

2 Fundamentos

2.1 Mecânica Quântica

2.1.1 Introdução

Quando a luz incide sobre uma superfície de metal, elétrons são arrancados desta mesma superfície. Este problema não pode ser perfeitamente explicado pela mecânica clássica devido a três fatores (Green01):

- Não importa o quão baixa é a intensidade da luz, os elétrons são arrancados instantaneamente da superfície do metal. De acordo com a física clássica, os elétrons precisariam acumular durante um certo tempo a energia necessária para escapar da superfície do metal, já que a luz deveria ter sua energia distribuída uniformemente ao longo da onda;
- Existe uma frequência de corte para o feixe incidente de luz, abaixo da qual nenhum elétron é emitido. De acordo com as leis de Maxwell, a energia transportada pela luz depende apenas da intensidade da onda eletromagnética e não diretamente da frequência;
- Se a energia cinética máxima dos elétrons emitidos for relacionada através de um gráfico com a frequência da luz incidente, uma linha reta será obtida. Como a energia deveria estar relacionada apenas com a amplitude da onda, de acordo com as leis de Maxwell, não deveria existir uma relação linear entre frequência e energia.

A explicação para estes fenômenos foi proposta em 1900 por Planck (Planck01, Green01) ao abandonar o conceito clássico para o comportamento da luz apenas como uma onda eletromagnética. Planck demonstrou que, se a energia da luz estiver concentrada em pacotes (ou *quanta*) de energia $E = h\nu$ (onde ν é a frequência da luz e h é a constante de Planck), os problemas descritos anteriormente podem ser, então, facilmente explicados. Através desta explicação, a luz passa a ter, em determinadas situações, um comportamento de uma partícula (chamada *fóton*).

2.1.2

Funções de Onda

A idéia de que a luz se comporta em alguns casos como uma onda e em alguns casos como uma partícula leva imediatamente à questão sobre se o fóton tem outras propriedades comuns às partículas. Como energia e massa estão relacionadas pela equação $E = mc^2$, o fóton (que se move com a velocidade da luz c) tem uma massa relativística $m = hv/c^2$. Portanto, como o fóton tem massa e velocidade, também deve ter um momento $p = mc = hv/c = h/\lambda$, onde λ é o comprimento de onda da luz.

De maneira inversa, assim como os fótons têm um momento característico h/λ , outras partículas (por exemplo, um elétron), que também têm um momento característico, devem possuir um comprimento de onda relacionado a este momento $\lambda = h/p$. Se uma partícula tem um comprimento de onda característico, então deve ser possível representar este comprimento de onda matematicamente da mesma maneira que uma onda é representada, ou seja, usando-se uma função de onda. Na física clássica, uma onda estacionária com comprimento de onda λ , propagando-se na direção positiva do eixo x , pode ser representada pela equação 2-1 (Green01).

$$\psi(x) = \exp(i2\pi x/\lambda) = \cos(2\pi x/\lambda) + i \sin(2\pi x/\lambda) \quad (2-1)$$

Onde $i = \sqrt{-1}$. Substituindo $\lambda = h/p$, obtém-se a equação (2-2).

$$\psi(x) = \exp(ipx/\hbar) \quad (2-2)$$

Onde $\hbar = h/2\pi$. A partir da função de onda, é possível determinar a probabilidade de que a partícula esteja em um determinado ponto do espaço ao se tentar observá-la. A densidade de probabilidade para a localização de uma partícula com função de onda ψ deve ser o quadrado da amplitude da função de onda, ou seja, $|\psi|^2$ (Gillespie73, Green01). Em uma dimensão, uma partícula, representada pela função de onda $\psi(x)$, terá uma densidade de probabilidade de ser encontrada no intervalo $x + dx$ igual a:

$$|\psi(x)|^2 dx = \psi^* \psi dx \quad (2-3)$$

Onde ψ^* representa o conjugado complexo de ψ . Em três dimensões, a densidade de probabilidade da partícula é $|\psi(\mathbf{r})|^2$ (onde \mathbf{r} é um vetor de três dimensões (x, y, z)). Isto significa que a probabilidade de se encontrar a partícula no elemento infinitesimal de volume $d\tau = dx dy dz$ no ponto \mathbf{r} é $|\psi(\mathbf{r})|^2 d\tau$. A integral deste valor sobre todo o espaço onde a partícula pode ser encontrada fornece a probabilidade de se encontrar esta partícula em algum lugar do espaço: conseqüentemente, este valor deve ser igual a 1 (equação 2-4).

$$\int_{-\infty}^{\infty} |\psi|^2 d\tau = 1 \quad (2-4)$$

As funções de onda introduzem um conceito importante e pouco intuitivo: a relação entre a função de onda e a localização da partícula é, por natureza, probabilística (Green01). Em um feixe de luz, por exemplo, a função de onda está espalhada por toda a frente de onda de forma homogênea mas, quando se tenta detectar os fótons, nota-se que os mesmos chegam de forma imprevisível ao detector, com uma densidade de probabilidade proporcional à intensidade do feixe de luz ou proporcional ao quadrado da amplitude da onda. Em outras palavras, o valor das propriedades da partícula que podem ser representadas por funções de onda (por exemplo, a posição da partícula) não pode ser conhecido à priori; sem efetivamente medir a propriedade que se deseja, tudo que se pode afirmar sobre esta propriedade é que ela tem diferentes probabilidades de assumir diferentes valores quando for observada. Assim, antes de se observar através de algum instrumento, por exemplo, a posição de um elétron confinado dentro de alguma região do espaço, o máximo que se pode determinar é a probabilidade de observá-lo em alguma região deste espaço de confinamento.

As funções de onda devem obedecer algumas propriedades. Estas propriedades são:

1. ψ deve ser finita;
2. ψ deve ser contínua;
3. ψ deve ser derivável duas vezes;
4. ψ deve ser integrável em todo o espaço.

2.1.3 Operadores Hermitianos

As informações contidas nas funções de onda são as únicas de que se pode dispor a respeito de um sistema físico microscópico. Como já foi visto, é possível manipular estas funções de onda de modo a se obter a distribuição de probabilidade espacial do sistema físico representado por esta função. No entanto, esta informação espacial não é a única que se deseja obter sobre o sistema físico. Deseja-se também ser capaz de medir quantidades como o momento, a energia e o momento angular de uma partícula. Apesar da função de onda não poder ser medida diretamente, todas essas propriedades mecânicas podem ser medidas diretamente e são chamadas, na mecânica quântica, de *observáveis* (em inglês, *observables*) (Gillespie73). Os observáveis são representados matematicamente por *operadores*. Alguns exemplos de observáveis são a posição e o momento de uma partícula.

Um operador é um objeto matemático que age sobre uma função transformando-a em outra função. Um exemplo é o operador $\hat{\partial}_x$, que realiza a

operação de “diferenciar com respeito a variável x ”. Este operador transforma uma função $f(x)$ em sua derivada. Neste trabalho, todos os operadores serão distinguidos de outros objetos matemáticos através do uso de um sinal $\hat{}$ sobre o operador.

Apesar de, em geral, os operadores transformarem uma função em outra, em alguns casos especiais um operador pode transformar uma função em um múltiplo dela mesma. Um exemplo é o operador $\hat{\partial}_x$ que transforma a função $e^{\alpha x}$ em $\alpha e^{\alpha x}$, ou seja, a mesma função multiplicada por uma constante α . A função que é transformada em um múltiplo dela mesma é chamada de autofunção (ou autovetor) e a constante multiplicativa introduzida pelo operador é chamada de autovalor.

Os conceitos de autofunção e autovalor são extremamente importantes na mecânica quântica. Isto se deve ao fato de que todos os observáveis são representados por operadores, sendo que os autovalores, os quais podem ser encontrados através da aplicação de um operador a uma função, são os únicos valores que podem efetivamente ser observados ao se fazer uma medida. Em outras palavras, qualquer medição de um observável deve encontrar um dos possíveis autovalores do operador (Gillespie73). Por exemplo, os níveis de energia de um átomo são autovalores do operador de energia; qualquer medida realizada para se determinar o nível de energia de um átomo deve encontrar um desses autovalores. No entanto, é importante destacar que, mesmo na mecânica quântica, os valores possíveis para os autovalores de um operador não precisam, necessariamente, ser discretos, ou seja, é possível que para certos observáveis, os autovalores do operador correspondente formem um conjunto contínuo.

Os operadores que representam quantidades observáveis devem respeitar duas restrições: linearidade e hermiticidade. Um operador \hat{A} é linear quando, para quaisquer valores constantes α e β , e para quaisquer funções de onda $\psi(x)$ e $\phi(x)$, a equação 2-5 for válida.

$$\hat{A}(\alpha\psi(x) + \beta\phi(x)) = \alpha\hat{A}\psi(x) + \beta\hat{A}\phi(x) \quad (2-5)$$

Um operador \hat{A} é hermitiano se, para quaisquer funções de onda $\psi(x)$ e $\phi(x)$, a equação 2-6 for válida.

$$\int_{-\infty}^{\infty} \psi^*(x)\hat{A}\phi(x)dx = \int_{-\infty}^{\infty} \hat{A}\psi^*(x)\phi(x)dx \quad (2-6)$$

Duas importantes conseqüências advêm destas restrições (Gillespie73): a primeira é que os autovalores de um operador hermitiano são números reais; a segunda é que as autofunções de um operador hermitiano correspondentes a autovalores diferentes, são ortogonais.

Supondo-se então que um observável \mathcal{A} é representado por um operador \hat{A} e que uma função de onda ψ é uma autofunção de \hat{A} com autovalor α (ou seja, $\hat{A}\psi = \alpha\psi$), então o sistema tem um valor definido α para o observável \mathcal{A} . Uma

medida de \mathcal{A} irá resultar no valor α com probabilidade 1. Por outro lado, se a função de onda ψ não for uma autofunção de \hat{A} , então uma medida de \mathcal{A} pode produzir qualquer um de diferentes autovalores do operador \hat{A} com diferentes probabilidades. O valor esperado de um observável \mathcal{A} , denotado por $\langle A \rangle$, é definido pela equação 2-7.

$$\langle A \rangle = \int \psi^* \hat{A} \psi d\tau \quad (2-7)$$

De modo a simplificar a notação matemática, pode-se usar a notação de Dirac para estes problemas (Green01). Neste caso, a função de onda ψ passa a ser representada pelo símbolo $|\psi\rangle$ (chamado *ket*). O conjugado complexo ψ^* da função de onda é representado pelo símbolo $\langle\psi|$ (chamado *bra*). As integrais do tipo $\int \phi^* \psi d\tau$ são representadas pelo símbolo $\langle\phi|\psi\rangle$ (é importante notar que este símbolo implica na integração da função, ou seja, $\langle\phi|\psi\rangle = \int \phi^* \psi d\tau$). Portanto, uma função de onda ψ está normalizada se $\langle\psi|\psi\rangle = 1$.

A principal característica que se deseja ressaltar com relação às funções de onda e aos operadores, é a possibilidade de se representar as funções de onda como uma expansão por autofunções. Em outras palavras, além de serem ortogonais, as autofunções de um operador hermitiano formam uma base ortonormal completa que permite escrever qualquer função de onda para o sistema físico como uma superposição (combinação linear) destas autofunções. Ou seja, se um operador \hat{A} possui um conjunto de autofunções $|\psi_i\rangle$ com autovalores α_i (ou seja, $\hat{A}|\psi_i\rangle = \alpha_i|\psi_i\rangle$), então qualquer função de onda ϕ válida para o sistema pode ser representada pela superposição de autofunções $|\psi_i\rangle$ conforme mostrado na equação 2-8.

$$\phi = \sum_i c_i |\psi_i\rangle \quad (2-8)$$

Onde $c_i^* c_i$ (ou $|c_i|^2$) é a probabilidade de se medir o autovalor α_i . Isto significa que, ao fazer uma medida de um estado quântico, a função de onda é projetada em uma das autofunções que formam a base ortonormal e o autovalor correspondente a esta autofunção é o valor observado. Em outras palavras, a função de onda pode ser escrita como uma combinação linear dos autovetores associados à um determinado operador. Os coeficientes associados aos autovetores nesta combinação linear permitem determinar a probabilidade de que a função de onda seja projetada sobre cada um dos autovetores quando a observação for realizada. Vale ressaltar que, como $|c_i|^2$ representa a probabilidade de um autovalor ser observado, e como a base ortonormal é completa, $\sum_i |c_i|^2 = 1$.

2.2

Um Exemplo de um Sistema Físico Quântico – O Modelo da Partícula na Caixa

Esta seção tem como objetivo mostrar um exemplo de um sistema físico quântico hipotético (hipotético no sentido de não ser possível reproduzir as condições ideais nas quais o experimento deve ser realizado) de modo a ilustrar os conceitos da mecânica quântica apresentados anteriormente. Além disso, o sistema físico aqui apresentado será usado, posteriormente, como um modelo físico que permite implementar o algoritmo genético com inspiração quântica apresentado neste trabalho.

Considere-se uma partícula (um elétron por exemplo) confinada dentro de uma caixa cujas paredes têm um potencial infinito. Além disso, o movimento desta partícula está limitado a uma dimensão (o sentido da largura da caixa). Um diagrama deste modelo pode ser visto na figura 2.1 (nesta figura, $V = \infty$ indica a região de potencial infinito).

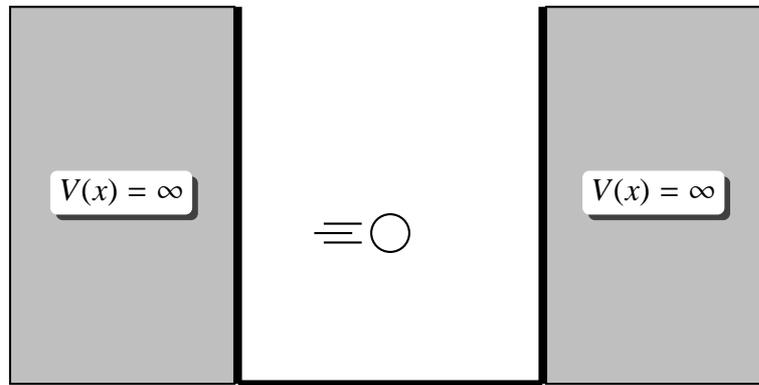


Figura 2.1: Diagrama do sistema físico hipotético da “partícula na caixa”

Para o problema unidimensional (ou seja, a partícula pode se mover apenas sobre o eixo x), pode-se usar a equação de Schrödinger independente do tempo (Green01) para calcular a função de onda. Esta equação pode ser escrita como em 2-9.

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \quad (2-9)$$

Onde \hbar é a constante de Planck reduzida, m é a massa da partícula, $\psi(x)$ é a função de onda estacionária da partícula (esta é a função que se deseja encontrar), $V(x)$ determina o valor da energia potencial em função da posição e E é a energia da partícula.

Supondo-se que se deseja encontrar a função de onda para uma partícula em uma situação onde as barreiras de potencial estejam posicionadas, respectivamente, em $x = 0$ e $x = L$, pode-se considerar que a equação anterior, dentro do intervalo $[0, L]$ para x assume a forma mostrada em 2-10, já que o potencial dentro dessa região é igual a zero.

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} = E\psi(x) \quad (2-10)$$

Esta equação diferencial é bem conhecida e sua solução é da forma:

$$\psi(x) = A \sin(kx) + B \cos(kx) \quad (2-11)$$

$$E = \frac{k^2\hbar^2}{2m} \quad (2-12)$$

Nestas equações, A e B podem ser quaisquer números complexos e o número k deve ser um número real (já que a energia E deve ser um número real). De modo a definir a solução específica para este problema, deve-se determinar as condições de contorno e encontrar os valores de A e B que satisfaçam essas condições. Como já foi mencionado anteriormente, a equação deve garantir que a função de onda seja contínua ao longo de todo o eixo real. Como a partícula não pode estar na região em que a barreira de potencial é infinita, deduz-se que, quando x tende a 0 ou quando x tende a L , a probabilidade $|\psi(x)|^2$ de encontrar uma partícula nesta região deve tender a zero. Isto está de acordo com a intuição física já que, ao se aproximar da barreira de potencial, a repulsão sobre a partícula aumenta. Assim, $\psi(0) = \psi(L) = 0$. De acordo com a equação 2-11, para que $\psi(0) = 0$ o termo com o cosseno deve desaparecer (já que $\sin(0) = 0$ e $\cos(0) = 1$) e, portanto, B deve ser igual a zero. Já para a situação em que $x = L$, $\psi(L) = A \sin(kL) = 0$. Para o caso não-trivial em que $A \neq 0$, $\sin(kL) = 0$ somente se $k = n\pi/L$, onde $n \in \mathbb{N}^+$.

Finalmente, para se encontrar o valor de A , deve-se lembrar que a função de onda deve ser normalizada, conforme mostrado na equação 2-4.

$$1 = \int_{-\infty}^{\infty} |\psi(x)|^2 dx = |A|^2 \int_0^L \sin^2(kx) dx = |A|^2 \frac{L}{2} \quad (2-13)$$

Rearranjando a equação anterior tem-se:

$$|A| = \sqrt{\frac{2}{L}} \quad (2-14)$$

Assim, A pode ser qualquer número complexo cujo valor absoluto $|A|$ seja igual a $\sqrt{2/L}$. Como diferentes valores de A produzem o mesmo estado físico, e para fins de simplificação, define-se $A = \sqrt{2/L}$.

Substituindo-se estes valores nas equações 2-11 e 2-12 chega-se as equações:

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \quad (2-15)$$

$$E_n = \frac{n^2\hbar^2\pi^2}{2mL^2} = \frac{n^2h^2}{8mL^2} \quad (2-16)$$

Onde $n \in \mathbb{N}^+$. O valor de n determina o número de picos e vales da senóide, como pode ser visto na figura 2.2.

A função de onda deve ser estacionária e, por este motivo, o número n deve ser inteiro e maior do que zero. Isto, juntamente com a equação 2-16, explica porque

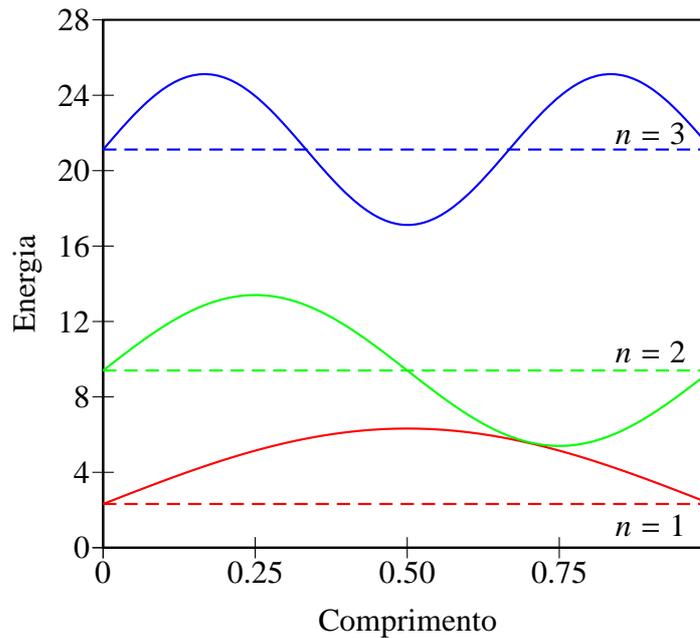


Figura 2.2: Gráfico mostrando os níveis de energia e a forma das funções de onda para $n = 1$, $n = 2$ e $n = 3$.

os níveis de energia da partícula são discretos.

Por sua vez, a função densidade de probabilidade $p(x)$ para este sistema físico é dada pela equação:

$$p(x) = |\psi(x)|^2 = \frac{2}{L} \sin^2\left(\frac{n\pi x}{L}\right) \quad (2-17)$$

O gráfico desta equação quando $n = 1$ é mostrado na figura 2.3.

O resultado observado neste gráfico é o mesmo esperado intuitivamente para o caso em que $n = 1$. A partícula tem mais probabilidade de ser observada na parte central da caixa do que nas bordas da mesma (onde a distância da barreira de potencial infinito começa a ser menor).

2.3

Computação Quântica

Um computador quântico é um dispositivo que faz uso direto de certos fenômenos da mecânica quântica para realizar operações com dados. Estes fenômenos permitem construir, em teoria, computadores que obedecem novas leis mais permissivas de complexidade computacional (Spector04). O termo “computação quântica” é, portanto, usado para descrever processos computacionais que se baseiam nestes fenômenos específicos e que são, assim, capazes de diminuir o esforço e a complexidade para se resolver determinados problemas. A principal perspectiva que se

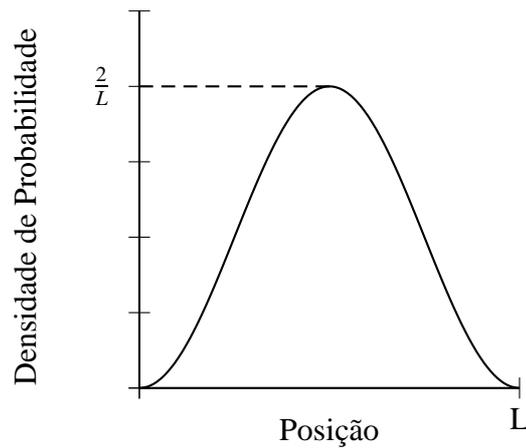


Figura 2.3: Gráfico da função densidade de probabilidade para o sistema da “partícula na caixa” quando $n = 1$

abre com o uso de computadores quânticos em relação ao seu poder de processamento é o fato de que “as possibilidades contam, mesmo que elas nunca venham a acontecer” (Spector04).

Em um computador clássico, o *bit* é a menor unidade de informação, podendo assumir os valores 0 ou 1. Em um computador quântico, a unidade de informação básica, chamada de *q-bit*, pode assumir os estados $|0\rangle$, $|1\rangle$ ou uma superposição dos dois estados. Esta superposição de dois estados é uma combinação linear dos estados $|0\rangle$ e $|1\rangle$ e pode ser representada pela equação 2-18.

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle \quad (2-18)$$

Onde $|\psi\rangle$ é o estado do *q-bit* e, conforme discutido na seção anterior, $|\alpha|^2$ e $|\beta|^2$ são as probabilidades de que os estados 0 ou 1 sejam observados quando o *q-bit* for medido.

Ao ser observado, o bit quântico será trazido para o nível clássico e o estado observado será o valor 0 ou 1. Esta superposição de estados oferece um grau de paralelismo incomparável aos computadores quânticos que, se bem explorado, permite que estes computadores realizem tarefas praticamente impossíveis de serem realizadas em computadores clássicos. Um exemplo de uma tarefa deste tipo é a fatoração inteira de grandes números. Neste problema, deseja-se encontrar os dois números primos cujo produto seja igual a um número x fornecido. Esta tarefa pode levar milhões de anos para ser resolvida em computadores clássicos usando os algoritmos conhecidos mais sofisticados. Em um computador quântico, no entanto, a mesma tarefa levaria apenas alguns segundos para ser concluída. Uma discussão mais detalhada sobre o funcionamento dos computadores quântico pode ser encontrada em (Rieffel00).

2.4

Algoritmos com Inspiração Quântica

Embora a computação quântica ofereça uma boa promessa em termos de capacidade de processamento, dois problemas impedem, atualmente, que a mesma se torne uma ferramenta útil: a dificuldade de se implementar um computador quântico e a dificuldade de se criar algoritmos que tirem proveito da capacidade de processamento destes computadores. Deste modo, as pesquisas na área de computação quântica se concentram, atualmente, em dois pontos:

- Desenvolvimento de novos algoritmos que sejam mais eficientes nos computadores quânticos do que os algoritmos equivalentes para computadores clássicos;
- Desenvolvimento de um *hardware* que torne viável o uso dos computadores quânticos.

No entanto, no artigo (Moore95), uma nova abordagem é proposta. Ao invés de desenvolver novos algoritmos para computadores quânticos ou de tentar viabilizar o uso destes, a idéia de *computação com inspiração quântica* é apresentada. O objetivo desta nova abordagem é criar algoritmos clássicos (capazes, portanto, de serem executados em computadores clássicos) que tirem proveito dos paradigmas da física quântica, de forma a melhorar o desempenho dos mesmos na resolução de problemas.

Uma metodologia para a formulação inicial de um algoritmo com inspiração quântica também é apresentada em (Moore95). Esta metodologia consiste dos seguintes passos:

1. O problema deve ter uma representação numérica ou, caso não tenha, um método para sua conversão em representação numérica deve ser empregado;
2. A configuração inicial deve ser determinada;
3. Uma condição de parada deve ser definida;
4. O problema deve poder ser dividido em sub-problemas menores;
5. O número de universos (ou estados de superposição) deve ser identificado;
6. Cada sub-problema deve ser associado a um dos universos;
7. Os cálculos nos diferentes universos devem ocorrer de forma independente;
8. Alguma forma de iteração entre os múltiplos universos deve existir. Esta interferência deve, ou permitir encontrar a solução para o problema, ou

fornecer informação para cada sub-problema em cada universo ser capaz de encontrá-la.

Um exemplo de um algoritmo com inspiração quântica é o “Algoritmo de Ordenação com Inspiração Quântica” (Moore95), baseado no algoritmo de “ordenação por seleção”. Este algoritmo é usado para resolver o problema de ordenação numérica onde deseja-se ordenar uma lista com n elementos em ordem ascendente ou descendente. O algoritmo funciona da seguinte maneira:

1. Denomina-se a lista que se quer ordenar por L_1 (esta lista deve conter pelo menos dois elementos) e inicializa-se uma segunda lista L_2 sem nenhum elemento;
2. Encontra-se o menor elemento (caso se esteja ordenando de forma ascendente) em L_1 e após remover este elemento da lista L_1 acrescenta-se o mesmo à lista L_2 ;
3. Repete-se o passo anterior até que L_1 esteja vazia. L_2 contém agora todos os elementos em ordem crescente.

Quando implementado de maneira eficiente, o tempo de execução deste algoritmo é da ordem de $O\left(\frac{1}{2}N(N-1)\right)$ (Knuth73), onde N é o tamanho inicial de L_1 . Por exemplo, caso se queira ordenar uma lista com 16 valores, serão necessários, em média, 120 buscas à lista para encontrar o menor elemento. Usando elementos inspirados na física quântica, este algoritmo pode ser melhorado da seguinte maneira (Moore95) (usa-se aqui, como exemplo, uma lista L_1 contendo 16 elementos $L_1 = (16, 14, 12, 3, 8, 4, 15, 7, 11, 6, 2, 5, 1, 10, 9, 13)$):

1. Fixar $n =$ número de elementos na lista = 16;
2. Fixar $u =$ número de superposições (universos) necessárias = $\sqrt{n} = 4$;
3. Fixar $l =$ número de elementos da lista em cada universo = $u = 4$;
4. Divide-se a lista L_1 em 4 partes: os primeiros quatro números são colocados no universo 1, os próximos 4 números são colocados no universo 2, e assim por diante. Desta forma, para este exemplo, pode-se construir a tabela 2.1 com os universos e os elementos que estão armazenados em cada um deles;
5. Usa-se o algoritmo de ordenação por seleção em cada um dos universos, separadamente. Deste modo, obtém-se a tabela 2.2.

u_0	16	14	12	3
u_1	8	4	15	7
u_2	11	6	2	5
u_3	1	10	9	13

Tabela 2.1: Números da lista a ser ordenada distribuídos pelos múltiplos universos do modelo com inspiração quântica

u_0	3	12	14	16
u_1	4	7	8	15
u_2	2	5	6	11
u_3	1	9	10	13

Tabela 2.2: Números da lista após uma primeira ordenação realizada em cada um dos 4 universos

6. Realiza-se uma *interferência diagonal*: os universos são rearranjados pegando-se as diagonais que têm, como primeiro elemento, os números da primeira coluna da matriz. Estes elementos são reordenados e reposicionados na matriz. Em outras palavras, o primeiro universo irá conter os elementos da diagonal principal (3, 7, 6, 13), o segundo universo irá conter os elementos formados pela diagonal que se inicia na primeira coluna da segunda linha (4, 5, 10, 16), o terceiro universo irá conter os elementos formados pela diagonal que se inicia na primeira coluna da terceira linha (2, 9, 14, 15) e, finalmente, o quarto universo irá conter os elementos formados pela diagonal que se inicia na primeira coluna da quarta linha (1, 12, 8, 11). A matriz obtida por essa interferência diagonal é novamente processada pelo algoritmo de ordenação, separadamente por universo. Assim, forma-se a tabela 2.3.

u_0	3	6	7	13
u_1	4	5	10	16
u_2	2	9	14	15
u_3	1	8	11	12

Tabela 2.3: Matriz de universos após a interferência diagonal

7. Em seguida, realiza-se a *interferência vertical*, onde cada coluna da matriz de universos é ordenada e transposta. Com isto, obtém-se a tabela 2.4.

u_0	1	2	3	4
u_1	5	6	8	9
u_2	7	10	11	14
u_3	12	13	15	16

Tabela 2.4: Matriz de universos após a interferência vertical

8. Checa-se o último valor do universo u_i e verifica-se se é menor ou igual ao primeiro valor de u_{i+1} para $i = 0, 1, \dots, n - 1$. Caso não seja, repete-se os passos 6 e 7 novamente até que esta condição seja satisfeita.

O tempo de execução deste algoritmo é da ordem de $\sqrt{N}[N + 2(\sqrt{N} - 1)]$ (Moore95). Portanto, uma lista com 16 elementos precisará, em média, de 88 buscas na lista usando-se a ordenação por seleção, o que é significativamente menor do que o número de buscas (120) necessárias para a execução do algoritmo sem inspiração quântica.

No artigo (Moore95), também é apresentado um algoritmo com inspiração quântica para determinar a solução para o problema do *15-puzzle*. Uma explicação mais detalhada sobre estes algoritmos e sobre os algoritmos com inspiração quântica pode ser encontrada em (Moore95).

2.5

Algoritmos Evolutivos com Inspiração Quântica – Representação Binária

O algoritmo evolutivo com inspiração quântica usando representação binária, que neste trabalho será designado pela sigla AEIQ- \mathcal{B} (*Quantum-Inspired Evolutionary Algorithm* – QEA na sigla original), foi proposto inicialmente em (Han00). Neste modelo, o algoritmo evolutivo proposto é caracterizado por um cromossomo, uma função de avaliação e uma dinâmica populacional. Entretanto, ao invés de uma representação binária convencional, este algoritmo usa uma representação especial que simula um cromossomo formado por q -bits. Esta representação é feita da seguinte forma: um q -bit é formado por um par de números (α, β) , como na equação 2-19.

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (2-19)$$

Onde $|\alpha|^2 + |\beta|^2 = 1$. O valor dado por $|\alpha|^2$ indica a probabilidade de que o q -bit terá o valor 0 quando for observado e o valor $|\beta|^2$ indica a probabilidade de que o q -bit terá o valor 1 quando for observado. Pode-se visualizar graficamente esta relação entre α e β , conforme mostrado na figura 2.4.

A partir da definição do q -bit acima, pode-se definir um indivíduo quântico formado por m q -bits, conforme a equação 2-20.

$$\left[\begin{array}{c|c|c|c} \alpha_1 & \alpha_2 & \dots & \alpha_m \\ \beta_1 & \beta_2 & \dots & \beta_m \end{array} \right] \quad (2-20)$$

De forma semelhante à mostrada na equação 2-19, $|\alpha_i|^2 + |\beta_i|^2 = 1$ onde $i = 1, 2, 3, \dots, m$.

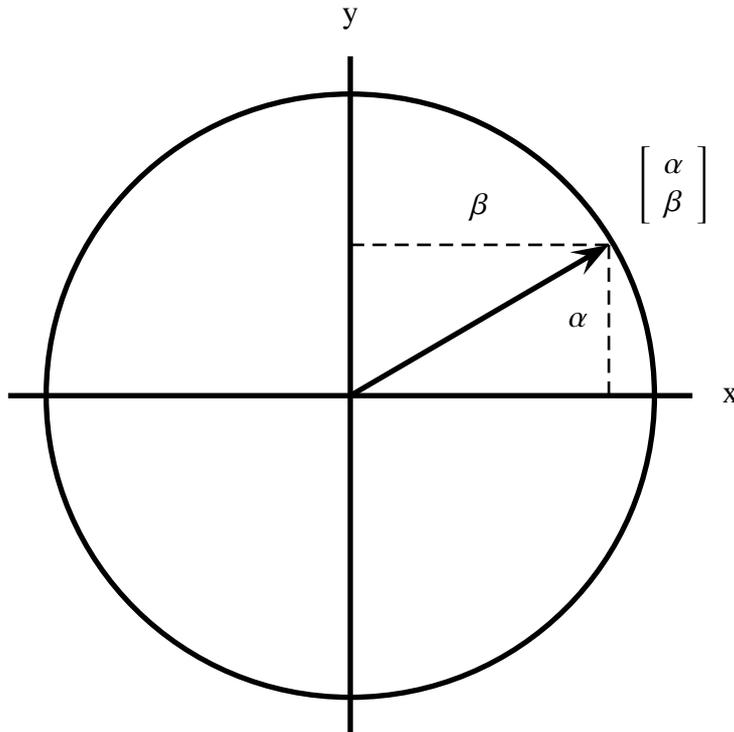


Figura 2.4: Representação gráfica de um círculo de raio unitário com as amplitudes de probabilidade de se observar os valores 0 e 1 para um q -bit qualquer, no caso em que α e β são reais.

Deste modo, cada indivíduo quântico representa uma superposição de indivíduos formados por m genes. Seja, por exemplo, um indivíduo quântico formado por 3 q -bits com os valores mostrados na equação 2-21.

$$\left[\begin{array}{c|c|c} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{\sqrt{3}}{2} \end{array} \right] \quad (2-21)$$

Para se calcular a *amplitude de probabilidade* (o módulo da amplitude de probabilidade ao quadrado é igual à probabilidade de se observar o estado) do estado $|000\rangle$, multiplica-se as amplitudes de probabilidade de se observar os estados 0 em cada um dos bits (α_1, α_2 e α_3). Para calcular a amplitude de probabilidade de se observar o estado $|001\rangle$, multiplica-se as amplitudes de probabilidade de se observar estes bits (α_1, α_2 e β_3). Estendendo-se isto para as outras possibilidades de observação dos estados, a superposição dos mesmos pode ser representada pela equação 2-22.

$$\frac{1}{4} |000\rangle + \frac{\sqrt{3}}{4} |001\rangle + \frac{1}{4} |010\rangle + \frac{\sqrt{3}}{4} |011\rangle + \frac{1}{4} |100\rangle + \frac{\sqrt{3}}{4} |101\rangle + \frac{1}{4} |110\rangle + \frac{\sqrt{3}}{4} |111\rangle \quad (2-22)$$

Este resultado significa que, cada um dos possíveis estados deste indivíduo, tem a probabilidade de ser observado mostrada na tabela 2.5.

O AEIQ- \mathcal{B} , proposto em (Han00), é definido como na figura 2.5. Neste

Estado	Probabilidade
000	$\frac{1}{16}$
001	$\frac{3}{16}$
010	$\frac{1}{16}$
011	$\frac{3}{16}$
100	$\frac{1}{16}$
101	$\frac{3}{16}$
110	$\frac{1}{16}$
111	$\frac{3}{16}$

Tabela 2.5: Probabilidades de observação de cada um dos possíveis estados do indivíduo quântico.

algoritmo, $Q(t)$ é uma população com um ou mais indivíduos quânticos. No passo de inicialização, estes indivíduos são inicializados de modo que os valores α_i e β_i (onde $i = 1, 2, 3, \dots, j$ e j é o comprimento do indivíduo quântico) sejam iguais a $\frac{1}{\sqrt{2}}$. Isto significa que, na geração inicial, todos os genes dos indivíduos terão a mesma probabilidade de gerarem os estados 0 ou 1.

A atualização da população, conforme indicado no algoritmo da figura 2.5, é realizada pelo operador $q - gate$, o qual é definido como a matriz de rotação dada pela equação 2-23.

$$U(\Delta\theta_i) = \begin{bmatrix} \cos(\Delta\theta_i) & -\sin(\Delta\theta_i) \\ \sin(\Delta\theta_i) & \cos(\Delta\theta_i) \end{bmatrix} \quad (2-23)$$

Esta matriz deve ser usada para multiplicar cada uma das colunas do indivíduo quântico, ou seja, os valores α_i e β_i são tratados como um vetor bidimensional e são rotacionados usando a matriz especificada. Em outras palavras, considerando-se o indivíduo quântico $Q(t) = \{(\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_i, \beta_i), \dots, (\alpha_j, \beta_j)\}$ (onde $i = 1, 2, 3, \dots, j$ e j é o comprimento do indivíduo quântico), a atualização deste indivíduo é feita de acordo com a equação 2-24.

$$\begin{bmatrix} \alpha'_i \\ \beta'_i \end{bmatrix} = U(\Delta\theta_i) \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} \quad (2-24)$$

Onde $i = 1, 2, 3, \dots, j$. Os valores de $\Delta\theta$ são definidos através de uma tabela, de modo que esta matriz de rotação seja capaz de modificar os valores de α_i e β_i aumentando as chances de que os indivíduos com as melhores avaliações sejam observados.

Finalmente, a estrutura $B(t)$ é usada para armazenar os melhores indivíduos gerados pelo algoritmo ao longo do processo evolutivo. Os últimos dois passos do algoritmo servem para armazenar os melhores indivíduos gerados na população atual com os melhores indivíduos criados nas gerações anteriores.

Este algoritmo com representação binária foi usado com sucesso em problemas de otimização combinatorial (Han00, Han02) e detecção de faces (Jang04),

```

iniciar
   $t \leftarrow 0$ ;
  inicializa  $Q(t)$ 
  gera  $P(t)$  observando estados de  $Q(t)$ 
  avalia  $P(t)$ 
  armazena as melhores soluções de  $P(t)$  em  $B(t)$ 
  enquanto não ocorrer condição de parada
     $t \leftarrow t + 1$ 
    gera  $P(t)$  observando estados de  $Q(t - 1)$ 
    avalia  $P(t)$ 
    atualiza  $Q(t)$  usando  $q$ -gate
    armazena as melhores soluções de  $B(t - 1)$  e  $P(t)$  em  $B(t)$ 
    armazena a melhor solução  $\mathbf{b}$  de  $B(t)$ 
  fim
fim

```

Figura 2.5: Pseudo-código do algoritmo evolutivo com inspiração quântica usando representação binária.

apresentando resultados superiores aos algoritmos genéticos convencionais em termos de tempo de convergência e qualidade das soluções encontradas. Uma descrição mais detalhada deste algoritmo pode ser encontrada em (Han00, Han02).

2.6

Algoritmos Culturais

Os algoritmos culturais são um tipo de algoritmo evolutivo que complementam a metáfora adotada pelos algoritmos genéticos tradicionais. Ao invés de usar apenas analogias dos conceitos de genética e de seleção natural, os algoritmos culturais são baseados também em teorias sociais e arqueológicas que modelam a evolução cultural. Estas teorias indicam que a evolução cultural pode ser vista como um processo de herança que ocorre em dois níveis: o nível micro-evolutivo e nível macro-evolutivo (Reynolds94).

No nível micro-evolutivo, os indivíduos são descritos em termos de características comportamentais (que podem ser socialmente aceitáveis ou não). Estas características são passadas de geração para geração usando uma série de operadores. No nível macro-evolutivo, os indivíduos são capazes de gerar "*mappas*" ou descrições gerais de suas experiências. Os *mappas* individuais podem ser agrupados através do uso de um conjunto de operadores (genéricos ou específicos para o problema). Os dois níveis evolutivos compartilham uma ligação através da qual podem se comunicar.

O objetivo do uso deste modelo em dois níveis é aumentar a velocidade de convergência do processo de otimização e permitir uma melhor resposta do algoritmo para uma gama maior de problemas.

2.6.1

Componentes de um Algoritmo Cultural

Os algoritmos culturais operam em dois espaços: o espaço de população, que modela as características comportamentais no nível micro-evolutivo, e o espaço de crenças, que modela os *mappas* de experiência no nível macro-evolutivo (Reynolds04). Inicialmente, os indivíduos no espaço de população são avaliados através do uso de uma função de desempenho *obj()*. Uma função de aceitação *acpt()* é então usada para determinar quais indivíduos do espaço de população irão influenciar o espaço de crenças. As características comportamentais destes indivíduos serão usadas para modificar as crenças existentes dentro do espaço de crenças, através de uma função de atualização *updt()*. Em seguida, as crenças que formam o espaço de crenças são usadas para influenciar a evolução da população através de uma função de influência *infl()*. Finalmente, alguns indivíduos da população são selecionados para a população da geração seguinte através de uma função de seleção *sel()*. As figuras 2.6 e 2.7 mostram, respectivamente, o algoritmo cultural em forma de pseudo-código e de um diagrama da estrutura geral de relacionamento entre os espaços de população e de crenças.

```

iniciar
     $t = 0$ ;
    inicializa espaço de crenças  $B^t$  e população  $P^t$ ;
    repetir
         $obj(P^t)$ ;
         $updt(B^t, acpt(P^t))$ ;
         $infl(B^t, P^t)$ ;
         $t = t + 1$ ;
         $P^t \leftarrow sel(P^{t-1})$ ;
    até que condição de parada alcançada;
fim

```

Figura 2.6: Pseudo-código do algoritmo cultural.

O espaço de crenças do algoritmo cultural é a região onde o conhecimento é armazenado. As cinco categorias básicas de conhecimento cultural, de acordo com (Reynolds04), são:

- *Conhecimento Normativo* – um conjunto de faixas de valores promissores dentro do espaço de busca que fornece padrões para os comportamentos individuais serem ajustados. O conhecimento normativo guia os indivíduos na direção das regiões promissoras, caso estes indivíduos já não estejam dentro destas regiões;
- *Conhecimento Situacional* – fornece um conjunto de “casos exemplares” que são úteis para a interpretação de experiências individuais específicas. Este

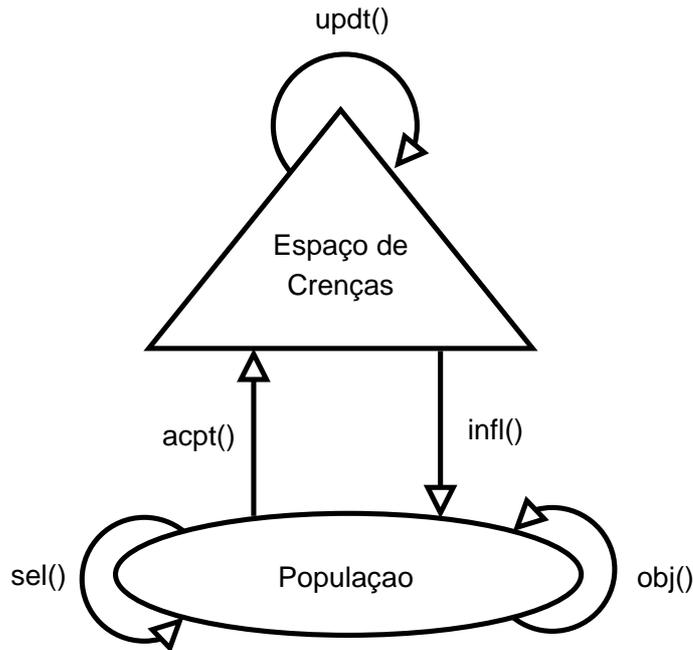


Figura 2.7: Diagrama da estrutura geral do algoritmo cultural.

tipo de conhecimento leva os indivíduos na direção dos melhores casos. Nos algoritmos genéticos convencionais (e também em várias implementações dos algoritmos culturais) este conhecimento é representado pelo melhor indivíduo da população;

- *Conhecimento Topográfico* – este conhecimento foi originalmente proposto para melhorar o conhecimento do processo de otimização sobre as estruturas locais do espaço de busca. O espaço de buscas é dividido em células e cada célula armazena o melhor indivíduo que está dentro dos seus limites. O conhecimento topográfico leva os indivíduos na direção dos melhores indivíduos locais;
- *Conhecimento do Domínio* – este tipo de conhecimento diz respeito ao domínio do problema que está sendo otimizado. Em geral, consiste em hibridizar, com o algoritmo cultural, alguma heurística particular ao problema que se quer otimizar;
- *Conhecimento Histórico* – o conhecimento histórico é usado, principalmente, em sistemas de aprendizado *online*. Ele é usado para armazenar o melhor indivíduo encontrado pelo algoritmo cultural antes de uma mudança na topologia da função de avaliação.

Os algoritmos culturais foram usados com sucesso em diversas áreas (Rychtycky03), como análise de fraudes em seguros de automóveis, precificação de estratégias para automóveis em um ambiente multi-agente e mineração de dados.

2.7

Neuro–Evolução

A neuro-evolução é um modelo híbrido que explora a potencialidade de duas diferentes áreas inspiradas em processos biológicos: Redes Neurais Artificiais (RNA) e Algoritmos Genéticos (AG). A idéia básica da neuro-evolução é buscar, automaticamente, a melhor configuração para uma rede neural usando algoritmos genéticos. Em outras palavras, a neuro-evolução combina a capacidade de generalização e aproximação de funções das redes neurais artificiais com um método eficiente de busca paralela. O objetivo dos algoritmos genéticos é melhorar os algoritmos de aprendizado, automatizando, total ou parcialmente, o processo de configuração da rede neural, bem como o processo de treinamento e atualização dos pesos da mesma.

Para a neuro-evolução funcionar, não é necessário que o sistema satisfaça nenhuma restrição em particular: o mesmo pode ser contínuo e parcialmente observável. No que concerne os métodos de aprendizado por neuro-evolução, a única exigência é que os mesmos possam, de alguma forma, ter as suas soluções candidatas avaliadas relativamente umas às outras. Se o problema é suficientemente regular, de modo que o mesmo possa ser resolvido, e a representação do fenótipo¹ é suficientemente poderosa, então estes métodos de aprendizado serão, muito provavelmente, capazes de encontrar uma solução.

As abordagens possíveis para os sistemas neuroevolutivos diferem uma da outra, basicamente, pelo modo como as mesmas codificam os pesos e a topologia das redes neurais nos cromossomos. Os cromossomos podem codificar qualquer informação relevante para a parametrização da rede neural, incluindo os pesos sinápticos, o número de camadas escondidas, a topologia da rede, a taxa de aprendizado, etc. A escolha do esquema de codificação tem um papel significativo na formação do espaço de buscas, no comportamento do algoritmo de busca e em como os genótipos devem ser transformados nos fenótipos (representação direta ou indireta).

No modelo descrito em (Gomez97) somente os pesos sinápticos são codificados, mantendo fixa a topologia e o número de neurônios da rede neural. Já os modelos SANE (*Symbolic, Adaptive Neuroevolution*) (Moriarty97), NEAT (*Neuroevolution of Augmenting Topologies*) (Stanley02) e ESP (*Enforced Sub-Populations*) (Gomez03) otimizam, além dos pesos sinápticos, a topologia e o número de neurônios da rede neural. Além disso, neste último trabalho, diversas aplicações de neuro–evolução em problemas de controle são apresentadas, com resultados promissores tanto com relação à capacidade de evoluir corretamente como com relação

¹Em biologia, o genótipo de um organismo vivo é constituído pelos genes e pela configuração dos mesmos, enquanto que o fenótipo é a expressão física do genótipo como, por exemplo, a cor dos olhos de uma pessoa.

ao desempenho do processo de otimização.

As principais desvantagens destes modelos estão relacionadas ao fato de que, ao usar modelos evolutivos baseados somente em uma população de indivíduos, o aprendizado do sistema pode ser muito lento, devido ao fato de que o espaço de buscas é mapeado pontualmente (cada indivíduo representa um único ponto no espaço de buscas). Isto pode retardar o processo de aprendizado e, conseqüentemente, tornar o aprendizado inviável em alguns sistemas que devem operar em tempo real ou que necessitem de atualizações *online*. Além disso, por se tratarem de algoritmos de busca global, os mesmos são mais ineficientes do que os algoritmos de busca baseados em gradiente. Finalmente, o número de pesos de uma rede neural cresce quadraticamente com o número de entradas. Isto implica no uso de indivíduos mais longos e, conseqüentemente, uma população maior para que a otimização seja bem-sucedida. O uso mais promissor de algoritmos genéticos para o treinamento de redes neurais foi feito usando-se os mesmos para encontrar os pesos iniciais da rede neural e posteriormente, através de um outro algoritmo, realizar o aprendizado propriamente dito (Skinner95). Este processo se mostrou eficiente pois o uso de algoritmos genéticos supre a deficiência que os algoritmos baseados em gradiente apresentam para fazer uma busca global eficiente (evitando mínimos locais), permitindo que os algoritmos de aprendizado das redes neurais façam a busca local de forma rápida (Skinner95).

O próximo capítulo detalha o funcionamento do algoritmo evolutivo com inspiração quântica e representação real, o seu uso para treinamento de redes neurais e também mostra como o modelo hipotético da partícula na caixa pode ser usado para simular este algoritmo.