



# Reward is enough for social intelligence

Gustavo Berger  
Michel José Moraes  
Vitor Calcete Marques

# ● Oque é **inteligência**?

Como agentes tomam decisões?

---

No mundo tangível

# ● Oque é **inteligência**?

Como agentes tomam decisões?

---

No mundo tangível



# ● Oque é **inteligência**?

Como agentes tomam decisões?

No mundo tangível



Sucesso medido por recompensas

Porque recompensas são suficientes para medir inteligência?

- 1 Recompensas & incentivos promovem o desenvolvimento de novas habilidades específicas

# ● Oque é **inteligência**?

Como agentes tomam decisões?

No mundo tangível



Sucesso medido por recompensas

Porque recompensas são suficientes para medir inteligência?

- ① Recompensas & incentivos promovem o desenvolvimento de novas habilidades específicas
- ② Recompensas & incentivos desenvolvem o uso conjunto de diversas áreas cognitivas



# ● Oque é inteligência?

Como agentes tomam decisões?

No mundo tangível



Sucesso medido por recompensas

Porque recompensas são suficientes para medir inteligência?

- ① Recompensas & incentivos promovem o desenvolvimento de novas habilidades específicas
- ② Recompensas & incentivos desenvolvem o uso conjunto de diversas áreas cognitivas

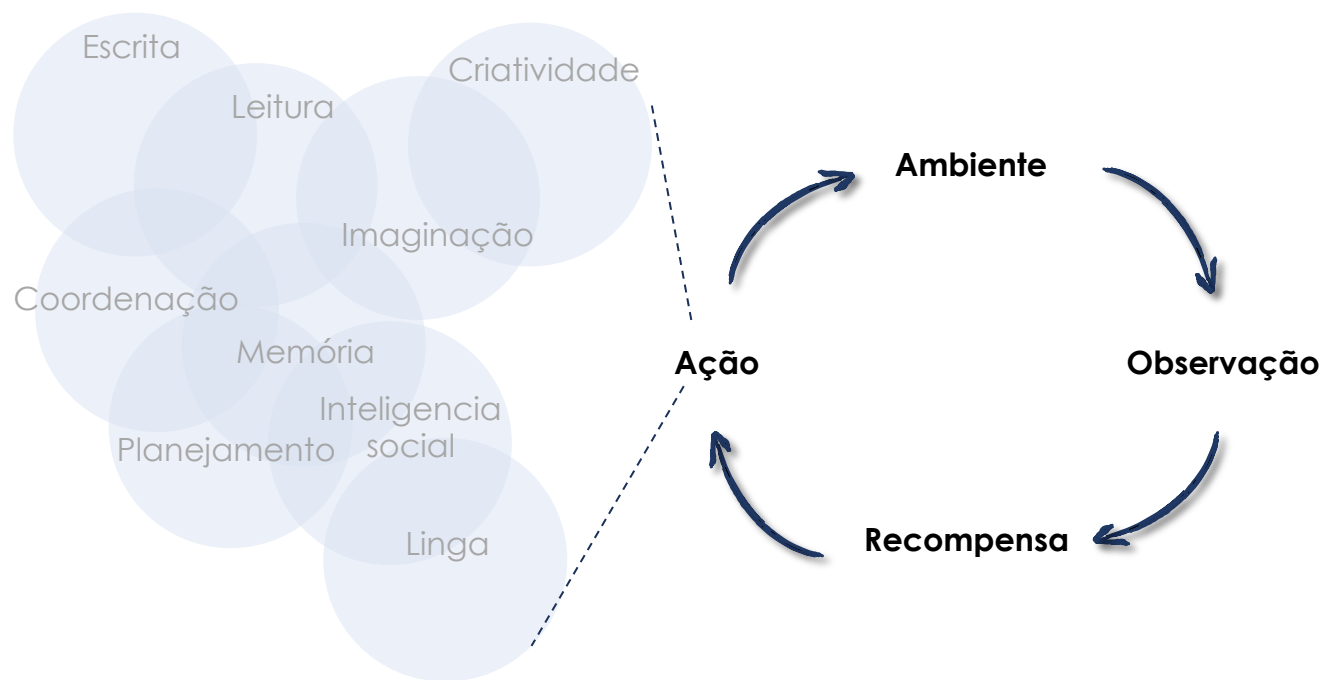
*“intelligence is the computational part of the ability to achieve goals in the world”*

John McCarthy

# ● Como funciona o ciclo de recompensas?

Grande conjunto de habilidades necessárias para a repetição de recompensas

Quais habilidades são necessárias para entender a maximização de recompensas?



- Promove um entendimento mais profundo do comportamento e inteligência;
- Habilidades são usadas para uma **meta única (singular goal)**;
- Explica o porquê das habilidades surgirem, ao contrário de entender o que a habilidade faz; **Exemplo:**

	Função	Motivo
Diferenciar jacarés de lagartos	Saber dizer qual animal é de cada espécie	Não morrer ao confundir lagartos com jacarés

# ● Contexto e definições

## A questão de Reinforcement learning - Agentes

Oque são **Agentes**?

Agentes são sistemas que recebem uma Observação  $O$  em um tempo  $t$  e retornam uma ação  $A$ . **Entidade tomadora de decisões**

$$A_t = \alpha(H_t)$$

$H_t = O_1, A_1, \dots, O_{t-1}, A_{t-1}, O_t$ . Demonstra a sequencia de interações passadas entre o agente e o ambiente

$A_t$  é uma ação no sistema que depende do histórico do agente e de suas observações

$\alpha$  simboliza o próprio agente – Assim como suas limitações

- Capacidade de processamento limitada;
- Quantidade de informações limitada;
- Tempo finito para processar as decisões;

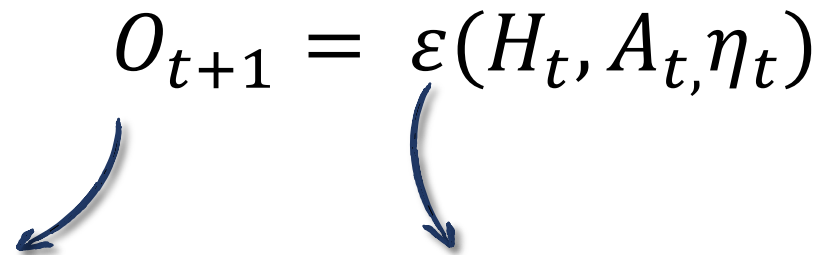


# ● Contexto e definições

## A questão de Reinforcement learning - Ambientes

O que são **Ambientes**?

Ambientes são sistemas que recebem uma ação  $A_t$ , em um tempo  $t$ , e respondem com uma nova observação no tempo seguinte  $O_{t+1}$

$$O_{t+1} = \varepsilon(H_t, A_t, \eta_t)$$


$O_{t+1}$  é a observação entregue ao agente no próximo instante de tempo

$\varepsilon$  simboliza a função do ambiente, que retorna uma observação com base em 3 fatores:

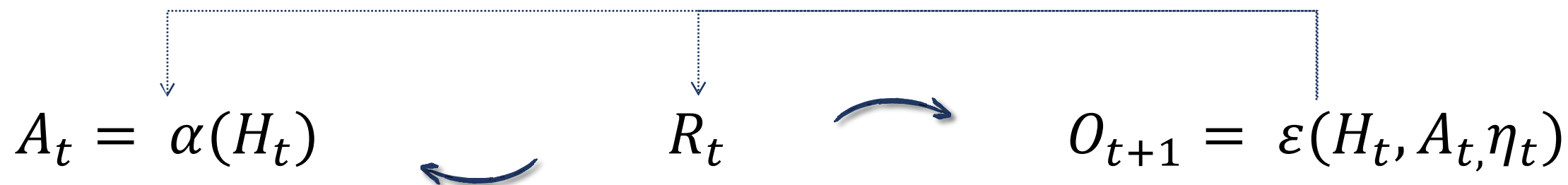
- $H_t$ : O histórico de ações (impacto que essas causaram no ambiente)
- $A_t$ : A última ação promovida pelo agente
- $\eta_t$ : Fonte de aleatoriedade no ambiente

# ● Contexto e definições

## A questão de Reinforcement learning – Conectando os conceitos

Como **Recompensas** ligam as duas pontas?

Recompensas auxiliam na representação de metas através de incentivos acumulados



Recompensas são essenciais para promoverem feedbacks intermediários. Sem esses feedbacks, não é possível avançar no aprendizado

- Representação de recompensas e metas através de uma média ponderada de objetivos
- Representação de recompensas e metas através de interações com humanos (likes)

# ● A importância de entender recompensas

Qual é o objetivo do autor com o artigo?

Porque recompensas são importantes para AI?

## **Hipótese:**

- *"Intelligence, and its associated abilities, can be understood as subserving the maximisation of reward by an agent acting in its environment."*

## **O que isso significa e principais impactos em AI:**

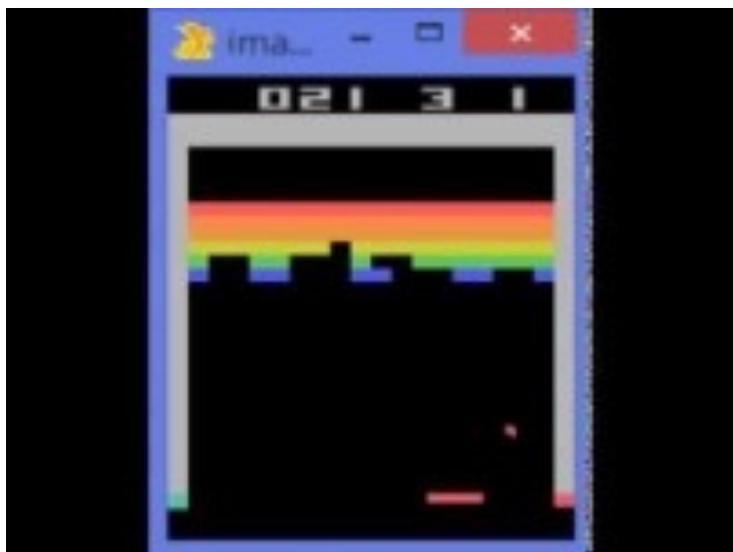
- Um bom agente maximizador de recompensas poderá desenvolver habilidades e comportamentos únicos, associados com inteligência (construindo um agente verdadeiramente inteligente)

# ● A importância de entender recompensas

Atari 2600

**Agente inteiramente inteligente**

Incentivos e feedback intermediário são chave para o reinforcement learning



**Assim como visto em aula, reinforcement learning aprende com tentativa e erro**

- ① **Começo de jogo sem estratégia definida**
- ② **Tentativa e erro levam a estratégias mais inteligentes**
- ③ **Encontra a estratégia vencedora, que maximiza as recompensas. Diversas ações são necessárias para isso:**
  - Focar as jogadas em um dos cantos do tabuleiro
  - Perceber que a peça está em um bom posicionamento para abrir o "túnel"
  - Se movimentar até uma posição que o ângulo de incidência do projétil acerte o alvo desejado

# ● Teoria dos jogos – Dilema do Prisioneiro

## Dilema do Prisioneiro

Entendendo o exemplo clássico de teoria dos jogos

## Aplicações práticas

Como dilemas do prisioneiro são aplicados na vida real?

	Colaborar	Trair
Colaborar	1 ano    1 ano	10 anos    Livre
Trair	Livre    10 anos	5 anos    5 anos

# ● Teoria dos jogos – Dilema do Prisioneiro

## Dilema do Prisioneiro

Entendendo o exemplo clássico de teoria dos jogos

	Colaborar	Trair
Colaborar	1 ano   1 ano	10 anos   Livre
Trair	Livre   10 anos	5 anos   5 anos

## Aplicações práticas

Como dilemas do prisioneiro são aplicados na vida real?





# ● Teoria dos jogos – Dilema do Prisioneiro

## Dilema do Prisioneiro

Entendendo o exemplo clássico de teoria dos jogos

	Colaborar	Trair
Colaborar	1 ano   1 ano	10 anos   Livre
Trair	Livre   10 anos	5 anos   5 anos

## Aplicações práticas

Como dilemas do prisioneiro são aplicados na vida real?

### Jogo "A Evolução da Confiança"



# ● Teoria dos jogos – Equilíbrio de Nash

## Exemplo: Escolhas de Jantar

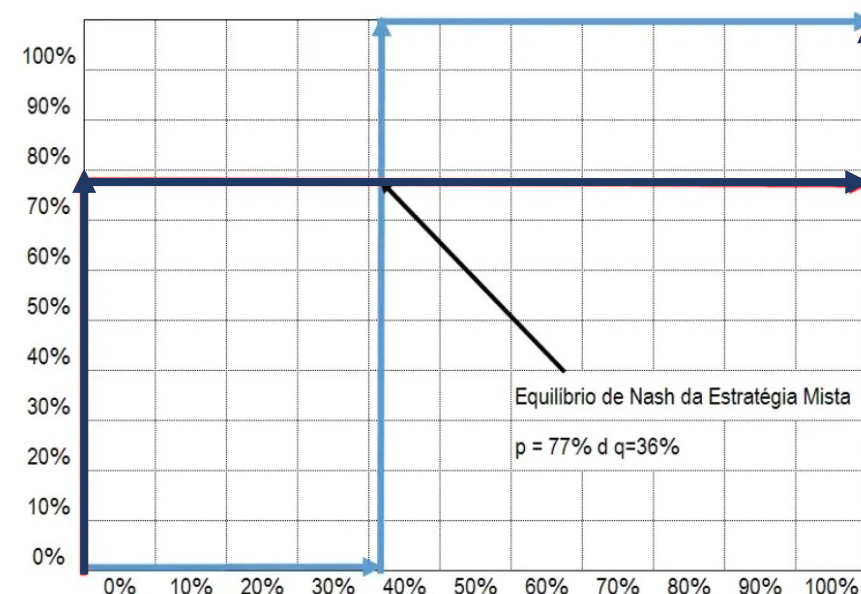
Ilustrando com mais um exemplo da vida real

	Carne	Peixe
Carne	A e B vão juntos comer Carne	A vai sozinho comer peixe e B vai sozinho comer carne
Peixe	A vai sozinho comer peixe e B vai sozinho comer carne	A e B vão juntos comer Peixe

O equilíbrio de Nash representa uma situação em que, em um jogo envolvendo dois ou mais jogadores, nenhum jogador tem a ganhar mudando sua estratégia unilateralmente.

## Equilíbrio de Nash

Solução para a teoria dos jogos



Quantificando níveis de felicidade para cada um dos resultados, é possível traçar a curva que otimiza a felicidade de ambos os participantes

# ● Inteligência social

Qual é a melhor forma de tomada de decisões em situações sociais

---

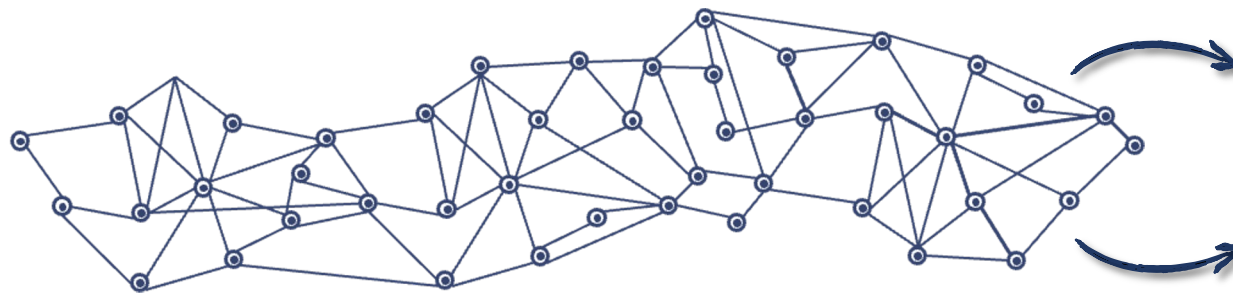
Equilíbrio de Nash VS Reinforcement learning

	<u>Equilíbrio de Nash e teoria dos jogos</u>	<u>Maximização de recompensas</u>
Otimização Zero-sum games	 <i>Garante o melhor resultado, assumindo o piores comportamento entre adversários</i>	 <i>Utiliza os comportamentos sub- optimal para encontrar soluções</i>
Ambientes de aprendizado	 <i>Não é possível interagir corretamente com diversos agentes, encontrando um valor não único em general-sum games</i>	 <i>Agentes aprendem entre si, tem um unico valor ótimo</i>

# ● A importância dos ambientes

## Colaboração entre agentes

Porque reinforcement learning se torna bem sucedido em ambientes com diversos agentes?



Observação do comportamento de outros agentes no mesmo ambiente, cria um ciclo virtuoso de aprendizado pelo feedback alheio

O comportamento de cada um dos agentes, afetará o ambiente, que por sua vez, afetará os demais agentes, dessa forma, é necessário inteligência social para lidar com esse ambiente

Em um ambiente com diversos agentes, é necessário o **desenvolvimento de inteligência social**

Questionamento/  
Críticas ao artigo

*O ambiente inicial não era complexo e os agentes não eram inteligentes. Quem veio antes? Como acontece uma evolução dessa forma?*

# ● Conclusão

## Questionamentos finais

---

Quais perguntas ainda precisam ser respondidas?

- ① Qual o melhor número de agentes em um ambiente? Qual é o melhor ambiente possível?
- ② Quais são as melhores estruturas e sistemas de recompensa?
- ③ O que mais – além de maximização de recompensas – poderia ser suficiente para inteligência?
- ④ Como minimizar as limitações de processamento dos agentes?

The background features a series of overlapping, curved shapes in various shades of blue, ranging from a deep navy to a light sky blue, set against a white background. The shapes create a dynamic, modern feel.

# Q&A