Lecture 13: Semiring Parsing
Ma 191c: Mathematical Models of Generative Linguistics

Matilde Marcolli

Caltech, Spring 2024

this part based on

- Matilde Marcolli, Robert C. Berwick, Noam Chomsky, *Syntax-semantics interface: an algebraic model*, arXiv:2311.06189

also included in the book:
Matilde Marcolli, Noam Chomsky, Robert C. Berwick, "Mathematical structure of syntactic Merge", MIT Press.

semiring parsing

- relation between grammars and semirings first developed by Chomsky–Schützenberger (1963)
- then commonly used framework of *semiring parsing* for *context free* grammars (or mildly context sensitive like TAGs)
- main setting: deduction rules of the form

$$\frac{A_1 \ldots A_k}{B} C_1 \ldots C_\ell$$

  - terms A_i (main conditions) are rules R of the grammar or input nonterminals
  - C_i are (non-probabilistic) Boolean side conditions
  - fraction notation means that if the numerator terms hold then the denominator term also does
- to main conditions one assigns semiring values, combine with semiring operations, obtain value for deduced output
- Question: what type of algebraic structure replaces this form of semiring parsing in Minimalism based on free symmetric Merge action on workspaces?

Roadmap:

- first step (warmup): revisiting the idea of Minimal Search as an example of Birkhoff factorization of a character of the Hopf algebra of workspaces with *target* a ring of (Laurent series of) Merge derivations
- second step: need to incorporate Merge derivations as *source* of parsing, this requires passing from Hopf algebras to Hopf algebroids (composition on matching source/target of Merge action)
- third step: *target* of parsing correspondingly needs to adapt from algebras/semirings to (a suitable notion of) algebroids/semiringoids with a suitable notion of Rota–Baxter structure
- fourth step: then "semiring parsing" becomes Birkhoff factorization again, but in this "-iod" setting

First warmup step: revisiting Minimal Search

Recall from earlier: different forms of Merge and action on workspaces

- **EM:** $F = T \sqcup T' \sqcup \hat{F} \mapsto F' = \mathfrak{M}(T, T') \sqcup \hat{F}$
- **IM:** $F = T \sqcup \hat{F} \mapsto F' = \mathfrak{M}(T_v, T/T_v) \sqcup \hat{F}$
- **SM(i):** $F = T \sqcup T' \sqcup \hat{F} \mapsto F' = \mathfrak{M}(T_v, T'_w) \sqcup T/T_v \sqcup T'/T'_w \sqcup \hat{F}$
- **SM(ii):** $F = T \sqcup T' \sqcup \hat{F} \mapsto F' = \mathfrak{M}(T, T'_w) \sqcup T'/T'_w \sqcup \hat{F}$
- **C/SM(iii):** $F = T \sqcup \hat{F} \mapsto F' = \mathfrak{M}(T_v, T_w) \sqcup T/(T_v \sqcup T_w) \sqcup \hat{F}$

$\hat{F}$ denotes part of the workspace that remains unaffected

All but EM and IM are eliminated by Minimal Search and effect on size of workspaces (Resource Restriction and Minimal Yield): these are not two different mechanisms but the same

size counting again

- $b_0(F)$ number of connected components, $\alpha(F)$ number of accessible terms, $\sigma(F) = b_0(F) + \alpha(F) = \#V(F)$ and $\hat{\sigma}(F) = b_0(F) + \sigma(F)$
- chain of Merge derivations

$$\Phi = \mathfrak{M}_{S_N,S'_N} \circ \cdots \circ \mathfrak{M}_{S_1,S'_1}$$

- size change measured by

$$\delta b_0 := b_0(F) - b_0(\Phi(F)), \quad \delta\alpha = \alpha(\Phi(F)) - \alpha(F),$$

$$\delta\sigma = \sigma(\Phi(F)) - \sigma(F)\,,$$

- so Minimal Yield conditions equivalent to

$$\delta b_0 \geq 0, \quad \delta\alpha \geq 0, \quad \delta\sigma = 1\,,$$

respectively "no divergence", "no information loss", "minimality of yield"

- weaker condition of "positive yield" $\delta\sigma \geq 0$
- also for $\Phi_0 : \pi_0(F) \to \pi_0(\Phi(F))$ (following in which component the root of each component ends up)

$$\delta \deg_a := (\deg(\Phi_0(a)) - \deg(a)) \quad \text{for } a \in \pi_0(F)$$

- "no complexity loss" principle: for all $a \in \pi_0(F)$

$$\delta \deg_a \geq 0$$

algebra of Merge derivations

- $\mathcal{DM}$ is the commutative associative $\mathbb{Q}$-algebra with the underlying vector space spanned by

$$\varphi_A = (F \stackrel{\mathfrak{M}_A}{\to} F')$$

$A \subset \mathcal{SO} \times \mathcal{SO}$ set of pairs (S, S') of syntactic objects

$$F \stackrel{\mathfrak{M}_{S_1,S_1'}}{\longrightarrow} F_1 \to \cdots F_{N-1} \stackrel{\mathfrak{M}_{S_N,S_N'}}{\longrightarrow} F'$$

all possible chains of Merge operations with $(S_i, S_i') \in A$

- algebra multiplication, for $\varphi_A = (F \stackrel{\mathfrak{M}_A}{\to} F')$ and $\varphi_B = (\tilde{F} \stackrel{\mathfrak{M}_B}{\to} \tilde{F}')$

$$\varphi_A \sqcup \varphi_B = (F \sqcup \tilde{F} \stackrel{\mathfrak{M}_{A\sqcup B}}{\to} F' \sqcup \tilde{F}')$$

- meaning of product: perform in parallel different Merge operations that affect different parts of a workspace
- unit empty forest mapped to itself

Laurent series and Rota-Baxter operator

- commutative associative algebra $\mathcal{A}$ and the algebra of Laurent series $\mathcal{A}[t^{-1}][[t]]$
- linear operator $R : \mathcal{A}[t^{-1}][[t]] \to \mathcal{A}[t^{-1}][[t]]$ that projects onto the polar part

$$R(\sum_{i=-N}^{\infty} a_i t^i) = \sum_{i=-N}^{-1} a_i t^i$$

- makes $(\mathcal{A}[t^{-1}][[t]], R)$ a Rota–Baxter algebra of weight -1
- Note: this is the way to "subtract divergences" in physics
- here consider $\mathcal{DM}[t^{-1}][[t]]$ Laurent series with coefficients in the algebra of Merge derivations $\mathcal{DM}$

Hopf algebra character: the generative process for F

- the map $\phi : \mathcal{H} \to \mathcal{DM}$,

$$\phi(F) = (L(F) \stackrel{\mathfrak{M}_{A(L(F),F)}}{\longrightarrow} F)$$

  assigning to a forest F the set $A(L(F), F)$ of all Merge derivations starting from the (multi)set $L(F)$ of individual lexical items and syntactic features to the forest F

- this defines a character (morphism of commutative algebras) from the Merge Hopf algebra $\mathcal{H}$ to algebra $\mathcal{DM}$ Merge derivations

Hopf algebra character: Laurent series version

- as above but with

$$\phi_t(F) = (L(F) \stackrel{\mathfrak{M}_{A(L(F),F)}}{\longrightarrow} F) \; t^{\delta(\mathfrak{M}_{A(L(F),F)})}$$

where δ is either δb_0 or $\delta\alpha$ or $\delta\sigma$

- morphism of commutative algebras

$$\phi_t : \mathcal{H} \to \mathcal{DM}[t^{-1}][[t]]$$

trivial factorization in this case of "full derivations"

- this character always takes values in the subalgebra

$$\mathcal{D}\mathcal{M}[[t]] = (1-R)\,\mathcal{D}\mathcal{M}[t^{-1}][[t]]$$

- if $\#L = \ell \geq 2$ then $\#V(T) = 2\ell - 1 = \sigma(T)$ and $\alpha(T) = 2\ell - 2$, with $b_0(T) = 1$, so that we have $\delta b_0 = \ell - 1 \geq 0$, $\delta\alpha = \ell - 2 \geq 0$, $\delta\sigma = \ell - 1 \geq 0$

Hopf algebra character: further refinement "partial derivations"

- for $T \in \mathfrak{T}_{\mathcal{SO}_0}$ let $\mathcal{F}_T \subset \mathfrak{F}_{\mathcal{SO}_0} \times \mathfrak{F}_{\mathcal{SO}_0}$ be the set of pairs (F, F') of forests F with $L(F) = L(F') = L(T)$ that are intermediate derivations for T: $\exists$ chain of Merge derivations

$$L(T) \stackrel{\mathfrak{M}_{S_1,S_1'}}{\longrightarrow} \cdots \stackrel{\mathfrak{M}_{S_i,S_i'}}{\longrightarrow} F \stackrel{\mathfrak{M}_{S_{i+1},S_{i+1}'}}{\longrightarrow} \cdots \stackrel{\mathfrak{M}_{S_j,S_j'}}{\longrightarrow} F' \stackrel{\mathfrak{M}_{S_{j+1},S_{j+1}'}}{\longrightarrow} \cdots \stackrel{\mathfrak{M}_{S_m,S_m'}}{\longrightarrow} T$$

- this includes previous case with $F = L(T)$ and $F' = T$
- then form a character $\psi_t : \mathcal{H} \to \mathcal{DM}[t^{-1}][[t]]$ with

$$\psi_t(T) = \sum_{(F,F')\in\mathcal{F}_T} (F \stackrel{\mathfrak{M}_{A(F,F')}}{\longrightarrow} F')\, t^{\delta(\mathfrak{M}_{A(F,F')})}$$

Birkhoff factorization and Minimal Search

- in the case of partial derivations the character in general has a nontrivial projection to the polar part $R\,\mathcal{DM}[t^{-1}][[t]] = \mathcal{DM}[t^{-1}]$, not just to the convergent part
- inductively constructed Birkhoff factorization of character ψ_t implements a form of Minimal Search
- Bogolyubov preparation

$$\tilde{\psi}_t(T) = \psi_t(T) + \sum \psi_{t,-}(F_{\underline{v}})\psi_t(T/F_{\underline{v}})$$

analyzes in parallel the Merge derivations of all accessible terms of T, ensuring that the undesirable forms of Merge violating the size contraints are progressively removed and only derivations with $\delta \geq 0$ retained at each step

Birkhoff factorization implementing Minimal Search

- just projecting with $(1-R)$ on $\phi(F)$ not sufficient to get rid of unwanted forms of Merge with $\delta < 0$
- *but...* Birkhoff factorization achieves that result
- $\psi_{t,+}(T) = (1-R)\tilde{\psi}_t(T)$ alg homom $\psi_{t,+} : \mathcal{H} \to \mathcal{DM}[[t]]$

$$\tilde{\psi}_t(T) = \psi_t(T) + \sum \psi_{t,-}(F_{\underline{v}})\psi_t(T/F_{\underline{v}})$$

- more details: if there is a term in $\psi_t(T)$ of the form $(F \to F')t^\delta$ where the derivation has $\delta < 0$ the forest F' will occur as a collection of accessible terms $F' = F_{\underline{v}}$ in T
- so in $\tilde{\psi}_t(T)$ the term $\psi_{t,-}(F_{\underline{v}})\psi_t(T/F_{\underline{v}})$ will contain a term $R(\psi_t(F'))\psi_t(T/F_{\underline{v}})$ which will contain a summand equal to $-(F \to F')t^\delta$
- has the effect of removing the unwanted derivation, while any term $(F \to F')t^\delta$ in $\psi_t(T)$ that only contains derivations with $\delta \geq 0$ is not cancelled by anything coming from the terms $\psi_{t,-}(F_{\underline{v}})\psi_t(T/F_{\underline{v}})$

case of "no complexity loss" principle

- similar procedure for constructing Hopf algebra character, but multivariables for connected components
- a set of variables t_λ for $\lambda \in \mathcal{SO}_0$
- trees $T \in \mathrm{Dom}(h)$ (with a head function)
- assign to each $a \in \pi_0(F)$ a variable $t_a := t_{\lambda(h(T_a))}$
- then set

$$\delta \deg_a(F \xrightarrow{\Phi} F') = \deg(\Phi_0(a)) - \deg(a)$$

$$\phi_t(F) = \sum_{\Phi: F \to F'} (F \xrightarrow{\Phi} F') \prod_{a \in \pi_0(F)} t_{h(T_a)}^{\delta \deg_a(\Phi)}$$

- character with values in $\mathcal{DM}[[t_\lambda]][t_\lambda^{-1}]$
- Birkhoff factorization retains terms with only "no complexity loss" derivations $\delta \deg_a \geq 0$ in all the accessible terms

Next step: Birkhoff factorization in algebroids

- refine the previous construction: in $\mathcal{DM}$ commutative product only accounts for "independent" derivations that affect different parts of workspace (hence commute)
- want to incorporate all derivations in the algebraic structure
- something that generalizes "derivation forest semirings" of context-free semiring parsing
- key idea: composing Merge transformations on workspaces is like composing arrows (morphisms of a category), composition only defined when target of first arrow is source of second one
- difference between a *group* (composition always defined) and a *groupoid* (composition defined with matching target/source)
- commutative Hopf algebras are "dual to groups" (group schemes)... the notion dual to groupoids is *Hopf algebroids*

commutative Hopf algebroid (dually groupoid scheme)

- pair of commutative algebras $\mathcal{A}^{(0)}$ and $\mathcal{H}^{(1)}$
- for any other commutative algebra $\mathcal{R}$, sets $\mathcal{G}^{(0)}(\mathcal{R}) = \mathrm{Hom}(\mathcal{A}^{(0)}, \mathcal{R})$ and $\mathcal{G}^{(1)}(\mathcal{R}) = \mathrm{Hom}(\mathcal{H}^{(1)}, \mathcal{R})$ are the objects and morphisms of a groupoid $\mathcal{G}$
- unpack this:
  - pair of commutative algebras $(\mathcal{A}^{(0)}, \mathcal{H}^{(1)})$ (functions on objects and arrows)
  - homomorphisms $\eta_s, \eta_t : \mathcal{A}^{(0)} \to \mathcal{H}^{(1)}$ give $\mathcal{H}^{(1)}$ the structure of a $\mathcal{A}^{(0)}$-bimodule (dual to source and target)
  - coproduct given by morphism of $\mathcal{A}^{(0}$-bimodules

$$\Delta : \mathcal{H}^{(1)} \to \mathcal{H}^{(1)} \otimes_{\mathcal{A}^{(0)}} \mathcal{H}^{(1)}$$

    (dual to composition of arrows in groupoid)
  - conjugation $S : \mathcal{H}^{(1)} \to \mathcal{H}^{(1)}$ (dual to inverse of morphisms in groupoid)
- bialgebroid: same without S, dual to a "semigroupoid" (small category) instead of groupoid

further properties

- counit $\epsilon : \mathcal{H}^{(1)} \to \mathcal{A}^{(0)}$ morphism of $\mathcal{A}^{(0}$-bimodules (dual to identity morphisms)
- $\epsilon\eta_s = \epsilon\eta_t = 1$ (identity morphisms have same source and target)
- $(1 \otimes \epsilon)\Delta = (\epsilon \otimes 1)\Delta = 1$ (composition with the identity morphism)
- $(1 \otimes \Delta)\Delta = (\Delta \otimes 1)\Delta$ (associativity of composition of morphisms)
- $S^2 = 1$ and $S\eta_s = \eta_t$ (inversion is an involution and exchanges source and target of morphisms)

- composition of a morphism with its inverse is identity morphism $\eta_t \epsilon = \mu(S \otimes 1)\Delta$ and $\eta_s \epsilon = \mu(1 \otimes S)\Delta$ $\mu : \mathcal{H}^{(1)} \otimes_{\mathcal{A}^{(0)}} \mathcal{H}^{(1)} \to \mathcal{H}^{(1)}$ extending the algebra multiplication $\mu : \mathcal{H}^{(1)} \otimes_{\mathbb{Q}} \mathcal{H}^{(1)} \to \mathcal{H}^{(1)}$
- $\Delta \eta_s = 1 \otimes \eta_s$, $\Delta \eta_t = \eta_t \otimes 1$ (source of composition of arrows is source of the first and target of composition is target of second)
- morphism $f : (\mathcal{A}_1^{(0)}, \mathcal{H}_1^{(1)}) \to (\mathcal{A}_2^{(0)}, \mathcal{H}_2^{(1)})$: algebra homomorphisms $f^{(0)} : \mathcal{A}_1^{(0)} \to \mathcal{A}_2^{(0)}$ and $f^{(1)} : \mathcal{H}_1^{(1)} \to \mathcal{H}_2^{(1)}$ with $f^{(0)} \circ \epsilon_1 = \epsilon_2 \circ f^{(1)}$, $f^{(1)} \circ \eta_{s,1} = \eta_{s,2} \circ f^{(0)}$, $f^{(1)} \circ \eta_{t,1} = \eta_{t,2} \circ f^{(0)}$, $f^{(1)} \circ S_1 = S_2 \circ f^{(1)}$, $\Delta_2 \circ f^{(1)} = (f^{(1)} \otimes f^{(1)}) \circ \Delta_1$

bialgebroid of Merge derivations (replaces "derivation forests")

- data $\mathcal{A}^{(0)} = (\mathcal{V}(\mathfrak{F}_{\mathcal{SO}_0}), \sqcup)$ and $\mathcal{H}^{(1)} = (\mathcal{DM}, \sqcup)$, define a bialgebroid
  - left and right $\mathcal{A}^{(0)}$-module structures (source and target)

$$\eta_s(F)\varphi_A = \begin{cases} \varphi_A & s(\varphi_A) = F \\ 0 & \text{otherwise} \end{cases} \qquad \eta_t(F)\varphi_A = \begin{cases} \varphi_A & t(\varphi_A) = F \\ 0 & \text{otherwise} \end{cases}$$

  - coproduct

$$\Delta(\varphi_A) = \varphi_A \otimes 1 + 1 \otimes \varphi_A + \sum_{\varphi_A = \varphi_{A_1} \circ \varphi_{A_2}} \varphi_{A_1} \otimes \varphi_{A_2}$$

where $\varphi_{A_2} = (F \stackrel{\mathfrak{M}_{A_2}}{\to} F')$ and $\varphi_{A_1} = (F' \stackrel{\mathfrak{M}_{A_1}}{\to} F'')$ with composition

$$\varphi_{A_1} \circ \varphi_{A_2} = (F \stackrel{\mathfrak{M}_{A_1 \circ A_2}}{\to} F'')$$

$\mathfrak{M}_{A_1 \circ A_2} = \mathfrak{M}_{A_1} \circ \mathfrak{M}_{A_2}$ set of all compositions of a chain of Merge derivations in set A_2 followed by one in A_1

- Note: coproduct of $\mathcal{V}(\mathfrak{F}_{\mathcal{SO}_0})$ (Hopf algebra of workspaces) now built into the arrows φ_A as part of Merge action

Rota-Baxter algebroids

- Merge bialgebroid as replacement of context-free derivation forests
- want then replacement of semiring parsing using again the Birkhoff factorization idea
- instead of $\phi : \mathcal{H} \to \mathcal{R}$ from Hopf algebra to Rota-Baxter algebra (or semiring) need analog from Hopf *algebroids* (or bialgebroid) to a suitable generalization of a Rota-Baxter algebra (or semiring) ... *algebroid* (*semiringoid*)
- Warning: different notions of *algebroid*, *semiringoid* are used in math, ours is motivated by compatibility with the notion of Hopf algebroid and bialgebroid
- so here algebroid like part of bialgebroid that does not involve coproduct
  - pair of commutative algebras $(\mathcal{A}, \mathcal{E})$
  - two morphisms $\eta_s, \eta_t : \mathcal{A} \to \mathcal{E}$ that make $\mathcal{E}$ bimodule over $\mathcal{A}$
  - morphism of $\mathcal{A}$-bimodules $\epsilon : \mathcal{E} \to \mathcal{A}$ with $\epsilon\eta_s = \epsilon\eta_t = 1_{\mathcal{A}}$

- if Hopf algebroids are like functions on a groupoid, bialgebroids on a small category (semigroupoid), what about algebroids? ... functions on a *directed graph*
- $(\mathcal{A}, \mathcal{E})$ commutative algebroid: for any commutative algebra $\mathcal{R}$ the sets $V(\mathcal{R}) = \mathrm{Hom}(\mathcal{A}, \mathcal{R})$ and $E(\mathcal{R}) = \mathrm{Hom}(\mathcal{E}, \mathcal{R})$ are sets of vertices and edges of a directed graph $G(\mathcal{R})$ with source and target maps $s, t : E(\mathcal{R}) \to V(\mathcal{R})$ determined by the morphisms $\eta_s, \eta_t : \mathcal{A} \to \mathcal{E}$ *(directed graph scheme)*
- also each vertex $v \in V(\mathcal{R})$ has a looping edge $e_v \in E(\mathcal{R})$ with $s(e_v) = t(e_v) = v$
- bialgebroid = case where the directed graph satisfies reflexivity and transitivity (small category); Hopf algebroid = case where reflexivity, transitivity, symmetry (groupoid)

Rota-Baxter structure on algebroids

- commutative algebroid $(\mathcal{A}, \mathcal{E})$ with pair of maps $R = (R_V, R_E)$ with $R_V \in \mathrm{End}(\mathcal{A})$ algebra hom and $R_E : \mathcal{E} \to \mathcal{E}$ linear

  $$R_E(\eta_s(a) \cdot \xi) = \eta_s(R_V(a)) \cdot R_E(\xi) \quad R_E(\eta_t(a) \cdot \xi) = \eta_t(R_V(a)) \cdot R_E(\xi)$$

  and $\epsilon \circ R_E = R_E \circ \epsilon$ with Rota–Baxter identity (weight -1)

  $$R_E(\xi) \cdot R_E(\zeta) = R_E(R_E(\xi) \cdot \zeta) + R_E(\xi \cdot R_E(\zeta)) - R_E(\xi \cdot \zeta)$$

  normalization $R_E(1_{\mathcal{E}}) = 0$ or $R_E(1_{\mathcal{E}}) = 1_{\mathcal{E}}$, for $1_{\mathcal{E}}$ the unit of the algebra $\mathcal{E}$
- Main example: functions on edges of a directed graph, with values in a Rota-Baxter algebra with Rota-Baxter operator acting only on coefficients of functions
  - G directed graph, $(\mathcal{R}, R)$ Rota–Baxter algebra
  - $\mathcal{A} = \mathbb{Q}[V_G]$ and $\mathcal{E} = \mathbb{Q}[E_G] \otimes_{\mathbb{Q}} \mathcal{R}$
  - morphisms $\eta_s, \eta_t : \mathcal{A} \to \mathcal{E}$ precomposition with $s, t : E_G \to V_G$
  - $R_V = \mathrm{id}$ and $R_E = 1 \otimes R$

semiringoids

- $(\mathcal{A}, \mathcal{E})$ two commutative semirings
- semiring homomorphisms $\eta_s, \eta_t : \mathcal{A} \to \mathcal{E}$ that make $\mathcal{E}$ bi-semimodule over $\mathcal{A}$
- bi-semimodule homomorphism $\epsilon : \mathcal{E} \to \mathcal{A}$ with $\epsilon\eta_s = \epsilon\eta_t = 1_{\mathcal{A}}$

Rota-Baxter semiringoid (weight $+1$)

- semiringoid $(\mathcal{A}, \mathcal{E})$ with semiring endomorphism $R_V : \mathcal{A} \to \mathcal{A}$ and an $R_E : \mathcal{E} \to \mathcal{E}$ morphism of $\mathbb{Z}_{\geq 0}$-semimodules with

$$R_E(\eta_s(a) \odot \xi) = \eta_s(R_V(a)) \odot R_E(\xi)$$

$$R_E(\eta_t(a) \odot \xi) = \eta_t(R_V(a)) \odot R_E(\xi)$$

and $\epsilon \circ R_E = R_E \circ \epsilon$, with Rota–Baxter relation of weight $+1$

$$R_E(\xi) \odot R_E(\zeta) = R_E(R_E(\xi) \odot \zeta) \boxdot R_E(\xi \odot R_E(\zeta))$$

Birkhoff factorization in algebroids/semiringoids

- $(\mathcal{A}^0, \mathcal{H}^{(1)})$ Hopf algebroid and $(\mathcal{A}, \mathcal{E})$ algebroid with Rota–Baxter structure (R_V, R_E) weight -1
- morphism $\Phi : (\mathcal{A}^{(0)}, \mathcal{H}^{(1)}) \to (\mathcal{A}, \mathcal{E})$ of algebroids
- have inductive construction of factorization

$$\Phi_{+,E}(f) = (\Phi_{-,E} \star \Phi_E)(f) = (\Phi_{-,E} \otimes \Phi_E)(\Delta f)$$

$$\Phi_{-,E}(f) = -R_E(\tilde{\Phi}_E(f))$$

$$\tilde{\Phi}_E(f) = \Phi_E(f) + \sum \Phi_{-,E}(f')\Phi_E(f'')$$

for $\Delta(f) = f \otimes 1 + 1 \otimes f + \sum f' \otimes f''$, and with $\Phi_{+,E}(f) = (1 - R_E)(\tilde{\Phi}_E(f))$

Similar for case of semiringoids

what does this mean?

- algebroid $(\mathcal{A},\mathcal{E}) \Leftrightarrow$ directed graph G
- $\Phi : (\mathcal{A}^{(0)},\mathcal{H}^{(1)}) \to (\mathcal{A},\mathcal{E})$ map of graphs $\alpha : G \to \mathcal{G}$ (where $\mathcal{G}$ also a category)
- take $f = \delta_\gamma$ for γ an arrow in $\mathcal{G}$
- if $R = \mathrm{id}$ (trivial RB structure weight -1)

$$\tilde{\Phi}_E(\delta_\gamma) = \sum_{e\in E_G\,:\,\alpha(e)=\gamma} \delta_e + \cdots + \sum_{e_1,\ldots,e_n\in E_G\,:\,\gamma=\alpha(e_1)\circ\cdots\circ\alpha(e_n)} \delta_{e_1}\cdots\delta_{e_n}$$

lists all the possible ways of obtaining γ as compositions of images of arrows in G

- for a weighted combination $\sum_i \lambda_i e_i$ in diagram G (eg probability) with R_E RB

$$\Phi_{E,-}(\delta_\gamma)(\sum_i \lambda_i e_i) = -(\sum_{\alpha(e)=\gamma} R_E(\lambda_e) + \sum_{\alpha(e_1)\circ\alpha(e_2)=\gamma} R_E(R_E(\lambda_{e_1})\lambda_{e_2}) + \cdots$$

$$+ \sum_{\alpha(e_1)\circ\cdots\circ\alpha(e_n)=\gamma} R_E(\cdots(R_E(\lambda_{e_1})\cdots)\lambda_{e_n}))$$

- case of Merge derivations: target functions on a graph values in semiring $(\mathbb{R} \cup \{-\infty\}, \max, +)$ with R ReLU or in Viterbi $([0,1], \max, \cdot)$ with $R = c_\lambda$ threshold
- map Φ assigns a possible *diagram of Merge derivations*
- *checking all possible ways of realizing some given chain of Merge derivations* γ *through compositions coming from the chosen diagram, weighted by elements in the given semiring and filtered by* R
  - $\mathcal{R} = (\mathbb{R} \cup \{-\infty\}, \max, +)$ with R =ReLU: all possibilities with weights of each step above the ReLU threshold
  - $\mathcal{R} = ([0,1], \max, \cdot)$ with the threshold $R = c_\lambda$ all possibilities with probabilities of each step above threshold
  - Boolean semiring $\mathcal{B} = (\{0,1\}, \max, \cdot)$ with $R = \mathrm{id}$: derivations γ realized through diagram G with truth values on each edge and composition of arrows = AND operation on truth values, different paths of derivations = OR operation on truth values

Example

- consider the chain of Merge derivations for the sentence "many people praise many people" (example from "Merge & SMT" §3.4) and choice of model diagram G for parsing