COMP-202A, Fall 2007, All Sections
Assignment 5 Background Information

This document contains background information about digital audio. This information is provided only for your general knowledge; reading this document is not necessary to successfully complete Assignment 5.

In the last 20 years, digital audio has grown to become ubiquitous. The Compact Disc (CD) has replaced compact audio cassettes as *the* standard means of music distribution on physical, tangible support, and CD players are now available with almost all home audio equipment, as well as a standard feature in cars. The last 5-7 years have seen the rise of portable digital music players such as Apple's iPod, whose interal storage mechanisms can hold thousands of digital audio tracks.

At the fundamental level, all digital audio works on the same basic principles, whether the audio tracks are distributed on a physical CD, as files over the Internet, or using any other means. Sound consists of waves that travel in a physical medium such as air; these waves have various frequencies, and the frequencies of these waves can be measured and represented as integers. Note that because a finite number of integers can only represent a finite number of frequencies, not all possible frequencies can be represented in this manner.

The basic process of constructing digital audio tracks involves measuring, or *sampling*, the frequencies of sound waves at fixed time intervals, and representing these frequencies as numbers using a fixed number of bits. To replay the constructed track, the audio equipment reads the numbers, and reproduces the frequencies they represent.

Better sound quality is achieved by increasing the number of samplings taken per time unit, or increasing the number bits used to represent frequencies, or both. For example, CD audio measures sound frequencies 44100 times per second and uses 16 bits to represent each sample. It also uses two *channels*; in other words, two devices sample the sound frequencies, and the measurements made by both devices are included in the digital audio track. Therefore, in theory, each second of music stored on a standard CD takes the same space as a text file containing 176400 ASCII characters (44100 samples / second $\cdot$ channel $\times$ 16 bits / sample $\times$ 2 channels $\div$ 8 bits / character). In practice, one second of music takes slightly less space than this, because there exists a way to get 16-bit sound quality using only 14 bits. Thus, CDs capable of holding more than 700 million characters can hold close to 80 minutes of music instead of 70 minutes.

Digital audio is useful to record, manipulate, and distribute sound. However, the storage size required to store even a second of digital audio has led to the design of *compression schemes*. In the context of digital audio, compression involves finding ways to minimize the space used to store a digital audio track while losing as little as possible of the information given by the sampled frequencies.

Compression schemes can be categorized in two broad classes: *lossless* compression schemes and *lossy* compression schemes. Lossless compression schemes attempt to minimize the space used to store a digital audio track without any loss of the information obtained from the sampled frequencies. Thus, if one applies a lossless compression scheme to a digital audio track and subsequently applies the corresponding decompression scheme, the recovered track will be exactly equal to the original.

On the other hand, lossy compression schemes minimize the space used to store a digital audio track while allowing some of the information given by the sampled frequencies to be lost. Thus, if one applies a lossy compression scheme to a digital audio track and subsequently applies the corresponding decompression scheme, the recovered track will *not* be equivalent to the original. Lossy compression schemes generally use models of human hearing to select the frequencies to discard, so that only frequencies that are inaudible to the vast majority of humans are discarded.

In the last few years, various audio compression schemes and formats have become available. MPEG-1 Audio Layer 3 (more commonly known as MP3) is undoubtedly the most common of the available lossy audio formats; it is supported by most (if not all) portable digital music players, as well as home and car stereo equipment, mobile phones, and various other electronic devices. Numerous software music players for use on general-purpose computers also support this format. Such software music players include Apple's iTunes, Microsoft's Windows Media Player, Real's Realplayer, Nullsoft's Winamp, AmaroK (the default music player in the KDE desktop environment), Rhythmbox (a common music player for the GNOME desktop environment), XMMS (a music player for Unix-like operating systems including Linux) and its various forks such as Audacious, and many, many more.

On the other hand, the fact that the MP3 format is patented in some countries has motivated the design of other, patent-free and royalty-free lossy audio formats such as Ogg Vorbis. Ogg Vorbis produces smaller files than other audio formats at equivalent or higher quality, but is not supported by many manufacturers of audio equipment. On the lossless compression front, FLAC (also patent-free and royalty-free) is supported by some electronic devices.

Finally, many audio formats have been extended to allow the storage of information about the audio track along with the sound frequencies which the track consists of. This information is stored in *metadata tags*, and examples of information such a tag can contain include the artist who produced it, the album on which it is included, its track number on that album, and the year in which it was released. One common format for including such tags with MP3 audio tracks is ID3, and many digital music players (both software and hardware) use the information in these tags to organize the user's audio tracks instead of the names of the files in which they are stored, or the order in which they appear on the internal storage mechanism.