

Attention-Based CNN in NLP

汇报人：张璐 汇报时间：2018.12.13





BCNN : Basic Bi-CNN

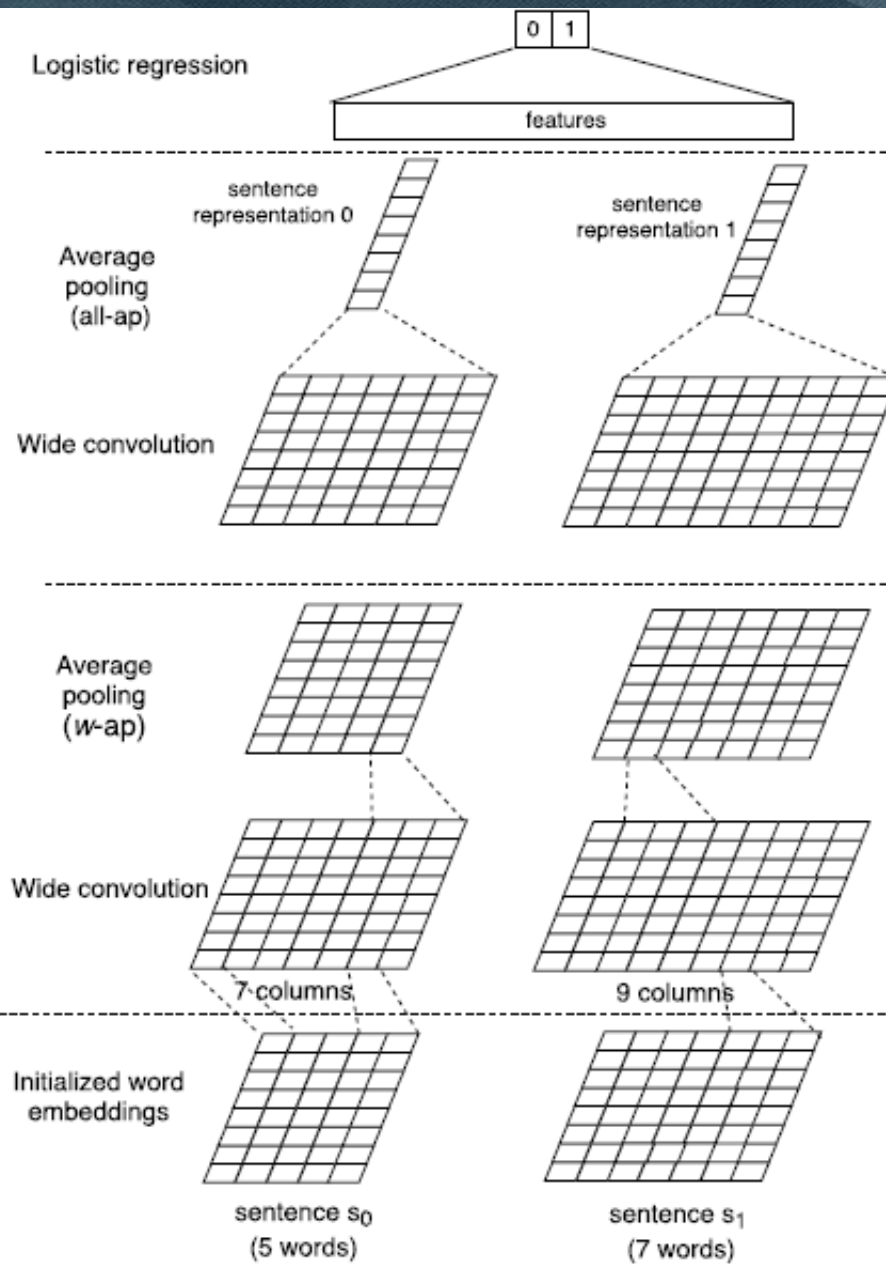
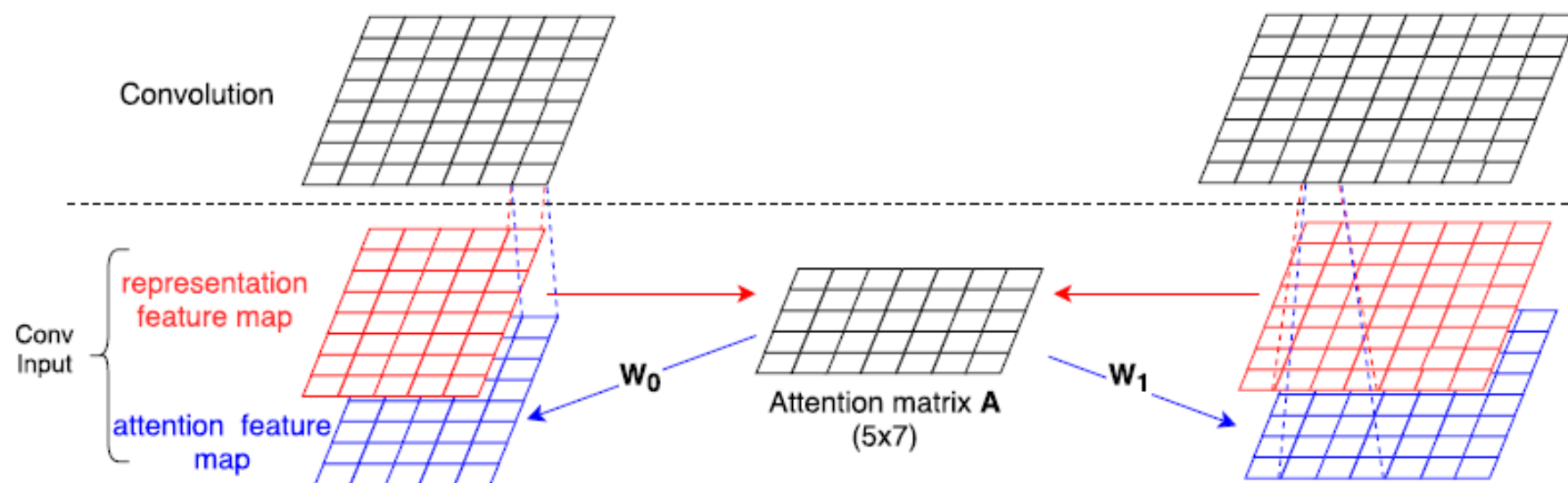


Figure 2: BCNN: ABCNN without Attention



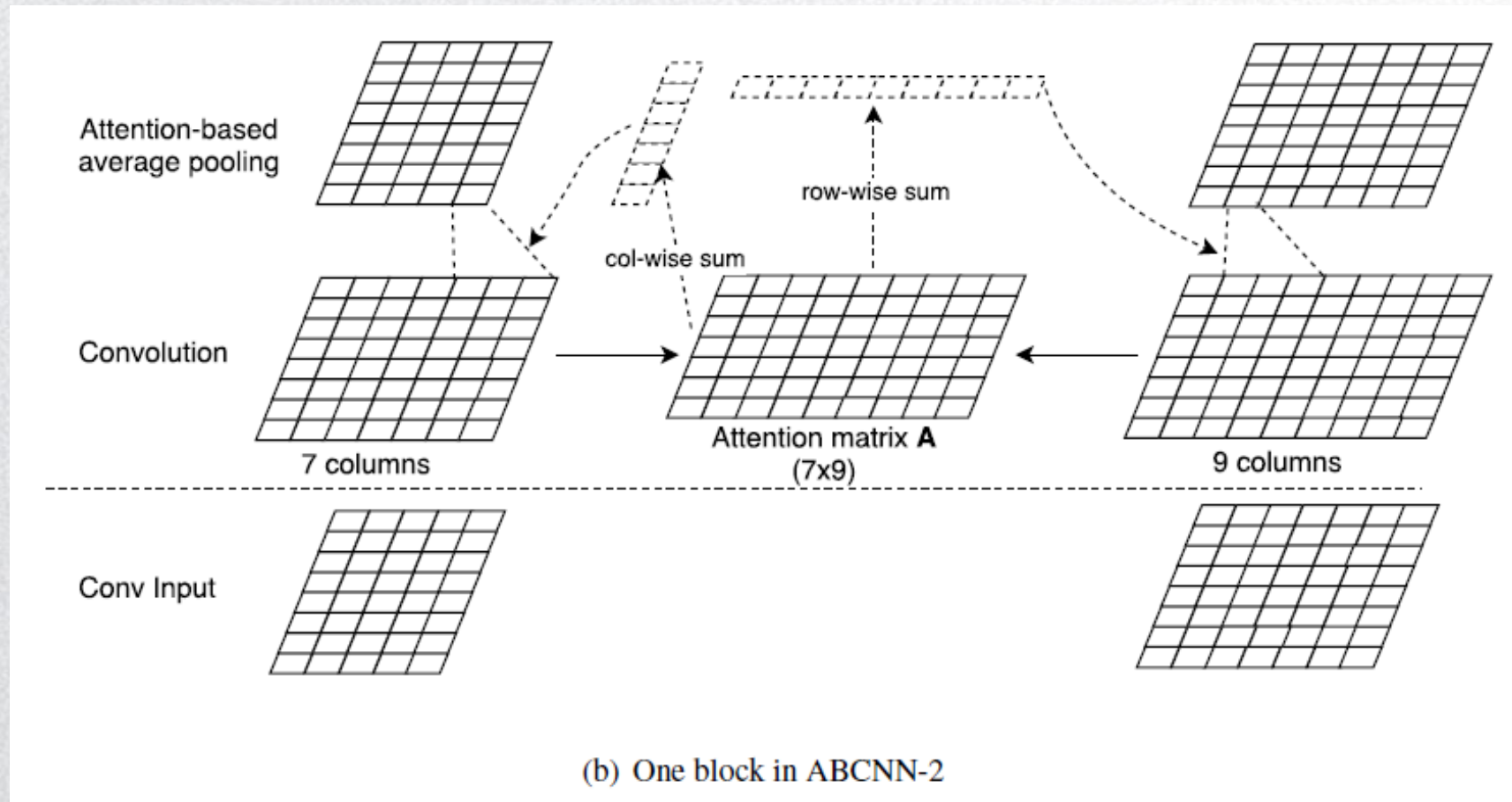
(a) One block in ABCNN-1

$$\mathbf{F}_{i,r} \in \mathbf{R}^{d \times s}$$

$$\mathbf{A}_{i,j} = \text{match-score}(\mathbf{F}_{0,r}[:, i], \mathbf{F}_{1,r}[:, j])$$

$$1/(1 + |x - y|)$$

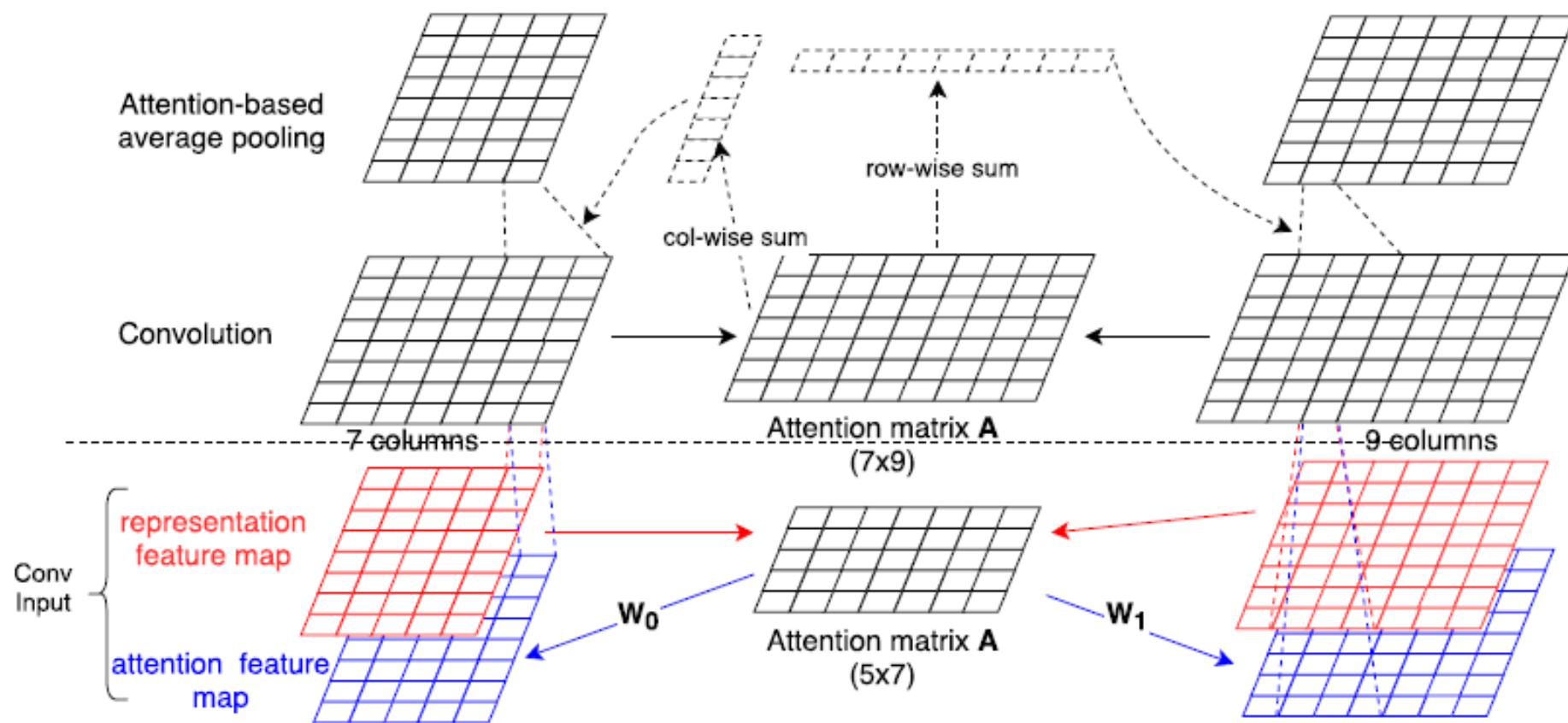
$$\mathbf{F}_{0,a} = \mathbf{W}_0 \cdot \mathbf{A}^\top, \quad \mathbf{F}_{1,a} = \mathbf{W}_1 \cdot \mathbf{A}$$



$$a_{0,j} = \sum \mathbf{A}[j, :]$$

$$a_{1,j} = \sum \mathbf{A}[:, j]$$

$$\mathbf{F}_{i,r}^p[:, j] = \sum_{|k=j:j+w} a_{i,k} \mathbf{F}_{i,r}^c[:, k], \quad j = 1 \dots s_i$$



(c) One block in ABCNN-3



参考资料:

ABCNN: Attention-Based Convolutional Neural
Network for Modeling Sentence Pairs

<https://arxiv.org/abs/1512.05193>

Attention-Based CNN in NLP

汇报人：张璐 汇报时间：2018.12.13

