



Universidade de São Paulo

Instituto de Matemática e Estatística

Algoritmos genéticos

Um algoritmo evolucionário para otimização

Aluno: Fabricio Kassardjian
Orientador: Anatoli Iambartsev



SUMÁRIO

1. Introdução

2. Algoritmos Genéticos

3. Implementação do Algoritmo e Resultados

a. Teste inicial

b. Modelo de Ising

4. Conclusões



1. INTRODUÇÃO

1A. ALGORITMOS EVOLUCIONÁRIOS

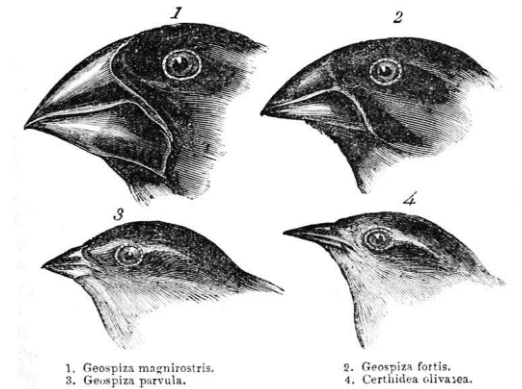
Computação Evolutiva

Baseada em processos encontrados na natureza de seleção, reprodução e mutação

Otimizar problemas

$$f : \Omega \rightarrow \mathbb{R}$$

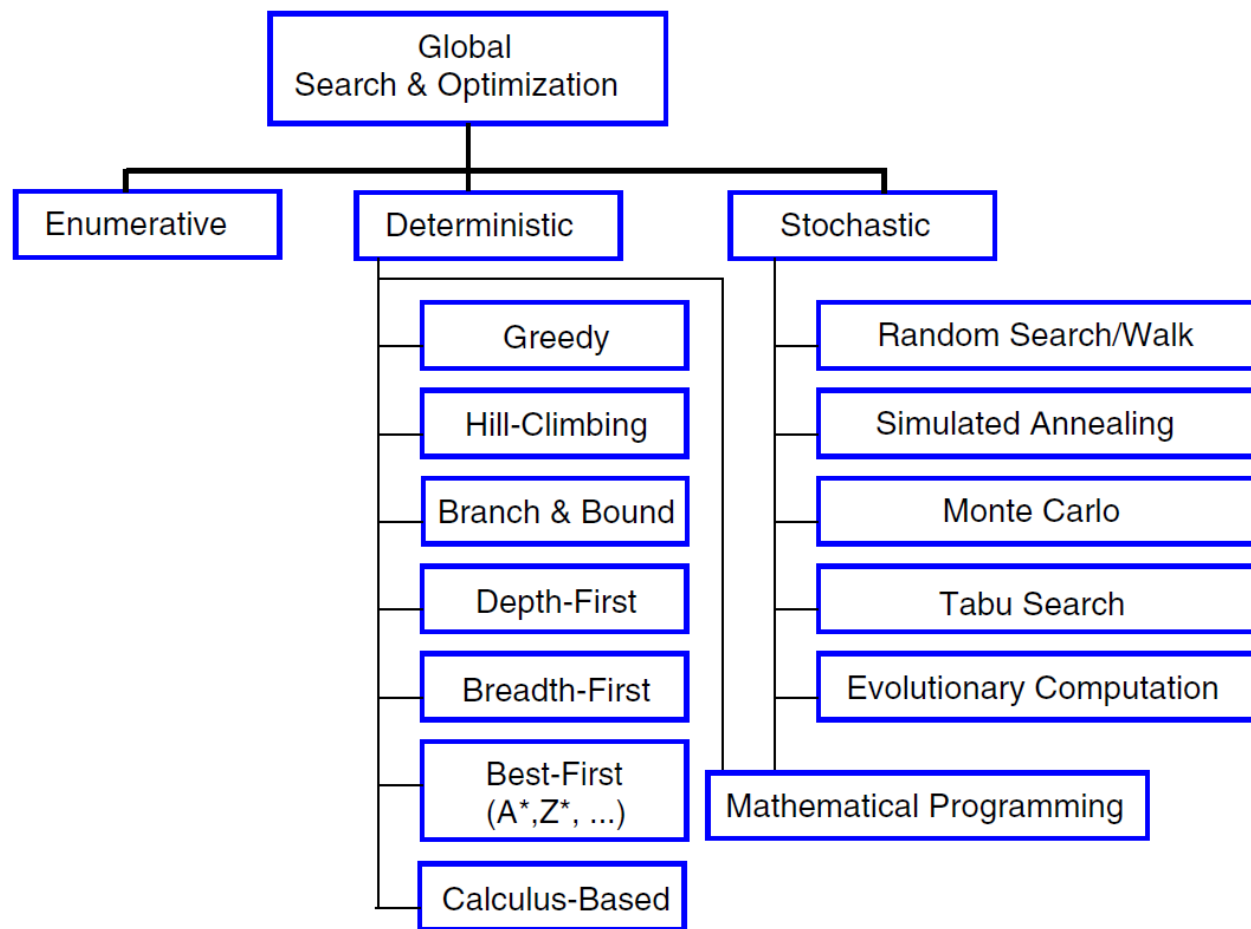
$$\arg \min_{x \in \Omega} f(x)$$



Créditos da imagem: "[Darwin's finches](#)," por John Gould (domínio público).



Classificação dos métodos





Biologia

Evolução das espécies / seleção natural – **Darwin**

Genética, compartilhamento de características entre os pais e descendentes – **Mendel**

Unidade de informação: **Gene**

Conjunto de genes: **Cromossomo**

Posição do gene no cromossomo: **Locus**

Valores possíveis de para um gene: **Alelos**

Interação entre genes: **Epistasia**

Conjunto específico de genes: **genótipo**

Características expressas pelo genótipo e ambiente: **fenótipo**



Genótipo

+



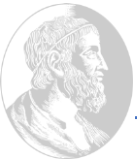
Meio ambiente

=



Fenótipo

<https://www.diferenca.com/genotipo-e-fenotipo/>



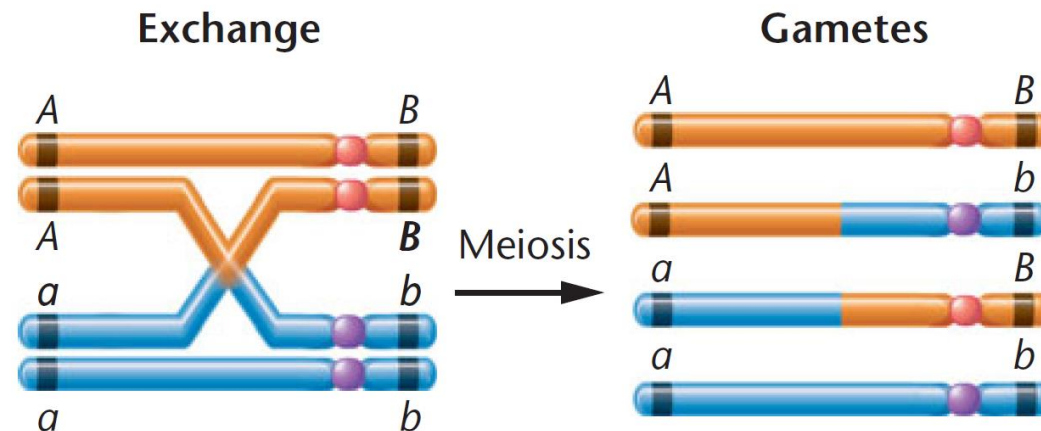
Reprodução

Organismos podem ser:

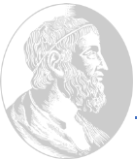
- Diplóides – com par de cromossomos, conceito de dominante-recessivo
- Haplóides – cromossomo simples

Tipo de reprodução:

- Assexuada – duplicação
- Sexuada – combinação de material genético de dois indivíduos + recombinação do material por *crossover*



Exemplo de reprodução com crossover - (KLUG et al., 2011)



Breve história...

Década de 40: inspiração na natureza para primeiros passos em IA

Década de 50: Sistemas adaptativos para solução de problemas difíceis

Década de 60: Rechenberg – estratégias evolucionárias – cromossomos com números reais para estudos de aerofólios

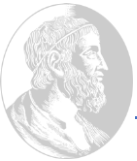
Décadas de 60 e 70: Holland – consolida os algoritmos genéticos

- Cria abstração da evolução biológica através de uma estrutura formal teórica
- Usou cromossomos binários: Alelos 1 ou 0
- Operadores de crossover, mutação e inversão



Algoritmo Genético por Holland

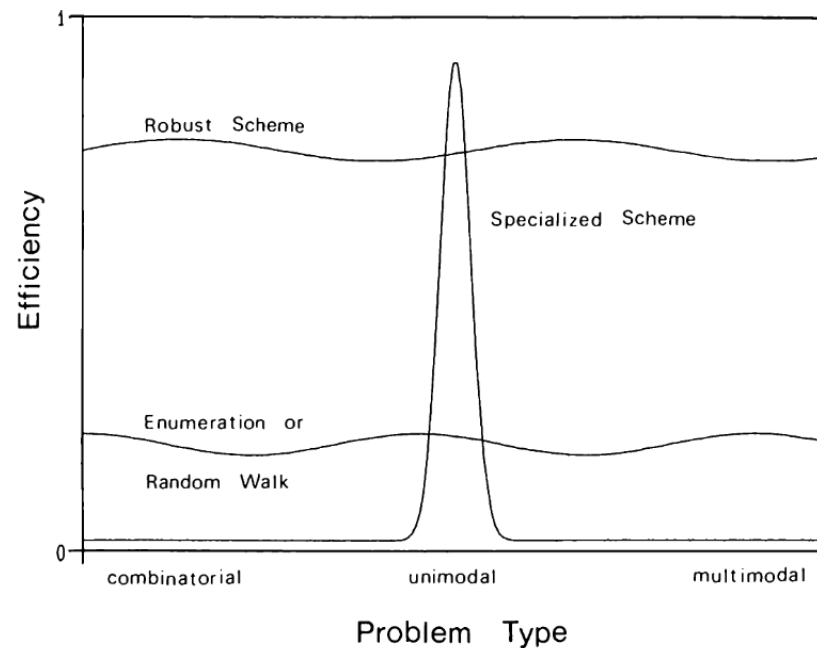
1. Seleção de um cromossomo na população de forma estocástica baseada nas avaliações de todos os cromossomos
2. Aplicações dos operadores genéticos sobre uma cópia do indivíduo selecionado em 1.
3. Seleção de outro cromossomo de forma aleatória com probabilidade igual para todos a ser substituído pelo novo cromossomo gerado em 2
4. Avaliar o novo cromossomo
5. Retonar ao 1



Teorema da inexistência de almoço grátis

Wolpert – Não existência de algoritmo universal para solução de problemas de otimização

Se houver um método específico para determinado problema definitivamente ele será mais eficiente que o GA.



Exemplo para NFL - ([GOLDBERG, 1989](#))



SUMÁRIO

1. Introdução

2. Algoritmos Genéticos

3. Implementação do Algoritmo e Resultados

a. Teste inicial

b. Modelo de Ising

4. Conclusões



Biologia x Algoritmo

- Cromossomo
- Gene
- Alelos
- *Locus*
- Genótipo
- Fenótipo
- Indivíduo / Solução candidata
- Codificação de valores, característica
- Valores possíveis para o gene
- Posição
- Estrutura com os valores
- Conjunto de parâmetros do problema

Todo algoritmo irá conter:

- Codificação
- População, função de avaliação e seleção
- Operadores genéticos



Esquemas

Usando o cromossomo com comprimento l e composto apenas por valores binários $\mathcal{A} = \{0, 1\}$

Define-se esquema H como uma máscara ou *template* com comprimento l e valores possíveis no conjunto $\mathcal{A}^+ = \{0, 1, *\}$

Exemplo com $l = 8$:

$$H = * * 11 * 0 * 1$$

Instâncias são cromossomos que se encaixam no esquema e formam o conjunto definido por $M(H)$

Exemplos de instâncias para H acima:

$$A_1 = 00110001 \text{ e } A_2 = 10111011$$



Esquemas

Ordem - $o(H)$: definido pelo quantidade de posições com valores definidos

$$o(* * 11 * 0 * 1) = 4$$

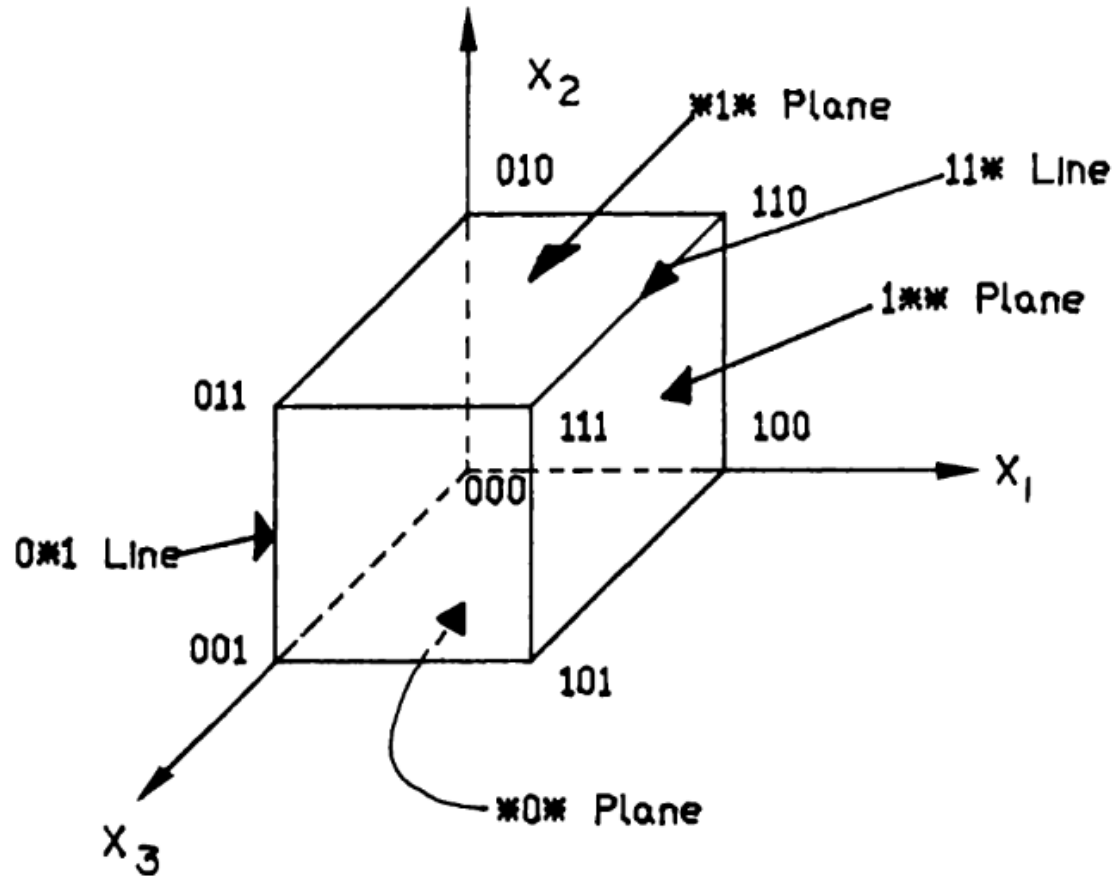
Tamanho - $\delta(H)$: definido pelo valor da ultima posição com valor definido menos a primeira com valor definido

$$\delta(* * 11 * 0 * 1) = 5$$



Esquemas

Representação em hiperplanos feita por Goldberg para $l = 3$



(GOLDBERG, 1989)



Codificação

Define como os parâmetros ou variáveis do problema serão representadas no algoritmo. Impor restrições quando possível

Mais básico é o binário: $\mathcal{A} = \{0, 1\}$

Podem ser de outros tipos:

- Com valores reais
- Definidos por ordem de ocorrência
- Grafos



População

Conjunto de soluções candidatas do problema

Tamanho é parâmetro do algoritmo

Exploitation (Aproveitamento)

testar uma região limitada mas promissora do espaço de soluções com a expectativa de melhorar uma solução já conhecida dentro dessa região

Exploration (Exploração)

testar uma região muito mais ampla do espaço de soluções, com a expectativa de encontrar novas soluções promissoras



População

Convergência genética

Elitismo

Estratégia $\mu + \lambda$

Steady state

Tamanho variável: tempo de vida e diversidade

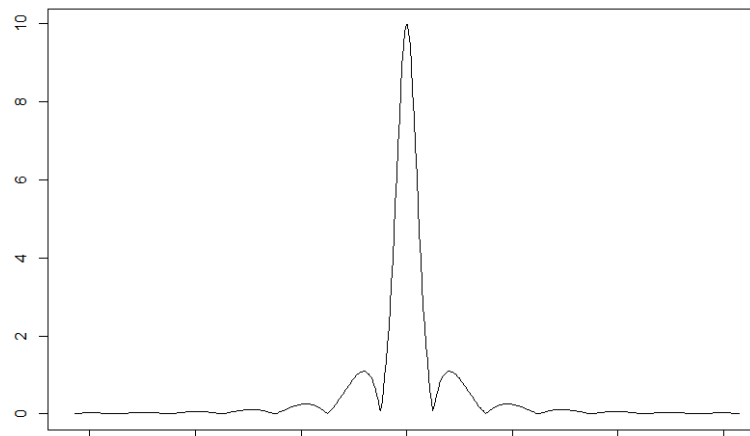
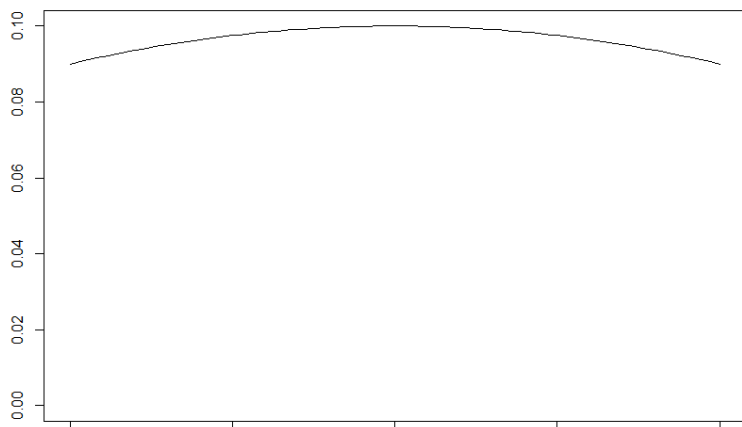
População inicial aleatória ou com espaço dividido



Avaliação

Função de avaliação:

$$f : \Omega \rightarrow \mathbb{R}^+$$



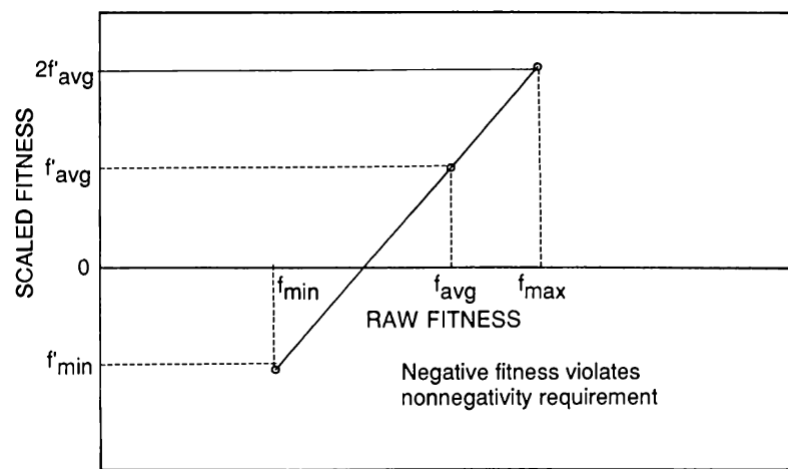
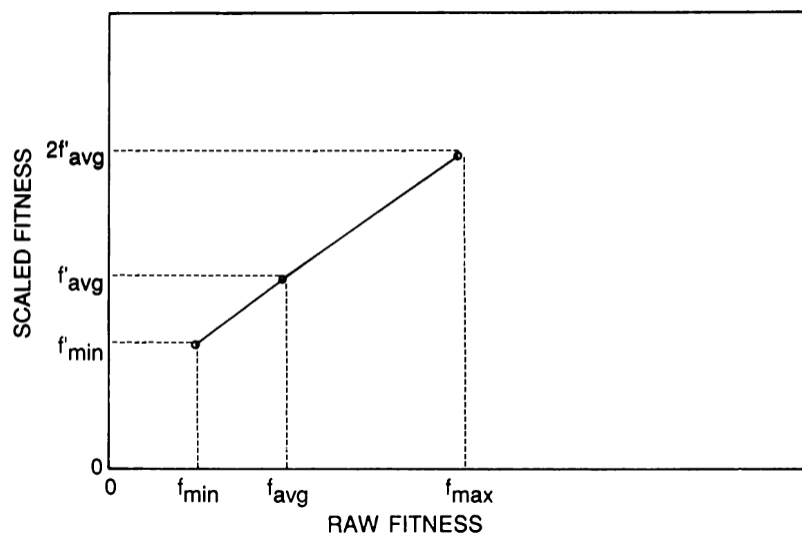
Define probabilidade de escolha:

$$p = \frac{f_i}{\sum f}$$



Variações na avaliação

normalização linear $f' = a * f + b$



(GOLDBERG, 1989)

Escalonamento sigma
$$E[i, t] = \begin{cases} 1 + \frac{f(i) - \bar{f}(t)}{2\sigma(t)} & , \text{ se } \sigma(t) \neq 0 \\ 1.0 & , \text{ se } \sigma(t) = 0 \end{cases}$$

Penalização de restrições do espaço de soluções



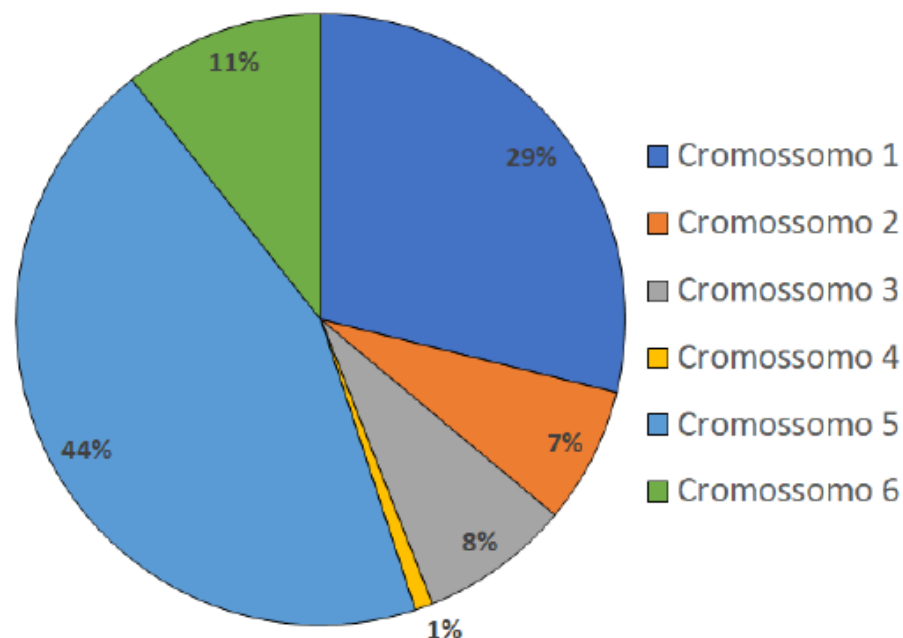
Seleção – Roleta Viciada

Sorteio de cromossomos para reprodução

Roleta Viciada:

$$p(i) = f(i) / \sum_{j=0}^n f(j)$$

	f	p
Cromossomo 1	152	28,84%
Cromossomo 2	38	7,21%
Cromossomo 3	42	7,97%
Cromossomo 4	5	0,95%
Cromossomo 5	234	44,40%
Cromossomo 6	56	10,63%
Total	527	100,00%





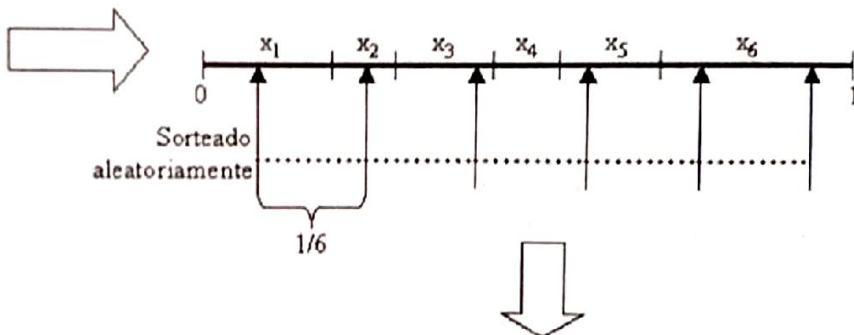
Seleção – Amostragem Estocástica Uniforme

Variação da roleta viciada

Sorteia um valor uma única vez

Indivíduos alinhados em uma reta continua, com segmentos proporcionais a sua probabilidade de seleção

Indivíduo	Fitness
x_1	200
x_2	100
x_3	150
x_4	100
x_5	150
x_6	300



Indivíduos sorteados: $x_1, x_2, x_3, x_5, x_6, x_6$

$$[\lfloor E[i, t] \rfloor, \lceil E[i, t] \rceil]$$

$$E[i, t] = \frac{f_i}{\sum f} \cdot n$$



Seleção – Torneio e *Ranking*

Torneio: Disputa entre dois ou mais cromossomos selecionados de maneira uniforme na população

Probabilidade de seleção do pior indivíduo: $1/N^k$

Ranking: Os cromossomos são ordenados e classificados de 1 a N

$$E[i, t] = \min + (\max - \min) \frac{\text{rank}(i, t) - 1}{N - 1}$$

Seguindo as restrições de $1 \leq \max \leq 2$ e $\min = 2 - \max$

Baker sugere usar o valor $\max = 1, 1$



Seleção de Boltzmann

Inspirado no resfriamento simulado (*simulated annealing*)

$$E[i, t] = \frac{e^{f_i/T}}{\frac{\sum_j e^{f_j/T}}{N}}$$

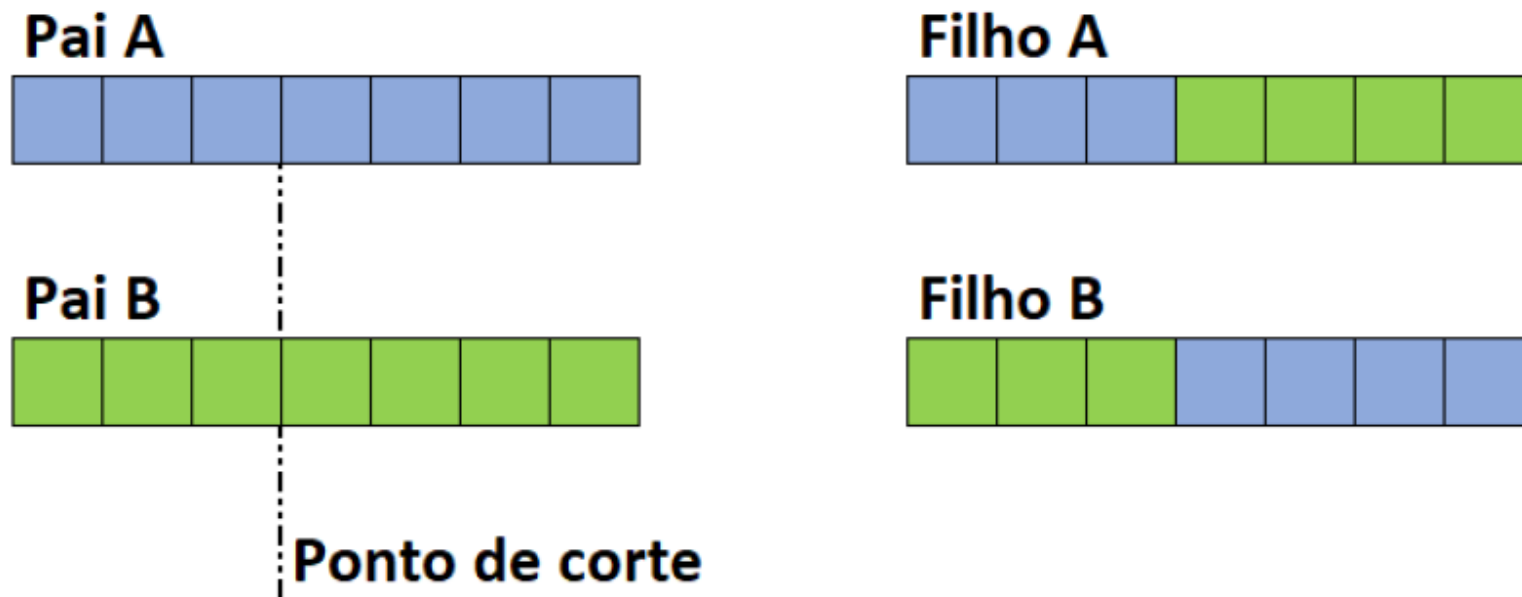


Crossover – 1 ponto

Responsável pela recombinação dos genes entre os pais e é o mais característico do GA

Parâmetro associado a aplicação do operador: p_c

Crossover de 1 ponto, define apenas um ponto de corte





Crossover de 2 pontos ou mais pontos

Sorteados dois ou mais pontos de corte no cromossomo

Pai A



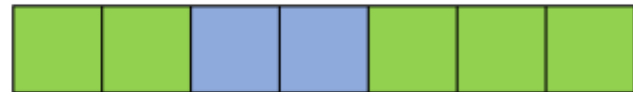
Pai B



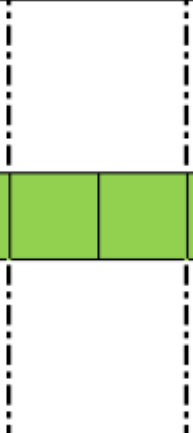
Filho A



Filho B



Pontos de corte

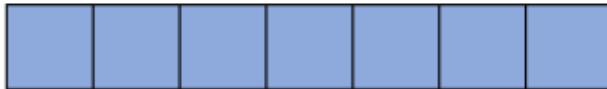




Crossover Uniforme

Sorteados gene a gene

Pai A



Filho A



Pai B



Filho B



1	0	0	1	0	1	1
---	---	---	---	---	---	---

bits sorteados



Crossover com codificação em ordem

Pai A

3	5	7	1	2	4	6
---	---	---	---	---	---	---

Pai B

2	1	6	4	7	5	3
---	---	---	---	---	---	---

Filho A

3	5	1	7	2	4	6
---	---	---	---	---	---	---

Filho B

2	1	4	6	7	5	3
---	---	---	---	---	---	---

Pontos de corte



Mutação

Responsável pela característica de *exploration*

Para cada gene é verificado se deve sofrer alteração com probabilidade dada por parâmetro do algoritmo p_m

O novo gene pode ser sorteado entre os alelos possíveis, ou apenas trocado pelo inverso em caso de apenas dois estados possíveis



Teorema fundamental dos esquemas

$$E[m(H, t + 1)] \geq \underbrace{\frac{\hat{u}(H, t)}{\bar{f}(t)}}_{\text{Seleção}} m(H, t) \underbrace{\left(1 - p_c \frac{\delta(H)}{l - 1}\right)}_{\text{Probabilidade sobrevivência crossover}} \underbrace{\left[(1 - p_m)^{o(H)}\right]}_{\text{Probabilidade sobrevivência mutação}}$$

Seleção

Probabilidade
sobrevivência
crossover

Probabilidade
sobrevivência
mutação

Define-se

$$\hat{u}(H, t) = \frac{\sum_{i \in M(H, t)} f(i)}{m(H, t)}$$

Valor esperado somente com seleção

$$E[m(H, t + 1)] = \frac{\sum_{i \in M(H, t)} f(i)}{\bar{f}(t)}$$



Paralelismo implícito

Holland afirma que a cada população de n cromossomos, intrinsicamente estão sendo avaliados uma quantidade bem maior de esquemas

Sequência binária de comprimento l é instância de 2^l esquemas

10 é instância de $**$, $*0$, $1*$ e 10

Quantidade de esquemas em determinada população é valor entre:

$$2^l \text{ e } n \cdot 2^l$$



Problemas associados ao GA

Carona (*hitchhiking*): genes vizinhos a genes que formam esquemas com boa avaliação tendem a ser favorecidos pelo algoritmo

Problemas enganadores (*deceptives*): Problemas que esquemas que não contemplem o máximo global com média de avaliação superior aos esquemas que contém. Função de avaliação apresenta picos cercados por vales.

Epistasia: genes interagem entre si e combinados podem produzir um resultado melhor ou pior do que individualmente.



SUMÁRIO

1. Introdução

2. Algoritmos Genéticos

3. Implementação do Algoritmo e Resultados

a. Teste inicial

b. Modelo de Ising

4. Conclusões



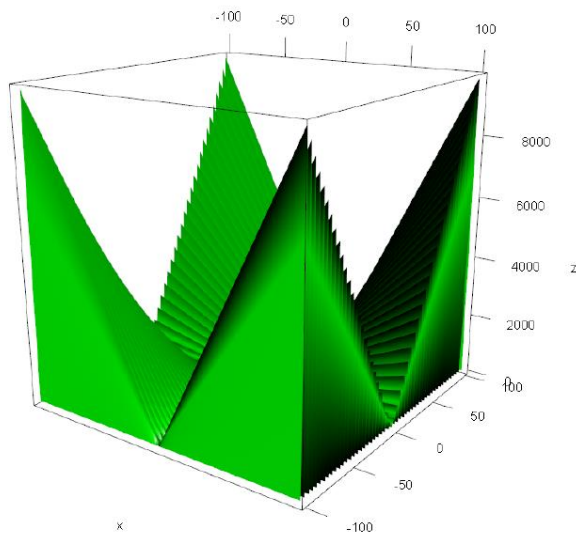
Definições

- x e y variáveis reais com valores no intervalo $[-100,100]$ usando codificação binária com 15 bits para cada
- Sequência binária no cromossomo é convertida para real para depois ser aplicada na função de avaliação.

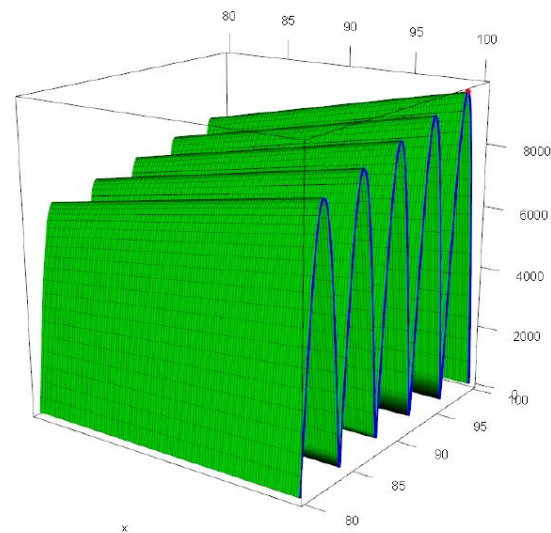
$$x = C \cdot i_x - 100 \quad C = \frac{200}{2^{15} - 1}$$

Função de avaliação:

$$f(x, y) = \left| x \cdot y \cdot \sin \left(\frac{\pi \cdot y}{4} \right) \right|$$



(a) Gráfico no domínio completo



(b) Gráfico ampliado em um dos máximos



Definições para os testes

- **Cromossomo**: vetor combinando as duas sequências binárias das variáveis e portando com $l=30$
- **Seleção**: roleta viciada
- **Operador de crossover**: usando ponto único
- **Operador de mutação**: inversão do bit
- Execução de 1000 testes com os parâmetros configurados e feito média dos resultados:
 - Tamanho da população: $n = 30; 50$
 - Iterações do algoritmo (quantidade de gerações): $T = 10; 100$
 - Probabilidades de crossover: $p_c = 0,8; 0,9$
 - Probabilidade de mutação: $p_m = 0,01; 0,05$



Resultados

Listagem mostrando operadores em funcionamento:

```
1  Geração: 1 | Média Avaliação: 2847,450535
2  #### CHILD #####
3  ID: 31
4  Fenotipo: x = -13,91949 | y = -94,32356
5  Avaliação: 1270,770525
6  #### PARENT 1 #####
7  ID: 11
8  Fenotipo: x = -13,91949 | y = -19,02829
9  Avaliação: 183,079928
10 Selecionado 1 vezes, 0,03
11
12 #### PARENT 2 #####
13 ID: 13
14 Fenotipo: x = 38,99960 | y = -94,28694
15 Avaliação: 3584,173001
16 Selecionado 6 vezes, 0,20
17
18 ### HERITAGE MAP ###
19 11111111111111111111222222222222
20 ### MUTATE MAP
21 00000000000000000000000000000000
```



Resultados

Parâmetros				Resultados GA	Resultados RS	#
n	T	p_c	p_m	Média \pm IC	Média \pm IC	GA > RS
30	10	0,8	0,01	8693,10 \pm 47,04	8419,40 \pm 39,28	659
50	10	0,8	0,01	9084,39 \pm 32,39	8736,45 \pm 32,43	716
30	100	0,8	0,01	9369,35 \pm 26,09	9352,98 \pm 15,41	578
30	10	0,8	0,05	8995,23 \pm 35,40	8481,60 \pm 38,72	741
30	100	0,8	0,05	9670,58 \pm 12,67	9367,58 \pm 15,16	852
30	10	0,9	0,01	8729,25 \pm 45,65	8453,47 \pm 39,33	623
30	10	0,9	0,05	8997,06 \pm 35,47	8504,56 \pm 38,22	735
50	100	0,8	0,05	9739,77 \pm 8,45	9454,56 \pm 12,07	927

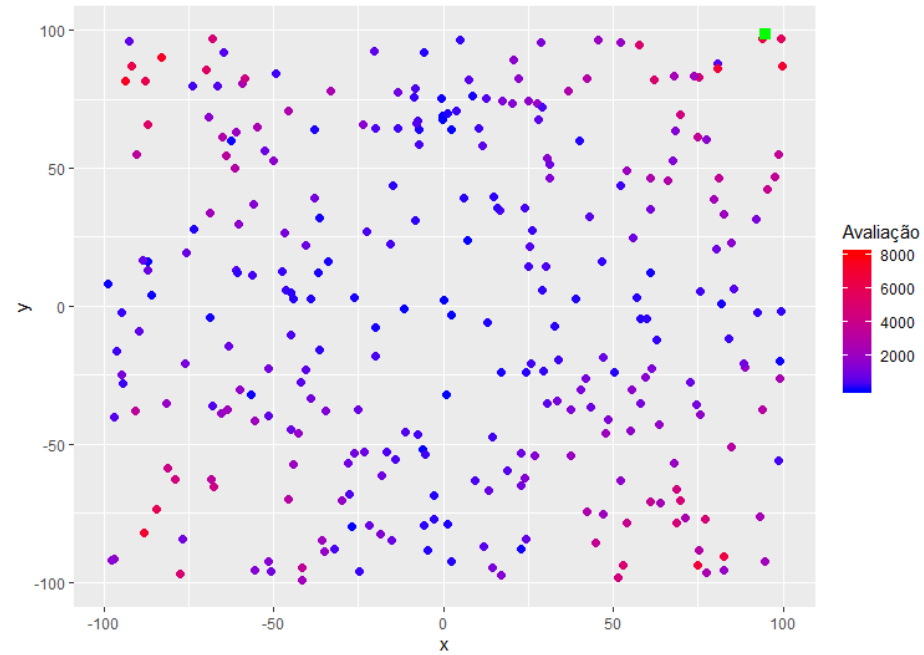


3. IMPLEMENTAÇÃO E RESULTADOS

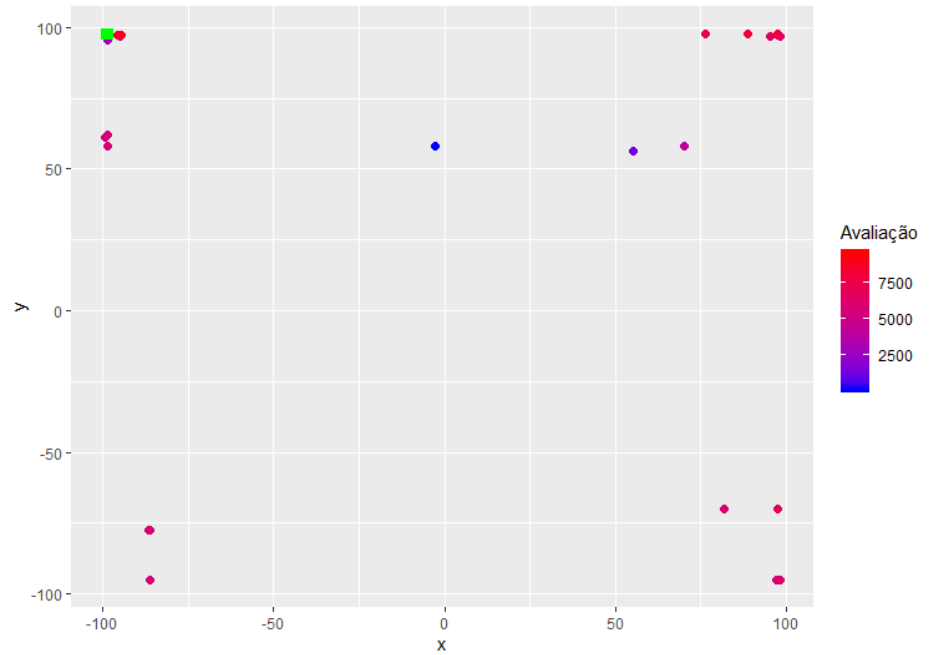
3A. TESTE INICIAL

Resultados

Random Search com 300 observações

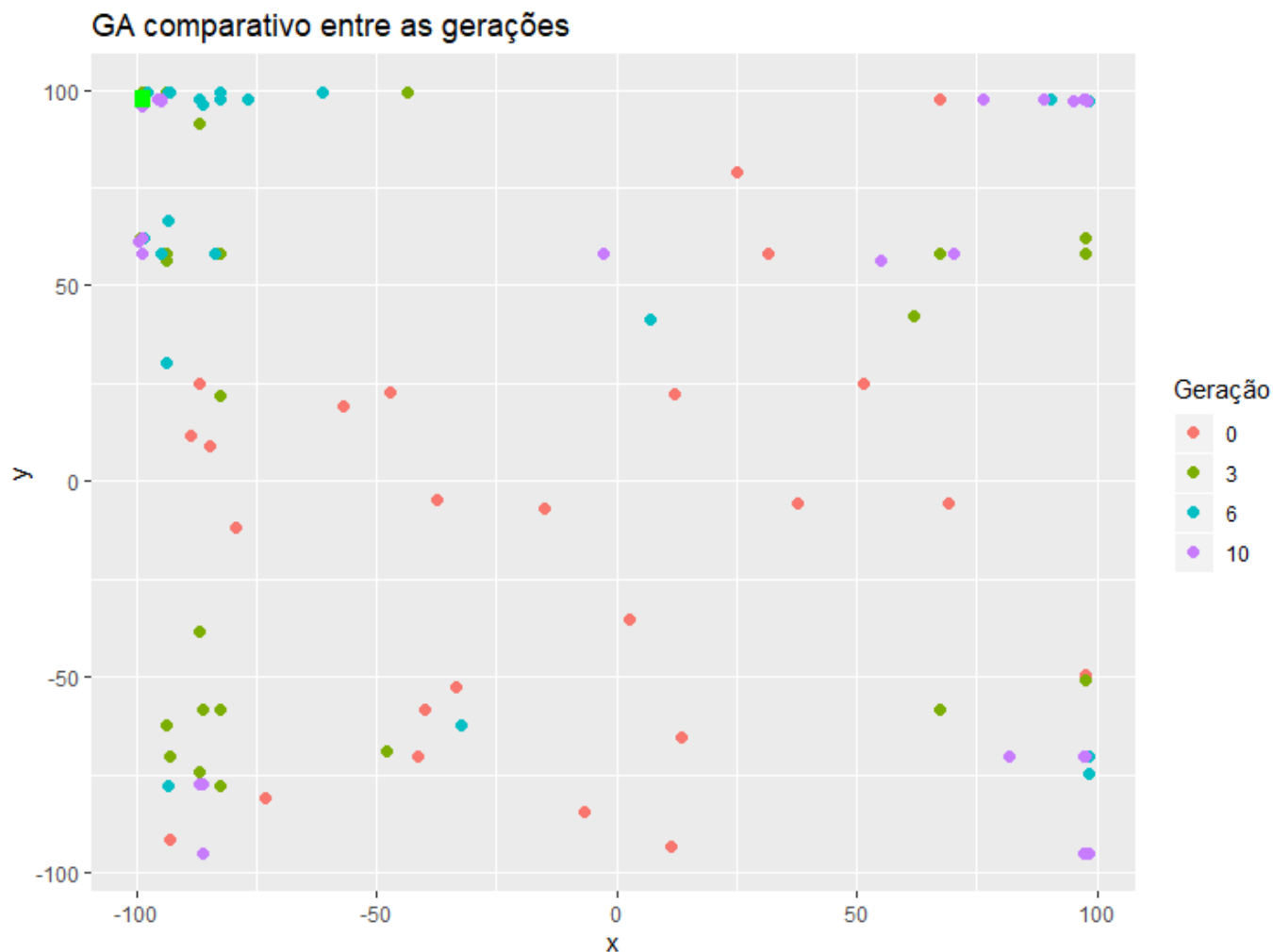


GA na geração 10 com 30 cromossomos



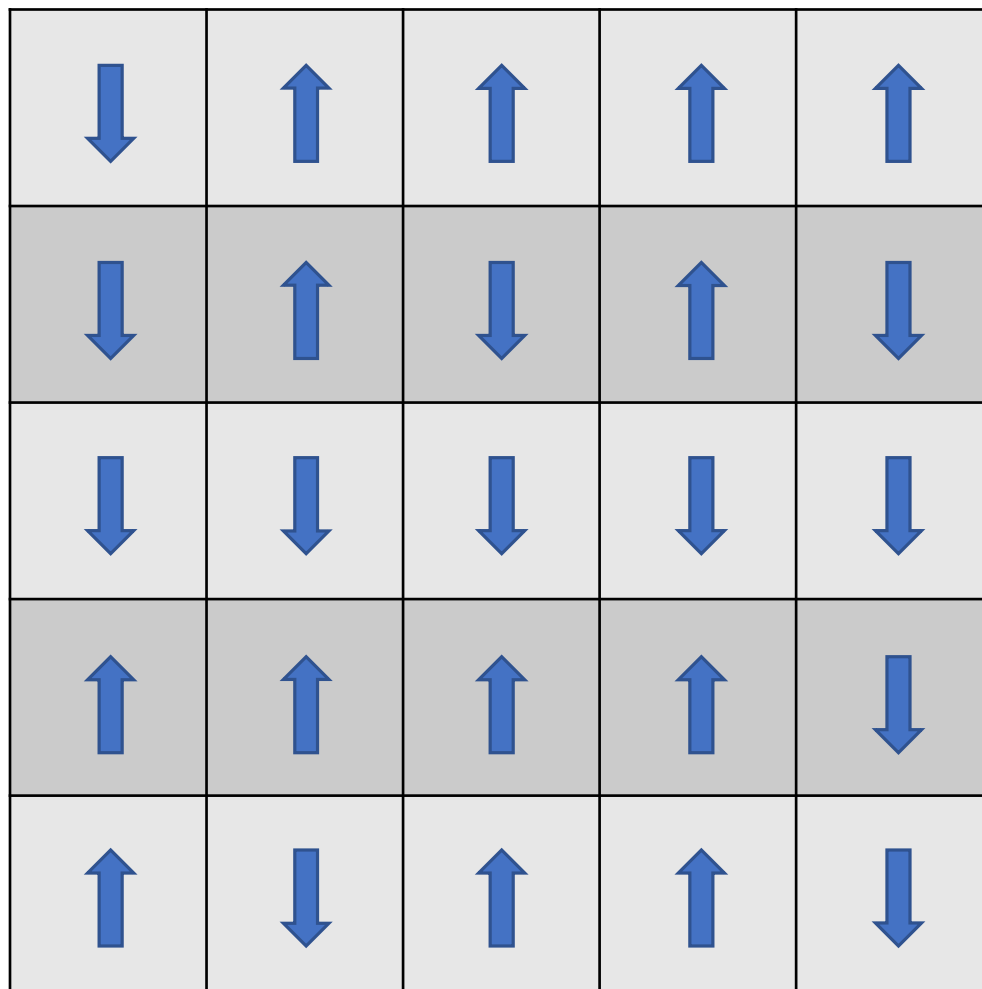


Evolução da população





Grade do modelo





Definição do modelo de Ising

$\sigma_s \in \{-1, +1\}$ spin na posição s da estrutura

$-J(s, s')\sigma_s\sigma_{s'}$ energia de interação entre dois spins

$-h\sigma_s$ energia de interação do spin com um campo magnético externo

Λ estrutura, ou grade, do modelo

$\Omega_\Lambda = \{-1, +1\}^\Lambda$ Espaço de possíveis estados do modelo

$\sigma \in \Omega_\Lambda$, $\sigma = (\sigma_s)_{s \in \Lambda}$ um possível estado

Hamiltoniano

$$\mathcal{H}(\sigma) = - \sum_{s, s' \in \Lambda} J(s, s') \sigma_s \sigma_{s'} - \sum_{s \in \Lambda} h \sigma_s$$



Modelo de Ising - Probabilidades

Probabilidade do modelo estar em determinado estado

$$P(\sigma) = \frac{e^{-\beta \mathcal{H}(\sigma)}}{Z(\beta)}$$

Onde

$$Z(\beta) = \sum_{\sigma \in \Omega_{\Lambda}} e^{-\beta \mathcal{H}(\sigma)}$$

e

$$\beta = 1/\kappa \mathcal{T}$$



Definições para os testes

$\Lambda = [1, 10]^2$ matriz 10 x 10

$J(s, s') = 1$ para s e s' vizinhos, 0 caso contrário e $h = 0$

A função de energia será dada por

$$\mathcal{H}(\sigma) = - \sum_{\langle s, s' \rangle} \sigma_s \sigma_{s'} = - \sum_{i=1}^{10} \sum_{j=1}^{10} \left(\mathbb{1}_{j \neq 10} \sigma_{i,j} \sigma_{i,j+1} + \mathbb{1}_{i \neq 10} \sigma_{i,j} \sigma_{i+1,j} \right)$$

$$\mathbb{1}_A = \begin{cases} 1 & , \text{ se } A \text{ for verdade} \\ 0 & , \text{ caso contrário} \end{cases}$$



Definições para os testes

cromossomo: vetor de inteiros com linhas da matriz concatenadas.

Exemplo:

$$\text{cromossomo} = [-1, 1, 1, 1, -1, 1, 1, 1, -1]$$

$$\text{fenótipo} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$



Definições para os testes

- **Seleção:** roleta viciada
- **Operador de crossover:** usando ponto único, dois pontos e uniforme
- **Operador de mutação:** sorteio do novo estado
- Execução de 1000 testes com os parâmetros configurados e feito média dos resultados:
 - Tamanho da população: $n = 30; 50$
 - Iterações do algoritmo (quantidade de gerações): $T = 10; 100$
 - Probabilidades de crossover: $p_c = 0,8; 0,9$
 - Probabilidade de mutação: $p_m = 0,01; 0,05$



Problema com função de avaliação

Sendo $\xi \in \theta_\Lambda \subset \Omega_\Lambda$ um estado pertencente a população

1ª função de avaliação:

$$f(\xi) = -\left(\mathcal{H}(\xi) + \min_{\sigma \in \Omega_\Lambda} \mathcal{H}(\sigma)\right)$$

2ª função de avaliação:

$$f(\xi) = e^{-\beta \mathcal{H}(\xi)}$$

Probabilidade de seleção de um cromossomo dada por:

$$Pr(\xi) = \frac{e^{-\beta \mathcal{H}(\xi)}}{\sum_{\sigma \in \theta_\Lambda} e^{-\beta \mathcal{H}(\sigma)}}$$

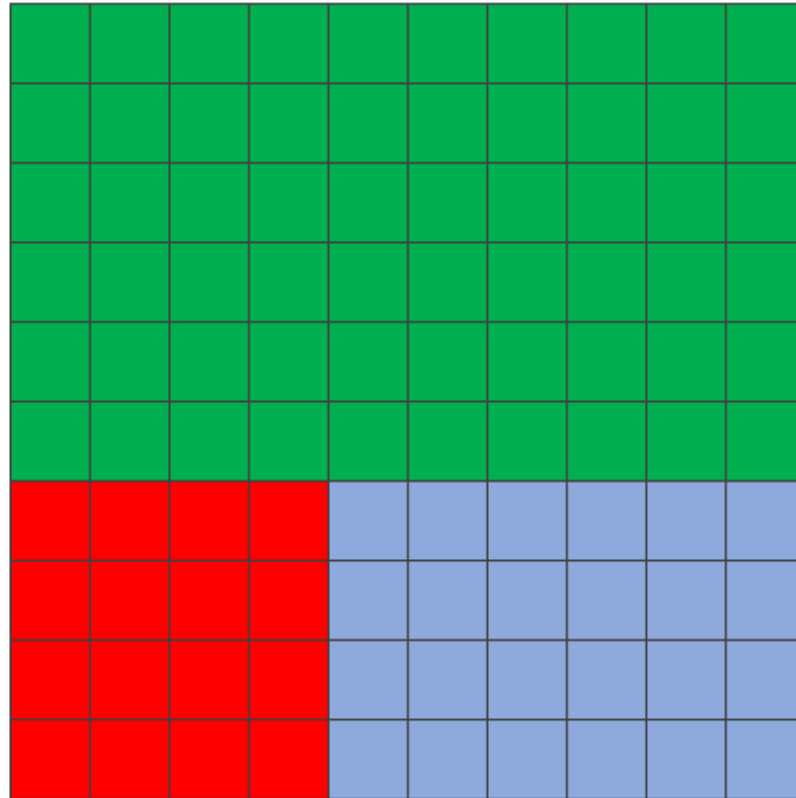


Variando tipo de crossover

Parâmetros				Resultados GA	Resultados RS	#
n	T	p_c	p_m	Média \pm IC	Média \pm IC	GA > RS
30	10	$0,8_1$	0,02	$-58,90 \pm 0,54$	$-39,05 \pm 0,33$	978
30	10	$0,8_2$	0,02	$-60,21 \pm 0,57$	$-38,98 \pm 0,32$	970
30	10	$0,8_u$	0,02	$-60,42 \pm 0,63$	$-39,39 \pm 0,35$	960
30	100	$0,8_1$	0,02	$-122,24 \pm 0,71$	$-48,54 \pm 0,27$	1000
30	100	$0,8_2$	0,02	$-126,05 \pm 0,75$	$-47,94 \pm 0,27$	1000
30	100	$0,8_u$	0,02	$-130,47 \pm 0,73$	$-48,23 \pm 0,28$	1000



Genótipo x crossover





Variando os parâmetros

Parâmetros				Resultados GA	Resultados RS	#
n	T	p_c	p_m	Média \pm IC	Média \pm IC	GA > RS
30	40	0,8	0,02	-105,07 \pm 0,72	-44,98 \pm 0,31	1000
30	40	0,8	0,05	-93,90 \pm 0,71	-44,63 \pm 0,29	1000
30	40	0,8	0,01	-101,69 \pm 0,74	-44,91 \pm 0,30	1000
40	40	0,8	0,02	-111,06 \pm 0,73	-46,24 \pm 0,30	1000
40	40	0,8	0,01	-109,59 \pm 0,73	-46,02 \pm 0,31	1000
30	40	0,95	0,02	-108,42 \pm 0,75	-44,91 \pm 0,31	1000
30	40	0,7	0,02	-101,26 \pm 0,69	-44,78 \pm 0,31	1000
30	40	0,8	0,02	-110,66 \pm 0,66	-44,81 \pm 0,30	1000
30	40	0,95	0,02	-113,69 \pm 0,70	-44,75 \pm 0,30	1000



Variando os parâmetros

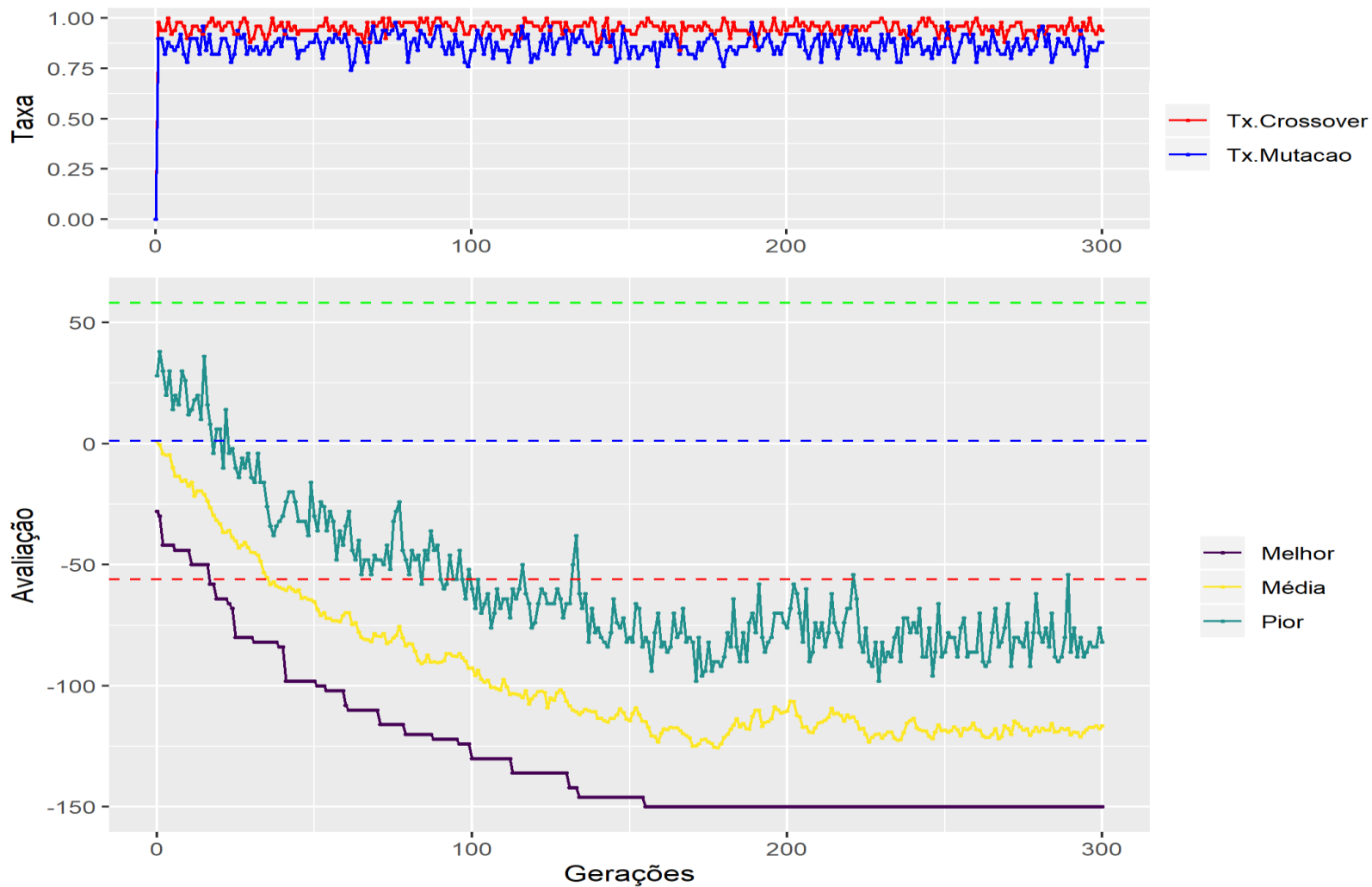
Parâmetros					Resultados GA	Resultados RS	#
n	T	p_c	p_m	β	Média \pm IC	Média \pm IC	GA > RS
30	40	0,95	0,02	0,05	-113,69 \pm 0,70	-44,75 \pm 0,30	1000
30	40	0,95	0,02	0,01	-90,78 \pm 0,66	-44,79 \pm 0,30	1000
30	40	0,95	0,02	0,1	-120,45 \pm 0,70	-44,95 \pm 0,30	1000



3. IMPLEMENTAÇÃO E RESULTADOS

3B. MODELO DE ISING

Teste 1 - $p_c = 0.950$ | $p_m = 0.020$ | $\beta = 0.010$

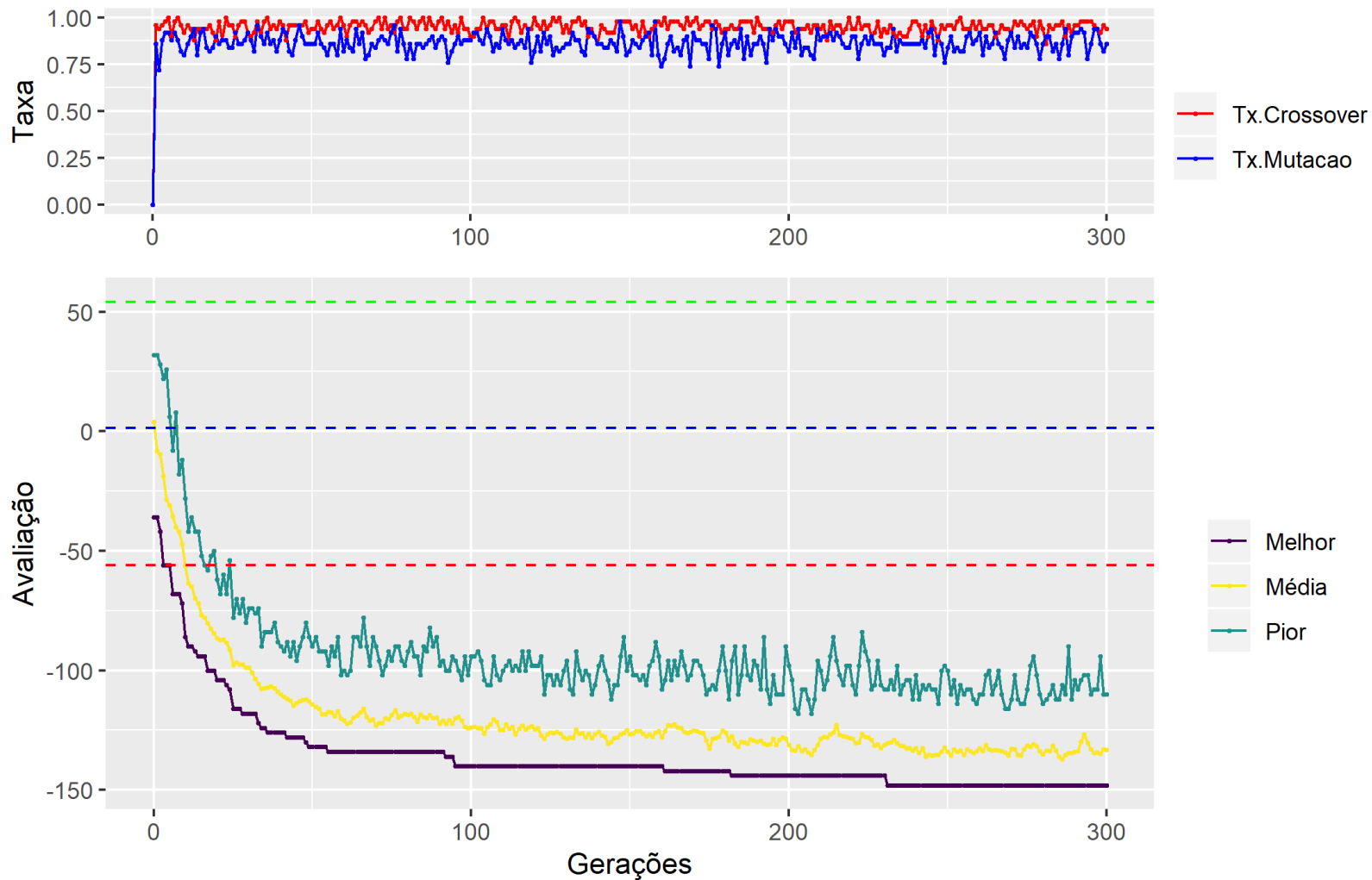




3. IMPLEMENTAÇÃO E RESULTADOS

3B. MODELO DE ISING

Teste 1 - $p_c = 0.950$ | $p_m = 0.020$ | $\beta = 0.050$

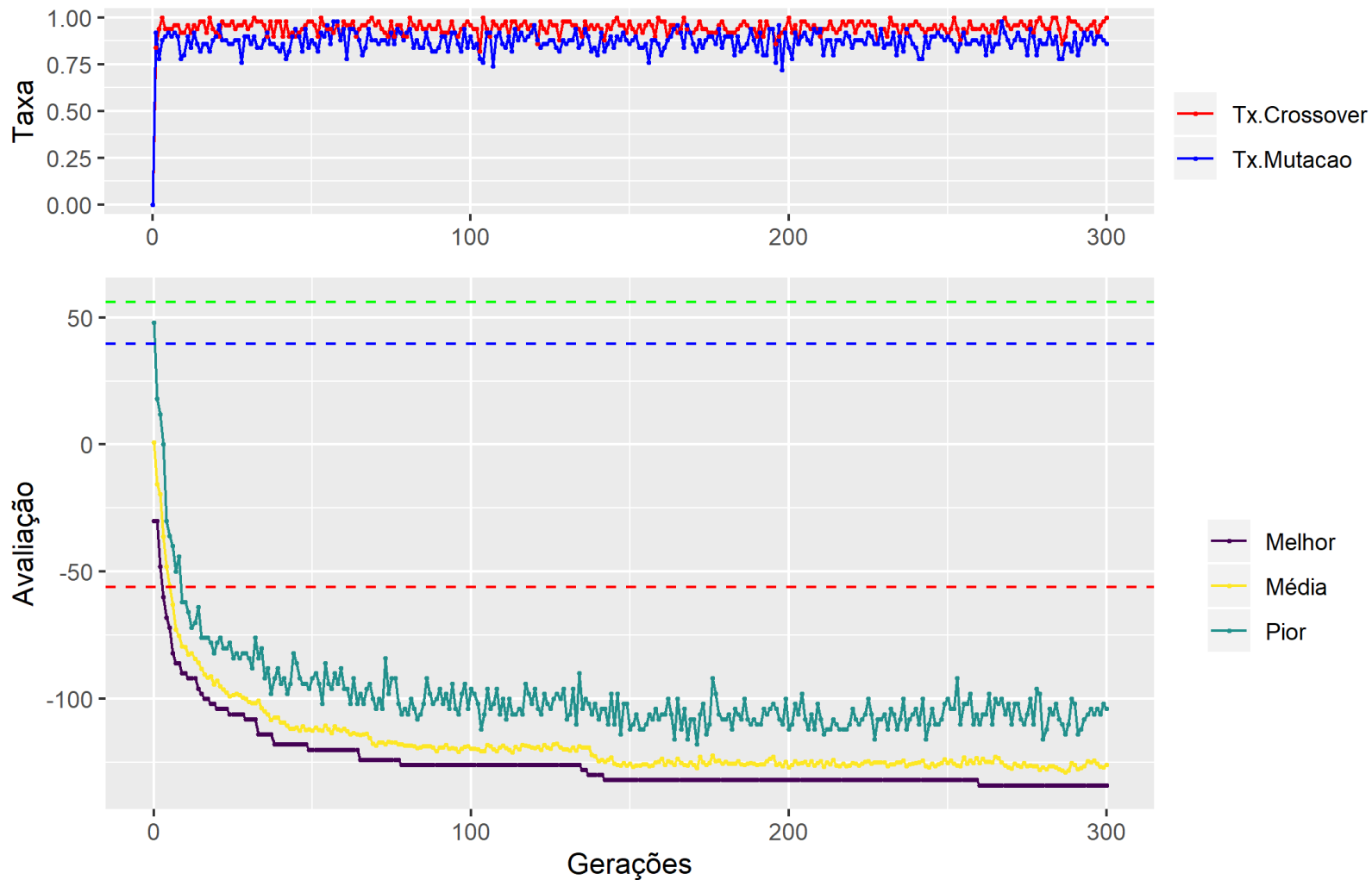




3. IMPLEMENTAÇÃO E RESULTADOS

3B. MODELO DE ISING

Teste 1 - $p_c = 0.950$ | $p_m = 0.020$ | $\beta = 0.200$





Comparando médias e variâncias

Com operador de mutação

Geração	0	10	20	30	40
Média	0	-47,49	-78,76	-96,46	-106,86
Variância	178,55	158,42	117,09	106,79	103,55

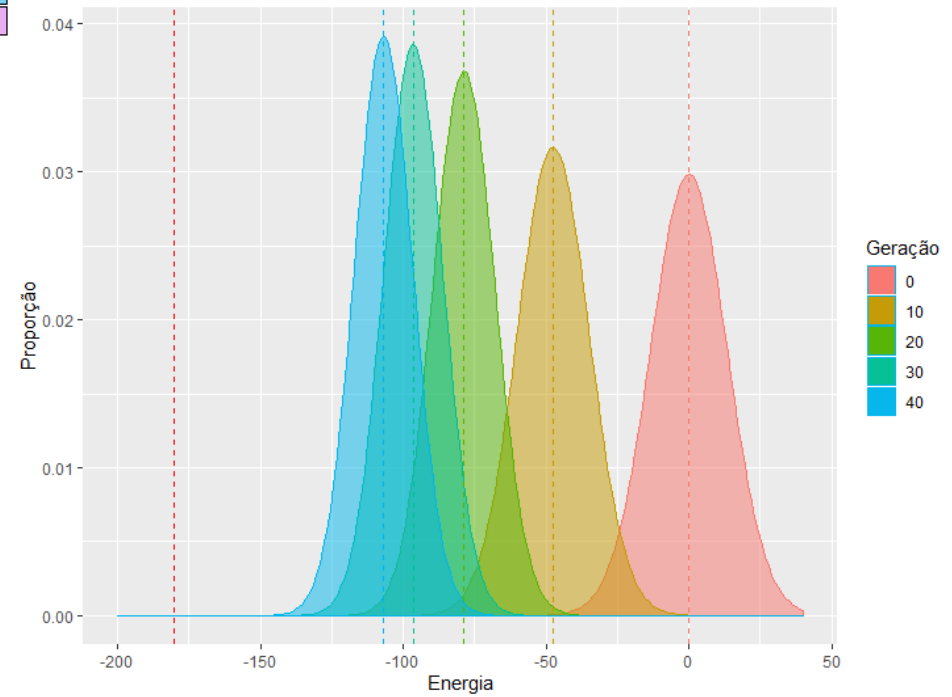
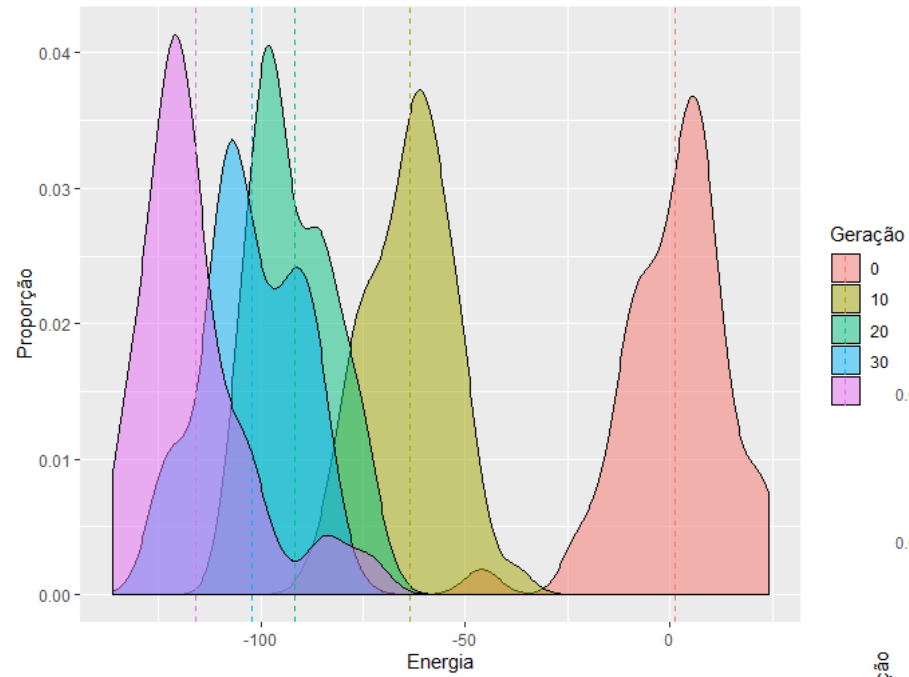
Sem operador de mutação

Geração	0	10	20	30	40
Média	0	-47,49	-78,76	-96,46	-106,86
Variância	178,55	158,42	117,09	106,79	103,55



3. IMPLEMENTAÇÃO E RESULTADOS

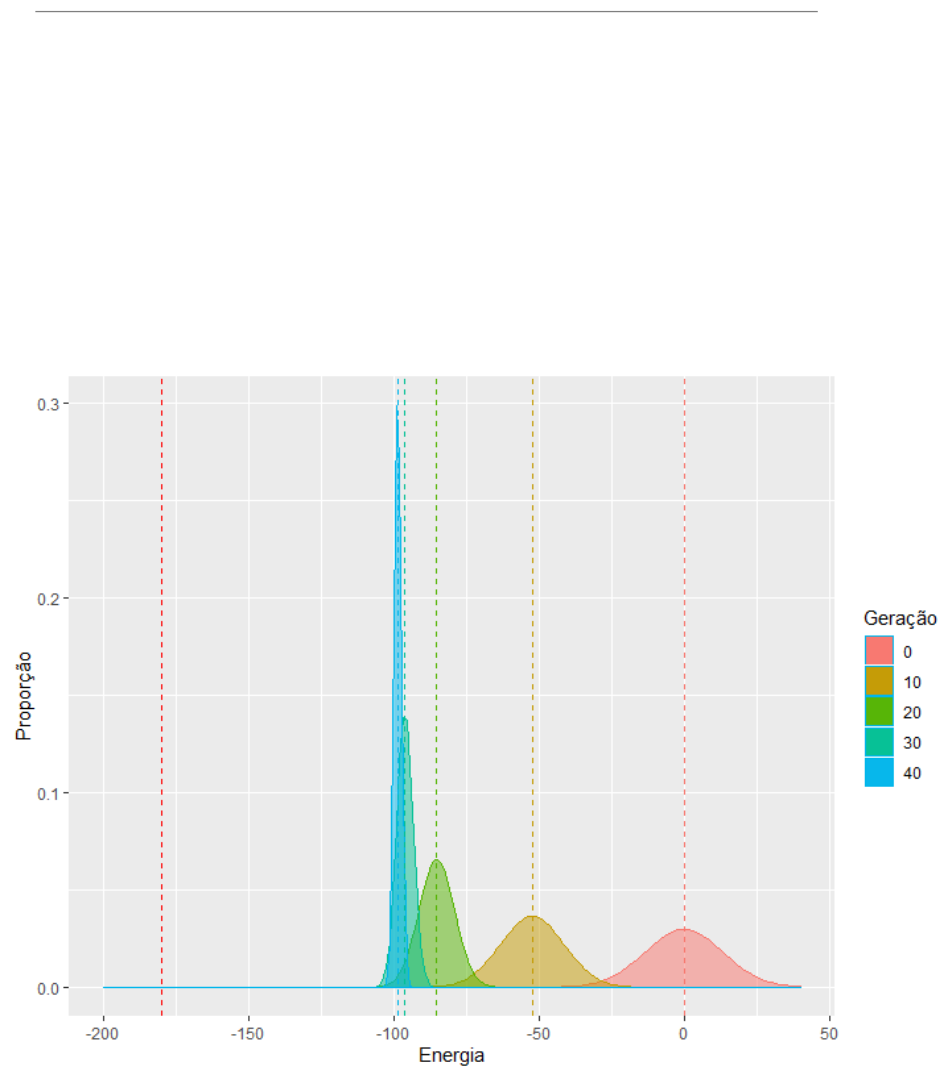
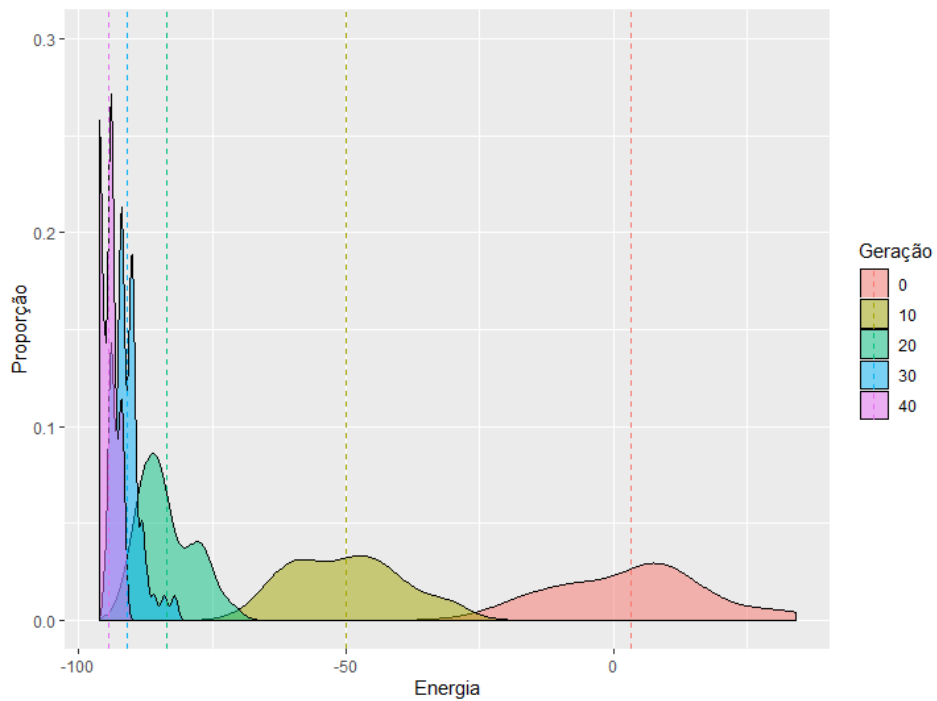
3B. MODELO DE ISING





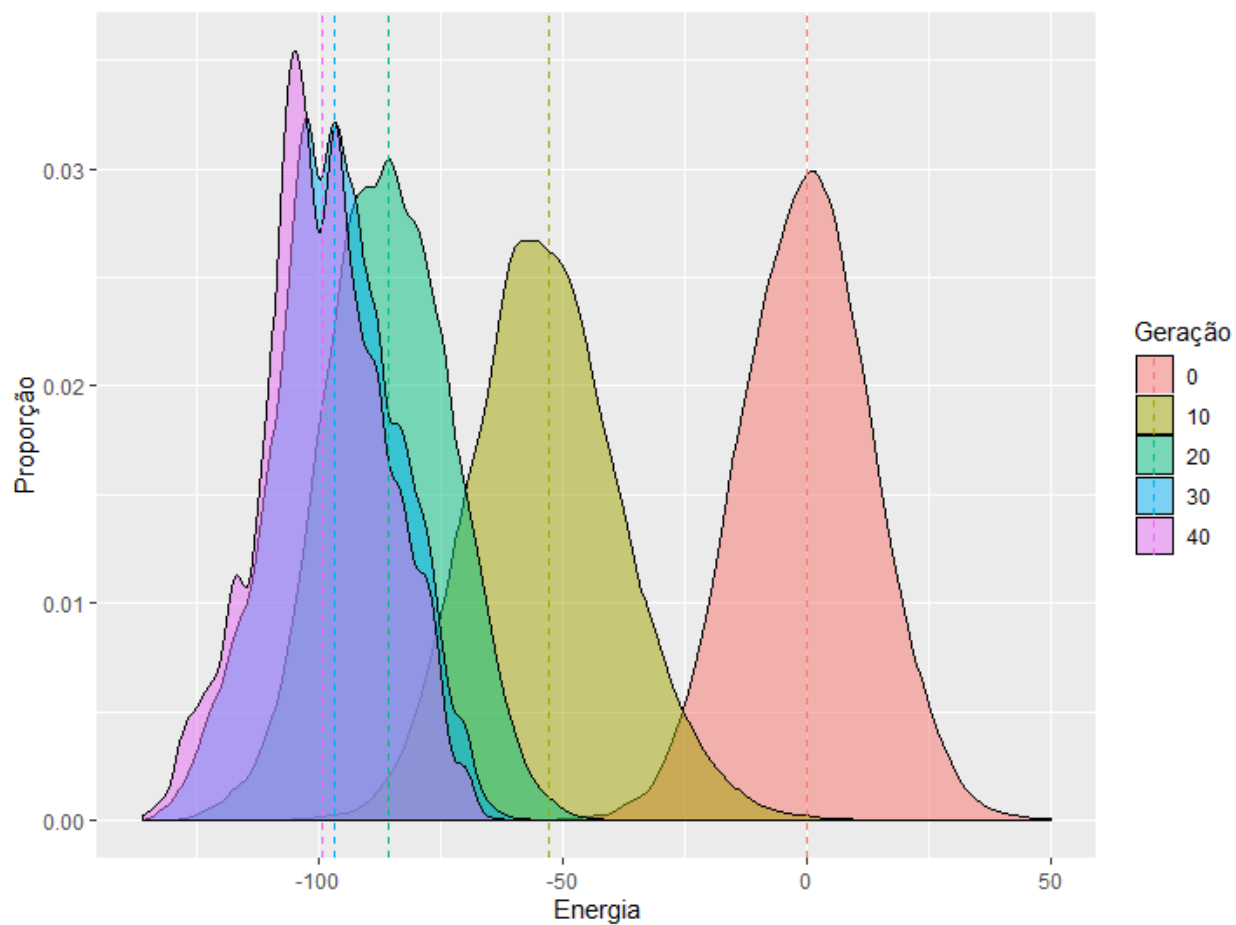
3. IMPLEMENTAÇÃO E RESULTADOS

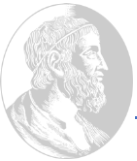
3B. MODELO DE ISING



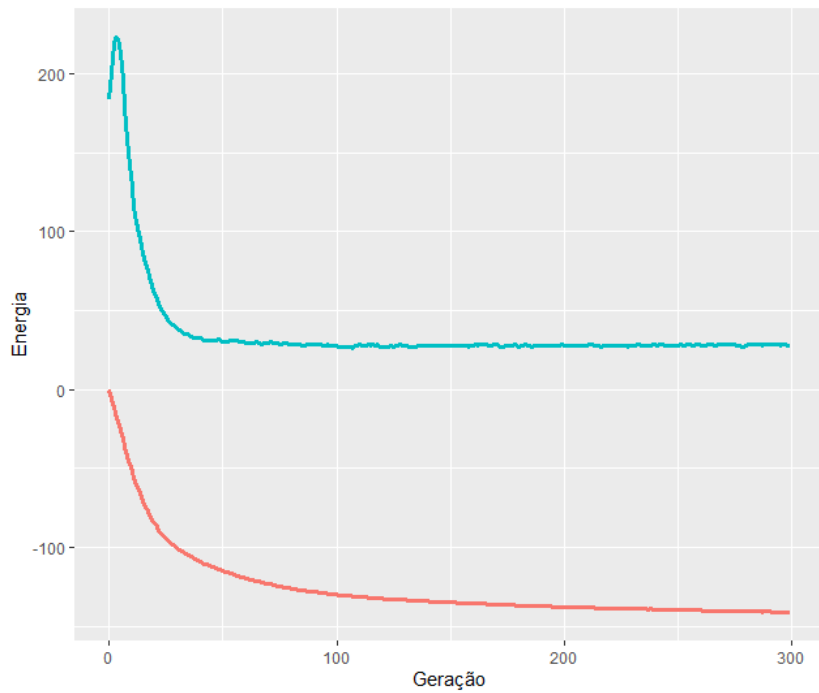


Distribuição de 500 testes nas gerações

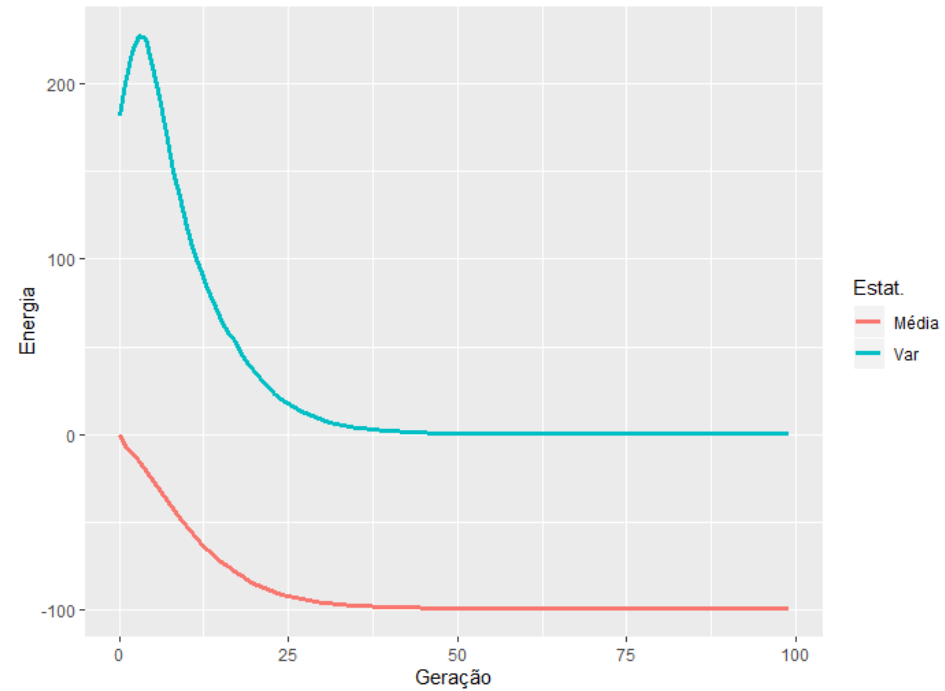




Progressão da média e variância



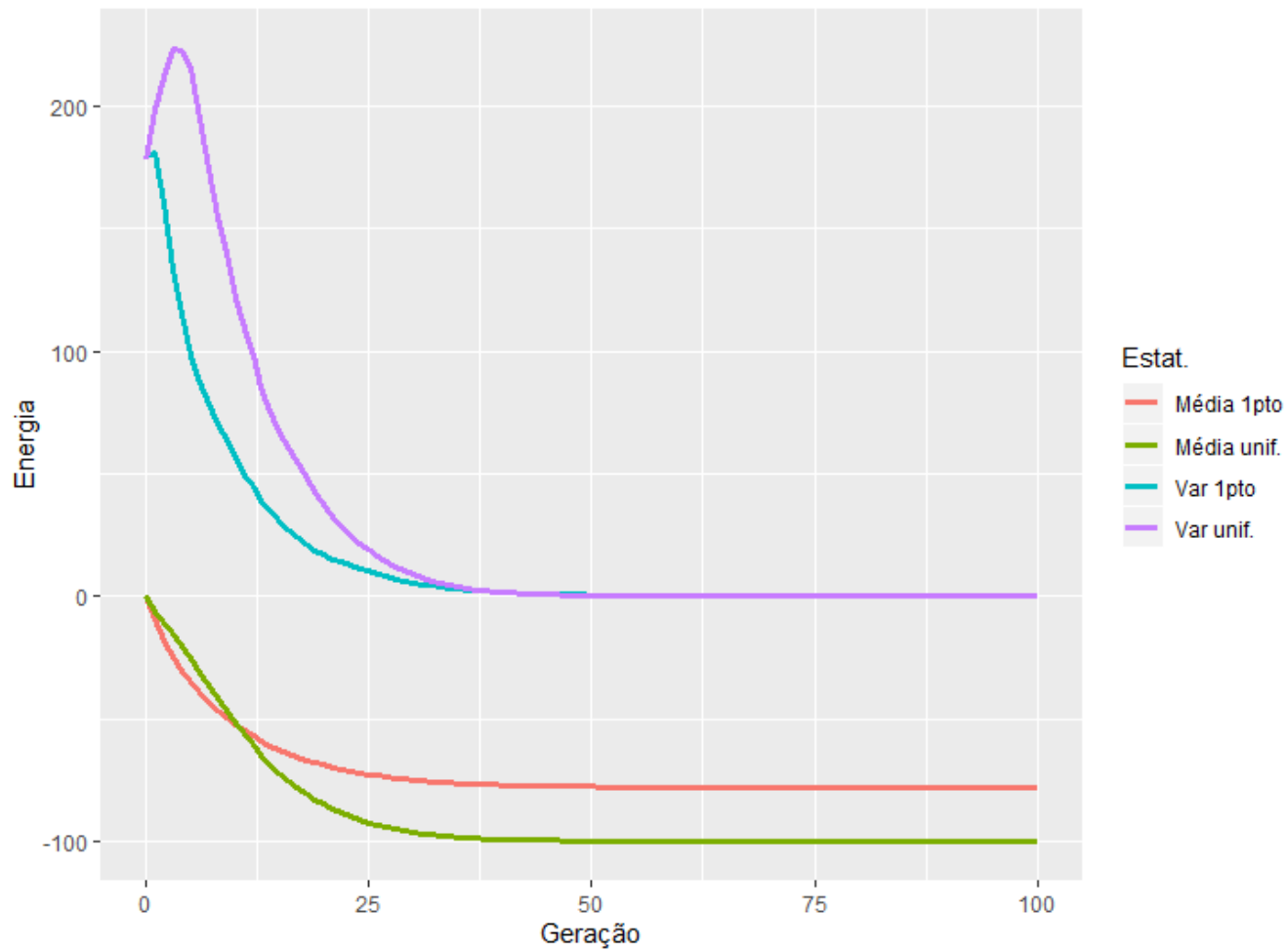
$$p_m = 0,005$$



$$p_m = 0$$



Comparando crossover



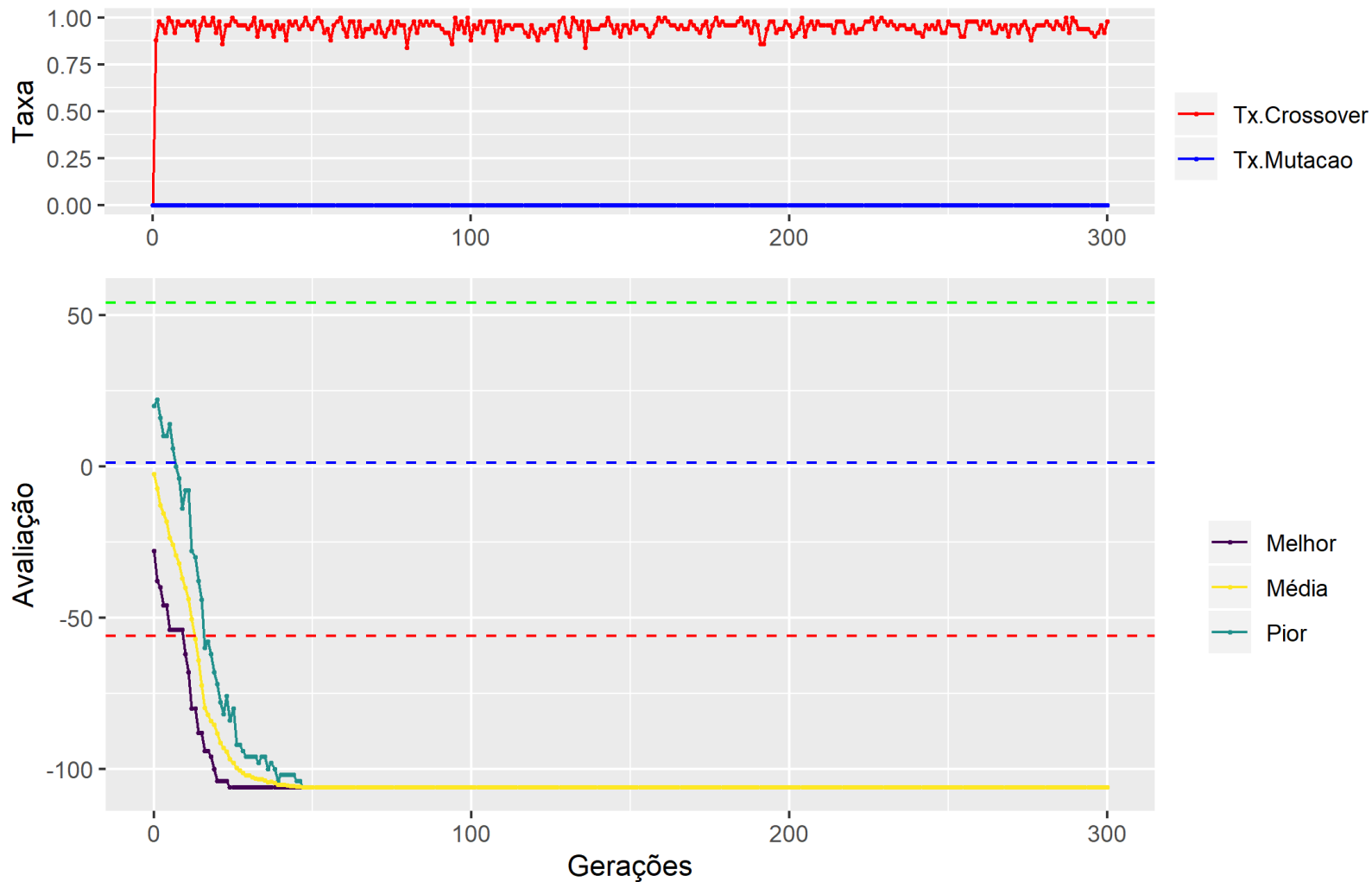
$$p_m = 0$$



3. IMPLEMENTAÇÃO E RESULTADOS

3B. MODELO DE ISING

Teste 1 - Probabilidades : Cross = 0.950 | Mut = 0.000



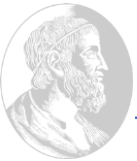


3. IMPLEMENTAÇÃO E RESULTADOS

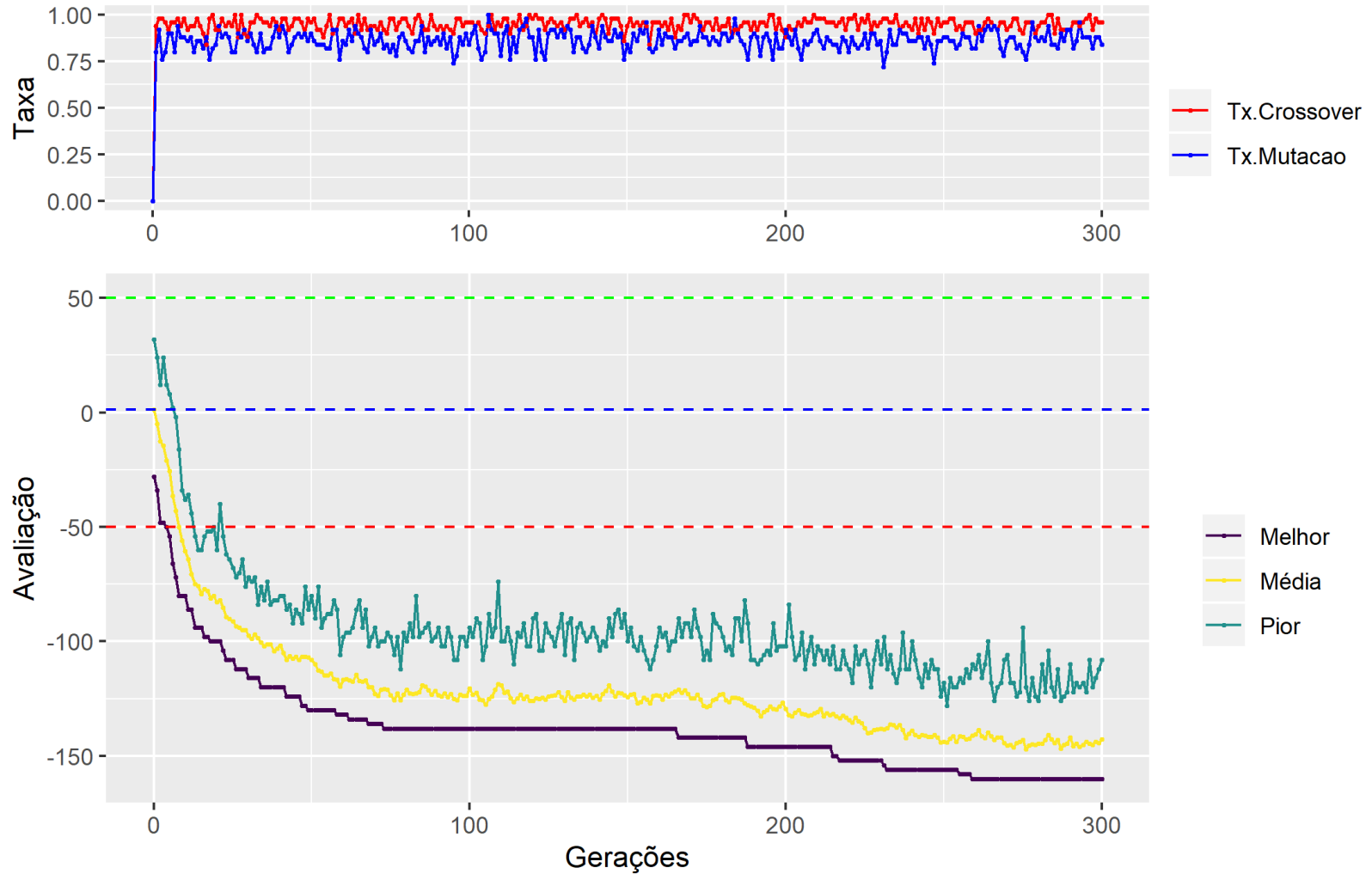
3B. MODELO DE ISING

Teste 1 - Probabilidades : Cross = 0.950 | Mut = 0.001





Teste 2 - Probabilidades : Cross = 0.950 | Mut = 0.020

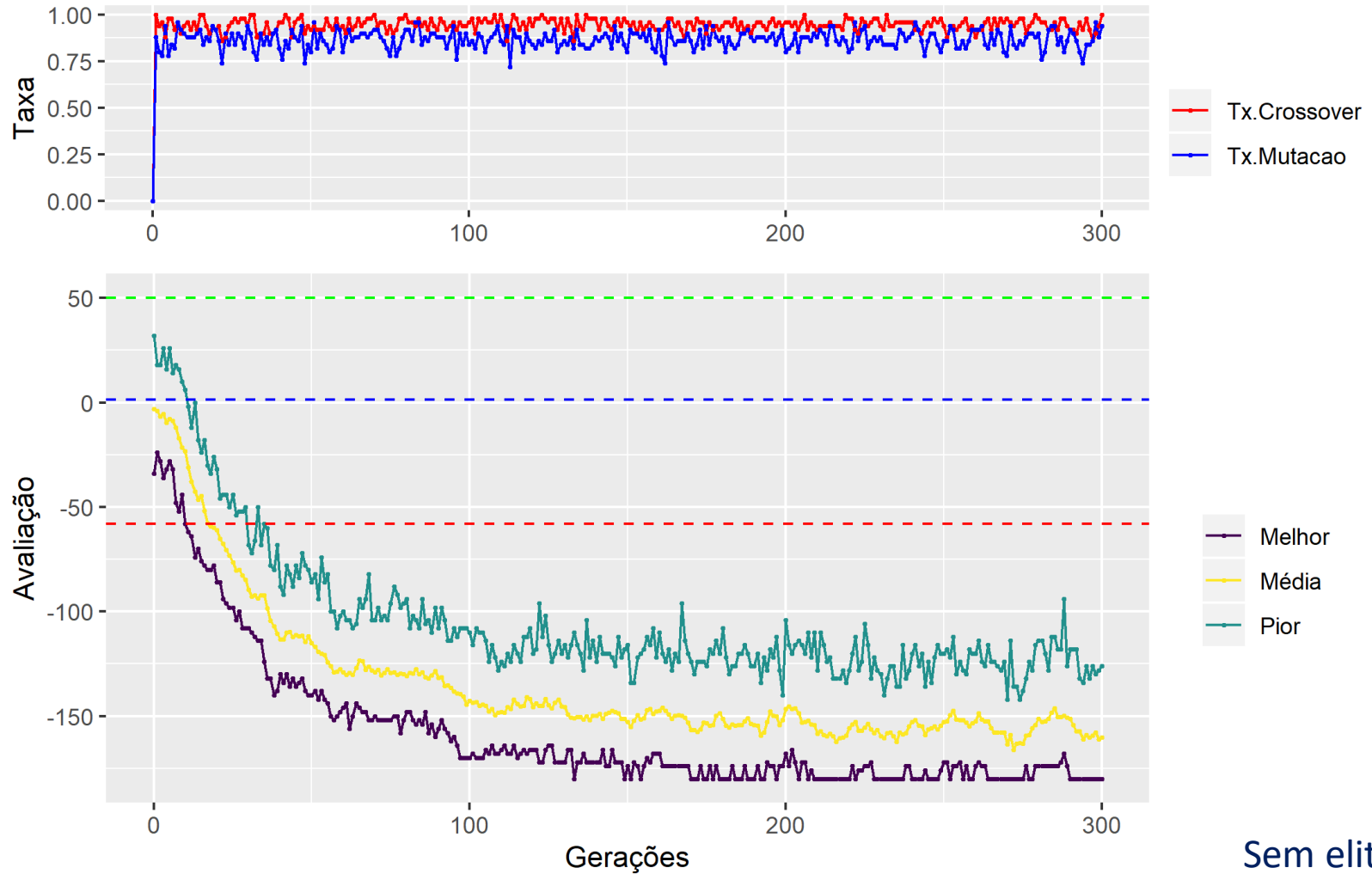




3. IMPLEMENTAÇÃO E RESULTADOS

3B. MODELO DE ISING

Teste 1 - Probabilidades : Cross = 0.950 | Mut = 0.020





SUMÁRIO

1. Introdução

2. Algoritmos Genéticos

3. Implementação do Algoritmo e Resultados

a. Teste inicial

b. Modelo de Ising

4. Conclusões



OBRIGADO!