

Реализация алгоритма построения представления группы по машине Тьюринга

Автор: Шамрай Максим Борисович

Научный руководитель: доцент, к. ф.-м.н. Григорьев С. В.

Рецензент: ведущий инженер ООО "Ланит-Терком" Смирнов К. К.

JetBrains Research, Programming Languages and Tools Lab
Санкт-Петербургский государственный университет
Системное программирование

25.04.2020

- Кроме иерархии Хомского, есть довольно много классов формальных языков, например, конъюнктивные и Булевы
- Не у всех есть критерий непредставимости языка в классе ($Conj \subseteq Bool$?)
- В последнее время все чаще прибегают к смежным дисциплинам для исследования языков
- Предлагается построить представление группы по языку, чтобы в дальнейшем можно было применять аппарат теории групп для исследований

Представление группы

Пусть Σ — конечный алфавит,
 $\Sigma^{-1} = \{a^{-1} \mid a \in \Sigma, aa^{-1} = a^{-1}a = 1_G\}$, тогда

- Σ^+ — свободная полугруппа
- Σ^* — свободный моноид
- $(\Sigma \cup \Sigma^{-1})^*$ — свободная группа

Представление группы

Пусть Σ — конечный алфавит,
 $\Sigma^{-1} = \{a^{-1} \mid a \in \Sigma, aa^{-1} = a^{-1}a = 1_G\}$, тогда

- Σ^+ — свободная полугруппа
- Σ^* — свободный моноид
- $(\Sigma \cup \Sigma^{-1})^*$ — свободная группа

$G = \langle A \mid R \rangle$ — представление группы

- $G = \langle a, b \mid a^3, b^2, (ab)^2 \rangle = \{\varepsilon, a, a^2, b, ab, a^2b\} = S_3$
- $G = \langle a \mid a^5 \rangle = Z_5$
- $G = \langle a, b \mid aba^{-1}b^{-1} \rangle$

- Представления групп описывают языки, которые могут быть заданы следующим выражением:
$$L(G) = \{\omega = 1_G \mid \omega \in (A \cup A^{-1})^*\}$$
- Построение представления группы по машине Тьюринга, которая распознает некоторый язык, было описано в статье¹

Теорема 1

Пусть $L \subseteq \Sigma^+$ язык, принимаемый машиной Тьюринга M , тогда существует конечно представленная группа $G(M) = \langle A \mid R \rangle$ и инъективное отображение $K : \Sigma^+ \rightarrow (A \cup A^{-1})^+$ такое что:

$$u \in L \iff K(u) = 1_G$$

¹Mark V. Sapir, Jean-Camille Birget and Eliyahu Rips "Isoperimetric and Isodiametric Functions of Groups" (2002)

Постановка задачи

Цель: Создать инструмент, с помощью которого можно будет рассматривать формальные языки как группы

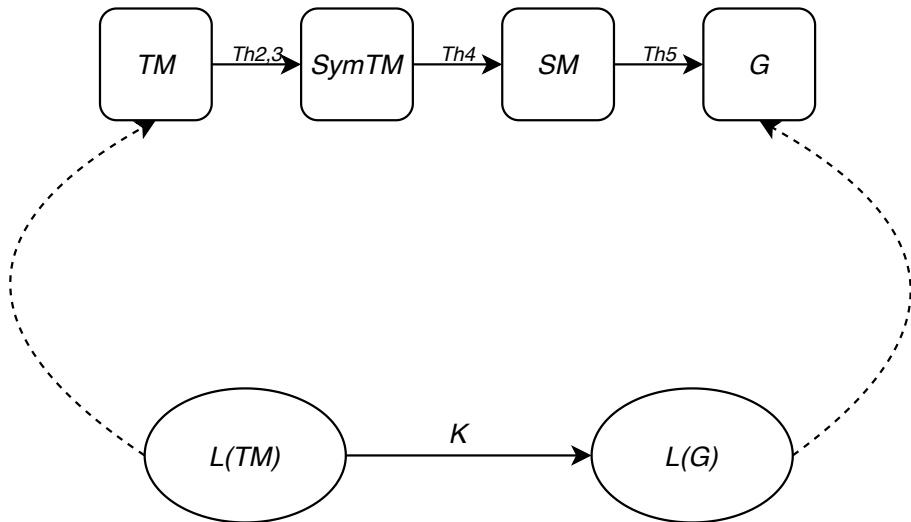
Задачи:

- 1 Реализовать алгоритм преобразования контекстно-свободной грамматики в машину Тьюринга
- 2 Реализовать алгоритм построения представления группы по машине Тьюринга
- 3 Разработать интерпретаторы промежуточных машин для проверки сохранения языка
- 4 Провести эксперименты

- Реализация магазинного автомата в терминах машины Тьюринга для построения машины Тьюринга из контекстно-свободной грамматики
- Построение представления группы из машины Тьюринга, основываясь на доказательствах теорем в статье², полностью повторяя их в алгоритме

²Mark V. Sapir, Jean-Camille Birget and Eliyahu Rips "Isoperimetric and Isodiametric Functions of Groups" (2002)

Схема построения представления группы



Симметризация недетерминированной машины Тьюринга

Теорема 2

Для любой машины Тьюринга M существует недетерминированная машина Тьюринга M' со следующими свойствами:

- M' симметричная
- Распознает тот же язык, что и M
- Каждая команда действует только на одной ленте
- Для сохранения языка добавляется лента, алфавитом которой являются команды машины Тьюринга
- Получившаяся симметричная машина Тьюринга может иметь много состояний и состоять из многих команд, что говорит о ее сложности
- Но при этом ее можно построить и для детерминированной, и для недетерминированной исходной машины Тьюринга

Симметризация детерминированной машины Тьюринга

Теорема 3

Для любой детерминированной машины Тьюринга M существует эквивалентная симметричная машина Тьюринга, полученная из M добавлениями команд τ^{-1} для каждой команды τ из M .

- Авторы теоремы E. Post и A. A. Markov (1947)
- Получившаяся симметричная машина Тьюринга гораздо легче машины, полученной по предыдущей теореме
- Но ее можно построить только для детерминированной машины Тьюринга

Построение представления группы

S-машина — система переписывания символов на ленте, которая поддерживает обратный алфавит.

Теорема 4

Для любой машины Тьюринга M' существует S-машина, которая симулирует M'

Теорема 5

Для любой S-машины существует соответствующая конечно-представленная группа

Нотация машины Тьюринга

Машина Тьюринга имеет k лент и k голов и может быть описана как шестиместный кортеж: $M = \langle X, \Gamma, Q, \Theta, \overline{s_1}, \overline{s_0} \rangle$, где

- X — входной алфавит.
- Γ — алфавит лент.
- $Q = \cup_{i=1}^k Q_i$ — множество состояний голов на лентах машины.
- Θ — множество команд машины.
- $\overline{s_1}$ — k -вектор начальных состояний машины.
- $\overline{s_0}$ — k -вектор конечных состояний машины.

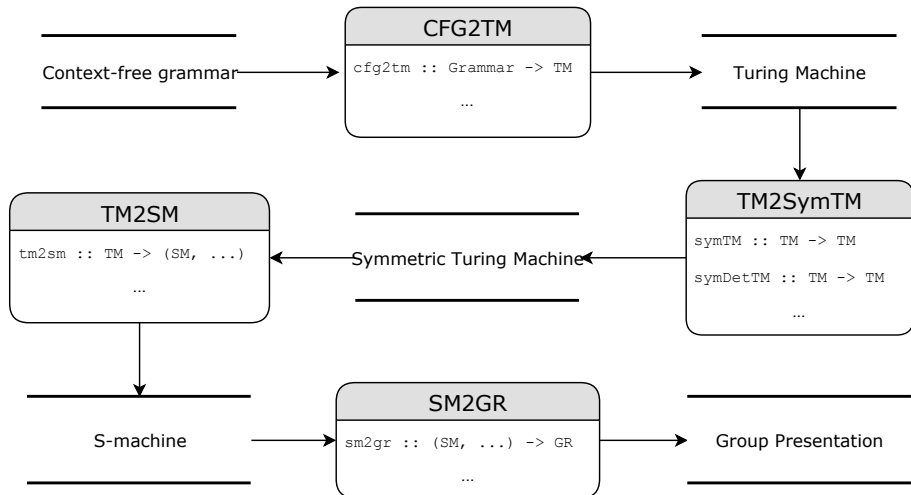
Команда одноленточной машины Тьюринга имеет вид:

$$uqv \rightarrow u'q'v'$$

где u, v, u', v' — ячейки, q, q' — состояния

- Необходимо проверять, сохраняется ли язык после каждого шага преобразования
- Так как авторы статьи используют свою эквивалентную нотацию машины Тьюринга, нами были разработаны интерпретаторы и S-машины, и машины Тьюринга
 - ▶ Дерево вычислений, корень — стартовая конфигурация
 - ▶ Обход дерева в ширину
 - ▶ В интерпретаторе S-машины используется множество пройденных конфигураций и построение дерева определенной высоты с выводом в DOT (graph description language)

Архитектура решения



Для оценки размера получившегося представления группы, запустили алгоритмы с недетерминированной и детерминированной симметризацией на трех грамматиках:

- one rule: $S \rightarrow a$

$$S \rightarrow AS$$

- a^* : $S \rightarrow \epsilon$
 $A \rightarrow a$

- Dyck:
 $S \rightarrow AC$
 $S \rightarrow \epsilon$
 $C \rightarrow SD$
 $D \rightarrow BS$
 $A \rightarrow a$
 $B \rightarrow b$

В таблице приведены мощности множеств

	Grammar			TM			
	Σ	N	R	X	Γ	Q	Θ
1 rule	1	1	1	1	3	6	5
a^*	1	2	3	1	4	8	10
Dyck	2	4	6	2	8	13	21

TM'				SM			G	
X	Γ	Q	Θ	Y	Q	Θ	A	R
1	16	318	248	16	119530	2888	121079	74401
1	28	595	476	28	402906	6386	406565	240613
2	58	1227	984	58	1711128	16858	1722169	1138799

³Используя алгоритм симметризации недетерминированных машин Тьюринга

В таблице приведены мощности множеств

	Grammar			TM			
	Σ	N	R	X	Γ	Q	Θ
1 rule	1	1	1	1	3	6	5
a^*	1	2	3	1	4	8	10
Dyck	2	4	6	2	8	13	21

TM'				SM			G	
X	Γ	Q	Θ	Y	Q	Θ	A	R
1	6	42	38	6	6232	503	6586	7673
1	8	76	76	8	16230	1026	16909	17697
2	16	167	166	16	69542	2841	71564	71645

⁴Используя алгоритм симметризации детерминированных машин Тьюринга

- Можно заметить, что размер получившихся представлений групп при симметризации в детерминированном случае гораздо меньше (для грамматики языка Дика более миллиона порождающих и отношений против 70 тысяч)
- А чем меньше получившееся представление группы, тем проще потом сказать о словах в группе, которые равны единице
- Из этого можно сделать вывод, что в случае детерминированных грамматик, нужно использовать симметризацию для детерминированных машин

- ➊ Реализован алгоритм преобразования контекстно-свободной грамматики в машину Тьюринга
- ➋ Реализован алгоритм построения представления группы по машине Тьюринга
- ➌ Разработаны интерпретаторы промежуточных машин для проверки корректности преобразований
- ➍ Проведен ряд экспериментов

Проблема слов

$$G = \langle A \mid R \rangle, \Sigma = A \cup A^{-1}$$

$$\phi : \Sigma^* \rightarrow G$$

$$W(G) = \phi^{-1}(1)$$

- $W(G)$ – регулярна $\iff G$ – конечна (Anisimov)
- $W(G)$ – контекстно-свободна $\iff \exists H < G$ – конечного индекса (Muller–Schupp)

