Aprendizado por Reforço

Thiago Cardoso
thiago.cardoso@hekima.com

HEKIMA

Big Data Analytics As a Service

THIAGO CARDOSO

@tncardoso

Bacharel e Mestre em
Ciência da Computação

CTO & **Co-founder**
@ hekima

Soluções customizadas de Big Data Analytics

HEKIMA

A Hekima é uma empresa de ***Big Data*** e ***Inteligência Artificial.***

Simplificamos a complexidade das aplicações de Big Data em soluções ready to use.

GREAT PLACE TO WORK®

DESTAQUE NO CAMINHO CERTO 2015

endeavor PROMESSAS

AWARDS
PARTNERS
COMMUNITIES

SanPedro valley

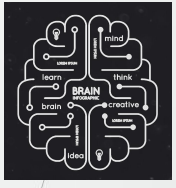
FIR CAPITAL

amazon web services | Partner Network

TECHNOLOGY PARTNER

Udemy

[Curso Data Science]



[Hekima developers blog]

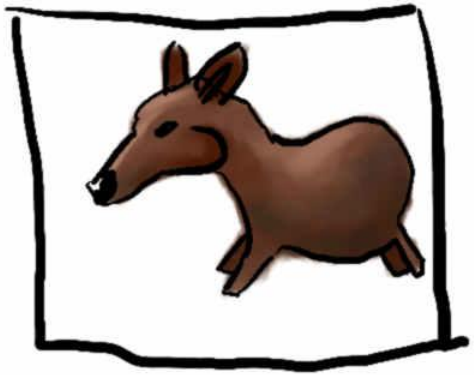
BIG DATA BUSINESS BY HEKIMA

[Big Data Business]

Alguns problemas são difíceis de serem traduzidos em programas

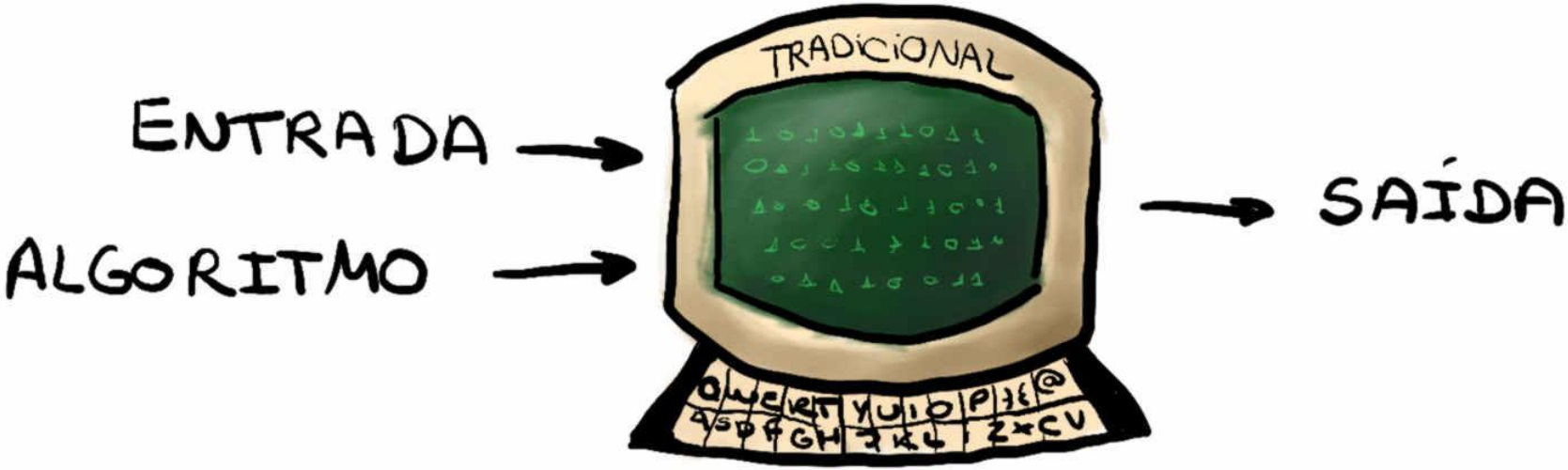
CARRO

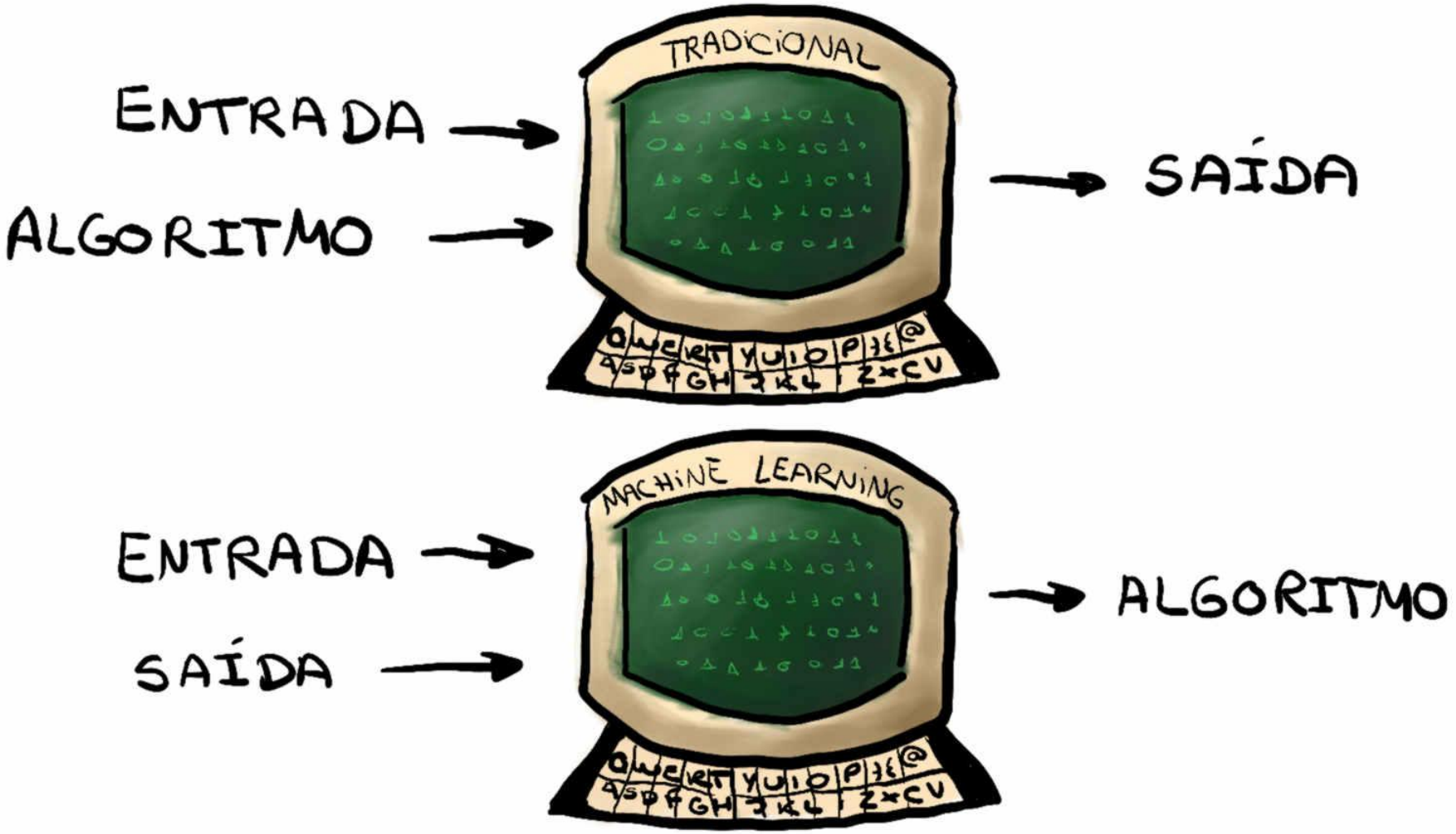
AVIÃO



CACHORRO

Mas temos diversos exemplos

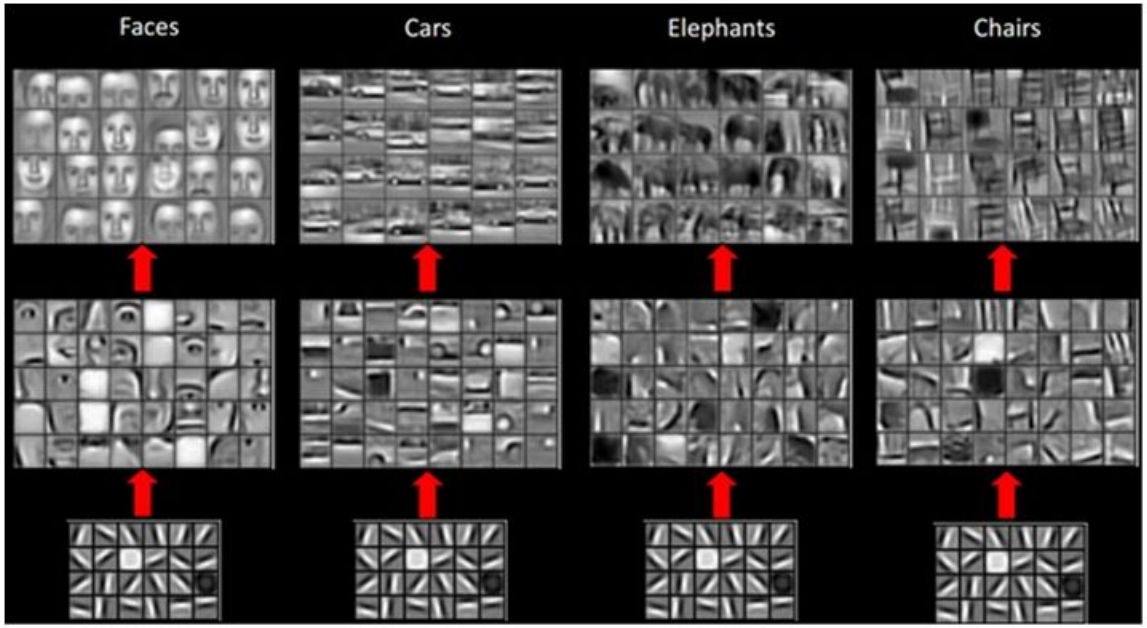

ENTRADA
ALGORITMO
TRADICIONAL
SAÍDA


TRADICIONAL
ENTRADA
ALGORITMO
SAÍDA
MACHINE LEARNING
ENTRADA
SAÍDA
ALGORITMO

“Big Data is not actually about the DATA

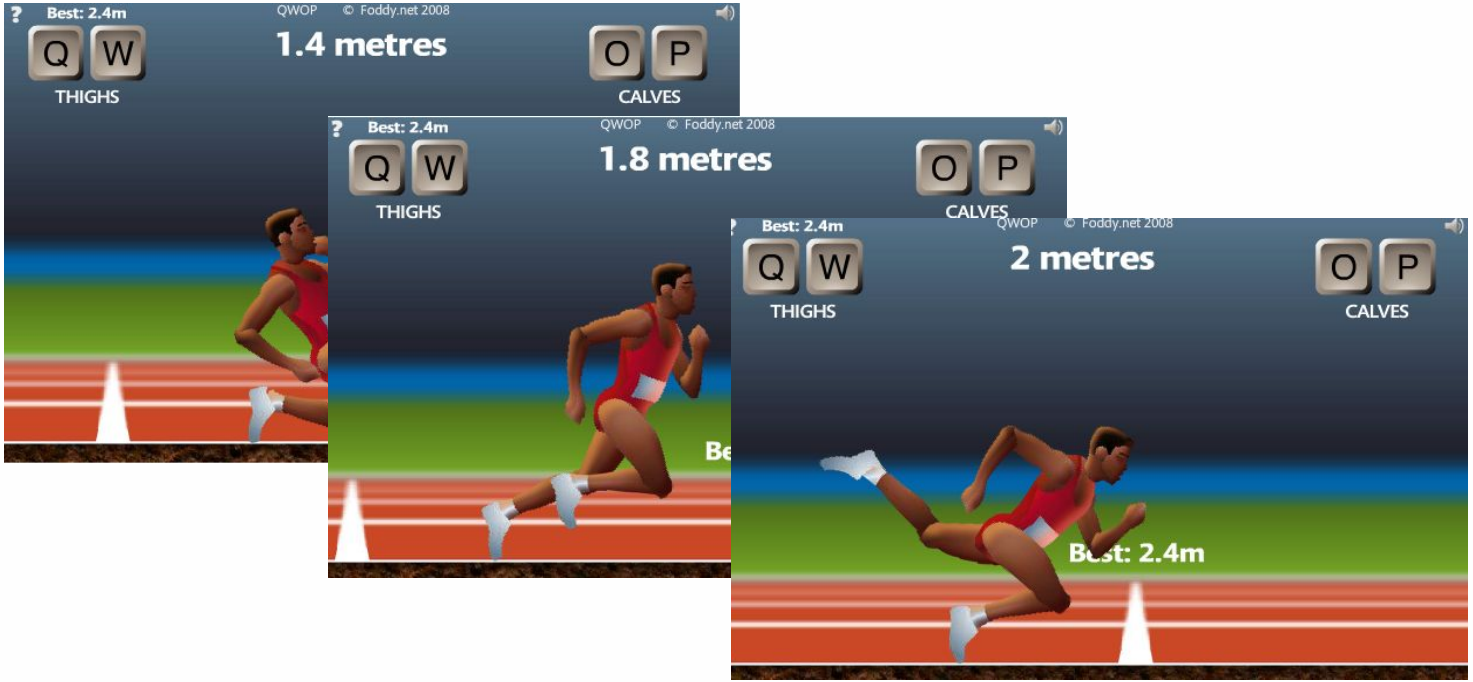
Gary King. Diretor Inst. for Quantitative Social Science de Harvard

Big Data nos permite aprender modelos



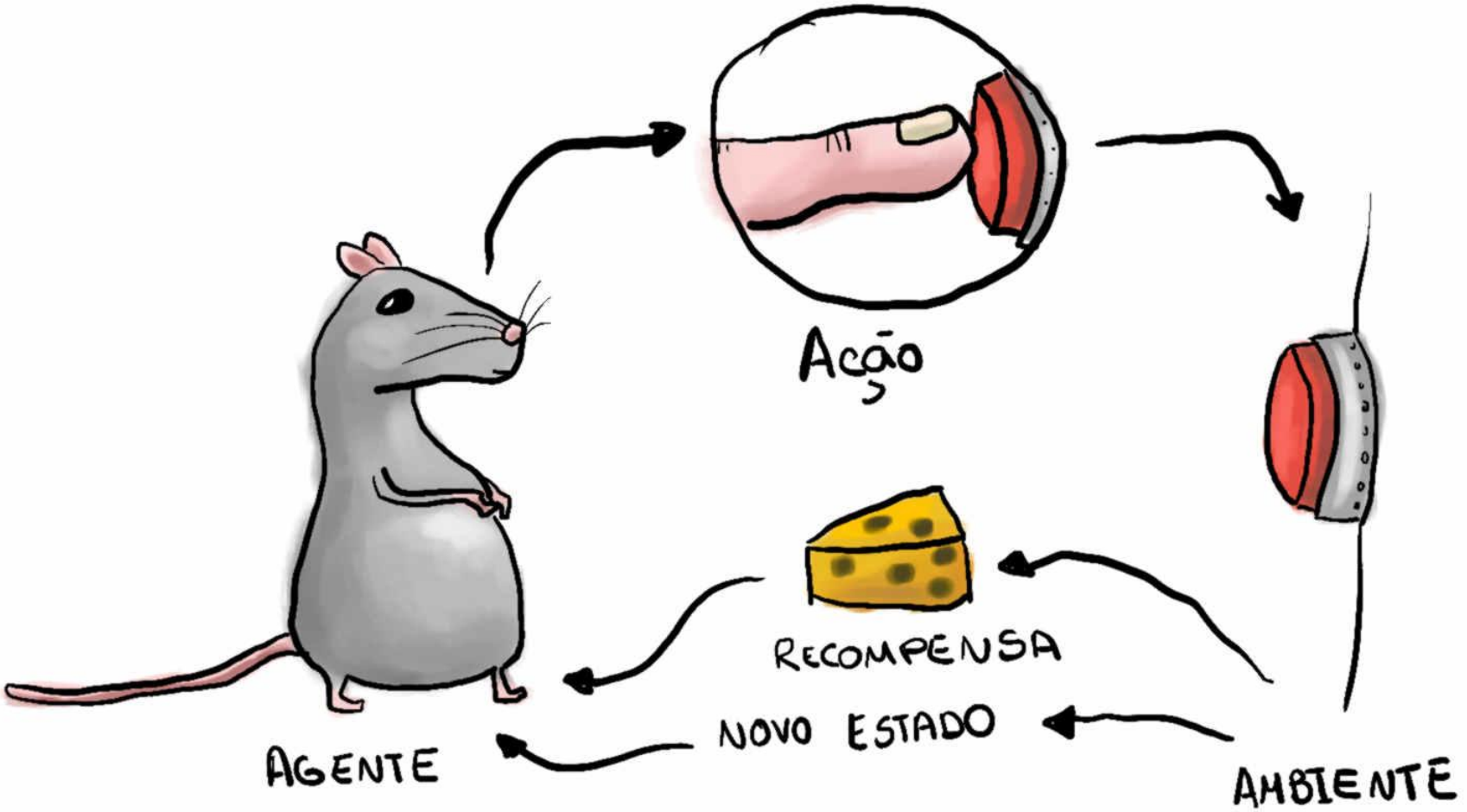
COMPLEXOS

Em algumas aplicações temos grandes quantidades de dados sem ocupar GBs



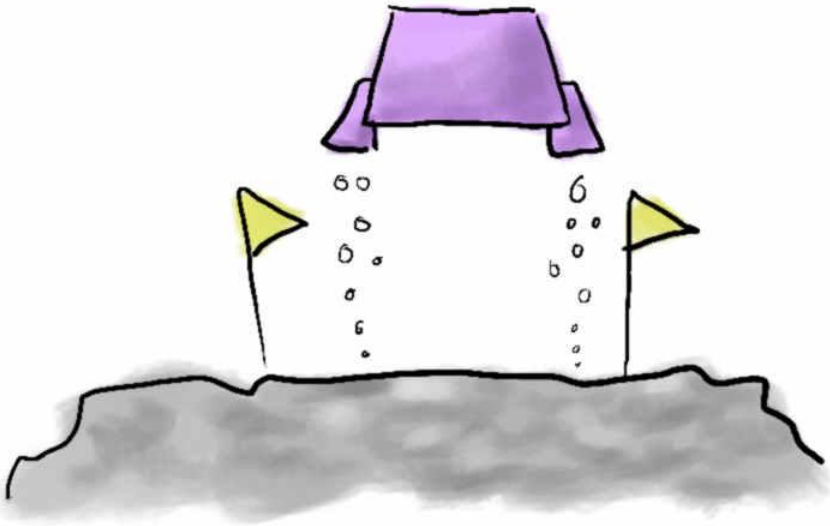
Reinforcement

Learning


Ação
RECOMPENSA
NOVO ESTADO
AGENTE
AMBIENTE

Alguns problemas de exemplo:

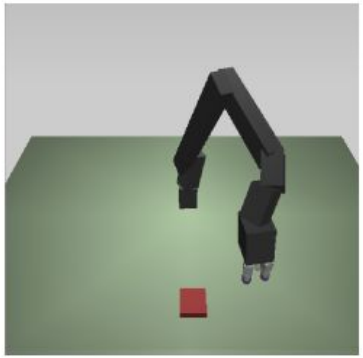
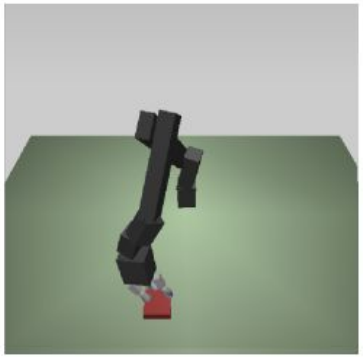
LUNAR LANDER

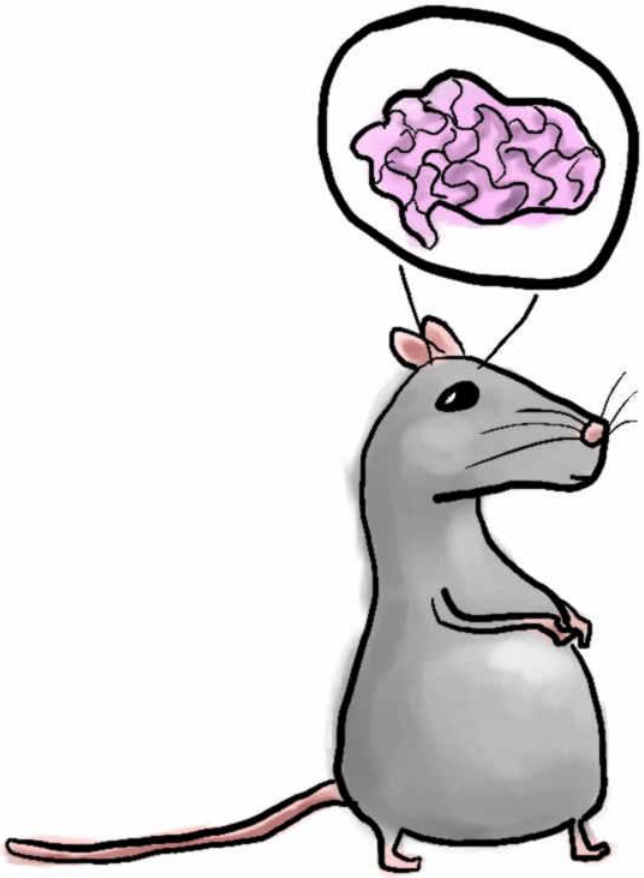


CARTPOLE

I KNOW
KUNG FU

sim-to-Real Robot
Learning from Pixels
with Progressive Nets



Qual ação devo escolher no atual estado

$\pi : S \rightarrow a$

- π: Política
- S: estado
- a: ação


VOU EXECUTAR A AÇÃO
QUE IRÁ ME EMPURRAR
PARA FRENTE!
valor esperado ação t =
r_{t+1}

Valor esperado da ação$_t$ =

r_{t+1} + r_{t+2} + r_{t+3}

r_{t+4} + r_{t+5} + r_{t+6} ...

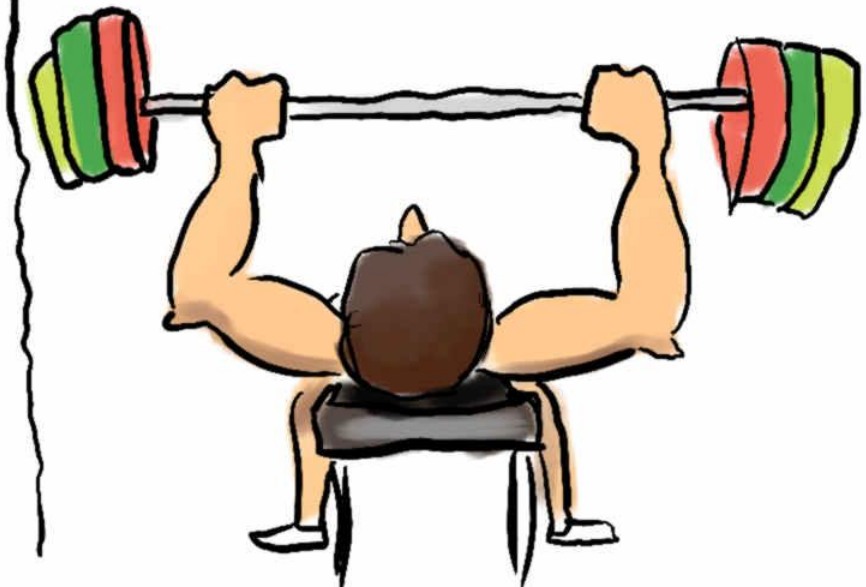
Parecido com a vida:

- TEMPO
- INTERAÇÃO
- RECOMPENSA
- Exploração

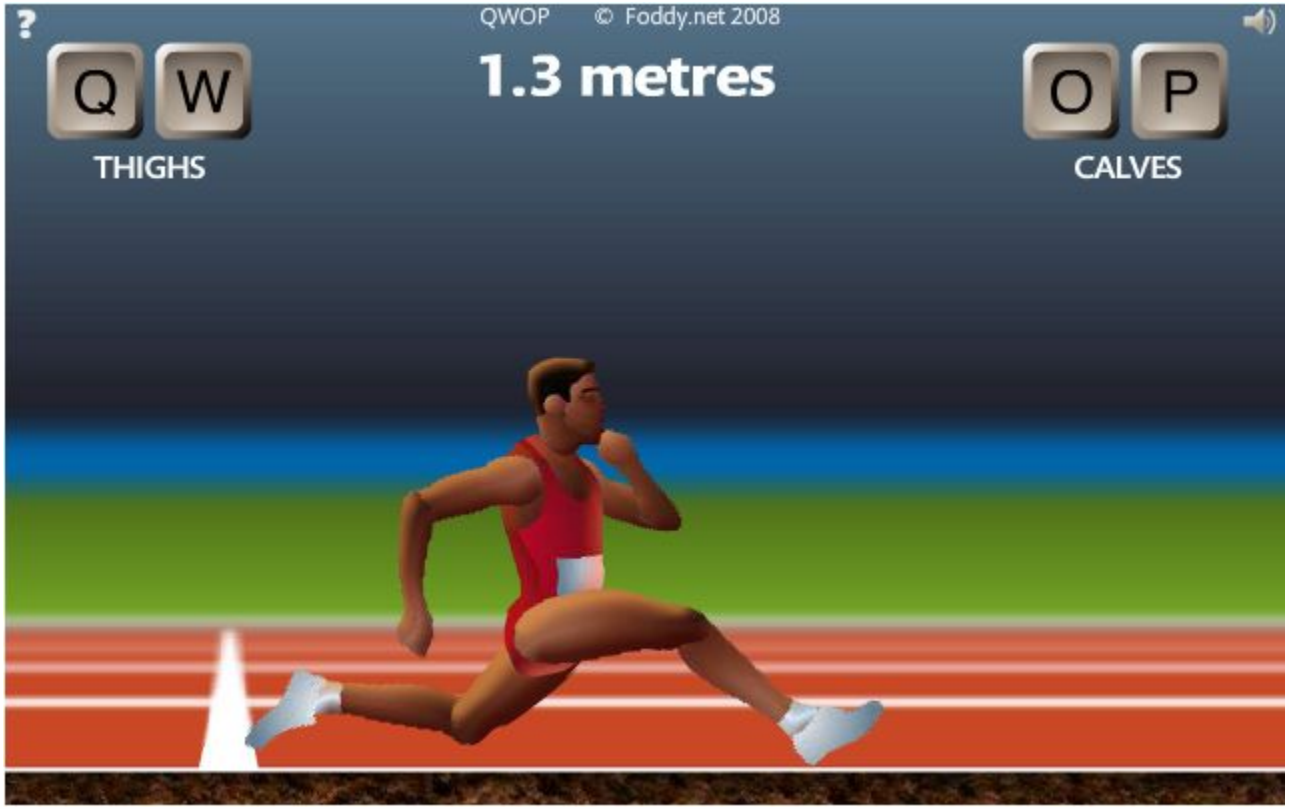
E até os mesmos dilemas:

RECOMPENSA DE CURTO PRAZO

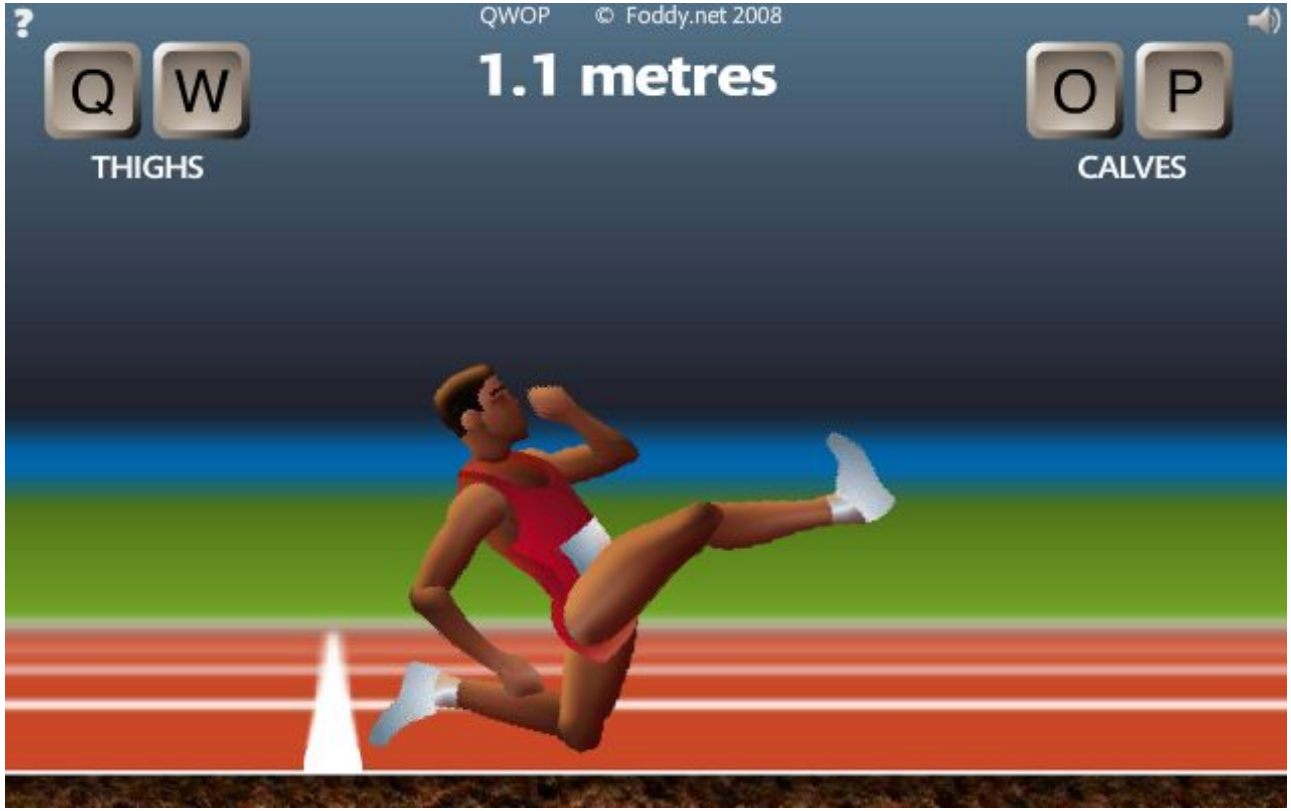
RECOMPENSA DE LONGO PRAZO



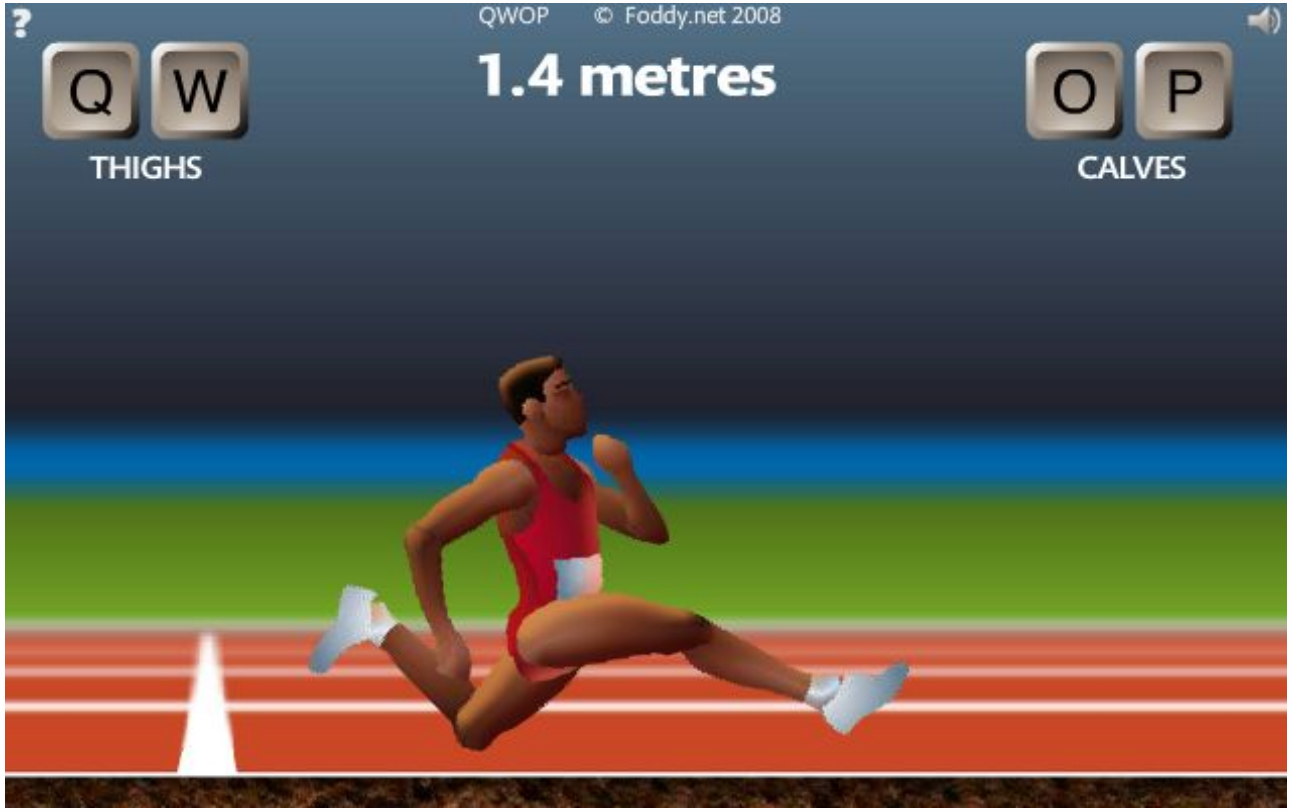
EXPLORATION VS EXPLOITATION



EXPLORATION VS EXPLOITATION



EXPLORATION VS EXPLOITATION



Como usar **RL** em problemas reais ?

DeepMind

- Diversos sensores espalhados pelo data center
- IA controla 120 variáveis (ex. cooler)
- **40%** de redução nos custos de arrefecimento

AlphaGo vs Lee Se-dol

Google's AlphaGo AI beats Lee Se-dol again to win Go series 4-1

By Sam Byford on March 15, 2016 05:00 am Email @345triangle

THE LATEST HEADLINES

Sigur Rós ar
streaming a
drive around

Finding Dor
for biggest d
opening for
film

Eli Roth is di
Death Wish
starring Bru

Stanford Autonomous Helicopter

Stanford Autonomous Helicopter

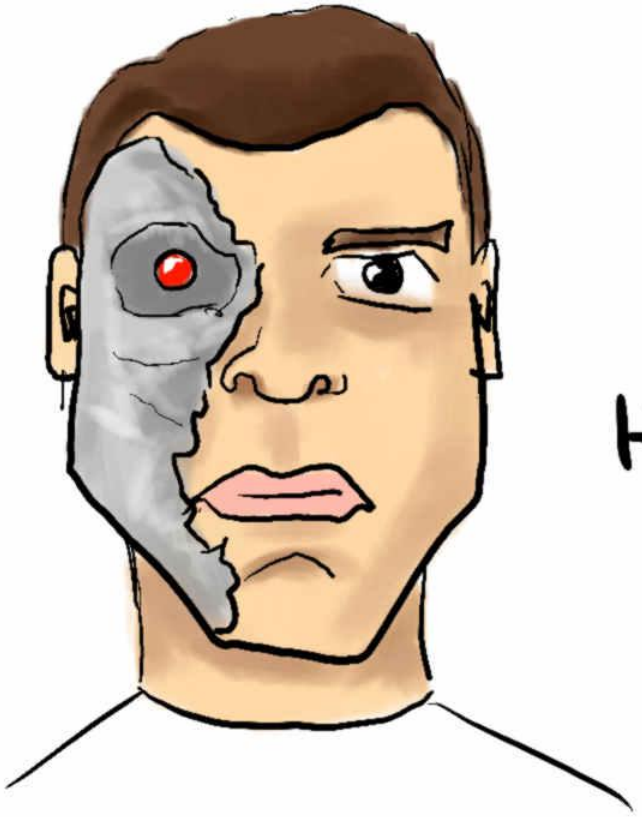
Stanford Autonomous Helicopter

Stanford Autonomous Helicopter

Stanford Autonomous Helicopter

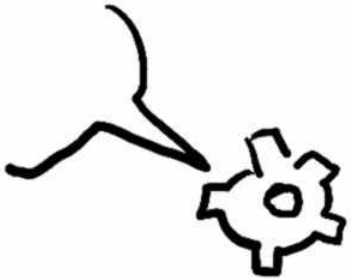
- CONTROLE DE TRÁFEGO AÉREO
- CONTROLE DE MOTORES E REATORES
- INTERFACE CÉREBRO - MÁQUINA
- ASSET MANAGEMENT
- ROBÓTICA
- ADMINISTRACÃO DE REDE
- ADMINISTRAÇÃO DE RECURSOS

DOMINAR A RAÇA HUMANA



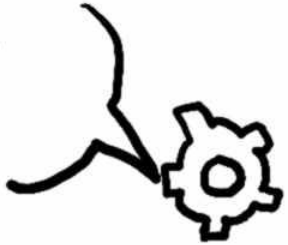
Vai chover, eu preciso de jaqueta?

Não

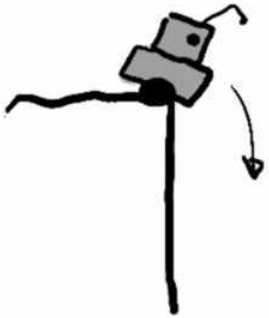


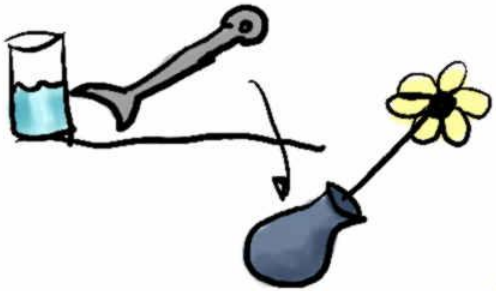
E de um guarda chuva?

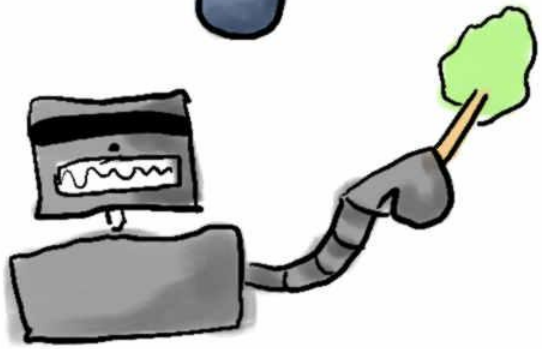
Você gosta de futebol?



- Exploração Segura
- Evitar efeito colateral negativo
- Hacking de recompensa







CONCLUSÕES:

- Alguns problemas podem gerar dados continuamente
- Reinforcement Learning: Agente, Estado, Ação
  Recompensa
- Queremos a ação que maximiza todo o
  reward Futuro
- Diversas oportunidades de pesquisa em
  segurança

Obrigado,

Thiago Cardoso

thiago.cardoso@hekima.com
twitter.com/tncardoso

HEKIMA

+ 55 (31) 2516 – 4812
contato@hekima.com

BELO HORIZONTE
Av. Álvares Cabral, 1315
Salas 7 e 8 - Lourdes
MG – Brasil | 30170-001

SÃO PAULO
Rua Casa do Ator, 919
2º andar - Vila Olímpia
SP – Brasil | 01404-000