Queen Mary
University of London

Lecture 2: Modelling musical creativity

Geraint A. Wiggins
Professor of Computational Creativity
CCLab
Queen Mary University of London

PROMOTING THE SCIENTIFIC EXPLORATION OF COMPUTATIONAL CREATIVITY

Overview

Queen Mary University of London

Aims:

- to motivate the scientific study of musical creativity
- to provide an approach to doing so
- to demonstrate the approach
- to place the model in the context of the Creative Systems Framework

A Science of Music?

Queen Mary University of London

Why would we want to study a science of music?

Music is

- an art form
- a cultural construct
- precious, even sacrosanct, to many people

What can science tell us that musicology cannot?

- Place of music in general cognition
- Contibution of musical behaviour to human developement
  - individual
  - cultural

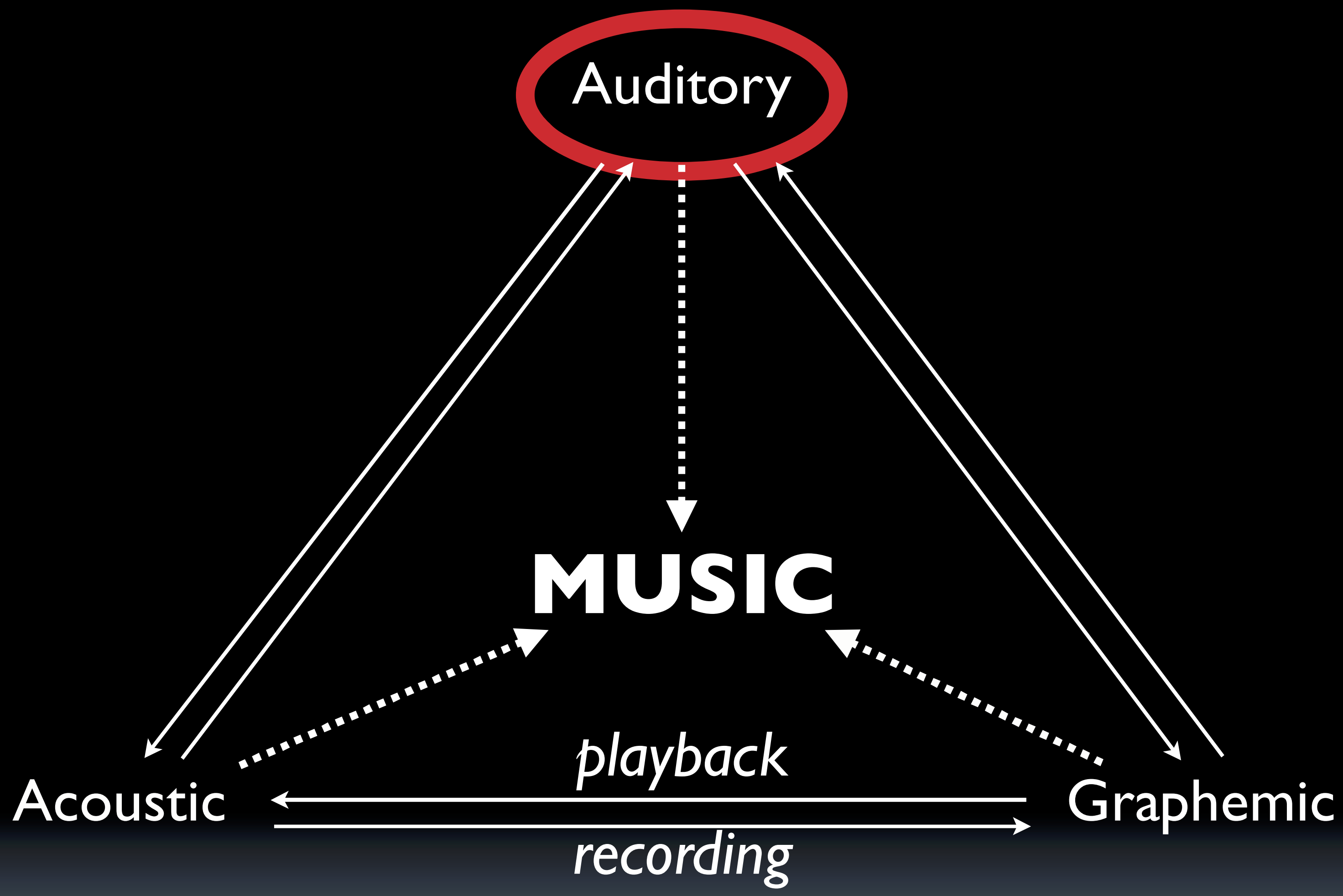
A Science of Music?

Queen Mary University of London

- Music seems to be uniquely a human faculty
  - There is no known human culture without music
  - Music is everywhere in every human culture
  - Music is irresistible to the majority of people
  - No other species has been shown to exhibit musical behaviour in the sense that humans do
- Yet, no known bio-evolutionary advantage is given by music
  - This needs to be explained, and not just wondered at!
- What is more, musical behaviour is a fundamental part of being human
  - So we need to understand music if we want to understand ourselves

Milton Babbitt's representational domains

Milton Babbitt (1965) proposed three *domains* of music representation

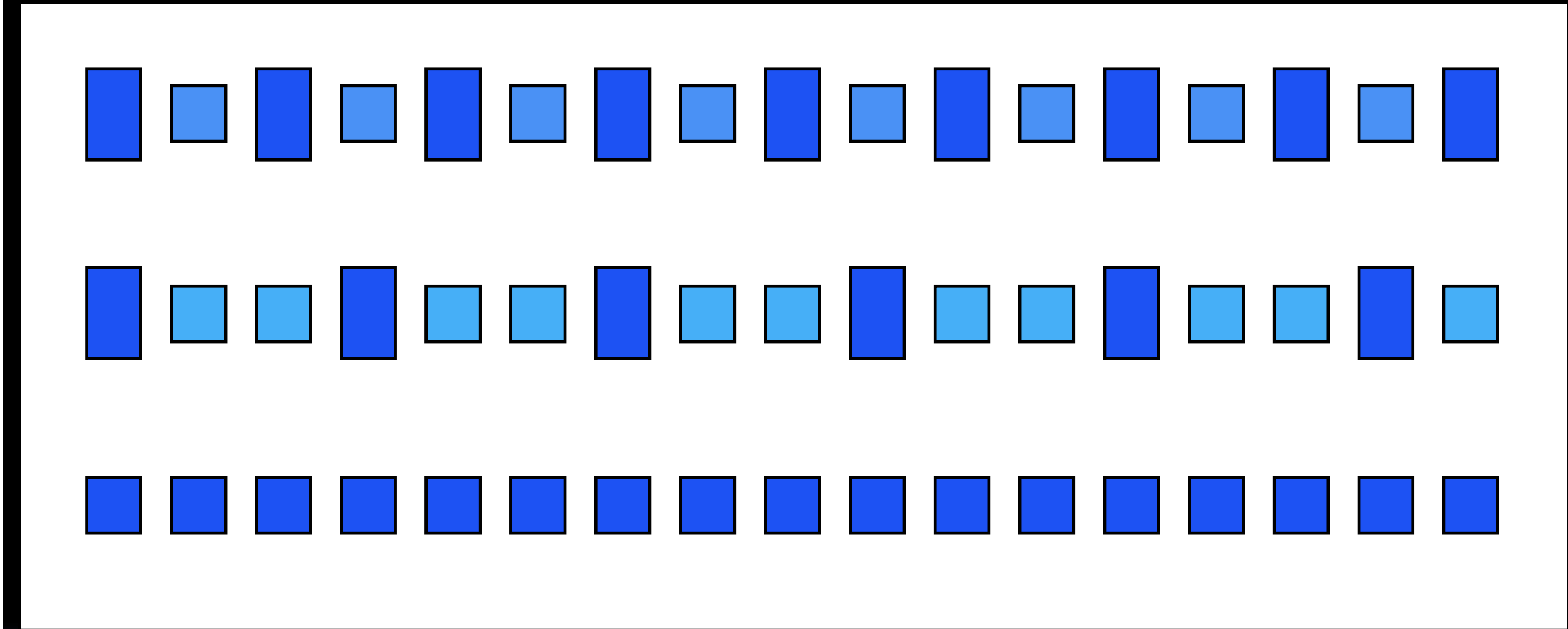


Music and meaning - a bonus for CC

Queen Mary
University of London

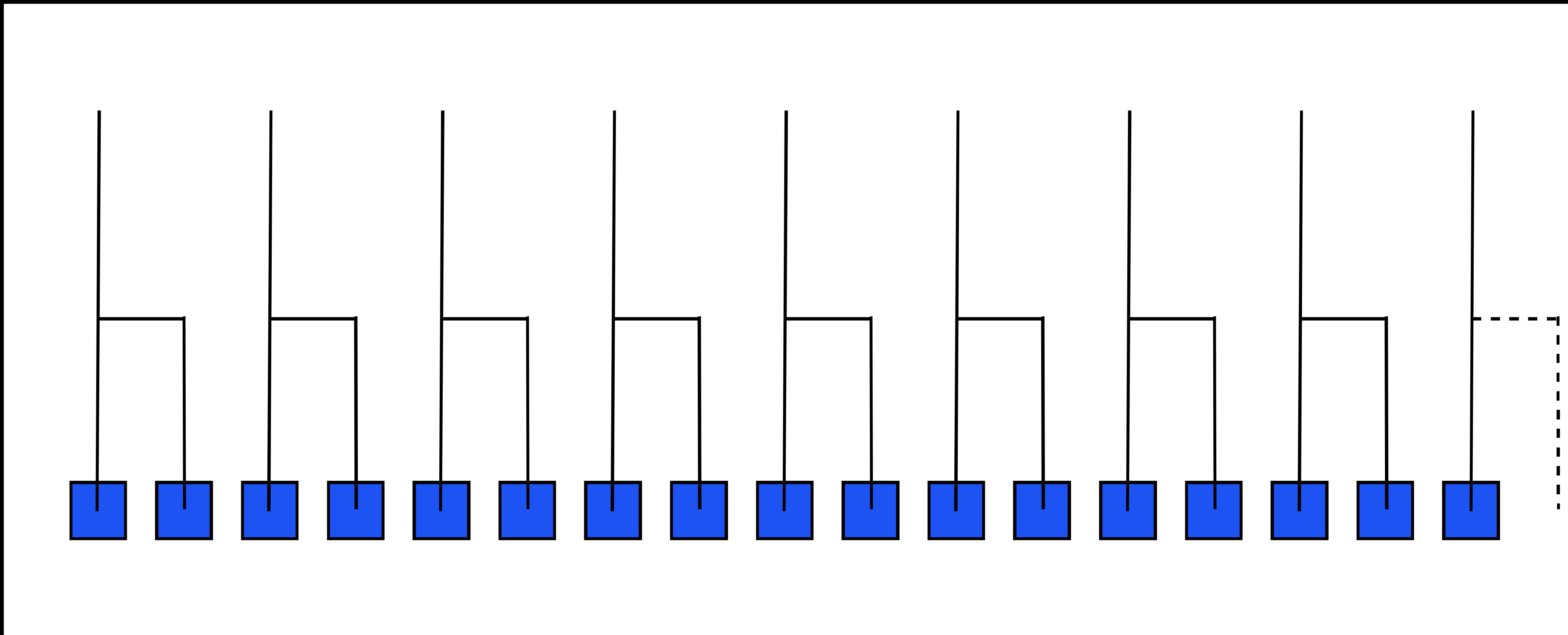
- Music is all about meaning
  - associations
  - emotional responses
- but it (usually) has no semantic content
  - so a musical creative system has a massively reduced framing problem
  - the frame is syntactic, not semantic
  - ie it's about style, intrinsic to music, not about an extrinsic world model

Feeling the beat

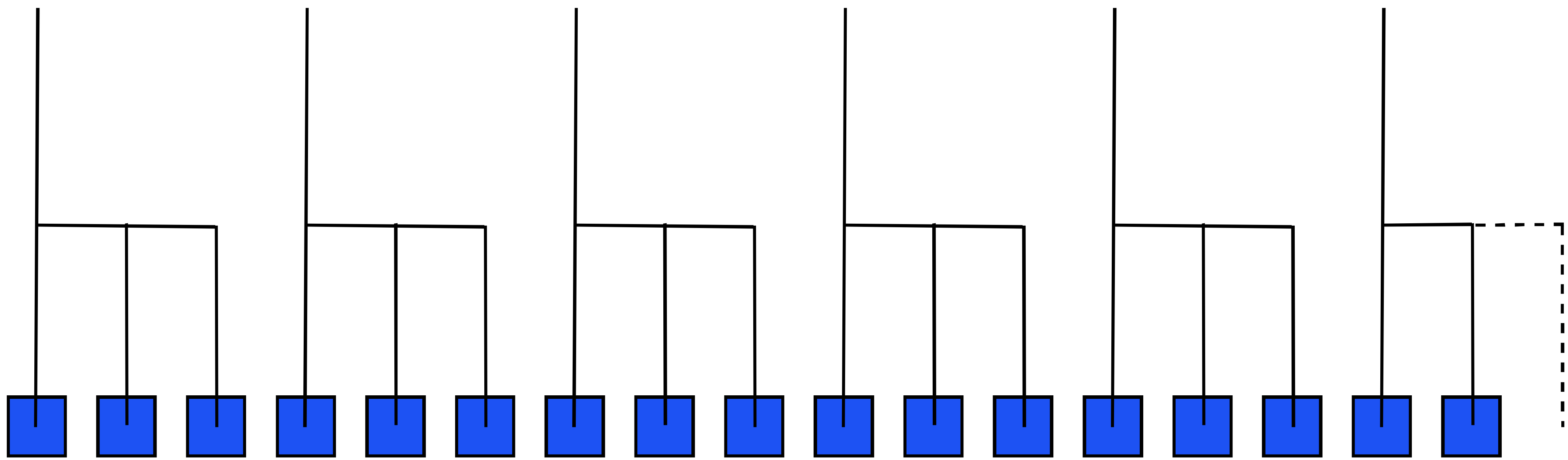


Feeling the beat

Queen Mary
University of London



Feeling the beat



Feeling the beat

Queen Mary
University of London

The pulse train demonstration shows that human listeners tend to hear rhythmic structure in sound...

...even when it isn't there

When we know it isn't there, we can manipulate our own perception, to hear either twos or threes

Why is there grouping?

Queen Mary University of London

Our senses are subject to a constant barrage of information

One way to help deal with this information is to break it into more easily manageable lumps

Broadly speaking, this is what is happening in grouping phenomena (and others)

“Chunking” is a general phenomenon at all levels throughout cognition

Grouping for music?

Queen Mary University of London

We make smaller and larger scale groupings too

- Small scale: harmonics *fuse* together to make notes
- Medium scale: notes fuse together to make chords
- Larger scale: chords work in sequence to make progressions

And some authors define music as "Organised Sound"

Grouping for music?

Queen Mary
University of London

But why?

It seems unlikely that we evolved the ability to group sounds because music was useful in evolutionary terms

- Music does not help find food
- Music does not protect against predators

Grouping for music?

Queen Mary
University of London

But the things that contribute to music perception do contribute to evolutionary fitness in other ways:

perceiving complex sound as tone colour
Is that the sound of a lion (lute)?

location of sound sources
Where is that tiger (trombone)?

grouping of separated rhythmic sounds
Do I hear footsteps (fandango)?

perceiving expressed emotion in sound
Is the other hominid pleased to see me?

Entrainment

Entrainment (in music) is the ability to perceive and duplicate a sequence of events in real time

- Chimpanzees and other primates are not capable of entrainment
- Until recently, it was thought that only humans could entrain, but...

Entrainment is fundamental to musical behaviour

- Musicians need to anticipate and reciprocate rhythm

What is more, humans really really like to entrain:

- tapping a foot to a beat
- clapping along with a song
- walking together in step

Hypothesis on entrainment

If entrainment and the associated affect was an early development (in human history),...

...then pairs and groups of humans would have enjoyed entraining together

In turn, this would be likely to increase social bonding...

...which is an evolutionary advantage for weak organisms like hominids...

...and which would in turn reinforce the genetic basis of the entraining behaviour

What is computational modelling of cognition?

Queen Mary
University of London

It is difficult to study minds

- you can't see them
- you can't stick electrodes in them
- their relationship with brains is almost completely unclear
- it is unethical to distort/deform them for testing purposes
- *etc*

Before the advent of computers, psychologists had two means of study:

- look at what happened when things went wrong
- make predictions from theory about what would happen in certain precise circumstances (hypotheses), and test them (experiments)

This is very time-consuming (decades, not hours), error-prone, and (in the first case) dependent on chance

What is computational modelling of cognition?

Queen Mary
University of London

With computers, however, new things become possible

We can write computer programs which embody theories and then test them to destruction (ethically!)

We can also make predictions by computer which can then be tested in experiments with humans

This can be much faster than the human-driven approach

It is more objective than the human-driven approach (so long as the program is written objectively)

What's the point?

Queen Mary University of London

This is really the only (ethical) way to understand how a cognitive phenomenon actually works

- duplicate it in an artificial system and test that to destruction
- if it matches human behaviour in all circumstances, it is a good model
- it's important to choose and stick to your level of abstraction

If you can write a program which embodies your theory, then your theory is fully worked through (a Very Good Thing)

How do we build a cognitive model?

Apply reductionist science!

- accept that most phenomena are too complex to understand all at once
- identify part(s) of the phenomenon that are (as) separable (as possible)
- be careful to use stimuli (music) that do not go beyond these boundaries
- remember that the resulting model is probably an oversimplification
- when you have understood the parts of the phenomenon, put them together, study the interactions between them, and test them in concert

This is quite different from, and antagonistic to, the holistic view usually taken in the humanities, but it is not incompatible

Human (musical) behaviour *must* be at the start and end of this process:

- theories behind the models come from observation of musical behaviour
- results from models are tested against musical behaviour

What are the limitations of cognitive modelling?

Queen Mary University of London

A model is only as good as

- the theory it embodies
- the computational implementation
- the input data
- the input and output data representation

We must always question and test (and re-test) results because of these potential sources of error

What are the limitations of cognitive modelling?

We can only take one small step at a time

- this science is in its infancy: we must not rush ahead and make mistakes

Therefore, we have to be satisfied with small, focused, isolated results

- we look at how a given aspect of something changes, given that everything else stays the same – an artificial situation

The results are only ever approximations

- we continue to refine models as our understanding improves

What are the requirements of a cognitive model?

Queen Mary
University of London

We must be careful to make the right *abstraction* of our data

- A representation based on a 12-note octave will not be able to model phenomena related to microtonal music
- A representation based on a 12-note octave will not be able to model phenomena related to conventional tonal tuning (eg playing into the key)

And that leads to thinking carefully about the abstraction of the model

- (NB cf: define universe or define conceptual space?)

A very good abstraction of Western Common Practice music: the score

- models categorical pitch and time perception (and tonality if need be)
- evolved over about 1,000 years to do this *well*
- not good for everything (eg no means of representing instrumental timbre)
- but very good at quite a lot!

Many cognitive models of music use (an equivalent of) score notation

Two kinds of cognitive model

Queen Mary University of London

Some models are **descriptive** (Wiggins, 2007, 2011)

they say what happens when stimuli are applied in each circumstance

the predict results in terms only of the application of rules

these rules may be complicated

these models do not explain WHY a cognitive effect is the way it is

they do explain WHAT the cognitive effect is, at the same level of abstraction as the representation they use

Some models are **explanatory** (Wiggins, 2007, 2011)

they give a **general underlying mechanism** by which a phenomenon occurs

they predict results using this mechanism

they explain WHY a cognitive effect is the way it is (possibly at some level of abstraction different from the representation)

Example 1: GTTM

Generative Theory of Tonal Music (Lerdahl & Jackendoff, 1983)

- "complete" theory of tonal music (actually not – still being updated)
- has 4 components, each being a set of rules, written in English
  - grouping
  - metre
  - time-span reduction
  - prolongation
- within each, there are two kinds of rule
  - fixed rules
  - preference rules
- "preference" rules without conditions for application mean that GTTM is not a computerisable theory
- therefore, it is not a rigorously objective model
- it is only a descriptive model, because no mechanism is given

Towards an explanatory model

Queen Mary University of London

Choice of domain

- We have to apply reductionist methodology (creativity, or even just music, is too complicated to model in one go)
- We should aim for ecological validity - so stimuli are as natural as possible
- We should aim to reduce "with the grain" of the domain

Choice of model

- Programmed, rule-based - probably not explanatory
- Learning based - possibly explanatory, depending on how it works
  - Needs to be unsupervised
- Therefore we need an underlying theory motivating how the model works

Learning-based models of music

Queen Mary University of London

There is plenty of evidence (all around us) that the ability to behave musically is universal(ly valued)

However, different (sub-)cultures have different musics

Music from other (sub-)cultures is often incomprehensible

- but it's based mostly on the same constructs
  - rhythm/meter
  - notes
  - patterns
  - repetition
  - timbre (tone colour)

Learning-based models of music

Queen Mary
University of London

All this suggests a collection of evolved perceptual mechanisms which combined to create music cognition

There is no reason at all why these things need to be otherwise connected

Then, we hypothesise, the musical experience is derived from the processing of percepts at this level by some general mechanism

However, the musical culture is learned *implicitly*

That is to say that a good model will not require explicit training

- in other words, we don't tell it what the outputs we expect are

A general perceptual computation

Queen Mary University of London

Problem: information is everywhere

There is too much information for even a brain to process

Possible solution: compress it as it arrives

We hypothesise that brains use *structural compression* to help manage the continuous information overload

So, for example, a chair is perceived as a chair, and not as a set comprising a seat, four legs and a back

Then, when we see many chairs, it is more efficient to represent them all as references to a definition of chairness, rather than as a detailed description of each one individually

Evolutionary perspective

Key evolutionary points

- organisms survive better if they can learn
- organisms survive better if they can anticipate
- organisms survive better if they can anticipate from what they learn
- organisms cannot be merely reactive
  - anticipation must be proactive – it must result in embodied action
- organisms must regulate cognitive resource – attention is expensive

Need to evolve a mechanism where organisms can learn without damage

- gives rise to "tension": emotional warning of uncertainty in the organism's world model
- cf. musical tension; narrative tension

Managing information in time

Queen Mary University of London

The same principle applies to things arranged in time as to things arranged in space

Music is often called "a time-based art" because it has no instantaneous existence (except *perhaps* in the minds of some highly skilled musicans)

And there are many other situations where the ability to react to sequences of events in time at a perceptual level is evolutionarily useful

The key property that admits this is...

EXPECTATION

Expectation allows us to deal with the world

- there is too much data out there to process in real time
- we need to manage it by predicting what comes next, so we have a chance to get ahead

Expectation works in many domains

- vision
- movement
- speech
- music

Managing information in time

Queen Mary
University of London

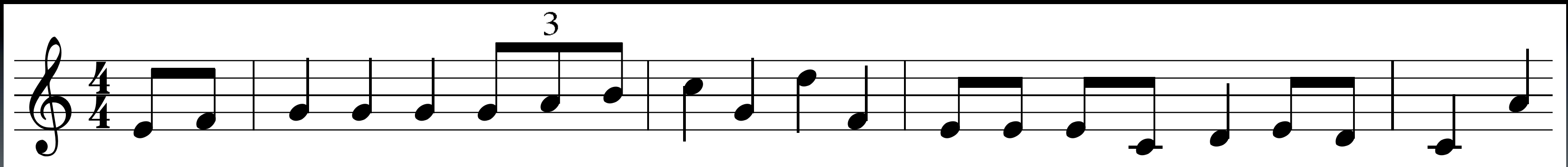
EXPECTATION

There must be a mechanism that

- learns from data
- predicts from data
- generalises from data (so it can deal with data it hasn't seen before)

It will also be able to enlist cognitive resources

- so that unexpected things can be dealt with



EXPECTATION

In speech and language understanding

it's easy to wreck a nice beach

Managing information in time

Knowing things can reduce the amount of information required to transmit information

Claude Shannon (1948) proposed a mathematical model of this idea, which he called "information theory", in the context of telecommunications engineering

It turns out that Shannon information theory works very well in a model of human perception based on (advanced) Markov Models

Skip Markov demo

Markov models

Queen Mary University of London

A very neat way to do this is by Markov Modelling

Each possible symbol in a stream of symbols is recorded, along with each of the contexts in which it is experienced (n-grams; n=2)

a b c b d e	* | a	a | b	b | c	c | b	b | d	d | e
a b c a b e	* | a	a | b	b | c	c | a	a | b	b | e
a b d b d e	* | a	a | b	b | d	d | b	b | d	d | e

Markov models

A very neat way to do this is by Markov Modelling

Each possible symbol in a stream of symbols is recorded, along with each of the contexts in which it is experienced (n-grams; n=2)

<table>
<tr><td>a b c b d e</td><td>* | a</td><td>a | b</td><td>b | c</td><td>c | b</td><td>b | d</td><td>d | e</td></tr>
<tr><td>a b c a b e</td><td>* | a</td><td>a | b</td><td>b | c</td><td>c | a</td><td></td><td>b | e</td></tr>
<tr><td>a b d b d e</td><td>* | a</td><td>a | b</td><td>b | d</td><td>d | b</td><td>b | d</td><td>d | e</td></tr>
<tr><td></td><td></td><td>a | b</td><td></td><td></td><td></td><td></td></tr>
</table>

Markov models

A very neat way to do this is by Markov Modelling

Each possible symbol in a stream of symbols is recorded, along with each of the contexts in which it is experienced (n-grams; n=2)

a b c b d e	* | a	a | b	b | c	c | a	d | b
a b c a b e	* | a	a | b	b | c	c | b	d | e
a b d b d e	* | a	a | b	b | d		d | e
		a | b	b | d		
			b | d		
			b | e		

Markov models

A very neat way to do this is by Markov Modelling

Each possible symbol in a stream of symbols is recorded, along with each of the contexts in which it is experienced (n-grams; n=2)

a b c b d e	* | a	a | b	b | c	c | a	d | b
a b c a b e		a | b	b | c	c | b	d | e
a b d b d e		a | b	b | d		d | e
		a | b	b | d		
			b | d		
			b | e		

Markov models

A very neat way to do this is by Markov Modelling

Each possible symbol in a stream of symbols is recorded, along with each of the contexts in which it is experienced (n-grams; n=2)

<table>
<tr><td>a b c b d e</td><td>* | a</td><td>a | b</td><td>b | c</td><td>c | a</td><td>d | b</td></tr>
<tr><td>a b c a b e</td><td></td><td></td><td>b | c</td><td>c | b</td><td>d | e</td></tr>
<tr><td>a b d b d e</td><td></td><td></td><td>b | d</td><td></td><td>d | e</td></tr>
<tr><td></td><td></td><td></td><td>b | d</td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td>b | d</td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td>b | e</td><td></td><td></td></tr>
</table>

Markov models

A very neat way to do this is by Markov Modelling

Each possible symbol in a stream of symbols is recorded, along with each of the contexts in which it is experienced (n-grams; n=2)

<table>
<tr><td>a b c b d e</td><td>* | a</td><td>a | b</td><td>b | c</td><td>c | a</td><td>d | b</td></tr>
<tr><td>a b c a b e</td><td></td><td></td><td>b | d</td><td>c | b</td><td>d | e</td></tr>
<tr><td>a b d b d e</td><td></td><td></td><td>b | e</td><td></td><td>d | e</td></tr>
</table>

Markov models

Queen Mary University of London

A very neat way to do this is by Markov Modelling

Each possible symbol in a stream of symbols is recorded, along with each of the contexts in which it is experienced (n-grams; n=2)

a b c b d e	* | a	a | b	b | c	c | a	d | b
a b c a b e			b | d	c | b	d | e
a b d b d e			b | e		

Markov models

* | a	a | b (1.0)	b | c (0.333)	c | a (0.5)	d | b
		b | d	c | b	d | e
		b | e (0.167)		

Markov models

Now, given a partial leftmost string, we can estimate the probability distribution of the next unseen symbol

	1.0	0.333	0.5	
* | a	a | b	b | c	c | a	d | b
		b | d	c | b	d | e
		0.167		
		b | e		

Markov models

Queen Mary University of London

Now, given a partial leftmost string, we can estimate the probability distribution of the next unseen symbol

*	a
* a	b
* a b	c or d or e
* a b c	a or b
* a b c b	c or d or e
* a b c b d	b or e
* a b c b d e	

	1.0	0.333	0.5	
* | a	a | b	b | c	c | a	d | b
		b | d	c | b	d | e
		0.167		
		b | e		

Markov models

Queen Mary
University of London

Now, given a partial leftmost string, we can estimate the probability distribution of the next unseen symbol

*	a
* a	b
* a b	c or d or e
* a b c	a or b
* a b c b	c or d or e
* a b c b d	b or e
* a b c b d e	

	1.0	0.333	0.5	
* | a	a | b	b | c	c | a	d | b
		b | d	c | b	d | e
		0.167 b | e		

p(a b c b d e) =
1.0 x 1.0 x 0.333 x 0.5 x 0.5 x 0.667 = 0.111

p(a b c a b e) =
1.0 x 1.0 x 0.333 x 0.5 x 1.0 x 0.167 = 0.028

p(a b d b d e) =
1.0 x 1.0 x 0.5 x 0.333 x 0.5 x 0.667 = 0.022

Markov models

Queen Mary University of London

Now, given a partial leftmost string, we can estimate the probability distribution of the next unseen symbol

*	a
* a	b
* a b	c or d or e
* a b c	a or b
* a b c b	c or d or e
* a b c b d	b or e
* a b c b d e	

	1.0	0.333	0.5	
* | a	a | b	b | c	c | a	d | b
		b | d	c | b	d | e
		0.167 b | e		

p(a b c b d e) =
1.0 x 1.0 x 0.333 x 0.5 x 0.5 x 0.667 = 0.111

p(a b c a b e) =
1.0 x 1.0 x 0.333 x 0.5 x 1.0 x 0.167 = 0.028

p(a b d b d e) =
1.0 x 1.0 x 0.5 x 0.333 x 0.5 x 0.667 = 0.022

p(a b d e) = 1.0 x 1.0 x 0.5 x 0.667 = 0.333

Ingredients for the model

Queen Mary University of London

We have

- an underlying mechanism (Markov models & Shannon information theory)
- a hypothesis to motivate it
- a creative subdomain (musical melody)
- a representation (essentially, the musical score)

We need

- data (The Essen Folksong Collection: 907 tonal folk melodies)
- a computational implementation (Pearce, 2005)

IDyOM: unexpectedness

Queen Mary
University of London

Shannon (1948) defines a measure of *entropy*, which has been interpreted in several ways in the literature

We interpret it in two ways here:

the entropy of a given note is a measure of its *unexpectedness in context*

the entropy of the distribution of an unseen note is a measure of the model's *uncertainty in context*

- *unexpectedness*, or *information content*, of a note with probability p is defined as

$$-\log_2 p$$

- Does this quantity model human listeners' estimates of their own perception of expectedness when listening?

Information Dynamics of Music

Queen Mary University of London

Middle layer of cognitive model of conscious musical experience

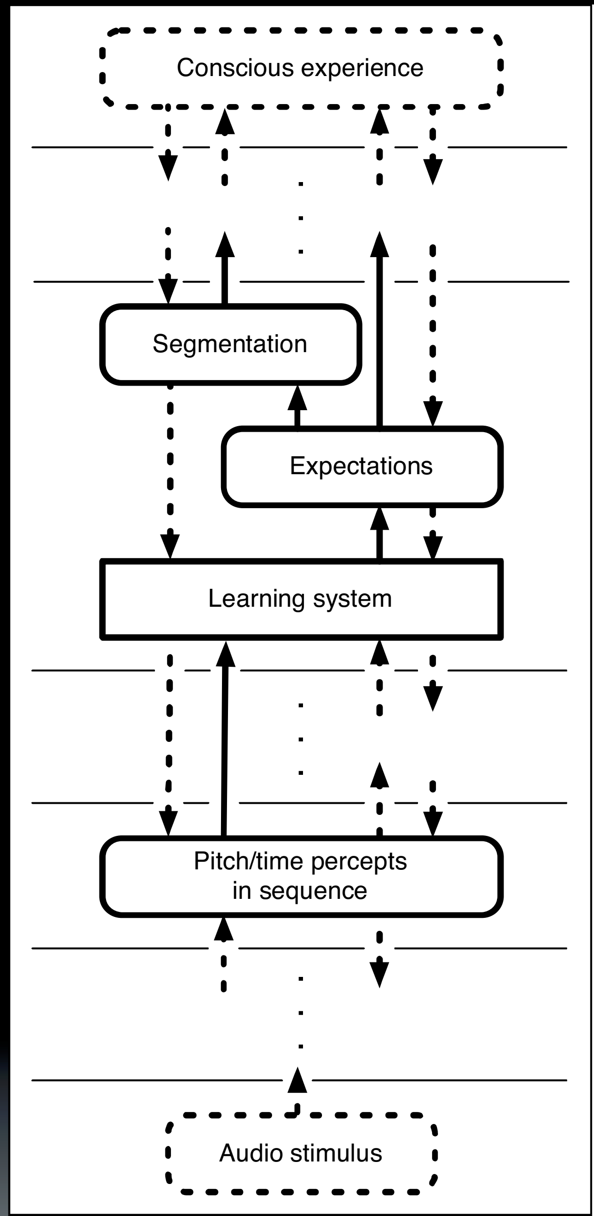
Unsupervised, implicit learning

Inputs are sequences of basic percepts

- notes, with pitch & time features
- derived percepts, e.g.,
  - interval
  - tonal centre

Outputs are

- distributions of predicted pitches
- information-theoretic derivatives of distributions



The IDyOM model

Queen Mary
University of London

IDyOM = Information Dynamics of Music

Model assembled and evaluated by Marcus Pearce (2005)

Uses Markov models as a simulation of perception of events in time-sequence

Clever implementation using suffix trees

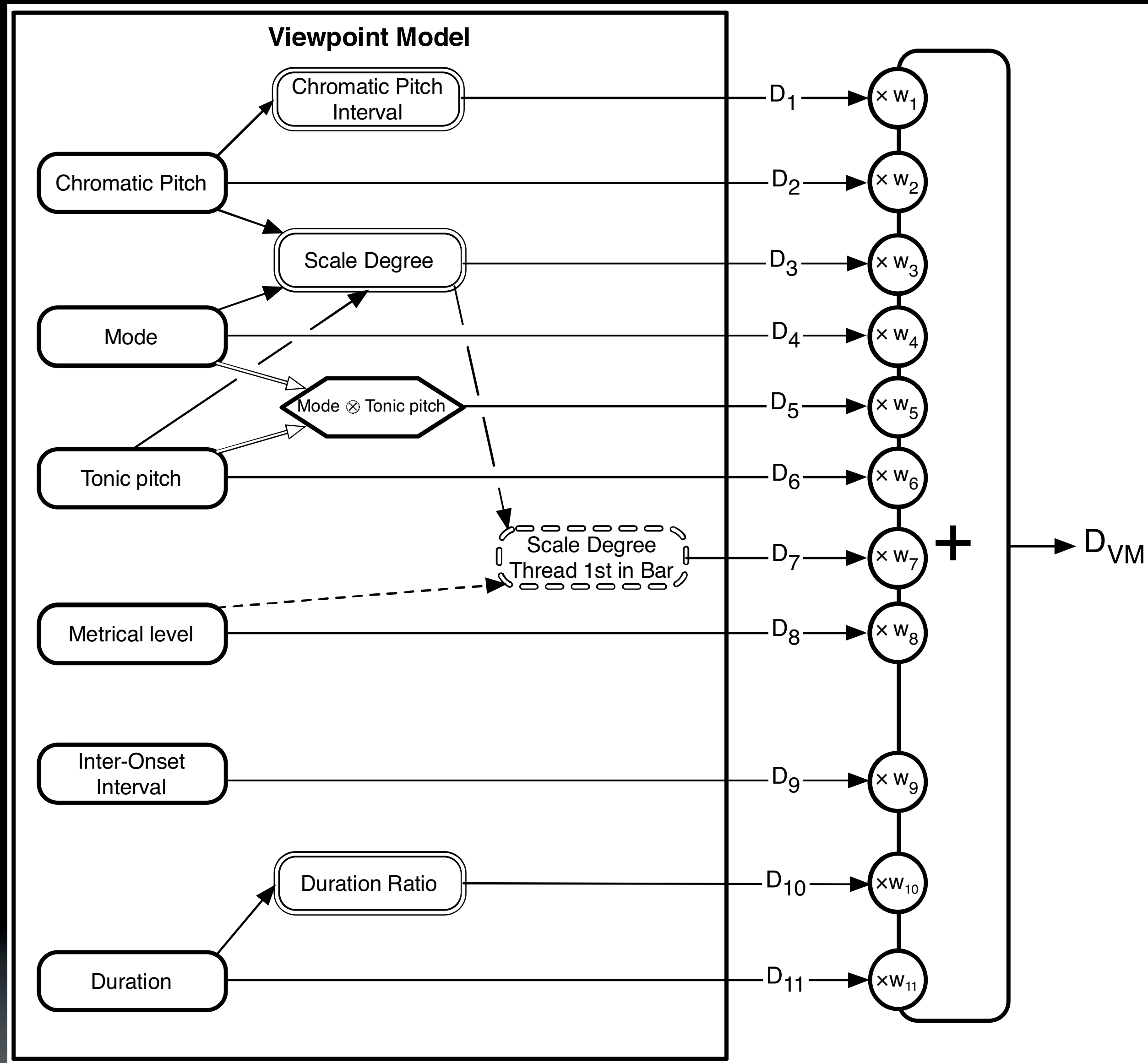
Implicit learning: learns the likelihood of each symbol appearing in a sequence from mere exposure, then predicts from this information

Estimates the likelihood of unseen symbols (uniform distribution)

Uses Shannon information theory to weight different components of the model as they contribute to a combined distribution

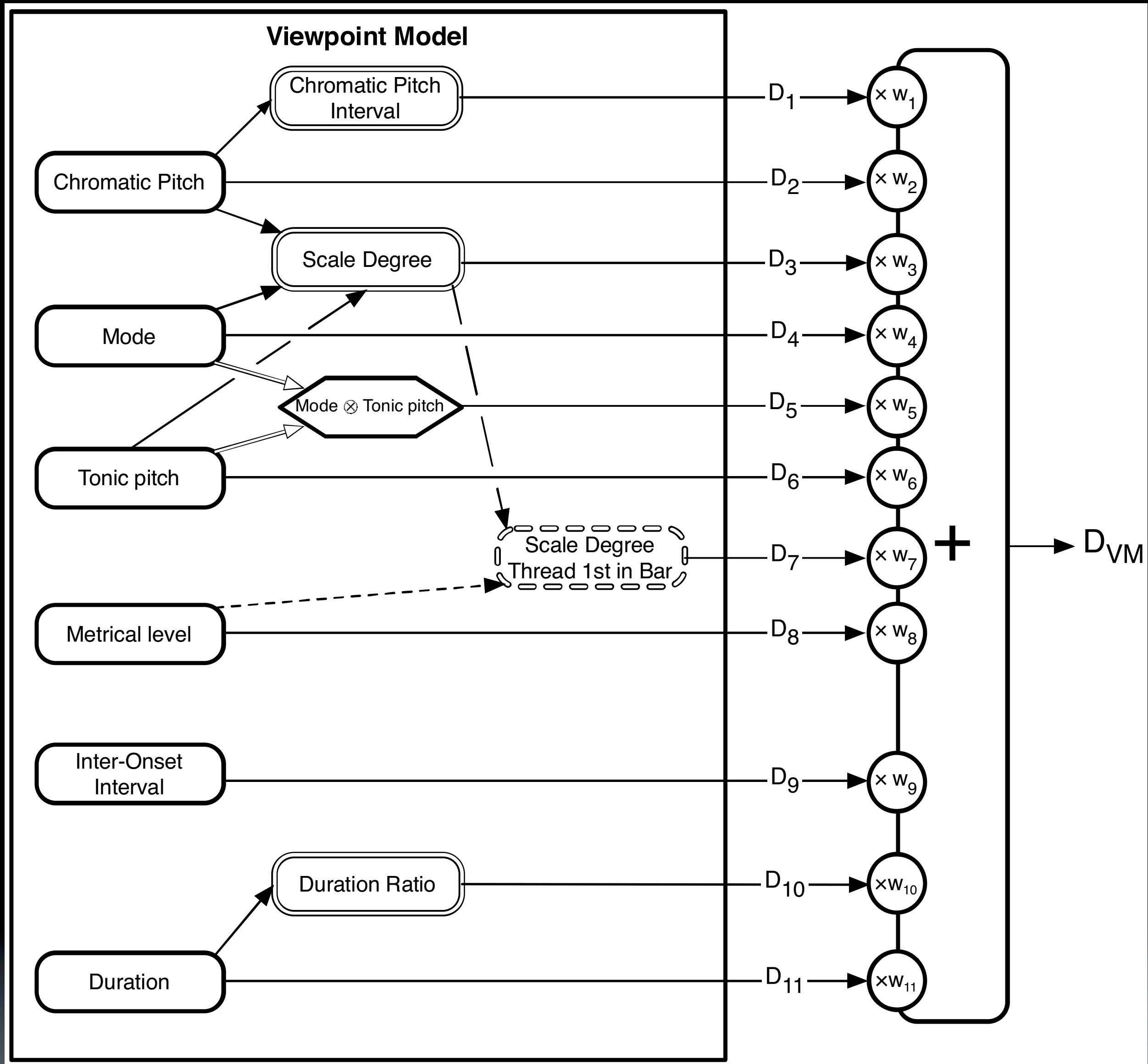
Efficient implementation of simple Markov chains

- but with multi-dimensional symbols
- select feature sequences (*viewpoints*)
  - basic
  - derived
- calculus of viewpoints
  - differentiation (delta)
  - cross-product (pairing)
  - thread (sub-sequence selection)



Information Dynamics of Music

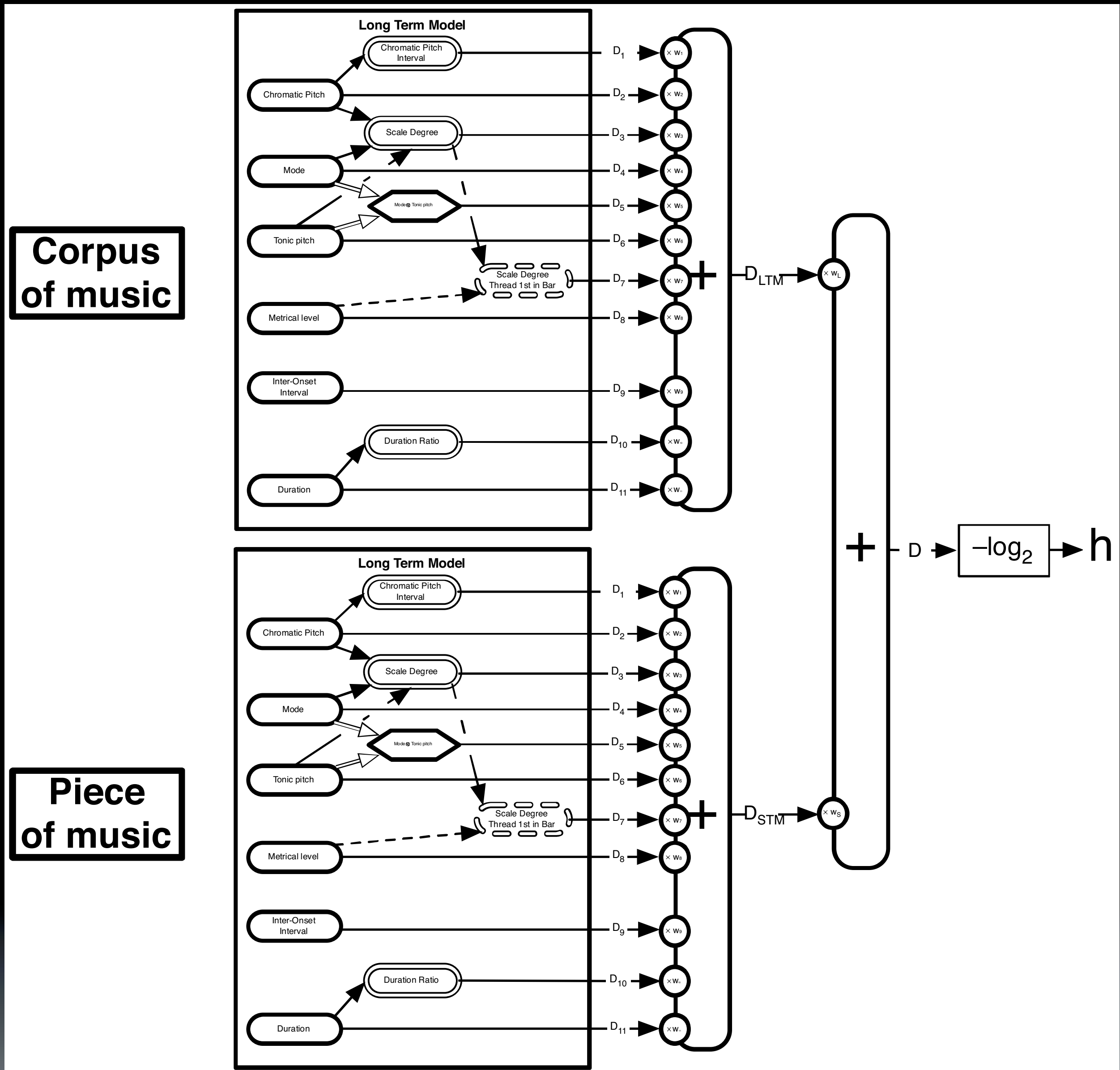
Queen Mary
University of London



Predictions made by

- matching current context with strings in memory
- all orders between 0 and maximum available
- all contribute to final distribution
- Feature predictions combined as linear sum weighted by entropy

Information Dynamics of Music



Combined outputs of two models

- one exposed to corpus of "enculturation" data
- one exposed only to current melody
- Combination is by entropic weighting, as before

Model is "optimised"

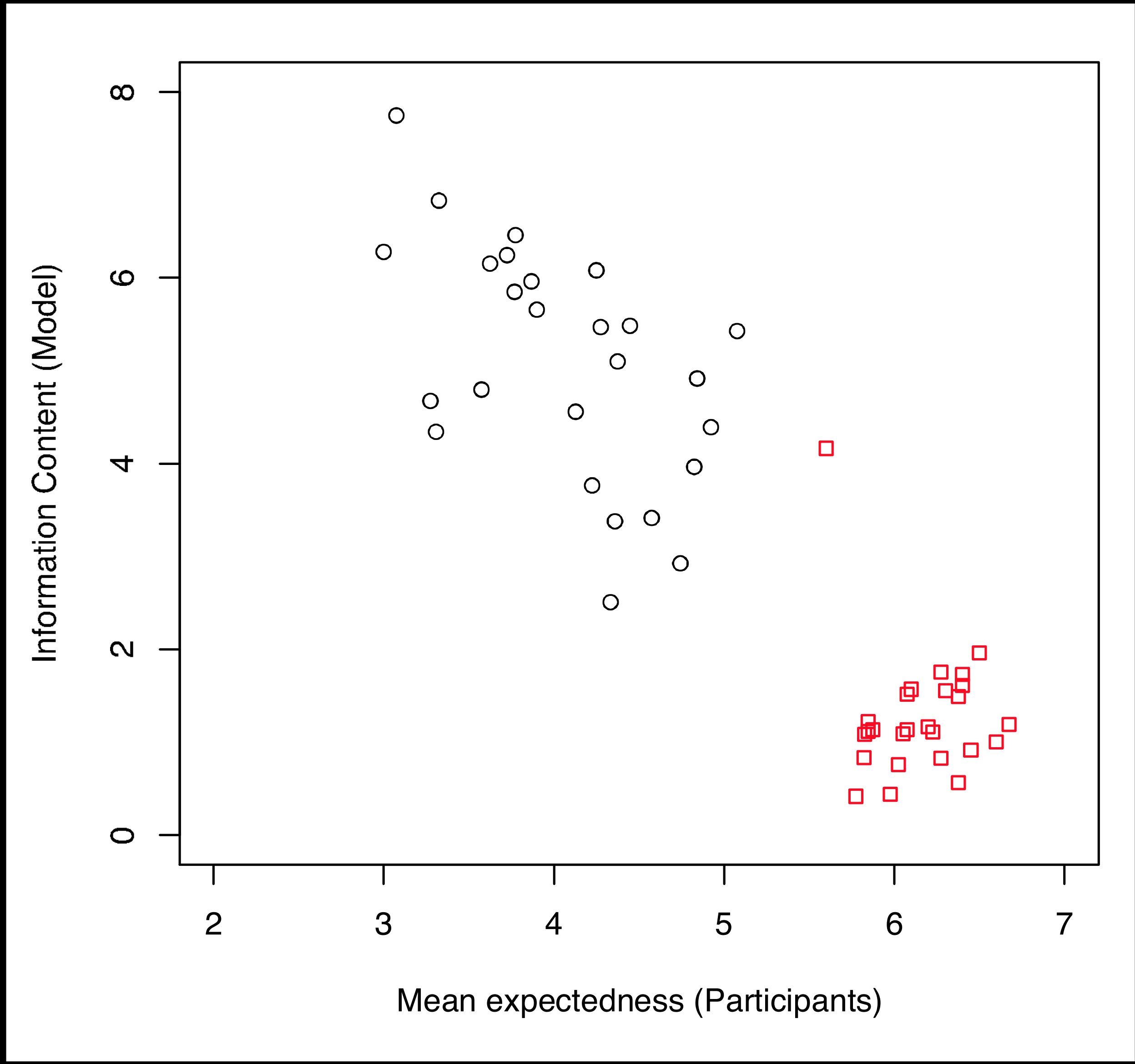
- inefficient viewpoints are discarded
- model with lowest average information content is used

Information Dynamics of Music

Queen Mary University of London

IDyOM predicts

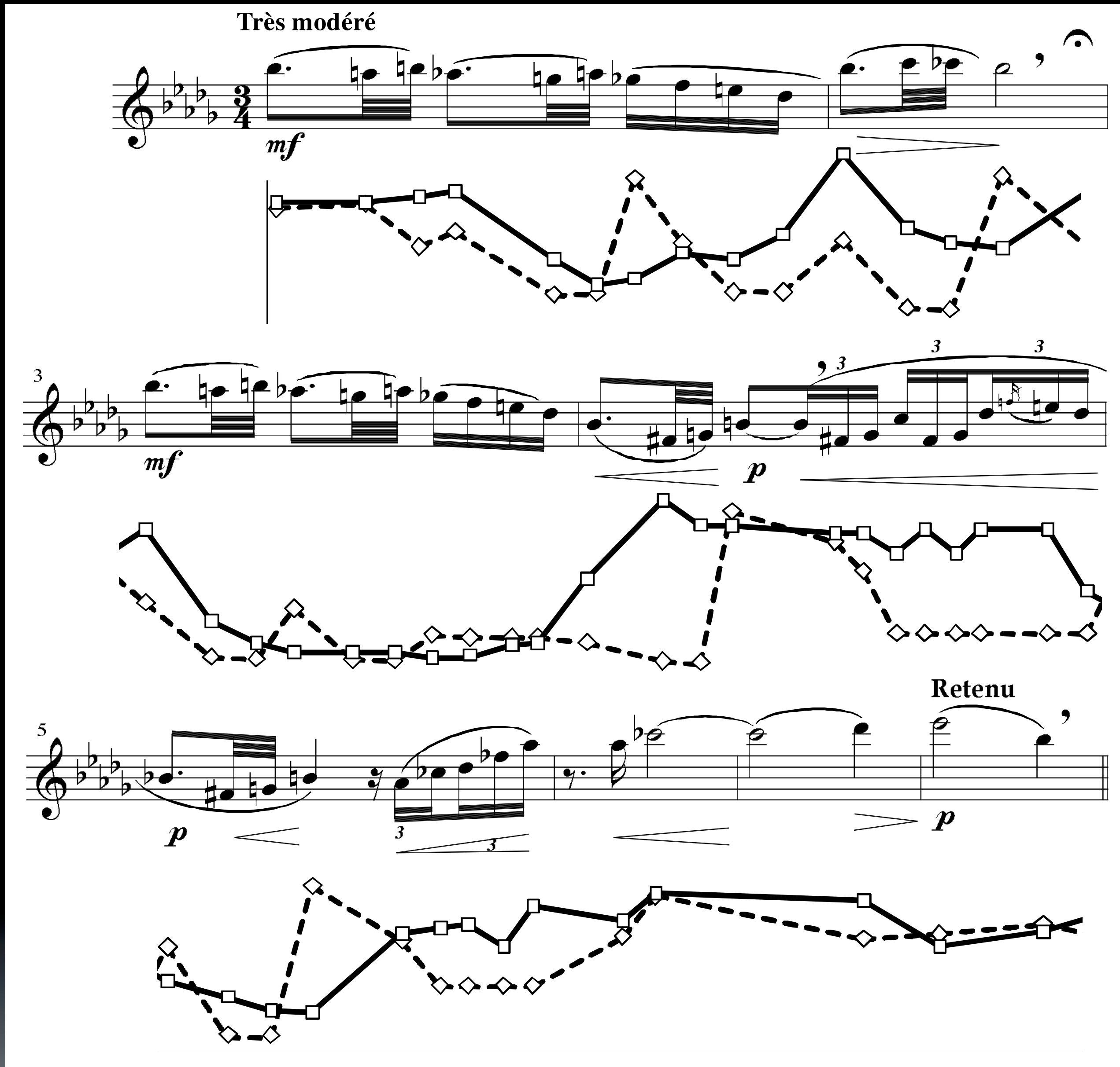
- listener's expectations of next note in melody
  - 4 studies; up to r=.91 correlation
  - 1 study; very high correlation with musicologists' predictions



Information Dynamics of Music

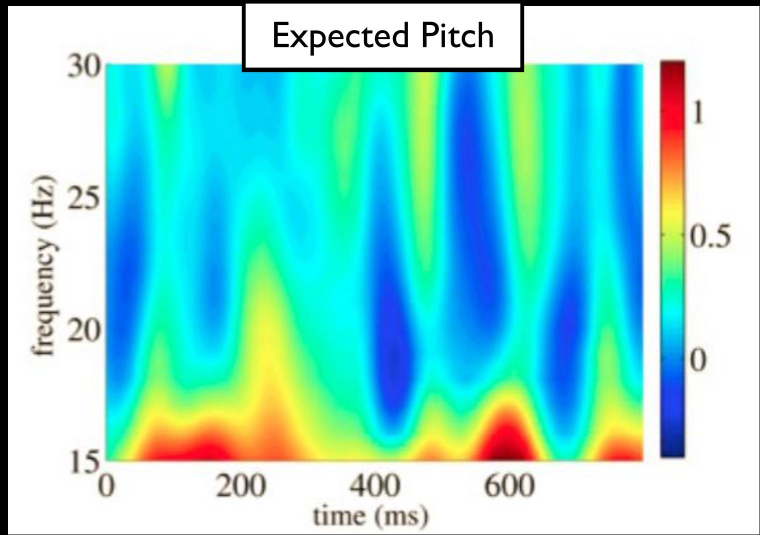
Queen Mary
University of London

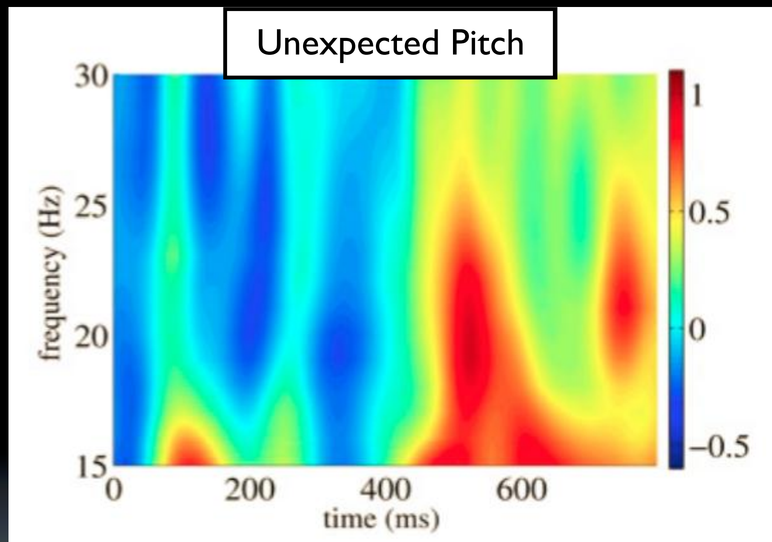
IDyOM predicts

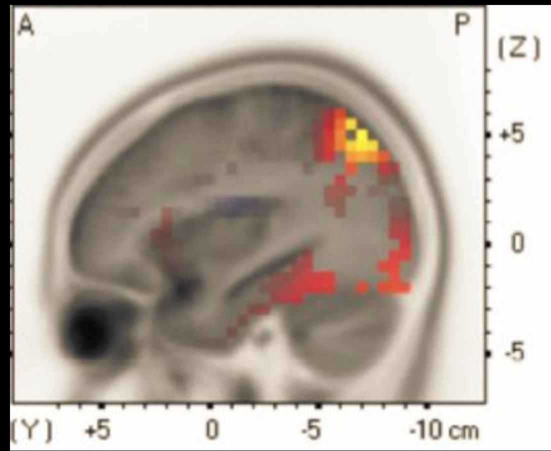
- listener's expectations of next note in melody
  - 4 studies; up to r=.91 correlation
  - 1 study; very high correlation with musicologists' predictions
- melodic segmentation
  - 2 studies;   = 0.58
  - vs musicologist judgements

Information Dynamics of Music

IDyOM predicts

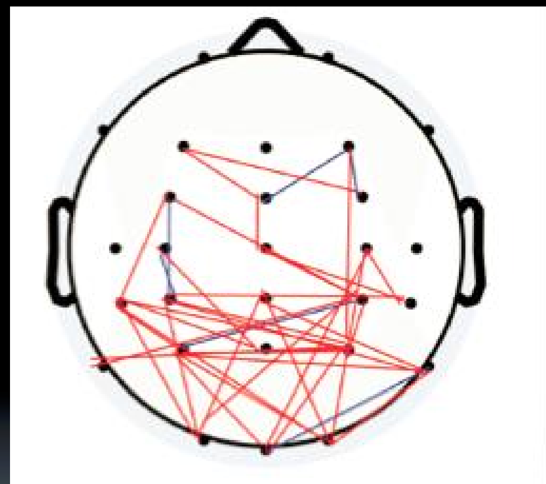
- listener's expectations of next note in melody
  - 4 studies; up to r=.91 correlation
  - 1 study; very high correlation with musicologists' predictions
- melodic segmentation
  - 2 studies; = 0.58
  - vs musicologist judgements
- neural activation with unexpectedness
  - centro-parietal region
  - strong sync. in beta-band







Unexpected – Expected



Interim conclusions

Queen Mary
University of London

IDyOM accounts for the vast majority of the data

- with no programmed musical knowledge
- with no training
- without advanced musical concepts, like harmony

This is evidence not only for IDyOM as a model of pitch perception but also for the Markovian idea of statistical perception in general

- there is similar evidence in computational linguistics

But it is possible to give further, stronger evidence

Meta-model

We introduce the concept of a *meta-model*

This is a model

which *reuses* an existing model

without changing it

to do something related but different

(meta = “beyond”)

- Such a meta-model constitutes strong evidence that the original model is doing something veridical

the model can do more than one thing

if these things are related, then this evidence that the model has captured the process involved in a strong way

IDyOM: boundary detection

Narmour (1990) hypothesises that musical phrasing is derived from the expectations generated and then realised or denied as a melody proceeds

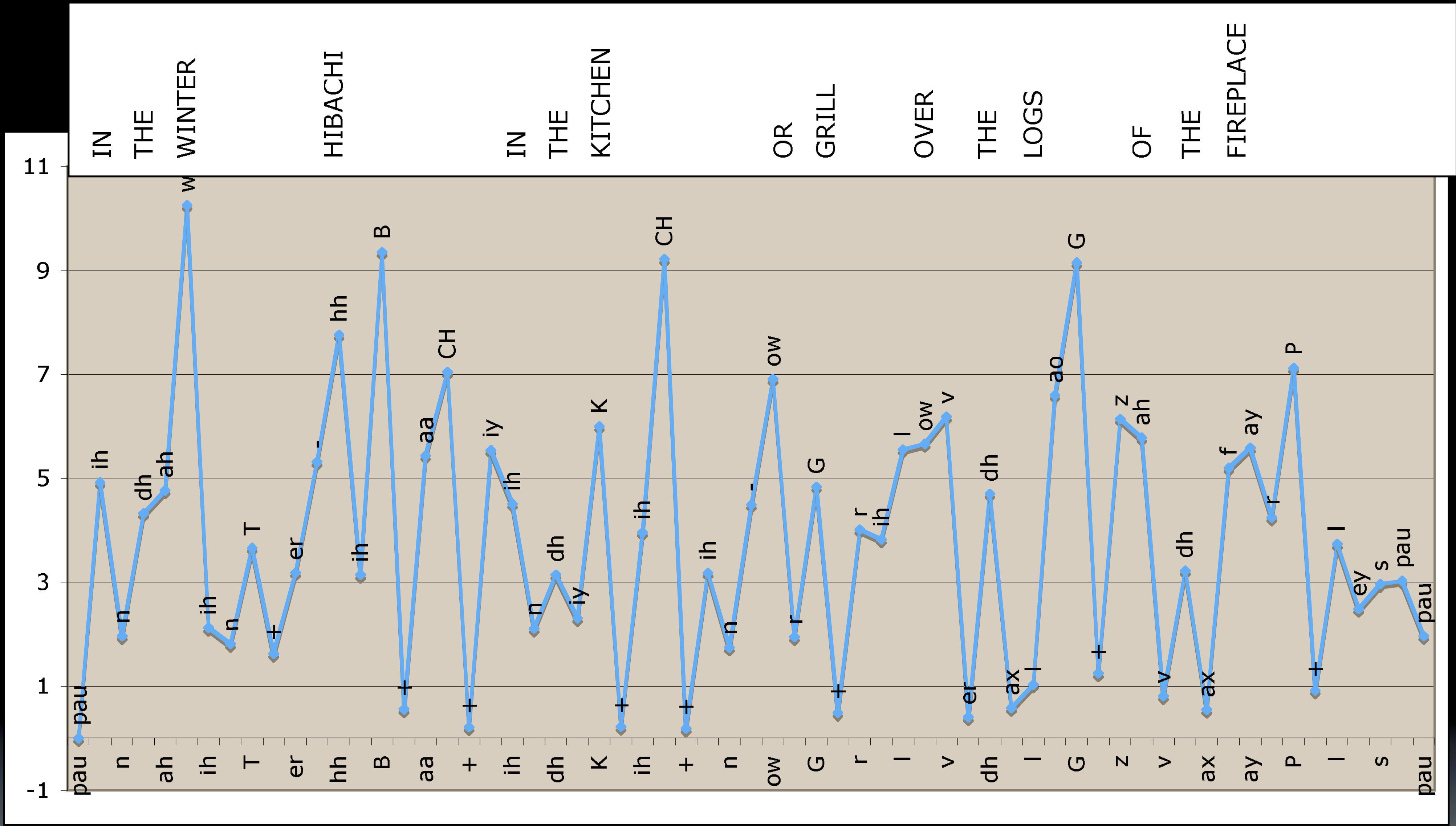
As a phrase ends, notes become more expected; when a new phrase starts, there are only weak expectations of what is to come next

So peak-picking in unexpectedness should allow us to predict boundaries

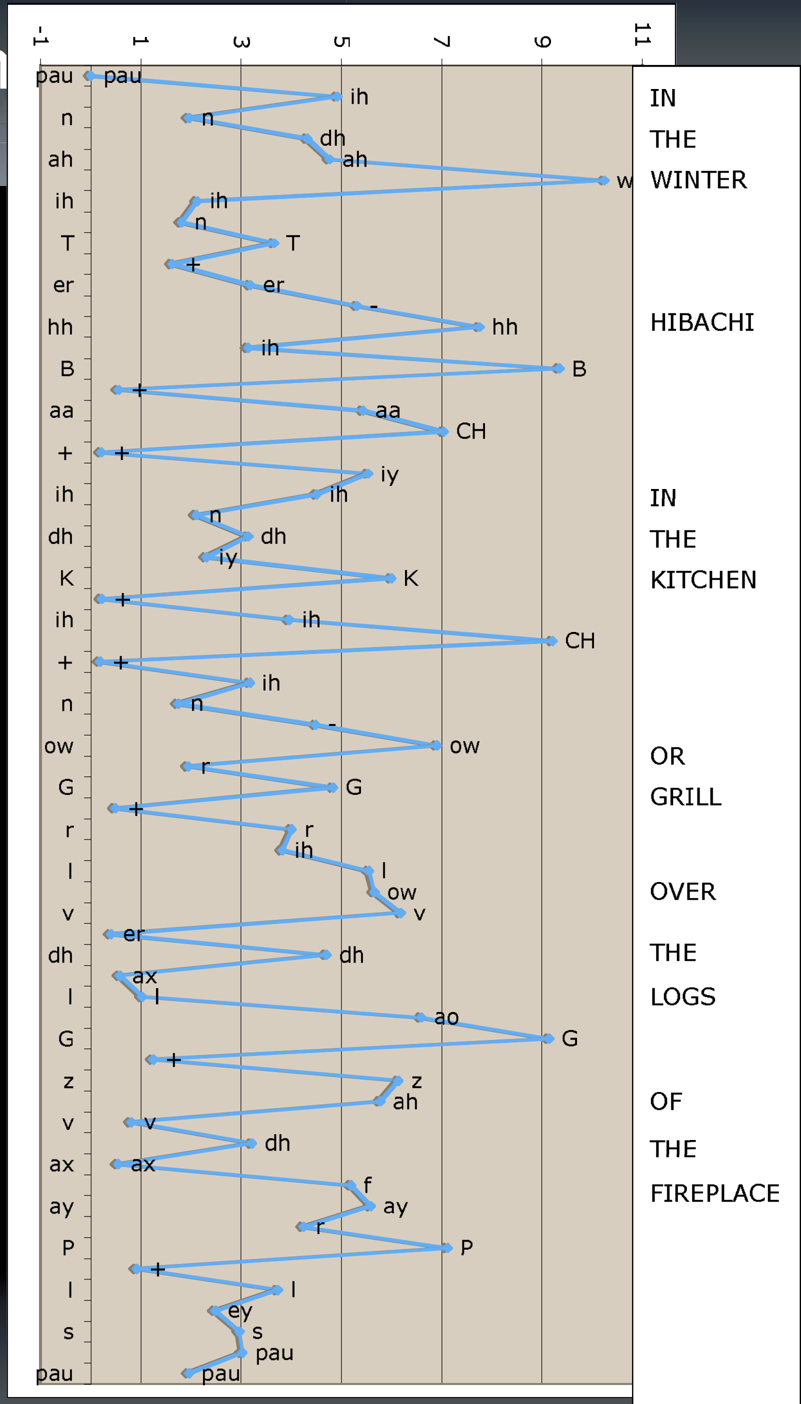
IDyOM: boundary detection

	Precision (1 - specificity)	Recall (sensitivity)	F1	d'	K
Grouper	0.67	0.87	0.76	2.94	0.73
GPR2a	0.95	0.55	0.70	2.93	0.68
LBDM2001	0.86	0.57	0.69	2.61	0.67
IDyOM	0.60	0.63	0.61	2.15	0.58
SimpleSeg.	0.25	0.35	0.29	0.99	0.22
GPR3a	0.16	0.37	0.22	0.69	0.12
always	0.08	1	0.15	-	-0.86
GPR2b	0.13	0.19	0.15	0.38	0.07
saffran.p.pitch	0.10	0.04	0.06	0.14	0.01
never	0	0	-	-	-0.04

Language segmentation



Language segmen



Queen Mary
University of London

Summary

Summary

Queen Mary University of London

Learned model and sequence retrieval method form **R**

Metropolis sampling forms **T**

Can perceptual models be expected reliably only to generate the things they perceive? (Relationship between **R** and **T**)

No

How can the quality of such generation be rigorously evaluated?

NB. Difference between

our evaluation of the scientific contribution

the system’s evaluation of its own outputs (**E**)

Context

Queen Mary
University of London

The models we are studying focus on the conceptual space, defined by the notional rule set

- in IDyOM this is learned, and stored in the LTM
- the advanced features of the retrieval system enhance the basic likelihoods with a kind of generalisation

This work does not attempt to address , the capacity of a creative system to introspect on its own output quality

The traversal strategy, , not really addressed in IDyOM

- uses a standard statistical optimisation method, metropolis sampling

Evaluation

A study of **IDyOM**'s melody generation

Motivation: Evaluating outputs of creative systems

Method: Consensual Assessment Technique

Question 1: Can a model of expectation be reliably used as a generative model?

Question 2: How can quality of output be improved?

Motivation

Queen Mary University of London

How can we objectively evaluate the outputs of creative systems?

“

We listened to a large number of the generated tunes, and they sounded quite good.” (********, 1998)

We need to be much more rigorous than this

Motivation

Queen Mary University of London

Aesthetic evaluation is (often) primary

But we can evaluate scientifically too, and if we can, we should

We can evaluate our generative systems in terms of engineering:

- are they reactive enough?
- are they reliable?

But what does it mean to evaluate an aesthetic output scientifically?

Motivation

Queen Mary University of London

We need to say what we mean by "good"

For example:

- how well does the music generator keep in time (if that's what it's meant to do)
- how well does the music generator match the implied harmony of the input melody (if that's what it's meant to do)

Motivation

Queen Mary University of London

This means that we need to understand

- the music we are generating
- its context
- its purpose

Just like mono-disciplinary science, we need to formulate precise research questions (which may be as much aesthetic as scientific) and precise tests to evaluate them

An artefact's (e.g., music) not being precisely specified does **not** mean that we can't ask precise questions about it

Motivation

But how?

Creative judgements are (desirably)

- subjective
- context-dependent

How, then, can we be rigorous in evaluating the success or otherwise of creative systems output?

The Consensual Assessment Technique (Amabile, 1998)

Consensual Assessment

Queen Mary University of London

Technique originally used for assessing the creative content of outputs

Used here to assess stylistic success of outputs

The task must be open-ended enough to permit considerable flexibility and novelty in the response

Response must be an observable product which can be rated by judges

Consensual Assessment

Queen Mary University of London

Judges must

- be experienced in the relevant domain;
- make independent assessments;
- assess other aspects of the products, such as technical accomplishment, aesthetic appeal or originality;
- make relative judgements of each product in relation to the rest of the stimuli;
- be presented with stimuli and provide ratings in orders randomised differently for each judge.

Consensual Assessment

Queen Mary University of London

Most importantly, in analysing the collected data, the inter-judge reliability of the subjective rating scales must be determined

If and only if reliability is high, we may correlate creativity ratings with other objective or subjective features of creative products

In our version, we replace "creativity" ratings with "stylistic success" ratings

Study - Background

Can this perceptual model generate music, as some researchers (eg Baroni, 1999) suggest it should?

Model's primary data is pitch, not rhythm, so

- create new melodies in existing rhythmic frameworks

Use human experts to rate outputs

- work in a well-established, formally-studied style
- analyse success of generation in terms of stylistic success

Use Consensual Assessment Technique to produce reliable consensus judgements

Study - Hypotheses

Queen Mary
University of London

Three generation systems:

A. pitch only;

B. interval from 1st note of piece; scale degree x note duration; 1st note in phrase

C. complex representation, including harmonic implications of melody

Three null hypotheses:

Each system can generate melodies rated as equally stylistically successful in the target style as existing, human-composed melodies.

Study - Participants

Judges were 16 music researchers or students

5 were male and 11 female

Their age range was 20–46 years (mean 25.9, SD 6.5)

They had been formally musically trained for 2–40 years (mean 13.8, SD 9.4)

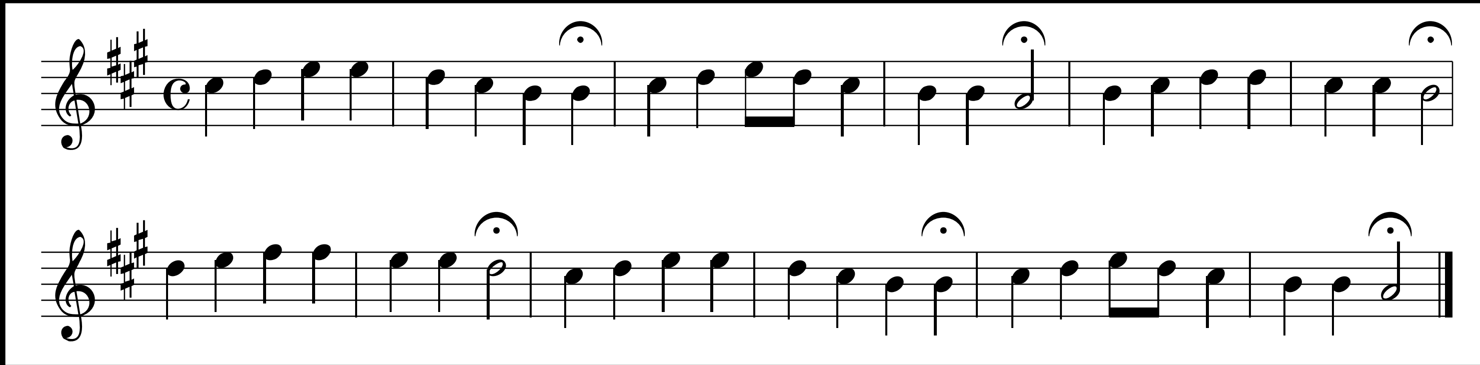
7 judges reported high familiarity with the chorale genre and nine were moderately familiar

All judges received a nominal payment, and worked for approximately an hour

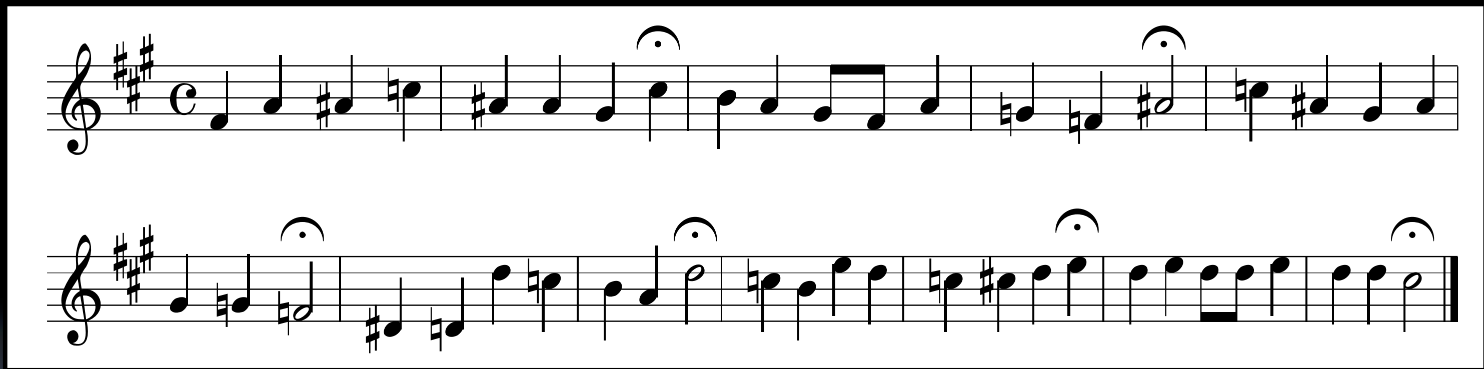
Study - Stimuli

Queen Mary
University of London

J. S. Bach: *Jesu, meiner Seelen Wonne* (BWV 359)



System A: *Jesu, meiner Seelen Wonne*



Study - Questions

Queen Mary
University of London

Questions:

- "How successful is this as a chorale melody?"
- Judges were advised to reflect such factors as
  - conformity to important stylistic features;
  - tonal organisation;
  - melodic shape and interval structure;
  - melodic form.
- and required to explain their answers
- "Do you recognise the melody?"

NB: participants not told that a computer was involved

- Moffat & Kelly (2006) demonstrate systematic bias against computers

Study - Results

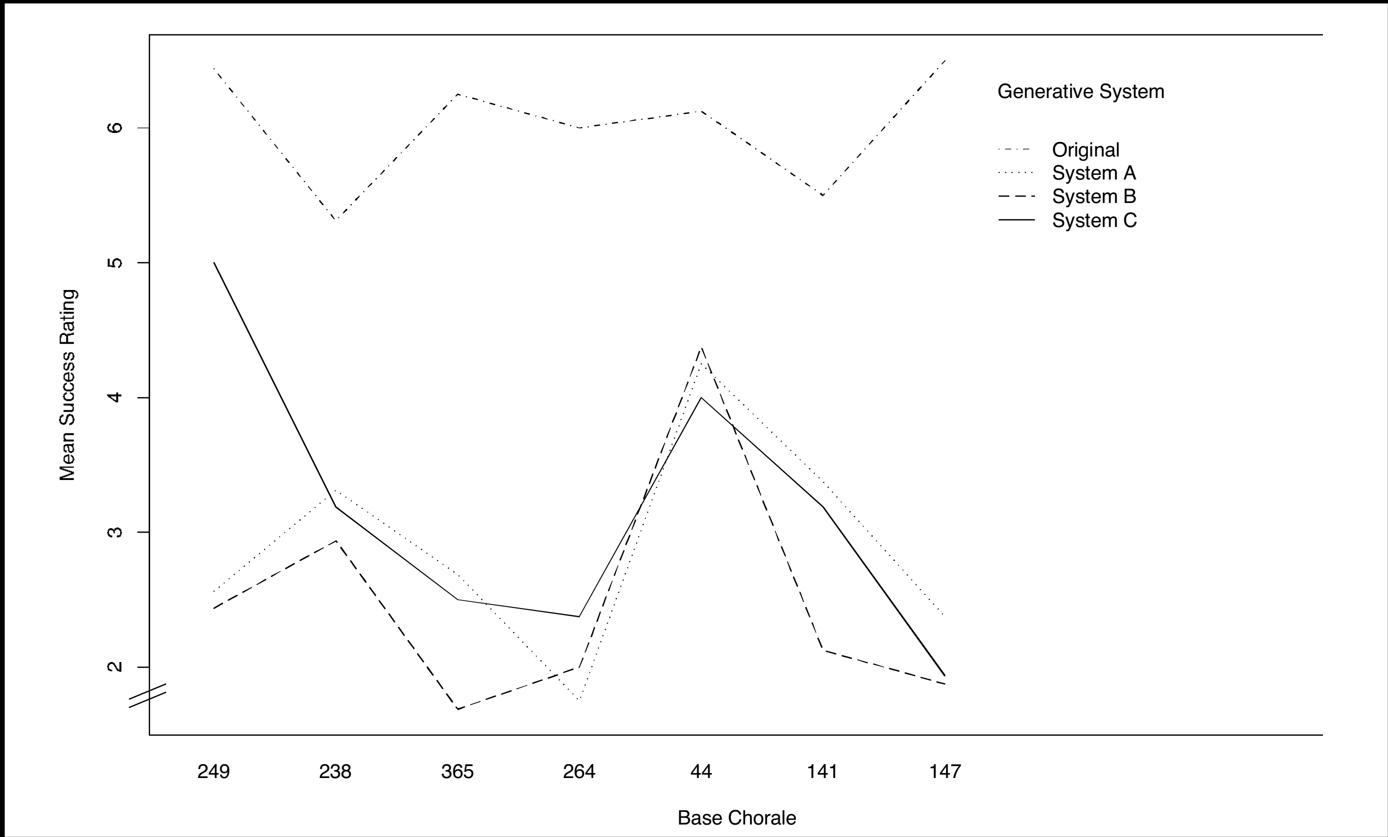
Queen Mary University of London

All but 2 of the 120 pairwise correlations between judges were significant ($p<0.05$) with a mean coefficient of $r(26)=0.65$ ($p<0.01$)

This high consistency warrants averaging the ratings for each stimulus across individual judges in subsequent analyses

Study - Results

Queen Mary
University of London



Friedman's Test shows significant intrasubject effect of system on ratings ($X^2(3)=32$, $p < 0.01$)

Wilcoxon rank sum test with Holm's Bonferroni correction show "original" ratings differ significantly from IDyOM's ($p < 0.03$)

Using the results

We examined the judges' explanations of their judgements

We developed a corresponding set of objective descriptors

We applied the descriptors in a multiple regression analysis

dependent variable: the rating scheme averaged across stimuli

Using the results

Queen Mary
University of London

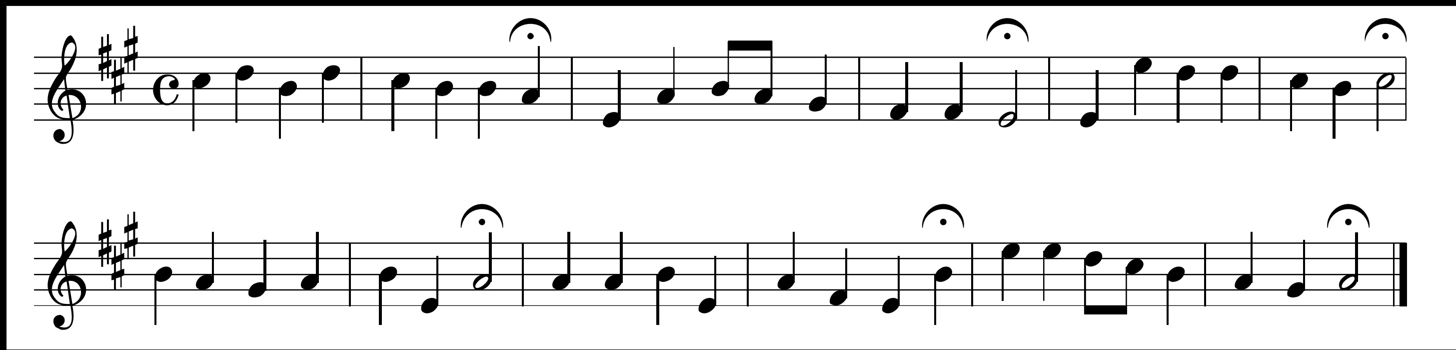
Objective features:

- Pitch Range
- Melodic Structure
- Tonal Structure
- Phrase Structure
- Rhythmic Structure

Redundant descriptors removed by backwards stepwise elimination

Using the results

System D: *Jesu, meiner Seelen Wonne*



Much more coherent example than earlier systems

But System D must be rigorously analysed in further cycles

What is going on here is transformational creativity:

the experimental results are being used to modify **R**

Summary of Results

Queen Mary
University of London

Systems A-D all fail to compare with original melodies

System D is, however, significantly better than Systems A-C

The CAT provides a rigorous way to assess the stylistic success of outputs of computational creative systems

Statistically significant agreement between expert judges is crucial in supporting the conclusions of the work

Skip to Summary

Summary

Two methods for establishing the quality of automated musical composition systems

- expert judgements on style
- non-expert judgements on meaning

Both are precisely specified (not just "good")

Two different techniques

- focused on particular knowledge, with articulate feedback on what is not right
- hidden amongst many different possible dimensions of quality, so as not to prime participants

It is up to researchers to establish what the dimensions of quality are for their own systems and their own work and to use them rigorously

In this way, perhaps we can also reach a more objective view of aesthetics

The Future

Many more models and many more studies are needed

Studying   is paramount (see, e.g., Guilford, Chikszentmihalyi for starting points)

Models that learn their rules are likely to be more credibly creative than programmed models

Music is a good domain for this study because it is uncluttered by semantics (in the language sense) so it can be looked at on the perceptual/cognitive level only

Summary

In this lecture we have looked at

a way of characterising creativity that is amenable to computational systems

a reason for, and a way of, studying music scientifically and objectively

a way of doing that study computationally, and getting more out of it in consequence

ways of evaluating what we do rigorously

We have now looked at

R: IDyOM perceptual model

E: Human evaluators; feedback in form of revised perceptual model

We have used a temporary make-do

T: Metropolis sampling

Acknowledgements

Queen Mary
University of London

This work is funded by UK Engineering and Physical Sciences Research Council grants

GR/S82220:
"Techniques and Algorithms for Understanding the Information Dynamics of Music"

EP/H01294X:

"Information and neural dynamics in the perception of musical structure"

EP/D038855:

"Modelling Musical Memory and the Perception of Melodic Similarity"

References

Queen Mary University of London

Lerdahl, F. & Jackendoff, R. (1983) **A Generative Theory of Tonal Music.** Cambridge, MA: MIT Press

Narmour, E. (1990) **The Analysis and Cognition of Basic Melodic Structures: The Implication-realisation Model**. Chicago: University of Chicago Press

Müllensiefen, D., Pearce, M.T., Wiggins, G.A. & Frieler, K. (2007) **Segmenting Pop Melodies: A Model Comparison Approach**. Proceedings of SMPC'07, Montreal, Canada

Pearce, M.T. & Wiggins, G.A. (2006) **Expectancy in melody: The influence of context and learning.** Music Perception, 23(5), 377–405

Potter, K., Wiggins, G.A. & Pearce, M.T. (2007) **Towards Greater Objectivity in Music Theory: Information-Dynamic Analysis of Minimalist Music.** Musicae Scientiae, 11(2), 295–324

Shannon, C. (1948) **A mathematical theory of communication.** Bell System Technical Journal, 27, 379-423, 623-56

Wiggins, G.A. (2007) **Models of Musical Similarity.** Musicae Scientiae, Discussion Forum 4a, 315–338

Pearce, M.T. & Wiggins, G.A. (2012) **Auditory Expectation: The Information Dynamics of Music Perception and Cognition**. Topics in Cognitive Science, in press.

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