Modeling Musical Anticipation

From the time of music to the music of time

Arshia Cont

Jointly supervised between
University of California in San Diego (UCSD), Music Department.
Center for Research in Computing and the Arts.
+
University of Paris VI, Ecole Doctorale de l'EDITE,
IRCAM-Centre Pompidou, Equipe IMTR.

Committee in charge

Alain DE CHEVEIGNÉ	ENS, Paris	chair
Shlomo DUBNOV	UCSD, Music	chair
David WESSEL	UC Berkeley, CNMAT	Reporter
Lawrence SAUL	UCSD, CSE	Reporter
Miller S. PUCKETTE	UCSD, Music	Reader
Philippe MANOURY	UCSD, Music	Reader
Jean-Pierre BRIOT	UPMC, LIP6	President

Thesis

Explicit consideration of anticipatory mechanisms, as observed in music cognition, within a computational framework, could

- Address "complex" problems in computer music,
- Reduce complexity of computation and design,
- and Provide access to temporal structures of music information as they unfold in (real) time, for creative applications (computer music).

Motivations

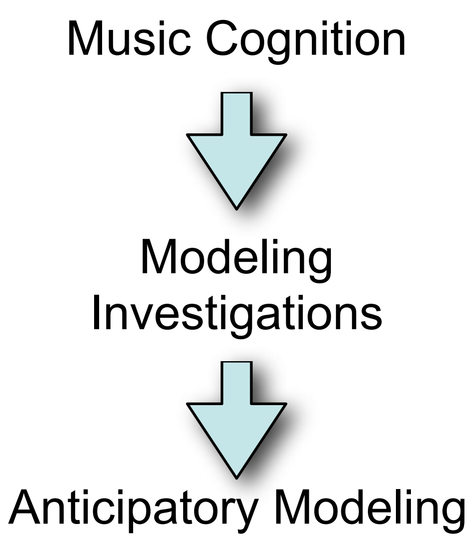
Role of *expectations* in musical experience

- In listening experience
  - Expectations imply mental representations in which our daily musical experience is being examined and updated.
  - Major responsibility for musical emotions
- In musical creativity
  - Meyer (1954): composition = choreography of musical expectations
  - Huron (2006): Demonstrates explicit cases of these "choreographies"
  - Grisey (1987): "A composer's reflections on musical time"
    - *The skin of time*
    - *From the time of music to the music of time...*
- No major consideration for *expectation* in computer music

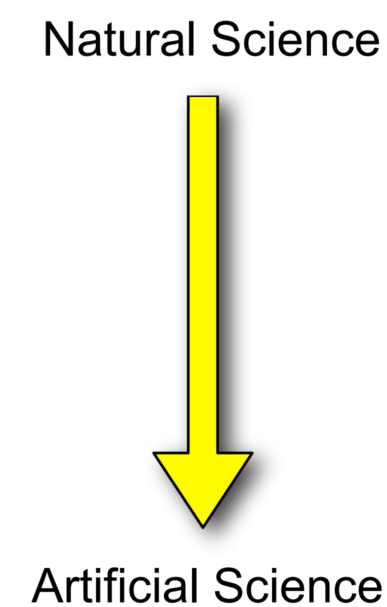
Approach

- A voyage from the domain of *natural science* to the *science of the artificial*
  - Define and clear out the context
  - From *modeling anticipation* to *anticipatory modeling*
  - *Modeling anticipation:* Research in music cognition literature for the study of musical behavior pertaining to *expectations*
  - *Anticipatory modeling:* A cognitive design principle for modeling artificial systems.
- Propose *anticipatory models* addressing three main preoccupations of *musical expectations:*
  - What to expect
  - How to expect
  - When to expect

PART (I)



> "I think that the search for a *universal* answer to the questions raised by musical experience will never be completely fulfilled; but we know that a question raised is often more significant than the answer received. Only a reckless spirit, today, would try to give a total explanation of music, but anyone who would never pose the problem is even more reckless."
>
> *Remembering the future*
> LUCIANO BERIO



From Modeling Anticipation to Anticipatory Modeling

Anticipation, Expectation, Prediction

<ins>*Definitions?*</ins>

- *Prediction*: Act of forecasting based on previously gained knowledge from an environment.
- *Expectation*: "a form of mental or corporeal belief that some event or class of events is likely to happen in the future." (Huron 2006)
- But *anticipation*...
  - Huron (2006): Anticipation is "a sub-product of expectation when the sense of appraisal for the expected future event is high."
  - Bharucha (96) calls this *yearning*.
  - Narmour (90) calls this *implication*.
  - Schmuckler (97): "Expectation is an anticipation of upcoming events based on information from past and present."

Psychology of Musical Expectation

Fact Constitution

1. *Expectations* entail mental representations, whether partial, accurate or fallible.
2. *Expectations* are *learned* through interactions with a surrounding environment *(auditory learning)*
   - The determinant factor for learning auditory phenomena is their stability in the surrounding environment
   - Statistical nature of auditory learning
3. Concurrent and Competitive Representations
   - Listeners appear to possess multiple representations for the same phenomena *(concurrency)*
   - *Expectations* are differentially favored depending on their predictive success *(competitive)*
4. *Expectations* lead to predictions, which by themselves evoke *actions* (physiological, mental or physical)
   - *Expectations* are always coupled with their consequent *actions*
   - In theory and also in practice... .

What is Anticipation?

We study the *activism* aspect of *expectation* under the term *Anticipation:*

> **Definition** *Anticipation* is an *action*, that a system takes as a result of *prediction*, based on current belief or expectations, including actions on its own internal state or belief.

- Implications:
  - *Expectations* demand constant *adaptation* and *interaction* with an environment
  - *Anticipation* is an activity of coupling with the environment.
  - *Anticipation* is not a process in the brain, but a kind of skillful activity and a mode of exploration of the environment.

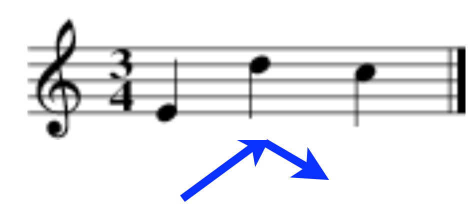
Modeling Investigations

- What is *modeling*?
  - In *natural sciences,* by modeling musical expectations researchers aim at
    - assessing a theory regarding one among many aspects of the psychology of musical expectations
- Things to be aware of...

  (A) Imperfect Heuristics
    - Recall "fact 1": Consequent mental representations out of auditory learning can be fallible.
      - We are selective and imperfect learners, constrained by all problems of induction.
    - Biological goal of expectation vs. Musical goal of expectation
      - Biological goal: Faulty heuristics ⇒ Potential danger
      - Musical goal: Every danger is... welcome!
    - Our senses are *biased* through the world!

Modeling Investigations

- Things to be aware of...
  (B) Naive Realism
  - Naive Realists consider senses as unbiased windows through the real world.
  - A sobering point for music theorists!
    - Considering that the structures seen in notations are the ones we experience, and what is experienced is what is seen in the notation.
    - The *majority* of music theoretic models of expectation undergo naive realism.
  - Example: The story of *post-skip reversal*



    - Common perception pattern among western-listeners
    - Adopted by many music theory researchers as an “rule” underlying melodic expectation (Narmour, Lerdahl, Margulis).
    - Not true!
      - Von Hipple and Huron (2000) show that post-skip reversal in music scores is the result of a less exciting phenomena.
      - Except in the music of *Giovanni Palestrina*, the initiator of the rule.

So... Should we “model” anticipation?

Revise the question!

Modeling Investigations

Reactive Frameworks

- Computational models based on *causality*
  - Action is the result of a *belief* based on the past and present
  - A universal framework:

    Given any mode of system behavior which can be described sufficiently accurately, there is a purely reactive system which exhibits precisely this behavior
  - Practical problems for modeling cognitive behavior
    - Representations are fallible! (no accurate description)
    - Not all forms of cognitive interactions can be transcribed or assumed as disposed
    - would not necessarily generalize to unknown situations.

Anticipatory Modeling

Definition

- *Anticipatory behavior:*
  A process or behavior that does not only depend on the past and present but also on predictions, expectations or beliefs about the future.
- Anticipatory Modeling is the *design process* for anticipatory systems:

> **Definition** An *Anticipatory System* is a system containing a predictive model of its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a later instant.

- In contrast to *Modeling Anticipation*:
  - No attempt to provide a universal framework for anticipatory behavior, but to provide models that anticipate.
  - Considers *Anticipation* as the fore-front concept in cognitive system design to achieve *complex* systems.

Anticipatory Modeling

Modeling Anticipation vs. Anticipatory Modeling

- *Modeling Anticipation*: An effort to explain musical behavior, as in *natural sciences*.
- *Anticipatory Modeling* :
  - A design process that addresses anticipatory behavior observed in music.
  - Complexity a result of adaptation... .
  - To avoid naive realism, and problems of imperfect heuristic

PART (II)

What to Expect

Music Information Geometry

Motivations

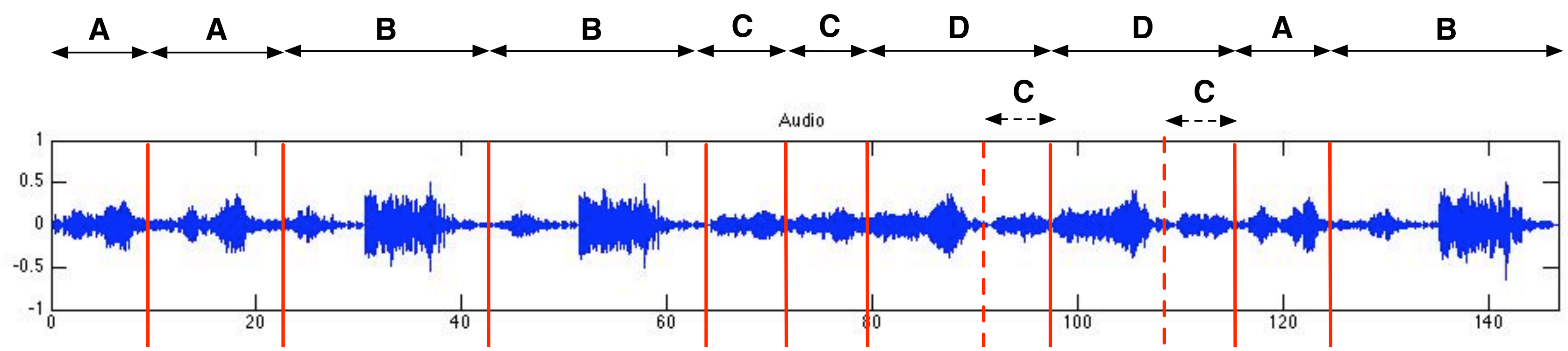
- To represent *relevant* and *stable* part of information arriving from the environment
- Classical Information Theory
  - Few answers to the *representation* and *fidelity* concerns
- Advances in Machine Learning
  - MIR Techniques based on measures of *self-similarity*
  - Lack of consideration for *time*
- Bring the two literatures together

⇨ *Information Geometry*

Information Geometry

Motivation

Consider the problem of extracting structures of musical information:



- Extracting this kind of information from a *symbolic score* is trivial, but a difficult problem in the *signal* domain.
  - Specially with no a priori knowledge of music
  - Worse in real time!
- Goal: To make this possible and more...
  - To obtain methods of access to structures of music signals
  - A general framework to fill in the following gap for musical applications:

signal ⟷ symbol

Information Geometry

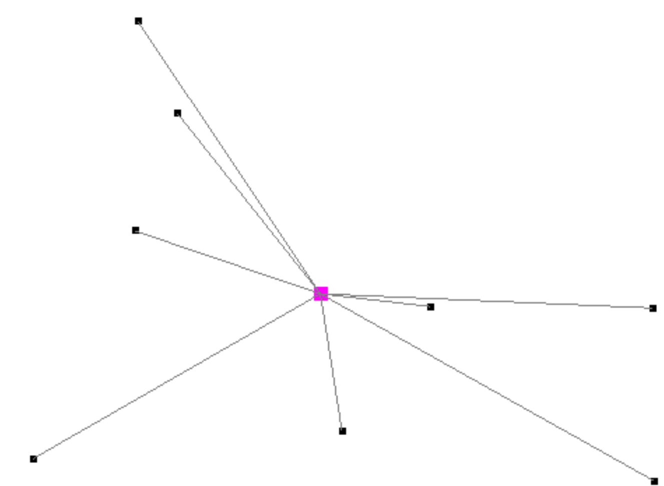
Intuition

- Consider the following geometry:
  - *Points* are probability distributions $p(x, \xi)$ (instead of *dots*)
  - *Distance* between two points is some measure of *information* between them
- Welcome to the world of *Information Geometry!*
  - Geometric manifolds with *information metrics* on *probability space*
    - Marriage of *Differential Geometry*, *Information Theory*, and *Machine Learning*
  - Considering probabilistic representations as well-behaved geometrical objects, with intuitive geometric properties
    - Spheres, lines (geodesics), rotations, volumes, lengths, angles, etc.
- Getting real...
  - Riemannian Manifolds over probability spaces with *Fisher Information* measure
  - Characterized by the type of employed *distance* (called *divergences)*
  - Our interest, canonical elements:
    - Space of *exponential distributions*
    - with *Bregman divergences*
      - Bijection between the two

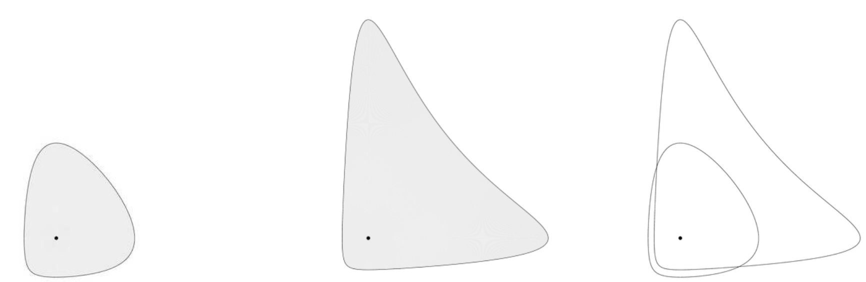
Information Geometry

Elements of Bregman Geometry

- *Bregman Centroids*
  - Significant property `(Thm 4.1)`
    - The "right type" centroid is independent of the choice of Bregman divergence and is equal to the mean:



- *Bregman Balls*
  - In analogy to Euclidean geometry, we can define balls using Bregman divs, centered at μ_k with radius R_k



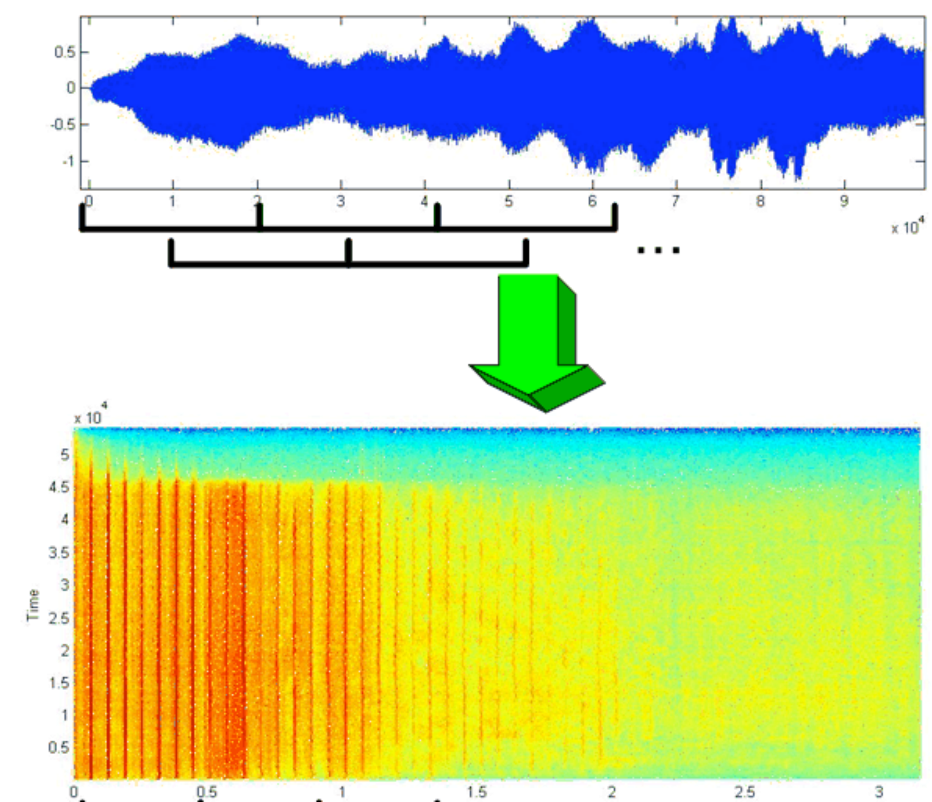
(a) (b) (c)

- *Bregman Information* of a random variable *X*
  - Defined as the expectation over divergences of all *points* from the centroid
  - Special cases: *variance, mutual information*

Music Information Geometry

General Framework

- Points = time domain windows of audio signal X_t, represented by their *frequency distributions* $S_t(\omega)$
  - Arriving incrementally / in real time
  - Corresponding to normalized log-scale Fourier transform amplitudes
  - Mapped to *Multinomial points* in the information geometry (one-to-one)
    - Corresponding Bregman divergence is *Kullback-Leibler divergence*
    - Therefore, Bregman Information is equivalent to *mutual information*



Goal: To capture, represent and qualify the information structure of audio data streams.

Music Information Geometry

Approach

- Do *not* formalize information content!
- Control *changes* of information content instead
  - Using some *metric d*, that gives rise to the notion of similarity:

> **Definition** Two entities $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1 \in \mathcal{X}$ are assumed to be *similar* if the information gain by passing from one representation to other is zero or minimal; quantified by $d_X(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1) < \epsilon$ which depends not on the signal itself, but on the probability functions $p_X(x; \boldsymbol{\theta}_0)$ and $p_X(\boldsymbol{x}; \boldsymbol{\theta}_1)$.

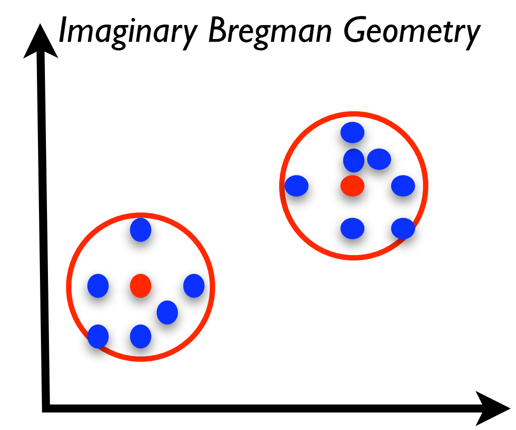
- Candidates for *d(.,.)*:
  - Information Rate (IR) of Dubnov (2005,2008)
    - *Data-IR*: For *stationary* signals
    - *Model-IR*: Between sub-sets of *quasi-stationary* signals

Music Information Geometry

Appoach

- Proposal: Use the bijected Bregman divergence of the information geometry of audio data streams
- Data-IR:
  - Is proven (mathematically) to be equal to *Bregman Information*
- Model-IR:
  - Requires *segmenting* audio stream into chunks.
  - Proposal:

Definition Given a dual structure manifold $(\mathcal{S}, g, \Delta^D, \Delta^{D^*})$ derived on a regular exponential family formed on data-stream X_k, a *model* θ_i consist of a set $\mathcal{X}_i = \{\boldsymbol{x}_k | k \in \mathcal{N}, \mathcal{N} \subset \mathbb{N}\}$ that forms a *Bregman Ball* $B_r(\boldsymbol{\mu}_i, R_i)$ with center $\boldsymbol{\mu}_i$ and radius R_i.

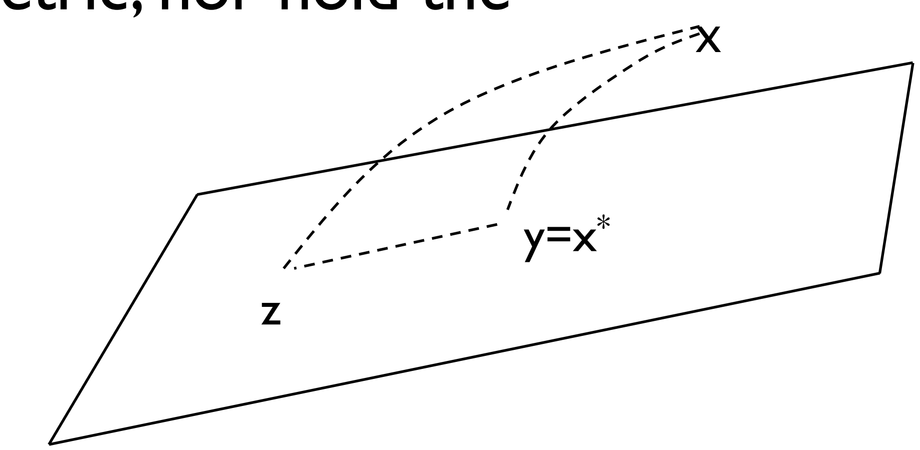


Music Information Geometry

From Divergence to Similarity Metric

- Further requirements for *d*:
  - *symmetric* $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$
  - and to hold the *triangular inequality* $d(\boldsymbol{x}, \boldsymbol{y}) \leq d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y})$

  to obtain equivalent classes.
- Problem: Bregman divergences are neither symmetric, nor hold the triangular inequality!
- Solutions: (Nielsen and Nock, 2007)



  a. Triangular inequality hold IFF *y* is the geometric projection of *x* onto the tangent plane passing through *zy*.

  b. In our geometry, the notions of max. likelihood and projection are ***equivalent***! `(Proposition 4.1)`

  c. Symmetrize Bregman divergence using a max. likelihood formulation!

We can approach both notions of symmetry and triangular inequality.

Music Information Geometry

Model Formation

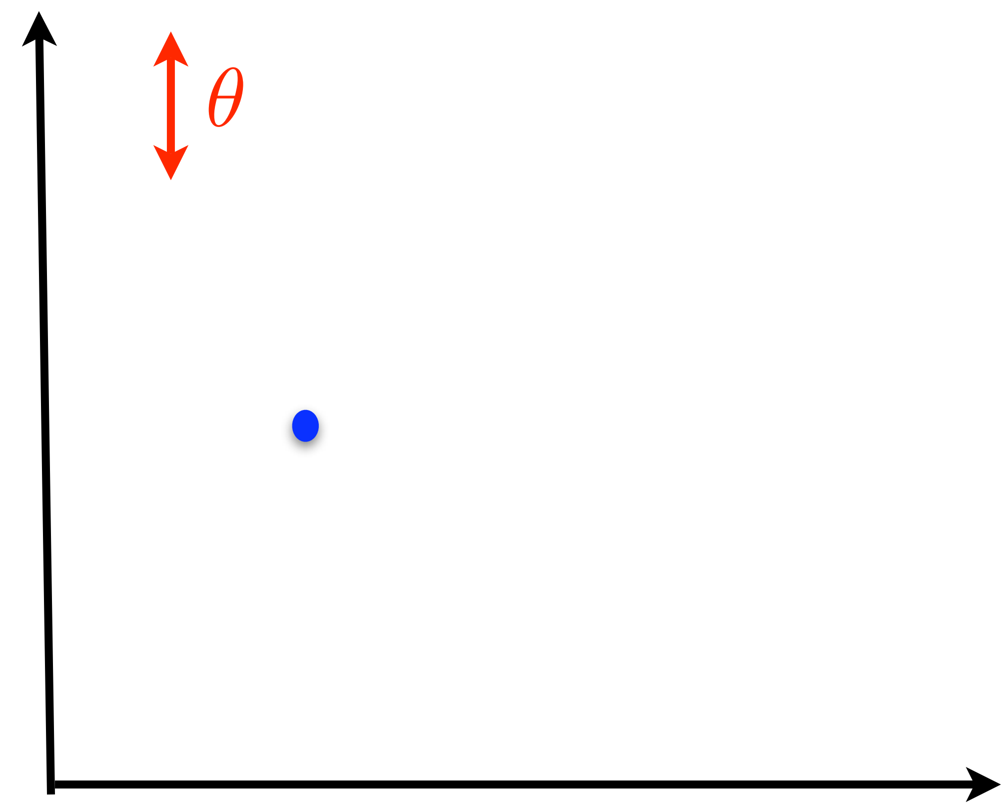
- Incrementally segment information regions into quasi-stationary chunks
  - A model is a Bregman ball whose information radius reveals the maximum distance in terms of *mutual information* within the ball.
  - Detect balls with *jumps* in information distance between a new point and a forming ball
  - Assume a fixed *information radius R* to detect jumps and account for continuity of information change.
  - *Computationally cheap:* Only comparing with the last

...Simulation...

Music Information Geometry

Incremental Segmentation

Imaginary Bregman Geometry



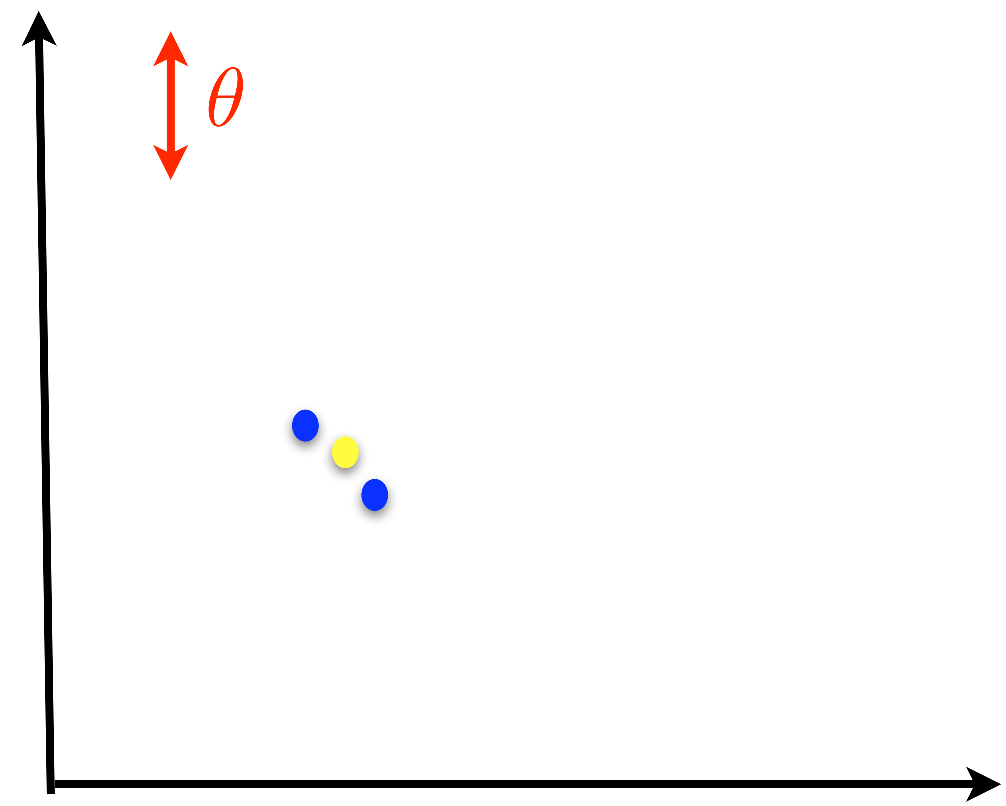
: Moving Centroid

REALTIME Scheduler
Time = 1
No event!

Music Information Geometry

Incremental Segmentation

Imaginary Bregman Geometry



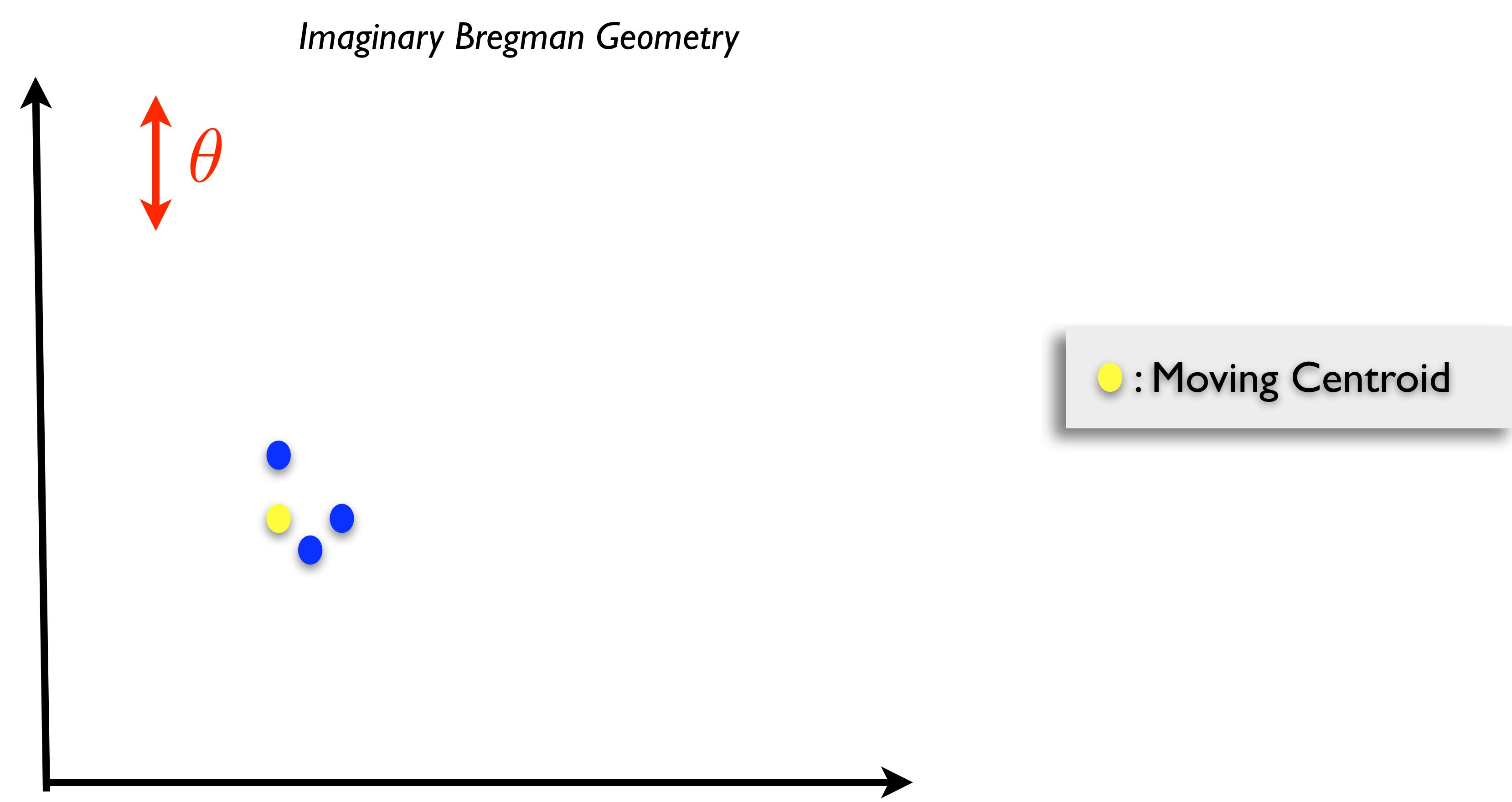
: Moving Centroid

REALTIME Scheduler

Time = **2**

Music Information Geometry

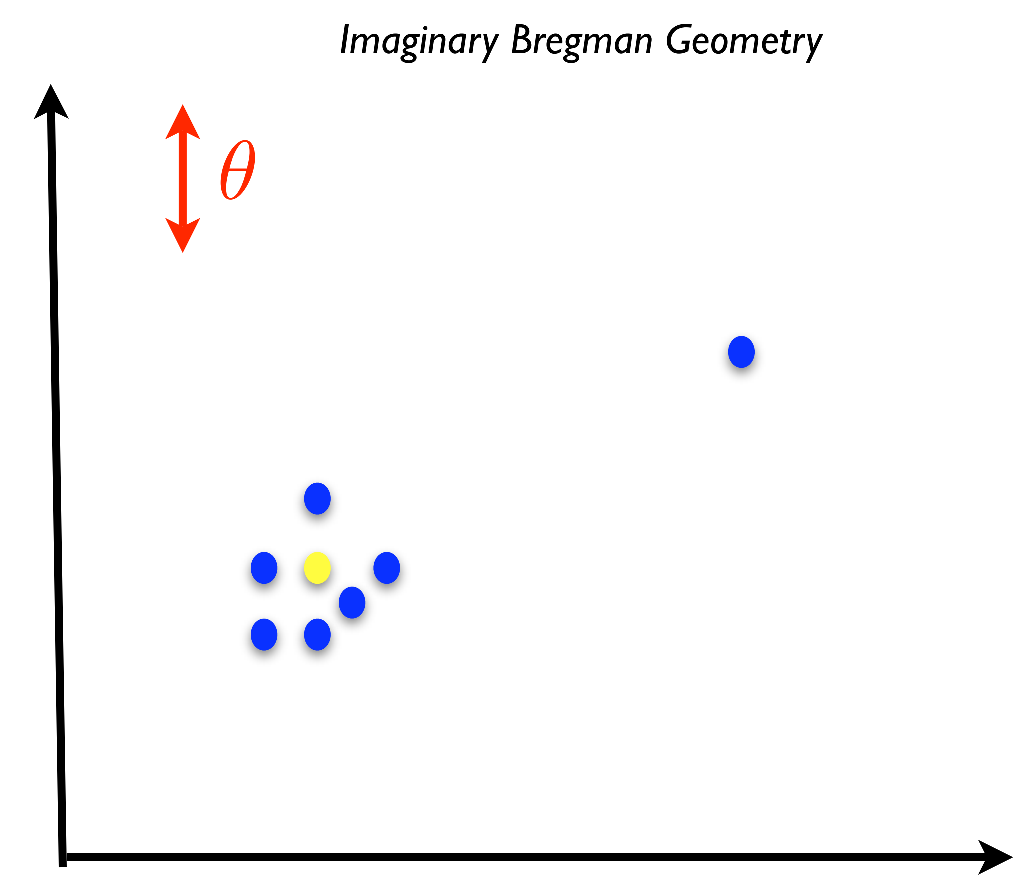
Incremental Segmentation



REALTIME Scheduler
Time = **3**

Music Information Geometry

Incremental Segmentation

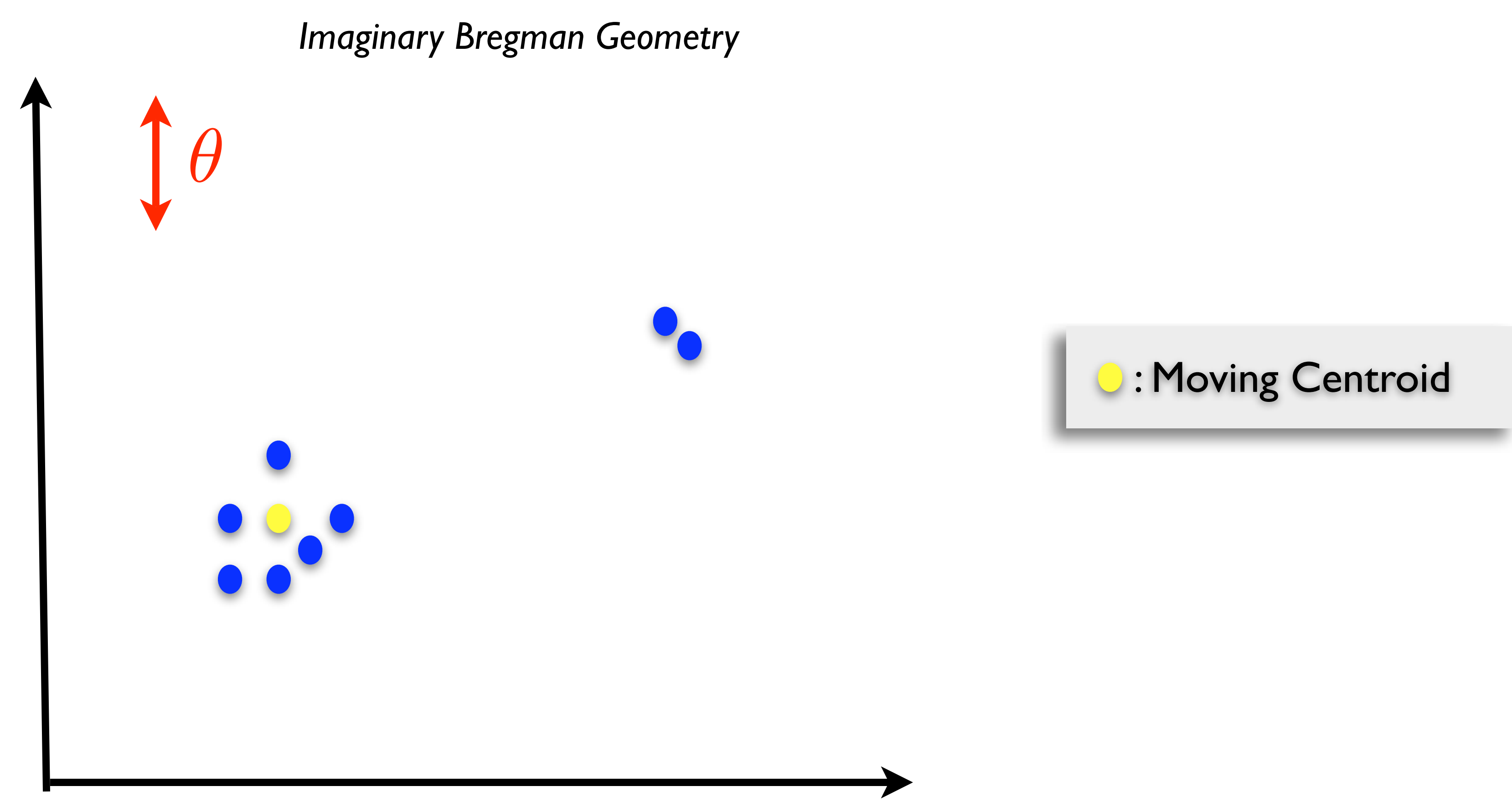


: Moving Centroid

REALTIME Scheduler
Time = **7**

Music Information Geometry

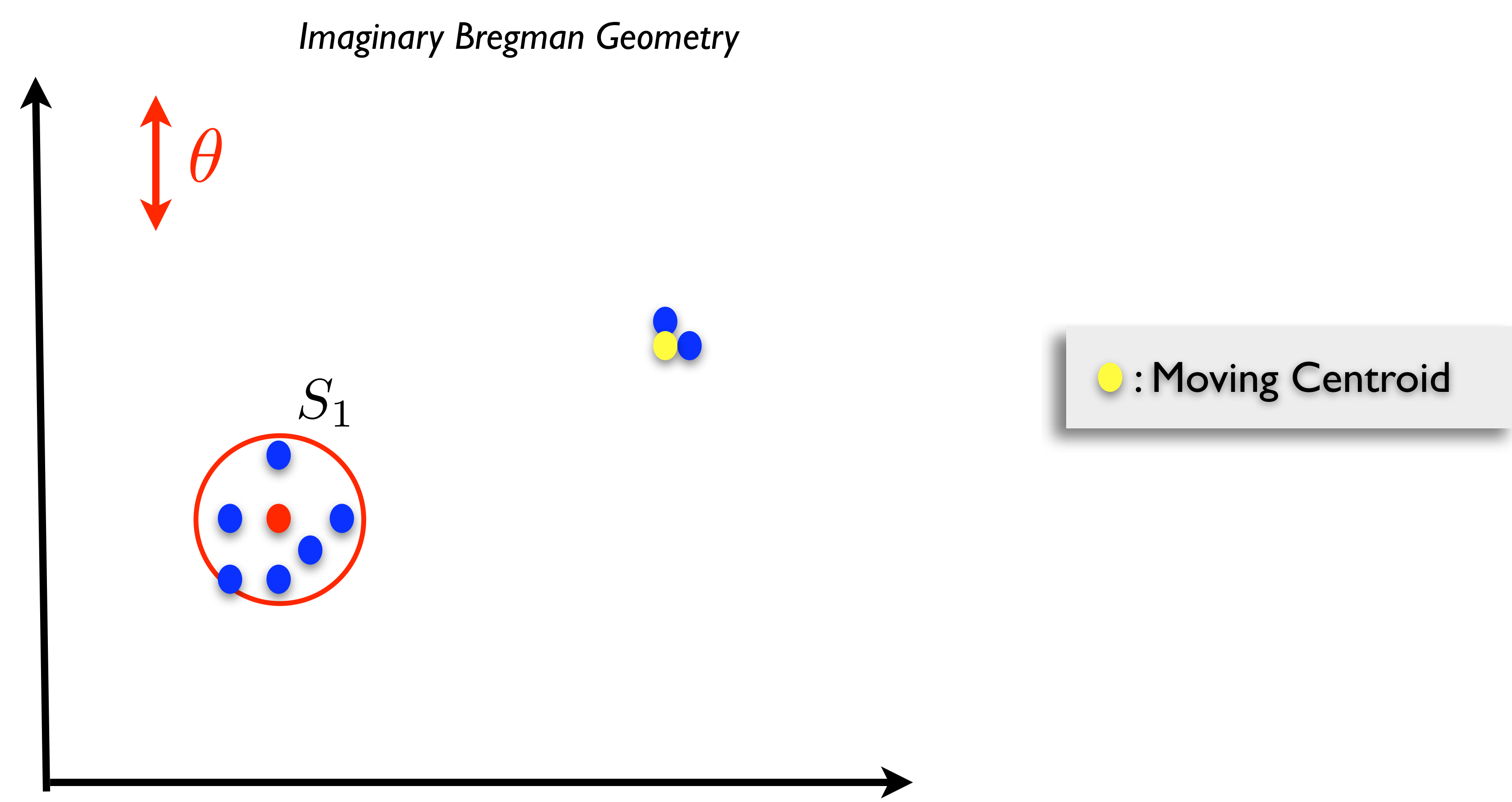
Incremental Segmentation



REALTIME Scheduler
Time = **8**

Music Information Geometry

Incremental Segmentation



REALTIME Scheduler

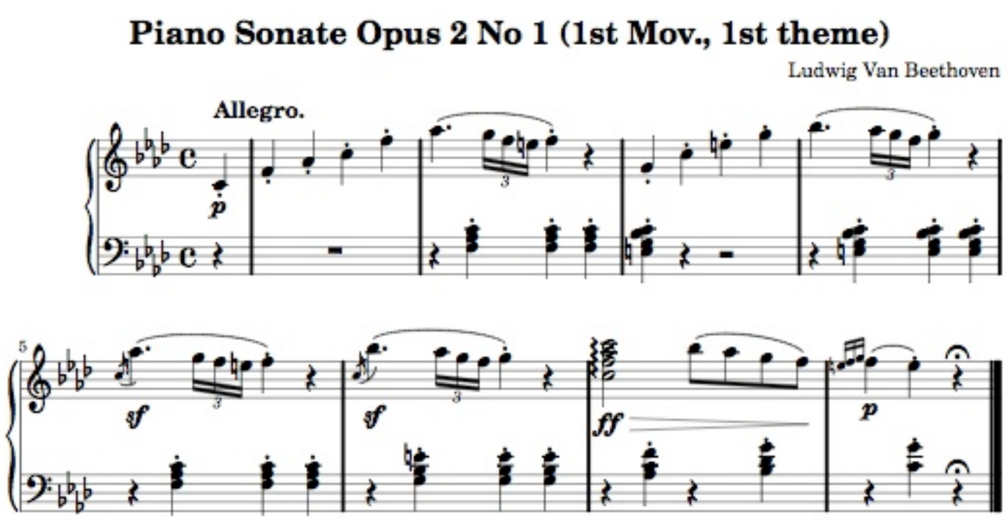
Time = **8**

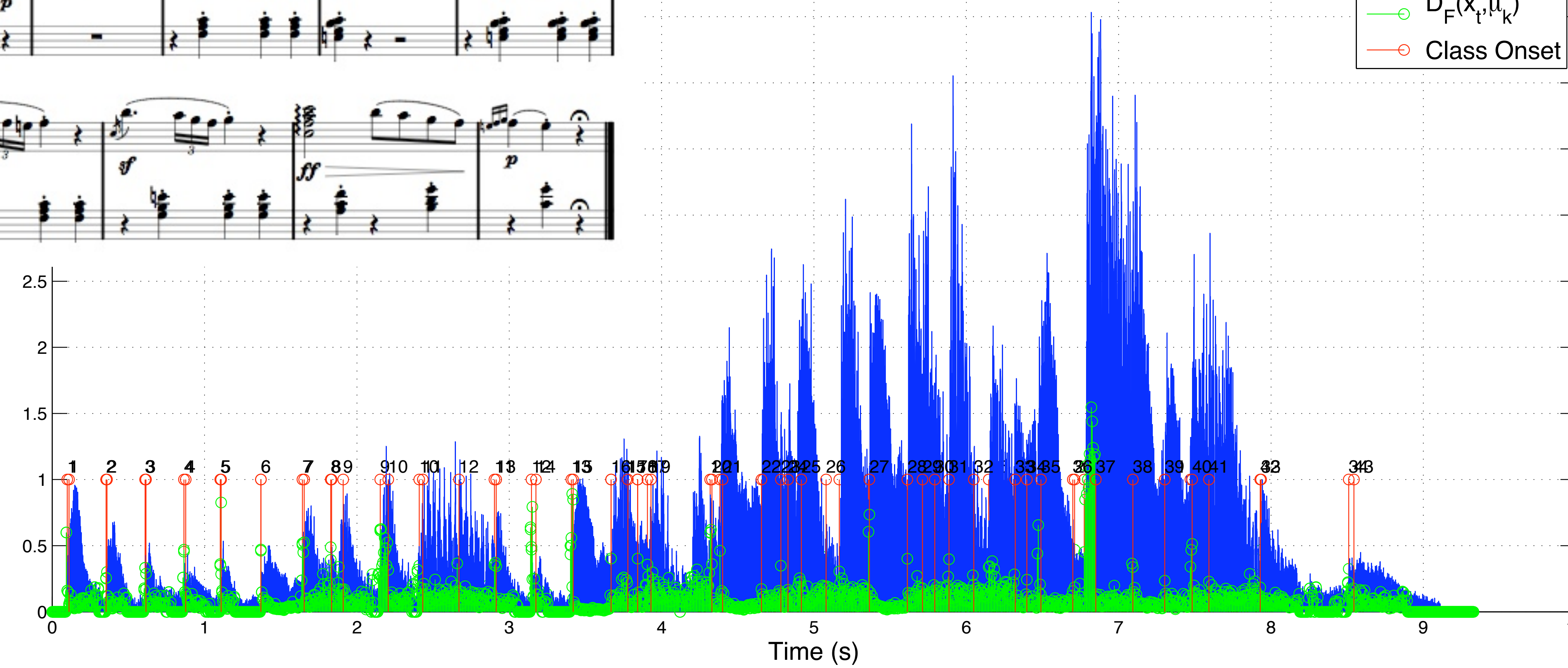
Model Formation

Music Information Geometry

Incremental Segmentation

- Sample Result: Beethoven's first piano sonata, first movement
  - performed by Friedrich Gulda (1958)





Methods of Information Access

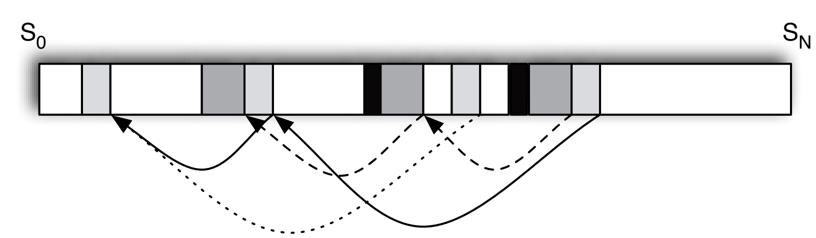
Incremental Structure Discovery

- Idea: The *models* in music information geometry provide instantaneous similarities between consequent models.
  - What about similarities between subsets of models at different time intervals?
  - What about grabbing long term regularities in the music signal?
- Literature of *Audio Structure Discovery* algorithms: Usually off-line and/or incorporate a priori beliefs over music structure
- Our goal:
  - Do it *online* and *incrementally* as audio signals arrive
  - Grab and learn regularities on-the-fly from the signal itself and without a priori knowledge
  - *Key for Anticipatory Modeling*: Grabbing stabilities and regularities of information in the environment

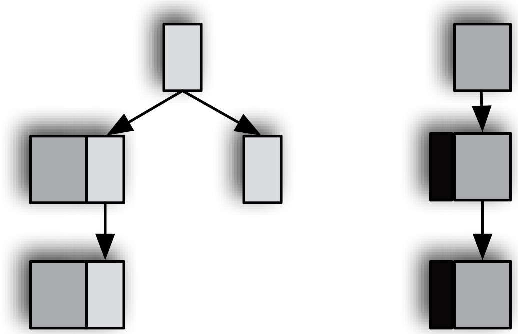
Methods of Information Access

Incremental Structure Discovery

- Proposal: Extend an existing algorithm in the symbolic domain to the continuous audio domain by passing through information geometry and *Models*.
- Point of departure: *Factor Oracles*
  - Used primarily on text and DNA data to detect repeating structures.
  - A finite-state automaton learned incrementally.
  - A state-space representation of repeating structures in a sequence



  - Provides *forest of suffix tree structures*

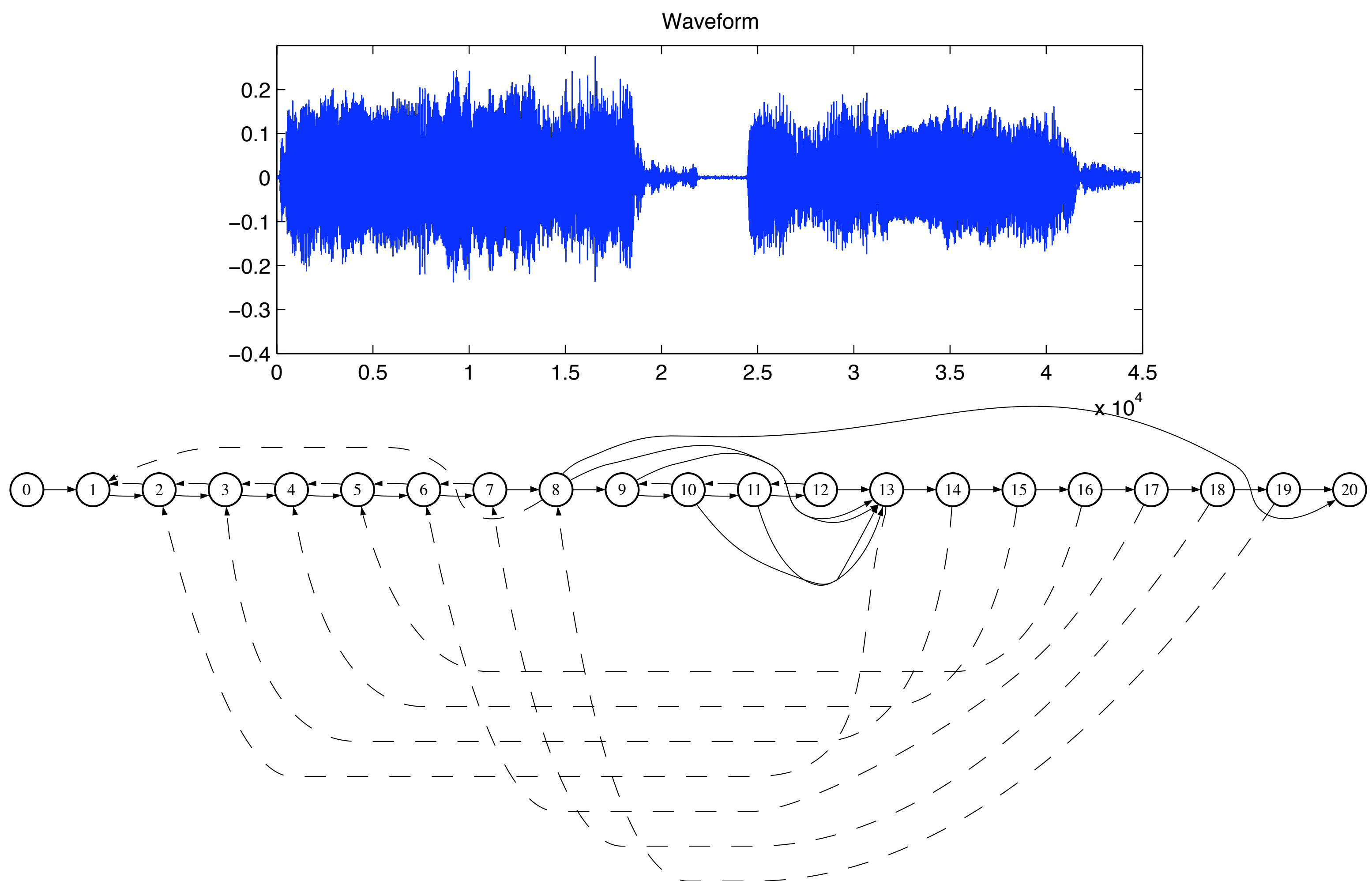


- *The beauty of MIG*
  - Keep the algorithm, replace symbols by *models* or *points* and equivalence by *similarity* in a music information geometry!

⇒ **Audio Oracle**

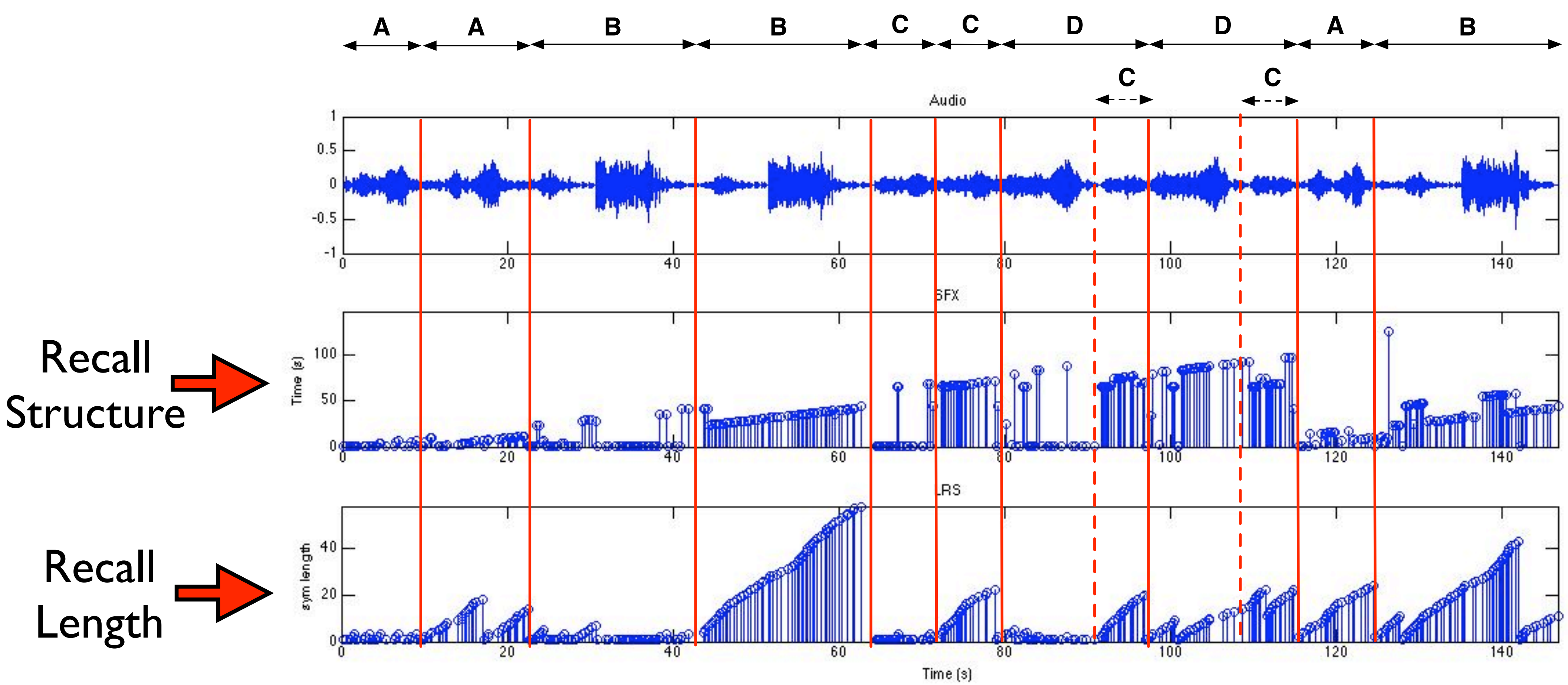
Methods of Information Access

- Audio Oracle results
  - On *points:* (each state=one analysis window)
    - *Natural bird uttering (natural repetition)*
    - *Using MFCC audio features on Multinomial music information geometry*



Methods of Information Access

- Audio Oracle results:
  - On *models*
    - *Beethoven's first Piano Sonata, Third Movement (Gulda, 1958)*
    - *Using Constant-Q amplitude spectrum on Multinomial music information geometry*
      - *150 seconds, > 9500 analysis frames, resulting to 440 states*

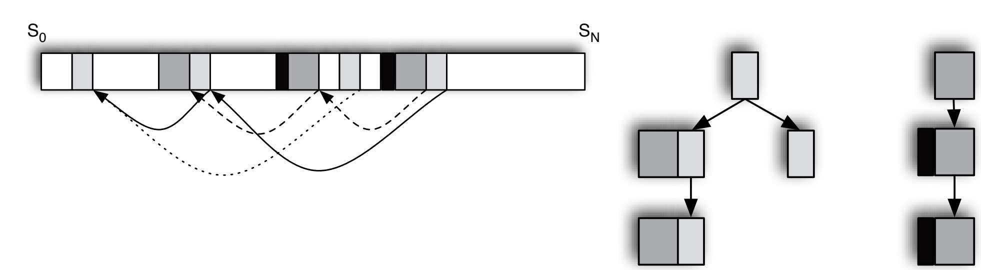


Methods of Information Access

Fast Information Retrieval

- Proposal: Compile an *search engine* over a database of audio and using an outside audio query
  - That is also capable of *recombining/reassembling* chunks of audio within a large target, to reconstruct the query.
- Related works: *Concatenative Synthesis, Unit Selection*
- Idea: Do not search on the audio itself but on *audio structures*
  - Audio Oracle as Meta data
  - (ab)use the long-term structures of Audio Oracle to maintain *perceptual continuity* of the results (access to long term structures)
- Simple *Dynamic Programming* algorithm:
  - Follow the *forest of suffix tree structures* to find the longest and best possible result
  - Maintains all the results (paths) at all times!

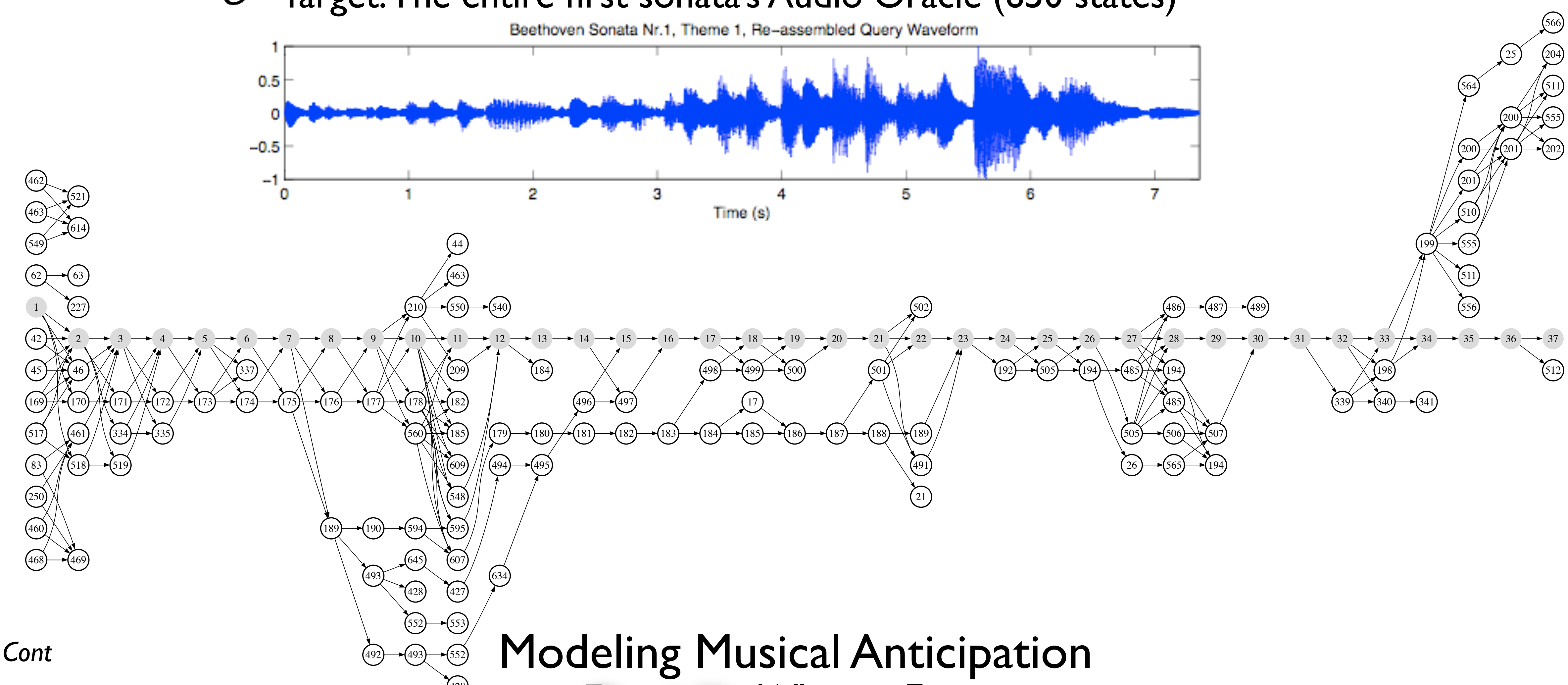
⇨ **Guidage**



Methods of Information Access

Guidage Results

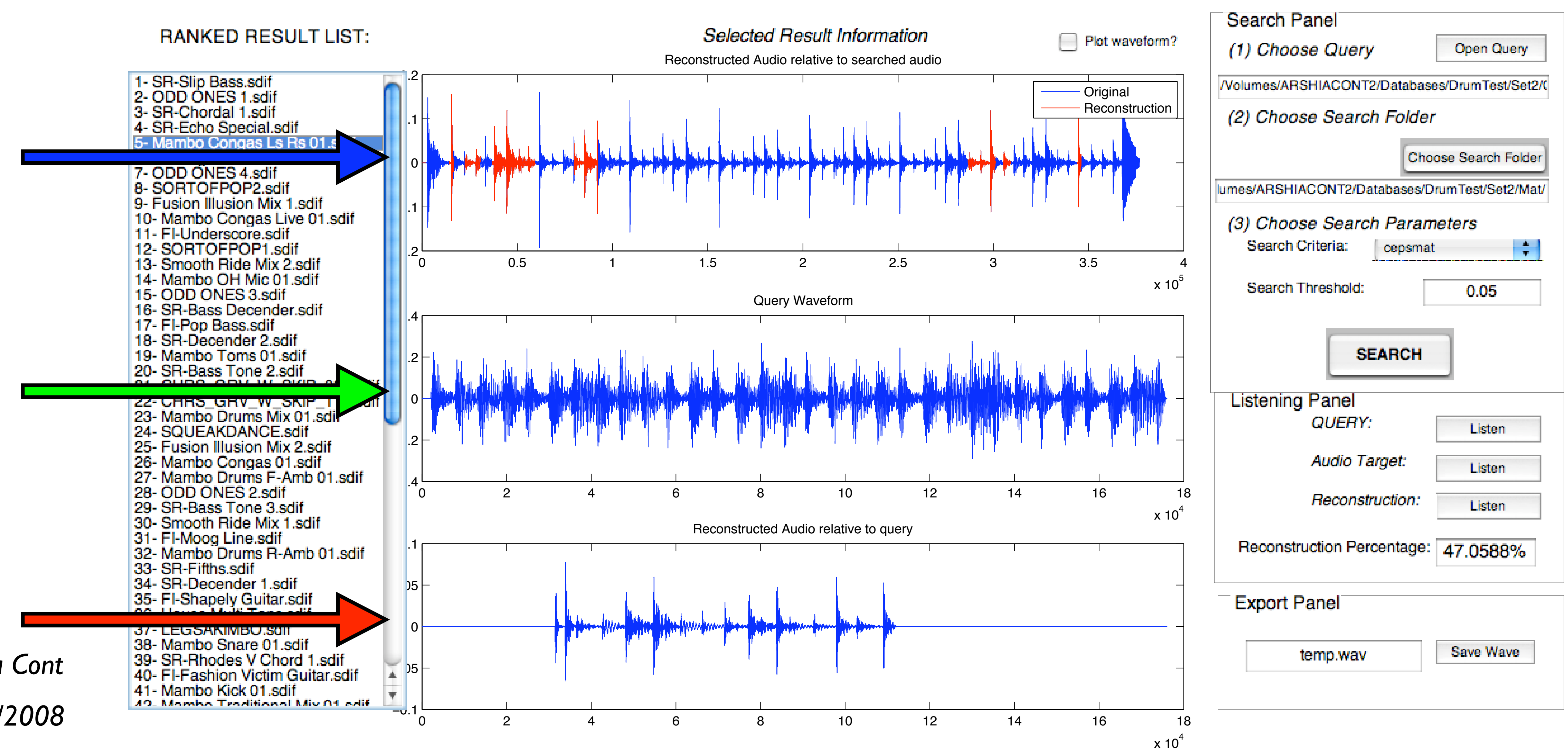
- Self-Similarity test
  - Task: Search for the *first theme* of the first Beethoven's sonata in the entire sonata.
    - Query: Audio of the first theme
    - Target: The entire first sonata's Audio Oracle (650 states)



Methods of Information Access

Guidage Results

- Database Search
  - Task: Find audio, or a recombination within a file that are similar to query
    - Query: African drum sample
    - Database: Loop database (*Kontakt*) 140 audio files, 200Mb, Mean duration 7s
      - Convergence time: 20s in *Matlab* on a 2.3Ghz unicore Intel machine



Arshia Cont

PART (III)

How to Expect

Adaptive and Interactive Learning

How what?

- We just saw how in interaction with an environment,
  - We can capture the regularities in the information structure,
  - represent it,
  - and have (fast) access to it.
- *Anticipation* is *expectations* or *beliefs* of a system bound to *actions*
  - We need to know *how* to act
  - We need to *learn* interactively the consequence of actions in the environment, take lessons, and adapt ourselves to new situations

Adaptive and Interactive Learning

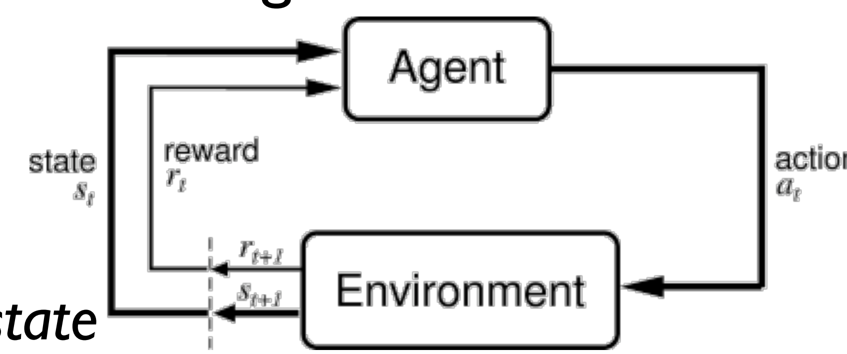
Automatic Improvisation

- We showcase this part on the problem of *Automatic Improvisation* and *Style imitation*
  - Existing systems are based on *predictions* on learned *context models*
  - We extend this to *Anticipation* through *anticipatory modeling*
  - *DISCLAIMER:*
    - **No** interest in imitating or recreating Bach!
    - To show that anticipatory learning provides ways to learn and act to gain long-term complex structures
    - With no a priori knowledge or incorporated rules
    - With presence of little data
    - With relative cheap computation and design
- In this part, we focus on *symbolic* data (MIDI signals, scores)

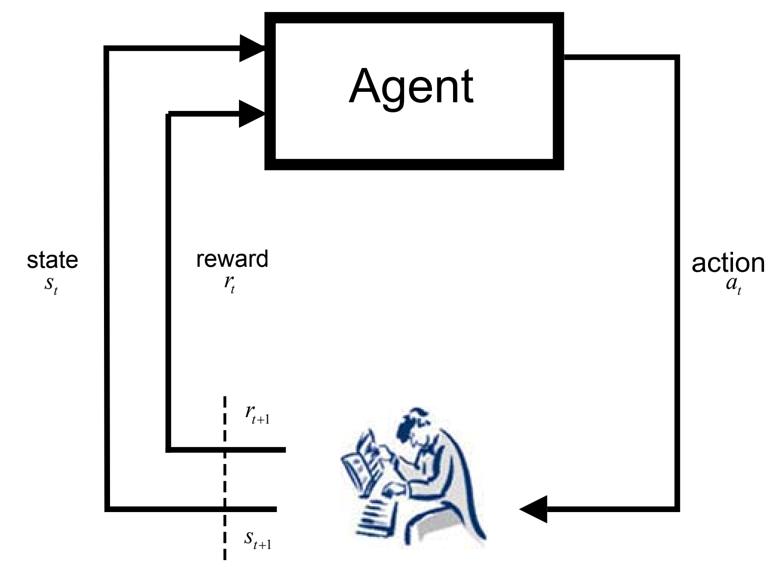
Adaptive and Interactive Learning

General Framework

- The system comprises of *agents* and its *environment*, interacting at all times. At each interaction cycle,
  - The system *perceives* the *state* of the environment
  - Select *actions* based on some *belief*
  - Based on this action, the environment might change *state*
  - and a *reward/punishment* signal *might* be sent to the agent.



- Design elements:
  - *Agent*: Computer improvisor,
    - Multiple agents [fact 3]
  - *Environment*: Human performer/Music Score
    - Dynamics of the environment ⇨ state transitions
    - Interactions ⇨ *rewards/guides* for learning
  - *Actions:* Music generations by the computer
  - *Policies:* Values associated to state-actions
    - used during generation/decision-making, learned during interactions



Adaptive and Interactive Learning

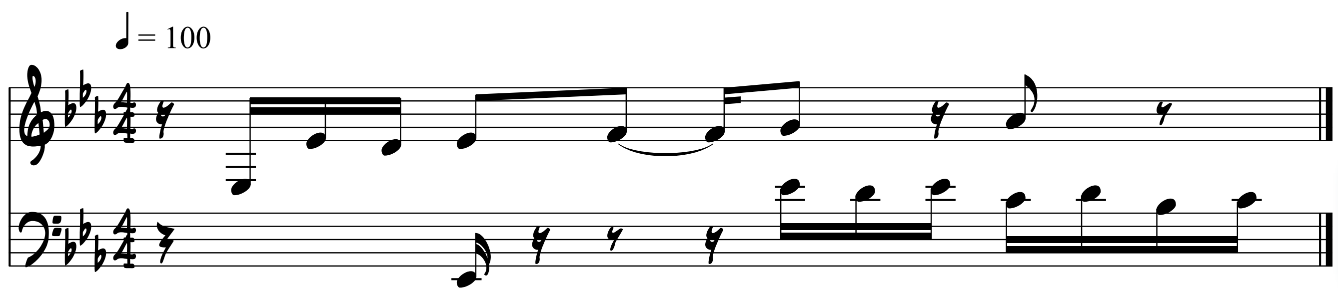
General Framework

- Problems to solve:
  1. How to represent the environment (memories/musical representations)?
     - ☑ Use *Factor Oracles* (symbolic data), *Audio Oracle* (continuous data) to model regularities in the environment
  2. How to incorporate interaction? (rewards)
     - ☑ Use *Guidage* to reinforce recurrent structures and retrieve regularities
       - At each interaction, reinforce the states in the *memory* that are factors of the new incoming sequence (from the environment).
  3. How to learn?
     - a. Learning the dynamics of the environment
       - ☑ Use the Oracle incremental updates
     - b. Learning *policies* for action decisions
       - *Active Learning* algorithm... .

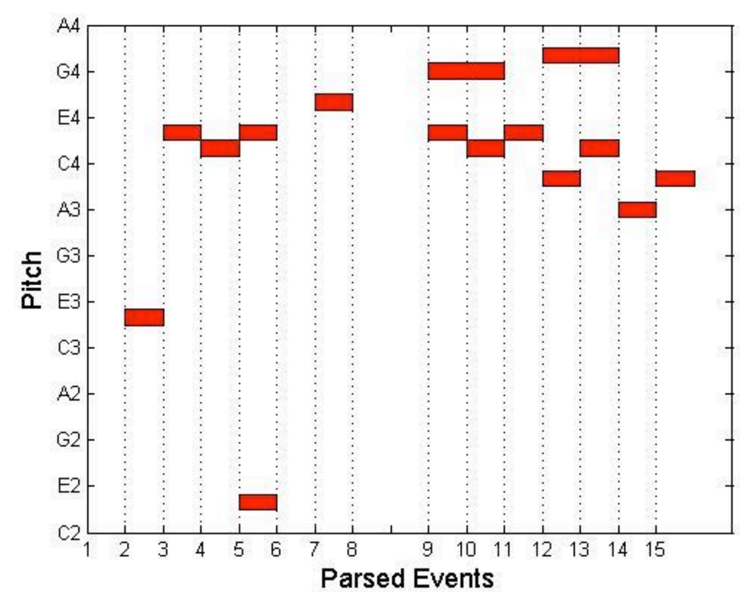
Adaptive and Interactive Learning

I. Musical Representation

- Using multiple attributes each treated as an independent sequence:



Quantization and Parsing

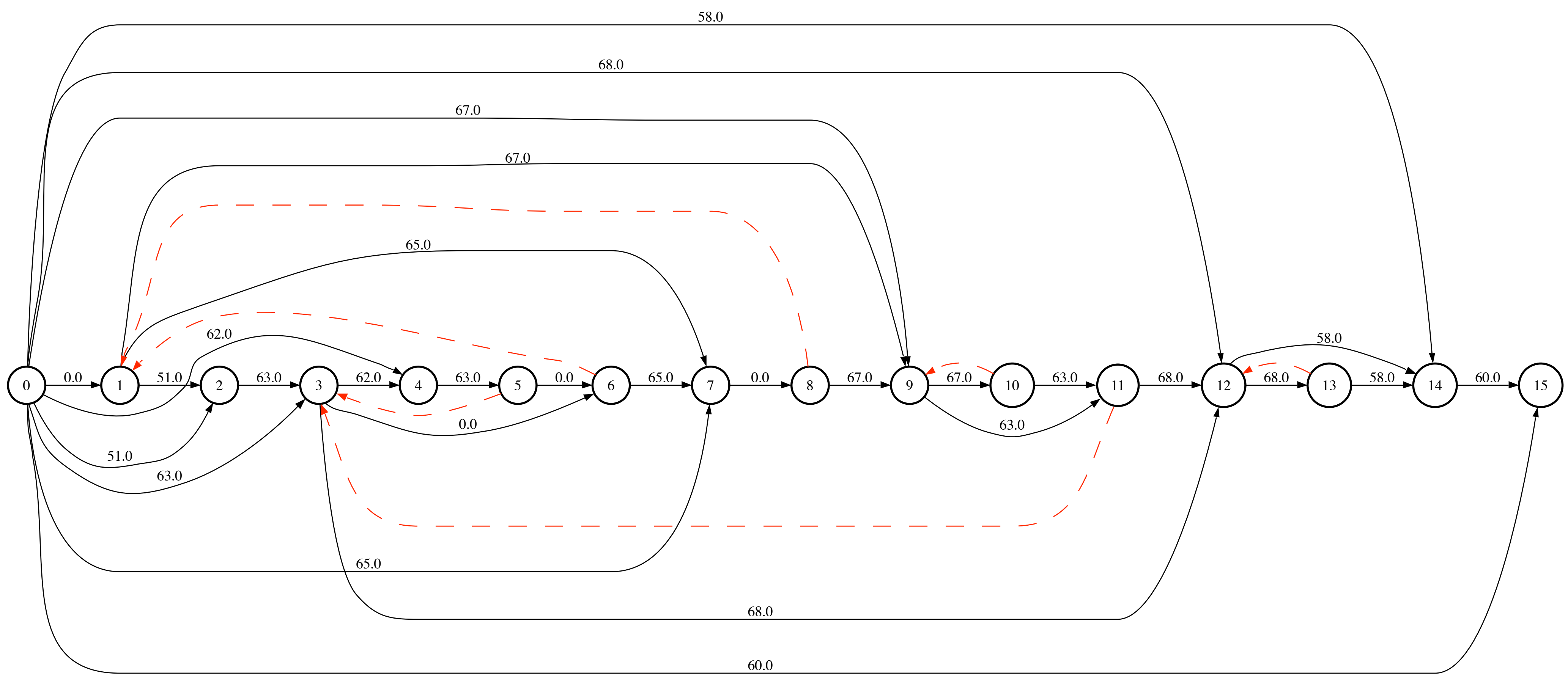


Feature Calculation

Event Number I_t	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}	I_{11}	I_{12}	I_{13}	I_{14}	I_{15}
MIDI Pitch (i_t^1)	0	51	63	62	63	0	65	0	67	67	63	68	68	58	60
Harmonic Interval	0	0	0	0	24	0	0	0	4	5	0	8	6	0	0
Duration (i_t^3)	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1

Adaptive and Interactive Learning

I. Musical Representation



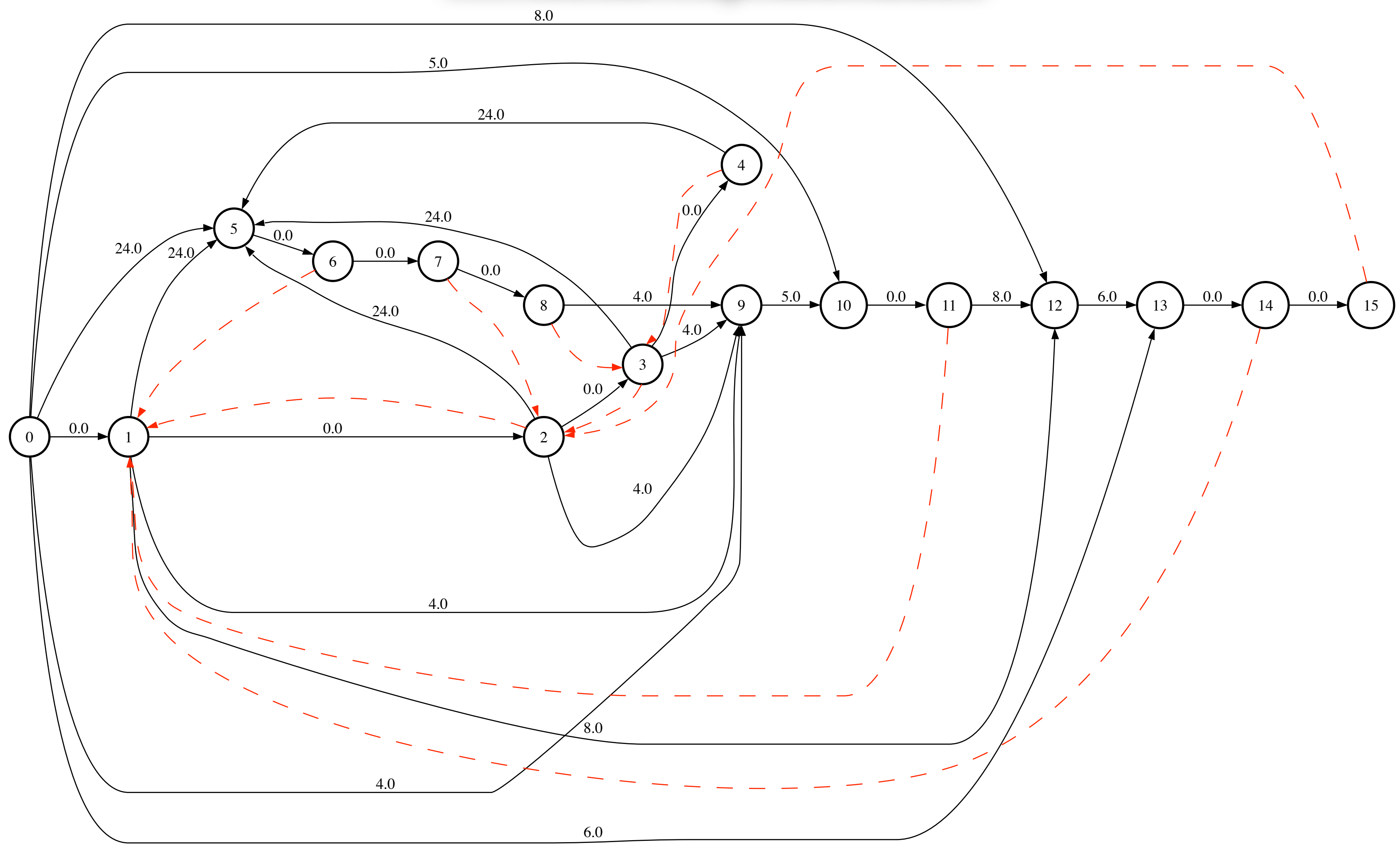
Event Number I_t	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}	I_{11}	I_{12}	I_{13}	I_{14}	I_{15}
MIDI Pitch (i_t^1)	0	51	63	62	63	0	65	0	67	67	63	68	68	58	60
Harmonic Interval	0	0	0	0	24	0	0	0	4	5	0	8	6	0	0
Duration (i_t^3)	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1

UCSD

ircam Centre Pompidou

Adaptive and Interactive Learning

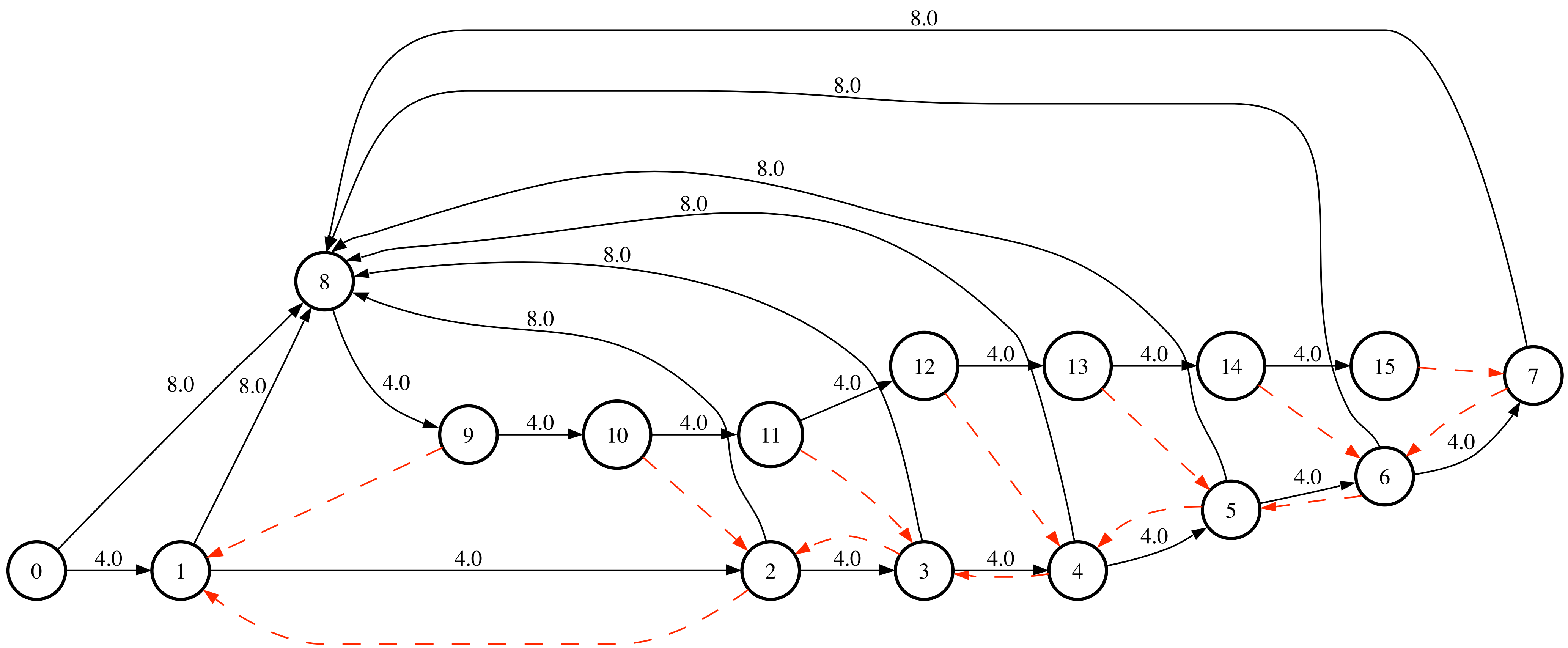
I. Musical Representation



Event Number I_t	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}	I_{11}	I_{12}	I_{13}	I_{14}	I_{15}
MIDI Pitch (i_t^1)	0	51	63	62	63	0	65	0	67	67	63	68	68	58	60
Harmonic Interval	0	0	0	0	24	0	0	0	4	5	0	8	6	0	0
Duration (i_t^3)	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1

Adaptive and Interactive Learning

I. Musical Representation

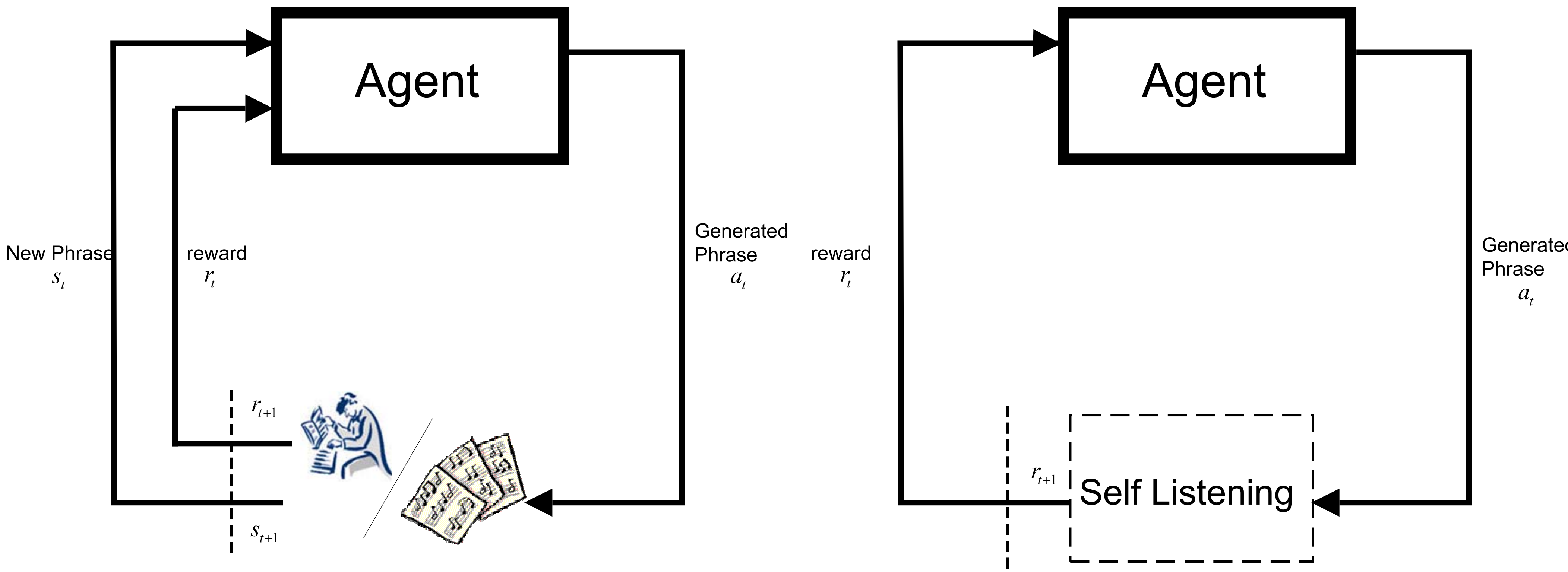


Event Number I_t	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}	I_{11}	I_{12}	I_{13}	I_{14}	I_{15}
MIDI Pitch (i_t^1)	0	51	63	62	63	0	65	0	67	67	63	68	68	58	60
Harmonic Interval	0	0	0	0	24	0	0	0	4	5	0	8	6	0	0
Duration (i_t^3)	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1

Adaptive and Interactive Learning

II. Interaction Modes

- Two modes:



  - "Interaction mode": (left) interacting with an environment, receiving rewards (or guides) and constructing knowledge.
  - "Self-listening mode": (right) During automatic generation. Reflecting on the changes in the environmental context caused by the system itself.

Adaptive and Interactive Learning

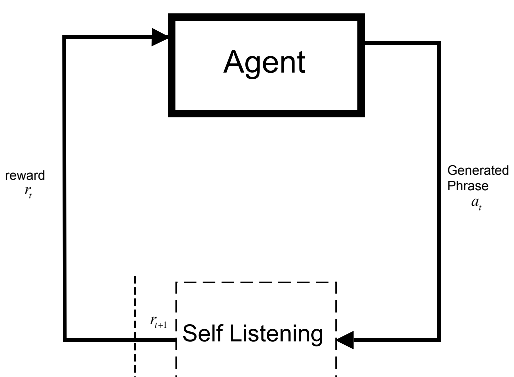
Anticipatory Learning

- Goal: To maximize *rewards* on each action by updating the policy of each state-action pair (reinforcement learning).
  - Rewards on a *future horizon:* $R(s_t) = \sum r(s_t, a_t) + \gamma r(s_{t+1}, a_{t+1}) + \cdots + \gamma^m r(s_{t+m}, a_{t+m}) + \cdots$
    - Predicting possible steps and evaluating them
    - Similar to the idea of *a rehearsing musician*
  - Updates on *selected* state-actions by *Guidage* at each interaction cycle
- *Competitive and Collaborative Learning*
  - Choose the winning agent at each cycle
  - Follow the winner for updates during in that episode.
  - Upon each update, influence relevant states in other agents

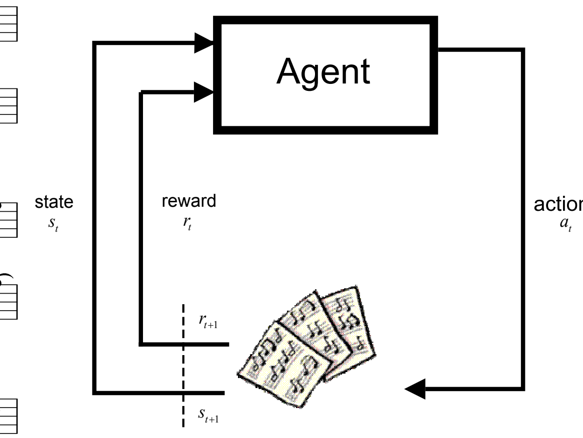
Adaptive and Interactive Learning

Sample Generation

- Learn sequentially on J.S.Bach's *"Two-part Invention, Book II, Nr. 3"*
- Generate (self-listening mode) using the gained knowledge



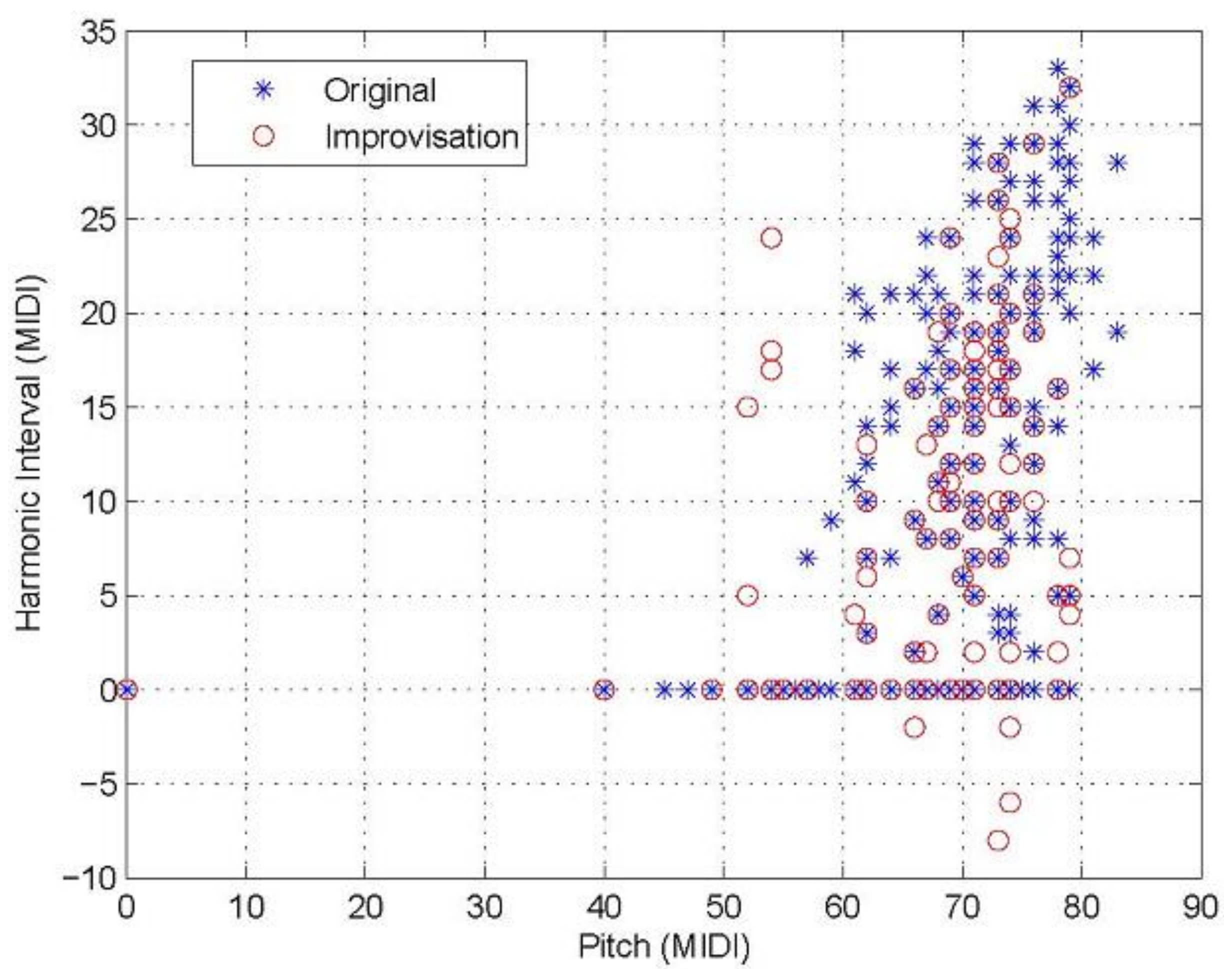
Improvisation Session after learning on Invention No.3 by J.S.Bach



Adaptive and Interactive Learning

Generalization

- As a result of collaborative learning and generation



- Preliminary evidence of long-term planning and complex structures
- No a priori knowledge of the domain
- Fast learning and generation
- Little data needed for training
- with many questions to ask and directions to pursue

PART (IV)

When to Expect

Anticipatory Synchronization

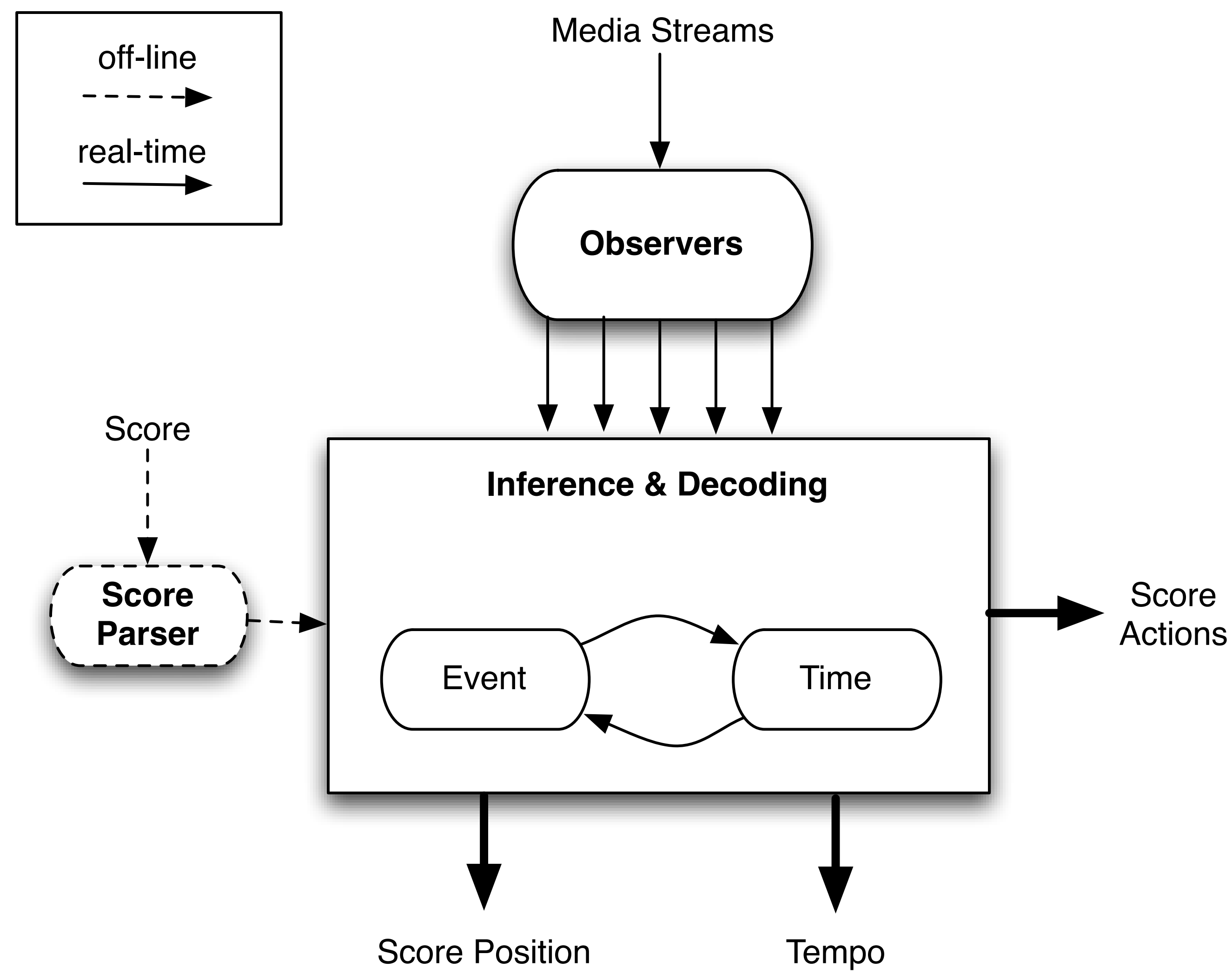
Motivations

- Technical
  - *Score Following* problem
    - Real Time Audio to Symbolic Score Synchronization
    - More focus on *acoustic features*, less emphasis on *time*
  - *Coordination* problem
    - Musicians employ various sources for synchronization
    - *Expectations about future* events play a role as important as the musical events themselves.
- Musical
  - Extend the score following paradigm
  - *At the Time of Composition:* Enable concurrent and flexible representations of events/time.
  - *At the Time of Performance:* Bring in the composition on stage by the virtue of interpretation of the written score.

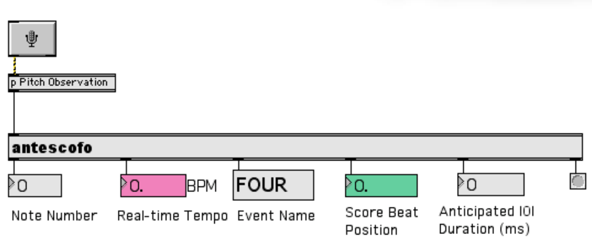
→ ***Towards a "writing" of Time and Interaction in Computer Music***

General Architecture

- Time of Composition:
  - One score containing *declarative* instrumental events and *electronic actions*
  - Concurrent Representations
  - Concurrent event time-scales
- Time of Performance:
  - Two collaborative and competitive agents: Event Agent and Tempo Agent
  - Interpreting and giving life to score parameters and written time structures

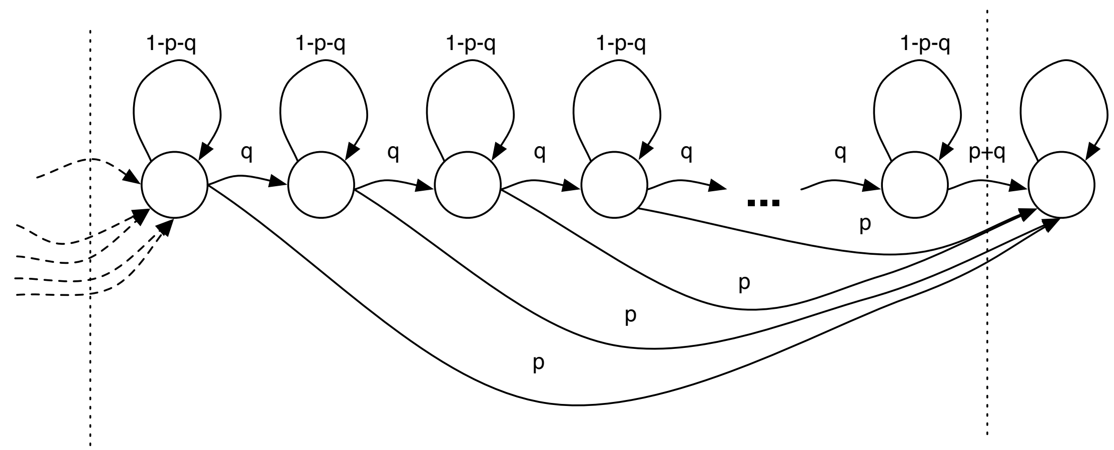


Antescofo



Probabilistic Models of Time

- *Concern:* Uncertain nature of time occupancy of musical events (neither deterministic nor arbitrary)
- Two mainstream views:
  1. *Non-Parametric Markov Occupancy Models*



  2. *Semi-Markov Occupancy Models*
     - One (and only one) state per macro-event!
     - But with an *Explicit* distribution $d_j(u)$ for time occupancy of each event *j*

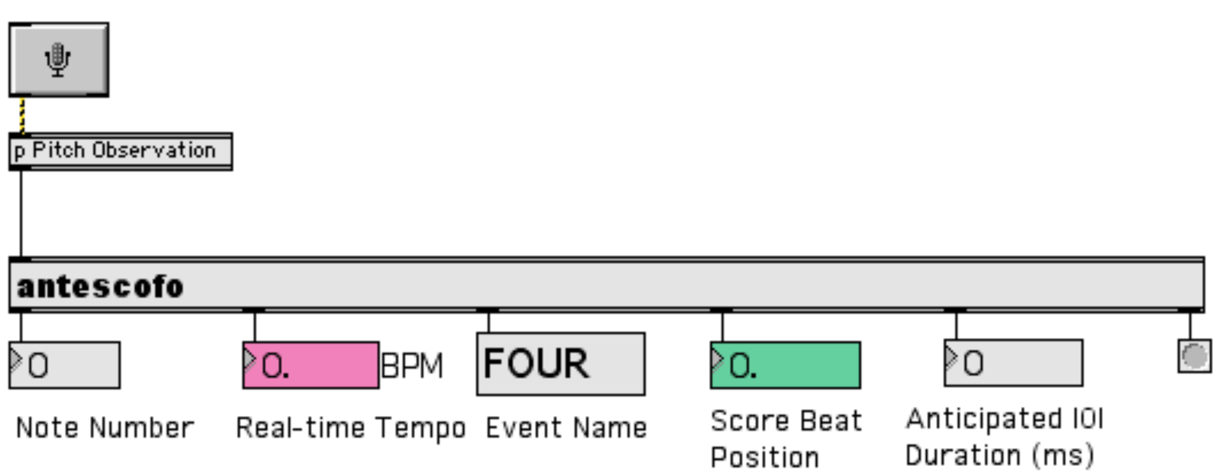
Probabilistic Models of Time

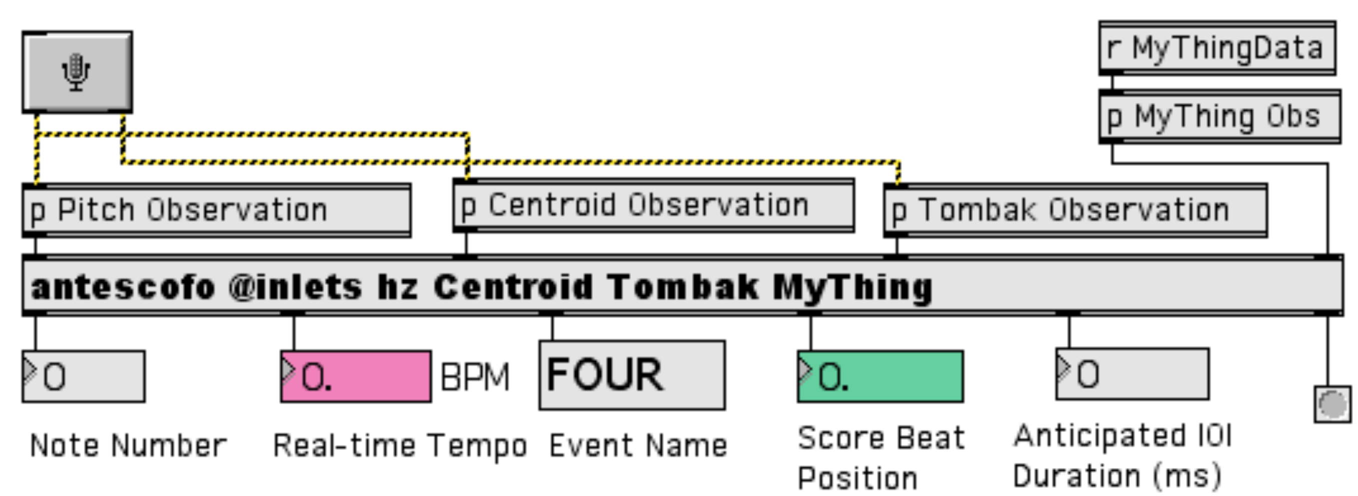
Proposal

- Use advantages of both worlds:
  - *Hybrid Markov/Semi-Markov Models*
- Collaborative and Competitive Inference
  - Use *predicted tempo* to anticipate (and prepare for) future events!
  - Coupling event (audio) and tempo agents instead of cascading!
  - Adaptive/online learning
- Advantages:
  - No need for offline training of the system
  - Reduce in number of parameters
  - No need for strong acoustic model!

Observers

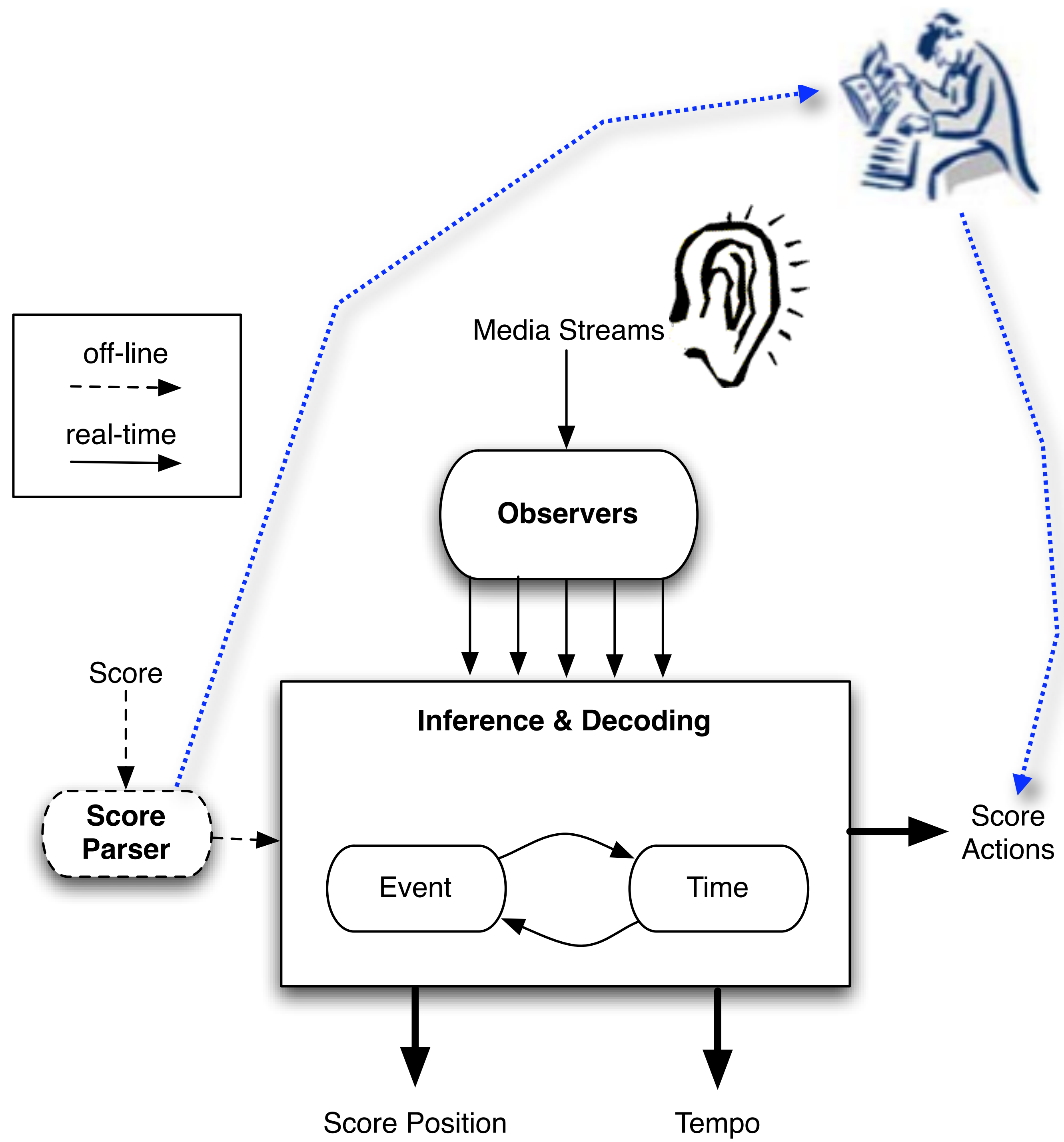
- Are the "eyes and ears" of the system during live performance!
  - Traditionally *pitch only*
  - Recently: Raw audio (audio matching), gesture (gesture following), Audio features, Video streams, etc.
- Proposal: Concurrent Observations
- Hard-coded Audio Observer for Polyphonic Events:
  - Compare real time audio spectrum to pre-constructed harmonic templates out of the score.
  - Normalized Audio Spectrum ⇨ Multinomial Dists ⇨ Use KL divs





Antescofo's Score Semantics

- Simple (and young) text-based and declarative language for writing of time and interaction
- Important Specificity:
  - Co-existence of instrumental score and electronic actions.
  - Bridging the gap between the *time of composition* to *time of performance*
  - *Actions* can be written in *relative time*, whose values are evaluated at run-time (live performance), coupled with real-time tempo



Antescofo

Antescofo has been used...

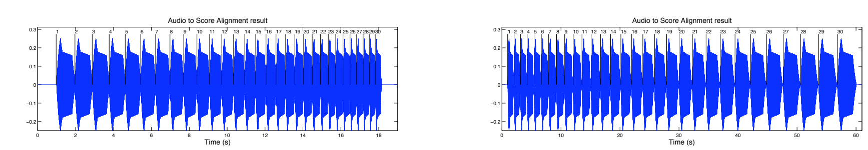
- *"... of Silence"*, Marco Stroppa, for Saxophone and chamber electronics *(Antescofo Premiere)*
- *"Anthèmes 2"*, Pierre Boulez, for Violin and Live electronics
- *"... Explosante-Fixe..."*, Pierre Boulez for flute, orchestra and electronics
  - LA Philharmonic
- *"Speakings"*, Jonathan Harvey for orchestra and live electronics
  - BBC Scotish Orchestra
- *and more to come...*

Try: http://cosmal.ucsd.edu/arshia/antescofo/

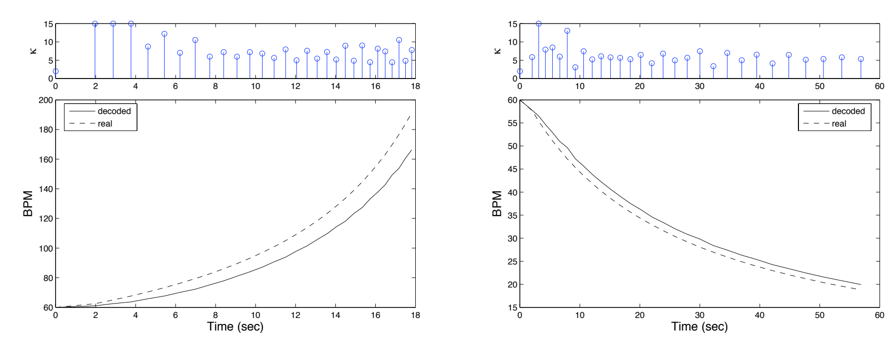
Evaluation

- Tempo Evaluation:
  - Use synthesized audio from score to attain milli-second tempo precision

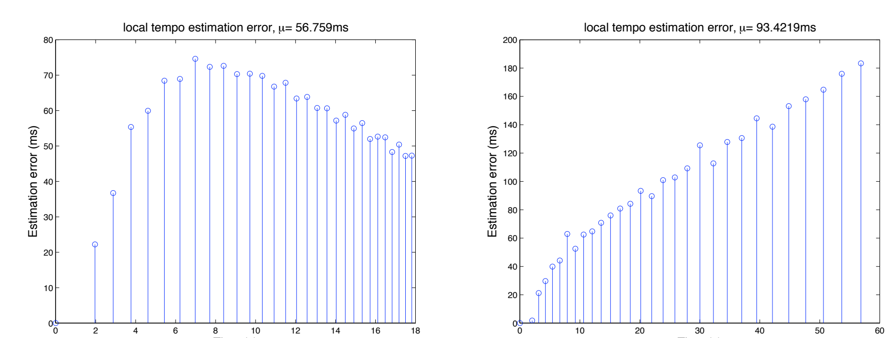
Accelerandos and decelerandos:



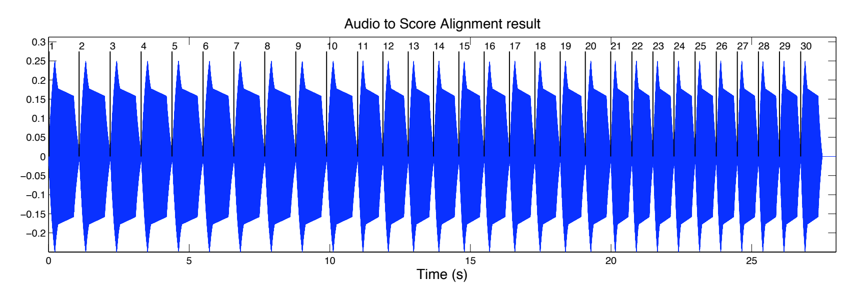
(a) Waveforms and alignments for accelerating (left) and decelerating (right) tempi



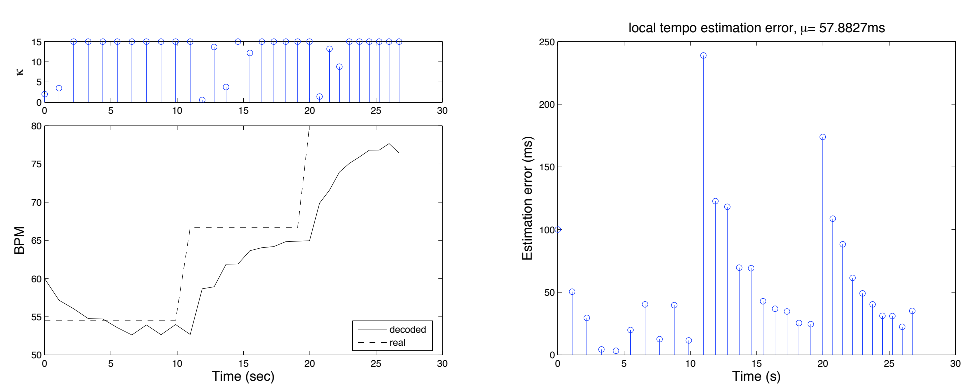
(b) Estimated and real tempi for acceleration and deceleration in BPM



Discrete Tempo Change during performance:



(a) Waveform and alignment result



Evaluation

- Alignment Evaluation:
  - MIREX 2006-08 Evaluation Campaign for Score Following
  - Augmented the database to contain polyphonic music

#	Piece name	Composer	Instr.	Files	Prefix	Events
1	Explosante-Fixe	P. Boulez	Flute	7	tx-sy	615
2	K. 370	Mozart	Clarinet	2	k370	1816
3	Violin Sonata 1	J.S. Bach	Violin	2	vs1-	2019
4	Fugue BWV.847	J.S. Bach	Piano	1	RA	225

  - Total Precision obtained = **91.49%**
    - Details in the manuscript

Contributions

I. From Modeling Anticipation to Anticipatory Modeling
   - Modeling Anticipation
   - Anticipatory Modeling
II. What to Expect?
   - Music Information Geometry
   - Methods of Information Access
III. How to Expect?
   - Anticipatory Learning in an environment
IV. When to Expect?
   - Anticipatory Synchronization
   - Towards a "writing" of time and interaction in computer music
V. Conclusions

- ☑ A formal of definition of *anticipation* destined for computational models of sound and music.
- ☑ A formal definition of *anticipatory modeling* inspired by music cognition.

- ☑ A mathematical framework for quantification and qualification of music information and content based on Information Geometry.
- ☑ An online algorithm for incremental clustering and structure discovery of music signals.
- ☑ A fast and online algorithm for unit selection over large databases of music based on users' audio query.

- ☑ An online adaptive and interactive learning framework achieving anticipatory planning strategies, based on Active Learning.
- ☑ A preliminary computational anticipatory framework for automatic style imitation and automatic improvisation.

- ☑ An anticipatory design for real-time audio to score alignment featuring coupled audio/tempo agents and capable of decoding real-time position as well as tempo of the performer for polyphonic music signals.
- ☑ A preliminary framework and language for writing of time and interaction destined for interactive mixed instrumental and live computer music repertoires

Thank you!

More Audio Examples:
http://cosmal.ucsd.edu/arshia/

Online Manuscript:
http://cosmal.ucsd.edu/arshia/papers/ArshiaCont_PhD.pdf