Theme And Variation Encodings with Roman Numerals (TAVERN): A New Data Set for Symbolic Music Analysis

Johanna Devaney, Claire Arthur, Nathaniel Condit-Schultz, and Kirsten Nisula
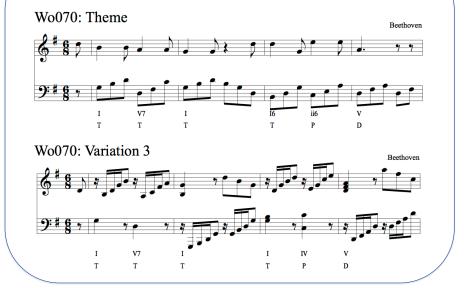School of Music, The Ohio State University

TAVERN

- 27 complete sets of theme and variations for piano composed between 1765 and 1810 by Mozart and Beethoven
- Includes harmonic and phrase-level function data
- Designed to facilitating the development of algorithms of automatic symbolic chord recognition and symbolic similarity, through a deeper understanding of the impact of **texture** on both of these tasks

Encoding Format

- Individual phrases are available as
  - 'analysis' files
  - 'score' files
  - joined 'analysis+score' files
- Score data encoded in **kern
- Chord labels encoded in **harm
- Function labels (T, P, D) encoded in newly defined **func format
- Analysis files also include a **comments spine where the analysts provided information about other possible analyses (where appropriate)

```
!!!COM: Beethoven
!!!OTL: 6 Variationen uber "Nel cor
!!!piu non mi sento" von G. Paisiello
!!!Variation: Theme a
**func  **harm  **kern  **kern
*M6/8   *M6/8   *M6/8   *M6/8
*G:     *G:     *G:     *G:
*tb8    *       *       *
8D      8V      8r      8dd
=       =       =       =
4.T     4.I     8GL     4b
.       .       8B      .
.       .       8dJ     8b
4.T     4.V7    8DL     4a
.       .       8F#     .
.       .       8cJ     8a
=       =       =2      =2
2.T     2.I     8GL     4g
.       .       8B      .
.       .       8dJ     8g
.       .       8BL     4r
.       .       8G      .
.       .       8DJ     8dd
.       .       =3      =3
4.T     4.I6    8BBL    4dd
.       .       8D      .
.       .       8GJ     8g
4.P     4.ii6   8CL     4ee
.       .       8E      .
.       .       8AJ     8ee
=       =       =4      =4
4.D     4.V/V   8DL     4.a
.       .       8F#     .
.       .       8AJ     .
4D      4V      8DL     8r
.       .       8F#     8r
*_      *_      *_      *_
```

Theme and Variation Example

Wo070: Theme — Beethoven

Wo070: Variation 3 — Beethoven

Phrase Model

- **Tonic** – at the beginning of a phrase it establishes the tonal center and at the end of a phrase signals the return of the tonal center
- **Pre-dominant** – prepares for the arrival of the dominant function
- **Dominant** – sets up an opposition to tonic, creating tension that is ultimately resolved by a return to tonic