Algorithm 2: CTC Loss and **softmax** gradient computation

Data: $\text{out}_{|\hat{A}| \times (\bar{W}_{padded}/4)}$ (result of softmax), where $|\hat{A}|$ (alphabet size);
$T = \bar{W}_{unpadded}/4$;
$label$ (encoded by alphabet);
$bl = 0$ (blank index)
begin

$\quad L = 2 \times len(label) + 1$
$\quad out_{unpadded} = zeros(|\hat{A}|, T)$
$\quad$**for** $t := 0$ **to** T **do**
$\quad\quad$**for** $i := 0$ **to** $|\hat{A}|$ **do**
$\quad\quad\quad out_{unpadded}[i][t] = out_{padded}[i][t]$

$\quad a = ComputeAlpha(out_{unpadded}, label, bl)$
$\quad out_{unpadded}^{flipped} = fliplr(out_{unpadded})$
$\quad label_{reversed} = reverse(label)$
$\quad b = ComputeAlpha(out_{unpadded}^{flipped}, label_{reversed}, bl)$
$\quad b = flipud(fliplr(b))$
$\quad ab = a * b$
$\quad lab = zeros(|\hat{A}|, T)$
$\quad$**for** $s := 0$ **to** S **do**
$\quad\quad i = max(0, floor(\frac{s-1}{2}))$
$\quad\quad$**if** s **mod** $2 = 0$ **then**
$\quad\quad\quad$**for** $t := 0$ **to** T **do**
$\quad\quad\quad\quad lab[bl][t] = lab[bl][t] + ab[s][t]$
$\quad\quad\quad\quad ab[s][t] = \frac{ab[s][t]}{out_{unpadded}[bl][t]}$
$\quad\quad$**else**
$\quad\quad\quad$**for** $t := 0$ **to** T **do**
$\quad\quad\quad\quad lab[label[i]][t] = lab[label[i]][t] + ab[s][t]$
$\quad\quad\quad\quad ab[s][t] = \frac{ab[s][t]}{out_{unpadded}[label[i]][t]}$

$\quad lh = zeros(T)$
$\quad$**for** $t := 0$ **to** T **do**
$\quad\quad lh[t] = \sum_{s=1}^{S} ab[s][t]$
$\quad loss = - \sum_{t=1}^{T} \ln lh[t]$
$\quad softmaxGrad = zeros(|\hat{A}|, (\bar{W}_{padded}/4))$
$\quad$**for** $t := 0$ **to** T **do**
$\quad\quad$**for** $i := 0$ **to** $|\hat{A}|$ **do**
$\quad\quad\quad softmaxGrad[i][t] = out_{unpadded}[i][t] - \frac{lab[i][t]}{out_{unpadded}[i][t]*lh[t]}$

$\quad$**return** $loss, softmaxGrad$