Segmentation and grouping structures in jazz chord sequences: An information-theoretic approach

HEDGES, TW; Wiggins, GA

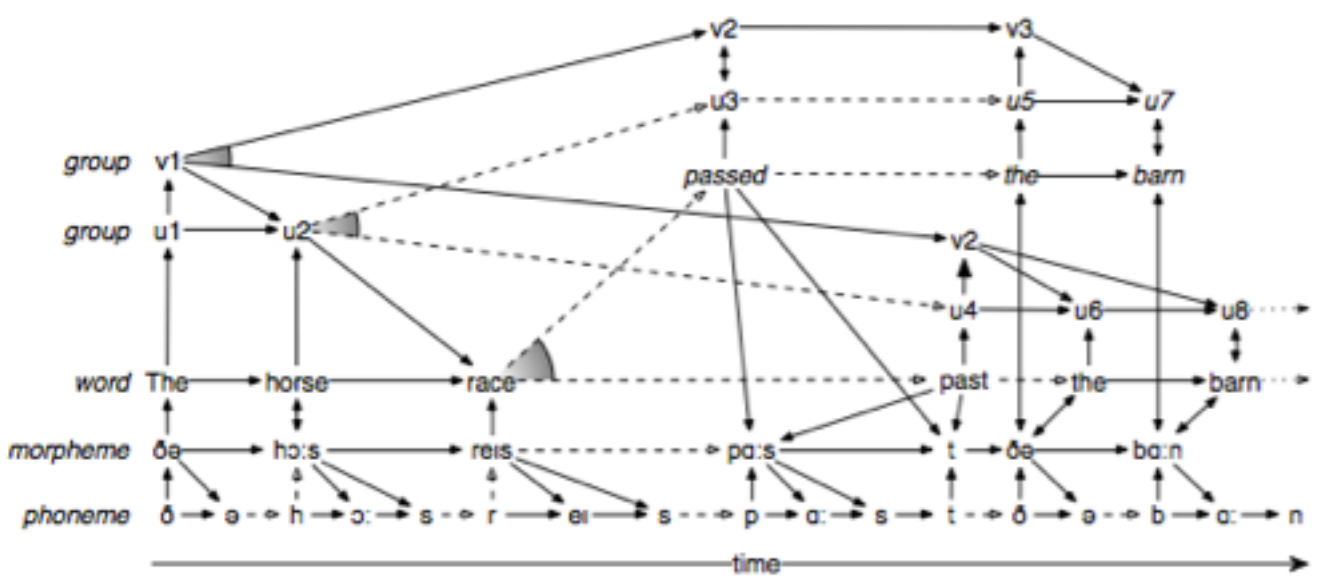
Segmentation and Grouping Structures in Jazz Chord Sequences: An Information-theoretic Approach.

Thomas W. Hedges
Queen Mary University of London
t.w.hedges@qmul.ac.uk

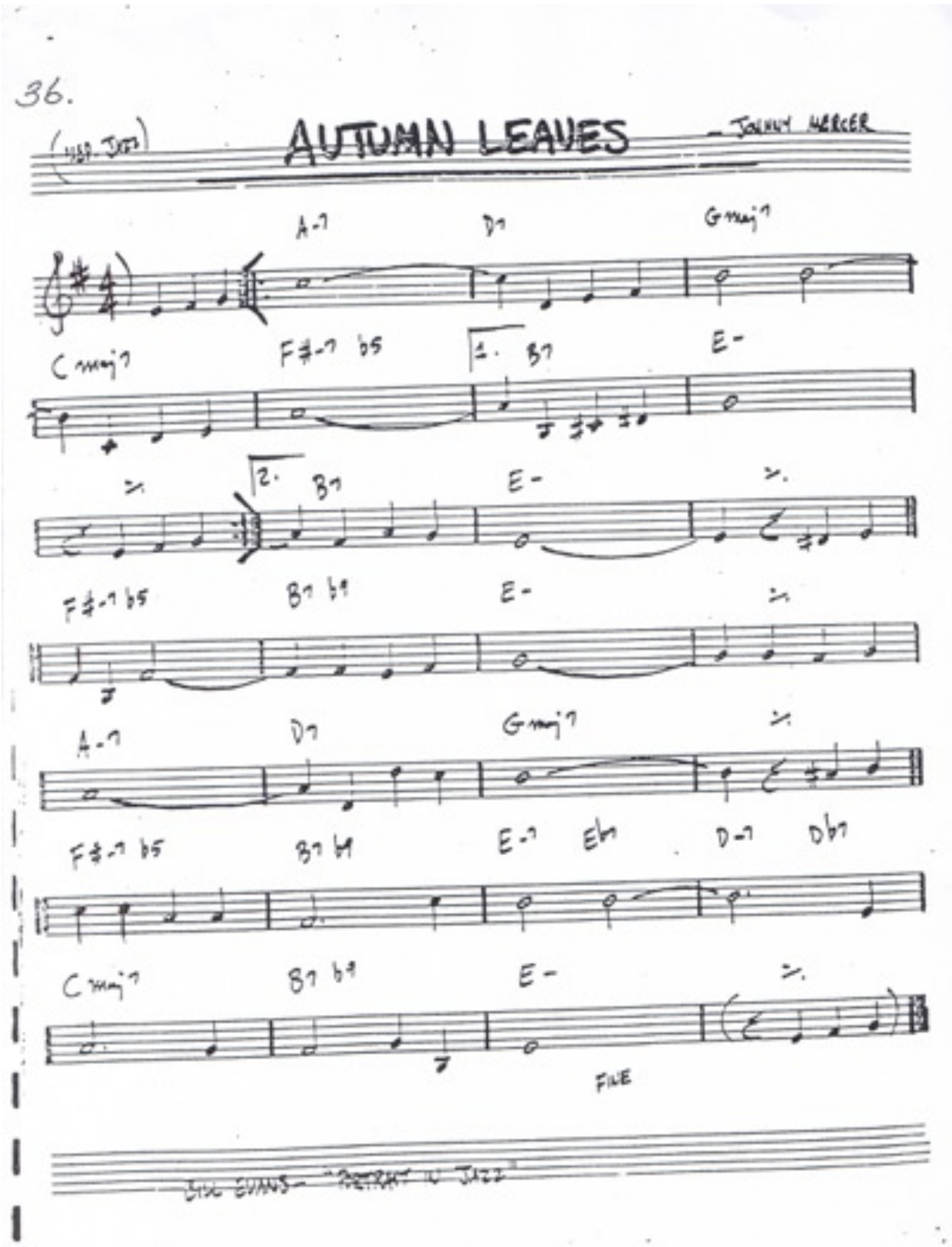
Geraint A. Wiggins
Queen Mary University of London
geraint.wiggins@qmul.ac.uk

Introduction

- How to model higher order / hierarchical structure with bottom-up, statistically driven models?
- Forth & Wiggins (2015) present IDyOT (Information Dynamics of Thinking), a cognitive architecture which expatiates IDyOM (Pearce 2005) to account for many aspects of human behaviour across multiple domains (language and music).
- The current research focusses on tonal harmony, in particular jazz.
- Segmentation is the first stage in these bottom-up models.



Jazz Leadsheets



- The entry point of the model is at the chord symbol level.
- 348 leadsheets (15,197 events) compiled by Pachet et al. (2013), taken from the Real Book vol. 1.
- Typical sequence learnt:

  Am7, D7, DM, CM, F#halfdim7, B7, Em

IDyOM: Statistical learning and modelling of the musical surface

- Information Dynamics Of Music (Pearce 2005)
- An unsupervised probabilistic model using variable order Markov models (PPM* - Cleary & Teahan 1997), interpolated smoothing (Cleary & Witten 1984, Moffat 1990) and multiple viewpoints weighted by entropy (Conklin & Witten 1995) to model expectation.

	BM^7	D^7	GM^7	$B\flat^7$	$E\flat M^7$
Root	11	2	7	10	3
ChordType	M	7	M	7	M
PosInBar	0	2	0	2	0
RootInt	$\perp$	3	5	3	5
ici	$\perp$	2	2	2	2
RootIntFiP	0	3	8	11	4
RootInt⊖FiB	$\perp$	$\perp$	8	$\perp$	8

Information theoretic segmentation

- Perceived segment boundaries before difficult to predict events (Pearce et al. 2010, Wiggins 2012, Griffiths et al. submitted).
- Difficulty to predict modelled by unexpectedness, defined by information content:

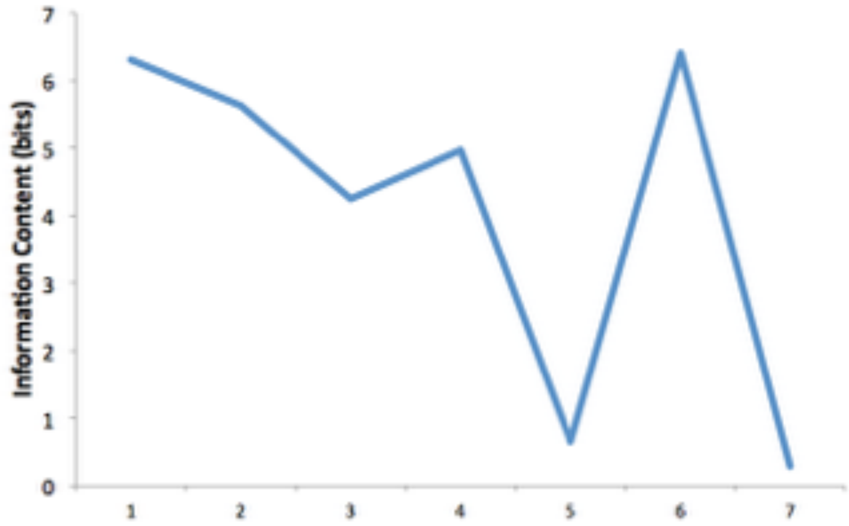
$$h\left(e_i|e_1^{i-1}\right) = -\log_2\ p\left(e_i|e_1^{i-1}\right)$$

- Place segment when before large rise in information content, when ratio between *h* of two adjacent events exceeds a threshold, *d*.

$$\frac{h\left(e_i|e_1^{i-1}\right)}{h\left(e_{i-1}|e_1^{i-2}\right)} > d$$

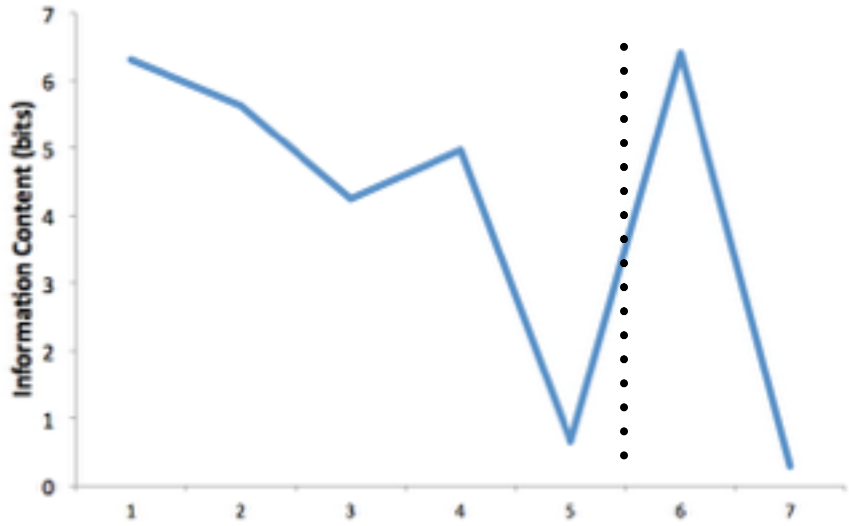
Information theoretic segmentation

- Information content profiles calculated with 10-fold cross validation.
- Viewpoint selected with forward stepwise selection algorithm.
- Viewpoints: ROOTINT⊗ICI, CHORDTYPE⊗ROOTINTTHRBAR, CHORDTYPE⊗ROOTINTFIP, POSINBAR⊗ROOTINTFIP, CHORDTYPE⊗POSINBAR, CHORDTYPE⊗ROOTINT, ROOT⊗CHORDTYPE



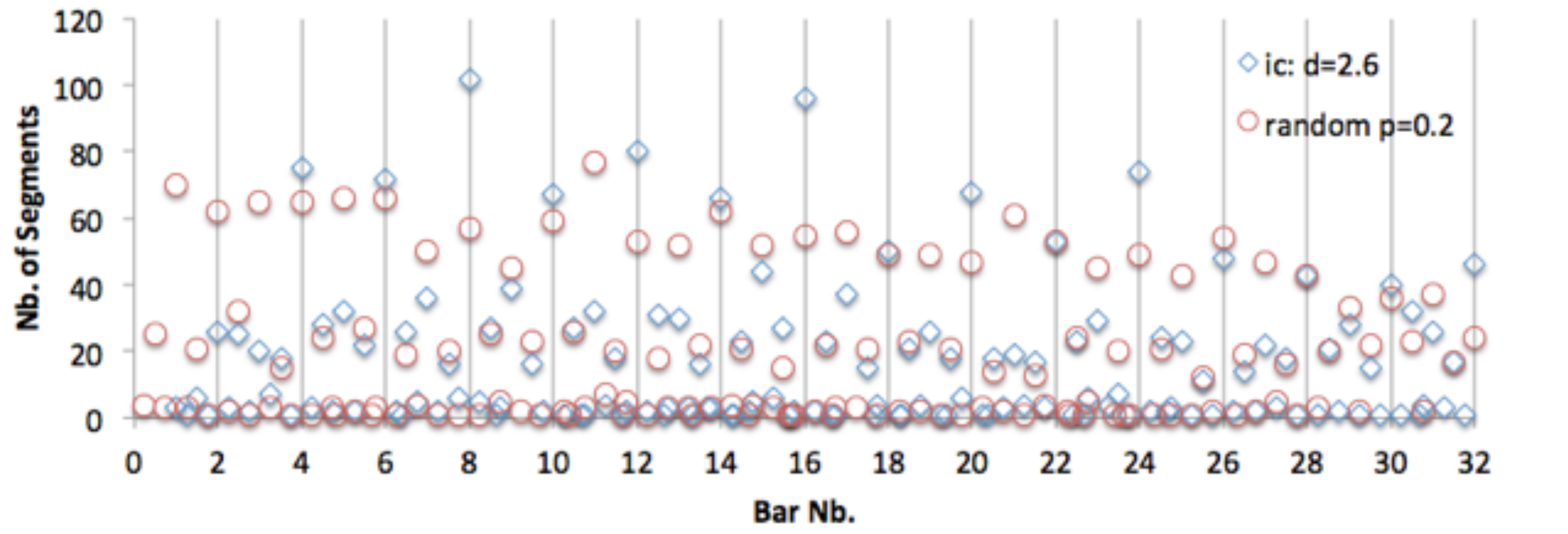
Information theoretic segmentation

- Information content profiles calculated with 10-fold cross validation.
- Viewpoint selected with forward stepwise selection algorithm.
- Viewpoints: ROOTINT⊗ICI, CHORDTYPE⊗ROOTINTTHRBAR, CHORDTYPE⊗ROOTINTFIP, POSINBAR⊗ROOTINTFIP, CHORDTYPE⊗POSINBAR, CHORDTYPE⊗ROOTINT, ROOT⊗CHORDTYPE

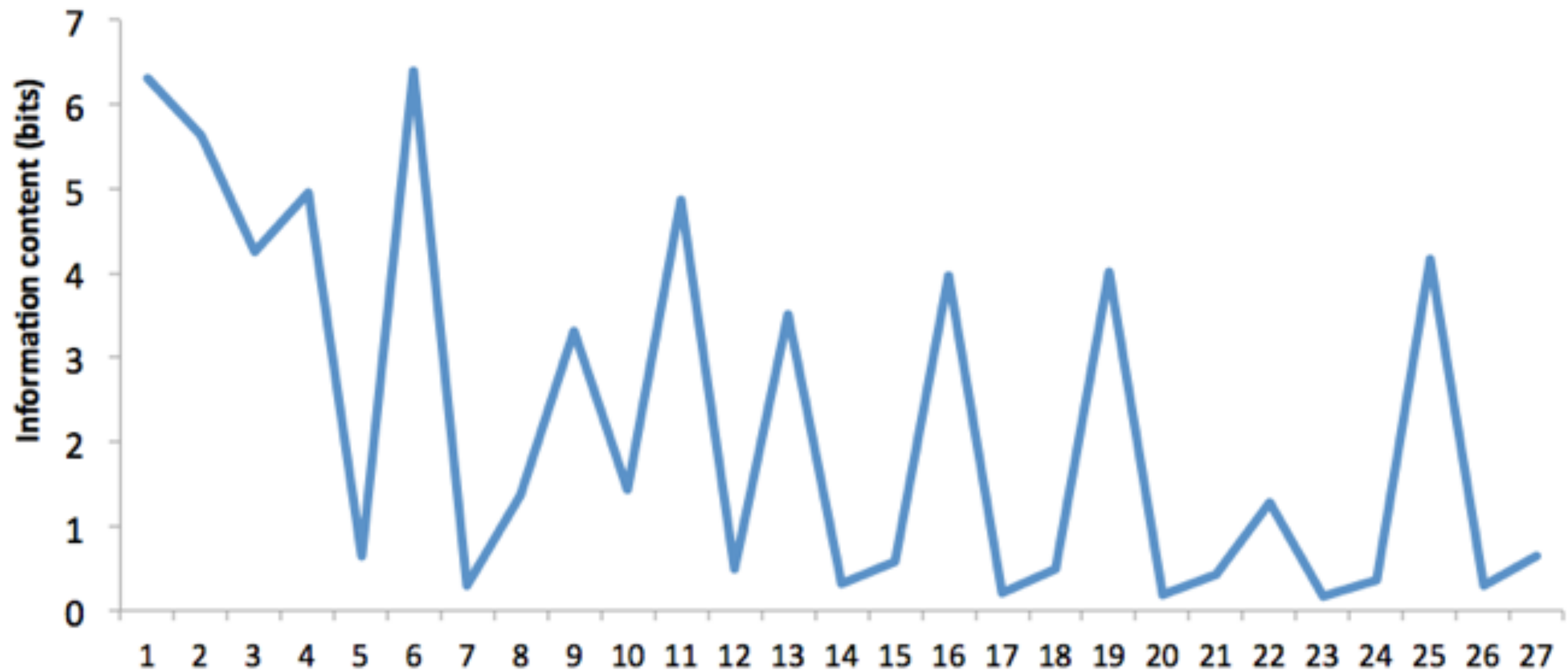


Phrase Structure

- No ground truth.
- Harmonic segmentation approximately reflect phrase structure.
- 4-bar phrases can be found segmenting with *d* = 2.6
    - κ: .24, accuracy: .76
- Random segmenter segments with *p* = .2
    - κ: .09, accuracy: .70



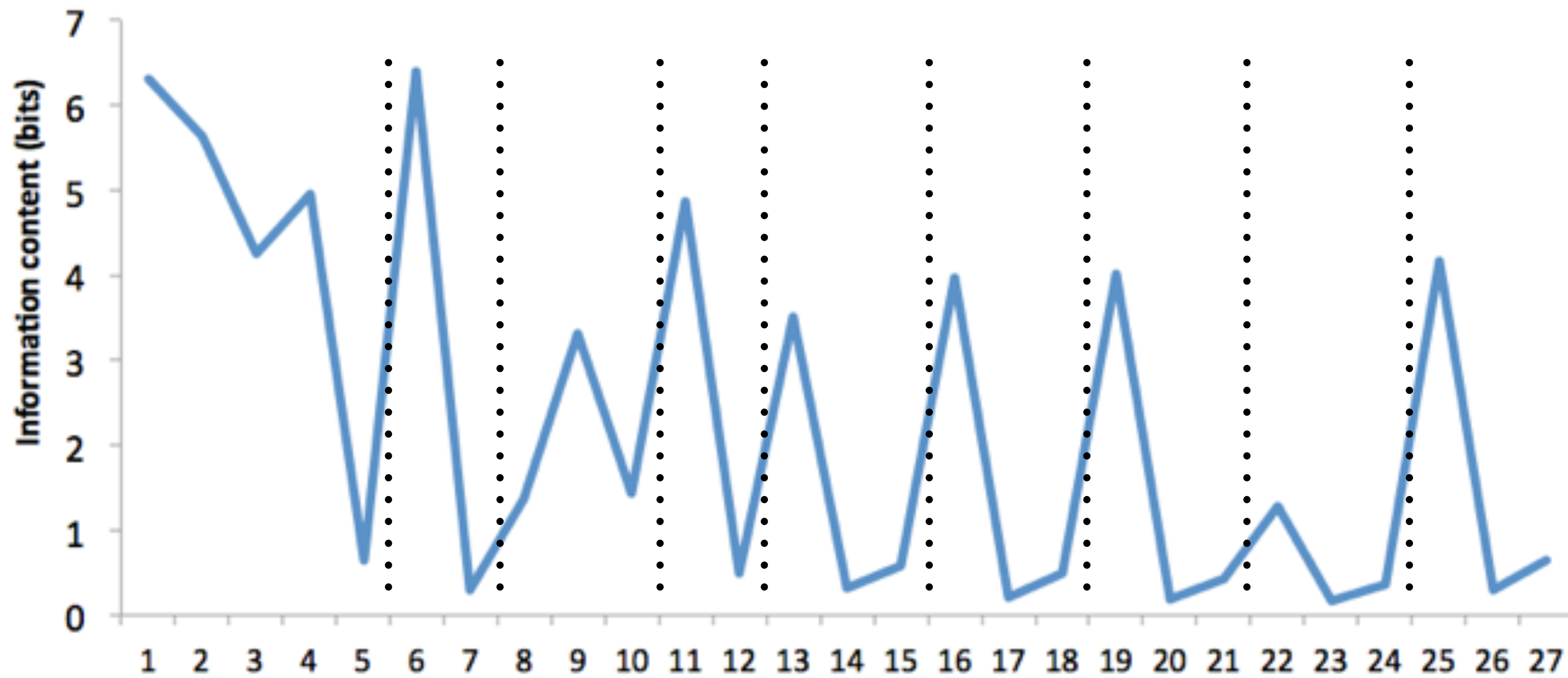
Giant Steps - John Coltrane



Threshold
$d = 2.6$

B	D7	G	Bb7	Eb		Am7	D7
G	Bb7	Eb	F#7	B		Fm7	Bb7
Eb		Am7	D7	G		C#m7	F#7
B		Fm7	Bb7	Eb		C#m7	F#7
B							

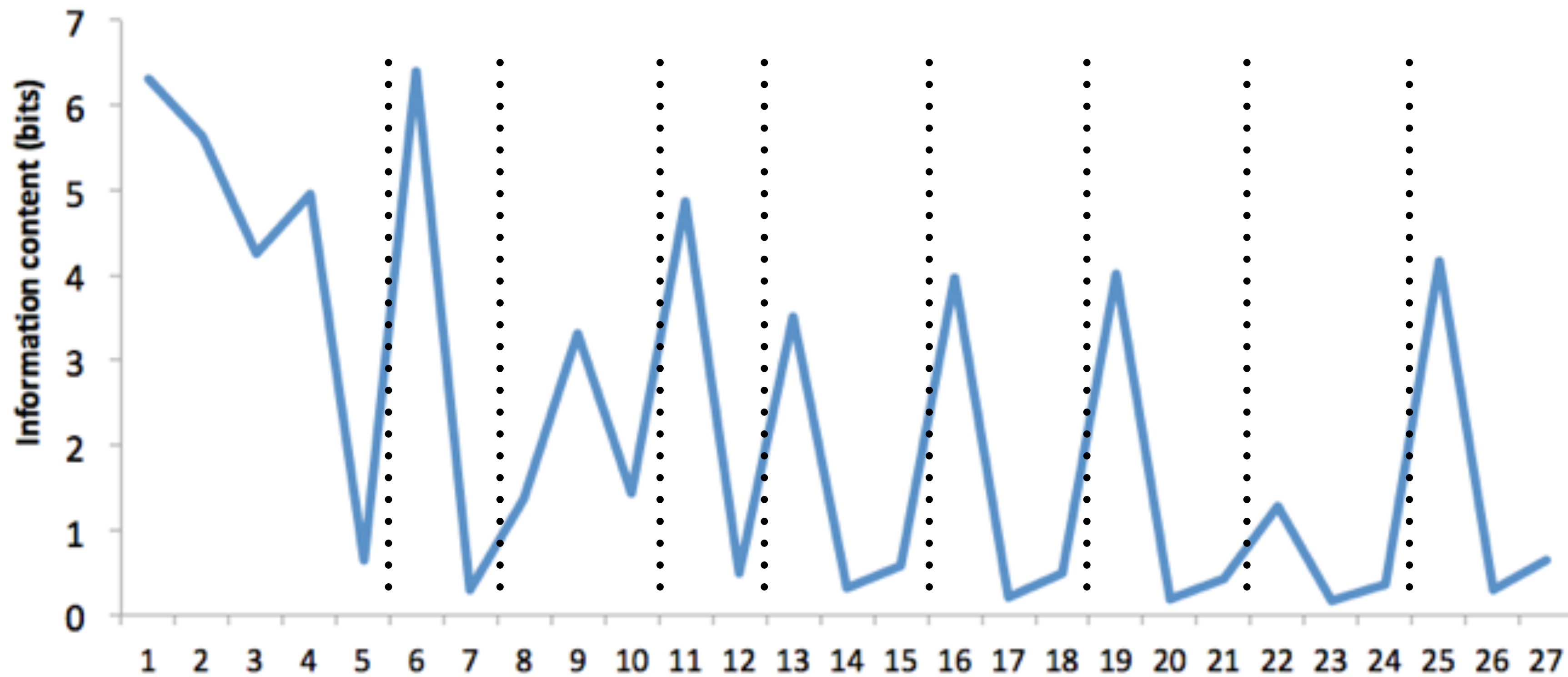
Giant Steps - John Coltrane



Threshold
$d = 2.6$

B D7	G Bb7	Eb	Am7 D7
G Bb7	Eb F#7	B	Fm7 Bb7
Eb	Am7 D7	G	C#m7 F#7
B	Fm7 Bb7	Eb	C#m7 F#7
B			

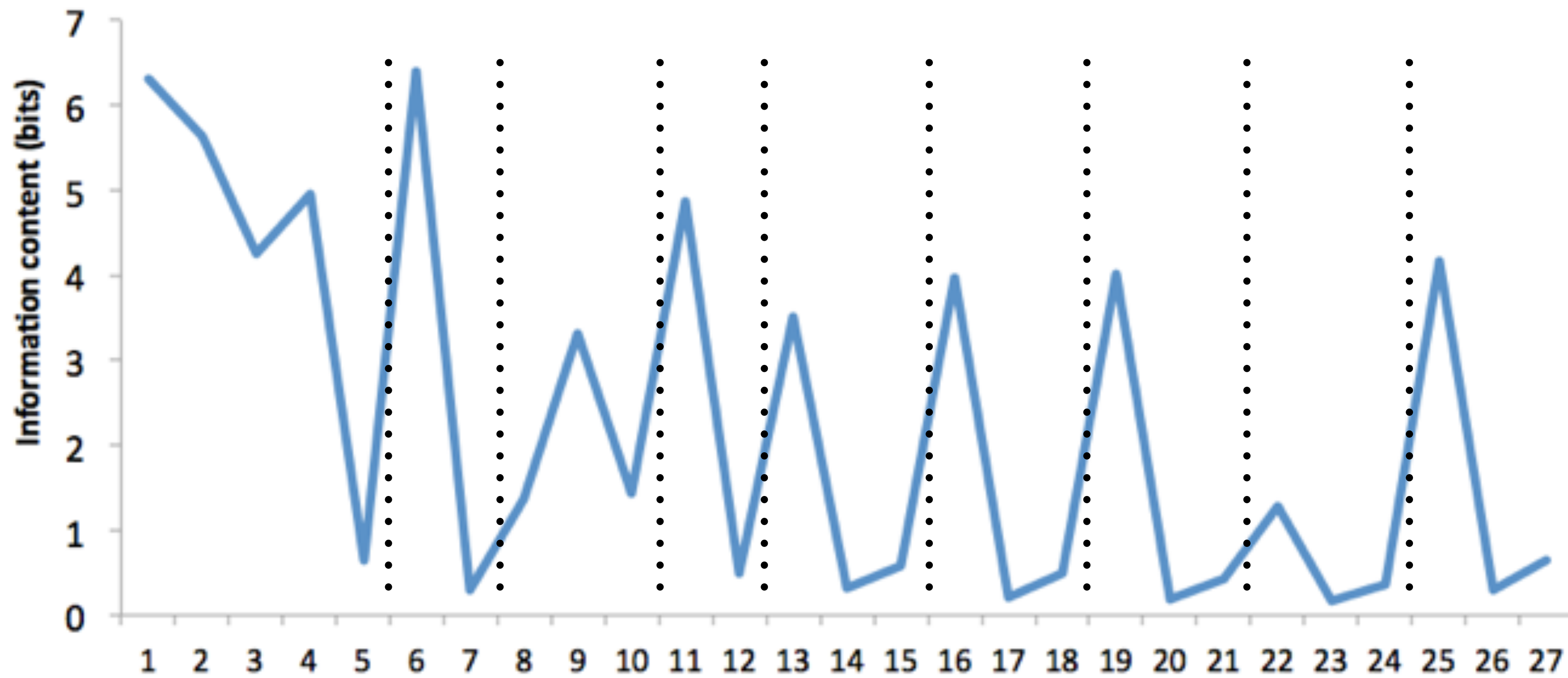
Giant Steps - John Coltrane



Threshold
d = 2.6

B D7	G Bb7	Eb	Am7 D7
G Bb7	Eb F#7	B	Fm7 Bb7
Eb	Am7 D7	G	C#m7 F#7
B	Fm7 Bb7	Eb	C#m7 F#7
B			

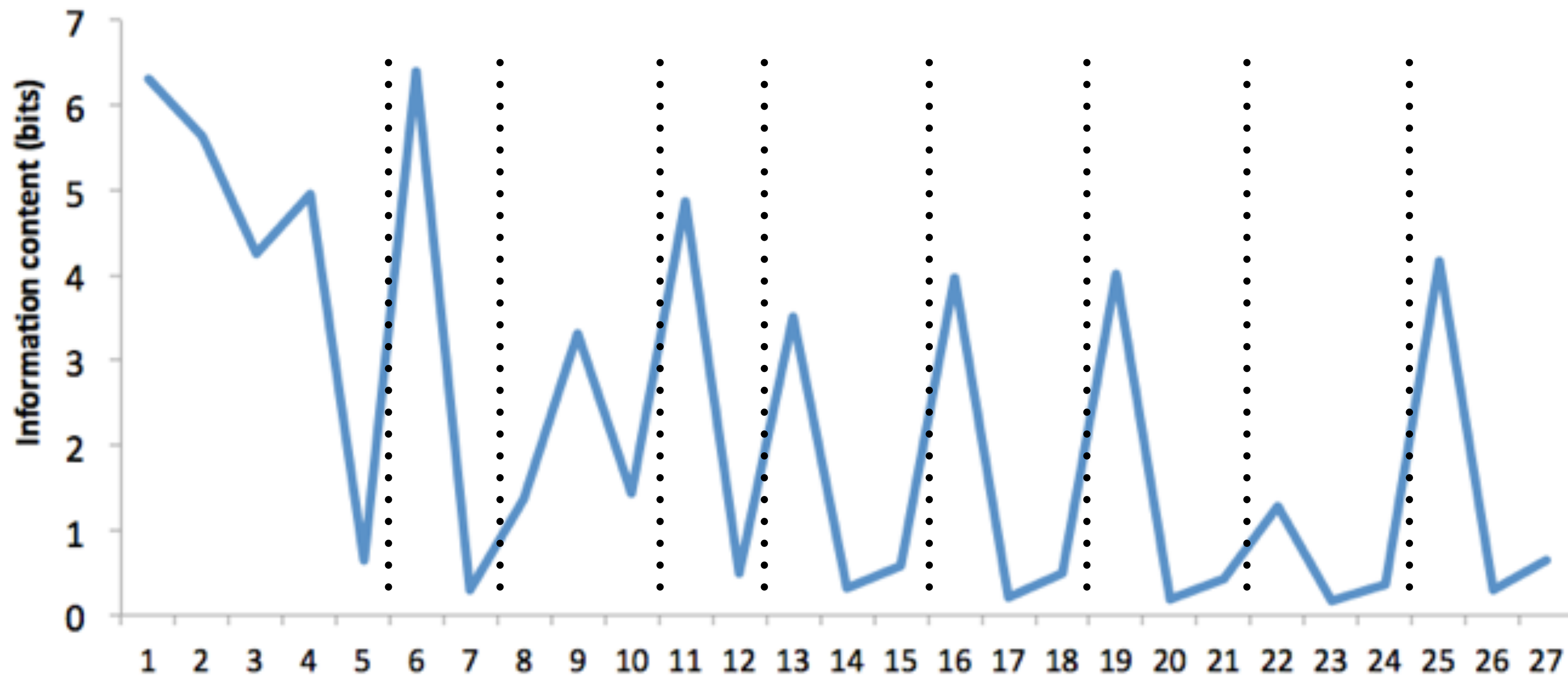
Giant Steps - John Coltrane



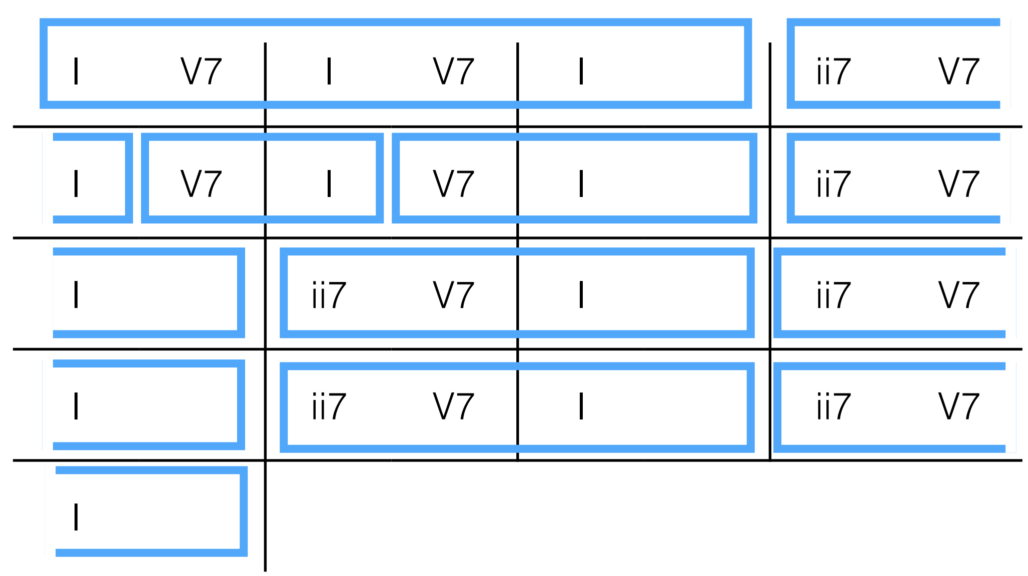
Threshold
$d = 2.6$

B D7	G Bb7	Eb	Am7 D7
G Bb7	Eb F#7	B	Fm7 Bb7
Eb	Am7 D7	G	C#m7 F#7
B	Fm7 Bb7	Eb	C#m7 F#7
B			

Giant Steps - John Coltrane

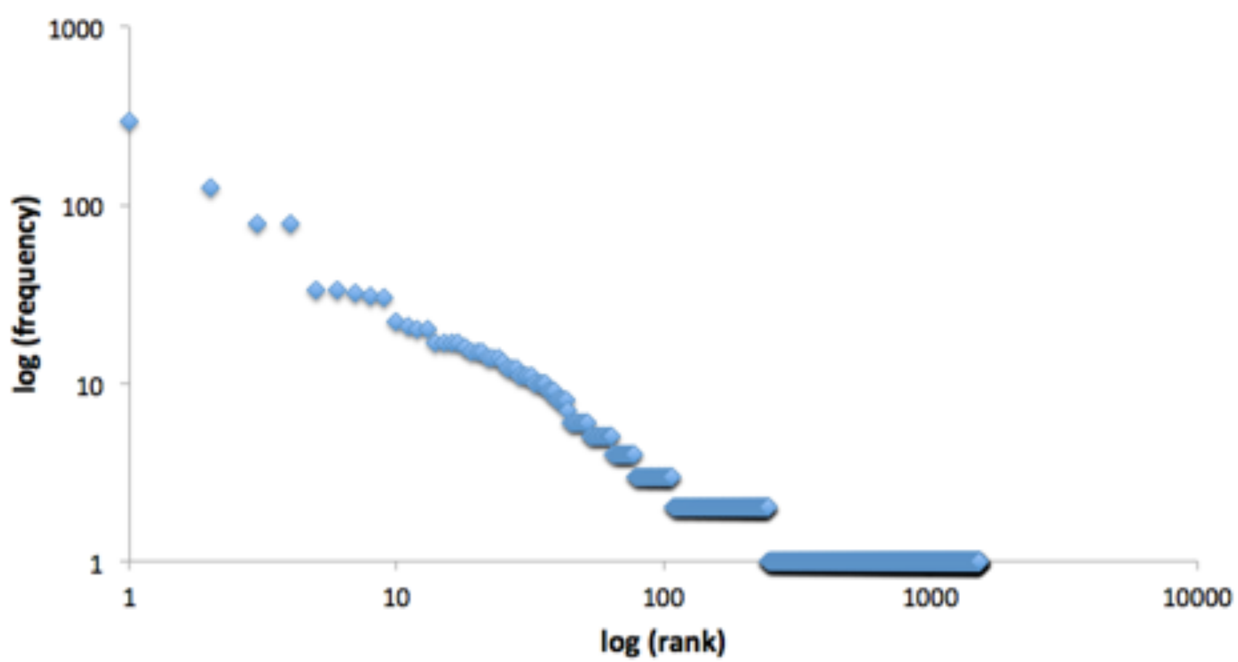


Threshold
$d = 2.6$



Segment Types

- Segmenting the whole corpus (15,197 chords) at $d = 2.6$
- 3,007 segment tokens
- 1,531 segment types (unique)



Rank	Frequency
1	297
2	125
3	79
4	78
5	33
6	33
7	32
8	31
9	30
10	22

Conclusion and Future Work

- Difficult to evaluate without a ground truth.
- Useful segmentations of jazz chord sequences.
- Rough phrase structure found.
- Key tonal-harmonic units identified (e.g. ii^7-V^7-I).
- No in-built knowledge of music theory or tonal harmony.
- Future work will compare with human segmentations of harmony.

References

Cleary, J. G., & Witten, I. (1984). Data compression using adaptive coding and partial string matching. Communications, IEEE Transactions on, 32(4), 396–402.

Cleary, J. G., & Teahan, W. J. (1997). Unbounded length contexts for PPM. The Computer Journal, 40(2 and 3), 67–75.

Conklin, D., & Witten, I. (1995). Multiple viewpoint systems for music prediction. Journal of New Music Research, 24(1), 51–73.

Griffiths, S., Forth, J., Purver, M., Mora McGinity, M., & Wiggins, G. A. (submitted). Segmentation of Natural Language using the IDyOM Model. Proceedings of the 37th Annual Conference of the Cognitive Science Society.

Pachet, F., Suzda, J., & Martín, D. (2013). A comprehensive online database of machine-readable leadsheets for jazz standards (pp. 275–280). Presented at the 14th International Society for Music Information Retrieval Conference, Curitiba, Brazil.

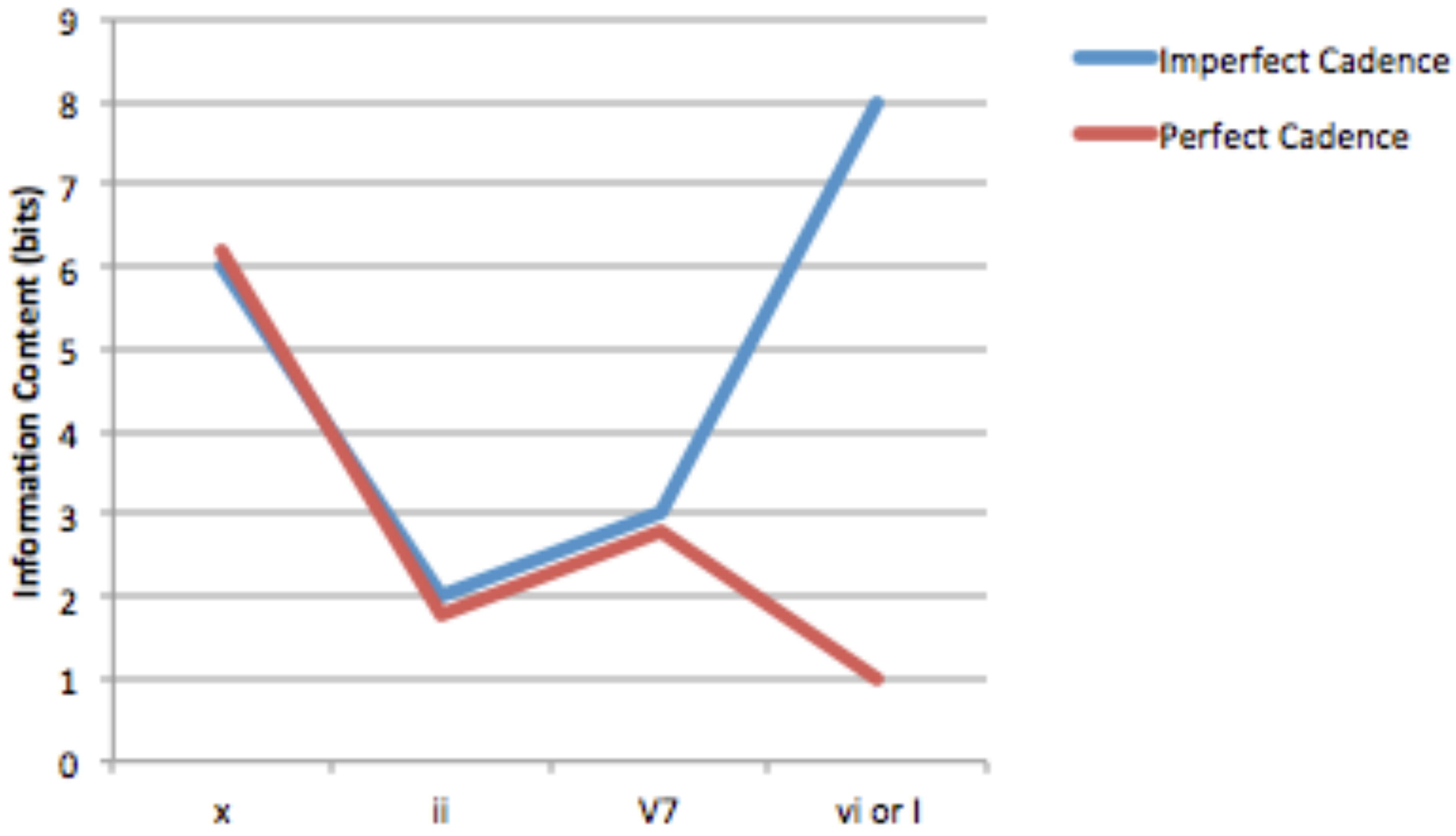
Moffat, A. (1990). Implementing the PPM data compression scheme. Communications, IEEE Transactions on, 38(11), 1917–1921. doi:10.1109/26.61469

Pearce, M. (2005). *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition*. PhD Thesis, City University, London.

Wiggins, G. A. (2012). "I let the music speak": cross-domain application of a cognitive model of musical learning. In P. Rebuschat & J. Williams, Statistical Learning and Language Acquisition (pp. 463–494). Amsterdam, NL: Mouton De Gruyter.

Wiggins, G. A., & Forth, J. (2015). IDyOT: A Computational Theory of Creativity as Everyday Reasoning from Learned Information. In *Computational Creativity Research: Towards Creative Machines* (Vol. 7, pp. 127–148). Atlantis Press.

cclab bonus: Imperfect Cadence Problem



cclab bonus: ’Round Midnight