Forward Probabilities

Natalie Parde
UIC CS 421

Observation Likelihood

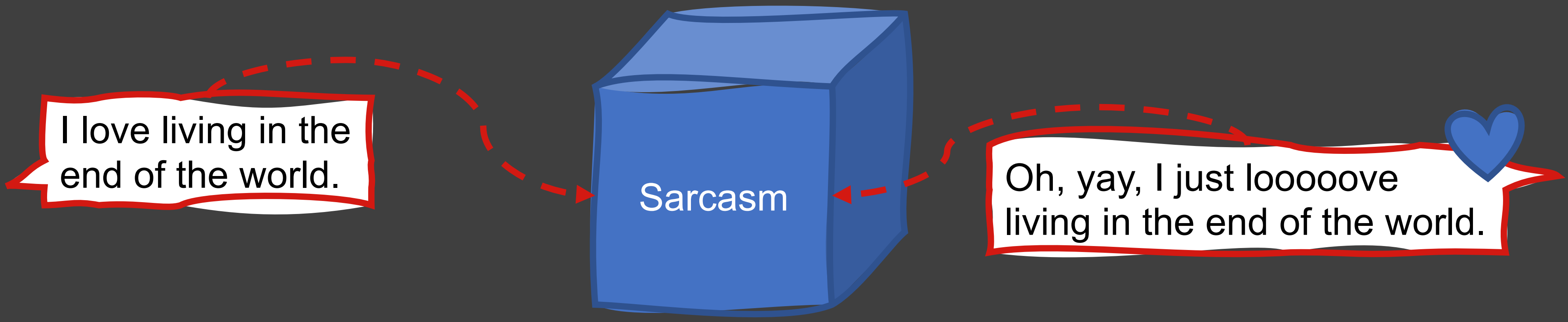
- Given a sequence of observations and an HMM, what is the probability that this sequence was generated by the model?
- Useful for two tasks:
  - Sequence classification
  - Selecting the most likely sequence

Sequence Classification

- Assuming an HMM is available for every possible class, what is the most likely class for a given observation sequence?
  - Which HMM is most likely to have generated the sequence?
- HMMs are commonly used in automated speech recognition (ASR) for this purpose
  - Given a set of sounds, what is the most likely word?

Most Likely Sequence

- Of two or more possible sequences, which one was most likely generated by a given HMM?



How can we compute the observation likelihood?

- Naïve Solution:
  - Consider all possible state sequences, *Q*, of length *T* that the model, λ, could have traversed in generating the given observation sequence, *O*
  - Compute the probability of a given state sequence from *A*, and multiply it by the probability of generating the given observation sequence for that state sequence
    - $P(O,Q \mid \lambda) = P(O \mid Q, \lambda) * P(Q \mid \lambda)$
  - Repeat for all possible state sequences, and sum over all to get $P(O \mid \lambda)$
- But, this is computationally complex!
  - $O(TN^T)$

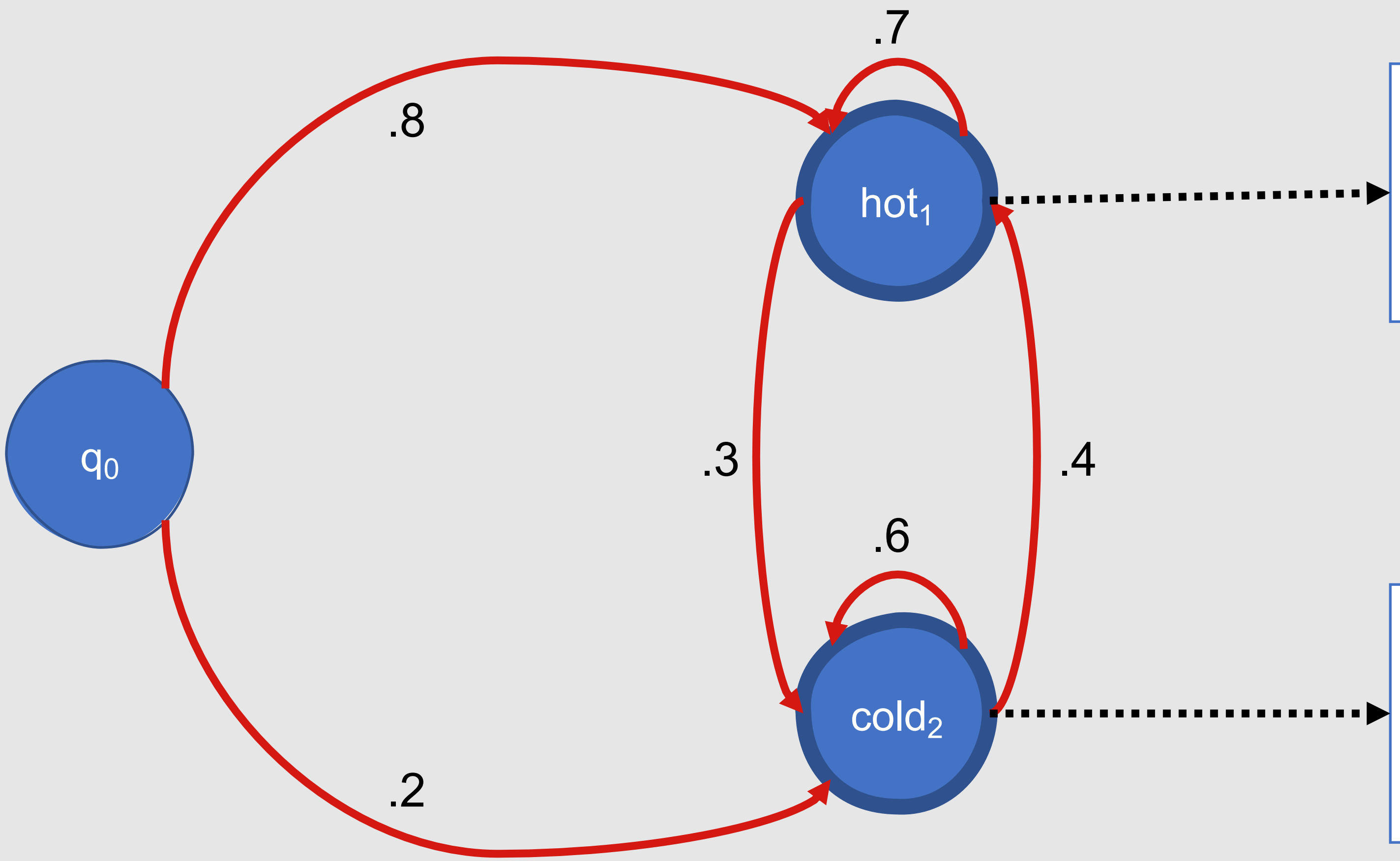
How can we compute the observation likelihood?

- Efficient Solution:
  - **Forward Algorithm:** Dynamic programming algorithm that computes the observation probability by summing over the probabilities of all possible hidden state paths that could generate the observation sequence.
  - Implicitly folds each of these paths into a single forward trellis
- Why does this work?
  - Markov assumption (the probability of being in any state at a given time *t* only relies on the probability of being in each possible state at time *t*-1).
- Works in $O(TN^2)$ time!

Sample Problem

- It is 2799 and you are a climatologist studying the history of global warming
- Unfortunately, you have no records of the weather in Baltimore for the summer of 2007, although you have some leading hypotheses of some key weather patterns, which you're representing using HMMs
- Fortunately, a major breakthrough occurs: you find Jason Eisner's diary, which lists how many ice cream cones he ate every day that summer
- You decide to focus on a three-day sequence:
    - Day 1: 3 ice cream cones
    - Day 2: 1 ice cream cone
    - Day 3: 3 ice cream cones

Current Leading HMM



B_1

$$\begin{bmatrix} P(1|hot_1) \\ P(2|hot_1) \\ P(3|hot_1) \end{bmatrix} = \begin{bmatrix} .2 \\ .4 \\ .4 \end{bmatrix}$$

B_2

$$\begin{bmatrix} P(1|cold_1) \\ P(2|cold_1) \\ P(3|cold_1) \end{bmatrix} = \begin{bmatrix} .5 \\ .4 \\ .1 \end{bmatrix}$$

How do you compute your forward probabilities?

- Let $\alpha_i(j)$ be the probability of being in state *j* after seeing the first *t* observations, given your HMM λ
- $\alpha_i(j)$ is computed by summing over the probabilities of every path that could lead you to this cell
  - $\alpha_i(j) = P(o_1, o_2 \ldots o_t, q_t = j|\lambda) = \sum_{i=1}^{N} \alpha_{t-1}(i) a_{ij} b_j(o_t)$
    - $q_t = j$ is the probability that the *t^th^* state in the sequence of states is state *j*
  - $\alpha_{t-1}(i)$: The previous forward path probability from the previous time step
  - a_{ij}: The transition probability from previous state q_i to current state q_j
  - $b_j(o_t)$: The state observation likelihood of the observed item o_t given the current state *j*

Formal Algorithm

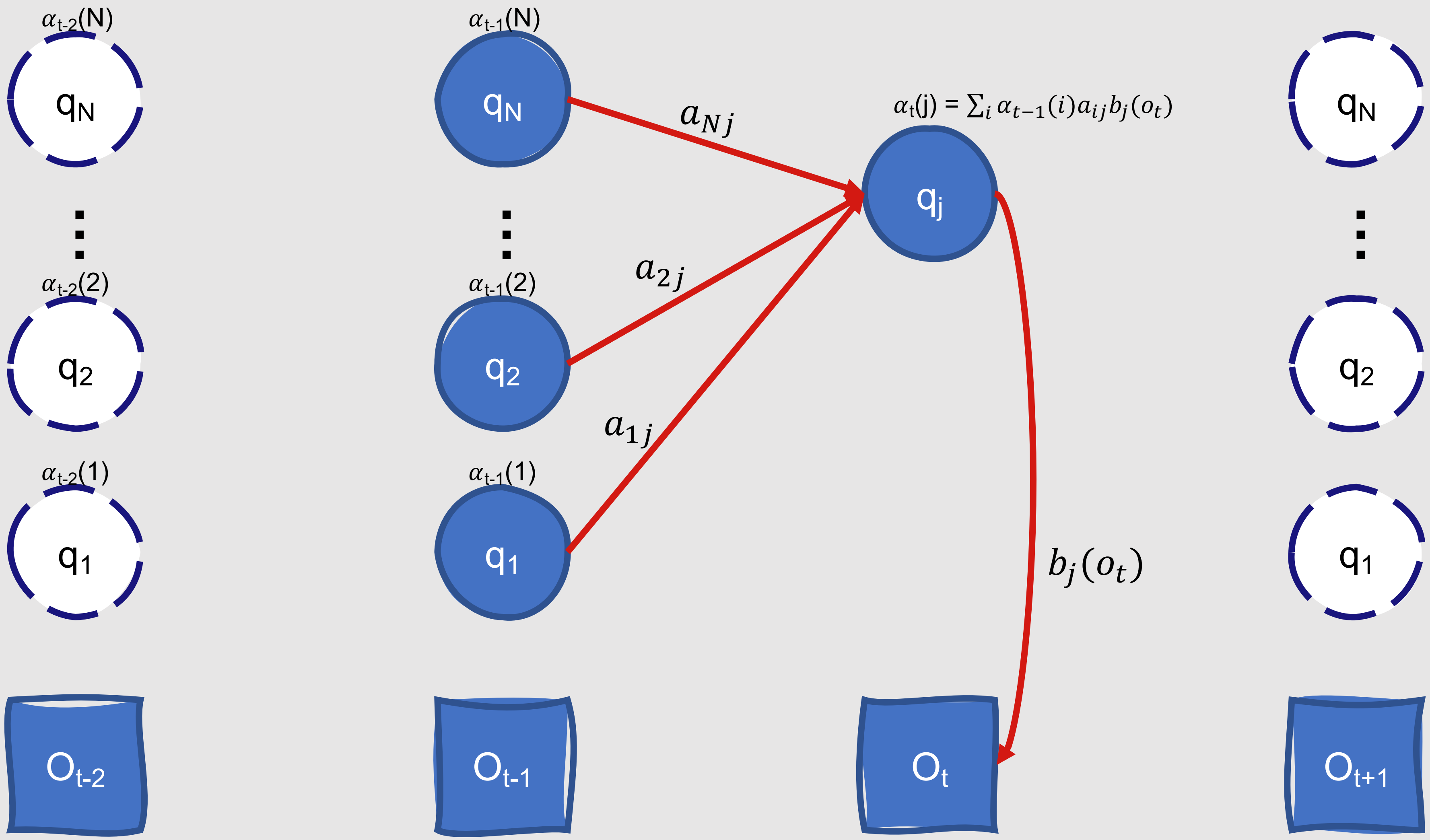
```
create a probability matrix forward[N+2,T]

for each state q in [1, …, N] do:
      forward[q,1] ← a_{0,q} * b_q(o_1)
for each time step t from 2 to T do:
      for each state q in [1, …, N] do:
            forward[q,t] ← Σ_{q'=1}^{N} forward[q',t-1] * a_{q',q} * b_q(o_t)
forwardprob ← Σ_{q=1}^{N} forward[q,T]
```

$$forward[q,1] \leftarrow a_{0,q} * b_q(o_1)$$

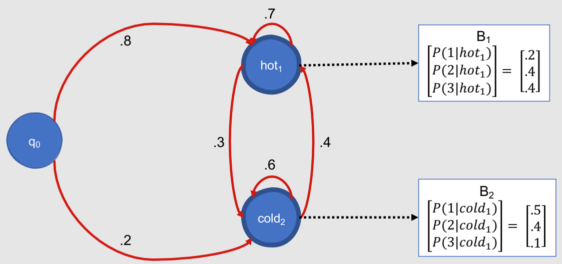
$$forward[q,t] \leftarrow \sum_{q'=1}^{N} forward[q',t-1] * a_{q',q} * b_q(o_t)$$

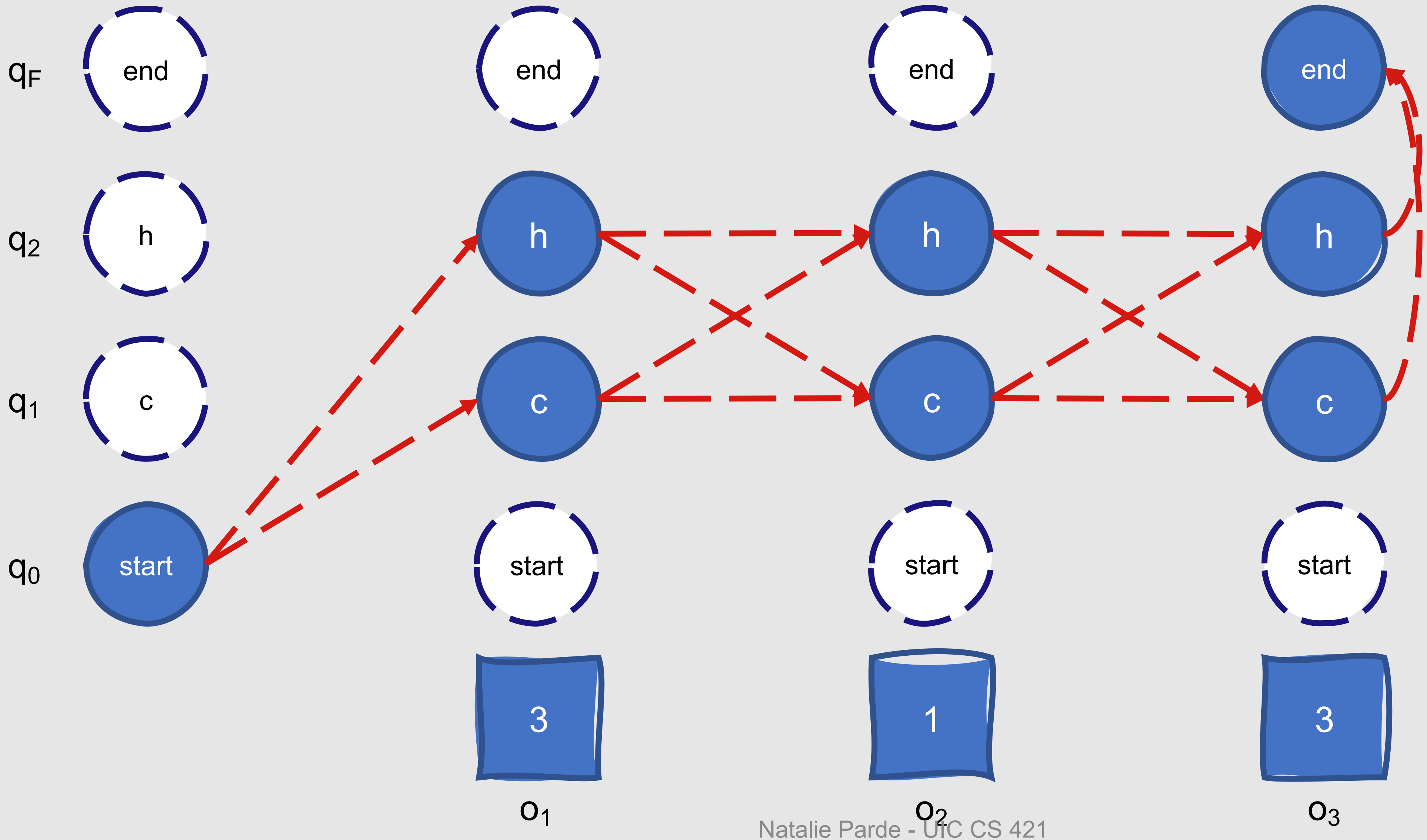
$$forwardprob \leftarrow \sum_{q=1}^{N} forward[q,T]$$

Forward Step

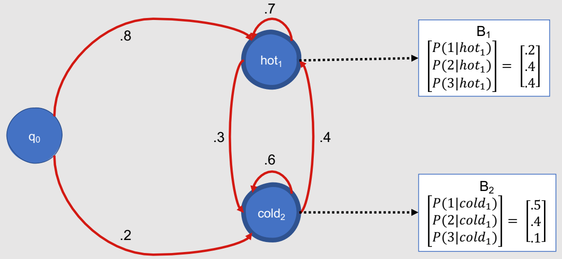


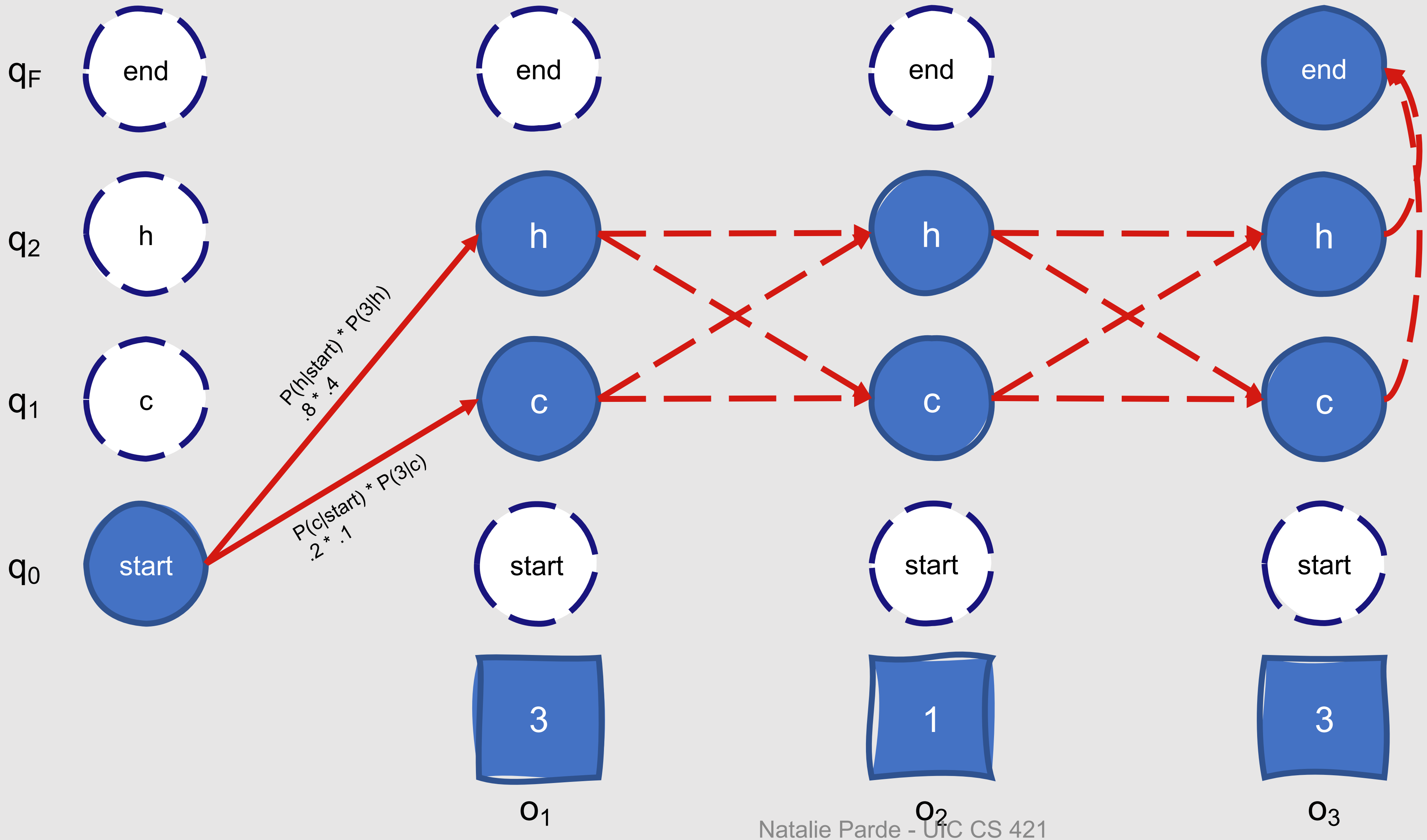
Forward Trellis



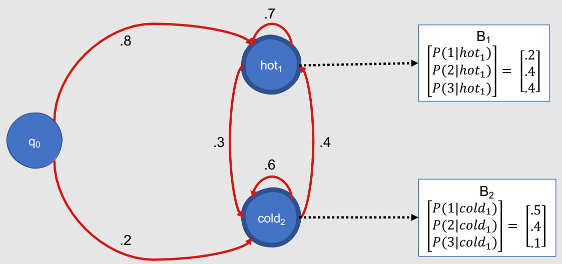


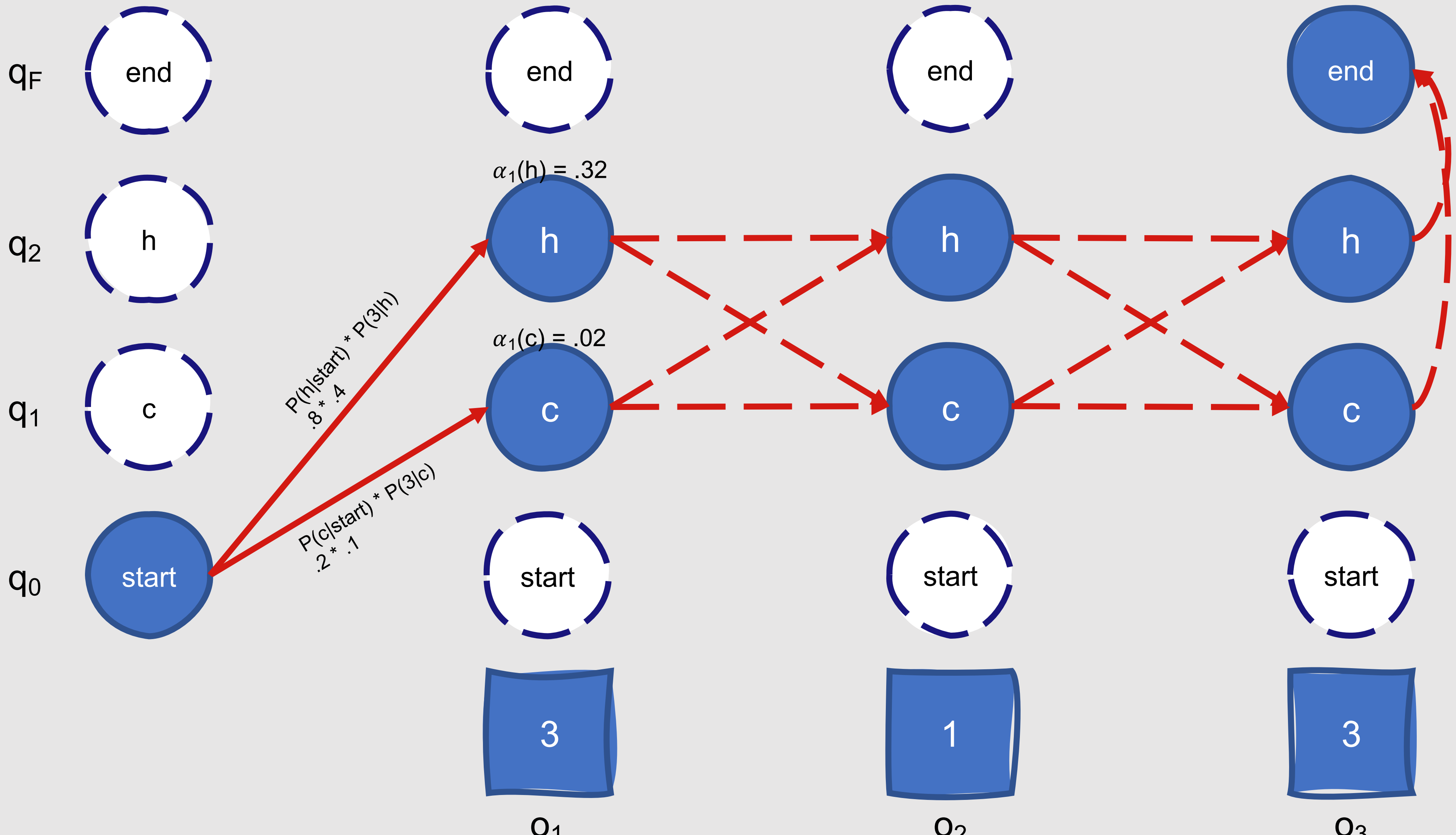
Forward Trellis



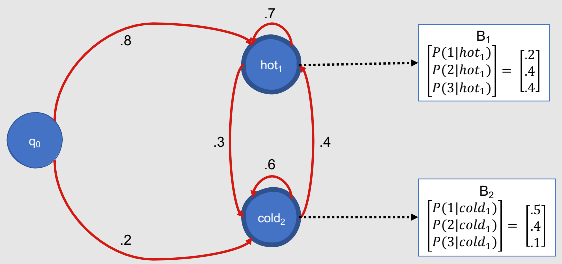


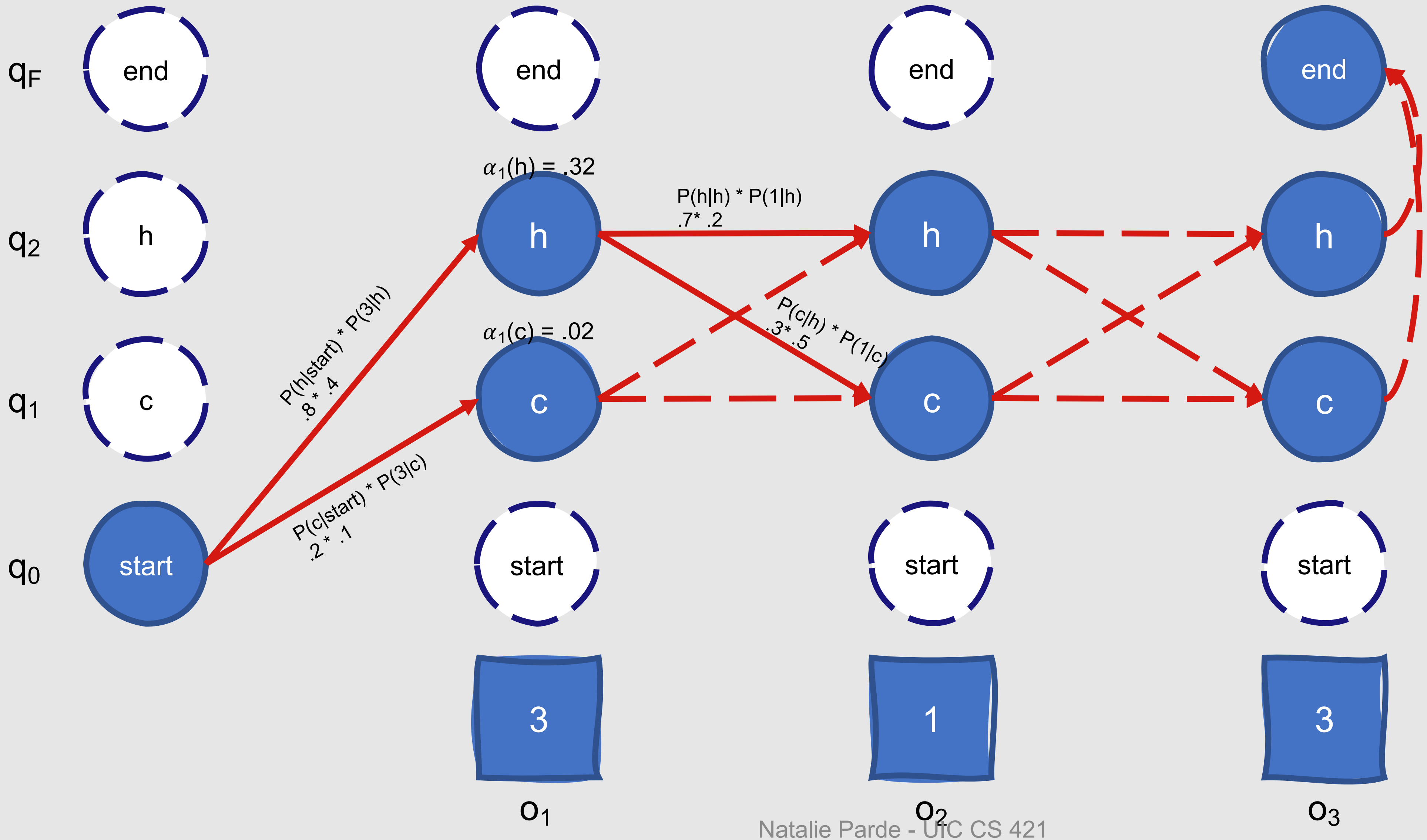
Forward Trellis



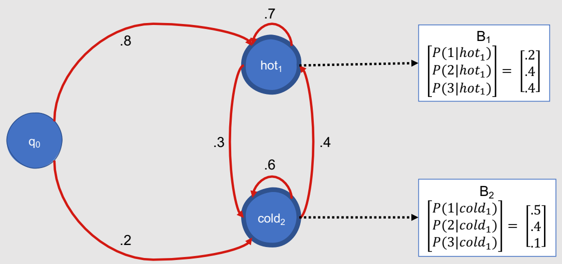


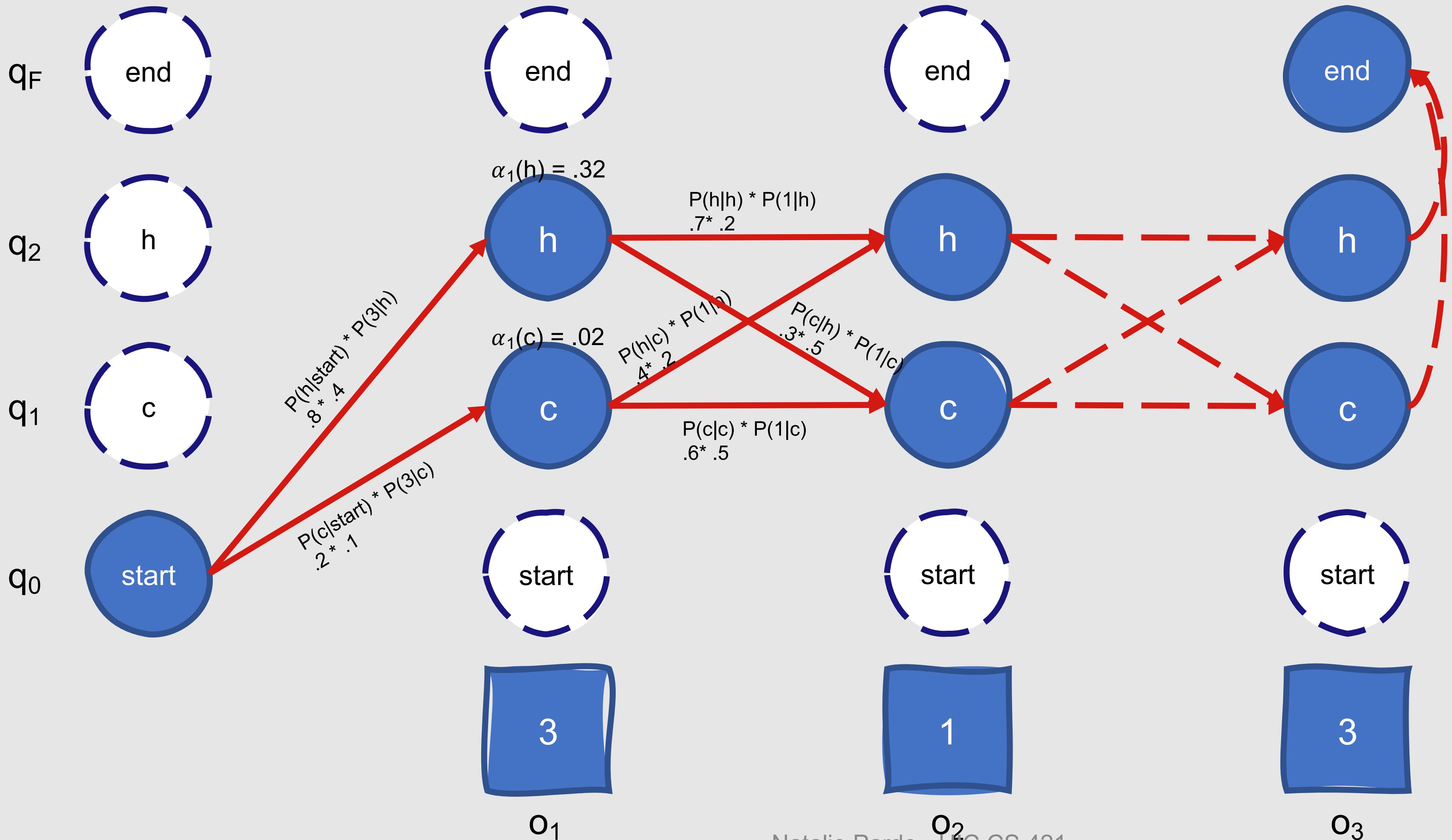
Forward Trellis



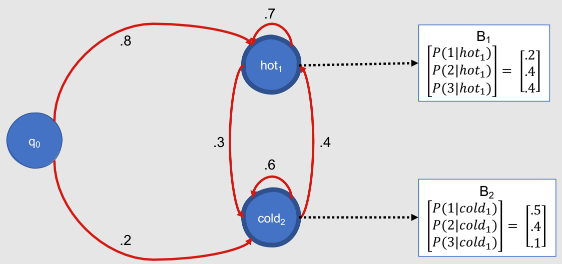


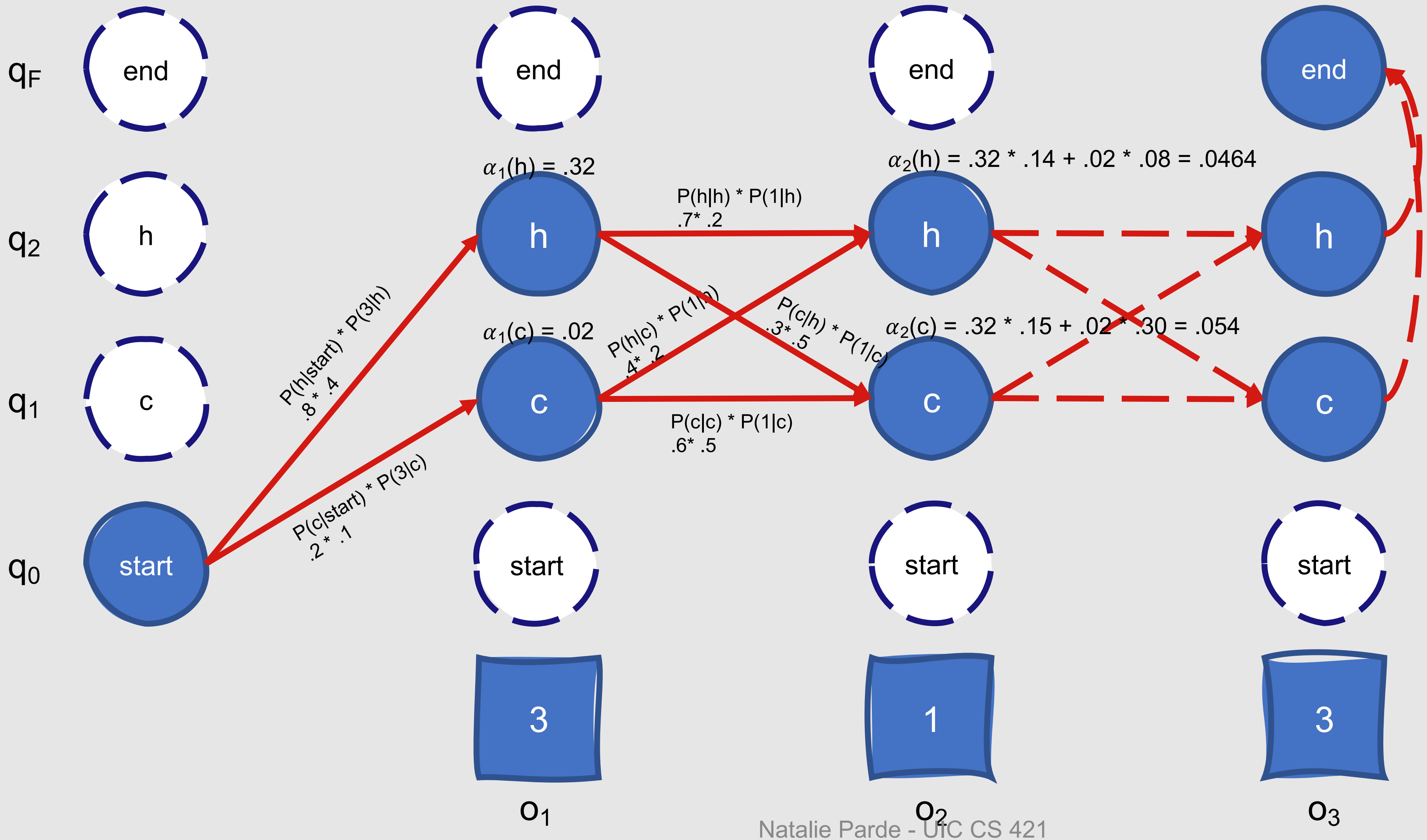
Forward Trellis



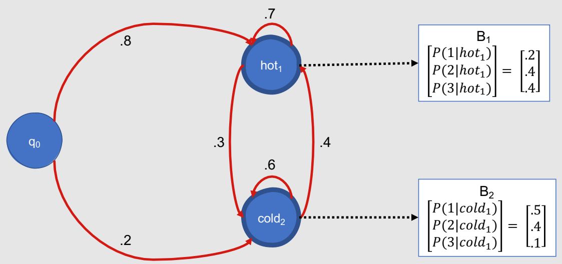


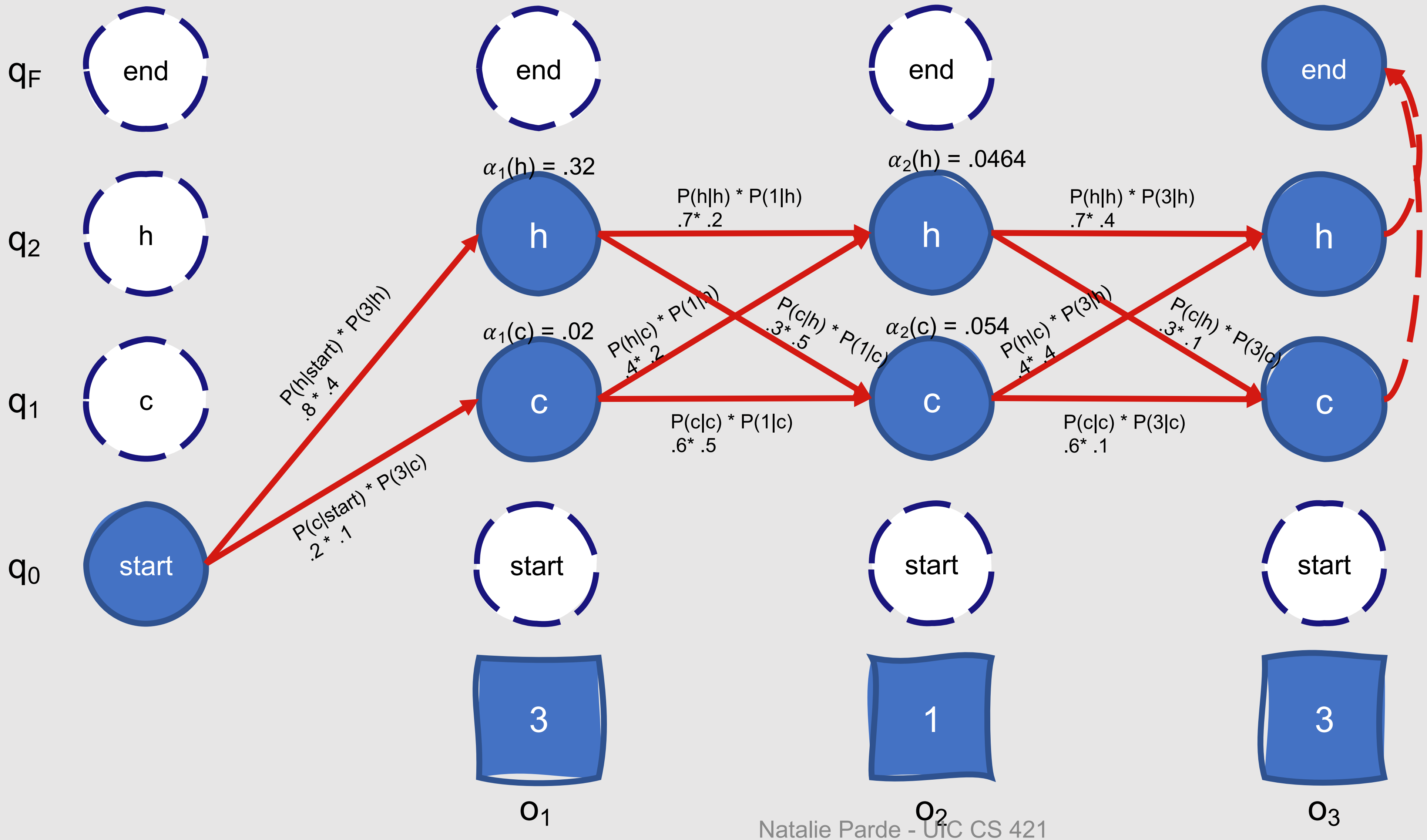
Forward Trellis



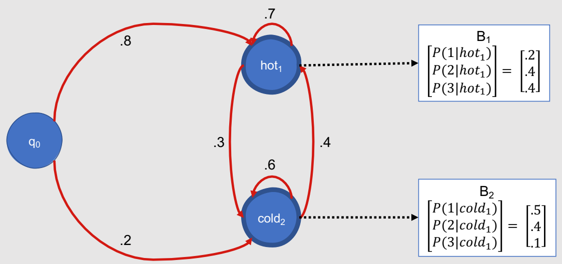


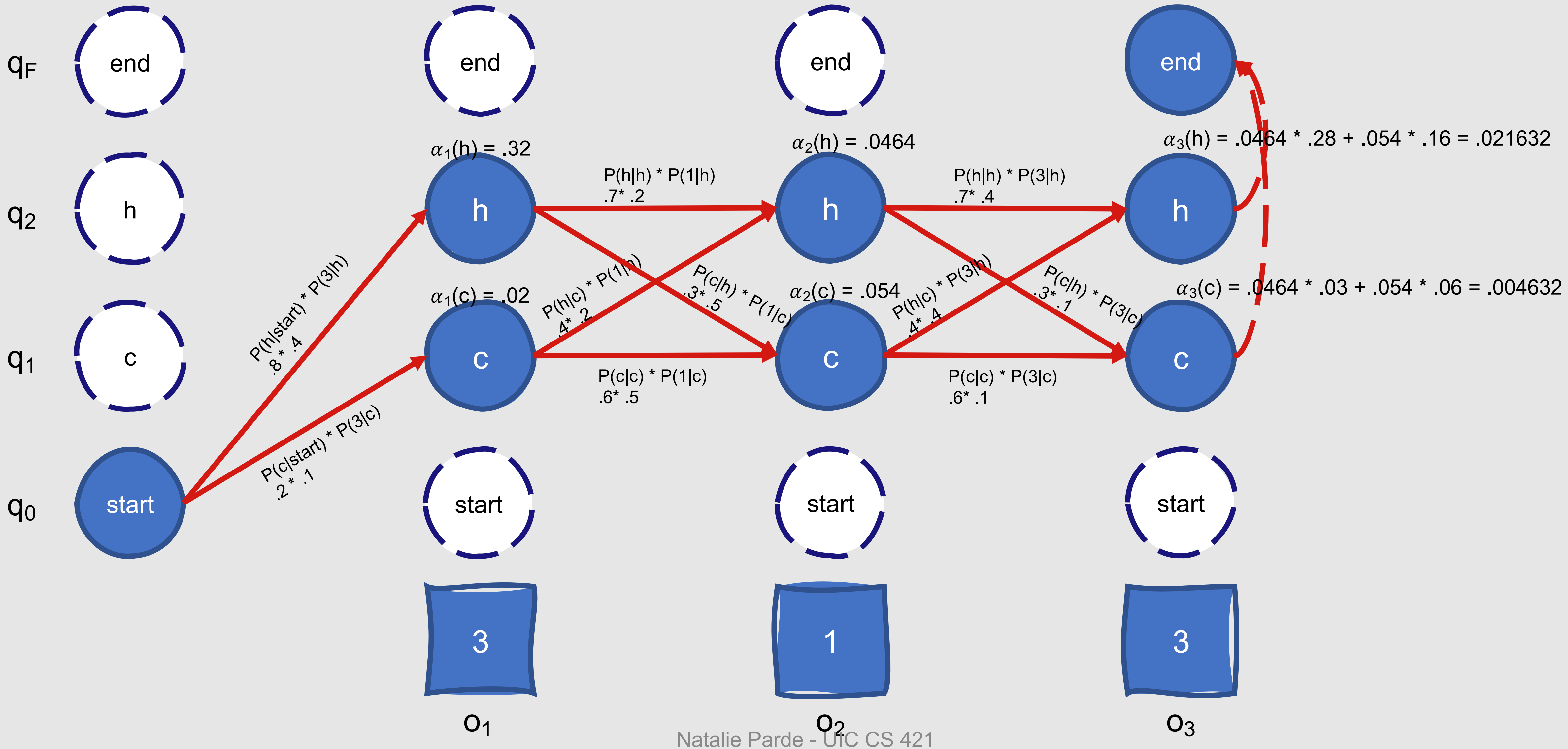
Forward Trellis





Forward Trellis





Forward Trellis