CS786A

Assignment 4

K. Siddarth
150312

1. Question 1

In this question, we store the encoding as a combination of the world vector and state m. We retrieve by comparing the current context vector with the encoding vectors. The average number of elements correctly recalled is 7.78, which was computed by averaging the number of elements recalled over 100 evaluations. The average scheduling load was around 17.44. We use dot product to compute strength of activation. Using an RBF kernel like activation metric provided a retrieval rate of 8.81

2. Question 2

Here we model the drift parameter as being sampled from a mixture of Gaussians with mixing probabilities of 0.5. The Gaussians have variance 1 and mean 0 and 1 respectively denoting the small and large change distributions.

By having all the encoding happen only in the last 10 time steps, we find that the scheduling load is maximum (500.0) but the retrieval rate is also very high, around 8.8. If we do the opposite, we can reduce the encoding load while still keeping the average number of successful retrievlas high. We must be careful to ensure that a majority of the scheduling time points are chosen close to the backend of the encoding period. This is done by having a gap of 99 between the first 5 elements, while only having a gap of 1 for the next 5. This way, the median is 99 and most of the encodings are recent. With this method we have an average encoding load of 5.1 while still having an average number of successful retrievals be 7.6

3. Question 3

Here, during the encoding period, we also store the delta values whenever we encode to memory. These delta values are then used to estimate the mixing probabilities, means and variances of the GMM using EM. This learnt distribution is then used to sample the new deltas during retrieval.

We achieved an average successful retrieval rate of 7.81 with an average encoding load of 5.1