Constant-delay Enumeration for Lorem Ipsum

Martín Muñoz ✉
Pontificia Universidad Católica de Chile

Cristian Riveros ✉
Pontificia Universidad Católica de Chile

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

2012 ACM Subject Classification Theory of computation → Database theory

Keywords and phrases Streams, query evaluation, enumeration algorithms.

1 Introduction

Write an introduction here.

2 Preliminaries

Documents, alphabet, and spans. We fix a finite alphabet Σ. A **document** is a finite string $d = a_1 \dots a_n \in \Sigma^*$. A **span** is a half-open interval $[i, j\rangle$ with $1 \leq i \leq j \leq |d| + 1$, it denotes the substring $d[i, j\rangle = a_i \dots a_{j-1}$. The sets of all spans of d is written $Spans(d)$. (Fagin-Document Spanners...)

Variables, markers and mappings. Let V be a finite set of **variables**. For each $x \in V$ we use two **variable markers** $[x$ open and $x\rangle$ close; $Markers_V = \{[x, x\rangle | x \in V\}$.(3361451). A **mapping** or valuation is a partial function.

$$\mu : dom(\mu) \subseteq V \to Spans(d).$$

Two mappings are **compatible** when they agree on their common variables. Mappings constitute thr basic tuples produced by our operators.

Document spanners. A document spanner P associates to every string d a finite set of mappings over some variable set $V = SVars(P)$. (Fagin-Document Spanners...) Intuitively, a spanner "extracts" all matches of a pattern as span relations.

Extended Variable Automata (eVA). REmatch compiles each REQL query to an extended variable-set automaton

$$E = (Q, q_0, F, \delta),$$

whose transitions are quadruples (q, a, S, q') with $a \in \Sigma \cup \{\#\}$ and a (possibly empty) set of markers S. While reading the $i-th$ symbol, the automaton outputs the pair (S, i); if $S = \varnothing$ nothing is produced. A run sequence

$$q_0 \xrightarrow{b_0/S_0} q_1 \xrightarrow{b_1/S_1} \dots \xrightarrow{b_n/S_n} q_{n+1}$$

that alternates variable transitions and letter transitions and respects marker nesting. The mapping defined by a valid accepting run is obtained by pairing every $[x$ with the corresponding $x\rangle$. Determinisation via a subset construction yields a deterministic eVA guaranteeing at most one accepting run per output sequence, a key invariant for output-linear enumeration. (Fagin-Document Spanners...)

Determinisation of extended VA. an eVA $E = (Q, q_0, F, \delta)$ is deterministic when

$$\delta : Q \times (\Sigma \cup (2^{Markers_V} \{\varnothing\})) \to Q$$

is a partial function: for every state q and pail (a, S) there is at most one outgoing transition (q, a, S, q'). Determinism guarantees that, for any document d and any output sequence, at most one accepting run produces it – a key property to avoid duplicates during enumeration.(paper). Every eVA con be turned into an equivalent deterministic eVA via the classical powerset method.

Variable-inclusion order. for two mappings μ, ν over the same document we write

$$\mu \preccurlyeq_{varinc} \nu \iff dom(\mu) \ and \ \forall x \in dom(\mu) : \mu(x) = \nu(x).$$

This partial order captures containment by identical spans plus extra variables and underlies the MAX selection strategy.

MAX operator. given a spanner P we define

$$[\![MAX(P)]\!]_d = \{\mu \in P(d) | \text{there is no } \nu \in P(d) \text{ with } \mu \preccurlyeq_{varinc} \nu\}$$

That is, we keep only those mappings that are maximal under $\preccurlyeq_{varinc}$. (GREZ)

3 Main results

Deterministic eVA. Let $E = (Q, q_0, F, \delta)$ be an eVA. Its pair-based subset construction yields the deterministic eVA

$$E_{\text{det}} = \big(Q_{\text{det}}, X_0, F_{\text{det}}, \Delta\big),$$

with

$$\begin{aligned} Q_{\text{det}} &= 2^Q, \\ X_0 &= \{q_0\}, \\ F_{\text{det}} &= \{X \subseteq Q \mid X \cap F \neq \varnothing\}, \\ \Delta(X, a, S) &= \{\, q' \mid \exists q \in X : \ (q, a, S, q') \in \delta\}. \end{aligned}$$

Selection strategy MAX. Given a deterministic extended variable automaton (eVA) $E_{\text{det}} = (2^Q, X_0, F_{\text{det}}, \Delta)$, we define a new eVA

$$E_{\max} = \big(Q_{\max}, (R_0, W_0), F_{\max}, \Delta_{\max}\big)$$

that accepts exactly the mappings that are maximal under variable inclusion, where $Q_{max} = 2^Q \times 2^Q$ and

$$Q_{\max} = \big\{ (R, W) \mid R, W \subseteq Q_{\text{det}} \text{ and } R \cap W = \varnothing \big\}.$$

Initial state. $(R_0, W_0) = (X_0, \varnothing)$.

Transition function. R represents a set of "current" states of E_{det} and W represents the set of states of E_{det} having a run that dominates the current run under variable inclusion. For $(R, W) \in Q_{\max}$, letter $a \in \Sigma \cup \{\#\}$, and marker set S:

$$\Delta_{\max}\big((R, W), a, S\big) \;=\; (R', W')$$

where

$$R' = \Delta(R, a, S) \setminus W' \quad \text{and} \quad W' = \begin{cases} \Delta(W, a, S) \;\cup\; \bigcup\limits_{S \neq \varnothing} \Delta(R, a, S') & \text{if } S' = \varnothing, \\ \Delta(W, a, S) \;\cup\; \bigcup\limits_{S \subset S'} \Delta(R, a, S') & \text{if } S \neq \varnothing. \end{cases}$$

Final states.

$$F_{\max} \;=\; \big\{ (R, W) \in Q_{\max} \;\big|\; R \cap F_{\text{det}} \neq \varnothing \text{ and } W \cap F_{\text{det}} = \varnothing \big\}.$$

4 Conclusions

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

References

A Proofs from Section 2

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

B Proofs of Section 3

B.1 Proof of Lemma ??

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

B.2 Proof of Theorem ??

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in

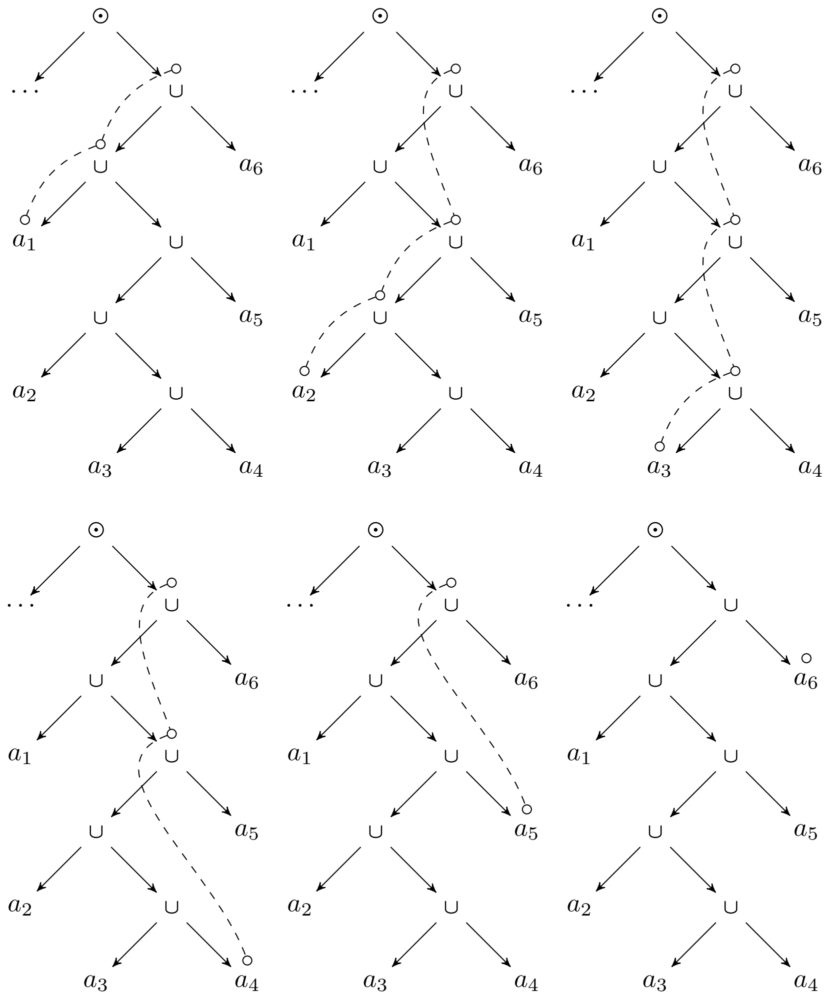

Figure 1 An example iteration of `trav` and `move`. The sequences of nodes joined by dashed lines represent a stack `St`, where the first one was obtained after calling `trav` over the topmost union node, and the following five are obtained by repeated applications of `move`(`St`).

voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco

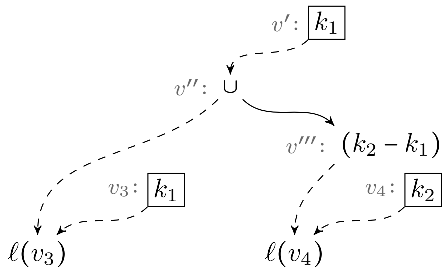

Figure 2 Gadget used in Theorem **??**.

laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

B.3 Proof of Proposition ??

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

▷ Claim 1. Fix $k \in \mathbb{N}$. Let $\mathcal{C}_k$ be the class of all duplicate-free and k-bounded D that satisfy the ϵ condition. Then one can solve the problem $\texttt{Enum}[\mathcal{C}_k]$ with output-linear delay and without preprocessing (i.e. constant preprocessing time).

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in

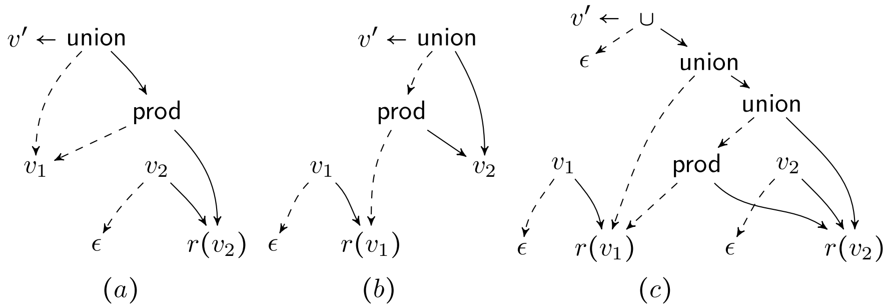

Figure 3 Gadgets for product as defined for an $\mathcal{D}$ with the ϵ-node.

voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.