

BiOWARE

***Automated Emotion:
Facial Animation in***



***Ben Cloward
Senior Technical Artist***

A Brief Introduction

I really enjoy key framing facial animation by hand

Bioware's products have lots of facial animation . . .

... So I should have a great time, right?

I had no idea what I was getting into.



BiOWARE

I love creating facial animation by hand. It's really fun to bring a character to life by keyframing mouth movements, facial expressions, and eye movements. I was really excited when I took the job at Bioware because I knew that Bioware games focused on story and I knew I'd get the chance to do lots of facial animation. I had no idea what I was getting into.

How Big Is This Project

Hundreds of thousands of lines of dialog

More dialog than all previous Bioware games put together

More dialog than any game ever made

Probably the largest voice-over project ever



BiOWARE

Star Wars: The Old Republic has hundreds of thousands of lines of dialog. That means that the game has more dialog than every previous Bioware game - put together. It has more dialog than any other game ever made - and it's probably the largest voice-over project ever - not only for games but also including movies and TV shows.

How Long Would It Take?

Working a normal schedule - 164 years

So we could ship the game by 2173

With 4 animators working 24/7 - still 10 years

We have a problem!



BiOWARE

I can generally do about 30 seconds of quality facial animation in about 4 hours. I did some math and I figure at that rate, it would take me around 164 years to complete the facial animation and our game would ship in the year 2,173. Even if we had 4 animators working around the clock non-stop all year, it would still take almost 10 years to complete just the facial animation. I think it's pretty obvious we have a problem.

The Problem

People aren't fast enough

Let the computer do it?

Computers can't animate

Animators want hand-keyed quality



BiOWARE

So our problem is that people aren't fast enough. Now the obvious solution is just - "Let the computer do it." There are lots of software packages out there that do automated lip sync and just about every game out there with a significant amount of voice over work uses automated lip sync. However, we also have a second problem. Our art director and lead animator want the facial animation to look as good as if it were hand animated and most of the computer-generated facial animation out there looks pretty flat and lifeless. So, we had to come up with a system that allows the computer to create facial animation that looks almost as good as hand-keyed animation, and requires no hand-retouching.

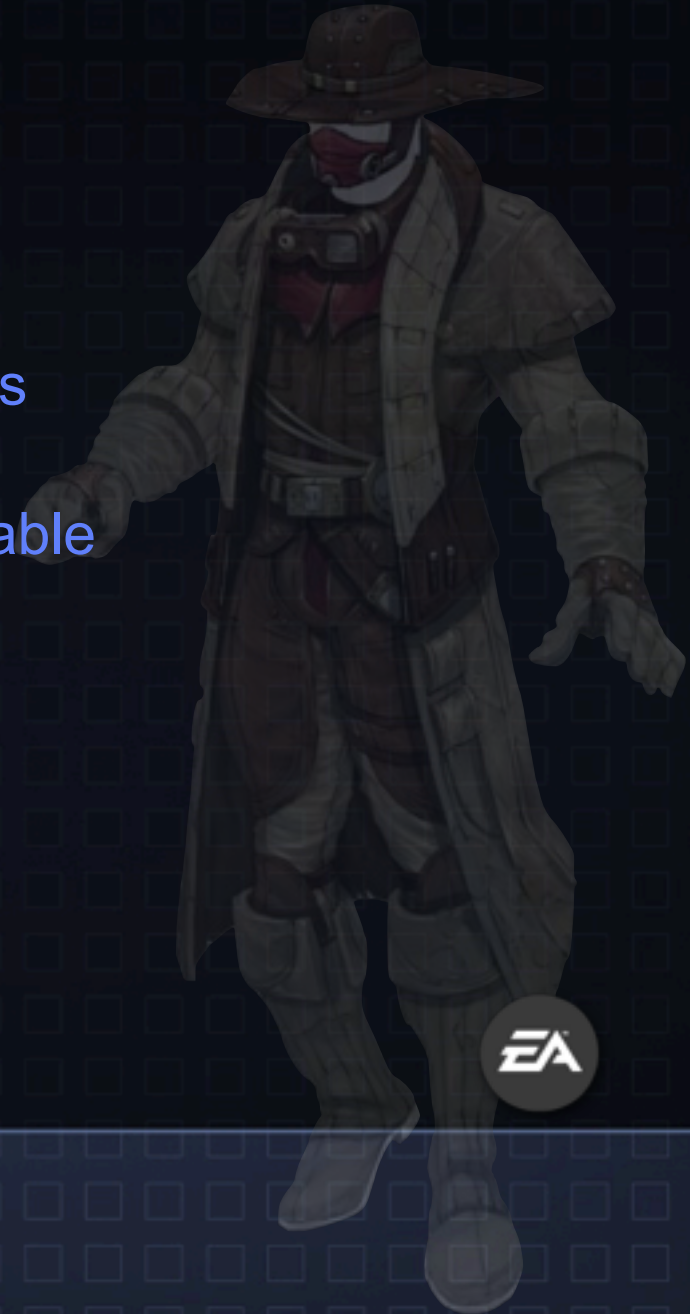
Our Solution

Break facial performance into key elements

Animators create many variations of elements

Computer combines elements using all available information

Computer speed + human quality = Win!



BiOWARE

EA

There are probably lots of good solutions to this problem. The one we ended up going with is to break down the facial performance into key elements - the facial expressions, the lip sync and phoneme shapes, and the head and eye movements. Then we get the animators to create sets of those elements. Finally, we teach the computer to combine those elements automatically and make intelligent choices about which elements to use based on the information available. This solution gives us the best of both worlds - it's fast and automated, and it also looks like it was hand-animated. Also mention that we started out with the system Bioware used for Mass Effect.

Talk Overview

Part 1: Creating Key Elements

- Phoneme Shapes
- Facial Expressions
- Head and Eye Movements

Part 2: Putting It All Together

BiOWARE



So, in this talk I'm going to cover our process of creating each of those key elements - the head and eye movement, the facial expressions, and the phoneme shapes and talk about some of the things we learned as we created them. Then I'm going to talk about the system we developed for automating the process of combining the elements.

Part 1: Creating Key Elements

Phoneme Shapes and Lip Sync

BiOWARE



So first - let's talk about creating just the mouth movements. The mouth movements are created by using phonemes shapes, but first I want to take a minute and talk about our facial rig. I personally like to use morph targets when I'm doing the facial animation for an individual character, because they give you more control over the exact shapes of the face. But for this project we have a wide variety of alien head shapes, and we wanted them to be able to all use the same facial animation. Using bones for the facial rig allows us to weight each head to the same animated bones.



But for this project we have a wide variety of alien head shapes, and we wanted them to be able to all use the same facial animation. Using bones for the facial rig allows us to weight each head to the same animated bones.



Our bones rig has twenty one bones representing facial muscles in the eye brows, eye lids, cheeks, lips, jaw, and tongue. Each head mesh is very carefully weighted to these bones in order to mimic the appearance of facial muscles.



Phonemes are mouth shapes that represent each of the basic sounds that the mouth makes during speech. Automated lip sync software combines these shapes to create lip sync animation.

Iteration

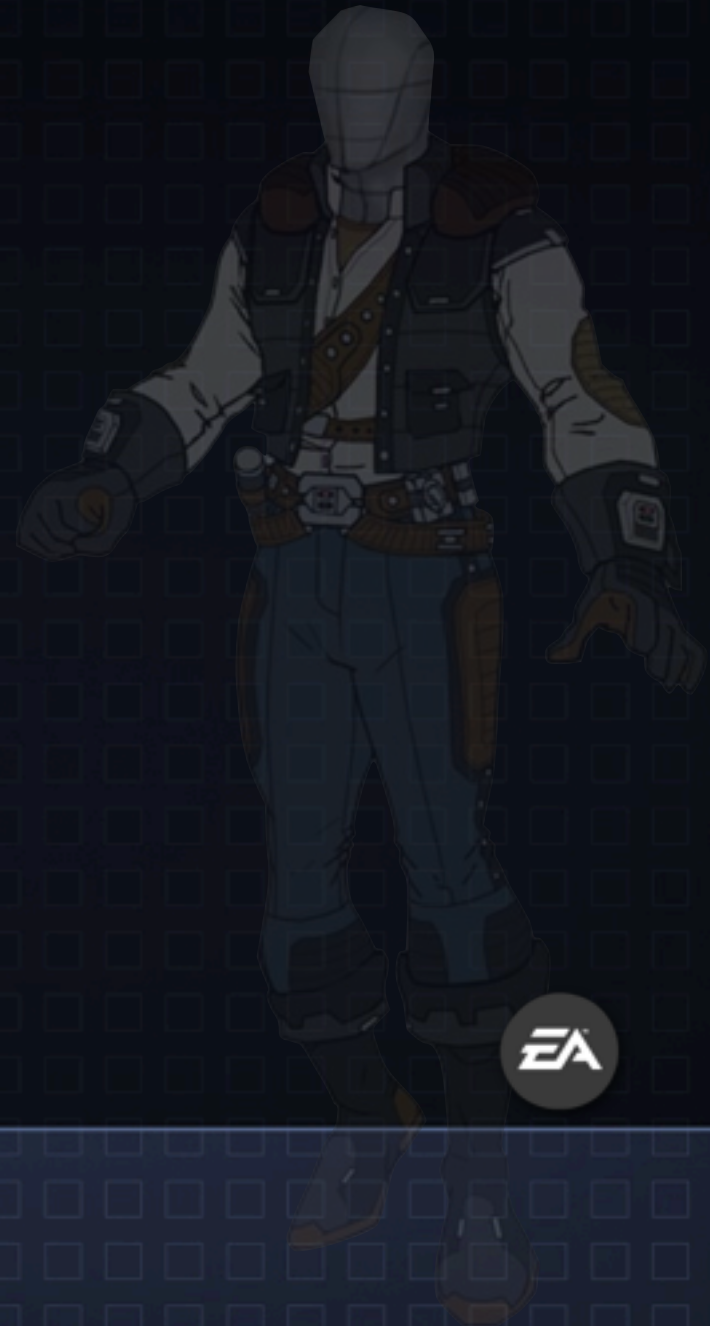
Make a set of shapes

Use them to generate animation

Find problems

Adjust shapes

Repeat until lip sync looks good



BiOWARE

EA

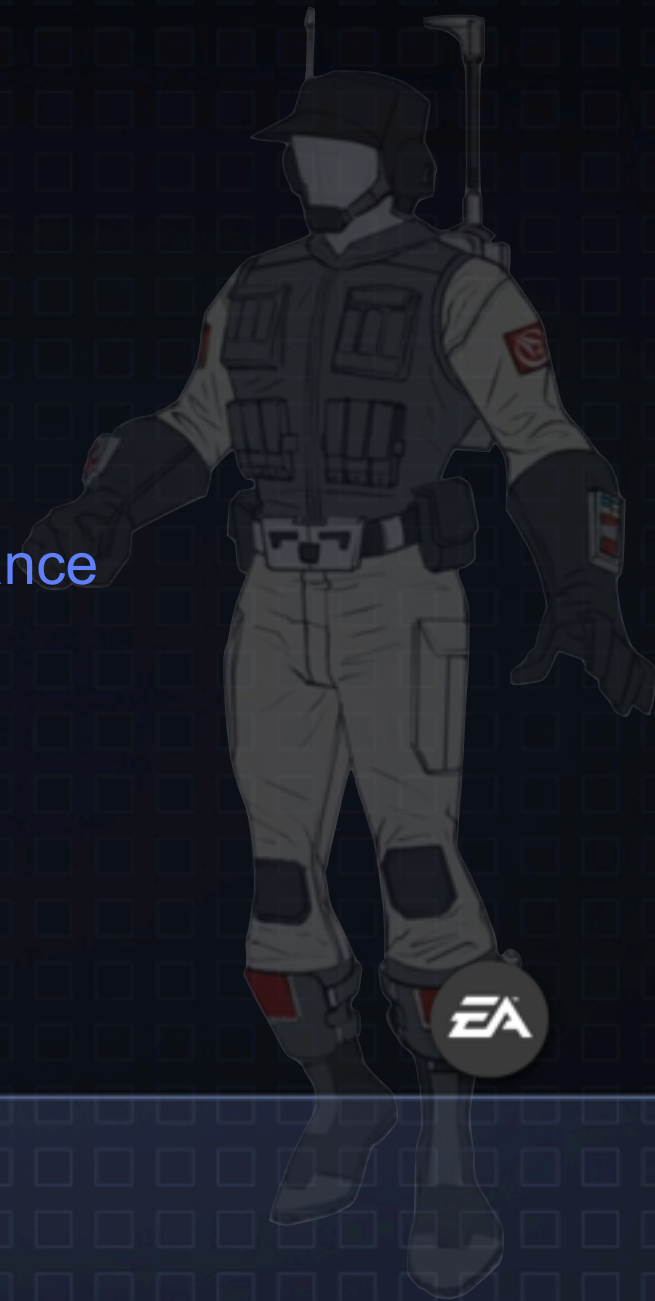
One of the things we learned early on is that making a good set of phonemes take a lot of iteration. Just because a single mouth shape looks correct by itself does not mean that it will work well when blended together with all of the other shapes. In order to create a set of phoneme shapes that worked well, we had to use them to create lots of sample animation. Then we'd find shapes that stuck out or didn't work quite right, and we'd tweak them. Then we'd create the sample animation again and look for more problems.

Mouth Corner Example

Mouth corners too jittery

Toning down the movement killed the performance

BiOWARE



One case where iteration really helped was with the corners of the mouth. With our first attempts, the mouth corners appeared really jittery and moved around too much. At first we tried toning down the variation in mouth corner poses from one phoneme to the next. This got rid of the jitters, but it killed the overall appearance of the phonemes and resulted in lip sync that looked really flat.

Solution

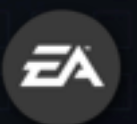
Eee
Phoneme

Ooo
Phoneme

Wide
Mouth

Narrow
Mouth

BiOWARE



Our second solution was to remove corner movement from the phonemes completely and separate it out into it's own shapes – corners in, corners out, corners up, and corners down. We assigned each phoneme a percentage weight value for how much corner in-out movement to apply and how much up-down movement to apply. The amount of in-out and up-down movement blended smoothly across the spectrum of phoneme shapes. This solution gave us much more uniform control over the mouth corner movement.

Sample Video #1

With a good set of phonemes our software was able to animate the mouth movements as our characters talked. Here's our first example video.

Sample Video #2

I think the speech animation looks pretty good at this point, but overall the character's performance is pretty flat. That works ok for a line of dialog like that one where the character is just delivering information. But this next video shows off the problem of using only lip sync animation. (show video) I think the problem is pretty obvious here. The voice actor's performance is filled with emotionally charged delivery and he's really bringing the story to life, but the face is totally lifeless. So we need our characters to show emotions that match the feeling in the audio.

Facial Expressions - Core Emotions

Disgust



Joy



Fear



Sadness



Anger



BiOWARE



We started by creating a set of facial expressions by posing our facial bones. These are our core expressions.

Disgust

Joy

Fear

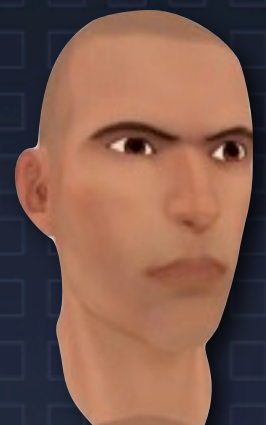
Sadness

Anger

Statement

Question

Exclamation



For the core emotions – anger, joy, sadness, fear, and disgust – we created three different versions: one for a statement, one for a question, and one for an exclamation. When people talk, they express the same emotion slightly different depending on the type of statement they’re making. We wanted the characters in our game to be able to do that too.

Special Expressions



BiOWARE



We also knew that we would have characters that would need some more subtle expressions like the sly expression on a smuggler, or a flattering or boot lick expression of a character who's trying to be persuasive. With our core expressions and our set of special expressions, we have a total of 27 possible facial expressions.



We tried to make each of the expressions show an emotion that was recognizable from a single frame and without the need of context clues. Some emotions can be mistaken for something else when taken out of context, so we had to avoid those.



Once we had created our set of 27 expressions, we made three variations of each giving us a total of 81 expressions. We needed the variations because we knew that the expressions would be used over and over again and we didn't want them to feel repetitive.

Subtle Mouth Changes

Mouth changes subtle

Additive with lip sync

Most expression in eyelids and brows

Mouth hints at expression



BiOWARE

EA

When making the expressions, we had to be careful to make changes to the mouth very subtle. The expressions are additive with the lip sync animation, so large changes in the mouth shape would distort the mouth during speech. We had to do most of the expression in the eyes and eye brows and only make small changes to hint at the expression in the mouth.

Sample Video #3

So here's an example video showing the results of the facial expressions. There's a before and an after for each test case.

Angry Phoneme Set

Second set of angry phonemes

Additive with the first set

Each expression controls the percentage of angry phonemes to apply

BiOWARE



When a character is angry, he tends to have very accentuated and large mouth movements. We weren't able to get these with our expressions and our base phoneme shapes. We decided to create a second set of phoneme shapes specifically for enraged speech. This set of phonemes is additive with the base set and we're able to dial in the amount of "enragement" to use by controlling the percentage of the angry phonemes. Each of our 81 expressions has a value in it that determines how much of the angry phonemes to use – so when we add an angry expression to the face, it also automatically starts using the angry phonemes.

Sample Video #4

Here's an example video showing results before and after adding in the angry phoneme shapes. The effect is subtle, so watch the mouth shapes closely.

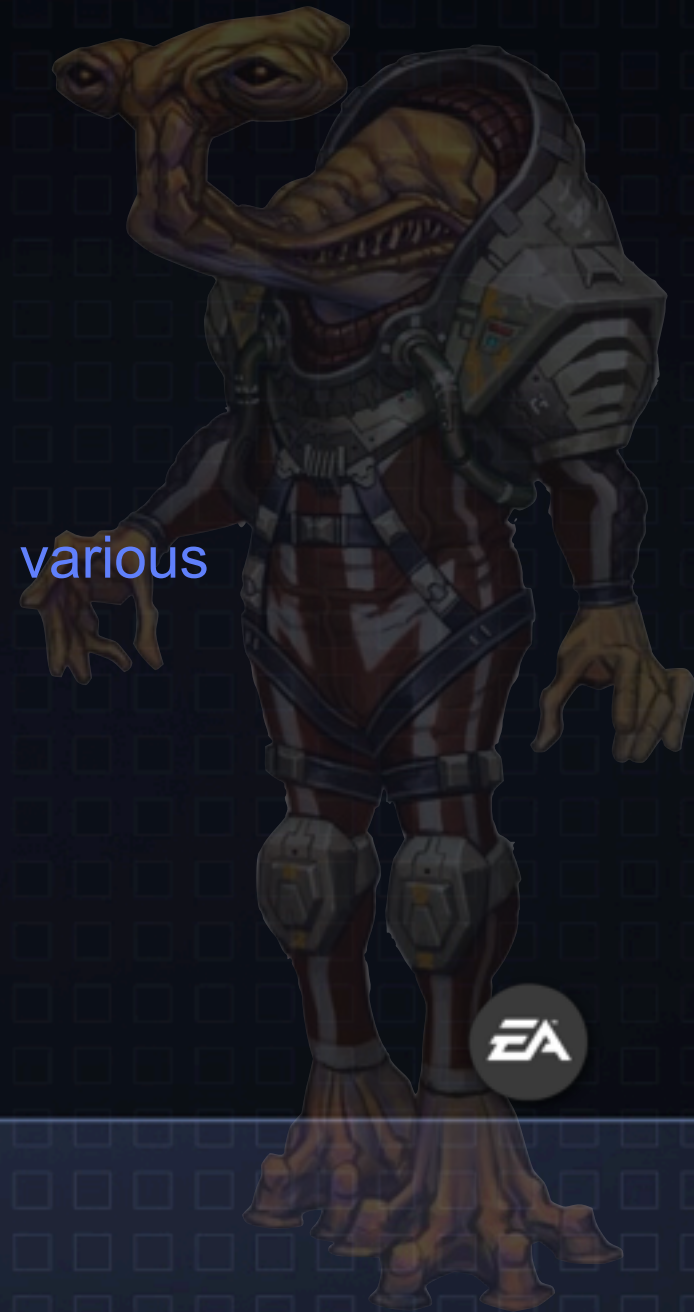
Head and Eye Movements

Animators create eye dart sequences

Animators create head movements to go with various phrase lengths and emotions

BiOWARE

EA



To round out the facial performance, we had the animators create a set of eye darting movements and head movements. We created head movements to go with many types of emotions and made several variations of each of these for different lengths of phrases.

Sample Video #5

Here's an example showing the head movements – before and after.

Part2: Putting it All Together

Phonemes



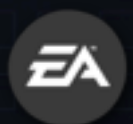
Expressions



Head Movements



BiOWARE



So now we have a set of phonemes and a second set for when the character is angry. We have a set of 81 facial expressions to draw from. And we have a large set of head and eye movements. If we just let the computer randomly pick expressions and head movements, we end up with chaos. We have to develop some rules that the computer can follow to more intelligently choose assets to combine for each phrase.

Available Information

Length of the phrase

Punctuation at the end of the phrase

Audio and text files

Character personality



BiOWARE



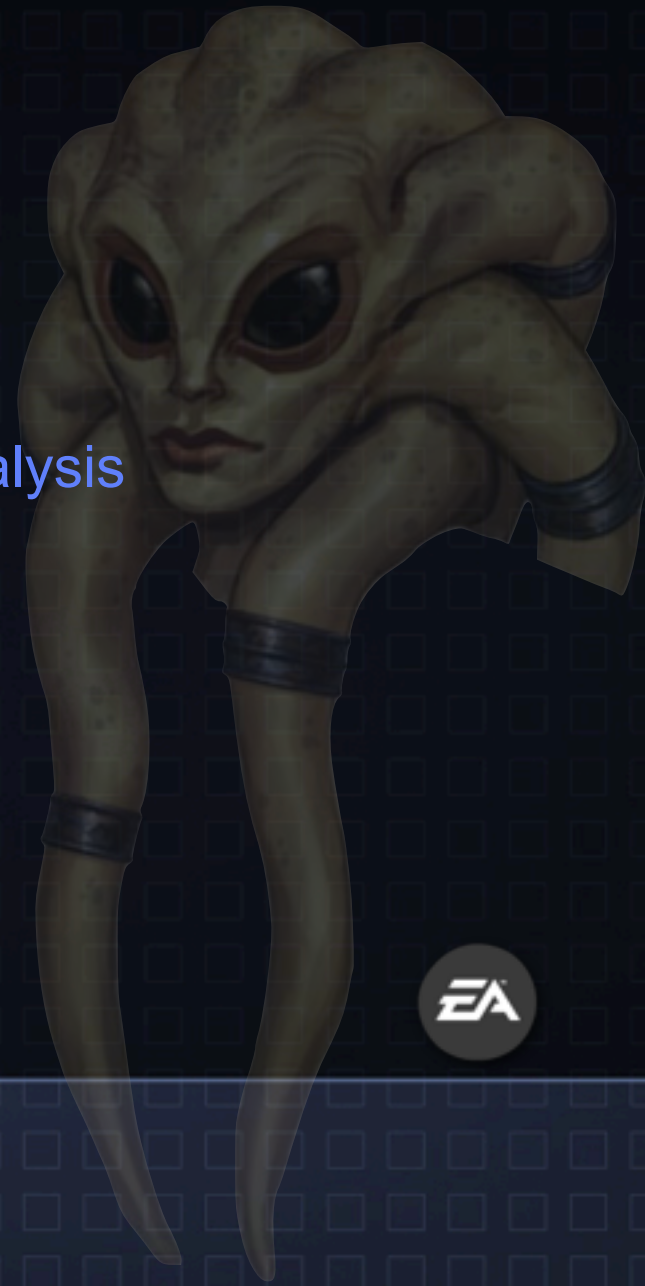
We have to use all of the available information when selecting which assets to combine to create the final facial animation. So what information do we know? We know the length of the phrase being spoken. We know the punctuation at the end of the phrase. We have the audio and text files available. As a final bit of the puzzle, we define a personality for each of our characters.

How Phoneme Shapes are Chosen

Computer analyzes both audio and text

Selects and combines phonemes based on analysis

Also adds subtle head and brow movement



BiOWARE



The software uses both the audio wave file and the text file together to select and combine the phonemes to create the lip sync animation. At this stage, it also adds subtle head and eye brow movements based on audio cues.

How Expressions are Chosen

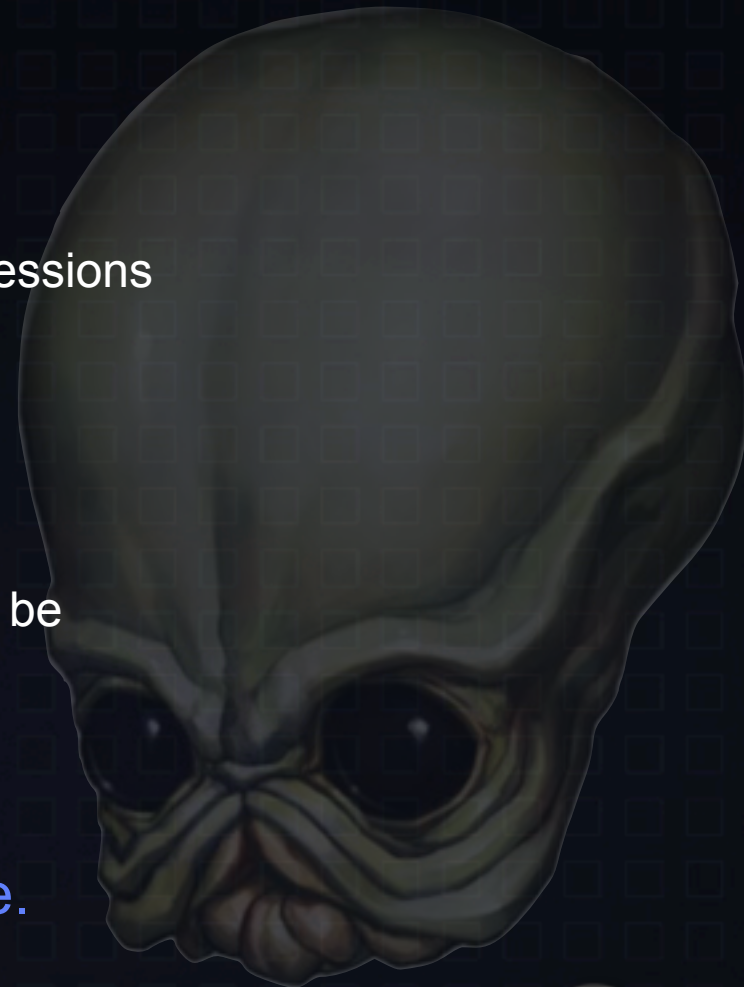
Total expression set is narrowed down based on character personality.

- An angry character might have a set of 8 expressions instead of the full 27.

The sub-set is further narrowed down based on phrase punctuation.

- Of those 8, the angry character has 3 that can be used for a question.

Of the remaining expressions, one is selected at random for the phrase.



BiOWARE



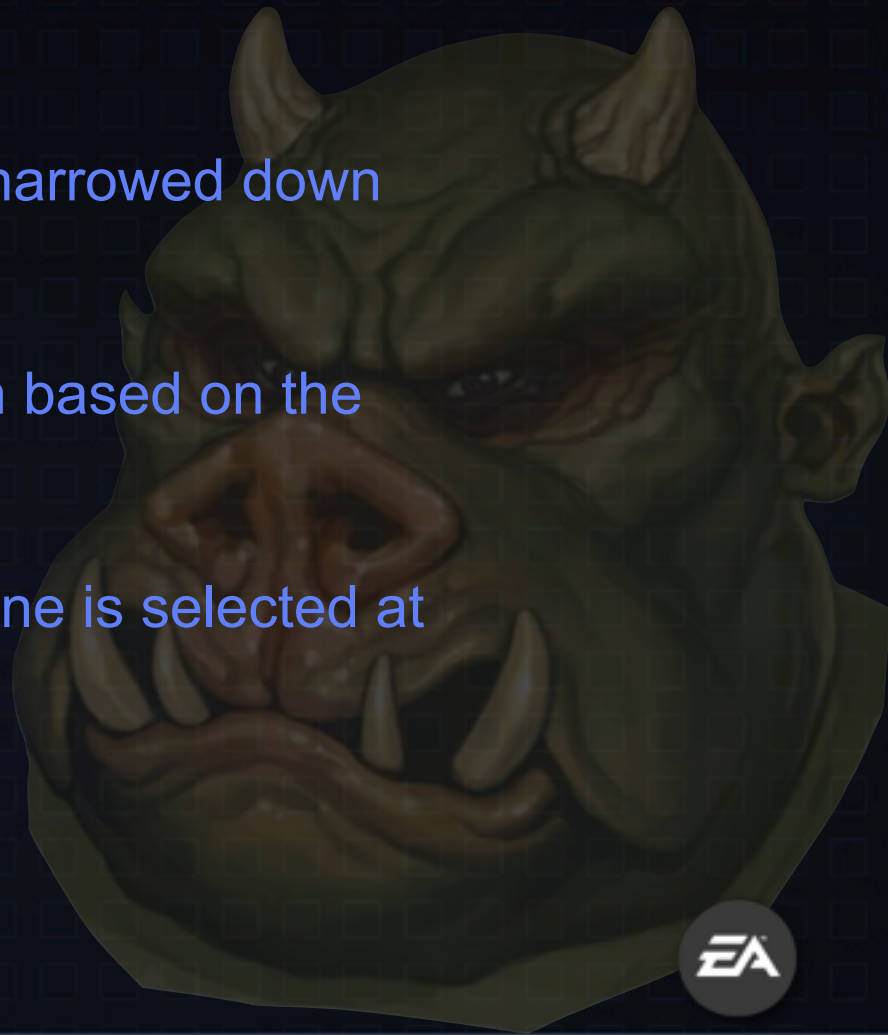
Based on the character's personality, we create a sub set of possible facial expressions that the character can use. We narrow the set of expressions down further using the type of punctuation at the end of the phrase. The the computer selects a random expression of the remaining available expressions.

How Head Movements are Chosen

Total number of head movements is narrowed down based on the length of the phrase

The sub-set is further narrowed down based on the facial expression that was selected.

Of the remaining head movements, one is selected at random for the phrase.



BiOWARE

EA

Head and eye movement animations are selected based on the length of the phrase and also based on which expression is chosen.

What the Computer Does

Pre-process text filter

- *Please! I am a fellow Sith. Judge me with an open mind, and grant me trial by combat - I beg you.*

Remove punctuation from phrases that are too short

- *Please I am a fellow Sith. Judge me with an open mind, and grant me trial by combat I beg you.*

Replace commas with periods

- *Please I am a fellow Sith. Judge me with an open mind. and grant me trial by combat I beg you.*

Replace punctuation with special symbols

BiOWARE

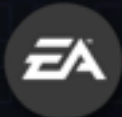


First, we run a pre-process text filter on the copy that the writers create. In this step, the computer first looks for phrases that are too short. Single word phrases like “Huh?” or “Yes” usually don’t animate well, so the filter removes the punctuation following phrases that are too short to combine them with the following phrases. Next, the filter replaces all commas with periods. Remember that we base our expressions on the final punctuation – so more final punctuation means we have more opportunities to change expressions.

Symbols Represent Expressions

. Base_Affirmation	# Sad_Disappointed	@ Angry_Mad	< Joy_Smile
? Base_Question	^ Sad_Question	* Angry_Question	> Joy_Question
! Base_Exclamation	\$ Sad_Miserable	% Angry_Enraged	= Joy_Excited
~ Fear_Worried	© Disgust_Conceited	´ Special_Surprised	— Special_Sly
{ Fear_Question	Æ Disgust_Question	^a Special_Amazed	fl Special_Bootlick
} Fear_Afraid	± Disgust_Distain	Δ Special_Ecstatic	β Special_Flatter
μ Special_False	« Special_Eager	≈ Special_Melancholy	

Please I am a fellow Sith \$ Judge me with an open mind ≈ and grant me trial by combat I beg you #



Finally, we replace all of our punctuation with symbols that represent the facial expression that should be applied to each phrase. The narrowing down process I described earlier is used to select an appropriate symbol.

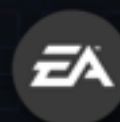
Sample Video #6

And once the text is processed, we pass it through the system and end up with something like this.

Automated Facial Animation



BiOWARE



Next, we pass our filtered text file and our audio file to FaceFX. The software processes the text file and the audio to create the lip sync animation. Our custom FaceFX plugin then uses the length of each phrase and the punctuation symbols at the end of each phrase to select appropriate head movements, eye movements, and facial expressions to add to each phrase.

Final Results

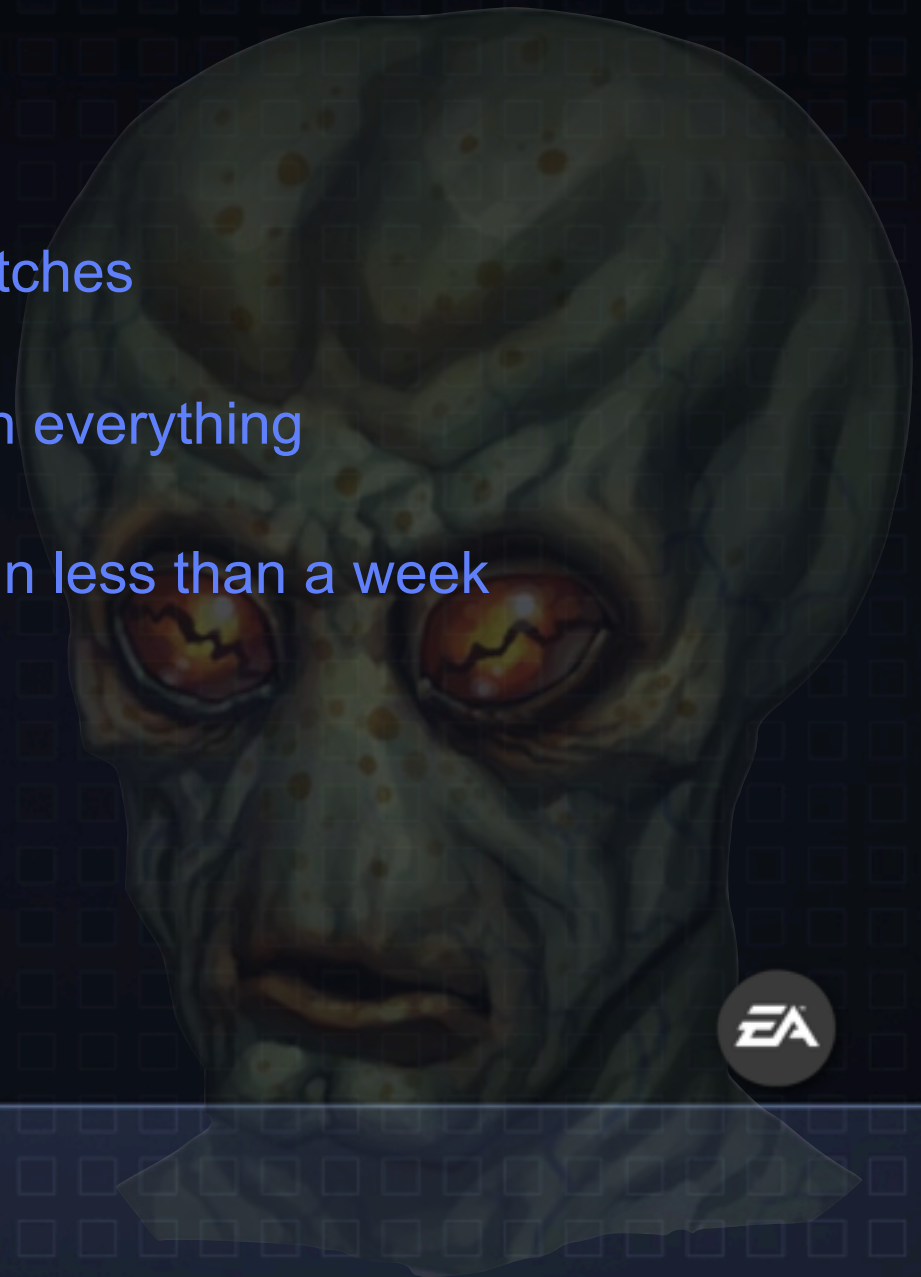
Processing can be done in large batches

Can make tweaks and then re-batch everything

Could generate all facial animation in less than a week

Much better than 164 years

BiOWARE



We pass all of this processing work to the computer in large batches. We create the facial animation for large sections of the game all at once. We have a dedicated machine with 8 processors and lots of RAM and we've worked with OC3 (the company that makes FaceFX) to streamline and multi-thread the processing as much as possible. Because the system requires no retouching by hand we're able to easily make tweaks to any element in the system – the expressions, the wording of the dialog, the head movement animations, etc, and then we can reprocess large portions of the game in just a few hours. The writers are free to change lines of dialog and throw in re-recorded dialog at any time. If we needed to, we could re-process all of the dialog in the game in less than a week. Compared with my initial estimate of 164 years, I think we came up with a pretty good solution.

Integration With Design Tools

Designer expression tweaking tool

Designers also have tools for full-body animation, camera placement, staging, and lighting

Complete package is a powerful system that makes the game possible

BiOWARE



Once the initial batch processing is done for a conversation, we also developed a tool that allows the designers to edit the results. In the tool, each phrase of dialog is displayed with a drop-down box below it. The box contains a list of all of the available facial expressions and its set to the expression that is currently being used for the phrase. If the designer wants to improve the choice of expression that the computer made, he can select a different expression from the list. Then with the push of a button, the dialog is instantly reprocessed with the new selection. This gives the designers the flexibility and control they need to get the results they want – but without requiring them to hand-touch every phrase.

When you combine this facial animation tool together with our other tools for full-body animations, and character, camera, and light placements, you get a powerful combination that allows the designers to put entire conversations together at an amazing rate. We've achieved our goal of creating a system that generates facial animation that comes close to hand-keyed quality, is fast and flexible, and works within the requirements of the rest of our team.

Sample Video #7

And finally, here's an example of what our facial animation looks like when it's applied to aliens with different head anatomy.

Credits and Thanks

Mark How - Lead Animator

Kiko Buyo - Senior Animator

Charles Beyer - Principle Programmer

Ben Cloward - Senior Technical Artist

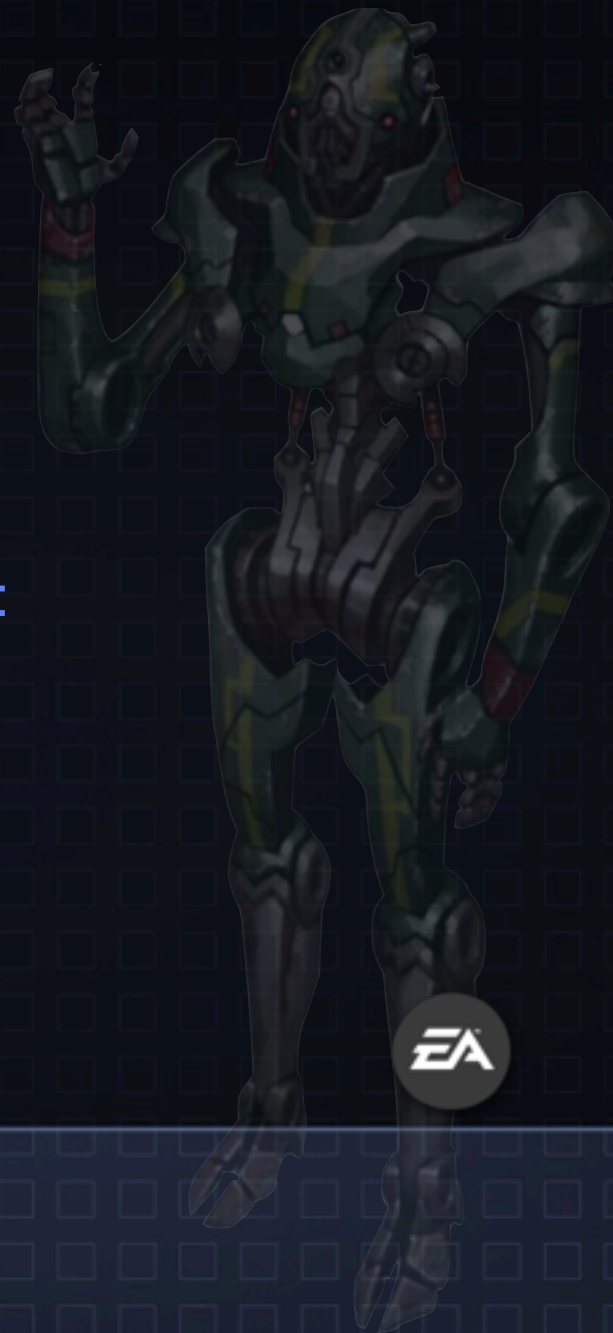
Original Facial Animation System in Mass Effect:

Ben Hindle

Rejean Poirier

Oliver Jauncey

BiOWARE



Questions?

Ben Cloward

bcloward@bioware.com

<http://www.bencloward.com>

BiOWARE

EA