# Teoria da Informação

**Charles Casimiro Cavalcante**

charles@gtel.ufc.br

Grupo de Pesquisa em Telecomunicações Sem Fio – GTEL
Programa de Pós-Graduação em Engenharia de Teleinformática
Universidade Federal do Ceará – UFC
http://www.gtel.ufc.br/∼charles

*"A principal função de um sistema de comunicação é reproduzir, exatamente ou de forma aproximada, uma informação proveniente de outro ponto diferente."*

Claude Shannon, 1948

# Conteúdo do curso

1. Revisão de probabilidade
2. Informação e Entropia
3. Codificação de fontes
4. Codificação e capacidade de canal
5. Complexidade de Kolmogorov
6. Funções de otimização
7. *Independent Component Analysis*

Parte V

# Introdução à Complexidade de Kolmogorov

# Considerações

- Kolmogorov, matemático russo, definiu em 1965 a **complexidade descritiva** de um objeto
- Até o momento, a descrição de uma variável aleatória $X$ com densidade de probabilidade $p_X(x)$
- Como $-\log\left[p_X(x)\right]$ define, segundo Shannon, a informação de $X = x$
- Complexidade descritiva de $X$ é $\lceil -\log\left[p_X(x)\right]\rceil$ para ser escrita por um código de Shannon
- Complexidade descritiva depende de $p_X(x)$!

# Considerações - cont.

- Kolmogorov definiu a complexidade algorítmica (descritiva) de um objeto como sendo o comprimento do menor programa binário que descreve o objeto
- Não utiliza a *densidade de probabilidade*
- Coincidentemente, o tamanho da menor representação binária é aproximandamente igual à entropia
- Complexidade algoritmica (descritiva) é um precursor da entropia

**Meta desta parte**

Fornecer fundamentos básicos para o entendimento da complexidade de Kolmogorov

# Máquina de Turing

- A partir de uma máquina $\mathcal{M}$, um programa $p$ e uma entrada $x$ calcula-se a saída $y$, ou seja,

$$\mathcal{M}_{p,x} = y \tag{115}$$

- $\mathcal{M}$ interpreta $p$ como uma descrição de $y$ na presença de uma informação lateral $x$
- Em outras palavras, $\mathcal{M}$ é um programa que transforma $x$ em $y$

# Definição

- Seja $x$ uma palavra binária de tamanho finito
- Seja $\mathcal{U}$ um computador universal
- $l(x)$ é o comprimento da palavra $x$
- $\mathcal{U}(p)$ denota a saída do computador $\mathcal{U}$ quando é apresentado o programa $p$

## Definição

A **complexidade de Kolmogorov** $K_{\mathcal{U}}(x)$ de uma palavra $x$ com relação a um computador universal $\mathcal{U}$ é definida por

$$K_{\mathcal{U}}(x) = \min_{p:\mathcal{U}(p)=x} l(p), \tag{116}$$

ou seja, o comprimento mínimo sobre todos os programas que fornecem $x$. Logo, $K_{\mathcal{U}}(x)$ é o menor comprimento da descrição de $x$ sobre todos os programas interpretados por $\mathcal{U}$

# Exemplo

- Seqüências (20 bits)
  11111111111111111111
  10110011010111011010

**Implementação da primeira**

```
FOR I=1 TO 20 PRINT 1
```

**Implementação da segunda**

```
PRINT 10110011010111011010
```

# Princípio de Occam

**William de Ockham (1285?-1349?)**

"...Entidades não devem ser multiplicadas desnecessariamente"

De uma forma análoga:

**Navalha de Occam**

Se há várias explicações igualmente válidas para um fato, então devemos escolher a mais simples

# Propriedades da complexidade e Kolmogorov

- Complexidade condicional: se o computador já conhece $l(x)$

$$K_{\mathcal{U}}(x|l(x)) = \min_{p:\mathcal{U}[p,l(x)]=x} l(p) \tag{117}$$

- Universalidade da complexidade de Kolmogorov: se $\mathcal{U}$ é um computador universal, então para qualquer outro computador $\mathcal{A}$

$$K_{\mathcal{U}}(x) \leq K_{\mathcal{A}}(x) + c_{\mathcal{A}} \tag{118}$$

- Complexidade condicional

$$K_{\mathcal{U}}(x|l(x)) \leq l(x) + c \tag{119}$$

- Limite superior

$$K_{\mathcal{U}}(x) \leq K_{\mathcal{U}}(x|l(x)) + 2\log\left[l(x)\right] + c \tag{120}$$