Introducing the Mazurka Project

Nick Cook and Craig Stuart Sapp

Music Informatics research group seminar

School of Informatics, City University, London

22 February 2007

the ***Sensationalized*** Story of

Introducing the Mazurka Project

An exclusive exposé by

Nick Cook and Craig Stuart Sapp

about their role (or lack thereof) in the alleged Hatto Hoax.

Music Informatics research group seminar

School of Informatics, City University, London

22 February 2007

Source material: mazurka recordings

Mazurka in G minor Op. 24, No. 1

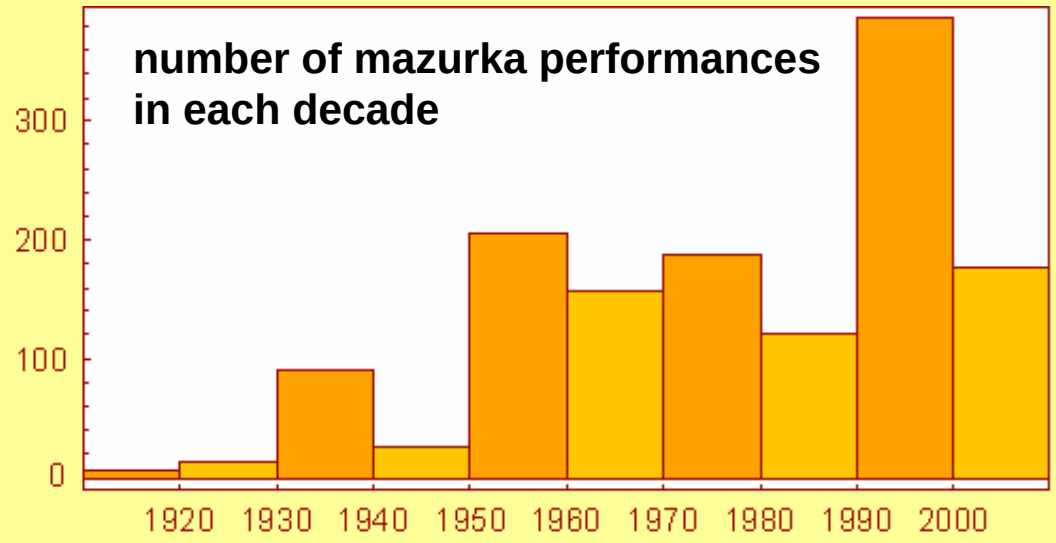
29 performances:

2:48	Ashkenazy (1981)	Decca 448 086-2
3:06	Biret (1990)	Naxos 8.550359
2:39	Block (1995)	ProPiano PPR224507
2:04	Brailowsky (1960)	Sony SB2K 63237
2:49	Chiu (1999)	HMX 2907352.53
2:30	Clidat (1994)	Forlane UCD16729
2:50	Cortot (1951)	Concert Artist 9180/12
2:44	Falvay (1989)	Naxos 8.550256
3:06	Fiorentino (1962)	Concert Artist 9200-2
3:01	Flière (1977)	Melodia 10 00439
1:41	François (1956)	EMI CZS 7 67413 2
3:12	Hatto (1997)	Concert Artist 9270/12
3:12	Indjic (2001)	Calliope 3321
2:37	Kapell (1951)	RCA 09026-68990-2
2:52	Luisada (1990)	DG 463054-2
3:11	Lushtak (2004)	Centaur CRC 2707
3:12	Magaloff (1977)	Phillips 426 817/29-2
2:45	Nezu (2005)	DUX KCh15-10
2:47	Pobłolcka (1999)	BeArTon CDB012/13
3:08	Rosen (1989)	Globe 5028
2:03	Rubinstein (1939)	Naxos 8.110656-57
3:32	Rubinstein (1952)	BMG 09026 63027-2
2:48	Rubinstein (1966)	BMG 09026-63050-2
2:57	Shebanova (2002)	DUX 0350/0351
3:45	Smith (1975)	EMI 724358576726
3:04	Ts'ong (1993)	Sony SB2K 53 246
2:59	Ts'ong (2005)	NIFC CD001
2:41	Tsujii (2005)	DUX KCh15-7
1:50	Uninsky (1959)	Philips 442 574-2

- **1,500+ recordings of 49 mazurkas**

  = 30 performances/mazurka on average

- **about 65 performers, 75 CDs**



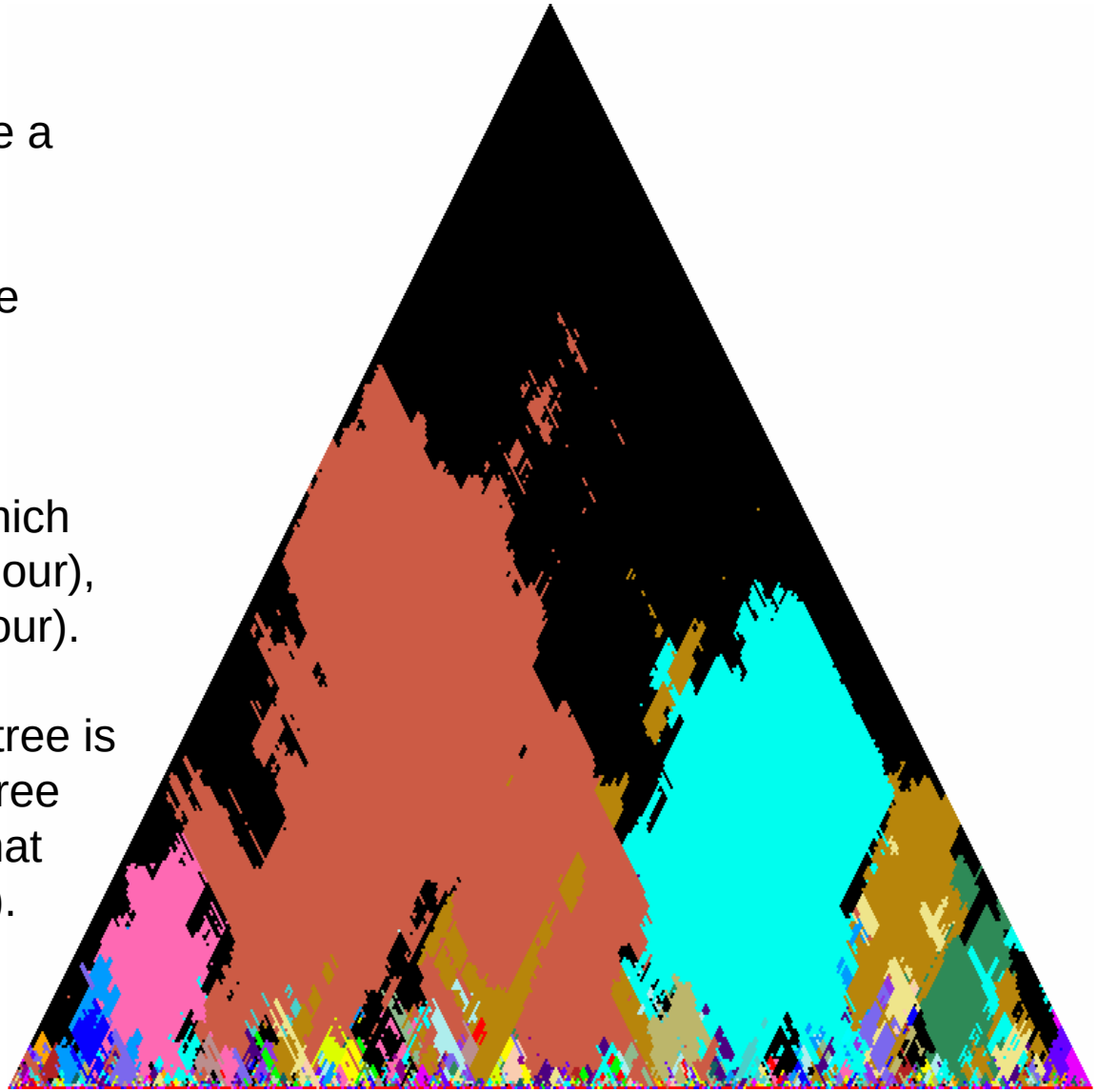
Performance-Correlation Timescapes

A multi-correlation timescape is like a photograph taken inside a forest.

The photographer is standing at the position of the performance being analyzed.

The colours are images of trees which are close by (large splotches of colour), or far away (small splotches of colour).

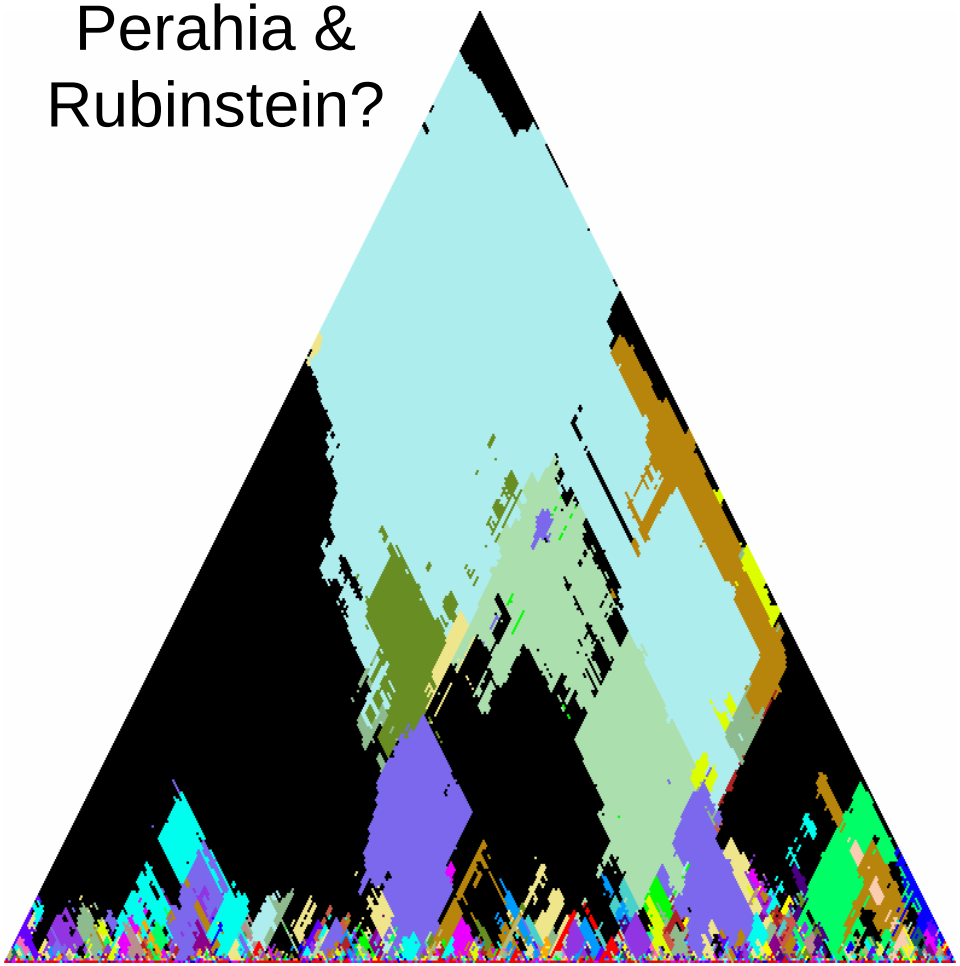
Black indicates that the performer tree is not particularly close to any other tree (but there could have been ones that died and have fallen to the ground).



Purpose of Timescapes

- To examine influences between performers

Perahia &
Rubinstein?



Rubinstein 1939

Uninsky 1971

Rubinstein 1966

Indjic 1988

Hatto 1997

Shebanova 2002

Magaloff 1977

Horowitz 1985

Horowitz 1971

Perahia 1994
Mazurka in A minor Op. 17, No. 4

for more: http://mazurka.org.uk/ana/pcor

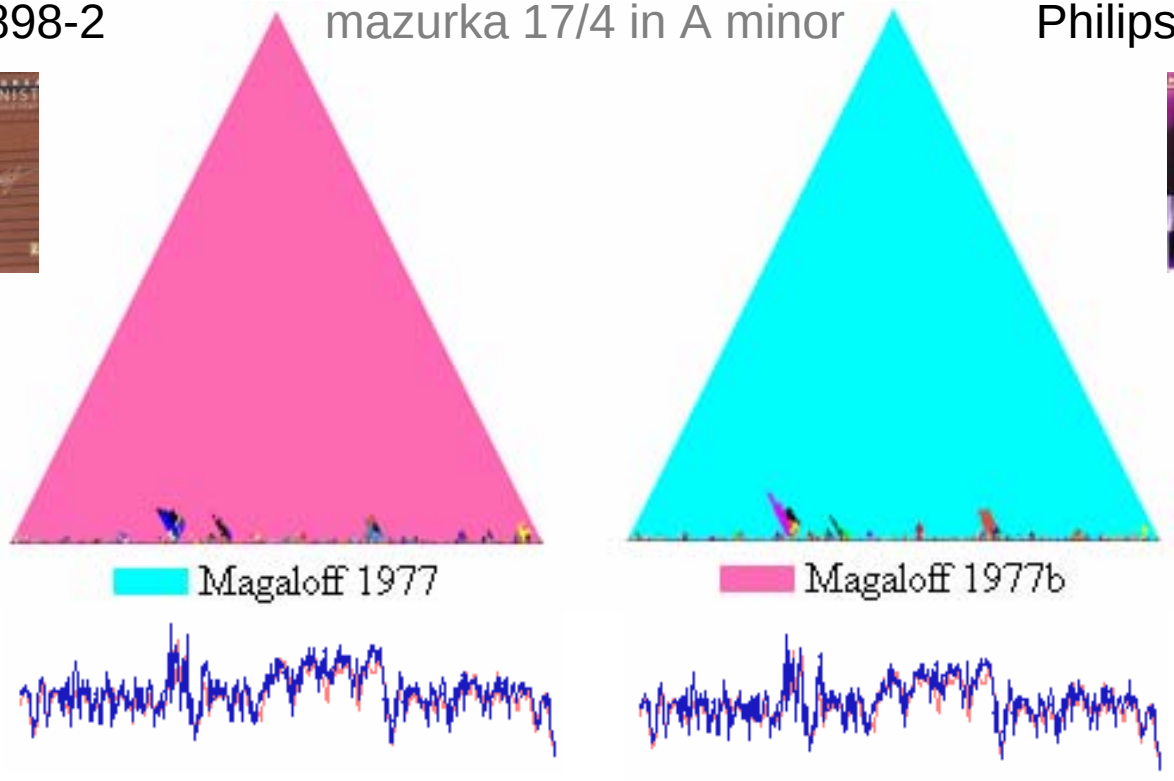
Boring Timescape Pictures

Occasionally we get over-exposed photographs back from the store, and we usually have to throw them in the waste bin.

The same performance by Magaloff on two different CD re-releases:

Philips 456 898-2

mazurka 17/4 in A minor



Philips 426 817/29-2

• Structures at bottoms due to errors in beat extraction, measuring limits in beat extraction, and correlation graininess.

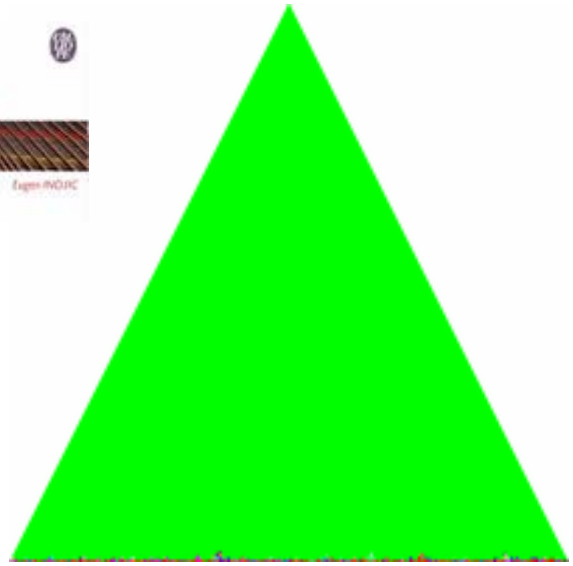
Boring Timescape Pictures?

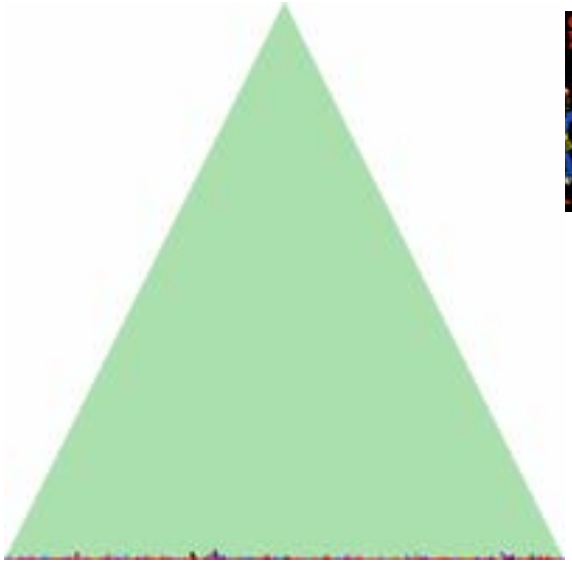
Two difference performances from two different performers on two different record labels from two different countries.

Calliope 3321

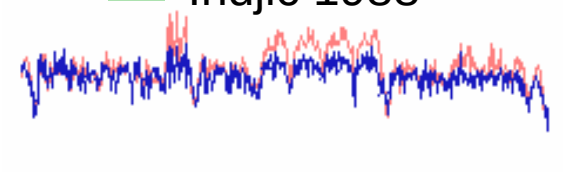
mazurka 17/4 in A minor

Concert Artist 20012

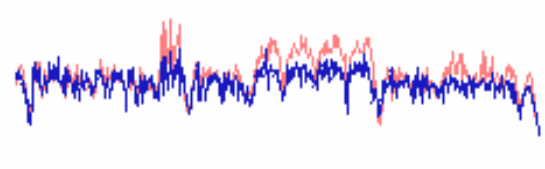




Indjic 1988



Hatto 1997



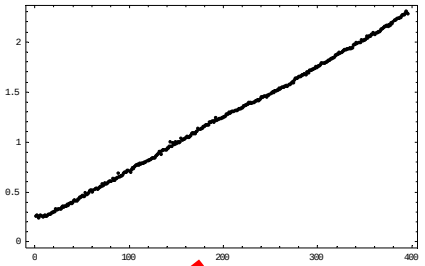
see: http://www.charm.rhul.ac.uk/content/contact/hatto_article.html

Beat-Event Timing Differences

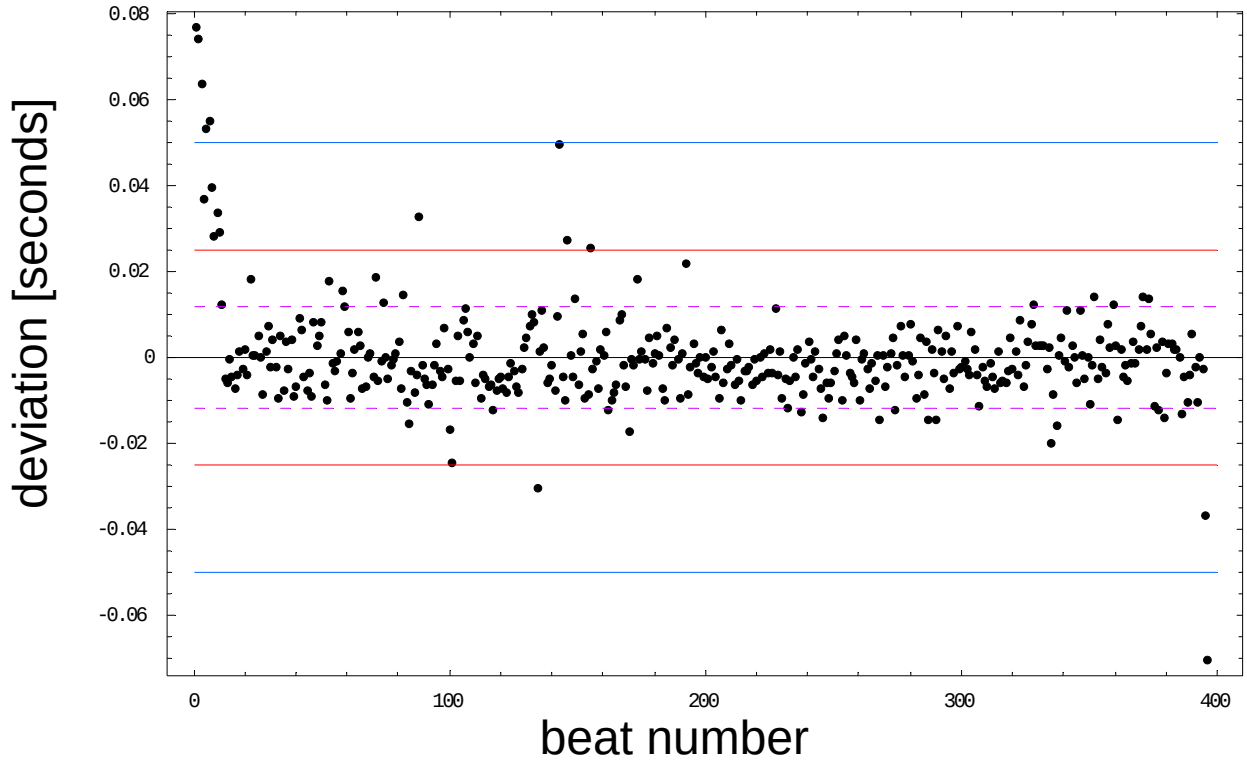
Hatto beat location times: 0.853, 1.475, 2.049, 2.647, 3.278, etc.

Indjic beat location times: 0.588, 1.208, 1.788, 2.408, 3.018, etc.

difference plot



Hatto / Indjic beat time deviations



remove 0.7% timeshift

Pataphysical[1] Analysis of Journalism

Reported:

> “Conclusive? Perhaps not, as there is still a one in 1000 possibility that Hatto made her own recordings, but certainly troubling.”

Actually:

> Based on reasonable scientific assumptions, the probability that the Hatto performance of Mazurka in A minor Op. 17 No. 4 is not the same performance as that of Indjic is less than

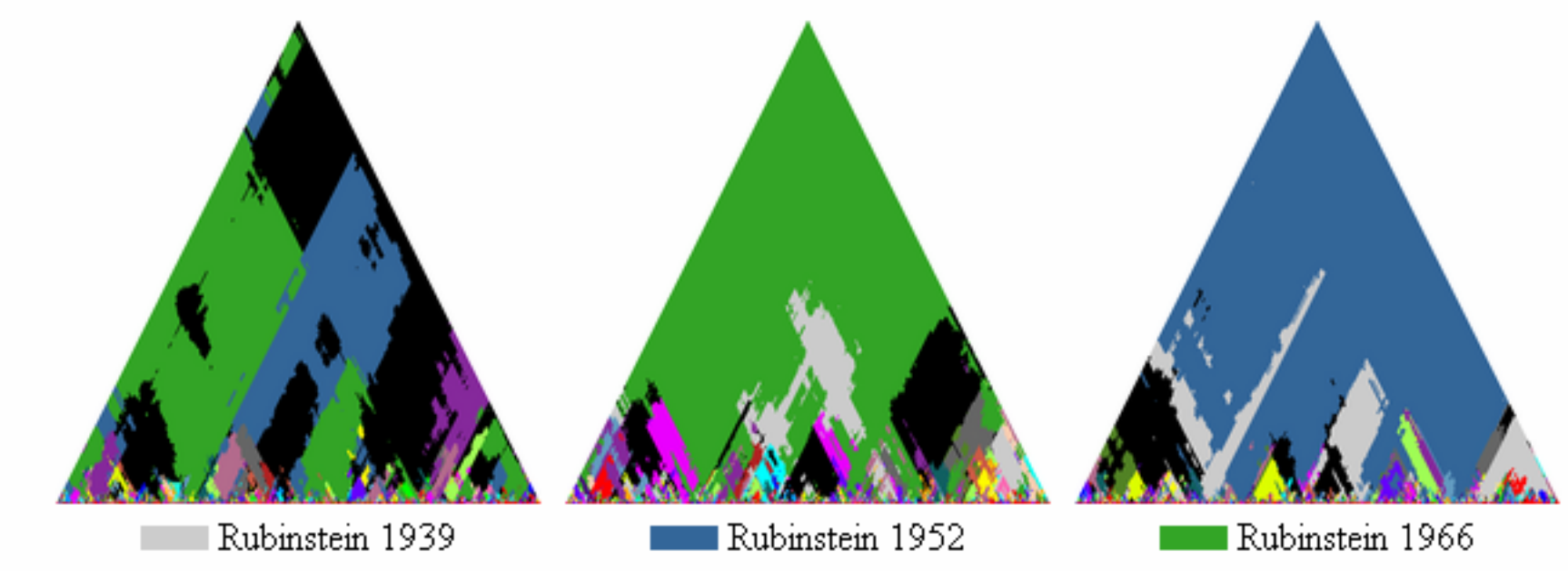
$$1 \text{ in } \sqrt{\text{googol}}$$

This probability is equivalent to 1 atom out of all atoms in a typical star

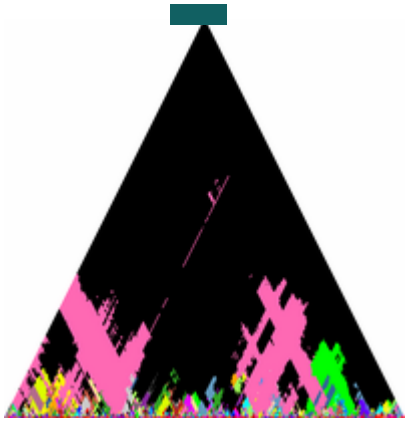
[1] MrT

Slightly More Interesting

Same work; same pianist; different performance

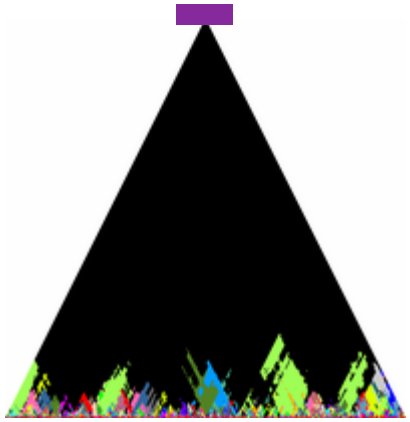


Yet More Interesting

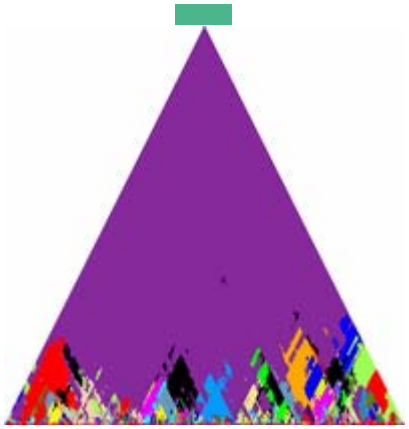


p

n

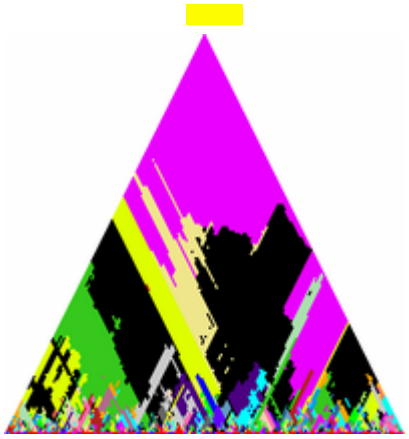


s

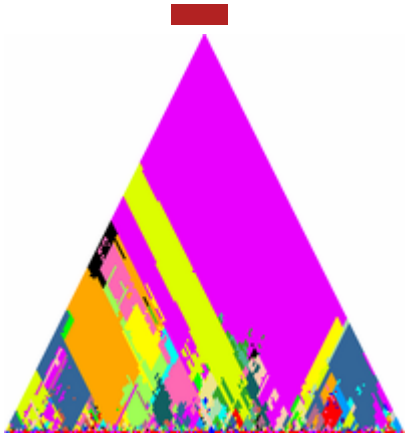


t

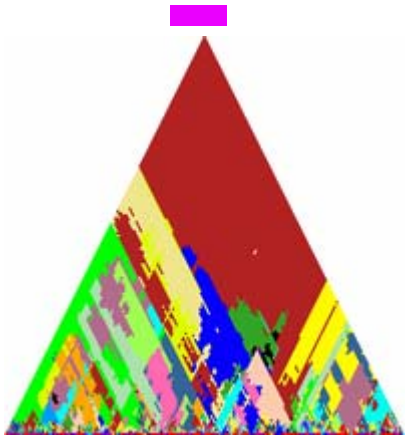
The conspiracy goes deeper? or Not?



c



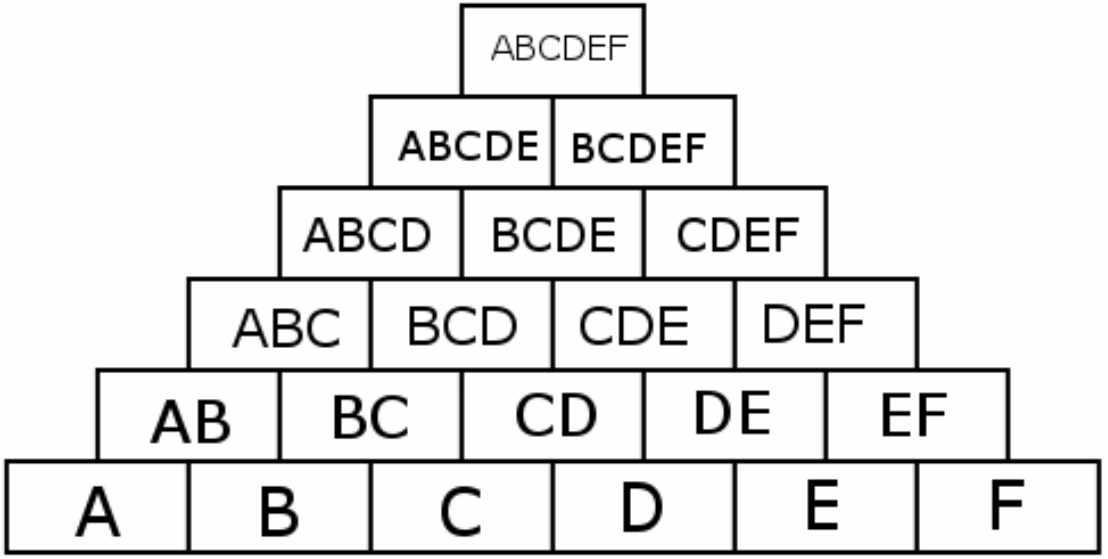
f



b

Timescape's Plotting Domain

- Examine the internal tempo/harmony structure of a performance/composition
- 1-D data sequences transformed into 2-D plot
- Example of a piece with 6 beats at tempos A, B, C, D, E, and F:



- In performance correlation timescapes, correlation measurements calculated for each cell. Also tempo average timescapes can be calculated.

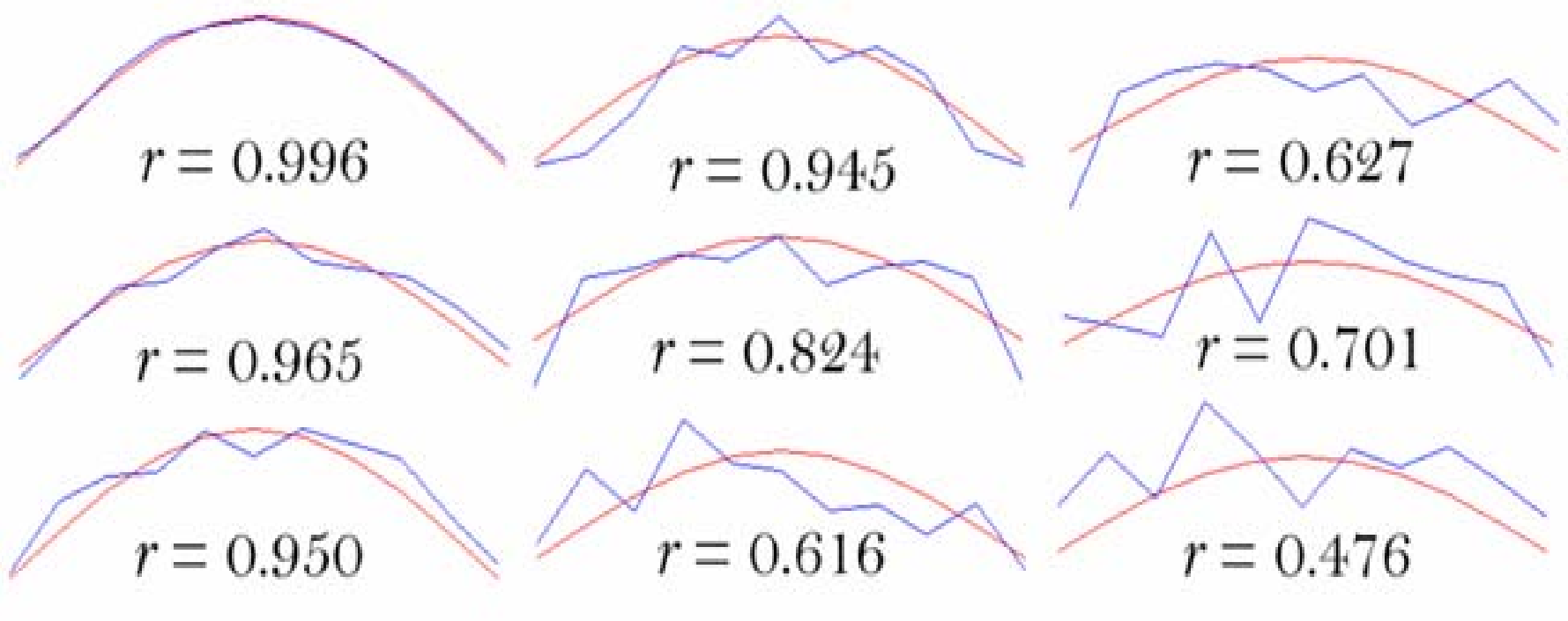
Under the Hood (Bonnet)

Pearson correlation:

$$\frac{\sum_i (x_i - \bar{x})\,(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

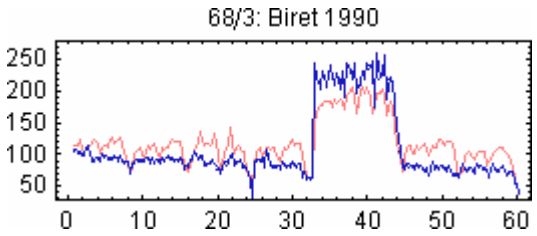
- **Measures how well two shapes match:**

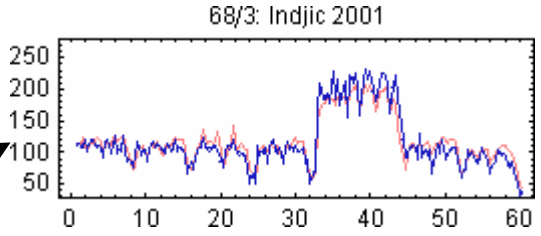
r = 1.0 is an exact match.
r = 0.0 means no relation at all.

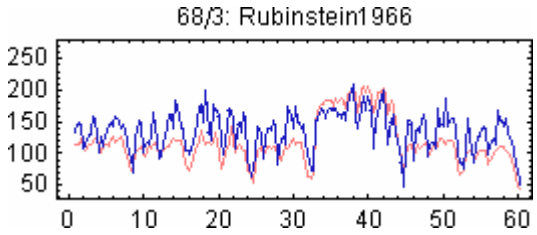


Overall performance correlations

	Bi	Br	Ch	Fl	In	Lu	R8	R6	Sm	Un
Biret	1.	0.92	0.81	0.83	0.95	0.85	0.62	0.5	0.55	0.86
Brailowsky	0.92	1.	0.81	0.86	0.91	0.84	0.66	0.55	0.65	0.85
Chiu	0.81	0.81	1.	0.86	0.86	0.81	0.76	0.74	0.67	0.89
Friere	0.83	0.86	0.86	1.	0.88	0.84	0.73	0.7	0.74	0.89
Indjic	0.95	0.91	0.86	0.88	1.	0.88	0.66	0.59	0.63	0.9
Luisada	0.85	0.84	0.81	0.84	0.88	1.	0.67	0.61	0.56	0.89
Rubinstein 1938	0.62	0.66	0.76	0.73	0.66	0.67	1.	0.77	0.62	0.75
Rubinstein 1966	0.5	0.55	0.74	0.7	0.59	0.61	0.77	1.	0.59	0.69
Smith	0.55	0.65	0.67	0.74	0.63	0.56	0.62	0.59	1.	0.64
Uninsky	0.86	0.85	0.89	0.89	0.9	0.89	0.75	0.69	0.64	1.





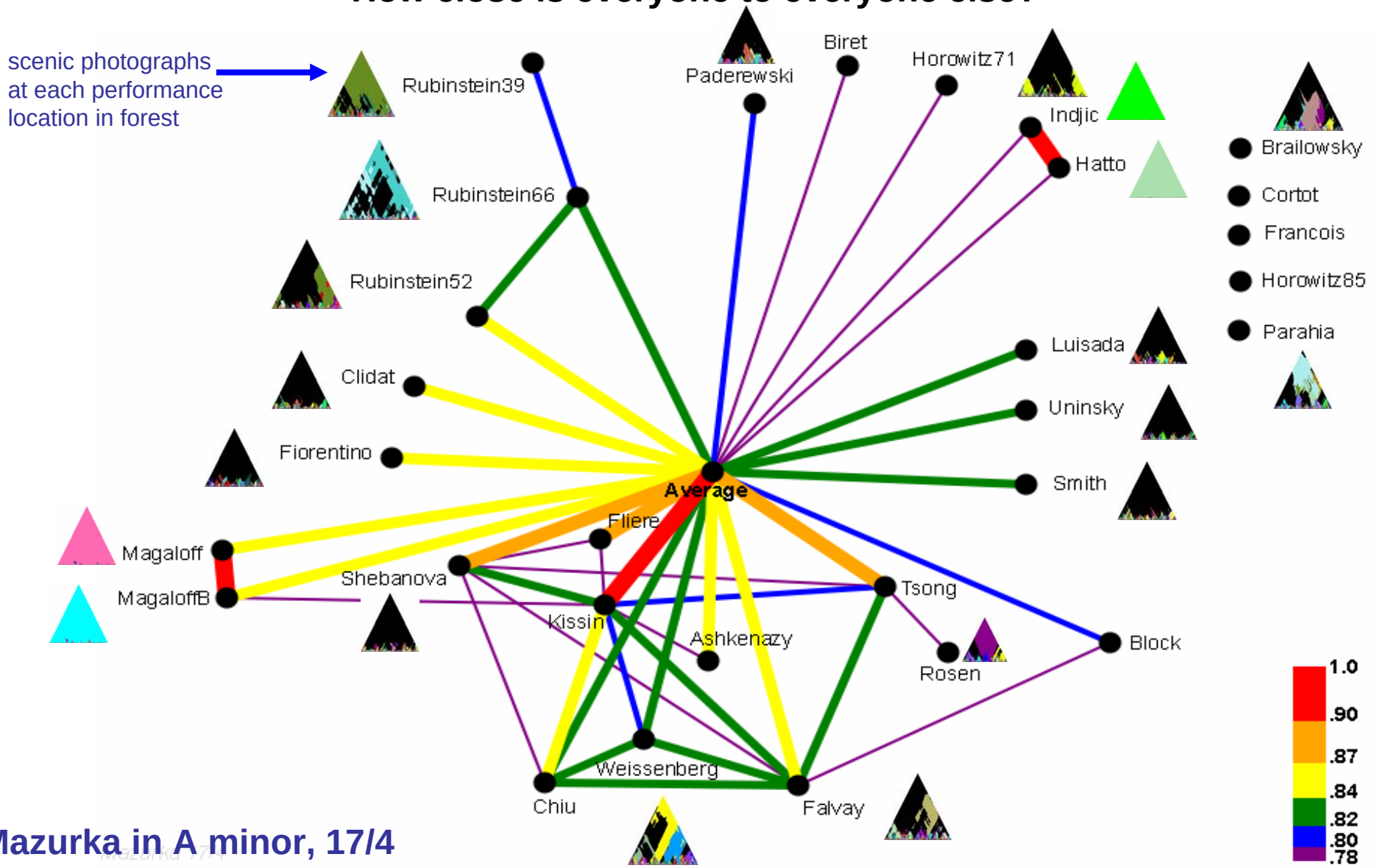


Highest correlation to Biret 1990

Lowest correlation to Biret 1990

Map of the Forest: Correlation Network

- **How close is everyone to everyone else?**



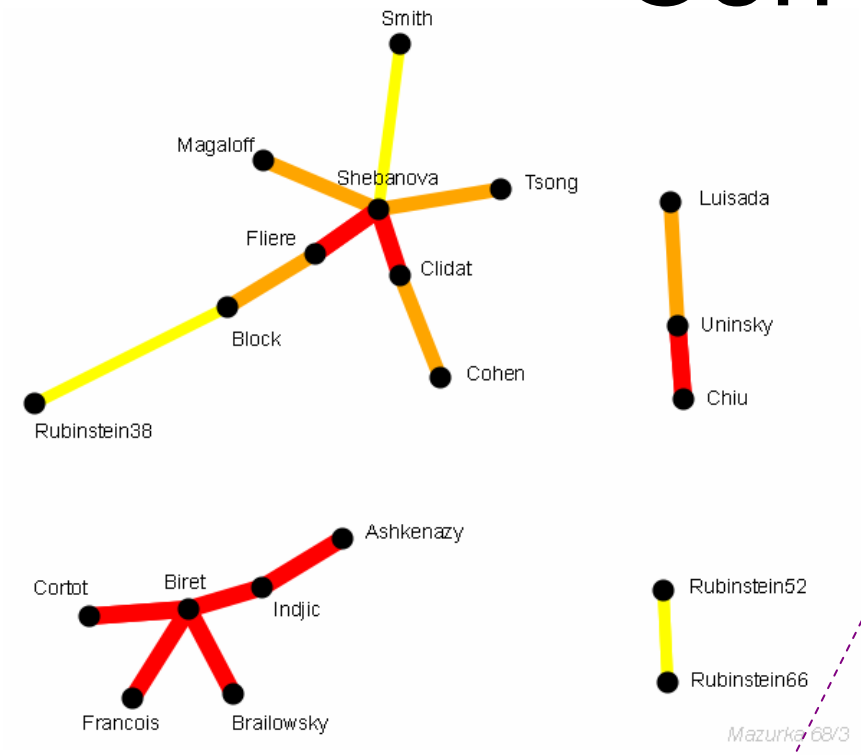
Mazurka in A minor, 17/4

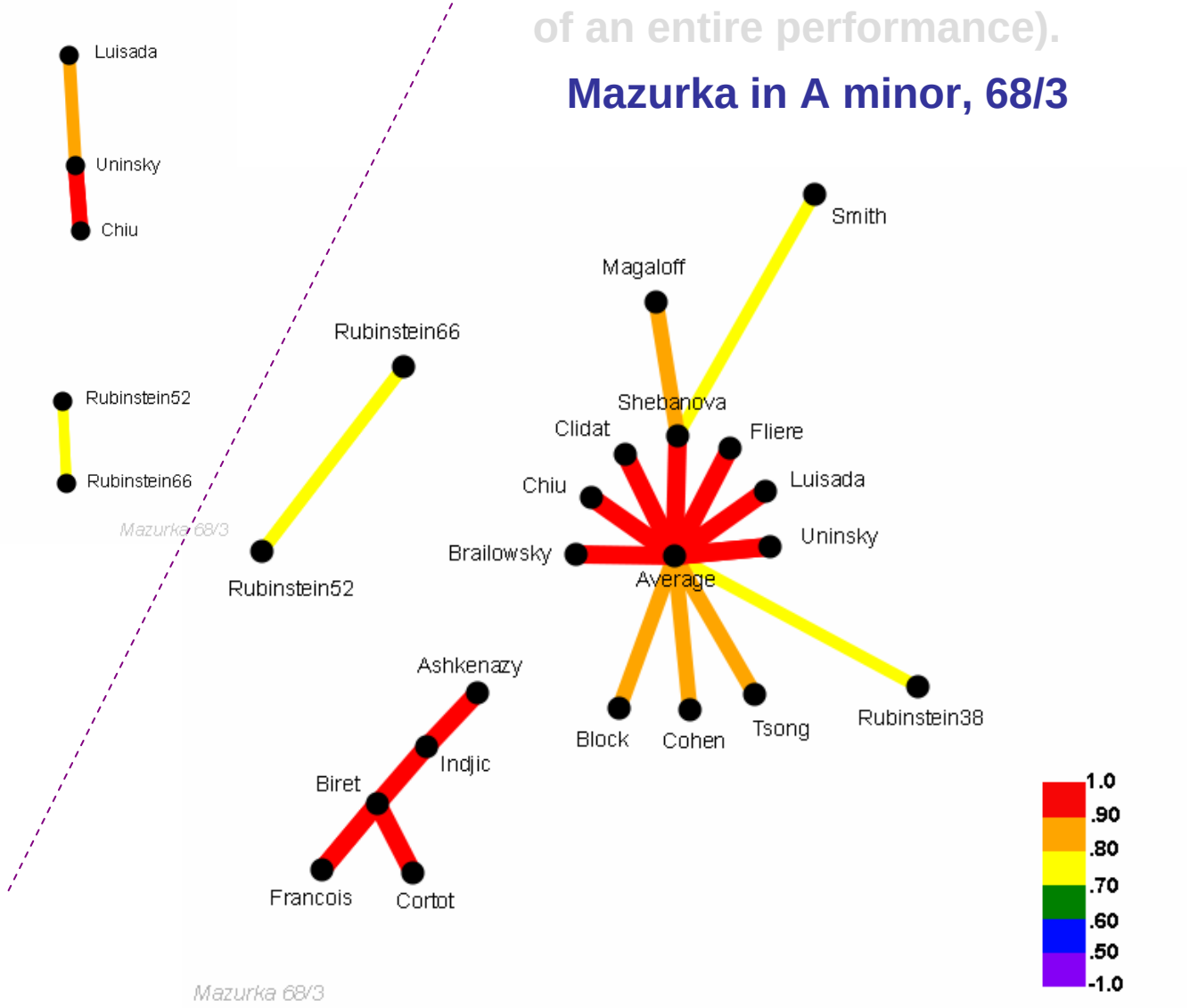
Correlation tree

- **Who is closest to whom?**

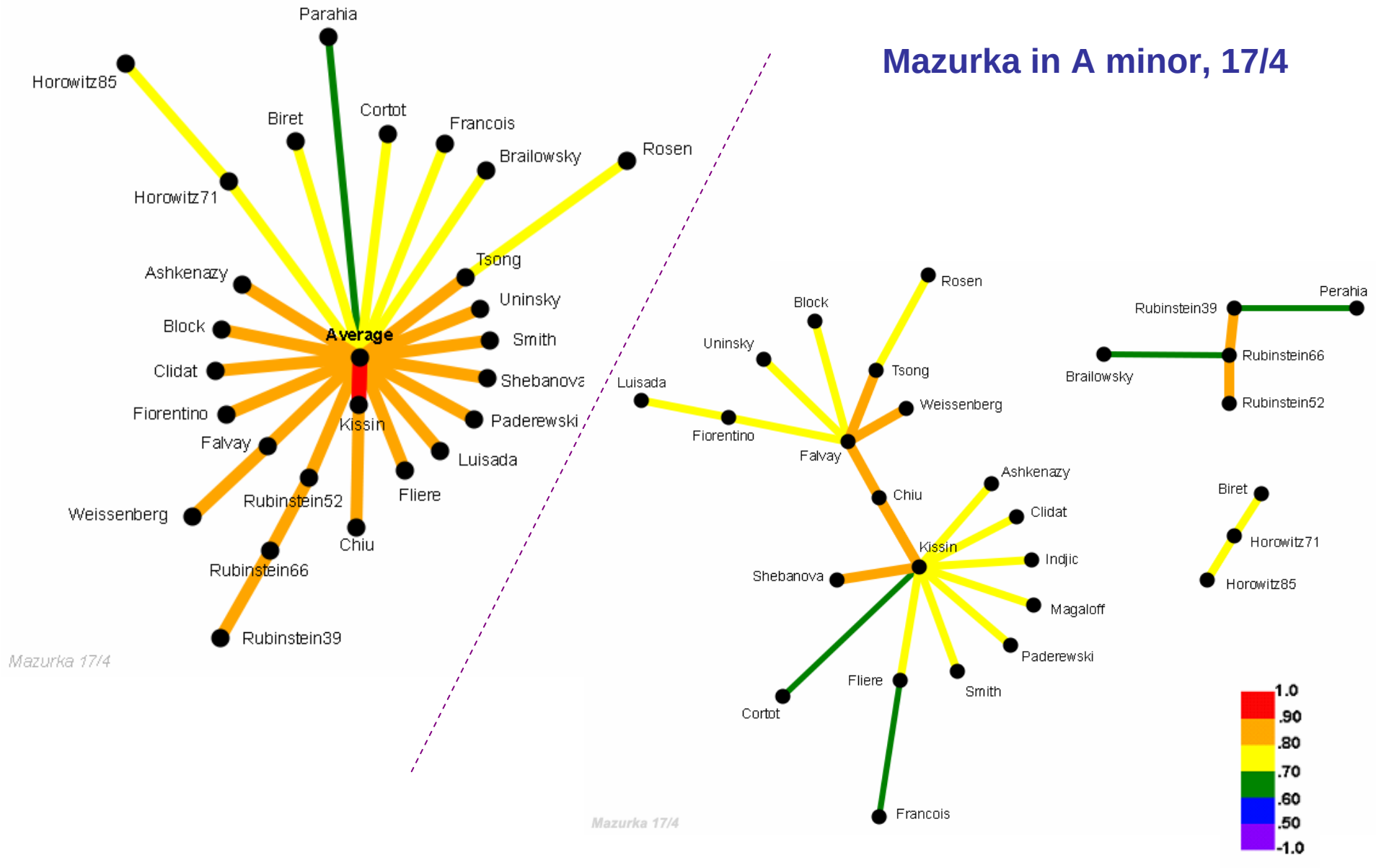
(with respect to beat tempos of an entire performance).

Mazurka in A minor, 68/3

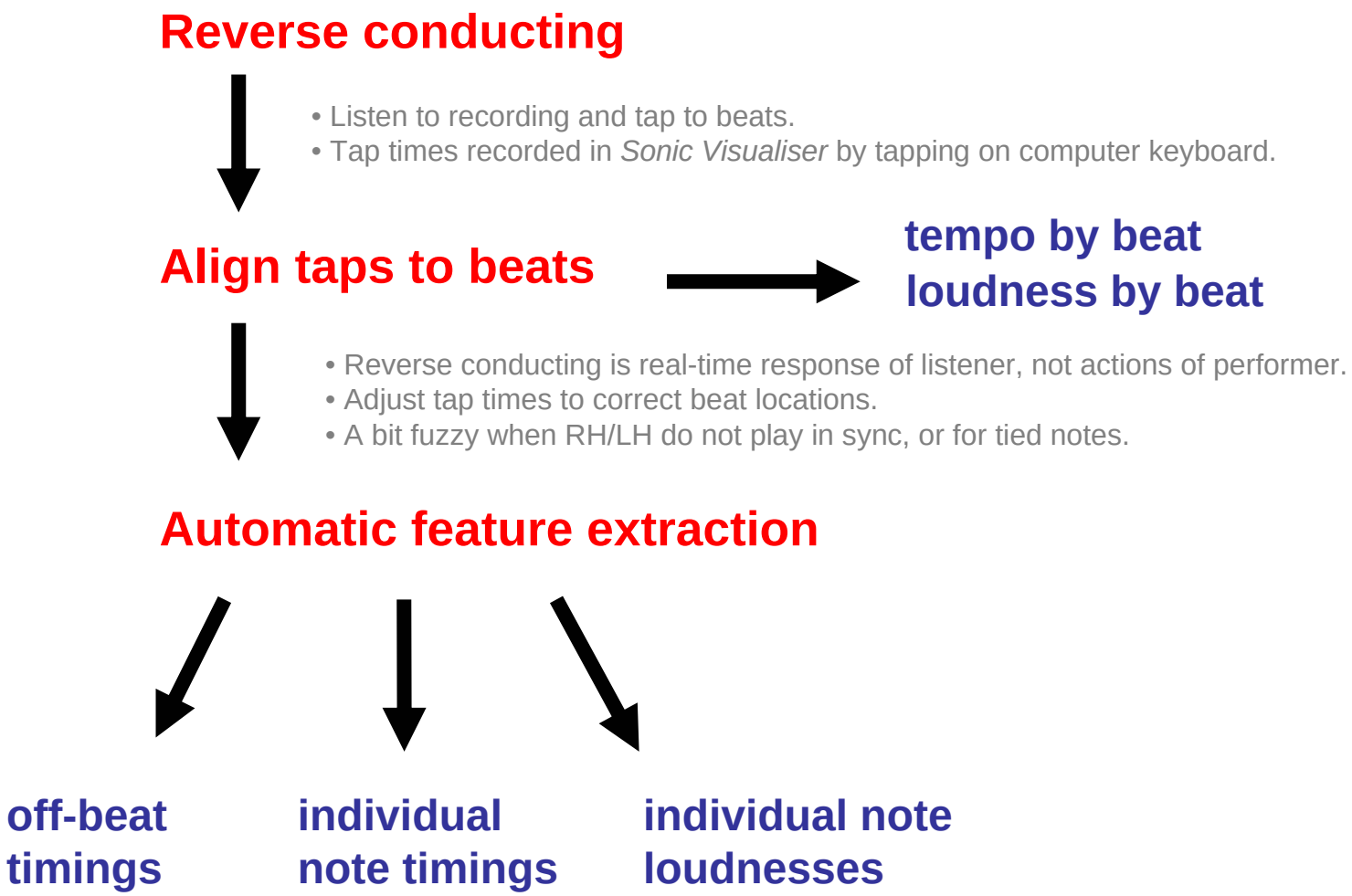




Correlation tree (2)

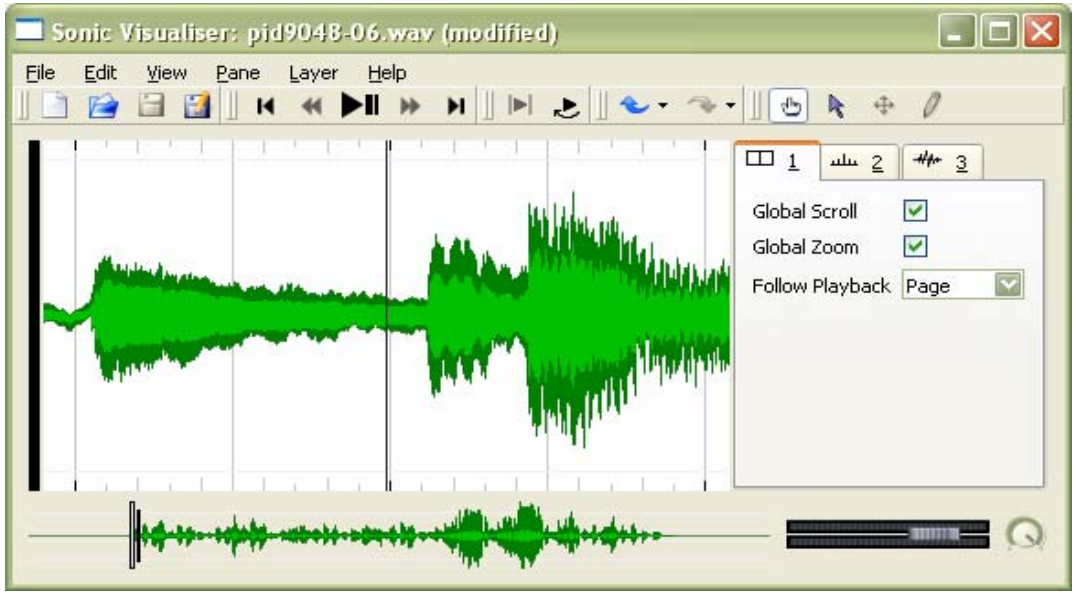


Performance data extraction



Sonic Visualiser

- Underlying data for timescapes extracted using Sonic Visualiser



Developed at the Center for Digital Music, Queen Mary, University of London

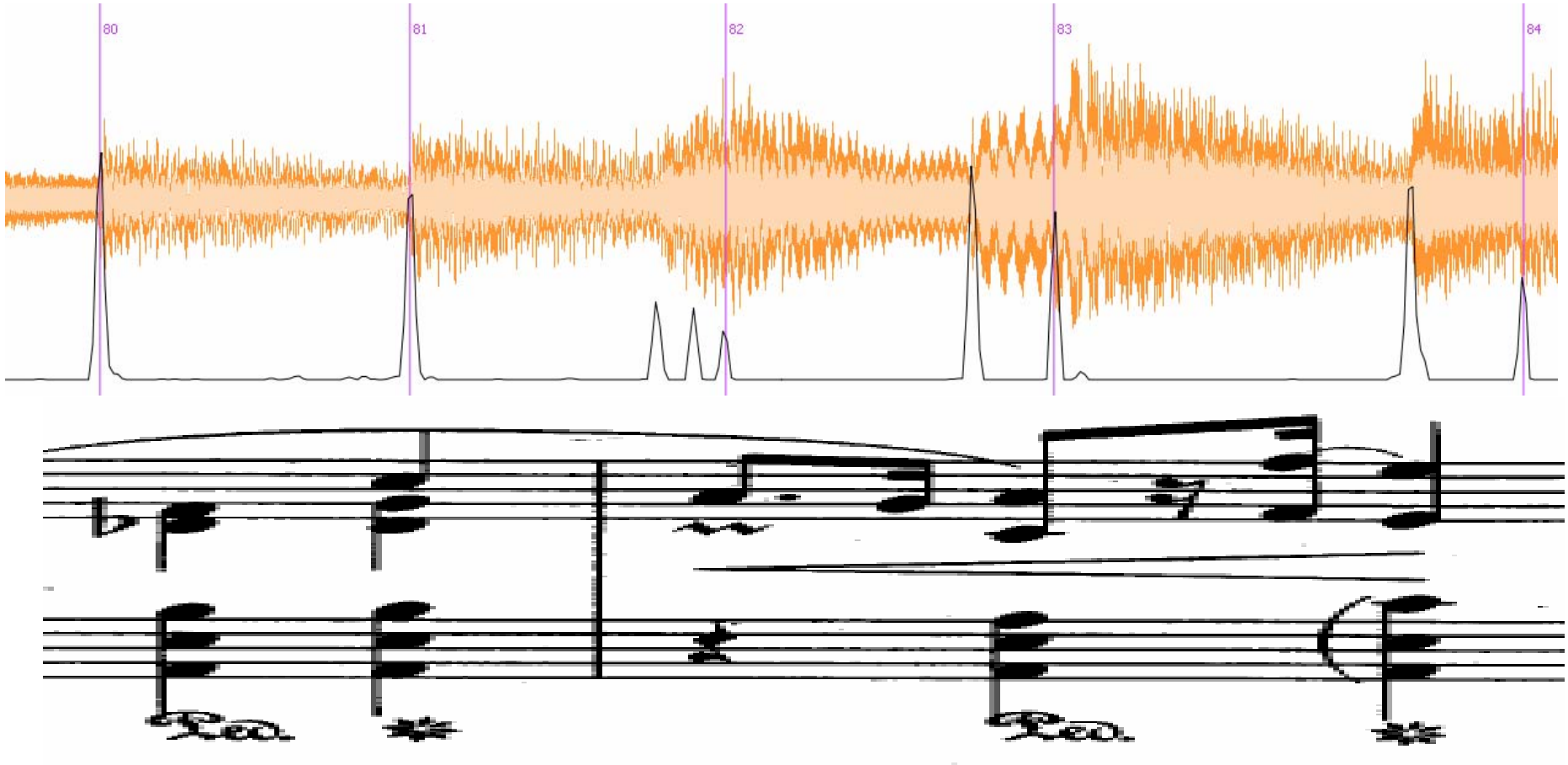
Main webpage for Sonic Visualiser: http://www.sonicvisualiser.org

Online documentation http://sv1.sourceforge.net/doc/reference/en

Analysis plugin resources: http://sv1.sourceforge.net/vamp.html

Mazurka plugins http://sv.mazurka.org.uk

*Mz*SpectralReflux plugin for SV



http://sv.mazurka.org.uk/MzSpectralReflux

http://mazurka.org.uk/cgi-bin/tapsnap = Move taps to nearest onset

Additional Slides

iTunes not Examing Musical Content

• iTunes does not examine musical content: only describes CD contents by track count and hash of track times.

• 1st CD identified by Brian Ventura using iTunes: Liszt's Transcendental Etudes, on or before 13 Feb 2007.

• Then 2nd CD of Rachmaninov Piano Concertos identified by Jed Distler on or before 15 Feb 2007 using iTunes.

• 4th CD identified of Brahms 2nd Piano Concerto identified by Steve de Mena on morning of 16 Feb 2007 or most likely before.

• 18 CD sources identified by morning of 21 February 2007 (methods vary)

references: http://www.classicstoday.com/features/021807-joycehatto.asp
http://www.gramophone.co.uk/newsMainTemplate.asp?storyID=2759&newssectionID=1
rec.music.classical.recording
http://en.wikipedia.org/Joyce_Hatto

iTunes / CDDB

http://en.wikipedia.org/wiki/Cddb

- Gramophone groups initial discovery was by using iTunes CDDB technology

John Oswald w Gilles Arteau & Émile Morin / .Parcours Scénographique

Tracks: 1
Total time: 60:10
Year: 1997
Disc-ID: classical / 020e1a01

1. Parcours Scénographique

Craig Sapp / Digital Whitenoise

Tracks: 1
Total time: 60:10
Year: 1999
Disc-ID: blues / 020e1a01

1. Whitenoise

Track/Time Hash is identical: 020e1a01

first six digits are hash of track times

last two digits are track count

discid problem usually resolved by placing in another category (classical v blues)

Purely Coincidental?

One of the major characters in the recent Hatto story is from this school:

> Andrew Rose from Pristine Audio has a music degree from the **City University of London**, and spent 14 years as a sound engineer at BBC Radio in London.

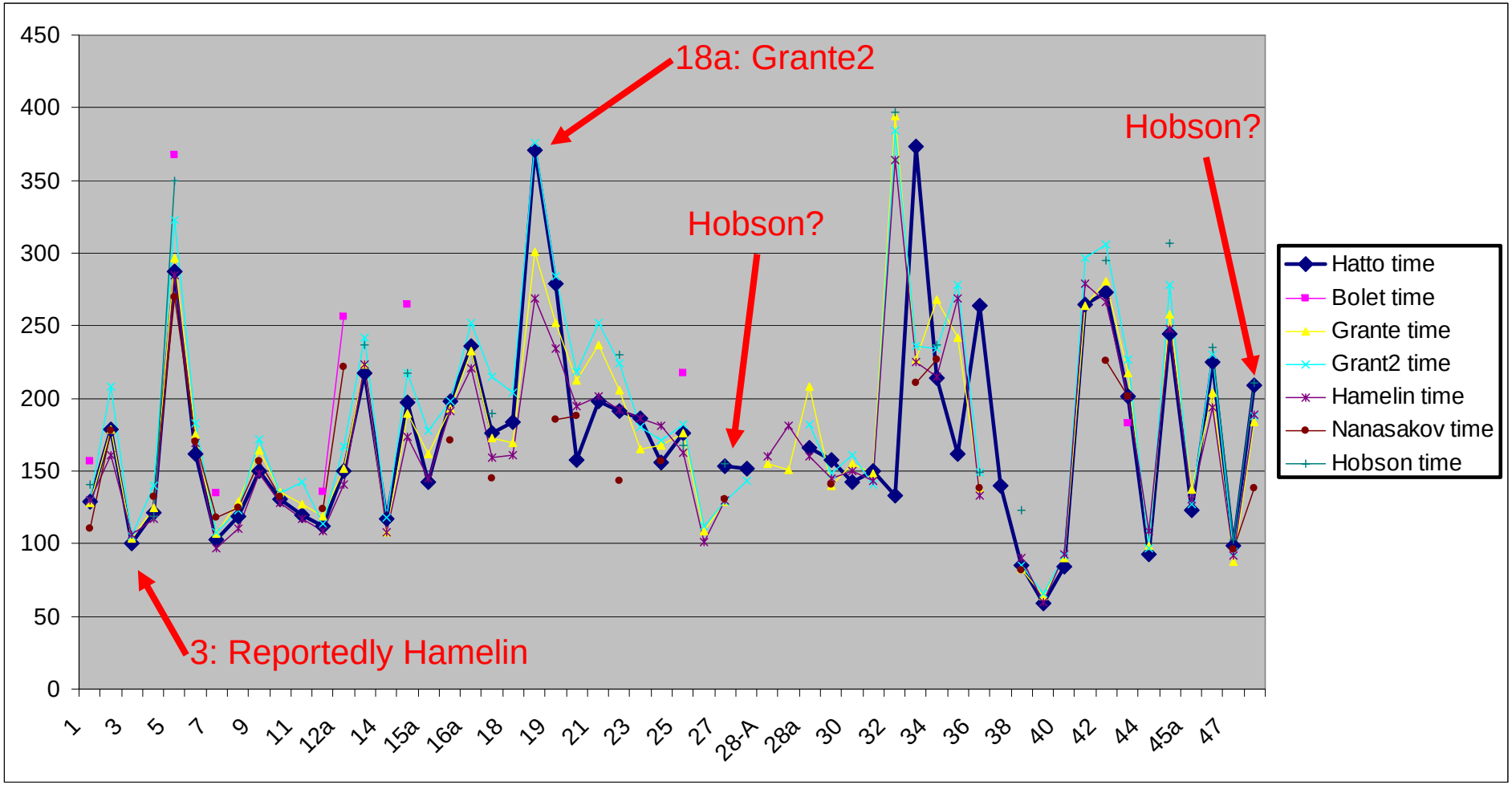
• 3rd CD identified by Gramophone group on or before 15 Feb 2007 was not found using iTunes. Andrew Rose of Pristine Audio verified first two results by expert analysis of audio content. He also discovered Godowsky/Chopin Studies to be a copy of a previous commercial release by Carlo Grante (Later it was reported by a contributor to rec.music.recording.classical that one of the tracks, #3, is by Marc Hamelin)

http://en.wikipedia.org/wiki/Leopold_Godowsky

> These [Studies] are so taxing even for virtuosi that only five have ventured to record the entire set: Geoffrey Douglas Madge, Carlo Grante, Marc-André Hamelin and Joyce Hatto.

So: Godowsky is the best choice for a manual search for a match.

Godowsky/Chopin Study track times



- Study 18a is the most suspicious coincidence.