ldots$, где $\alpha_0 \in J, \alpha_i \in K$. Значит, в $\mathcal{A}$ есть последовательность состояний $q_0, q_1, q_2 \ldots$ такая, что $s(q_l, q_{l+1}, \alpha_l)$ истинно для всех l и $t(q_j, q_{j+1}, \alpha_j)$ истинно для бесконечного множества индексов j. Пусть β– любое слово из JK^ω. Тогда $\beta = \beta_0\beta_1\beta_2 \ldots$, где $\beta_0 \in J, \beta_i \in K$ при $i > 0$. Имеем $\alpha_i \sim \beta_i$ для всех i, откуда $s(q_l, q_{l+1}, \beta_l)$ истинно для всех l и $t(q_j, q_{j+1}, \beta_j)$ истинно для бесконечно многих j. Значит, $\beta \in L$. □

Теорема (Рамсей, конечный вариант)**.** *Пусть k – любое число. Найдется такое $R = R(k)$, что в любом графе с R вершинами есть либо полный подграф с k вершинами, либо пустой подграф с k вершинами.*

Доказательство. Пусть $r(m, n)$ – наименьший размер графа, у которого всегда есть либо полный подграф с m вершинами либо пустой подграф с n вершинами. Тогда $R(k) = r(k, k)$. Ясно, что $r(2, n) = n,\ r(m, 2) = m$. Докажем, что $r(m, n) \leqslant r(m, n-1) + r(m-1, n)$. Возьмем граф с $N = r(m, n-1) + r(m-1, n)$ вершинами и пусть v – любая вершина. Рассмотрим множество G всех вершин, смежных с v. Если в G по крайней мере $r(m-1, n)$ вершин, то все хорошо. В

противном случае в исходном графе по крайней мере $r(m, n-1)$ вершин, не смежных с v и опять все хорошо. $\square$

Пример. Задача про знакомства: $R(3) = 6$.

Обобщение. Пусть k, c любые числа. Тогда $\exists R = R(k, c)$, что если ребра полного графа с R вершинами покрасить в c цветов, то найдется одноцветный полный граф с k вершинами.

Пример. $R(3, 3) = 17$.

Обобщение. Пусть k, c, s – любые числа. Тогда $\exists R = R(k, c, s)$, что в любом подмножестве из R элементов покрасить все s-подмножеств в c цветов, то найдется k- подмножество, у которого все s-подмножества будут покрашены в один цвет.

Лемма 3. *Пусть $\alpha \in \Sigma^\omega$. Тогда существуют $\sim$-классы J, K такие, что $\alpha \in JK^\omega$.*

Доказательство. Пусть $i < j$. Обозначим через $\alpha[i, j]$ кусок слова α от i-ой позиции до $j-1$-й позиции. Покрасим все 2-х элементные подмножества в $\mathbb{N}$ красками, соответствующими $\sim$-классам так: $\{i, j\}$,где $i < j$ – это $\sim$-классы слова $\alpha[i, j]$. По теореме Рамсея в $\mathbb{N}$ есть бесконечное подмножество $\{i_1 < i_2 < i_3 < \ldots\}$, у которого все 2-х элементные подмножества покрашены одним цветом. Пусть K – это $\sim$-класс, который отвечает этому цвету $\alpha[i_1, i_2], \alpha[i_2, i_3], \ldots \in K$. Пусть Y - это класс слова $\alpha[1, i_1]$. Тогда $\alpha = \alpha[1, i_1]\alpha[1, i_2] \ldots \in JK^\omega$. $\square$

Утверждение. Если $L \subseteq \Sigma^\omega$ — распознаваемый язык, то $\Sigma^\omega \setminus L$ — распознаваемый язык.

Доказательство. По лемме 3 для любого $\alpha \in \Sigma^\omega \setminus L$ существуют $\sim$-классы J, K, такие что $\alpha \in JK^\omega$. Значит, $\Sigma^\omega \setminus L$ лежит в объединении множеств JK^ω. Если бы $JK^\omega \cap L \neq \varnothing$, то по лемме 2 множество $JK^\omega \subseteq L$, но JK^ω содержит слова из дополнения языка L. Получили, что $JK^\omega \cap L = \varnothing$. Отсюда имеем, что объединение множеств JK^ω лежит $\Sigma^\omega \setminus L$. Значит дополнение языка L – это объединение множеств JK^ω. По лемме 1 объединение конечное и J, K — регулярные языки. По предложению $\Sigma^\omega \setminus L$ — распознаваемо. $\square$

Лекция 7 (19.04.2012).

Следствие 1. *Монадическая теория 2-го порядка множества $\mathbb{N}$ с отношением «следует за» разрешима.*

S2S - то же разрешима (деревья).

Теперь под формулами понимаем формулы первого порядка.

Любую формулу первого порядка можно заменить на эквивалентную ей префиксную формулу, т.е. формулу вида:

$\forall x\ \exists y\ \forall z\ \exists w\ \Phi(x, y, z, \ldots, w) \leftarrow$ бесквантороная часть.

$\neg(\forall x\ \Phi(x)) \sim \exists x\ \neg\Phi(x)$,

$\neg(\exists x\ \Phi(x)) \sim \forall x\ \neg\Phi(x)$,

$\forall x\ \Phi(x)\,\&\,\forall x\ \Psi(x) \sim \forall x\ (\Phi(x)\,\&\,\Psi(x))$,

$\exists x\ \Phi(x) \vee \exists x\ \Psi(x) \sim \exists x\ (\Phi(x) \vee \Psi(x))$,

$\forall x\ \Phi(x) \vee \forall x\ \Psi(x) \nsim \forall x\ (\Phi(x) \vee \Psi(x))$,

$\forall x\ \Phi(x) \vee \forall x\ \Psi(x) \sim \forall x \forall y\ (\Phi(x) \vee \Psi(y))$,

$\exists x\ \Phi(x) \,\&\, \exists x\ \Psi(x) \nsim \exists x\ (\Phi(x) \,\&\, \Psi(x))$,
$\exists x\ \Phi(x) \,\&\, \exists x\ \Psi(x) \sim \exists x \exists y\ (\Phi(x) \,\&\, \Psi(y))$.

5. Игры Эренфойхта-Фрессе.

Определение. Сложность формулы (c):
Если Φ – атомарная формула, $c(\Phi) = 0$,
$c(\neg\Phi) = c(\Phi)$,
$c(\Phi \,\&\, \Psi) = \max\{c(\Phi), c(\Psi)\}$,
$c(\exists x\ \Phi) = c(\Phi) + 1$.

Предложение. *Пусть v – конечный набор переменных. Тогда существует только конечное число формул первого порядка в префиксной форме, у которых сложность $< c$, а все переменные лежат в v.*

Доказательство. Число атомарных формул ограничено в терминах числа предикатных символов и числа подмножеств в v. Дальше можно применить индукцию по c. □

Пусть w_1, w_2 - v-слова
$(a_1, X_1)(a_2, X_2)\dots(a_n, X_n), X_i \cap X_j = \varnothing,$ при $i \neq j, \bigcup\limits_{i=1}^{n} X_i = v$.

$w_1 \sim_r w_2 \rightleftarrows w_1$ и w_2 удовлетворяют одному и тому же набору формул сложности $\leqslant r$.

Из предложения следует, что $\sim_r$ эквивалентность конечного индекса.

Определение. Определим r-раундовую игру на (w_1, w_2). Алиса хочет доказать, что w_1 и w_2 различны, а Боб хочет доказать противоположное. У каждого игрока r жетонов. На i-ом ходу Алиса кладет жетон z_i на букву в одном из слов. В ответ Боб должен положить свой жетон z с тем же номером на какую-то букву в другом слове. Когда игра заканчивается, у нас есть два $(v \cup \{z_1, \dots z_n\})$-слова w_1' и w_2'. Боб выигрывает, если для каждой атомарной формулы α:

$$\alpha \text{ выполнена на } w_1' \iff \alpha \text{ выполнена на } w_2'.$$

В противном случае Алиса выигрывает.

Лекция 8 (26.04.2012).

$u \sim_r v \rightleftharpoons u$ и v удовлетворяют одним и тем же формулам сложности $\leqslant r$.

Теорема. $u \sim_r v \iff$ *Боб имеет выигрышную стратегию в r-раундовой игре на (u, v).*

Доказательство. Необходимость. Индукция по r. Если $r = 0$, то Боб сразу выиграл, т.к. w_1 и w_2 удовлетворяют одним и тем же формулам. Пусть утверждение верно для $r - 1$ и не верно для r. Это значит, что для какой-то пары слов (u, v) со свойством $u \sim_r v$ у Алисы есть выигрышная стратегия. Позволим Алисе сделать первый ход этой стратегии. Можно считать, что она ставит жетон на букву из слова u. В результате получится такая структура u', что при любом ответе Боба получится такая пара (u', v'), на которой Алиса имеет выигрышную стратегию в игре с $r - 1$ раундом. По предположению индукции $u' \nsim_{r-1} v'$. Пусть Φ – это конъюнкция нормальных форм всех формул сложности $< r$, которые выполнены в u'. Тогда Φ не выполнена в v' при любом

способе получения v'. Это значит, что формула $\exists z_1\ \Phi$ не выполнена на v, но она выолнена на u. Отсюда $u \nsim_r v$. Противоречие.

Достаточность. Индукция по r. Если $r = 0$, то $u \sim_0 v$. Пусть утверждение верно для $r-1$ и предположим, что у Боба есть выигрышная стратегия на (u, v), но $u \nsim_r v$. Последнее означает, что существует формула Φ сложности $\leqslant r$ такая, что она выполнена на u, но не выполнена на v. Можно считать, что Φ имеет вид $\exists z\ \Psi$, где Ψ – формула сложности $\leqslant r-1$. Если Алиса положит свой жетон на то место, которое делает истинным формулу Ψ на u, то Боб может отметить какое-то место в слове v. Получится новая пара слов (u', v'), в которой у Боба есть выигрышная стратегия в $r-1$ ходов. По предположению индукции имеем, что u' и v' удовлетворяют одним и тем же формулам сложности $\leqslant r-1$. В частности, v' удовлетворяет Ψ. Но тогда слово v удовлетворяет формуле $\exists z\ \Psi = \Phi$. Противоречие. □

Следствие. *С помощью теоремы можно доказать, что*
$(\mathbb{R}, <)$ *и* $(\mathbb{Q}, <)$ *удовлетворяют одним и тем же формулам 1-го порядка.*

Пример. $V = \varnothing, w_1 = ab, w_2 = baa$. Отношение — отношение $=$.
В нораундовой игре выигрывает Боб.
В двухраундовой игре выигрывает Алиса:
Алиса $w_2 = \underset{z_1}{b}\underset{z_2}{a}a$, Боб $w_1 = \underset{z_1}{a}\underset{z_2}{b}$.
Соответствующая формула $\exists x\ \exists y\ (\neg(x = y)\,\&\,Q_a x\,\&\,Q_a y)$.

Пример. 2)w_1 – слово, которое кончается на a,w_2 – слово, которое кончается на b. Отношение $<$.

Алиса выигрывает за два раунда. Первым ходом она ставит z_1 на последнюю букву в w_1. Если в w_2 нет буквы a, то Боб сразу проиграл. Если есть, он вынужден поставить z_1 на эту букву. Тогда Алиса ставит свой жетон на последнюю букву в w_2. Боб отвечает и проигрывает.

В w_2 $z_1 < z_2$, в w_1 $z_2 < z_1$.
$\exists x\ (\forall y\,\neg(x < y)\,\&\,Q_b x)$.

$\mathrm{Reg} = \mathrm{MSO}(<) = \mathrm{MSO}(y = x + 1)$
$\mathrm{Reg} \supsetneqq \mathrm{FO}(<) \supsetneqq \mathrm{FO}(y = x + 1)$

Теорема. *Множество слов четной длины не лежит в FO(<).*

Доказательство. О/п. Допустим, что есть формула Φ, которая определяет множество слов четной длины. Пусть $c(\Phi) = r$. Тогда Φ выполнена на $a^k \iff k$ чётно. Покажем, что $a^{2^r} \sim_r a^{2^r-1}$. Мы построим выигрышную стратегию для Боба r-раундовой игры на паре (a^k, a^{k+1}) при $k \geqslant 2^r - 1$. Это делается индукцией по r. При $r = 1$ любой ответ Боба – выигрышный. Пусть $r > 1$ и утверждение верно при $r-1$. Допустим, что Алиса положила z_1 на одно из двух слов, получив структуру $a^s(a, \{z_1\})a^t$. Тогда либо $s \leqslant \frac{k-1}{2}$, либо $t \leqslant \frac{k-1}{2}$. Считаем, что $s \leqslant \frac{k-1}{2}$, если иначе выполняем все симметрично. Боб ставит свой жетон на $(s+1)$-ую букву 2-го слова: $a^s(a, \{z_1\})a^{t'}$, где $t' = t+1$ или $t' = t-1$. Так как $2^{r-1} - 1 \leqslant k = min\{t, t'\} + s + 1 \leqslant min\{t, t'\} + \frac{k-1}{2} + 1$, получаем, что $min\{t, t'\} \geqslant \frac{k-1}{2} \geqslant 2^{r-1} - 1$. По предположению индукции Боб имеет выигрышную стратегию в $(r-1)$ раундовой игре на $(a^t, a^{t'})$. Если теперь Алиса ставит жетон на i-ую букву с $i \leqslant s$, тогда Боб отвечает ходом на i-ую букву второго слова. Иначе Боб действует согласно своей выигрышной стратегии для пары

$(a^t, a^{t'})$. Докажем, что Боб выигрывает. В результате получается две $(z_1, \ldots, z_r)$ - структуры. Надо показать, что они удовлетворяют одинаковым атомарным формулам $z_i < z_j$. Пусть первая структура удовлетворяет этой формуле. Если z_i и z_j появляются среди первых $(s+1)$ буквы. То они стоят на тех же позициях во второй структуре. Если они стоят среди последних $t(t')$ букв, то можно считать, что они появились в результате какой-то игры на паре $(a^t, a^{t'})$, а поскольку Боб придерживался выигрышной стратегии, $z_i < z_j$ выполнено и во втором слове. Если z_i - в голове, а z_j - в хвосте, то и во втором слове это верно. □

Лекция 9(03.05.2012).

$FO(x = y+1) \subseteq FO(<) \subset Reg = MSO(<) = MSO(x = y+1)$

6. Локально барьерно тестируемые языки.

Рассмотрим на множестве слов Σ^* отношение $\approx_r^k$.

$w_1 \approx_r^k w_2 \iff$ выполняются следующие три условия:

1) у w_1 и w_2 одинаковые приставки длины $k-1$,

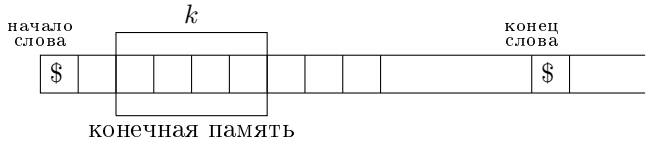
2) у w_1 и w_2 одинаковые суффиксы длины $k-1$,

3) если w – слово длины k, то число вхождений слова w в w_1 в качестве фактора равно числу вхождений слова w в w_2 в качестве фактора, если оба этих числа $< r$; либо оба этих числа $\geqslant r$.

Определение. Если $w_1 = uwv$, где $u, v \in \Sigma^*$, то w — фактор слова w_1.

Пример. В слове *abbaabbba* слово *bb* входит 3 раза.

Сканирующее устройство:



Два слова находятся в отношении $w_1 \approx_r^k w_2$, если сканирующее устройство не может их отличить.

$w_1 \approx_r^k w_2$ – это эквивалентность конечного индекса.

Определение. Язык L называется локально барьерно тестируемым, если он является объединением классов $\approx_r^k$ для некоторых r и k.

Мы хотим доказать, что именно эти языки соответствуют $FO(x = y+1)$.

Лемма. *Любой локально барьерно тестируемый язык задаётся формулой теории первого порядка с отношением* $(x = y+1)$.

Доказательство. Достаточно доказать, что для всех k и r каждый $\approx_r^k$ класс задаётся $FO(x = y+1)$. Выпишем формулу соответствующую условиям 1),2),3).

1) Пусть u – какое-то фиксированное слово длины $k-1, u = a_1 a_2 \ldots a_{k-1}$.

$\exists x_1 \exists x_2 \ldots \exists x_{k-1}(First(x_1) \,\&\, x_2 = x_1{+}1 \,\&\, \ldots \,\&\, x_{k-1} = x_{k-2}{+}1 \,\&\, \underset{1 \leqslant i \leqslant k-1}{\&} Q_{a_i} x_i)$,

где $First(x)$ – сокращение для формулы $\neg\exists y(x = y+1)$.

2) Аналогично.

3) Пусть $v = a_1 a_2 \dots a_k$ – слово длины k и t – число меньше r.

$$\exists x_{11} \exists x_{12} \dots \exists x_{1k} \exists x_{21} \dots \exists x_{tk} (\mathop{\&}_{1 \leqslant i \leqslant t} \mathop{\&}_{1 \leqslant j \leqslant k-1} (x_{ij+1} = x_{ij} + 1) \,\&$$

$$\mathop{\&}_{1 \leqslant i < j \leqslant t} (x_{i1} \neq x_{j1}) \,\& \mathop{\&}_{1 \leqslant i \leqslant t} \mathop{\&}_{1 \leqslant j \leqslant k-1} Q_{a_j} x_{ij}).$$

□

Теорема. *Языки из $FO(x = y + 1)$ – это в точности локально барьерно тестируемые языки.*

Доказательство. Пусть $\sim_r$ – это отношение, определяемое правилом

$w_1 \sim_r w_2 \iff$ на w_1 и w_2 выполнены одни и те же $\mathrm{FO}(x = y + 1)$-формулы сложности $\leqslant r$. Докажем, что если $R = 3^r$ и $w_1 \approx^r_{3R} w_2$, то $w_1 \sim_r w_2$, т.е. у Боба есть выигрышная r-раундовая стратегия.

Множество $\{i, \dots, j\}$ назовём носителем фактора. Два фактора разделены, если их носители не пересекаются. Все неразделенные факторы объединим – это объединение множества исходных факторов. Допустим, что i раундов уже сыграли. Это значит, что $z_1, \dots, z_i$ уже поставлены в каждом слове. Для каждой позиции m,на k-ую поставлен жетон в слове $w_j, (j = 1$ или $2)$. Рассмотрим фактор $w_j[m - 3^{r-i} + 1, m + 3^{r-i}]$. Если $m - 3^{r-i} + 1 < 1$ или $m + 3^{r-i} > |w_j| + 1$, то соответственно рассматриваем префикс $w_j[1, m + 3^{r-i}]$ или суффикс $w_j[m - 3^{r-i} + 1, |w_j|]$. Рассмотрим объединение в каждом слове. Пусть это объединение состоит из $k_{i,j}$ факторов $u_{i,j,1}, \dots, u_{i,j,k_{i,j}}$. Объясним, как Боб должен играть, чтобы для каждого $p, 1 \leqslant p \leqslant j$ фактор $u_{i,1,n}$, в котором стоит жетон z_p в первом слове был равен фактору $u_{i,2,n}$, в котором этот же жетон стоит во втором слове, и при этом z_p стоит в обоих факторах на той же позиции. Индукция по i. Если $i = 0$, доказывать нечего. Допустим, что утверждение верно для i. Пусть есть выигрышная r-раундовая стратегия. Пусть Алиса поставила z_{i+1} на какую-то позицию m слова w_1. Если фактор $v = w_1[m - 3^{r-i-1}, m + 3^{r-i-1}]$ целиком попал в один из факторов $u_{i,1,n}$, то Боб просто ставит свой жетон на соответствующую позицию в $u_{i,2,n}$. Если это не так, то v разделен от факторов $u_{i+1,1,h}$, которые содержат жетоны $z_1, \dots, z_i$. В слове w_2 берем фактор нужной длины, отделенный от всех факторов вида $u_{i+1,2,h}$ и ставим z_{i+1} в его середину. □

Лекция 10 (10.05.2012).

Следствие. *$FO(y = x + 1) \subset FO(<)$.*

Доказательство. Пусть $\Sigma = \{a, b, c\}$ и $L = a^* b a^* c a^*$. Докажем, что $L \in FO(<)$. Просто напишем формулу:

$$\exists x \; \exists y \; (x < y) \; Q_b x \,\& \, Q_c y \,\& \, \forall z \; (x \neq z \,\& \, y \neq z \to Q_a z).$$

Для доказательства того, что $L \notin FO(y = x + 1)$, воспользуемся теоремой. Допустим, что язык L является объединением $\approx^k_r$-классов для каких-то r и

k. Рассмотрим $a^kba^kca^k$ и $a^kca^kba^k$. Видно, что они находятся в одном $\approx_k^r$-классе (при любом r). У них одинаковые слова длины k: $a^{l-1}ba^{k-l}, a^{l-1}c^{k-l}$. Но $a^kba^kca^k \in L$, а $a^kca^kba^k \notin L$. Противоречие. □

7. Алгебраическая характеризация $FO(<)$.

Определение. Если L - язык в Σ^*, то $\sim_L$(синтаксическая конгруэнция) определяется так:

$w_1 \sim_L w_2 \iff (\forall u, v \in \Sigma^*\ uw_1v \in L \iff uw_2v \in L)$.

Синтаксический моноид — это $\Sigma^*/\sim_L$.

Моноид апериодический, если все подгруппы тривиальны или, что эквивалентно, $\mathcal{H}$-тривиальны или, что эквивалетно, $x^k = x^{k+1}$ для любого x и для некоторого k.

Теорема (Макнотон)**.** *Язык задаётся формулой первого порядка с предикатом $<$ тогда и только тогда, когда его синтаксический моноид – апериодический.*

Доказательство. Необходимость. Индукция по построению формулы.

База индукции. Формула атомарна. Если $uw^2v \models Q_ax$, то x появляется либо в u, либо в v. Но тогда $uw^3v \models Q_ax$ и обратно, если $uw^3v \models Q_ax$, то $uw^2v \models Q_ax$. Отсюда uw^2v и uw^3v одновременно принадлежат или не принадлежат L. Поэтому $w^3 \sim w^2$ и в $\Sigma^*/\sim\ x^3 = x^2$ для любого x. Значит, моноид языка $L(Q_ax)$ апериодический. Если $uw^2v \models x < y$, то x и y появляются либо в u, либо в v. Но тогда $uw^3v \models x < y$ и обратно, если $uw^3v \models x < y$, то $uw^2v \models x < y$.

Шаг индукции. Пусть Φ и Ψ – такие формулы, что $M(L(\Phi))$ и $M(L(\Psi))$ – апериодические. Докажем, что и $M(L(\Phi\,\&\,\Psi))$ апериодический. Пусть k такое, что $x^k = x^{k+1}$ для всех $x \in M(L(\Phi))$ и всех $x \in M(L(\Psi))$. Возьмем любые слова u, w, v. Имеем $w^{k+1} \sim_{L(\Phi)} w^k$, откуда $uw^{k+1}v \models \Phi \iff uw^kv \models \Phi$ и $w^{k+1} \sim_{L(\Psi)} w^k$, откуда $uw^kv \models \Psi \iff uw^kv \models \Psi$. Поэтому $uw^{k+1}v \models \Phi\,\&\,\Psi \iff uw^kv \models \Phi\,\&\,\Psi$. Отсюда $w^k \sim_{L(\Phi\,\&\,\Psi)} w^{k+1}$ и $x^k = x^{k+1}$ в $M(L(\Phi\,\&\,\Psi))$. С отрицанием тоже легко. Допустим, что $uw^{2k+1}v \models \exists x\Phi$. Можно x куда-то «положить» так, что получится слово, удовлетворяющее Φ: $(a, X) \to (a, X \cup \{x\})$. Это слово можно представить в виде $u'w^kv' \models \Phi$. Поэтому и $u'w^{k+1}v' \models \Phi$. Отсюда следует, что обратное тоже верно

$$uw^{2k+1}v \in L(\exists x\Phi) \iff uw^{2k+2}v \in L(\exists x\Phi).$$

Следовательно, $w^{2k+1} \sim_{L(\exists x\Phi)} w^{2k+2}$, т.е. в $M(L(\exists x\Phi))$ выполнено равенство $x^{2k+1} = x^{2k+2}$.

Достаточность. Нужно доказать, что если моноид языка L апериодический, то $L \in FO(<)$. Мы воспользуемся теоремой Шютценберже, согласно которой, L - беззвездный язык. Формула строится индукцией по построению беззвездного языка из одноэлементных и единственный сложный элемент − произведение. Пусть $L_1 = L(\Phi), L_2 = L(\Psi)$ как построить формулу, задающую L_1, L_2? Нужно научиться по данной формуле Φ строить новую формулу Φ', такую что $w \models \Phi' \iff$ у w есть приставка, удовлетворяющая Φ. □

Лекция 11 (17.05.2012).

$L_1L_2 = \{w | \exists u \in L_1\ \exists v \in L_2\ w = uv\}$.

Пусть w – слово. Скажем, что φ выполнена на w между i и j, где $i < j$, если $w = a_1 a_2 \dots a_i a_{i+1} \dots a_{j-1} a_j \dots a_n$ и слово $a_i a_{i+1} \dots a_{j-1}$ удовлетворяет формуле φ.

Предложение (о регуляризации)**.** *Для любой формулы φ теории 1-го порядка с отношением $<$ существует формула $\varphi(x, y)$ с двумя новыми свободными переменными x и y такая, что для любого слова w и для любых i, j таких, что $1 \leqslant i < j \leqslant |w|$, слово w удовлетворяет $\varphi(i, j) \iff w$ удовлетворяет φ между i и j.*

Доказательство. Для атомарной формулы φ полагаем $\varphi(x, y) = \varphi$.

Если $\varphi = \neg\psi$, то $\varphi(x, y) = \neg\psi(x, y)$.

Если $\varphi = \psi_1 \vee \psi_2$, то $\varphi(x, y) = \psi_1(x, y) \vee \psi_2(x, y)$.

Если $\varphi = \exists z \psi$, то $\varphi(x, y) = \exists z((x \leqslant z) \& (z < y) \& \psi(x, y))$. □

$\exists x_1 \exists x_2 \dots \exists x_n (x_1 < x_2 < x_3 < \dots < x_n \,\&\, Q_a x_1 \,\&\, Q_a x_2 \,\&\, \dots \& Q_a x_n)$.

Точечная иерархия.

Языки уровня 1 — кусочно-тестируемые (нет перемен кванторов).

Языки уровня k — это булева комбинация языков вида $L_1 a_1 L_2 a_2 \dots L_n a_n L_{n+1}$, где $a_1, a_2, \dots, a_n \in \Sigma$, а $L_1, L, \dots, L_{n+1}$ – языки уровня $\leqslant k - 1$. Это в точности языки, задаваемые формулами первого порядка с $k - 1$ переменой кванторов.

$FO(x = y + 1) \subsetneq FO(<)$.

Теорема (Боке, Пэн)**.** *Язык принадлежит к классу $FO(x = y + 1)$ тогда и только тогда, когда его синтаксический моноид M апериодический и для любых e, e', s, s', s'' из образа языка в синтаксическом моноиде, где e и e' – идемпотенты $ese's'es'' = es''e's'ese'$.*