CMPT 413: Computational Linguistics

HMM2: N-grams versus HMMs

Anoop Sarkar

http://www.cs.sfu.ca/~anoop

N-grams versus HMMs

- Bigram model: “British Left Waffles on Falkland Islands”
  - P(British) P(Left | British) P(Waffles | Left) P(on | Waffles) P(Falkland | on) P(Islands | Falkland)
- HMM: “British Left Waffles on Falkland Islands”
  - HMM for state sequence: (N, N, V, P, N, N)
  - P(N) P(British | N) P(N | N) P(Left | N) P(V | N) P(Waffles | V) P(P | V) P(on | P) P(N | P) P(Falkland | N) P(N | N) P(Islands | N)

N-grams versus HMMs

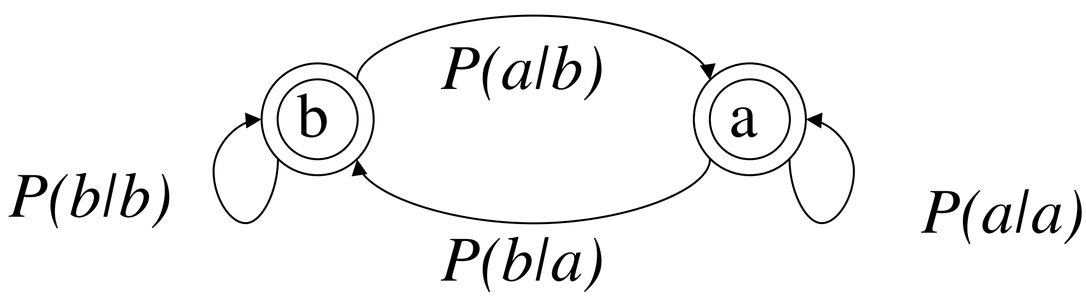
- An n-gram model is called a Markov model or Markov chain
- Hidden Markov Models add the notion of a "Hidden" state
- N-grams directly model the probability of a sequence of observations: $P(w_j \mid w_i)$
- HMMs use a more abstract state representation: $P(X_j \mid X_i)\ P(w_j \mid X_j)$

Markov Chains vs. HMMs

- For observation sequence *babaa*

  $i.e: o_1=b, o_2=a, \ldots, o_5=a$
- Compute *P(babaa)* using a bigram model

  $P(b)*P(a|b)*P(b|a)*P(a|b)*P(a|a)$
- Equivalent Markov chain:

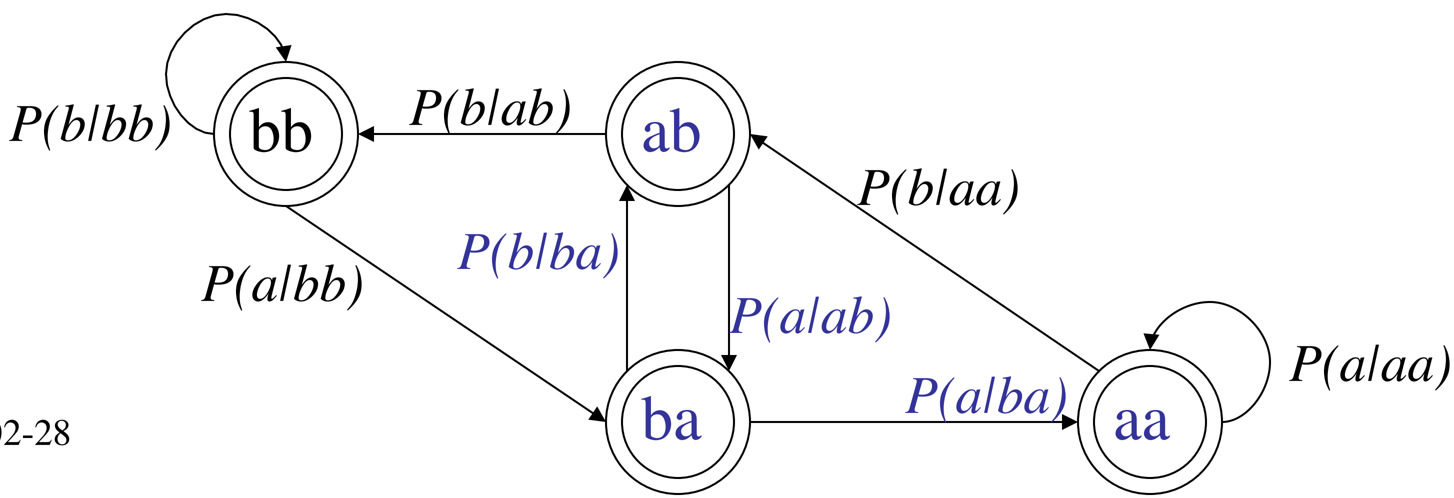


Markov Chains vs. HMMs

- For observation sequence *babaa*

  i.e: $o_1=b$, $o_2=a$, …, $o_5=a$
- Compute *P(babaa)* using a trigram model

  *P(ba)*P(b|ba)*P(a|ab)*P(a|ba)*
- Equivalent Markov chain:

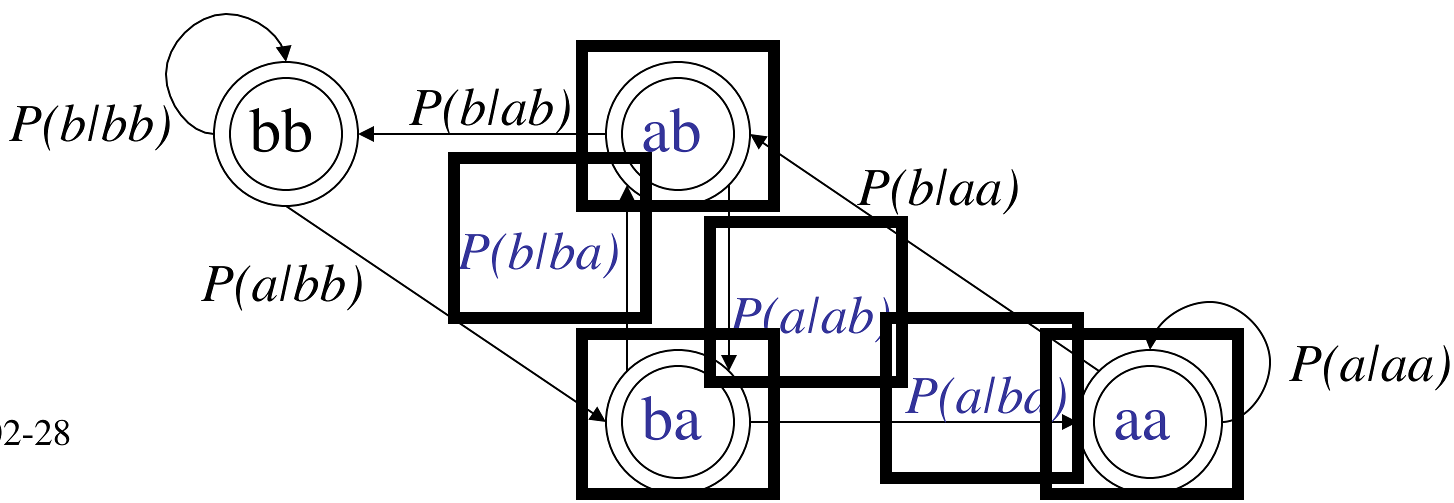


Markov Chains vs. HMMs

- For observation sequence *babaa*

  $i.e{:}\ o_1=b,\ o_2=a,\ \ldots,\ o_5=a$
- Compute *P(babaa)* using a trigram model

  $P(ba) * P(b|ba) * P(a|ab) * P(a|ba)$
- Equivalent Markov chain:



Markov Chains vs. HMMs

- Given an observation sequence
  $\mathbf{O}=(o_1, \ldots, o_t, \ldots, o_T)$
- An *n*th order Markov Chain or *n*-gram model computes the probability
  $P(o_1, \ldots, o_T)$
- An HMM computes the probability
  $P(X_1, \ldots, X_T, o_1, \ldots, o_T)$ where the state sequence is *hidden*

Properties of HMMs

- Markov assumption

$$P(X_t = s_i \mid \ldots, X_{t-1} = s_j)$$

- Stationary distribution

$$P(X_t = s_i \mid X_{t-1} = s_j) = P(X_{t+l} = s_i \mid X_{t+l-1} = s_j)$$

CMPT 413: Computational Linguistics

HMM2: N-grams versus HMMs

Anoop Sarkar

http://www.cs.sfu.ca/~anoop