

Future Technologies for Game Sound Design

Nicolas Fournel – Principal Audio Programmer

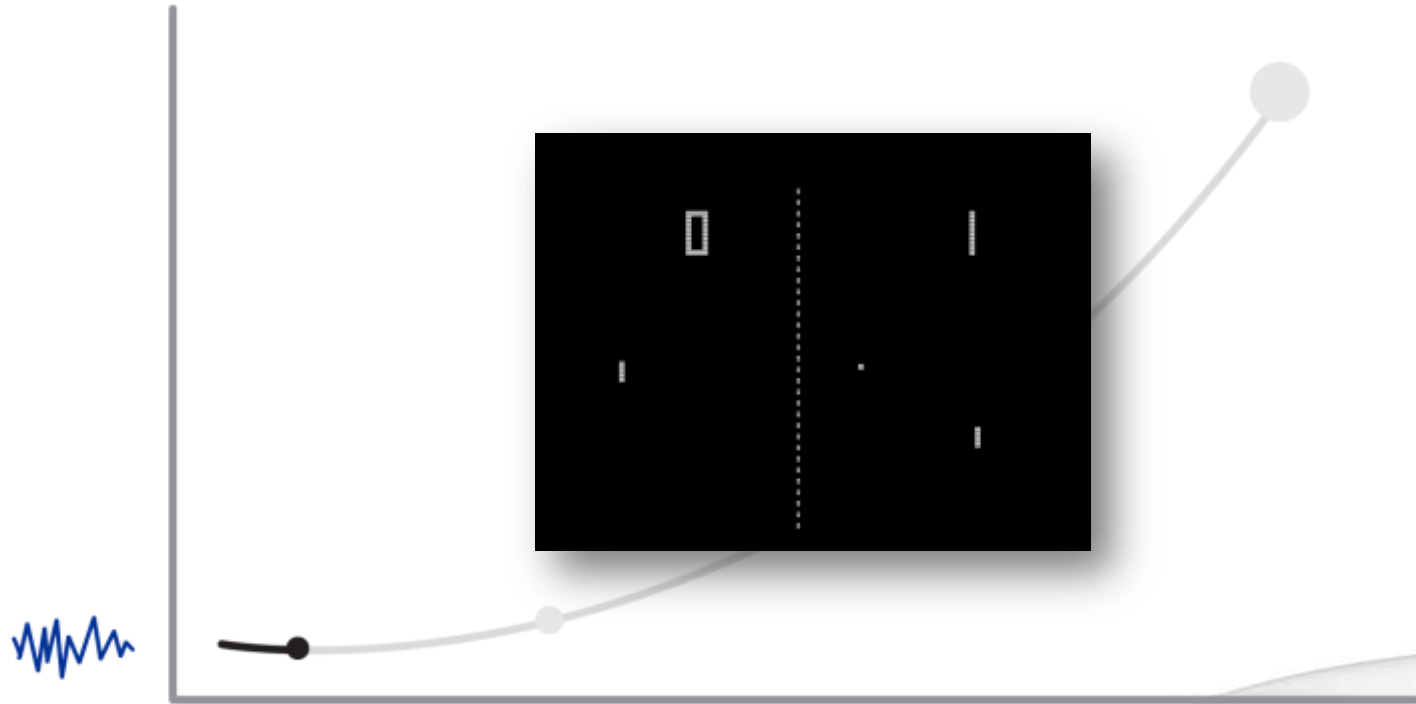
Sony Computer Entertainment Europe



Worldwide Studios
Creative Services Group



Game Audio Evolution



Worldwide Studios
Creative Services Group



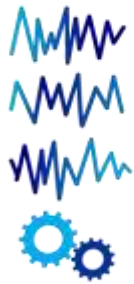
Game Audio Evolution



Worldwide Studios
Creative Services Group



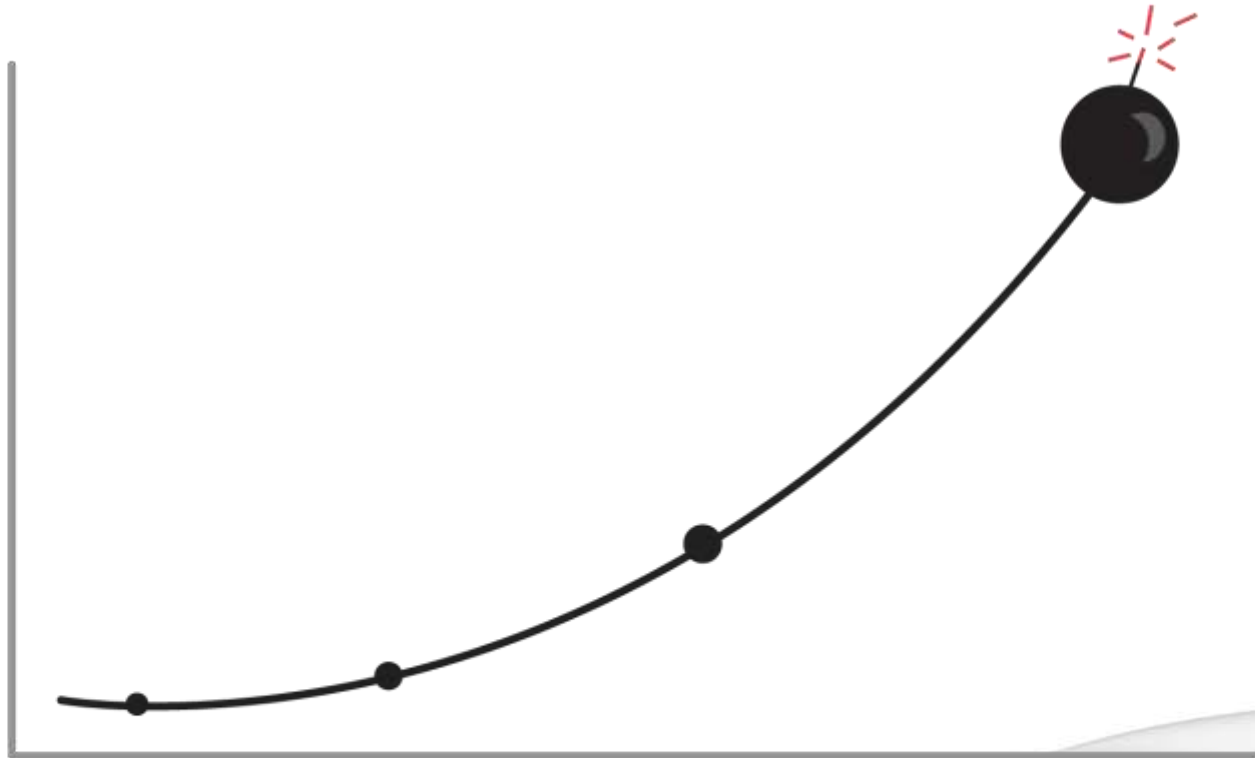
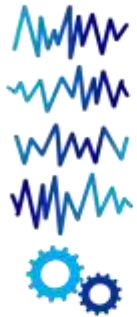
Game Audio Evolution



Worldwide Studios
Creative Services Group



Game Audio Evolution



Worldwide Studios
Creative Services Group



Overview

- Audio Analysis
- Procedural Audio
- Combining them



Worldwide Studios
Creative Services Group



Audio Analysis



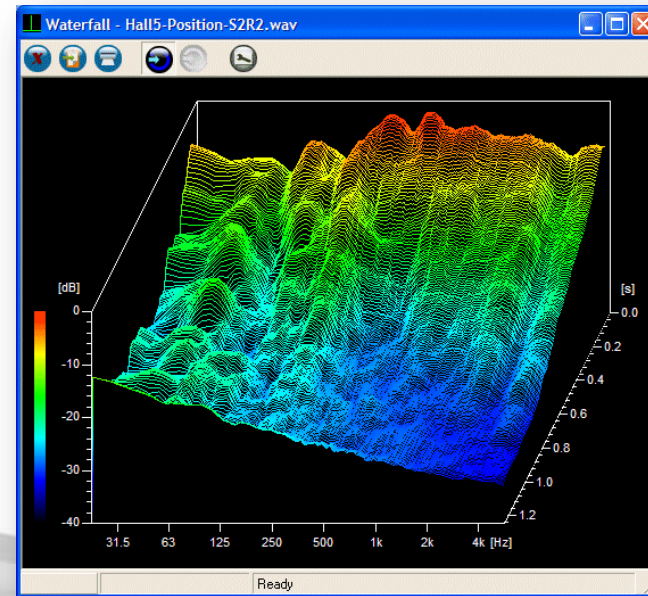
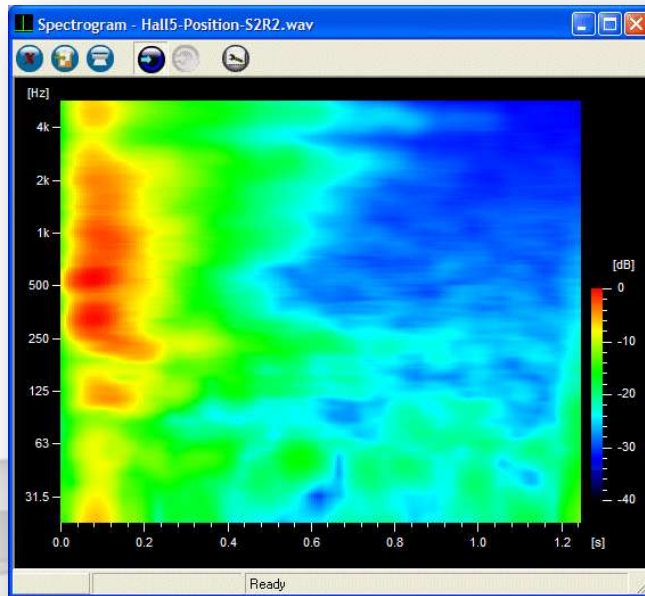
Worldwide Studios
Creative Services Group



Audio analysis is more than...

RMS (Root Mean Square)

FFT (Fast Fourier Transform)



Worldwide Studios
Creative Services Group



More fun with spectra

Other features based on the spectrum

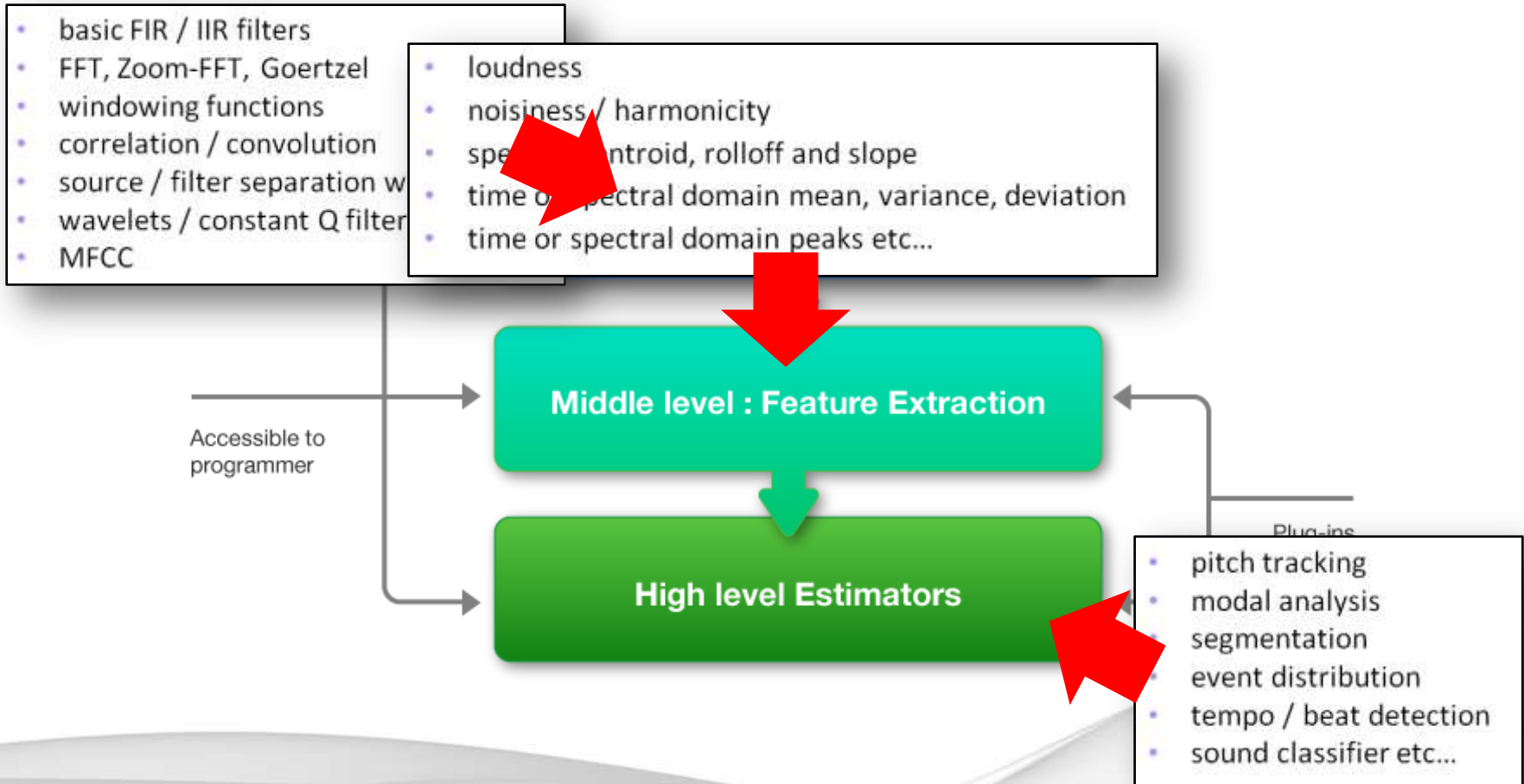
- Spectral Flux, Rolloff, Centroid, Flatness, Kurtosis etc...
- Noisiness / inharmonicity / even and odd harmonics

Other ways to detect spectral content

- Goertzel algorithm
- Constant Q filter banks
- Wavelets



SCEE's AFEX



AFEX Tool

The screenshot shows the AFEX Tool interface with the following components:

- Files to analyze:** A list of audio files including AHa1.wav through AHa6.wav, and various vibrato samples like DarknessVibrato1.wav, DarknessVibrato2.wav, KatieVibrato1.wav, KatieVibrato2.wav, and NewJack_Greatest.wav. The file 'D:\TestSamples\singing pitch detection\aha6.wav' is selected.
- Singing Pitch Graph:** A line graph with 'Vibrato (Hz)' on the y-axis (0.000 to 250.000) and 'Time (s)' on the x-axis (0.000 to 4.000). A blue line represents 'Pitch' and a yellow line represents 'Vibrato'. The pitch starts at approximately 100,000 Hz, rises to about 200,000 Hz at 1.5 seconds, and remains stable until 4.5 seconds. The vibrato is near zero.
- Marker Properties:** A table showing properties for the selected marker:

Name	Pitch
Color	RoyalBlue
Type	Line
Marker	None
Marker size	5
- Analyses to perform:** A list containing 'Singing Pitch' with 'Add...', 'Remove', and 'Clear' buttons below it.
- Selected analysis parameters:**

Pitch minimum	80
Pitch maximum	1100
Vibrato minimum rate	3
Vibrato maximum rate	7
Vibrato maximum amplitude	50
- Input data:**

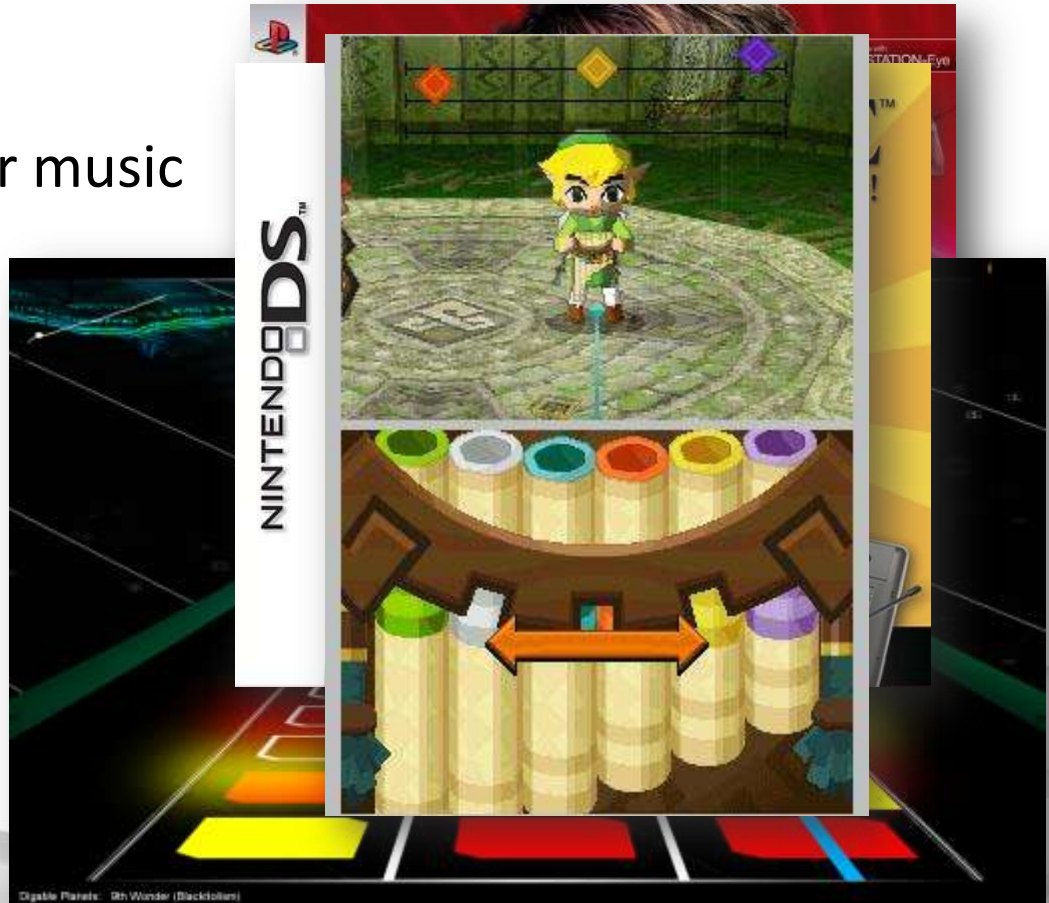
Length	4096
Slide	2048
Window	Rectangular
- Footer:** 'aha6.wav: 16-bit, mono, 44100 Hz, 4.853 seconds'



Game Design Applications

- Singing games
- Levels generated by user music
- Voice recognition
- Sound classification

Even more important with social gaming !

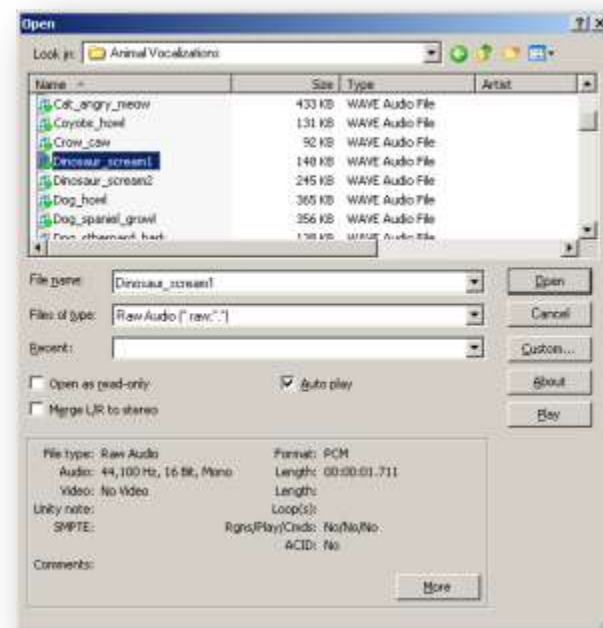
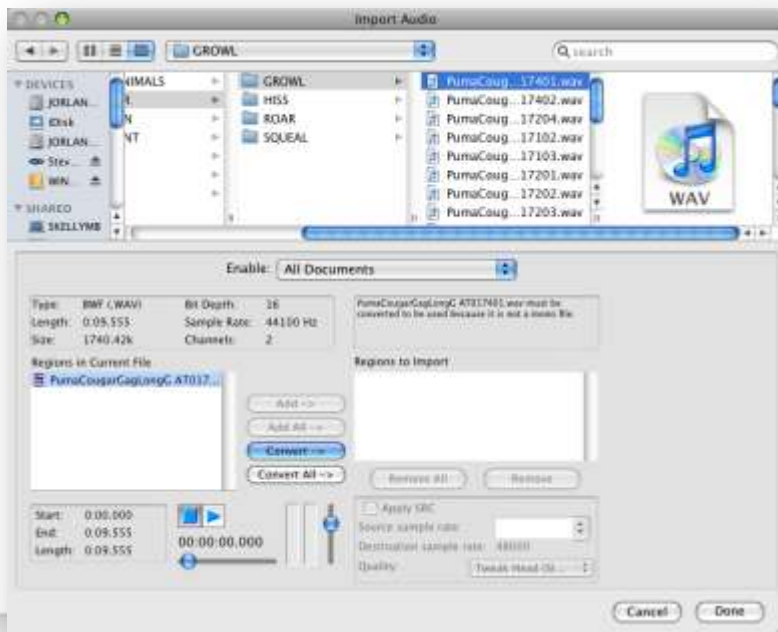


Worldwide Studios
Creative Services Group



Selecting sample files

Typical sample files browsers:



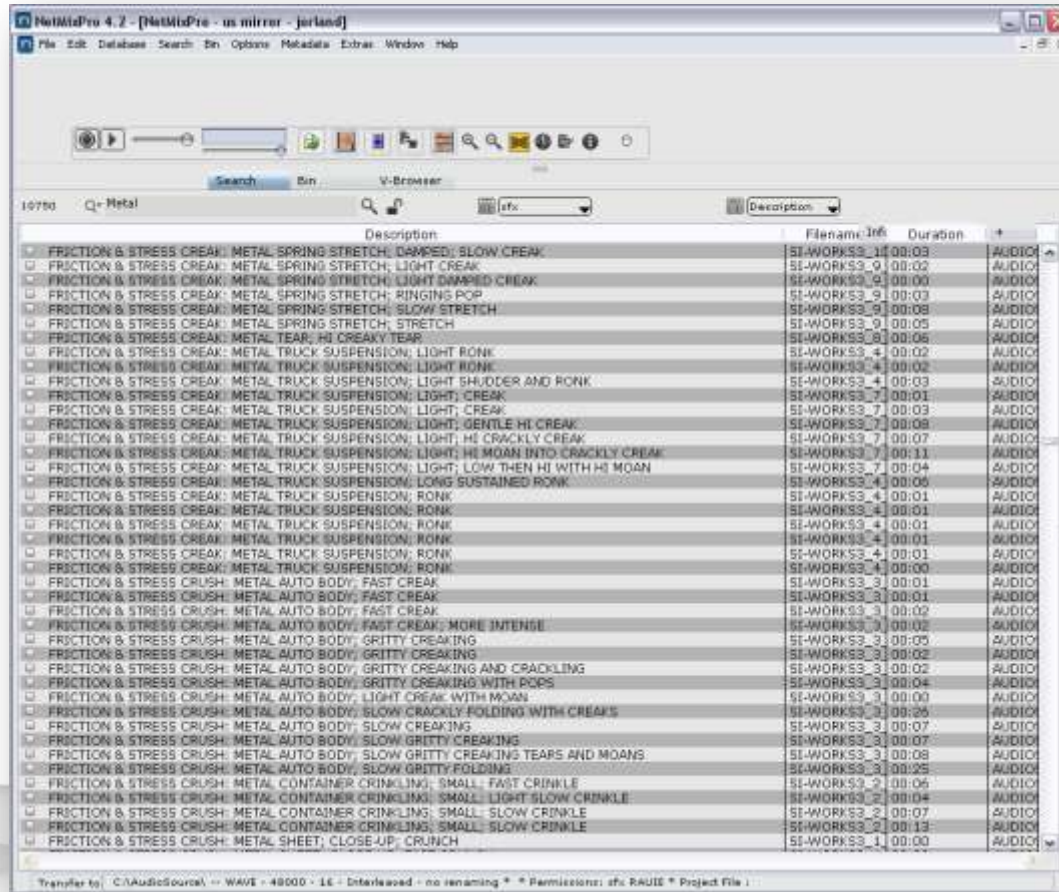
[AFEX Browser](#)



Worldwide Studios
Creative Services Group



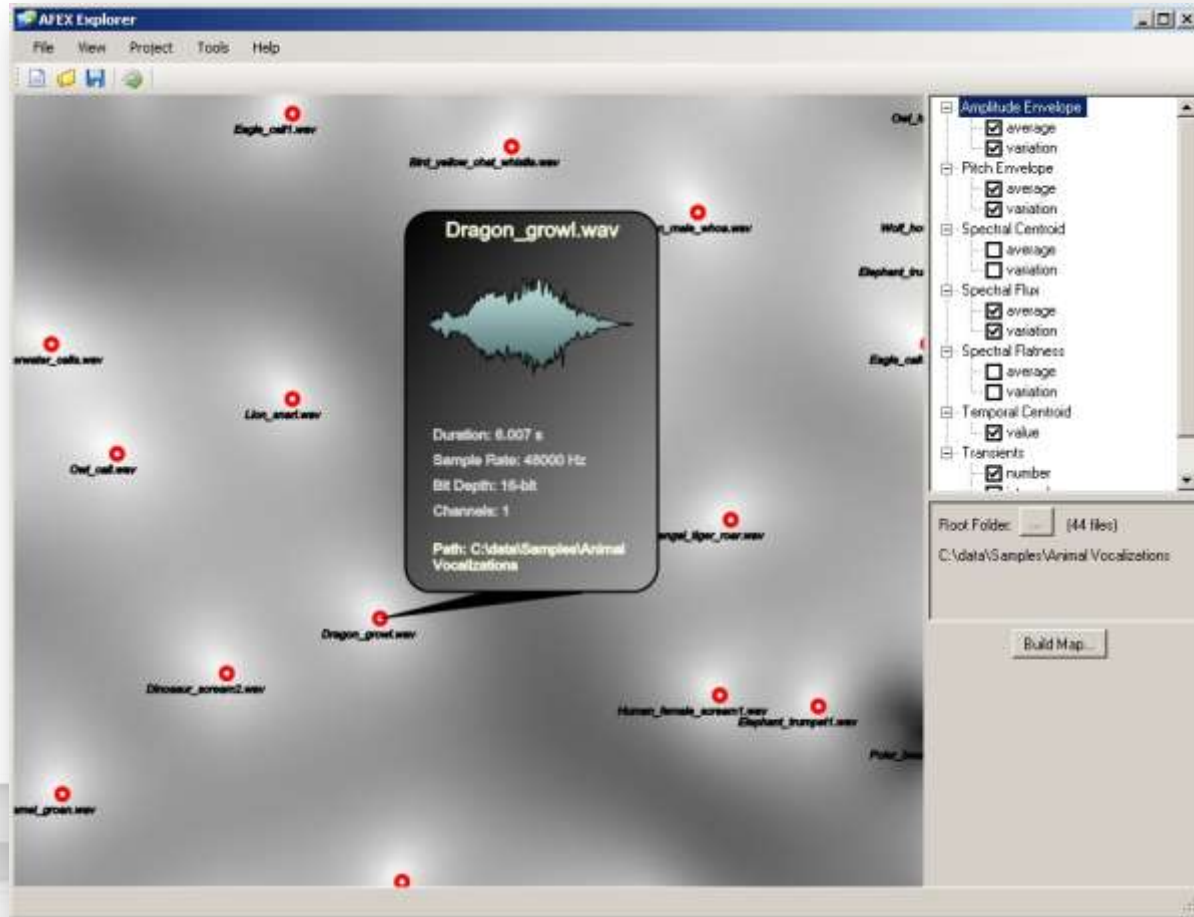
Browsing sound effects libraries



Worldwide Studios
Creative Services Group



Example: AFEX Explorer



Dicto: checking dialogue assets

Dicto - Dialogue Checking Tool - Slider

File Tools Help

language German (de)

Missing Files | Extra Files | Wrong Format | Different Durations | Different Levels

File	Ref. Duration	Loc. Duration	Variation %	Padding In	Padding Out
D:\Slider\ALL GER\KR_yoromeki_1.wav	3.014	0.609	-79.790	0.000	0.004
D:\Slider\ALL GER\TB_ATK_s_8.wav	0.489	0.676	79.032	0.033	0.070
D:\Slider\ALL GER\TB_BREAK_2.wav	0.703	1.227	74.382	0.031	0.047
D:\Slider\ALL GER\TB_chakuchi_0.wav	0.612	1.640	167.770	0.114	0.024
D:\Slider\ALL GER\TB			104.011	0.117	0.202
D:\Slider\ALL GER\TB			121.958	0.101	0.196
D:\Slider\ALL GER\TB			93.693	0.006	0.053
D:\Slider\ALL GER\TB			72.362	0.056	0.055
D:\Slider\ALL GER\TB			470.613	0.069	0.067
D:\Slider\ALL GER\TB			74.194	0.016	0.022
D:\Slider\ALL GER\TB			-51.502	0.016	0.017
D:\Slider\ALL GER\TB			260.415	0.019	0.041
D:\Slider\ALL GER\TB			64.483	0.086	0.253
D:\Slider\ALL GER\TB			99.680	0.183	0.162
D:\Slider\ALL GER\TB			114.385	0.148	0.262
D:\Slider\ALL GER\TB			74.021	0.259	0.065
D:\Slider\ALL GER\TB			73.182	0.069	0.113
D:\Slider\ALL GER\TB			70.318	0.098	0.066
D:\Slider\ALL GER\TB_Re_0.wav	1.836	3.077	67.576	0.150	0.574
D:\Slider\ALL GER\TB_Re_1.wav	1.573	3.029	92.573	0.231	0.287
D:\Slider\ALL GER\TB_SIDE_STP_0.wav	0.260	0.441	69.506	0.015	0.037

Comparison Settings

Reference language: French (fr)

Compare with...

- German (de)
- Italian (it)
- Japanese (jp)
- Polish (pl)
- Portuguese (pt)
- Russian (ru)
- Spanish (es)

Report...

- Missing Files
- Extra Files
- Wrong Format
- Different Durations
- Different Levels

Audio Analysis Settings...

padding duration **1**

padding threshold **-36**

peak level variation **12**

RMS variation **12**

size variation **50**

padding duration
Maximum amount of silence allowed at the beginning or end of a file, expressed in seco...

Start...



Some of the games Dicto helped



Worldwide Studios
Creative Services Group



Audio analysis at run-time

- Most audio engines are...



deaf !

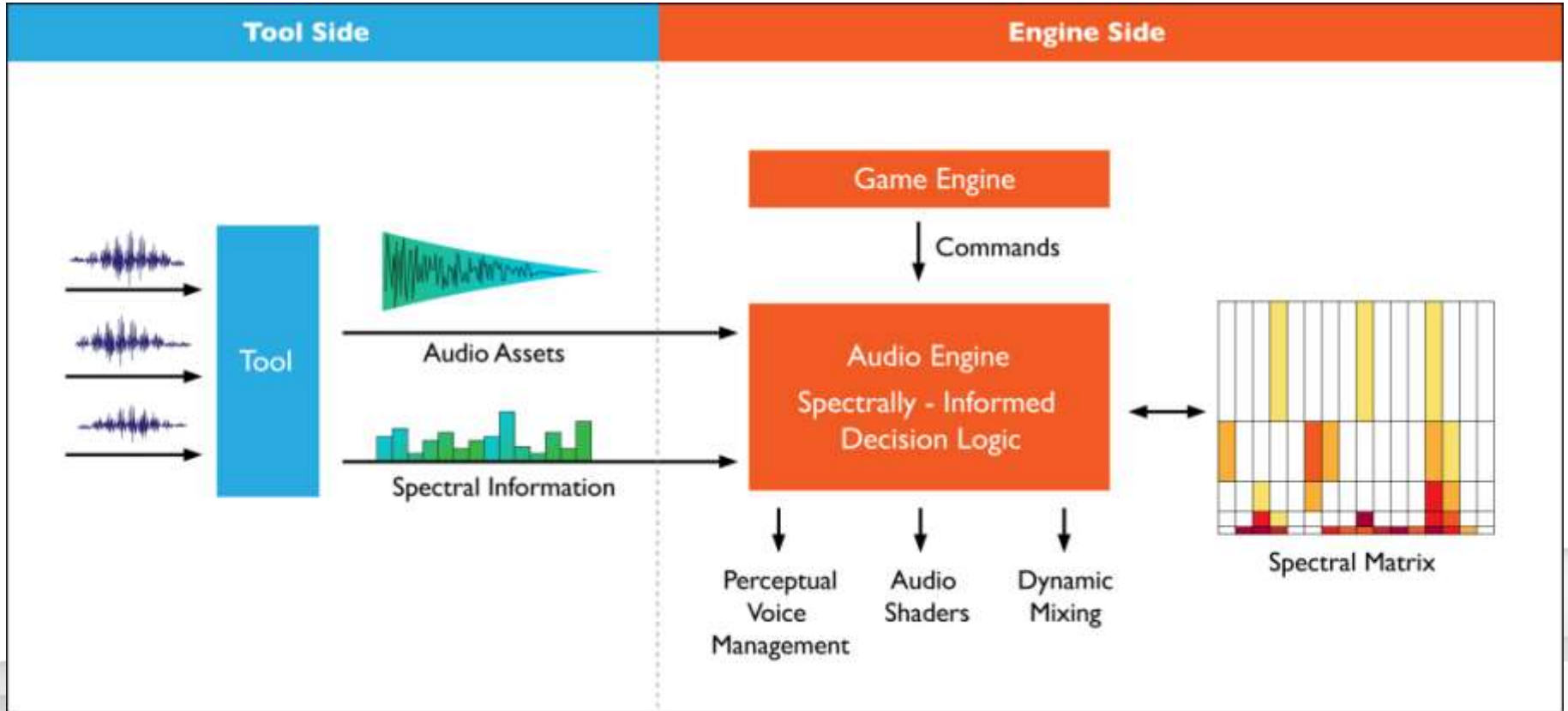
They make decisions that impact the audio output of a game without knowing what they are playing.



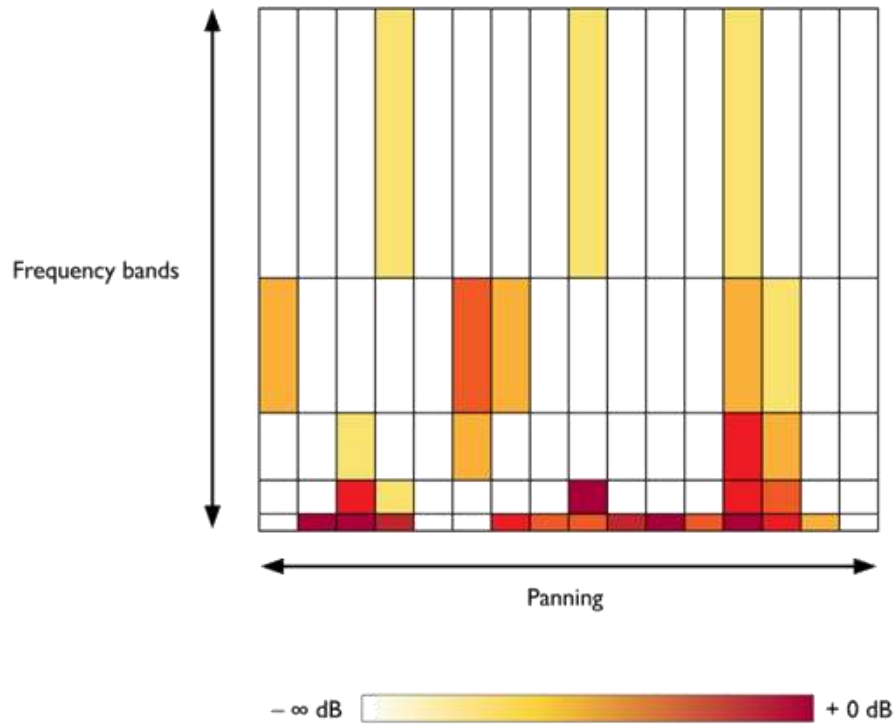
Worldwide Studios
Creative Services Group



Audio Engines



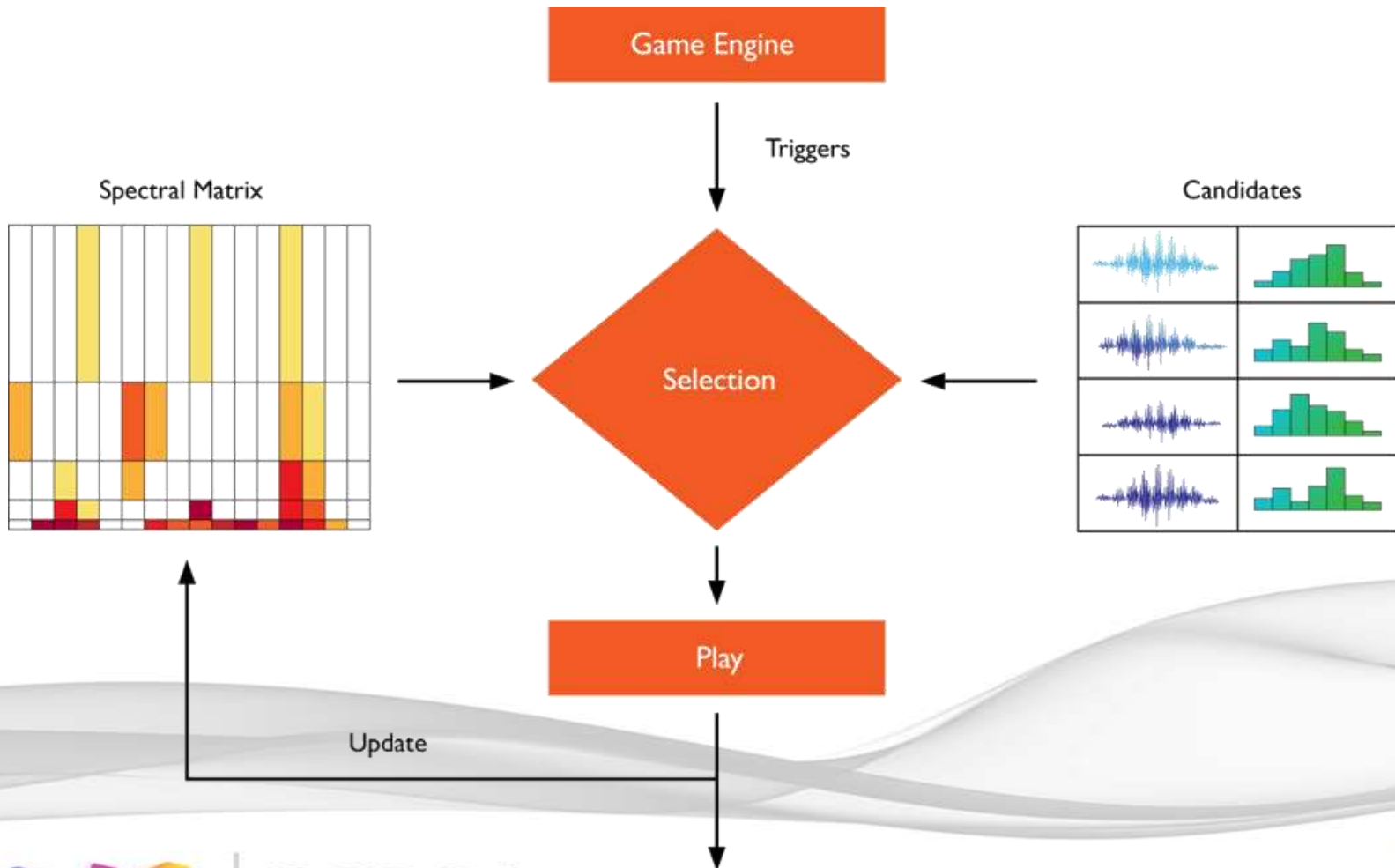
Spectral Matrix



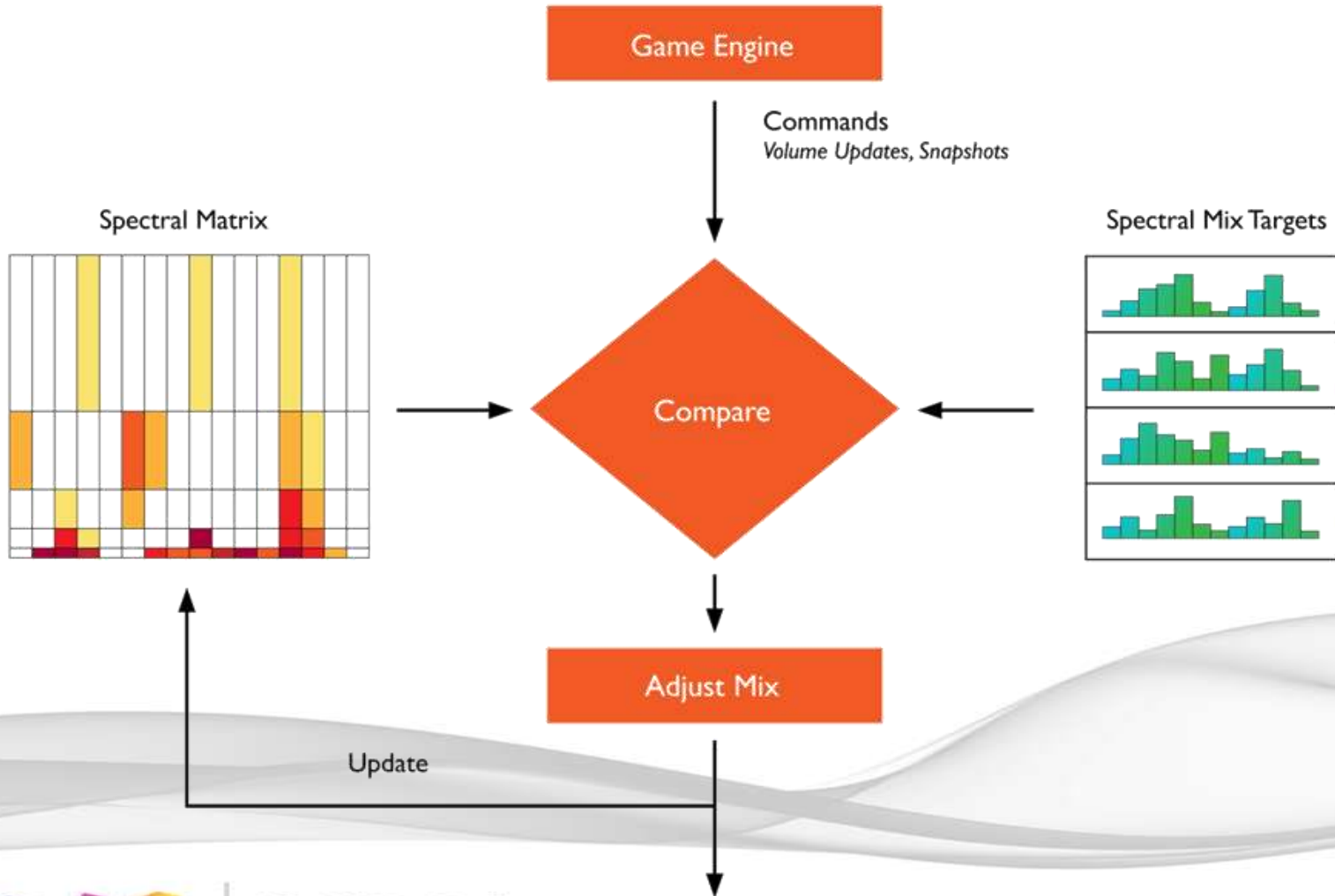
Worldwide Studios
Creative Services Group



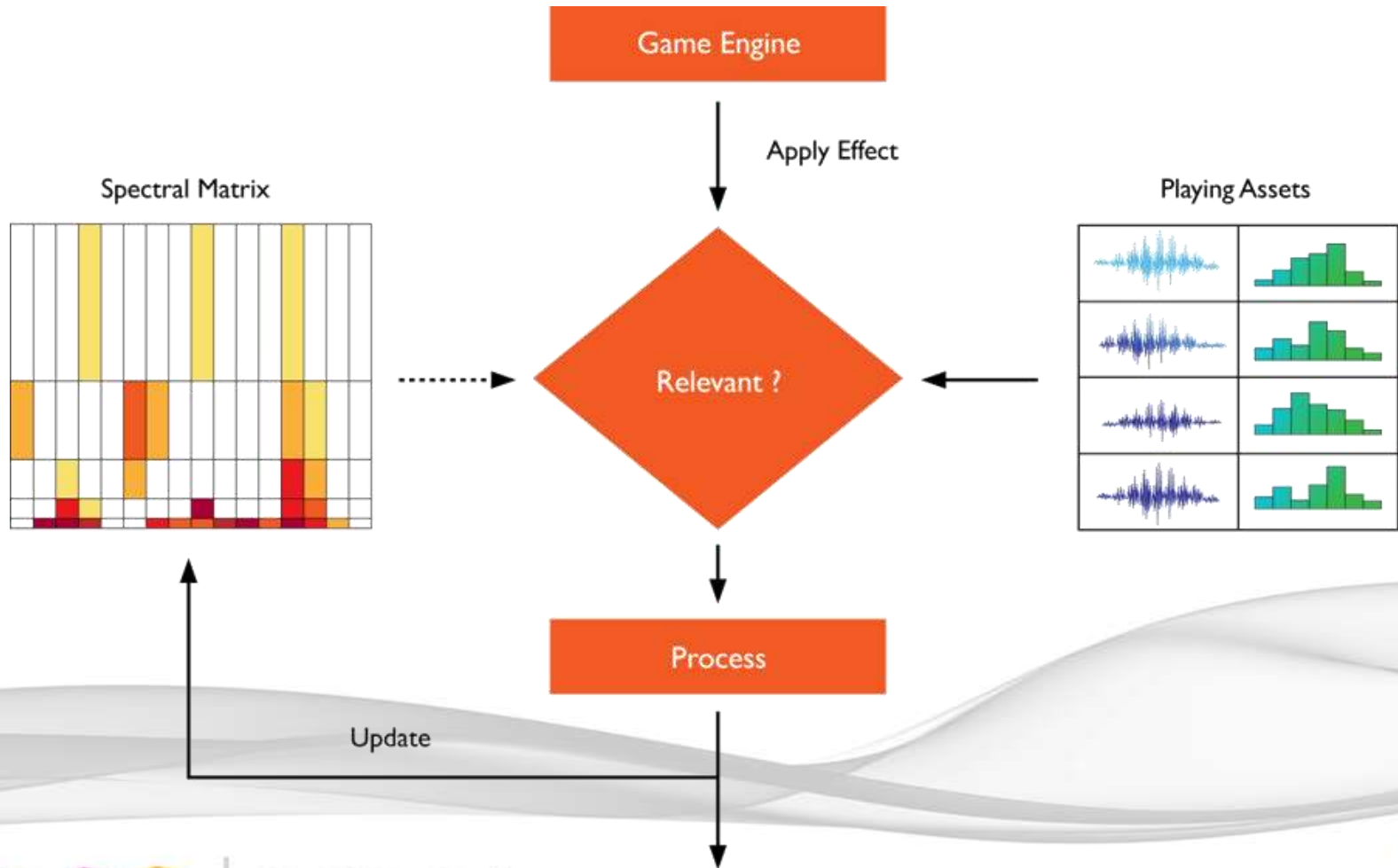
Perceptual Voice Management



Dynamic Mixing



Audio Shaders



Procedural Audio



Worldwide Studios
Creative Services Group



What is Procedural Audio ?

For sound effects:

- Real-time sound synthesis
- With control parameters coming from other sub-systems
- Examples of existing systems:
 - WWISE SoundSeed (Impact and Wind / Whoosh)
 - AudioGaming
 - SCEE's Spark

Applies also to music (algorithmic composition) and dialogue
(Phonetic Arts)



When to use PA ?

Good candidates:

- Repetitive (e.g. footstep, impacts)
- Large memory footprint (e.g. wind, ocean waves)
- Require a lot of control (e.g. car engine, creature vocalizations)
- Highly dependent on the game physics (e.g. rolling ball, sounds driven by motion controller)
- Just too many of them to be designed (vast universe, user-defined content...)



Two approaches to Procedural Audio

Bottom-Up:

- examine how the sounds are physically produced
- write a system recreating them

Top-Down

- analyse examples of the sound we want to create
- find the adequate synthesis system to emulate them



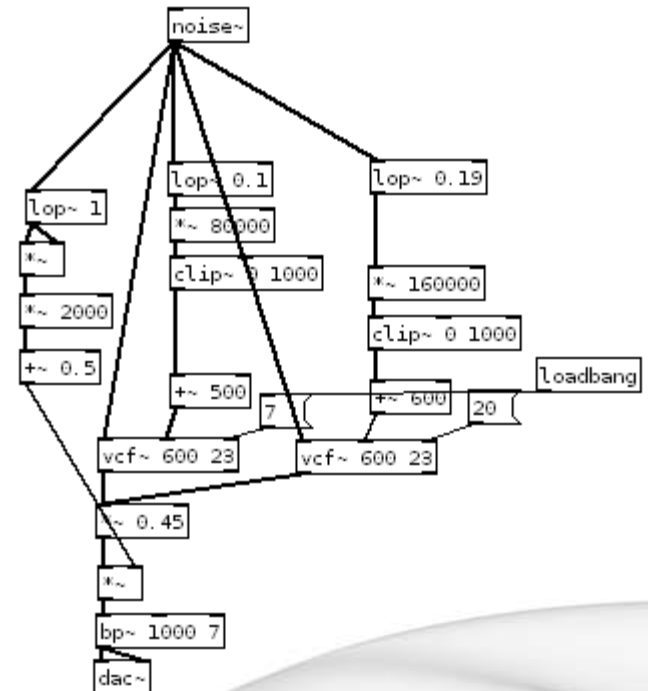
Procedural Model Example : Wind

Good example of bottom-up versus top-down design

- Computational fluid dynamics to generate aerodynamic sound (Dobashi / Yamamoto / Nishita)

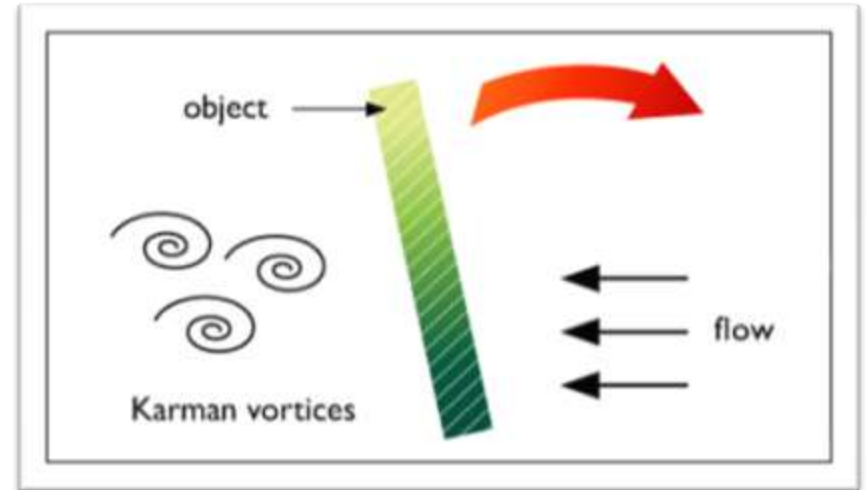


- Noise generator and bandpass filters ([Subtractive synthesis](#))



Procedural Model Example : Whoosh

- Karman vortices are periodically generated behind the object (primary frequency of the aerodynamic sound)
- Using classic subtractive synthesis is cheaper
- Ideal candidate for motion controllers



Procedural Model Example :Whoosh

Heavenly Sword:

- about 30 Mb of whooshes on disk
- about 3 Mb in memory at all times

Recorded whooshes



Subtractive synthesis (SoundSeed)



Aerodynamics computations



Worldwide Studios
Creative Services Group



Procedural Model Example

Water / Bubbles

Physics of a bubble is well-known

- Impulse response = damped sinusoid
- resonance frequency based on radius
- Energy loss based on simple thermodynamic laws
- Statistical distributions used to generate streams / rain
- Impacts on various surfaces can be simulated

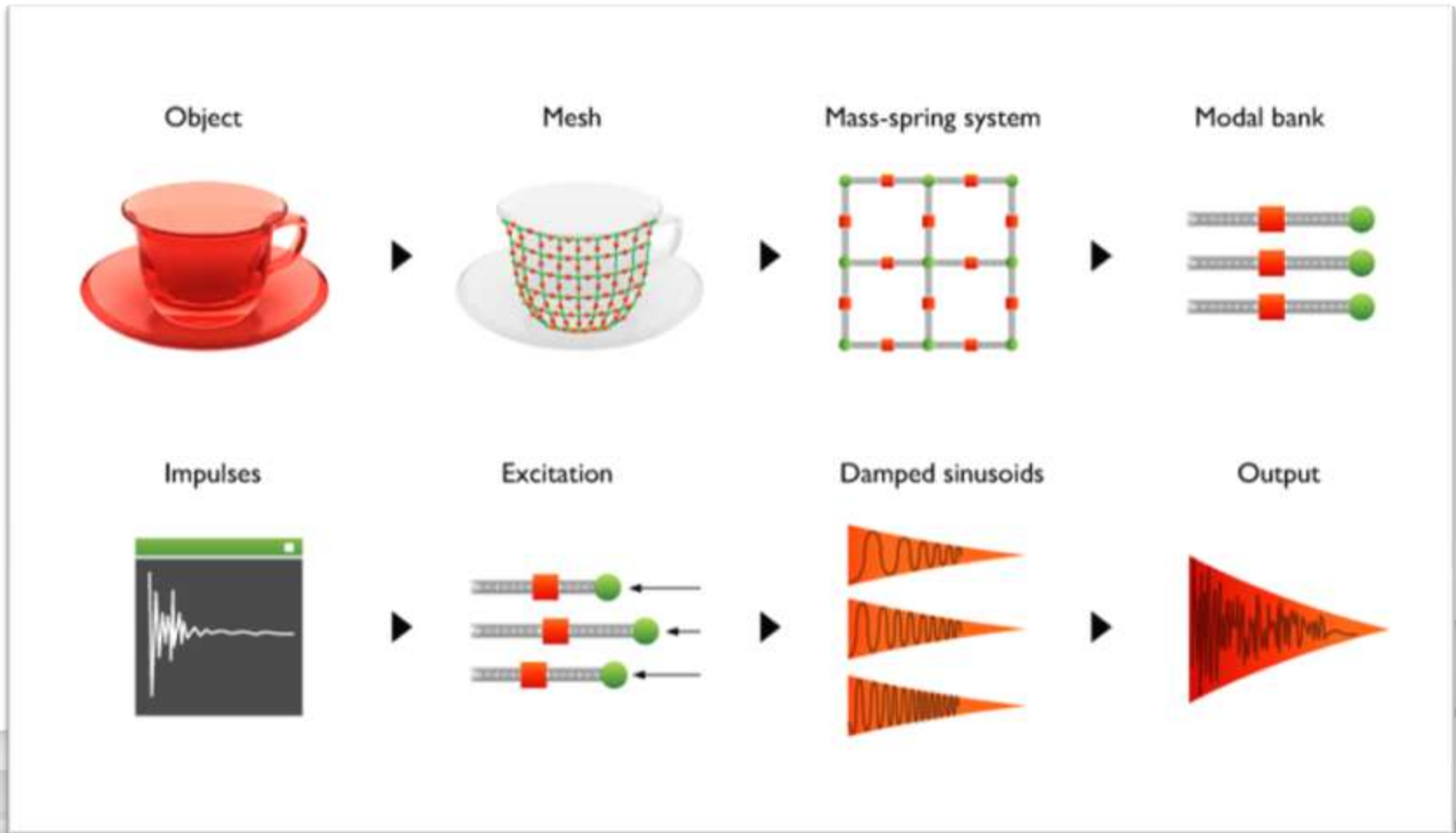
[Bubbles generated with procedural audio](#)



Worldwide Studios
Creative Services Group



Procedural Model Example : Solids



Procedural Model Example : Solids

Other solutions for the analysis part:

- LPC analysis
Source – Filter separation
- Spectral Analysis
Track modes, calculate their frequency, amplitude and damping

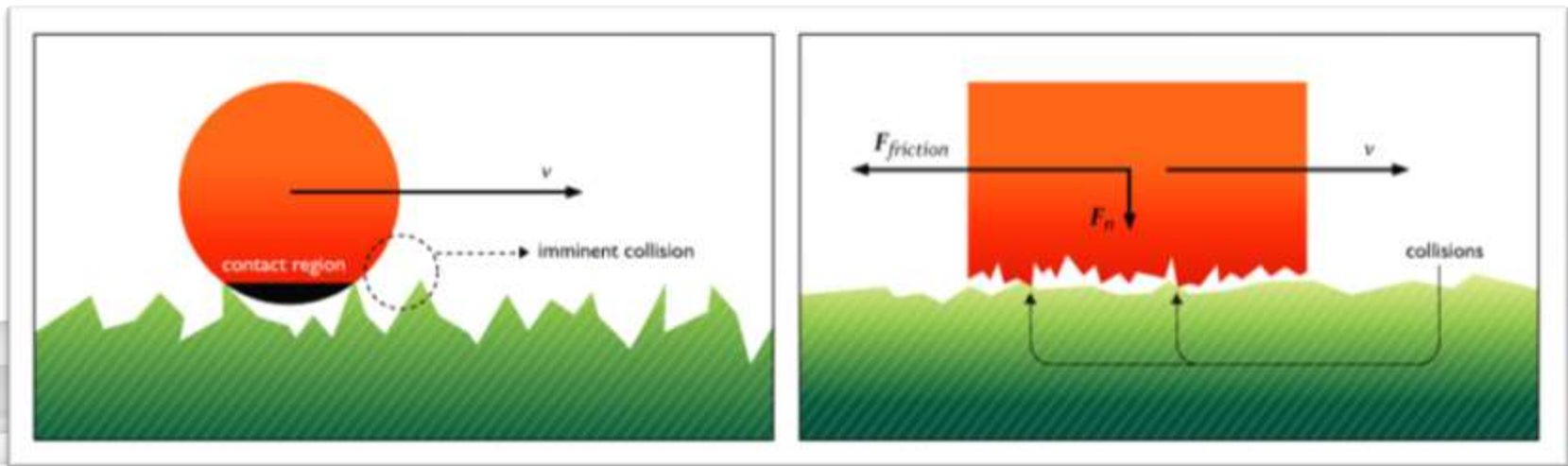


Procedural Model Example : Solids

Different excitation signals for:

- Impacts (hitting)
- Friction (scraping / rolling / sliding)

Interface with game physics engine / collision manager



Worldwide Studios
Creative Services Group



Procedural Model Example : Solids

“Physics” bank for Little Big Planet on PSP:

- 85 waveforms
- 60 relatively “complex” Scream scripts
- Extra layer of control with more patches (using with SCEA’s Xfade tool)

[Impacts generated by procedural audio](#)

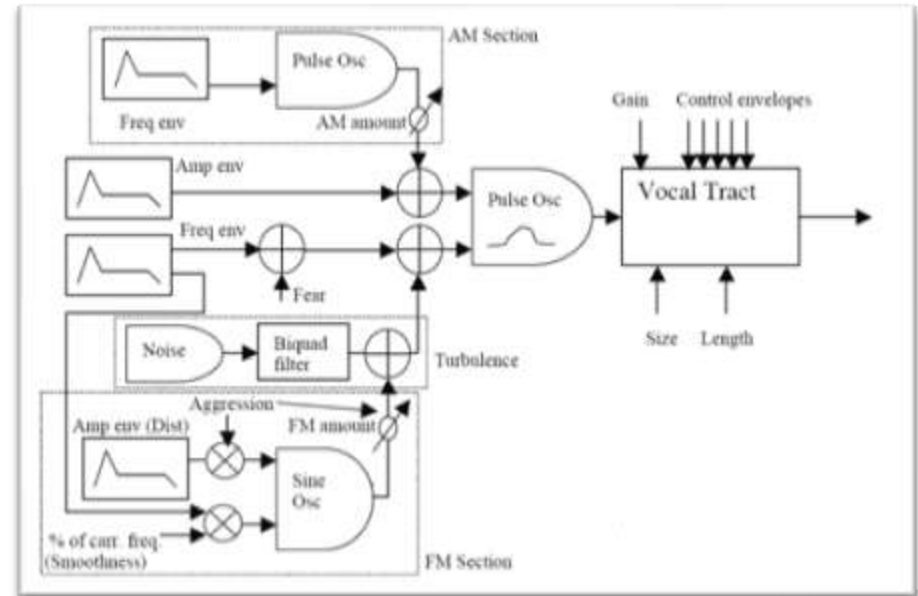


Worldwide Studios
Creative Services Group



Procedural Model Example : Creature

- Physical modelling of the vocal tract (Kelly-Lochbaum model using waveguides)
- Glottal oscillator



Procedural Model Example : Creature

Eye Pet vocalizations:

- Over a thousand recordings of animals
- 634 waveforms used
- In 95 sound scripts

Eye Pet waveforms



Synthasaurus



Worldwide Studios
Creative Services Group



Pros

- Offers a lot of advantages compared to static sounds (non repetitive, dynamic, lots of control parameters)
- One model = many sounds !
- Some models can be implemented very easily
 - Impacts / contacts
 - Footsteps
 - Air / Water
- Procedural audio is not necessarily more CPU expensive



Cons

- Not a solution for everything
- It is still harder to implement and to debug
- Mostly due to lack of:
 - trained sound designers / programmers / testers
 - adapted tools / run-time
 - ready-to-use models



Combining Procedural Audio & Audio Analysis



Worldwide Studios
Creative Services Group



One of the main tasks of
sound designers:

Transforming static sounds
into dynamic ones



Worldwide Studios
Creative Services Group



Implementation with Scripting

Current scripting solutions:

- randomization of assets
- volume / pan / pitch variations
- streaming for big assets

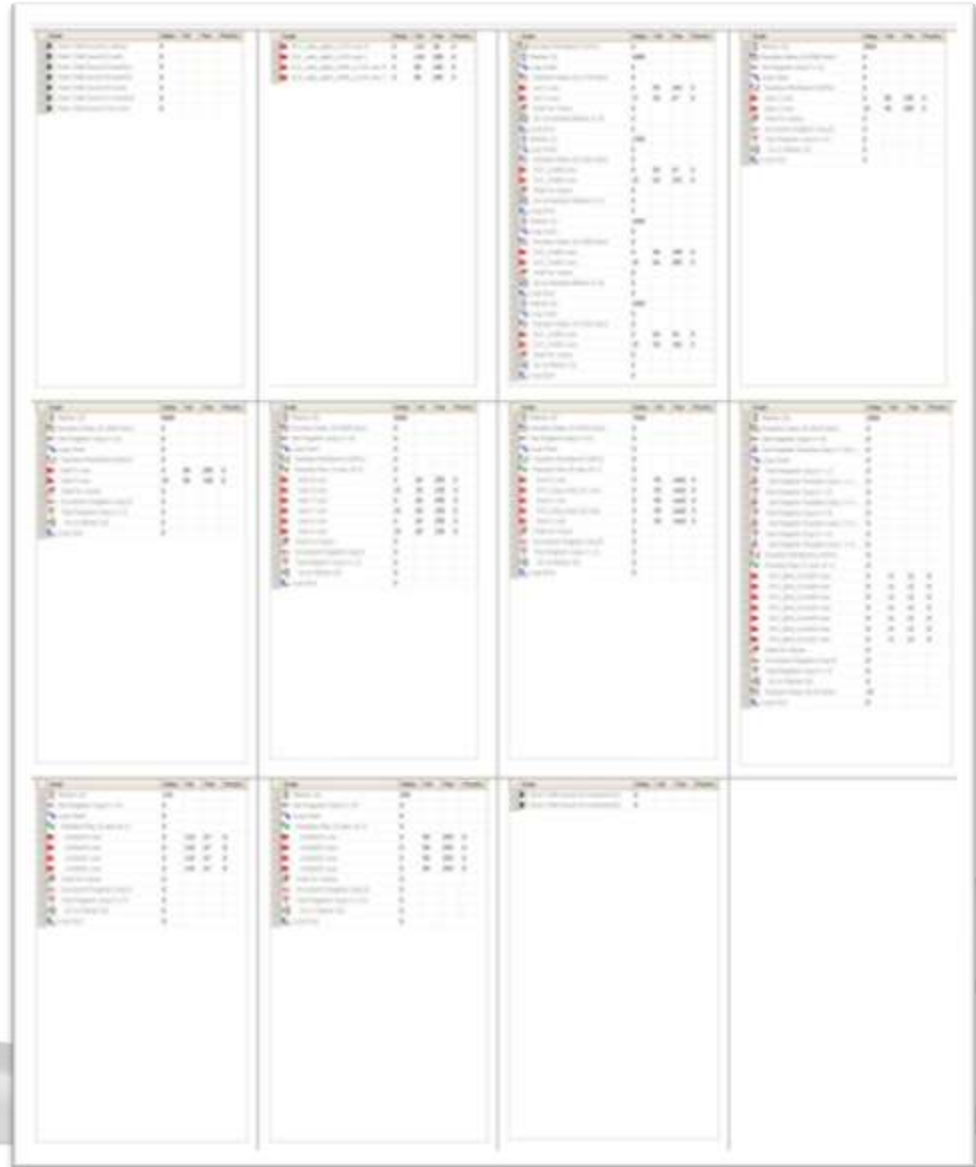
Remaining issues:

- no timbral modifications
- still uses a lot of resources (memory or disk)
- not really dynamic



A “simple” patch in Sony Scream Tool:

- 11 concurrent scripts
- each “grain” has its own set of parameters

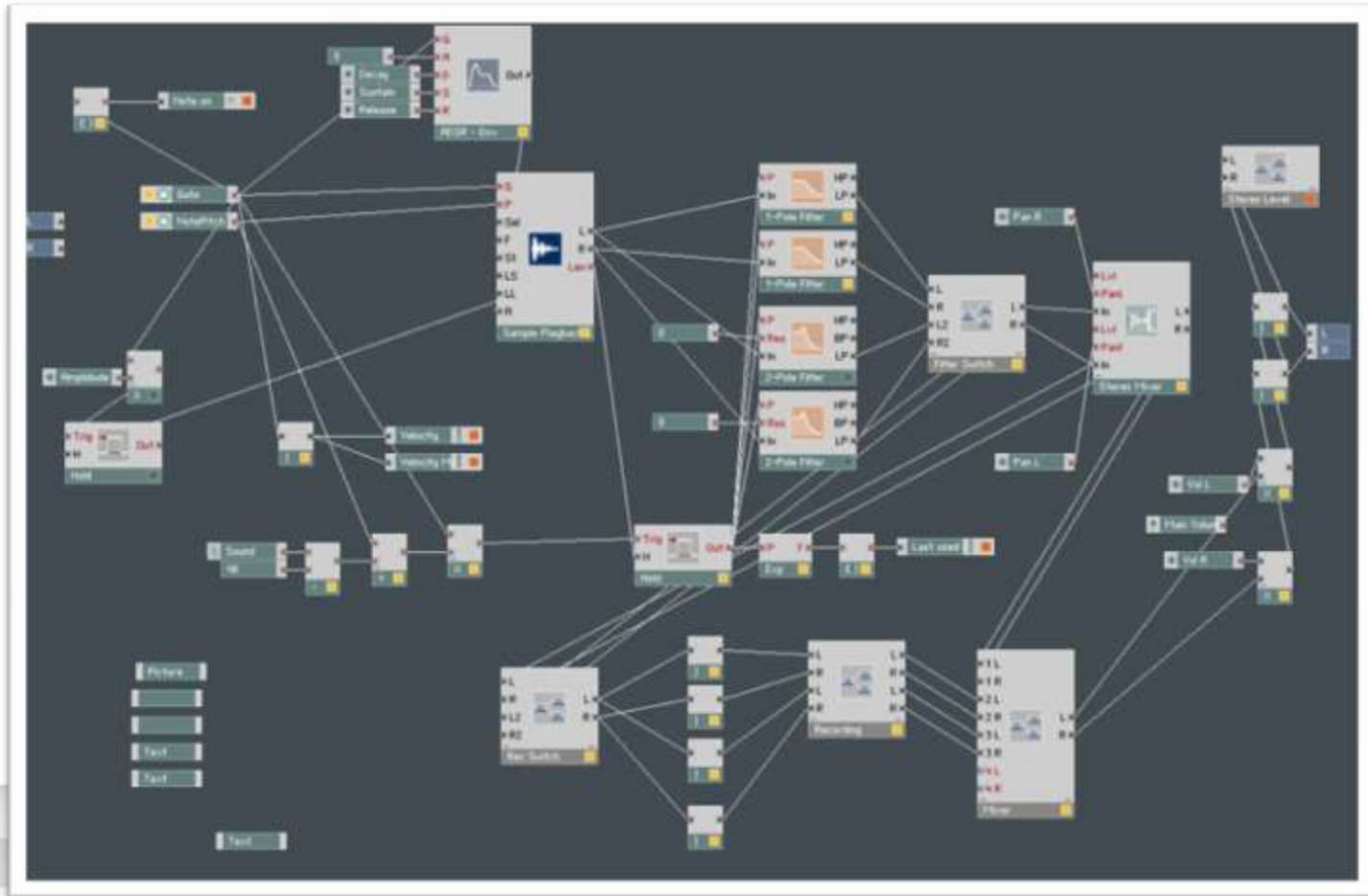


Implementation with Patching

- Tools such as Pure Data / MAX MSP / Reaktor
- Better visualisation of flow and parallel processes and where the control parameters arrive in the model
- Sometimes hard to understand due to the granularity of operators
- Requires a PhD in mechanics, animal anatomy, physics etc..



A “simple” patch in Reaktor...



Worldwide Studios
Creative Services Group



Another solution

Vendors of ready-to-use Procedural Audio models:

- easy to use but...
- limited to available models
- limited to what parameters they allow
- limited to the idea the vendor has of the sound

Examples:

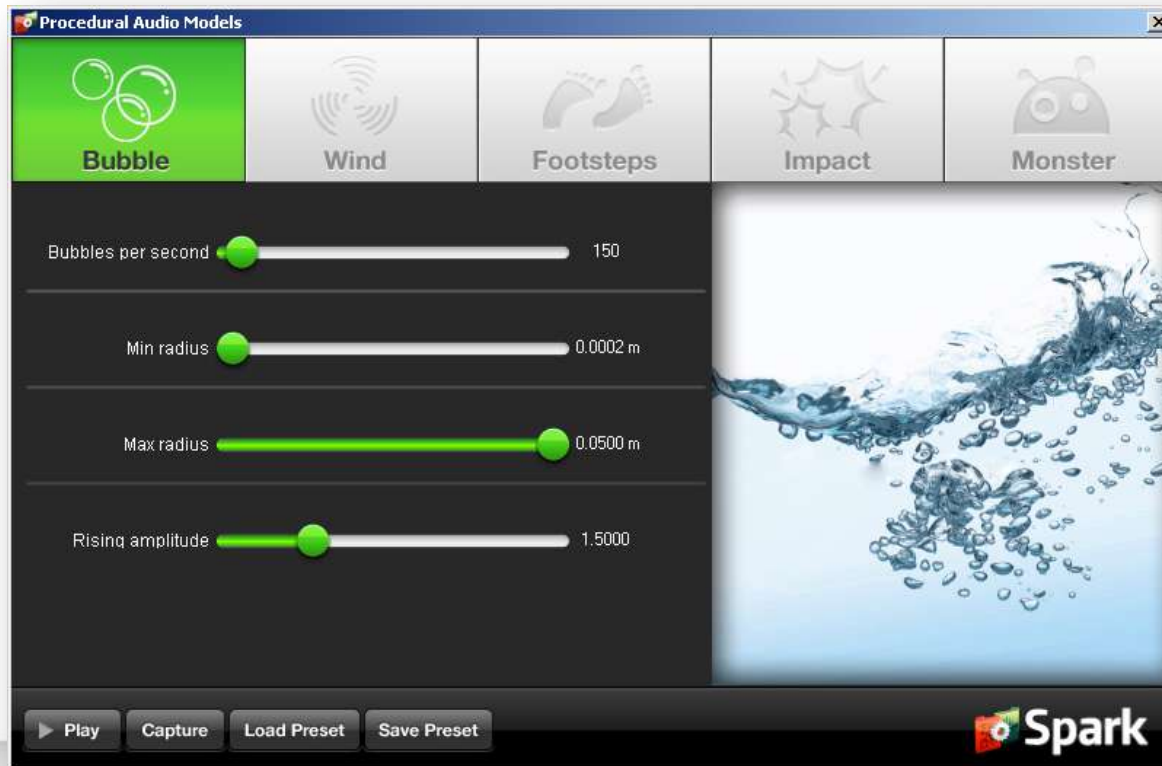
- Staccato Systems already in 2000...
- WWISE SoundSeed series
- SCEE's Spark



Worldwide Studios
Creative Services Group



Spark



Worldwide Studios
Creative Services Group



Going further...

Need for higher-level tools that let the designer:

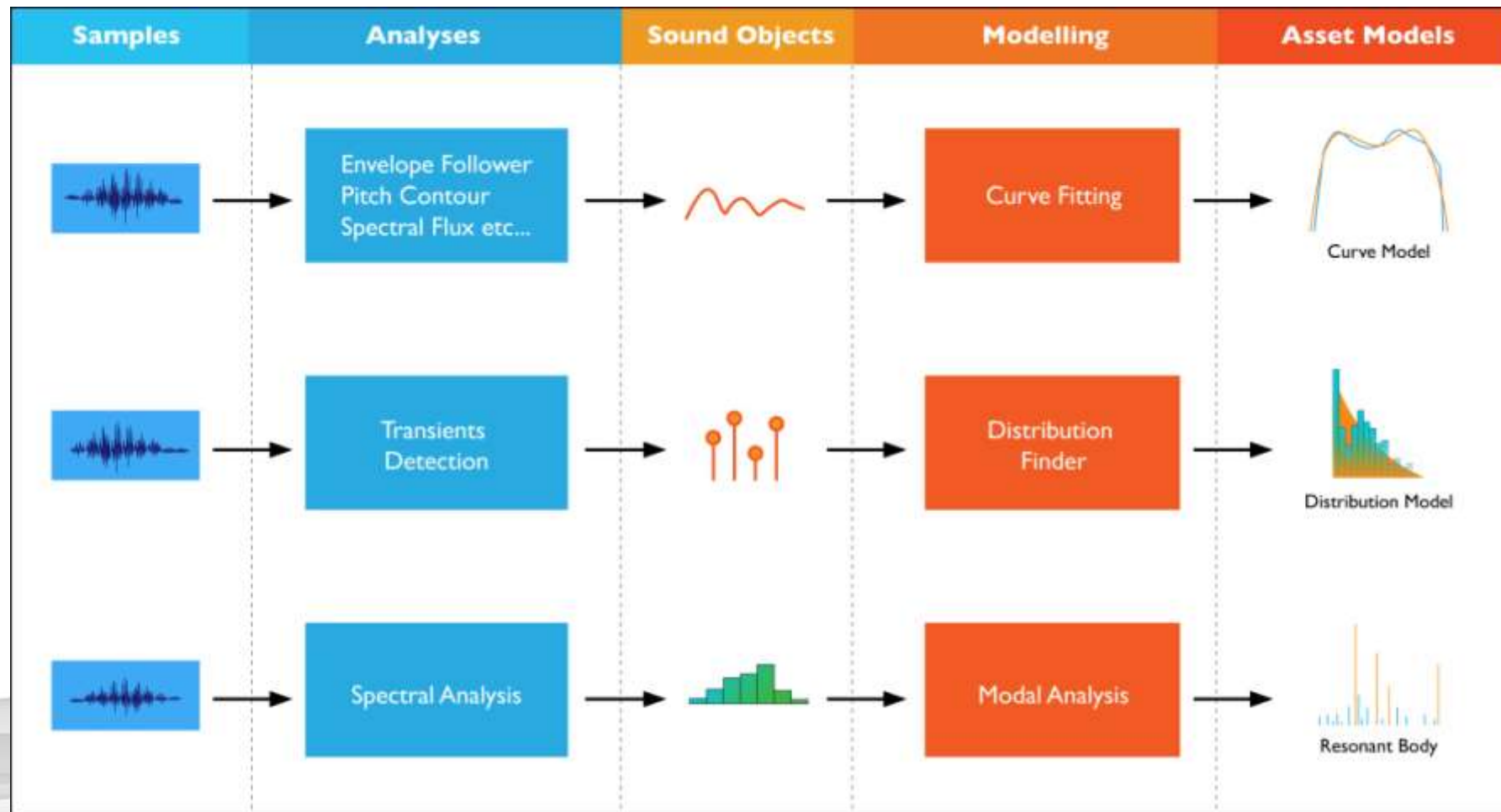
- create its own dynamic model of a sound
- specify its own control parameters
- without having an extensive knowledge of synthesis / sound production mechanisms
- without having to rely on third party models



Worldwide Studios
Creative Services Group



Think asset models, not assets

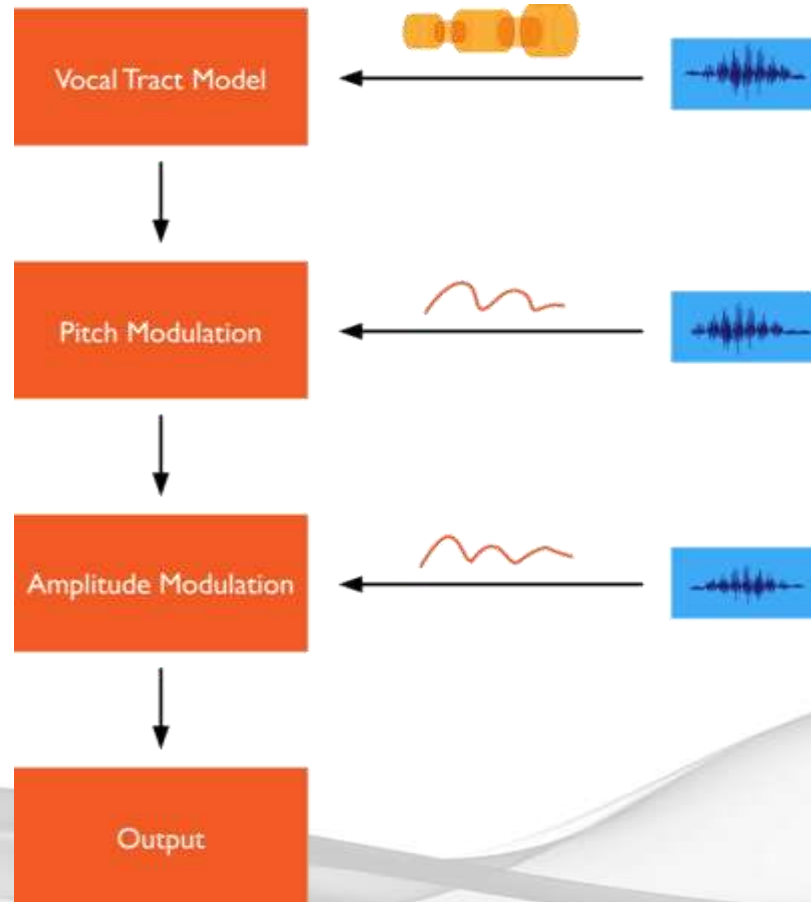


Because we are using analysis...

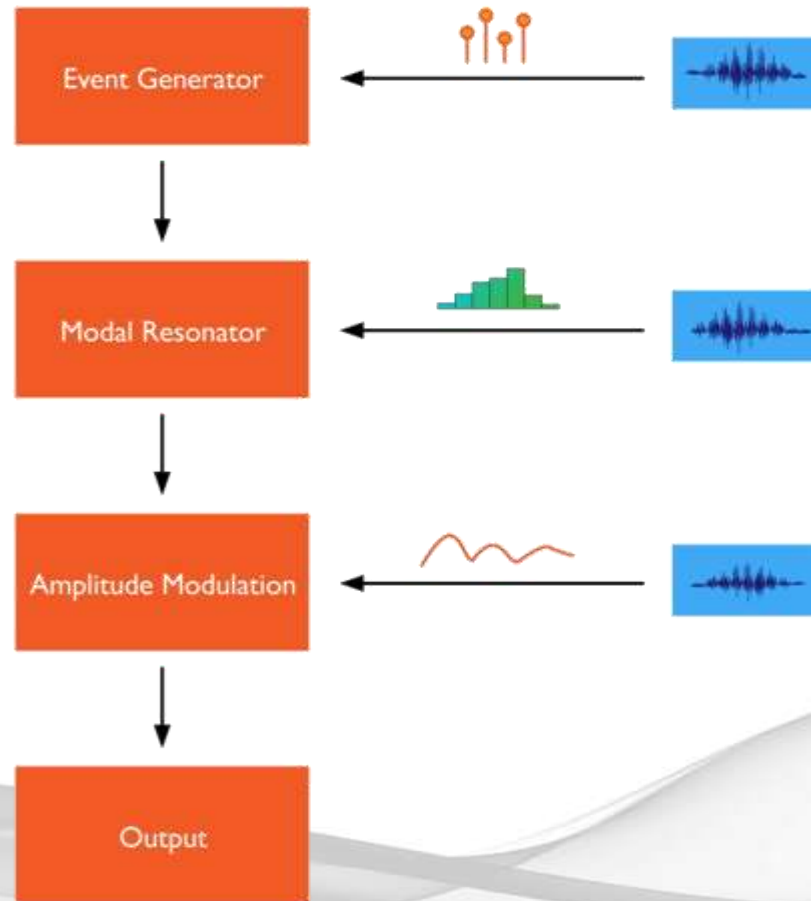
- We can use our own sounds as basis for a model
- The modules can implement more complex behaviours
- We can have a smaller number of modules



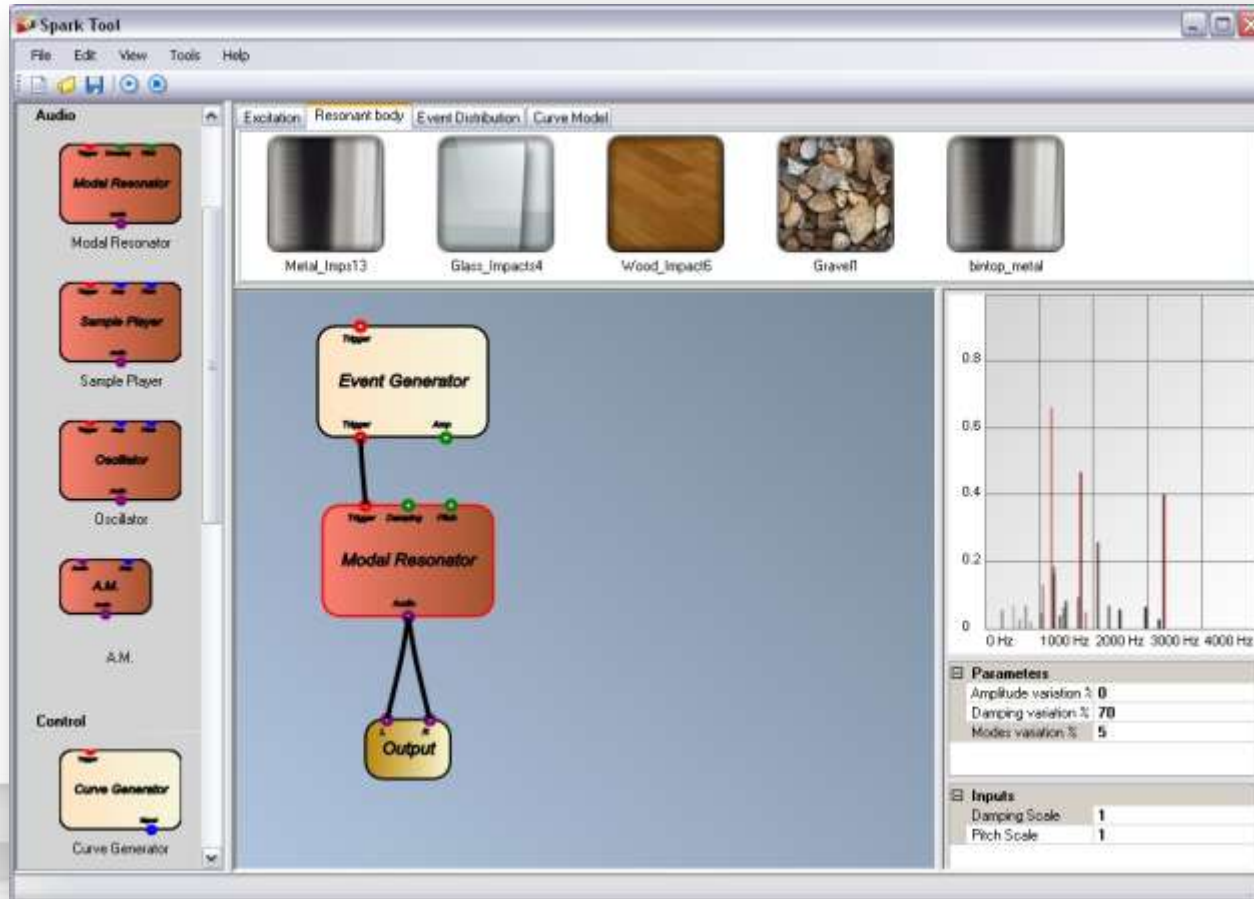
Creature Vocalisations



Debris / impacts



Spark Tool = Spark + AFEX



Worldwide Studios
Creative Services Group



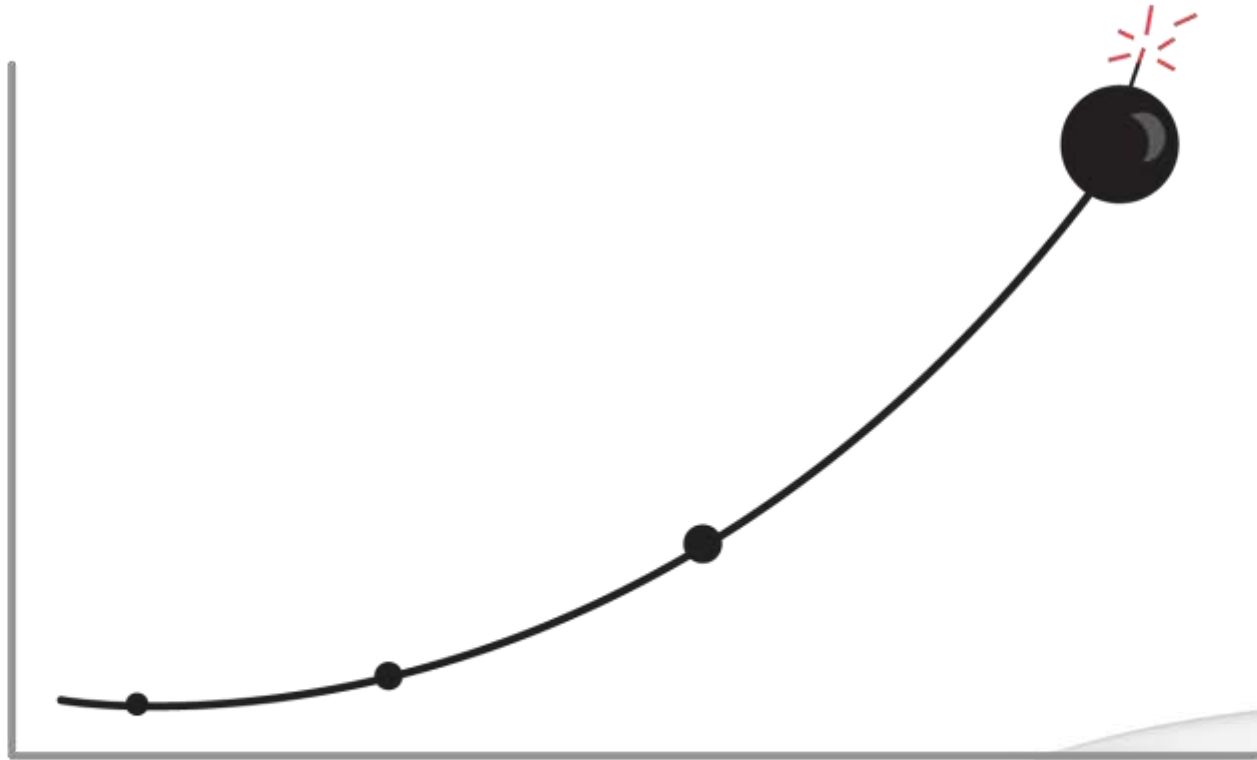
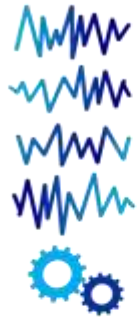
Conclusions



Worldwide Studios
Creative Services Group



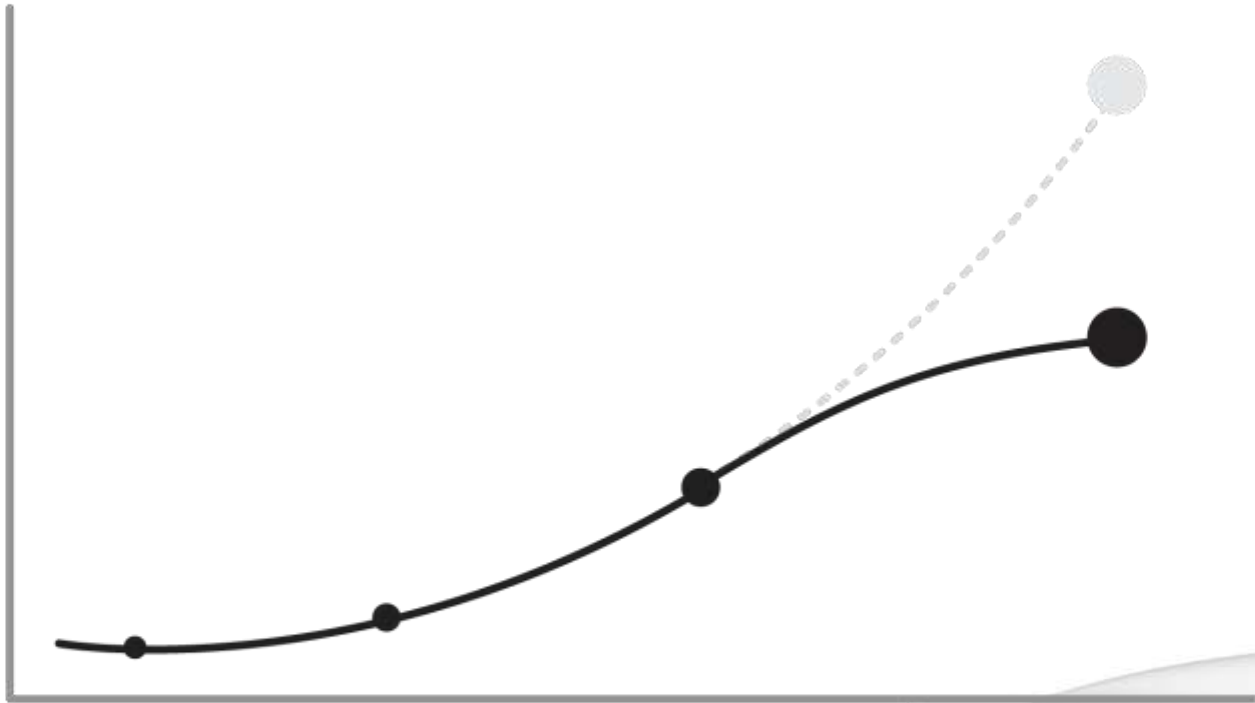
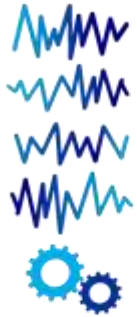
How do we diffuse that bomb ?



Worldwide Studios
Creative Services Group



Smarter Audio Pipelines & Engines !



Worldwide Studios
Creative Services Group



Future Technologies for Game Sound Design

- Data-driven as much as possible
- Better integration with other game sub-systems
- Knowing your data is the key to smarter tools and engines (audio analysis)
- Generate audio at run-time when it makes sense (procedural audio using models created with audio analysis)





Thank you!

Any questions?

Contact: nicolasfournel@yahoo.com
nicolas_fournel@scee.net



Worldwide Studios
Creative Services Group

