(i.e. parents come before children.) v_N denotes the variable we're trying to compute derivatives of (e.g. loss).

Let v_1, \ldots, v_N be a topological ordering of the computation graph

backward pass $\begin{bmatrix} \overline{v_N}=1\\ & \text{For } i=N-1,\dots,1\\ & \overline{v_i}=\sum_{j\in\operatorname{Ch}(v_i)}\overline{v_j}\,\frac{\partial v_j}{\partial v_i} \end{bmatrix}$

forward pass
$$egin{bmatrix} ext{For } i=1,\ldots,N \ ext{Compute } v_i ext{ as a function of } ext{Pa}(v_i) \end{bmatrix}$$

forward pass
$$v_i = v_i$$
 Compute v_i as a function of $\operatorname{Pa}(v_i)$