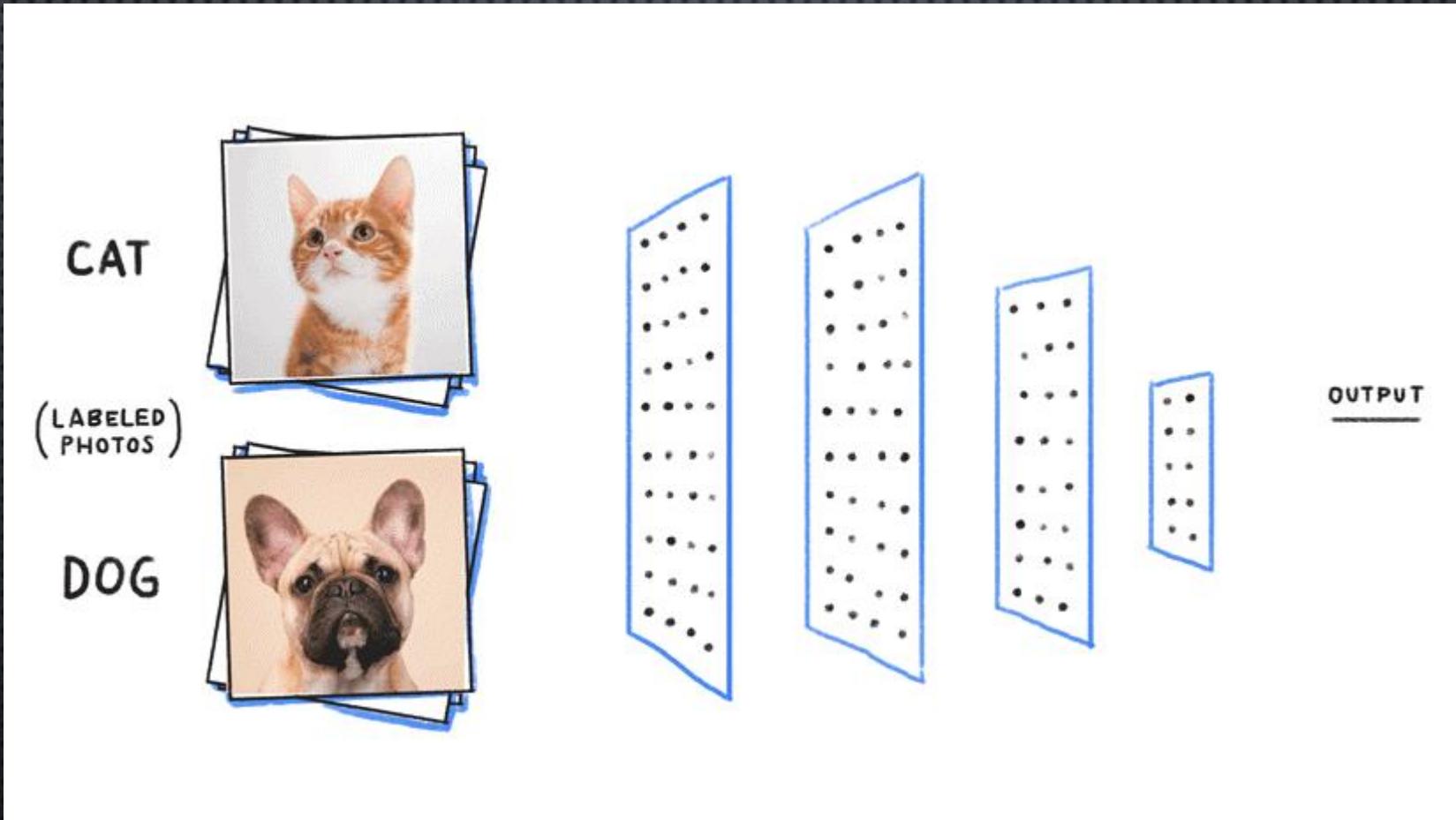


# NEURAL TURING MACHINES

JOHN THEO – 18/06

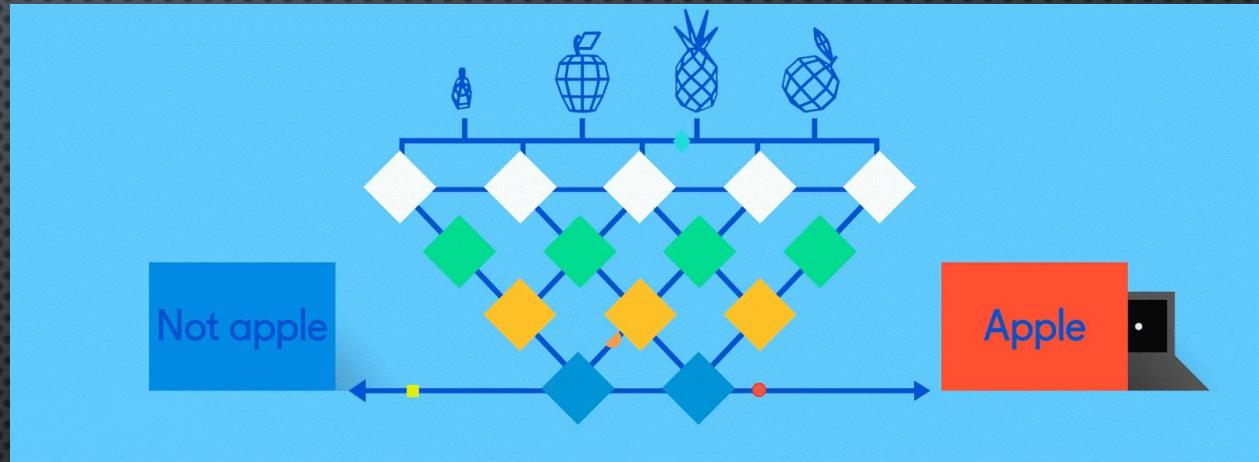
# CONTEXTO

Redes Neurais mapeam dados em uma representação abstrata



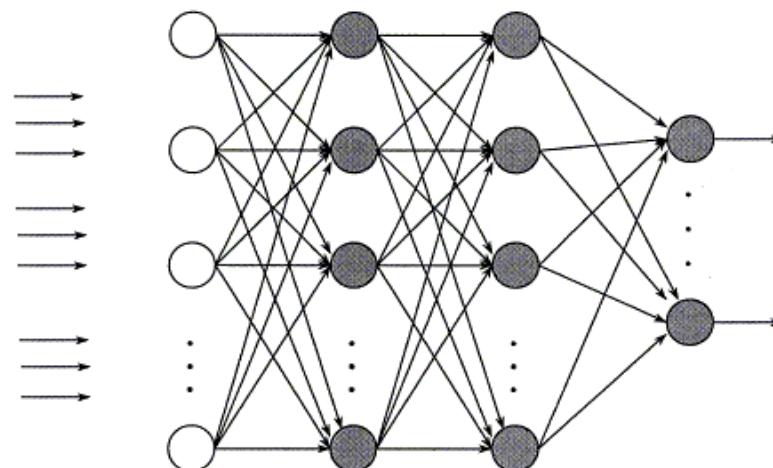
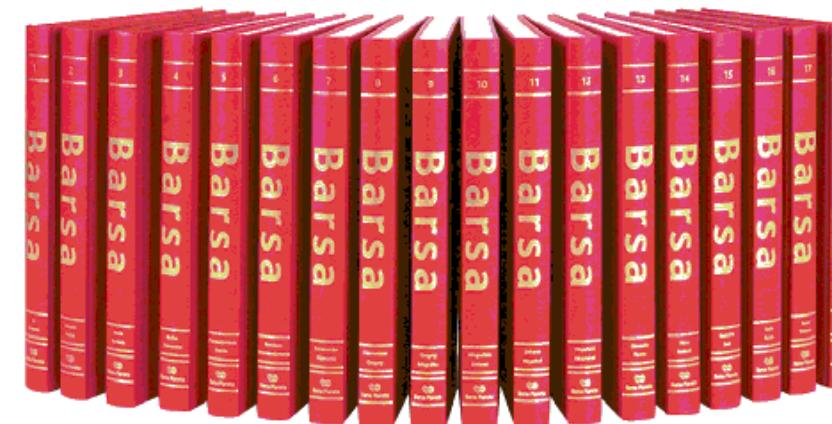
<https://www.google.com/about/main/machine-learning-qa/img/cat-dog-flow-horizontal.gif>

- RNNs TEM EFICIENCIA E FLEXIBILIDADE.
  - MODELAGEM SEQUENCIAL COM O CONCEITO DE MEMÓRIA.
- **PATTERN:** CAMADAS DIFERENCIÁVEIS
  - OTIMIZAÇÕES BASEADAS EM GRADIENTE PERMITEM MELHORIA NAS PREDIÇÕES



# Inserir conhecimento não é trivial:

- Overfitting
- Complexidade da Rede(layers/parâmetros)
- Generalização



Simkhada PP et al.  
*J Health Humanities* 2013, Vol 13 (No. 1), 1-5  
Writing an academic paper for publication

## ■ Guest Editorial

### Writing an academic paper for publication

PP Simkhada<sup>1</sup>, E van Teijlingen<sup>2</sup> & V Hundley<sup>3</sup>  
<sup>1</sup>SCHARR, University of Sheffield, Sheffield, England, UK & Visiting Professor at Manmohan Memorial Institute of Health Sciences, Tribhuvan University, Nepal

<sup>2</sup>School of Health & Social Care, Bournemouth University, England, UK & Visiting Professor at Manmohan Memorial Institute of Health Sciences, Tribhuvan University, Nepal

<sup>3</sup>School of Health & Social Care, Bournemouth University, England, UK

#### Background

This paper on writing for publication aims to help new authors getting started with the process. The paper starts by asking the basic question: "Why should one write an academic paper?" The paper

the profile and status of: (a) oneself; (b) your department or institution; and (c) your academic discipline. Helping junior staff and postgraduate students to publish their research may act as a motivating factor for both, and finally, writing may

<http://cultura.estadao.com.br/blogs/curiocidade/wp-content/uploads/sites/520/2012/03/barsa-2012.jpg>

<http://jblomo.github.io/datamining290/slides/img/2-layer-nn.gif>

<http://blogs.bournemouth.ac.uk/research/files/2015/09/Acad-writing.jpg>

- **REDES NEURAIS DE MEMÓRIA AUMENTADA**
  - ADIÇÃO DE UMA MEMÓRIA INDEPENDENTE QUE AGE COMO UMA BASE DE CONHECIMENTO
  - **WORKING MEMORY** (NEUROCIÊNCIA)
  - NEURAL TURING MACHINES | NEURAL RANDOM ACCESS MACHINES | NEURAL GPU
- A **ESTRUTURA** É CRUCIAL PARA MANTER A ESCRITA E LEITURA RÁPIDAS
- LSTMS PODEM TER MEMÓRIA, PORÉM O CONHECIMENTO ESTÁ DISTRIBUIDO



- **META-APRENDIZADO**

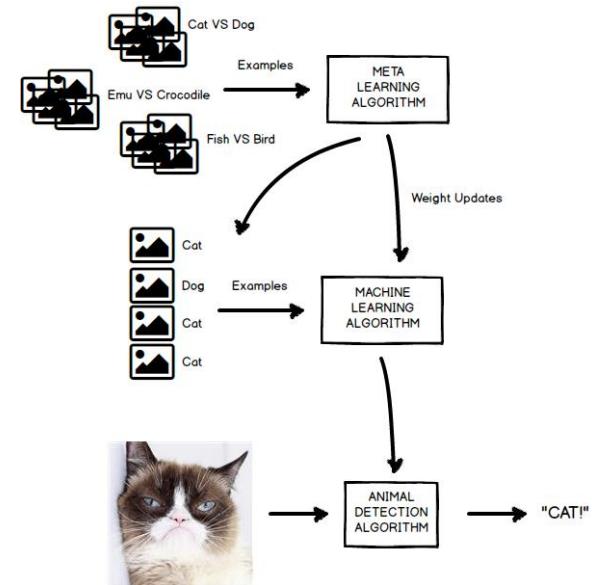
- SHORT-TERM → PREDIZER X
- LONG-TERM → APRENDER A PREDIZER X → CONSEGUIR PREDIZER Y,Z,W
- NAO É NECESSÁRIO TREINAR EXAUSTIVAMENTE UMA REDE PARA REALIZAR UMA TAREFA ÚNICA

## Meta Learning with Memory-Augmented Neural Networks

Adam Santoro  
Google DeepMind  
Sergey Bartunov  
Google DeepMind, National Research University Higher School of Economics (HSE)  
Matthew Botvinick  
Dawn Wigstrom  
Timothy Lillicrap  
Google DeepMind

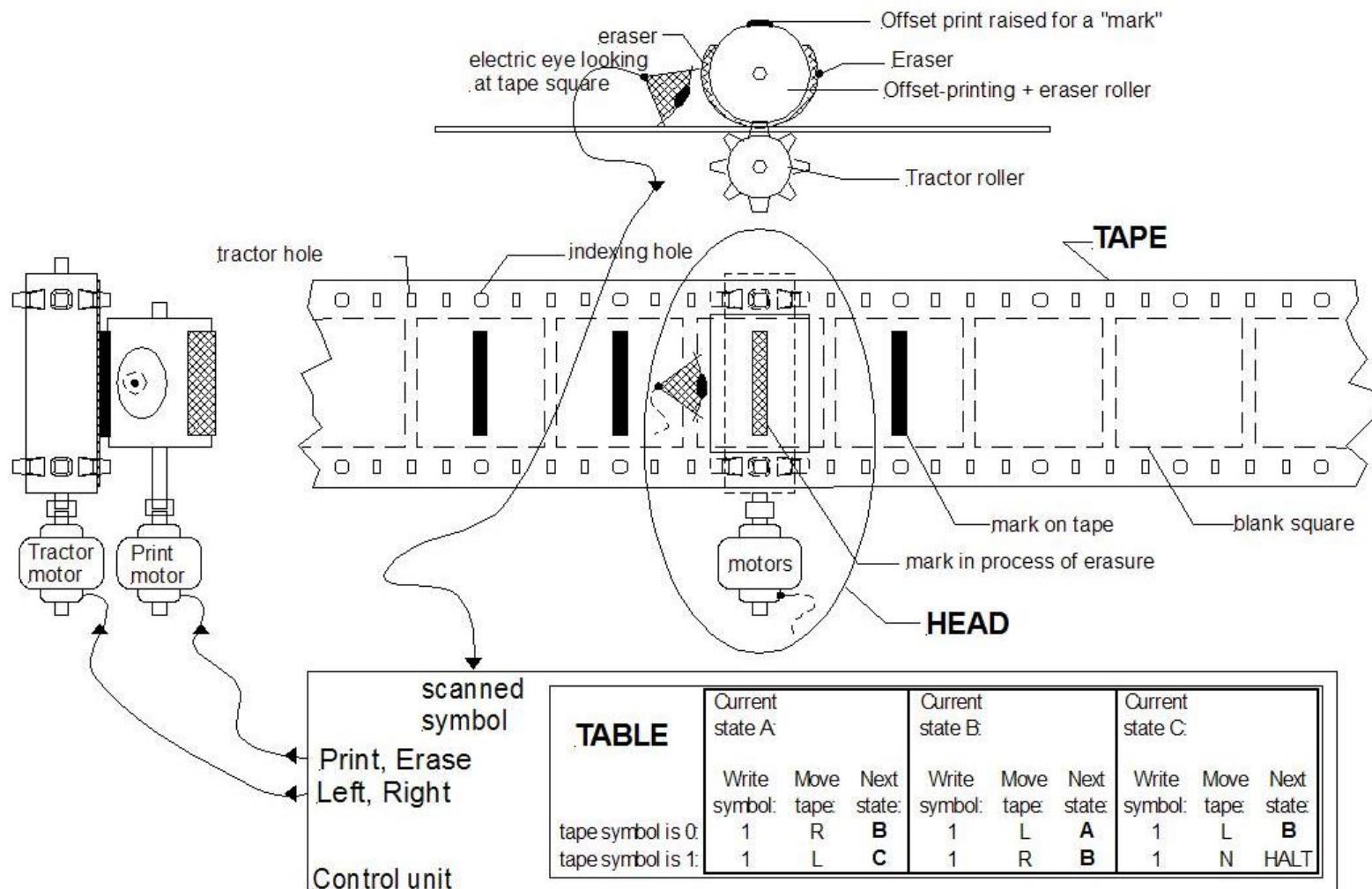
ADAMSANTORO@GOOGLE.COM  
SHOS@SHOS.IN  
BOTVINICK@GOOGLE.COM  
WIGSTROM@GOOGLE.COM  
COUNTZERO@GOOGLE.COM

ICML 2016, citation: 16  
Katy@DataLab  
2017.03.28



# DEFINIÇÃO

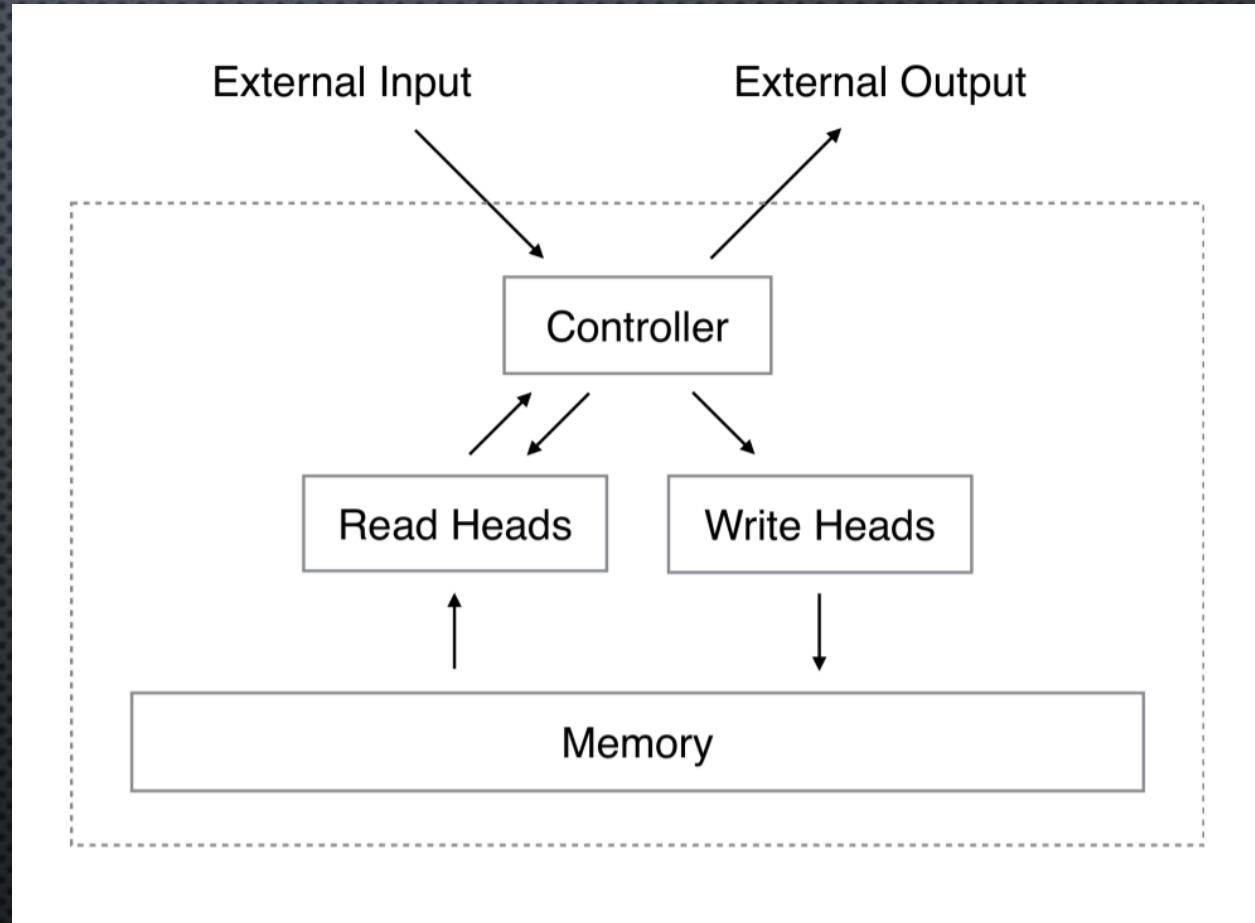
# MÁQUINA DE TURING



A fanciful mechanical Turing machine's TAPE and HEAD. The TABLE instructions might be on another "read only" tape, or perhaps on punch-cards. Usually a "finite state machine" is the model for the TABLE.

# NEURAL TURING MACHINE

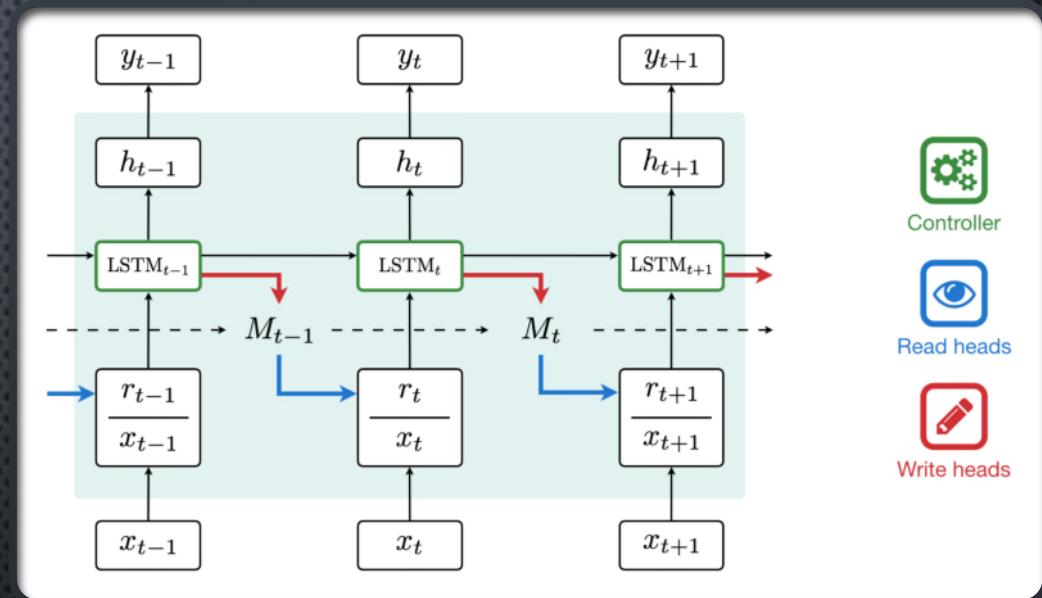
- CONTROLADOR
- CABEÇA(LEITURA/ESCRITA)
- MEMÓRIA



[https://rylanschaeffer.github.io/content/research/neural\\_turing\\_machine/ntm\\_architecture.png](https://rylanschaeffer.github.io/content/research/neural_turing_machine/ntm_architecture.png)

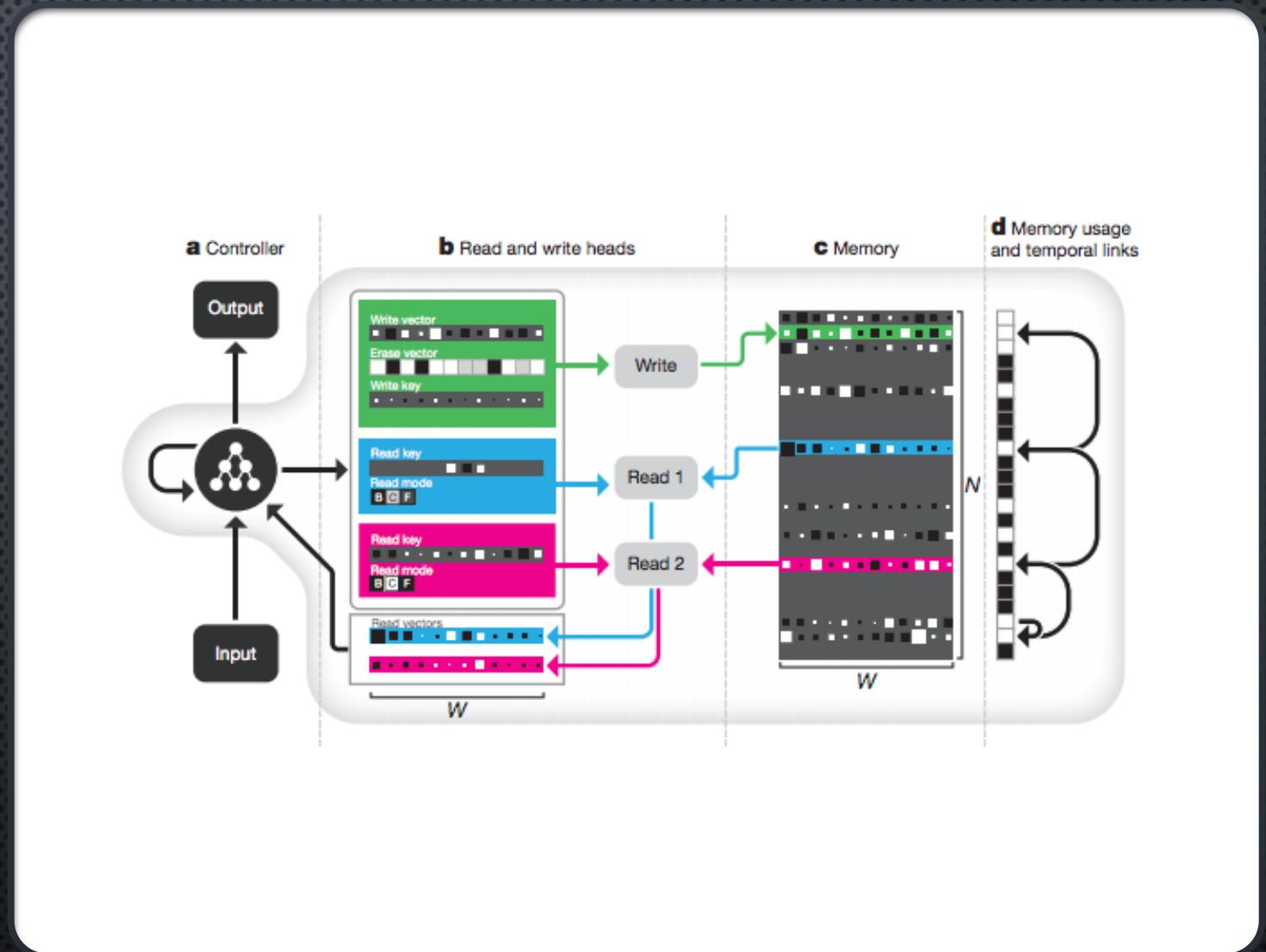
# CONTROLADOR

- É UMA REDE NEURAL QUE PROVÊ UMA REPRESENTAÇÃO DA ENTRADA
- LÊ/ESCREVE O **HEAD** PARA INTERAGIR COM A MEMÓRIA
- **NÃO** É IDENTICO AO QUE ESTÁ NA MEMÓRIA, MAS SIM UMA **FUNÇÃO** PARA REPRESENTAÇÃO
- TIPO DO CONTROLADOR É CRUCIAL:
  - FEED-FOWARD: MAIS RÁPIDO, TRANSPARENTE, MENOS EXPRESSIVO (COMPUTATION/TIME\_STEP)
- A CADA PASSO, DADA UMA ENTRADA, RESPONDE COM 4 VETORES:
  - VETOR DE SAÍDA: REFERENTE A SAÍDA CRUA
  - VETOR DE LEITURA: UTILIZADO PELA CABEÇA DE LEITURA PARA ENDEREÇAR UM ITEM DA MEMÓRIA
  - VETOR DE REMOÇÃO: UTILIZADO PARA SELECIONAR E APAGAR UM ELEMENTO DA MÉMÓRIA
  - VETOR DE ESCRITA: UTILIZADO PARA ESCREVER NA MEMÓRIA
- TODOS OS VETORES SÃO NORMALIZADOS [0,1]

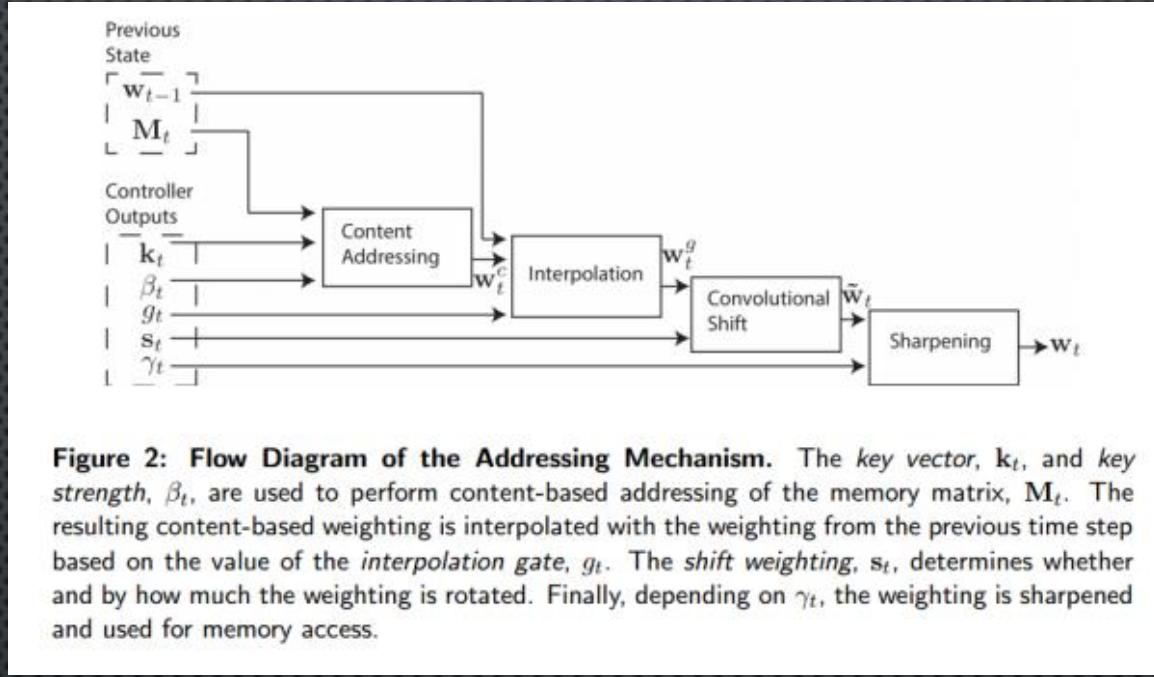


# MECANISMO DE LEITURA/ESCRITA

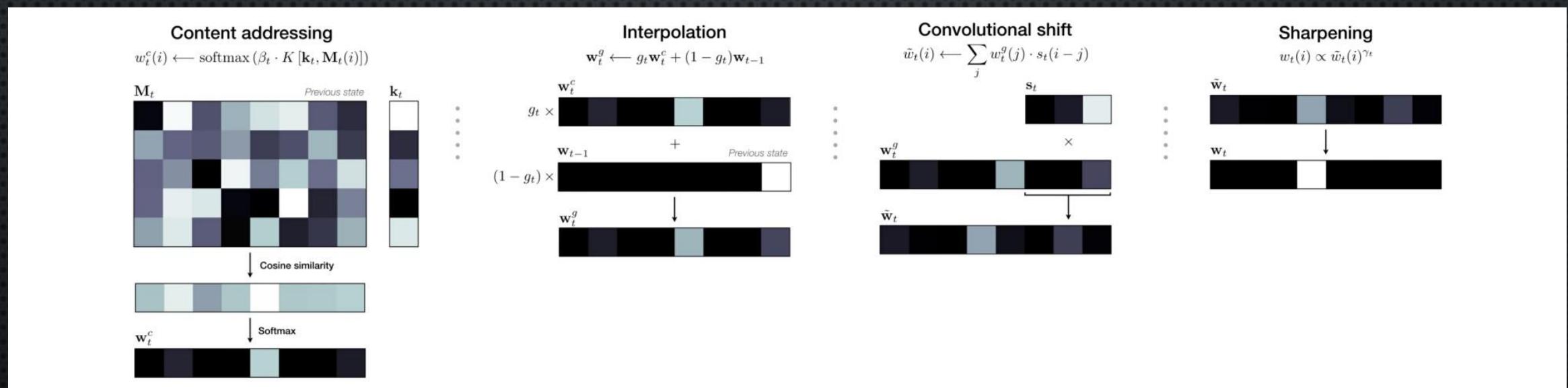
- FUNCIONA POR ATENÇÃO(ATTENCION) OU POR FOCO
- ATENÇÃO → LOCATION
- FOCO → CONTENT
- CADA CABEÇA É CONTROLADA POR SEU PRÓPRIO VETOR DE PESOS ATUALIZADO A CADA PASSO
- CADA VALOR NO VETOR DE PESOS CORRESPONDE AO GRAU DE INTERAÇÃO COM CADA POSIÇÃO DE MEMÓRIA
- DADO O VETOR(LEITURA/ESCRITA/REMOÇÃO) REALIZA A SIMILARIDADE COM A POSIÇÃO DE MEMÓRIA E ATUA
- NO CASO DE REMOÇÃO E ESCRITA, REMOVE PRIMEIRAMENTE PARA DEPOIS ESCREVER
- **POR UTILIZAR UM MECANISMO ANALOGO À MAQUINA DE TURING AO FINAL DO TREINAMENTO É POSSÍVEL INTERPRETAR O “ALGORTIMO” QUE A REDE APRENDEU**



<https://qph.ec.quoracdn.net/main-qimg-29b05040bbbada74b05c986e10f49567>



<https://adriancolyer.files.wordpress.com/2016/03/ntm-fig-2.png>



[https://cdn-images-1.medium.com/max/2000/1\\*k3IuW2bN7oCZEh1GtxuwhA.jpeg](https://cdn-images-1.medium.com/max/2000/1*k3IuW2bN7oCZEh1GtxuwhA.jpeg)

# APLICAÇÕES

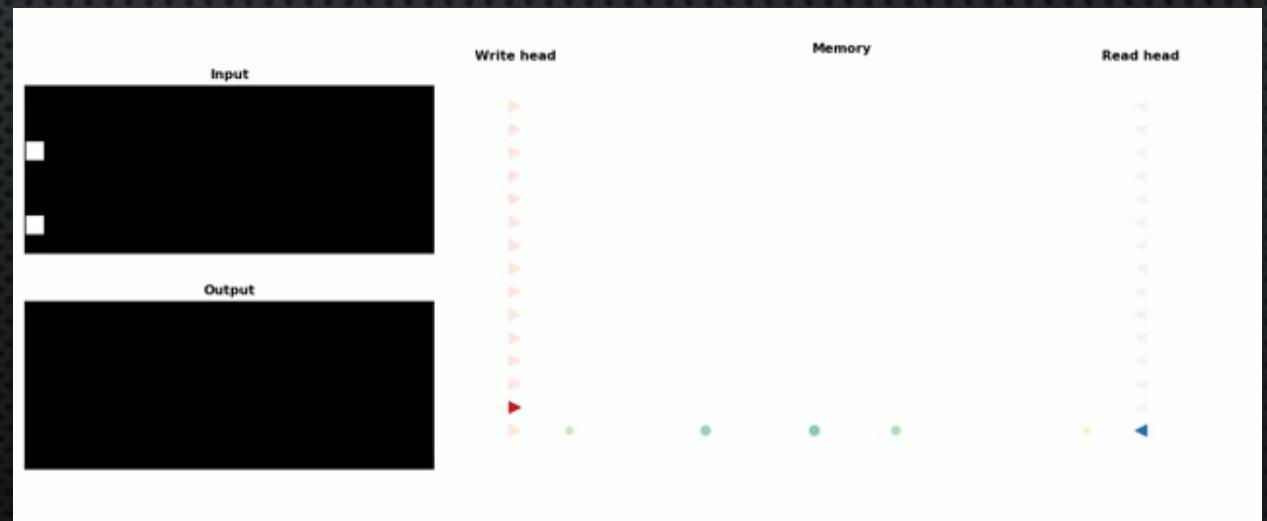
Possível Perceber que o Conhecimento Adquirido pela NTM refere-se a **Manipulação da Memória**

## OPERAÇÕES SIMPLES, PERMITINDO GENERALIZAÇÕES:

- CÓPIA
- CÓPIA REPETIDA
- LEMBRANÇA ASSOCIATIVA
- DYCK WORDS



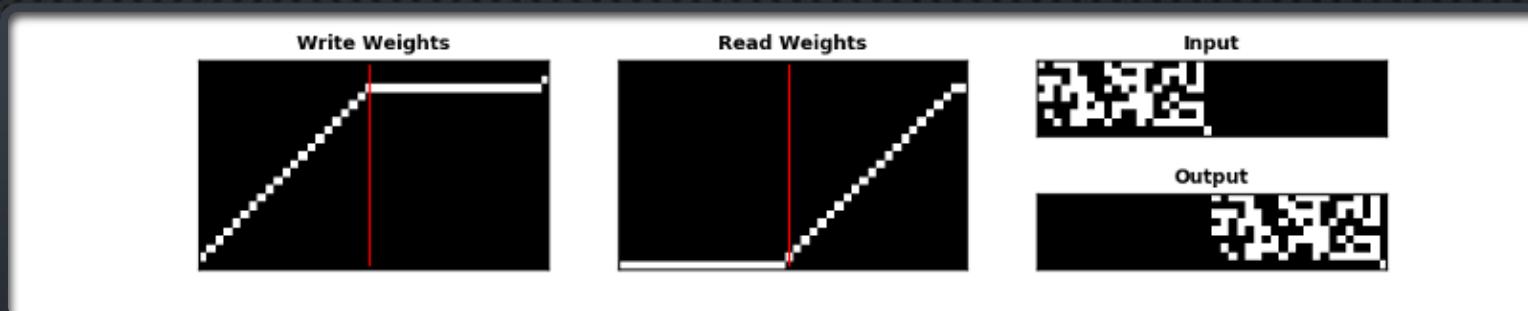
<https://thumbs.gfycat.com/WelllitInferiorAndeancondor-mobile.mp4>



<https://thumbs.gfycat.com/CarefulSecondaryHorsemouse-mobile.mp4>

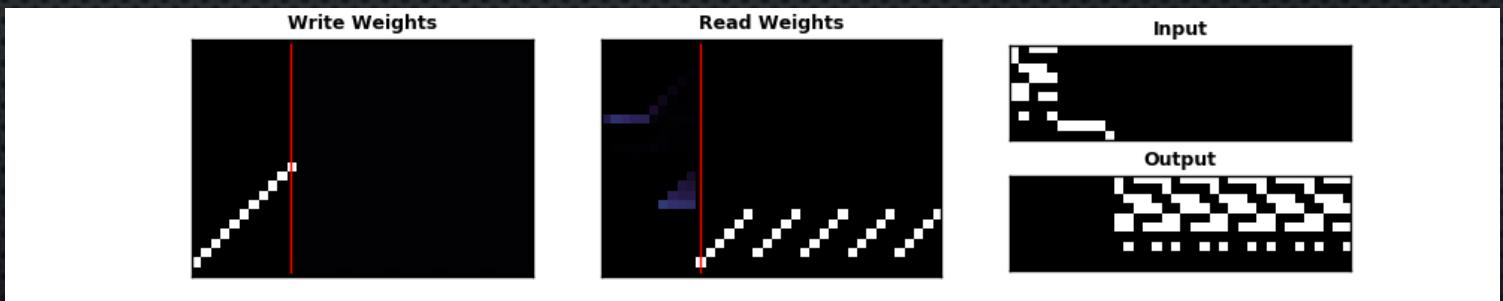
# CÓPIA

- REPRODUZIR A SEQUENCIA DE ENTRADA COMO SAÍDA
- LÊ A ENTRADA E ESCREVE A SAÍDA CORRESPONDENTE, **UM ITEM POR VEZ**
- NTM NÃO APRENDE APENAS A COPIAR ENTRADAS DE TAMANHO SEMELHANTE, MAS **GENERALIZA**.
- APRENDE O **ALGORITMO DE CÓPIA** E PODE ESTENDER PARA TAMANHOS VARIADOS



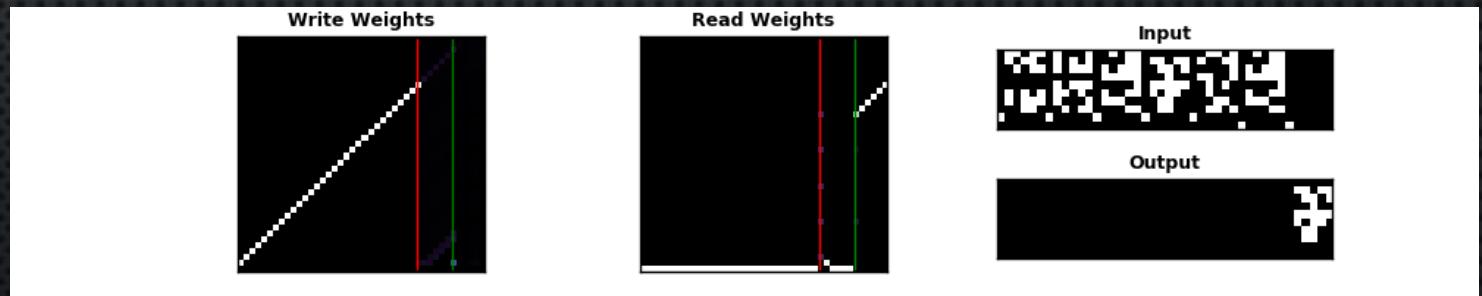
# CÓPIA REPETIDA

- AUMENTO DE COMPLEXIDADE. OPERAÇÃO **FOR**
- REPRODUZIR A SEQUENCIA DE ENTRADA EM UM NÚMERO ESPECÍFICO DE VEZES NA SAÍDA
- NTM LÊ A ENTRADA CRUA E NÃO TEM INFORMAÇÕES DE ONDE ESTÁ O DADO E NEM MESMO A QUANTIDADE DE VEZES A SEREM REPETIDAS
- A “SUPERVISÃO” OCORRE DURANTE O TREINAMENTO ONDE O MODELO UTILIZA A SAÍDA DESEJADA PARA “DESCOBRIR” A INFORMAÇÃO RELEVANTE



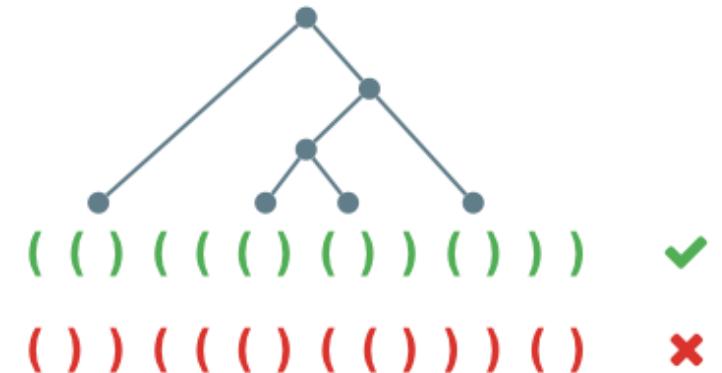
# LEMBRANÇA ASSOCIATIVA

- APRENDER UM SISTEMA DE PERGUNTA/RESPOSTA
- APROVEITANDO O MECANISMO DE ENDEREÇAMENTO (UTILIZADO PELAS CABEÇAS), VERIFICAR SE A NTM É CAPAZ DE APRENDER ALGO PARECIDO COM PONTEIROS (LINGUAGEM C)
- DADA UMA SEQUÊNCIA DE ITENS SEPARADOS POR UM DELIMITADOR, DESCOBRIR O ITEM SUBSEQUENTE AO ITEM QUESTIONADO.
- NTM LÊ A ENTRADA CRUA E NÃO TEM INFORMAÇÕES DE ONDE ESTÁ O DADO.
- NTM NECESSITA APRENDER QUE EXISTE UMA FLAG QUE DELIMITA A ENTRADA
- LINHA VERMELHA → FIM DO DADO
- LINHA VERDE → FIM DA QUERY



# DYCK WORDS

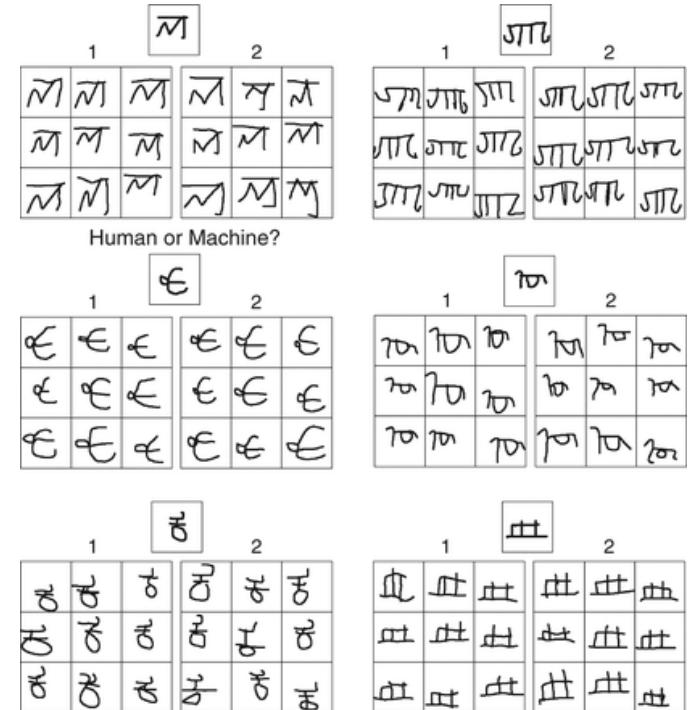
- APRENDER A CLASSIFICAR “PALAVRAS” DE PARENTHESES BALANCEADOS
- UM ALGORITMO SIMPLES DE PILHA PODE RESOLVER O PROBLEMA
- DEPENDÊNCIA LONG-TERM: EM RNNs PODE SER INVÍAVEL POIS NECESSITA AUMENTAR COMPLEXIDADE DA REDE(LAYERS)
- NTM CONSEGUE CODIFICAR ALGO ANÁLOGO AO ALGORITMO
- NTM MANIPULA A MEMÓRIA DE MANEIRA NÃO LINEAR POIS USA A CABEÇA DE LEITURA PARA MANTER STATUS DA PILHA
- PARANTESES ABERTO → CABEÇA PARA DIREITA(CIMA)
- PARANTESES FECHADO → CABEÇA PARA A ESQUERDA(BAIXO)



# ONE-SHOT LEARNING

- APRENDER COM UMA AMOSTRA OU COM UM NÚMERO REDUZIDO DE AMOSTRAS
- OMNIGLOT DATASET:
  - 1600 CARACTERES
  - 50 ALFABETOS
  - 20 AMOSTRAS DE CADA CARACTER
- ACURÁCIA

MODEL	INSTANCE (% CORRECT)					
	1 <sup>ST</sup>	2 <sup>ND</sup>	3 <sup>RD</sup>	4 <sup>TH</sup>	5 <sup>TH</sup>	10 <sup>TH</sup>
HUMAN	34.5	57.3	70.1	71.8	81.4	92.4
FEEDFORWARD	24.4	19.6	21.1	19.9	22.8	19.5
LSTM	24.4	49.5	55.3	61.0	63.6	62.5
MANN	<b>36.4</b>	<b>82.8</b>	<b>91.0</b>	<b>92.6</b>	<b>94.9</b>	<b>98.1</b>



One-shot Learning with Memory-Augmented Neural Networks

Adam Santoro

Google DeepMind

Sergey Bartunov

Google DeepMind, National Research University Higher School of Economics (HSE)

ADAMSANTORO@GOOGLE.COM

SBOS@SBOS.IN

Matthew Botvinick

Daan Wierstra

Timothy Lillicrap

Google DeepMind

BOTVINICK@GOOGLE.COM

WIERSTRA@GOOGLE.COM

COUNTZERO@GOOGLE.COM

# REFERÊNCIAS

- **QUORA EXPLANATION** - [HTTPS://WWW.QUORA.COM/HOW-DOES-THE-DEEPMIND-DNC-DIFFERENTIABLE-NEURAL-COMPUTER-COMPARE-TO-LSTMS-AND-RNNs](https://www.quora.com/How-does-the-DeepMind-DNC-Differentiable-Neural-Computer-compare-to-LSTMs-and-RNNs)
- **HOW TO LEARN FROM LITTLE DATA** - [HTTPS://WWW.YOUTUBE.COM/WATCH?V=tChcZPBTTA](https://www.youtube.com/watch?v=tChcZPBTTA)
- **NTM-LASAGNE** - [HTTPS://MEDIUM.COM/SNIPS-AI/NTM-LASAGNE-A-LIBRARY-FOR-NEURAL-TURING-MACHINES-IN-LASAGNE-2CDCE6837315](https://medium.com/snips-ai/ntm-lasagne-a-library-for-neural-turing-machines-in-lasagne-2cdce6837315)
- **NEURAL TURING MACHINES(POSTAGEM BLOG)** - [HTTPS://BLOG.ACOLYER.ORG/2016/03/09/NEURAL-TURING-MACHINES/](https://blog.acolyer.org/2016/03/09/neural-turing-machines/)
- **ONE-SHOT LEARNING WITH MEMORY-AUGMENTED NEURAL NETWORKS** -  
[HTTPS://ARXIV.ORG/PDF/1605.06065.PDF](https://arxiv.org/pdf/1605.06065.pdf)
- **NEURAL TURING MACHINE(DEEPMIND)** - [HTTPS://ARXIV.ORG/PDF/1410.5401.PDF](https://arxiv.org/pdf/1410.5401.pdf)