

Trabalho 1 – Áudio

OBJETIVO

Geral: Aplicar o conhecimento teórico aprendido nas aulas em atividades práticas e, através dos resultados quantitativos e qualitativos obtidos, avaliá-los criticamente.

Específicos:

- Conhecer diversos softwares usados na área de multimídia, em especial no campo do Áudio;
- Comparar diferentes formatos de compressão lossless (sem perdas) e lossy (com perdas) aplicados a áudio;
- Julgar a qualidade subjetiva de diferentes formatos de áudio lossy;
- Avaliar os formatos de áudio mais adequados para as mais diversas situações.

GRUPOS

O trabalho deverá ser feito em dupla, podendo haver no máximo um trio (caso “sobre” um aluno(a) sem uma dupla). A escolha é livre. Por motivos didáticos, esse trabalho **não poderá ser feito individualmente**.

DESENVOLVIMENTO

O desenvolvimento do trabalho deverá seguir os princípios básicos do Método Científico e consistirá, majoritariamente, em vários testes envolvendo **Áudio** em Multimídia. O resultado será exclusivamente um relatório a ser entregue em formato PDF através do local correspondente na sala do AVA.

Atenção: Todos os testes que envolvem áudio devem ser feitos com um **fone de ouvido adequado** (conforme já orientado previamente). O ideal seria que ambos os membros da dupla usassem o mesmo fone, mas em virtude da situação de isolamento social, esse requisito não poderá ser atingido. Porém, esteja atento a esse fato: fones de ouvidos muito diferentes podem levar a conclusões diferentes. Tente usar fones próximos em termos de qualidade.

Em resumo, você fará, para cada atividade, o seguinte:

- **Execução dos testes** solicitados e anotar as **observações quantitativas e qualitativas**.
- **Análise dos resultados** orientando-se, em parte, pelas informações dadas da própria atividade. No entanto, não se limite a elas. Você é livre para fazer outras observações/questionamentos/análises, sem fugir do assunto principal da atividade, é claro.
- Diante da sua análise, talvez você acabe com uma teoria em mãos. Se sim, pense numa forma de confirmá-la com novos testes. Ou talvez você simplesmente esteja “vendo”, na prática, um ou mais aspectos abordados nas aulas teóricas (neste caso, não se esqueça de tomar nota disso e escrever no relatório, já que isso contribuirá com o aumento da sua nota.). No método científico, suas afirmações precisam ser comprovadas experimentalmente ou referenciadas na teoria já consolidada. Caso contrário, são apenas hipóteses (que também são importantes, mas tem menos força que uma afirmação feita com respaldo teórico/experimental).

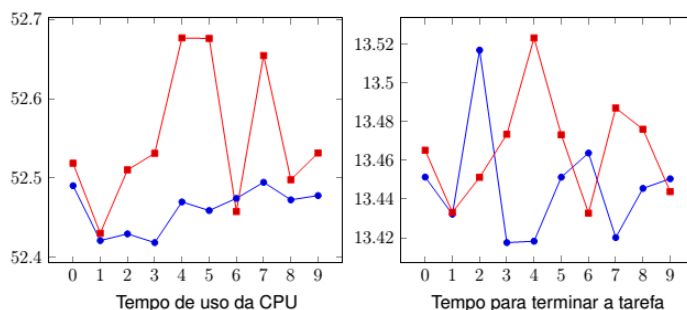
Está em dúvida sobre algum assunto? Pesquise. Google está aí para isso. Converse com seus colegas. Eles também têm olhos/ouvidos e talvez perceberam algo que você não foi capaz naquele momento.

Por fim, algo **MUITO IMPORTANTE**: **Jamais divida as atividades do trabalho onde cada um “faz a sua parte”, em especial as que envolvem ouvir arquivos de áudio**. A maioria das atividades aqui propostas envolve processos mentais que são inerentemente subjetivos e por isso todos devem fazer. E, após isso, escrever as análises em conjunto ou, se divergirem (como de fato vai acontecer em alguns casos), ambas devem estar descritas no relatório.

CONSTRUÇÃO DO RELATÓRIO

Deixo, a seguir, algumas normas e conselhos que recomendo que vocês sigam ao montarem o relatório:

- Siga o máximo que puder **as normas da ABNT**, como devem ser praticamente todos os relatórios redigidos num curso superior. Mantenha a coerência nas fontes, gráficos e tabelas quanto ao tamanho, cor, tipo, etc. É simples: Não torne seu trabalho uma colcha de retalhos! **Olha só que coisa feia que é isso....**
- **Siga sempre norteado pelas atividades.** Separe cada uma delas num “capítulo” a parte, evitando misturá-los. Não troque a ordem delas, pois podem existir co-dependências.
- Seja criativo, porém claro e objetivo ao mostrar os resultados. Mesmo que opte por usar gráficos (o que pode ser bom!), **SEMPRE COLOQUE UMA TABELA com os valores obtidos nos testes.** Ou seja, gráficos são apenas uma ferramenta de análise dos dados, eles não substituem a boa e tradicional tabela com os dados originais. Evite também múltiplas tabelas quando todos os resultados dos testes puderem ser colocados numa única tabela. Isso facilita enormemente a análise dos dados.
- **Adote legendas** para cada gráfico/imagem que você inserir. Fica mais fácil quando, no texto, você fizer referência a elas. Além de facilitar a compreensão de quem está lendo (eu!).
- **Observe a escala ao criar gráficos.** Um aluno, num certo trabalho, colocou os seguintes gráficos:



Olhando rapidamente, dá a entender que houve uma variação de uso de CPU e de tempo muito significativa entre os testes. Mas um olhar um pouco mais atento e você vai observar que a escala (eixo Y) do primeiro gráfico vai de 52,4 a 52,7 (portanto, apenas 0,57% de variação). O segundo gráfico é análogo.

Portanto, não faça como algumas reportagens tendenciosas que mudam a escala para parecer que houve uma variação muito maior que a real. Seja justo(a).

- Sempre que convir, **calcule e exiba a diferença percentual entre os dados.** Nos casos envolvendo compressão de dados (com ou sem perdas), siga a convenção de **adotar o 100% como sendo a referência** (um áudio sem compressão, por exemplo) e o arquivo comprimido (que será, naturalmente, menor) como sendo $TAM_COMPRIMIDO/TAM_ORIGINAL$. Por exemplo, se a música original tinha 50 MB e a comprimida tinha 30 MB, o percentual será 30/50, ou seja, temos um arquivo comprimido que corresponde a 60% do original (ou, em outras palavras, houve uma redução de 40% no tamanho). Não altere a referência (em relação ao que você está comparando), a menos que seja necessário.
- **Ao usar tabelas contendo valores numéricos, evite misturar unidades diferentes numa mesma coluna.** Ou seja, normalize os valores para uma única unidade a fim de facilitar comparações. Ex: Se há valores em MB e KB, normalize para conter apenas MB ou apenas KB (a depender da magnitude dos valores).
- **Use a quantidade de casas decimais adequadas.** Em resultados de testes ou cálculos de porcentagem em tabelas, evite usar muitas casas decimais onde não há necessidade. Avalie a magnitude dos números. Na maioria das vezes, em números com parte inteira diferente de zero, bastam duas ou três casas decimais (as vezes, até, nenhuma). Fazer isso deixa as tabelas mais “leves” de se analisar.
- **Pense duas vezes antes de colar um print da janela dos programas.** Sim, só use imagens que sejam úteis para ilustrar o que o grupo quer dizer. Não coloque prints simplesmente para mostrar o resultado numérico dado pelo programa. Isso polui o texto e dificulta a análise. *Além de eu achar que vocês estão com preguiça!*
- **Evite começar muito bem e perder o fôlego, terminando mal o trabalho.** Isso mesmo. Frequentemente encontro uma série de trabalhos que começam muito bem, fazendo análises corretas e, ao que parece, acaba faltando tempo e o grupo limita-se a apresentar os resultados dos testes com pouca ou nenhuma interpretação dos resultados. Isso é ruim, demonstra mal planejamento do grupo em relação ao tempo e um professor mais experiente consegue perceber isso facilmente.

PONTUAÇÃO

O referido relatório será avaliado de 0 a 100 pontos e contribuirá com 20% da nota semestral. Cada atividade será corrigida conforme o valor atribuído a ela.

ENTREGA

A data limite para entrega do trabalho, via campo específico no AVA, é o dia 22/07/2020 (**quarta-feira**) até as **23:59**.

O trabalho deverá ser entregue em formato PDF cuja nomenclatura do arquivo deverá, OBRIGATORIAMENTE, ser

[FSM 2020-1] Trabalho 1 (Áudio) - Aluno1, Aluno2.

Siga estritamente as regras acima, ou seu trabalho poderá ser desclassificado. Facilite a vida do professor!

----- ### -----

ATIVIDADES

Atividade 1: Entropia na prática

Teoria relacionada à atividade: Teoria da Informação / Entropia / Compressão de dados

Pontuação: 12 pontos

Elaborar um programa de computador escrito em C (devido ao maior controle sobre a manipulação de binários) que gere um arquivo binário com exatos 10 MB, formado por números inteiros não negativos de 8 bits (faça as contas a fim de determinar quantos números são necessários para gerar esse tamanho de arquivo) no intervalo 0 – 255 (ou seja, **unsigned char**, em C). Gere o arquivo de três formas distintas:

1. Modo completamente aleatório. Use algum mecanismo de geração de números aleatórios (oi, função rand...) e preencha o arquivo totalmente com números aleatórios dentro de toda a faixa inteira não negativa de 8 bits (0 – 255).
2. Modo aleatório restrito: Faça praticamente o mesmo que o acima, porém restrinja a faixa numérica para apenas 10% (aproximadamente) da faixa de valores representáveis por um **int8** (inteiro de 8 bits). Ou seja, a faixa numérica que antes variava uniformemente de 0 – 255, agora irá variar, digamos, de 0 – 25.
3. Modo não aleatório: Use qualquer estratégia que você quiser para gerar números que não sejam aleatórios e também não sejam igualmente distribuídos (em outras palavras, alguns valores devem aparecer bem mais do que outros, ou seja, **eles não devem ser equiprováveis**). Mescle cadeias de caracteres repetidos com outras variações. Seja criativo!

Observação importante: Certifique-se de “*printar*” esses números no arquivo em **modo binário**. Ou seja, não trabalhe com o arquivo em modo texto. Caso contrário, você comprometerá toda a questão. (Em C, você pode usar o `fwrite`).

Depois de gerados esses três arquivos, use qualquer compressor de dados genérico (Winrar¹/7zip/etc.) e comprima cada um dos arquivos. Monte uma tabela com os percentuais de compressão (*ratio* = tamanho comprimido / tamanho original).

Agora, a luz da *Teoria da Informação*, **disserte** sobre os resultados. Alguns pontos em que você deve dar atenção são: Como foi a compressão de um arquivo completamente aleatório? Ao restringir o intervalo de variação, os resultados mudaram? Por quê? E no caso do arquivo não aleatório, a compressão “deu certo”?

Desafio bônus 1 (Nível: SSJ): Use mais de um software de compressão e avalie se os resultados são equiparáveis entre os diferentes algoritmos usados por esses programas (ou seja, pode-se, ou não, observar comportamento semelhante, independente do algoritmo de compressão?).

Desafio bônus 2 (Nível: SSJ2): Estime (calcule) a entropia de cada um dos arquivos e faça um paralelo com os resultados obtidos.

Atividade 2: Áudio Binaural

Teoria relacionada à Atividade: Múltiplos Canais de Áudio

Pontuação: 5 pontos

¹ Baixe a última versão do Winrar em www.rarlab.com/download.htm

Binaural - literalmente significa "possuindo ou sendo relacionado às duas orelhas". A audição binaural, juntamente com a filtragem de frequências, permite aos animais determinar a direção da origem dos sons. É uma técnica de gravação e reprodução sonora bastante interessante, pois, com apenas dois microfones, é possível criar o efeito de som ambiente. Para a gravação são colocados dois microfones acoplados à cabeça de um manequim. Os microfones devem ser colocados na posição das orelhas. A banda Pink Floyd tem um álbum gravado com essa técnica, *The Final Cut*. E o futuro PS5 adotará um processamento de áudio muito sofisticado com finalidade semelhante.

Um detalhe importante é que o efeito não é perceptível quando reproduzido em alto-falantes. Para percebê-lo claramente é necessário ouvi-lo através de fones de ouvido.

Nesta atividade, ouça cada um dos arquivos presentes dentro da pasta "Áudio Binaural" da seguinte forma:

1. Coloque os fones de ouvido e ajuste adequadamente o volume;
2. Abaix a cabeça e feche os olhos. Isso é importante para evitar a interferência sensorial (dos olhos, por exemplo).
3. Ouça tudo, sem ser interrompido e sem "viajar" (mantenha o foco/concentração).
4. Enquanto você ouve, peça a outra pessoa para te observar e anotar as suas reações físicas. A ideia é observar, com isso, as reações involuntárias do corpo (especialmente as expressões faciais) que se mostram, ao longo do processo evolucionário humano, tão importantes à sobrevivência da espécie.

Pronto. **Agora relate suas experiências!**

Gostou? Quer continuar a experiência? Visite www.binauralairwaves.com

Atividade 3: Avaliando formatos de compressão de áudio sem perdas

Teoria relacionada à Atividade: Compressão de Áudio sem Perdas (*Lossless*)

Pontuação: 20 pontos

3A) **[16 pontos]** Por definição, não faz sentido falar em qualidade de áudio quando se trata de compressão sem perdas. No entanto, algo muito interessante acontece quando se usa o mesmo algoritmo lossless em diferentes gêneros musicais. Para verificar, na pasta "Músicas sem Compressão", você encontrará os seguintes arquivos (formato WAVE Stereo (sem compressão)):

- Avicii - Wake Me Up.wav
- Bee Gees - Words.wav
- Bruce Springsteen - Nebraska.wav
- Coldplay - Midnight.wav
- Coldplay - Paradise.wav
- Coldplay - Princess Of China [Feat. Rihanna].wav
- Coldplay - Sunrise.wav
- Death Cab For Cutie - I Will Follow You Into The Dark.wav
- Eagles - Hotel California.wav
- Ed Sheeran - I See Fire.wav
- Embrace - Gravity.wav
- Goo Goo Dolls - We'll Be Here (When You're Gone).wav
- Lasgo - Intro.wav
- Maroon 5 - Unkiss Me.wav
- Oasis - Supersonic.wav
- Puddle Of Mudd - Blurry.wav
- Shout Out Louds - Illusions.wav
- The Album Leaf - Window.wav

Você precisará também dos arquivos da pasta "Ruído Branco", que são:

- Ruído Branco (ATENÇÃO, audio mono - 1 canal apenas).wav
- Ruído Branco (ATENÇÃO, audio stereo - 2 canais).wav

Antes de fazer qualquer tipo de compressão, ouça todas as músicas, sem muito compromisso com prestar atenção a cada detalhe. Tome nota de duas músicas que você considera de mais baixa complexidade e de outras duas que você considera de mais alta complexidade.

Agora, você deve avaliar a taxa de compressão (*ratio* = tamanho comprimido / tamanho original) de todos esses arquivos comprimindo-os² nos seguintes formatos: **FLAC -8**, **Monkey's Audio (Extra High)** e **Apple ALAC**. Para efeito de comparação, adicionalmente, comprima cada arquivo original (WAV) no formato **RAR** (normal) com o Winrar x64 que pode ser baixada diretamente do site do autor (www.rarlab.com). Elabore uma tabela com os resultados obtidos.

Em sua primeira análise, e sendo muito honesto, responda: dentre as duas músicas que você acreditou ter **menor complexidade**, alguma delas teve o menor bitrate da tabela (ou seja, você estava certo)? E para o caso de **maior complexidade**?

De posse dessa tabela, faça agora uma análise mais completa. **Ouça atentamente cada um dos arquivos a fim de conhecer as características musicais** de cada um deles (Ah... e aprecie as músicas, Giraldele gosta!). Apesar da “trilha sonora” estar *levemente* polarizada de acordo com o gosto musical do professor, as músicas não foram escolhidas ao acaso (sim, dá MUITO trabalho encontrar as músicas adequadas). Ao ouvir, procure observar o espectrograma e o analisador de espectro. Eles são capazes de mostrar com muita clareza qual a intensidade com que as frequências se apresentam ao longo da música, o quão amplo é o espectro (ou seja, qual é a predominância dos sons de alta, média ou baixa frequência), os trechos onde o som tem características **estacionárias** (variações mais suaves, como a voz e instrumentos não percussivos) ou **transientes** (variações mais bruscas, típicas de instrumentos percussivos, como a bateria e determinados *samples* eletrônicos). Transientes geram uma reta vertical no espectrograma, manifestando quase todas as frequências bruscamente e por um breve período. Quando isso se repete, o espectrograma fica parecendo um pente (de pentear cabelo).

Depois de ouvir, observe:

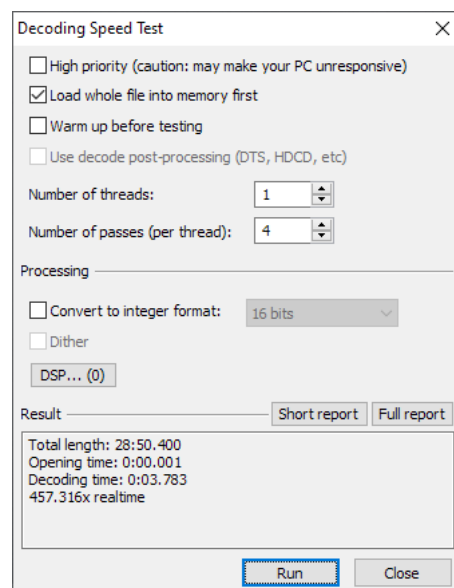
- O *bitrate* (ou, talvez melhor seria avaliar o *ratio*, que é normalizado) de cada um dos arquivos comprimidos em cada um dos formatos lossless (incluindo o RAR). Justifique os valores obtidos de acordo com a teoria estudada, gênero/estilo de cada música e as orientações dadas acima. Analise a eficiência dos formatos lossless projetados para arquivos de áudio frente ao RAR, que é genérico.
- E esse Ruído Branco? Que resultados foram esses? Bug? DISSERTE. Ouviu bem? **DISSERTE** sobre esse fato que já era previsto, se você entendeu bem a teoria estudada.

3B) [4 pontos] Em trabalhos anteriores, analisaríamos nessa subatividade as velocidades de compressão e descompressão de cada algoritmo de compressão de áudio lossless. No entanto, os computadores modernos estão tão rápidos e dotados de vários núcleos (o que permite várias threads de compressão simultâneas) que faz pouco sentido discutir a velocidade de **compressão**. É fato que alguns algoritmos (como o Monkey's Audio) são consideravelmente mais pesados que outros (como o FLAC). Mas a tarefa de compressão é feita uma única vez na grande maioria dos casos.

Já a velocidade de **descompressão** é outra história! Em seu uso corriqueiro, o áudio precisará ser descomprimido todas as vezes que se deseja ouvi-lo. E, mesmo que os computadores atuais consigam fazer isso com absoluta tranquilidade, outro fator tem maior importância: algoritmos de descompressão mais complexos (mais “pesados”) demandarão mais da CPU, o que certamente impacta bastante no consumo energético (bateria) de dispositivos móveis, como smartphones³.

Assim sendo, vamos tomar uma das músicas e efetuar a medição, seguindo os passos abaixo:

1. Abra o foobar2000, arraste a música “Eagles - Hotel California” em WAV, FLAC, Monkey's Audio e ALAC para a janela principal. O motivo de escolhermos essa música é simplesmente porque é a de maior duração. Todas as demais apresentarão resultado semelhante já que foram comprimidas com os mesmos parâmetros.



² Para convertê-las, use Foobar2000, clicando com o botão direito do mouse sobre a(s) música(s) e selecionando Convert. Você pode escolher o preset adequado na lista de presets que eu já criei previamente. Mas, a princípio, recomendo que clique na última opção (com os 3 pontinhos) para entender o que está acontecendo. Há os mesmos presets do lado esquerdo. Duplo-clique neles ou clique + Load carrega as configurações do preset que podem ser conferidas na parte direita da janela. Explore um pouco para entender as opções. Não é nada de outro mundo, mas se você precisar alterar algum parâmetro para algum outro teste, estará familiarizado. Cuidado para não sobrescrever algum preset acidentalmente e prejudicar seu trabalho.

³ Em dispositivos portáteis específicos para música (ao estilo “MP3 Player”), geralmente a decodificação é feita via hardware, o que impacta menos no tempo de bateria. Mas, mesmo nesses casos, a complexidade do decoder continua relevante. Porém, em menor grau.

2. Antes de fazer cada teste, certifique-se de fechar todos os programas que puder, a fim de não influenciar no resultado.
3. Para cada arquivo, clique com o botão direito sobre ele e abra *Utilities > Decoding Speed Test*. Ajuste os parâmetros segundo a janela acima.
4. Para cada teste, anote o valor exibido na última linha de *Result*, um valor no formato “número x realtime”. Por exemplo, num determinado teste obtive *457.316x realtime*⁴. Isso quer dizer que o sistema em questão pode decodificar o arquivo de áudio aproximadamente 457 vezes mais rápido que o tempo real (ou seja, caso ele fosse tocado “normalmente”). Quanto mais rápida for a decodificação, menor é o esforço computacional envolvido no processo.

Visto que diferentes formatos têm diferentes complexidades, será natural e esperado que alguns sejam mais “leves” do que outros. Seu objetivo é exatamente mensurar isso. De posse desses valores, **discuta** os resultados.

Atividade 4: Qualidade de Áudio *Lossy*

Teoria relacionada à Atividade: Compressão de Áudio com Perdas (*Lossy*)

Pontuação: 46 pontos

4A) **[10 pontos]** O primeiro objetivo dessa atividade é bem simples. Você deverá comprimir todas as músicas presentes na pasta “Músicas sem Compressão” para dois formatos de áudio *lossy* muito populares: **MP3** (Lame -V2) e **AAC** (Apple VBR Q 91). Monte uma tabela com os bitrates médios de cada música em cada um dos formatos.

Agora, ouça cada uma das músicas em MP3, observando o **espectrograma** (se vc der duplo-clique ele vai para tela cheia), o **analisador de espectro** (as barrinhas) a **variação do bitrate** (em texto, na barra inferior esquerda) e o **que você está ouvindo**, é claro! A variação dinâmica (o “volume”, essencialmente) e dada pelas duas barras verticais da esquerda inferior (lembre-se: 0 dB é a referência de volume máximo em áudio). Tome nota das observações individuais feitas para cada música. O motivo dessa atividade é ajudar a treinar os seus ouvidos um pouco. O professor não é músico, só está usando músicas para te mostrar algumas características importantes no campo do áudio. Se quiserem descontrair um pouco e fazerem comentários sobre a música (sem perder o foco da análise científica) também pode. Humanos gostam de música desde sempre e isso tem forte influência neurológica (além de cultural, claro). Abaixo eu faço algumas observações muito úteis, que justificam as minhas escolhas:

- **Avicii - Wake Me Up:** Música eletrônica bem animada para começar! Cheia de estímulos sonoros. Quero que observem algo interessante a partir do instante 1:10 até 1:57. Um instrumento eletrônico faz uma espécie de “refrão” bem repetitivo. E outros sons vão entrando e preenchendo cada vez mais o espectro. No entanto, você continua ouvindo separada e perfeitamente o instrumento a que me refiro. A pergunta é: Por quê? Teorize, sem medo de errar.
- **Bee Gees - Words:** Impressionado com o bitrate dessa música? Por que ele é assim? (Dica: repare a correlação existente entre os dois canais de áudio e a época que essa música foi gravada. Além disso, os instrumentos usados). Consegue ouvir o ruído de fundo da gravação? Dá pra “vê-lo” no espectrograma poluindo quase todo o espectro (na versão sem perdas é mais fácil ver).
- **Bruce Springsteen - Nebraska:** Essa faixa é muito interessante por uma razão bem simples: ela praticamente não tem transientes (não há instrumentos percussivos, que normalmente são os que marcam a música com transientes). Ou seja, é toda muito “bem comportada”, características estacionárias do começo ao fim. Uma gaita desenha um espectrograma lindo e cheio, muito cheio de frequências bem marcadas e com todos os seus harmônicos.
- **Coldplay - Midnight:** Essa daqui está aqui porque... você já sabe, não é? Mas não é só por isso não. Ela é rica em uma base feita por um baixo muito bonito a partir de 1:20 (se você não estiver ouvindo o “peso” desse baixo, seus fones definitivamente não são bons, simples assim! Dentre todas as músicas, essa é com a maior e mais clara presença de baixos e sub-baixos). A forte presença de frequências baixas é simples de ser comprovada pelo vermelho bem evidente no espectrograma. Olhando pro analisador de espectro, nas primeiras barras da esquerda, dá para ver claramente quando o baixista muda as notas musicais. Esse “peso” é praticamente todo perdido quando você ouve suas músicas em dispositivos de reprodução (fones ou caixas acústicas) de baixa qualidade. Arrisco-me a afirmar que muita gente sequer sabe o quanto de riqueza musical está perdendo. No momento 2:00 esses sub-baixos saem por completo da música por algum tempo e, é como se toda a música de repente ficasse mais “leve” (basta observar o espectrograma). E a voz do Chris Martin foi alterada usando um sintetizador, deixando-a semi-robótica (Se arrisca a dizer por que nós conseguimos perceber isso?).

⁴ Atenção: o ‘ponto’ aqui é o separador de decimal (a nossa vírgula)

- **Coldplay - Paradise:** Essa você deve conhecer, foi um *single* muito popular do álbum *Mylo Xyloto*. Quero que você observe um fenômeno interessante: repare o bitrate médio do trecho 0:00 a 0:34 e a complexidade da música via espectrograma, agora, no trecho imediatamente seguinte, a drástica mudança no espectrograma. E o bitrate, como ficou no instante da transição e nos segundos seguintes? Você esperava algo diferente? Tente explicar.
- **Coldplay - Princess Of China [Feat. Rihanna]:** Essa aqui você pode, nesse primeiro momento, só curtir. ☺ Depois, tenho um desafio bem simples para ela, juntamente com a música do Maroon 5.
- **Coldplay - Sunrise:** Essa é, na verdade, um prelúdio do álbum *Everyday Life* (2019). Quase toda feita com um trio de cordas (baixo, cello e violino). O espectrograma é interessantíssimo na parte do violino (a partir de 1:28). Quero que você tente descrever o que você está vendo e ouvindo, do ponto de vista das frequências e do “traçado” que elas fazem no espectrograma. Ora as linhas ficam mais ou menos retas, ora elas oscilam de modo quase perfeito. A troca das notas/acordes é facilmente vista (e ouvida).
- **Death Cab For Cutie - I Will Follow You Into The Dark:** Música calma, levemente percussionada e com bitrate baixo. Faça apenas as observações básicas solicitadas.
- **Eagles - Hotel California:** Essa é uma obra de arte. Quero que você observe a riqueza de cada instrumento e a espacialidade da gravação. Cada ouvido seu ouvirá sons diferentes. Stereo bem definido e marcante. Você, mais a frente, usará um editor de áudio para separar apenas o trecho de aplausos do final da música a fim de efetuar testes de transparência.
- **Ed Sheeran - I See Fire:** O motivo dessa faixa estar aqui é que vocês percebiam o quão rica, espectralmente, pode ser a voz humana quando gravada adequadamente. Repare que a zona de maior energia se concentra em frequências mais baixas conforme estudado. Porém, existem muitas componentes frequenciais com significativa energia na parte alta do espectro (altas frequências)⁵. Sim, a voz humana é muito rica. E a do Ed é maneira, não? Compare os bitrates das regiões que contém só a voz dele, só instrumentos, e ambos.
- **Embrace - Gravity**⁶: Essa música é um verdadeiro “achado” para testes envolvendo codecs de áudio *lossy*. A extrema importância dessa faixa se mostrará evidente nos testes de transparência, logo a seguir. Eu prometo! A princípio, apenas ouça com atenção. E, talvez, se emocione, né?⁷.
- **Goo Goo Dolls - We'll Be Here (When You're Gone):** Essa faixa do Goo Goo Dolls é bem complexa em termos frequenciais. Repare o bitrate e sua variação ao longo de toda a faixa e você entenderá.
- **Lasgo - Intro:** Já nos primeiros segundos dessa música, simplesmente conferindo o espectrograma que é desenhado, é possível entender por que essa faixa está aqui. Melodia construída com sintetizadores digitais, numa vibe anos 80, promove uma análise muito rica do ponto de vista de complexidade do sinal, do espectrograma e do bitrate necessário para manter uma boa qualidade aproximadamente transparente. Quase todo o espectro é preenchido, até os extremos agudos estão aí.
- **Maroon 5 - Unkiss Me:** Essa música é outro achado com uma característica muito interessante. Repare o bitrate logo nos primeiros 7 segundos! E o que acontece depois? Por quê? Consegue explicar?
- **Oasis - Supersonic:** Faixa com espectro bem “sujo” e irregular. Muitos instrumentos distorcidos. E com forte uso de compressores⁸ (repare que o volume é alto o tempo todo, praticamente. Há pouca variação de volume). Poderá ser útil em testes de transparência e artefatos mais a frente.
- **Puddle Of Mudd - Blurry:** Essa faixa é bem marcada pela bateria e um vocal evidente. O refrão é muito complexo frequencialmente, mas o bitrate não vai para a estratosfera. Pelo contrário, mantém-se normalmente abaixo do bitrate médio. Por que será? E sim, também faz forte uso de compressores.
- **Shout Out Louds - Illusions:** Ouçam um pouco de música Indie e as suas experimentações que sempre acontecem. Música “complexa” em termos instrumentais, mas tanto o bitrate médio como as variações, são bem comportados.
- **The Album Leaf - Window:** Essa é uma banda de post-rock. Além do professor curtir muito o som deles, ele quer que você observe o espectrograma a partir instante 0:16 em diante. Para cada acorde tocado (conjunto de 3 notas), o espectrograma e o analisador de espectro ficam tão bem marcados que é impossível não notar as notas que compõem o acorde “subindo” praticamente ao mesmo tempo e com quase a mesma intensidade. Notas são frequências, coloque isso na cabeça, aqui mostrado de um jeito bem nítido. A partir de 2:00, baixos e sub-baixos aparecem com bastante intensidade.

⁵ Um teste interessante e simples de ser feito é usar um filtro passa-baixa conforme comentamos nas aulas, a fim de cortar as frequências mais altas (No Foobar2000 em Preferences > Playback > DSP Manager ative o DSP IIR Filter, clique nos ... e adicione o filtro lowpass, ajustando a frequência de corte). A voz continuará bem inteligível, mas perderá o seu “brilho”. É bem bacana.

⁶ Curiosidade: A banda inglesa Embrace gravou essa música Gravity em 2004. Mas, advinha quem compôs essa música para eles? Há... e os compositores originais também gravaram essa música como B-Side. E, na interpretação deles, eu considero a música mais depressivamente linda do mundo. ☺

⁷ Curiosidade: Essa música é composição do Coldplay e foi cedida para a banda britânica Embrace, que gravou sua própria versão. A versão do Coldplay você só deve ouvir se tiver emocionalmente bem, hehehe. É, provavelmente, a música mais triste que já ouvi na vida. Mas, belíssima!

⁸ Compressores musicais, não de arquivo, ok? Atenção ao contexto em que essa palavra é usada.

Desafio bônus 1 (Nível: kaioken 1x): use o mp3plot, fornecido na pasta de softwares para gerar um gráfico de distribuição do bitrate para o caso dos arquivos em MP3. Use esse gráfico na sua análise. As barras verticais indicam a distribuição e a barra horizontal a distribuição do bitrate ao longo do tempo. (Eu incluí um .bat com os ajustes. Basta alterar o nome do arquivo internamente e executar o .bat)

Desafio bônus 2 (Nível: SSJ): Repare o espectrograma dos primeiros 10 segundos do começo de **Coldplay - Princess Of China [Feat. Rihanna]** e os primeiros 7 segundos de **Maroon 5 - Unkiss Me**. As características espectrais de ambas parecem semelhantes, não? Ambas parecem simplesmente um conjunto de ruídos. Mas, se sim, por que os bitrates desses trechos são tão diferentes?

4B) [15 pontos] O próximo objetivo é simples de ser entendido, mas um pouco difícil de ser executado por um iniciante destreinado. Volto a dizer: é difícil, mas você logo pega o jeito. Outros já conseguiram, logo, você também pode conseguir! Você deverá encontrar o ajuste mínimo a partir do qual o encoder gerará um áudio que é transparente para você. E você fará isso para os formatos MP3 e AAC, ambos em VBR (Variable Bitrate). Em outras palavras, até que ponto você pode comprimir (abaixar o bitrate) sem perder a transparência (para você, claro). **Não há forma melhor de realmente assimilar esse assunto se não for assim, praticando!**

O LAME MP3 simplificou muito seus ajustes VBR. É basicamente $-Vx$, onde $x = 0$ até 9, em ordem **decrecente** de qualidade, conforme visto na teoria (número menor = maior qualidade). O ajuste médio recomendado para transparência é de $-V2$. Mas, muito provavelmente, você já não conseguirá perceber diferença mesmo para um ajuste acima disso ($-V4$ ou até mais). Atenção para a escala: a “pior qualidade”/menor arquivo é o $-V9$ e a “melhor qualidade”/maior arquivo é o $-V0$.

Já o Apple AAC, no modo VBR, usa um parâmetro chamado `--tvbr x`, onde x pode ser um dos seguintes valores: 0, 9, 18, 27, 36, 45, 54, 64, 73, 82, 91, 100, 109, 118, 127. A ordem é **crescente** de qualidade (número maior = maior qualidade). O ajuste inicial recomendado é de `--tvbr 64` (que, para o padrão CD, dá um bitrate aproximado de 128 kbps)⁹. Comece com ele, mas faça o contrário do MP3, ou seja, vá diminuindo o valor a fim de aumentar o grau de compressão (diminuir o tamanho do arquivo, descartando mais informações, claro).¹⁰ Confuso? Anteriormente era mais complicado usar o encoder da Apple fora do iTunes. Mas, facilitou muito nas versões mais novas do Foobar2000, já que ele apresenta uma janela com as configurações (ao invés de ser necessário usar comandos de texto). Basta, no conversor, optar pelo AAC (Apple), setar o modo do bitrate e o valor do parâmetro (que, na barra Quality é expresso como bitrate médio ou parâmetro Q). Os modos suportados são:

- **Variable Bit Rate (VBR):** que é conhecido, internamente, como `tvbr (true vbr)`, observe o parâmetro Q com os valores mencionados acima quando você move o ajuste de qualidade.
- **Constrained VBR e Average Bit Rate (ABR):** basicamente um modo de bitrate médio. Não me pergunte a diferença, nesse encoder, de CVBR para ABR, pois eu provavelmente teria muita dificuldade em explicar. Numa resposta simples, CVBR fará o mesmo que ABR, ou seja, permitirá que o bitrate varie, porém mantendo o valor médio no arquivo. A diferença refere-se a um fator técnico um tanto complexo chamado *bit-reservoir*, que no modo CVBR pode ser muito maior que no modo ABR.¹¹
- **Constant Bit Rate (CBR):** Bitrate constante. Dispensa maiores explicações.

Mas, como saber se está transparente ou não? Você precisa fazer testes. E para isso você usará a metodologia de **Teste Cego** (Blind Test) chamada ABX. O Foobar2000 tem uma excelente e simples ferramenta para isso. Basta carregar na *playlist* o arquivo original (ou comprimido sem perdas) e o *lossy* a ser testado. Para cada música, selecione o original e o lossy (só segurar ctrl). Depois, clique com o botão direito sobre alguma delas e escolha Utilities > ABX Tracks... Deixe todas as opções desmarcadas, ajuste o parâmetro numérico para 10 tentativas e clique em OK¹². Agora, observe que a tela abaixo aparece. A e B são os arquivos na ordem conhecida. X e Y são aleatórios (ora serão A e B, ora serão B e A, respectivamente). Só há duas alternativas: [X is A, Y is B] ou [Y is A, X is B]. Você pode apertar quaisquer botões de Playback antes de decidir em Choice. Ouça atentamente A, B, X ou Y. Você pode trocar entre quaisquer delas e a qualquer momento. Pode ouvir outros trechos mudando

⁹ Há, aqui, uma pequena confusão (talvez devido a aproximação numérica). O qaac fala que o fator é 63, mas o Foobar2000 informa 64. Para não dar confusão, mantive 64 em tudo. E, mais uma coisa... AAC já é praticamente transparente em bitrate médio de 128 kbps, por isso iniciaremos com esse valor.

¹⁰ Para maiores informações, visite o site do qaac, que é um aplicativo que o foobar2000 usa internamente e que permite o acesso ao Apple AAC Codec mesmo fora do iTunes/Quicktime. Link: <https://github.com/nu774/qaac/wiki/Encoder-configuration>

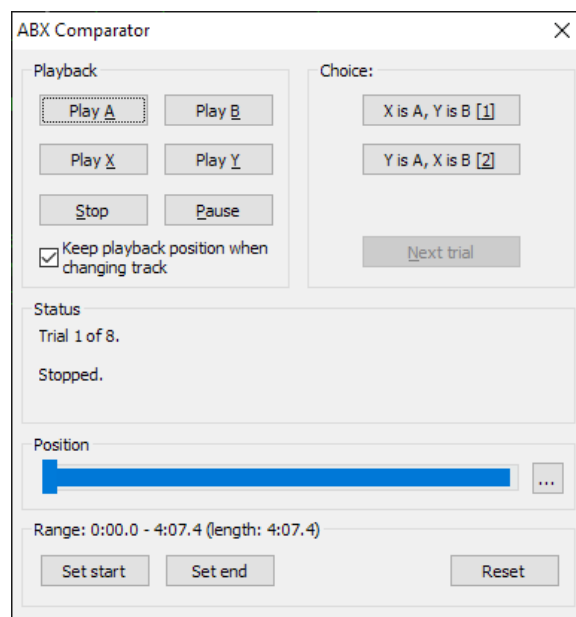
¹¹ Generally, such modes describe the order of magnitude of the time duration over which an encoder produces a constant bit consumption for a certain target bit-rate. Within that duration, the encoder may distribute bits unequally across frame boundaries. The time duration is controlled by the size of the encoder's bit-reservoir.

- *Hard CBR: no bit-reservoir, every frame (21.5-43 ms) has the same bit consumption (e.g. -hard-cbr in Opus)*
- *Soft CBR: small bit-reservoir, for example one with the size given in the MPEG standard (e.g. Fraunhofer CBR)*
- *ABR: large bit-reservoir, maybe up to a few seconds. Entirely up to the developer (e.g. ABR in Nero's encoder?)*
- *CVBR: very large bit-reservoir, maybe up to a minute. Entirely up to the developer (e.g. CVBR in Apple's encoder?)*
- *TVBR: maximum bit-reservoir, every frame and channel may use between ~10 and 6144 bits (e.g. Apple's or Fraunhofer's VBR)*

¹² Se você quiser fazer um treino antes dos testes, marque a opção Training mode.

o slide em position. Já decidi quem é X e Y em relação as referências A e B? Só marcar sua escolha e clicar em Next Trial (Próximo teste). O teste só te informará seus acertos (ou erros) no relatório final.

Após terminar o último teste, ele abrirá um log. Dois valores te interessam muito aí: o seu número de acertos (expressos em “x acertos”/“y tentativas”) e o p-value, que é um parâmetro estatístico de significância do resultado obtido. Para o caso em questão, **quanto menor for o p-value, mais confiável é o resultado obtido** (ou seja, menor é a probabilidade dele ter ocorrido por qualquer outra razão que não seja o fato de você realmente ouvido as diferenças e feito as escolhas corretas¹³). Quanto mais testes são feitos e mais acertos você tem, menor é o p-value (com 8 acertos em 8 testes ele já chega a $1/2^8 = 0,00390625$ ou 0,39%). Assim, se p-value for muito baixo (menor que 1%) e você tiver acertado bastante, significa que o ajuste **não é transparente** para você. Ou seja, há diferenças audíveis entre o *lossy* e o *lossless* e você foi capaz de perceber isso facilmente. Isso indica que você precisará aumentar a qualidade (escolhendo um fator de compressão menos agressivo, com maior bitrate). Caso contrário, ou seja, você não está conseguindo perceber as diferenças (por isso está errando nos testes), você pode diminuir ainda mais a qualidade/bitrate até que as diferenças sejam perceptíveis aos seus ouvidos... Entendeu?



O que fazer então? Brincar com algumas faixas musicais presentes na pasta “Músicas sem Compressão”. Visto serem muitas músicas e em semestres anteriores o professor quase ter sofrido um atentado, resolvi selecionar um subconjunto com apenas 5 faixas, que são as que você usará em todos os testes dessa sub-questão e da próxima. Abaixo existem quatro grupos e uma faixa obrigatória. Você deve escolher **uma faixa de cada grupo** (A, B, C e D) + **a obrigatória** (que é... obrigatória!). Fique a vontade para escolher outras, adicionalmente, se quiser:

- **Grupo A:** Bee Gees - Words / Bruce Springsteen - Nebraska / Coldplay - Sunrise / Death Cab For Cutie - I Will Follow You Into The Dark / Eagles - Hotel California / Ed Sheeran - I See Fire.
- **Grupo B:** Avicii - Wake Me Up / Coldplay - Paradise / Coldplay - Princess Of China [Feat. Rihanna] / Maroon 5 - Unkiss Me / Shout Out Louds – Illusions.
- **Grupo C:** Coldplay - Midnight / Lasgo - Intro / The Album Leaf - Window.
- **Grupo D:** Goo Goo Dolls - We'll Be Here (When You're Gone) / Oasis - Supersonic / Puddle Of Mudd - Blurry.
- **Obrigatória:** Embrace - Gravity (Ela tem alguns pontos que exibem mais facilmente as imperfeições dos codecs (logo após o piano, ponto onde entra a bateria, tem muitas transientes complicadas), principalmente no formato MP3)

Faça testes suficientes com ambos os formatos. Monte uma tabela resumizando os resultados.

Sim, para cada uma das 5 músicas escolhidas, comece com o parâmetro de MP3 -v2 e AAC Q 64, tidos como transparentes para a maioria das pessoas, nos mais diversos tipos de música. **Faça o teste ABX até você atingir muitos acertos e um p-value abaixo de 5%** (o que te permite errar no máximo 1 vez em 10 tentativas)¹⁴. Quando você chegar nesse ponto, quer dizer que as diferenças são bastante perceptíveis. Daí é possível inferir que o seu nível de transparência está no teste anterior. Captou? Quando se está no nível de transparência, os acertos são meros chutes, já que as diferenças são praticamente imperceptíveis pra você.

E a partir disso, **disserte**. Tipo... **MUITO**. Repare as distorções no vocal e principalmente em instrumentos de percussão com espectro mais amplo, como os pratos. O som que o dedos fazem ao se arrastarem sobre as cordas do violão. Notas dos pianos que deveriam aparecer soar limpas, mas não estão. Essa é, provavelmente, a principal atividade desse trabalho! Tanto que ela vale mais. Gaste tempo nela.

IMPORTANTE: Esse teste é muito cansativo. Quando se sentir cansado, pare, saia do computador, tome uma água, um banho, sei lá... mas dê uma pausa.

¹³ Em outras palavras, quer dizer que você pode rejeitar a hipótese de ter sido por chute.

¹⁴ Por exemplo, se usando MP3 -v2 você não conseguiu observar diferenças (o teste vai mostrar isso, conforme já explicado), crie um novo arquivo usando o ajuste -v3 e teste novamente (**CUIDADO:** Use sempre o arquivo original, sem perdas!). Ainda não viu diferenças? Repita... em algum ponto até -v9 as diferenças serão bem nítidas no caso do MP3. No caso do AAC, que o parâmetro desce, você pode chegar até Q 0.

Desafio bônus (Nível: SSJ): Use qualquer editor de áudio¹⁵ e separe o trecho final de **Eagles - Hotel California** que contém os aplausos (e o “Thank you...”). Lembre-se de usar o áudio original, sem perdas, e gravar o trecho isolado em WAV (no editor). No Foobar2000, converta e efetue testes de transparência em MP3 e AAC apenas nesse trecho. Ouça atentamente e disserte sobre o assunto.

4C) **[10 pontos]** O ponto nesta subatividade é diferente: Avaliar a performance do MP3, AAC-LC e HE-AAC e do Opus em bitrate baixo. Para isso, você fará os três conjuntos de testes básicos:

	MP3	AAC-LC	HE-AAC (LC + SBR)	Opus
1	48 kbps (--abr 64)	48 kbps (--cvbr 48)	48 kbps (--he --cvbr 64)	48 kbps (--cvbr 48)
2	64 kbps (--abr 64)	64 kbps (--cvbr 64)	64 kbps (--he --cvbr 64)	64 kbps (--cvbr 64)
3	96 kbps (--abr 96)	96 kbps (--cvbr 96)	96 kbps (--he --cvbr 96)	96 kbps (--cvbr 96)

Os presets já estão todos criados (basta selecionar e apertar Load)! Você pode editá-los para ver como são construídos. Alguns precisam ser feitos manualmente, mas não é coisa de outro mundo. Não se preocupe se o bitrate do arquivo final não for exatamente igual aos acima. Isso é normal em codificação ABR (em caso de dúvidas, pergunte ao professor).

Use as mesmas 5 faixas que você escolheu na atividade 4B e gere os arquivos com os parâmetros acima. Sim, são 12 versões diferentes para cada uma das 5 faixas escolhidas! Antes de julgar, pesquise e entenda o que estão envolvidos nos dois perfis AAC diferentes (LC e LC+SBR). O formato Opus também entra na jogada e você entenderá ao fazer os testes.

Não há necessidade de montar tabela para essa subatividade. O que você tem que fazer é, para cada música, ouvir cada um dos arquivos codificados e **dissertar!** Separe suas observações por música. Tome nota de coisas como: apesar do espectrograma estar possivelmente “muito” alterado, a música soa perceptivelmente muito diferente? Existem artefatos? Se sim, eles incomodam muito? O som se tornou “metalizado” ou “abafado”? A enorme redução no tamanho da música, compensa a possível baixa qualidade? Como o MP3 se saiu frente aos outros formatos? Em quais cenários você usaria essas tecnologias que se saíram melhor em testes com bitrate baixo? Na sua coleção pessoal de músicas? Você vê algum cenário dentro de Sistemas de Informação em que esses codecs seriam adequados?

4D) **[6 pontos]** Agora, faremos algo diferente. Vamos codificar arquivos de áudio contendo apenas voz com os formatos HE-AAC e Opus, ambos em bitrate baixíssimo. O objetivo aqui é encontrar o que provê melhor qualidade nessa categoria (voz em bitrate baixíssimo). Pode-se dizer que ambos os formatos são adequados, mas talvez um possa ser melhor do que outro e você vai descobrir isso experimentalmente. :-) A ideia é que o formato preserve o máximo possível da qualidade da voz e obrigatoriamente mantenha a inteligibilidade¹⁶.

Você recebeu na pasta “Áudio contendo apenas voz” um arquivo produzido pelo professor na qual são lidas sentenças padronizadas de uma lista chamada “Listas de Sentenças em Português” (proposto por Costa MJ, 1998) e são usadas na área de Audiologia. Se tiver curiosidade, coloquei um artigo em PDF na mesma pasta.

O áudio tem 45 segundos e foi gravado usando um microfone de lapela, num arquivo sem compressão (WAV), em modo mono, com taxa de amostragem de 48 KHz e 16 bits por amostra, mais do que o suficiente para voz.

Você deve converter o arquivo original para:

- HE-AAC¹⁷, modo ABR, em 16, 24 e 32 kbps¹⁸.
- Opus, com otimizações para voz (speech), modo CVBR, em 8, 16, 24 e 32 kbps.

Fez? Ouça com atenção! Compare os formatos. Disserte sobre qual formato você escolheria e em qual bitrate já estaria adequado. Justifique sua escolha, proponha um cenário, enfim... **DISSERTE!** Agora, dê um jeito de obter um arquivo original de mensagem de voz do Whatsapp do seu próprio smartphone, vasculhando os arquivos pasta “Media” dentro da pasta real do Whatsapp em qualquer dispositivo Android (iOS não deve permitir isso). Você pode usar o app Airdroid para isso. Depois, carregue o arquivo no foobar2000 e conferir as propriedades dele. E aí? 😊

¹⁵ Recomendo o Audacity (www.audacityteam.org)

¹⁶ Sim, são coisas diferentes: é possível ter uma voz inteligível, porém com baixa qualidade, como alguns radiocomunicadores.

¹⁷ Ao reproduzir o arquivo codificado em HE-AAC, o foobar2000 irá exibi-lo “erroneamente” como se fosse um arquivo stereo. Ignore isso. Não afeta em nada os resultados obtidos. É uma questão chata e complexa de tentar explicar aqui. Se ainda assim quiser saber o motivo, boa aventura: <https://sites.google.com/site/qaacpage/news/aboutmonoheencoding>

¹⁸ Apesar de teoricamente ser possível configurar o encoder para HE-AAC, ABR 8 kbps, na prática ele não se mostrou capaz de gerar um arquivo com bitrate tão baixo. Assim sendo, o menor bitrate para este formato será o de 16 kbps.

4E) [5 pontos] Para finalizar essa atividade, vamos testar as velocidades de descompressão de todos os formatos usados, nos cenários de bitrate normal e bitrate baixo:

- Compressão de músicas em geral, bitrate “normal”: MP3 (LAME -v2); AAC-LC (Apple VBR Q 91)
- Compressão de músicas em geral, bitrate “baixo”: MP3 64 kbps (--abr 64); AAC-LC 64 kbps (--cvbr 64); HE-AAC 64 kbps (--he --cvbr 64); Opus 64 kbps (--cvbr 64).

Para tanto, tome como referência a música “Eagles - Hotel California” em WAV e comprima em cada um dos formatos acima. **Cuidado** para não se perder com os arquivos (dê nomes apropriados). Agora, basta decodificar cada um deles nos mesmos moldes do que foi feito na atividade 3B. Discuta os resultados.

Atividade 5: Transcoding em Áudio

Teoria relacionada à Atividade: Compressão de Áudio com Perdas (Lossy)

Pontuação: 12 pontos

O termo *Transcoding* é bastante usado em Multimídia. Consiste, essencialmente, na conversão digital-para-digital, de um formato para outro. É especialmente útil em dois cenários: quando o dispositivo de destino não suporta determinado formato ou quando se deseja converter para outro formato mais eficiente e/ou com menor bitrate.

A operação de *transcoding* geralmente envolve a decodificação de um formato, seguida da codificação para o outro formato¹⁹. São raros os casos onde a conversão pode ocorrer de forma direta, sem a necessidade de se decodificar primeiro.

Na área de Áudio, que é o que nos interessa no momento, temos basicamente 3 situações, já a quarta possível situação não faz o menor sentido:

- Lossless → Lossless: É uma operação **SEMPRE** segura do ponto de vista da qualidade de áudio. Apesar da mudança na representação, a **mensagem** (a onda sonora) de ambos os arquivos é **matematicamente idêntica**. Possíveis razões para essa mudança: compatibilidade, maior eficiência de um formato mais novo, gosto pessoal, etc.
- Lossless → Lossy: É uma operação que envolve apenas **um nível de perdas**. É uma operação muito comum quando se deseja a conversão para outro formato com perdas exatamente para se conseguir diminuir ainda mais o bitrate (por razões de armazenamento ou taxa de transmissão pela internet), ou porque o dispositivo de destino não aceita o formato lossless. FLAC → MP3 ou ALAC → AAC são possíveis exemplos.
- Lossy → Lossy: É uma operação que deve ser evitada sempre que possível, pois envolve **dois níveis de perdas**. Isso porque o arquivo origem já veio de algum outro formato e sofreu perdas em sua geração. A conversão para outro formato com perdas vai implicar em novas perdas/distorções. Considere que sempre haverá perdas e/ou distorções no sinal, mesmo que aparentemente sejam imperceptíveis. E isso **vale até para situações onde os formatos são iguais**. Mais ainda, **até quando tem o mesmo bitrate** (já que, de um bitrate menor para um maior não faz sentido!). Evidentemente, quando, além disso, há diminuição do bitrate, as perdas se acentuam. É algo análogo ao que acontece quando se tira uma fotocópia (Xerox) de uma fotocópia.
- Lossy → Lossless: **Não faz absolutamente NENHUM sentido essa conversão**. As perdas envolvidas no processo lossy são irreversíveis e a mudança para um formato lossless simplesmente aumentará o tamanho do arquivo, já que ele terá exatamente a mesma qualidade do arquivo que deu origem a ele.

Dito tudo isso, chegou a vez de você avaliar, qualitativamente, e perceber que o professor realmente tem razão.

O procedimento, apesar de meio chato, é simples e deve ser aplicado as mesmas 5 músicas selecionadas na atividade anterior:

- Carregue as músicas originais (wav) no Foobar2000.
- Converta-as para MP3 -V2. Renomeie a pasta com as faixas codificadas para “Conversão 1”.
- Limpe a *Playlist* e carregue agora todas as músicas novamente, mas use as que já foram convertidas para MP3, presentes na “Conversão 1”. Converta-as novamente para MP3 -V2. Observe que o Foobar2000 te avisará dessa operação problemática. Aceite assim mesmo. Renomeie a pasta resultante para “Conversão 2”.
- Repita os procedimentos acima pelo menos até “Conversão 6”.
- Agora ouça, em ordem crescente de conversões, o que acontece com a qualidade. **Disserte** sobre o assunto. Para se guiar, você pode usar algumas das dicas/questionamentos que dei na questão anterior para nortear sua análise.

¹⁹ Em áudio, seria algo como FLAC → WAV → MP3, por exemplo. O formato intermediário é um tipo sem compressão.

Desafio bônus 1 (Nível: SSJ): Será que o mesmo acontece com o formato AAC? Será que, neste aspecto, ele se comporta pior ou melhor que o MP3? Só testando para saber, não é? 😊

Desafio bônus 2 (Nível: SSJ): Especificamente no caso do MP3, e se a conversão for feita com bitrates de entrada diferentes dos de saída (pra mais ou pra menos)? O que será que acontece? (Já vi gente convertendo um MP3 de 128 para 320 kbps para “aumentar” (sic) a qualidade. Nesse dia eu quase descreditei na humanidade...)

Desafio bônus 3 (Nível: SSJ God): Aponte razões teóricas (baseadas no funcionamento interno do MP3) que justifiquem a significativa perda de qualidade em *transcoding* sucessivos, mesmo usando exatamente os mesmos parâmetros de compressão.

Atividade 6: Descobrindo se o arquivo de áudio é um *Lossless* de verdade.

Teoria relacionada à Atividade: Compressão de Áudio sem Perdas (*Lossless*) e Compressão de Áudio com Perdas (*Lossy*)

Pontuação: 5 pontos

Ter em mãos um arquivo FLAC ou mesmo um WAV não é garantia de que você possui um arquivo que representa fielmente o áudio original. Como assim? Simples... Esse “fake FLAC” pode ser na verdade advindo do processo CD Original > WAV > MP3 > WAV > FLAC. Percebeu a compressão *lossy* aí no meio?

Felizmente, há formas científicas de comprovar, com boa dose de certeza, de que estamos diante de um WAV original ou de um áudio que passou por processamento com perdas no meio do caminho.

Alguns softwares já foram criados para isso. Mas talvez o principal algoritmo seja o Aucdtect²⁰ que é usado pelo Tau Analyzer²¹ fornecido no pacote de softwares. No entanto, o Tau Analyzer verifica exclusivamente CD-Audio.

Outro software mais prático para essa atividade é o Audio Identifier, fornecido na pasta de softwares. Apesar do inconveniente dele não funcionar bem em diretórios que contém caracteres acentuados (ATENÇÃO com isso!), ele é muito simples de ser usado. Basta abrir o programa rodando o AI.exe, encontrar a pasta com os arquivos WAV (sim, se estiver em FLAC ou qualquer outro lossless, tem que converter para WAV antes) e clicar com o direito sobre a faixa e acionar o Scan with AuCDtect. Pronto... o resto você vai entender.

Nesta atividade você fará alguns testes:

1. Converter todas as músicas da pasta “Músicas sem Compressão” para MP3 –V2 e então de volta para WAV. Cuidado para não sobrescrever os arquivos originais! Agora, teste esses WAVs *lossy* com o AI.
2. Agora, a fim de continuar submetendo o programa a testes, repita o teste 1 com MP3 na mais “alta qualidade”, em CBR 320 Kbps. Submeta ao AI novamente. Ele teve alguma dificuldade em identificar? Por que estaríamos fazendo esse tipo de teste?
3. Por fim, faça o mesmo procedimento dos testes 1 e 2 com o AAC, ou seja, use um parâmetro de compressão de menor bitrate (logo, maior descarte de informações) e outro de maior bitrate (mais qualidade). Quais os resultados?

É isso. Sumarize os resultados e comente-os.

Outros links interessantes:

- <http://www.maurits.vdschee.nl/fakeflac/>
- <http://blog.pkh.me/p/6-las-lossy-audio-spotter.html>

²⁰ http://tausoft.org/wiki/True_Audio_Checker_Algorithm

²¹ http://tausoft.org/wiki/True_Audio_Checker_Downloads

CONCLUSÃO

Pronto, chegou a hora de você responder da maneira mais abrangente possível a simples questão:

O que você pode aprender com esse trabalho?

Todos os três trabalhos que faremos nessa disciplina têm componentes altamente subjetivos, como as percepções individuais de qualidade, transparência e artefatos de compressão (que incomodam mais alguns indivíduos que outros). Assim sendo, não dá para fazer uma boa conclusão sem vocês abrindo a boca e se expressando, colocando em palavras.

Pois é... esse é um dos únicos trabalhos que você fará com o Giraldele em que não há conclusão escrita. **A conclusão de vocês será exclusivamente em vídeo.** Vocês farão uma videoconferência entre vocês e comentarão todas as percepções que tiveram acerca do trabalho. Deverão gravar essa videoconferência (sugiro o Google Duo e algum software que grave a tela do celular, ou do computador). Uma vez gravado o vídeo, subam para o Youtube e coloquem o link na conclusão do seu trabalho. Só isso.

Não vou fixar um tempo mínimo, apenas um máximo: 30 minutos. O mínimo eu vou contar com o bom senso de vocês em entregarem realmente uma conclusão, e nada de *“esse trabalho foi ohhh... super interessante, legal, importante pra nossa vida profissional”*. Quero observações pontuais, resultantes das atividades feitas. **Não só dos resultados obtidos nas atividades, mas da experiência em fazê-las.** Convém fazer um mini roteiro dos pontos que vocês irão abordar, para lembrarem. Mas, **jamais**, em hipótese alguma, gravem um vídeo lendo um texto, mesmo que tenha sido escrito por vocês. Não há mistério aqui: é apenas um vídeo de vocês comentando o trabalho. Quanto mais honesto for, melhor será.

Vençam a timidez, a vergonha, o medo de falar. Em maior ou menor grau, o sucesso da vida profissional de vocês dependerá do desenvolvimento dessa habilidade. Cada um de vocês é único. Basta pensar nos milhares de atores bem sucedidos que existem no mundo e ver que cada um tem sua própria maneira de se expressar. Você não precisa tentar copiar ninguém. Você precisa praticar para se tornar cada vez melhor nisso... em **se fazer entender**.

Caprichem nessa **conclusão**, pois ela terá um peso enorme na avaliação feita pelo professor (e, consequentemente, na nota). A importância dela é tamanha que seu trabalho simplesmente **NÃO SERÁ ACEITO sem a conclusão**. Não há uma pontuação definida para ela. Eu simplesmente vou começar a correção do trabalho de vocês assistindo a esse “vídeo conclusão” e em seguida corrigirei o relatório em si influenciado por essa conclusão. Entenderam o recado?

Acredite, o professor gastou muito tempo pensando em cada aspecto desse trabalho. E sim, ele fez todos os testes possíveis antes de te solicitar. Corresponda a altura!

Bom Trabalho!