Attention based models in End-to-End ASR

Exploration of Attention in ESPNET toolkit
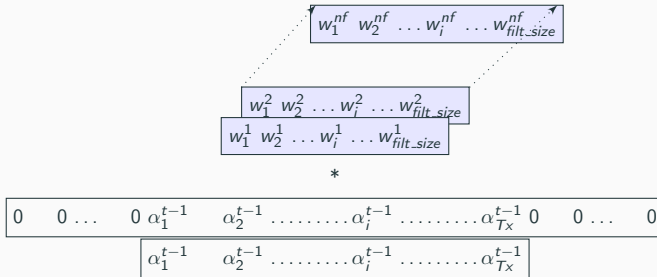
Shreekantha Nadig

November 8, 2018

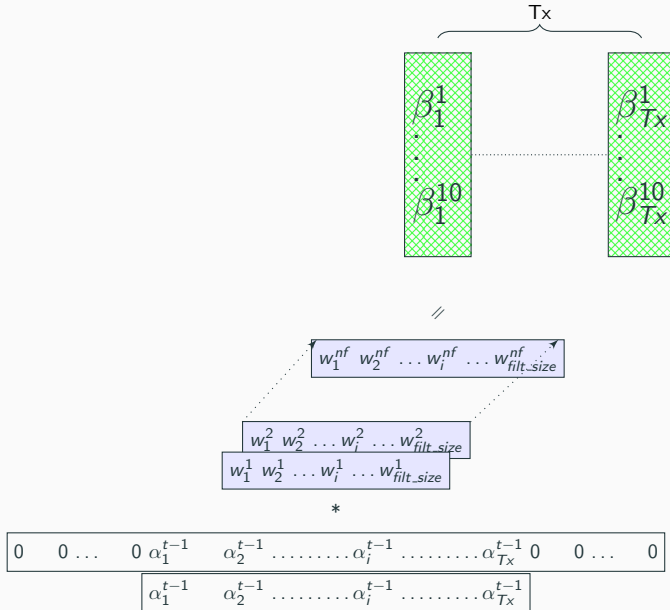International Institute of Information Technology - Bangalore

Introduction

$$\beta_1^1 \quad \beta_2^1 \ldots\ldots\ldots \beta_i^1 \ldots\ldots\ldots \beta_{Tx}^1$$

$$\|$$

$$w_1^1 \ w_2^1 \ldots w_i^1 \ldots w_{filt\_size}^1$$

$$*$$

$$0 \quad 0 \ldots \quad 0 \ \alpha_1^{t-1} \quad \alpha_2^{t-1} \ldots\ldots\ldots \alpha_i^{t-1} \ldots\ldots\ldots \alpha_{Tx}^{t-1} \ 0 \quad 0 \ldots \quad 0$$

$$\boxed{\beta_1^2 \quad \beta_2^2 \ldots\ldots\ldots \beta_i^2 \ldots\ldots\ldots \beta_{Tx}^2}$$

$$\|$$

$$\boxed{w_1^2 \ w_2^2 \ \ldots \ w_i^2 \ \ldots \ w_{filt\_size}^2}$$

$$*$$

$$\boxed{0 \quad 0 \ldots \quad 0 \ \alpha_1^{t-1} \quad \alpha_2^{t-1} \ldots\ldots\ldots \alpha_i^{t-1} \ldots\ldots\ldots \alpha_{Tx}^{t-1} \ 0 \quad 0 \ldots \quad 0}$$

Tx

$$\beta_1^1$$
$$\vdots$$
$$\beta_1^{10}$$

$$\beta_{Tx}^1$$
$$\vdots$$
$$\beta_{Tx}^{10}$$

$$//$$

$$w_1^{nf} \; w_2^{nf} \ldots w_i^{nf} \ldots w_{filt\_size}^{nf}$$

$$w_1^2 \; w_2^2 \ldots w_i^2 \ldots w_{filt\_size}^2$$

$$w_1^1 \; w_2^1 \ldots w_i^1 \ldots w_{filt\_size}^1$$

$$*$$

$$0 \quad 0 \ldots \quad 0 \; \alpha_1^{t-1} \quad \alpha_2^{t-1} \ldots \ldots \alpha_i^{t-1} \ldots \ldots \alpha_{Tx}^{t-1} \, 0 \quad 0 \ldots \quad 0$$

$$\alpha_1^{t-1} \quad \alpha_2^{t-1} \ldots \ldots \alpha_i^{t-1} \ldots \ldots \alpha_{Tx}^{t-1}$$

2/3