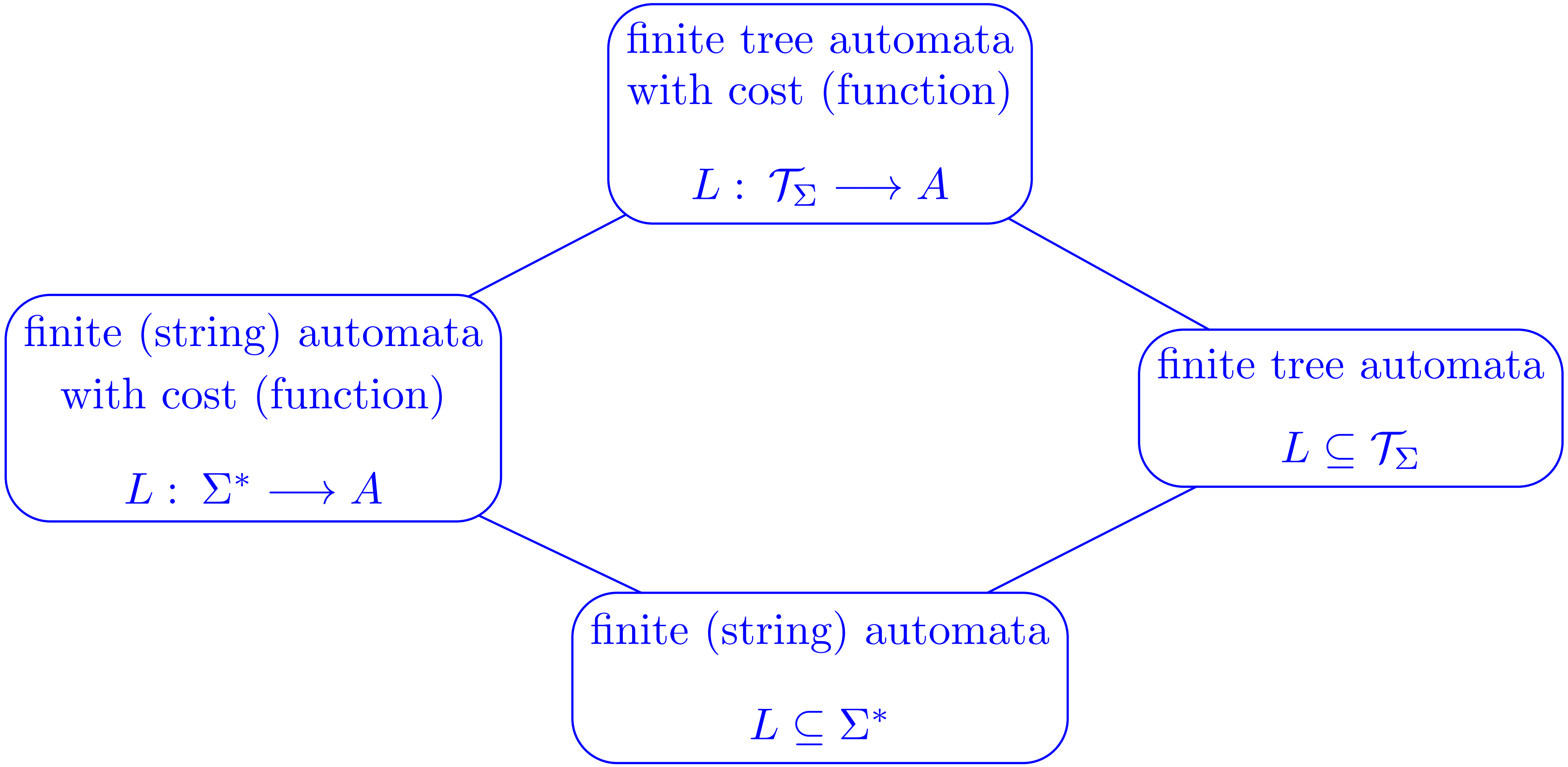
Boundedness of tree automata with polynomial cost function

Andreas Maletti

July 4, 2003

1. Introduction and Motivation
2. Tree Automata and Cost Functions
3. Monotonic Semirings
4. Boundedness Results

Generalization Hierarchy



Semirings

Definition: A *semiring* is an algebraic structure $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$, where

- A is the *carrier set*,
- $\oplus$ and $\odot$ are *associative*, i.e. $a \otimes (b \otimes c) = (a \otimes b) \otimes c$ with $\otimes \in \{\oplus, \odot\}$,
- $\oplus$ is *commutative*, i.e. $a \oplus b = b \oplus a$,
- $\mathbf{0}$ and $\mathbf{1}$ are the *unit elements* of addition and multiplication, respectively, i.e. $\mathbf{0} \oplus a = a \oplus \mathbf{0} = a$ and $\mathbf{1} \odot a = a \odot \mathbf{1} = a$,
- $\odot$ *distributes over* $\oplus$, i.e. $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$ and $(b \oplus c) \odot a = (b \odot a) \oplus (c \odot a)$ and
- $\mathbf{0}$ is *absorbing*, i.e. $\mathbf{0} \odot a = a \odot \mathbf{0} = \mathbf{0}$.

Semiring Examples

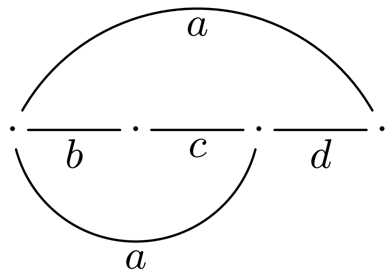
- the *semiring of natural numbers* $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$,
- the *arctic semiring* $\mathbb{A} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$,
- the *tropical semiring* $\mathbb{T} = (\mathbb{N} \cup \{+\infty\}, \min, +, +\infty, 0)$,
- the *subset semiring* $\mathbb{F} = (\mathcal{P}_f(\mathbb{N}), \cup, \boldsymbol{+}, \emptyset, \{0\})$ with

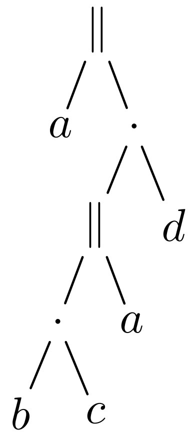
$$A \boldsymbol{+} B = \{\, a + b \mid a \in A, b \in B \,\},$$

- the *boolean semiring* $\mathbb{B} = (\{\bot . \top\}, \vee, \wedge, \bot, \top)$.

Series-Parallel Graphs

Let $\Sigma = \{\|^{(2)}, \cdot^{(2)}, a^{(0)}, b^{(0)}, c^{(0)}, d^{(0)}\}$. The term $a \parallel (((b \cdot c) \parallel a) \cdot d)$ corresponds to the graphical representations





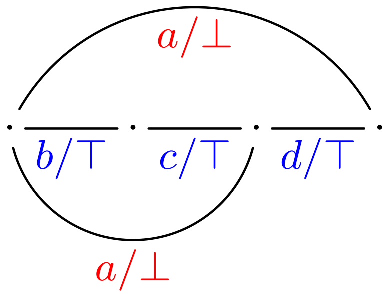
The leftmost node is called *source*, whereas the rightmost one is called *sink*. Assume we apply costs as follows:

$$c(G_1 \parallel G_2) = c(G_1) \oplus c(G_2) \qquad \text{and} \qquad c(G_1 \cdot G_2) = c(G_1) \odot c(G_2)$$

Series-Parallel Graphs (cont'd)

- in the *Boolean semiring* $\mathbb{B}$ with $c(b) = c(c) = c(d) = \top$ and $c(a) = \bot$:

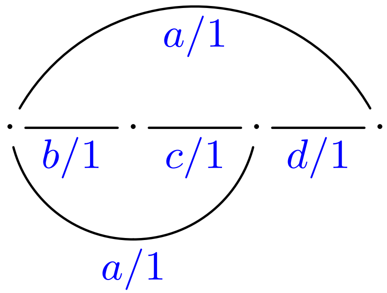
$$c(G) = \top \iff \text{there is a path from source to sink without edges labelled } a \text{ in } G$$



$$\begin{aligned} & c(a \parallel (((b \cdot c) \parallel a) \cdot d)) \\ = \quad & c(a) \vee (((c(b) \wedge c(c)) \vee c(a)) \wedge c(d)) \\ = \quad & \bot \vee (((\top \wedge \top) \vee \bot) \wedge \top) = \top \end{aligned}$$

- in the *semiring of natural numbers* $\mathbb{N}$ with $c(a) = c(b) = c(c) = c(d) = 1$:

$$c(G) = \text{ the number of different paths from source to sink in } G$$

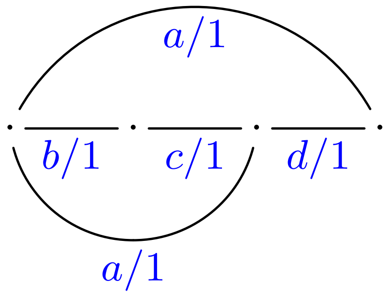


$$\begin{aligned} & c(a \parallel (((b \cdot c) \parallel a) \cdot d)) \\ = \quad & c(a) + (((c(b) \cdot c(c)) + c(a)) \cdot c(d)) \\ = \quad & 1 + (((1 \cdot 1) + 1) \cdot 1) = 3 \end{aligned}$$

Series-Parallel Graphs (cont'd)

- in the *arctic semiring* $\mathbb{A}$ with $c(a) = c(b) = c(c) = c(d) = 1$:

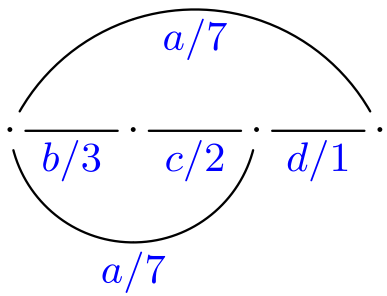
  $c(G) =$ the number of edges in a longest path from source to sink (*critical path*) in G



$$\begin{aligned} & c(a \parallel (((b \cdot c) \parallel a) \cdot d)) \\ = \; & \max(c(a), \max(c(b) + c(c), c(a)) + c(d)) \\ = \; & \max(1, \max(1 + 1, 1) + 1) = 3 \end{aligned}$$

- in the *tropical semiring* $\mathbb{T}$ with $c(a) = 7$, $c(b) = 3$, $c(c) = 2$ and $c(d) = 1$:

  $c(G) =$ the length of a shortest path from source to sink in G

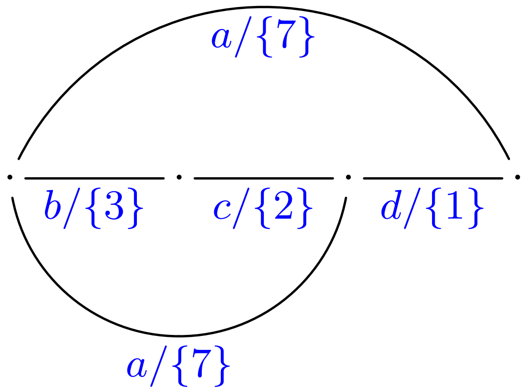


$$\begin{aligned} & c(a \parallel (((b \cdot c) \parallel a) \cdot d)) \\ = \; & \min(c(a), \min(c(b) + c(c), c(a)) + c(d)) \\ = \; & \min(7, \min(3 + 2, 7) + 1) = 6 \end{aligned}$$

Series-Parallel Graphs (cont'd)

- in the *subset-semiring* $\mathbb{F}$ with $c(a) = \{7\}$, $c(b) = \{3\}$, $c(c) = \{2\}$ and $c(d) = \{1\}$:

$$c(G) = \text{ the set of all path lengths from source to sink in } G$$



$$\begin{aligned} & c(a \parallel (((b \cdot c) \parallel a) \cdot d)) \\ = \quad & c(a) \cup (((c(b) + c(c)) \cup c(a)) + c(d)) \\ = \quad & \{7\} \cup (((\{3\} + \{2\}) \cup \{7\}) + \{1\}) = \{7, 6, 8\} \end{aligned}$$

Those computations can be incorporated into finite tree automata.

Tree Automata

Definition: A *tree automaton* is a quadruple $M = (Q, \Sigma, \delta, F)$, where

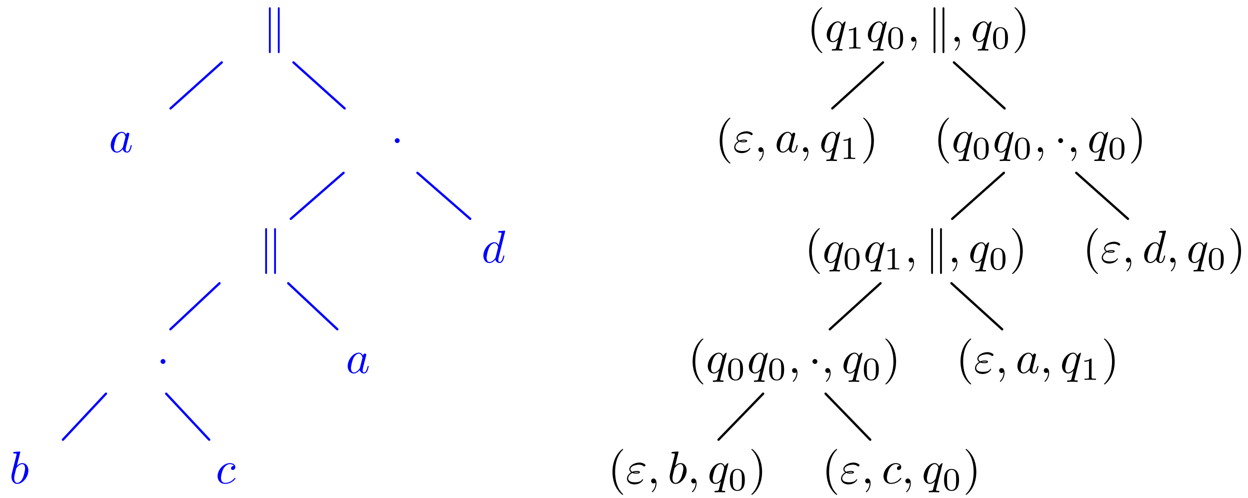
- Q is a finite, non-empty set of *states*,
- Σ is a ranked alphabet of *input symbols*,
- $\delta \subseteq \bigcup_{k \in \mathbb{N}} Q^k \times \Sigma \times Q$ is a set of *transitions* and
- $F \subseteq Q$ is the set of *final states*.

Example: Let $\Sigma = \{\|^{(2)}, \cdot^{(2)}, a^{(0)}, b^{(0)}, c^{(0)}, d^{(0)}\}$ be as before, $Q = \{q_0, q_1\}$, $F = \{q_1\}$ and the transitions are specified in the following tables.

symbol	(q_0, q_0)	(q_0, q_1)	(q_1, q_0)	(q_1, q_1)
$\|$	q_0	q_0	q_0	q_1
$\cdot$	q_0	q_1	q_1	q_1

symbol	ε
a	q_1
b, c, d	q_0

Computation of a Tree Automaton



Since this is the only possible computation tree for the given input tree and $q_0 \notin F$, the input tree is rejected, i.e. does not belong to the (tree) language accepted by the tree automaton.

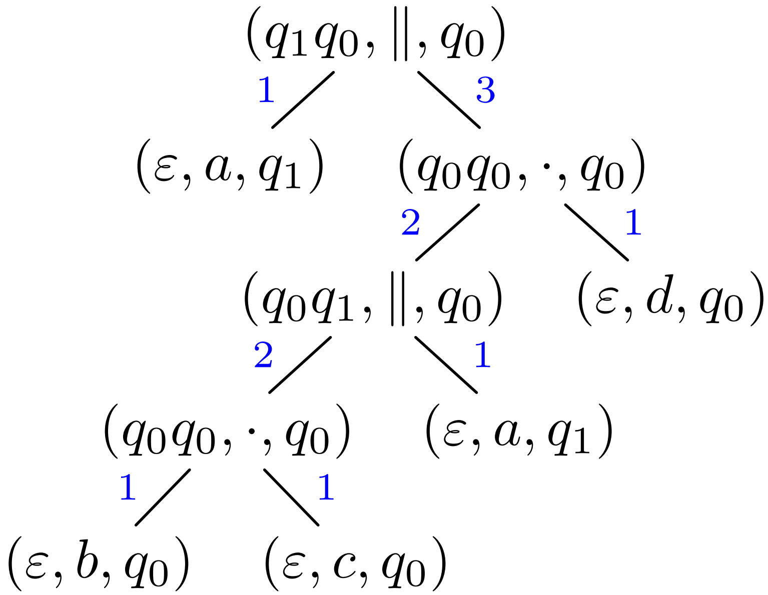
Generally speaking: This tree automaton accepts series-parallel graphs, in which every path from the source to the sink contains at least one a.

Cost Function

Definition: Given a tree automaton $M = (Q, \Sigma, \delta, F)$, a *cost function for M* over semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ is a mapping $c: \ \delta \longrightarrow A[X]$.

Example: using the arctic semiring $\mathbb{A}$

$$
\begin{aligned}
1 &= c(\varepsilon, a, q_1) = c(\varepsilon, b, q_0) \\
&= c(\varepsilon, c, q_0) = c(\varepsilon, d, q_0) \\
\max(x_1, x_2) &= c(q_0q_0, \|, q_0) = c(q_0q_1, \|, q_0) \\
&= c(q_1q_0, \|, q_0) = c(q_1q_1, \|, q_1) \\
x_1 + x_2 &= c(q_0q_0, \cdot, q_0) = c(q_0q_1, \cdot, q_1) \\
&= c(q_1q_0, \cdot, q_1) = c(q_1q_1, \cdot, q_1)
\end{aligned}
$$



Monotonic Semirings

Definition: A semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ is called *monotonic*, iff

- there is a partial order $(A, \preceq)$ such that
- $a \preceq a \odot b$ and $b \preceq a \odot b$ for every $a, b \in A \setminus \{\mathbf{0}\}$,
- $a \preceq a \oplus b$ and
- $a \prec a \odot a$ for every $a \notin \{\mathbf{0}, \mathbf{1}\}$.

Examples:

- Semiring of natural numbers $\mathbb{N}$,
- Arctic semiring $\mathbb{A}$,
- (Finite language) semiring $\mathbb{L} = (\mathcal{P}_f(\Sigma^*), \cup, \circ, \emptyset, \{\varepsilon\})$ with the common operations of union and concatenation.

Naturally Ordered and Additively Idempotent Semirings

Observation: Let $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ be an *additively idempotent* semiring, i.e. the equality $\mathbf{1} \oplus \mathbf{1} = \mathbf{1}$ holds. Then $\mathcal{A}$ is monotonic, if $a \prec a \odot a$ for every $a \notin \{\mathbf{0}, \mathbf{1}\}$, where $\preceq \subseteq A \times A$ is defined by

$$a \preceq b \quad \Longleftrightarrow \quad a \oplus b = b.$$

Observation: Let $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ be a *naturally ordered* semiring, i.e. the relation $\sqsubseteq \subseteq A \times A$ with

$$a \sqsubseteq b \quad \Longleftrightarrow \quad (\exists c \in A): \; a \oplus c = b$$

is a partial order over A. Then $\mathcal{A}$ is monotonic, if for every $a \notin \{\mathbf{0}, \mathbf{1}\}$ the condition $a \prec a \odot a$ holds.

Star Search

Definition: The *star* of a semiring element $a \in A$ is defined as:

$$a^* = \lim_{n \to \infty} \sum_{i=0}^{n} a^i \qquad \text{(compare } \Sigma^* = \lim_{n \to \infty} \bigcup_{i=0}^{n} \Sigma^i \text{ and reflexive, transitive closure).}$$

Example: The star of $\mathbf{0}$ exists in any semiring and is always $\mathbf{1}$.

- in the semiring of natural numbers $\mathbb{N}$: no more stars exist
- in the arctic semiring $\mathbb{A}$: 0^* exists
- in the tropical semiring $\mathbb{T}$: 0^* exists
- in the subset semiring $\mathbb{F}$: $\{0\}^*$ exists
- in the boolean semiring $\mathbb{B}$: $\top^*$ exists
- in the (finite language) semiring $\mathbb{L}$: $\{\varepsilon\}^*$ exists

Some More Properties of Semirings

Definition: A monoid $\mathcal{A} = (A, \otimes, \mathbf{1})$ is *periodic*, if for every element $a \in A$ there exist $i, j \in \mathbb{N}$ with $i < j$ and $a^i = a^j$.

Example: Every additively idempotent semiring is additively periodic with $i = 1$ and $j = 2$.

Definition: A monoid $\mathcal{A} = (A, \otimes, \mathbf{1})$ is *locally finite*, if for every finite $B \subseteq A$ also $\langle B \rangle$ is finite.

Example: Every additively idempotent semiring is additively locally finite.

Observation: Given a semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$, $\mathcal{A}$ is additively locally finite, iff $\mathcal{A}$ is additively periodic.

Observation: Given a semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ where $\mathbf{1}^*$ exists. Then $\mathcal{A}$ is additively periodic. Moreover, on monotonic semirings: $\mathbf{1}^*$ exists, iff $\mathcal{A}$ is additively periodic.

Finitely Factorizing Semirings

Definition: A semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ is *finitely factorizing*, if both monoids $(A, \oplus, \mathbf{0})$ and $(A \setminus \{\mathbf{0}\}, \odot, \mathbf{1})$ are finitely factorizing, i.e. given a monoid $(B, \otimes, \mathbf{1})$ for every $b \in B$ the set $\left\{ (c, d) \in B^2 \mid b = c \otimes d \right\}$ is finite.

Example: The semirings $\mathbb{N}$, $\mathbb{A}$ and $\mathbb{L}$ are finitely factorizing, while $\mathbb{T}$ and $\mathbb{F}$ are not.

E-states

Definition: Let $M = (Q, \Sigma, \delta, F)$ be a tree automaton with cost function $c\colon\ \delta \longrightarrow A[X]$ over a semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ and $E \subseteq A$. The set of *E-states* of M, denoted Q_E, is

$$Q_E = \{\, q \in Q \mid \text{for every } q\text{-computation } \psi\text{: } c(\psi) \in E \,\}\,.$$

Lemma: For monotonic semirings we can effectively determine $Q_{\{\mathbf{0}\}}$ and $Q_{\{\mathbf{0},\mathbf{1}\}}$.

Example: The set of all $\{\mathbf{0}\}$-states $Q_{\{\mathbf{0}\}} \subseteq Q$ can be computed as follows:

- Set $Q_0 = Q$.
- For every $n \in \mathbb{N}$ set

$$Q_{n+1} = Q_n \setminus \left\{ q \in Q_n \;\middle|\; \begin{array}{l} (\exists \tau = (q_1 \ldots q_k, \sigma, q) \in \delta)\big(\exists m \in \mathrm{mon}(c(\tau))\big) \\ (\forall j \in \mathrm{var}(m))\,:\ q_j \in Q \setminus Q_n \end{array} \right\}.$$

Then $Q_{\{0\}} = Q_\omega$.

Boundedness

Classical notion of boundedness fails, since there are several semirings (e.g. $\mathbb{T}$) which possess a maximal element (w.r.t. some partial order). Every tree automaton with cost function over such a semiring would then be bounded.

Definition: A tree automaton $M = (Q, \Sigma, \delta, F)$ with cost function $c : \delta \longrightarrow A[X]$ over a semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ is *bounded*, if

$$c(M) = \{ c(\psi) \mid \psi \text{ is an accepting computation of } M \}$$

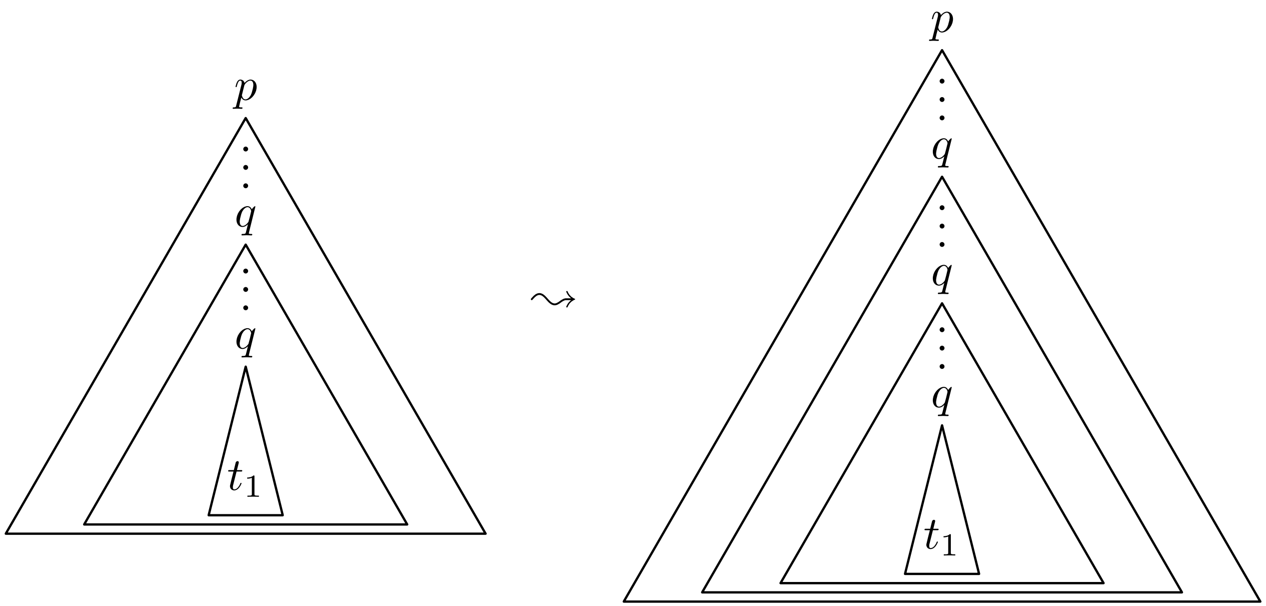
is finite.

Observation: Every tree automaton with cost function over a finite semiring is bounded.

Boundedness Result

Theorem: Let $M = (Q, \Sigma, \delta, F)$ be a tree automaton with cost function $c: \delta \longrightarrow A[X]$ over a finitely factorizing and monotonic semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$. M is bounded, iff for every q-q-computation ψ with $q \notin Q_{\{\mathbf{0},\mathbf{1}\}}$ either

- $c(\psi) = x_1 + a$ for some $a \in A$ and ($\mathbf{1}^*$ exists or $a = \mathbf{0}$) or
- $c(\psi)$ is a constant.



Boundedness Result (cont'd)

Rationale: The cases $c(\psi) = x_1$ and $c(\psi) = a$ for some $a \in A$ are straightforward. Let $c(\psi) = x_1 + a$, thus $\mathbf{1}^*$ exists. It follows that $\mathcal{A}$ is additively periodic, hence additively locally finite.

Rationale: Let $c(\psi) = bx_1$ with $b \notin \{\mathbf{0}, \mathbf{1}\}$. By $b \prec b \odot b$ pumping yields unboundedness.

Rationale: Let $c(\psi) = x_1^2$. By $q \notin Q_{\{\mathbf{0},\mathbf{1}\}}$ and $b \prec b \odot b$ pumping again yields unboundedness.

Examples

Example: Let $M = (Q, \Sigma, \delta, F)$ be a tree automaton with cost function $c: \delta \longrightarrow \mathbb{N}[X]$ over the semiring $(\mathbb{N}, +, \cdot, 0, 1)$. M is bounded if and only if for every q-q-computation ψ with $q \notin Q_{\{0,1\}}$ either

- $c(\psi) = a$ for some $a \in \mathbb{N}$ or
- $c(\psi) = x_1$ holds.

Example: Let $M = (Q, \Sigma, \delta, F)$ be a tree automaton with cost function $c: \delta \longrightarrow (\mathbb{N} \cup \{-\infty\})[X]$ over the arctic semiring $(\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$. M is bounded if and only if for every q-q-computation ψ with $q \notin Q_{\{0,1\}}$ either

- $c(\psi) = a$ for some $a \in (\mathbb{N} \cup \{\infty\})$ or
- $c(\psi) = \max(x_1, c_0)$ for some $c_0 \in (\mathbb{N} \cup \{-\infty\})$ holds.

Remaining Questions

- Can we decide the property required for all q-q-computations?
- Can we also characterize boundedness by some property which is based on single transitions rather than q-q-computations?
- Which properties of monotonic semirings are obsolete when restricting ourselves to tree automata with linear cost functions?
- Can we characterize unboundedness of tree automata with cost function over certain semirings which are not finitely factorizing?
- Can we establish sufficient or necessary criteria for boundedness/unboundedness with less restrictions on the semiring?