



# От синтаксического анализа графов к системам матричных уравнений

Юлия Сусанина

JetBrains Research, Programming Languages and Tools Lab  
Санкт-Петербургский Государственный Университет

14.12.2019

# Context-Free Path Querying (CFPQ)

- КС-грамматика  $G = (N, \Sigma, R)$   
 $\mathcal{L}(G_S) = \{\omega \mid S \Rightarrow_G^* \omega\}, S \in N$

# Context-Free Path Querying (CFPQ)

- КС-грамматика  $G = (N, \Sigma, R)$   
 $\mathcal{L}(G_S) = \{\omega \mid S \Rightarrow_G^* \omega\}, S \in N$
- Орграф  $D = (V, E, \sigma), \sigma \subseteq \Sigma, E \subseteq V \times \sigma \times V$   
 $m\lambda n$  — путь из  $m$  в  $n$  в графе  $D$ ,  $\lambda$  — слово данного пути

# Context-Free Path Querying (CFPQ)

- КС-грамматика  $G = (N, \Sigma, R)$   
 $\mathcal{L}(G_S) = \{\omega \mid S \Rightarrow_G^* \omega\}, S \in N$
- Орграф  $D = (V, E, \sigma), \sigma \subseteq \Sigma, E \subseteq V \times \sigma \times V$   
 $m\lambda n$  — путь из  $m$  в  $n$  в графе  $D$ ,  $\lambda$  — слово данного пути
- $R_A = \{(m, n) \mid m\lambda n \text{ — путь в } D, \lambda \in \mathcal{L}(G_A)\}$

# Context-Free Path Querying (CFPQ)

- КС-грамматика  $G = (N, \Sigma, R)$   
 $\mathcal{L}(G_S) = \{\omega \mid S \Rightarrow_G^* \omega\}, S \in N$
- Орграф  $D = (V, E, \sigma), \sigma \subseteq \Sigma, E \subseteq V \times \sigma \times V$   
 $m\lambda n$  — путь из  $m$  в  $n$  в графе  $D$ ,  $\lambda$  — слово данного пути
- $R_A = \{(m, n) \mid m\lambda n \text{ — путь в } D, \lambda \in \mathcal{L}(G_A)\}$

- Использование достижений линейной алгебры
- Приближенные методы для ускорения вычислений
- Развитие и постоянное улучшение библиотек и программных пакетов, связанное с развитием искусственного интеллекта
- Возможность применения параллельных вычислений

- Создание подхода, основанного на методах линейной алгебры и вычислительной математики, к задаче CFPQ

# Сведение CFPQ к решению матричных уравнений (1)

$$S \rightarrow aSb \mid ab$$

$$T_E \in \mathbb{M}^{|V| \times |V|} : (T_E)_{ij} = 1 \iff (i, j) \in R_E \quad \forall E \in (N \cup \Sigma)$$



# Сведение CFPQ к решению матричных уравнений (1)

$$S \rightarrow aSb \mid ab$$

$$T_E \in \mathbb{M}^{|V| \times |V|} : (T_E)_{ij} = 1 \iff (i, j) \in R_E \quad \forall E \in (N \cup \Sigma)$$

$\Downarrow$

$$\{T_S^k\} : \begin{aligned} T_S^0 &= \mathbf{0} \\ T_S^{k+1} &= T_a T_S^k T_b + T_a T_b \end{aligned}$$

$T_S^\infty$  – наименьшее решение  $T_S = T_a T_S T_b + T_a T_b$

## Сведение CFPQ к решению матричных уравнений (2)

$\Downarrow$

$$\{\mathcal{T}_S^k\} : \begin{aligned} \mathcal{T}_S^0 &= \mathbf{0} \\ \mathcal{T}_S^{k+1} &= \epsilon(T_a \mathcal{T}_S^k T_b + T_a T_b) \end{aligned}$$

$\mathcal{T}_S^\infty$  – наименьшее решение  $\mathcal{T}_S = \epsilon(T_a \mathcal{T}_S T_b + T_a T_b)$ ,  
где  $\epsilon$  т.ч.  $\mathcal{T}_S^k \leq \mathbf{1} \quad \forall k$

## Сведение CFPQ к решению матричных уравнений (2)

$\Downarrow$

$$\{\mathcal{T}_S^k\} : \begin{aligned} \mathcal{T}_S^0 &= \mathbf{0} \\ \mathcal{T}_S^{k+1} &= \epsilon(T_a \mathcal{T}_S^k T_b + T_a T_b) \end{aligned}$$

$\mathcal{T}_S^\infty$  – наименьшее решение  $\mathcal{T}_S = \epsilon(T_a \mathcal{T}_S T_b + T_a T_b)$ ,  
где  $\epsilon$  т.ч.  $\mathcal{T}_S^k \leq \mathbf{1} \quad \forall k$

$$(\mathcal{T}_S^{k+1})_{ij} > 0 \iff (\mathcal{T}_S^{k+1})_{ij} = 1$$

$$\text{ceil}(\mathcal{T}_S^\infty) = T_S^\infty$$

- Линейный случай

- ▶ линейные уравнения специального типа  
(уравнения Сильвестра:  $AXB + CXD = F$ )
- ▶ линейные системы  $Ax = b$

- Нелинейный случай

- ▶ метод Ньютона

$$X = G(X) \Rightarrow F(X) = X - G(X) = \mathbf{0}$$

$$X_{i+1} = X_i - (F'(X_i))^{-1}F(X_i) \iff \begin{cases} F'(X_i)H_i = -F(X_i) \\ X_{i+1} = X_i + H_i \end{cases}$$

- *SciPy*
  - ▶ sSLV — решение разреженной системы линейных уравнений
  - ▶ dNWT — нахождение корней уравнения, метод Ньютона
- Результаты сравнения с матричным алгоритмом (в мс)

Ontology	V	dNWT	sSLV	dGPU	sCPU	sGPU
bio-meas	341	284	35	276	91	24
people-pets	337	73	49	144	38	6
funding	778	502	184	1246	344	27
wine	733	791	171	722	179	6
pizza	671	334	161	943	256	23

- **ACM SIGMOD 2020 Student Research Competition:**  
Yuliya Susanina. Context-Free Path Querying via Matrix Equations.  
(принят во второй тур соревнования)

(+) *Результаты работы в первом полугодии*

- **CIBB 2019:**  
Yuliya Susanina, Anna Yaveyn and Semyon Grigorev. Modification of Valiant's Parsing Algorithm for String-Searching Problem.  
(доклад, ожидается публикация)
- **Журнал «Труды ИСП РАН»:**  
Сусанина Ю.А., Григорьев С.В. Модификация алгоритма Валианта для задачи поиска подстрок.  
(ождается публикация)

## Дальнейшие планы

- Эффективная реализация предложенного подхода с использованием специализированных библиотек и параллельных вычислений
- Определение подклассов полиномиальных уравнений, решение которых может быть сведено к CFPQ
- Попытка построить взаимное сведение между CFPQ и решением соответствующих подклассов уравнений
- Планируется публикация результатов на GRADES-NDA 2020