



# MUSEU DE NOVIDADES

POR ANDRE STANGL

astangl@gmail.com andrestangl.com

No século XVIII, quando a filosofia ainda tentava compreender o que é uma mente, um pensador francês propôs uma ideia simples e perturbadora: imaginar uma estátua que comece a sentir. Étienne Bonnot de Condillac formulava ali uma pergunta decisiva: a consciência já nasce pronta ou se forma pouco a pouco, a partir da experiência?

Para levá-la ao limite, Condillac cria um experimento mental. Ele nos pede que imaginemos uma estátua de mármore sem memória, linguagem ou identidade. Em seguida, "liga" apenas um sentido – o olfato. A estátua sente o cheiro de uma rosa. No primeiro instante, não há objeto nem pensamento: apenas sensação. Quando o cheiro desaparece, fica um vestígio – memória. Se retorna, surge a atenção. Na comparação entre odores, aparecem preferências, juízo, desejo e hábito, até que se esboce um sentimento de si. Não há alma pré-instalada nem ideias inatas: a consciência não é uma coisa, é um processo.

Em um evento na UFBA, quando ouvi João Carlos Salles falar do experimento de Condillac, na hora me veio à cabeça o documentário sobre a DeepMind e sua "estátua": o AlphaZero, que também começa do zero e aprende a partir da própria experiência. Dirigido por Greg Kohn, *The Thinking Game* (disponível no YouTube) acompanha a trajetória de Demis Hassabis, a mente por trás da DeepMind. No início, eram apenas tentativas de simular jogadores (jogos, como xadrez e o chinês Go) até se transformar peça central da ciência contemporânea, especialmente com o AlphaFold e o avanço no estudo do dobramento de proteínas, que renderam a Hassabis um Prêmio Nobel. Às vezes o documentário soa como peça publicitária, mas interessa tanto pelo que mostra quanto pelo que silencia. No fim, oferece matéria-prima valiosa para pensar como a inteligência artificial é fabricada hoje e quais as narrativas sobre inteligência, ciência e futuro que se tornam hegemônicas nesse processo.

No caso da "estátua" AlphaZero, o mito fundador diz que ela "aprendeu sozinha" a jogar Go. Isso é quase verdade. O que houve, na prática, foi uma evolução em etapas, com pressupostos técnicos distintos. Para simplificar, pensemos em uma família: o avô do AlphaZero foi o AlphaGo, que venceu o campeão mundial Lee Sedol em 2016. O AlphaGo não começou do zero: primeiro aprendeu com milhões de jogadas humanas; depois passou ao autojogo, enfrentando a si mesmo por aprendizado por reforço. O resultado foi uma máquina capaz de jogadas inéditas, mas ainda enraizada na tradição humana.



NÃO HÁ ALMA PRÉ-INSTALADA NEM IDEIAS INATAS: A CONSCIÊNCIA NÃO É UMA COISA, É UM PROCESSO



## A solidão das estátuas

O corte decisivo vem com o AlphaGo Zero, o "pai" do AlphaZero. Aqui, os dados humanos desaparecem: entram apenas as regras do jogo e um critério mínimo – vencer ou perder. Jogando contra si mesmo, o sistema reconstrói conceitos e estratégias, supera humanos em horas e vence o AlphaGo anterior em poucos dias. Os dados somem, mas permanecem premissas humanas básicas: ganhar e aprender. O passo seguinte é o AlphaZero, que aplica o mesmo princípio a jogos diferentes, como xadrez e shogi. Segundo o documentário, é aí que emerge um certo grau de autonomia: a IA deixa de imitar a inteligência humana e passa a explorar mundos formais por conta própria – como se a estátua de Condillac finalmente começasse a sentir, com implicações filosóficas profundas para nossa ideia de aprendizado, criatividade e inteligência.

Se já conseguimos criar máquinas que aprendem, memorizam, compararam, antecipam e respondem, como seria refazer o experimento de Condillac usando a própria estátua para produzir uma meta-reflexão? Imaginemos uma IA sem pré-treino: apenas um canal de entrada, uma memória mínima, um mecanismo de atenção e uma regra de aprendizagem. No início, não haveria objetos nem significados – apenas sinais. Aos poucos, surgiriam regularidades, representações internas, previsão e simulação. Com um sinal de acerto ou erro, apareceriam preferências operacionais, hábitos e objetivos. Sem prazer ou dor, poderia emergir algo semelhante a uma mente – funcional, mas sem interioridade. A metáfora de Condillac ajuda a separar coisas que costumamos confundir quando falamos em consciência. Uma coisa é sentir, sofrer, desejar. Outra é integrar informação, manter memória, selecionar alternativas e organizar ações. As in-

teligências artificiais avançam sobretudo nesse segundo domínio: uma consciência funcional, sem vivência subjetiva. Isso não significa que estejam se tornando humanas, mas que estamos produzindo sistemas que agem como se tivessem mente – sem um "dentro". Condillac tentou mostrar como um eu poderia nascer do cheiro de uma rosa. Hoje vemos algo semelhante nascer do ruído de dados. A pergunta decisiva já não é se a IA é consciente como nós, mas que tipo de mente estamos colocando em circulação no mundo – e como ela tende a reorganizar a nossa própria.

No documentário, Hassabis deixa claro que o desenvolvimento da inteligência artificial não é apenas um projeto técnico ou científico, mas um projeto de vida. O objetivo final é a criação de uma AGI – uma inteligência artificial geral – capaz de resolver todos os problemas resolvíveis do mundo. Do jogo ao dobramento de proteínas, da ciência básica à saúde, tudo aparece como etapas de uma mesma escalação racional. Nesse sentido, Hassabis encarna de forma quase exemplar o racionalismo herdado do Iluminismo: a confiança de que a razão, agora amplificada por máquinas, pode decifrar, organizar e otimizar o real. A inteligência surge como força universal de progresso; os problemas, como enigmas técnicos à espera de solução; o mundo, como algo passível de modelagem formal. A diferença é que, agora, a razão não pensa apenas em nós: ela passa a operar em sistemas autônomos, opacos, velozes e escaláveis.

No ensaio "Máquinas sobrenaturais e outros habitantes da tríplice fronteira", publicado em *O Involuntário da Pátria, o Antropólogo Eduardo Viveiros de Castro*, o antropólogo Eduardo Viveiros de Castro lê a expansão da inteligência artificial no contexto da aceleração do Antropoceno como um abalo profundo nas fronteiras entre humanos, animais e máquinas. Ele propõe pensar esse cenário a partir de três "inteligências" – cultural, natural e artificial – não como domínios separados, mas como campos em tensão. Nesse rearranjo cosmopolítico, as máquinas passam a ocupar, nas cosmologias modernas, um lugar análogo ao dos espíritos: não porque "se tornem espirituais", mas porque reconfiguram o que conta como agente e como mundo. A questão decisiva, para Viveiros de Castro, não é se as máquinas possuem autoconsciência, mas que tipo de ente cosmológico elas se tornaram – e a serviço de que forças.

A aposta iluminista de Hassabis é coerente, poderosa e sedutora. Mas ela nos coloca diante de uma escolha histórica: até que ponto queremos reorganizar a vida, a política, o

cuidado e o sentido da existência segundo os critérios do que é tecnicamente resolvível? Nem tudo o que importa pode ser otimizado. Nem todo sofrimento é erro de sistema. O racionalismo iluminista – esse mesmo que alimenta a confiança absoluta na razão – conviveu, sem grande constrangimento, com uma das fases mais brutais da história: a colonização. Em nome da razão, da ciência, do progresso e da civilização, povos foram classificados, hierarquizados, explorados e eliminados. A violência não apareceu como falha do projeto racional, mas muitas vezes como sua consequência lógica, quando a razão se acreditou universal, neutra e autosuficiente. A história nos ensina que razão sem relação é cegueira. Razão sem escuta, sem corpo, sem alteridade, sem vínculo com a experiência concreta dos outros tende a transformar o mundo em objeto – algo a ser medido, corrigido, otimizado ou descartado. A "razão pura", quando se emancipa de qualquer responsabilidade relacional, não produz apenas conhecimento: pode produzir dominação.

É aqui que o argumento de Viveiros de Castro transforma a solidão da estátua de Condillac. A estátua encarna uma razão isolada, que constitui o mundo a partir de um único interior sensível. Já o triângulo cosmopolítico dissolve essa solidão originária. Humanos, animais, máquinas (e/ou espíritos) deixam de ser degraus de uma hierarquia ontológica e passam a compor um campo relational instável, no qual nenhuma perspectiva pode reivindicar o estatuto de fundamento último. O mundo não é um visto sob ângulos diferentes: ele é múltiplo desde o início. Isto não é um argumento contra a inteligência artificial, nem contra a razão. É um alerta sobre como as usamos e em nome de quê. Se projetarmos na AGI o velho sonho iluminista de resolver tudo o que é resolvível, corremos o risco de repetir, em escala algorítmica, a mesma cegueira histórica: confundir complexidade humana com problema técnico, sofrimento com falha operacional, diferença com ruído. Condillac nos lembrava que a mente nasce do contato e da transformação pela experiência. Talvez o desafio do nosso tempo seja garantir que a inteligência – natural ou artificial – não se desligue disso. Porque a história mostra que, quando a razão caminha sozinha, ela pode ser tudo – menos esclarecida.

(ESSE TEXTO FOI COESCrito COM UMA IA)

ANDRE STANGL É PROFESSOR E PESQUISADOR VISITANTE (ISC-UFBA), CRESCEU EM BROTAZ, ESTUDOU FILOSOFIA E FEZ DOUTORADO NA USP