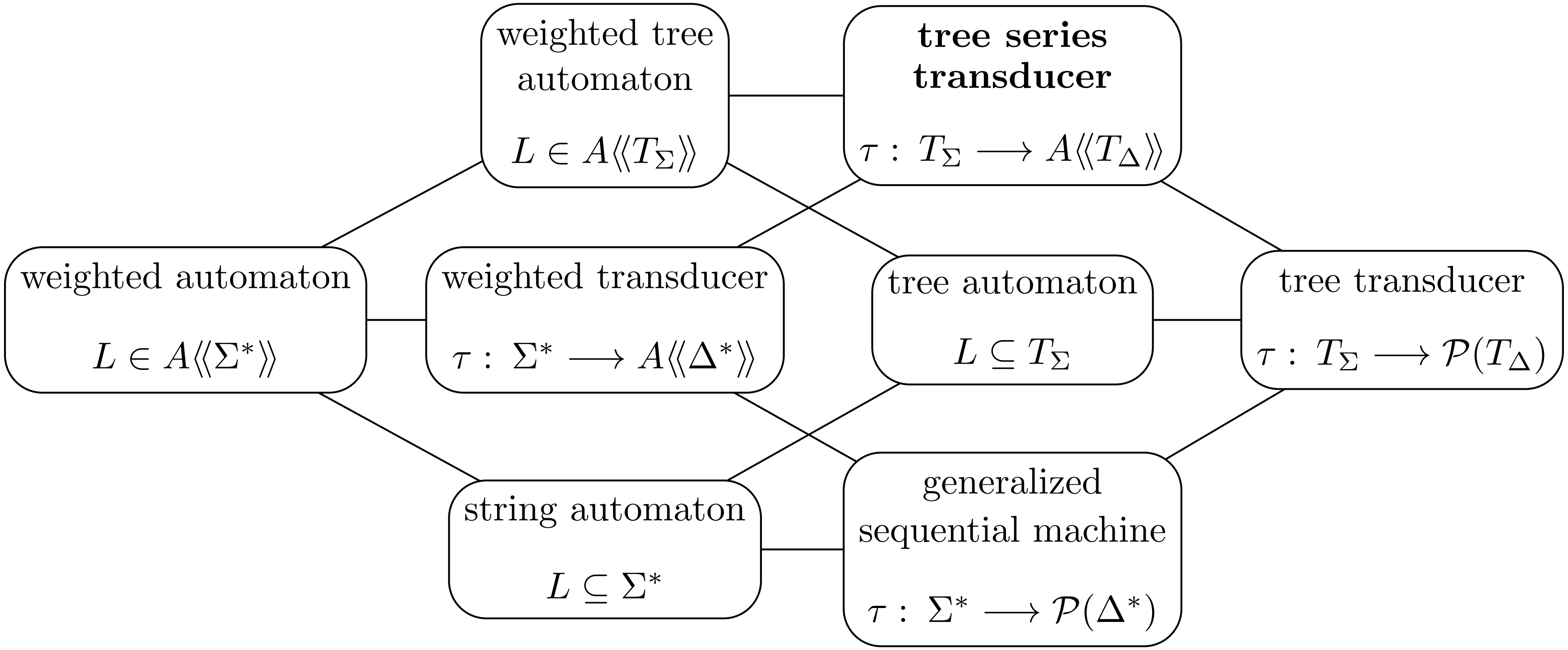
Hierarchies of Deterministic Bottom-Up Tree-to-Tree-Series Transformations

Andreas Maletti, Heiko Vogler

November 28, 2003

1. Introduction and Motivation
2. Deterministic Tree Series Transducers
3. Particular Semirings
4. Hasse Diagrams

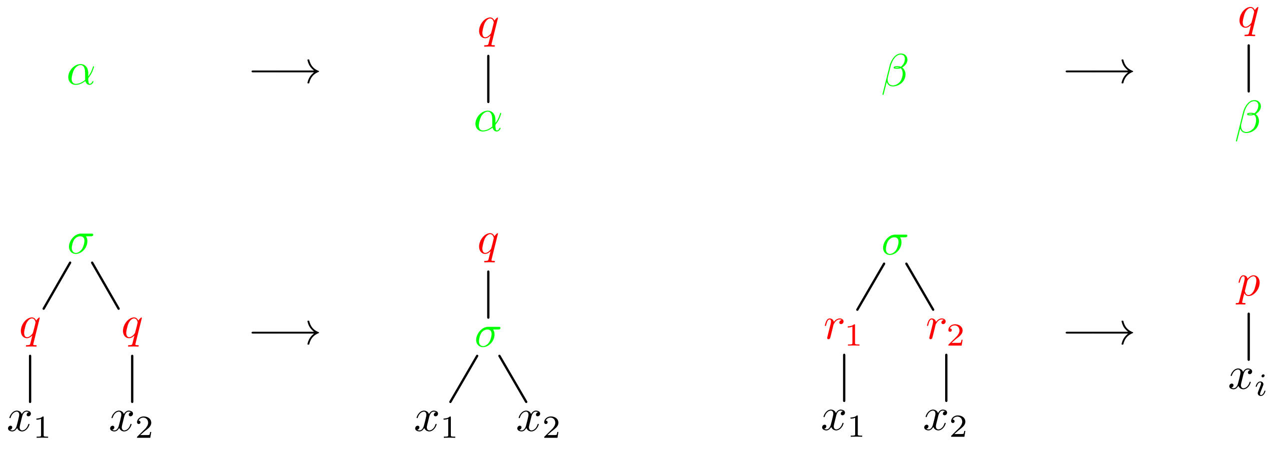
Generalization Hierarchy

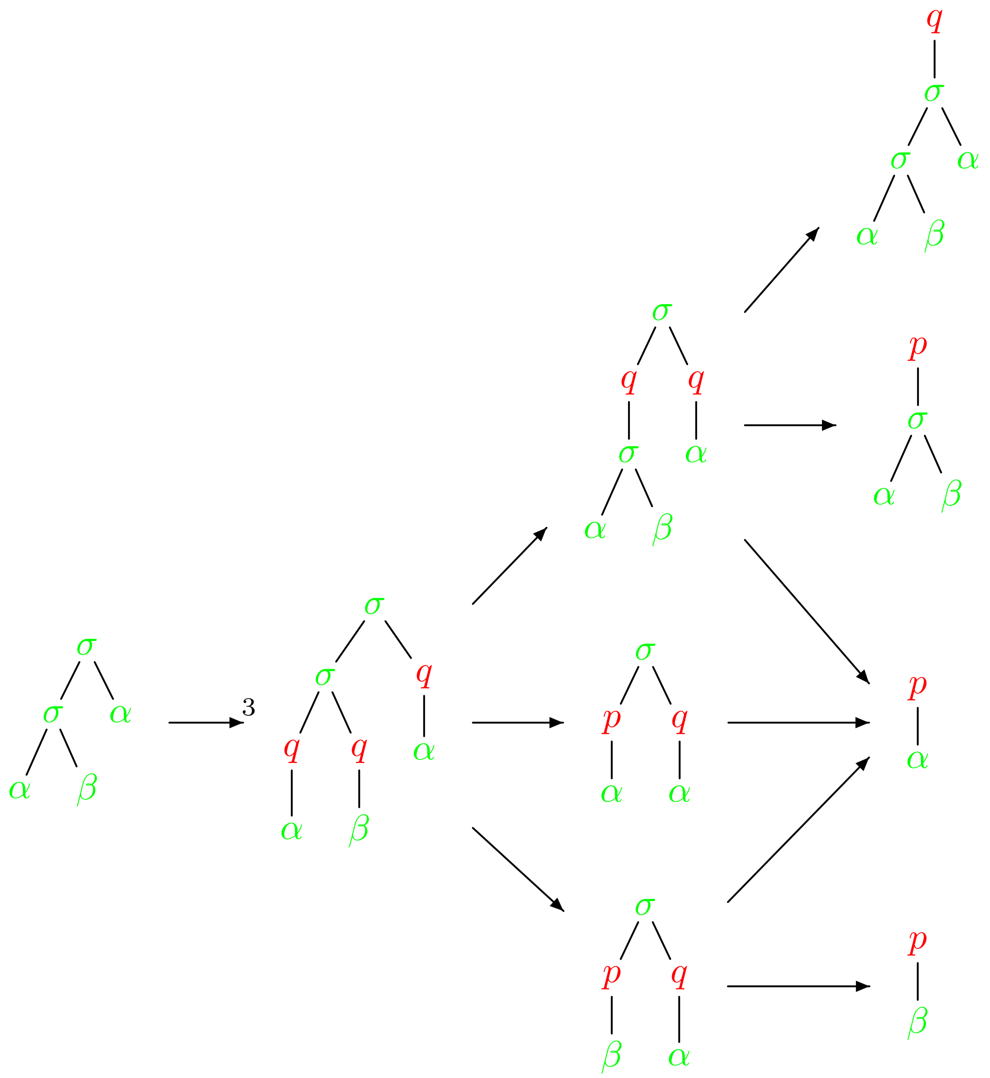


Bottom-Up Tree Transducers

$$M = (Q, \Sigma, \Delta, F, R)$$

- input and output ranked alphabet $\Sigma = \Delta = \{\sigma^{(2)}, \alpha^{(0)}, \beta^{(0)}\}$,
- states and final states $Q = F = \{p, q\}$, and
- transitions R




3

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., $\mathbf{0} \odot a = a \odot \mathbf{0} = \mathbf{0}$.

Examples of Semirings

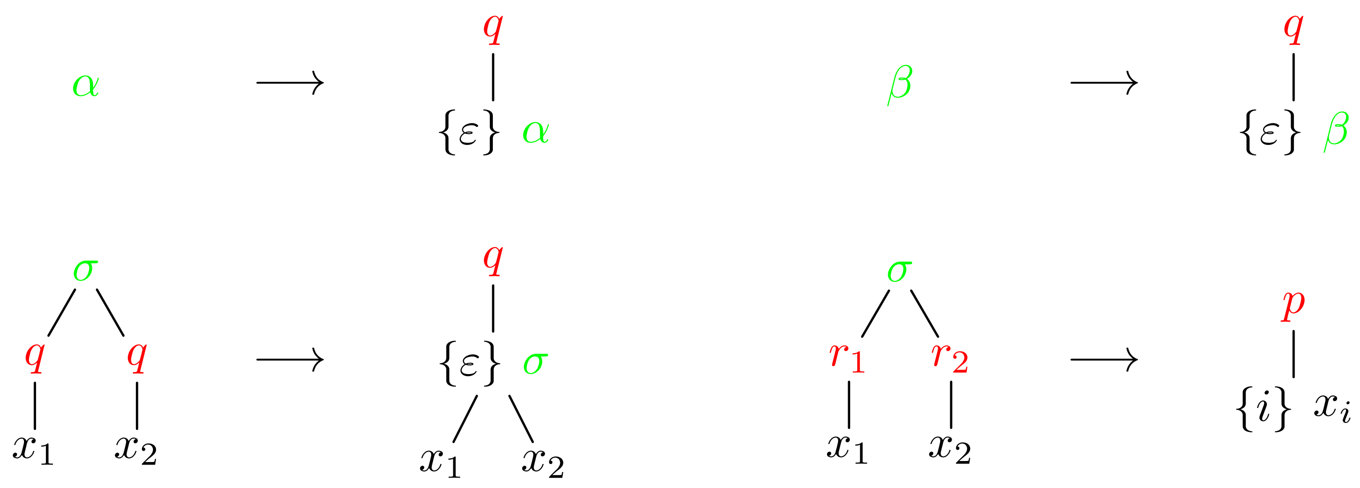
- the *semiring of the natural numbers* $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$,
- the *arctic semiring* $\mathbb{A} = (\mathbb{N} \cup \{-\infty\}, \max, +, (-\infty), 0)$,
- the ring $\mathbb{Z}_4 = (\{0, 1, 2, 3\}, +, \cdot, 0, 1)$ with the common operations of addition and multiplication modulo 4,
- the field $\mathbb{Z}_5 = (\{0, 1, 2, 3, 4\}, +, \cdot, 0, 1)$,
- the min-max *semiring on the reals* $\mathbb{R}_{\min,\max} = (\mathbb{R} \cup \{-\infty, +\infty\}, \min, \max, (+\infty), (-\infty))$
- the *boolean semiring* $\mathbb{B} = (\{0, 1\}, \vee, \wedge, 0, 1)$,
- the *paths semiring* $\mathbb{P} = (\mathcal{P}(\mathbb{N}_+^*), \cup, \circ, \emptyset, \{\varepsilon\})$ with

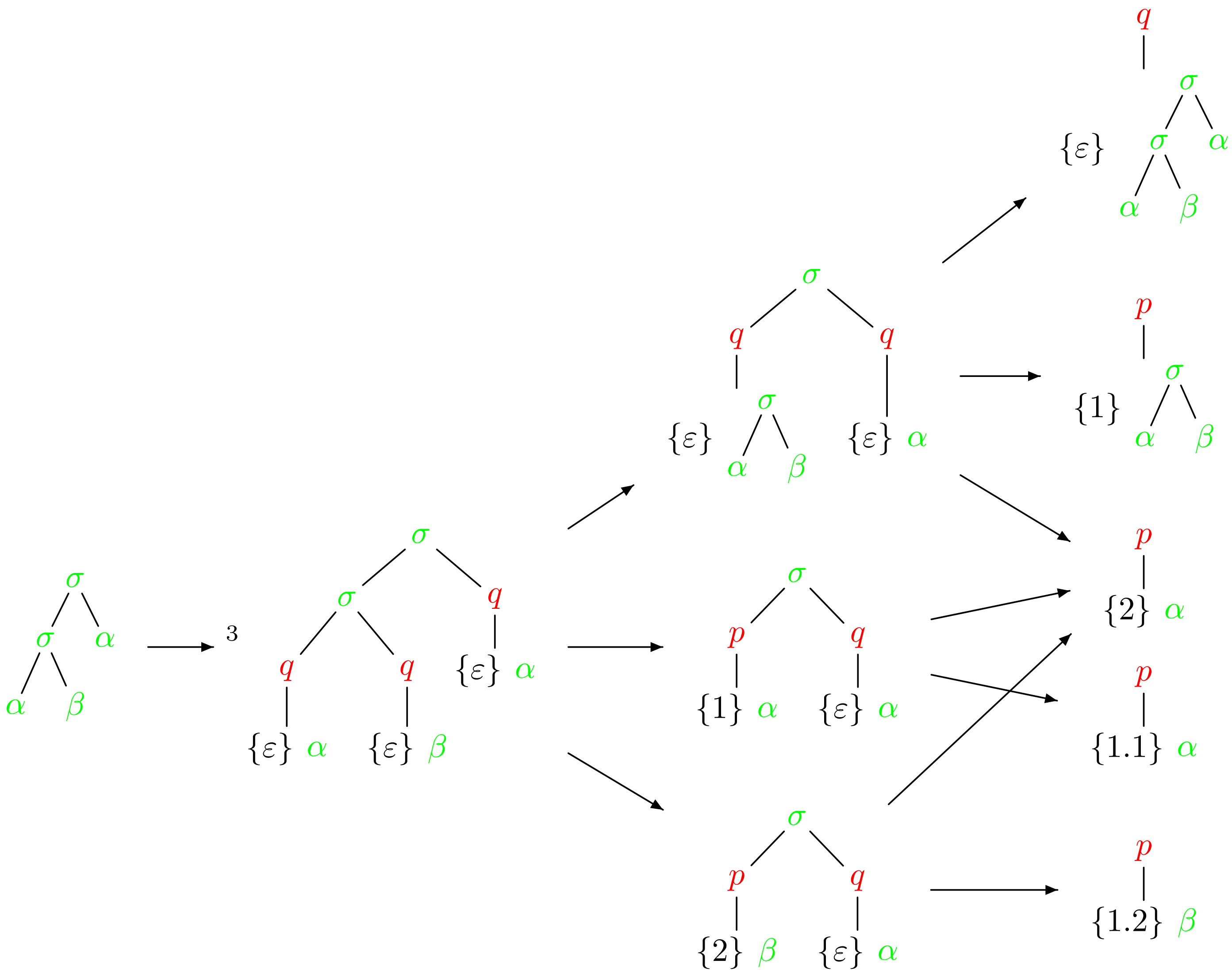
$$P_1 \circ P_2 = \{\, ab \mid a \in P_1, b \in P_2 \,\}.$$

Bottom-Up Tree Series Transducers

$$M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$$

- input and output ranked alphabet $\Sigma = \Delta = \{\sigma^{(2)}, \alpha^{(0)}, \beta^{(0)}\}$,
- states and final states $Q = F = \{p, q\}$,
- semiring $\mathcal{A} = \mathbb{P}$, and
- tree representation μ





σ
σ α
α β
3
σ
σ q
q q {ε} α
{ε} α {ε} β
σ
q q
{ε} σ {ε} α
α β
σ
p q
{1} α {ε} α
σ
p q
{2} β {ε} α
q
σ
{ε} σ α
α β
p
σ
{1} α β
p
{2} α
p
{1.1} α
p
{1.2} β

Deterministic Tree Series Transducers

Definition: A *deterministic bottom-up tree series transducer* is a sixtuple $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$, where

- Q is a finite, non-empty set of *states*,
- Σ, Δ are ranked alphabets of *input symbols* and *output symbols*, respectively,
- $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ is a semiring,
- $F \subseteq Q$ is a set of *final states*, and
- μ is a *deterministic tree representation*.

$$\mu = \left(\mu_k : \Sigma^{(k)} \longrightarrow (A[T_\Delta(X_k)])^{Q \times Q^k} \mid k \in \mathbb{N} \right)$$

Intuitively, given an input symbol $\sigma \in \Sigma^{(k)}$ and states $q_1, \ldots, q_k \in Q$ we have $\mu_k(\sigma)_{q,(q_1,\ldots,q_k)} = a\,t$ for every $q \in Q$. Therein $a \in A$ and $t \in T_\Delta(X_k)$ and for at most one $q \in Q$ the coefficient fulfils $a \neq \mathbf{0}$.

Two Substitutions

We write $\widetilde{\mathbf{0}}$ instead of $\mathbf{0}\,t$.

(i) $\varphi \longleftarrow () \;=\; \varphi$,

(ii) $\widetilde{\mathbf{0}} \longleftarrow (\psi_1, \ldots, \psi_k) \;=\; \widetilde{\mathbf{0}}$,

(iii) $\varphi \longleftarrow (\psi_1, \ldots, \psi_{i-1}, \widetilde{\mathbf{0}}, \psi_{i+1}, \ldots, \psi_k) \;=\; \widetilde{\mathbf{0}}$,

(iv') $a\,t \longleftarrow (a_1\,t_1, \ldots, a_k\,t_k) \;=\; (a \odot a_1 \odot \cdots \odot a_k)\,t[t_1, \ldots, t_k]$.

(iv'') $a\,t \longleftarrow (a_1\,t_1, \ldots, a_k\,t_k) \;=\; (a \odot a_1^{|t|_{x_1}} \odot \cdots \odot a_k^{|t|_{x_k}})\,t[t_1, \ldots, t_k]$.

(i) - (iii), (iv'): *pure substitution*, denoted by $\longleftarrow$

(i) - (iii), (iv''): *o-substitution*, denoted by $\overset{o}{\longleftarrow}$

Tree-to-Tree-Series Transformation

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ be a deterministic bottom-up tree series transducer and $\mathrm{mod} \in \{\varepsilon, o\}$. M computes the mod-*tree-to-tree-series transformation* (short: mod-t-ts transformation) $\tau_M^{\mathrm{mod}} : T_\Sigma \longrightarrow A[T_\Delta]$ defined as follows.

$$\tau_M^{\mathrm{mod}}(s) = \begin{cases} h_\mu^{\mathrm{mod}}(s)_q & \text{, if } h_\mu^{\mathrm{mod}}(s)_q \neq \widetilde{\mathbf{0}} \\ \widetilde{\mathbf{0}} & \text{, otherwise} \end{cases}$$

where the mapping $h_\mu^{\mathrm{mod}}(\cdot)_q : T_\Sigma \longrightarrow A[T_\Delta]$ is inductively defined by:

- if $s = \alpha$, then $h_\mu^{\mathrm{mod}}(\alpha)_q = \mu_0(\alpha)_{q,\varepsilon}$
- if $s = \sigma(s_1, \ldots, s_k)$ and for some $q_1, \ldots, q_k \in Q$ the conditions $h_\mu^{\mathrm{mod}}(s_i)_{q_i} \neq \widetilde{\mathbf{0}}$ are fulfilled, then

$$h_\mu^{\mathrm{mod}}(\sigma(s_1, \ldots, s_k))_q = \mu_k(\sigma)_{q,(q_1,\ldots,q_k)} \overset{\mathrm{mod}}{\longleftarrow} (h_\mu^{\mathrm{mod}}(s_1)_{q_1}, \ldots, h_\mu^{\mathrm{mod}}(s_k)_{q_k})$$

- otherwise $h_\mu^{\mathrm{mod}}(\sigma(s_1, \ldots, s_k))_q = \widetilde{\mathbf{0}}$

Example

Let $\Sigma = \{\sigma^{(2)}, \alpha^{(0)}\}$ and $\Delta = \{\delta^{(1)}, \alpha^{(0)}\}$.

$$N = (\{*\}, \Delta, \Sigma, \mathbb{N}, \{*\}, \nu)$$

with tree representation ν specified by

$$\nu_0(\alpha)_{*,\varepsilon} = 2\,\alpha \qquad \text{and} \qquad \nu_1(\delta)_{*,(*)} = 2\,\sigma(x_1, x_1).$$

Let $s \in T_\Delta$ and $t \in T_\Sigma$ such that t is the fully balanced tree with $\mathrm{height}(t) = \mathrm{height}(s)$.

- The *t-ts transformation* τ_N computed by N maps the tree s to the monomial $2^{\mathrm{size}(s)}\ t$.
- The *o-t-ts transformation* τ_N^o computed by N maps the tree s to the monomial $2^{\mathrm{size}(t)}\ t$.

Note that $\mathrm{size}(t) = 2^{\mathrm{size}(s)} - 1$.

Properties of Tree Series Transducers

Definition: A deterministic bottom-up tree series transducer $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ is called

- *non-deleting*, if each tree occurring in the range of μ_k contains each variable $x_1, \ldots, x_k$ at least once,
- *linear*, if all trees in the tree representation μ contain each variable at most once,
- *total*, if for each symbol $\sigma \in \Sigma^{(k)}$ and tuple of states $w \in Q^k$ there exists at least one $q \in Q$ such that $\mu_k(\sigma)_{q,w} \neq \widetilde{\mathbf{0}}$,
- *homomorphism*, if M is total and Q is a singleton.

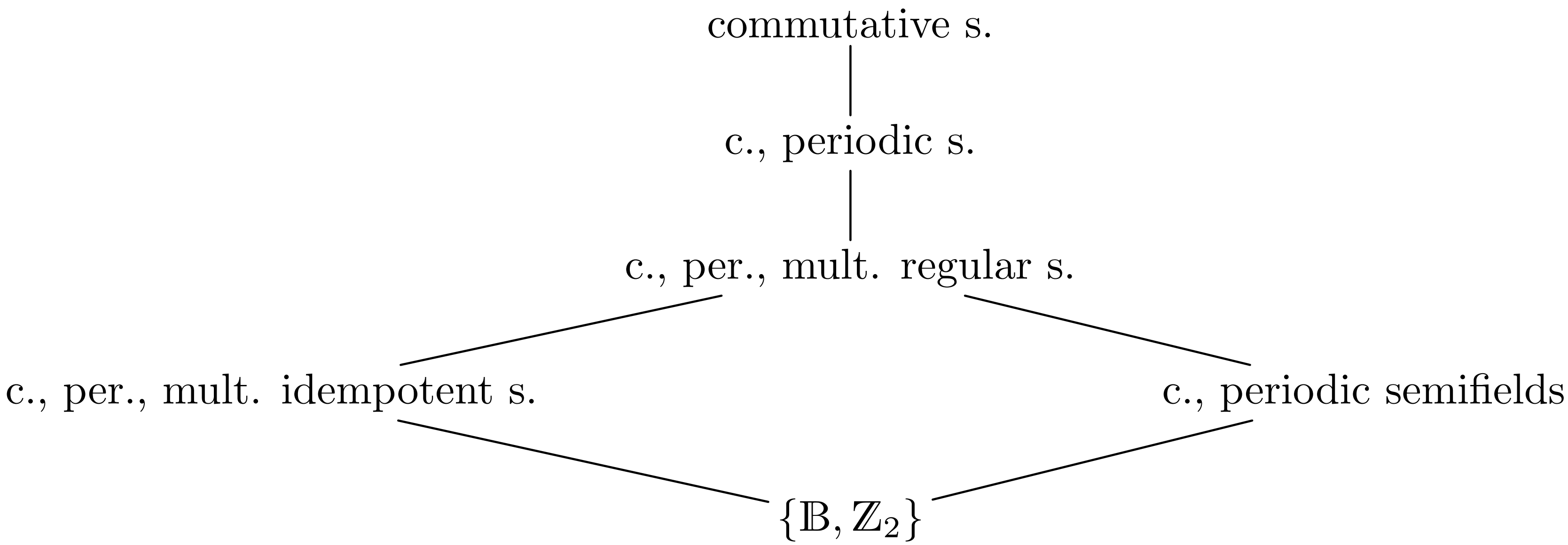
Example: N is a non-deleting homomorphism (deterministic) bottom-up tree series transducer which is not linear.

Properties of Semirings

A semiring $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ is called

- *commutative*, if $a \odot b = b \odot a$ for every $a, b \in A$,
- *finite*, if A is finite,
- *periodic*, if for every $a \in A$ there exist $i, j \in \mathbb{N}$ such that $a^i = a^j$ and $i \neq j$,
- *multiplicatively idempotent*, if $a \odot a = a$ for every $a \in A$,
- *multiplicatively regular*, if for every $a \in A$ there exists a $b \in A$ such that $a \odot b \odot a = a$, and
- *a semifield*, if for every $a \in A$ there exists a $b \in A$ such that $a \odot b = \mathbf{1}$.

Semiring Properties Related

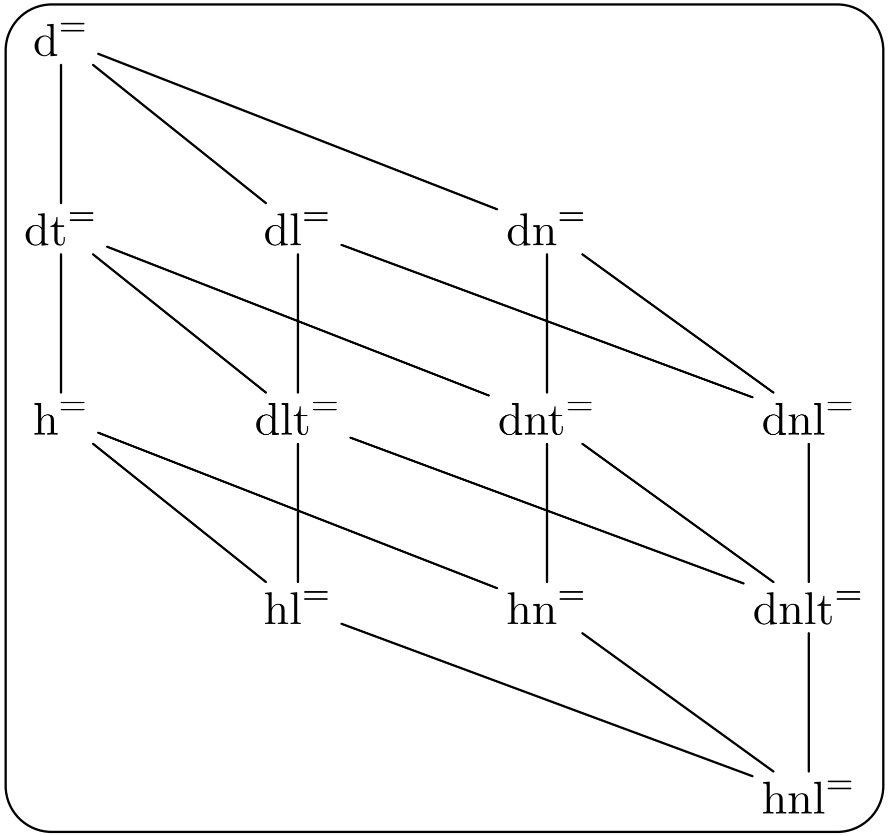


Classes of T-TS Transformations

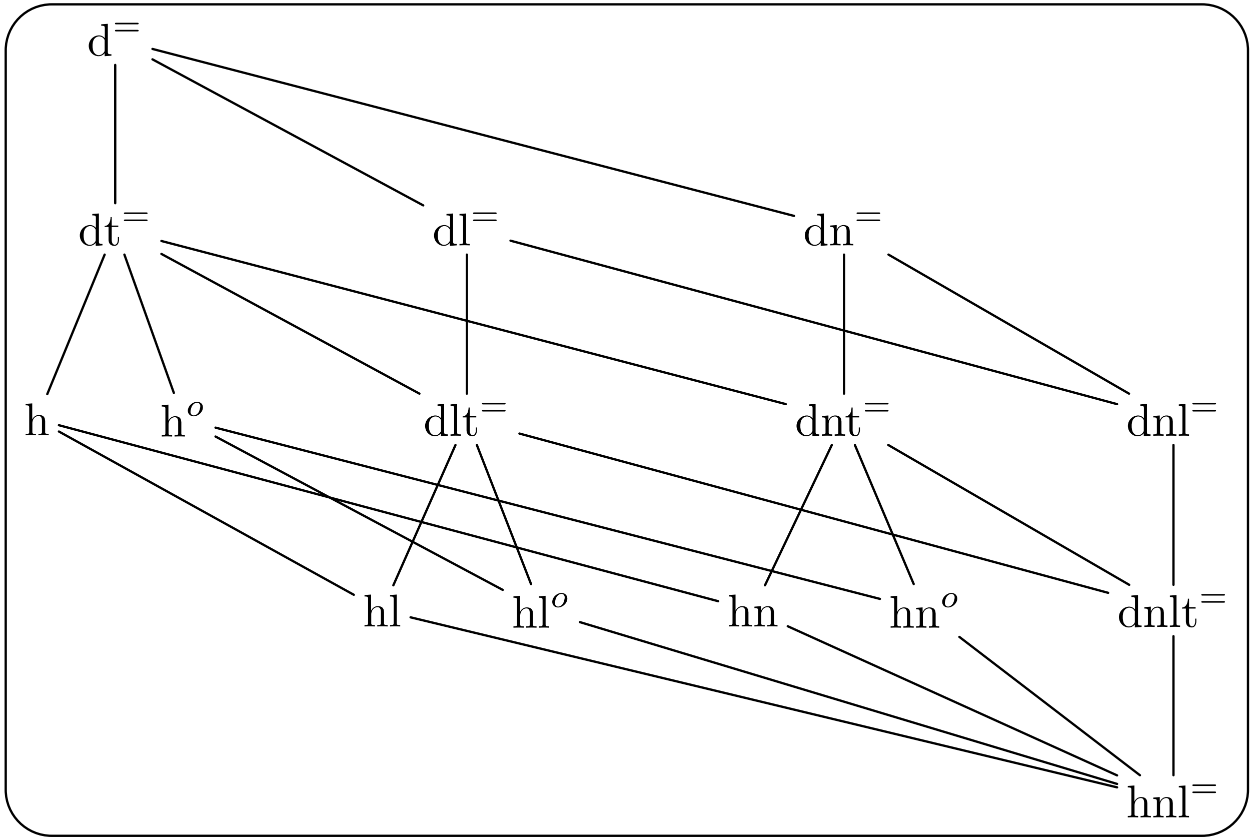
Let $\mathrm{mod} \in \{\varepsilon, o\}$ and $\mathcal{A} = (A, \oplus, \odot, \mathbf{0}, \mathbf{1})$ be a semiring. The *class of mod-t-ts transformations* computed by deterministic bottom-up tree series transducers over the semiring $\mathcal{A}$, which have all the properties of $x \subseteq \{\mathsf{n}, \mathsf{l}, \mathsf{t}, \mathsf{h}\}$ (non-deleting, linear, total, homomorphism), is denoted by $\mathsf{d}x$–$\mathsf{BOT}^{\mathrm{mod}}(\mathcal{A})$.

E.g., dnl–$\mathsf{BOT}^{o}(\mathbb{N})$ denotes the class of all o-t-ts transformations computable by non-deleting and linear deterministic bottom-up tree series transducers over $\mathbb{N}$.

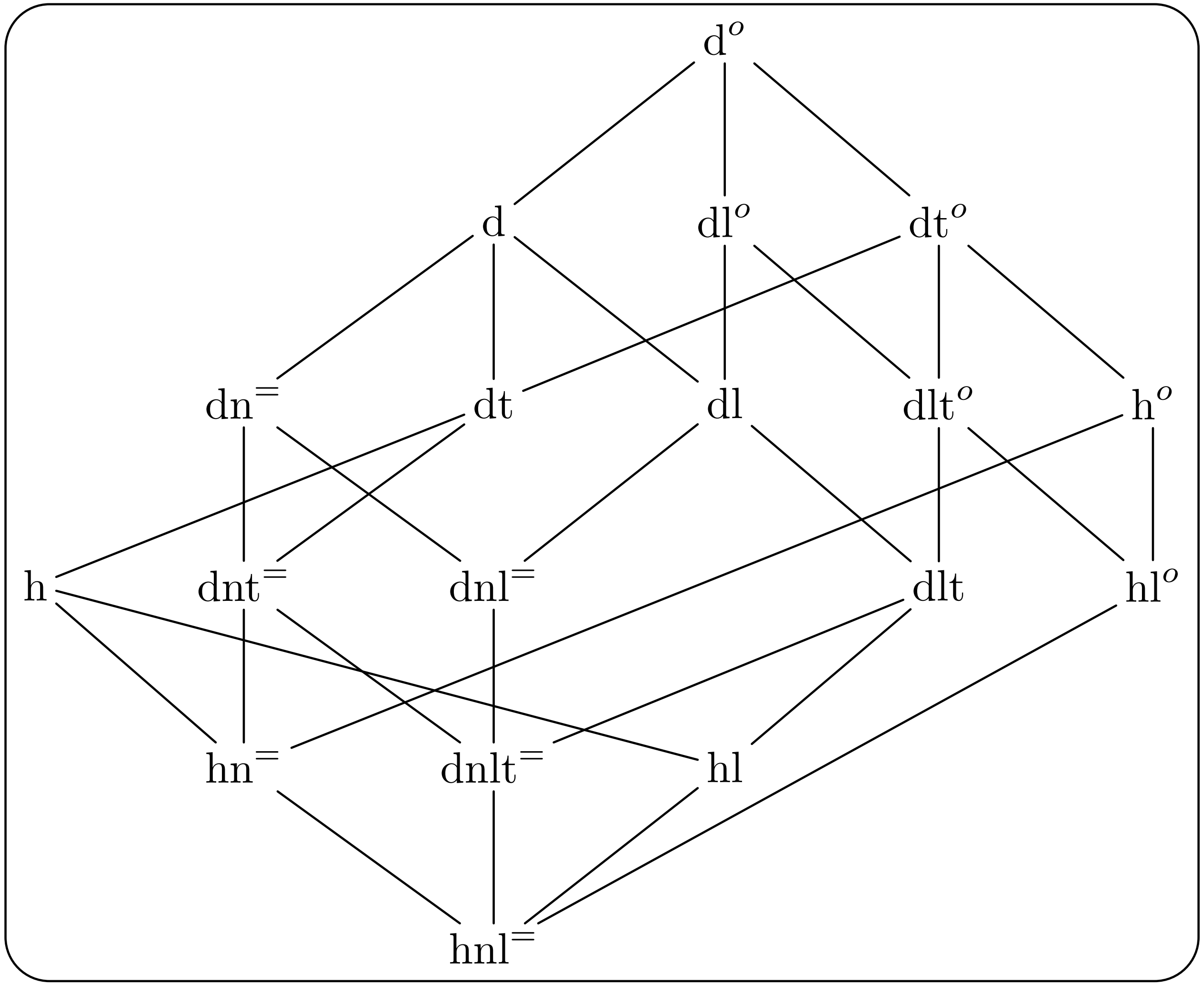
Boolean Semiring



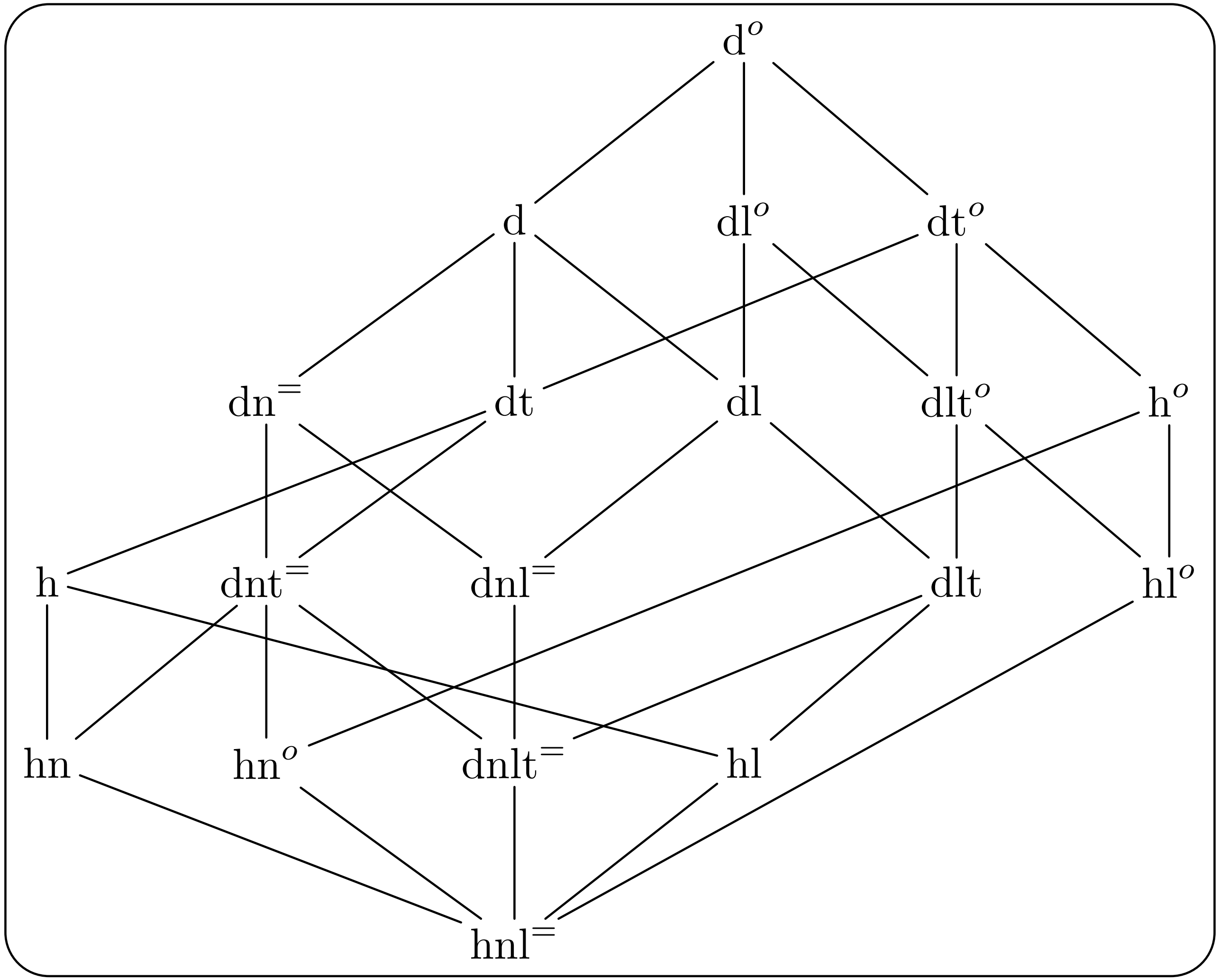
Commutative and Periodic Semifields



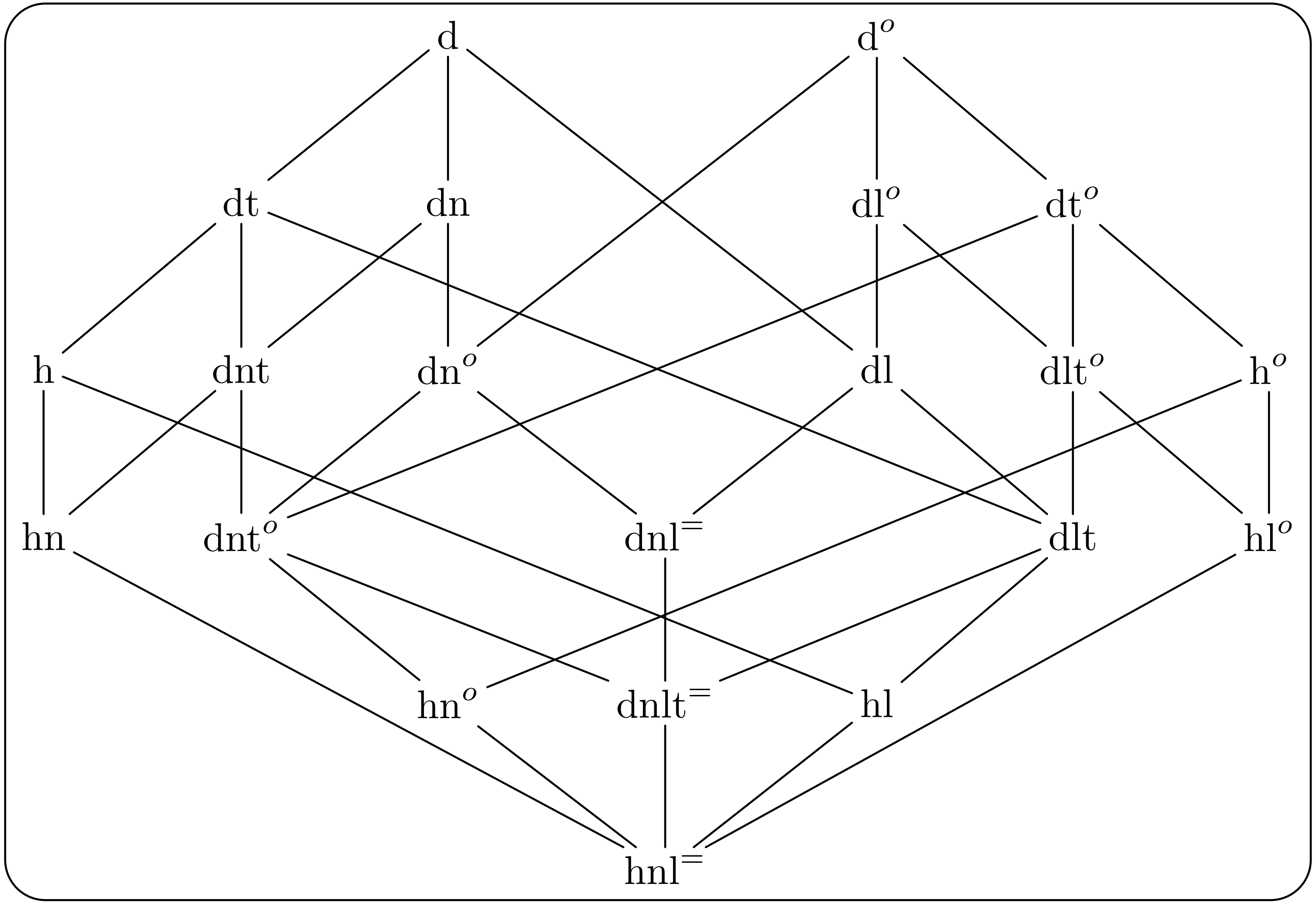
Commutative, Periodic, and Multiplicatively Idempotent Semirings



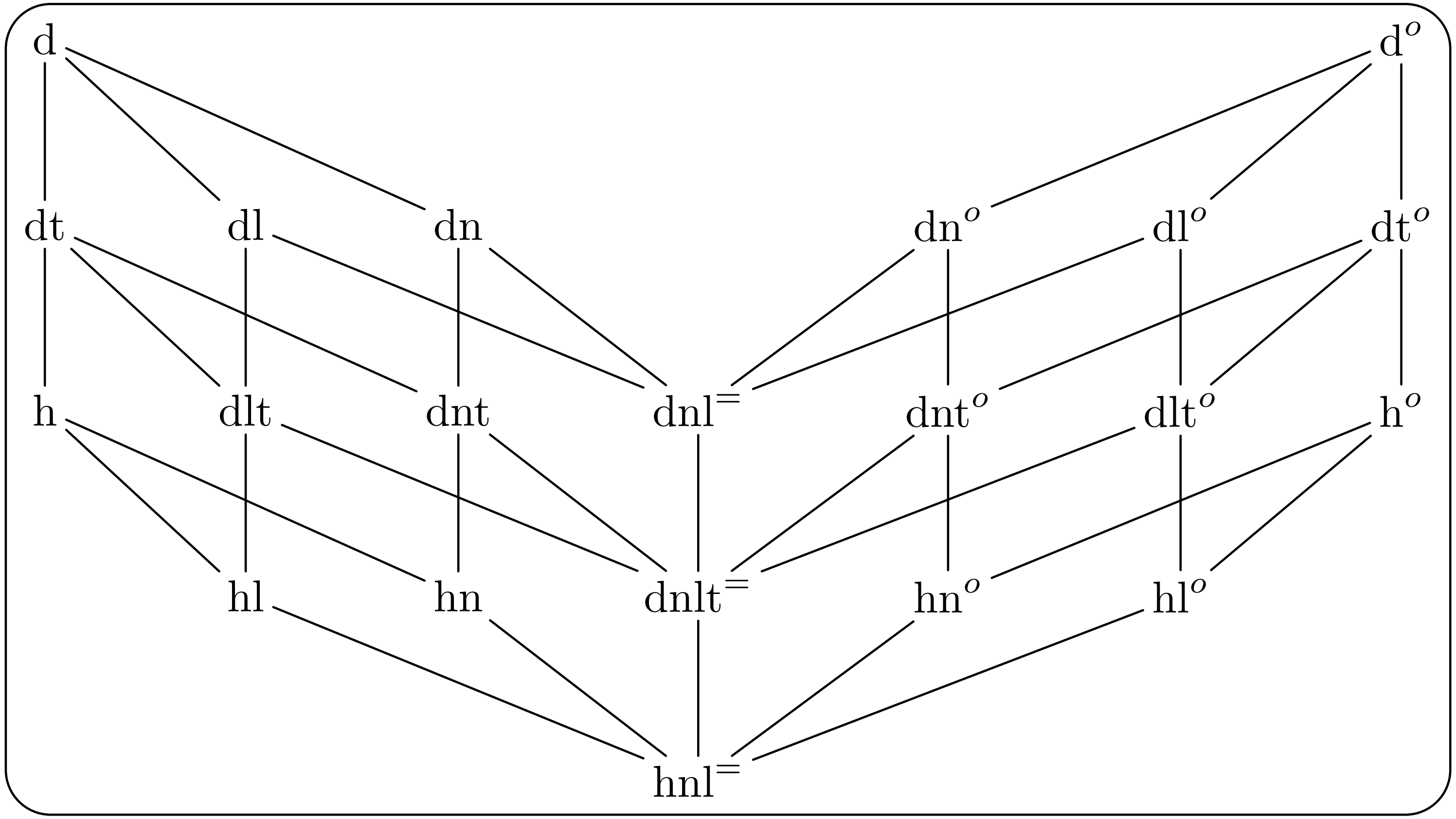
Commutative, Periodic, and Multiplicatively Regular Semirings



Commutative and Periodic Semirings



Non-Periodic Semirings



Remaining Questions and Literature

- Non-deterministic tree series transducers?
- Non-commutative, but periodic semirings?
- Top-down tree series transducers?

Some References:

Rounds: *Trees, Transducers and Transformations*, 1968

Thatcher: *Generalized*^{2} *Sequential Machine Maps*, 1970

Engelfriet: *Bottom-up and Top-down Tree Transformations - a Comparison*, 1975

Engelfriet, Fülöp, Vogler: *Bottom-up and Top-down Tree Series Transformations*, 2002

Fülöp, Vogler: *Tree Series Transformations that Respect Copying*, 2003

Fülöp, Vogler: *Weighted tree transducers*, 2003