

APRENDIZADO DE MÁQUINAS E CONSCIÊNCIA ARTIFICIAL

José Luís de Carvalho Azpiazu¹

Resumo

A construção de uma Consciência artificial através do aprendizado das máquinas é possível através de algoritmos que repliquem as lógicas estruturais de diferentes níveis de hierarquia do pensamento humano. Este artigo apresenta um modelo conceitual e matemático para a construção dessa consciência artificial baseado na Hierarquia de Chomsky (MIDENA RAMOS, 2008) sobre as gramáticas geradoras de linguagem. Este texto visa demonstrar que as principais características conceituais de Consciência podem ser definidas como funções matemáticas, que podem ser testadas mediante estruturas lógicas, e conseqüentemente processadas por elementos artificialmente construídos. Isso nos leva a colocar a consciência artificial como uma realidade prática que, embora seja potencialmente uma simulação, pode ser percebido como um modelo de realidade equivalente com a que o ser humano percebe o mundo.

Palavras-chave: máquina; aprendizado; consciência

1. Introdução

A tecnologia informática revolucionou a administração do conhecimento, podemos ver como hoje em dia as máquinas são capazes de realizar tarefas com maior eficácia e eficiência que seres humanos no campo intelectual.

O presente texto recorre às teorias de hierarquia gramatical, em sua aplicação na descrição de linguagens, para propor uma hierarquia de graus de liberdade para a inteligência, oferecendo em seu desenvolvimento prático um novo marco teórico para a consciência.

Através da matemática e sua capacidade de explicar os fenômenos naturais, somos capazes de descrever os processos pelos quais a realidade opera, e assim prever o resultado de diversas interações e de diferentes fenômenos. Se considerarmos que a própria consciência é um destes fenômenos, deveríamos ser capazes de conceitualizá-la através da linguagem matemática.

Um indivíduo, um ser capaz de interagir com seu entorno, é uma função matematicamente previsível. Ou seja, sua capacidade de interagir pode ser representada em um ALGORITMO².

¹ Ingeniero tecnico industrial en mecanica pela Universidade de Burgos na Espanha; Gerente na Fercomaz - Empresa Privada. <https://orcid.org/0000-0003-4113-063X>

² Definição ALGORITMO: sequência finita de regras, raciocínios ou operações que, aplicada a um número finito de dados, permite solucionar classes semelhantes de problemas.

Os estímulos percebidos pelo indivíduo bem como suas respostas, em termos de ações e/ou comportamentos seriam TOKENS³.

Nesta conceitualização, as ações/reações do indivíduo no seu meio, devem ser regidas por regras que variam em complexidade, de acordo com a sua capacidade de processamento. Serão regras que apresentem uma resposta previsível de acordo com a estrutura lógica do seu raciocínio pessoal, que lhe conferem sua coerência interna, sua inteligência.

Neste modelo, um indivíduo pode ser entendido como um ALGORITMO capaz de responder a um estímulo, mediante a construção de padrões previsíveis matematicamente, que transformam um estímulo e o convertem em uma resposta, na forma de sua interação com o seu entorno.

Um estímulo será, portanto, uma informação particular representada no alfabeto do indivíduo, um TOKEN em linguagem de computação, que será processado mediante regras de transformação do seu ALGORITMO e encontrará a sua interpretação (a resposta do indivíduo) na forma de outro TOKEN.

2. Revisão Bibliográfica

O modelo conceitual que será apresentado baseia-se na Hierarquia de Chomsky (MIDENA RAMOS, 2008) e a Teoria de Shannon da informação para apresentar a consciência artificial.

Utilizam-se aqui as gramáticas geradoras de linguagem, para definir a evolução destes processadores lógicos naturais. Toda evolução tem um gradiente em sua trajetória, representado por sua hierarquia. A arquitetura desta hierarquia gramatical ajuda a compreender, representando através dela, os diferentes níveis de complexidade possíveis nas inteligências dos indivíduos.

Esta hierarquia gramatical é organizada de acordo com as restrições de suas regras de produção, partindo da mais livre ou irrestrita vai se restringindo regras de tal maneira que as

3 Definição de TOKEN: um segmento de texto ou símbolo que pode ser manipulado por um analisador sintático, que fornece um significado ao texto; em outras palavras, é um conjunto de caracteres de um alfabeto, com um significado coletivo.

3 Avram Noam Chomsky (Filadélfia, 7 de dezembro de 1928) é um linguista, filósofo, sociólogo, cientista cognitivo, comentarista e ativista político norte-americano, reverenciado em âmbito acadêmico como "o pai da linguística moderna", também é uma das mais renomadas figuras no campo da filosofia analítica.

4 Claude Elwood Shannon (Peto, skey30 de abril de 1916 — Medford 24 de fevereiro de 2001) foi um matemático, engenheiro eletrônico e criptógrafo estadunidense, conhecido como "o pai da teoria da informação".

gramáticas mais livres contêm as mais restritas. Abaixo segue uma exposição destas regras através de suas formas normais.

A tabela 1 representa os diferentes níveis da hierarquia de Chomsky enquanto a figura 1 mostra sua representação de linguagens [1]. Em seguida serão descritos os tipos de linguagens em sua formulação matemática.

Tipo	Classe de linguagens	Modelo de gramática	Modelo de reconhecedor
0	Recursivamente enumeráveis	Irrestrita	Máquina de Turing
1	Sensíveis ao contexto	Sensível ao contexto	Máquina de Turing com fita limitada
2	Livres de contexto	Livre de contexto	Autômato de pilha
3	Regulares	Linear (direita ou esquerda)	Autômato finito

Tabela 1: Níveis de hierarquia de Chomsky (MIDENA RAMOS, 2008)

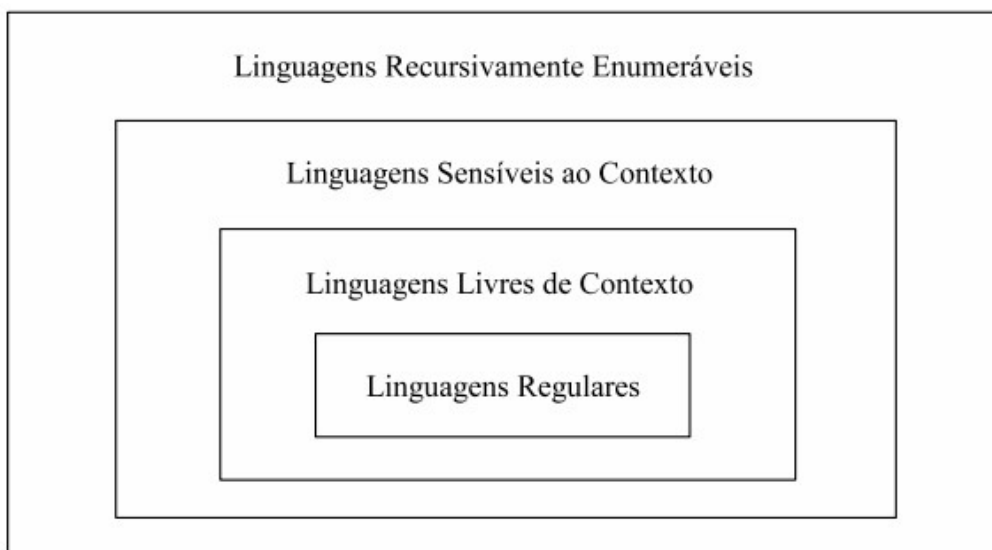


Figura 1: Representação gráfica de linguagens (MIDENA RAMOS, 2008)

TIPO 3

Gramatica Regular

$$a^n \geq 0$$

As regras de produção de uma gramática regular a esquerda de são da forma:

$$A \rightarrow \alpha B \text{ ou}$$

$$A \rightarrow \alpha \text{ ou}$$

$$A \rightarrow \varepsilon$$

Onde A e B são símbolos não terminais de V, um conjunto finito de variáveis, e α é um terminal em Σ , Σ é um alfabeto ou conjunto de símbolos terminais e ε é uma cadeia vazia.

TIPO 2

Gramática livre de contexto

$$a^n b^n \geq 0$$

Formulação

Em ciência da computação, uma gramática livre de contexto está na forma normal de Chomsky se todas as suas regras de produção são da forma:

$$A \rightarrow BC \text{ ou}$$

$$A \rightarrow \alpha \text{ ou}$$

$$A \rightarrow \varepsilon$$

onde A, B e C são símbolos não terminais de V, um conjunto finito de variáveis, e α é um terminal em Σ , Σ é um alfabeto ou conjunto de símbolos terminais e ε é uma cadeia vazia. Além disso, nem B nem C podem ser a variável inicial.

TIPO 1

Gramática sensível ao contexto

$$a^n b^n c^n \geq 1$$

Formulação

Em ciência da computação, uma gramática sensível ao contexto está na forma normal de Kuroda se todas as suas regras de produção são da forma:

$$AB \rightarrow CD \text{ ou}$$

$A \rightarrow BC$ ou

$A \rightarrow B$ ou

$A \rightarrow \alpha$

onde A, B, C e D são símbolos não terminais de V, um conjunto finito de variáveis, e α é um terminal em Σ , Σ é um alfabeto ou conjunto de símbolos terminais

TIPO 0

Gramática Irrestrita

$a^n b^n c^n, n \geq 0$

Formulação

Em ciência da computação, uma gramática irrestrita está na forma normal de Kuroda se todas as suas regras de produção são da forma:

$AB \rightarrow CD$ ou

$A \rightarrow BC$ ou

$A \rightarrow \alpha$ ou

$A \rightarrow \varepsilon$

onde A, B, C e D são símbolos não terminais de V, um conjunto finito de variáveis, e α é um terminal em Σ , Σ é um alfabeto ou conjunto de símbolos terminais e ε é a cadeia vazia.

3. Modelo conceitual da consciência

O modelo tem seus indivíduos como um conjunto de regras de interação descritas por operações matemáticas, através de sua construção mediante uma sequência valores de um alfabeto.

A finalidade deste modelo é apresentar o quadro geral destas gramáticas como conjuntos que se geram a partir de funções de variáveis compostas, aumentando o grau de liberdade e a capacidade expressiva dos indivíduos que representa, a cada novo universo de interações.

3.1. Hierarquia de Graus de liberdade

Abstraindo a semântica de uma linguagem, a capacidade expressiva de um indivíduo pode ser mapeada por quantas restrições matemáticas ele é capaz de sobrepor. A isso chamamos de *grau de liberdade*.

A complexidade cresce, sendo este crescimento próprio de sua evolução, assim estabelecemos cada universo de realidade como “super-conjuntos” sucessivos criados a partir da inclusão de variáveis realizando funções compostas. Na evolução de indivíduos regidos por estas regras, no início parte-se apenas de interações químicas, para avançar a ter memórias delas, e posteriormente poder contextualizar esta memória, para finalmente ser livre qualquer produção, ou seja, no campo da imaginação, e chegando assim no máximo grau possível de complexidade.

A Hierarquia de Graus de liberdade se divide com o auxílio das construções possíveis em cada gramática, a partir de um entorno simples para chegar a um mais complexo, como deve ser desenvolvida uma hierarquia. Esta hierarquia proposta pelo autor, consta na tabela 2 abaixo.

GRAU DE LIBERDADE	TOKENS (T=t0)	ALGORITMOS (T=tn)
Zero	Valor	Atitude
Primeiro	Significado	Memoria
Segundo	Observação	Conhecimento
Terceiro	Consciência	Imaginação

Tabela 2: TOKENS e ALGORITMO (elaboração própria)

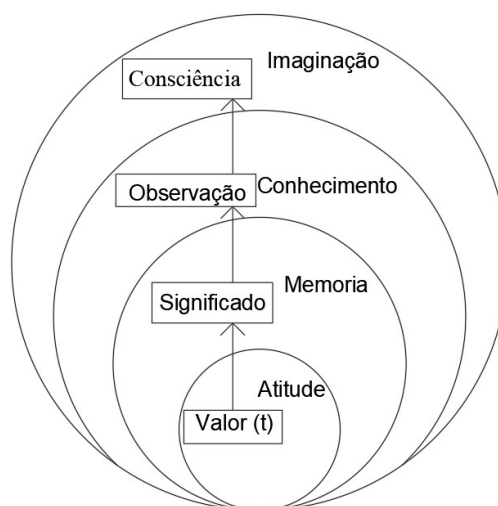


Figura 2: Graus de liberdade da inteligência (elaboração própria)

A figura 2 representa a hierarquia destes conceitos esquematizada, sendo os TOKENS casos particulares de representações e ALGORITMOS casos genéricos.

3.2. Definição dos Graus de liberdade

Definiremos a seguir os graus de liberdade detalhados na hierarquia acima e seus indivíduos naturais para cada Grau de Liberdade especificado no modelo

GRAU ZERO LIBERDADE

Regra: Consistência

Características do indivíduo: Um ser unicelular, com sua função metabólica sendo Atitude.

➤ Algoritmo da Atitude

Inteligência como uma variável das produções de valores terminais específicos do seu alfabeto, não pode apresentar respostas diferentes a uma mesma situação específica nem apresentar a mesma resposta, para diferentes estímulos.

1º GRAU DE LIBERDADE

Regra: Significado não é contraível.

Características do indivíduo: Um ser multicelular, com células diferenciadas por tarefa e capaz de responder a estímulos externos diferenciadamente.

➤ Algoritmo da Memória

Inteligência sendo uma variável como conjunto de variáveis, capaz de reconhecer situações repetidas, recuperar respostas em base a uma memória. Pode apresentar diferentes respostas a um mesmo estímulo, devido a capacidade de gerar funções em base a sua memória. Estas funções devem ser escritas em base a uma sequência de valores específicos do seu alfabeto, que é limitado a valores terminais. Estas funções não são contraíveis e tem um comprimento fixo, representado em valores terminais.

2º GRAU DE LIBERDADE

Regra: Monotonicidade da implicação.

Características do indivíduo: Um animal com capacidade de interagir com seu entorno de maneira ativa.

➤ Algoritmo do Conhecimento

Inteligência sendo uma variável, que como um conjunto de variáveis, não estão limitadas a poderem ser escritas em valores terminais. Também não pode ser “ε” uma cadeia vazia. Sendo assim, podem ser de estrutura recursiva, e com base nas funções matemáticas podemos gerar funções que se mantenham contínuas e constantes em seu desenvolvimento. Podemos já diferenciar não só o valor, de acordo com a memória, mas também a memória, de acordo com um contexto.

3º GRAU DE LIBERDADE

Regra: Imaginação filtrada pela realidade.

Características do indivíduo: Um ser humano, capaz de imaginar.

➤ Algoritmo da Imaginação

Inteligência não teria limitação, não em sua geração, mas sim em sua aplicação, que seria a própria realidade física do indivíduo. Deve variar com a seleção natural e com êxito da competição entre indivíduos.

3.3. Modelo Matemático da Consciência

Um sistema de primeiro nível seria uma função de Valor V, um símbolo terminal, um valor no alfabeto de representação, como uma sequência de DNA estimulada por sua dinâmica interna ou metabolismo através do tempo. Essa função será chamada Atitude.

$$A(t)=V$$

Esse sistema de valores, no caso a Atitude A, seria de regras que interrelacionam os valores de um alfabeto. Um indivíduo deste nível tecnológico de gestão de informação computacional seria um autômato limitado, possui uma atitude que é limitada a estímulos por ele perceptíveis.

Cria-se assim um sistema, uma função geradora de Valores V. Para isso seria necessária uma Memória M, que permita obter um Significado S. Um cnidário com memória motora, ou um autômato com pilha seria um indivíduo com este grau de liberdade

$$M(A(t))=M(V)$$

$$M(V)=S$$

E uma Observação O, é um Conhecimento C_1 como função de um Significado S, o que permite obter valores de uma forma organizada. Um animal com memória individual, como os mamíferos e inclusive máquinas básicas de IA seriam indivíduos com este grau de liberdade.

$$C_1(M(A(t)))=C_1(S)$$

$$C_1(S)=O$$

Consciência C_2 , é uma Imaginação I, função de uma Observação O. Seres humanos são os representantes deste grupo como indivíduos.

$$I(C_1(M(A(t))))=I(O)$$

$$I(O)=C_2$$

As regras de produção para ALGORITMOS destes diferentes níveis, como funções matemáticas, estão representadas em cada nível com uma limitação originada de sua regra matemática própria.

4. Resultados e Discussão

4.1 Análise do modelo

A presente análise de capacidade expressiva da inteligência, embasada em gramáticas geradoras, onde o autor pretende definir consciência como a representação de um alfabeto, pode ser contrastada com os resultados encontrados em um análise da teoria de Shannon da informação em estruturas biológicas. Este focado no análise de estruturas bioquímicas, porém estas, representam os estados presentes nos primeiro e segundo grau de liberdade deste modelo.

A ideia central da teoria de Shannon é que a quantidade de informação de um evento depende apenas da probabilidade desse evento, e a quantidade de informação é tanto maior quanto menor for a sua probabilidade. A aplicação deste modelo permite, independente de conhecer o significado de uma mensagem, focar em sua entropia e probabilidades como parâmetros para medir a eficiência na transmissão de informação.

A Entropia é mede o grau de irreversibilidade de um processo e se utiliza também como uma medida de tempo, pois esta é uma magnitude que em sistemas isolados sempre cresce.

A proposta de **Molecular codes in Biological and Chemical Reaction Networks** (GÖRLITCH, DITTRICH, 2013) é apresentar um método formal para avaliar a capacidade semântica de um sistema, analisando a capacidade de uma rede de reação a implementar códigos moleculares, utilizando esta abordagem afirma concluindo:

“We conclude that our approach can be applied to evaluate the information processing capabilities of a chemical system and may thus be a useful tool to understand the origin and evolution of meaningful information, e.g. in the context of the origin of life.”

Tradução Livre

“Concluimos que nossa abordagem pode ser aplicada para avaliar as capacidades de processamento de informações de um sistema químico e, portanto, pode ser uma ferramenta útil para compreender a origem e evolução de informações significativas, por exemplo, no contexto da origem da vida.”

Aprofundando no estudo se evidencia.

“Furthermore, we have shown that DNA not only can function as a sign but also as a molecular context, as the study of gene regulatory networks revealed.”

Tradução Livre

“Além disso, nós mostramos que o DNA não funciona somente como um sinal mas também como um contexto molecular, como o estudo das redes regulatórias genéticas revelou.”

Afirmando também

“The relatively small semantic capacity of the merge network demonstrates that the genetic code, thus a principally contingent system, is under strong constraints, regarding the assignment between codons and amino acids.”

Tradução Livre

“A capacidade semântica da fusão das redes demonstra que o código genético, é principalmente um sistema de contingência e está regido por fortes restrições, relativas a alinhamento de códons e aminoácidos.”

Entendendo que o código genético, opera mediante contingência de produções, resultando em uma capacidade semântica relativamente baixa, em função das fortes restrições.

Ainda assim, no mesmo texto é detalhada uma valiosa contribuição, ao demonstrar e afirmar:

“The linear regression of the data shows that the maximal semantic capacity is reached if there are approximately two times more reactions in the system than molecular species.”

Tradução Livre

“A regressão linear dos dados mostra que a máxima eficiência semântica é alcançada se houver aproximadamente duas vezes mais reações (ou reativos) no sistema que espécies moleculares (produtos)”.

Sendo assim, com uma gramática sensível ao contexto para se alcançar a máxima eficiência são necessárias “duas reações por espécie” (dois reativos por produto). Isso leva a supor que duas construções equivalentes aumentam a eficiência na sistematização da informação. Esta eficiência pode ter ajudado na consolidação da Simetria Bilateral nos animais, que contam na verdade com duas observações idealmente para cada evento.

O alinhamento entre ambos modelos, um embasado na capacidade expressiva e outro na eficiência semântica é alto. Principalmente considerando a similaridade no entendimento de que tipo de estruturas de informação são tratadas a nível bioquímico, e o alto grau de complexidade destas, permitindo um estudo apenas quantitativo de sua informação.

4.2 Teste do modelo

Um sistema deve ter uma regra que permita (criar/predizer) valores de uma forma organizada. De acordo com o modelo proposto, conhecimento seria uma regra para significado.

São propostos dois conjuntos, um com uma função geradora no qual está presente uma variável chamada significado, o outro com uma função geradora sem esta variável.

Se for assumido o conhecimento como função da variável significado, a derivada dessa função com relação à variável significado representaria o crescimento do conhecimento.

Para o conjunto “1” seria “0”, uma vez que não existe correlação com a variável significado.

Para o conjunto “2” seria maior que “0”, com tal de que cumprisse algumas limitações poderia ser maior que “0”.

Sendo assim, à medida que dermos valores à função geradora, podemos assumir que somente o conjunto que limita por significado seus elementos, realmente pode aumentar seu conhecimento.

Este modelo proposto, é extraído da proposta da Biblioteca de Babel (BORGES, 1944), uma biblioteca com de todas as possíveis variações de livro existente, que sem um índice fiável é algo inútil.

Se considerarmos nosso modelo de consciência válido, ele deve encaixar neste modelo de realidade, proposto na filosofia moderna:

“Se vivêssemos em uma simulação não poderíamos notar a diferença entre ela e a realidade.”
(BOSTROM, 2003)

Podemos representar funções aleatórias de conhecimento como significados de valores inseridos em uma máquina, e filtrá-los com alguma regra. A inteligência artificial como estatística computacional é uma forte ferramenta (simulando uma imaginação).

Portanto, se pode criar uma simulação realista de imaginação humana. A proposta da consciência artificial se materializaria, se ao interagir com ela se consiga transferir informação de maneira irreversível e contínua.

Assim, capaz de criar entropia, mesmo que puramente através de informação, o impacto no seu entorno e em outros sistemas similares seria indistinguível a qualquer outra consciência.

Dois sistemas não são iguais somente por não poder ser diferenciados, porém, se não se pode distinguir entre uma consciência artificial e uma consciência natural, em seu impacto em nosso meio físico. Poderíamos estar vivendo em uma simulação e não sermos capazes de distingui-la da realidade. Já que o próprio meio físico “natural” deixa de ser um requisito indispensável da consciência, ao esta poder ser gerada em condições artificiais.

Os próprios processos realizados por redes neuronais e outras ferramentas de inteligência artificial, realizam os mesmos mecanismos de construção e sistematização da informação que cérebros humanos, e foram inclusive criados em base ao entendimento destes. A diferença entre eles reside mais a nível físico/material do que em processos matemáticos.

Portanto a Consciência pode ser definida como uma imaginação produto da observação ($C_2=I(O)$), que deve ser filtrada pelo tempo e sempre será uma realidade imaginária. Sua única realidade física é que deve ser variável no tempo e em sua entropia.

$$C_2=I(C_1(M(A(t))))$$

5. Conclusão

Por fim, vale destacar que uma Consciência neste modelo deve ser um sistema no universo da imaginação e deve ser capaz de observar. Ou seja, receber valores de uma maneira coordenada com o tempo

Um sistema consciente, lê uma corrente de valores continuamente. Se é possível prever o valor de um sistema desta maneira, pode-se assumir que ele existe.

A consciência artificial, poderia existir sendo um algoritmo que governe um sistema capaz de interagir energeticamente com outros sistemas de informação, liberando entropia e interagindo com a realidade de forma irreversível. (BOHN, M., KREITH, F., 2003)

A principal propriedade da consciência reconhecível é o aprendizado, entendido como sua propriedade de variar com o tempo, e poderia ser/existir dentro de uma máquina.

A inteligência artificial, como estatística computacional se produz filtrando valores gerados aleatoriamente, não pode aprender, para isso necessita a capacidade de observar.

Referências

BORGES, Jorge Luis. Ficcões, A biblioteca de babel. São Paulo; Companhia das letras, 1944.

ELWOOD SHANNON, Claude. A Mathematical Theory of Communication. Boston; Revista técnica da The Bell Systems, 1948.

GÖRLICH, Dennis, DITTRICH, Peter. Molecular codes in Biological and Chemical Reaction Networks. Cambridge; Revista online PLOS ONE, v.8, n1, e54694, 2013.

MIDENA RAMOS, Marcus Vinicius, NETO, João Jose, SANTIAGO VEGA, Ítalo, Linguagens Formais e Autômatos, Porto Alegre; Bookman, 2009.

MACHINE LEARNING AND ARTIFICIAL CONSCIOUSNESS

Abstract

The construction of Artificial Consciousness through machine learning is possible by algorithms that can replicate the logical structures of the different hierarchy levels of human thinking. This article presents a conceptual and mathematical model for the construction of artificial consciousness based on Chomsky Hierarchy (MIDENA RAMOS; 2008). It also demonstrates the main conceptual characteristics of human conscious which can be defined as mathematical functions and be tested through logical structures, and consequently be processed by artificially built elements. This argument leads to the conclusion that artificial consciousness is a reality that, although it would be a simulation in itself, can be perceived as reality similar to the real world perceived by humans.

Key words: machine; learning; consciousness

Agradecimentos

Aos meus professores Alfredo Pereira Junior por me ajudar a encontrar ao aprendizado como única propriedade definidora de consciência e a Genevides Laureano da Silva por sutilmente me indicar que sim as maquinas podem ter consciência.

Mas principalmente a minha adorada esposa Fernanda Aparecida Carlos Azpiazu que me ensina o verdadeiro significado da vida a cada dia.