Algorithm 1: Averaged Structured Perceptron

Input: Training sentences: $\{s_i\}_{i=1}^{N}$

$\theta \leftarrow 0$

for $t \leftarrow 1 \ldots T$ **do**

  for $i \leftarrow 1 \ldots N$ **do**

    $(\hat{g}_i, \hat{u}_i) = \arg\max_{g_i, u_i} \Phi(g_i, u_i, s_i, \mathcal{KB}) \cdot \theta$

    if $(u_i^+, g_i^+) \neq (\hat{u}_i, \hat{g}_i)$ **then**

      $\theta \leftarrow \theta + \Phi(g_i^+, u_i^+, s_i, \mathcal{KB}) - \Phi(\hat{g}_i, \hat{u}_i, s_i, \mathcal{KB})$

return $\frac{1}{T}\sum_{t=i}^{T} \frac{1}{N}\sum_{i=1}^{N} \theta^i$