Hidden Markov Transformer (HMT) Basics and Combined Objective for HMT with SCST and Dynamic Wait-k

Hidden Markov Transformer (HMT) Basics

The **Hidden Markov Transformer (HMT)** combines the principles of Hidden Markov Models (HMMs) and Transformers for sequence modeling. The key mathematical formulations are as follows:

Joint Sequence Probability in HMT

$$P(X, Z) = P(Z) \prod_{t=1}^{T} P(x_t \mid z_t) \tag{1}$$

where:

- $X = (x_1, x_2, \ldots, x_T)$: Observed sequence (e.g., source or target sequence),
- $Z = (z_1, z_2, \ldots, z_T)$: Hidden states (latent variables),
- $P(Z)$: Transition probability between hidden states,
- $P(x_t \mid z_t)$: Emission probability of observing x_t given z_t.

Transformer Attention for Transition

The transition probability $P(Z)$ is modeled using self-attention:

$$A_{ij} = \text{softmax}\left(\frac{Q_i K_j^\top}{\sqrt{d_k}}\right) \tag{2}$$

where:

- Q, K, V: Query, key, and value matrices for self-attention,
- d_k: Dimensionality of the hidden states.

Combined Objective for HMT with SCST and Dynamic Wait-k

The overall objective integrates **Self-Critical Sequence Training (SCST)**, **HMT probabilities**, and the **Wait-k policy**:

HMT Loss

$$\mathcal{L}_{HMT} = \mathbb{E}_Z \left[-\sum_{t=1}^{T} \log P(x_t \mid z_t) \right] \tag{3}$$

SCST Loss

The SCST reward $R(\theta)$ is computed as:

$$R(\theta) = \sum_{t=1}^{T} (r_t - b_t) \log P(y_t \mid x; \theta) \tag{4}$$

where:

- r_t: Reward at step t (e.g., BLEU score),
- b_t: Baseline reward (computed using greedy decoding).

The SCST loss is:

$$\mathcal{L}_{SCST} = -\sum_{t=1}^{T} (r_t - b_t) \log P(y_t \mid x; \theta) \tag{5}$$

Dynamic Wait-k Selection

The Wait-k function dynamically determines the source tokens to read before decoding:

$$g(t; k) = \min(k + t - 1, |Z|) \tag{6}$$

where:

- k: Number of source tokens to "wait",
- t: Current decoding step,
- $|Z|$: Length of the latent states.

To optimize k based on translation quality and latency:

$$k^* = \arg\max_k [\text{BLEU} - \alpha \cdot AL] \tag{7}$$

where:

- $AL = \frac{1}{\tau} \sum_{t=1}^{\tau} [g(t) - t - 1] \cdot \frac{|y|}{|x|}$ AL: Average Lagging (latency metric),

α: Hyperparameter balancing quality and latency.

Final Combined Loss

$$\mathcal{L} = \mathcal{L}_{HMT} + \lambda \mathcal{L}_{SCST} \tag{8}$$

where λ is a weighting factor for reinforcement learning.