Topic 5 - Character Animation

- We **communicate** using our **bodies**. When we talk, we typically use gestures maybe to point something out or just to emphasize what we're saying
- We **express emotions** through our **facial expression**
- Making **eye contact** is one of the most **important social signals**
- Our speech is the explicit meaning of the conversation, and all of the other body language is structured around the rhythms of speech. Our **tone of voice** says a lot as well.

Introduction to skeletal Animation

- **Human bodies** are represented by a **polygon mesh**
- We could create a **compound object** with separate objects for our hand, lower arm, and upper arm and this is what a **early computer games and computer graphics did**
- **Movement** actually comes down to the movement of the **skeleton.**
- **The skeleton** consists **of a number of rigid bones linked together** by rotational joints
- This skeleton needs to be attached to the mesh, which means that we need to say which bone each vertex is attached to. A process called **rigging**.

Forward Kinematics

- A **cascade of components** that **can be transformed theirs positions and rotations** to produce an animation
- Skeletal animation works through a process called Forward Kinematics or FK
- FK is just handled automatically by the way transforms work

Inverse Kinematics

- IK is an **algorithm** which **takes the position where you want the hands** and does a **load of calculations** to work out what the rotations of all the joint should be to get your hand in the right position

Posture and Gesture

Gestures are movements we make while talking, mostly with our hands, but sometimes with our heads or other body parts. Gestures can be meaningful. They can add extra meaning beyond the words we say. For example, some gestures have very definite meanings, like the okay sign of thumbs up. These are called iconic gestures. Iconic gestures can vary a lot between cultures. Nodding and shaking our heads can mean yes and no in Western cultures. But in India, head shaking gestures mean agreement.

We use gestures mainly to:

- Add extra meaning beyond words we say
- Point at things
- Emphasis on what we say

All of the gestures we've been talking about are very closely linked to speech. Sometimes they can substitute speech, but they mostly accompany and enhance it. The timing at rhythm of gestures is closely synced to that of speech. That means if you want characters to do gestures, then we need to create them together with speech. If we're generating speech for the chatbot, then we need an algorithm to dynamically generate gestures to go with it, which can be pretty hard.

Posture on the other hand, is the pose of the whole body when it doesn't move. Posture is less of a directly communicative channel. But it can express a lot of things. For example, emotion. We can easily recognize a sad, dejected posture and distinguish it from a pro posture.

Blinkin and Lip Sync Animation

- Maintaining emotional facial expressions can be hard and time-consuming. Motion capture data from one person might not apply very well onto a different phase.
- Blinking and lip-sync are the two very important features. It is quite creepy to talk to a human-looking avatar who does not blink and it's really weird and could be confusing to interact with an avatar who talks without opening and closing their mouth.
- The **average duration of a human blink** is between 0.1-0.4 seconds. If you are rendering your animation at 60 frames per second, you will have to have the whole blinking finished in between 6-24 frames.
- That is, for a really slow blink, you should set frame 1 for eyes open, frame 12 for eyes closed and frame 24 for eyes open again. It's also useful to know that we blink about **15-20 times per minute**. That's every 3-4 seconds.
- The basic implementation with **lip-sync** works in a similar way. When the avatar is talking, we just need **to set her mouth to be continuously opening and closing** until she stops talking.
- We can make it more realistic, but try **to match her mouth shape to her speech**. Each sound of phoneme we produce has a distinctive mouth shape, which is called a **viseme**

Which of the following is true about lip sync?

- Visemes are visual versions of basic sounds called phonemes
- Blend shapes representing the mouth shapes when we make certain sounds are called visemes

Which of the following could be a realistic blinking?

- A blink of 10 frames with the frame rate of 75 frames per second (Fast)
- A blink of 40 frames with the frame rate of 120 frames per second (Slow)

Gaze animation

- The eyes are perhaps the most important social signal of all. Making eye contact is a form of social connection. It shows that we're paying attention to someone and can indicate a range of social meanings from intimacy to aggression
- When we say emulate gaze behavior, there are only way to **two states**, the character is **either looking at the other player or not**

Questions

A conversation state machine of a character could have three states, they are:

- Talking
- Listening
- Neither.

Which of the following are true about eye gaze?

- We are quite sensitive when someone is looking at us, but we pay less attention to all the other places they are looking. So when we look away, looking around a random place is fine.
- We look at our conversation partner more when we are listening and a bit less when we are talking