E9 261

09-03-2016

Recap ...

Solution to HMM-GMM re-estimation

$$\pi_i = \frac{\sum_{e=1}^{E} \gamma_i^e(1)}{E}$$

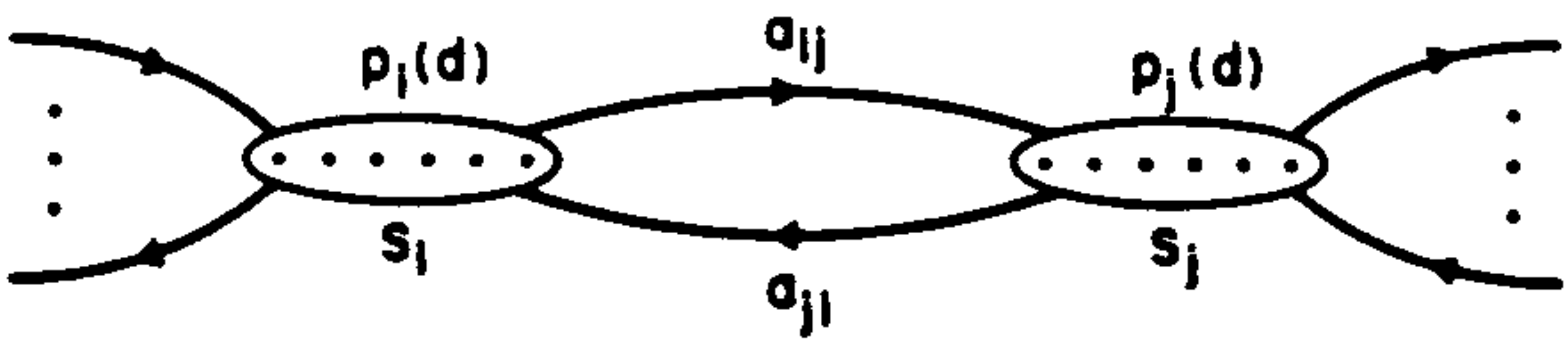
$$c_{i\ell} = \frac{\sum_{e=1}^{E} \sum_{t=1}^{T_e} \gamma_{i\ell}^e(t)}{\sum_{e=1}^{E} \sum_{t=1}^{T_e} \gamma_i^e(t)}$$

$$\mu_{i\ell} = \frac{\sum_{e=1}^{E} \sum_{t=1}^{T_e} \gamma_{i\ell}^e(t) o_t^e}{\sum_{e=1}^{E} \sum_{t=1}^{T_e} \gamma_{i\ell}^e(t)}$$

Left to right versus Ergodic HMMs

Other Considerations

- Implementation issues
  - Scaling
- Explicit duration modeling in HMM



Forward Recursion in Duration Model HMM

$$\alpha_t(i) = \sum_q \sum_d \pi_{q_1} \cdot p_{q_1}(d_1) \cdot P(\mathbf{o}_1 \mathbf{o}_2 \ldots \mathbf{o}_{d_1} | q_1)$$
$$\cdot a_{q_1 q_2} p_{q_2}(d_2) P(\mathbf{o}_{d_1+1} \ldots \mathbf{o}_{d_1+d_2} | q_2) \ldots$$
$$\cdot a_{q_{r-1} q_r} p_{q_r}(d_r) P(\mathbf{o}_{d_1+d_2+\cdots+d_{r-1}+1} \ldots \mathbf{o}_t | q_r) \quad (6.67)$$

where the sum is over all states q and all possible state durations d. By induction we can write $\alpha_t(j)$ as

$$\alpha_t(j) = \sum_{i=1}^{N} \sum_{d=1}^{D} \alpha_{t-d}(i) a_{ij} p_j(d) \prod_{s=t-d+1}^{t} b_j(\mathbf{o}_s) \quad (6.68)$$

where D is the maximum duration within any state. To initialize the computation of $\alpha_t(j)$ we use

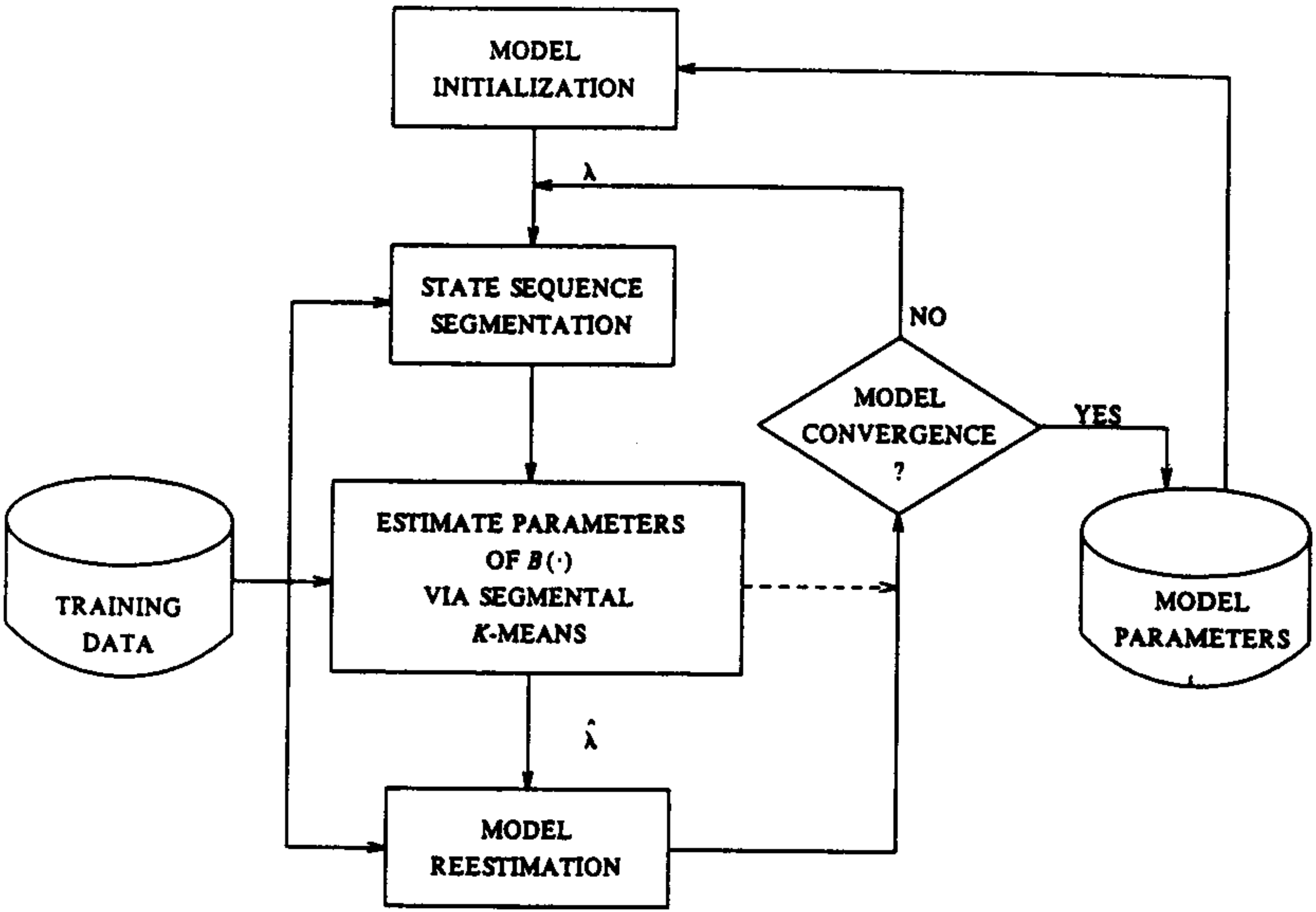
$$\alpha_1(i) = \pi_i p_i(1) \cdot b_i(\mathbf{o}_1) \quad (6.69a)$$

$$P(\mathbf{O}|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

Other Considerations

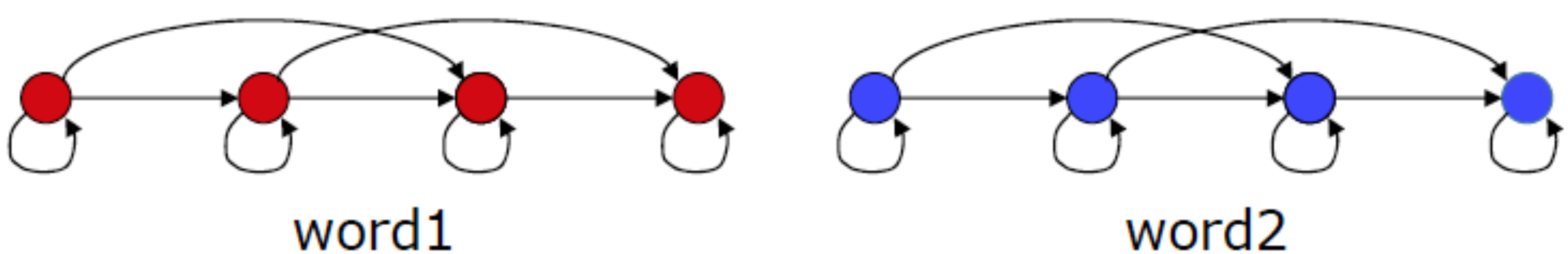
- Implementation issues
  - Scaling
- Explicit duration modeling in HMM
- Comparison of HMMs
- ML versus Bayesian Estimation
- Multiple observation sequence
- Initialization - Flat start

Segmental k-means Algorithm



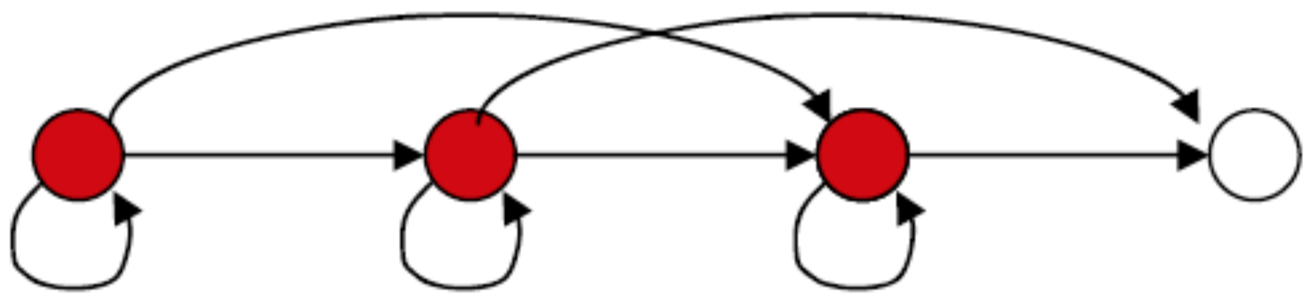
Dealing with Continuous Speech

- Isolated Word Recognition is limited
    - Need to deal with string of words
- Word sequences modeled with HMMs which are composed of word HMMs
- Given two words which are both Bakis topology

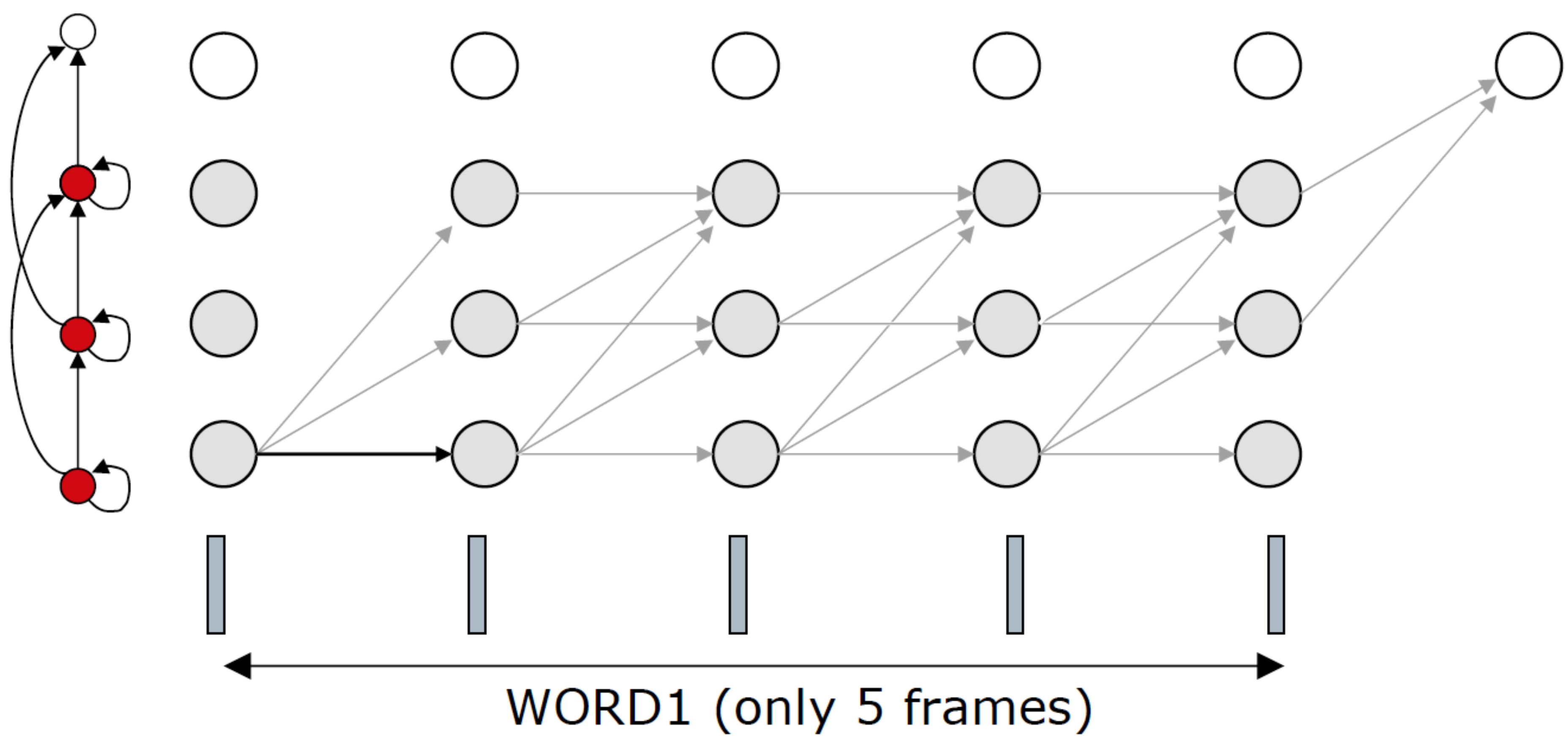


"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

Introducing Non-emitting states



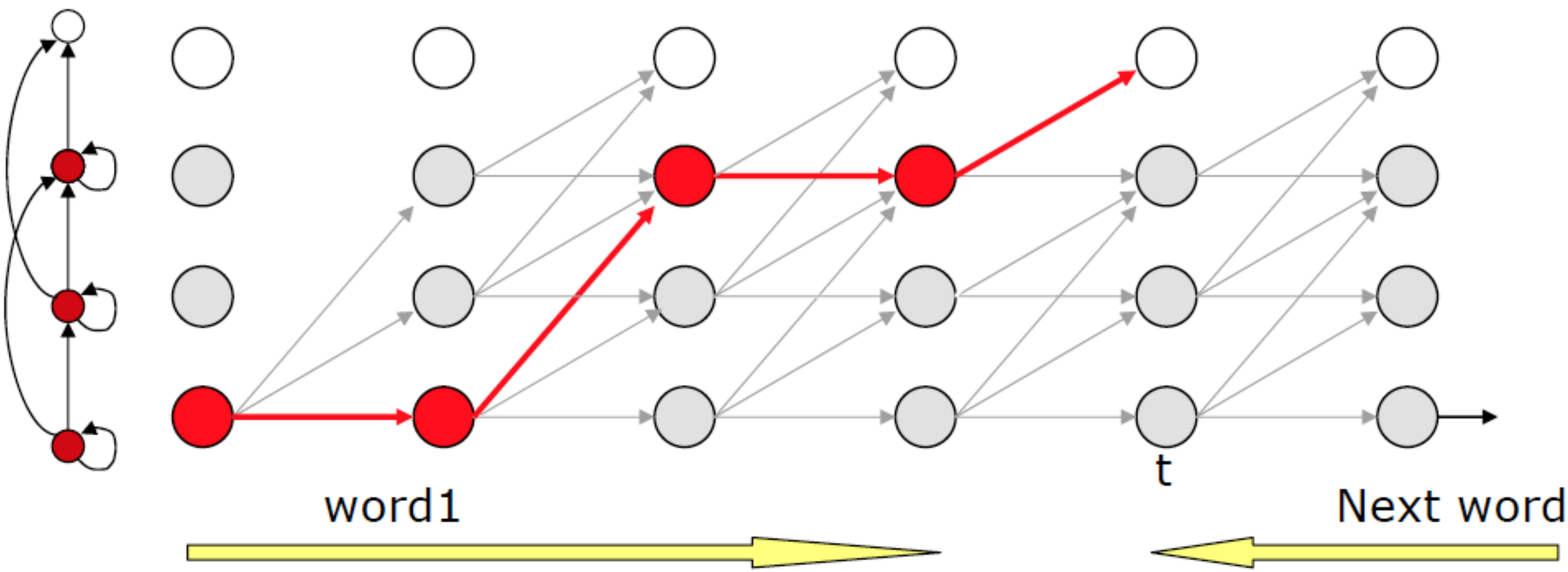
- Non-emitting or null states where no observation is emitted



"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

Introducing Non-emitting states

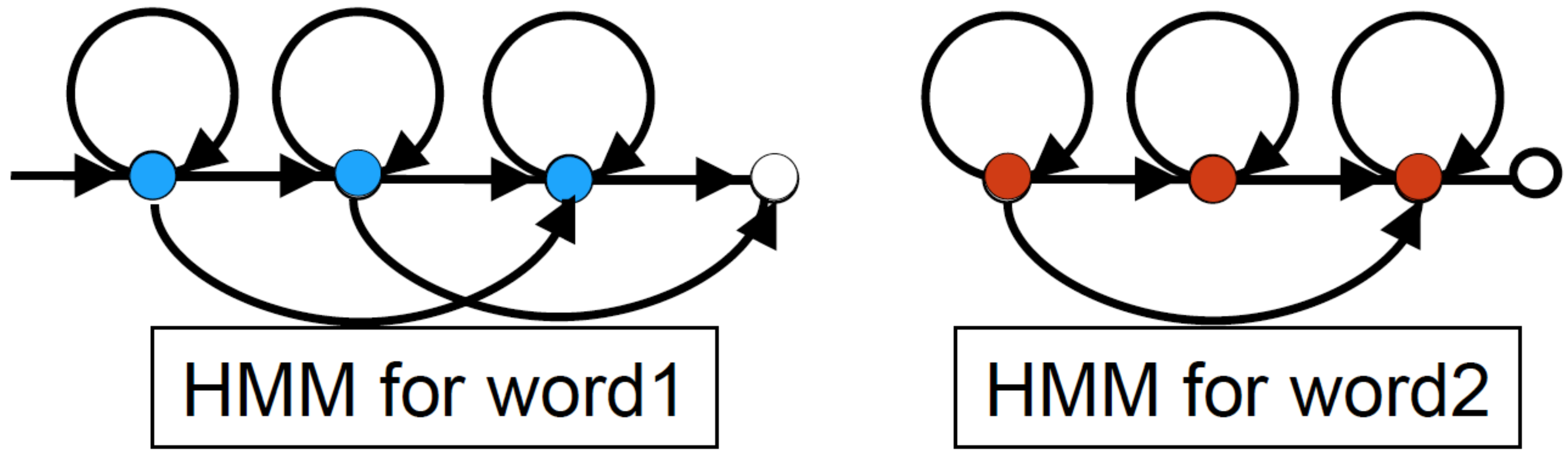
- Slight modifications to forward and backward recursion

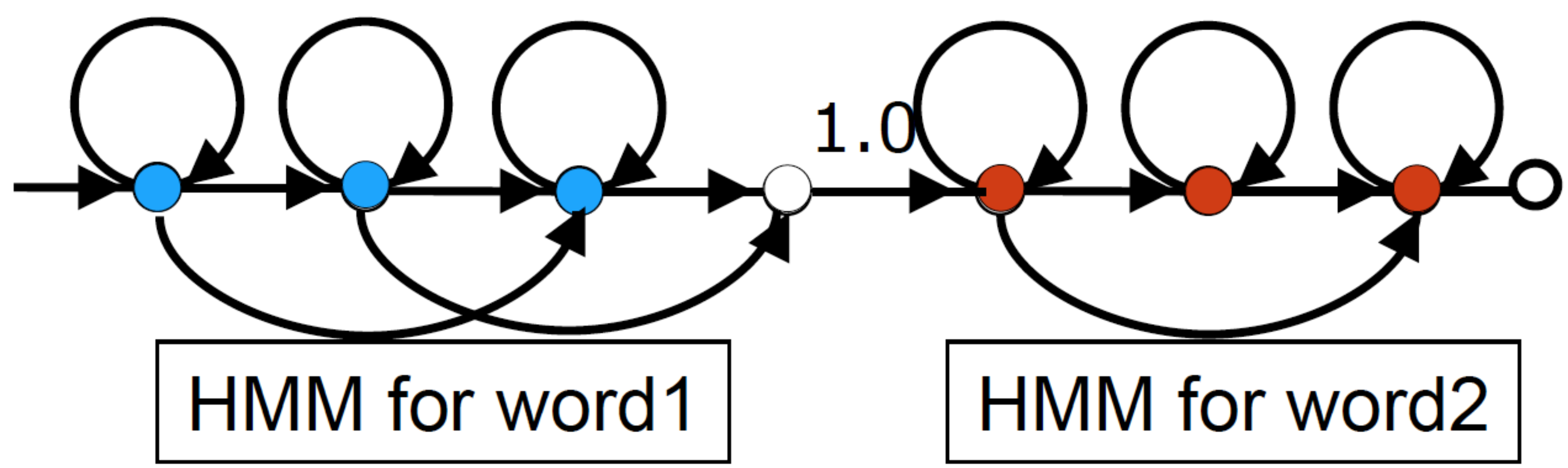


- Probability of reaching state N+1 at time t is equivalent to the probability of exiting word1 at time t

"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

Connecting Word HMMs

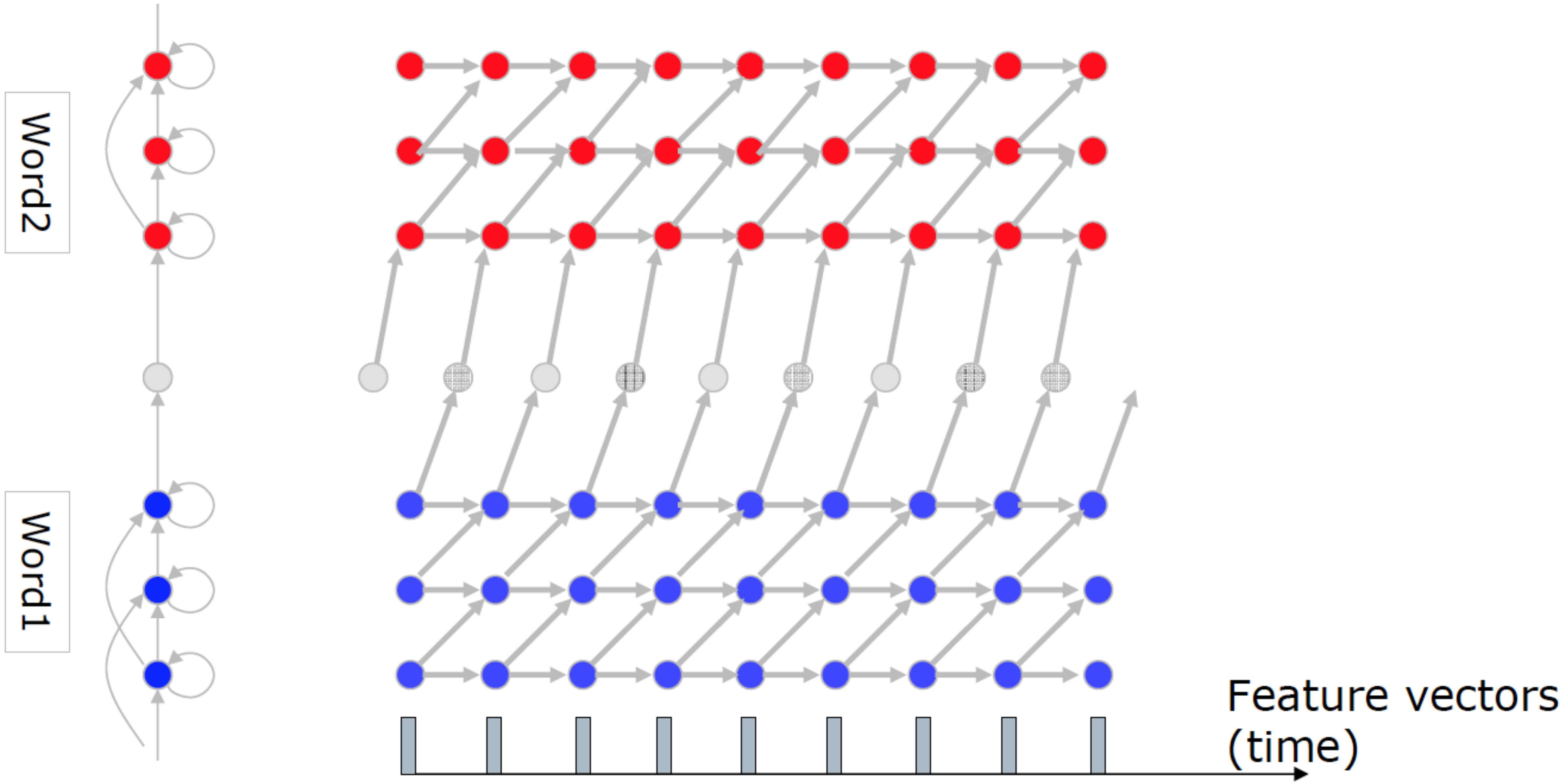




- One can also introduce the probability of word2 following word1 in this case.

"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

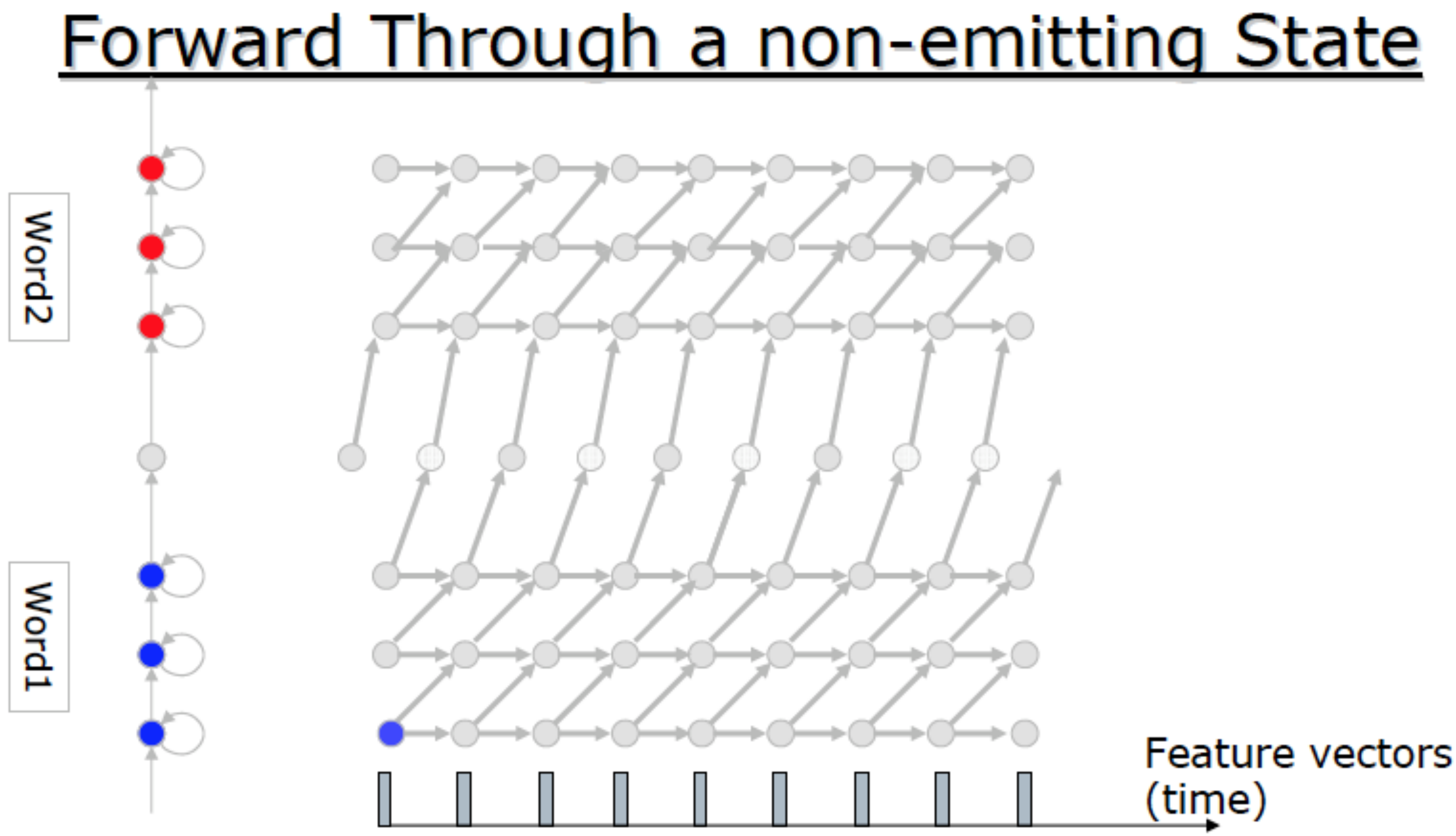
Trellis with connected word models



- Null states don't have time instant associated so they happen between two time instants.

"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

Trellis with connected word models

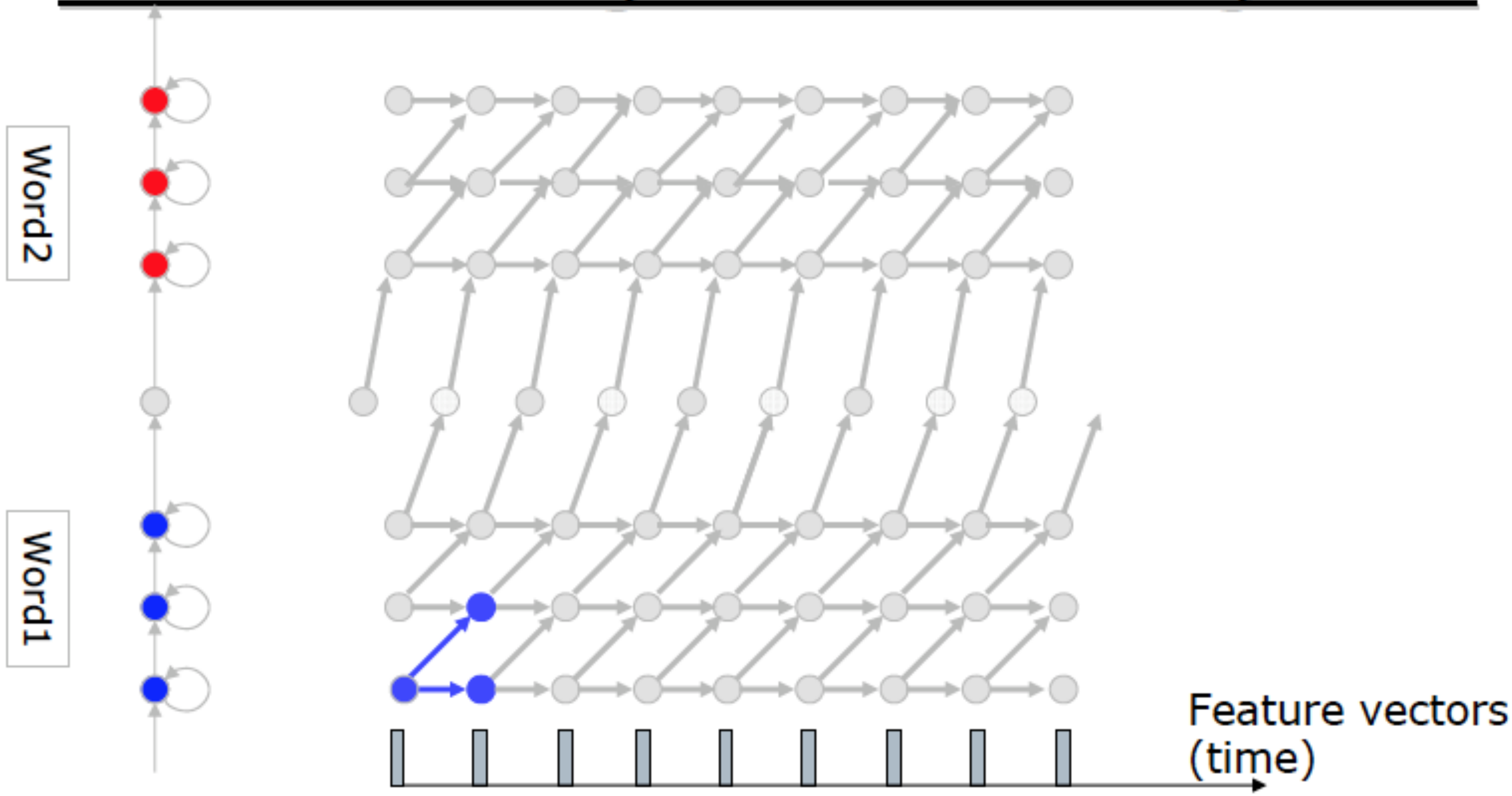


- At the first instant only one state has a non-zero forward probability

“Dealing with Connected Words”, Rita Singh and Bhiksha Raj, CMU.

Trellis with connected word models

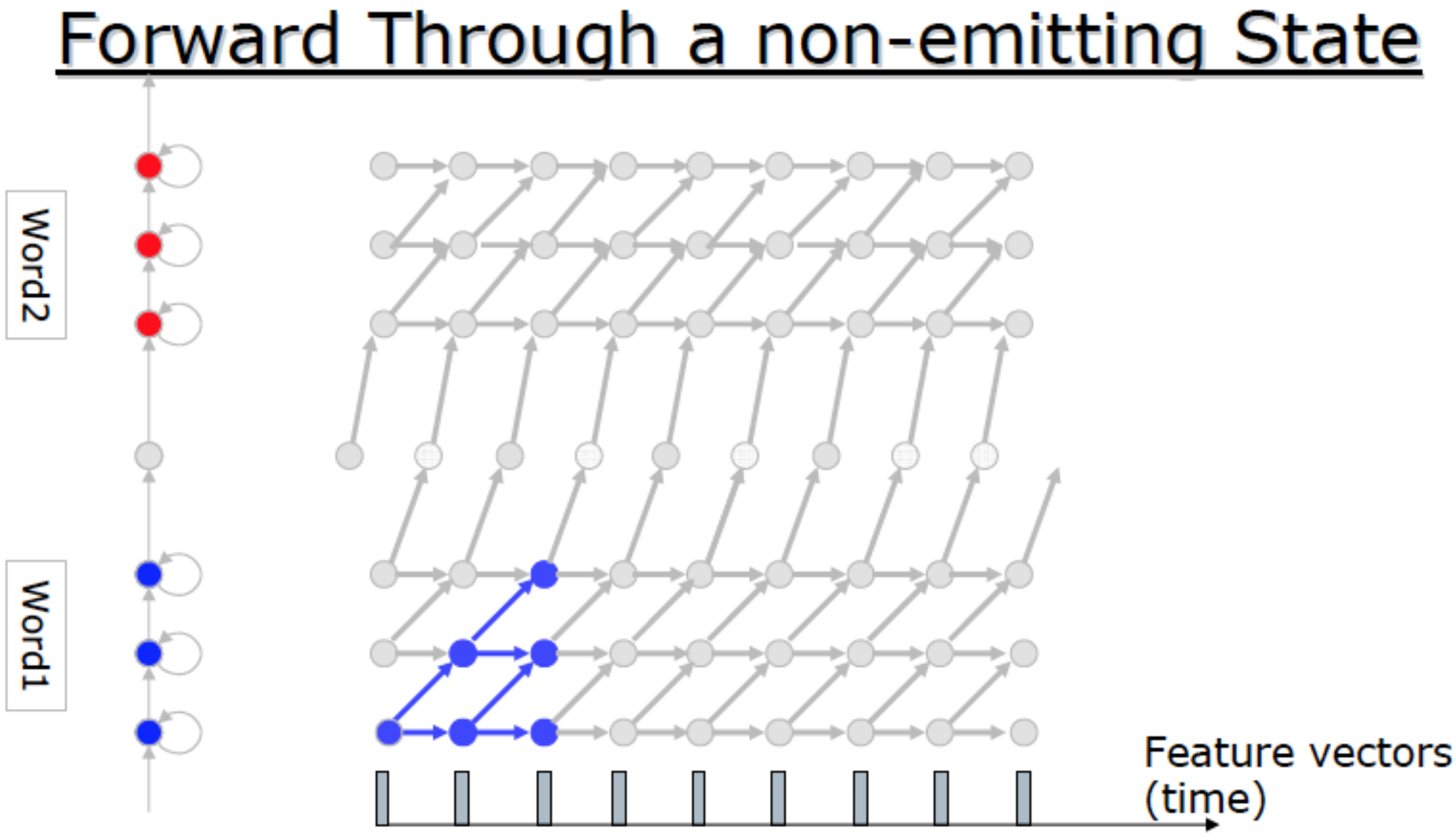
Forward Through a non-emitting State



- From time 2 a number of states can have non-zero forward probabilities
  - Non-zero alphas

"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

Trellis with connected word models

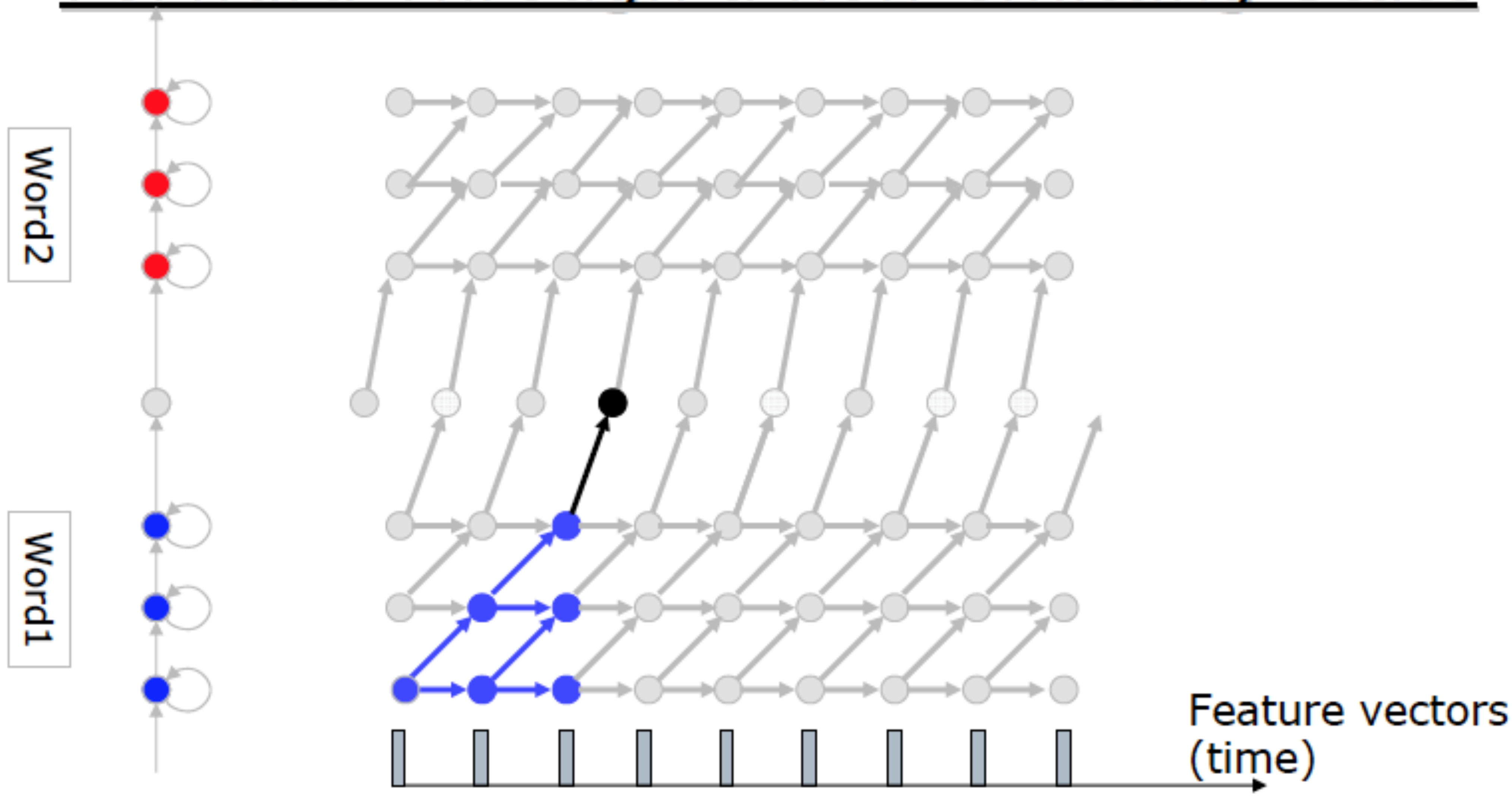


- From time 2 a number of states can have non-zero forward probabilities
  - Non-zero alphas

"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

Trellis with connected word models

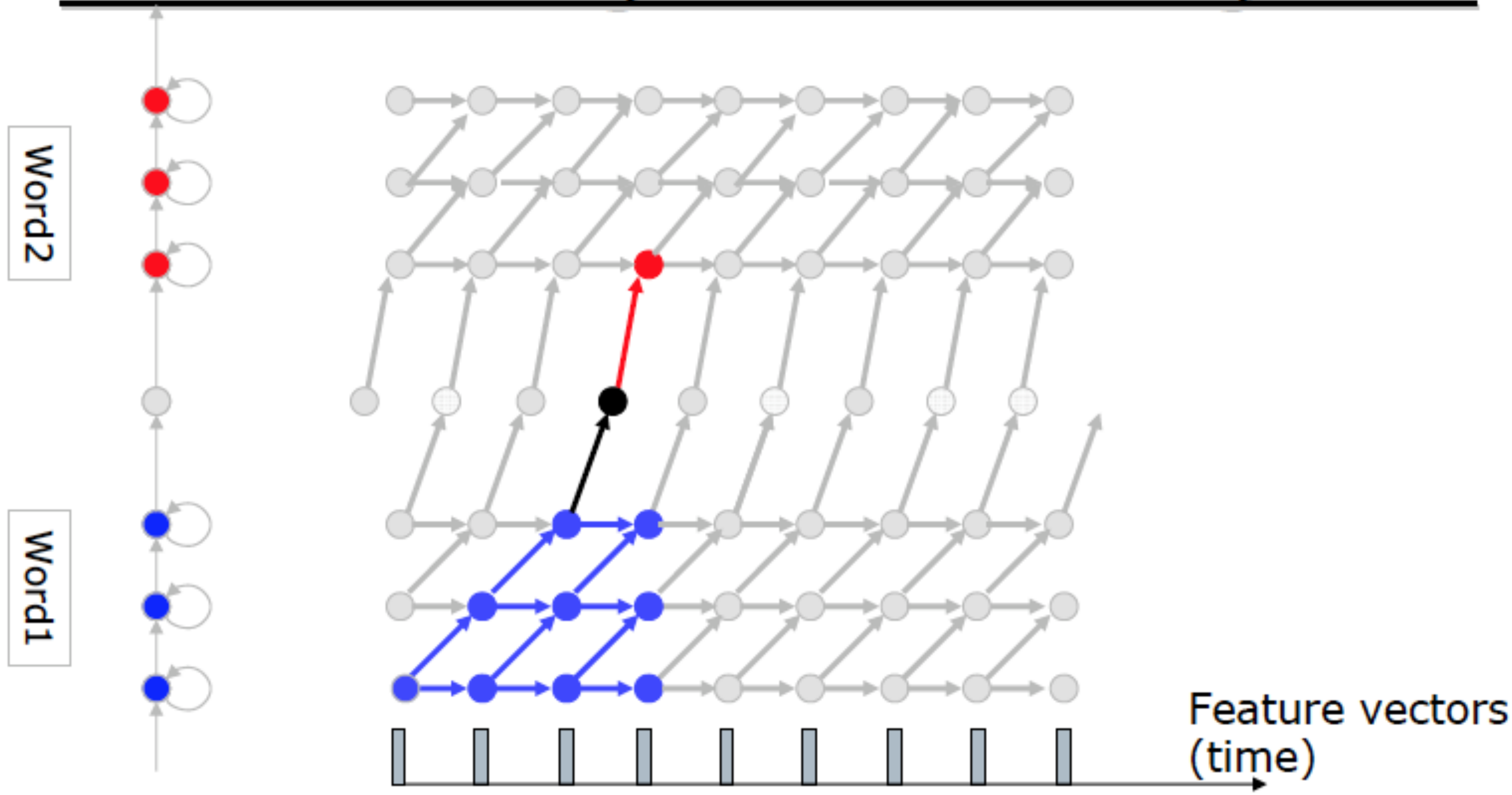
Forward Through a non-emitting State



- Between time 3 and time 4 (in this trellis) the non-emitting state gets a non-zero alpha

Trellis with connected word models

Forward Through a non-emitting State

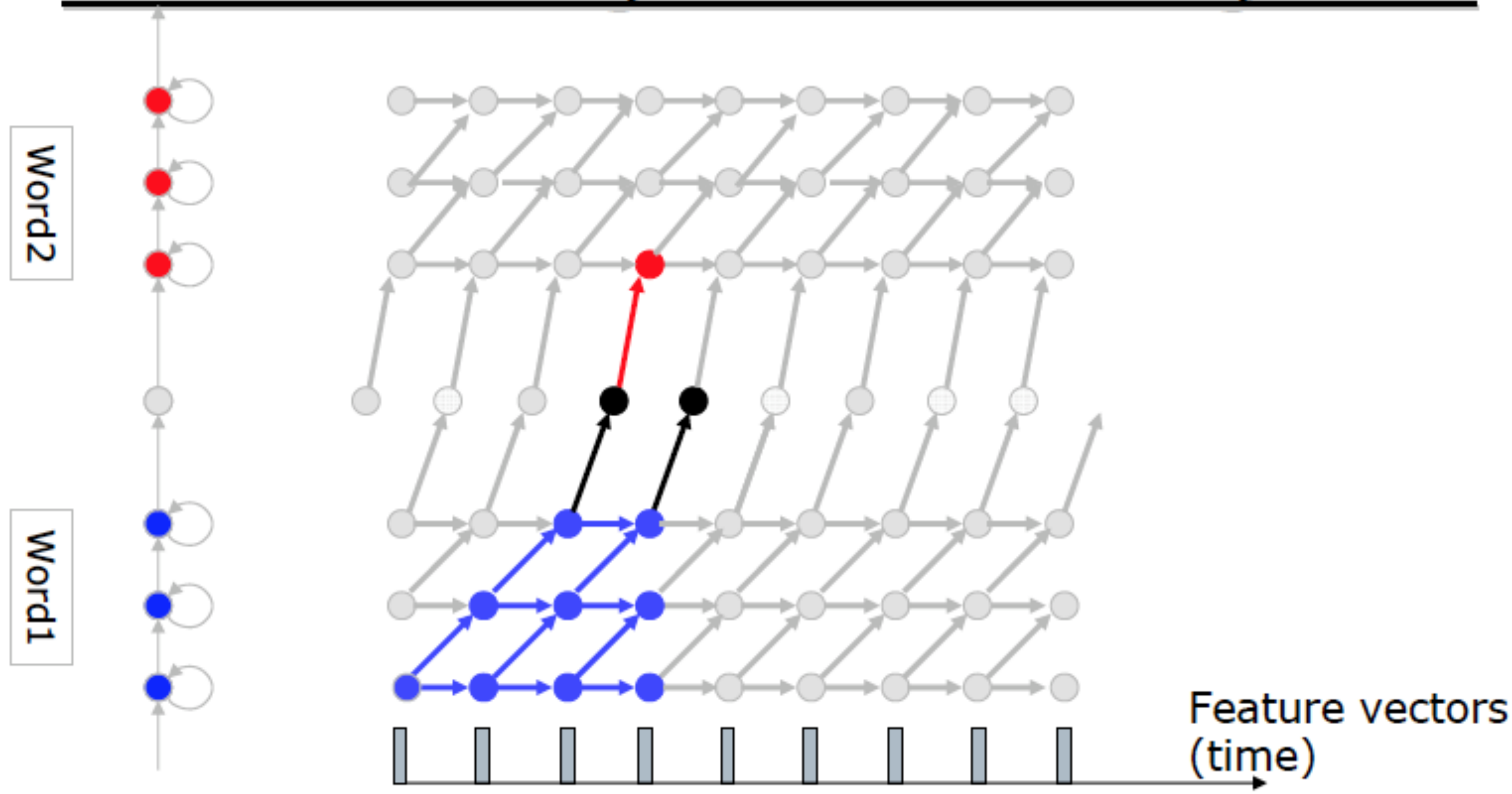


- At time 4, the first state of word2 gets a probability contribution from the non-emitting state

"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

Trellis with connected word models

Forward Through a non-emitting State

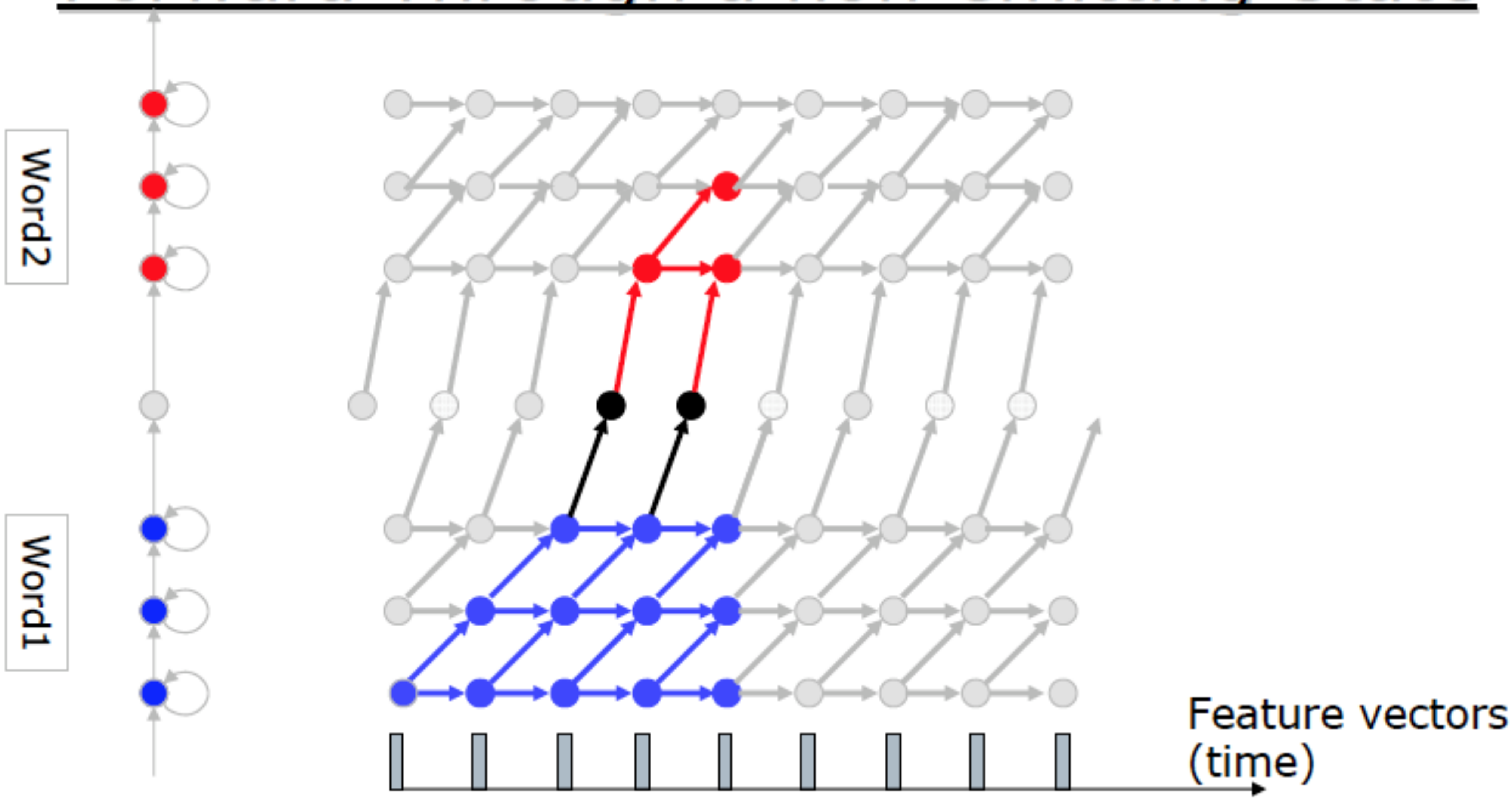


- Between time4 and time5 the non-emitting state may be visited

"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.

Trellis with connected word models

Forward Through a non-emitting State



- At time 5 (and thereafter) the first state of word 2 gets contributions both from an emitting state (itself at the previous instant) and the non-emitting state

"Dealing with Connected Words", Rita Singh and Bhiksha Raj, CMU.