



Samuel Liló Abdalla

Matemática + Biologia = Algoritmos Genéticos

Sistemas Inteligentes

A Inteligência Computacional é uma área da ciência que busca, através de técnicas inspiradas na Natureza, Procura desenvolver sistemas inteligentes que imitam aspectos do comportamento humano, tais como: aprendizado, percepção, raciocínio, evolução e adaptação.

Técnica	Inspiração
Redes Neurais	Neurônios biológicos
Algoritmos Genéticos	Evolução biológica
Lógica Fuzzy	Proc. linguístico
Sistemas Especialistas	Inferência Humana

Problemas de Otimização Combinatória

A **Otimização Combinatória** é um ramo da ciência da computação que estuda problemas de otimização em conjuntos.

Problemas de otimização, na sua forma geral, têm como objetivo a maximização ou minimização (determinação de máximo ou mínimo) de uma função definida sobre um certo domínio. Os problemas de otimização combinatória apresentam um conjunto finito de candidatas a solução; além disso, em geral pode – se listar os seus elementos e também testar se um dado elemento pertence a esse conjunto. No entanto, a idéia ingênua de testar todos os elementos deste domínio na busca da melhor solução mostra-se inviável na prática, mesmo para instâncias de tamanho moderado.

Todos os problemas de otimização combinatória surgem naturalmente em aplicações práticas, tais como o projeto de redes de telecomunicação, o empacotamento de objetos em containeres, o roteamento de veículos, etc. Outras áreas de aplicação incluem a estatística (análise de dados), a economia (matrizes de entrada/saída), a física (estados de energia mínima), a biologia molecular (alinhamento de DNA e proteínas, inferência de padrões), etc.

Como exemplos clássicos de problemas de otimização combinatória podemos citar o problema do caixeiro viajante, o problema da mochila e o problema da floresta de Steiner.

Existem muitas classificações possíveis para o problema de otimização, e algumas delas apresentarão métodos exatos e eficientes de resolução. Outras levarão à necessidade de métodos não-exatos (**heurística**), uma vez que sua formulação e/ou resolução exatas levariam a uma complexidade intratável.

A utilização dos Algoritmos genéticos tem – se mostrado uma técnica poderosa na busca de soluções para estes problemas. Este artigo abordará essa técnica de busca, que foi inspirada na teoria da Evolução.

Um Breve Histórico

Até meados do século 19, os naturalistas acreditavam que cada espécie havia sido criada separadamente por um ser supremo (divindade criadora) ou através de geração espontânea. Os trabalhos do naturalista **Carolus Linnaeus** sobre a classificação biológica de organismos despertaram o interesse pelas similaridades entre certas espécies, intuitivamente acreditou – se na existência de uma relação entre elas.

Outros trabalhos influenciaram os naturalistas em direção à teoria da seleção natural, tais como os de **Jean Baptiste Lamarck**, que sugeriu uma teoria evolucionária no "uso e desuso" de órgãos; e de **Thomas Robert Malthus**, que propôs que doenças e escassez de alimentos eram efeitos, provocados pelo meio ambiente, que limitavam o crescimento populacional.

Depois de mais de 20 anos de observações e experimentos, **Charles Darwin** apresentou em 1858 sua teoria de evolução através de seleção natural, paralelamente com outro naturalista inglês **Alfred Russel Wallace**. No ano seguinte, Darwin publica o seu *On the Origin of Species by Means of Natural Selection* com a sua teoria completa, sustentada por muitas evidências colhidas durante suas viagens a bordo do Beagle (Cão de origem inglesa do tipo sabujo, de porte médio, onde até hoje é usado em caça à lebre e à raposa).

“Quanto melhor um indivíduo se adaptar ao seu meio ambiente, maior será sua chance de sobreviver e gerar descendentes.”

(DARWIN, 1859)

Este trabalho influenciou de forma avassaladora o futuro não apenas da Biologia, Botânica e Zoologia, mas contribuiu com grande influência sobre o

pensamento religioso, filosófico, político e econômico da época. A teoria da evolução e a computação nasceram praticamente na mesma época: O matemático **Charles Babbage**, um dos fundadores da computação moderna e amigo pessoal de Darwin desenvolveu sua máquina analítica em 1833. Ambos provavelmente estariam surpresos e orgulhosos com a ligação entre estas duas áreas.

Por volta de 1900, o trabalho de **Gregor Mendel**, desenvolvido em 1865, sobre os princípios básicos de herança genética em conjunto com as idéias de Darwin e Wallace sobre a seleção natural, deram origem ao princípio básico da Genética Populacional.

“A variabilidade entre indivíduos em uma população de organismos que se reproduzem é produzida pela mutação e pela recombinação genética”

Este princípio foi desenvolvido durante os anos 30 e 40, por biólogos e matemáticos de importantes centros de pesquisa. Nos anos 50 e 60, muitos biólogos começaram a desenvolver simulações computacionais de sistemas genéticos. Entretanto, foi **John Holland** quem iniciou os primeiros desenvolvimentos concretos sobre o tema, Holland publicou em 1975 seu livro *Adaptation in Natural and Artificial Systems*, hoje considerado a Bíblia de Algoritmos Genéticos.

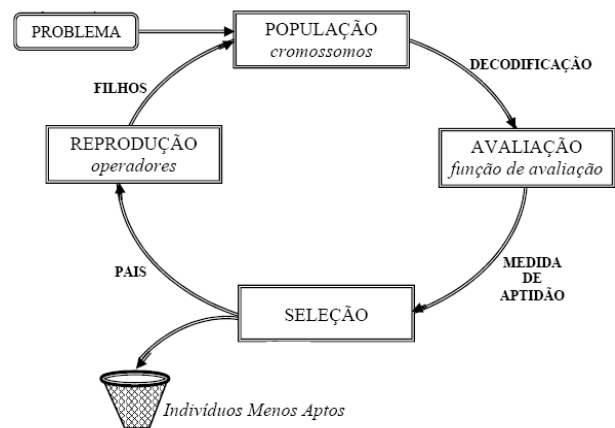
“Algoritmos genéticos são modelos matemáticos computacionais que imitam os mecanismos da Evolução natural para resolver problemas de otimização”

(JOHN HOLLAND, 1970)

Pode – se dizer que a matemática e biologia se uniram para viver felizes para sempre, desde então, os algoritmos genéticos vêm sendo largamente aplicados com sucesso em diversos problemas de otimização e aprendizado de máquina.

Os componentes Básicos dos Algoritmos Genéticos

- Um **problema** para ser resolvido pelo algoritmo;
- Um método para codificar soluções do problema através de **cromossomos**;
- Uma **função de avaliação** que mede quão bem, cada solução é capaz de resolver o problema;
- Um método para criar a **população inicial** de cromossomos;
- Um conjunto de **parâmetros** para o algoritmo genético;
- Um conjunto de **operadores** que atuam no processo de reprodução;



Cromossomos

A representação da solução é orientada como string (seqüência de vários caracteres simples) na estrutura do problema e deve descrever o espaço de busca em termos suas características.

Codificação Binária

- É a mais comum devido a sua simplicidade
- Cada cromossomo é uma string de bits – 0 ou 1
 - Crom: A = 1011001011
 - Crom: B = 1111110000

- Exemplo de uso: problema da mochila
- Codificação: Cada bit diz se um elemento está ou não na mochila

Codificação Real

- Mais usado em problemas de ordenação
- Cada cromossomo é uma string de números reais que representa uma posição numa sequência

- Crom A: 1 5 3 2 6 4 7 9 8
- Crom B: 8 5 6 7 2 3 1 4 9

- Exemplo de uso: problema do caixeiro viajante
- Codificação: os cromossomos descrevem a ordem em que o caixeiro irá visitar as cidades

Função de Avaliação

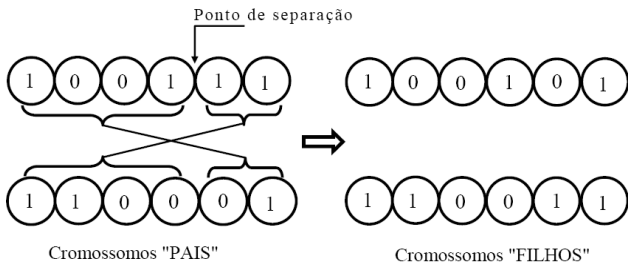
Verifica a aptidão das soluções produzidas, atribuindo uma nota a cada uma delas.

Operadores Genéticos

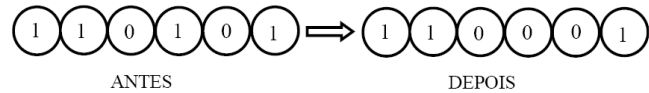
Atuam no processo de reprodução:

1. Crossover
2. Mutação
3. Inversão

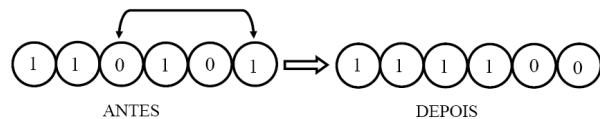
Crossover: Executa troca de partes dos cromossomos pais para produzir os cromossomos filhos.



Mutação: Introduz aleatoriamente modificações na informação genética.



Inversão: Inverte a posição (ordem) de dois elementos escolhidos aleatoriamente em um cromossomo.

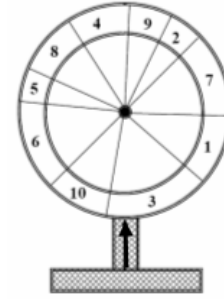


Seleção Natural

- Método da Roleta
- Método do Torneio
- Amostragem Universal Estocástica

Método da roleta:

- Coloca-se os indivíduos em uma roleta, dando a cada um uma "fatia" proporcional à sua aptidão relativa.
- Depois se roda a agulha da roleta. O indivíduo em cuja fatia a agulha parar permanece para a próxima geração
- Repete-se o sorteio quantas vezes forem necessárias para selecionar a quantidade desejada de indivíduos



Exemplo 1. O problema das 8 Rainhas.

É possível colocar 8 Rainhas em um tabuleiro de xadrez de modo que nenhuma fique em casa guardada por outra?

Este problema foi proposto pela primeira vez na revista Schachzeitung em 1848. Por volta de 1850 o matemático Johann Karl Friedrich Gauss (1775 - 1855) e o astrônomo Heinrich Schumacher (1780 - 1850) descobriram 12 soluções fundamentais que por rotação e reflexão dão origem a um total de 92 soluções distintas.

Encontrar soluções para esse problema não é tão fácil quanto parece à primeira vista pois existem:

$$C_8^{64} = \frac{64!}{8!(64-8)!} = 4426165368$$

maneiras distintas de dispor as 8 rainhas no tabuleiro. Se for utilizado o critério de que cada coluna deve conter uma, e apenas uma rainha então a quantidade de possibilidades diminui para

$$8^8 = 16777216$$

Se for utilizado o critério de que cada coluna e cada fileira deve conter uma, e apenas uma rainha, então a quantidade de possibilidades diminui para

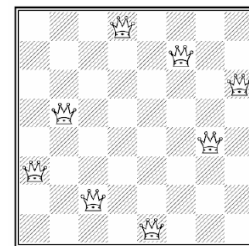
$$8! = 40320$$

Com esse último critério, pode-se representar cada uma das disposições como uma permutação do conjunto {1, 2, 3, 4, 5, 6, 7, 8}. Ou seja, começamos com um cromossomo.

Utilizando como principal operação genética a inversão chega-se numa solução:

$$\{3, 5, 2, 8, 1, 7, 4, 6\}$$

De acordo com o cromossomo apresentado temos:



Como estaremos falando de cromossomos, vale a pena relembrar da conversão de um número decimal para a base binária.

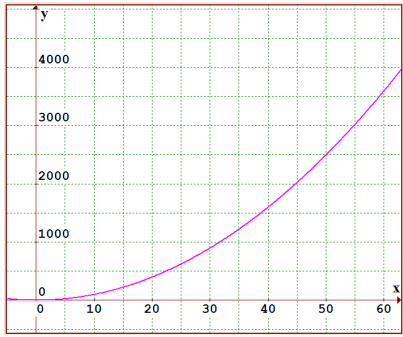
Se o número for inteiro, divide – se sucessivamente o número decimal por 2 e os quocientes que vão sendo obtidos até que o quociente seja 0 ou 1. A sequência de todos os restos obtidos dispostos na **ordem inversa** representa o número binário.

Exemplo: Conversão do número 10 em binário.

$$\begin{array}{r} 10 \ | 2 \\ \hline 0 \ 5 \ | 2 \\ \hline 1 \ 2 \ | 2 \\ \hline 0 \ 1 \ | 2 \\ \hline 1 \ 0 \end{array} \qquad 10_{(10)} = 1010_{(2)}$$

Exemplo 2. AG na busca por um ponto máximo da função: $f(x) = x^2$, sujeito as seguintes restrições: $\{ x \in \mathbb{Z} \text{ tq } 1 \leq x \leq 63 \}$.

graficamente temos:



Solução:

1) cromossomos binários com 6 bits

0 = 000000 --- 63 = 111111

2) Aptidão

Pode ser a própria função objetivo:

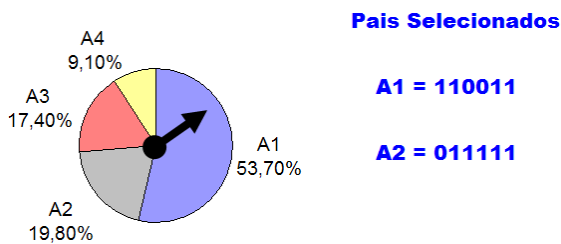
Exemplo:

Aptidão(010100) = $f(010100) = f(20) = 400$

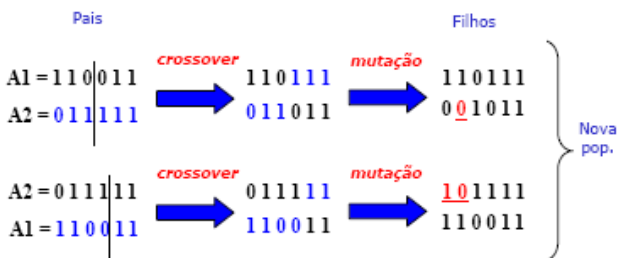
3) Cria-se uma população inicial de forma aleatória, como mostra a tabela abaixo:

cromossomos	x	f(x)	Probabilidade proporcional a aptidão
A1 = 110011	51	2601	53,7 %
A2 = 011111	31	961	19,8 %
A3 = 011101	29	841	17,4 %
A4 = 010101	21	441	9,1 %

4) Seleção Natural proporcional a aptidão (método da roleta)



5) Crossover e Mutação para obtenção da primeira geração



6) Análise da primeira Geração

cromossomos	x	f(x)	Probabilidade proporcional a aptidão
A1 = 110111	55	3025	29,0%
A2 = 110011	51	2601	24,9%
A3 = 110011	51	2601	24,9%
A4 = 101111	47	2209	21,2%

7) Obtenção de outras gerações

Segunda geração			
	filhos	x	f(x)
1	110111	55	3025
2	110000	48	2304
3	101011	43	1849
4	101111	47	2209

Terceira geração			
	filhos	x	f(x)
1	110111	55	3025
2	101111	47	2209
3	001011	11	121
4	100001	33	1089

Quarta geração			
	filhos	x	f(x)
1	110111	55	3025
2	101111	47	2209
3	111111	63	3969
4	101111	47	2209

Quinta geração			
	filhos	x	f(x)
1	111111	63	3969
2	111111	63	3969
3	111111	63	3969
4	101111	47	2209

Como já era previsto o valor máximo da função é $f(63) = 3969$, observado claramente na quinta geração.

Algumas Aplicações Práticas dos Algoritmos Genéticos

Os Algoritmos Genéticos, desde os conceitos básicos realizados por Holland, vêm sendo utilizados em várias áreas de pesquisa tais como:

- Controle de Sistemas Dinâmicos;
- Indução e Otimização de Bases de Regras;
- Encontrar Novas Topologias Conexionistas:
 - Engenharia de Sistemas Neurais Artificiais;
 - Modelagem de Estruturas Neurais Biológicas;
- Simulação de Modelos Biológicos;
- Evolução Interativa de Imagens;
- Composição Musical.

Em todas as áreas os AG's tem mostrado excelentes resultados, como descritos a seguir:

a) A AIS (Barcelona, Espanha) utilizou um sistema apoiado em AGs e Sistemas Especialistas (SEs) para programar os Jogos Paraolímpicos de 1992 já que nas Olimpíadas os atletas são organizados em duas grandes classes, masculino e feminino, e os competidores paraolímpicos são divididos em mais de 100 (cem) classes, segundo certas restrições médicas.

b) Um sistema em construção na *New Mexico State University* descreve imagens faciais de criminosos a partir de testemunhas do crime, utilizando AGs. O sistema tem se mostrado mais efetivo na produção de imagens aprimoradas de criminosos do que qualquer outra técnica de obtenção de informação de imagens.

c) Sponsler (1989) mostrou um sistema protótipo desenvolvido para avaliar a aplicabilidade dos Algoritmos Genéticos na otimização da programação do telescópio espacial *Hubble*.

d) No Hospital universitário da UFSC, em Florianópolis, os Algoritmos Genéticos foram utilizados para auxiliar na elaboração de uma escala de trabalho dos médicos plantonistas da maternidade. O objetivo pretendido foi o de auxiliar na solução da escala de trabalho dos médicos, bem como diminuir o esforço e o desgaste humanos para a confecção do plantão.

e) Blanchard mostrou em congresso sobre inteligência computacional o caso da *US West*, uma companhia regional de telecomunicações do estado do Colorado, que vem usando um sistema baseado em AGs que possibilita projetar, em duas horas, redes óticas especializadas, trabalho que levaria seis meses utilizando especialistas humanos. O sistema produz resultados ainda 10% (dez por cento) melhores que os realizados pelo homem. A companhia estima que o sistema possibilitará uma economia de 100 milhões de dólares até o final do século.

Para saber mais:

- IEEE transactions on Industrial Electronics, VOL 43, N° 5, October 1996 --
- Genetic Algorithms: Concepts e Applications – K.F. Man , K.S Tang and S. Kwong
- http://www2.dem.inpe.br/ijar/gene1.html

Samuel Liló Abdalla – Engenheiro Eletrônico pela UNESP – Matemático – Pós-graduado em Computação Aplicada pelo ITA - Professor de Matemática e Computação