



1. The weights of the observation encoder is shared across the network.
2. Weighted sum of the encoded sequence.
3. Program embeddings are randomly initialized.

