



# Compressione

## Parte 2

---

Prof. Filippo Milotta  
milotta@dmi.unict.it



# Fattori di compressione per codifiche basate su PCM

## ■ Dipendono dalla implementazione della PCM:

### □ IMA ADPCM: **4a1 (75%)** ~ circa 50 kBps

- Con specifica G.721: 16 o 32 kBps (bit-rate)
- Con specifica G.723: 24 kBps (bit-rate)

### □ ACE/MACE ADPCM: **2a1 (50%)**

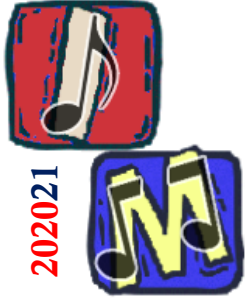
## ■ Ma soprattutto dipende dall'utente, che stabilisce in base alle necessità la fedeltà vs compressione del segnale

Queste codifiche con compressione sono di tipo *lossy*

IMA: Interactive Multimedia Association, usato in MS Windows

ACE/MACE è la compressione APPLE

Il bit-rate è il tasso di trasferimento



# Compressione percettiva (Entropia percettiva)

- J.D. Johnston ha fissato un limite teorico alla comprimibilità di un segnale se si vuole ottenere una codifica trasparente
  - **Codifica trasparente:** è una codifica compressa (*lossy*) che permette una riproduzione non distinguibile dal segnale originale non compresso
- Tale limite è di circa 2.1 bit / campione

Studio della  
*Entropia Percettiva*



# Compressione percettiva e Codifica trasparente – Esempio

## ■ CD Audio

- Tasso di campionamento: **44,1kHz**
- PCM lineare **16 bit**: **44,1kHz** \* **16** = **705,6kbps**

- Compressione a **64kbps**

- E' una codifica compressa trasparente?

Il bit rate compresso è minore di quello non compresso (705,6kbps). Questo significa che comprimendo, ogni campione verrà codificato con meno bit di quelli iniziali...

- Ogni campione verrà campionato con **64.000 / 44.100 = 1,45 bit / campione**

...non più 16 bit / campione, ma 1,45 bit / campione

- $1,45 < 2,1 \rightarrow \rightarrow \rightarrow$  Codifica NON trasparente



# Codifica trasparente

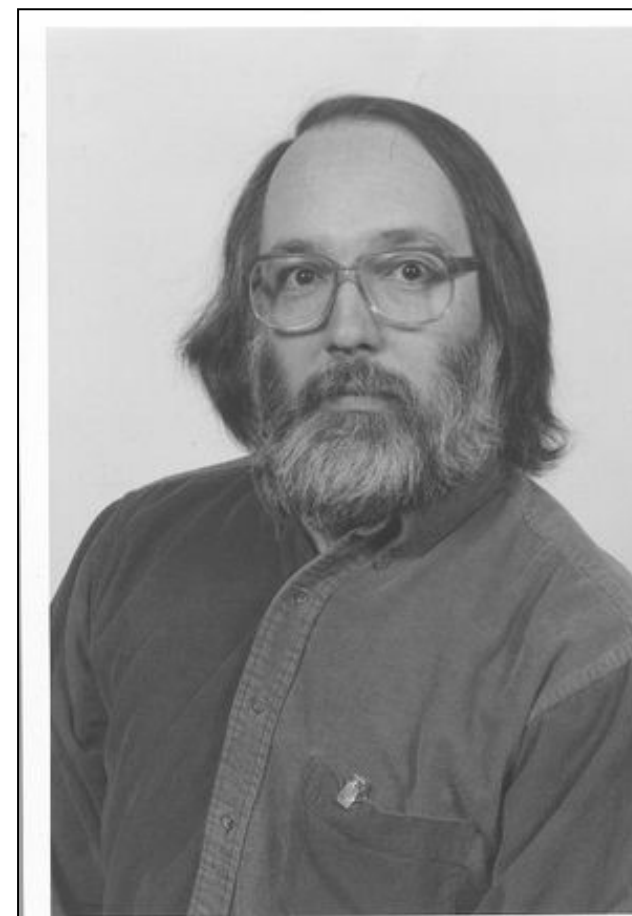
- La trasparenza non è una proprietà necessaria delle codifiche di compressione
- E' più che altro una conseguenza diretta del bit-rate di compressione scelto
- → Cioè, tipicamente non si può scegliere il bit-rate, perché è dettato dalla strumentazione
  - Fissato il bit-rate è però possibile dire se la codifica sarà trasparente

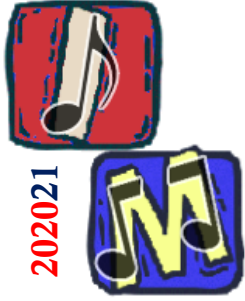


# James D. Johnston

(?? – ancora in vita?)

- Noto come *Il padre delle codifiche di compressione di tipo percettivo*
- Responsabile di numerose codifiche all'interno dei formati MP3 e MPEG-2. Lavorò per 26 anni nei Bell Labs. Dopo aver lasciato i Bell Labs si è trasferito alla Microsoft.





# La tecnica **Compansion**

*(Cap. 3.6 – Pag.130 – IV edizione)*

- **Compansion = Compression + Expansion**  
(intesi proprio come operatori dinamici)
  - Compressione in fase di registrazione
  - Espansione in riproduzione
- → Utilizzata negli schemi di **compressione di tipo percettivo**
- Ideata dalla Dolby negli anni '60-'70 per risolvere i problemi di SNR sui nastri magnetici

*Dal testo: il rumore, in particolar modo il fruscio delle cassette audio, non è più di ampiezza costante e indipendente dal segnale: ora è più forte quando il segnale è più forte, ed è più debole quando il segnale è più debole*



# Compressione di tipo percettivo

- Negli schemi di compressione di tipo percettivo vengono impiegate numerose tecniche, combinate in vari modi
  - Abbiamo appena (ri)visto la *Compansion*
- a questa aggiungiamo:
  - la *quantizzazione non uniforme* (in  **$\mu$ -law** e **A-law**)
  - la codifica differenziale *ADPCM*
  - e altre 4 tecniche principali... →

Prendiamo in considerazione anche le **debolezze dell'udito umano** :

- La THQ (Threshold to Quiet)
- Le Bande Critiche e il mascheramento





# Compressione di tipo percettivo

- Si basa su 4 tecniche principali:
  1. Block Coding
  2. Transform Coding
  3. Sub-band Coding
  4. Huffman Coding

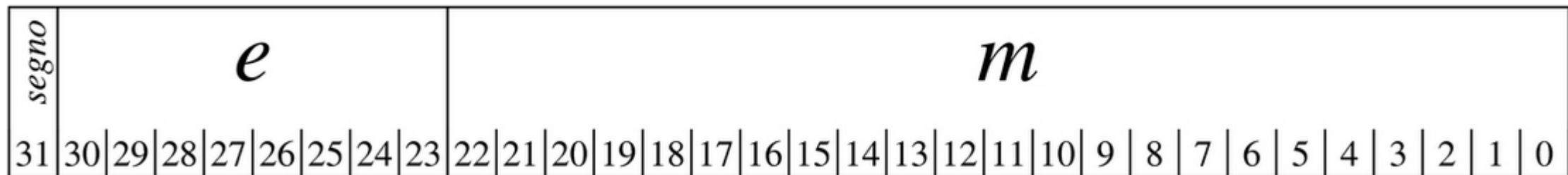


# Block Coding

## Codifica per blocchi

- La quantizzazione non uniforme si può vedere come una codifica a virgola mobile
  - Esponente e mantissa
  
- Nelle tracce audio ci si aspetta che l'esponente vari pochissimo
  - Si può codificare l'esponente una volta sola per blocco

Es.:  $+ 10^{-e=-8} \times (m = 123456) = +0,00123456$





# Block Coding

## Codifica per blocchi – Pre-echi

- Problema dei pre-echi
  - Dovuto principalmente a transitori impulsivi
- Si può risolvere in 2 modi:
  - Ridurre la durata dei blocchi
  - Usare blocchi di durata variabile in base all'andamento dei transitori, per circoscrivere i rumori impulsivi

Introducono forti cambiamenti di scala (ordine di grandezza)  
→ Rendono impossibile utilizzare un unico esponente per tutto il blocco

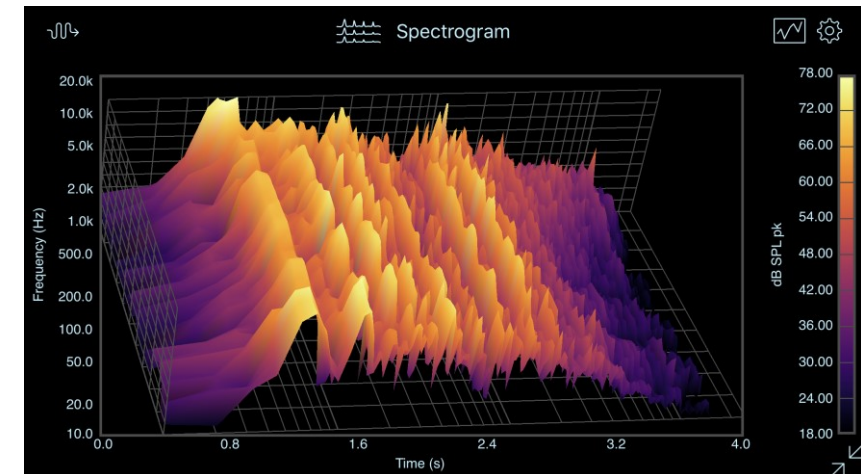
In modo che sia più probabile che tutte le intensità dentro un blocco siano dello stesso ordine di grandezza



# Transform Coding

## Codifica nel dominio delle frequenze

- Il segnale audio nel dominio delle frequenze tende a variare meno rispetto al dominio dello spazio  
(→ *Spettrogramma*)



- Al posto della DFT applichiamo trasformate efficienti come la FFT o la DCT
  - La DCT è da preferire

Perché la DFT e la FFT utilizzano numeri complessi, mentre la DCT solo numeri reali, e le funzioni di base sono tutte (e solo) sinusoidi  
[Pag 168 – Fig.4.12]

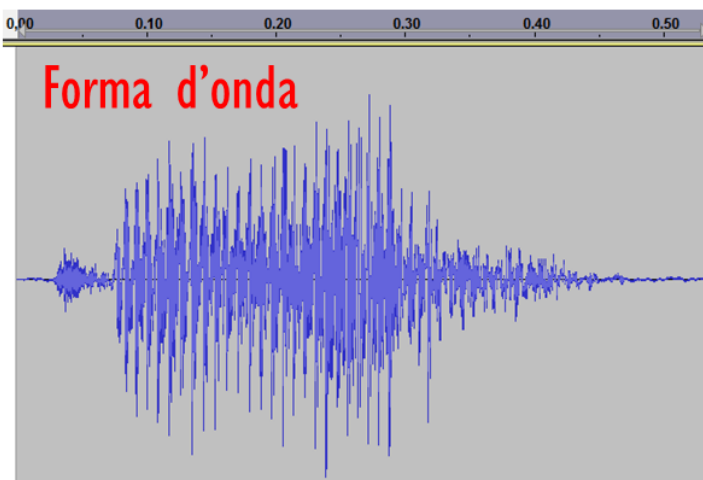
FT: Fourier Transform  
DFT: Discrete FT  
FFT: Fast FT  
DCT: Discrete Cosine Transform  
MDCT: Modified DCT



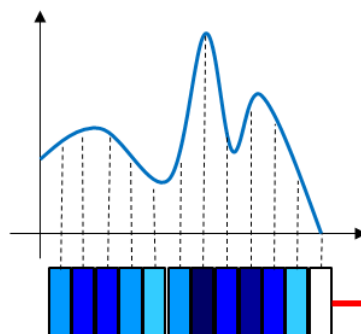
# Ripasso – Acustica – Parte 5



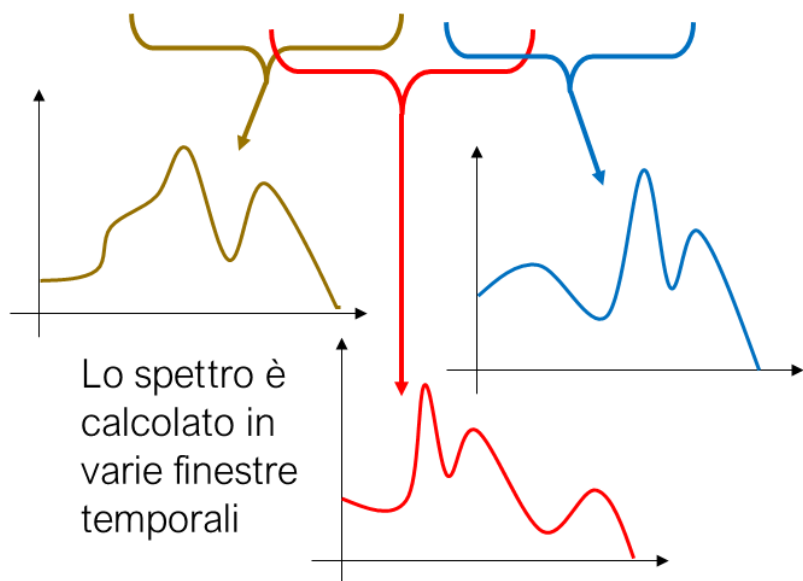
## Altre rappresentazioni dello spettro



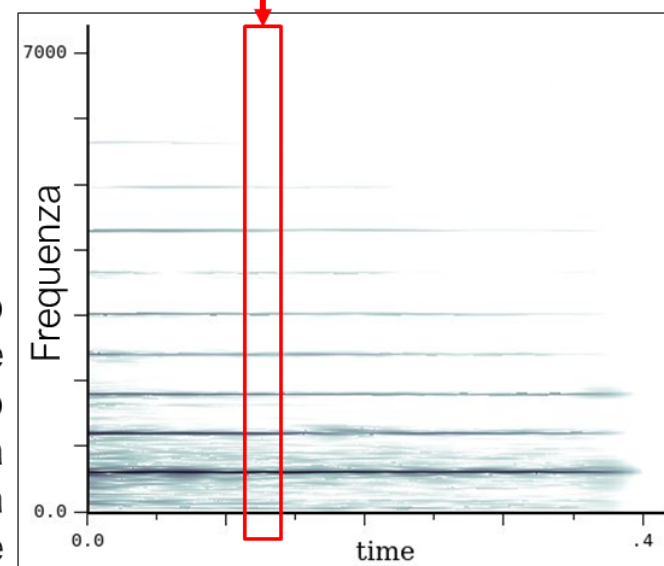
Spettro (continuo)



Se lo vedessimo... dall'alto  
 Colori più scuri → Valori più alti  
 Colori più chiari → Valori più bassi



Ogni colonna nello spettrogramma è relativa ad uno spettro calcolato da una finestra temporale



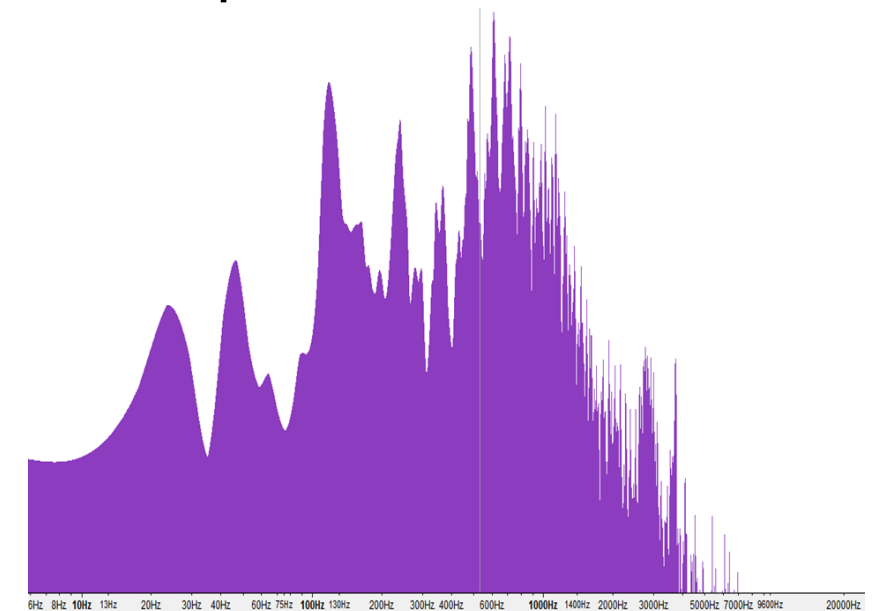


# Transform Coding

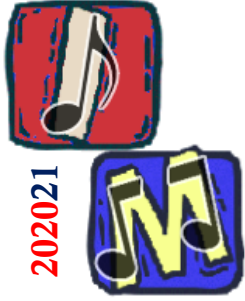
## Codifica nel dominio delle frequenze

- Vantaggiosa se applicata a blocchi con bassa gamma dinamica
- Per evitare i pre-echi
  - Si calcola la trasformata su intervalli sovrapposti per il 50%
    - Con questo metodo si ottiene però il doppio dei campioni necessari
- La compressione viene quindi applicata nel dominio delle frequenze,
  - Sui coefficienti delle trasformate
  - Sugli spettri (→ Sub-band Coding)

**Windowing**  
cioè sovrapposizione per finestre



Spettro (continuo)



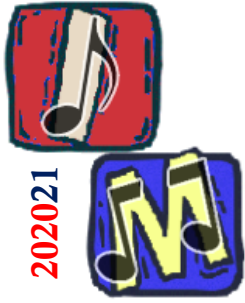
# Sub-band Coding

## Codifica per sottobande

- Analogamente alla codifica a blocchi →
- Divide lo spettro di frequenze in sottobande codificate in maniera individuale
  - Le sottobande con gamma dinamica ristretta possono essere codificate con meno bit
- Il processo di ***band-splitting*** non è semplice e richiede il giusto compromesso fra complessità di splitting e tasso di compressione

Ad esempio, si potrebbe applicare una suddivisione in base alle bande critiche (come nella MDCT)





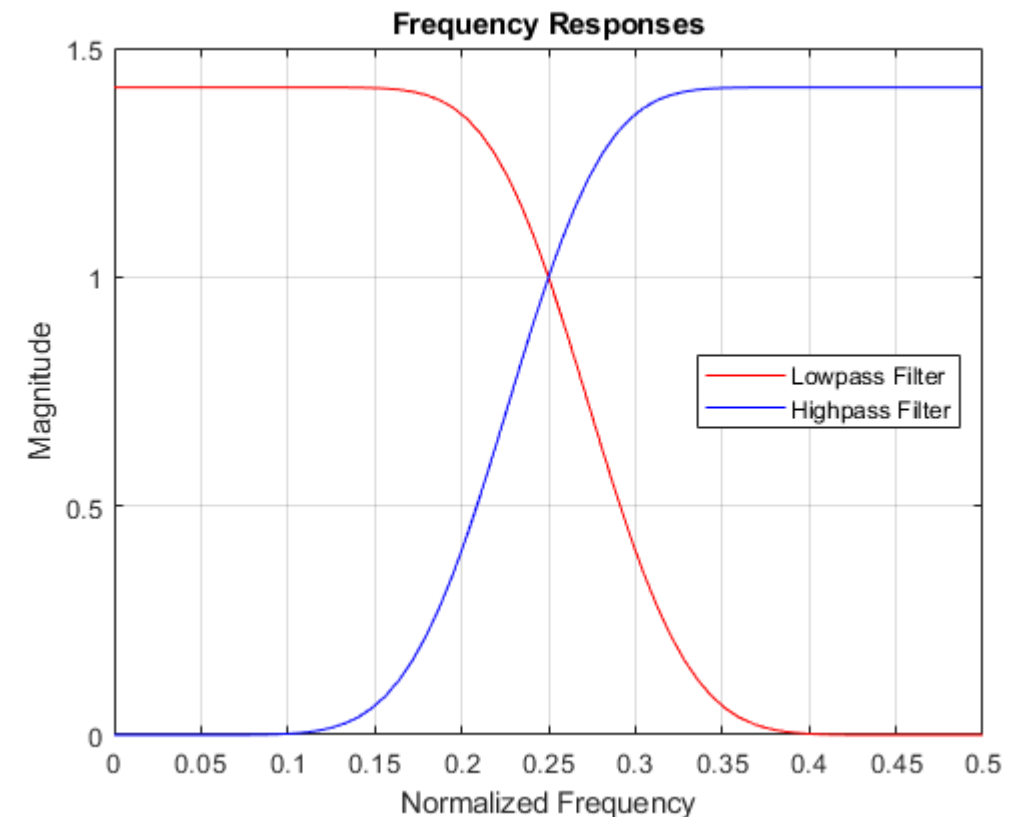
# Sub-band Coding

## Codifica per sottobande – QMF

Filtraggio a specchio di quadratura

### ■ QMF: Quadrature Mirror Filtering

- Si considerano 2 segnali separati: basse e alte frequenze
- I filtri QMF possono essere usati in cascata e possono operare in polifase (cioè in parallelo)
- Le due bande devono avere la stessa grandezza







# Huffman Coding

## Compressione di Huffman

- Codifica ottimale  
(→ si avvicina al limite di Shannon)
- Codici senza prefissi
- Compressione Lossless
  
- Algoritmo greedy:
  1. Selezione di due caratteri con frequenze minime
  2. Sostituzione dei due caratteri con uno fittizio la cui frequenza è la somma delle precedenti due
  3. Ripetere fino a ottenere il carattere con frequenza 1



# Huffman Coding

## Compressione di Huffman

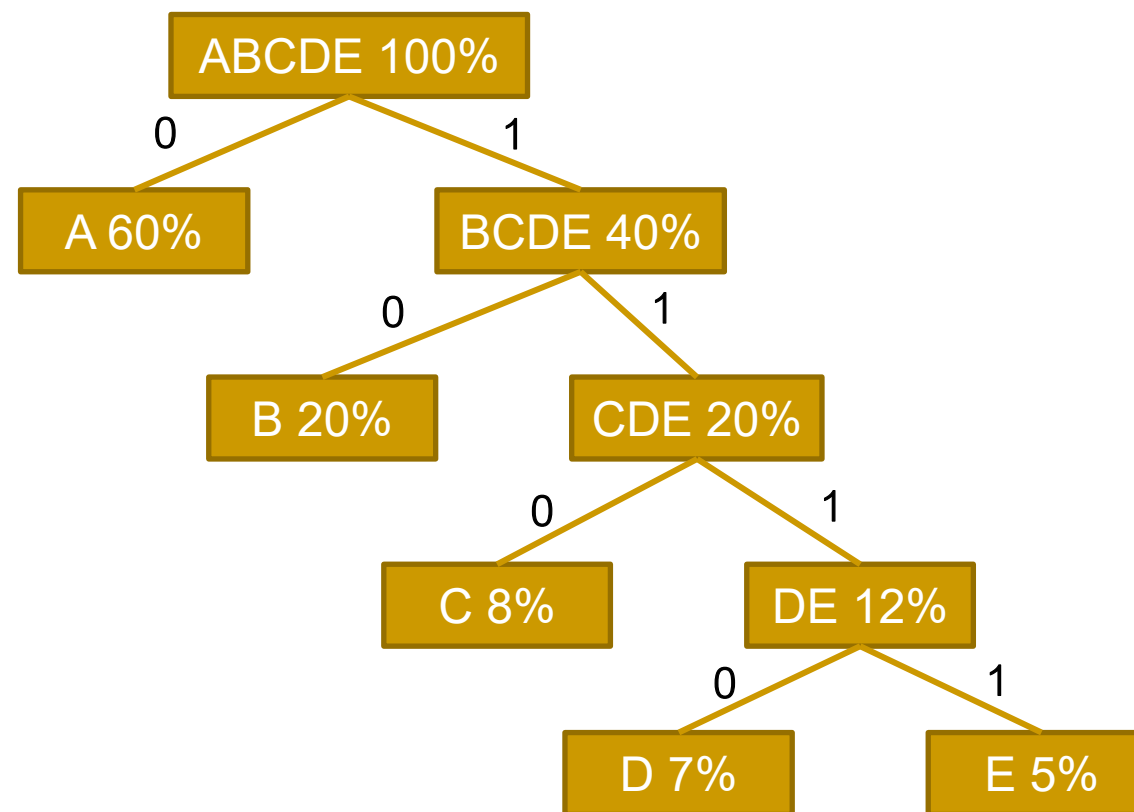
- A questo punto si ottiene un albero, la cui forma può variare in base alle scelte prese
- Si etichettano i rami binari con 0 e 1
- Si assegnano le codeword alle foglie leggendo dalla radice le etichette dei rami
  - Caratteri frequenti avranno codeword brevi
  - Caratteri rari avranno codeword lunghe



# Huffman Coding

## Esempio

Simboli e frequenze iniziali: A 60% B 20% C 8% D 7% E 5%



Simboli e codifiche finali:

A: 0, B: 10, C: 110, D: 1110, E: 1111



# Schema generale di compressione di tipo percettivo

1. **Block-Coding:** Segmentazione della traccia audio in frame *quasi-stazionari* di 2-50 msec
  - Quasi-stazionari: con transitori poco variabili
2. **Transform Coding:** si passa all'analisi nel dominio delle frequenze
3. **Sub-band Coding** opzionale, se si vuole ulteriormente ottimizzare la codifica del range dinamico
4. **Rimozione delle ridondanze** tramite codifiche lossy (ADPCM) o lossless (Huffman)



# Approfondimenti

- ***[EN] Paper: Johnston's limit to compression***  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.422.1835&rep=rep1&type=pdf>
- ***[EN] J.D. Johnston brief biography***  
[https://ethw.org/James\\_D.\\_Johnston](https://ethw.org/James_D._Johnston)
- ***[EN] A tutorial on MPEG/Audio compression***  
<https://www.icg.isy.liu.se/courses/tsbk35/material/mpegaud.pdf>