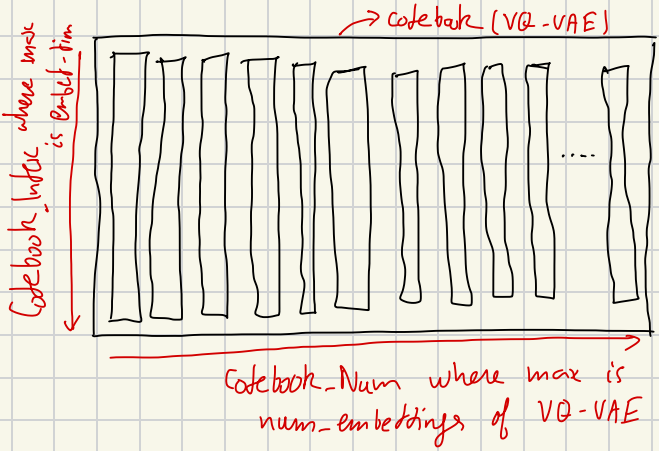
VQ-VAE Custom Environment :-

⊕ State → A latent space vector of length embed_dim
↳ Reset → randomly initialize a latent space vector

S_t = [latent vector, action_history]

⊕ Action → Tuple of 2 values :- (Codebook_Num, Codebook_Index)

↓
We will add an extra codebook which when selected would mean "terminate" action.



⊕ Reward → A combination of delta change in accuracies (given by the surrogate model for timestamp t)
↳ we might want to add a penalty for larger network sizes.

⊕ Functions :-

① Reset → Randomly initialize a state vector of length embed_dim.
→ Reset action history
→ Reset observation variables

② Step → ⓐ Check to make sure terminate codebook_num is not selected.
ⓑ Extract codebook_num of replace value of state vector with codebook vector at codebook_index.
ⓒ Append action in previous actions
ⓓ Calculate Reward → Decode the network at get accuracy from surrogate.

(e) Check if max_allowed_action number reached?
(f) Set done accordingly and return observations.

→ Things that would need to be defined while initializing the custom environment :-

(1) embed_dim → the codebook vectors and state vector dimension

(2) num_embeddings → # of vectors in the codebook + 1 for terminate action.

(3) Surrogate Model → Used in reward calculation, to get the accuracy for a model.

(4) Decoder → Used to convert the state vector into a representation understood by the surrogate model for reward calculation

(5) max_allowed_actions → # of steps after which the env should terminate.

(6) Codebook → The codebook learned by the VQ-VAE model

(7) num_previous_actions → number of previous actions to keep track of in the action_history.