Playing with Cases:
Rendering Expressive Music Performance with Case-Based Reasoning

Ramon Lopez de Mantaras
Artificial Intelligence Research Institute (IIIA)
Spanish National Research Council (CSIC)
http://www.iiia.csic.es/~mantaras

To Max Mathews (11/13/1926 – 4/21/2011), Computer music pioneer

Robert S. Engelmore Lecture, IAAI 2011, San Francisco

Outline

- About expressive music
- Some AI-based approaches to music performance
- About CBR
- SaxEx
- TempoExpress
- Conclusions and future work (GuitarLab)

About expressive music

- When we, humans, perform music, the result is never a literal rendering of what is written in the score: We *deviate* from the score. As far as performance *deviations* are intentional, they are commonly thought of as conveying expressivity

- Music elicits emotions and feelings and the more expressive is the music the stronger are these emotions and feelings (J. Sloboda, Psychology of Music 19, 1991; Boold, Zatorre et al. Nature Neuroscience 2(4), 1999)

- Mothers talk and sing to infants using a cooing tone and higher pitch than when interacting with adults and this helps pre-linguistic infants regulate their emotional states (L. Trainor, Nature 453, 29 May 2008)

AI-based expressive music performance

- Johnson ('92)
  - Rule-Based System to determine the note duration and articulation of fugues from "The well tempered clavier"
- Director Musices (Bresin'01)
  - Rules & Neural Networks for tempo, dynamics and articulation "Chopin Op.28 N.7"
- START and YQX projects (Widmer '99-)
  - Data Mining & ML to discover performance patterns (rules) and use them to render expressive performance "Mozart K280"

Also: Dannenberg, Hazan & Ramirez, Suzuki, Pachet, Reck-Miranda, Wiggins, de Poli,...

AI approach to music at the IIIA

Human performers use musical knowledge, not explicit in the score, that is very difficult to verbalize.

That an expressive effect is applied only once does not mean it is insignificant" (J. Sundberg et al., Music Perception 9, 1991)

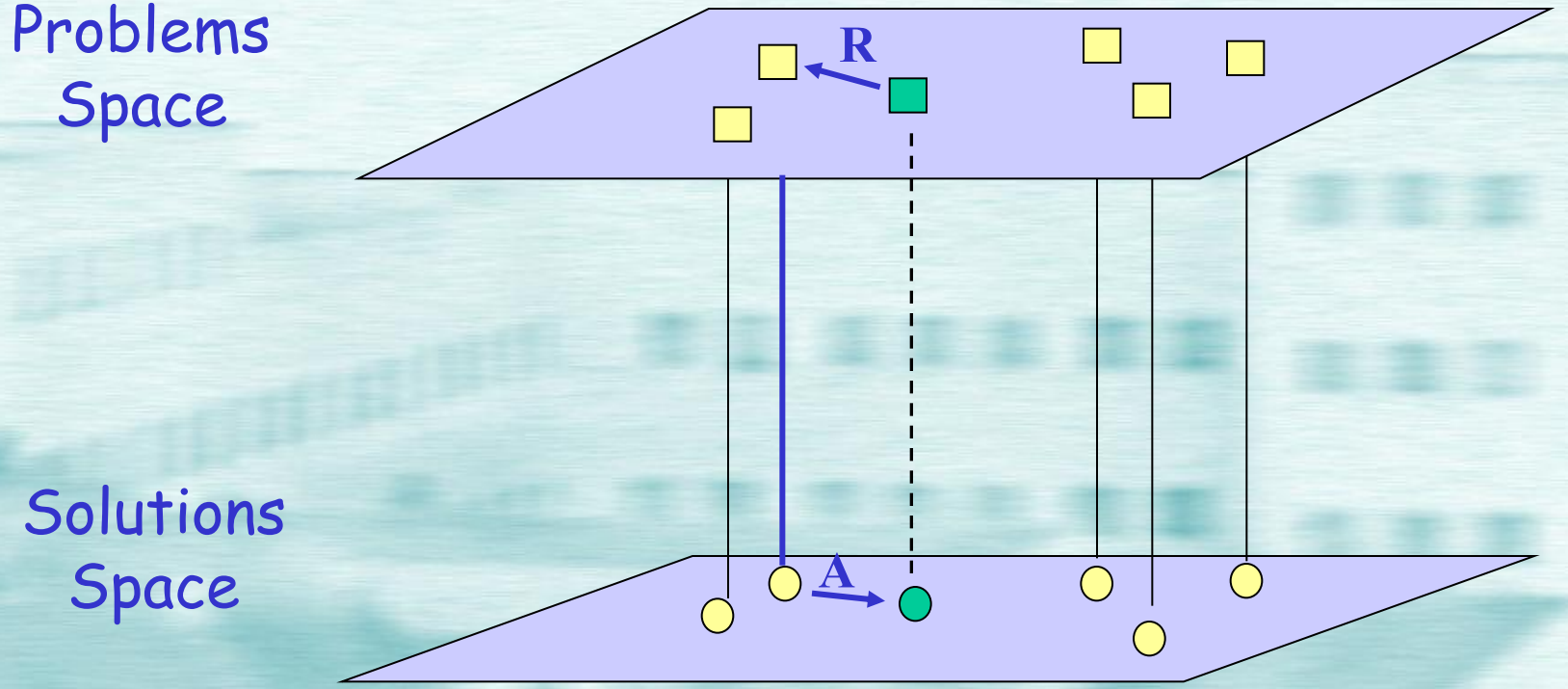
Alternative: Directly use the knowledge implicit in examples from recordings of human performances:

Imitative Expressive performance using CBR:

- SaxEx: Arcos, L. de Mantaras, and Serra
- TempoExpress: Grachten, Arcos, and L. de Mantaras

About Case-Based Reasoning

- Problems are solved *adapting* solutions of *similar* previously solved problems



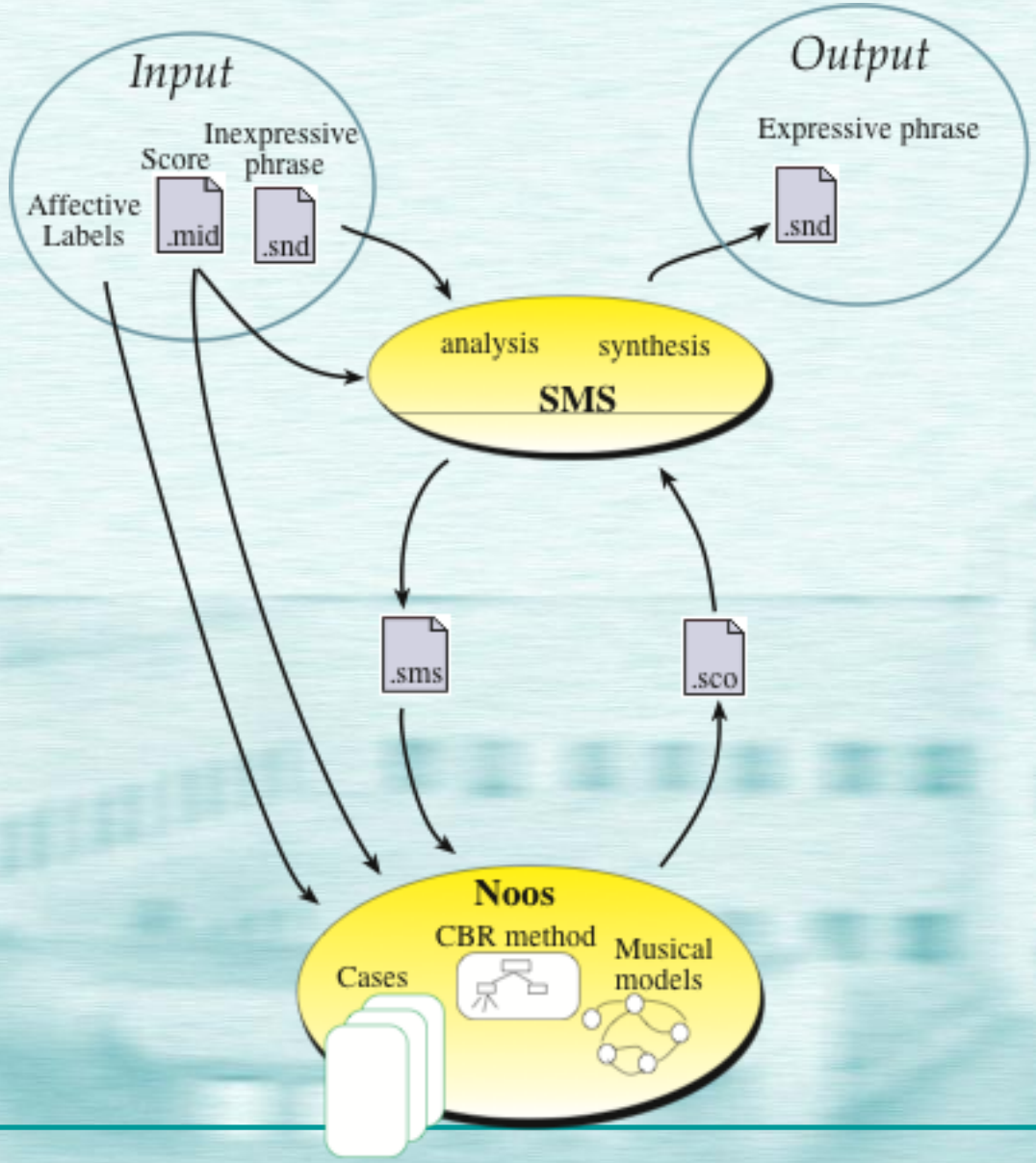
SaxEx

Goal:

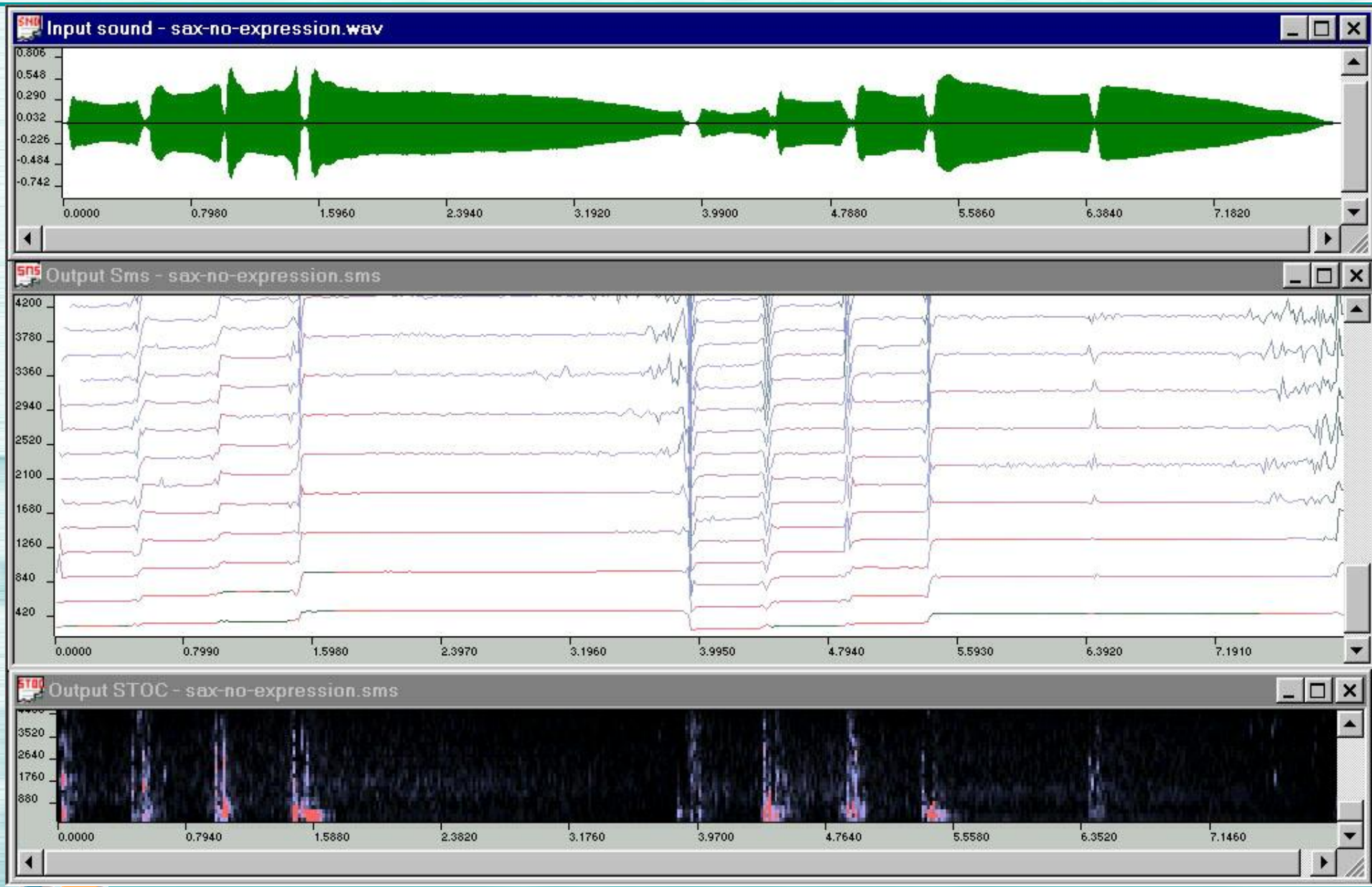
- Transforming, through reasoning from precedents (via CBR and music knowledge), an inexpressive melody into an expressive one, in the context of jazz ballads.

- These precedents are monophonic examples recorded by a professional jazz saxophonist

SaxEx components



SMS snapshot



Case knowedge components

- Musical Knowledge
  - Narmour's implication-realization music perception model.
  - Lerdahl & Jackendoff's GTTM (metrical structure, time-span reduction & prolongational structure).
- Score
- Performance representation (solution description)

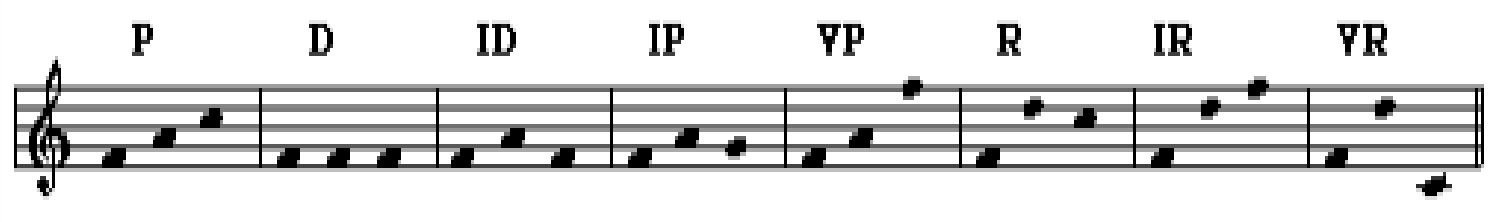
SOLUTION

sound transformation operations on 5 expressive ressources:

-eg: high ***dynamics***, medium ***rubato***, very-legato ***articulation***, medium ***vibrato***, standard ***attack***.

Musical knowledge

- Implication/Realization (I/R) *(E. Narmour, "The analysis and cognition of basic melodic structures", Univ. Chicago Press, 1990)*
  - Basic patterns:

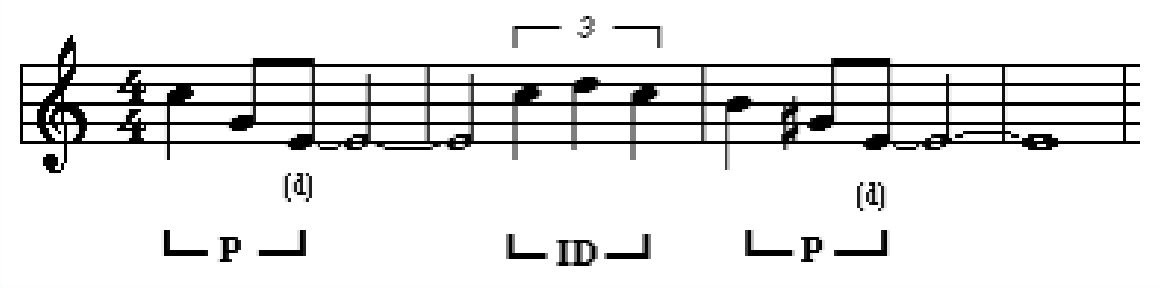


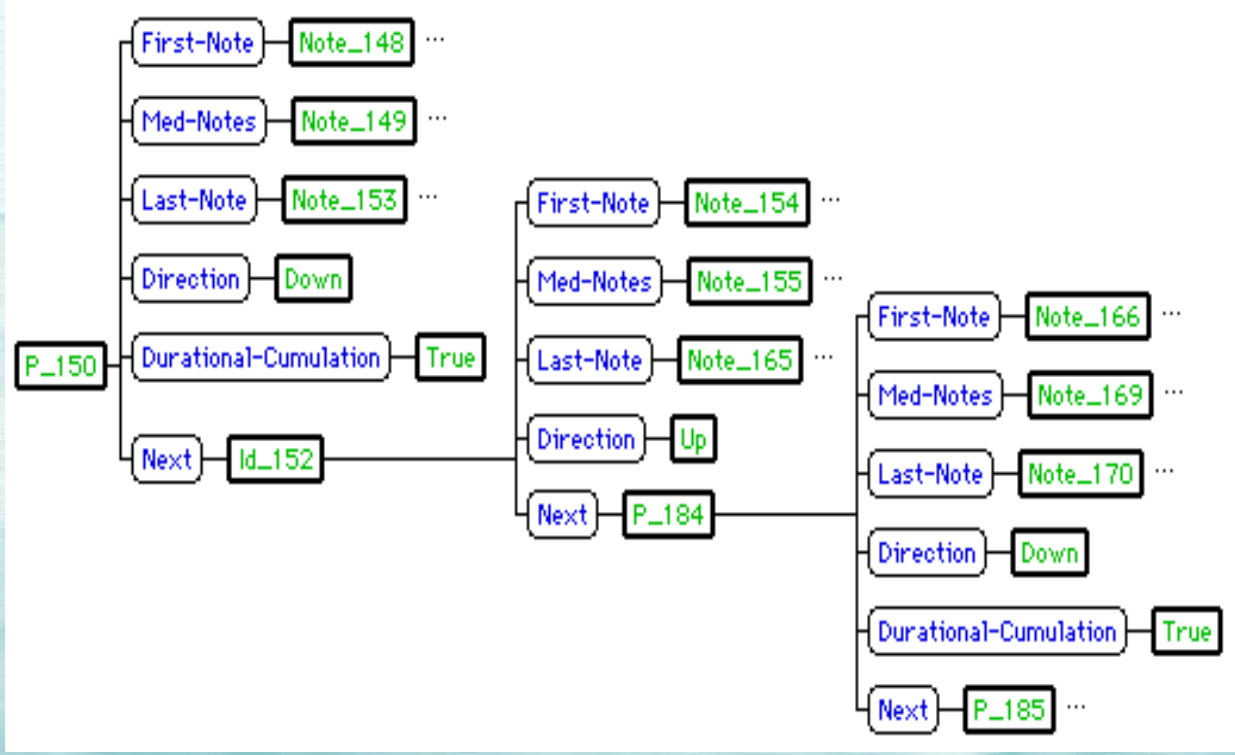
  - Melodic direction, durational cumulation
- GTTM theory *(F. Lerdahl & R. Jackendoff, "A Generative Theory of Tonal Music", MIT Press 1996)*

Music is built from notes and rules that assemble the notes into sequence and organize them into three hierarchical structures:

- Metrical structure (metrical strength of notes)
- Time-span reduction (relative importance of notes within phrases or sub-phrases)
- Prolongational reduction (tensions, relaxations)

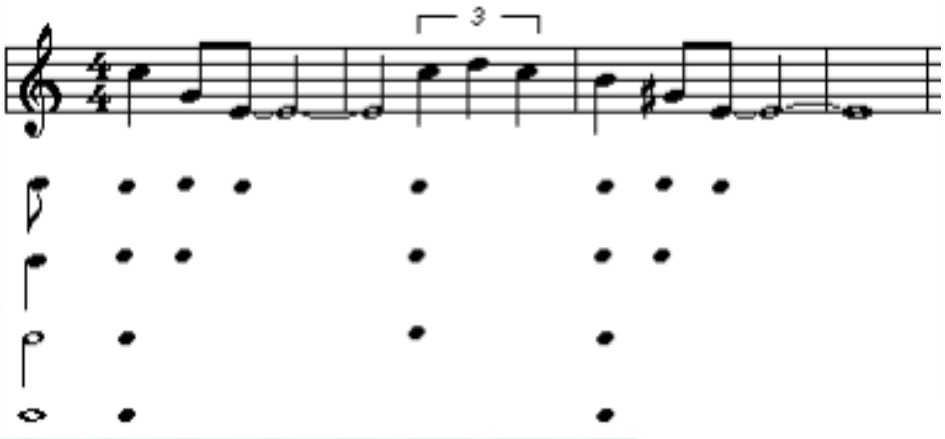
Musical Knowledge: I/R Model



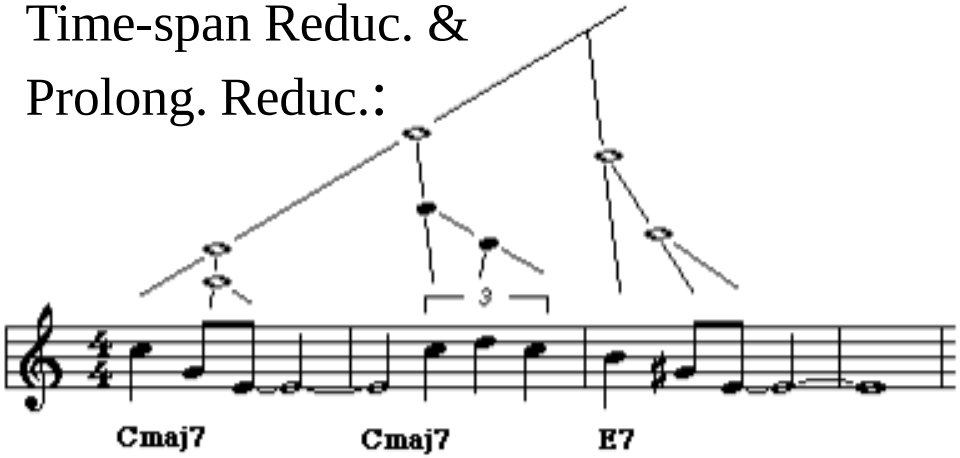


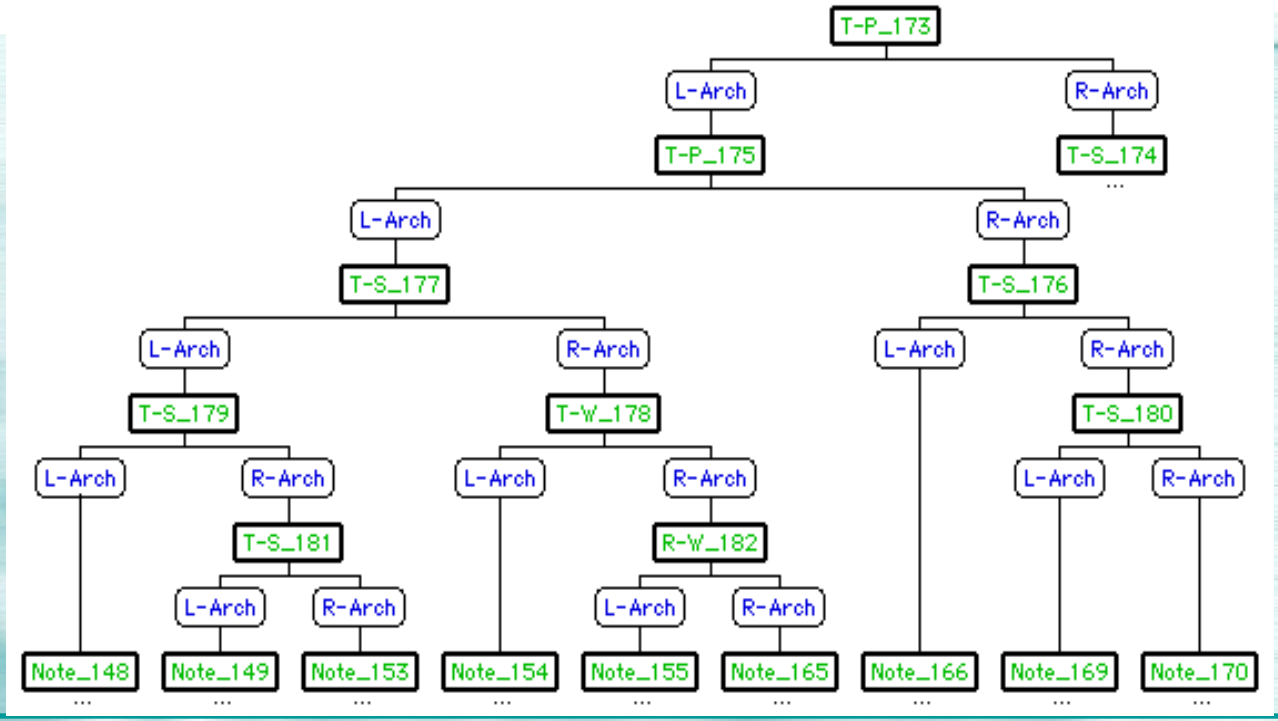
Musical Knowledge GTTM Theory

Metrical strength:

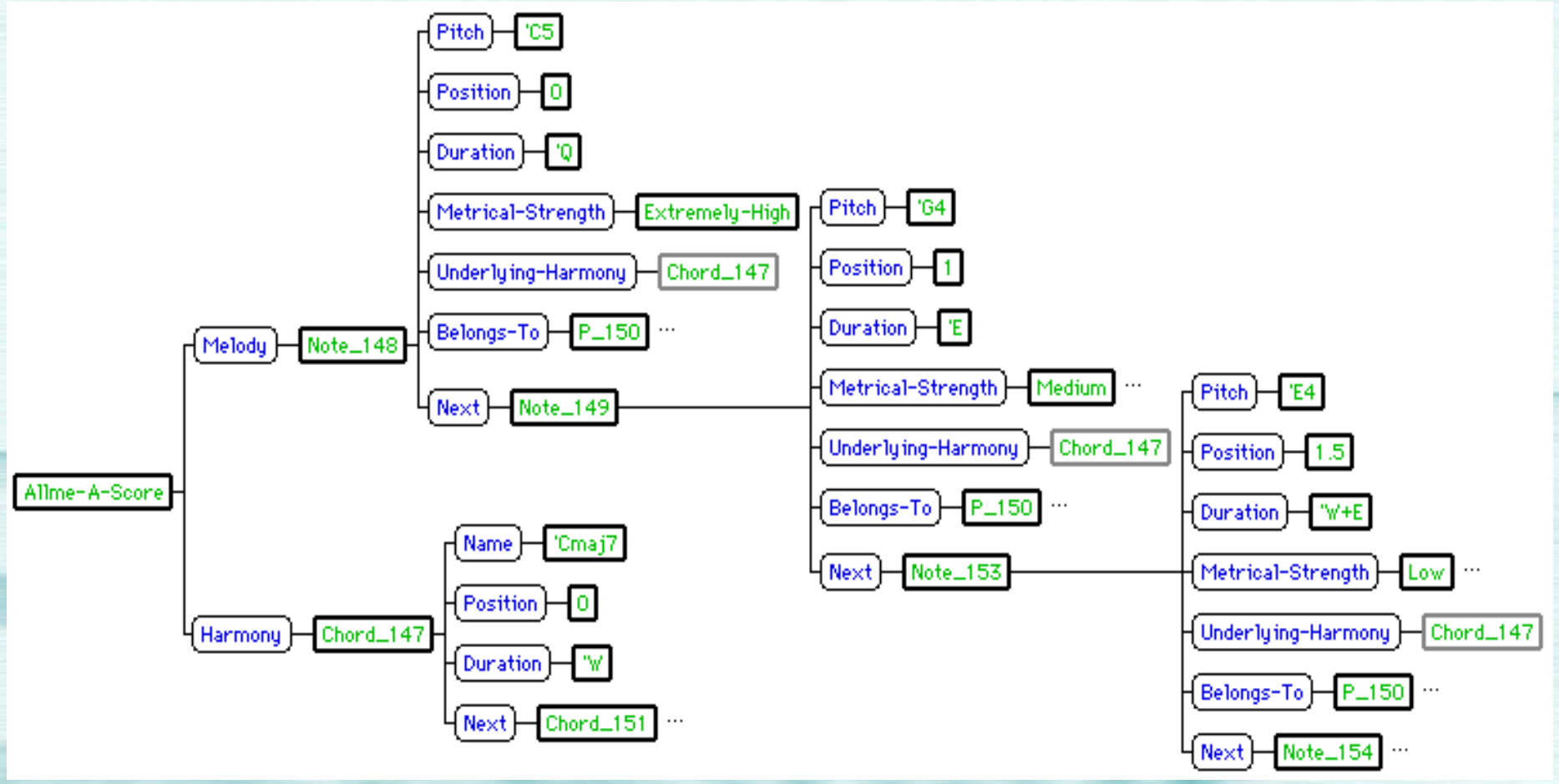


Time-span Reduc. &
Prolong. Reduc.:

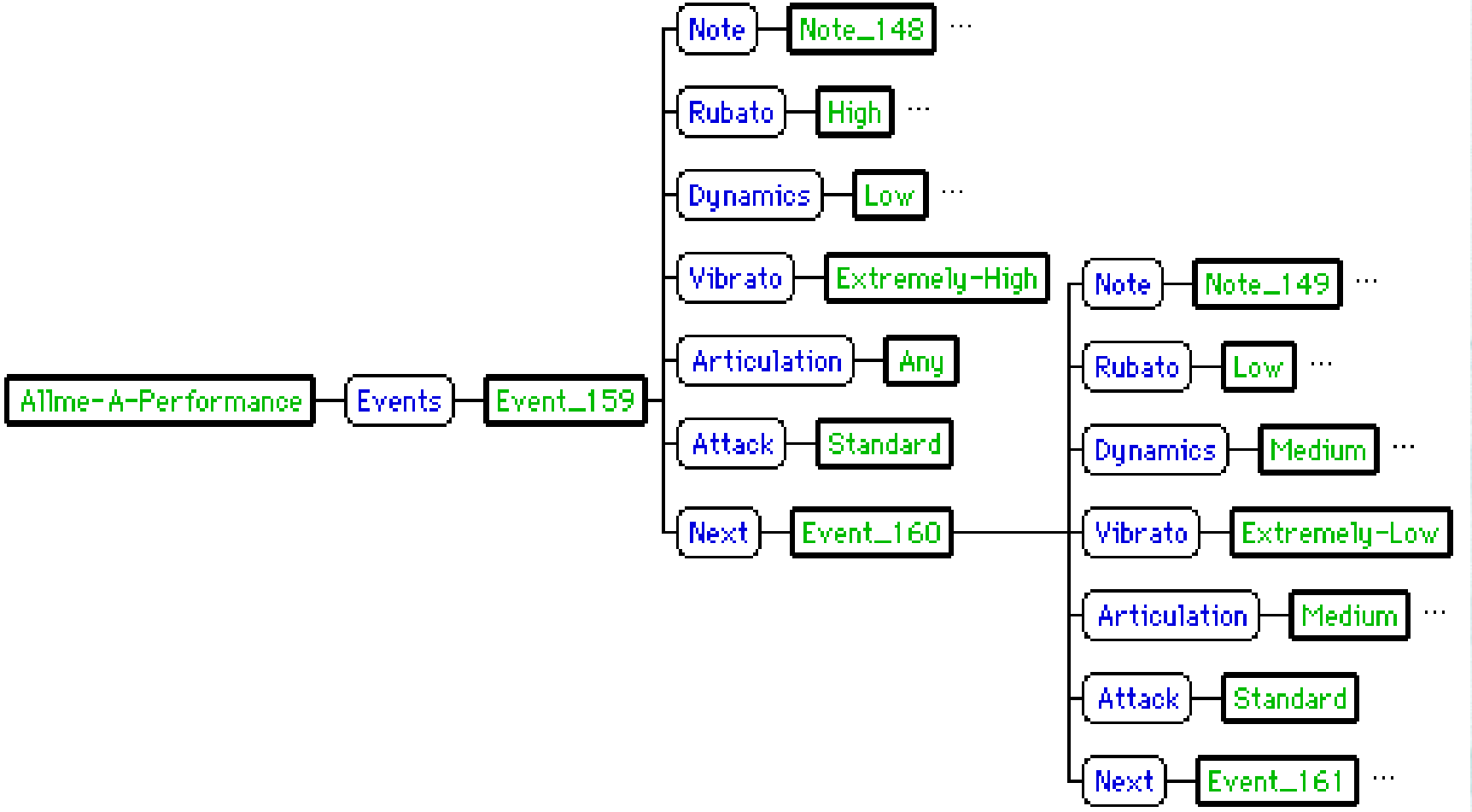




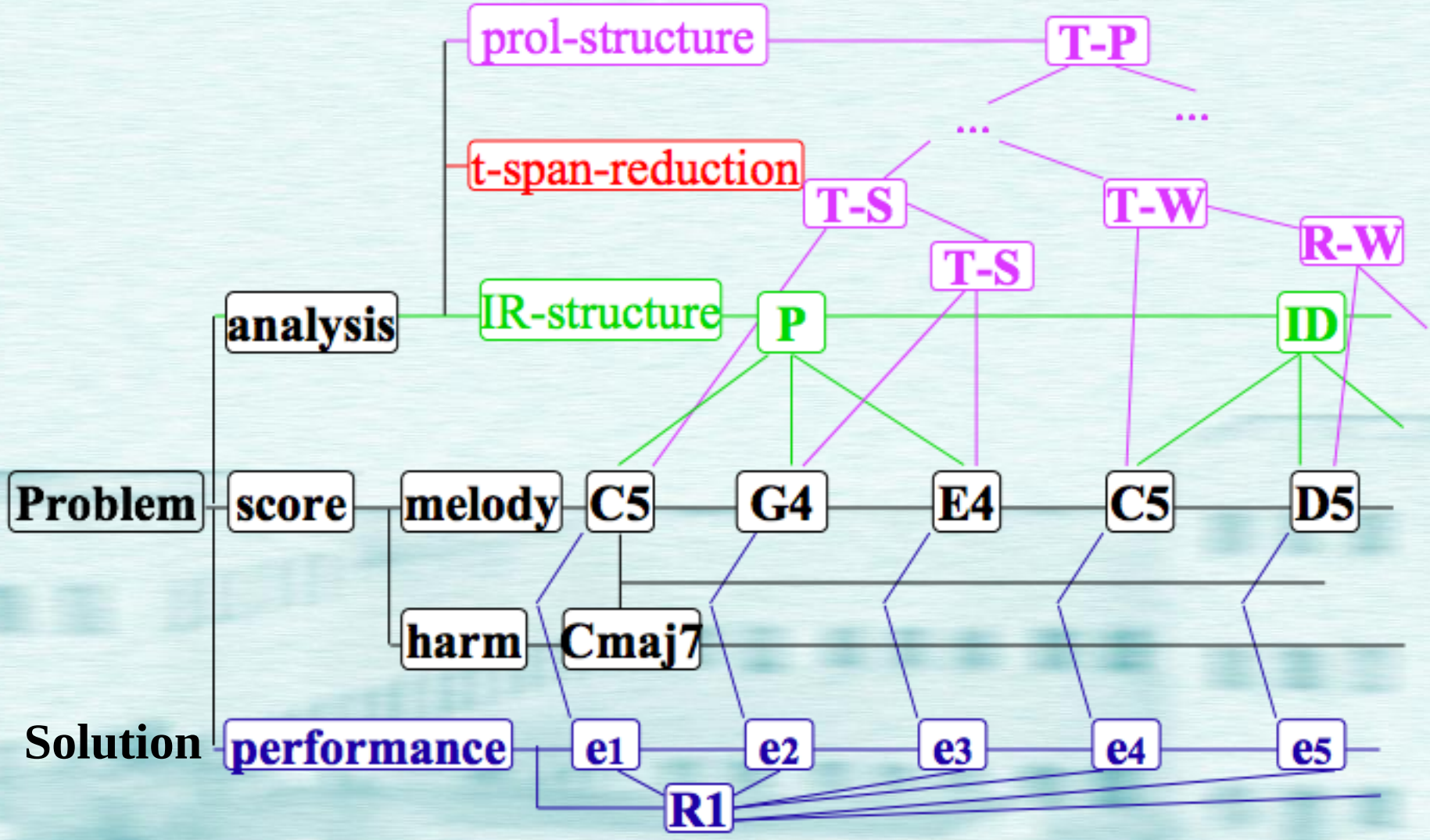
Score


All-me-A-Score
Melody
Note_148
Pitch
'C5
Position
0
Duration
'Q
Metrical-Strength
Extremely-High
Underlying-Harmony
Chord_147
Belongs-To
P_150
Next
Note_149
Pitch
'G4
Position
1
Duration
'E
Metrical-Strength
Medium
Underlying-Harmony
Chord_147
Belongs-To
P_150
Next
Note_153
Pitch
'E4
Position
1.5
Duration
'W+E
Metrical-Strength
Low
Underlying-Harmony
Chord_147
Belongs-To
P_150
Next
Note_154
Harmony
Chord_147
Name
'Cmaj7
Position
0
Duration
'W
Next
Chord_151

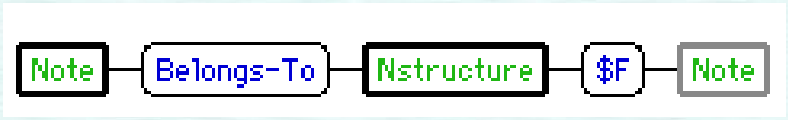
Performance (Case solution)

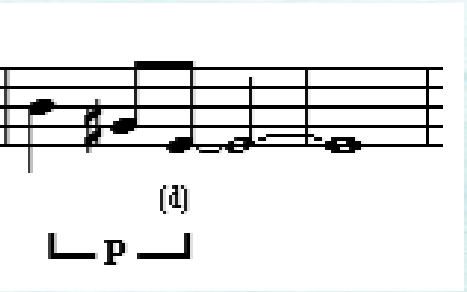


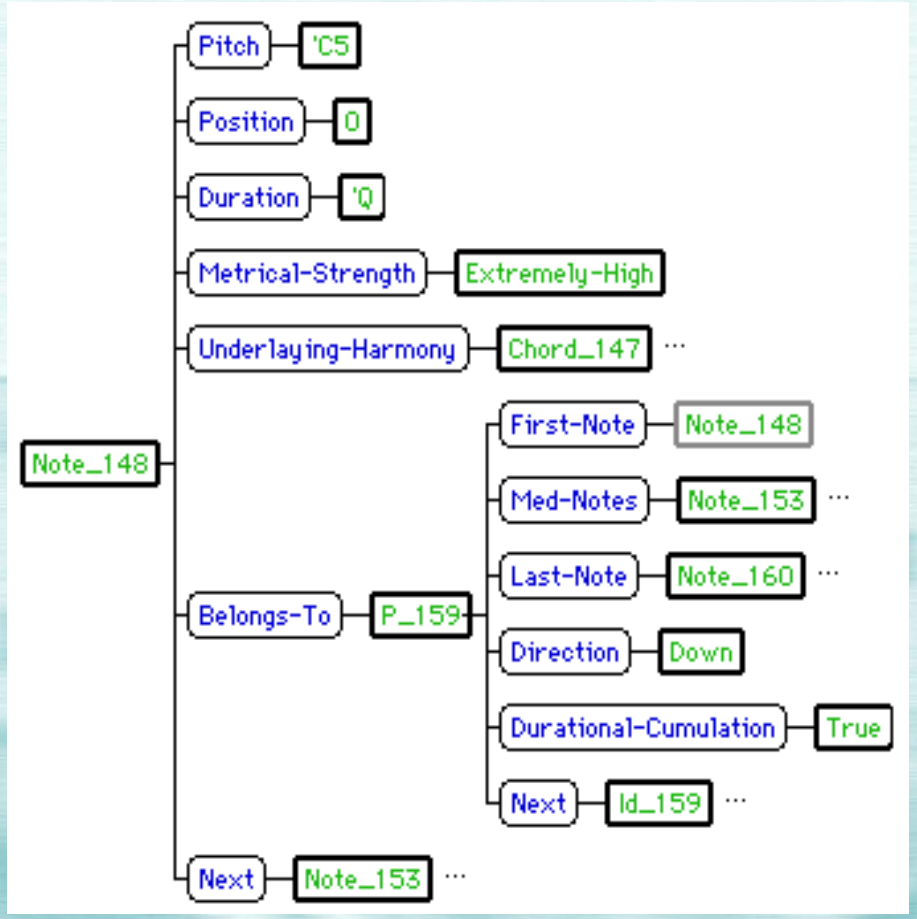
Case components relations



Knowledge-intensive Retrieval





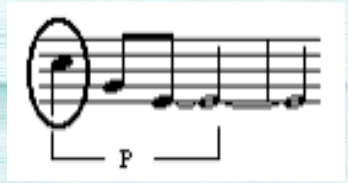


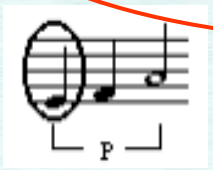
CBR: Simple retrieval example

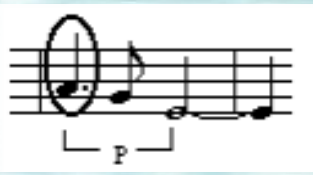
Search
(I/R pattern)

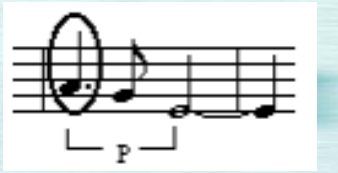
Selection
(direction, durational cumulation, metrical strength,...)

Problem





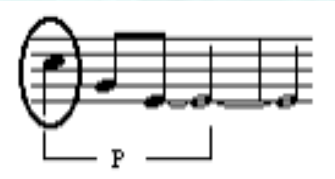




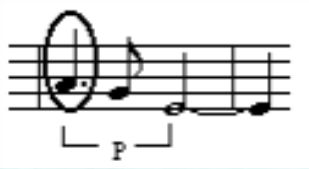
Case-Base Memory

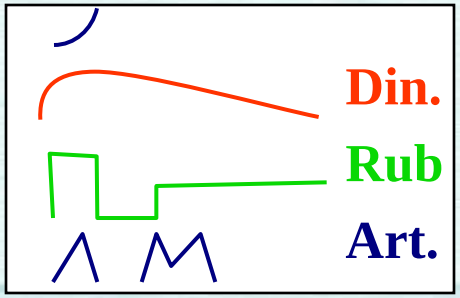
CBR: Reuse (one precedent selected)

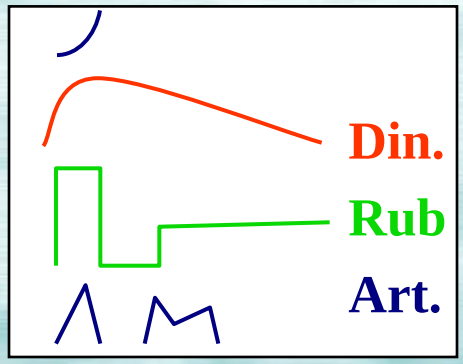
Problem



Precedent Case







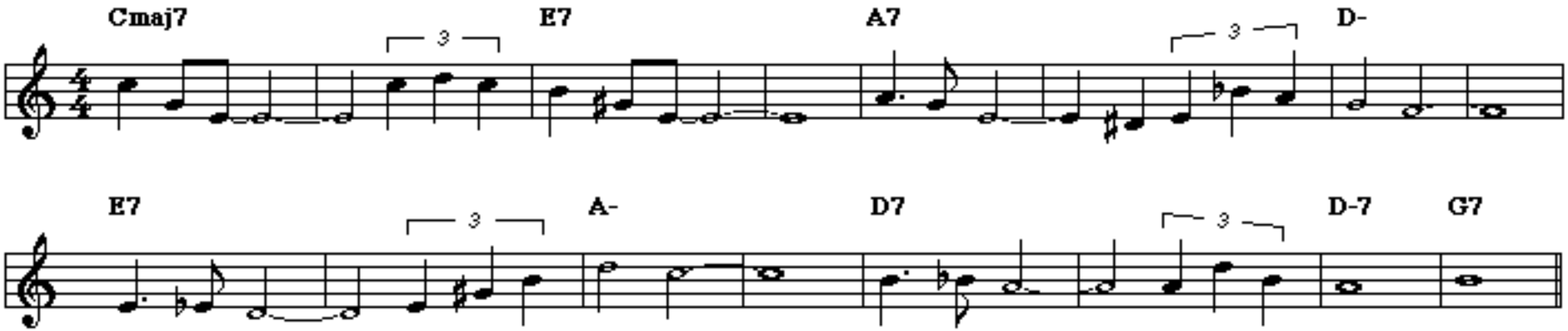
CBR: Reuse criteria

(If more than one precedent is left after the selection step)

- Majority
- Minority
- Continuity
- Diversity
- Random

SaxEx Example

All of me



"Inexpressive" Input phrase

Affective values

SaxEx

"Joyful"

"Sad"

From SaxEx to TempoExpress

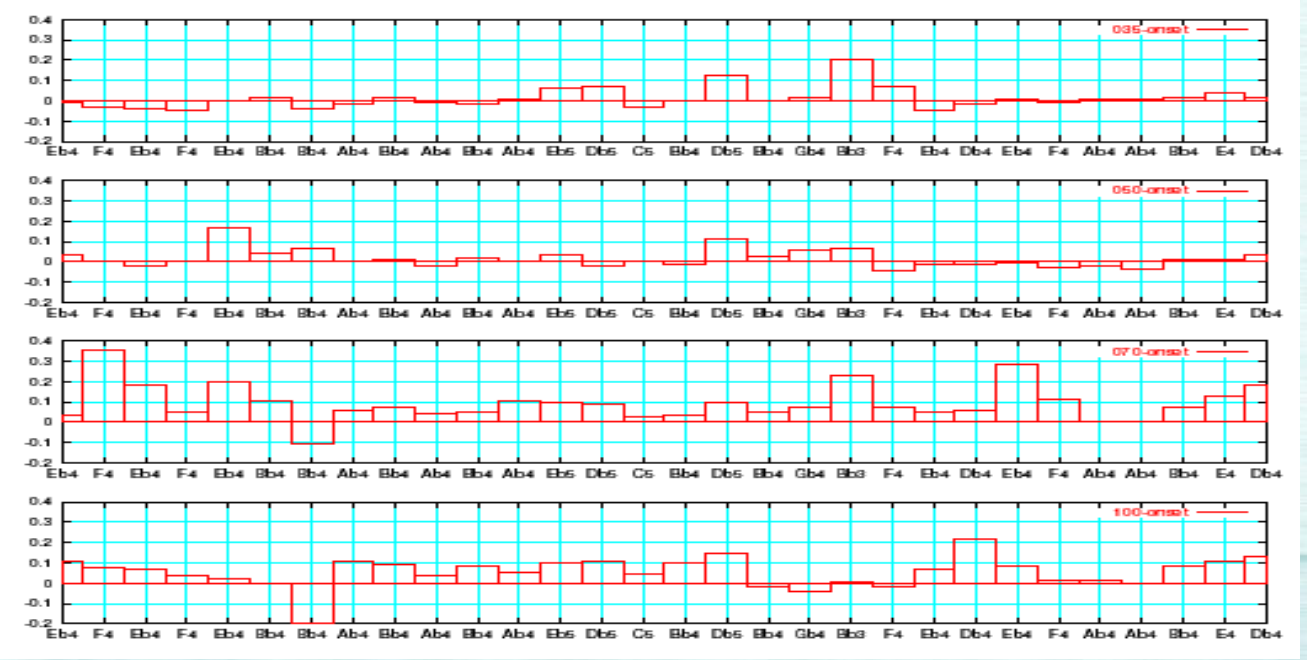
The effect of tempo on expressivity is an important research issue:

-It has been argued (Repp, Psychological Research 56, 1994) that temporal aspects of performance scale uniformly when tempo changes (*Uniform Time Stretching*)
-Counter-evidence to the validity of *Uniform Time Stretching* has been provided (Fridberg et al, Music Perception 19(3), 2002, Timmers et al, Music Perception 20(1), 2002, **Honing** Psychology of Music, 2006)

"Listeners are capable to distinguish, above chance level, between a recording of a performance that has been uniformly time stretched and an original recording"
Why?

Musical explanation: Expressivity is a result of the conception of the music by the performer, and this conception changes with tempo [Desain & Honing, Psychological Research 56, 1994]

Onset deviations at different tempos (Body and Soul A1)



slower tempo

faster tempo

i.e. timing of notes changes w.r.t. beat (Desain & Honing, Psychological Research 56, 1994)

Other expressive phenomena (e.g. ornamentations, consolidations, fragmentations) usually also change as a function of the tempo

TempoExpress

Goal:

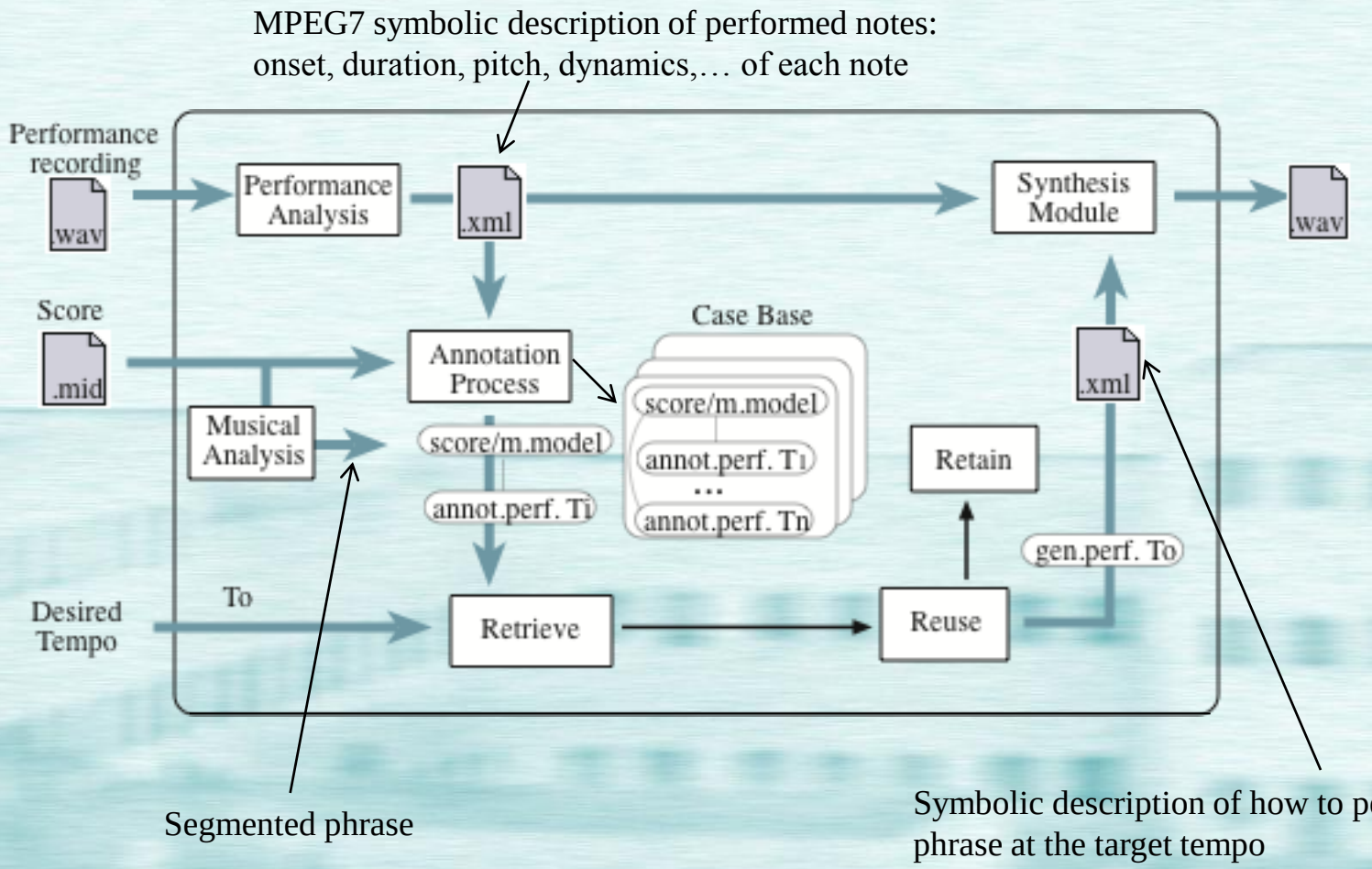
- Changing the original performing tempo of a melody, preserving expressiveness, in the context of jazz standards (not only ballads!).

Application:

Audio editing software for Video / Audio post-production for commercials (video constrains audio)

Tempo-Express

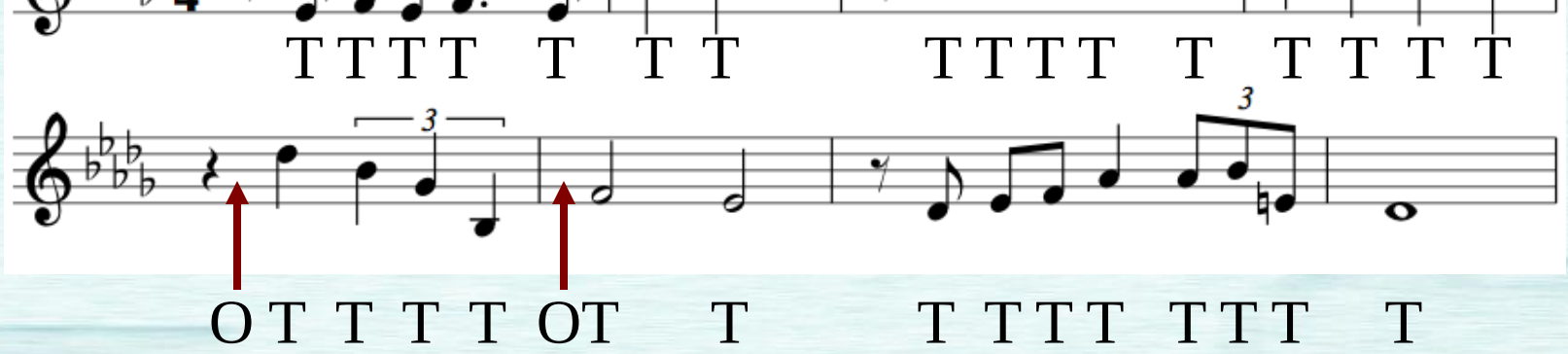
- Architecture



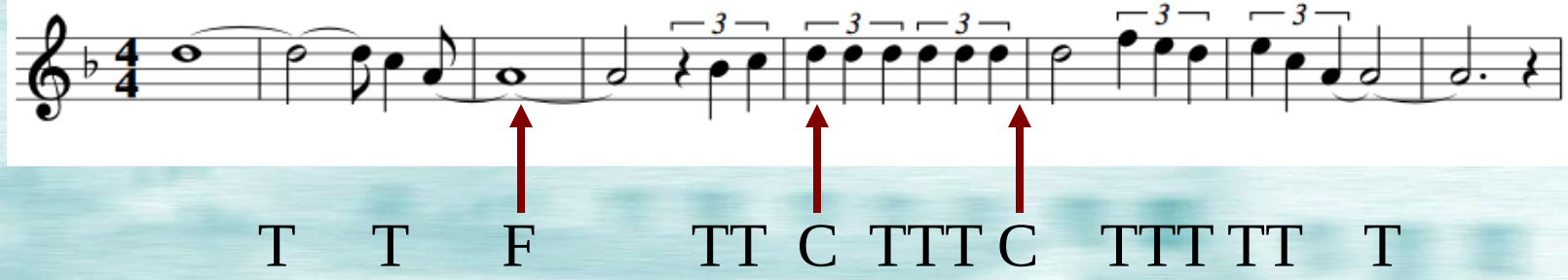
Annotation (examples)

Based on edit distance between the score and the symbolic description of the performed notes

Body and Soul

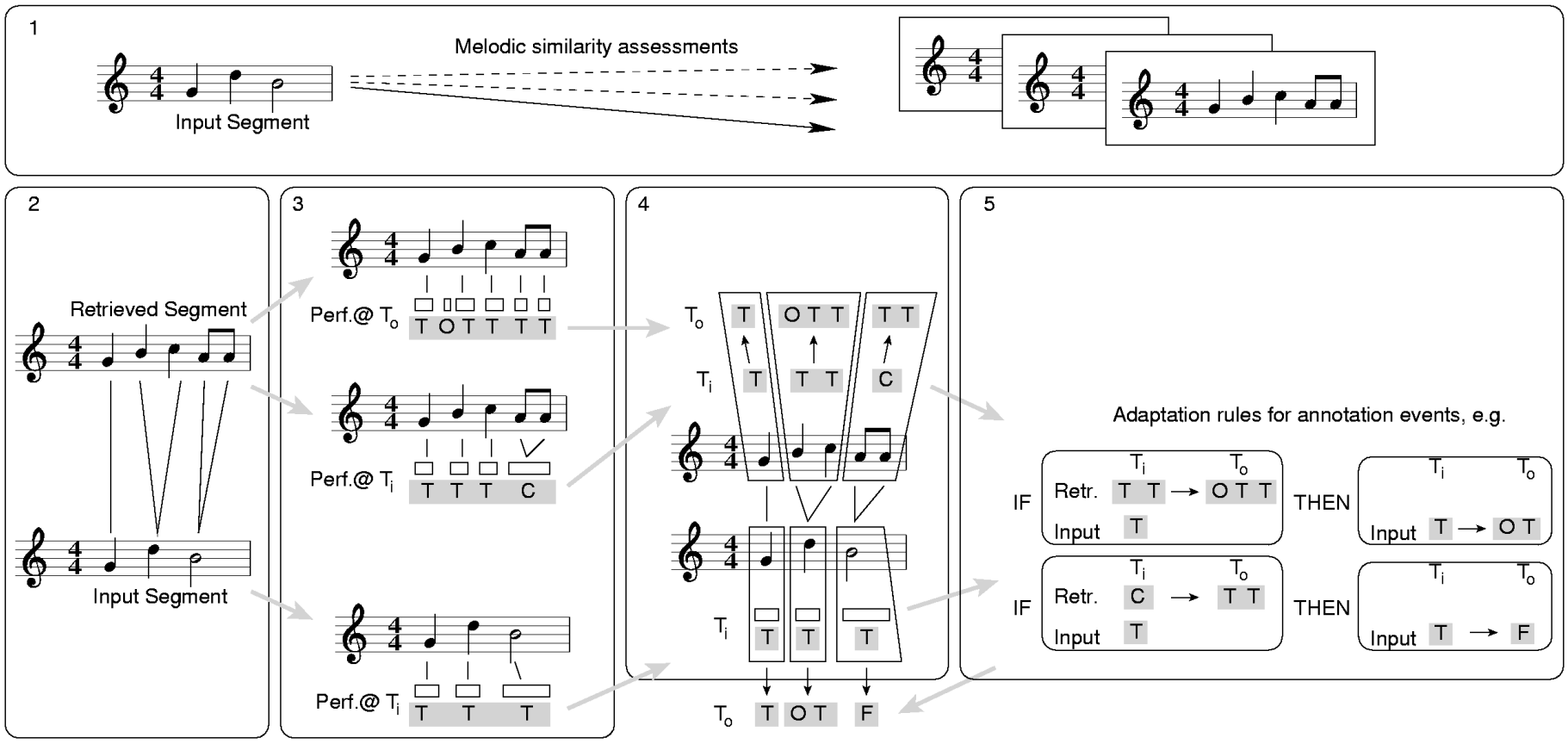


Once I Loved



TempoExpress: Retrieval and Reuse

1) Cases filtering by tempo
2) Cases filtering by edit distance on the I/R description
3) Next:



Experimental comparison with UTS

- 14 different monophonic phrases (64 I/R melodic segments) from jazz standards performed each at 12 different tempos by a professional musician containing a total of over 4000 performed notes
- 8448 tempo-transformation problems in the case base (all the pairwise tempo combinations of performances for each segment: 64X12X11)

UTS versus expressive performance

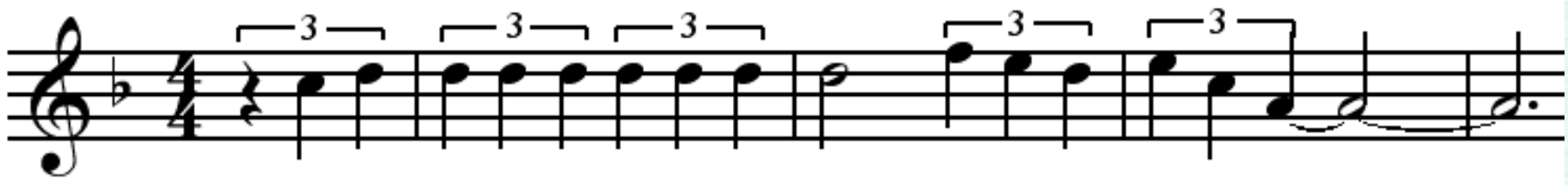
Up Jumped Spring

Tempo 180 bpm

Uniform Time Stretching (90 bpm)

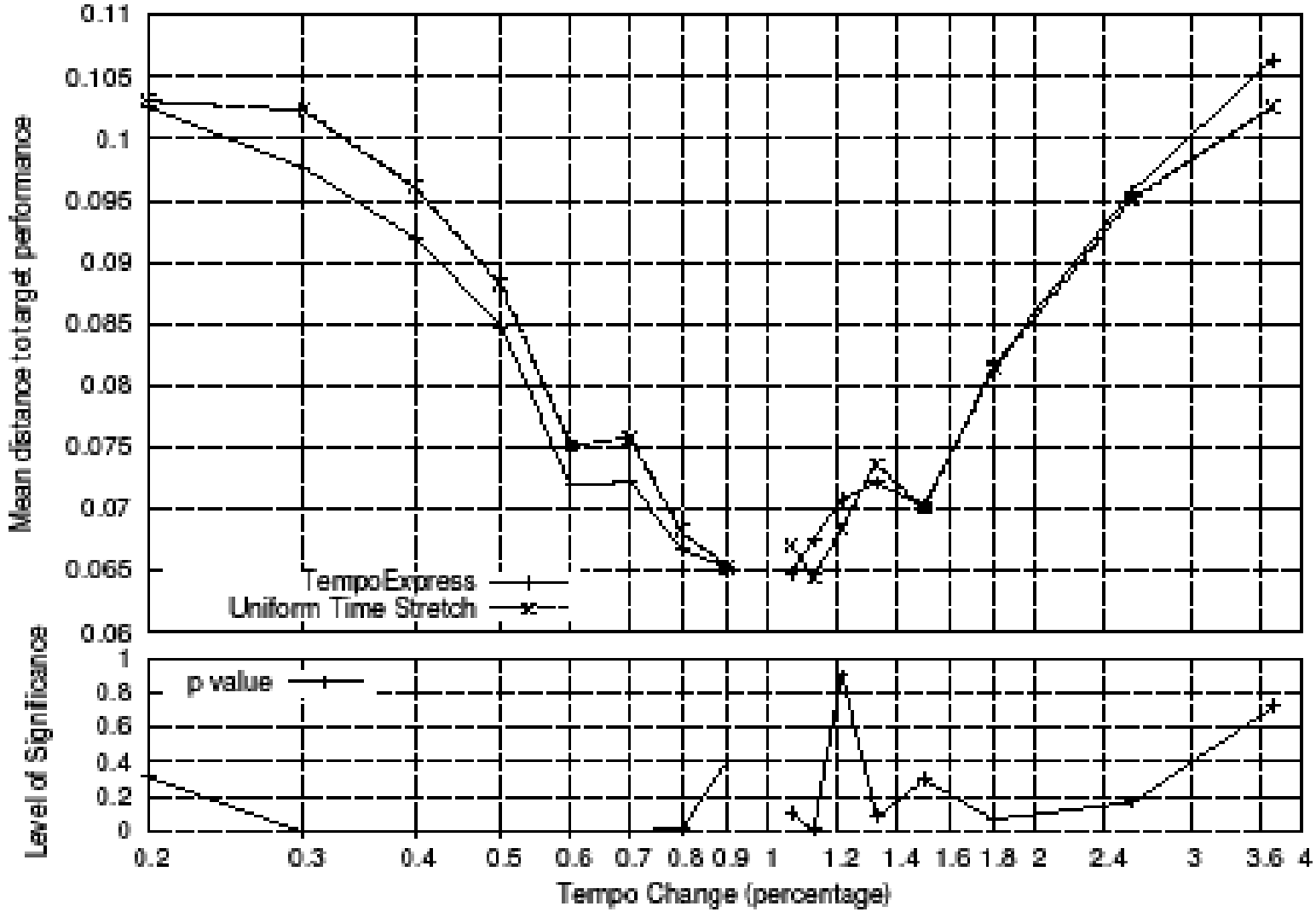
performance (90 bpm)

UTS versus CBR & Performance



55 bpm --------> 100 bpm

- Uniform Time Stretching
- CBR
- Human



TempoExpress vs. UTS as a function of the ratio between target tempo and source tempo. The lower plot shows the probability of incorrectly rejecting the hypothesis (that there is no difference between TempoExpress and UTS) for the Wilcoxon signed-rank test.

Conclusions & Future

- SaxEx successfully imitates human expressive performances
- SaxEx as a pedagogical tool:
  - Users can experiment with the system
  - Helps understanding how to use the different expressive resources
- TempoExpres: an application to automatic tempo changes that outperforms UTS particularly when changing to slower tempos

FUTURE:

- Study expressivity by relating audio and gesture (GuitarLab project)
  - Music is conceived by our brains, ***played through our bodies***, perceived through our sensory organs and then interpreted by our brains

GuitarLab project

- Gesture Acquisition and analysis system: focuses on the gestures of the left hand. Captures both macro-scale changes (i.e. the presence of fingers in frets) and micro-scale changes (i.e. vibrato) in player's movements.

- Audio-Gesture analysis: Obtain an annotated performance relating the gesture to the audio → Towards learning to control "augmented instruments" (the connection between what you do and the sound that comes out)

Collaborators acknowledgment

- Josep-Lluis Arcos

- Maarten Grachten

- Xavier Serra