Условные нижние оценки для задачи поиска путей с контекстно-свободными ограничениями

Истомина Александра

Научный руководитель: Григорьев Семён Вячеславович

Санкт-Петербургский государственный университет

16 июня 2023 г.

Определения

Контекстно-свободная грамматика — это четверка $G = (N, \Sigma, P, S)$, где:

- N — конечное множество нетерминалов
- Σ — конечный алфавит
- P — конечное множество правил вывода
- $S \in N$ — стартовый нетерминал

Шаг вывода — это операция $\alpha A \beta \Rightarrow \alpha \gamma \beta$, где $\alpha, \gamma, \beta \in (\Sigma \cup N)^*$, а $A \to \gamma \in P$.

$\mathcal{L}(G) = \{w \in \Sigma^* | S \Rightarrow \ldots \Rightarrow w\}$ — язык, задаваемый грамматикой.

Пример (Язык Дика на k типах скобок, Dyck-k)

- $N = \{S\}$
- $\Sigma = \{(_i,)_i\}, \forall i = 1, \ldots, k$
- *Правила вывода:* $S \to \epsilon | SS | (_1 S)_1 | \ldots | (_k S)_k$, *где* ϵ *— это пустая строка.*

Определения

Определение (Задача контекстно-свободной достижимости, CFL reachability)

Дано: *ориентированный граф* $D = (V, E, L)$ *с пометками на ребрах* $L \subseteq \Sigma$, *КС грамматика* $G = (N, \Sigma, P, S)$.
Надо: *понять, есть ли путь между вершинами, такой что пометки на ребрах образуют слово из* $\mathcal{L}(G)$.
Варианты: all-pairs/s-t

Грамматика обычно предполагается фиксированной.
Сложность алгоритма считается от $n = |V|$ — количества вершин в графе.

История

Известные алгоритмы:

[Melski, Reps 1997] $\mathcal{O}(n^3)$

[Chaudhuri 2008] $\mathcal{O}(n^3/\log n)$

Вопрос: Существует ли алгоритм со временем работы $\mathcal{O}(n^{3-\epsilon}), \epsilon > 0$?

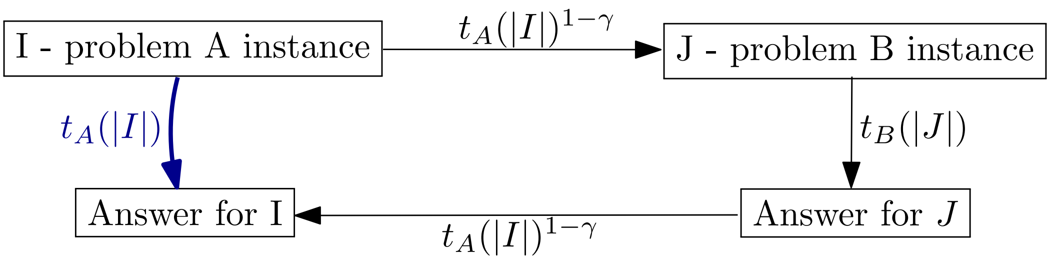
Способы решения:

- Придумать более быстрый алгоритм
- Найти нижнюю оценку на сложность

Условные нижние оценки

Условная нижняя оценка — верна, если верна некоторая гипотеза
Получаются с помощью сведений от одной задаче к другой.

Сведение (A, t_A) $\rightarrow$ (B, t_B):



Цель: найти задачу X с гипотезой об ее временной сложности и свести ее к задаче CFL reachability.

Популярные задачи и гипотезы

Гипотеза (BMM)

Перемножить две булевы матрицы размера $n \times n$ *нельзя комбинаторным алгоритмом за время* $\mathcal{O}(n^{3-\epsilon}), \epsilon > 0$.

Гипотеза (SAT)

(N)SETH - определить выполнимость булевой формулы нельзя (ко-недетерминированно) за время $\mathcal{O}(2^{(1-\epsilon)n}), \epsilon > 0$.

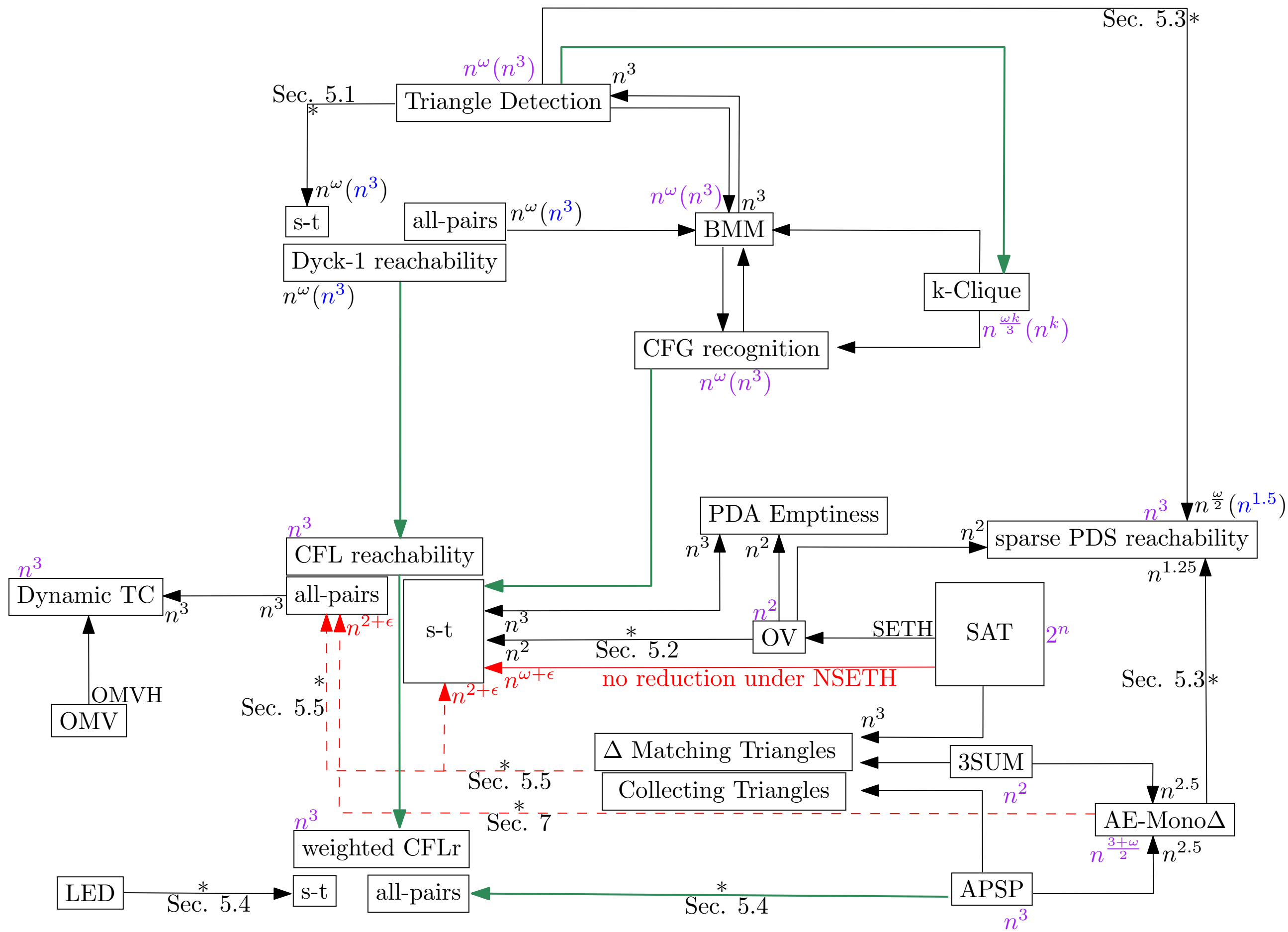
Гипотеза (APSP)

Найти наименьшие расстояния между всеми парами вершин во взвешенном графе нельзя за время $\mathcal{O}(n^{3-\epsilon}), \epsilon > 0$.

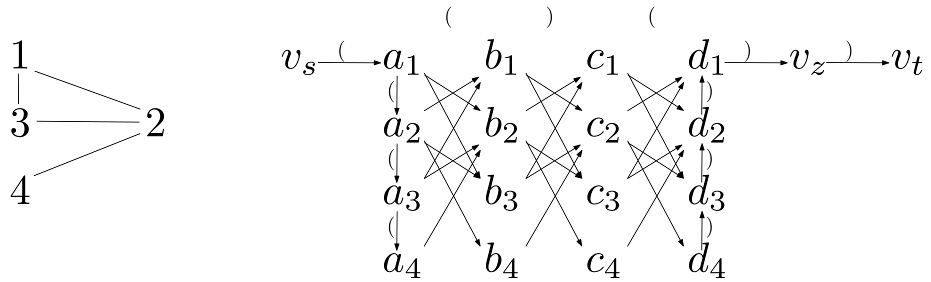
Гипотеза (3SUM)

Определить, есть ли в массиве тройка чисел, суммирующаяся в 0, нельзя за время $\mathcal{O}(n^{2-\epsilon}), \epsilon > 0$.

Карта сведений



- (Triangle Detection, n^ω) $\to$ (Dyck-1 reachability, n^ω)



- (OV, n^2) $\to$ (s-t CFL reachability, n^2)

X

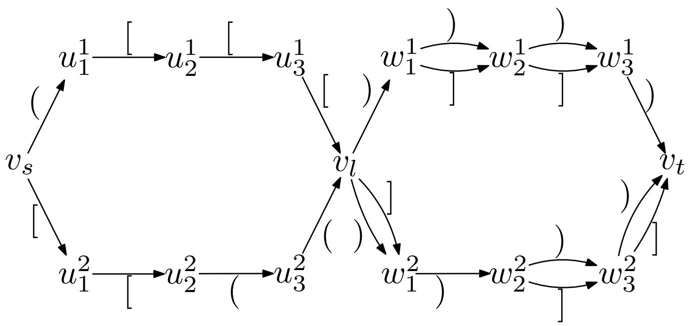
$(0,1,1,1)$

$(1,1,0,0)$

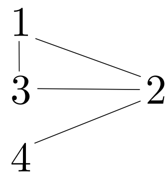
Y

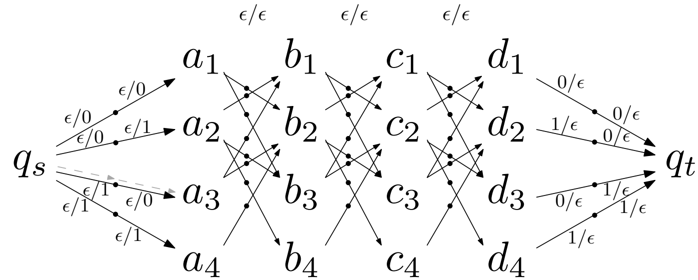
$(1,0,0,1)$

$(0,0,1,0)$

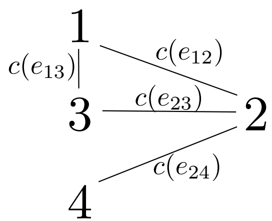


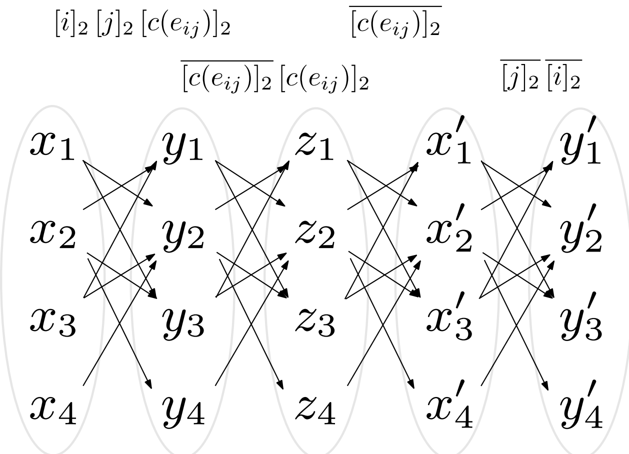
- (Triangle Detection, n^3) $\to$ (sparse PDS reachability with stack bound $\lceil \log n \rceil$, $n^{1.5}$)





- (AE-MonoΔ, $n^{2.5}$) $\to$ (sparse PDS reachability with stack bound $4\lceil \log n \rceil$, $n^{1.25}$)

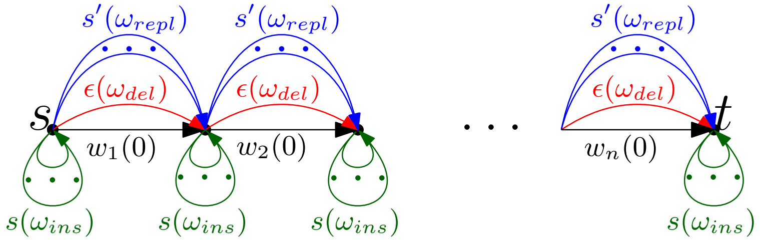




Результаты: Сведения

- (LED, n^c) $\rightarrow$ (weighted s-t CFL reachability, n^c), $c > 1$

$w_1 w_2 \ldots w_n$



Следствие

Для задачи weighted s-t CFL reachability не существует алгоритма со временем работы $\mathcal{O}(n^{3-\epsilon}), \epsilon > 0$, если APSP гипотеза верна.

Результаты: Потенциальные сведения

Теорема

Если существует сведение (Collecting Triangles/Δ Matching Triangles, n^3) $\to$ (CFL reachability, n^c), $c > 2$, то опровергается хотя бы одно из следующих утверждений:

- $\omega = 2$
- *NSETH*
- *Существует сведение (all-pairs CFL reachability, n^c) $\to$ (s-t CFL reachability, $n^{c'}$), $c' > 2$.*

Результаты: Ограниченные пути

Теорема (Schepper, 2018)

Существует алгоритм для задачи s-t CFL reachability на ациклических графах со временем работы $\mathcal{O}(n^{\omega})$.

Следствие

Существует алгоритм для задачи all-pairs CFL reachability на ациклических графах со временем работы $\mathcal{O}(n^{\omega})$.

Следствие

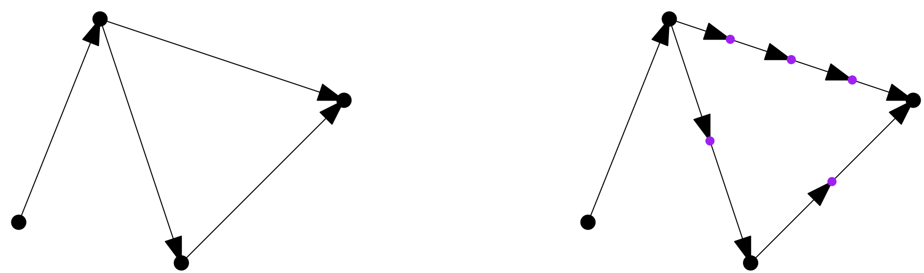
Существует алгоритм для задачи all-pairs CFL reachability со временем работы $\tilde{\mathcal{O}}(n^{c}), c < 3$, *находящий все* $\mathcal{L}$*-пути длины не более чем* $k = \tilde{\mathcal{O}}(n^{c'})$, *где* $c' < \frac{3}{\omega} - 1$.

Результаты: Техника длинных ребер

Случай

Входной граф G это подразбиение графа G'. В задаче all-pairs CFL reachability нас интересуют только исходные вершины рграфа G'.

Длинное ребро e' графа G — это множество ребер, полученных делением одного ребра G'
Длина e' $=$ количество делений $+ 1$



G' $\qquad\qquad$ G

Хотим записать на ребре G', например, цвет, или номер вершины, или номер ребра, и т.д.

Результаты: Техника длинных ребер

Теорема

Пусть есть алгоритм для задачи all-pairs CFL reachability с временной сложностью $\mathcal{O}(n^c \cdot poly(|G|)) = \mathcal{O}(n^c)$. *Тогда существует алгоритм для задачи all-pairs CFL reachability на графах с длинными ребрами длины* $k = o(\log n)$ *с временной сложностью* $\mathcal{O}(n^{c+c'}), \forall c' > 0$.

Теорема

Если аналогичная теорема верна для графов с ребрами длины $k = \Omega(\log n)$, *то верно хотя бы одно из следующих утверждений:*

- *В сведении используется граф, структура которого "сильно отличается" от входного графа*
- *Обе гипотезы APSP и 3SUM неверны.*

Итог

- Получены новые сведения к задаче CFL reachability и близким к ней задачам.
- Доказана лучшая оценка на сложность задачи CFL reachability для случая путей ограниченной длины.
- Рассмотрена техника длинных ребер: ее возможности и ограничения.

Спасибо за внимание!