Lecture 9: Comparison with Older Minimalism
Ma 191c: Mathematical Models of Generative Linguistics

Matilde Marcolli

Caltech, Spring 2024

this part based on

- Matilde Marcolli, Robert C. Berwick, Noam Chomsky, *Old and new Minimalism: a Hopf algebra comparison*, arXiv:2306.10270

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

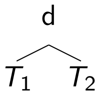
Comparison with Old Minimalism: Stabler's computational minimalism

- constructing I/E Merge directly on planar trees
- including labeling and domains for applicability based on labels
- Hopf algebras of planar binary rooted trees (Loday–Ronco Hopf algebra)
  1. no workspaces: only work with trees (not forests) makes compatible product and coproduct structure more difficult
  2. partially defined Merge operations (feature checking) introduces further layers of algebraic structure
- role of I/E Merge in terms of Loday–Ronco Hopf algebra
- different structures for Internal and External Merge (not coming from same operation)
  - Internal Merge determines system of right-ideal coideals (weak notion of quotient)
  - External Merge determines partially defined operated algebra

planar binary rooted trees

- $\mathcal{V} = \oplus_{k\geq 0}\mathcal{V}_k$ vector space spanned by planar binary rooted tree, k = number of non-leaf vertices ($k+1$ leaves)
- now will have labeling of internal vertices also D_V set of labels
- for $d \in D_V$ grafting operator $\wedge_d$

$$\wedge_d : \mathcal{V}\otimes\mathcal{V}\to\mathcal{V}, \quad T_1\otimes T_2 \mapsto T = T_1 \wedge_d T_2 =$$



- $S\backslash T$ (S under T) grafting root of T to rightmost leaf of S
- T/S (S over T) grafting the root of T to leftmost leaf of S
- $T_1 \wedge_d T_2 = T_1/S\backslash T_2$ with S planar binary tree with single non-leaf vertex decorated by $d \in D_V$
- each planar rooted tree is $T = T_\ell \wedge_d T_r$ (left and right subtrees below root)

Loday–Ronco Hopf algebra of planar binary rooted trees $\mathcal{H}_{LR}$

- vector space $\mathcal{V}_k$ spanned by planar binary rooted trees T with k internal vertices (hence $k+1$ leaves)

$$\dim \mathcal{V}_k = (\# D_V)^k \frac{(2k)!}{k!(k+1)!}$$

$\# D_V$ cardinality of set D_V of vertex labels

- graded vector space $\mathcal{V} = \oplus_{k \geq 0} \mathcal{V}_k$ with $\mathcal{V}_0 = \mathbb{Q}$
- given label $d \in D_V$, grafting operator $\wedge_d$

$$\wedge_d : \mathcal{V} \otimes \mathcal{V} \to \mathcal{V}, \quad T_1 \otimes T_2 \mapsto T = T_1 \wedge_d T_2$$

with $\wedge_d : \mathcal{V}_k \otimes \mathcal{V}_\ell \to \mathcal{V}_{k+\ell-1}$

- attaching the two roots v_{r_1} of T_1 and v_{r_2} of T_2 to a single root vertex v labelled by $d \in D_V$

Loday–Ronco Hopf algebra $\mathcal{H}_{LR}$

- vector space $\mathcal{V} = \oplus_{k\geq 0}\mathcal{V}_k$ with $\mathcal{V}_0 = \mathbb{Q}$
- multiplication and a comultiplication inductively by degrees
- trees $T = T_\ell \wedge T_r$ and $T' = T'_\ell \wedge T'_r$ with product

$$T \star T' = T_\ell \wedge (T_r \star T') + (T \star T'_\ell) \wedge T'_r$$

- coproduct

$$\Delta(T) = \sum_{j,k} (T_{\ell,j} \star T_{r,k}) \otimes (T'_{\ell,n-j} \wedge T'_{r,m-k}) + T \otimes \bullet$$

with $T = T_\ell \wedge T_r$ and $\Delta(T_\ell) = \sum_j T_{\ell,j} \otimes T'_{\ell,n-j}$ and $\Delta(T_r) = \sum_k T_{r,k} \otimes T'_{r,m-k}$ for $T_\ell \in \mathcal{V}_n$ and $T_r \in \mathcal{V}_m$

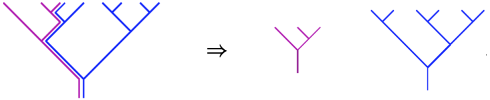
- antipode on graded bialgebras inductively

$$S(X) = -X - \sum S(X')X''$$

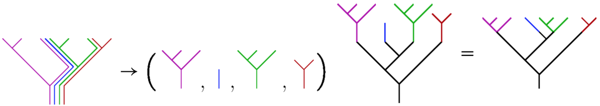
for $\Delta(X) = X \otimes 1 + 1 \otimes X + \sum X' \otimes X''$ lower deg X', X''

Graphical form of Loday-Ronco product/coproduct

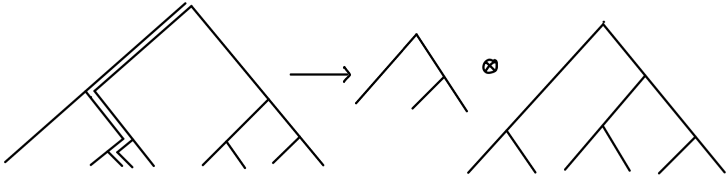
- product and coproduct defined inductively by degrees
- can also see graphically
  - coproduct sum $\Delta(T) = \sum T' \otimes T''$ over all decompositions of tree along paths from one of leaves to root



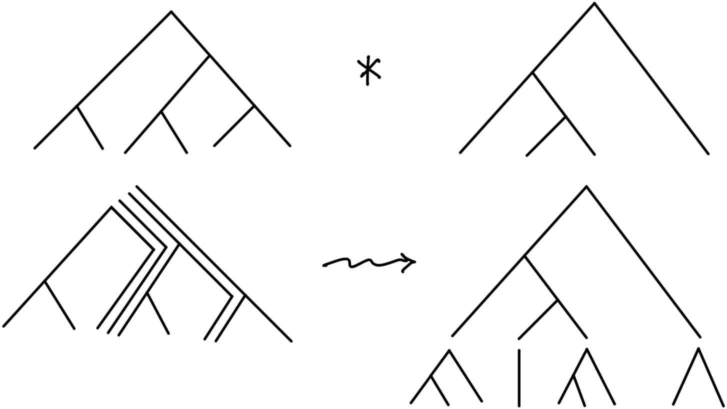
  - product $T \star T' = \sum_{(T_0,\dots,T_n)} \gamma(T_0,\dots,T_n;T')$ using same decompositions of first tree into as many pieces as leaves of second tree then grafting to leaves



  - antipode inductively constructed by degrees



terms of Loday-Ronco coproduct



terms of Loday-Ronco product

Hopf algebra comparison

(Noncommutative) Connes-Kreimer Hopf algebra $\mathcal{H}^{nc}_{CK}$

- Hopf algebra of planar rooted forests (not necessarily binary)
- (noncommutative) algebra freely generated by the planar rooted trees T
- coproduct: sum over all admissible cuts

$$\Delta(T) = T \otimes 1 + 1 \otimes T + \sum_C \pi_C(T) \otimes \rho_C(T)$$

- grading: planar rooted trees with k internal vertices
- antipode defined inductively on graded bialgebras

Relation to Loday-Ronco

- map of Hopf algebras $\phi : \mathcal{H}_{CK}^{nc} \to \mathcal{H}_{LR}$
- maps unit $1 \in \mathcal{H}_{CK}^{nc}$ (empty tree) to binary tree consisting of single root vertex $\bullet$
- maps single vertex tree $\bullet$ in $\mathcal{H}_{CK}^{nc}$ to binary tree with a single internal vertex (one root and two leaves)
- otherwise maps

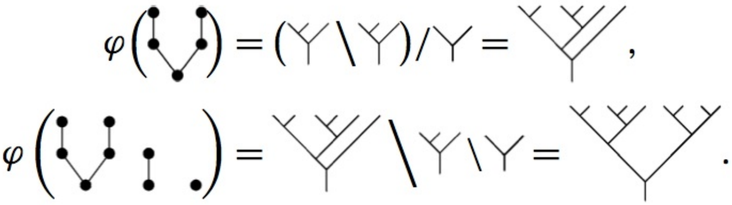
$$\phi(T) = \phi(F(T))/\phi(\bullet)$$

  with $F(T)$ forest obtained by removing root of T and $/$ is the concatenation operation grafting root of $\phi(F(T))$ to left leaf of $\phi(\bullet)$
- for a forest $F = T_1 \cdots T_n$ in $\mathcal{H}_{CK}$ image

$$\phi(F) = \phi(T_1)\backslash\phi(T_2)\backslash \cdots \backslash\phi(T_n)$$

  with $\backslash$ the other concatenation operation grafting root of $\phi(T_{i+1})$ to rightmost leaft of $\phi(T_i)$
- compatible with product and coproduct and antipode

Example



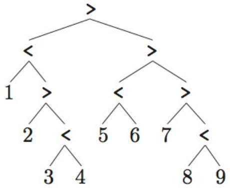
Some references

- M. Aguiar, F. Sottile, *Structure of the Loday–Ronco Hopf algebra of trees*, Journal of Algebra, Vol.295 (2006) 473–511
- J.L. Loday, M. Ronco, *Hopf algebra of the planar binary trees*, Adv. Math. 139 (1998) N.2, 293–309

Stabler's Computational Minimalism

- example of old formulation of Minimalism

  (Stabler's formulation also known for relation to formal languages)
- planar binary rooted trees with labels:
  - leaves labelled by lexical items and syntactic features $X \in \{N, V, A, P, C, T, D, \ldots\}$
  - also "selector" features σX for *head* selecting a phrase XP
  - can also have labels that are strings (ordered finite sets) $\alpha = X_0 X_1 \cdots X_r$ of syntactic features
  - labels "licensor" ω and "licensee" $\bar{\omega}$
  - internal vertices labelled by $\{>, <\}$ following *head* of subtree



External and Internal Merge: combinatorial structure

- External Merge

$$\mathcal{E}(T_1 \otimes T_2) = \begin{cases} \bullet \wedge T_2 & T_1 = \bullet \\ T_2 \wedge T_1 & \text{otherwise,} \end{cases}$$

- Internal Merge

$$\mathcal{I}(T) = \pi_C(T) \wedge \rho_C(T)$$

  C elementary admissible cut of T with $\rho_C(T)$ pruned tree containing root of T and $\pi_C(T)$ part severed by cut (elementary cut: tree not forest)

Note: admissible cuts are *not* the Loday-Ronco Hopf algebra coproduct now: there is a relation to the Loday-Ronco coproduct but is more involved

External Merge: domain

- $T[\alpha]$ for tree where head label starts with α
- domain of External Merge

$$\mathrm{Dom}(\mathcal{E}) = \mathrm{span}_{\mathbb{Q}}\{(T_1[\beta], T_2[\alpha]) \,|\, \beta = \sigma\alpha\}$$

- for $\alpha = X_0 X_1 \cdots X_r$ or $\alpha = \sigma X_0 X_1 \cdots X_r$ take $\hat{\alpha} = X_1 \cdots X_r$
- External Merge

$$\mathcal{E}(T_1[\sigma\alpha], T_2[\alpha]) = \begin{cases} T_1[\widehat{\sigma\alpha}] \wedge_< T_2[\hat{\alpha}] & |T_1| = 1 \\ T_2[\hat{\alpha}] \wedge_> T_1[\widehat{\sigma\alpha}] & |T_1| > 1 \end{cases}$$

Internal Merge: domain

- tree $T[\alpha]$ where $\alpha = X_0 \cdots X_r$ or $\alpha = \sigma X_0 \cdots X_r$ or $\alpha = \omega X_0 \cdots X_r$ or $\alpha = \bar{\omega} X_0 \cdots X_r$
- domain of Internal Merge

$$\mathrm{Dom}(\mathcal{I}) = \mathrm{span}_{\mathbb{Q}} \left\{ T[\alpha] \,|\, \exists T_1[\beta] \subset T[\alpha], \text{with } \begin{array}{l} \beta = \bar{\omega} X_0 \hat{\beta}, \\ \alpha = \omega X_0 \hat{\alpha} \end{array} \right\}$$

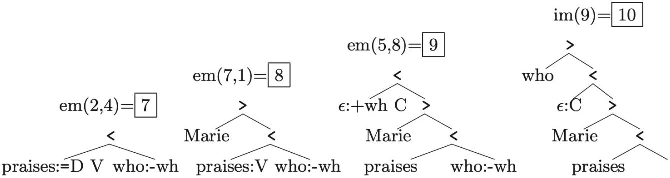
- Internal Merge (Stabler's notation and admissible cuts notation)

$$\mathcal{I}(T[\alpha]) = T_1^M[\hat{\beta}] \wedge_> T\{T_1[\beta]^M \to \emptyset\} = \pi_C(T) \wedge_> \rho_C(T)$$

Note: there are issues with EM and IM in this form (unlabelable exocentric constructions) producing $\{XP, YP\}$ results (observation by Riny Huijbregts)

Examples from Stabler

0	Pierre::D	who::D -wh	4
1	Marie::D	ϵ::=V +wh C	5
2	praises::=D =D V	knows::=C =D V	6
3	ϵ::=V C		



domains under iteration (compounding problem)

- iterations of the internal merge, $\mathcal{D}_{N+1} \subset \mathcal{D}_N$ with $\mathcal{D}_N := \mathrm{Dom}(\mathcal{I}^N)$
  1. N (complete) subtrees $T_1, \ldots, T_N$ in T
  2. $T_1^M, \ldots, T_N^M$ maximal projections of subtrees (also complete subtrees)
  3. subtrees T_i^M are disjoint.

$$\mathrm{Dom}(\mathcal{I}^N) = \left\{ T[\alpha] \,|\, \exists T_1[\beta^{(1)}], \ldots, T_N[\beta^{(N)}] \right.$$

$$\left. \text{with} \quad \begin{array}{l} (1), (2), (3) \text{ are satisfied} \\ \beta_0^{(1)} = \bar{\omega} X_0, \ldots, \beta_0^{(N)} = \bar{\omega} X_{N-1} \\ \alpha = \omega X_0 \omega X_1 \cdots \omega X_{N-1} \cdots \end{array} \right\}$$

$$\mathcal{I}^{\#C}(T[X]) = \bigwedge^{1+\#C} \left(\pi_C(T)[\hat{\mathbb{Y}}] \; \rho_C(T)[\hat{X}^N] \right)$$

$$\pi_C(T)[\hat{\mathbb{Y}}] = T_N^M[\hat{\beta}^{(N)}] \cdots T_1^M[\hat{\beta}^{(1)}]$$

label $[\hat{\alpha}^N]$ of tree $\rho_C(T)$: what remains of original label X after removing initial terms $\omega X_0 \omega X_1 \cdots \omega X_{N-1}$

feature checking complexity

- vertex labeling: strings $\alpha = X_0 X_1 \dots X_r$ of syntactic features
- set Σ_ℓ strings of length ℓ with $\#\Sigma_\ell = \mathfrak{s}^\ell$ with $\mathfrak{s}$ total number of syntactic features
- $\Sigma_\ell(a_0 \dots a_r)$ sequences starting with $a_0, \dots, a_r$
- counting formula for planar binary rooted trees with k internal vertices labeled by Σ_ℓ

$$d_{k,\ell} = (\#\Sigma_\ell)^k \frac{(2k)!}{k!(k+1)!} = \mathfrak{s}^{k\,\ell} \frac{(2k)!}{k!(k+1)!}$$

- trees with given label $\alpha \in \Sigma_\ell$ at root vertex and arbitrary labels elsewhere

$$d_{k,\ell}(\alpha) = (\#\Sigma_\ell)^{k-1} \frac{(2k)!}{k!(k+1)!} = \mathfrak{s}^{(k-1)\,\ell} \frac{(2k)!}{k!(k+1)!}$$

feature checking: one application of Internal Merge

- necessary condition defining domain

$$\mathcal{R}_{\mathcal{I}} := \{(\alpha, \beta) \in \Sigma_\ell \times \Sigma_\ell \,|\, \alpha = \omega X_0 \hat{\alpha}\,, \quad \beta = \bar{\omega} X_0 \hat{\beta}\}$$

- dimension of $\mathcal{D}_{1,k,\ell}(\alpha) = \mathcal{D}_1 \cap \mathcal{V}_{ling,k,\ell}(\alpha)$ for $\mathcal{D}_1 = \mathrm{Dom}(\mathcal{I})$

$$d_{\mathcal{I},k,\ell}(\alpha) = (\mathfrak{s}^{(k-1)\,\ell} - \mathfrak{s}^{(k-1)(\ell-2)}(\mathfrak{s}^2 - 1)^{k-1}) \frac{(2k)!}{k!(k+1)!}$$

- because with root label $a_0 a_1 \dots$ fixed counting all possible ways of having (at least) one of the vertices labeled by $\Sigma_\ell(a_0 a_1)$
- same as all the assignments of labels in Σ_ℓ not all of them in the complement $\Sigma_\ell \smallsetminus \Sigma_\ell(a_0 a_1)$ and

$$\mathfrak{s}^\ell - \mathfrak{s}^{\ell-2} = \mathfrak{s}^{\ell-2}(\mathfrak{s}^2 - 1) = \#(\Sigma_\ell \smallsetminus \Sigma_\ell(a_0 a_1))$$

feature checking: repeated applications of Internal Merge

- necessary condition defining domain

$$\mathcal{R}_{\mathcal{I}^N} = \left\{ (\alpha, \beta_1, \ldots, \beta_N) \,\middle|\, \begin{array}{l} \alpha = \alpha = \omega X_0 \omega X_1 \cdots \omega X_{N-1} \cdots \\ \beta_1 = \bar{\omega} X_0, \ldots \\ \cdots \\ \beta_N = \bar{\omega} X_{N-1} \end{array} \right\}$$

- some counting functions

$$S_N(a,b) := \binom{k-1}{N} b^N (a-b)^{k-1-N}$$

$$S_{N,k}(a,b) := S_N(a,b) + S_{N+1}(a,b) + \cdots + S_{k-1}(a,b) \leq a^{k-1}$$

- $S_N(a,b)$ counts number of label assignments to a set of $k-1$ points where N of them have labels in a set $B \subset A$ with $b = \#B$, $a = \#A$ and the remaining $k-1-N$ have labels in the complement $A \smallsetminus B$

- dimension of $\mathcal{D}_{N,k,\ell}(\alpha) = \mathcal{D}_N \cap \mathcal{V}_{ling,k,\ell}(\alpha)$, with $D_N = \mathrm{Dom}(\mathcal{I}^N)$ is

$$d_{\mathcal{I},k,\ell,N}(\alpha) = S_{N,k}(\mathfrak{s}^\ell, \mathfrak{s}^{\ell-2N}(\mathfrak{s}^{2N} - 1))\frac{(2k)!}{k!(k+1)!}$$

- again use

$$\mathfrak{s}^\ell - \mathfrak{s}^{\ell-2N} = \mathfrak{s}^{\ell-2N}(\mathfrak{s}^{2N} - 1) = \#(\Sigma_\ell \smallsetminus \Sigma_\ell(a_0 \dots a_{2N-1}))$$

- counting all the possible ways in which among the $k-1$ labels assigned to root vertices at least N are not in the complement of $\Sigma_\ell(a_0 \dots a_{2N-1})$

  (analysis of complexity of computational implementations of Minimalism, see also Indurkya 2020, 2021; also Berwick succintness result compared to formal language description)

Internal Merge and coproduct and product in $\mathcal{H}_{LR}$

- coproduct $\Delta(T) = \sum T' \otimes T''$ decompositions: in each one side contains *head* of tree T (both if on boundary line of cut)
- if *head* of T and *head* of $\pi_C(T)$ same side then that side is in $\mathrm{Dom}(\mathcal{I})$
- so pieces of the coproduct are in $\mathrm{Dom}(\mathcal{I}) \otimes \mathcal{H}_{ling} + \mathcal{H}_{ling} \otimes \mathrm{Dom}(\mathcal{I})$
- other terms (different sides) are in $\mathcal{H}_{ling} \otimes \mathcal{H}_{ling}$
- T in $\mathrm{Dom}(\mathcal{I})$ and C elementary admissible determined by label condition; set of partitions

$$\mathcal{P}_{\mathcal{I}}(T) = \{ T = (T', T'') \,|\, \begin{array}{l} (h(T) \in T' \text{ and } h(\pi_C(T)) \in T') \text{ or} \\ (h(T) \in T'' \text{ and } h(\pi_C(T)) \in T'') \end{array} \}$$

- modify coproduct

$$\Delta_{\mathcal{I}}(T) := \sum_{(T',T'') \in \mathcal{P}_{\mathcal{I}}(T)} T' \otimes T''$$

and remains same outside of $\mathrm{Dom}(\mathcal{I})$

- with this modified coproduct $\mathrm{Dom}(\mathcal{I})$ is a *coideal* of the coalgebra $\mathcal{H}_{ling}$

$$\Delta_{\mathcal{I}}(\mathrm{Dom}(\mathcal{I})) \subset \mathrm{Dom}(\mathcal{I}) \otimes \mathcal{H}_{ling} + \mathcal{H}_{ling} \otimes \mathrm{Dom}(\mathcal{I})$$

- modified product $\star_{\mathcal{I}}$ on $\mathcal{H}_{ling}$: for trees T, T' where T' has $n+1$ leaves: decompositions where *head* of T in component grafted to *head* of T'

$$\mathcal{P}_{\mathcal{I}}(T, T') := \{(T_0, \ldots, T_n) \,|\, h(T) \text{ and } h(\pi_C(T)) \in T_{h(T')}\}$$

$$T \star_{\mathcal{I}} T' = \sum_{(T_0,\ldots,T_n)\in\mathcal{P}_{\mathcal{I}}(T,T')} \gamma(T_0, \ldots, T_n; T')$$

- $h(T \star_{\mathcal{I}} T') = h(T)$ as *head* of each $\gamma(T_0, \ldots, T_n; T')$ same as the *head* of T
- component $T_{h(T')}$ is in $\mathrm{Dom}(\mathcal{I})$ when $T \in \mathrm{Dom}(\mathcal{I})$ so $T \star_{\mathcal{I}} T'$ also in $\mathrm{Dom}(\mathcal{I})$
- $\mathrm{Dom}(\mathcal{I})$ right-ideal of algebra $(\mathcal{H}_{ling}, \star_{\mathcal{I}})$

$$\mathrm{Dom}(\mathcal{I}) \star_{\mathcal{I}} \mathcal{H}_{ling} \subset \mathrm{Dom}(\mathcal{I})$$

- not left-ideal

[planar trees so noncommutative product, and left and right ideals differ]

- form of Internal Merge $\mathcal{I}(T \star_{\mathcal{I}} T') =$

$$\sum_{(T_0,\ldots,T_n)\in\mathcal{P}_{\mathcal{I}}(T,T')} \pi_C(T_{h(T')}) \wedge_> \gamma(T_0,\ldots,\rho_C(T_{h(T')}),\ldots,T_n; T')$$

- internal merge $\mathcal{I}$ defines a right $(\mathcal{H}_{ling}, \star_{\mathcal{I}})$-module given by the cosets

$$\mathcal{M}_{\mathcal{I}} := \mathrm{Dom}(\mathcal{I})\backslash\mathcal{H}_{ling}$$

- combined with iteration of domains: $\mathcal{D}_{N+1}\backslash\mathcal{D}_N$ determines a coideal in the coalgebra $\mathcal{D}_{N+1}\backslash\mathcal{H}_{ling}$
- this gives a projective system of right-module coalgebras

$$\mathcal{M}_{\mathcal{I}^N} := \mathrm{Dom}(\mathcal{I}^N)\backslash\mathcal{H}_{ling}$$

- quotient right-module coalgebras or "generalized quotients" of Hopf algebras: suitable notion of quotients in the case of noncommutative Hopf algebras

Old External Merge and "operated algebras"

- context: Rota's "operated algebras" program (algebras together with linear operators satisfying polynomial constraints (eg Rota-Baxter ops, Leibniz rule, etc) ... version for binary operations
- $\wedge_\Omega$-algebra: algebra $(\mathcal{A}, \star)$ with binary operations $\wedge_\alpha$, $\alpha \in \Omega$

$$a \star b = a_1 \wedge_\alpha (a_2 \star b) + (a \star b_1) \wedge_\alpha b_2$$

where $a = a_1 \wedge_\alpha a_2$ and $b = b_1 \wedge_\alpha b_2$

- if also Hopf algebra: cocycle condition

$$\Delta(a \wedge_\alpha b) = (a \wedge_\alpha b) \otimes 1 + (\star \otimes \wedge_\alpha) \circ \tau(\Delta(a) \otimes \Delta(b))$$

- External Merge in Stabler's Minimalism is a cocycle $\wedge_\Omega$-algebra structure on the LR Hopf algebra
- Conclusion: the implementation of Merge at the level of planar trees introduces significant complications in algebraic structure compared to free symmetric Merge followed by Externalization (confirmed also by analysis of computational implementations, Indurkya, Berwick...)