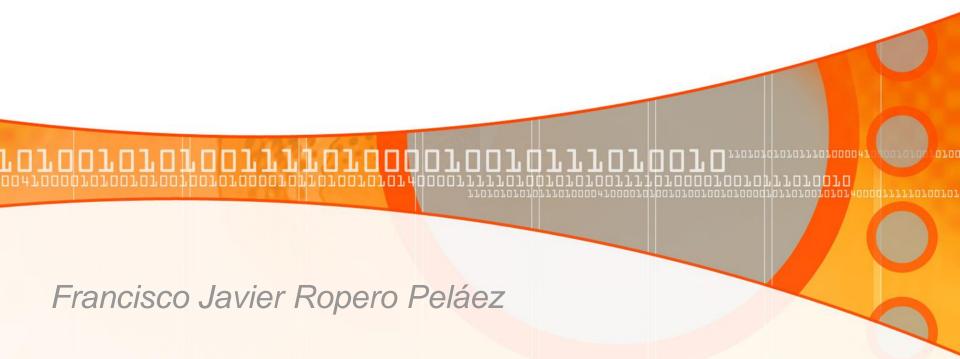
Algoritmos genéticos

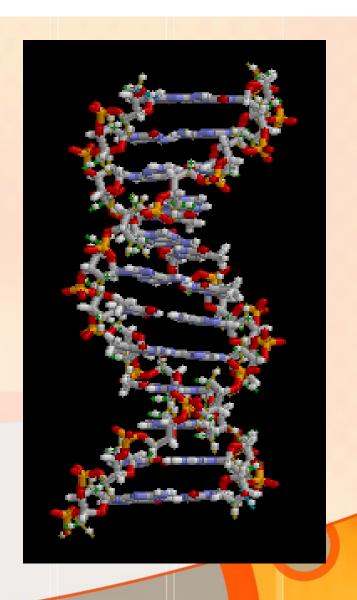


INTRODUÇÃO AO CROMOSSOMO/DNA

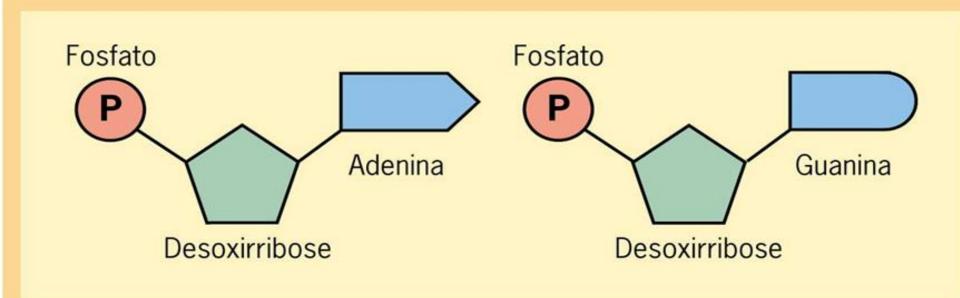


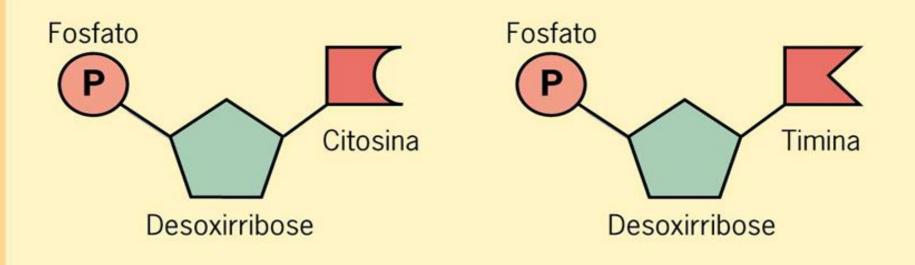
DNA (I)

- Watson e Crick descobriram a estrutura de dupla hélice do DNA
- No DNA aparecem uma seqüência de informações escritas com três letras: A,T,C,G (Adenina, timina, citosina e guanina)
- Cada três letras conformam uma palavra chamada códon.
- Cada "degrau" do DNA está formado por dois nucleotídeos
- Cada nucleotídeo é uma molécula formada por um açúcar, um fosfato e uma base nitrogenada.

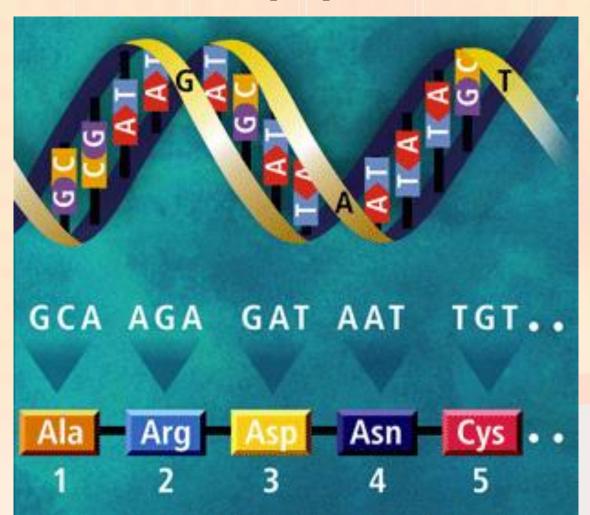


Os nucleotídeos do DNA





DNA (II)

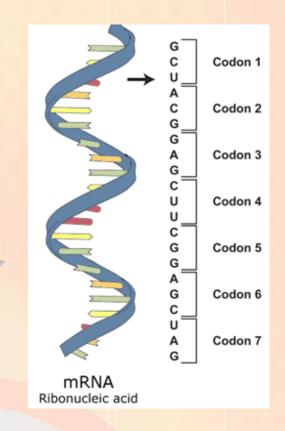


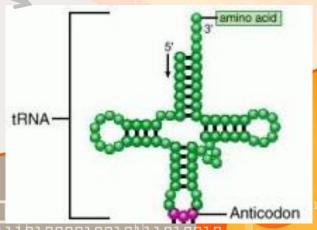
- Cada

 nucleotídeo se
 liga com o seu
 nucleotídeo
 complementar
- Assim a Adenina se liga com a
 Timina e a
 Citosina com a
 Guanina/

No DNA está a lista de instruções para a construção de um ser vivo

- As fábricas das proteínas dos seres vivos são os ribossomos.
- Para fabricar as proteínas a "fábrica" (os ribossomos) precisam:
- A) De uma cópia das instruções para fabricar a proteína (o chamado RNA mensageiro)
- B) Das peças (os aminoácidos) com as quais a proteína será construídas
- Ver:http://www.youtube.com/watch?v=NJ xobgkPEAo





Diferentes tamanhos de DNA

- Mycoplasma genitalium: 577.000 a 590.000 pares de bases
- Humano 3.156.105.057 pares de bases (bp)
- Pandora virus 2.470.000 bp





Main page Contents

Featured content

Current events

Random article

Donate to Wikipedia

Wikipedia store

Interaction

Help

About Wikipedia

Community portal Recent changes

Contact page

Tools

What links here Related changes Article Talk Read Edit View history Search

Billion

From Wikipedia, the free encyclopedia

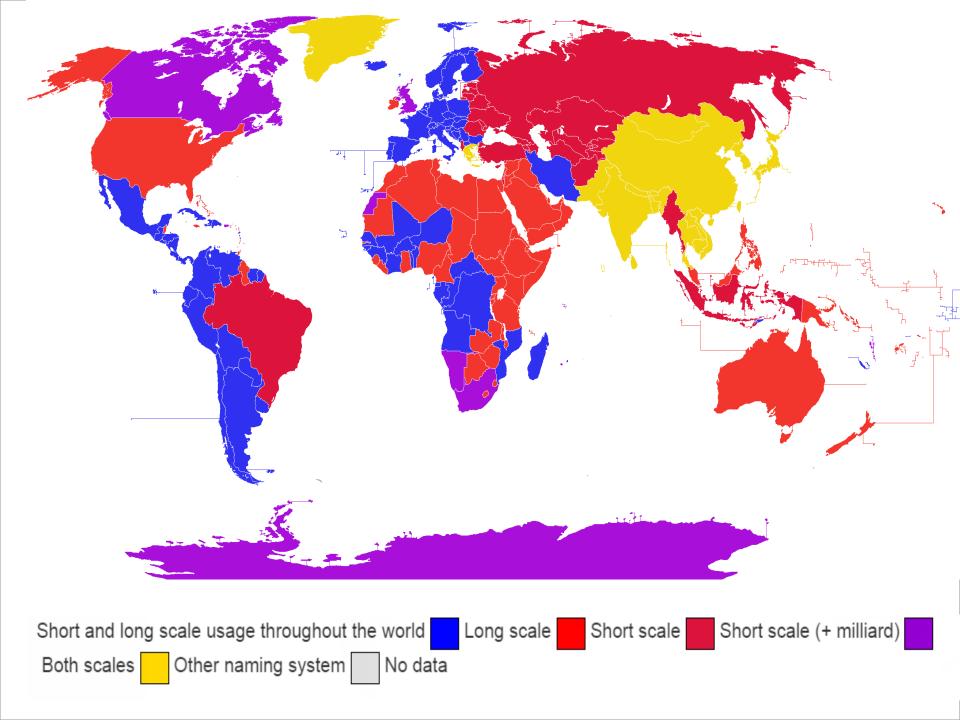
For other uses, see Billion (disambiguation).

A billion is a number with two distinct definitions:

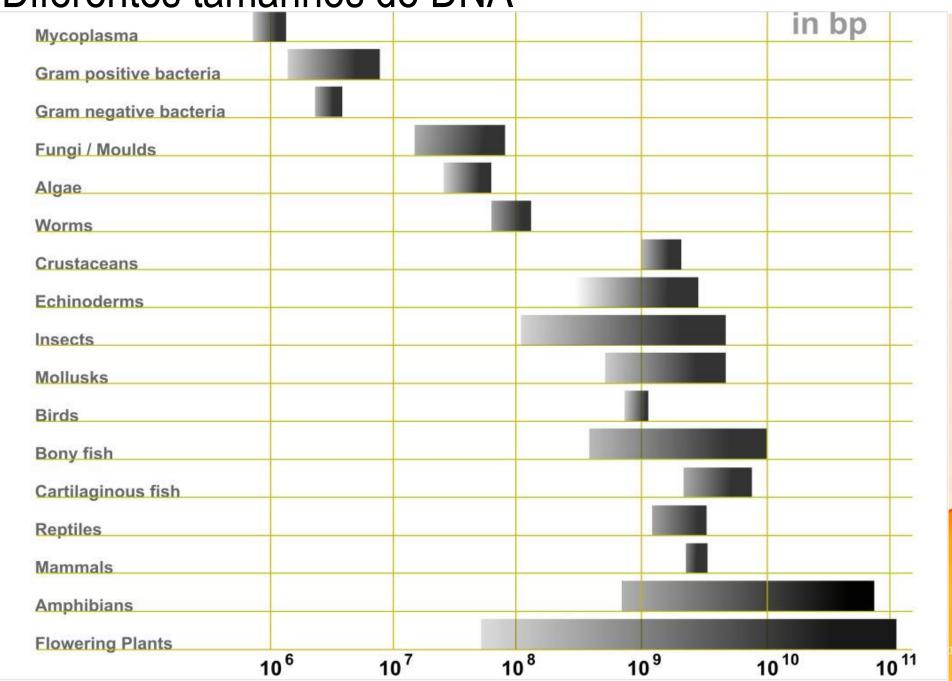
- 1,000,000,000, i.e. one thousand million, or 10⁹ (ten to the ninth power), as defined on the short scale. This is now generally the meaning in both British and American English.^{[1][2]}
- 1,000,000,000,000, i.e. one million million, or 10¹² (ten to the twelfth power), as defined on the long scale. This is one thousand times larger than the short scale billion, and equivalent to the short scale trillion.

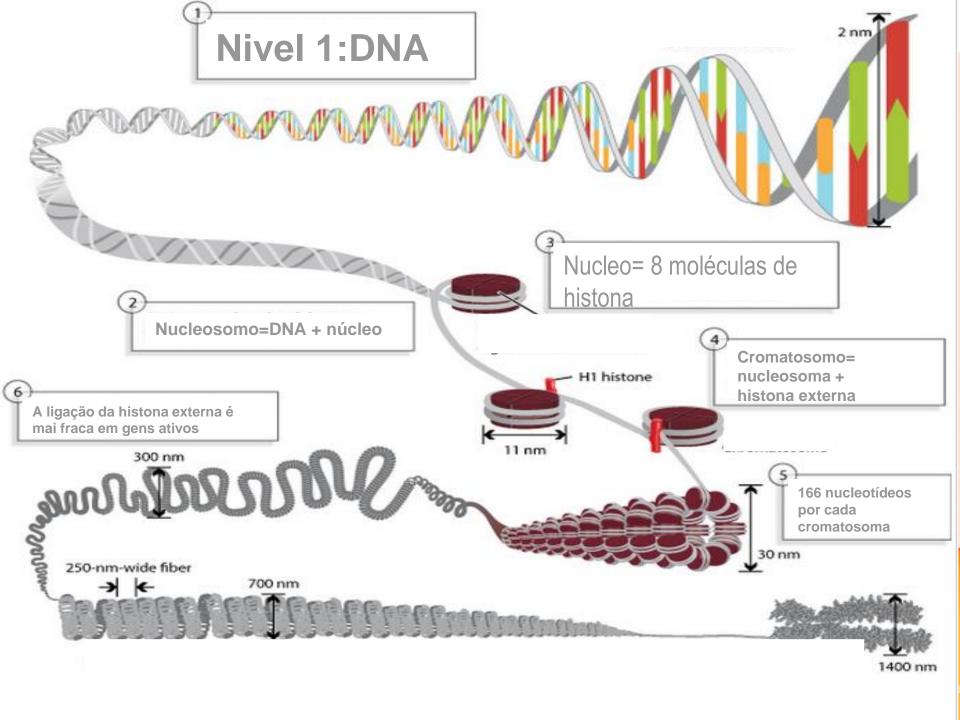
American English always uses the short scale definition but British English has employed both versions. Historically, the United Kingdom used the long scale billion but since 1974 official UK statistics have used the short scale. Since the 1950s the short scale has been increasingly used in technical writing and journalism, although the long scale definition still enjoys common usage.^[3]

Other countries use the word *billion* (or words cognate to it) to denote either the long scale or short scale billion. For details, see Long and short scales – Current usage.



Diferentes tamanhos de DNA





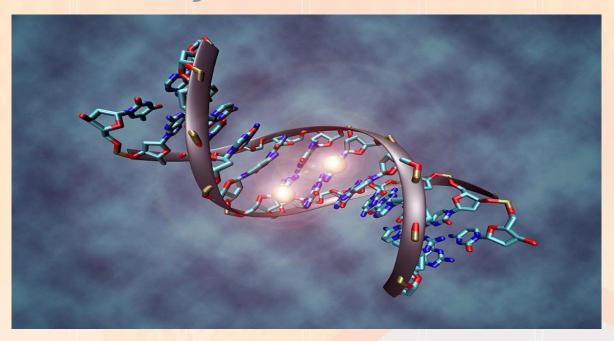
Epigenetica:

- No cromossomo existem informações ativas (cromatosomos fáceis de abrir).
- Também existem informações reservadas (em cromatosomos) difíceis de abrir.
- Eventos externos podem enfraquecer historias externas permitindo que possam ser lidas novas informações.
- Assim um evento externo pode que pais, filhos, netos, etc heredem uma nova caraterística.
- Exemplo: 11-S



Epigenetica (II):

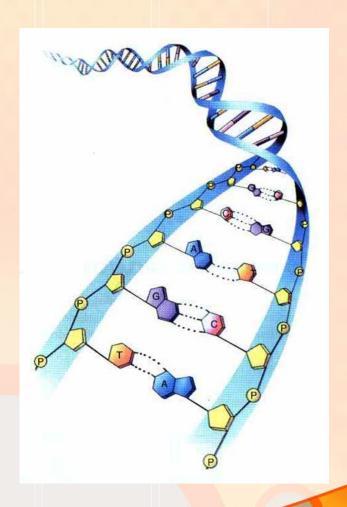
 Um outro método de desativar genes é pela metilação do DNA



 Além do enfraquecimento das histonas externas, a ativação e desativação de nucleotídeos específico pode acontecer por causa da metilação de bases nitrogenadas específicas.

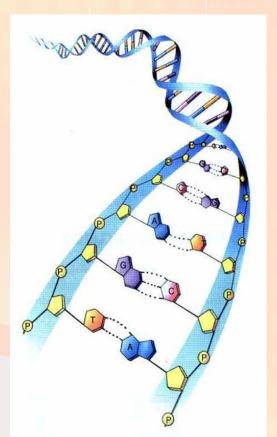
Introdução ao DNA, RNA

- No DNA encontrasse a informação genética
- É um código formado por quatro letras (que correspondem as quatro bases nitrogenadas): A,T,C,G
- A: Adenina, T: Timina, C: Citosina e G: Guanina
- Uma sequência de três destas letras pode nomear cada um dos 20 aminoácidos dos seres humanos mas o chamado STOP códon que fica no final de cada sequência.
- No RNA a timina é substituída por Uracila



Introdução ao DNA, RNA (II)

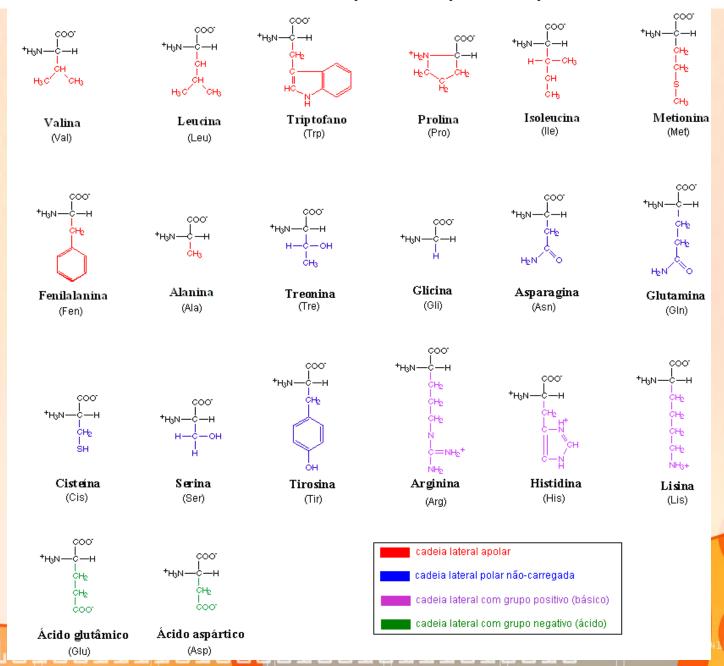
- Cada códon é uma palavra de três letras e identifica um aminoácido.
- Seria possível escrever 4x4x4=64 palavras de três letras, ou seja poderia haver 64 aminoácidos.
- Contudo existem vários codons "sinônimos" para cada aminoácido. Por exemplo, a arginina (ARG) pode ser codificada como: CGA, CGC, CGG, CGU, AGA, AGG
- Portanto existe muita redundância no código genético
- Os aminoácidos mais freqüentes onde podem acontecer mais erros têm maior redundância
- O STOP codon tem muita redundância (escreve-se como UAA, UAG, UCA) porque embora seja escasso é muito importante





Aminoácido (%)	Codificação	Aminoácido (%)	Codificação
ARG (5,28)	CGA,CGC,CGG, CGU,AGA,AGG	HIS (2,36)	CAC CAU
LEU (9,56)	CUA, CUC, CUG, CUU, UUA, UUG	GLU (6,82)	GAA, GAG
SER (7,25)	UCA, UCC, UCG, UCU, AGC, AGU	ASP (5,07)	GAC, GAU
THR (5,68)	ACA, ACC,ACG,ACU	TYR (3,13)	UAC, UAU
PRO (5,67)	CCA,CCC,CCG,CC U	CYS (2,44)	UGC, UGU
ALA (6,99)	GCA,GCC,GCG,GC U	PHE (3,84)	UUC, UUU
GLY (7,10)	GGA,GCC,GCG,G CU	ILE (4,50)	AUA,AUC, AUU
VAL (6,35)	GUA,GUC,GUG,GU U	MET (2,23)	AUG
LYS (5,71)	AAA, AAG	TRP (1,38)	UGG
ASN (3,92)	AAC, AAU	STOP (0,24)	UAA,UAG, UGA
GLN (4,47)	CAA, CAG		

Os vinte aminoácidos que compõe as proteínas



010

Aminoácido (%)	Codificação	Aminoácido (%)	Codificação
ARG (5,28)	CGA,CGC,CGG, CGU,AGA,AGG	HIS (2,36)	CAC CAU
LEU (9,56)	CUA, CUC, CUG, CUU, UUA, UUG	GLU (6,82)	GAA, GAG
SER (7,25)	UCA, UCC, UCG, UCU, AGC, AGU	ASP (5,07)	GAC, GAU
THR (5,68)	ACA, ACC,ACG,ACU	TYR (3,13)	UAC, UAU
PRO (5,67)	CCA,CCC,CCG,CC U	CYS (2,44)	UGC, UGU
ALA (6,99)	GCA,GCC,GCG,GC U	PHE (3,84)	UUC, UUU
GLY (7,10)	GGA,GCC,GCG,G CU	ILE (4,50)	AUA,AUC, AUU
VAL (6,35)	GUA,GUC,GUG,GU U	MET (2,23)	AUG
LYS (5,71)	AAA, AAG	TRP (1,38)	UGG
ASN (3,92)	AAC, AAU	STOP (0,24)	UAA,UAG, UGA
GLN (4,47)	CAA, CAG		

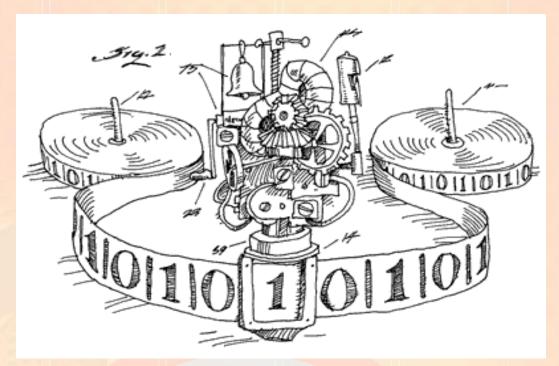
Ação das Enzimas

- ➤ Grupo de substâncias orgânicas de natureza normalmente protéica (existem também enzimas constituídas de RNA as ribozimas), com atividade intra ou extracelular que têm funções catalisadoras, catalisando reações químicas que, sem a sua presença, dificilmente aconteceriam.
- > As enzimas corrigem erros nas cadeias de DNA.
- Contribuem a diminuir o desordem, imprevisibilidade ou entropia da informação genética.
- > Para isto precisam de energia.
- A nível computacional o DNA equivaleria á fita da máquina de Turing, e o enzima equivaleria a cabeça de leitura e escritura da máquina



A Máquina de Turing o primeiro computador teórico:

- Consistia numa máquina que se mexia por uma fita sem fim.
- A máquina lia, escrevia, e apagava símbolos na fita.
- A máquina de Turing possuía "computabilidade universal" ou seja poderia simular a a lógica de qualquer algoritmo computacional.

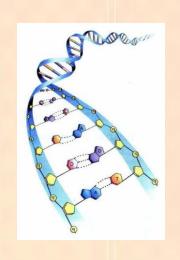


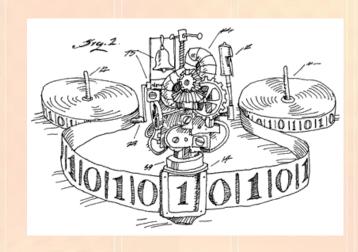


Exemplo de programa da máquina de Turing para contagem decimal

Instruction I _i :	If machine reads:	Should write:	Then shift machine:	And go to instruction Ii
l ₁	Nothing	Nothing	Leftwards	l ₂
	1	1	Rightwards	l ₁
	2	2	Rightwards	I 1
	3	3	Rightwards	I 1
	4	4	Rightwards	I 1
	5	5	Rightwards	l ₁
	6	6	Rightwards	I 1
	7	7	Rightwards	I 1
	8	8	Rightwards	l ₁
	9	9	Rightwards	l ₁
	0	0	Rightwards	I 1
12	Nothing	1	Rightwards	l ₁
	0	1	Rightwards	I 1
	1	2	Rightwards	l ₁
	2	3	Rightwards	l ₁
	3	4	Rightwards	l ₁
	4	5	Rightwards	l ₁
	5	6	Rightwards	l ₁
	6	7	Rightwards	l ₁
	7	8	Rightwards	I ₁
	8	9	Rightwards	l ₁
	9	0	Leftwards	l ₂

DNA e Máquina de Turing







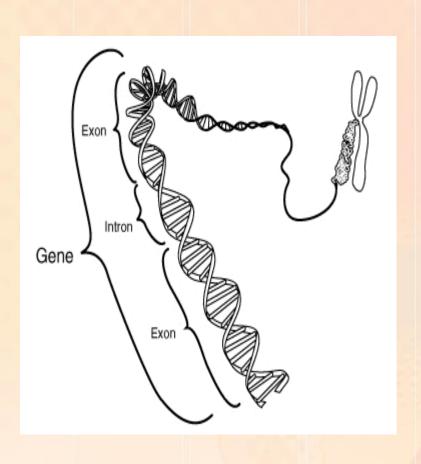
- A fita sem fim de Turing têm certa analogia com a fita do nosso DNA.
- A própria máquina de Turing e semelhante as enzimas que percorrem o DNA lendo, apagando e escrevendo informações no DNA.

O DNA é robusto

- Existem várias razões que impedem a degradação da informação genética contida no DNA:
- No DNA a Adenina liga com o complementar (Timina) e a Citosina com o complementar (Guanina) e vice versa.
- Se falta um nucleotídeo num degrau do DNA, olhando o nucleotídeo complementar podemos saber qual falta.
- Por outro lado as enzimas, se movimentando toda hora por cima do DNA, vão reparando os defeitos do código.

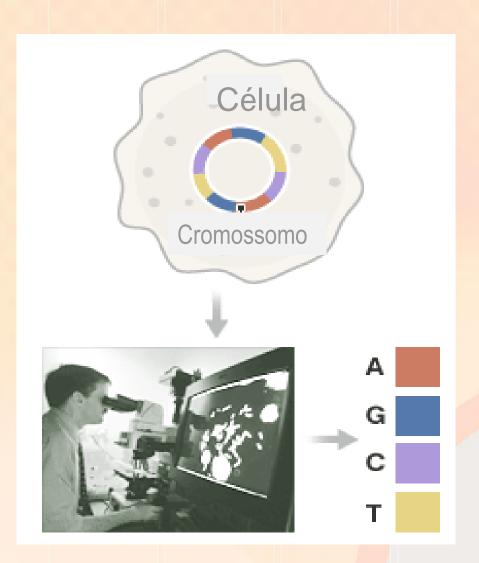


O Gene e a informação genética

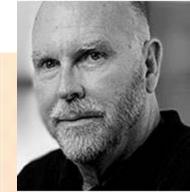


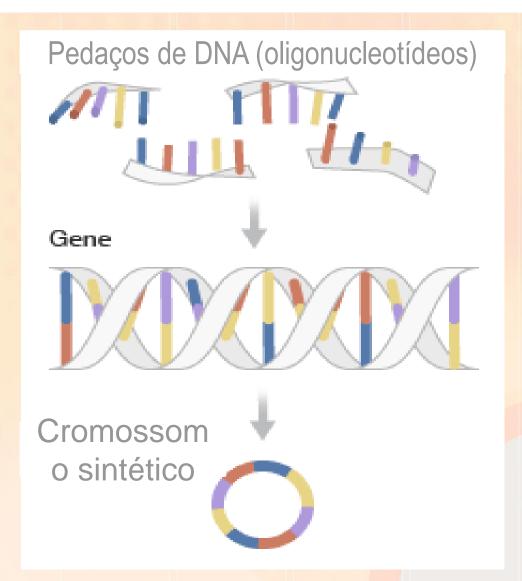
Segmento de um cromossomo a que corresponde um código distinto, uma informação para produzir uma determinada proteína ou para controlar uma característica, por exemplo, a cor dos olhos.

Craig Venter e o cromossomo sintético



Recentemente (20/5/2010), Craig Venter e colaboradores desenharam uma fita de ADN inteiramente no computador (ADN sintético)

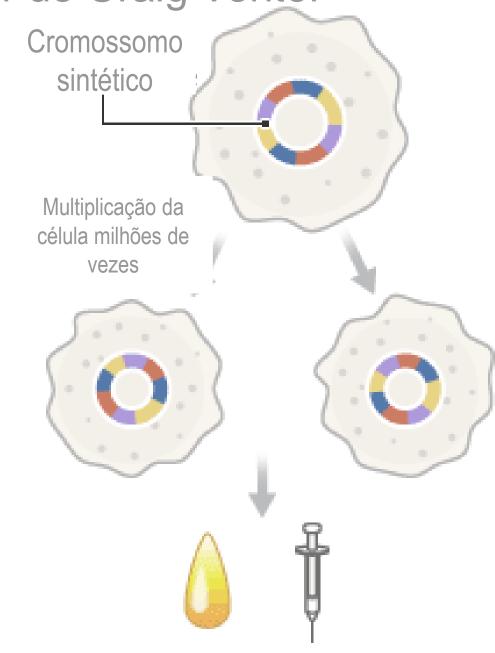




A sequência programada no computador foi implementada Interligando oligonucleotídeos

A célula: o computador de Craig Venter

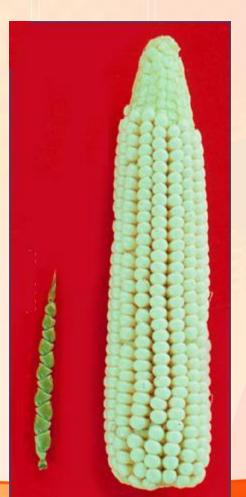
- A célula com o cromossomo sintético foi capaz de se replicar milhões de vezes.
- ➤ Uma mudança foi feita no DNA para que todas as "células" fossem azuis.
- Segundo Craig Venter (Maio 2010), "a fita de DNA sintético pode ser comparada ao programa de um computador, sendo a célula o próprio computador".



Utilização dos princípios da genética pelo ser humano

- O ser humano vem aproveitando as leis da evolução genética desde milênios para o próprio beneficio.
- Exemplo: Domesticação de animais e plantas.
- Atualmente a informação genética é muito importante para conhecer e enfrentar as enfermidades.
- Estudaremos os chamados algoritmos genéticos: técnicas computacionais para otimização de processos.

 Desde o início da civilização o ser humano utilizou os princípios da seleção natural para o próprio beneficio (Exemplo: obtenção de animais domesticados e novos tipos de plantas)

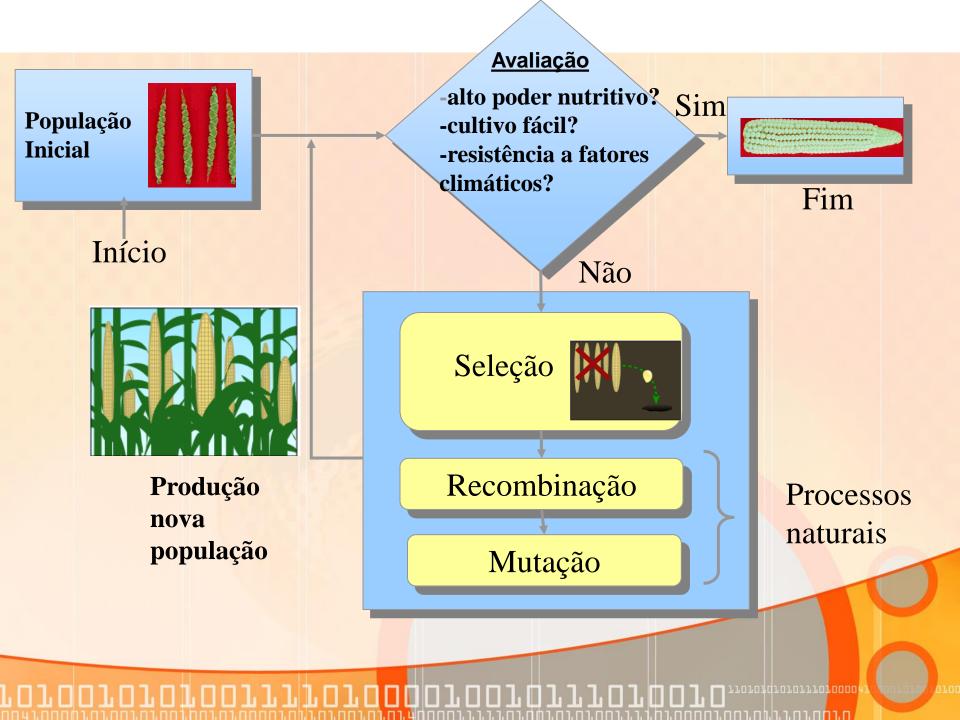


- Exemplo:
 Desenvolvimento do milho a partir do teosinto (4,000 a.C.)
- Alguns dos princípios presentes no desenvolvimento do milho aparecem nos algoritmos genéticos.

Introdução aos algoritmos genéticos:

- Técnica de otimização baseada nos processos de seleção natural
- Os melhores indivíduos (segundo determinados critérios) são selecionados a partir de população inicial.
- Esses indivíduos são re-combinados (cruzados).
- Uma nova geração é produzida.
- È feita uma nova avaliação para encontrar os melhores indivíduos segundo os critérios especificados.

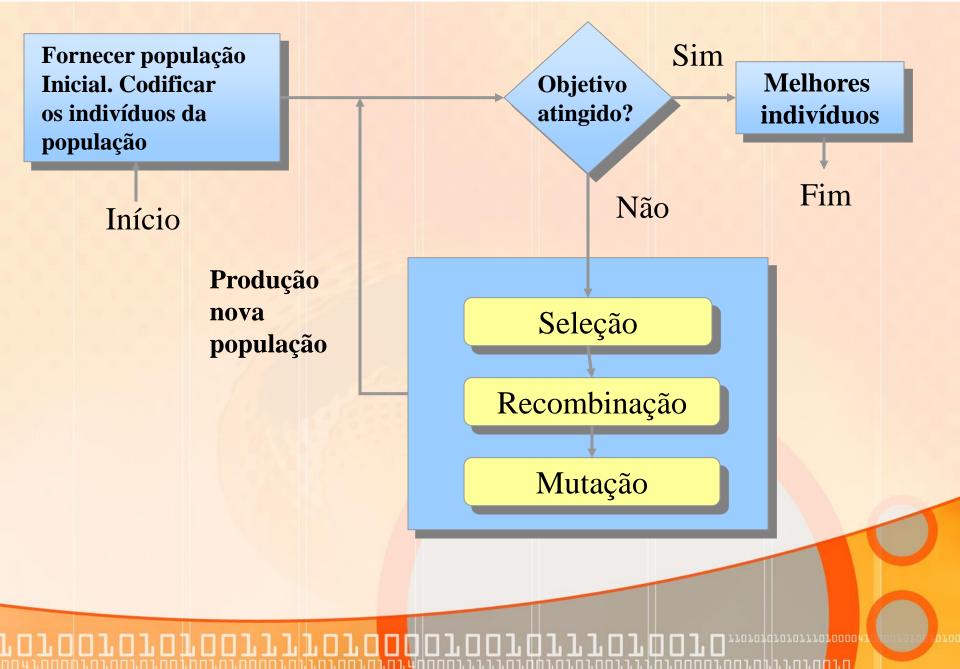




Introdução aos algoritmos genéticos:

- Técnica de otimização baseada nos processos de seleção natural.
- Os melhores indivíduos (segundo determinados critérios) são selecionados a partir de população inicial.
- Esses indivíduos são re-combinados (cruzados).
- Uma nova geração é produzida.
- È feita uma nova avaliação para encontrar os melhores indivíduos segundo os critérios especificados.

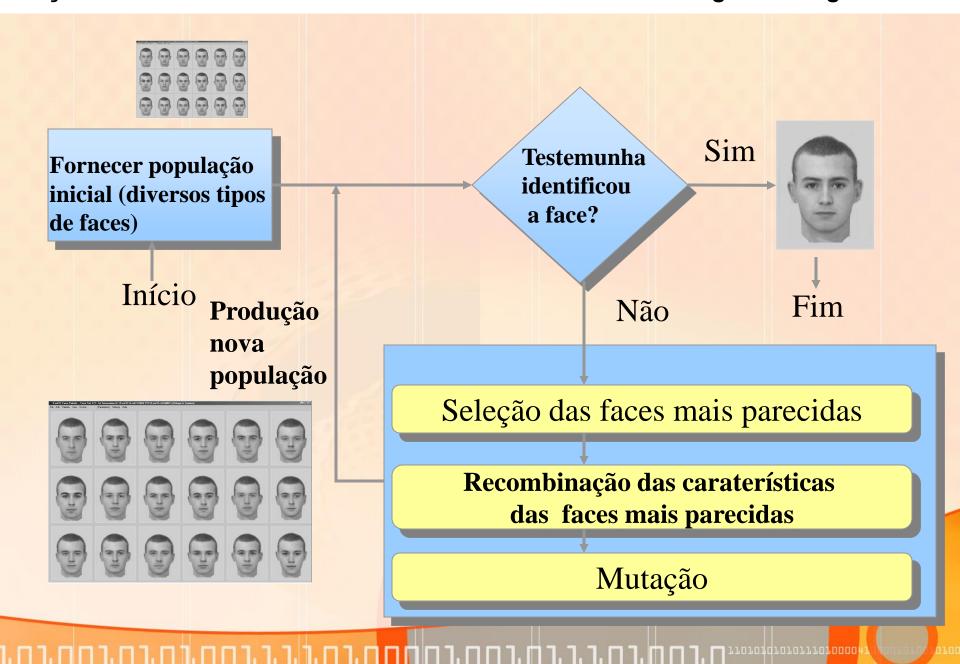
Estrutura de um algoritmo genético



Algumas aplicações dos algoritmos genéticos

- · Ajuste dos pesos em redes neurais artificiais.
- Problema do caixeiro viajante.
- Desenho de hardware ("evolvable hardware")
- Otimização de redes elétricas ou hidráulicas.
- Identificação facial de criminosos.

Criação de retrato falado de criminosos usándo sistemas algoritmos genéticos



Etapas nos algoritmos genéticos

- Codificação indivíduos.
- Avaliação dos indivíduos (função de ajuste)
- Seleção dos indivíduos que vão produzir a seguinte geração.
- Recombinação.
- Mutação



<u>Codificação</u>

- Cada indivíduo é identificado por uma seqüência de carateres (binária ou não) chamada cromossomo. Cada carater é chamado de alelo.
- A sugestão (Lawrence Davis) é usar a codificação que seja mais natural para o problema e que possa ser utilizada no algoritmo genético.
- Exemplo1: Se o problema é encontrar o máximo de uma função de tres variaveis f(x,y, z) cada variavel pode ser um número de 12 digitos binários. O cromosomo teria 36 digitos binarios.
- Exemplo 3: No problema do caixeiro viajante cada cidade estaria representada por uma letra. O cromossomo de um possível itinerário seria representado por uma sequencia de letras: GDAFHEBC

Avaliação dos indivíduos (função objetivo)

- Esta função $V_i=g(x_i)$ é definida pelo usuario dependendo da aplicação (x_i) é cada indivíduo)
- Os V_i são normalizados :

$$v_i = \frac{V_i}{\sum_{i=1}^{N} V_i}$$

Seleção dos indivíduos do grupo de acasalamento.

- Ao invés de selecionar somente os indivíduos com maior v_i, o processo de seleção permite escolher outros indivíduos para preservar a variabilidade populacional e evitar a convergência prematura.
- Para isto o v_i é transformado num valor de probabilidade de escolha.
- Existem vários métodos para selecionar os indivíduos do grupo de acasalamento.

roroorororororo

Métodos usados para seleção dos indivíduos do grupo de acasalamento.

Individu o nº	V _i (de maior para menor)	Soma
4	0,5	0,5
1	0,3	0,8
2	0,15	0,95
3	0.03	0,98
5	0,02	1

- Método da roleta
- Amostragem estocástico universal
- Método do "ranking"
- Uma vez selecionado o grupo de acasalamento selecionasse casais aleatoriamente dentro de grupo até esgotar os indivíduos do grupo.

<u>Recombinação</u>

- Pedaços dos cromossomos de cada casal são intercambiadas para produzir a nova geração (dois filhos por casal)
- Existem diversos métodos de recombinação ou cruzamento: cruzamento de um ponto, de dois pontos, uniforme, recombinação intermediária, PMX (partially matched crossover), etc.
- A eleição do tipo de recombinação depende do tipo de problema

Métodos de recombinação

- · Cruzamento de um ponto.
- Cruzamento de dois pontos.
- Cruzamento uniforme.
- Recombinação intermediária.
- PMX (partially matched crossover),

<u>Mutação</u>

- Evita convergência prematura em direção para um mínimo local.
- Consiste na mudança aleatória de um dos alelos do cromossomo.
- A probabilidade de mutação é normalmente muito pequena p_m=0.001

Referências:

Genetic Algorithms – A tutorial A.A.R. Townsend

http://citeseerx.ist.psu.edu/viewdoc/download? doi=10.1.1.184.3148&rep=rep1&type=pdf