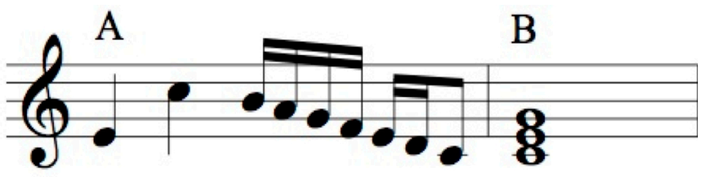
DEVELOPING A LANGUAGE MODEL AND REPRESENTATION OF SYMBOLIC MUSIC DATA FOR CLASSIFICATION AND SIMILARITY TASKS

Johanna Devaney, The Ohio State University

1. GOAL

- Develop a symbolic representation that
  - provide an estimate of the structural significance of each note
  - works for a range of musical textures
  - captures temporal relationships
  - facilitates the analysis of multiple levels of musical structure
  - is computationally tractable
- For example the representation should facilitate the analysis of both A and B as C major chords



2. PARALLELS WITH SPEECH REC.

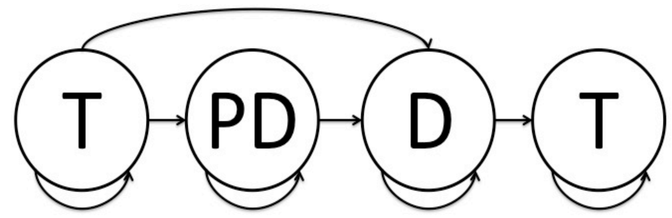
Speech recognition	↔	Higher-level music analysis
"Language model"	↔	Functional analysis
Phonemes	↔	Harmonic analysis
MFCCs and other representations	↔	Representation
Acoustic signal	↔	Musical surface

3. "LANGUAGE MODEL" FOR MUSIC

- A model of phrase-level function and its relationship to roman numeral labels can be used as a "language model"
- Phrases are musical statements built from the ordered presentation of three harmonic functions
  - tonic, pre-dominant, and dominant functions
- Phrases end with a cadence
  - remain on the dominant function for a half cadence
  - return to the tonic function for an authentic or a deceptive cadence

4. PILOT STUDY

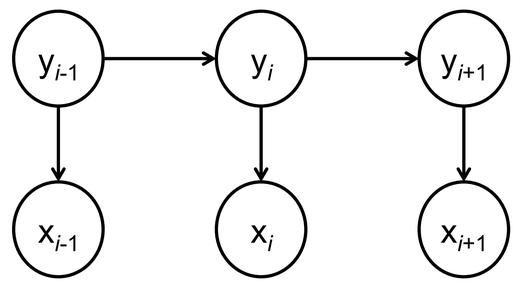
- **Purpose**: to evaluate the relationship between chord labels and functions
- **Method**: 3-state single-layer hidden-Markov model



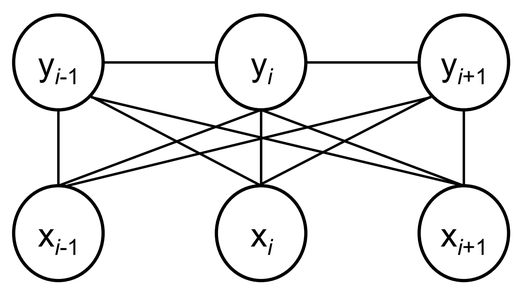
- **Dataset**: Labeled musical examples in Stephen Laitz's *The Complete Musician* textbook (2011)
  - E.g., Chords Labels: I V43 I6 IV V I
    Functions: T - - - - - - - - PD D T
- **Evaluation:** 80/20 split of textbook data (128 items)
  - 94.3% overall accuracy
    - tonic 93% pre-dominant 93%, dominant 89%

5. CURRENT WORK

- **GOAL:** Given the musical surface, jointly infer the chords and the functions
- **APPROACH:** Use the phrase model to constrain the space and sequence of possible chords
- Conditional random fields (CRFs) are a better option than HMMs for this task
  - HMM observations are isolated, only containing information about its own label



  - CRFs encode information about temporal context, including features from previous time frames and feature functions about the transitions between states



- Features capture pitch and metrical information
  - E.g., bass note, chord root, duration, metrical strength
- Feature functions capture information about relationships between chords and functions
  - E.g., a V43 chord preceded by a I chord is more likely to have a tonic label rather than a dominant label