

Proposta de um Neurônio Quântico

Wilkerson de Lucena Andrade
Herman Martins Gomes

Bruno Coitinho Araújo
Joseana Macêdo Fechine

*Universidade Federal de Campina Grande, Departamento de Sistemas e Computação
Av. Aprígio Veloso s/n, 58109-970 Campina Grande PB
{wilker, brunoc, hmg, joseana}@dsc.ufcg.edu.br*

Abstract

Quantum neural networks are emerging as a new paradigm in the field of Quantum Computing. This paradigm may exhibit some interesting properties that, in theory, can suppress some of the problems usually found in classical neural networks. In the past decade, some quantum neural network models have been proposed, however not all agree with regards to the concepts incorporated from the Quantum Mechanics area, neither with regards to the hardware/software requirements for implementation. Within this context, the objective of this work is to present a review of the existing models as well as to propose a new model for a quantum neuron based on the most promising approach identified in the review. Some initial simulations of this model using a Quantum Circuit simulator are also presented.

1. Introdução

É possível encontrar inúmeras evidências na literatura [Shor, 1997], [Deutsch, 1989], [Feynman, 1982], [Nielsen et al, 2000] de que a incorporação de conceitos quânticos pode minimizar ou eventualmente eliminar problemas inerentes aos modelos de computação clássica, uma vez que a Computação Quântica tem como principais características o processamento e a transmissão de dados armazenados em estados quânticos de uma forma muito mais eficiente que os modelos de computação convencionais. A Computação Quântica tem potencial para resolver problemas que são considerados “intratáveis” por máquinas clássicas atuais, ou seja, solucionar problemas que ainda não foram resolvidos devido a grande dificuldade computacional.

De uma forma geral, uma rede neural é um modelo desenvolvido com a finalidade de modelar o processamento que ocorre em estruturas neurais biológicas ao realizar uma tarefa particular ou uma

determinada função delegada [Haykin, 1999]. Embora a gama de problemas que podem ser resolvidos com a ajuda de uma rede neural seja bastante grande, há limitações que impedem a resolução, de forma eficiente, de muitos desses problemas. Por exemplo, uma limitação refere-se ao tempo de treinamento de redes neurais, que, dependendo do algoritmo de aprendizagem (*Backpropagation*, na maioria dos casos) tende a ser muito lento. Algumas vezes são necessários milhares de ciclos para se chegar a níveis de erros aceitáveis, principalmente se estiver sendo simulado em computadores seriais, pois o processador deve calcular as funções para cada unidade e suas conexões separadamente, o que pode ser problemático em redes muito grandes, ou com grande quantidade de dados.

Considerando a natureza multidisciplinar e as limitações de Redes Neurais Artificiais, aliadas à hipótese de que as sinapses entre neurônios poderiam ser mediadas por fenômenos quânticos [Penrose, 1994], foi natural o surgimento do interesse de pesquisas em Redes Neurais que incorporam conceitos de Física Quântica. Redes Neurais Quânticas é uma área promissora no campo da computação e informação quânticas. No entanto, ainda há pouco entendimento dos componentes essenciais de uma Rede Neural Artificial baseada em técnicas e conceitos quânticos [Gupta et al., 2001].

O propósito do presente trabalho é, portanto, apresentar uma breve descrição dos trabalhos relevantes na área de em Redes Neurais Quânticas, realizando uma análise crítica dos modelos existentes. Em seguida, será proposto um novo modelo de neurônio quântico (o qual, em etapas posteriores deste trabalho, será estendido para um modelo de Rede Neural Quântica). A proposição do modelo tomou como base a análise realizada. Ao final é apresentada uma breve descrição das simulações desse neurônio realizadas utilizando um simulador de circuitos quânticos.

2. Princípios de Computação Quântica

A Computação Quântica é baseada em princípios da Mecânica Quântica, tais como emaranhamento, superposição e descoerência. A seguir, serão apresentados alguns desses princípios.

Um fato interessante na Mecânica Quântica é a existência de partículas (fótons, elétrons, etc) em mais de um estado diferente ao mesmo tempo. Isto é chamado de fenômeno da Superposição Linear. Na realidade, quando se realiza uma medição, ou seja, quando se aplica um operador ao estado, chega-se a um valor real representando o estado da partícula após a medição. Esta operação "colapsa" um estado de superposição em um estado simples, ou seja, o sistema, antes composto por diversos estados, será reduzido a um dos seus estados elementares possíveis. Do ponto de vista matemático, para criar um estado de superposição são utilizados operadores lineares simples como o operador Hadamard, que recebe um *qubit* como entrada, e o coloca em estado de superposição [Nielsen et. Al. 2000].

Um sistema composto por partículas é dito emaranhado quando não é possível representar o seu estado como o produto tensorial (útil para resolver equações lineares em que a incógnita é uma matriz.) dos estados componentes, ou seja, não é possível expressar o estado do sistema na forma de uma equação apropriada. Quando um estado composto de dois subsistemas está emaranhado, não é possível atribuir um estado quântico puro para cada subsistema sozinho. É possível escrever estados quânticos misturados para cada subsistema considerado sozinho.

Os fenômenos da coerência e descoerência estão intimamente ligados à superposição linear. Um sistema quântico é dito ser coerente se este está em superposição linear com seus estados básicos. Por outro lado, descoerência é o processo pelo qual um sistema quântico decai para um estado clássico através de sua interação com o ambiente, ou seja, há uma perturbação do sistema, reduzindo-o a um único estado.

A formalização matemática da Computação Quântica baseia-se na representação e manipulação de estados quânticos em um espaço de Hilbert, o qual pode ser definido como um espaço vetorial complexo e linear, munido de produto interno, e completo em relação à norma definida por esse produto interno. O espaço vetorial é complexo porque os componentes de um vetor de estado têm valores complexos, ou seja, são números do tipo $a + bi$, em que a e b são reais e $i = \sqrt{-1}$. É linear porque a soma de vetores e a

multiplicação de vetores por números produzem vetores no mesmo espaço.

Normalmente, os vetores nesse espaço obedecem à notação de Dirac, comumente utilizada na área, sendo representados através dos símbolos $\langle \varphi |$ (bra) e $|\varphi \rangle$ (ket).

Como na Computação Clássica, a Computação Quântica possui uma unidade elementar de informação, o *qubit*, definido como sendo a superposição de dois estados independentes $|0\rangle$ e $|1\rangle$ (base ortogonal), denotado por $\alpha|0\rangle + \beta|1\rangle$, em que α e β são números complexos tais que $|\alpha|^2 + |\beta|^2 = 1$.

De uma forma geral, vetores de estado são definidos através de uma escolha particular de vetores bases e um conjunto de números complexos, os quais indicam as contribuições de cada componente base para o vetor de estado completo. Um sistema quântico simples de dois estados pode, por definição, estar em um de dois possíveis estados. Conseqüentemente, o vetor de estado desse sistema tem exatamente dois componentes

3. Trabalhos relacionados

Nos últimos anos, com o crescente interesse por Mecânica Quântica, o desenvolvimento de áreas a ela relacionadas tem sido intenso. Dessa forma, seguindo a mesma tendência, a área de Redes Neurais Quânticas tem sido fonte de vários trabalhos que exploram os conceitos e ferramentas da Mecânica Quântica para superar limitações e problemas existentes, propondo novos modelos e introduzindo novos conceitos.

Dentre algumas das características atribuídas a Redes Neurais Quânticas, pode-se citar a capacidade exponencial de memória (resultante da possibilidade de uma partícula estar em mais de um estado ao mesmo tempo, aumentando, assim, a quantidade de informação que pode ser armazenada num sistema quântico) [Ventura et al., 1998], a rápida capacidade de aprendizagem (se os vetores de pesos forem representados como uma superposição linear de estados, o processo de aprendizagem de múltiplos padrões pode ocorrer paralelamente em cada um dos estados), a eliminação do problema do esquecimento catastrófico (sendo a rede treinada independentemente para cada novo padrão, não há o problema de esquecimento de padrões anteriormente aprendidos, na presença de novos padrões de treinamento), soluções utilizando redes de uma única camada para problemas linearmente inseparáveis [Menner et al., 1995], dentre outros.

Existem opiniões bastante díspares sobre o que seja uma Rede Neural Quântica. Muitos pesquisadores aplicam suas próprias analogias e estabelecem conexões divergentes entre os conceitos relativos à Mecânica Quântica e às Redes Neurais Artificiais [Ventura et. al., 2000], sintetizados na Figura 1.

Ainda não há, a princípio, uma correspondência direta entre os conceitos apresentados na Figura 1. Na verdade, a tarefa de estabelecer uma analogia entre os conceitos é uma das tarefas mais árduas no momento de concepção de modelos de Redes Neurais Quânticas [Ventura et al., 2000].

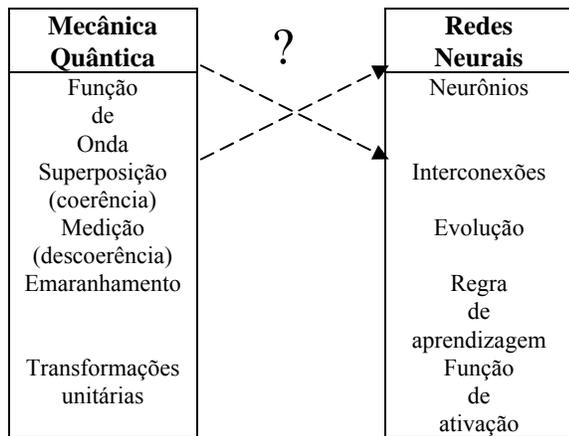


Figura 1. Conceitos da Mecânica Quântica versus Conceitos de Redes Neurais Artificiais.

A seguir, será realizada uma breve descrição dos modelos estudados observando características como: formato da representação da entrada, propagação da informação pela rede, unidade de processamento utilizada, custo e tecnologia de implementação. Uma análise comparativa mais detalhada pode ser encontrada em [Herbster et al, 2004].

O modelo proposto por Altaisky [Altaisky, 2004] usa dispositivos ópticos e sua entrada é representada sob vetores complexos. As conexões e a informação são representadas através de portas quânticas. Para a implementação, é necessária a aquisição de instrumentos ópticos. O custo de implementação é alto devido ao hardware necessário.

Um outro modelo usa Circuitos Quânticos [Gupta et al., 2001]. A entrada é representada através de vetores complexos, a informação flui de forma progressiva, as conexões estão sob a forma de ligações diretas e a unidade de processamento é uma porta quântica. Para a implementação é necessário pelo menos um simulador de circuitos quânticos e por isso o custo associado é baixo.

O modelo de [Behrman et al., 2000] utiliza uma tecnologia chamada de Quantum Dot Molecules. A entrada é representada sob vetores complexos, a informação flui de forma progressiva ou retroativa, as conexões são representadas por valores reais, a unidade de processamento é o estado quântico temporal. Para implementação é necessário o domínio da tecnologia de Quantum Dot Molecules, que eleva o custo e dificulta a implementação.

Um outro modelo [Narayanan et al., 2000] usa como ponto de partida o experimento da dupla fenda. A entrada da rede é representada sob vetores reais, a informação flui de forma progressiva, os pesos das conexões são valores reais e a unidade de processamento é o neurônio. Para implementação é necessário apenas o uso de simuladores de redes neurais clássicas, que apresentam custo de aquisição baixo.

A partir da análise dos trabalhos acima, é possível concluir que uma implementação através de Circuitos Quânticos seria o modelo mais promissor por dois motivos: primeiro, já há experiência suficiente nesta área e, segundo, há a viabilidade de simulação, pois já existem muitos simuladores de Circuitos Quânticos desenvolvidos.

3. Proposta do modelo

O ponto de partida para a modelagem da Rede Neural baseada nos conceitos da Computação Quântica é o neurônio de McCulloch-Pitts (Figura 2).

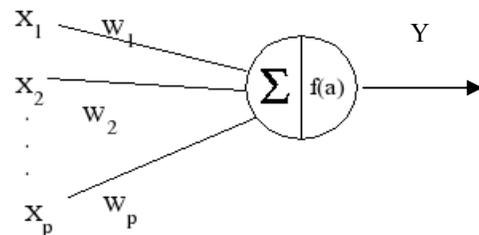


Figura 2: Neurônio de McCulloch-Pitts.

X_1, X_2, \dots, X_p representam as entradas da rede, W_1, W_2, \dots, W_p representam os pesos das conexões, Σ representa o somatório dos pesos das conexões multiplicado pelas respectivas entradas da rede, $f(a)$ representa a função de ativação e Y representa a saída da rede.

A operação de um neurônio na rede pode ser descrita através dos seguintes passos:

- Sinais são apresentados à entrada;
- Cada sinal é multiplicado por um peso que indica sua influência na saída do neurônio;
- É calculada a soma ponderada dos sinais que produzem um nível de atividade;
- Se este nível excede um limite (*threshold*) o neurônio produz uma saída.

Baseando-se no neurônio descrito, a seguir é apresentada a proposta de um neurônio análogo usando Circuitos Quânticos.

O neurônio quântico proposto foi modelado fazendo-se uma analogia com o neurônio de McCulloch-Pitts. A modelagem foi construída utilizando circuitos quânticos com diversas portas quânticas bem estabelecidas, como Hadamard e Toffoli (CCNOT).

A seguir será apresentado um modelo quântico de duas entradas, ilustrado na Figura 3, em que:

- E1 e E2 são as entradas do neurônio. E1 e E2 são qubits ($|0\rangle$ ou $|1\rangle$).
- W1 e W2 são qubits que representam os pesos das conexões.
- X é um sinal externo que afeta o limiar. X também é um estado quântico.

O modelo foi simulado utilizando o simulador de circuitos quânticos ZENO [Cabral et al.], desenvolvido numa dissertação de mestrado da Pós-Graduação em Informática da UFCG. A operação desse novo neurônio pode ser descrita através dos seguintes passos:

Passo 1. Inicialização dos estados quânticos: Os pesos das conexões (W_1 e W_2) recebem o estado $|0\rangle$, antes de serem colocados em superposição usando portas Hadamard.

Passo 2 – Operações 1-4. É realizado o produto entre os pesos das conexões e as entradas ($W_1 \times E_1 + W_2 \times E_2$). Isso é feito através de uma porta Toffoli (CCNOT), efetuando-se um AND entre cada entrada e seu respectivo peso (os resultados ficam nos *qubits* (6) e (7) na Figura 3). Uma vez com as entradas ponderadas, efetuamos a soma das parcelas.

Passo 3 – Operações 5-10. O limiar utilizado para comparação é determinado a partir do valor do *qubit* X. Se X é igual a zero, usa-se um limiar igual a 1. Com X igual a um, o limiar torna-se dois. Dependendo do

resultado do passo anterior, tem-se as alterações de valores locais, para ser possível a comparação com o limiar no próximo passo.

Passo 4 – Operação 11-12. A comparação do resultado do Passo 2 com o limiar, finalmente, é realizado. Assim, o neurônio produzirá a saída adequada.

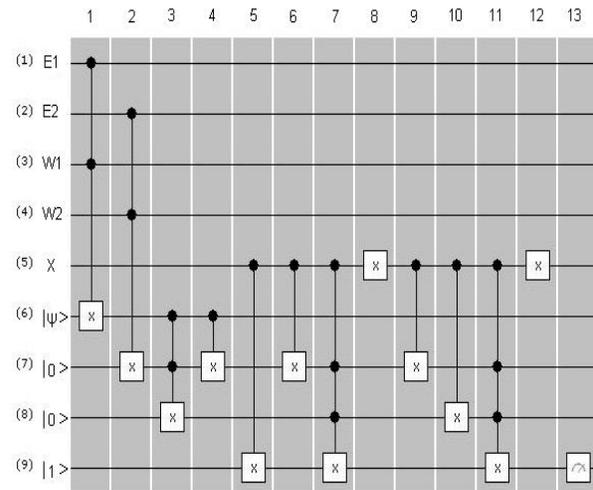


Figura 3. Circuito que implementa um neurônio quântico.

Inicialmente, os pesos do neurônio são postos em superposição através da porta Hadamard. Cada operação executada sobre o resultado do Passo 2 descrito acima (a ponderação das entradas), realiza não apenas uma, mas todas as possíveis ponderações entre as entradas e pesos possíveis, devido à superposição. Ao final do processo, o *qubit* representante da resposta do neurônio é medido, dando fim à superposição do estado quântico do neurônio, “colapsando-se” em um valor, dentre os resultados possíveis. É importante o ganho do ponto de vista computacional que um neurônio quântico sugere, ao realizar o processamento de todos os pesos paralelamente.

Nem sempre a resposta encontrada é aquela desejada. A probabilidade de ocorrência de um resultado depende das amplitudes da superposição dos pesos. Amplitudes “bem escolhidas” fazem o neurônio realizar a computação correta.

Um exemplo de execução da rede: fazendo as entradas E_1 e E_2 terem os valores $|0\rangle$ e $|1\rangle$, respectivamente. E fazendo X valer $|1\rangle$. Os pesos (W_1 e W_2), após passarem pelas portas Hadamard, valerão $(|0\rangle + |1\rangle) / 2^{1/2}$. O Passo 2, citado acima, resultará em $0 \times W_1 + 1 \times W_2$, ou seja, o estado geral do sistema (considerando todos os seus 9 *qubits*) será de: 0.5

$|110010001\rangle + 0.5 |110110101\rangle + 0.5 |111011101\rangle + 0.5 |111111011\rangle$. Ao término do Passo 3, o estado geral será de: $0.5 |110000100\rangle + 0.5 |110100000\rangle + 0.5 |111001000\rangle + 0.5 |111101111\rangle$, revelando que um dos pesos fez uma possível resposta do neurônio ser “um”, enquanto com os outros três pesos a resposta seria “zero”, com 25% de probabilidade para cada um. Finalmente, depois de comparado com o valor do limiar (que será 2, como determinado pelo Passo 3), obtém-se a resposta $0.5 |110010100\rangle + 0.5 |110110000\rangle + 0.5 |111011000\rangle + 0.5 |111111111\rangle$ do neurônio. Naturalmente, após a medição (operação 13), apenas um valor será o resultante.

4. Conclusões

O presente artigo apresentou uma visão geral das principais abordagens que utilizam de princípios da Mecânica Quântica aplicadas à área de Redes Neurais. Realizou uma breve revisão dos modelos existentes, culminando na proposição de um novo modelo que foi simulado como um circuito de portas quânticas básicas.

O neurônio proposto utiliza-se de propriedades quânticas para seu funcionamento. Sendo a mais importante destas propriedades o processamento paralelo, que permite que as entradas do neurônio sejam avaliadas todas de uma vez, levando em consideração uma combinação (superposição) linear de pesos iniciais, que podem conter uma solução para o problema do aprendizado da rede.

O modelo de neurônio apresentado neste artigo é ainda bastante primitivo, tratando-se, portanto, de um modelo inicial para uma rede neural quântica completa.

Trabalhos futuros deverão especificar uma regra para as alterações das amplitudes dos pesos possíveis (regra de aprendizagem do neurônio) e estratégias para conexão de múltiplos neurônios em rede (arquiteturas e regras de aprendizagem da rede).

Referências bibliográficas

[Altaisky, 2004] Altaisky, M. V. Quantum neural network. Joint Institute for Nuclear Research, Russia. Technical report, Available online at the Quantum Physics repository: http://arxiv.org/PS_cache/quant-ph/pdf/0107/0107012.pdf, last accessed 31/5/2004.

[Behrman et al., 2000] Behrman, E. C., Nash, L. R., Steck, J. E., Chandrashekar, V. G., Skinner, S. R. Simulations of

Quantum Neural Networks. Information Sciences, 128(3-4):257-269, 2000.

[Cabral et al.] Cabral, G. E., Lula, B., Lima, A. F. Zeno: a New Graphical Tool for Design and Simulation of Quantum Circuits. In Proceedings of SPIE Quantum Information and Computation III - Defense and Security Symposium, 2005, Orlando.

[Deutsch, 1989] Deutsch, D. Quantum computational networks, Proceedings of the Royal Society of London, A 425, pages 73-90, 1989.

[Feynman, 1982] Feynman, R.P. Simulating physics with computers, Int. J.Theor.Phys. vol. 21, pages 467-488, 1982.

[Gupta et al., 2001] Gupta, S., Zia, R. K. P. Quantum Neural Networks. Journal of Computer and System Sciences, vol. 63, pages 355-383, 2001.

[Haykin, 1999] Haykin, S. Neural Networks: A Comprehensive Foundation. 2a ed., New Jersey: Prentice Hall, 1999.

[Herbster et al, 2004] Harbster, R. F., Andrade, W. L., Gomes, H. M., Machado, P. D. L. O Estado da Arte em Redes Neurais Artificiais Quânticas. Revista de Iniciação Científica da SBC, ano IV, número IV, 2004.

[Menneer et al., 1995] Menneer, T., Narayanan, A. Quantum-inspired neural networks. Technical Report R329, Department of Computer Science, University of Exeter, UK., 1995

[Narayanan et al., 2000] Narayanan, A., Menneer, T. Quantum artificial neural networks architectures and components. Information Sciences, 128 (3-4):231-255, 2000.

[Nielsen et al, 2000] Nielsen, M. A., Chuang, I. L. Quantum Computation and Quantum Information. Cambridge, UK, Cambridge University Press, 2000.

[Penrose, 1994] Penrose, R. Shadows of the Mind. Vintage Science, 1994.

[Shor, 1997] Shor, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM Journal on Computing, 26(5):1484-1509, October, 1997.

[Ventura et al., 2000] Ezhov, A., Ventura, D. Quantum Neural Networks, in Future Directions for Intelligent Systems and Information Science, Ed. N. Kasabov, Physica-Verlag, 2000.

[Ventura et al., 1998] Ventura, D., Martinez, T. Quantum associative memory with exponential capacity, Proceedings of the International Joint Conference on Neural Networks, pages.509-513, 1998.