Лекция 1

1.1 Теорема Клини

Предполагается, что слушатель знаком с основами теории формальных языков и конечных автоматов в пределах стандартного университетского курса дискретной математики. В частности, предполагается известной теорема Клини о том, что класс языков над данным конечным алфавитом Σ, распознаваемых конечными детерминированными автоматами, совпадает с классом *рациональных* языков над Σ, т. е. с наименьшим классом языков, который

а) содержит пустой язык и все языки вида $\{a\}$, где $a \in \Sigma$;

б) вместе с любым языком L содержит его *итерацию* L^*, т. е. множество всевозможных конечных произведений слов из L (включая пустое произведение, которое считается равным пустому слову 1);

в) вместе с любыми двумя языками L и K содержит их теоретико-множественное объединение $L \cup K$ и их *произведение* LK, т. е. множество всевозможных произведений слов из L на слова из K.

Мы будем считать известным задание рациональных языков *регулярными выражениями* и будем пользоваться такими заданиями. Например, выражение $(ab)^*$ задает итерацию одноэлементного языка $\{ab\}$.

1.2 Беззвездные языки

Среди операций, используемых в определении рациональных языков, итерация является наиболее «сложной», так как фактически описывает некоторый бесконечный процесс:

$$L^* = \{1\} \cup L \cup L^2 \cup L^3 \cup \cdots \cup L^n \cup \dots .$$

Действительно ли она необходима? Ясно, что просто удалить итерацию из определения рациональных языков нельзя, поскольку все остальные операции не могут произвести бесконечный язык из конечных языков. Но, может быть, можно заменить итерацию какой-нибудь более простой операцией, которая тем не менее может произвести бесконечный язык из конечных? Например, таким свойством обладает операция взятия *дополнения*. Напомним,

что из теоремы Клини вытекает, что класс рациональных языков замкнут относительно взятия дополнений. Дадим соответствующее определение.

Определение 1.1. Класс *беззвездных* (star-free) языков над данным конечным алфавитом Σ – это наименьший класс языков, который

а') содержит пустой язык, язык $\{1\}$ и все языки вида $\{a\}$, где $a \in \Sigma$;

б') вместе с любым языком L содержит его дополнение L^C;

в) вместе с любыми двумя языками содержит их объединение и их произведение.

Вопрос, который мы обсуждали выше, можно теперь сформулировать так: верно ли, что любой рациональных язык является беззвездным? Ответ на этот вопрос отрицателен – есть языки, которые не являются беззвездными. В качестве примера можно привести языки $(a^2)^*$ и $\{aba, b\}^*$. Мы докажем это позже, после того, как разовьем соответствующую технику.

Естественным образом возникает *проблема беззвездности*: как по данному языку над конечным алфавитом узнать, является ли он беззвездным. Эту проблему решил в 1966 г. Шютценберже[1]. Отметим, что проблема беззвездности далеко не тривиальна: если язык задан каким-то регулярным выражением, явно использующим итерацию *, это еще не означает, что язык не является беззвездным.

Пример 1.1. Рациональный язык $(ab)^*$ над алфавитом $\Sigma = \{a, b\}$ на самом деле является беззвездным. Действительно, несложно проверить, что

$$(ab)^* = (\varnothing^C a \cup b\varnothing^C \cup \varnothing^C a^2 \varnothing^C \cup \varnothing^C b^2 \varnothing^C)^C.$$

В самом деле, с учетом того, что $\varnothing^C = \Sigma^*$, выражение в правой части описывает в точности множество всех слов, которые

- не оканчиваются на a;
- не начинаются с b;
- не содержат двух вхождений буквы a подряд;
- не содержат двух вхождений буквы b подряд.

Ясно, что это множество состоит из пустого слова и всевозможных слов, которые начинаются с a, оканчиваются на b и в которых вхождения букв a и b чередуются. Но это в точности описание множества $(ab)^*$.

Упражнение 1.1. *Доказать, что языки* $\{ab, ba\}^*$ *и* $(a(ab)^*b)^*$ *являются беззвездными.*

[1]Marcel-Paul (Marco) Schützenberger (1920–1996) – французский математик, сделавший существенный вклад в развитие компьютерных наук. Свою первую научную степень он получил по медицине в 1948 г., его диссертация была отмечена призом Французской академии медицины. Вторую диссертацию по теории информации Шютценберже защитил в 1953 г. Математические интересы Шютценберже были очень широки и включали теорию автоматов, теорию формальных языков, теорию информации. Также его можно по праву считать одним из основоположников комбинаторики слов, которая начала бурно развиваться с выходом в 1983 году одноименной книги, написанной под псевдонимом M. Лотэр (M. Lothaire) Шютценберже в соавторстве с учениками.

1.3 Кусочно тестируемые языки

Определение 1.2. Язык над данным конечным алфавитом Σ называется *кусочно тестируемым*, если он может быть получен с помощью конечного числа операций объединения, пересечения и дополнения из языков вида $\Sigma^* a_1 \Sigma^* a_2 \Sigma^* \dots \Sigma^* a_k \Sigma^*$, где $a_i \in \Sigma$.

Можно определить класс кусочно тестируемых языков и с помощью соответствующих распознавателей – так называемых *автоматов-гидр*. (Напомним, что гидрой в греческой мифологии называлось многоглавое чудовище, см. рис. 1.1.)

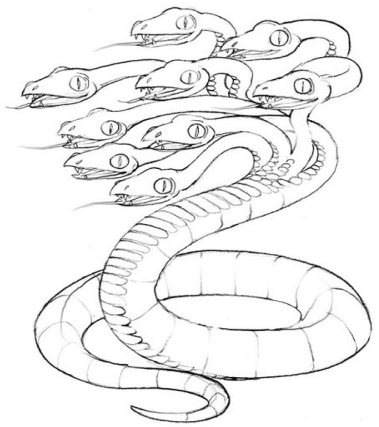

Рис. 1.1: Девятиглавая гидра

Автомат-гидра с h головками – это устройство, в состав которого входят:

- потенциально бесконечная лента, разделенная на ячейки, в которых могут быть вписаны буквы некоторого конечного алфавита Σ;
- h читающих головок, которые могут передвигаться вдоль ленты независимо друг от друга, но с сохранением взаимного порядка (первая головка всегда остается самой левой и т. д.), причем каждая головка может считывать символ из обозреваемой ей ячейки;
- конечной read-only памяти, которая содержит два списка слов длины $\leq h$ над Σ: список паролей и список запретов.

Автомат-гидра *принимает* слово $w \in \Sigma^*$, если он находит в w один из паролей и при этом не обнаруживает в w ни одного из запретов. В противном случае он *отвергает* w. Например, автомат, изображенный на рис. 1.2, принимает слово, написанное на ленте (AmpleUglyElkByRumba), поскольку находит в этом слове пароль (algebra), но не находит в нем запрещенного слова (erotica). Язык $L \subseteq \Sigma^*$ *распознается* автоматом-гидрой, если данный автомат принимает в точности те слова, которые принадлежат L. Несложно

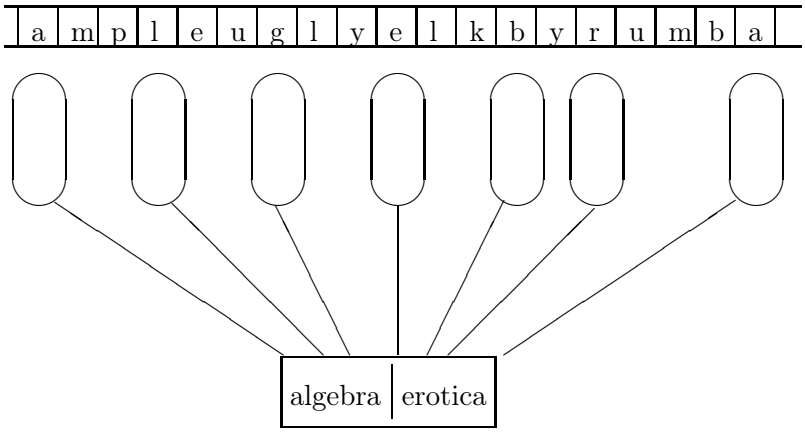

Рис. 1.2: Автомат-гидра с семью головками

понять, что класс языков, распознаваемых автоматами-гидрами, совпадает с классом кусочно тестируемых языков.

Пример 1.2. Язык $\Sigma^* ab \Sigma^*$ является кусочно тестируемым тогда и только тогда, когда $\Sigma = \{a, b\}$.

Утверждение «тогда» понятно, так как $\Sigma^* ab \Sigma^* = \Sigma^* a \Sigma^* b \Sigma^*$ – если в слове от букв a и b есть вхождение a, предшествующее вхождению b, то есть и вхождение a, непосредственно предшествующее вхождению b. А вот утверждение «только тогда» мы пока доказать не можем.

Возникает *проблема кусочной тестируемости*: как по данному языку над конечным алфавитом узнать, является ли он кусочно тестируемым. Эту проблему решил в 1972 г. Саймон[2].

1.4 Алгебраический метод

Результаты Шютценберже и Саймона объединяет одно важное обстоятельство: в обоих случаях ответ был получен в терминах алгебры, а точнее — теории полугрупп. А именно, с каждым формальным языком L некоторым каноническим образом связан алгебраический объект $M(L)$, именуемый *синтаксическим моноидом* этого языка. Если язык L рационален, то его синтаксический моноид $M(L)$ конечен и его можно эффективно вычислить по явному заданию L конечным автоматом или регулярным выражением. Оказалось, что такие свойства языка, как беззвездность и кусочная тестируемость, «транслируются» при соответствии $L \mapsto M(L)$ в некоторые

[2]Imre Simon (1943–2009) – бразильский математик и информатик венгерского происхождения. Проблему кусочной тестируемости Саймон решил в своей диссертации, выполненной под руководством Януша Бжозовского. Дальнейшие его исследования были связаны в основном с комбинаторикой слов и теорией автоматов. В связи с некоторыми вопросами этих областей Саймон ввел в рассмотрение множество действительных чисел с операциями $x \oplus y = \min\{x, y\}$ и $x \otimes y = x + y$. Этот объект получил название *тропическое полукольцо*, «тропическое» – в честь места, где жил автор этого понятия.

свойства моноида $M(L)$, причем интересно то, что соответствующие беззвездности и кусочной тестируемости свойства весьма естественны с точки зрения абстрактной теории полугрупп и изучались в ней задолго до того, как выяснилось их значение для теории формальных языков[3]. Для каждого конечного моноида наличие или отсутствие этих свойств можно установить с помощью некоторых несложных вычислений; отсюда получаются основанные на алгебре алгоритмы для распознавания беззвездности и кусочной тестируемости.

1.5 Полугруппы

> Я приветствую полугруппу,
> где бы я ее ни встретил,
> а встречается она повсюду.
> Впрочем, от друзей я слышал,
> что в математике попадаются
> объекты, отличные
> от полугрупп.
>
> Эйнар Хилле [3]

Полугруппой называется непустое множество S, на котором определена бинарная операция, удовлетворяющая закону ассоциативности:

$$\forall a, b, c \in S\colon (ab)c = a(bc). \tag{1.1}$$

В записи закона ассоциативности (1.1) операция «обозначена» подразумеваемой точкой, т.е. так, как обычно обозначается умножение. Мы будем называть эту операцию умножением.

Поскольку никаких других требований, помимо (1.1), вся теория полугрупп состоит из следствий закона ассоциативности! Довольно удивительно, что на столь простой основе удалось развить столь содержательную теорию, богатую красивыми результатами и полезными приложениями.

Единица полугруппы S – это такой элемент 1, что $a1 = 1a = a$ для любого элемента $a \in S$. Для полугруппы S через S^1 будем обозначать полугруппу S с единицей, возможно присоединенной. Это значит, что $S^1 := S$, если в S есть единица; в противном случае, $S^1 := S \cup \{1\}$ и умножение на S^1 продолжает умножение на S так, чтобы «свежий» символ 1 играл роль единицы. Полугруппу с единицей называют *моноидом*.

[3]Здесь проявляется та удивительная закономерность, которую Вигнер [2] назвал «непостижимой эффективностью математики» и которую Бурбаки [1] сформулировали следующим образом: «... математика представляется скоплением абстрактных форм, причем определенные аспекты реальности как будто бы в результате предопределения укладываются в некоторые из этих форм».

1.6 Отношения Грина

Отношениями Грина[4] называются следующие бинарные отношения на произвольной полугруппе S:

1. $a\mathscr{R}b \Leftrightarrow aS^1 = bS^1$. Это означает, что $\exists\ u, v \in S^1\colon a = bu,\ b = av$, т.е. элементы a и b делят друг друга справа (aS^1 – *главный правый идеал*, порожденный элементом a).

2. $a\mathscr{L}b \Leftrightarrow S^1a = S^1b$. Это означает, что $\exists\ x, y \in S^1\colon a = xb,\ b = ya$, т.е. элементы a и b делят друг друга слева (aS^1 – *главный левый идеал*, порожденный элементом a).

3. $a\mathscr{H}b \Leftrightarrow aS^1 = bS^1,\ S^1a = S^1b$, т.е. $\mathscr{H} = \mathscr{R} \cap \mathscr{L}$.

4. $a\,\mathscr{J}\,b \Leftrightarrow S^1aS^1 = S^1bS^1$. Это означает, что $\exists\ u, v, x, y \in S^1\colon a = ubv,\ b = xay$ (S^1aS^1 – *главный идеал*, порожденный элементом a).

Напомним, что *отношение эквивалентности* – это рефлексивное, симметричное и транзитивное бинарное отношение.

Упражнение 1.2. *Отношения Грина являются отношениями эквивалентности.*

Предпорядок – это рефлексивное и транзитивное бинарное отношение. Бывают полезны связанные с отношениями Грина предпорядки:

1. $a \leq_{\mathscr{R}} b \Leftrightarrow aS^1 \subseteq bS^1$.

2. $a \leq_{\mathscr{L}} b \Leftrightarrow S^1a \subseteq S^1b$.

3. $a \leq_{\mathscr{H}} b \Leftrightarrow aS^1 \subseteq bS^1,\ S^1a \subseteq S^1b$.

4. $a \leq_{\mathscr{J}} b \Leftrightarrow S^1aS^1 \subseteq S^1bS^1$.

Бинарное отношение ρ на полугруппе S называется *стабильным справа* (*слева*), если $a\ \rho\ b \Rightarrow ac\ \rho\ bc$ (соответственно $a\ \rho\ b \Rightarrow ca\ \rho\ cb$) для любых $a, b, c \in S$.

Предложение 1.1. *Отношения $\leq_{\mathscr{L}}$ и $\mathscr{L}$ стабильны справа, а отношения $\leq_{\mathscr{R}}$ и $\mathscr{R}$ – стабильны слева.*

Доказательство. $a \leq_{\mathscr{L}} b \Rightarrow a = ub$ для некоторого $u \in S^1$. Умножим на c справа: $ac = ubc \Rightarrow ac \leq_{\mathscr{L}} bc$. Все остальные утверждения проверяются аналогично. □

Если α и β бинарные отношения, то

$$\alpha\beta := \{(x, y) \mid \exists z\colon (x, z) \in \alpha, (z, y) \in \beta\}.$$

[4] James Alexander (Sandy) Green (1926–2014) – английский математик. С 18 лет вместе с другими юными талантами Великобритании работал в Блетчли-парке над взломом немецких шифров. Отношения, позднее названные его именем, ввел и изучил в 1951 г. в своей диссертации, выполненной под руководством Филиппа Холла и Дэвида Рисса.

Предложение 1.2. $\mathscr{L}\mathscr{R} = \mathscr{R}\mathscr{L}$.

Доказательство. Пусть $a \mathscr{L}\mathscr{R} b$. Тогда существует элемент $c \in S$ такой, что $a \mathscr{L} c$ и $c \mathscr{R} b$, т.е. $\exists\, u, v, x, y \in S^1 \colon a = uc,\ c = va,\ c = bx,\ b = cy$. Через d обозначим $ay = ucy = ub$. Покажем, что $a \mathscr{R} d$ и $d \mathscr{L} b$. Действительно, в силу предложения 1.1 $a \mathscr{L} c \Rightarrow ay \mathscr{L} cy \Rightarrow d \mathscr{L} b$. Аналогично, в силу предложения 1.1 $c \mathscr{R} b \Rightarrow uc \mathscr{R} ub \Rightarrow a \mathscr{R} d$. Получили, что $a \mathscr{R}\mathscr{L} b$, т.е. $\mathscr{L}\mathscr{R} \subseteq \mathscr{R}\mathscr{L}$. Аналогично получаем обратное включение. Таким образом, $\mathscr{L}\mathscr{R} = \mathscr{R}\mathscr{L}$. □

Следствие 1.1. *Отношение $\mathscr{D} := \mathscr{L}\mathscr{R}$ является наименьшим отношением эквивалентности, содержащим $\mathscr{L}$ и $\mathscr{R}$ одновременно.*

Доказательство. Ясно, что $\mathscr{L} \subseteq \mathscr{D}$ и $\mathscr{R} \subseteq \mathscr{D}$. Покажем, что $\mathscr{D}$ является отношением эквивалентности:

1. Рефлексивность – очевидно.
2. Симметричность – сразу следует из того, что $\mathscr{L}\mathscr{R} = \mathscr{R}\mathscr{L}$.
3. Транзитивность – пусть $a \mathscr{D} b$ и $b \mathscr{D} c$, тогда $a \mathscr{L} x \mathscr{R} b \mathscr{L} y \mathscr{R} c$ для некоторых x и y. Отсюда $x \mathscr{R}\mathscr{L} y \Rightarrow x \mathscr{L}\mathscr{R} y$, т.е. $x \mathscr{L} z \mathscr{R} y$ для некоторого z. Следовательно, $a \mathscr{L} z$ и $z \mathscr{R} c$, откуда $a \mathscr{L}\mathscr{R} c$.

Так как любое отношение эквивалентности, содержащее $\mathscr{L}$ и $\mathscr{R}$, содержит $\mathscr{L}\mathscr{R}$, заключаем, что $\mathscr{D}$ – наименьшее отношение эквивалентности с этим свойством. □

Таким образом, включения между отношениями Грина на произвольной полугруппе описываются следующей диаграммой:

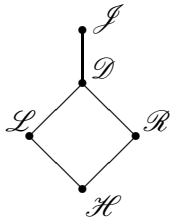

Рис. 1.3: Включения между отношениями Грина

В конкретных полугруппах некоторые (или даже все) из отношений Грина могут совпадать. Например, так происходит, если умножение коммутативно.

Упражнение 1.3. *В любой группе G все отношения Грина совпадают с универсальным отношением $G \times G$.*

Литература

[1] Н. Бурбаки. Очерки по истории математики. М.: Мир., 1965.

[2] Е. Вигнер. Непостижимая эффективность математики в естественных науках. Успехи физических наук. 1968. Т.94, №3. С.535–546.

[3] Э. Хилле. Функциональный анализ и полугруппы. М.: Изд. иностранной литературы, 1951.