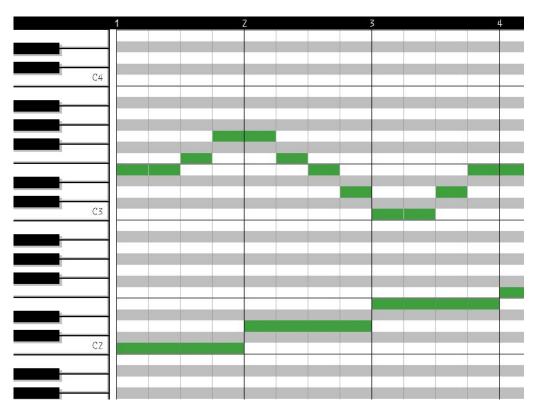
Symbolic Music Representation in Computers

- ABC Notation
- Piano Roll
- MIDI format
- Music XML

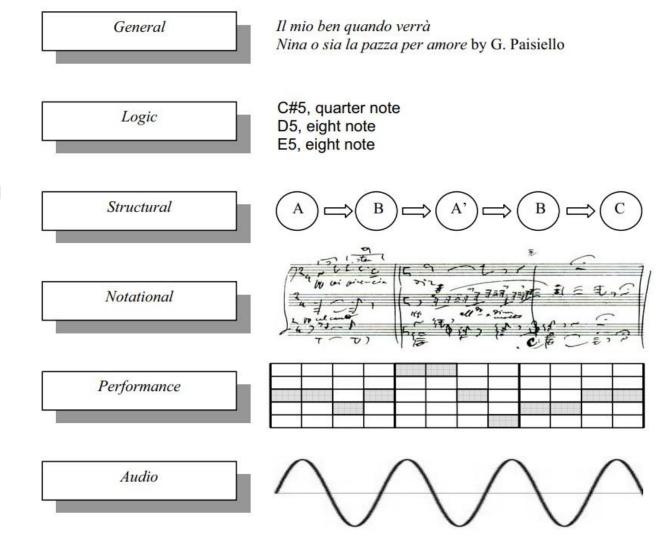


IEEE 1599: Music XML

- Recommended Practice Dealing With Applications and Representations of Symbolic Music Information Using the XML Language.
- Music needs to be represented with something that goes beyond unreadable, binary standards for audio, such as WAV and MP3, which are not music standards—they are audio standards. Musical aspects beyond audio must be represented in human-readable form, such as symbols and characters.
- This has always been the case for music scores in classical music, and for music notation in all civilizations for at least 40 centuries. This is also the case with other symbols, such as the harmonic grid in jazz, and also in other written codes, as in non-Western music.

IEEE 1599: Music XML

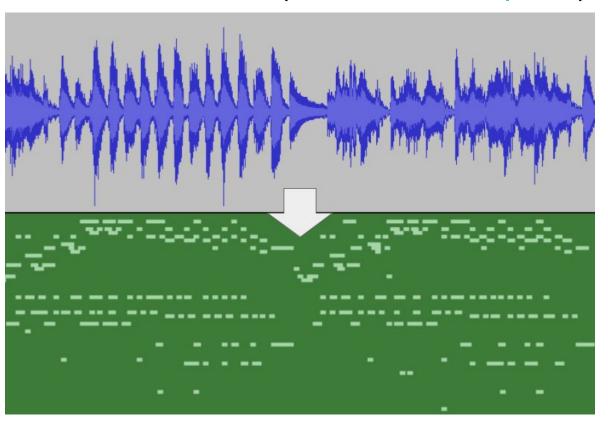
The main distinguishing features of the IEEE 1599 technology are the use of *symbols* to represent music, and the concept of *layers*.



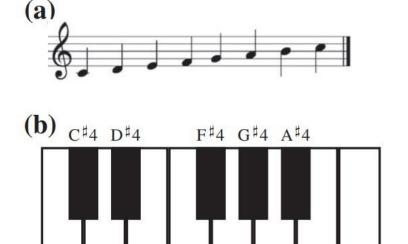
Examples



Music Information Retrieval (Piano Transcription)



Music Representation



(c)

Note	p	$f_{\text{MIDI}}(p)$	$f_{\text{MIDI}} \ (p - 0.5)$	$f_{\text{MIDI}} \ (p + 0.5)$	bandwidth	Q
C4	60	261.63	254.18	269.29	15.11	17.31
C ^{\$4}	61	277.18	269.29	285.30	16.01	17.31
D4	62	293.66	285.30	302.27	16.97	17.31
D ^{\$4}	63	311.13	302.27	320.24	17.97	17.31
E4	64	329.63	320.24	339.29	19.04	17.31
F4	65	349.23	339.29	359.46	20.18	17.31
F#4	66	369.99	359.46	380.84	21.37	17.31
G4	67	392.00	380.84	403.48	22.65	17.31
G ^{\$4}	68	415.30	403.48	427.47	23.99	17.31
A4	69	440.00	427.47	452.89	25.41	17.31
A#4	70	466.16	452.89	479.82	26.93	17.31
B 4	71	493.88	479.82	508.36	28.53	17.31
C5	72	523.25	508.36	538.58	30.23	17.31

FIGURE 27.2

C4

D4

E4

F4

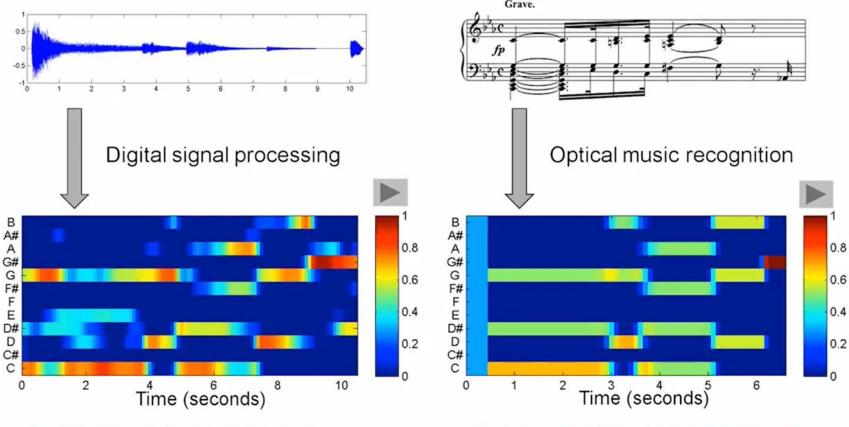
G4

A4

B4

(a) Musical score of a C major scale (C4 to C5). (b) Part of piano keyboard with keys ranging from C4 to C5. (c) MIDI pitch p, center frequency $f_{\text{MIDI}}(p)$, cutoff frequencies $f_{\text{MIDI}}(p-0.5)$ and $f_{\text{MIDI}}(p+0.5)$ to the left and right, bandwidth, and Q factor.

Convert into common mid-level feature representation



Audio chroma representation

Image chroma representation