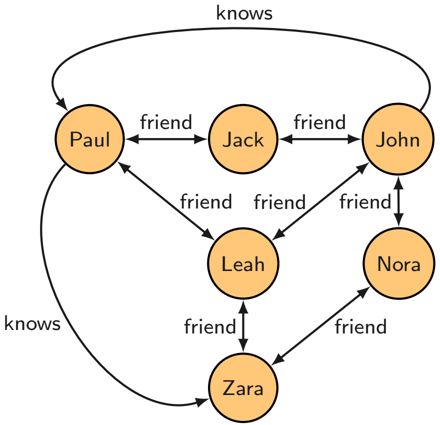
A Polynomial-Time Approximation Algorithm for Counting Words Accepted by an NFA

Marcelo Arenas

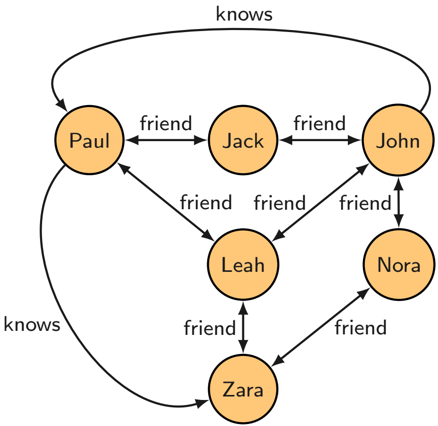
PUC & IMFD Chile

Joint work with Luis Alberto Croquevielle, Rajesh Jayaram and Cristian Riveros

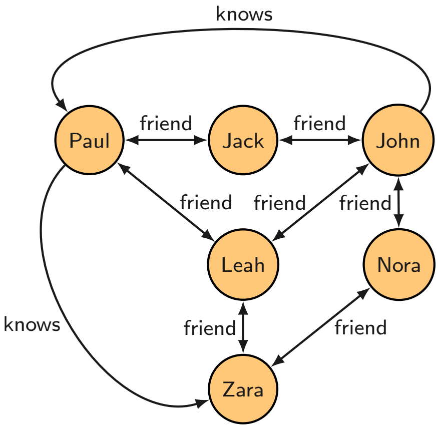
A graph database *G*

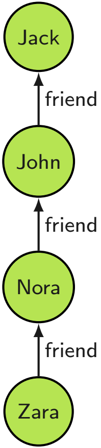


A query over G: (friend + knows)*

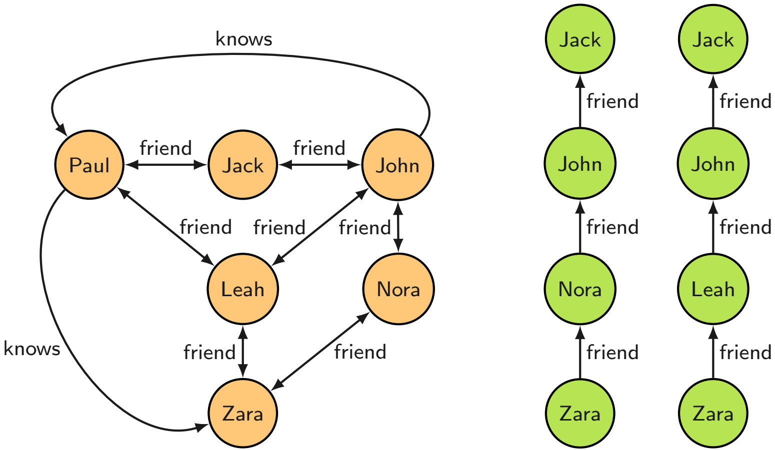


A query over G: (friend + knows)*





A query over G: $(\text{friend} + \text{knows})^*$



The length n of paths as a parameter

Two fundamental problems:

- COUNT(G, r, n): count the number of paths p in G such that p conforms to regular expression r and the length of p is n
  - n is given in unary as 0^n
- GEN(G, r, n): generate uniformly at random a path p in G such that p conforms to r and the length of p is n

COUNT is #P-complete

Only approximate solutions are possible

- Best known approximations work in quasi-polynomial time

COUNT is #P-complete

Only approximate solutions are possible

- Best known approximations work in quasi-polynomial time

Our goal is to construct an FPRAS $\mathcal{B}$ for COUNT

- For every G, r, n and error $\varepsilon \in (0,1)$:

$$\mathbf{Pr}\left(\left| \frac{\text{COUNT}(G,r,n) - \mathcal{B}(G,r,n,\varepsilon)}{\text{COUNT}(G,r,n)} \right| \leq \varepsilon \right) \quad \geq \quad \frac{3}{4}$$

- $\mathcal{B}$ works in time poly$(\|G\|, \|r\|, n, \frac{1}{\varepsilon})$

COUNT can be reduced to the following problem

Input	:	An NFA A, a length n given in unary and $\varepsilon \in (0, 1)$
Output	:	Number of words w such that $w \in \mathcal{L}(\mathcal{A})$ and $|w| = n$

COUNT can be reduced to the following problem

Input	:	An NFA A, a length n given in unary and $\varepsilon \in (0,1)$
Output	:	Number of words w such that $w \in \mathcal{L}(\mathcal{A})$ and $|w| = n$

$A = (Q, \{0,1\}, \Delta, I, F)$

- Q is a finite set of states
- $\Delta \subseteq Q \times \{0,1\} \times Q$ is the transition relation
- $I \subseteq Q$ is a set of initial states
- $F \subseteq Q$ is a set of final (accepting) states

The problem to solve

Assuming $\mathcal{L}_n(A) = \mathcal{L}(A) \cap \{0,1\}^n$

The task is to compute a number N that is a $(1 \pm \varepsilon)$-approximation of $|\mathcal{L}_n(A)|$:

$$(1-\varepsilon)|\mathcal{L}_n(A)| \ \leq \ N \ \leq \ (1+\varepsilon)|\mathcal{L}_n(A)|$$

Besides, number N has to be computed in time poly$(m, n, \frac{1}{\varepsilon})$ with $m = |Q|$

First component: unroll automaton A

Construct A_{unroll} from A:

- for each state $q \in Q$, include copies q^0, q^1, ..., q^n in A_{unroll}
- for each transition $(p, a, q) \in \Delta$ and $i \in \{0, 1, \ldots, n-1\}$, include transition (p^i, a, q^{i+1}) in A_{unroll}

Besides, eliminate from A_{unroll} unnecessary states: each state q^i is reachable from an initial state p^0 ($p \in I$)

Second component: a sketch to be used in the estimation

Define $\mathcal{L}(q^i)$ as the set of strings w such that there is a path from an initial state p^0 to q^i labeled with w

- Notice that $|w| = i$

Besides, define for every $X \subseteq Q$:

$$\mathcal{L}(X^i) = \bigcup_{q \in X} \mathcal{L}(q^i)$$

Second component: a sketch to be used in the estimation

Define $\mathcal{L}(q^i)$ as the set of strings w such that there is a path from an initial state p^0 to q^i labeled with w

- Notice that $|w| = i$

Besides, define for every $X \subseteq Q$:

$$\mathcal{L}(X^i) \;=\; \bigcup_{q \in X} \mathcal{L}(q^i)$$

Then the task is to compute an estimation of $|\mathcal{L}(F^n)|$

Second component: a sketch to be used in the estimation

Let $\kappa = \lceil \frac{nm}{\varepsilon} \rceil$

Second component: a sketch to be used in the estimation

Let $\kappa = \lceil \frac{nm}{\varepsilon} \rceil$

We maintain for each state q^i:

- $N(q^i)$: a $(1 \pm \kappa^{-2})^i$-approximation of $|\mathcal{L}(q^i)|$
- $S(q^i)$: a multiset of uniform samples from $\mathcal{L}(q^i)$ of size $2\kappa^7$

Data structure to be inductively computed:

$$\text{sketch}[i] \quad = \quad \{N(q^j), S(q^j) \mid 0 \leq j \leq i \text{ and } q \in Q\}$$

The algorithm template

1. Construct A_{unroll} from A
2. For each state $q \in I$, set $N(q^0) = |\mathcal{L}(q^0)| = 1$ and $S(q^0) = \mathcal{L}(q^0) = \{\lambda\}$
3. For each $i = 1, \ldots, n$ and state $q \in Q$:
   (a) Compute $N(q^i)$ given sketch[$i-1$]
   (b) Sample polynomially many uniform elements from $\mathcal{L}(q^i)$ using $N(q^i)$ and sketch[$i-1$], and let $S(q^i)$ be the multiset of uniform samples obtained
4. Return an estimation of $|\mathcal{L}(F^n)|$ given sketch[n]

Computing an estimation $N(F^n)$ of $|\mathcal{L}(F^n)|$

We use notation $N(X^i)$ for an estimation $|\mathcal{L}(X^i)|$

- Such an estimation is not only needed in the last step of the algorithm, but also in the inductive construction of sketch[i]

Notice that $|\mathcal{L}(X^i)| = \sum_{p \in X} |\mathcal{L}(p^i)|$ does not hold in general

Computing an estimation $N(F^n)$ of $|\mathcal{L}(F^n)|$

We use notation $N(X^i)$ for an estimation $|\mathcal{L}(X^i)|$

- Such an estimation is not only needed in the last step of the algorithm, but also in the inductive construction of sketch[i]

Notice that $|\mathcal{L}(X^i)| = \sum_{p \in X} |\mathcal{L}(p^i)|$ does not hold in general

But the following holds, given a linear order $<$ on Q:

$$|\mathcal{L}(X^i)| \;=\; \sum_{p \in X} \Big|\mathcal{L}(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i)\Big|$$

Computing an estimation $N(X^i)$ of $|\mathcal{L}(X^i)|$

We have that:

$$|\mathcal{L}(X^i)| \quad = \quad \sum_{p \in X} \left| \mathcal{L}(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i) \right|$$

Computing an estimation $N(X^i)$ of $|\mathcal{L}(X^i)|$

We have that:

$$
\begin{aligned}
|\mathcal{L}(X^i)| &= \sum_{p \in X} \left| \mathcal{L}(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i) \right| \\
&= \sum_{p \in X} \left| \mathcal{L}(p^i) \right| \frac{\left| \mathcal{L}(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i) \right|}{\left| \mathcal{L}(p^i) \right|}
\end{aligned}
$$

Computing an estimation $N(X^i)$ of $|\mathcal{L}(X^i)|$

We have that:

$$
\begin{aligned}
|\mathcal{L}(X^i)| &= \sum_{p \in X} \left| \mathcal{L}(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i) \right| \\
&= \sum_{p \in X} \left| \mathcal{L}(p^i) \right| \frac{\left| \mathcal{L}(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i) \right|}{\left| \mathcal{L}(p^i) \right|}
\end{aligned}
$$

So we will use the following approximation:

$$
N(X^i) = \sum_{p \in X} N(p^i) \frac{\left| S(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i) \right|}{\left| S(p^i) \right|}
$$

Computing an estimation $N(X^i)$ of $|\mathcal{L}(X^i)|$

$N(X^i)$ can be computed in polynomial time in the size of sketch[i]

- $S(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i)$ is constructed by checking for each $w \in S(p^i)$ whether w is not in $\mathcal{L}(q^i)$ for every $q \in X$ with $q < p$

Computing an estimation $N(X^i)$ of $|\mathcal{L}(X^i)|$

$N(X^i)$ can be computed in polynomial time in the size of sketch$[i]$

- $S(p^i) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^i)$ is constructed by checking for each $w \in S(p^i)$ whether w is not in $\mathcal{L}(q^i)$ for every $q \in X$ with $q < p$

What guarantees that $N(X^i)$ is a good estimation of $|\mathcal{L}(X^i)|$?

The main property to maintain

$\mathcal{E}(i)$ holds if for every $p \in Q$ and $X \subseteq Q$:

$$\left| \frac{\left|\mathcal{L}(p^i) \setminus \bigcup_{q \in X} \mathcal{L}(q^i)\right|}{\left|\mathcal{L}(p^i)\right|} - \frac{\left|S(p^i) \setminus \bigcup_{q \in X} \mathcal{L}(q^i)\right|}{\left|S(p^i)\right|} \right| < \frac{1}{\kappa^3}$$

The use of the main property

Proposition

If $\mathcal{E}(i)$ holds and $N(p^i)$ is a $(1 \pm \kappa^{-2})^i$-approximation of $|\mathcal{L}(p^i)|$ for every $p \in Q$, then $N(X^i)$ is a $(1 \pm \kappa^{-2})^{i+1}$-approximation of $|\mathcal{L}(X^i)|$ for every $X \subseteq Q$

The use of the main property

Proposition

If $\mathcal{E}(i)$ holds and $N(p^i)$ is a $(1 \pm \kappa^{-2})^i$-approximation of $|\mathcal{L}(p^i)|$ for every $p \in Q$, then $N(X^i)$ is a $(1 \pm \kappa^{-2})^{i+1}$-approximation of $|\mathcal{L}(X^i)|$ for every $X \subseteq Q$

$\mathcal{E}(0)$ holds and $N(p^0)$ is a $(1 \pm \kappa^{-2})^0$-approximation of $|\mathcal{L}(p^0)|$ for every $p \in Q$

- Recall that $N(p^0) = |\mathcal{L}(p^0)|$ and $S(p^0) = \mathcal{L}(p^0)$ for every $p \in Q$

The use of the main property

Proposition

If $\mathcal{E}(i)$ holds and $N(p^i)$ is a $(1 \pm \kappa^{-2})^i$-approximation of $|\mathcal{L}(p^i)|$ for every $p \in Q$, then $N(X^i)$ is a $(1 \pm \kappa^{-2})^{i+1}$-approximation of $|\mathcal{L}(X^i)|$ for every $X \subseteq Q$

$\mathcal{E}(0)$ holds and $N(p^0)$ is a $(1 \pm \kappa^{-2})^0$-approximation of $|\mathcal{L}(p^0)|$ for every $p \in Q$

- Recall that $N(p^0) = |\mathcal{L}(p^0)|$ and $S(p^0) = \mathcal{L}(p^0)$ for every $p \in Q$

Then $N(X^0)$ is a $(1 \pm \kappa^{-2})$-approximation of $|\mathcal{L}(X^0)|$ for every $X \subseteq Q$

The use of the main property

For each state $p \in Q$ and $b = 0, 1$, define:

$$R_b(p^1) \quad = \quad \{q^0 \mid (q^0, b, p^1) \text{ is a transition in } A_{unroll}\}$$

The use of the main property

For each state $p \in Q$ and $b = 0, 1$, define:

$$R_b(p^1) \quad = \quad \{q^0 \mid (q^0, b, p^1) \text{ is a transition in } A_{unroll}\}$$

Then $\mathcal{L}(p^1) \ = \ \mathcal{L}(R_0(p^1)) \cdot \{0\} \uplus \mathcal{L}(R_1(p^1)) \cdot \{1\}$

- So that $|\mathcal{L}(p^1)| = |\mathcal{L}(R_0(p^1))| + |\mathcal{L}(R_1(p^1))|$

The use of the main property

For each state $p \in Q$ and $b = 0, 1$, define:

$$R_b(p^1) \quad = \quad \{q^0 \mid (q^0, b, p^1) \text{ is a transition in } A_{unroll}\}$$

Then $\mathcal{L}(p^1) \;=\; \mathcal{L}(R_0(p^1)) \cdot \{0\} \uplus \mathcal{L}(R_1(p^1)) \cdot \{1\}$

- So that $|\mathcal{L}(p^1)| = |\mathcal{L}(R_0(p^1))| + |\mathcal{L}(R_1(p^1))|$

Hence, given that $N(R_b(p^1))$ is a $(1 \pm \kappa^{-2})$-approximation of $|\mathcal{L}(R_b(p^1))|$ for $b = 0, 1$:

$$N(R_0(p^1)) + N(R_1(p^1)) \text{ is a } (1 \pm \kappa^{-2})\text{-approximation of } N(p^1)$$

The use of the main property: a summary

$\mathcal{E}(0)$ holds and $N(p^0)$ is a $(1 \pm \kappa^{-2})^0$-approximation of $|\mathcal{L}(p^0)|$ for every $p \in Q$

The use of the main property: a summary

$\mathcal{E}(0)$ holds and $N(p^0)$ is a $(1 \pm \kappa^{-2})^0$-approximation of $|\mathcal{L}(p^0)|$ for every $p \in Q$

$$\Downarrow$$

$N(X^0)$ is a $(1 \pm \kappa^{-2})^1$-approximation of $|\mathcal{L}(X^0)|$ for every $X \subseteq Q$

The use of the main property: a summary

$\mathcal{E}(0)$ holds and $N(p^0)$ is a $(1 \pm \kappa^{-2})^0$-approximation of $|\mathcal{L}(p^0)|$ for every $p \in Q$

$$\Downarrow$$

$N(X^0)$ is a $(1 \pm \kappa^{-2})^1$-approximation of $|\mathcal{L}(X^0)|$ for every $X \subseteq Q$

$$\Downarrow$$

$N(p^1) = N(R_0(p^1)) + N(R_1(p^1))$ is a $(1 \pm \kappa^{-2})^1$-approximation of $N(p^1)$ for every $p \in Q$

The use of the main property: a summary

$N(p^1)$ is a $(1 \pm \kappa^{-2})^1$-approximation of $|\mathcal{L}(p^1)|$ for every $p \in Q$

The use of the main property: a summary

$\mathcal{E}(1)$ holds and $N(p^1)$ is a $(1 \pm \kappa^{-2})^1$-approximation of $|\mathcal{L}(p^1)|$ for every $p \in Q$

The use of the main property: a summary

$\mathcal{E}(1)$ holds and $N(p^1)$ is a $(1 \pm \kappa^{-2})^1$-approximation of $|\mathcal{L}(p^1)|$ for every $p \in Q$

$$\Downarrow$$

$N(X^1)$ is a $(1 \pm \kappa^{-2})^2$-approximation of $|\mathcal{L}(X^1)|$ for every $X \subseteq Q$

The use of the main property: a summary

$\mathcal{E}(1)$ holds and $N(p^1)$ is a $(1 \pm \kappa^{-2})^1$-approximation of $|\mathcal{L}(p^1)|$ for every $p \in Q$

$$\Downarrow$$

$N(X^1)$ is a $(1 \pm \kappa^{-2})^2$-approximation of $|\mathcal{L}(X^1)|$ for every $X \subseteq Q$

$$\Downarrow$$

$N(p^2) = N(R_0(p^2)) + N(R_1(p^2))$ is a $(1 \pm \kappa^{-2})^2$-approximation of $N(p^2)$ for every $p \in Q$

The use of the main property: a summary

$\mathcal{E}(1)$ holds and $N(p^1)$ is a $(1 \pm \kappa^{-2})^1$-approximation of $|\mathcal{L}(p^1)|$ for every $p \in Q$

$$\Downarrow$$

$N(X^1)$ is a $(1 \pm \kappa^{-2})^2$-approximation of $|\mathcal{L}(X^1)|$ for every $X \subseteq Q$

$$\Downarrow$$

$N(p^2) = N(R_0(p^2)) + N(R_1(p^2))$ is a $(1 \pm \kappa^{-2})^2$-approximation of $N(p^2)$ for every $p \in Q$

$$\Downarrow$$

$$\ldots$$

The final result

Proposition

If $\mathcal{E}(i)$ *holds for every* $i \in \{0, 1, \ldots, n\}$, *then* $N(F^n)$ *is a* $(1 \pm \varepsilon)$*-approximation of* $|\mathcal{L}(F^n)|$

The final result

Proposition

If $\mathcal{E}(i)$ *holds for every* $i \in \{0, 1, \ldots, n\}$, *then* $N(F^n)$ *is a* $(1 \pm \varepsilon)$*-approximation of* $|\mathcal{L}(F^n)|$

The issue then is to maintain property $\mathcal{E}(i)$

- Multisets $S(q^i)$ of uniform samples play a central role on this

Sampling from a state

We need to construct the multiset $S(q^i)$ of uniform samples

Recall that:

- $S(q^i)$ contains $2\kappa^7$ words from $\mathcal{L}(q^i)$
- $S(q^i)$ is computed assuming that $N(q^i)$ and $\text{sketch}[i-1] = \{N(q^j), S(q^j) \mid 0 \leq j \leq i-1\}$ have already been constructed

To recall

1. Construct A_{unroll} from A
2. For each state $q \in I$, set $N(q^0) = |\mathcal{L}(q^0)| = 1$ and $S(q^0) = \mathcal{L}(q^0) = \{\lambda\}$
3. For each $i = 1, \dots, n$ and state $q \in Q$:
   (a) Compute $N(q^i)$ given sketch$[i-1]$
   (b) Sample polynomially many uniform elements from $\mathcal{L}(q^i)$ using $N(q^i)$ and sketch$[i-1]$, and let $S(q^i)$ be the multiset of uniform samples obtained
4. Return an estimation of $|\mathcal{L}(F^n)|$ given sketch$[n]$

Sampling from q^i

To generate a sample in $\mathcal{L}(q^i)$, we construct a sequence w^i, w^{i-1}, ..., w^1, w^0 such that

- $w^i = \lambda$
- $w^j = b_j w^{j+1}$ with $b_j \in \{0, 1\}$
- $w^0 \in \mathcal{L}(q^i)$

Sampling from q^i

To generate a sample in $\mathcal{L}(q^i)$, we construct a sequence w^i, w^{i-1}, ..., w^1, w^0 such that

- $w^i = \lambda$
- $w^j = b_j w^{j+1}$ with $b_j \in \{0, 1\}$
- $w^0 \in \mathcal{L}(q^i)$

To choose $w^{i-1} = bw^i$, construct for $b = 0, 1$:

$$P_b^i = \{p^{i-1} \mid (p^{i-1}, b, q^i) \text{ is a transition in } A_{unroll}\}$$

Sampling from q^i

P_0^i and P_1^i are sets of states at layer $i-1$

We can use the following estimations:

$$N(X^{i-1}) \quad = \quad \sum_{p \in X} N(p^{i-1}) \frac{\left| S(p^{i-1}) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^{i-1}) \right|}{\left| S(p^{i-1}) \right|}$$

Sampling from q^i

P_0^i and P_1^i are sets of states at layer $i-1$

We can use the following estimations:

$$N(X^{i-1}) \quad = \quad \sum_{p \in X} N(p^{i-1}) \frac{\left| S(p^{i-1}) \setminus \bigcup_{q \in X \,:\, q < p} \mathcal{L}(q^{i-1}) \right|}{\left| S(p^{i-1}) \right|}$$

We choose $b \in \{0, 1\}$ with probability:

$$\frac{N(P_b^i)}{N(P_0^i) + N(P_1^i)}$$

We could have started from a set of states

The previous procedure works for every set of states P^i:

$$P^i_b \quad = \quad \{p^{i-1} \mid \exists r^i \in P^i : (p^{i-1}, b, r^i) \text{ is a transition in } A_{unroll}\}$$

In particular, we applied the procedure for $P^i = \{q^i\}$

We could have started from a set of states

The previous procedure works for every set of states P^i:

$$P^i_b \quad = \quad \{p^{i-1} \mid \exists r^i \in P^i : (p^{i-1}, b, r^i) \text{ is a transition in } A_{unroll}\}$$

In particular, we applied the procedure for $P^i = \{q^i\}$

The following recursive procedure summarizes the previous idea:

$$\textbf{Sample}(i, \{q^i\}, \lambda, \varphi_0)$$

It uses sets of states $P^i = \{q^i\}$, P^{i-1}, ..., P^1, P^0 and an initial probability φ_0

The sampling algorithm

Sample(j, P^j, w^j, φ)

1. If $j = 0$, then with probability φ return w^0, otherwise return **fail**
2. Compute $P_b^j = \{p^{j-1} \mid \exists r^j \in P^j : (p^{j-1}, b, r^j)$ is a transition in $A_{unroll}\}$ for $b = 0, 1$
3. Choose $b \in \{0, 1\}$ with probability $p_b = \dfrac{N(P_b^j)}{N(P_0^j) + N(P_1^j)}$
4. Set $P^{j-1} = P_b^j$ and $w^{j-1} = bw^j$
5. Return **Sample**$(j - 1, P^{j-1}, w^{j-1}, \frac{\varphi}{p_b})$

The key observation

Let $x = x_1 \cdots x_i$ be word in $\mathcal{L}(q^i)$

The key observation

Let $x = x_1 \cdots x_i$ be word in $\mathcal{L}(q^i)$

We have that:

$$
\begin{aligned}
&\mathbf{Pr}(\text{the output of } \textbf{Sample} \text{ is } x)\\
&\quad = \mathbf{Pr}(w^0 = x \wedge \text{ the last call to } \textbf{Sample} \text{ does not fail})\\
&\quad = \mathbf{Pr}(\text{the last call to } \textbf{Sample} \text{ does not fail} \mid w^0 = x) \cdot \mathbf{Pr}(w^0 = x)\\
&\quad = \left(\left(\prod_{j=1}^{i} \frac{N(P^j_{x_j})}{N(P^j_0) + N(P^j_1)} \right)^{-1} \cdot \varphi_0 \right) \cdot \left(\prod_{j=1}^{i} \frac{N(P^j_{x_j})}{N(P^j_0) + N(P^j_1)} \right)\\
&\quad = \varphi_0
\end{aligned}
$$

The value of the initial probability φ_0

Proposition

Assume that $\mathcal{E}(j)$ holds for each $j < i$. If w is the output of **Sample**$(i, \{q^i\}, \lambda, \frac{e^{-5}}{N(q^i)})$*, then*

- $\varphi \in (0,1)$ *in every recursive call to* **Sample**
- $\mathbf{Pr}(w = \mathbf{fail}) \leq 1 - e^{-9}$
- $\mathbf{Pr}(w = x) = \dfrac{e^{-5}}{N(q^i)}$ *for every* $x \in \mathcal{L}(q^i)$

The value of the initial probability φ_0

Proposition

Assume that $\mathcal{E}(j)$ holds for each $j < i$. If w is the output of **Sample**$(i, \{q^i\}, \lambda, \frac{e^{-5}}{N(q^i)})$, *then*

- $\varphi \in (0, 1)$ *in every recursive call to* **Sample**
- $\mathbf{Pr}(w = \mathbf{fail}) \leq 1 - e^{-9}$
- $\mathbf{Pr}(w = x) = \dfrac{e^{-5}}{N(q^i)}$ *for every* $x \in \mathcal{L}(q^i)$

Hence, conditioned on not failing, **Sample**$(i, \{q^i\}, \lambda, \frac{e^{-5}}{N(q^i)})$ returns a uniform sample from $\mathcal{L}(q^i)$

The last step: bounding the probability of breaking the main assumption

Recall that $\mathcal{E}(i)$ holds if for every $q \in Q$ and $X \subseteq Q$:

$$\left| \frac{\left|\mathcal{L}(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)\right|}{\left|\mathcal{L}(q^i)\right|} - \frac{\left|S(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)\right|}{\left|S(q^i)\right|} \right| < \frac{1}{\kappa^3}$$

We know that $\mathcal{E}(0)$ holds.

The last step: bounding the probability of breaking the main assumption

Recall that $\mathcal{E}(i)$ holds if for every $q \in Q$ and $X \subseteq Q$:

$$\left| \frac{\left|\mathcal{L}(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)\right|}{\left|\mathcal{L}(q^i)\right|} - \frac{\left|S(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)\right|}{\left|S(q^i)\right|} \right| < \frac{1}{\kappa^3}$$

We know that $\mathcal{E}(0)$ holds. We need to compute a lower bound for:

$$\mathbf{Pr}\left(\bigwedge_{j=0}^{n} \mathcal{E}(j) \right)$$

Bounding the probability of breaking $\mathcal{E}(i)$

Assume that $\bigwedge_{j=0}^{i-1} \mathcal{E}(j)$ holds

Let $q \in Q$ and $S(q^i)$ be a multiset of $2\kappa^7$ samples from $\mathcal{L}(q^i)$ computed by calling **Sample**$(i, \{q^i\}, \lambda, \frac{e^{-5}}{N(q^i)})$

- Each element of $S(q^i)$ is obtained by repeatedly calling **Sample** until the output is different from **fail**

Assume that $S(q^i) = \{w_1, \ldots, w_t\}$ with $t = 2\kappa^7$

Bounding the probability of breaking $\mathcal{E}(i)$

Let $X \subseteq Q$, and Y_i be a Bernoulli random variable for $i \in \{1, \ldots, t\}$:

$$Y_i = 1 \quad \text{if and only if} \quad w_i \in \left(\mathcal{L}(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i) \right)$$

Bounding the probability of breaking $\mathcal{E}(i)$

Let $X \subseteq Q$, and Y_i be a Bernoulli random variable for $i \in \{1, \ldots, t\}$:

$$Y_i = 1 \quad \text{if and only if} \quad w_i \in \left(\mathcal{L}(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i) \right)$$

We have that:

$$\begin{aligned}
\mathbb{E}[Y_i] &= \frac{|\mathcal{L}(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)|}{|\mathcal{L}(q^i)|} \\
\sum_{j=1}^{t} Y_i &= |S(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)| \\
t &= |S(q^i)|
\end{aligned}$$

By using Hoeffding's inequality

$$\mathbf{Pr}\left(\left|\frac{|S(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)|}{|S(q^i)|} - \frac{|\mathcal{L}(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)|}{|\mathcal{L}(q^i)|}\right| \geq \frac{1}{\kappa^3} \;\middle|\; \bigwedge_{j=0}^{i-1} \mathcal{E}(j)\right) \leq 2e^{-4\kappa}$$

By using Hoeffding's inequality

$$\mathbf{Pr}\left(\left| \frac{|S(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)|}{|S(q^i)|} - \frac{|\mathcal{L}(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)|}{|\mathcal{L}(q^i)|} \right| \geq \frac{1}{\kappa^3} \;\middle|\; \bigwedge_{j=0}^{i-1} \mathcal{E}(j) \right) \leq 2e^{-4\kappa}$$

By taking the union bound:

$$\mathbf{Pr}\left(\exists q \in Q \, \exists X \subseteq Q \left| \frac{|S(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)|}{|S(q^i)|} - \frac{|\mathcal{L}(q^i) \setminus \bigcup_{p \in X} \mathcal{L}(p^i)|}{|\mathcal{L}(q^i)|} \right| \geq \frac{1}{\kappa^3} \;\middle|\; \bigwedge_{j=0}^{i-1} \mathcal{E}(j) \right) \leq 2e^{-2\kappa}$$

The conclusion

Rewriting the previous result:

$$\mathbf{Pr}\left(\mathcal{E}(i) \;\middle|\; \bigwedge_{j=0}^{i-1} \mathcal{E}(j)\right) \;\geq\; 1 - e^{-2\kappa}$$

We conclude that:

$$\mathbf{Pr}\left(\bigwedge_{j=0}^{n} \mathcal{E}(j)\right) \;\geq\; 1 - e^{-\kappa}$$

The complete algorithm

Input: NFA $A = (Q, \{0,1\}, \Delta, I, F)$ with $m = |Q|$, length n given in unary and error $\varepsilon \in (0,1)$

The complete algorithm

Input: NFA $A = (Q, \{0,1\}, \Delta, I, F)$ with $m = |Q|$, length n given in unary and error $\varepsilon \in (0,1)$

1. If $\mathcal{L}_n(A) = \emptyset$, then return 0

The complete algorithm

Input: NFA $A = (Q, \{0,1\}, \Delta, I, F)$ with $m = |Q|$, length n given in unary and error $\varepsilon \in (0,1)$

1. If $\mathcal{L}_n(A) = \emptyset$, then return 0
2. Construct A_{unroll} and set $\kappa = \lceil \frac{nm}{\varepsilon} \rceil$

The complete algorithm

Input: NFA $A = (Q, \{0,1\}, \Delta, I, F)$ with $m = |Q|$, length n given in unary and error $\varepsilon \in (0,1)$

1. If $\mathcal{L}_n(A) = \emptyset$, then return 0
2. Construct A_{unroll} and set $\kappa = \lceil \frac{nm}{\varepsilon} \rceil$
3. Remove each state q^i from A_{unroll} that is not reachable from an initial state in I^0

The complete algorithm

Input: NFA $A = (Q, \{0,1\}, \Delta, I, F)$ with $m = |Q|$, length n given in unary and error $\varepsilon \in (0,1)$

1. If $\mathcal{L}_n(A) = \emptyset$, then return 0
2. Construct A_{unroll} and set $\kappa = \lceil \frac{nm}{\varepsilon} \rceil$
3. Remove each state q^i from A_{unroll} that is not reachable from an initial state in I^0
4. For each $q^0 \in I^0$, set $N(q^0) = 1$ and $S(q^0) = \{\lambda\}$

The complete algorithm

5. For each layer $i = 1, \ldots, n$ and state q^i in A_{unroll}:

The complete algorithm

5. For each layer $i = 1, \ldots, n$ and state q^i in A_{unroll}:

   5.1 Set $R_b = \{p^{i-1} \mid (p^{i-1}, b, q^i) \text{ is a transition in } A_{unroll}\}$ for $b = 0, 1$

The complete algorithm

5. For each layer $i = 1, \ldots, n$ and state q^i in A_{unroll}:

   5.1 Set $R_b = \{p^{i-1} \mid (p^{i-1}, b, q^i) \text{ is a transition in } A_{unroll}\}$ for $b = 0, 1$

   5.2 Set $N(q^i) = N(R_0) + N(R_1)$

The complete algorithm

5. For each layer $i = 1, \ldots, n$ and state q^i in A_{unroll}:

   5.1 Set $R_b = \{p^{i-1} \mid (p^{i-1}, b, q^i) \text{ is a transition in } A_{unroll}\}$ for $b = 0, 1$

   5.2 Set $N(q^i) = N(R_0) + N(R_1)$

   5.3 Set $S(q^i) = \emptyset$. Then while $|S(q^i)| < 2\kappa^7$:

The complete algorithm

5. For each layer $i = 1, \ldots, n$ and state q^i in A_{unroll}:

   5.1 Set $R_b = \{p^{i-1} \mid (p^{i-1}, b, q^i)$ is a transition in $A_{unroll}\}$ for $b = 0, 1$

   5.2 Set $N(q^i) = N(R_0) + N(R_1)$

   5.3 Set $S(q^i) = \emptyset$. Then while $|S(q^i)| < 2\kappa^7$:

      5.3.1 Run **Sample**$(i, \{q^i\}, \lambda, \frac{e^{-5}}{N(q^i)})$ until it returns $w \neq$ **fail**, and at most $c(\kappa) \in \Theta(\log(\kappa))$ times

The complete algorithm

5. For each layer $i = 1, \dots, n$ and state q^i in A_{unroll}:
   5.1 Set $R_b = \{p^{i-1} \mid (p^{i-1}, b, q^i) \text{ is a transition in } A_{unroll}\}$ for $b = 0, 1$
   5.2 Set $N(q^i) = N(R_0) + N(R_1)$
   5.3 Set $S(q^i) = \emptyset$. Then while $|S(q^i)| < 2\kappa^7$:
      5.3.1 Run **Sample**$(i, \{q^i\}, \lambda, \frac{e^{-5}}{N(q^i)})$ until it returns $w \neq$ **fail**, and at most $c(\kappa) \in \Theta(\log(\kappa))$ times
      5.3.2 If $w =$ **fail**, then return 0 (failure event)

The complete algorithm

5. For each layer $i = 1, \ldots, n$ and state q^i in A_{unroll}:

   5.1 Set $R_b = \{p^{i-1} \mid (p^{i-1}, b, q^i) \text{ is a transition in } A_{unroll}\}$ for $b = 0, 1$

   5.2 Set $N(q^i) = N(R_0) + N(R_1)$

   5.3 Set $S(q^i) = \emptyset$. Then while $|S(q^i)| < 2\kappa^7$:

      5.3.1 Run **Sample**$(i, \{q^i\}, \lambda, \frac{e^{-5}}{N(q^i)})$ until it returns $w \neq$ **fail**, and at most $c(\kappa) \in \Theta(\log(\kappa))$ times

      5.3.2 If $w =$ **fail**, then return 0 (failure event)

      5.3.3 Set $S(q^i) = S(q^i) \cup \{w\}$ (recall that $S(q^i)$ allows duplicates)

The complete algorithm

5. For each layer $i = 1, \ldots, n$ and state q^i in A_{unroll}:

   5.1 Set $R_b = \{p^{i-1} \mid (p^{i-1}, b, q^i) \text{ is a transition in } A_{unroll}\}$ for $b = 0, 1$

   5.2 Set $N(q^i) = N(R_0) + N(R_1)$

   5.3 Set $S(q^i) = \emptyset$. Then while $|S(q^i)| < 2\kappa^7$:

      5.3.1 Run **Sample**$(i, \{q^i\}, \lambda, \frac{e^{-5}}{N(q^i)})$ until it returns $w \neq$ **fail**, and at most $c(\kappa) \in \Theta(\log(\kappa))$ times

      5.3.2 If $w =$ **fail**, then return 0 (failure event)

      5.3.3 Set $S(q^i) = S(q^i) \cup \{w\}$ (recall that $S(q^i)$ allows duplicates)

6. Return $N(F^n)$ as an estimation of $|\mathcal{L}_n(A)|$

The complete algorithm: final comments

The probability that the algorithm returns a wrong estimate is at most $\frac{1}{4}$

- Considering $c(\kappa) = \lceil \frac{2+\log(4)+8\log(\kappa)}{\log(1-e^{-9})^{-1}} \rceil$

The algorithm runs in time $\text{poly}(m, n, \frac{1}{\varepsilon})$

Final remarks

- The algorithm also provides a randomized polynomial-time algorithm for GEN
  - Such an algorithm can also be obtained from [Jerrum, Valiant & Vazirani 1986]

- COUNT is SpanL-complete under parsimonious reductions. We conclude that each function in SpanL admits an FPRAS
  - SpanL is the class of functions computable as $|S|$, where S is the set of output values returned by an NL Turing machine

The complete version of the paper can be found at
https://arxiv.org/abs/1906.09226

Thanks!