Block 3: Prediction

Below are exercises accompanying the lecture on prediction. Building on your GHMM implementation from Block 2, you will now compute belief states and next-token vectors, and use them to probe the predictive structure of canonical processes (Z1R, Mess3, RRXOR). Spend ~30 minutes on these exercises; start with the core set, and progress to the extension if you have time.

Core exercises.

1. Using your programming language of choice, implement a function that takes in:
    a. The transition matrices of a GHMM (a three-tensor)
    b. The initial vector
    c. A sequence of tokens

    and outputs the corresponding: belief state and next-token (prediction) vector.
2. Test your function on the zero-one-random (Z1R) process when initialised in the state $\eta^{(\varnothing)}$ = [1/3 1/3 1/3], and compare to the results of the worked example on Slide 12.
3. Visualise belief state geometry and the next-token geometry for the Mess3 process initialised in the uniform distribution. Does it look like the theoretical prediction in Fig. 1 of Shai et al?
4. Derive the belief state update rule on Slide 5.
5. Show that the matrix elements of the matrix taking beliefs to next-token probabilities are given by expression stated on Slide 7.

Extension exercises.

1. Evaluate whether the linear map from beliefs to observation probabilities is invertible for:
    a. RRXOR
    b. Mess3
2. Complete the MSP of the Z1R process (explored in the lecture) by annotating the tokens with the correct emission probabilities.
3. Compute the MSP for the zero-random-random process.
4. For an element $a \in \mathcal{X}^*$, the corresponding *causal state equivalence class* is given by

    $$[a]_{\mathrm{C}} := \{x \in \mathcal{X}^* \,|\, \Pr(y \,|\, a) = \Pr(y \,|\, x),\ \forall\, y \in \mathcal{X}^*\}.$$

    Likewise, the corresponding *belief state equivalence class* is given by

    $$[a]_{\mathrm{B}} := \{x \in \mathcal{X}^* \,|\, \eta^{(a)} = \eta^{(x)}\}.$$

    How are these sets related to each other?