Reading: The first half of the lecture is taken from Chapt 1 of Gregory, *Game Engine Architecture*. The second half comes from any modern computer graphics text.

Computer Game Architecture: A large computer game is a significant technical undertaking, involving a large number of interacting components. Of course, not all computer games require the same level of complexity. Different genres of games require different capabilities. The combination of components used for a simple casual 2-dimensional game (e.g., Tetris, Mahjong, or Bejeweled) is very different from a high-end 3-dimensional first-person shooter game (e.g., Unreal, Call of Duty, or Halo).

In order to show the contrast of requirements among different games, here are some examples of a few common game genres and the typical technical elements required to implement these games.

First-Person Shooters (FPS): These games involve simulating a relatively slow entity moving on foot through a highly structured world, typified by passageways and corridors. Modern versions can include more complex types of locomotion, such as climbing. Typical features of these games include:

- real-time rendering of large 3-dimensional virtual worlds
- highly responsive camera control/aiming mechanics
- high fidelity animations of player’s virtual arms and weapons
- availability of a separate (through the scope) camera mode for accurate aiming
- a range of hand-held weapons
- forgiving player character motion and a loose collision model, which gives the feeling of floating,
- high fidelity animations of non-player characters (both enemies and team members)
- multiplayer capabilities

In order to achieve real-time rendering of complex environments, these games frequently limit player movement to well-defined regions, and use various culling techniques that allow the game to draw only the small amount of the total world, which is visible to the player. This is among the most challenging genres from the perspective of the complexity and scale of the technologies required.

Platform Games: This genre is typified by games like *Donkey Kong*, *Super Mario Bros* and *Sonic the Hedgehog*. The player controls an animated character that is seen externally, that is, from a third-person perspective. This character moves through an obstacle course in pursuit of a goal. This may involve running, jumping, swinging, and climbing. Features of these games include:

- moving platforms, ladders, ropes, trellises and generally a variety of locomotion modes
- puzzle-like environmental elements
- third-person camera follows the player character
- (particularly for 3-dimensional platform games) complex camera collision system that ensures that the player character is always in view
In contrast to FPS games, accuracy in locomotion is a key element to the game play. Players must accurately time their jumps, or combine running with jumping to avoid specific obstacles. In order to keep the player character continuously in view, it may be necessary to “cheat” on the geometry by allowing the camera to see through obstacles (e.g., walls) that are very close to the character, since they would otherwise block the player’s ability to see the character.

**Fighting Games:** These involve two players controlling humanoid characters that battle one another. Common examples include *Street Fighter*, *Tekken*, and includes sports boxing games like *Fight Night*. Features of these games include:

- a variety of fighting animations
- accurate hit detection
- input system capable of complex button/joystick combinations
- simple backgrounds
- simple camera motion focused on the center of the action

**Racing Games:** These games involve simulating driving a vehicle (typically an automobile) on a track. These include realistic games, such as *Need for Speed* and *Gran Turismo*, and cartoonish ones, such as *Mario Kart*. Like FPS games, these are highly linear. (Note that this does not include games like *Grand Theft Auto* (GTA). Although GTA involves driving simulation, it is not considered racing since the driving is not on a track. Games like GTA where the player can explore vast regions of the world are referred to as open world games.) Some features of racing games include:

- real-time rendering of highly detailed backgrounds that are moving very rapidly through the field of view (which means that various tricks can be used to render them, since the viewer cannot focus on them for long)
- the AI for nonplaying vehicles can be simplified due to the need to stick to the racing circuit
- camera may follow the vehicle from a third-person perspective or may be situated behind the steering wheel from a first-person perspective
- to accurately render the effects of collisions, vehicle modeling may need to be flexible to render dented or scratched vehicle bodies

**Computer Game Engine Architecture:** One way to better understand the software structure underlying a generic game is to understand the structure of a typical game engines. Game engines arose in the mid-1990s. In particular, the software for the popular game *Doom* provided a separation between:

- core game components (such as the rendering system, collision detection system, audio system)
- art assets (models, textures, animations)
- rules of play

This separation made it easy for users to modify, or “modding”, the game, and provided a framework for adding new elements. This model was extended to other games, including *Quake*, *Unreal*, and *Unreal Tournament* (all FPS games). At some point, these simple “modding systems” became generic enough that it was possible to implement a wide variety of very different games based on a common core set of components, the game engine. Examples of modern game engines include the *Unreal Engine*, Microsoft’s *XNA Game Studio*, and *Ogre3D*. 
Game engines vary along a spectrum of ease of use and flexibility. Simple game engines can generate only a single type of game, but are generally easy to pick up and use. Complex game engines can generate a great variety of games, but it can take quite a bit of time to master their various capabilities.

The following is a summary of the basic components of a modern game engine. We think of the engine as being designed in a number of layers, ranging from the lower levels (hardware and operating system) up to the higher levels (game specific entities like rules). Here is a summary of the levels, from low to high.

**System:** This is the hardware and operating system on which the game runs, such as general personal computers (running, say, Microsoft Windows, Linux, or Mac OS), game consoles (XBox, Playstation, Wii), or mobile devices (handheld game consoles or smart-phones).

**Third-Party SDKs and Middleware:** These are libraries and software development toolkits (SDKs), usually provided from a third party. Examples include graphics (OpenGL and DirectX), physics (Havok, PhysX), basic algorithms and data structures (C++ STL, Boost++), character animation (Granny), networking support (Unix sockets).

**Platform Independence Layer:** Since most games are developed to run on many different platforms, this layer provides software to translate between game specific operations and their system-dependent implementations.

**Core System:** These include basic tools necessary in any software development environment, including assertion testing, unit testing, memory allocation/deallocation, mathematics library, debugging aids, parsers (e.g., for xml-based input files), file I/O, video playback.

**Resource Manager:** Large graphics programs involve accessing various resources, such as geometric models for characters and buildings, texture images for coloring these geometric models, maps representing the game’s world. The job of the resource manager is to allow the program to load these resources. Since resources may be compressed to save space, this may also involve decompression.

**Rendering Engine:** This is one of the largest and most complex components of any real-time 3-dimensional game. This involves all aspects of drawing, and may involve close interaction with the graphics processing unit (GPU) for the sake of enhanced efficiency.

**Low-Level Renderer:** This comprises the most basic elements of producing images. As we will see below when we discuss OpenGL, your program interacts with the GPU by asking it to render objects. Each object may be a single triangle (or generally a polygon) or a mesh consisting of many triangular elements. Objects are specified according to their coordinates in 3-dimensional space. Your program also informs the GPU what colors (or what image textures) to apply to these objects, where lights are positioned, and where the camera is positioned.

It is then the job of the GPU to perform the actual rendering (projection, coloring, shading) of the objects. In particular, it determines where each object projects onto the 2-dimensional image plane, which objects are visible and which are hidden from view, what is the color and brightness of each object. Your program needs to convey all this information to the GPU. In our discussion of OpenGL, we will discuss how this is done. This also includes elements like displaying text messages and subdividing the window into subwindows (called viewports) for the purposes of showing status information or maps.
Graphics Device Interface: Since the graphics device requires updating 30–60 times per second, but some operations (such as displaying messages to the user) occurs at a significantly different time scale (of multiple seconds), these components shield the programmer from some of the low-level timing issues when dealing with the graphics system.

Scene Graph: Game entities are naturally organized into hierarchical structures. This is true for dynamic and static objects. For example, a human body consists of a head, torso, arms, legs; an arm consists of a hand, lower-arm, and upper-arm; a hand consists of fingers. Thus, there is a natural structure in the form of a rooted tree.

In general, all the entities of the games world can be represented in a large tree, where the root represents the entire world, and the nodes of the tree implicitly represent the subtrees that are descended from these nodes. This makes it possible to perform operations easily on an entire portion of the tree. For example, we could “render the objects rooted at node $u$” or “rotate the object rooted at node $v$.” We can also create multiple instances, as in “create 200 instances of the zombie object at node $z$.”

This software component is responsible for creating, modifying, rendering, manipulating, and deleting elements of the scene graph. Another feature of using a scene graph is that it allows us to remove, or cull, entities that are not visible to the camera. For example, if the camera is located in a room represented by some node $v$, we need only render the objects lying within this room or that are visible through the room’s doors and windows. Because game worlds are so large and complex, efficient rendering demands that we only attempt to draw the things that might be visible to the camera.

Visual Effects: This includes support for a number of complex effects such as:
- particle systems (which are used for rendering smoke, water, fire, explosions)
- decal systems (for painting bullet holes, damage scratches, powder marks from explosions, foot prints, etc)
- complex lighting, shadowing, and reflections
- image processing effects

Others: This includes visual elements of the user interface, such as displaying menus or debugging aids (for the developer) and video playback for generating the back story.

Collisions and Physics: These components simulate the movement of objects over time, detect when objects collide with one another, and determine the appropriate response in the event of a collision (like knocking down the houses where the little pigs live). Except in very simple physical phenomena, like a free-falling body, physical simulation can be very difficult. For this reason, it is often handled by a third-party physics SDK.

Animation: While the game may direct a character to move from one location to another, the job of the animation system is to make this motion look natural, for example, by moving the arms and legs in a manner that is consistent with a natural walking behavior. The typical process for most animals (including humans) involves developing a skeletal representation of the object and wrapping flesh (which is actually a mixture of skin and clothing) around this skeleton. The skeleton is moved by specifying the changes in the angles of the various joints. This approach is called a skin and bones representation.

Input Handlers: These components process inputs from the user, including keyboard, mouse, and game controller inputs. Some devices also provide feedback to users (such as the vibration in
some game controllers). Modern vision based systems, such as the XBox Kinect, add a whole new level of complexity to this process.

**Audio:** Audio components handle simple things like playback for background music, explosions, car engines, tire squealing, but they may also generate special audio effects, such as stereo effects to give a sense of location, echo effects to give a sense of context (inside versus outside), and other audio effects (such as creating a feeling of distance by filtering out high-frequency components).

**Multiplayer/Networking:** Multiplayer and online games require a number of additional supporting components. For example, multiplayer games may have a split-screen capability. Online games require support for basic network communication (serialization of structures into packets, network protocol issues, hiding network latency, and so on). This also includes issues in the design of game servers, such as services for maintaining registered users and their passwords, matching players with one another, and so on.

**Gameplay Foundation System:** The term **gameplay** refers to the rules that govern game play, that is, the player’s possible actions and the consequences of these actions. Since this varies considerably from one game to another, designing a system that works for a wide variety of games can be quite daunting. Often, these systems may be designed to support just a single genre of games, such as FPS games. There are a number of subcomponents:

**Game Worlds and Object Models:** This constitutes the basic entities that make up the game’s world. Here are some examples:

- static background objects (buildings, terrain, roads)
- (potentially) dynamic background objects (rocks, chairs, doors and windows)
- player characters (PCs) and non-player characters (NPCs)
- weapons and projectiles
- vehicles
- graphics elements (camera and lights)

The diversity of possible game objects is a major challenge to programmers trained in object-oriented methods. Objects share certain universal qualities (e.g., they can be created and destroyed), common qualities (e.g., they can be rendered), and more specialized qualities (e.g., gun-like objects can be shot, chair-like objects can be sat on).

**Event System:** Game objects need to communicate with one another. This is often handled by a form of message passing. When a message arrives, we can think of it as implicitly signaling an event to occur, which is to be processed by the entity. Game objects can register their interest in various events, which may affect their behavior. (For example, characters need to be made aware of explosions occurring in their vicinity. When such an explosion occurs, the system informs nearby characters by sending them a message... “you’re toast, buddy.”)

**Scripting System:** In order to allow game developers to rapidly write and test game code, rather than have them write in a low-level programming language, such as C++, it is common to have them produce their code in a scripting language, like Python. A sophisticated scripting system can handle the editing of scripts and reloading a script into a game while it is running.

**Artificial Intelligence:** These components are used to control the behavior of non-player characters. They are typically modeled as AI agents. As we will discuss later, an agent is an
object that can sense its environment, maintain a memory (or state), and can respond in potentially complex ways depending on the current input and state.

**Game Specific Systems:** This is a catch-all for any components that are so specific to a given game that they don’t fall into one of the other categories. This may include aspects like the mechanics controlling a player’s state, the algorithms by which a camera moves throughout the world, the specific goals of the game’s non-player characters, the properties of various weapons, and so on.

**Interactive 3-dimensional Graphics:** In order to get a quick jump into game development, we will start by discussing the basic elements of interactive computer graphics systems. This will require that we understand a bit about how graphics processing units work and will lead us eventually into a discussion of OpenGL.

Anyone who has played a computer game is accustomed to interaction with a graphics system in which the principal mode of rendering involves 3-dimensional scenes. Producing highly realistic, complex scenes at interactive frame rates (at least 30 frames per second, say) is made possible with the aid of a hardware device called a graphics processing unit, or GPU for short. GPUs are very complex things, and we will only be able to provide a general outline of how they work.

Like the CPU (central processing unit), the GPU is a critical part of modern computer systems. It has its own memory, separate from the CPU’s memory, in which it stores the various graphics objects (e.g., object coordinates and texture images) that it needs in order to do its job. Part of this memory is called the frame buffer, which is a dedicated chunk of memory where the pixels associated with your monitor are stored. Another entity, called the video controller, reads the contents of the frame buffer and generates the actual image on the monitor. This process is illustrated in schematic form in Fig. 1.

Traditionally, GPUs are designed to perform a relatively limited fixed set of operations, but with blazing speed and a high degree of parallelism. Modern GPUs are programmable, in that they provide the user the ability to program various elements of the graphics process. For example, modern GPUs support programs called vertex shaders and fragment shaders, which provide the user with the ability to fine-tune the colors assigned to vertices and fragments.

Recently there has been a trend towards what are called general purpose GPUs (GPGPUs), which can perform not just graphics rendering, but general scientific calculations on the GPU. Since we are interested in graphics here, we will focus on the GPUs traditional role in the rendering process.
The *Graphics Pipeline*: The key concept behind all GPUs is the notion of the *graphics pipeline*. This is a conceptual tool, where your user program sits at one end sending graphics commands to the GPU, and the frame buffer sits at the other end. A typical command from your program might be “draw a triangle in 3-dimensional space at these coordinates.” The job of the graphics system is to convert this simple request to that of coloring a set of pixels on your display. The process of doing this is quite complex, and involves a number of stages. Each of these stages is performed by some part of the pipeline, and the results are then fed to the next stage of the pipeline, until the final image is produced at the end.

Broadly speaking the pipeline can be viewed as involving four major stages. (This is mostly a conceptual aid, since the GPU architecture is not divided so cleanly.) The process is illustrated in Fig. 2.

![Fig. 2: Stages of the graphics pipeline.](image.png)

**Vertex Processing:** Geometric objects are introduced to the pipeline from your program. Objects are described in terms of vectors in 3-dimensional space (for example, a triangle might be represented by three such vectors, one per vertex). In the vertex processing stage, the graphics system *transforms* these coordinates into a coordinate system that is more convenient to the graphics system. For the purposes of this high-level overview, you might imagine that the transformation projects the vertices of the three-dimensional triangle onto the 2-dimensional coordinate system of your screen, called *screen space*.

This stage involves other tasks as well. For one, *clipping* is performed to snip off any parts of your geometry that lie outside the viewing area of the window on your display. Another operation is *lighting*, where computations are performed to determine the colors and intensities of the vertices of your objects. (How the lighting is performed depends on commands that you send to the GPU, indicating where the light sources are and how bright they are.)

**Rasterization:** The job of the rasterizer is to convert the geometric shape given in terms of its screen coordinates into individual pixels, called *fragments*.

**Fragment Processing:** Each fragment is then run through various computations. First, it must be determined whether this fragment is *visible*, or whether it is hidden behind some other fragment. If it is visible, it will then be subjected to coloring. This may involve applying various coloring textures to the fragment and/or color blending from the vertices, in order to produce the effect of smooth shading.

**Blending:** Generally, there may be a number of fragments that affect the color of a given pixel. (This typically results from translucence or other special effects like motion blur.) The colors of these fragments are then blended together to produce the final pixel color. The final output of this stage is the *frame-buffer image*.