

The log-loss framework

- ▶ Algorithm A predicts a sequence c^1, c^2, \dots, c^T over alphabet $\Sigma = \{1, 2, \dots, k\}$
- ▶ The prediction for the c^t th is a distribution over Σ :
 $\mathbf{p}_A^t = \langle p_A^t(1), p_A^t(2), \dots, p_A^t(k) \rangle$