Coverage mechanism

Coverage vector:

$$c^j = \sum_{j'=0}^{j-1} p^{j'}$$

Modified attention:

$$e_i^j = w^T \tanh(W_h h_i + W_s s_j + w_c c_i^j + b_{attn})$$

Coverage loss:

$$covloss_j = \sum_i \min(p_i^j, c_i^j)$$