Hidden Variable Structures for Training and Decoding with Graphical Models in ASR

Geoff Zweig

CLSP WS01 Talk

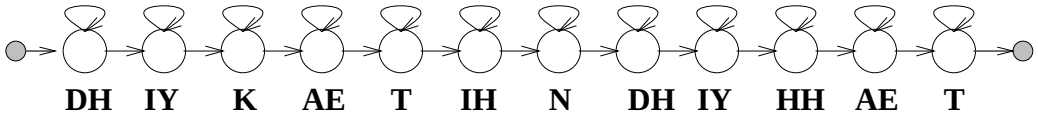
July 6, 2001

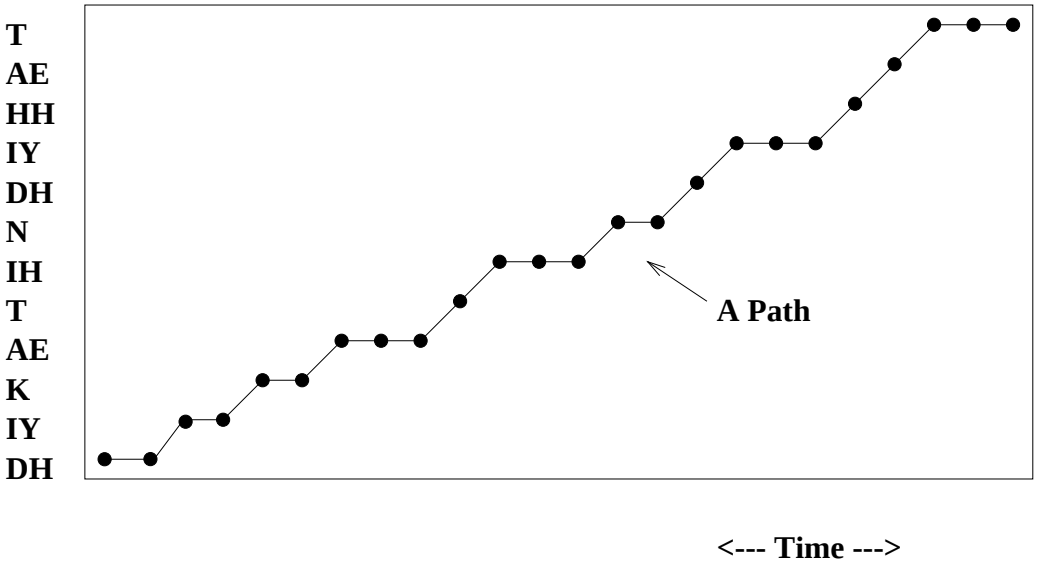
Outline

- HMMs: The Point of Departure
  - A graph for training.
  - A graph for decoding.
- The Key to Reasoning
  - Paths in an HMM
  - Variable assignments in a GM
- The Simplest Training Structure
- Decoding with a Bigram LM
- Training with Optional Silence
- Decoding with a Trigram LM
- Training with Lexical Variants
- Articulatory Networks

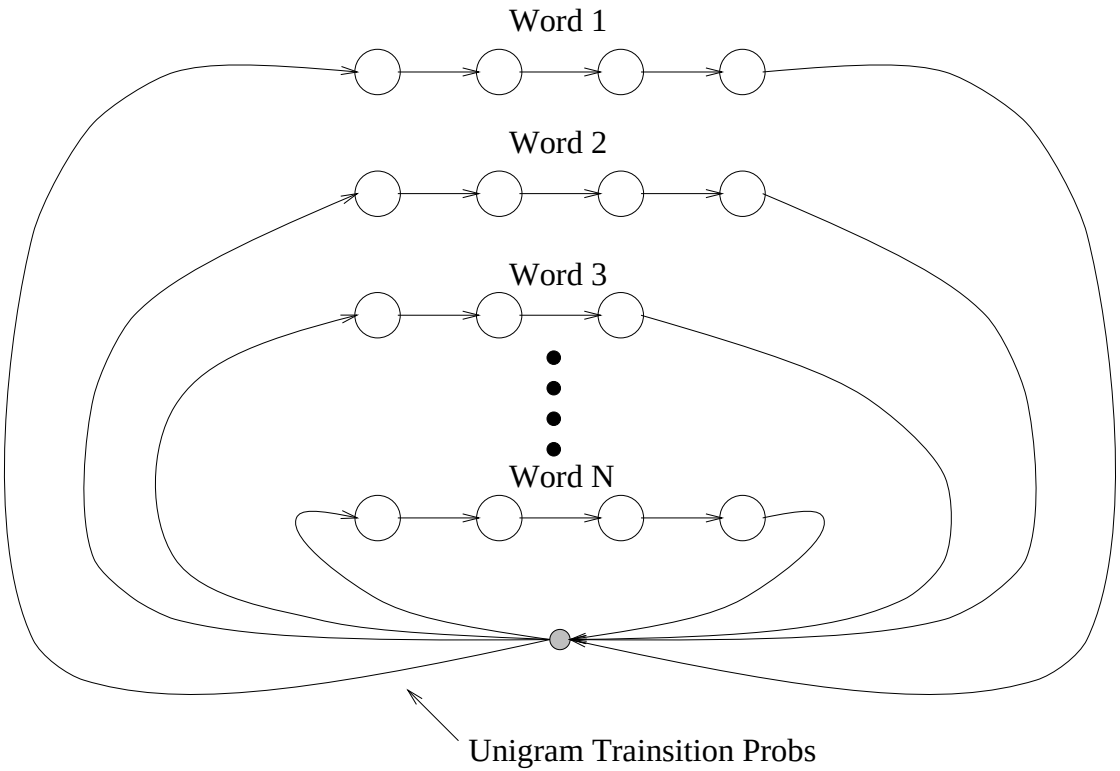
HMM Training

The cat in the hat.





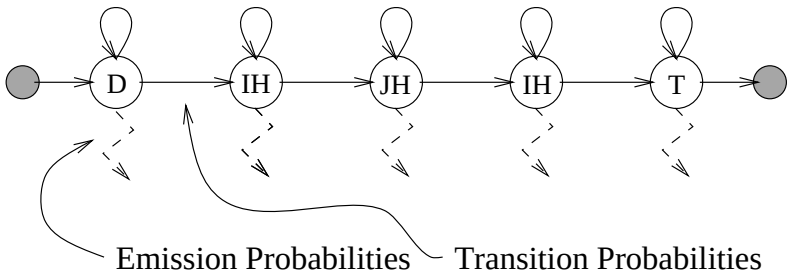
HMM Decoding



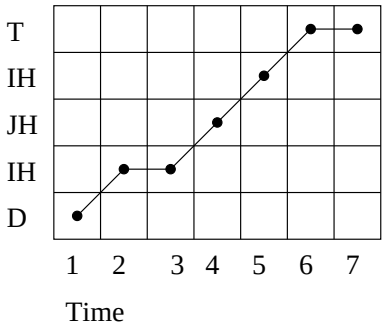
;

The Key To Reasoning: Paths & Assignments

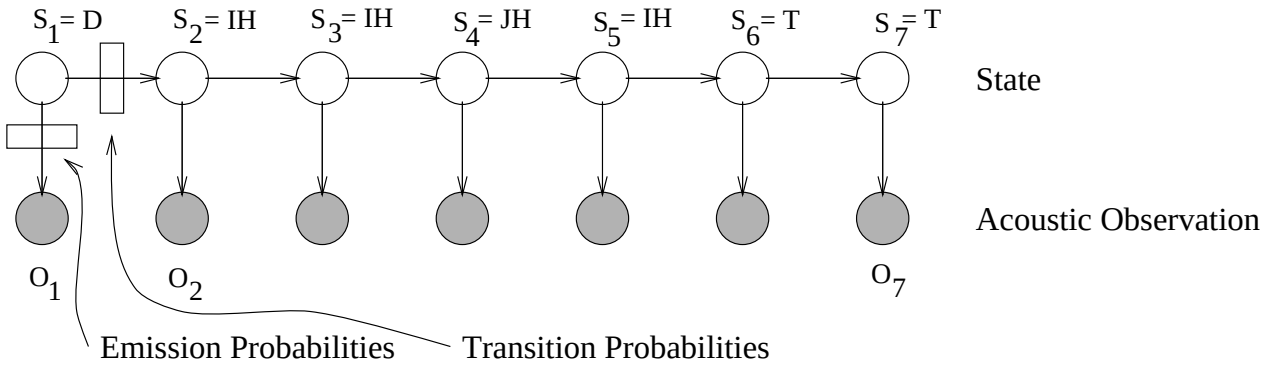
HMM



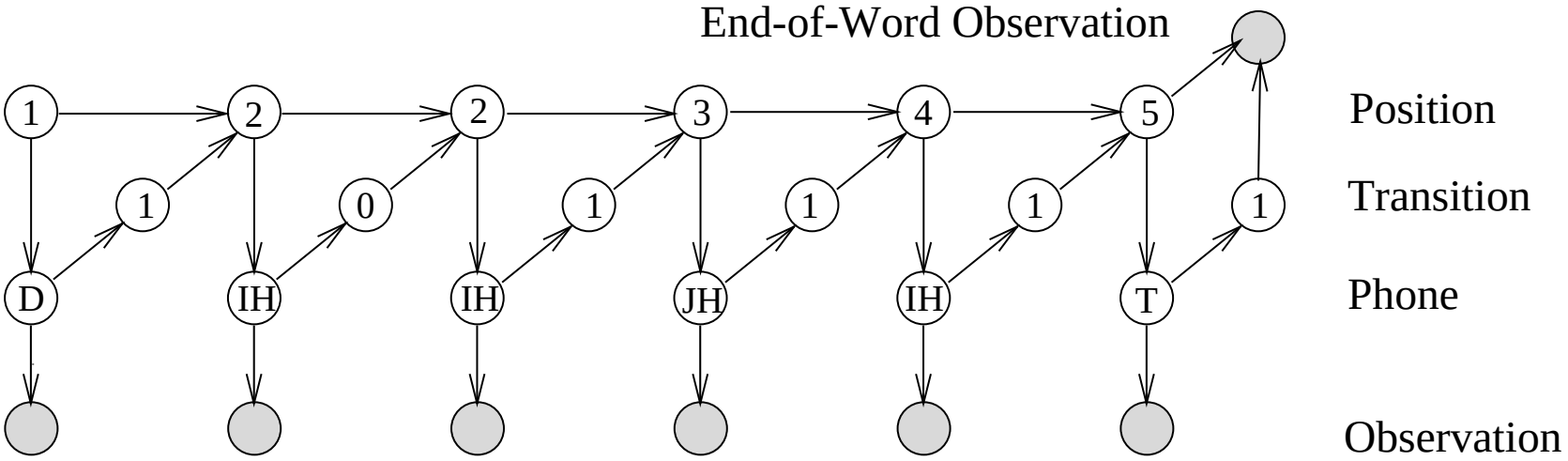
HMM Grid:
7-Frame Utterance



BN

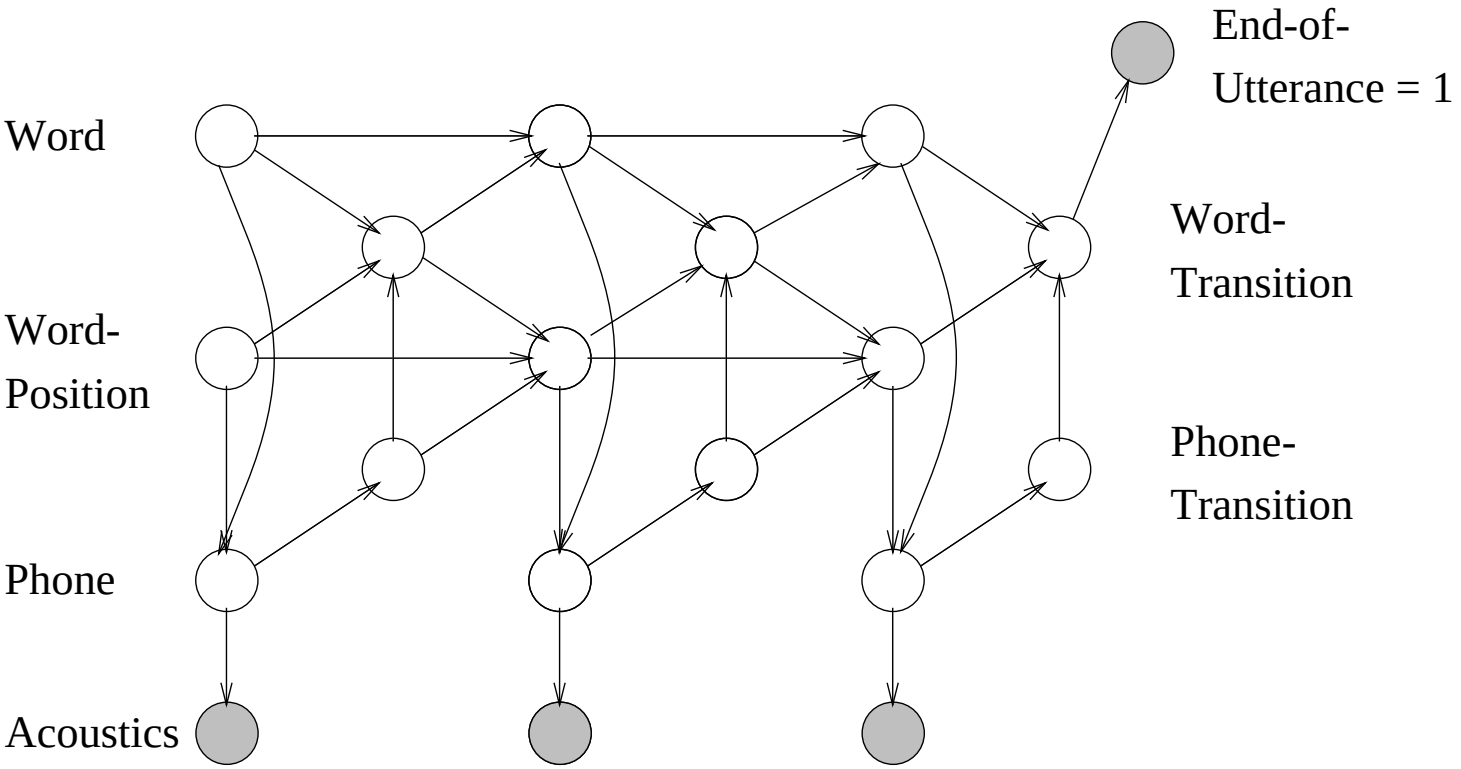


The Simplest Training Structure



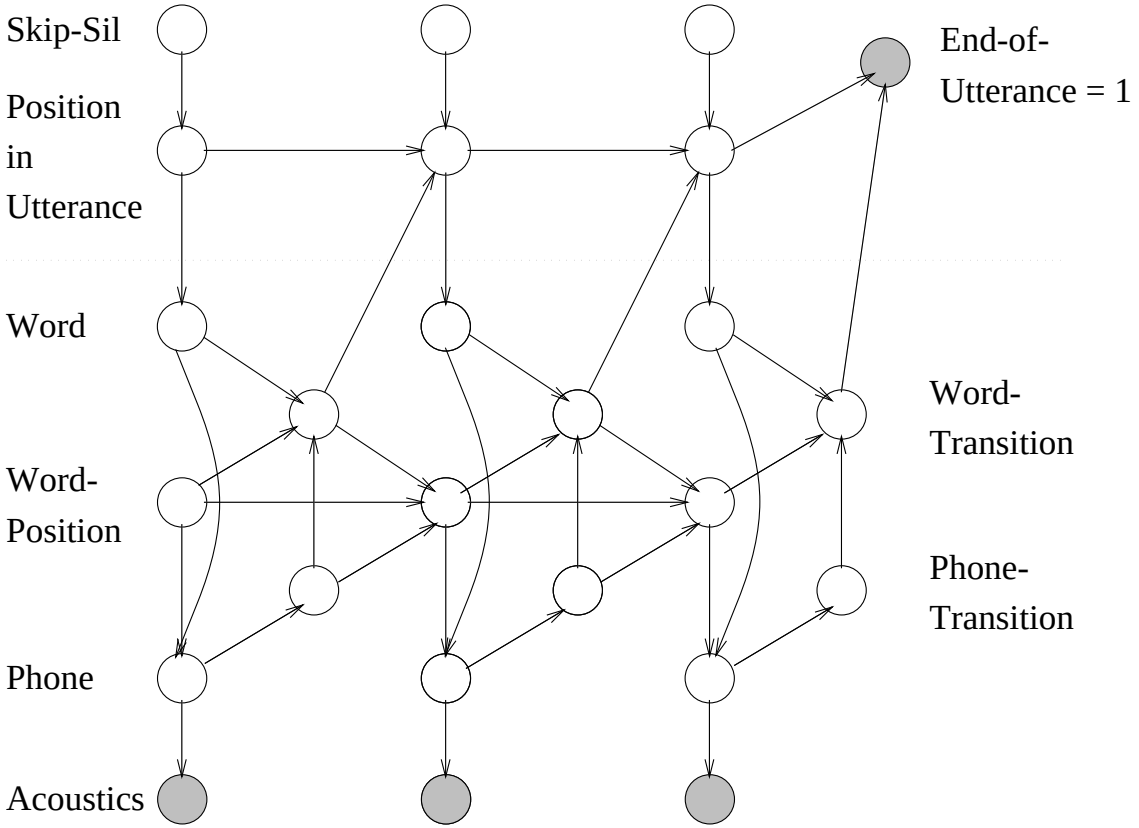
- Each BN assignment $\equiv$ path through HMM.
- Sum over BN assignments $\equiv$ sum over HMM paths.
- Max over BN assignments $\equiv$ max over HMM paths.

Decoding with A Bigram LM

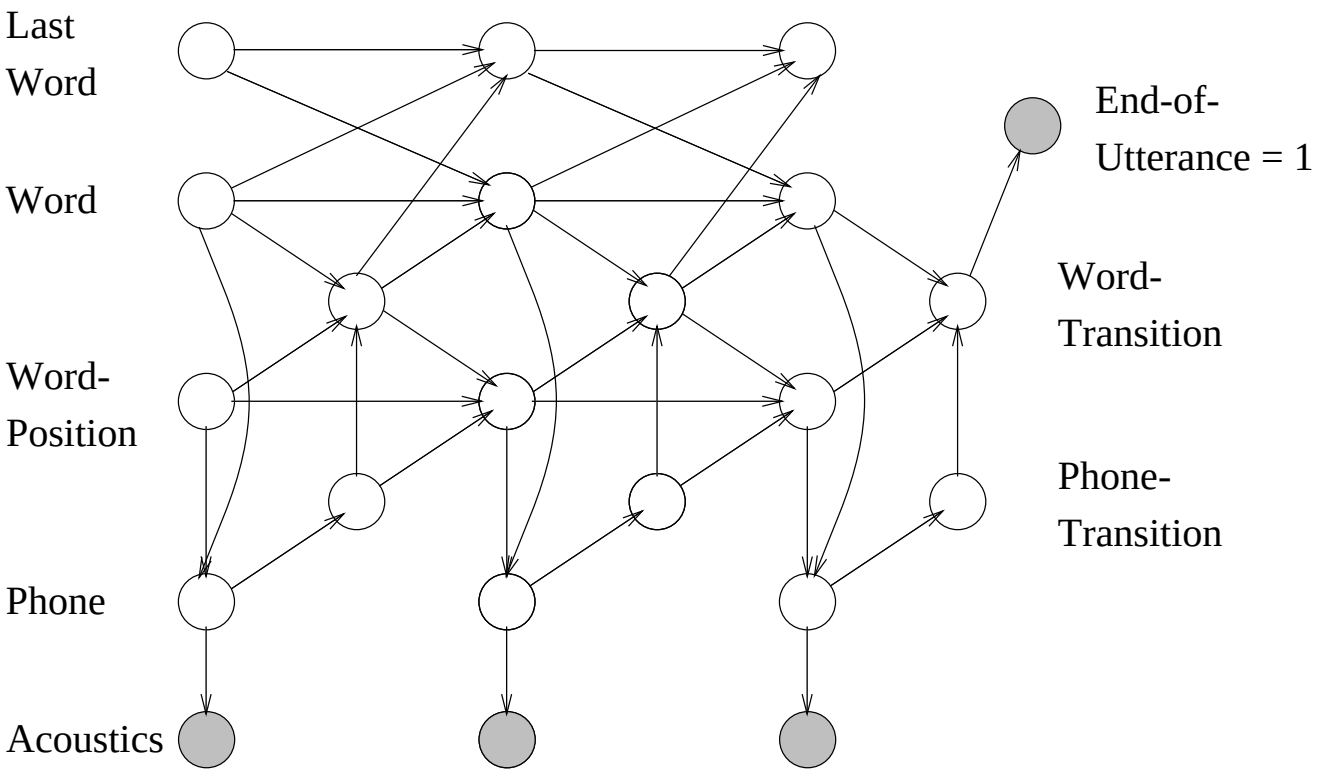


Training with Optional Silence

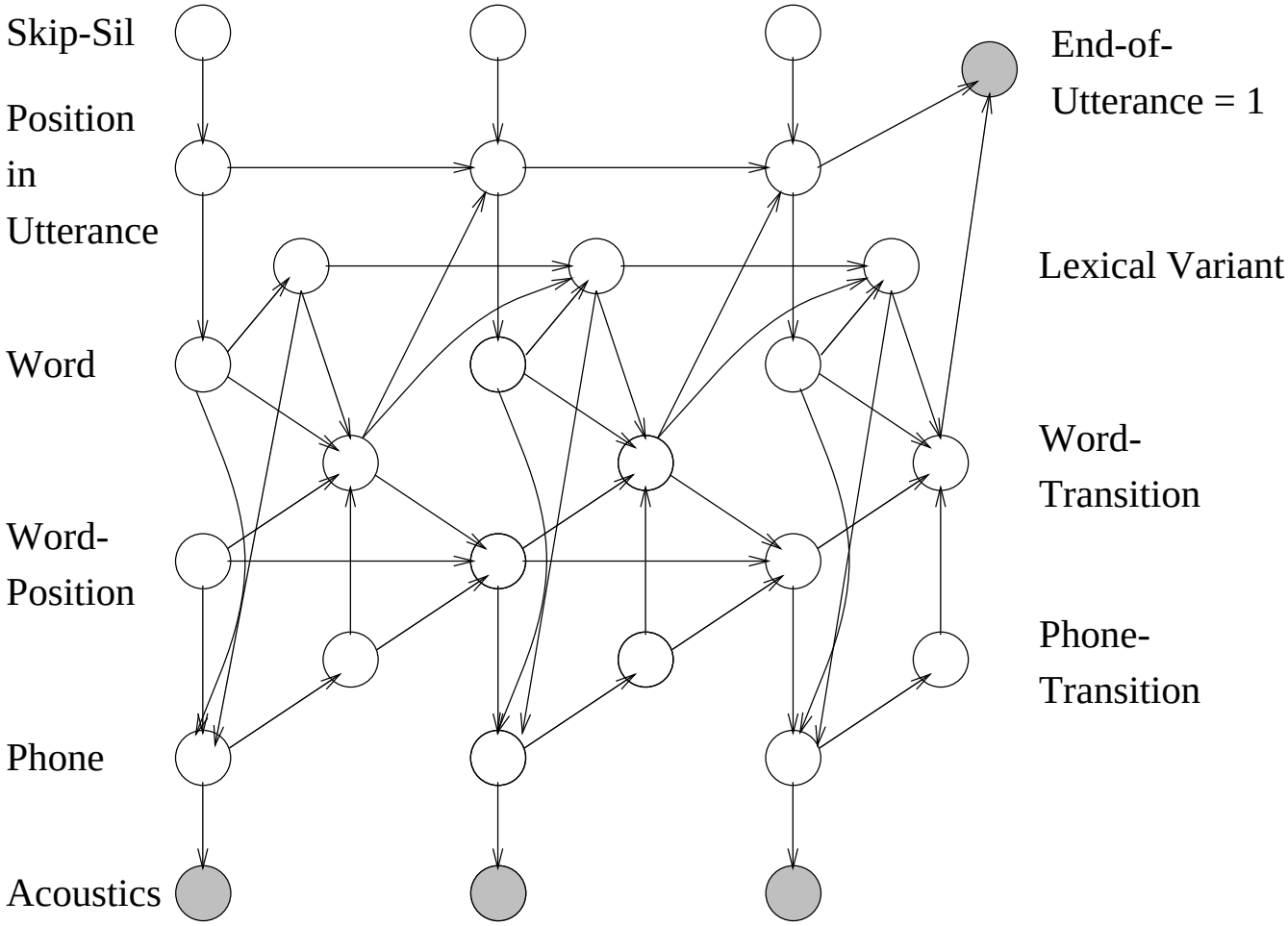
sil (0) The (1) sil (2) cat (3) sil (4) in (5) sil (6) the (7) sil (8) hat (9) sil (10)



Decoding with A Trigram LM



Training with Lexical Variants



Subtleties and the Key To Reasoning

An Example Subtlety:
- What happens on a word transition in the final silence?
- Stay put? Then paths with too many transitions contribute.

e.g. 1-state, 10-frames. Transitions *should* be:
0 0 0 0 0 0 0 0 0 1
But:
0 1 1 1 1 1 1 1 1 1
would be counted

Always reason in terms of enumerating out all the concrete instantiations. Add variables and set up CPTs as appropriate.

Articulatory Networks