CSE 576: Topics in NLP

Consider the following notations for a Hidden Markov Model with states and observations denoted by q and o respectively, given model parameters λ;
$a_{ij} = P(q_{t+1} = s_j | q_t = s_i)$
$b_j(k) = P(o_k | q_t = s_j)$
$\alpha_t(j) = P(o_1, o_2, .., o_t, q_t = s_j | \lambda)$
$\beta_t(i) = P(o_{t+1}, o_{t+2}, .., o_T | q_t = s_i, \lambda)$

Provided this,

1. If $\zeta_t(i,j)$ denotes the probability of being in state s_i at time t and s_j at time t+1 given all observations and the model i.e. $\zeta_t(i,j) = P(q_{t+1} = s_j, q_t = s_i | O, \lambda)$,

   then show that $\zeta_t(i,j) = \frac{\alpha_t(i) \cdot a_{ij} \cdot b_j(o_{t+1}) \cdot \beta_{t+1}(j)}{P(O|\lambda)}$

2. Prove that $\beta_t(i) = \sum_{j=1}^{N} a_{ij} \cdot b_j(o_{t+1}) \cdot \beta_{t+1}(j)$ where $1 \leq i \leq N$ and $1 \leq t \leq T$.
   (N denotes total number of states and T represents total time instances)

3. If $\gamma_t(j)$ denotes probability of being in state s_j at time t given all observations and the model i.e. $\gamma_t(j) = P(q_t = s_j | O, \lambda)$,

   then show that $\gamma_t(j) = \frac{\alpha_t(j) \cdot \beta_t(j)}{P(O|\lambda)}$

(Show all intermediate steps and provide brief justification of each step for all three questions. Submit a handwritten/digital copy as a PDF file on Canvas.)