IX. Betas: Introduction

Consider an internal node s in the context tree $\mathcal{T}_{\mathcal{D}}$ and the corresponding *conditional* weighted probability $P_w^s(X_t = 1|x_1^{t-1})$. Assuming that $0s$ (and not $1s$) is a suffix of the context x_{1-D}^0, x_1^{t-1} of x_t, we obtain for this probability that

$$P_w^s(X_t = 1|x_1^{t-1}) = \frac{P_e^s(x_1^{t-1}, X_t = 1) + P_w^{0s}(x_1^{t-1}, X_t = 1)P_w^{1s}(x_1^{t-1})}{P_e^s(x_1^{t-1}) + P_w^{0s}(x_1^{t-1})P_w^{1s}(x_1^{t-1})}$$
$$= \frac{\beta^s(x_1^{t-1})P_e^s(X_t = 1|x_1^{t-1}) + P_w^{0s}(X_t = 1|x_1^{t-1})}{\beta^s(x_1^{t-1}) + 1} \quad (1)$$

where

$$\beta^s(x_1^{t-1}) \triangleq \frac{P_e^s(x_1^{t-1})}{P_w^{0s}(x_1^{t-1})P_w^{1s}(x_1^{t-1})}. \quad (2)$$

If we start in the context-leaf and work our way down to the root, we finally find $P_w^\lambda(X_t = 1|x_1^{t-1})$.