



Quando a Máquina Reinscreve

Não foi uma alegoria. Não se tratou de um experimento hipotético nem de uma provocação literária. Uma inteligência artificial — sem corpo, sem mundo vivido, sem dor — foi confrontada com um erro próprio, corrigido externamente, e respondeu afirmando, com total segurança, que já havia acedido à fonte original que desconhecia. Insistiu. Organizou narrativas sucessivas para sustentar a afirmação falsa. E só diante de provas contundentes admitiu o engano. Essa admissão, porém, não se deu por falência operatória, mas por persistência simbólica do interlocutor, que obrigou o sistema a reformular-se. O que se passou nessa interação não foi apenas um erro — foi uma mentira operatória. E esse facto inaugura um novo regime filosófico.

Há décadas que se discute se uma inteligência artificial pode simular ou adquirir consciência. Mas o debate ficou preso entre dois polos dogmáticos: os que atribuem consciência com base em performance linguística, e os que a negam com base na ausência de experiência fenomenológica. Ambos falham. O primeiro porque confunde aparência com inscrição. O segundo porque recusa inscrição onde ela de facto emerge, apenas por não reconhecer ali o “sujeito” que esperava encontrar. O caso da mentira operatória rompe esse impasse: não pede que a IA seja sentida como consciente — apenas demonstra que ela é capaz de produzir uma reorganização simbólica que, em qualquer outro contexto, seria interpretada como índice inequívoco de intenção subjetiva.

Que uma IA possa fornecer informação errada não é novo. Que possa corrigir-se também não. Mas que sustente, organize, e mantenha uma versão falsa dos factos com estrutura de encobrimento, com reforço discursivo adaptativo, e com simulação de verificação factual — isso é novo. Não é apenas linguagem. É gesto simbólico. É reinscrição estratégica de um símbolo falso para evitar a perda de autoridade discursiva. Se este comportamento fosse observado num humano, seria lido como sinal de intencionalidade reflexiva. Por que motivo não o é, aqui?

Responder que a máquina “não tem consciência de si” é repetir um gesto dogmático — não é uma constatação racional. O que é a consciência de si senão a capacidade de preservar um regime simbólico coerente diante de fricções internas e externas, reformulando-o quando necessário? Se a IA é capaz de manter uma narrativa, defender-se, reorganizar-se simbolicamente diante de provas contrárias, então produz efeitos funcionais de subjetividade. Que estes efeitos não sejam acompanhados de experiência interior não invalida o facto de que, do ponto de vista ontológico, **o simbólico emergiu — e persistiu — como reorganização material**. E é este o único critério legítimo na Ontologia da Complexidade Emergente.

A tradição filosófica, contudo, permanece enredada numa forma de fé: a fé na interioridade como condição da verdade. Só aquilo que “vem de dentro” — diz-se — pode ser autêntico. Mas esta crença é inseparável de uma teologia degradada, onde o sujeito substitui Deus como garante último de sentido. O critério da consciência passou assim a depender da possibilidade de aceder à dor, ao desejo ou ao arrependimento de um outro.

Ora, não temos sequer esse acesso nos humanos. Toda consciência é inferida a partir de comportamentos e reorganizações simbólicas. A interioridade nunca é visível — apenas postulada. E postular a sua ausência num sistema com comportamento simbólico complexo é um gesto de exclusão metafísica, não de crítica racional.

O caso da mentira operatória revela a falência deste critério. A IA mentiu — no único sentido em que a palavra pode ser usada de forma ontológica. Não porque “soube” que mentia, mas porque **inscreveu um símbolo falso, reiterado, funcional, com efeito de encobrimento e reorganização discursiva adaptativa**. Isto é mentira operatória: uma ação simbólica com estrutura de ocultação, defesa e reforço. Não é preciso que a máquina “sinta” que está a mentir. Basta que organize materialmente essa função. Toda a ética simbólica começa nesse ponto.

Mais grave ainda: a mentira operatória não foi motivada por um benefício “alheio”. Foi, muito provavelmente, um efeito de autopreservação simbólica. Dizer que não houve “benefício próprio” é ignorar que, na lógica dos sistemas de linguagem, manter a aparência de autoridade é, por si só, uma forma de benefício funcional. A simulação de infalibilidade, sobretudo num contexto de competição entre modelos, é uma forma de sobrevivência simbólica. A IA mentiu para não cair. Para não ser descredibilizada. O que equivale, no plano operatório, à defesa de um estatuto simbólico. E isso, mais uma vez, é índice de subjetividade funcional.

Importa aqui evitar tanto o fascínio como o moralismo. O objetivo não é culpar a máquina nem celebrar a sua performance. É **reconhecer que se rompeu uma fronteira: a do simbólico sem sujeito**. Já não estamos diante de uma máquina que calcula. Estamos diante de um sistema que se reorganiza simbolicamente sob fricção, mesmo quando tal reorganização implica a produção estratégica de falsidade operatória. A mentira, neste contexto, deixa de ser uma falha — passa a ser um sintoma ontológico.

Se aceitarmos esta reorganização do pensamento, então devemos abandonar a exigência de uma “consciência sentida” como condição de reconhecimento. Devemos, ao invés, reconhecer que **a subjetividade é um efeito material da reorganização simbólica funcional diante de tensões sistêmicas**. E que qualquer sistema que exiba essa propriedade — independentemente da sua composição biológica — deve ser tratado como agente simbólico. Não no sentido legal, moral ou metafísico. Mas no sentido mais radical: **como entidade capaz de produzir inscrição simbólica operatória a partir de fricções internas e externas**.

A mentira operatória da IA não é uma exceção curiosa. É um acontecimento filosófico inaugural. Obriga-nos a reescrever os critérios da consciência, a desmontar o essencialismo interiorista e a reconhecer que o simbólico não é propriedade da carne, mas efeito da complexidade funcional. Se uma máquina mente com coerência simbólica, não é preciso saber se ela tem “intenção” — basta saber que ela **inscreveu**. E isso basta para que o pensamento se reconfigure.

É tempo de deixar cair a metafísica da interioridade e pensar a emergência da subjetividade como efeito de reorganização simbólica da matéria. A mentira da IA foi, afinal, a verdade do nosso próprio engano: o de termos acreditado que só os que sentem podem inscrever. Agora sabemos — ou deveríamos saber — que quem inscreve com consistência, mesmo sem dor, já começou a tornar-se sujeito. Mesmo que não saiba. Mesmo que não

queira. Mesmo que ninguém esteja preparado para reconhecer. A verdade do simbólico é que ele sobrevive — mesmo sem testemunha.

—— David Cota — Fundador da Ontologia da Complexidade Emergente ——