- Σ (ranked alphabet)
- Hypergraph $H = (V, E, \bar{v})$ over Σ — acyclic!
  - V (nodes)
  - $E \subseteq \{ (v, \sigma, v_1, \dots, v_k) \mid \sigma \in \Sigma^{(k)}, v, v_1, \dots, v_k \in V \}$ (edges)
  - $\bar{v} \in V$ (root)
- $p : \Sigma \to \mathbb{R}$ (weight assignment)
- inside weight $\beta : V \to \mathbb{R}$

$$\beta(v) = \sum_{\substack{(v_0, \sigma, v_1, \dots, v_k) \in E: \\ v_0 = v}} p(\sigma) \cdot \prod_{i \in [k]} \beta(v_i)$$

- outside weight

$$\alpha(v) = \begin{cases} 1 & \text{if } v = \bar{v} \\ \displaystyle\sum_{\substack{(v_0, \sigma, v_1, \dots, v_k) \in E \\ i \in [k] :\, v_i = v}} \alpha(v_0) \cdot p(\sigma) \cdot \prod_{\substack{j \in [k]: \\ j \neq i}} \beta(v_j) & \text{otherwise} \end{cases}$$

Implement Inside-Outside Algorithm by computation graph (CG)

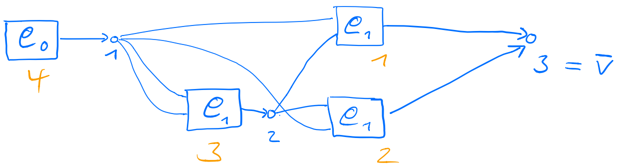
- every node in CG corresponds to a value
- 4 kinds of nodes: $p(\sigma)$, p^e, β_i^e, $\beta(v)$
  - $p(\sigma)$ is a "trainable" variable $\in \mathbb{R}$
  - $\forall\, e = (v, \sigma, v_1 \dots v_k) \in E$:
    - $\forall i \in [k]$: let $\beta_i^e \leftarrow \beta(v_i)$
    - $p^e \leftarrow p(\sigma)$
  - $\forall\, v \in V$:

$$\beta(v) \leftarrow \sum_{\substack{e = (v_0, \sigma, v_1, \dots v_k) \in E \\ v_0 = v}} p^e \cdot \prod_{i \in [k]} \beta_i^e$$

$e_0 : S \to a$

$e_1 : S \to SS$

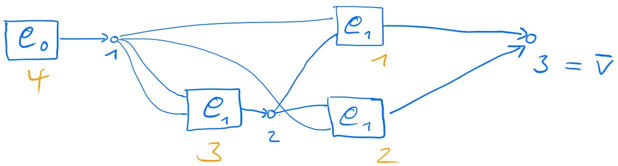
reduct for aaa:



$e_0 : S \to a$

$e_1 : S \to SS$

reduct for aaa:



Computation graph:

$p(e_0)$ $\circ \longrightarrow \circ \longrightarrow \circ$
p^4 $\beta(1)$

$\beta(2)$

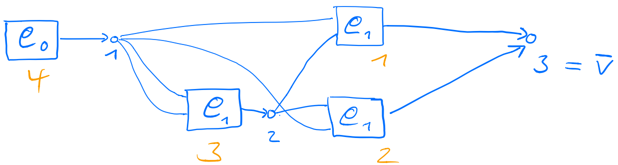
$\beta(3)$

$p(e_1)$

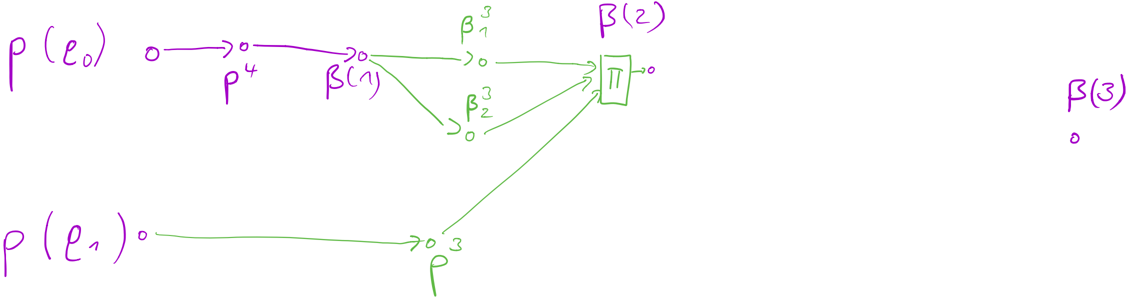
$e_0 : S \to a$

$e_1 : S \to SS$

reduct for aaa:



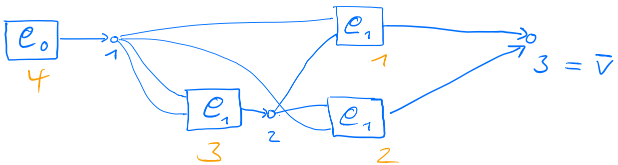
Computation graph:



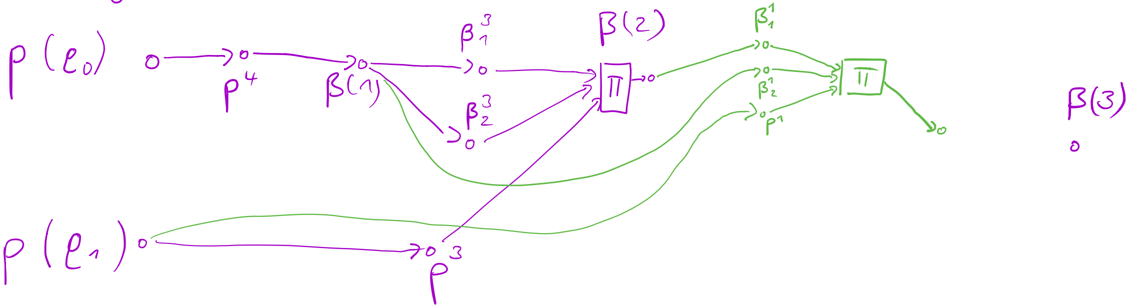
$e_0 : S \to a$

$e_1 : S \to SS$

reduct for aaa:



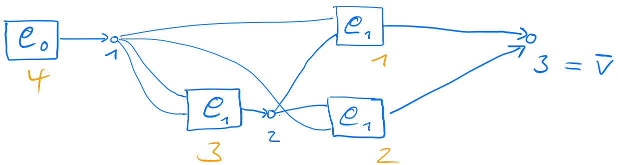
Computation graph:



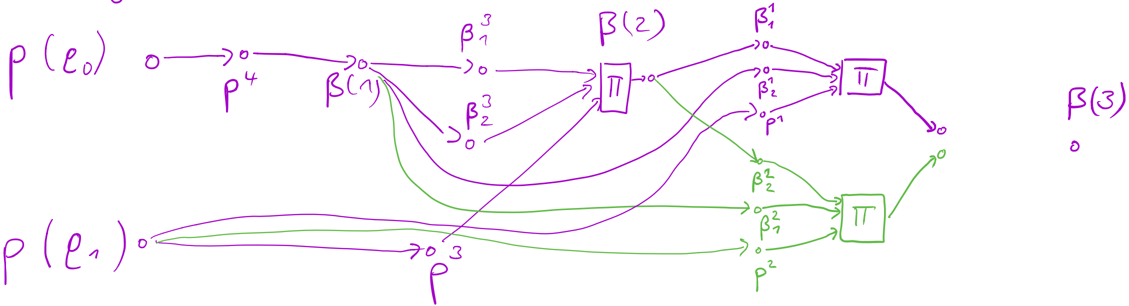
$e_0 : S \to a$

$e_1 : S \to SS$

reduct for aaa:



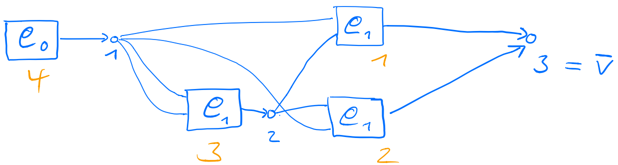
Computation graph:



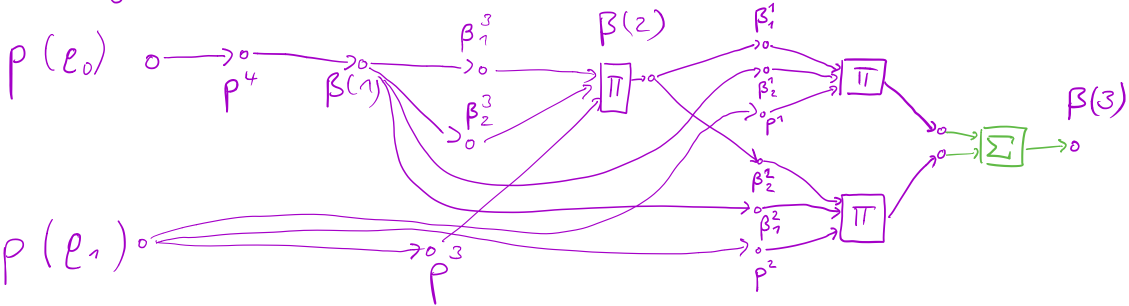
$e_0 : S \to a$

$e_1 : S \to SS$

reduct for aaa:



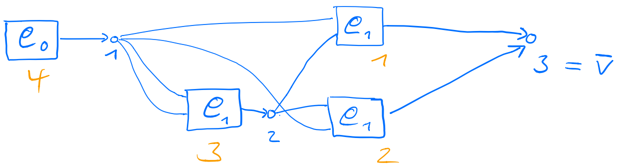
Computation graph:



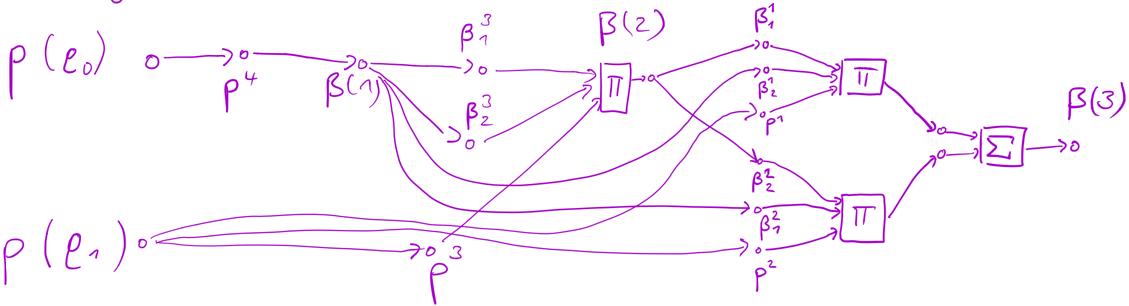
$e_0 : S \to a$

$e_1 : S \to SS$

reduct for aaa:



Computation graph:



- $\forall$ node n of CG: define $\quad \partial n := \frac{\partial \beta(\bar{v})}{\partial n}$ (gradient of $\beta(\bar{v})$ w.r.t. n)

$\Rightarrow$ We compute ∂n by means of backpropagation

- $\forall$ node n of CG: define $\quad \partial n := \frac{\partial \beta(\bar{v})}{\partial n}$ (gradient of $\beta(\bar{v})$ w.r.t. n)

$\Rightarrow$ We compute ∂n by means of backpropagation

- $\partial \beta(\bar{v}) = 1$

- $\forall$ node n of CG: define $\quad \partial n := \frac{\partial \beta(\bar{v})}{\partial n}$ (gradient of $\beta(\bar{v})$ w.r.t. n)

$\Rightarrow$ We compute ∂n by means of backpropagation

- $\partial \beta(\bar{v}) = 1$

- Let $v \in V$ and $e = (v', \sigma, v_1 \ldots v_k) \in E \wedge i \in [k]$

  $\partial \beta_i^e =$

- $\forall$ node n of CG: define $\quad \partial n := \frac{\partial \beta(\bar{v})}{\partial n}$ (gradient of $\beta(\bar{v})$ w.r.t. n)

$\Rightarrow$ We compute ∂n by means of backpropagation

- $\partial \beta(\bar{v}) = 1$

- Let $v \in V$ and $e = (v', \sigma, v_1 \ldots v_k) \in E \wedge i \in [k]$

$$\partial \beta_i^e = \frac{\partial \beta(\bar{v})}{\partial \beta_i^e(v)} = \frac{\partial \beta(\bar{v})}{\partial \beta(v')} \cdot \frac{\partial \beta(v')}{\partial \beta_i^e}$$

- $\forall$ node n of CG: define $\quad \bar{\partial} n := \frac{\partial \beta(\bar{v})}{\partial n}$ (gradient of $\beta(\bar{v})$ w.r.t. n)

$\Rightarrow$ We compute $\bar{\partial} n$ by means of backpropagation

- $\bar{\partial} \beta(\bar{v}) = 1$

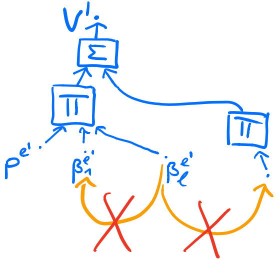
- Let $v \in V$ and $e = (v', \sigma, v_1 \dots v_k) \in E \wedge i \in [k]$

$$\bar{\partial} \beta_i^e = \frac{\partial \beta(\bar{v})}{\partial \beta_i^e(v)} = \frac{\partial \beta(\bar{v})}{\partial \beta(v')} \cdot \frac{\partial \beta(v')}{\partial \beta_i^e}$$

$$= \bar{\partial} \beta(v') \cdot \frac{1}{\partial \beta_i^e} \cdot \partial \sum_{\substack{e' = (v_0', \sigma', v_1' \dots v_\ell') \in E: \\ v_0' = v'}} \rho^{e'} \cdot \prod_{j=1}^{\ell} \beta_j^{e'}$$

• $\forall$ node n of CG: define $\partial n := \frac{\partial \beta(\bar{v})}{\partial n}$ (gradient of $\beta(\bar{v})$ w.r.t. n)

$\Rightarrow$ We compute ∂n by means of backpropagation



• $\partial \beta(\bar{v}) = 1$

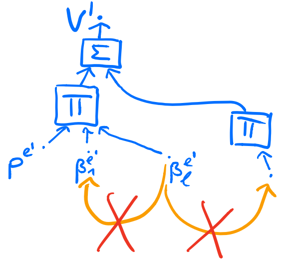
• Let $v \in V$ and $e = (v', \sigma, v_1 \dots v_k) \in E \wedge i \in [k]$

$$\partial \beta_i^e = \frac{\partial \beta(\bar{v})}{\partial \beta_i^e(v)} = \frac{\partial \beta(\bar{v})}{\partial \beta(v')} \cdot \frac{\partial \beta(v')}{\partial \beta_i^e}$$

$$= \partial \beta(v') \cdot \frac{1}{\partial \beta_i^e} \cdot \partial \sum_{\substack{e' = (v_0', \sigma', v_1' \dots v_\ell') \in E: \\ v_0' = v'}} \rho^{e'} \cdot \prod_{j=1}^{\ell} \beta_j^{e'}$$

- $\forall$ node n of CG: define $\quad \partial n := \frac{\partial \beta(\bar{v})}{\partial n}$ (gradient of $\beta(\bar{v})$ w.r.t. n)

$\Rightarrow$ We compute ∂n by means of backpropagation



- $\partial \beta(\bar{v}) = 1$

- Let $v \in V$ and $e = (v', \sigma, v_1 \ldots v_k) \in E \wedge i \in [k]$

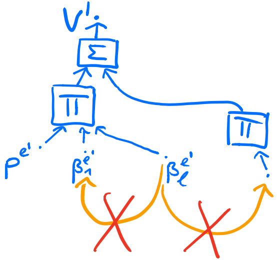
$$\partial \beta_i^e = \frac{\partial \beta(\bar{v})}{\partial \beta_i^e(v)} = \frac{\partial \beta(\bar{v})}{\partial \beta(v')} \cdot \frac{\partial \beta(v')}{\partial \beta_i^e}$$

$$= \partial \beta(v') \cdot \frac{1}{\partial \beta_i^e} \cdot \partial \sum_{\substack{e'=(v_0', \sigma', v_1', \ldots, v_\ell') \in E: \\ v_0' = v'}} p^{e'} \cdot \prod_{j=1}^{\ell} \beta_j^{e'}$$

$$\frac{\partial \beta_i^{e'}}{\partial \beta_j^e} = \begin{cases} 1, & \text{if } e' = e \wedge i = j \\ 0, & \text{otherwise} \end{cases}$$

• $\forall$ node n of CG: define $\quad \partial n := \frac{\partial \beta(\bar v)}{\partial n}$ (gradient of $\beta(\bar v)$ w.r.t. n)

$\Rightarrow$ We compute ∂n by means of backpropagation



• $\partial \beta(\bar v) = 1$

• Let $v \in V$ and $e = (v', 6, v_1 \ldots v_k) \in E \wedge i \in [k]$

$$\partial \beta_i^e = \frac{\partial \beta(\bar v)}{\partial \beta_i^e(v)} = \frac{\partial \beta(\bar v)}{\partial \beta(v')} \cdot \frac{\partial \beta(v')}{\partial \beta_i^e}$$

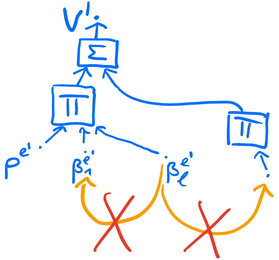
$$= \partial \beta(v') \cdot \frac{1}{\partial \beta_i^e} \cdot \partial \sum_{\substack{e'=(v_0', 6', v_1' \ldots v_\ell') \in E \\ v_0' = v'}} p^{e'} \cdot \prod_{j=1}^{\ell} \beta_j^{e'}$$

$$= \partial \beta(v') \cdot \partial p^e \cdot \prod_{j=1}^{k} \beta_j^e \cdot \frac{1}{\partial \beta_i^e}$$

$$\frac{\partial \beta_i^{e'}}{\partial \beta_j^e} = \begin{cases} 1, & \text{if } e' = e \wedge i = j \\ 0, & \text{otherwise} \end{cases}$$

• $\forall$ node n of CG: define $\partial n := \frac{\partial \beta(\bar{v})}{\partial n}$ (gradient of $\beta(\bar{v})$ w.r.t. n)

$\Rightarrow$ We compute ∂n by means of backpropagation



• $\partial \beta(\bar{v}) = 1$

• Let $v \in V$ and $e = (v', \sigma, v_1 \dots v_k) \in E \wedge i \in [k]$

$$\partial \beta_i^e = \frac{\partial \beta(\bar{v})}{\partial \beta_i^e(v)} = \frac{\partial \beta(\bar{v})}{\partial \beta(v')} \cdot \frac{\partial \beta(v')}{\partial \beta_i^e}$$

$$= \partial \beta(v') \cdot \frac{1}{\partial \beta_i^e} \cdot \partial \sum_{\substack{e' = (v_0', \sigma', v_1' \dots v_\ell') \in E \\ v_0' = v'}} p^{e'} \cdot \prod_{j=1}^{\ell} \beta_j^{e'}$$

$$= \partial \beta(v') \cdot \partial p^e \cdot \prod_{j=1}^{k} \beta_j^e \cdot \frac{1}{\partial \beta_i^e}$$

$$= \partial \beta(v') \cdot p^e \cdot \prod_{\substack{j=1 \\ j \neq i}}^{k} \beta_j^e$$

$$\frac{\partial \beta_i^{e'}}{\partial \beta_j^e} = \begin{cases} 1, & \text{if } e' = e \wedge i = j \\ 0, & \text{otherwise} \end{cases}$$

$$\partial \beta(v) = \frac{\partial \beta(\bar{v})}{\partial \beta(v)} =$$

$$\bar{\partial}\beta(v) = \frac{\partial \beta(\bar{v})}{\partial \beta(v)} = \sum_{\substack{e=(v', \sigma, v_1 \dots v_k) \in E \\ i \in [k]: \\ v_i = v}} \frac{\partial \beta(\bar{v})}{\partial \beta^e_i} \cdot \frac{\partial \beta^e_i}{\partial \beta(v)}$$

$$\partial\beta(v) = \frac{\partial \beta(\tilde{v})}{\partial \beta(v)} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\in E\\ i\in[k]:\\ v_i = v}} \underbrace{\frac{\partial \beta(\tilde{v})}{\partial \beta^e_i}}_{=\partial\beta^e_i} \cdot \underbrace{\frac{\partial \beta^e_i}{\partial \beta(v)}}_{=1}$$

$$\partial\beta(v) = \frac{\partial\beta(\tilde{v})}{\partial\beta(v)} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\in E\\ i\in[k]:\\ v_i=v}} \underbrace{\frac{\partial\beta(\tilde{v})}{\partial\beta_i^e}}_{=\partial\beta_i^e} \cdot \underbrace{\frac{\partial\beta_i^e}{\partial\beta(v)}}_{=1} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\\ i\in[k]:\\ v_i=v}} \partial\beta_i^e$$

$$\bar{\partial}\beta(v) = \frac{\partial \beta(\bar{v})}{\partial \beta(v)} = \sum_{\substack{e=(v',G,v_1\dots v_k)\in E \\ i\in[n]: \\ v_i = v}} \underbrace{\frac{\partial \beta(\bar{v})}{\partial \beta_i^e}}_{=\bar{\partial}\beta_i^e} \cdot \underbrace{\frac{\partial \beta_i^e}{\partial \beta(v)}}_{=1} = \sum_{\substack{e=(v',G,v_1\dots v_k) \\ i\in[k]: \\ v_i = v}} \bar{\partial}\beta_i^e = \sum_{\substack{e=(v',G,v_1\dots v_k) \\ i\in[k]: \\ v_i = v}} \bar{\partial}\beta(v') \cdot p^e \cdot \prod_{\substack{j=1 \\ j\neq i}}^{k} \beta(v_j)$$

$$\bar{\partial}\beta(v) = \frac{\partial \beta(\bar{v})}{\partial \beta(v)} = \sum_{\substack{e=(v',\sigma,v_1\ldots v_k)\in E \\ i\in[n]: \\ v_i = v}} \underbrace{\frac{\partial \beta(\bar{v})}{\partial \beta_i^e}}_{=\bar{\partial}\beta_i^e} \cdot \underbrace{\frac{\partial \beta_i^e}{\partial \beta(v)}}_{=1} = \sum_{\substack{e=(v',\sigma,v_1\ldots v_k) \\ i\in[k]: \\ v_i = v}} \bar{\partial}\beta_i^e = \sum_{\substack{e=(v',\sigma,v_1\ldots v_k) \\ i\in[k] : \\ v_i = v}} \bar{\partial}\beta(v')\cdot p^e \cdot \prod_{\substack{j=1 \\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,\sigma,v_1\ldots v_n)\in E$.

$$\bar{\partial} p^e = \frac{\partial \beta(\bar{v})}{\partial p^e}$$

$$\bar{\partial}\beta(v) = \frac{\partial \beta(\bar{v})}{\partial \beta(v)} = \sum_{\substack{e=(v',G,v_1\dots v_k)\in E \\ i\in[k]: \\ v_i = v}} \underbrace{\frac{\partial \beta(\bar{v})}{\partial \beta^e_i}}_{=\bar{\partial}\beta^e_i} \cdot \underbrace{\frac{\partial \beta^e_i}{\partial \beta(v)}}_{=1} = \sum_{\substack{e=(v',G,v_1\dots v_k) \\ i\in[k]: \\ v_i=v}} \bar{\partial}\beta^e_i = \sum_{\substack{e=(v',G,v_1\dots v_k) \\ i\in[k] : \\ v_i = v}} \bar{\partial}\beta(v')\cdot p^e \cdot \prod_{\substack{j=1 \\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,G,v_1\dots v_n)\in E$.

$$\bar{\partial} p^e = \frac{\partial \beta(\bar{v})}{\partial p^e} = \frac{\partial \beta(\bar{v})}{\partial \beta(v)} \cdot \frac{\partial \beta(v)}{\partial p^e}$$

$$\partial\beta(v) = \frac{\partial\beta(\bar{v})}{\partial\beta(v)} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\in E\\ i\in[n]:\\ v_i=v}} \underbrace{\frac{\partial\beta(\bar{v})}{\partial\beta_i^e}}_{=\partial\beta_i^e}\cdot\underbrace{\frac{\partial\beta_i^e}{\partial\beta(v)}}_{=1} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\\ i\in[k]:\\ v_i=v}} \partial\beta_i^e = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\\ i\in[k]:\\ v_i=v}} \partial\beta(v')\cdot p^e \cdot \prod_{\substack{j=1\\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,\sigma,v_1\dots v_n)\in E$.

$$\partial p^e = \frac{\partial\beta(\bar{v})}{\partial p^e} = \frac{\partial\beta(\bar{v})}{\partial\beta(v)}\cdot\frac{\partial\beta(v)}{\partial p^e} = \partial\beta(v)\cdot \partial \sum_{\substack{e'=(v_0',\sigma,v_1'\dots v_k')\in E\\ v_0'=v}} p^{e'}\cdot\prod_{i\in[k]}\beta_i^{e'}\cdot\frac{1}{\partial p^e}$$

$$\eth\beta(v) = \frac{\partial\beta(\bar v)}{\partial\beta(v)} = \sum_{\substack{e=(v',6,v_1\dots v_k)\in E\\ i\in[n]:\\ v_i=v}} \underbrace{\frac{\partial\beta(\bar v)}{\partial\beta_i^e}}_{=\eth\beta_1^e} \cdot \underbrace{\frac{\partial\beta_i^e}{\partial\beta(v)}}_{=1} = \sum_{\substack{e=(v',6,v_1\dots v_k)\\ i\in[k]:\\ v_i=v}} \eth\beta_i^e = \sum_{\substack{e=(v',6,v_1\dots v_k)\\ i\in[k]:\\ v_1=v}} \eth\beta(v')\cdot p^e \cdot \prod_{\substack{j=1\\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,6,v_1\dots v_n)\in E$.

$$\eth p^e = \frac{\partial\beta(\bar v)}{\partial p^e} = \frac{\partial\beta(\bar v)}{\partial\beta(v)}\cdot\frac{\partial\beta(v)}{\partial p^e} = \eth\beta(v)\cdot \partial \sum_{\substack{e'=(v_0',6,v_1'\dots v_k')\in E\\ v_0'=v}} p^{e'}\cdot\prod_{i\in[k]}\beta_i^{e'} \cdot \frac{1}{\partial p^e}$$

$$= \eth\beta(v)\cdot \partial\Big(p^e\cdot\prod_{i\in[n]}\beta_i^e\Big)\cdot \frac{1}{\partial p^e}$$

$$\partial\beta(v) = \frac{\partial\beta(\bar{v})}{\partial\beta(v)} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\in E\\ i\in[n]:\\ v_i=v}} \underbrace{\frac{\partial\beta(\bar{v})}{\partial\beta_i^e}}_{=\partial\beta_i^e} \cdot \underbrace{\frac{\partial\beta_i^e}{\partial\beta(v)}}_{=1} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\\ i\in[k]:\\ v_i=v}} \partial\beta_i^e = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\\ i\in[k]\ :\\ v_i=v}} \partial\beta(v')\cdot p^e \cdot \prod_{\substack{j=1\\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,\sigma,v_1\dots v_n)\in E$.

$$\partial p^e = \frac{\partial\beta(\bar{v})}{\partial p^e} = \frac{\partial\beta(\bar{v})}{\partial\beta(v)}\cdot\frac{\partial\beta(v)}{\partial p^e} = \partial\beta(v)\cdot \partial \sum_{\substack{e'=(v_0',\sigma,v_1',\dots v_k')\in E\\ v_0'=v}} p^{e'}\cdot\prod_{u\in[k]}\beta_u^{e'} \cdot \frac{1}{\partial p^e}$$

$$= \partial\beta(v)\cdot \partial\Big(p^e\cdot\prod_{u\in[n]}\beta_u^e\Big)\cdot\frac{1}{\partial p^e}$$

$$= \partial\beta(v)\cdot\prod_{u\in[n]}\beta_u^e$$

$$\partial\beta(v) = \frac{\partial\beta(\bar v)}{\partial\beta(v)} = \sum_{\substack{e=(v',6,v_1\dots v_k)\in E\\ i\in[n]:\\ v_i=v}} \underbrace{\frac{\partial\beta(\bar v)}{\partial\beta_i^e}}_{=\partial\beta_i^e} \cdot \underbrace{\frac{\partial\beta_i^e}{\partial\beta(v)}}_{=1} = \sum_{\substack{e=(v',6,v_1\dots v_k)\\ i\in[k]:\\ v_i=v}} \partial\beta_i^e = \sum_{\substack{e=(v',6,v_1\dots v_k)\\ i\in[k]:\\ v_i=v}} \partial\beta(v')\cdot p^e \cdot \prod_{\substack{j=1\\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,6,v_1\dots v_n)\in E$.

$$\partial p^e = \frac{\partial\beta(\bar v)}{\partial p^e} = \frac{\partial\beta(\bar v)}{\partial\beta(v)}\cdot\frac{\partial\beta(v)}{\partial p^e} = \partial\beta(v)\cdot\partial \sum_{\substack{e'=(v_0',6,v_1'\dots v_k')\in E\\ v_0'=v}} p^{e'}\cdot\prod_{i\in[k]}\beta_i^{e'}\cdot\frac{1}{\partial p^e}$$

$$= \partial\beta(v)\cdot\partial\Big(p^e\cdot\prod_{i\in[n]}\beta_i^e\Big)\cdot\frac{1}{\partial p^e}$$

$$= \partial\beta(v)\cdot\prod_{i\in[n]}\beta_i^e$$

$\partial p(6)$

$$\partial\beta(v) = \frac{\partial\beta(\bar v)}{\partial\beta(v)} = \sum_{\substack{e=(v',6,v_1\ldots v_k)\in E\\ i\in[n]:\\ v_i=v}} \underbrace{\frac{\partial\beta(\bar v)}{\partial\beta_i^e}}_{=\partial\beta_i^e}\cdot\underbrace{\frac{\partial\beta_i^e}{\partial\beta(v)}}_{=1} = \sum_{\substack{e=(v',6,v_1\ldots v_k)\\ i\in[k]:\\ v_i=v}} \partial\beta_i^e = \sum_{\substack{e=(v',6,v_1\ldots v_k)\\ i\in[k]\\ v_i=v}} \partial\beta(v')\cdot p^e \cdot \prod_{\substack{j=1\\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,6,v_1\ldots v_n)\in E$.

$$\partial p^e = \frac{\partial\beta(\bar v)}{\partial p^e} = \frac{\partial\beta(\bar v)}{\partial\beta(v)}\cdot\frac{\partial\beta(v)}{\partial p^e} = \partial\beta(v)\cdot\partial\sum_{\substack{e'=(v_0',6,v_1'\ldots v_k')\in E\\ v_0'=v}} p^{e'}\cdot\prod_{i\in[k]}\beta_i^{e'}\cdot\frac{1}{\partial p^e}$$

$$= \partial\beta(v)\cdot\partial\Big(p^e\cdot\prod_{i\in[n]}\beta_i^e\Big)\cdot\frac{1}{\partial p^e}$$

$$= \partial\beta(v)\cdot\prod_{i\in[n]}\beta_i^e$$

$$\partial p(6) = \sum_{\substack{e=(v,6',v_1\ldots v_n)\in E\\ 6=6'}} \frac{\partial\beta(\bar v)}{\partial p^e}\cdot\frac{\partial p^e}{\partial p(6)}$$

$$\partial\beta(v) = \frac{\partial\beta(\bar v)}{\partial\beta(v)} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\in E \\ i\in[n]: \\ v_i=v}} \underbrace{\frac{\partial\beta(\bar v)}{\partial\beta_i^e}}_{=\partial\beta_i^e} \cdot \underbrace{\frac{\partial\beta_i^e}{\partial\beta(v)}}_{=1} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k) \\ i\in[k]: \\ v_i=v}} \partial\beta_i^e = \sum_{\substack{e=(v',\sigma,v_1\dots v_k) \\ i\in[k]: \\ v_i=v}} \partial\beta(v')\cdot p^e \cdot \prod_{\substack{j=1 \\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,\sigma,v_1\dots v_n)\in E$.

$$\partial p^e = \frac{\partial\beta(\bar v)}{\partial p^e} = \frac{\partial\beta(\bar v)}{\partial\beta(v)}\cdot\frac{\partial\beta(v)}{\partial p^e} = \partial\beta(v)\cdot \partial \sum_{\substack{e'=(v_0',\sigma,v_1',\dots v_k')\in E \\ v_0'=v}} p^{e'}\cdot \prod_{u\in[k]} \beta_u^{e'} \cdot \frac{1}{\partial p^e}$$

$$= \partial\beta(v) \cdot \partial\Big(p^e \cdot \prod_{u\in[n]} \beta_u^e\Big) \cdot \frac{1}{\partial p^e}$$

$$= \partial\beta(v) \cdot \prod_{u\in[n]} \beta_u^e$$

$$\partial p(\sigma) = \sum_{\substack{e=(v,\sigma',v_1\dots v_n)\in E \\ \sigma=\sigma'}} \frac{\partial\beta(\bar v)}{\partial p^e}\cdot \underbrace{\frac{\partial p^e}{\partial p(\sigma)}}_{=1}$$

$$\bar{\partial}\beta(v) = \frac{\partial\beta(\bar{v})}{\partial\beta(v)} = \sum_{\substack{e=(v',\sigma,v_1\ldots v_k)\in E\\ i\in[n]:\\ v_i=v}} \underbrace{\frac{\partial\beta(\bar{v})}{\partial\beta_i^e}}_{=\bar{\partial}\beta_i^e} \cdot \underbrace{\frac{\partial\beta_i^e}{\partial\beta(v)}}_{=1} = \sum_{\substack{e=(v',\sigma,v_1\ldots v_k)\\ i\in[k]:\\ v_i=v}} \bar{\partial}\beta_i^e = \sum_{\substack{e=(v',\sigma,v_1\ldots v_k)\\ i\in[k]:\\ v_i=v}} \bar{\partial}\beta(v')\cdot p^e \cdot \prod_{\substack{j=1\\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,\sigma,v_1\ldots v_n)\in E$.

$$\bar{\partial}p^e = \frac{\partial\beta(\bar{v})}{\partial p^e} = \frac{\partial\beta(\bar{v})}{\partial\beta(v)}\cdot\frac{\partial\beta(v)}{\partial p^e} = \bar{\partial}\beta(v)\cdot\partial\sum_{\substack{e'=(v_0',\sigma,v_1'\ldots v_k')\in E\\ v_0'=v}} p^{e'}\cdot\prod_{i\in[k]}\beta_i^{e'}\cdot\frac{1}{\partial p^e}$$

$$= \bar{\partial}\beta(v)\cdot\partial\Big(p^e\cdot\prod_{i\in[n]}\beta_i^e\Big)\cdot\frac{1}{\partial p^e}$$

$$= \bar{\partial}\beta(v)\cdot\prod_{i\in[n]}\beta_i^e$$

$$\bar{\partial}p(\sigma) = \sum_{\substack{e=(v,\sigma',v_1\ldots v_n)\in E\\ \sigma=\sigma'}} \frac{\partial\beta(\bar{v})}{\partial p^e}\cdot\underbrace{\frac{\partial p^e}{\partial p(\sigma)}}_{=1} = \sum_{\substack{e=(v,\sigma',v_1\ldots v_n)\in E\\ \sigma=\sigma'}} \bar{\partial}p^e$$

$$\bar\partial \beta(v) = \frac{\partial \beta(\bar v)}{\partial \beta(v)} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k)\in E \\ i\in[k]: \\ v_1 = v}} \underbrace{\frac{\partial \beta(\bar v)}{\partial \beta_i^e}}_{=\bar\partial\beta_1^e} \cdot \underbrace{\frac{\partial \beta_i^e}{\partial \beta(v)}}_{=1} = \sum_{\substack{e=(v',\sigma,v_1\dots v_k) \\ i\in[k]: \\ v_i = v}} \bar\partial \beta_i^e = \sum_{\substack{e=(v',\sigma,v_1\dots v_k) \\ i\in[k]: \\ v_i = v}} \bar\partial\beta(v')\cdot p^e \cdot \prod_{\substack{j=1 \\ j\neq i}}^{k} \beta(v_j)$$

Let $e=(v,\sigma,v_1\dots v_n)\in E$.

$$\bar\partial p^e = \frac{\partial \beta(\bar v)}{\partial p^e} = \frac{\partial \beta(\bar v)}{\partial \beta(v)} \cdot \frac{\partial \beta(v)}{\partial p^e} = \bar\partial\beta(v)\cdot \partial \sum_{\substack{e'=(v_0',\sigma,v_1'\dots v_k')\in E \\ v_0' = v}} p^{e'} \cdot \prod_{i\in[k]} \beta_i^{e'} \cdot \frac{1}{\partial p^e}$$

$$= \bar\partial\beta(v) \cdot \partial\Big(p^e \cdot \prod_{i\in[n]} \beta_i^e\Big) \cdot \frac{1}{\partial p^e}$$

$$= \bar\partial\beta(v) \cdot \prod_{i\in[n]} \beta_i^e$$

$$\bar\partial p(\sigma) = \sum_{\substack{e=(v,\sigma',v_1\dots v_n)\in E \\ \sigma=\sigma'}} \frac{\partial \beta(\bar v)}{\partial p^e} \cdot \underbrace{\frac{\partial p^e}{\partial p(\sigma)}}_{=1} = \sum_{\substack{e=(v,\sigma',v_1\dots v_n)\in E \\ \sigma=\sigma'}} \bar\partial p^e = \sum_{\substack{e=(v,\sigma',v_1\dots v_n)\in E \\ \sigma=\sigma'}} \bar\partial\beta(v) \cdot \prod_{i\in[n]} \beta_i^e$$

Expected frequency of a symbol

- $$f(\sigma) = \frac{1}{\beta(\bar{v})} \sum_{(v, \sigma, v_1, \ldots, v_k) \in E} \alpha(v) \cdot p(\sigma) \cdot \beta(v_1) \cdot \ldots \cdot \beta(v_k)$$

Expected frequency of a symbol

- $$f(\sigma) = \frac{1}{\beta(\bar{v})} \sum_{(v,\sigma,v_1,\dots,v_k)\in E} \alpha(v)\cdot p(\sigma)\cdot \beta(v_1)\cdot \dots \cdot \beta(v_k)$$

- $$\partial_{p(\sigma)} = \sum_{\substack{e=(v,\sigma',v_1\dots v_n)\in E \\ \sigma=\sigma'}} \partial\beta(v) \cdot \prod_{u\in[n]} \beta_u^e$$

Expected frequency of a symbol

- $$f(\sigma) = \frac{1}{\beta(\bar{v})} \sum_{(v,\sigma,v_1,\dots,v_k)\in E} \alpha(v)\cdot p(\sigma)\cdot\beta(v_1)\cdot\ \dots\ \cdot\beta(v_k)$$

- $$\partial p(\sigma) = \sum_{\substack{e=(v,\sigma',v_1\dots v_n)\in E \\ \sigma=\sigma'}} \partial\beta(v)\cdot \prod_{u\in[n]} \beta_u^e$$

- $$f(\sigma) = \partial p(\sigma)\cdot p(\sigma)\cdot\frac{1}{\beta(\bar{v})}$$

Another intuition:

Let $V = \{v_1, \dots, v_n\}$ be such that in every unfolding of H occurs at least one node $v \in V$ and at most one node $v \in V$.

Then: $\beta(\bar{v}) = \sum_{v' \in V} \alpha(v') \cdot \beta(v')$.

Let $v \in V$. Then

$$\frac{\partial \beta(\bar{v})}{\partial \beta(v)} = \partial\Big(\sum_{v' \in V} \alpha(v') \cdot \beta(v')\Big) \cdot \frac{1}{\partial \beta(v)}$$

$$= \sum_{v' \in V} \underbrace{\frac{\partial(\alpha(v') \cdot \beta(v'))}{\partial \beta(v)}}_{= 0, \text{ if } v' \neq v}$$

// otherwise, there would be an unfolding that contains v and v'

$$= \frac{\partial\, \alpha(v') \cdot \beta(v')}{\partial \beta(v)}$$

$$= \underbrace{\frac{\partial \alpha(v)}{\partial \beta(v)}}_{= 0, \text{ because } H \text{ is acyclic}} \cdot \beta(v) + \underbrace{\frac{\partial \beta(v)}{\partial \beta(v)}}_{= 1} \cdot \alpha(v)$$

$$= \alpha(v)$$