

TPTurbo equations

Notation change from whiteboard figure:

$$X^{\rightarrow} \rightarrow X^{-}; X^{\leftarrow} \rightarrow X^{+}; \quad T_t^{\leftrightarrow} \rightarrow T_t^{\pm}; \quad \hat{h}_t^{\leftrightarrow} \rightarrow L_t^{\pm}; \quad h_t^{\leftrightarrow} \rightarrow h_t^{\pm}; \quad h_t \rightarrow H_t$$

New notation (not fully kosher since the *names* of the arguments to the function, not just their values, matter):

$$“z = f_{\mathbf{W}}(u_{\tau}, v_{\tau}, \dots)” \equiv “z = \sigma(W_{zu}u_{\tau} + W_{zv}v_{\tau} + \dots)”;$$

$$“L = \mathbb{L}_{\mathbf{W}}(u_{\tau}, v_{\tau}, \dots)” = “L \text{ is the state variable of an LSTM unit receiving input } W_{Lu}u_{\tau} + W_{Lv}v_{\tau} + \dots”$$

$$“L^* = \mathbb{L}_{\mathbf{W}}^*(u_{\tau}, v_{\tau}, \dots)” = “L^* \text{ are the other variables of an LSTM unit receiving input } W_{L^*u}u_{\tau} + W_{L^*v}v_{\tau} + \dots” \text{ [needs work]}$$

$$W_{u^{\pm}v^{\pm}} = [W_{u^{-}v^{-}}; W_{u^{+}v^{+}}] \text{ (corresponding to } [W_{uv}^{\rightarrow}; W_{uv}^{\leftarrow}])$$

Parameters: $\mathbf{W} = [F; R; W_{a^F \pm h^{\pm}}; W_{a^F \pm x}; W_{a^R \pm h^{\pm}}; W_{a^R \pm x}; W_{L^{\pm} h^{\pm}}; W_{L^{\pm} x}; W_{L^* \pm h^{\pm}}; W_{L^* \pm x}]$ (note that e.g. $W_{a^F \pm h^{\pm}} = [W_{a^F-L^{-}}; W_{a^F-T^{-}}; W_{a^F+L^{+}}; W_{a^F+T^{+}}]$)

Equations:

$$x_t = [\text{GLOVE}(w_t); \text{Char-CNN}(w_t)]; \quad w_t = \text{'context' [source text] word at time } t$$

everything below applies equally to the word-sequence q_t of the query

$$H_t = [h_t^{+}; h_t^{-}]$$

$$h_t^{+} = [L_t^{+}; T_t^{+}]; \quad h_t^{-} = [L_t^{-}; T_t^{-}]$$

$$T_t^{\pm} = (Fa_t^{F\pm})(Ra_t^{R\pm})^{\top}$$

$$a_t^{F\pm} = f_{\mathbf{W}}(h_{t\pm 1}^{\pm}, x_t); \quad a_t^{R\pm} = f_{\mathbf{W}}(h_{t\pm 1}^{\pm}, x_t)$$

$$L_t^{\pm} = \mathbb{L}_{\mathbf{W}}(h_{t\pm 1}^{\pm}, x_t); \quad L_t^{*\pm} = \mathbb{L}_{\mathbf{W}}^*(h_{t\pm 1}^{\pm}, x_t)$$