FUNDAMENTOS DE COMUNICAÇÃO DE DADOS

Licenciatura em Engenharia Informática

Departamento de Informática
Universidade do Minho

2021-2022

EQUIPA DOCENTE

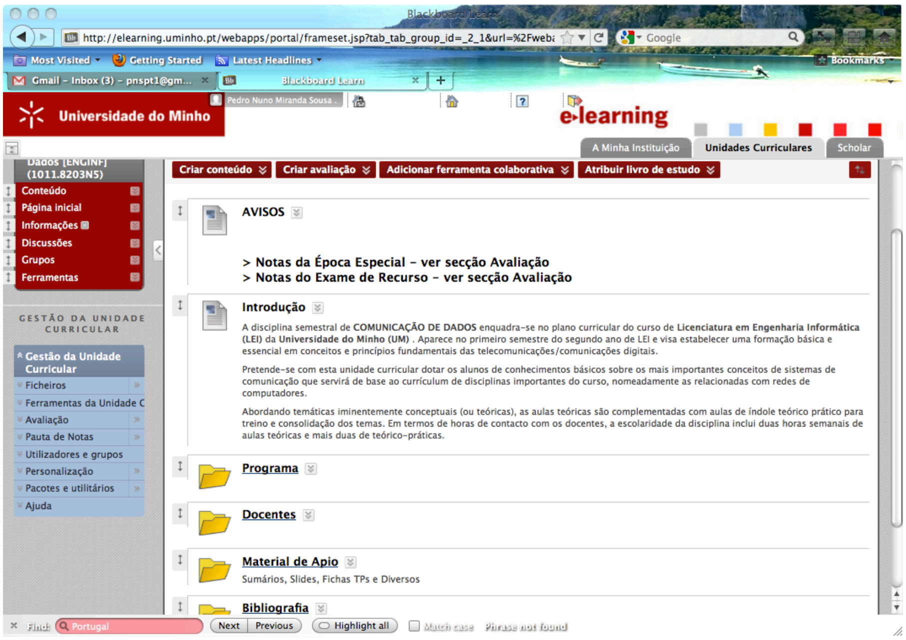
- **Pedro Sousa**
  `pns@di.uminho.pt`
  `253 604 436`
  (Docente Responsável: Teóricas + 3 Turnos TPs)
- **Bruno Dias**
  (2 Turno TP)

INFORMAÇÕES E MATERIAL DE APOIO À UNIDADE CURRICULAR

• Aceder à plataforma de e-learning da Universidade do Minho

Algum documento que necessite password: FCD2122

Pré-inscrição BB: FCD2122



Introdução

A unidade curricular **FUNDAMENTOS COMUNICAÇÃO DE DADOS** enquadra-se no plano curricular do curso da Licenciatura em Engenharia Informática da Universidade do Minho. Aparece no primeiro semestre do segundo ano de LEI e visa estabelecer uma formação básica e essencial em conceitos e princípios fundamentais da área das Comunicações Digitais/Telecomunicações.

Pretende-se com esta unidade curricular dotar os alunos de conhecimentos básicos sobre os mais importantes conceitos de sistemas de comunicação que servirá de base ao curriculum de outras disciplinas relacionadas com a área das redes de computadores / comunicações por computador .

Abordando temáticas iminentemente conceptuais/teóricas, as aulas teóricas são complementadas com aulas de índole teórico prático para treino e consolidação dos temas.

BIBLIOGRAFIA

- ***Fundamentos das Telecomunicações***
  V. Freitas, Universidade do Minho.

  [algum material complementar poderá ser facultado ao longo do semestre]

- ***Principles of Communications, 5th Edition***
  R. Ziemer, W. Tranter, John Wiley & Sons.

- ***Communication Systems,***
  A. Bruce Carlson, McGraw-Hill Series

Fundamentos de Comunicação de Dados
Licenciatura em Engenharia Informática
Departamento de Informática, Universidade do Minho

AVALIAÇÃO

- **Regime de Avaliação**
  2 Testes de Avaliação (T1,T2)
  - » *em regime de avaliação periódica distribuídos ao longo do semestre*
  - » *mais informações sobre os testes serão posteriormente anunciadas*
  - » **Nota Final [0.5*T1 + 0.5*T2]**

- **Exame:** os alunos sem aproveitamento (i.e. nota final < 10) podem efectuar uma prova final de avaliação na data definida para o efeito pelo Conselho de Cursos.

DATAS dos Testes de Avaliação (<u>confirmar no calendário LEI</u>)

Universidade do Minho

Conselho Pedagógico da EEUM

Calendário Escolar: Ano Letivo 2021/2022

1º Ciclo de Estudos, Ciclo de Estudo Integrado, 2º Ciclo e 3º Ciclo de Estudos

		Semana	2ª Feira	3ª Feira	4ª Feira	5ª Feira	6ª Feira	Sábado
	1	04/10 a 09/10						
	2	11/10 a 16/10						
	3	18/10 a 23/10						
	4	25/10 a 30/10						
	5	01/11 a 06/11						
	6	08/11 a 13/11						
	7	15/11 a 20/11						
	8	22/11 a 27/11						
sem aulas →	9	29/11 a 04/12		1º Teste 30 Nov				
	10	06/12 a 11/12						
	11	13/12 a 18/12						
		20/12 a 25/12						
		27/12 a 01/01						
última semana de aulas →	12	03/01 a 08/01						
	13	10/01 a 15/01				2º Teste 13 Jan		
	14	17/01 a 22/01				Publicação das notas da avaliação periódica ou contínua com		
		24/01 a 29/01	5 dias úteis antes da data do exame final (RAUM 142.º, p. 9)					
		31/01 a 05/02				Recurso 3 Fev		
		07/02 a 12/02						
	1	14/02 a 19/02						

TURNOS T + TPs

- T1, T2
- TP1, TP2, TP3, TP4, TP5

Respeitar inscrições efetuadas via DC (**<u>não enviar emails ao docente a pedir troca de turnos!</u>**)

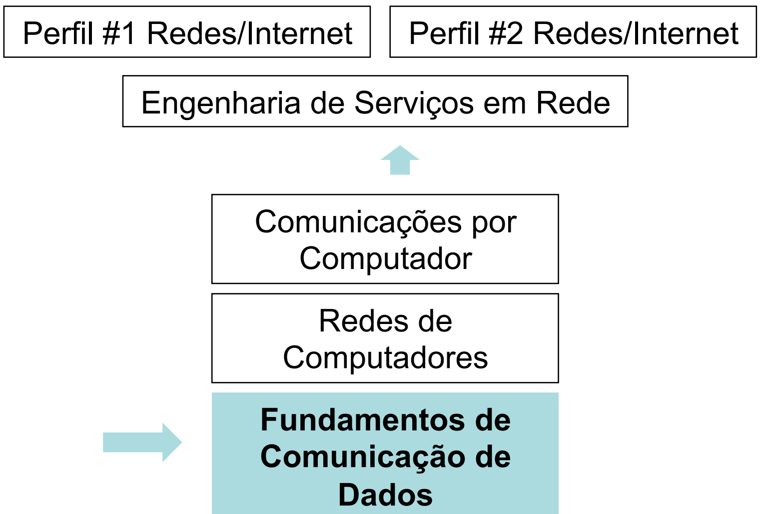
- Material obrigatório para as TPs
  - Sebenta/capítulo da disciplina
  - Máquina calculadora
  - Ficha de exercícios (disponibilizada no elarning na respectiva semana)
- Início das aulas TPs – semana 11 Outubro
- Regras de funcionamento T & TPs / Faltas / etc.

TURNOS T + TPs

Fundamentos de Comunicação de Dados
Licenciatura em Engenharia Informática
Departamento de Informática, Universidade do Minho

Área das Redes / Comunicações / Internet

Enquadramento da UC no LEI/MEI ...



PROGRAMA RESUMIDO

I. Teoria da Informação
II. Digitalização
III. Multiplexagem
IV. (Cap. Introdução) + Análise de Sinais
V. Análise de Sistemas de Transmissão
VI. Códigos para Controlo de Erros (+ breve introdução a ruído e erros)

PROGRAMA DETALHADO

I Teoria da Informação

Introdução
Informação, informação própria, entropia e débito de informação
Codificação da fonte
Rendimento e compressão obtida por um código
Códigos de Shannon-Fano e exemplos de exercícios
Codificação por blocos
Codificação de fontes discretas sem memória
Codificação de fontes com memória (de 1ª ordem)
Exemplo de aplicações

PROGRAMA DETALHADO

II Digitalização

Conceitos prévios: Largura de Banda de um sinal; Banda de Transmissão e Ritmo máximo de símbolos num sistema

Analógico versus Digital (sinais, transmissão, ...)

Frequência de Amostragem

Teoria da Amostragem

Quantização Uniforme e não uniforme

Ruído de Quantização

Conversão Analógico a Digital

PCM e Ruído em PCM (...e questões relacionadas...)

Breve referência a outras técnicas de digitalização

PROGRAMA DETALHADO

III Multiplexagem

Por Divisão do Tempo (TDM)

Organização das tramas, Tramas PCM

Hierarquias de Multiplexagem (e.g. PDH, SONET, ...)

Por Divisão estatística do Tempo (TDM estatístico)

Modelo M/D/1 e formulas associadas, probabilidade de sobrelotação e perda; atraso em fila, exemplo de problemas

Por Divisão de Frequência (FDM)

Técnicas combinadas

Outras técnicas de Multiplexagem

PROGRAMA DETALHADO

Cap1: Sinais e Sistemas de Comunicação; Limitações Fundamentais: Ritmo de Nyquist e Lei de Hartley Shannon; Capacidade de um canal

IV Análise de Sinais

Sinais periódicos e não periódicos

Análise Espectral de Sinais: Séries de Fourier

Passagem do domínio das frequências para o domínio do tempo (e vice-versa)

Potência de um sinal

Energia e Largura de Banda de um Sinal

Teorema de Parseval

Modulação e codificação de sinais (Teorema de Modulação)

PROGRAMA DETALHADO

V Análise de Sistemas de Transmissão

Transmissão e Filtragem de Sinais

Função de Transferência (e representação gráfica)

Largura de Banda de Transmissão

Atenuação/Distorção do Sinal

Ganhos e Perdas de Potência

Filtros reais e ideais

Filtros (sistemas) de ordem superior: Filtros de Butterworth

PROGRAMA DETALHADO

VI Códigos para Controlo de Erros

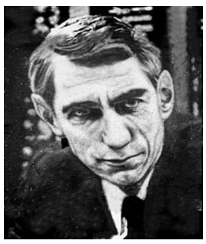
Ruído e erros (implicações para a transmissão/dados)
Códigos Lineares de Bloco
Rendimento de um código
Distancia mínima de um código
Capacidade de controlo de erros de um código
Geração de Códigos Cíclicos Sistemáticos
Exemplos de exercícios
Circuitos Codificadores de Códigos e analise e funcionamento

I. TEORIA DA INFORMAÇÃO

Em termos gerais a Teoria da Informação é uma teoria que aborda várias temáticas relacionadas com sistemas de comunicação, transmissão de dados, informação, codificação, compressão de dados, ruído, correção de erros, entre outras...

Claude Shannon Engenheiro/Matemático/ Investigador Americano é reconhecido como sendo o "pai" da Teoria da Informação

Também apresentou importantes contributos noutras áreas: e.g. circuitos digitais, criptografia, inteligência artificial, digitalização...



Claude Shannon
[April 30, 1916 – February 24, 2001]

I. TEORIA DA INFORMAÇÃO

Teorema Fundamental da Teoria de informação

*"Dado um **canal de comunicação** e uma **fonte de informação** cujo débito de informação não excede a capacidade do canal, existe um código tal que a informação pode ser transmitida através do canal com uma frequência de erros arbitrariamente pequena, apesar da presença de ruído no canal."*

Fundamentos de Comunicação de Dados
Licenciatura em Engenharia Informática
Departamento de Informática, Universidade do Minho

I. TEORIA DA INFORMAÇÃO

Teoria de informação estuda 4 problemas fundamentais:

– A medida de informação produzida por uma fonte ...

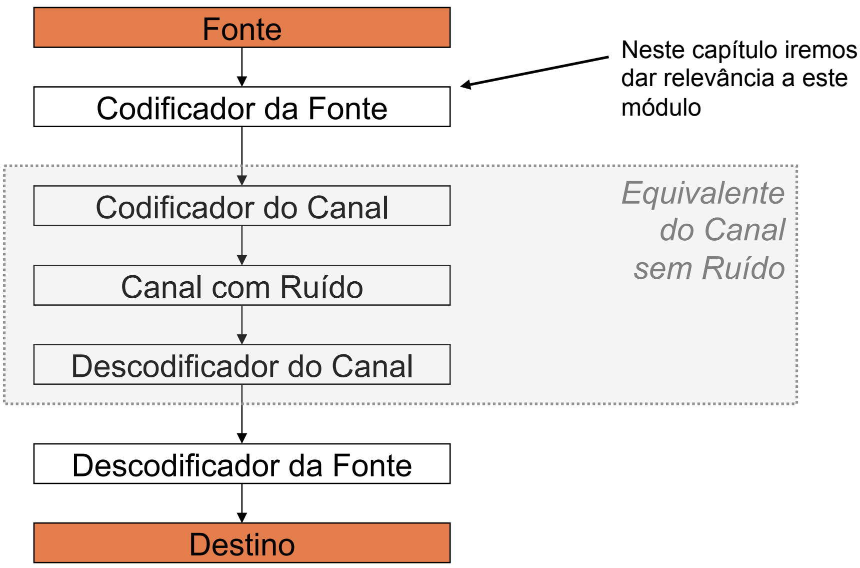
– A codificação eficiente da fonte ...

– A capacidade do canal ...

– A codificação do canal para controlo de erros ...

I. TEORIA DA INFORMAÇÃO

Sistema de Comunicação com codificação da fonte e do canal



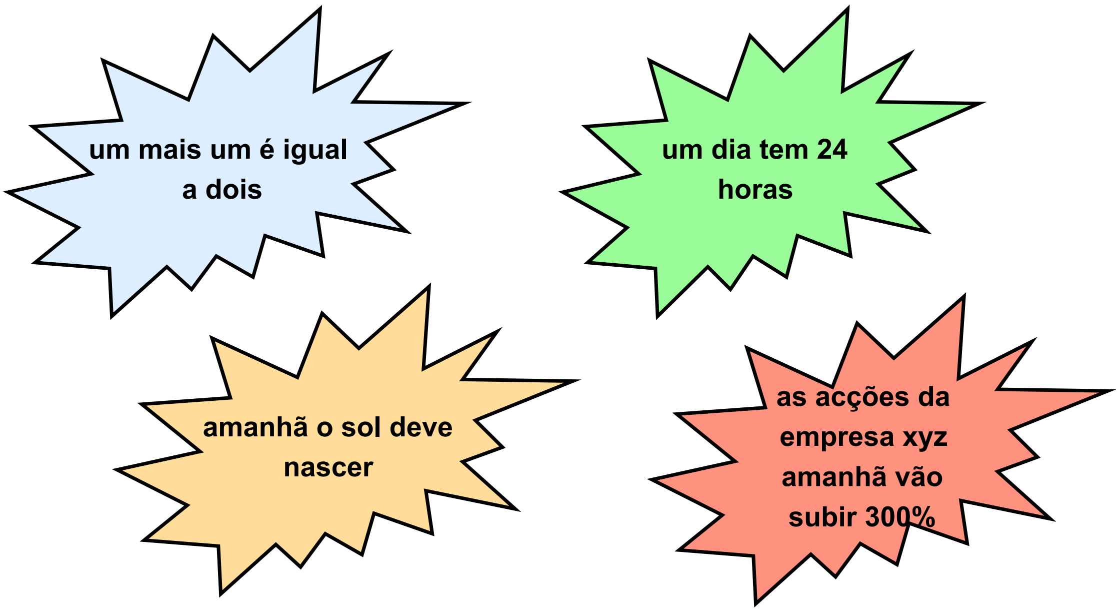
Fundamentos de Comunicação de Dados
Licenciatura em Engenharia Informática
Departamento de Informática, Universidade do Minho

I. TEORIA DA INFORMAÇÃO

- Estudo da produção e transferência de informação
- Relevância na informação da mensagem em si e não dos sinais utilizados para a transmitir
- **Informação**: (no contexto das tele/comunicações)

  "*objecto imaterial útil produzido por uma fonte que tem de ser transmitido para um determinado destino*"

I. TEORIA DA INFORMAÇÃO



I. TEORIA DA INFORMAÇÃO

- Como definir uma **medida de informação** ?
  - relacionada com o grau de incerteza do destinatário relativamente à mensagem que vai receber
  - relacionada com a probabilidade da ocorrência da mensagem
  - vai ser definida como uma função que leva em conta essa probabilidade f(Pi)

I. TEORIA DA INFORMAÇÃO

- Informação própria de uma mensagem X_i:

$$\boxed{I_i = f(P_i)}$$

- Propriedades:

(i) $f(P_i) \geq 0 \quad \text{para} \quad 0 \leq P_i \leq 1$ → informação nunca é negativa

(ii) $\lim_{P_i \to 1} f(P_i) = 0$ → se a probabilidade for 1, a informação é 0, logo não serve para nada

(iii) $f(P_i) > f(P_j) \quad \text{para} \quad P_i < P_j$ → - probabilidade → + informação; + probabilidade → - informação

(iv) $f(P_i P_j) = f(P_i) + f(P_j)$

↳ sucessão de mensagens independentes

I. TEORIA DA INFORMAÇÃO

- Adoptar uma função (função logarítmica negativa) que satisfaz estas propriedades:

$$\boxed{-\log_b()}$$

- A base adoptada define a unidade de medida de informação
- base=2 na teoria de informação
- logo a unidade correspondente é o *bit*

I. TEORIA DA INFORMAÇÃO

Bit como unidade de medida de informação

O bit é a quantidade de informação necessária para escolher uma entre duas alternativas igualmente prováveis ou, a quantidade de informação contida numa mensagem emitida por uma fonte capaz de emitir apenas duas mensagens distintas e equiprováveis.

Portanto, e por definição, a quantidade de informação, ou informação própria, I_i numa mensagem x_i é dada por:

$$\boxed{I_i \stackrel{def}{=} \log_2 \frac{1}{P_i} \quad bits}$$

→ medida de informação

→ probabilidade da mensagem ou do acontecimento

I. TEORIA DA INFORMAÇÃO

- Assumir uma fonte que emite uma série de símbolos $X = \{x_1, ..., x_m\}$ com probabilidades $\{P_1.....P_m\}$

- **Entropia**: informação média (por símbolo) gerada pela fonte

→ probabilidade da mensagem x
→ informação própria da mensagem x

$$\boxed{\mathcal{H}(X) \stackrel{def}{=} \sum_{i=1}^{m} P_i \, I_i \;=\; \sum_{i=1}^{m} P_i \log_2 \frac{1}{P_i} \quad bits/símbolo}$$

I. TEORIA DA INFORMAÇÃO

- Quais os **limites para a entropia** de uma fonte?
- Valor que depende:
  - das probabilidades dos símbolos da fonte e
  - da cardinalidade (m)

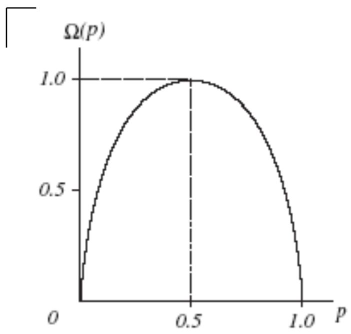
→ quantidade de símbolos gerados

$$\boxed{0 \leq \mathcal{H}(X) \leq \log_2 m}$$

I. TEORIA DA INFORMAÇÃO

- **Exemplo 1**: Fonte binária (m=2); P_1=p e P_2=1 -p ; entropia?

$$\mathcal{H}(X) = \Omega(p) \stackrel{def}{=} p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$$



I. TEORIA DA INFORMAÇÃO

- **Débito de Informação**
  - indica o débito médio de informação por segundo
  - assumindo que a fonte produz r_s símbolos por segundo:

$$\boxed{\mathcal{R} \stackrel{def}{=} r_s\, \mathcal{H}(X) \quad bits/seg}$$

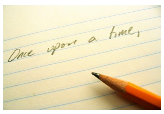
$$\boxed{\mathcal{H}(X) \stackrel{def}{=} \sum_{i=1}^{m} P_i\, I_i \;=\; \sum_{i=1}^{m} P_i\, \log_2 \frac{1}{P_i} \quad bits/símbolo}$$

I. TEORIA DA INFORMAÇÃO

- **Exemplo 2**: Fonte emite 2000 símbolos/seg de um alfabeto de 4 símbolos (m=4) com probabilidades:

x_i	P_i	I_i
A	$1/2$	1
B	$1/4$	2
C	$1/8$	3
D	$1/8$	3

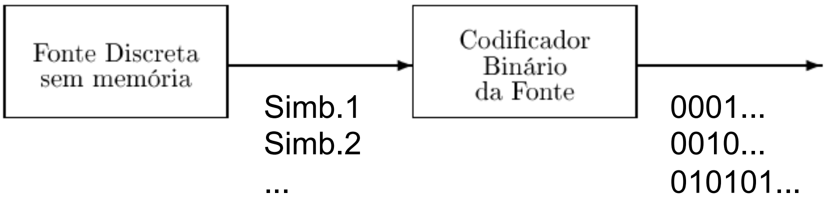
- Entropia?
- Débito de informação?



$$\mathcal{H}(X) = \frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{8} \times 3 = 1.75 \quad \text{bits/símb}$$

$$\mathcal{R} = 2000 \times 1.75 = 3500 \quad \text{bits/seg}$$

I. TEORIA DA INFORMAÇÃO



- N_i - comprimento da palavra de código correspondente ao símbolo i
- **Comprimento médio do código**:

$$\overline{N} = \sum_{i=1}^{m} P_i N_i \quad \text{dig bin/símbolo}$$

I. TEORIA DA INFORMAÇÃO

- **Rendimento do código**

$$\rho = \frac{\mathcal{H}(X)}{\overline{N}} \leq 1$$

- **Compressão obtida numa codificação**

$$c = \frac{N_f - \overline{N}}{N_f} \times 100\ \%$$

codificação com um **código de comprimento fixo mínimo**

A 00
B 01
C 10
D 11
↳ no mínimo 2 bits
$N_f = 2$

I. TEORIA DA INFORMAÇÃO

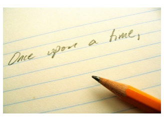
- **Como obter códigos?**
  - existem várias alternativas com diferentes desempenhos
  - os códigos necessitam de ser decifráveis (e.g. desigualdade de kraft apresentada na secção códigos óptimos)

$$\boxed{\mathrm{Kr} = \sum_{i=1}^{m} 2^{-N_i} \leq 1}$$

  - melhores códigos -> melhores rendimentos

I. TEORIA DA INFORMAÇÃO

- **Exemplo**: diferentes codificações para uma fonte que gera quatro símbolos (entropia 1.75 bits/símbolo) – Comprimentos médios e rendimentos dos códigos?



x_i	P_i	Código I	Código II	Código III	Código IV
A	1/2	00	0	0	0
B	1/4	01	1	01	10
C	1/8	10	10	011	110
D	1/8	11	11	0111	111
$\overline{N}$		2.0	1.25	1.875	1.75
Kr		1.0	1.50	0.9375	1.0

rendimento 88%

menor que a entropia!!
mas código não decifrável
BA = C => 10 = 10

código em vírgula
melhor que código I

código em árvore
que neste caso
tem rendimento =
100%

I. TEORIA DA INFORMAÇÃO

- **Códigos de Shannon-Fano / *Huffman e outras variantes***
  - Podem ser usados para construir códigos decifráveis
  - Geram códigos de comprimento variável
  - Geram códigos com "*bom*" rendimento
  - Algoritmos para geração de códigos? – vamos analisar unicamente um dos algoritmos mais simples para construção de códigos deste tipo
    - » Códigos de *Shannon-Fano*

I. TEORIA DA INFORMAÇÃO

- **Códigos de Shannon-Fano** (nota: em alguma bibliografia estes códigos são também por vezes associados aos Códigos de Huffman, mas na realidade estes últimos são uma evolução dos primeiros, e usam uma técnica distinta – corrigir na pp. 208 -)

(1) Ordenar os símbolos por ordem decrescente de probabilidade;

(2) Dividir o conjunto assim ordenado em dois subconjuntos tais que a soma das probabilidades em cada um deles seja o mais aproximadamente possível igual a metade da soma das probabilidades no conjunto anterior. Manter a ordenação.

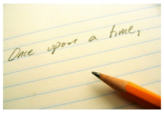
(3) O dígito seguinte do código binário dos símbolos do primeiro dos sub-conjuntos é o **0** e o dos do outro é o **1**;

(4) Se os sub-conjuntos contêm um só elemento, a codificação terminou para esses sub-conjuntos;

(5) Repetir para cada um dos restantes sub-conjuntos (passo 2.)

I. TEORIA DA INFORMAÇÃO

Códificação da fonte - **Exemplo**: aplicar o algoritmo anterior para codificar a fonte com oito símbolos (m=8)



x_i	A	B	C	D	E	F	G	H
P_i	0.50	0.15	0.15	0.08	0.08	0.02	0.01	0.01

→ 1º ordenar

x_i	P_i	Passos de codificação 1	2	3	4	5	6	Código
A	0.50	0						0
B	0.15	1	0	0				100
C	0.15	1	0	1				101
D	0.08	1	1	0				110
E	0.08	1	1	1	0			1110
F	0.02	1	1	1	1	0		11110
G	0.01	1	1	1	1	1	0	111110
H	0.01	1	1	1	1	1	1	111111
$\mathcal{H}(X) = 2.15$								$\overline{N} = 2.18$

→ $0{,}5\times1+0{,}15\times3+0{,}15\times3+0{,}08\times3+\ldots$

Entropia?

Código?

Comprimento médio?

Rendimento?

$\frac{H(x)}{\overline{N}} = 0{,}98 = 98\%$

Compressão ?

$C = \frac{N_f - \overline{N}}{N_f} = \frac{3-2{,}18}{3} \times 100\% = 97\%$

↳ $\log_2 8 = 3$ ↳ m

↳ $\log_2 m < 3$, m = 8

$\overline{N} < H(x) \Rightarrow$ nunca pode acontecer

I. TEORIA DA INFORMAÇÃO

– Codificação por blocos

→ geralmente melhoria da codificação da fonte, melhor rendimento

- agrupar símbolos da fonte e proceder à sua codificação
- daí a noção de "bloco"
- blocos de K símbolos
- normalmente leva a melhorias no rendimento do código...
- ... e na compressão obtida

I. TEORIA DA INFORMAÇÃO

– **Exemplo:**

- Fonte que emite símbolos de um alfabeto X com apenas dois símbolos X={A,B}; $P_A = 0.8$ e $P_B = 0.2$. (entropia = 0.722 bits/símbolo)
- Se se codificarem dois símbolos de cada vez temos um novo alfabeto Y={AA,AB,BA,BB}
- $P_{ij} = P_i * P_j$
  - por se tratar de uma fonte sem memória
  - ou seja, símbolos estatisticamente independentes
- código de *Shannon-Fano* para Y (blocos de K=2)?

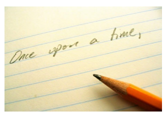
I. TEORIA DA INFORMAÇÃO

- Tabela das probabilidades/palavras de código

Código?

Comprimento médio?

y_i	P_{y_i}	Código
AA	0.64	0
AB	0.16	11
BA	0.16	100
BB	0.04	101
$\overline{N}_2$		1.56



- para uma codificação K=1 comprimento médio do código era?
- logo ....

I. TEORIA DA INFORMAÇÃO

normalizar

$\overline{N}_2 = 1.560$ *dígitos binários/símbolo*$_Y$

$\overline{N} = \frac{\overline{N}_2}{2} = 0.780$ *dígitos binários/símbolo*$_X$

y_i	P_{y_i}	Código
AA	0.64	0
AB	0.16	11
BA	0.16	100
BB	0.04	101
$\overline{N}_2$		1.56

Rendimento e compressão obtidos com (K=2) ?

$$\rho = \frac{\mathcal{H}(X)}{\overline{N}} = \frac{0.722}{0.780} = 0.926$$

$$c = \frac{N_f - \overline{N}}{N_f} \times 100 = \frac{1-0.780}{1} = 22\ \%$$

Rendimento e compressão obtidos com (K=1) (sem blocos) ?

0.722

0%

I. TEORIA DA INFORMAÇÃO

- **Rendimento e compressão obtidos com (K=3) ?**
  - experimentar.... melhor rendimento e compressão?
- **O que está a acontecer aos comprimentos médios dos códigos?**
  - à medida que K aumenta $\overline{N}$ tem tendência a diminuir; matematicamente isto é expresso na seguinte expressão:

$$\mathcal{H}(X) \leq \overline{N} < \mathcal{H}(X) + \frac{1}{K}$$

I. TEORIA DA INFORMAÇÃO

Um dos teoremas fundamentais da Teoria da Informação

Toda a fonte de informação caracterizada por um valor da entropia $\mathcal{H}(X)$ bits/símbolo, pode ser codificada em binário de tal forma que o comprimento médio do código, $\overline{N}$, é limitado por

$$\mathcal{H}(X) \leq \overline{N} \leq \mathcal{H}(X) + \epsilon$$

Na codificação por blocos está-se a fazer $\epsilon = \frac{1}{K}$.

- código ideal será aquele em que ε=0; na prática nem sempre é possível sendo satisfatório um código que possua bom rendimento

Fundamentos de Comunicação de Dados
Licenciatura em Engenharia Informática
Departamento de Informática, Universidade do Minho

I. TEORIA DA INFORMAÇÃO

Fontes com memória

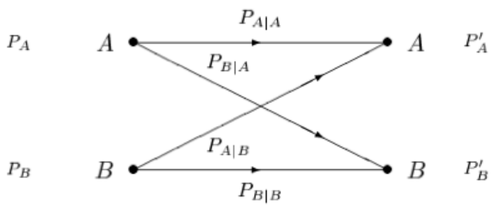
- Por vezes a probabilidade de emissão de um determinado símbolo depende dos símbolos anteriormente emitidos
- Fontes com memória de primeira ordem
  - fonte só se *lembra* do símbolo precedente
  - noção de probabilidade condicional
  - probabilidade de um símbolo ter ocorrido depois de um outro símbolo da fonte

I. TEORIA DA INFORMAÇÃO

Fontes com memória de primeira ordem

- **P (x_i | x_j)** - probabilidade de o símbolo x_i ser escolhido depois do símbolo x_j
- **P (x_i x_j)** - se for interpretado como a probabilidade da ocorrência de x_j e posteriormente x_i **:**

P (x_i x_j) = P(x_j) * P (x_i | x_j) ← ...para a construção da tabela de blocos de símbolos



$$
\begin{aligned}
P'_A &= P_A \cdot P_{A|A} + P_B \cdot P_{A|B} \\
P'_B &= P_A \cdot P_{B|A} + P_B \cdot P_{B|B}
\end{aligned}
$$

I. TEORIA DA INFORMAÇÃO

Fontes com memória

Como se calcula a entropia para fontes com memória de primeira ordem?

- Entropia condicional relativamente ao símbolo x_j

$$\boxed{\mathcal{H}(X|x_j) \stackrel{def}{=} \sum_{i=1}^{m} P(x_i|x_j)\, \log_2 \frac{1}{P(x_i|x_j)}}$$

- Entropia real de uma fonte de primeira ordem

$$\boxed{\mathcal{H}(X) = \sum_{j=1}^{m} P(x_j)\, \mathcal{H}(X|x_j)}$$

I. TEORIA DA INFORMAÇÃO

Fontes com memória

• Quando as probabilidades condicionais de uma fonte com memória reduzem significativamente o valor da entropia face ao seu valor máximo:

- *a fonte diz-se redundante*

- possibilidade de codificar a fonte com códigos mais eficientes (i.e. comprimento médio do código mais próximo da entropia real da fonte)

I. TEORIA DA INFORMAÇÃO

• Processos de **codificação da fonte** estudados no contexto da Teoria da Informação

- Levam em conta o **grau de incerteza** da fonte para tentar

- Tentam **retirar a redundância** produzida pela fonte

- Daí se designarem por mecanismos de **compressão da fonte**

I. TEORIA DA INFORMAÇÃO

- **Processos de Codificação da Fonte** (... algumas considerações adicionais)

  - Aplicabilidade em processos de transmissão de dados, em mecanismos/algoritmos de compressão, etc.

  - Códigos Shannon-Fano são métodos muito simples/ básicos de compressão (lossless)... foram abordados mais pela sua importância histórica no contexto da Teoria da Informação

  - Existem inúmeros mecanismos de compressão alternativos que assumem diferentes estratégias e com diferentes desempenhos

I. TEORIA DA INFORMAÇÃO

Mecanismos/algoritmos de compressão

Lossless

Algoritmos que permitem recuperar todos os dados originais a partir do ficheiro comprimido

Lossy

Algoritmos que só permitem recuperar uma aproximação aos dados originais

- Diferentes algoritmos de compressão são desenhados para lidar com determinados tipos de dados tendo em consideração algumas características específicas desses dados

I. TEORIA DA INFORMAÇÃO

Mecanismos/algoritmos/formatos de compressão (só alguns exemplos...)

Lossless *

General: RLE, LZ78, LZW, LZF, DEFLATE, bzip2, LZMA, Brotli, etc...

Audio: ALAC, ATRAC, DST, FLAC, RealPlayer, TTA, WavPack, WMA lossless, etc..

Graphics: PNG, TIFF, TGA, PCX, ILBM, JBIG2, etc...

Video: Dirac lossless, FFV1, H.264 lossless, etc...

Lossy *

Graphics: JPEG, JPEG2000, DjVu, JBIG2, PGF, etc...

Video: Motion JPEG, H.264/MPEG-4 AVC, Dirac, VC-1, H.265/HEVC, etc ...

Audio: AAC, ADPCM, ATRAC, Dolby Digital (AC-3), MP2, MP3, Musepack, Ogg Vorbis, WMA lossy, etc ...

Speech: Adaptive Multi-Rate, Codec2, Speex, etc ...

* alguns dos algoritmos/formatos mencionados tambem suportam a variante alternativa lossless/lossy

I. TEORIA DA INFORMAÇÃO

- Além da codificação da fonte a Teoria da Informação também aborda questões relacionadas:
  - Com o canal de comunicação.... e.g. **Capacidade do Canal**
  - Com a **Codificação do Canal** (iremos estudar mais tarde nesta unidade curricular)

I. TEORIA DA INFORMAÇÃO

Transmissão de Informação: o canal

(secção 8.4 da sebenta)

- Aborda a transmissão de informação em canais de comunicação
- *Não iremos abordar esta parte da matéria em detalhe...*
- .... mas iremos mais tarde utilizar a fórmula da Capacidade do Canal que é demonstrada nesta secção

Fundamentos de Comunicação de Dados
Licenciatura em Engenharia Informática
Departamento de Informática, Universidade do Minho

I. TEORIA DA INFORMAÇÃO

Transmissão de Informação: o canal

(breve referência)

- Capacidade do Canal - dependente de vários fatores, nomeadamente (conceitos mais tarde abordados):
  - potência do sinal (S)
  - potência do ruído (N)
  - banda de transmissão (B_T)
- Considerada uma das equações mais importantes no contexto da telecomunicações

I. TEORIA DA INFORMAÇÃO

Transmissão de Informação: o canal
(breve referência)

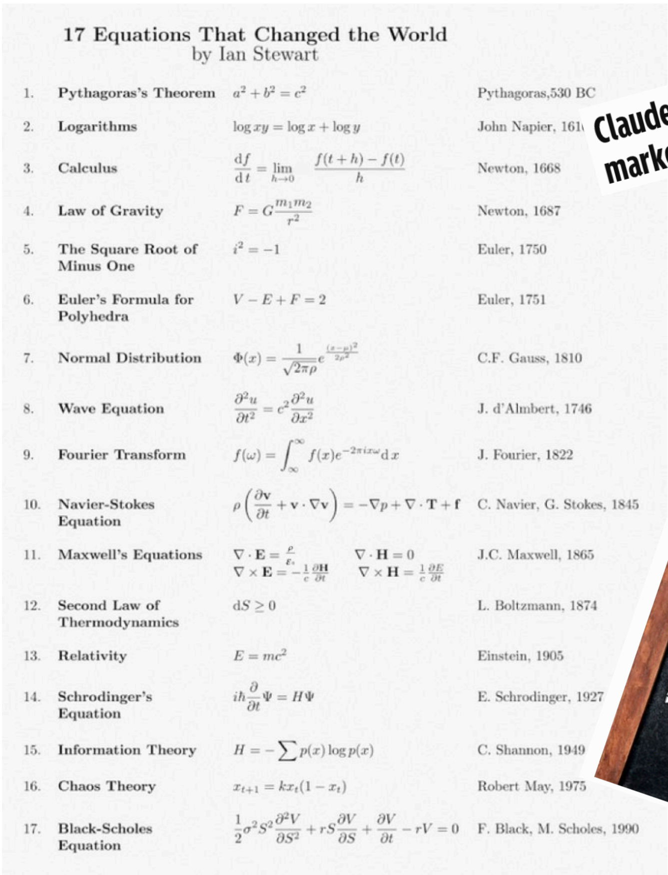
- Capacidade do Canal > .. mede informação ... <

$$C = B_T \log_2 \left(1 + \frac{S}{N}\right) \quad bits/s$$

I. TEORIA DA INFORMAÇÃO

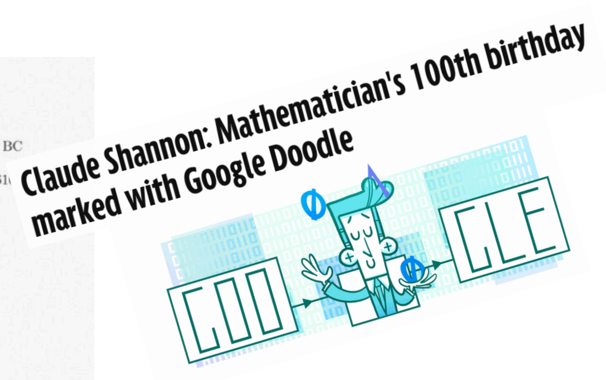
Teorema Fundamental da Teoria de informação

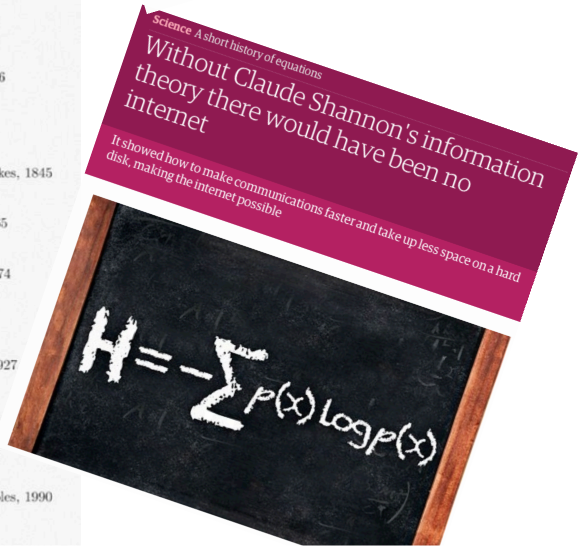
*"Dado um **canal de comunicação** e uma **fonte de informação** cujo **débito de informação** não excede a **capacidade do canal**, existe um **código** tal que a informação pode ser transmitida através do canal com uma frequência de erros arbitrariamente pequena, apesar da presença de ruído no canal."*



17 Equations That Changed the World
by Ian Stewart

1.	Pythagoras's Theorem	$a^2+b^2=c^2$	Pythagoras,530 BC
2.	Logarithms	$\log xy = \log x + \log y$	John Napier, 161
3.	Calculus	$\frac{df}{dt} = \lim_{h\to 0} \frac{f(t+h)-f(t)}{h}$	Newton, 1668
4.	Law of Gravity	$F = G\frac{m_1 m_2}{r^2}$	Newton, 1687
5.	The Square Root of Minus One	$i^2=-1$	Euler, 1750
6.	Euler's Formula for Polyhedra	$V-E+F=2$	Euler, 1751
7.	Normal Distribution	$\Phi(x) = \frac{1}{\sqrt{2\pi\rho}} e^{\frac{(x-\mu)^2}{2\rho^2}}$	C.F. Gauss, 1810
8.	Wave Equation	$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$	J. d'Almbert, 1746
9.	Fourier Transform	$f(\omega) = \int_{\infty}^{\infty} f(x)e^{-2\pi i x\omega} dx$	J. Fourier, 1822
10.	Navier-Stokes Equation	$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p + \nabla\cdot\mathbf{T} + \mathbf{f}$	C. Navier, G. Stokes, 1845
11.	Maxwell's Equations	$\nabla\cdot\mathbf{E} = \frac{\rho}{\epsilon_0}$, $\nabla\cdot\mathbf{H}=0$, $\nabla\times\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t}$, $\nabla\times\mathbf{H} = \frac{1}{c}\frac{\partial E}{\partial t}$	J.C. Maxwell, 1865
12.	Second Law of Thermodynamics	$dS \geq 0$	L. Boltzmann, 1874
13.	Relativity	$E=mc^2$	Einstein, 1905
14.	Schrodinger's Equation	$i\hbar\frac{\partial}{\partial t}\Psi = H\Psi$	E. Schrodinger, 1927
15.	Information Theory	$H = -\sum p(x)\log p(x)$	C. Shannon, 1949
16.	Chaos Theory	$x_{t+1} = kx_t(1-x_t)$	Robert May, 1975
17.	Black-Scholes Equation	$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} - rV = 0$	F. Black, M. Scholes, 1990





Correcção - p. 210:

A codificação por blocos conduz tendencialmente a um código óptimo, isto é, com $K \to \infty$ tem-se $\boxed{\overline{N}} \to \mathcal{H}(X)$, $\rho \to 1$ e $c \to c_{max}$. De facto, para a codificação por blocos, a desigualdade 8.13 escreve-se

$$\boxed{K * \mathcal{H}(X) \leq \overline{N}_K < K * \mathcal{H}(X) + 1}$$

donde, dividindo por K e tendo em atenção que a entropia da fonte não se altera com a codificação, se obtém

$$\mathcal{H}(X) \leq \frac{\overline{N}_K}{K} < \mathcal{H}(X) + \frac{1}{K}$$

ou, visto que $\overline{N} = \frac{\overline{N}_K}{K}$,

$$\mathcal{H}(X) \leq \overline{N} < \mathcal{H}(X) + \frac{1}{K}$$

Podemos agora enunciar um dos teoremas fundamentais da Teoria da Informação embora não procedamos à sua demonstração geral:

Correcção - p. 208:

Corrigir títulos da secção e exemplo:

Secção 8.2.3 – Códigos de *Shannon-Fano*

Exemplo 8.4 – Codificação de *Shannon-Fano*