CSC 170 – Introduction to Computers and Their Applications

Lecture #2 – Digital Audio Basics

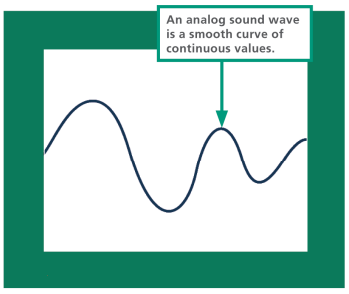
Digital Audio Basics

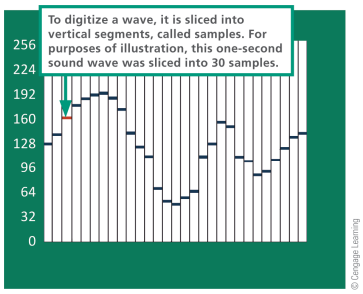
- ***Digital audio*** is music, speech, and other sounds represented in binary format for use in digital devices.
- Most digital devices have a built-in microphone and audio software, so recording external sounds is easy.

Digital Audio Basics

- To digitally record sound, samples of a sound wave are collected at periodic intervals and stored as numeric data in an audio file.
- Sound waves are sampled many times per second by an ***analog-to-digital converter***.
- A ***digital-to-analog converter*** transforms the digital bits into analog sound waves.

Digital Audio Basics





SAMPLE	SAMPLE HEIGHT (DECIMAL)	SAMPLE HEIGHT (BINARY)
1	130	10000010
2	140	1000110
3	160	10100000
4	175	10101111

The height of each sample is converted into a binary number and stored. The height of sample 3 is 160 (decimal), so it is stored as its binary equivalent—10100000.

Digital Audio Basics

- ***Sampling rate*** refers to the number of times per second that a sound is measured during the recording process.
- Higher sampling rates increase the quality of the recording but require more storage space.

Digital Audio File Formats

- A digital file can be identified by its type or its file extension, such as Thriller.mp3 (an audio file).
- The most popular digital audio formats are: **AAC, MP3, Ogg, Vorbis, WAV, FLAC,** and **WMA**.

Digital Audio File Formats

AUDIO FORMAT	EXTENSION	ADVANTAGES	DISADVANTAGES
AAC (Advanced Audio Coding)	.aac, .m4p, or .mp4	Very good sound quality based on MPEG-4; lossy compression; used for iTunes music	Files can be copy protected so that use is limited to approved devices
MP3 (also called MPEG-1 Layer 3)	.mp3	Good sound quality; lossy compression; can be streamed over the Web	Might require a standalone player or browser plugin
Ogg Vorbis	.Ogg	Free, open standard; lossy compression; supported by some browsers	Slow to catch on as a popular standard; part of Google's WebM format

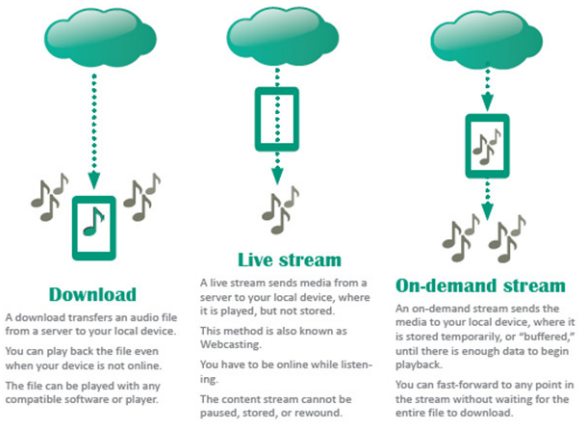
Digital Audio File Formats

AUDIO FORMAT	EXTENSION	ADVANTAGES	DISADVANTAGES
WAV	.Wav	Good sound quality; supported in browsers without a plugin	Audio data is stored in raw, noncompressed format, so files are very large
FLAC (Free Lossless Audio Compression)	.flac	Excellent sound quality; lossless compression	Open source format support ported by many devices
WMA (Windows Media Audio)	.wma	Lossy or lossless compression; very good sound quality; used on several music download sites	Files can be copy protected; requires an add-on player for some devices

Digital Audio File Formats

- ***Ripping*** is a slang term that refers to the process of importing tracks from a CD or DVD to your computer's hard disk.
- The technical term for ripping music tracks is ***digital audio extraction***.
- A ***download*** copies a file from a private network or Internet server to a local device.
- Audio files can be acquired as a ***live stream*** or ***on-demand stream*** in addition to downloads.

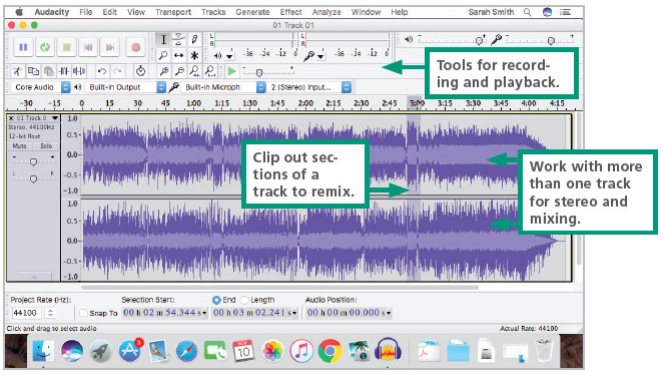
Digital Audio File Formats



Digital Audio File Formats

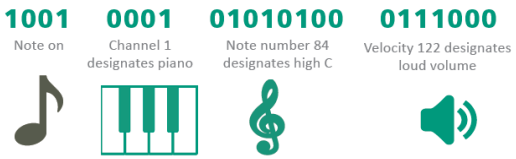
- To play a digital audio file, you must use some type of audio software, such as:
  - **Audio players**: small standalone software application or mobile app.
  - **Audio plugins**: software that works in conjunction with your computer's browser to manage and play audio from a Web page.
  - **Audio software**: general-purpose software and apps used for recording, playing, and modifying audio files, such as iTunes, Windows Media Player, and Audacity.

Digital Audio File Formats



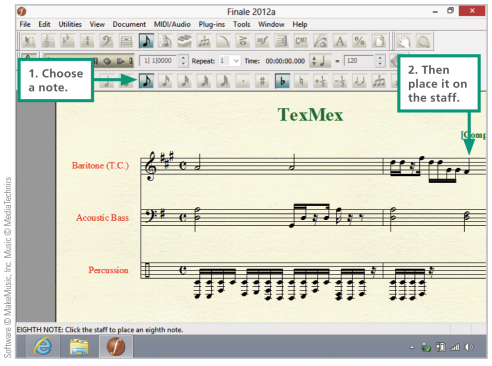
MIDI

- MIDI (Musical Instrument Digital Interface) specifies a standard way to store music data for synthesizers, electronic MIDI instruments, and computers.
- MIDI messages are instructions that specify the pitch of a note, the point at which the note begins, the volume of the note, etc.
- An MIDI message may look like this:



MIDI

- Music composition software with MIDI support makes it easy to place notes on a screen-based music staff then play back the composition on a MIDI keyboard or through the speakers of a digital device.



Digitized Speech

- ***Speech synthesis*** is the process by which machines produce sound that resembles spoken words.
- ***Speech recognition*** (or voice recognition) refers to the ability of a machine to understand spoken words.

Digitized Speech

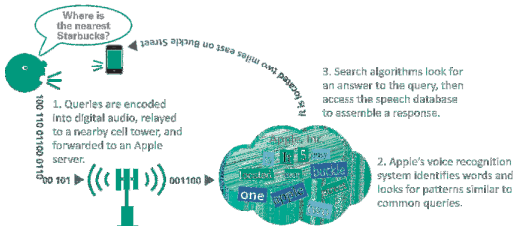
- Speech recognition software analyzes the sounds of your voice and converts each word into groups of phonemes (basic sound units).
- The software then compares the groups to the words in a digital dictionary to find a match.
- When a match is found, the software can display the word on the screen or use it to carry out a command.

Digitized Speech

The voices of Siri were created by live actors, but they couldn't record every possible response to user queries.

Instead they record the words and sentences from a script. Although the script may appear nonsensical, the sentences are rich in phonemes.

When the recordings are complete, a team of linguists analyzes and tags various speech components and loads them into a database. The process is called concatenative speech synthesis.



1. Queries are encoded into digital audio, relayed to a nearby cell tower, and forwarded to an Apple server.

2. Apple's voice recognition system identifies words and looks for patterns similar to common queries.

3. Search algorithms look for an answer to the query, then access the speech database to assemble a response.