



O QUE TEMOS DENTRO, FORA E SOBRE NOSSAS CABEÇAS?

Tópicos atuais da relação entre
Filosofia e Ciências Biológicas:
uma abordagem interdisciplinar

João Cortese

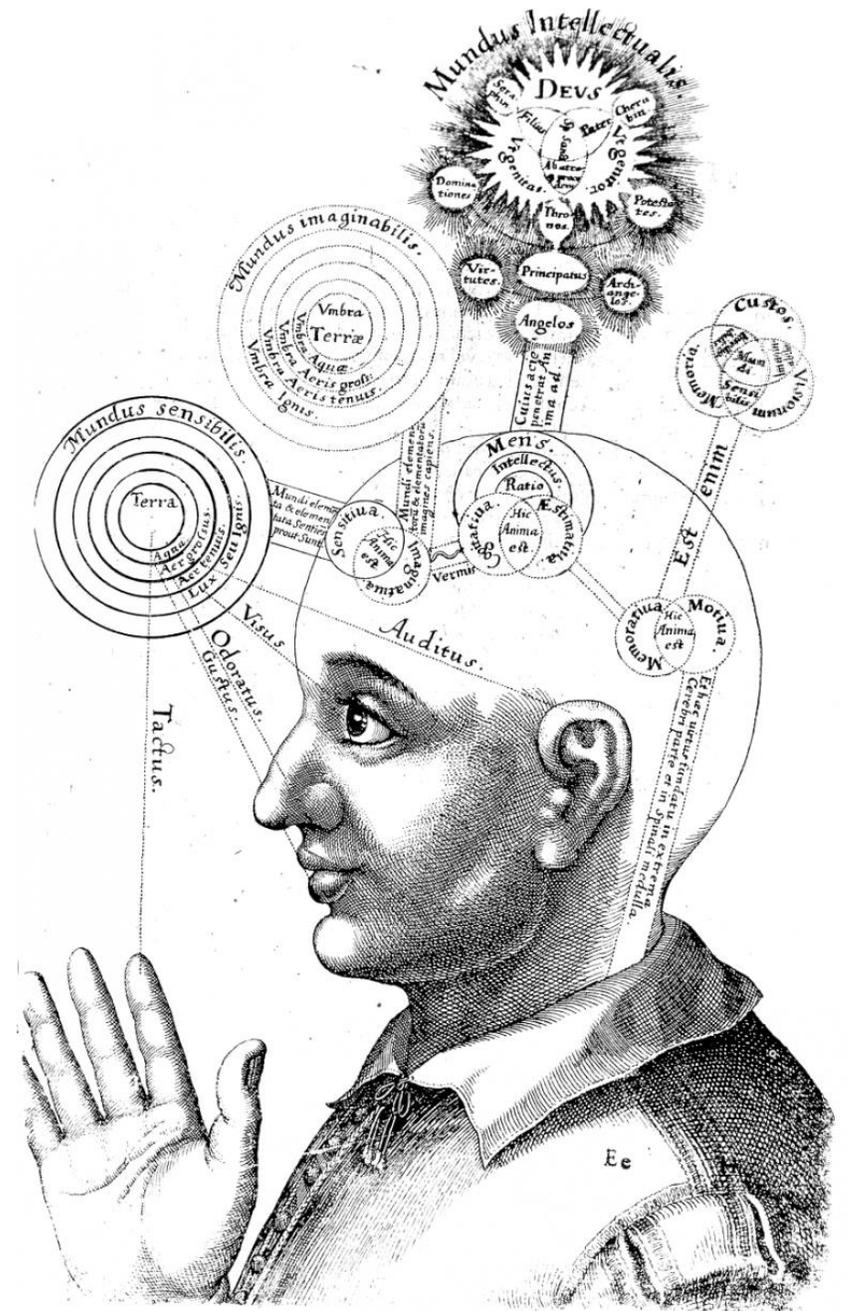
QUESTÕES EPISTEMOLÓGICAS: SUJEITO E OBJETO

objeto: do latim *ob* + *iacio*,
“jogar adiante”

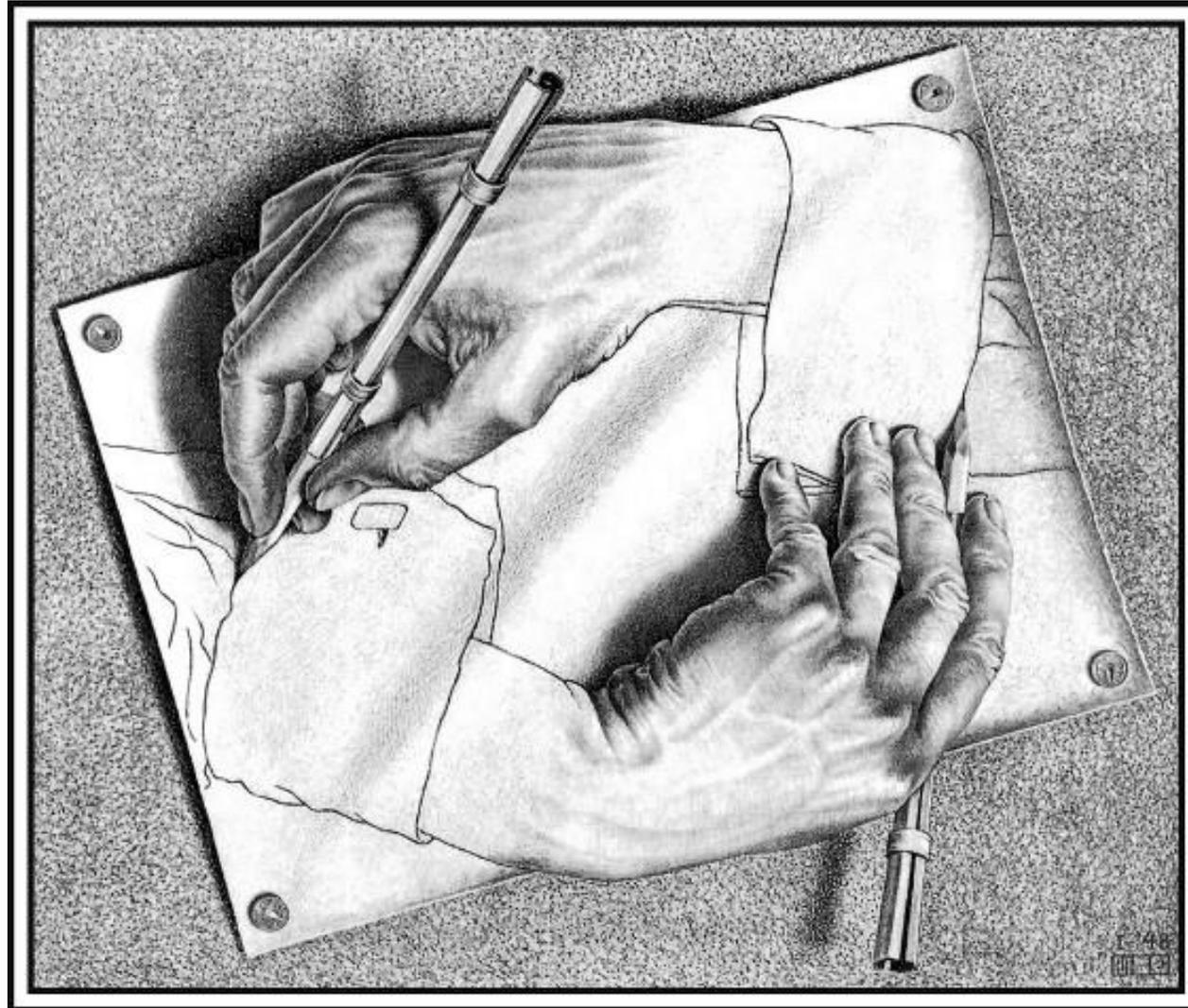
grego *parabolé*:

“jogar ao lado” → comparar

sujeito: do latim *subjectus*,
“colocado abaixo”



QUESTÕES METODOLÓGICAS: REFLEXIVIDADE



M. C. Escher, 1898-
1972



**É POSSÍVEL QUESTIONAR SOBRE O QUE TEMOS
DENTRO DE NOSSAS CABEÇAS?**

É POSSÍVEL A MIM DISCUTIR O MEU LIVRE-ARBÍTRIO?

LUCRÉCIO, *DE RERUM NATURA*

“Se todo movimento está interconectado, o novo surgindo do antigo em uma determinada ordem - se os átomos nunca desviarem de modo a originar algum movimento novo que romperá os laços do destino, a sequência eterna de causa e efeito - qual é a fonte do livre arbítrio possuído pelas coisas vivas em toda a terra?”

LIVRE-ARBÍTRIO?

A evolução implicaria a *ausência* de livre-arbítrio humano?

Francis Galton, primo de Darwin: “our ability to choose our fate is not free, but depends upon our biological inheritance”

Darwin nunca se pronunciou explicitamente sobre a questão, mas em um caderno de 1838 (quando tinha 29 anos), se referiu ao livre-arbítrio como “uma desilusão geral”

The descent of man: “At the moment of action, man will no doubt be apt to follow the stronger impulse; and though this may occasionally prompt him to the noblest deeds, it will more commonly lead him to gratify his own desires at the expense of other men.” → mas a indicação desta **tensão** não parece ser contrária à existência do livre-arbítrio

Impressão de livre-arbítrio

&

Causa e efeito → Determinismo

→ como conjugar ambos? É mais fácil defender a *ausência* de livre-arbítrio dentro da ciência?

Samuel Johnson: “All theory is against the freedom of the will; all experience for it.”

John Locke: “I cannot have a clearer perception of any thing than that I am free.”

Mas seria o livre-arbítrio apenas uma *ilusão* da percepção?

DESCARTES – UM DUALISMO

Para Descartes, a liberdade da vontade é “auto-evidente”, uma noção inata a nós

Porém, Descartes via o corpo e o cérebro de maneira mecanicista

Mas uma máquina não poderia ter livre-arbítrio

Solução para esta tensão: um local para sediar o eu racional → **a glândula pineal**

A única parte do cérebro que não é dupla. Ponto de interação entre o corpo de um lado e a mente e a alma de outro. O livre-arbítrio estaria além da capacidade de investigação natural.

Dualismo abandonado por boa parte dos filósofos e neurocientistas

MAS SE NÃO TEMOS LIVRE-ARBÍTRIO...

Por que/como convencer alguém sobre algo? Ex. conservação da biodiversidade

Por que ensinar valores cívicos?

Então a própria reflexão sobre livre-arbítrio... Não é fruto do livre-arbítrio! Como abordar a questão?

LIVRE-ARBÍTRIO E EVOLUÇÃO

A evolução se baseia em processos físicos → predictibilidade → livre-arbítrio como ilusão?

No entanto, mesmo os que rejeitam o livre-arbítrio parecem viver na prática como se ele reinasse sobre nós (David Barash). Por que o livre-arbítrio tem tanta importância?

Roy Baumeister: “We cannot break the laws of physics, but we can act in ways that go far beyond physical causation”

Os que crêem que o livre-arbítrio é ilusório argumentam que o sentido de responsabilidade individual que ele acarreta teria sido selecionado evolutivamente

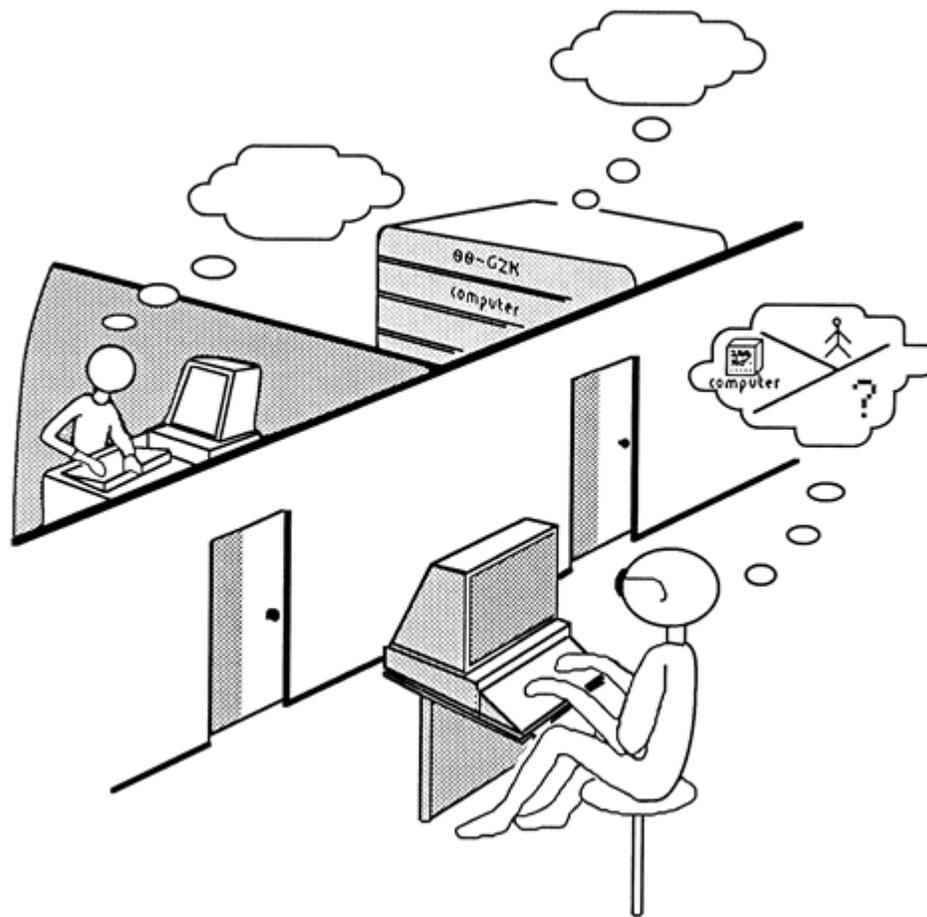
Raymond Tallis: se o livre-arbítrio é uma ilusão, como é que ele foi selecionado positivamente? Se for uma ilusão, ela é auto-realizadora



UMA INTELIGÊNCIA ARTIFICIAL PODE TER UM ESTATUTO ÉTICO PRÓPRIO?

Uma máquina que passasse em um teste interativo deveria ser tratada de maneira ética?

POR MEIO DE UM CHAT (TESTE DE TURING)



VOL. LIX. No. 236.]

[October, 1950

MIND

A QUARTERLY REVIEW

OF

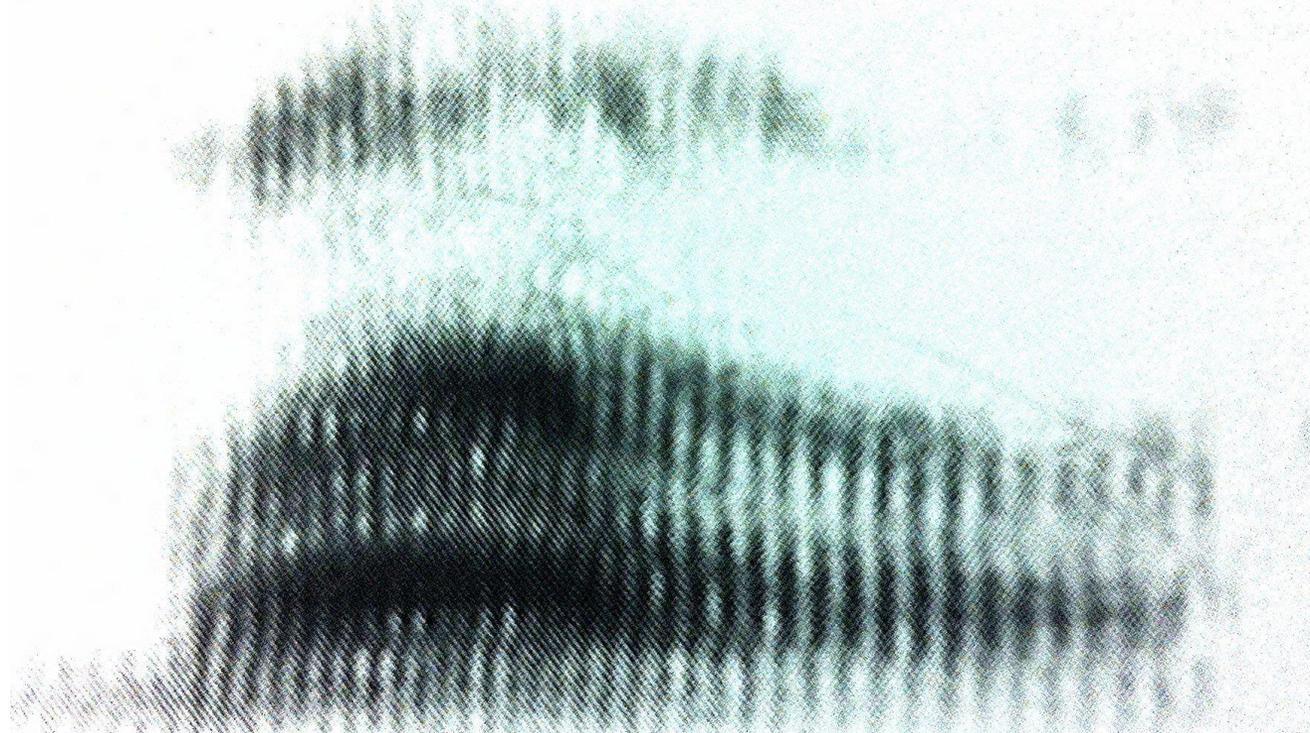
PSYCHOLOGY AND PHILOSOPHY



I.—COMPUTING MACHINERY AND
INTELLIGENCE

By A. M. TURING

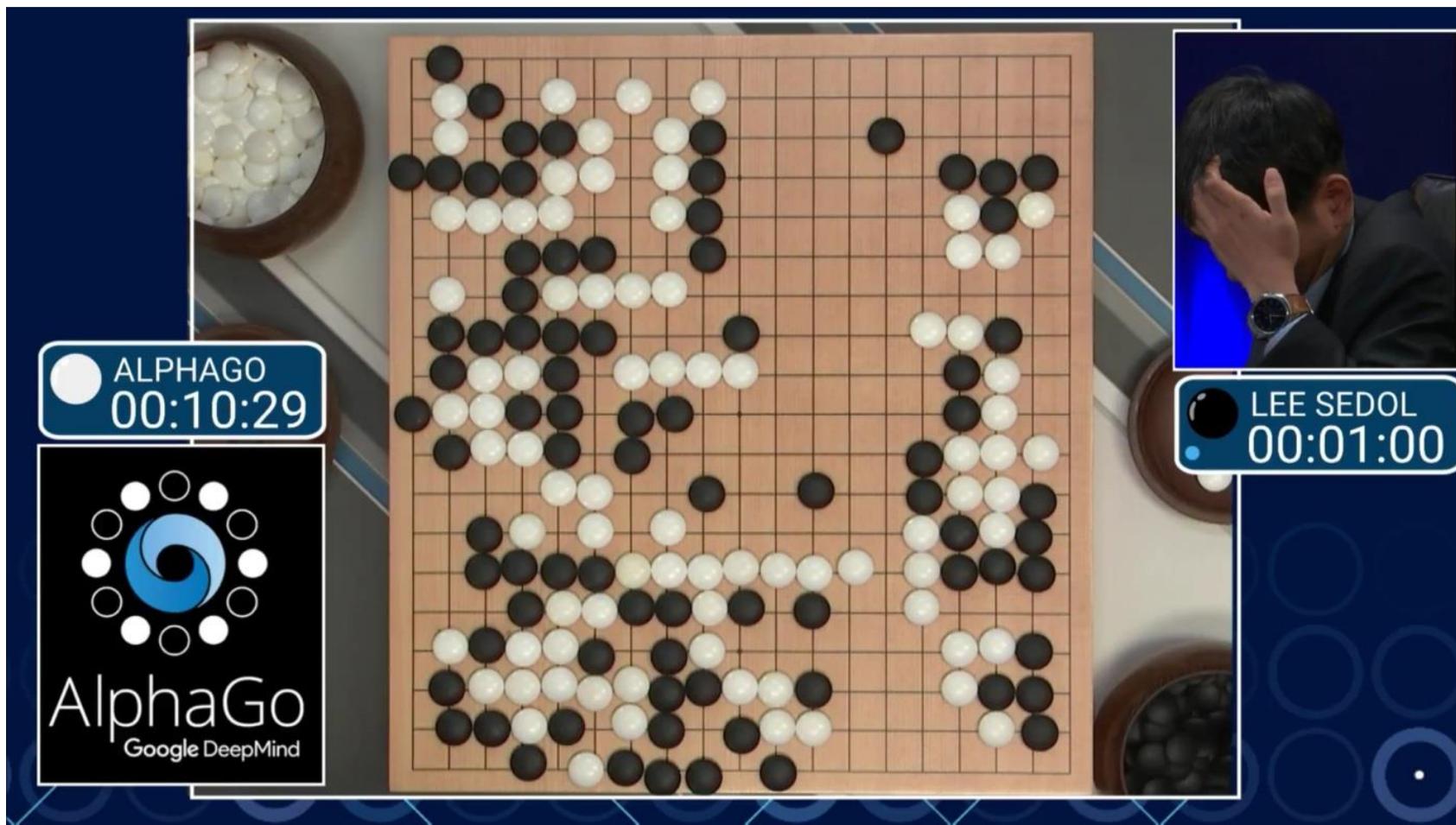
FALANDO POR ÁUDIO



A spectrogram for "whoa." (Lorenzo Tlacaoel / CC BY 2.0)

Gershgorn, D. "Google's voice-generating AI is now indistinguishable from humans", *Quartz Media*, 26 de dezembro de 2017. <https://qz.com/1165775/googles-voice-generating-ai-is-now-indistinguishable-from-humans/>. Ver Shen, Jonathan, et al. "Natural TTS Synthesis by Conditioning WaveNet on Mel Spectrogram Predictions." *arXiv preprint arXiv:1712.05884* (2017).

VENCENDO NO GO



THE NEXT REMBRANDT (2016)

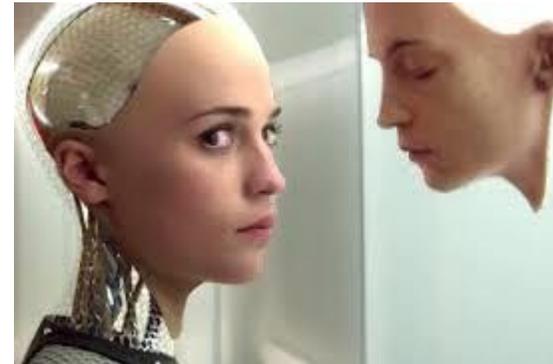




ROBOT THEATER



ROBÔ SOPHIA E ALÉM



etc., etc., etc.

ou seja: potencialmente, máquinas poderiam
passar em qualquer teste interacional adaptável a elas

O BEHAVIORISMO

B. F. Skinner (1904 – 1990)

Uma psicologia comportamental, que não se compromete com as *causas* das ações humanas

Condicionamento: reforço ou inibição

A MODELAGEM DO COMPORTAMENTO *FUNCIONA*

A teoria do behaviorismo “foi usada, por exemplo, para ensinar pombas a jogar tênis de mesa. Quem diria ser possível ensinar pombas a jogar tênis de mesa?” (Dalrymple 2017, p. 28).

“Na verdade, não estou certo de que dá para ensinar: o tênis de mesa envolve mais que meramente bater uma bola de um lado a outro por cima de uma pequena rede esticada sobre uma mesa verde, apesar de ser impressionante que uma pomba tenha mais essa habilidade. Entre outras coisas, o tênis de mesa real abrange o desejo de vencer algo abstruso como um jogo, um desejo que é difícil de acreditar que as pombas tenham; e que também saibam as regras. Provavelmente nenhuma pomba sabe registrar os pontos nem celebrar a vitória ao chegar aos 21 pontos. **Em outras palavras, não demonstrará nenhum sinal de comportamento que indique que compreendeu o significado do que está fazendo”** .

(Dalrymple 2017, p. 28)

MAS E A INTENCIONALIDADE?

“O que começou como metodologia tornou-se ontologia. Um adágio antigo do diagnóstico médico diz que a ausência de evidências nem sempre é a evidência da ausência, mas os behavioristas ignoram esse sábio chamado para a modéstia. Em vez disso, começaram a acreditar que estímulo e resposta era só o que havia na vida humana, que tudo que é humano pode ser explicado dessa maneira. Embora risível, isso foi levado extremamente a sério por muitos”. (Dalrymple 2017, p. 28)

**É QUIXOTESCO PRESSUPOR QUE POSSUÍMOS ALGO
ALÉM DO QUE O COMPORTAMENTO PODE REVELAR?**



O QUE O SER HUMANO TEM “A MAIS”?

Qualia?

A ética não como uma questão de fazer, mas como uma questão de *ser* – mas ser o que?

Pensar sobre a inteligência artificial é a outra face de se pensar sobre a inteligência humana. O estudo da inteligência artificial e de sua diferença ao homem faz apelo a uma antropologia fundamental

O aspecto puramente interativo não pode dar uma resposta sobre a constituição de um agente ético, *nem no caso dos homens e nem no caso das máquinas.*

A QUEM CABE O ÔNUS DA PROVA?

Deve-se mostrar que as máquina podem pensar / têm estatuto moral

OU

Deve-se mostrar que o homem possui algo além do que aquilo que pode ser encontrado na máquina?

A CONSCIÊNCIA: UMA QUESTÃO CIENTÍFICA?

“Resposta” científica ao problema da consciência

A consciência emerge de um cérebro “construído” ou não? *Blue Brain Project.*

Questão não respondida cientificamente

BLOOM E HARRIS 2018

But how will we know if our machines become conscious?

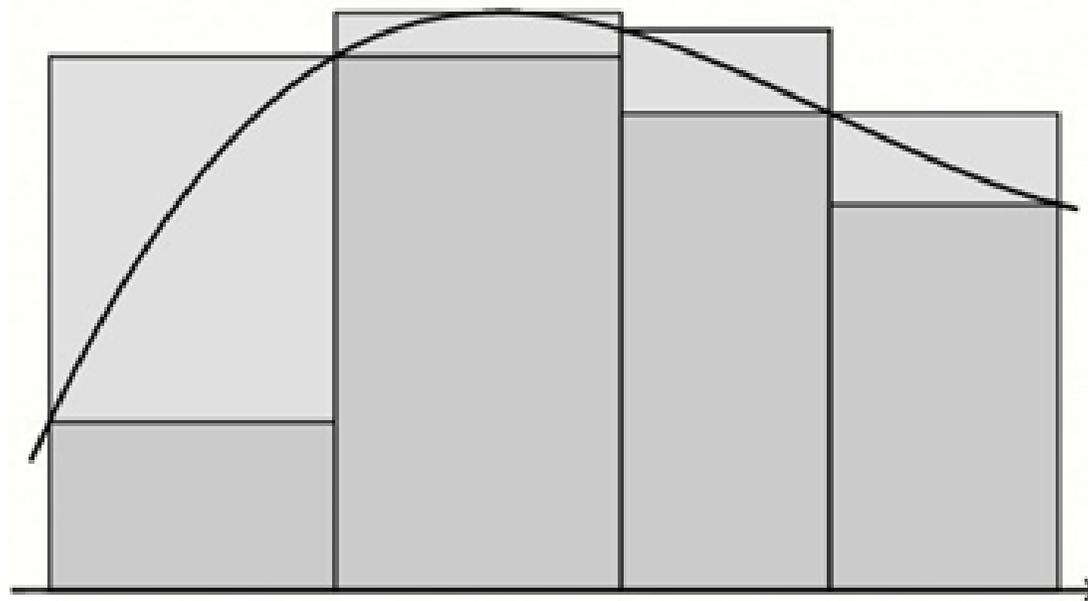
Descartes argued that one's own consciousness is beyond any possibility of doubt. In the case of others, we're never absolutely sure. Many of us have entertained, if only for a moment, the idea that everyone else might be a zombie: laughing, crying, complaining, rejoicing, but with no one home. Perhaps scientists will eventually discover the signature of consciousness, and then we will be able to test for it in our robots, as well as in animals and one another. **But it is certain that we will build machines that seem conscious long before we arrive at that point.**

INTERAÇÃO

“Se me arrisco a olhar pela janela os homens passando pela rua, não falho em dizer, ao vê-los, que vejo homens... e, todavia, o que vejo dessa janela senão chapéus e capas, que cobrem fantasmas ou bonecos que se movem somente por meio de molas? Mas considero que sejam homens realmente e, portanto, compreendo, unicamente pelo poder do juízo que reside em minha mente, o que acredito ver com meus olhos”.

(Descartes, *Segunda Meditação*)

APROXIMAÇÃO NAS MATEMÁTICAS



INFINITO X INDEFINIDO

“o infinito” X “tão grande quanto se queira” [indefinido]

Uma teoria do erro:

Se a diferença é “menor do que uma grandeza dada qualquer”, logo as duas quantidades comparadas são “iguais”

O controle do erro permite uma aproximação indefinida, aceita na prática como uma solução válida.

Aristóteles, *Física* (III, 1 1): os matemáticos não precisam do infinito atual para seus cálculos, mas apenas de um finito tão grande quanto se queira

UM EXEMPLO SIMPLES

$$10 \div 3 = 3,333\dots$$

$$3 \times 3,333\dots = 9,999\dots$$

MAS

$$3 \times 3,333\dots = 1$$

DENNETT (2013): “COMPUTADORES HUMANOS”



Dryden Flight Research Center E49-0053 Photographed 10/49
Early "computers" at work. NASA photo



É assim que Turing (1936, p. 251) declara que “**podemos agora construir uma máquina para fazer o trabalho deste computador [humano]**”. Nesta passagem, diz Dennett (2013, p. 571), “vemos a redução de *todas as computações possíveis* a um processo sem uma mente [*mindless*]”.

Sim, mas por outro lado podemos dizer que este processo sem mente repete *apenas* todas as computações possíveis. A questão, no fundo, é saber o que é passível de computação.

BOSTROM & YUDKOWSKI (2011/2014) SOBRE A ÉTICA DA INTELIGÊNCIA ARTIFICIAL

Princípio da Não-Discriminação do Substrato: “Se dois seres têm a mesma funcionalidade e a mesma experiência consciente, e diferem apenas no substrato de sua aplicação, então eles têm o mesmo *status moral*”.

Princípio da Não-Discriminação da Ontogenia: “Se dois seres têm a mesma funcionalidade e mesma experiência de consciência, e diferem apenas na forma como vieram a existir, então eles têm o mesmo *status moral*”

IRREDUTIBILIDADE DO HOMEM

A despeito da “equivalência” comportamental

Infinito humano X **indefinido** da máquina

Qual ética, dada a distinção ontológica?

CONTATO (1997)



“Você não é real, nada disso é real; quando eu estava inconsciente, você extraiu meus pensamentos, minhas lembranças”.

“Achamos que assim seria mais fácil para você”, respondem *eles*, sob a forma do seu pai.

O *outro*, enquanto *outro*, sempre se apresenta de maneira mais diferente do que o *eu* gostaria de crê-lo

E. LÉVINAS (1906-1995)

A ética *precede* a ontologia

O outro é infinitamente transcendente a mim

Trata-se, segundo Lévinas, de fazer uma escolha moral: tratar o outro como um *sujeito moral* ou como um *objeto*? Colocar-se face a face

A alteridade é irredutível – mas uma máquina seria um outro?

A MÁQUINA APRESENTA UMA ALTERIDADE AO HOMEM?

O *Deep Blue* jogava xadrez *melhor* do que seus programadores (mas apresentava ele um modo original de pensar?)

No caso do AlphaGo,
parece que sim – os
humanos aprendem
com o modo de jogar
da máquina



ENCONTRAMOS UM “TU” NA TECNOLOGIA?

A máquina como um *Tu* ou o homem como um *isso*?

AND YET...

Mesmo que não haja um estatuto ético ontológico, vimos que pode fazer sentido tratar eticamente a um robô ou a uma IA

Kant e os animais

Uma “moralidade derivada” em relação a um objeto

UM EXEMPLO PROSAICO

Um garoto trata *Siri* de maneira indelicada, insultando-a quando lhe ordena que forneça alguma informação. A mãe lhe repreende: “filho, não fale assim com ela”. O filho responde: “mas mãe, é só uma máquina”.

UM MODO DE PENSAR A QUESTÃO

Ética “ontológica” – *Eu-Tu* – Há alguém aí? Com quem eu falo?

Agente moral, paciente moral

X

Ética derivada – trato com um *isso* – O que está aí? Com o que eu falo?

“Objeto de moralidade”

ONDE ESTÃO OS LIMITES DA VIDA HUMANA?



Rayna meets a "robot".