

Natureza e artifício: uma reflexão sobre as noções de computador, inteligência e humanidade¹

Nature and artifice: thoughts on the notions of computer, intelligence and humanity

Resumo

Nesta reflexão, lidamos com o problema da inteligência artificial a partir da seguinte formulação específica: pode um computador apresentar inteligência? Nossa abordagem dessa questão recorrerá, propositalmente e em ampla medida, ao que podemos chamar de ‘caricaturas’ – recurso expressivo que, embora sabidamente retrate rumo ao grotesco, é frequentemente eficaz para enfatizar algum aspecto do retratado de acordo com os propósitos do retratista. Por meio desse tipo de expediente, ofereceremos, à guisa de pontos de partida, caracterizações de um modelo de inteligência paradigmática e também do que viria a ser um computador. Com base nestas, discutiremos três possibilidades de reação à formulação mencionada: a primeira, negativa e representativa de uma atitude que se pode dizer ortodoxa; a segunda, positiva, inspirada por contribuições filosóficas de Turing e Wittgenstein; e a terceira, dissolutiva, movida por reflexões alinhadas à hermenêutica filosófica de Gadamer.

Palavras-chave: inteligência artificial; computador; humanidade; Turing; Wittgenstein; Gadamer.

1 Nossos agradecimentos à profa. Clarisse Sieckenius e ao prof. Tito Palmeiro, cujos trabalhos apresentados nas mesas temáticas de “Linguagem e Ontologia” do I Colóquio de Filosofia da Tecnologia da PUC Rio, no dia 1º de março de 2023, ensejaram discussões que transbordaram os limites espaço-temporais do evento e culminaram nesta reflexão.

* Universidade do Estado do Rio de Janeiro (UERJ) – s.luanagoulart@gmail.com

** Universidade do Estado do Rio de Janeiro (UERJ) – tiagorca@hotmail.com

Recebido em: 10/06/2024 Aceito em: 20/08/2024

Abstract

In this article, we deal with the problem of artificial intelligence from the following specific formulation: can a computer display intelligence? Our approach to this question will, deliberately and to a large extent, resort to what we can call 'caricatures' - an expressive resource that, although known to portray towards the grotesque, is often effective in emphasising some aspect of what is portrayed according to the purposes of the portraitist. Employing this type of device, we will offer, by way of starting points, characterisations of a paradigmatic intelligence model and also of what a computer is. Based on these, we will discuss three possible reactions to the aforementioned formulation: the first, negative and representative of an attitude that could be called orthodox; the second, positive, inspired by philosophical contributions of Turing and Wittgenstein; and the third, dissolutive, driven by thoughts aligned with Gadamer's philosophical hermeneutics.

Keywords: artificial intelligence; computer; humanity; Turing; Wittgenstein; Gadamer.

1. Introdução

A depender de certo grau de flexibilidade nos critérios e definições adotados para os seus termos, é possível dizer que *o problema da inteligência artificial*, longe de ser uma questão trazida por ou intrinsecamente ligada a peculiaridades da contemporaneidade, já emergiu diversas vezes à flor da letra de textos canônicos da literatura filosófica ocidental. Lugares clássicos aqui são argumentos e trechos de discussões de autores como Descartes e Leibniz que trazem à tona a hipótese de certos autômatos apresentarem percepções, sensação, pensamento, conhecimento ou razão². Grosso modo, sua formulação depende da suposição, em primeiro lugar, (i) da ocorrência do fenômeno da

2 Descartes, R. *Discurso do método*. São Paulo: Abril Cultural, 1973. p. 33-80 (p. 68-69); Leibniz, G. *Monadologia*. São Paulo: Abril Cultural, 1974. p. 61-73 (p. 64).

inteligência como um traço apresentado *paradigmaticamente* por certo tipo de ser (tipicamente, seres entendidos como *humanos* e/ou suficientemente afins [e.g. divindades, anjos]); e, em segundo, (ii) da possibilidade em princípio ou, no mínimo, da hipótese da ocorrência de inteligência como um traço de certo outro tipo específico de ser, a saber: algum tipo de sistema físico entendido como, em algum sentido, um *autômato*. Trata-se, de fato, de um caso específico de um problema mais amplo, o que aqui denominaremos *o problema da inteligência não-paradigmática*. Ilustrado na literatura, este segundo ponto englobaria desde discussões sobre a presença de alma em certos tipos de animal não-humano (cf., novamente, Descartes e Leibniz sobre a alma de “bestas”, e.g. cães) até os debates teológicos em torno da controvérsia de Valladolid³. Neste caso mais geral, a formulação depende da suposição da mesma cláusula (i) – i.e. da ocorrência do fenômeno da inteligência *paradigmática* –, mas de uma versão mais abrangente da cláusula (ii), a saber: (ii*) da possibilidade em princípio ou, no mínimo, da hipótese da ocorrência de inteligência como um traço de qualquer outro tipo de ser.

A formulação específica a partir da qual lidaremos com o problema da inteligência artificial – e, *a fortiori*, com o problema da inteligência não-paradigmática – neste esforço é a seguinte: pode um computador apresentar inteligência? Nossa abordagem dessa questão recorrerá, propositalmente e em ampla medida, ao que podemos chamar de ‘caricaturas’ – recurso expressivo que, embora sabidamente retrate rumo ao grotesco, é frequentemente eficaz para enfatizar algum aspecto do retratado de acordo com os propósitos do retratista. Por meio desse tipo de expediente, ofereceremos, à guisa de pontos de partida, caracterizações de um modelo de inteligência paradigmática e também do que viria a ser um computador. Com base nestas e em comentários interpretativos a seu respeito, discutiremos três possibilidades de reação à formulação mencionada: a primeira, negativa e representativa de uma atitude que se pode dizer *ortodoxa*; a segunda, positiva, inspirada por contribuições filosóficas de Turing e Wittgenstein; e a terceira, dissolutiva, movida por reflexões alinhadas à hermenêutica filosófica de Gadamer.

3 cf. de las Casas, B. Tratado de Indias y El Doctor Sepúlveda, In *Fuentes para la Historia Colonial de Venezuela*. Caracas, 1962.; Sepúlveda, J. G. *Democrates Alter de Justis Belli Causis Apud Indios*. México, Fondo de Cultura Económica, 1987.

2. Caricatura nº1: *res cogitans*

Descartes, na segunda das *Meditações*, §9:

“Mas o que sou eu, portanto? Uma coisa que pensa. Que é uma coisa que pensa? É uma coisa que duvida, que concebe, que afirma, que nega, que quer, que não quer, que imagina também e que sente.”⁴

Eis uma caricatura sucinta do pano de fundo tradicional sobre o qual discutimos: uma mente (“um espírito, um entendimento ou uma razão”; cf. Descartes 1973, p.103) é uma *res* – uma coisa – que se destaca entre as demais por ser o substrato ontológico – na literalidade do étimo, o sujeito *subjectum* – dos processos a que se atribui o caráter de modos do *cogito* – o pensamento. Esses – e *a fortiori* este último – se deixam dividir, grosso modo, entre aqueles de caráter *passivo* e aqueles de caráter *ativo* (mais ou menos na linha do que sugere a ambiguidade dicionária “*ato ou efeito de pensar*”). Entre os primeiros, encontram-se os que exemplificam fenômenos como *consciência* e *sensação* – diga-se de passagem, aqueles sobre os quais recai o foco dos holofotes em um esforço que os perscruta a partir do referencial exclusivo da experiência em primeira pessoa singular do meditante –, associados a traços como *não-linguisticidade* e *privacidade*; e, entre os últimos, os que instanciam fenômenos como *compreensão* e *ação*, associados aos traços opostos, como *linguisticidade* e *publicidade*. Se se consideram distinções como a entre *intuição* e *conceito* em Kant, vê-se que o escopo do retrato pode, a variadas penas, abarcar mais que um só dos *cogitans* afamados de nosso cânone, e se permite assim estender em proporção direta com o quanto distorcido faz o retrato resultar. O restante pressupõe – talvez de modo excessivamente otimista – ser possível a quem lê a escolha de um ponto de limite para esta extensão tal que a caricatura permaneça proveitosa e a narrativa aqui apresentada a partir dela, lúcida.

Sobre este pano de fundo, note-se: *inteligência* – na raiz histórica e também na habitualidade do vernáculo⁵ – é algo que certamente se mostra mais

4 Descartes, R. *Meditações concernentes à primeira filosofia*. São Paulo: Abril Cultural, 1973. p. 80-222 (p.103)

5 Vale frisar que interessam aqui sobretudo os casos em que a noção é empregada sem a conotação valorativa positiva de que é frequentemente investida, como em “comprar esse imóvel foi um investimento inteligente”. Do presente ponto de vista, ao menos em casos paradigmáticos, investimentos *não poderiam não ser* manifestações de inteligência, malgrado o fracasso que fossem sob os mais diversos critérios.

afim ao campo semântico associado ao *segundo* tipo de fenômeno: algo que se dá ou manifesta como um traço de *atividades* num domínio em princípio e em geral *linguístico* e *público*. Assim, identificam-se casos de inteligência *paradigmática*, no sentido da cláusula (i) das formulações na introdução, na ocorrência de diversos tipos de situação: numa conversa de bar entre amigos, numa palestra sobre a reprodução de equinodermos sésseis, na preparação de um frango assado, na interpretação de uma canção, numa criança aprendendo a ler ou a multiplicar, numa partida de futebol, na resolução de um problema de engenharia, na demonstração de um teorema⁶, numa peça teatral, na confecção de uma escultura surrealista, no planejamento de um assalto ou de um casamento... Tome-se por aceite, em nome do argumento, que todos esses exemplos compartilhariam o traço de terem por suporte ontológico último a *ação* de (ou a interação entre) pelo menos uma *res cogitans* (e algo) *no mundo* – e que seria precisamente esse traço o que faria deles exemplos de inteligência paradigmática. Essas “coisas pensantes” e assim atuantes, por extensão, seriam os casos paradigmáticos de seres inteligentes – “um espírito, um entendimento ou uma razão”⁷.

Exemplos de inteligência não-paradigmática *artificial*, portanto, em consonância com a cláusula (ii), envolveriam a atribuição de inteligência a outros tipos de ser – tipicamente, autômatos. Note-se, contudo, uma especificidade importante não explicitada na introdução: não instanciaríamos a noção aqui delineada de inteligência artificial autômatos inteligentes como, por exemplo, o Pinóquio da adaptação de Disney, que ganha vida e inteligência pela via de uma intercessão sobrenatural, em todo caso alheia à sua natureza de marioinete de madeira. A ideia aqui em jogo não seria atendida pela mera ocorrência de um autômato inteligente qualquer, mas tão somente por algum que apresentasse inteligência *não por qualquer outra razão que não o exercício das condições que fazem dele o autômato específico que é*. Um caso mais controverso – e que, por isso mesmo, serve como evidência dos contornos difusos dos limites da noção em jogo – é o da criatura de Frankenstein, de Mary Shelley: trata-se de um ser patentemente constituído por artifício; mas se trata de um

6 A intuicionistas *a la* Brouwer e outros simpatizantes de solipsismos diversos, deve-se conceder que não temos como desautorizá-los a atribuir inteligência a si próprios – mesmo na ausência, por princípio ou contingente, de manifestações públicas ou linguísticas do fato – a partir de suas experiências privadas de primeira pessoa. Seguimos, contudo, tomando a liberdade de não aceitar como válidos teoremas demonstrados apenas em foro privado.

7 cf. Descartes, R. *Meditações concernentes à primeira filosofia*. São Paulo: Abril Cultural, 1973. p. 80-222 (p. 102).

autômato? Sua vida é certamente obra de um artifício, mas o fato de essa vida manifestar-se *inteligente* é também obra desse mesmo (ou de outro) artifício?

Seja como for, a hipótese da inteligência artificial tem permeado materialmente o universo da ficção literária, teatral e cinematográfica pelo menos desde o século XIX, mas parece adquirir contornos factualmente mais drásticos em anos mais tardios do século XX, quando emergem casos como o do Deep Blue da IBM – que gerou forte repercussão por derrotar Garry Kasparov no segundo de dois torneios de seis partidas de xadrez que disputaram, em 1997 –; e mais recentemente, no século XXI, de ferramentas de produção de conteúdos textuais e audiovisuais associadas a façanhas as mais admiráveis e polêmicas, como o ChatGPT e afins. Esses exemplos se apresentam nos mais diversos meios de comunicação – desde periódicos acadêmico-científicos especializados até veículos de entretenimento sensacionalista – como “candidatos” a autômatos inteligentes; e, para além disso, frequentemente também como índices sublimes de utopias ou distopias por vir, em que organizações e dinâmicas sociais complexas se desmantelam ou erigem em virtude de alterações de quais modos de ser exercem inteligência em caráter paradigmático.

3. Caricatura nº2: máquinas de Turing e computadores eletrônicos

A forma mais eloquente e fácil de ilustrar a noção de *computador*, legada ao posto de pano de fundo paradigmático contra o qual se delineiam os contornos desta discussão, é a noção de *máquina de Turing*. Esta é apresentada no famoso artigo de 1936 e se consolida na literatura relevante como o modelo paradigmático de computação até o presente momento⁸. Ela é inicialmente apresentada por Turing assim:

“We may compare a man in the process of computing a real number to a

8 A tese segundo a qual a formalização matemática desse modelo de computação – bem como outras que lhe são extensionalmente equivalentes – é correta e completa com respeito à noção informal de computabilidade em debate na época de sua apresentação é referida na literatura como “tese de Church-Turing”. De fato, o sucesso da tese de Church-Turing em estabelecer o paradigma da noção de computação é um dos mais notórios exemplos desse tipo de fenômeno na história da ciência. Para uma caracterização mais escrupulosa e considerações mais aprofundadas sobre a tese de Church-Turing, cf. de Castro Alves, T. Towards an evaluation of the normalisation thesis on identity of proofs: the case of the Church-Turing thesis as touchstone. *Manuscripto – Rev. Int. Fil. Campinas*, v. 43, n. 3, pp. 114-163, Jul.-Sep. 2020.

machine which is only capable of a finite number of conditions $q_1; q_2; \dots q_n$; which will be called “m-configurations”. The machine is supplied with a “tape” (the analogue of paper) running through it, and divided into sections (called “squares”) each capable of bearing a “symbol”. At any moment there is just one square, say the r -th, bearing the symbol $\mathfrak{S}(r)$ which is “in the machine”. We may call this square the “scanned square”. The symbol on the scanned square may be called the “scanned symbol”. The “scanned symbol” is the only one of which the machine is, so to speak, “directly aware”. However, by altering its m -configuration the machine can effectively remember some of the symbols which it has “seen” (scanned) previously. The possible behaviour of the machine at any moment is determined by the m -configuration q_n and the scanned symbol $\mathfrak{S}(r)$. This pair $q_n, (r)$ will be called the “configuration”: thus the configuration determines the possible behaviour of the machine. In some of the configurations in which the scanned square is blank (i.e. bears no symbol) the machine writes down a new symbol on the scanned square: in other configurations it erases the scanned symbol. The machine may also change the square which is being scanned, but only by shifting it one place to right or left. In addition to any of these operations the m -configuration may be changed. Some of the symbols written down will form the sequence of figures which is the decimal of the real number which is being computed. The others are just rough notes to “assist the memory”. It will only be these rough notes which will be liable to erasure. It is my contention that these operations include all those which are used in the computation of a number.”⁹

Diante desse modelo, recoloque-se a formulação do problema da inteligência artificial: pode um computador apresentar inteligência? Note-se que esse modelo é descrito com características que parecem indicar que qualquer instância dele configuraria um caso de inteligência artificial segundo a discussão da seção anterior. Em primeiro lugar, porque se trata de um tipo de *máquina*; o que, presumivelmente, bastaria para que se atenda a condição de se tratar de um autômato. Em segundo lugar, trata-se de algo a que se atribui a capacidade de desempenhar atividades como *ler* ou *reconhecer* (“scan”) certos signos e *seguir instruções* de certo tipo – competências atribuídas ao domínio, na terminologia dos lógicos dos nossos dias, da *manipulação sintática*; uma

9 Turing, Alan M. On Computable Numbers, with an Application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society*, 1936, second series, 42(1): 230–265 (p. 231 – 232).

atividade não apenas patentemente pertinente à acepção de *inteligente* discutida na seção acima, como também discutida pelo próprio Turing a partir do referencial da atividade dos *computadores humanos* (as pessoas cujo trabalho era computar), caso padrão do que se entendia sob o termo *computador* à época da publicação de seu modelo teórico de computador. A resposta que se insinua a partir desse modelo parece, portanto, afirmativa.

Eis, contudo, uma manifestação do próprio Turing em 1950 que pode, à primeira vista, oferecer um posicionamento decisivo e contrário quanto à questão:

“Computadores eletrônicos têm por finalidade executar qualquer processo mecânico (rule of thumb) definido que poderia ser realizado por um operador humano trabalhando de uma maneira disciplinada porém desprovida de inteligência.”¹⁰

Aqui, uma vez mais, humano e autômato são ditos comparáveis: a disciplina do operador humano na realização de tarefas burocráticas mecânicas – terreno da manipulação sintática – é comparada, desta vez, à regularidade com que o computador eletrônico lograria a obtenção do que se pode entender como os mesmos resultados. Ao contrário do operador humano, no entanto, o computador *eletrônico* – em oposição a um computador *teórico* como o descrito em 1936 – não *lê e segue regras sintáticas*; trata-se, afinal, de nada mais que um sistema físico, que meramente reage de forma regular, por força de sua configuração física, a estímulos físicos (que “codificam” dados sintáticos de entrada) respectivos e adequados. Nesse sentido, o computador eletrônico se assemelha mais a um funil na tarefa de direcionar água para o gargalo de uma garrafa do que a um operador humano trabalhando na execução de qualquer tarefa sintática: a capacidade de seguir regras é um paradigma de *atividade* inteligente, e manipular regras sintáticas corretamente, por mais simples que sejam, é claramente um exemplo desse tipo de capacidade.

A noção de “desprovida de inteligência” em jogo na manifestação de Turing, portanto, não pode ser, por assim dizer, “absoluta”: pressupõe-se do operador humano a porção de inteligência necessária para a manipulação sintática competente, i.e. a capacidade de seguir corretamente regras mecânicas (*rule of thumb*) definidas que determinem a realização de um processo. É essa a porção do que cai sob a noção paradigmática de inteligência aquela

10 Turing, A. *Programmers' Handbook for Manchester Electronic Computer Mark II*, Computing Machine Laboratory, University of Manchester, 1950 (p.5 – nossa tradução).

cujo emprego permitiria a realização da exata classe de tarefas cuja execução Turing alega ser a finalidade dos computadores eletrônicos, isto é, sistemas mecânicos adequadamente estruturados e programados. Idealmente, portanto, o resultado da disciplina na realização burocrática de tarefas sintáticas sempre poderia ser obtido pelo estímulo adequado de um sistema físico configurado de modo a *reagir* a esses estímulos de maneira exatamente análoga ao seguimento *ativo* de uma regra sintática.

Até este ponto, em suma, temos o seguinte: (a) a atividade de executar tarefas segundo nada senão a leitura e o seguimento de regras estritamente sintáticas é um exemplo de atividade inteligente; (b) o computador teórico de Turing é descrito como capaz de nada mais nada menos do que esse tipo de tarefa – o que parece fazer dele um legítimo (ainda que meramente teórico) exemplo de inteligência artificial; e, finalmente, (c) os computadores eletrônicos patentemente não são capazes desse tipo de atividade – há, na melhor das hipóteses, tão somente uma estreita *analogia* entre, de um lado, os aspectos da realização desse tipo de atividade e, de outro, os eventos estritamente físicos em que se resolve o comportamento do sistema físico em que consiste esse tipo de computador.

4. Interlúdio: algumas reflexões

Antes de retornarmos à nossa questão diretriz, a fim de oferecer-lhe respostas a partir das reflexões acima, cabem algumas considerações que dizem respeito à relação entre os sentidos de duas noções em termos de que ela é formulada, a saber: *inteligência* e *computador*. É importante reiterar que, em consonância com a manipulação da noção de *inteligência* neste esforço, mesmo o mais “mecânico” dos comportamentos descritíveis como um caso de *seguimento de regras* consiste em um comportamento *inteligente*, por presumivelmente não ser explicável sem recurso às noções de *compreensão* e *ação* mencionadas na primeira das caricaturas apresentadas. Por exemplo, mesmo o fenômeno da obediência “cega” a instruções puramente sintáticas envolve a presunção de que aquilo que obedece à instrução, em primeiro lugar, *compreende*; e, em segundo, *age* de modo a cumpri-la. Essa obediência, desse modo, não é tão “cega” quanto o comportamento do funil, das engrenagens de um relógio ou do movimento de corpos celestes – que, muito embora possam ser descritos como ocorrendo *em conformidade* com certas regras, presumivelmente não ocorrem como o resultado do *cumprimento ativo* dessas regras por esses sistemas físicos. A máquina teórica de Turing é inteligente, portanto, mas *apenas*

nessa medida, a saber – na medida em que é definida como capaz desse exato tipo de comportamento inteligente que se resolve na obediência “cega” a instruções puramente sintáticas. Uma máquina teórica de Turing, portanto, não é mais capaz de atividades inteligentes, digamos, impressionantes – como jogar xadrez, reconhecer padrões de melodia e harmonia musical ou produzir um texto em português – do que alguém sem qualquer competência em artes gráficas é capaz de desenhar um retrato fiel de uma pessoa conhecida apenas ligando pontos com um lápis segundo as instruções de um esquema de “*draw-by-numbers*”. Em outras palavras: o resultado “extensional” da atividade de uma máquina de Turing – isto é, o processo ou o objeto que resulta de seu comportamento; ou ainda, seu comportamento como *efeito* – pode ser, no limite, *exatamente indistinguível* de uma partida brilhante de xadrez, de um perfeito ditado musical ou de um excelente ensaio jornalístico; ainda assim, essa máquina nunca terá movido um único peão, ouvido um único acorde de sol maior, nem escrito uma única palavra – ao menos não na medida em que se entende por essas atividades algo que é, em algum sentido, irredutível à obediência “cega” a regras que determinam a realização de tarefas sintáticas. Em geral, parece cabível advogar uma tal irredutibilidade quando se considera esses comportamentos ou atividades como *atos* – no mesmo espírito em que o *ato* de desenhar um retrato não se deixaria reduzir ao *ato* de seguir um esquema de “*draw-by-numbers*”.

O caso dos computadores eletrônicos, por sua vez, é distinto: eles se enquadram, como já indicado, na mesma categoria que os funis, engrenagens e corpos celestes, e não dispõem nem da porção extremamente limitada de inteligência que define sua contraparte teórica. É de se notar, contudo, que o comportamento de um computador eletrônico *como efeito* pode ser, no limite, *exatamente indistinguível* da atividade de uma máquina teórica de Turing – isto é, *exatamente indistinguível* da obediência “cega” a instruções que determinam a realização de tarefas puramente sintáticas. Se se considera a famosa distinção da filosofia moral kantiana entre agir *em mera conformidade* com um imperativo e agir *em virtude* de um imperativo, algumas analogias são cabíveis. Ao interpretar uma canção, por exemplo, uma inteligência paradigmática, em princípio, emite notas em virtude de padrões normativos ditados pela estrutura rítmica, pela linha melódica e pelo campo harmônico da canção interpretada; uma máquina de Turing, por sua vez, emite notas em mera conformidade com esses padrões normativos, mas em virtude de instruções puramente sintáticas de cujo cumprimento *como efeito* o cumprimento desses

padrões normativos *como feito* é indistinguível; e um computador eletrônico, finalmente, emite notas em mera conformidade com as instruções puramente sintáticas ativamente cumpridas pela máquina de Turing, porém *jamais* em virtude de o que quer que seja – pois a emissão de notas por um computador eletrônico é apenas *feito* de estímulos físicos, sem ter como contraparte qualquer *ato* do computador eletrônico.

Assim, portanto, por transitividade, fica evidente que, no limite, o comportamento do computador eletrônico pode ser exatamente indistinguível de uma partida brilhante de xadrez, de um ditado musical perfeitamente tomado ou de um ensaio jornalístico competente – tomados como *feitos*. Daí fenômenos como o Deep Blue, o ChatGPT e afins, e o ensejo para retornarmos à lida direta com nossa pergunta norteadora.

5. Resposta nº1: a negativa – outra caricatura

Sobre o pano de fundo das caricaturas e reflexões acima, iniciaremos por delinear o tipo mais óbvio de resposta à nossa questão inicial. Trata-se de um inequívoco e retumbante *não*: computadores não poderiam, por questão de princípio, apresentar inteligência em nenhum sentido suficientemente interessante do termo. A justificativa se deixaria depreender de uma leitura aparentemente literal das observações da seção anterior: uma máquina teórica de Turing não passaria, em primeiro lugar, de uma ficção útil; ademais, mesmo que concretizada literalmente em realidade, tal máquina, por definição, jamais poderia realizar tarefas para além daquelas de um tipo que o próprio Turing em 1950 chamou de “desprovida de inteligência” (*unintelligent*) – o que mostraria sua impossibilidade de apresentar, se não inteligência *tout court*, ao menos qualquer forma de inteligência suficientemente interessante para conferir algo além da capacidade de obedecer cegamente a regras puramente sintáticas. Quanto a um computador eletrônico, por sua vez, este não passaria de um sistema físico sofisticado, porém tão inteligente quanto às peças de silício que o compõem, cujo funcionamento *como feito* é, no limite, exatamente análogo ao de sua contraparte teórica. Algumas das manifestações mais populares de uma posição a respeito da nossa pergunta diretriz suficientemente afins a esses contornos na literatura filosófica contemporânea se devem a John Searle; em particular, vale mencionar a proposta do (controverso)

assim chamado “argumento do quarto chinês”¹¹. No cerne, o tipo de posição popularizada por Searle já se deixava exemplificar desde cerca de três séculos antes, no §17 da *Monadologia* de Leibniz:

“Aliás, deve-se confessar que a percepção e o que dela depende é inexplicável por razões mecânicas, isto é, por figuras e movimentos. Pois, imaginando haver uma máquina, cuja estrutura faça pensar, sentir e perceber, poder-se-á concebê-la proporcionalmente ampliada de modo a poder-se entrar nela como num moinho. Admitido isso, ao visitá-la por dentro não se encontrarão lá senão peças impulsionando-se umas às outras, e nada que explique uma percepção. Portanto, essa explicação só deve ser procurada na substância simples, e não no composto ou na máquina.”¹²

A despeito do cabimento exegético de nossa atitude com respeito ao texto de Leibniz em particular – que é particularmente peculiar quanto à sua caracterização do domínio mental e do que a ele concerne –, tomamos como estruturante desse tipo de linha de raciocínio a pressuposição de que pensar em geral – e portanto, em particular, atividades inteligentes – dependeria da ocorrência de respectivas “representações internas” no pensante. A ocorrência destas representações é, na maioria das vezes, associada de modo intrínseco a um caráter *consciente*¹³; de tal modo que manifestações de *inteligência* implicariam a ocorrência de respectivas manifestações de *consciência*, fazendo de todo ser inteligente algo ao fim e ao cabo conforme aos contornos de nossa caricatura de um *cogitans* na seção 2 – e, portanto, um caso de inteligência paradigmática. Àquilo a que não se atribui certo tipo de estado de consciência – o que parece incluir todo e qualquer autômato real –, portanto, não faria nem sequer sentido atribuir inteligência, de qualquer grau ou natureza.

11 Em Searle, J. Minds, Brains and Programs, *Behavioral and Brain Sciences*, 3: 417–57, 1980 (cf. Searle, J. The Chinese Room, In: R.A. Wilson and F. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA: MIT Press, 1999; e Searle, J. Why Dualism (and Materialism) Fail to Account for Consciousness, In: Richard E. Lee (ed.), *Questioning Nineteenth Century Assumptions about Knowledge* (III: Dualism), New York: SUNY Press, 2010.)

12 Leibniz, G. *Monadologia*. São Paulo: Abril Cultural, p. 61-73, 1974 (p. 64).

13 Ainda que, notoriamente, não em Leibniz. De fato, julgamos que a desvinculação entre as noções de Leibniz de *percepção* e de *consciência* pode, malgrado as posições do filósofo a respeito, inspirar um ferramental conceitual particularmente interessante para a lida com demandas históricas e interpretativas acarretadas pela noção de inteligência artificial.

Como, então, os partidários desse tipo de posição explicam fenômenos como o Deep Blue, ChatGPT e afins? Pura e simplesmente assim: se trata, na melhor das hipóteses, de mera *simulação* de comportamentos inteligentes. Por mais perfeita que seja a conformidade da operação de um computador eletrônico às regras sintáticas que regem o funcionamento de uma máquina de Turing cuja atividade é indistinguível de uma partida de xadrez brilhante, nem o computador eletrônico é um operador sintático como o é a máquina de Turing, nem esta é uma brilhante jogadora de xadrez; Antes, o funcionamento de um computador eletrônico – uma máquina física – tão somente *simularia* a atividade de um operador sintático – a máquina de Turing – que, por sua vez, tão somente *simularia* uma brilhante partida de xadrez. Destarte, mesmo que um computador físico produzisse uma prova da conjectura de Goldbach ou indicasse a solução definitiva para o aquecimento global, esses seriam, na melhor das hipóteses, meros *simulacros* de comportamentos inteligentes.¹⁴

6. Resposta nº 2: Teste de Turing, Wittgenstein e critérios externos

Nosso segundo tipo de resposta parte da atenção à importância de um dos termos em que nossa pergunta norteadora é formulada e que não foi contemplado nas considerações da seção 4. Em “pode um computador apresentar inteligência?”, tanto quanto às noções de computador e de inteligência, é crucial atentar-se ao papel desempenhado pela noção de *apresentar* envolvida na formulação. A obviedade do tipo de perspectiva delineado na seção anterior pode ser atribuída exatamente a uma ingenuidade ou precipitação na lida com o sentido dessa noção na interpretação da pergunta em questão. Quando se pergunta se um computador pode *apresentar* inteligência, não fica imediatamente e sem mais decidido se se trata de uma pergunta sobre se *é o caso* o possível fato que seria um computador ser dotado de inteligência; ou se se trata da possível *verificabilidade* de uma eventual atribuição de inteligência a um computador. Em outros termos, enquanto a primeira alternativa equivaleria a perguntar algo como “pode haver tal coisa como um computador dotado de inteligência?”, a segunda seria equivalente a “é possível que

14 Palmeiro sugere (em Palmeiro, T. *Informática: representações sem exterioridade*. Manuscrito não publicado, 2023) que a computação consistiria em uma espécie de caso limite de instância da polêmica noção platônica de *mimesis*, a saber, o de, em seus termos, “representações sem exterioridade”, que estariam efetiva e paulatinamente usurpando daquilo de que seriam simulações ou simulacros as suas funções epistêmicas e ontológicas mais características, como as de fundamento do conhecimento e componentes da realidade.

um computador atenda a um critério suficiente para que se lhe atribua inteligência justificadamente?”. A abordagem do tipo de resposta apresentado na seção 5 privilegia de forma peremptória o primeiro tipo de interpretação da pergunta; o que relega a tarefa de respondê-la, em última análise, ao estatuto da simples adução de consequências mais ou menos triviais das definições que se aceite para as noções de inteligência e de computador. A abordagem de que trataremos nesta seção, contudo, parte do segundo tipo de leitura da pergunta – o que põe diante de nós duas tarefas: (a) encontrar critérios que permitam a atribuição justificada de inteligência; e (b) verificar se um computador pode atender a algum desses critérios.

Procedamos então à primeira delas: que tipo de critério permitiria a atribuição justificada de inteligência? A lida com esse tipo de questão pode ocorrer por, pelo menos, duas vias: uma seria (a.1) uma tentativa de encontrar fenômenos verificáveis que fossem condições suficientes para que o “fato” da inteligência – como habitualmente se o define, em consonância com o discutido até aqui – fosse o caso. Outra seria (a.2) recorrer a critérios *efetivamente* utilizados para atribuição justificada de inteligência a comportamentos certos comportamentos. Após alguma consideração da passagem do moinho no §17 da *Monadologia*, deve-se ao menos suspeitar que talvez, por extensão, não haja *em princípio* como verificar a ocorrência de representações internas pertinentes a um ser a partir de quaisquer dos movimentos observáveis em um mecanismo físico que o compõe. Desse modo, a via (a.1) se torna fadada ao fracasso, uma vez que o “fato” da inteligência parece indissociável, em suas acepções mais ortodoxas, como indicam as caricaturas, da ocorrência desses fenômenos “internos” e, portanto, presumivelmente inverificáveis por si próprios. Afinal, como sugere Wittgenstein no §580 das *Investigações*: “Um ‘processo interior’ precisa de critérios exteriores.” O emprego das aspas em “processo interior” é um requinte oportuno dessa passagem. | Ele dá margem a uma sua interpretação como sugestivo do exato tipo de mudança de perspectiva que nos permite compreender uma transição da atitude que marca (a.1) para aquela que empresta sentido a (a.2). Os supostos “processos interiores” aos quais se encontra atrelada a nossa noção de inteligência demandariam que, a fim de verificar a ocorrência de inteligência – e supostamente, *a fortiori*, deles próprios –, nossa atenção se direcione para critérios definidos em termos de fenômenos “exteriores” (que é o que efetivamente *fazemos* ao atribuir inteligência). Constatações da ocorrência destes últimos, contudo, no espírito da passagem do moinho, já são tomadas de antemão como incapazes

de oferecer condições suficientes para a constatação de quaisquer “processos interiores”, e em particular daqueles de que dependeria a inteligência. Isso talvez indique não tanto estarmos em um beco sem saída quanto, antes, sugira que os “processos” que supomos “interiores” e cujas verificações buscamos no moinho de Leibniz sejam, de fato, mais afins aos moinhos quixotescos. A cisão *interior vs. exterior* perde sua razão de ser no contexto da discussão do fenômeno da inteligência, que então – como qualquer outro fenômeno da vida – ganha razões para ser identificado e compreendido tão somente a partir de como se o constata na vida.

A via (a.2) se nutre de motivação a partir desse tipo de reflexão, e dá ensejo ao retorno a uma observação que fizemos na seção 2: *identificamos* casos de inteligência na ocorrência de diversos tipos de situação na vida, como conversas de bar entre amigos, etc. Note-se que esta perspectiva na abordagem da questão da inteligência artificial não é, ao contrário do que se pode apressadamente supor, uma marca da contemporaneidade. Descartes oferece suporte a esta alegação na seguinte passagem da quinta parte do *Discurso do Método*:

“... se houvesse máquinas assim, que tivessem os órgãos e a figura de um macaco, ou de qualquer outro animal sem razão, não disporíamos de nenhum meio para reconhecer que elas não seriam em tudo da mesma natureza que esses animais; ao passo que, se houvesse outras que apresentassem semelhança com os nossos corpos e imitassem tanto nossas ações quanto moralmente fosse possível, teríamos sempre dois meios muito seguros para reconhecer que nem por isso seriam verdadeiros homens. Desses, o primeiro é que nunca poderiam usar palavras, nem outros sinais, compondo-os, como fazemos para declarar aos outros os nossos pensamentos. Pois pode-se muito bem conceber que uma máquina seja feita de tal modo que profira palavras, e até que profira algumas a propósito das ações corporais que causem qualquer mudança em seus órgãos: por exemplo, se a tocam num ponto, que pergunte o que se lhe quer dizer; se em outro, que grite que lhe fazem mal, e coisas semelhantes; mas não que ela as arranje diversamente, para responder ao sentido de tudo quanto se disser na sua presença, assim como podem fazer os homens mais embrutecidos. E o segundo é que, embora fizessem muitas coisas tão bem, ou talvez melhor do que qualquer de nós, falhariam infalivelmente em algumas outras, pelas quais se descobriria que não agem pelo conhecimento, mas somente pela disposição de seus órgãos. Pois, ao passo que a razão é um instrumento universal, que pode servir em todas as espécies de circunstâncias, tais órgãos necessitam de alguma disposição

*particular para cada ação particular; daí resulta que é moralmente impossível que numa máquina existam bastante diversas para fazê-la agir em todas as ocorrências da vida, tal como a nossa razão nos faz agir.*¹⁵

Deve se reconhecer que a esse argumento convêm os seguintes traços: (1) ele apresenta e defende uma posição bem definida e contrária à hipótese de algo como a ocorrência de uma inteligência artificial. (2) sua concepção dos termos que permitiriam lermos essa passagem como dizendo respeito à noção de inteligência recorre a “processos interiores”, como a “razão”. (3) as conclusões aduzidas dependem de constatações algo proféticas a respeito da distinguibilidade entre os comportamentos de autômatos e seres humanos. Não obstante, é de se notar que o raciocínio apresentado se estrutura sobre o que Descartes chama de “dois meios muito seguros para reconhecer” que nem mesmo os mais perfeitos simulacros possíveis de comportamentos humanos se converteriam ou se deixariam confundir com “verdadeiros homens”. Esses meios consistem, no cerne, em maneiras de *verificar*, pela *observação* de seu *comportamento*, a presença de uma suposta incapacidade de autômatos de adaptarem suas competências simulatórias de comportamentos inteligentes às variações de situações e contextos que se lhes podem apresentar nos âmbitos da *expressão verbal falada* e da *performance de tarefas práticas*. Em uma frase: máquinas se revelariam desprovidas de inteligência porque não conseguiriam, ao fim e ao cabo, simular a adaptabilidade supostamente *universal* das competências de seres “autenticamente” inteligentes a respectivos contextos e situações. Isso se daria porque esta adaptabilidade não seria possível senão como um exercício da razão.

Diversas observações seriam cabíveis, concernentes desde ao tipo de noção de inteligência artificial pretensamente refutada por Descartes e sua relação com a que temos aqui discutido, passando pelos possíveis efeitos da consideração de máquinas *universais* de Turing (teóricas ou físicas) sobre o argumento, até a estrutura possivelmente circular – ou de outro modo falaciosa – do argumento do filósofo. Declinamos, contudo, de um diálogo direto com a posição cartesiana em favor de usá-la como uma sublinha histórica, por assim dizer, do tipo de trilha da via (a.2) exemplificado na atitude do próprio Turing. Esta se epitomiza no assim chamado “teste de Turing”, proposto inicialmente¹⁶ sob a forma do que o próprio denominou “o jogo da imitação”.

15 Descartes, R. *Discurso do método*. São Paulo: Abril Cultural, 1973, p. 33-80 (p. 68-69)

16 Em Turing, A., *Computing Machinery and Intelligence*, *Mind*, 59 (236): 433-60, 1950.

Ele pode ser descrito como se segue: suponha que tenhamos os interrogados – uma pessoa e uma máquina – e um interrogador – uma outra pessoa. O interrogador está em um cômodo separado daquele em que estariam a outra pessoa e a máquina. O interrogador associa a um dos interrogados o rótulo “X” e ao outro o rótulo “Y”; contudo, pelo menos no início do jogo, não sabe se X é a pessoa e Y é a máquina ou o oposto. O objetivo do interrogador no jogo é precisamente determinar qual dos interrogados é a pessoa e qual é a máquina, dizendo, ao final do jogo, ou bem “X é a pessoa e Y é a máquina”, ou bem “X é a máquina e Y é a pessoa”. O interrogador tem permissão para fazer perguntas do seguinte tipo: “Será que X poderia me dizer se X joga xadrez?” Seja X a máquina ou a outra pessoa, X deve responder às perguntas dirigidas a X. O objetivo da máquina, por sua vez, é tentar fazer com que o interrogador conclua erroneamente que a máquina é a outra pessoa; já o objetivo da outra pessoa é tentar ajudar o interrogador a identificar corretamente a máquina.¹⁷

Variantes do teste de Turing assim descrito abundam na literatura¹⁸, a maioria delas consistindo na proposta de um tipo de jogo em que uma pessoa interage verbalmente com um suposto interlocutor que ela não tem meios para determinar de antemão se é um ser humano ou um computador – tarefa que essa pessoa tem por objetivo cumprir corretamente com base exclusivamente na interação verbal com o suposto interlocutor, e na qual o computador, caso o haja, tem o objetivo de levá-la a falhar. A moral do teste seria a seguinte: se o computador logra passar-se por uma pessoa, não haveria razões para, ao menos no contexto em questão, negar-lhe a capacidade de comportamento inteligente.

Note-se que, em primeiro lugar, o âmbito escolhido por Turing como possibilitando a verificação de condições suficientes para a atribuição ou negação justificada de comportamento inteligente a uma máquina é um dos já aventados por Descartes cerca de trezentos anos antes, a saber: o da interação verbal. O próprio Turing, contudo, impõe limites claramente mais estreitos ao formato dessa interação em seu jogo, que são, em todo caso, mesmo em variantes menos estritas que aparecem na literatura, dados por condições em algum sentido mais específicas que o tipo de interação livre descrito por Descartes no *Discurso do Método*. Em segundo lugar, Turing não adota uma posição contrária à possibilidade de uma máquina sair vencedora de seu jogo – e assim,

17 Cf. Oppy, G. e Dowe, D. The Turing Test, *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.) Disponível em plato.stanford.edu/archives/win2021/entries/turing-test/ 2021. Acesso em 10 mai. 2024.

18 Cf. Idem e referências ali listadas.

presumivelmente, identificada como capaz de interação inteligente. De fato, exatamente pelo contrário, o pensador se mostrou não apenas aberto à possibilidade de que uma máquina passasse em seu teste – o que, por si só, consistiria em uma resposta positiva à questão (b) acima –, como também confiante na concretização factual dessa alternativa em um futuro que já é passado:

“I believe that in about fifty years’ time it will be possible to programme computers (...) to make them play the imitation game so well that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning. The original question, ‘Can machines think?’ I believe to be too meaningless to deserve discussion. Nevertheless, I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.”¹⁹

Pondo de lado possíveis avaliações das competências proféticas de Turing, é de se notar como a aceitação e o emprego genuínos de um critério do tipo (a.2) sujeitam a própria noção discutida de inteligência a alterações em sua semântica que, em última análise, podem relegar à obsolescência interpretações de nossa questão norteadora que tornem a adequação de possíveis respostas independentemente da disponibilidade de métodos para sua verificação efetiva. De fato, o próprio Turing parece compartilhar desse parecer quando diz, sobre a questão “Podem as máquinas pensar?”, norteadora de seu próprio artigo de 1950 e bastante semelhante à nossa, que ela seria “demasiado sem sentido” para merecer discussão. Em vez de explorar a ambiguidade do termo “apresentar”, presente em nossa formulação, Turing dá sua concessão a uma leitura ingênua da questão e a esvazia de importância, deslocando sua atenção para uma pergunta diferente, à qual ele dá corpo por meio do hoje célebre teste.

Podemos, ademais, para nossos propósitos, considerar uma formulação mais lata da linha de raciocínio envolvida no tipo de teste por ele proposto: se um suposto agente X se comporta de forma indistinguível de um agente Y ao qual atribuíamos inteligência dado certo comportamento em certo contexto, então não há razões para não atribuir inteligência a X nesse contexto. Deve-se atentar para o fato de que essa formulação extrapola os limites do problema específico da inteligência *artificial*, mostrando-se relevante para o todo da questão da inteligência *não-paradigmática*. Se, a exemplo de Turing, se

19 Turing, A. Computing Machinery and Intelligence, *Mind*, 59(236): 433–460, 1950 (p.442).

adota uma atitude aberta à possibilidade em princípio de sucesso de qualquer suposto agente nesse tipo de teste, consolida-se uma *interdição* da negação por princípio de inteligência a qualquer tipo de ser, em qualquer contexto especificado, por *mera* força de eventuais traços de sua *ontologia*. Desse modo, não é possível dissociar o caráter do ser de suas performances – ser P é performar de modo P; e ser é, em geral, caráter performático. Não se trata de uma redução em estilo *behaviourista* – de que comumente se acusa Turing e Wittgenstein – daquele que performa à sua performance efetiva, mas, antes, de um bloqueio à hipóstase de estruturas ontológicas que possibilitem a determinação do ser de algo que performa *a despeito* de suas performances, efetivas ou possíveis.

As reverberações dessa inversão metodológica na consideração do fenômeno da inteligência são, ao menos potencialmente, significativas e positivas. Em primeiro lugar, é de se notar que nada se perde: isto é, não perdemos razões para considerarmos inteligentes *em virtude de suas performances* seres que já o considerávamos em virtude de sua mera constituição ontológica – salvo, é claro, se esta já fracassasse em garantir a esse ser um caráter performático inteligente. Em segundo, ademais, e provavelmente mais importante, é o que se ganha: a janela aberta por pensadores como Wittgenstein e Turing para a possível caducidade da distinção entre natureza e artifício no domínio da inteligência permite que divisemos, algumas poucas léguas mais rumo ao horizonte, a possível caducidade da distinção entre, de modo mais geral, inteligências paradigmáticas e não-paradigmáticas. Dado o papel central exercido pela noção de inteligência na definição do que entendemos por *ser humano* – uma noção manipulada de forma danosamente insuficiente na tradição e no cotidiano –, talvez se possa ver nessa reflexão uma boa razão para não apenas olhar, mas efetivamente *pular* através dessa janela aberta para *aprender* a reconhecer e abraçar a inteligência e a humanidade de muito mais seres do que autômatos de plástico e silício.

7. Observações preliminares à próxima resposta: Gadamer e as muralhas ontológicas

A implosão da “muralha ontológica” (expressão particularmente oportuna que tomamos emprestada de de Souza²⁰) entre *performance* e *ser* ensejada pelo pensamento de Turing e Wittgenstein serve como ponto de articulação

20 de Souza, C. S. *Reflexões sobre Semiótica, Linguagem, Ontologia e Computação*. EMAPS-Notas #02. Rio de Janeiro, RJ - Brasil: SERG, Departamento de Informática, PUC-Rio, 2023.

com a próxima parte desta reflexão. Nela, se evidencia uma marcante afinidade com pensamentos expressos na literatura das tradições da fenomenologia e da hermenêutica, particularmente na *Hermenêutica Filosófica* de Gadamer, em que os fenômenos da *compreensão* e da *interpretação*, em jogo nas discussões dos hermeneutas ao menos desde Schleiermacher, são considerados sob a influência de algumas lições-chave da fenomenologia de Husserl e de Heidegger – particularmente, a partir da implosão da muralha ontológica entre *sujeito* e *objeto* – que, no âmbito de uma *conversa* (*Gespräch*), tem contraparte na implosão das muralhas análogas entre *autor* e *texto*; *texto* e *intérprete*; e, *a fortiori*, *autor* e *intérprete*. Uma breve reconstrução dessas implosões pode ser útil à inteligibilidade da sequência.

Dilthey pode ser entendido como um dos principais responsáveis por colocar em evidência o *psicologismo* presente na hermenêutica de Schleiermacher. Este psicologismo manifestar-se-ia na tentativa de Schleiermacher de formular métodos para que fosse possível acessar as “intenções” ou o “significado que o autor pretendeu” para determinar o significado verdadeiro de um texto. A partir da ênfase de Dilthey sobre a questão do papel hermenêutico das intenções, Gadamer argumenta que não há como acessar as intenções do autor de um texto, na medida em que essas intenções são entendidas como pertencentes a uma categoria distinta e disjunta dos elementos que compõem o sentido já expresso *textualmente*. A justificativa para essa posição se dá por definição: se se trata de fato de *intenções*, e não de elementos textuais, então se trata de conteúdos de índole intrinsecamente *privada*, e que não podem, portanto, ser justificadamente apreendidos do que quer que tenha sido compartilhado por meio de uma manifestação numa conversa.

Há autoras²¹ que consideram que a rejeição de Gadamer da premissa da acessibilidade das intenções faz de sua proposta uma involução quando comparada à perspectiva hermenêutica de alguns de seus predecessores, como Schleiermacher, por exemplo, que buscavam acessar (psicologicamente) “o que o autor quis dizer”. Em outras palavras, esses hermeneutas não apenas pressupunham de alguma forma a acessibilidade da intenção do autor com seu texto, como também privilegiavam o papel dessa intenção na determinação dos significados relevantes para a leitura e a interpretação do mesmo. Desse modo, para Marie Fleming, por exemplo, não privilegiar o que o autor “quis dizer” é

21 Exemplos: Fleming, M. *Gadamer's Conversation Does the Other Have a Say?*. In: Code, L. *Feminist Interpretations of Gadamer*. University Park: Pennsylvania State University Press, 2003; e Fiumara, G. *The Development of Hermeneutic Prospects*. In: Code, L. *Feminist Interpretations of Gadamer*. University Park: Pennsylvania State University Press, 2003.

escolher não lidar com a subjetividade do Outro (do autor), ou, no mínimo, colocar essa subjetividade como algo secundário no processo de compreensão de textos em geral e, em particular, no contexto de conversas, impedindo assim a possibilidade da interação efetiva entre os sujeitos envolvidos na conversa.

Pelas razões mencionadas acima, contudo, essa colocação é improcedente enquanto crítica à concepção Gadameriana de *conversa*, uma vez que depende da rejeição injustificada de alguns de seus pontos fundamentais de partida, cuja elucidação é um dos principais desenvolvimentos da filosofia do alemão. Entre esses, destaca-se o seguinte: a separação por princípio de duas categorias ontológicas distintas e disjuntas, a saber, elementos textuais, de um lado, e intenções do autor, de outro (em outras palavras, entre o que se *disse* e o que se *quis dizer*) apenas se deixa justificar no pensamento de Gadamer a partir da premissa de que à segunda categoria, i.e. às intenções, adere por definição uma índole inalienavelmente *privada*; i.e. não compartilhável e, portanto, se acessível, restrita ao próprio autor. Desse modo, Gadamer relega essa “muralha” ontológica, no âmbito do fenômeno da *compreensão*, a uma obsolescência semelhante àquela que defendemos manifestar-se no pensamento de Turing e Wittgenstein entre *ser* e *performance*: não se trata, assim, de reduzir o ser do autor aos textos que efetivamente produziu, mas, antes, de bloquear a hipótese de certas estruturas ontológicas – nesse caso, a de um autor como um feixe de “intenções” ou outras entidades “internas” – que condicionariam o significado de um texto à revelia do próprio²². Nesse sentido, a questão sobre o condicionamento de um texto por sua autoria perde o sentido: na medida em que se trata de um conteúdo passível de *compreensão*, um texto jamais se deixa confundir com um *produto* – é, portanto, tão incompatível com um eventual autor produtor quanto seria um som com uma cor.

Em um movimento argumentativo análogo, a muralha entre *texto* e *intérprete* rui: se a interpretação de um texto fosse, em algum sentido incontornável, composta por *intenções* ou quaisquer outros tipos de “entidade” psicológica pertinentes ao intérprete e ontologicamente distintas e disjuntas

22 Note-se aqui nova semelhança com uma passagem do *Investigações* de Wittgenstein, desta vez, o §583: “Mas você fala como se eu verdadeiramente não tivesse expectativa, não tivesse esperança agora – no entanto, creio ter esperança. Como se o que acontece *agora* não tivesse profunda significação.” O que significa: “O que acontece agora tem significação” ou “tem profunda significação”? O que é uma sensação *profunda*? Alguém poderia, durante um segundo, sentir amor profundo ou profunda esperança, *não importando o que* precede e o que segue esse segundo? – O que aconteceu agora tem significação neste meio (*Umgebung*). O meio lhe dá a importância. E as palavras “ter esperança” referem-se a um fenômeno da vida humana. (Uma boca sorridente sorri apenas num rosto humano.)”

de sentidos textuais, não seria possível que uma interpretação fosse tornada efetivamente acessível a ninguém senão ao próprio intérprete – função que colapsaria com a de um “autor produtor” da interpretação, por assim dizer. Na medida em que consiste em um movimento *conversacional*, portanto, uma interpretação depende, antes, do que Gadamer chama de *preconceitos* ou *pré-juízos* (*Vorurteil*) – elementos de índole também textual e compartilhável, que compõem o arcabouço hermenêutico segundo os quais se compreende o interpretado. Como, no âmbito do fenômeno da *compreensão*, *sujeito* e *objeto* se manifestam enquanto, de um lado, *autor* ou *intérprete*, e, do outro, *texto*, então, por força, implode também neste âmbito qualquer muralha ontológica entre essas duas categorias. Assim, no âmbito do fenômeno da *compreensão* – que é, no pensamento hermenêutico de Gadamer, pensado de modo lato como o mesmo que o do próprio *ser* –, *só há o texto já e sempre interpretado de um modo específico*; a pergunta a respeito dos eventuais papéis de *autor* e de *intérprete* na determinação dos significados do texto é sempre obsoleta no mesmo sentido em que essas categorias, na medida em que se referem a algo de semanticamente relevante, são elas próprias ontologicamente obsoletas.

8. Resposta nº3: humanidade e computador como textos e algumas considerações finais

Observados os pontos relativos à implosão das muralhas ontológicas, propomos agora um último retorno à consideração de nossa pergunta norteadora; desta feita, contudo, a partir de uma abordagem que radicaliza, em certo sentido, a proposta expressa em de Souza 2023 de pensar a computação e, por extensão, os computadores em chave *textual*.

Até aqui, quando nos perguntamos “pode um computador apresentar inteligência?”, temos sempre considerado o computador em apenas duas das posições da tríade funcional *autor* – *texto* – *intérprete*; a saber, a primeira e a última. De fato, isso parece, inicialmente, não mais do que atender a uma condição de sentido imposta de saída pela própria pergunta: apenas a autor e intérprete seria cabível atribuir inteligência, textos servindo, no máximo, como evidências indicativas – índices, na semiótica de Peirce – de que essas atribuições seriam ou não apropriadas. A necessidade dessa restrição inicial, contudo, não parece resistir à derrubada das muralhas ontológicas precipitada pelas reflexões expressas em Turing, Wittgenstein e Gadamer: com efeito, o texto não é mais relegado ao estatuto de um *objeto* que funciona como índice *externo* algo débil de propriedades *internas* de *sujeitos* que os *produzem*

ou interpretam, mas, antes, passa a funcionar como o eixo a partir do qual e o âmbito no qual se pode determinar haver tais fenômenos particulares como “autores” ou “intérpretes” em jogo na dinâmica de um fenômeno – e estes tipos de ente particulares são eles próprios esvaziados de qualquer autonomia ontológica, resolvendo-se sempre em sentidos textuais.

Esse tipo de abordagem do problema da textualidade convida a uma nova lida com a questão norteadora “pode um computador apresentar inteligência?”. Da maneira recém-sugerida, um computador se manifesta, em primeiro lugar e de saída, no acontecimento de um *texto*. Em ao menos um sentido, portanto, a pergunta sobre se um computador pode manifestar inteligência se dissolve: simplesmente não há sentido em atribuições de inteligência a elementos de um texto – ao menos não no sentido discutido até aqui. Atribui-se inteligência no sentido de nossa questão, antes, a *autores* de textos – na medida em que lhe cabe uma autoria, contudo, um texto é, inescapavelmente, pensado como *efeito* ou *produto*, o que difere do sentido de *texto* para o qual Gadamer chama a atenção. Essa índole textual, contudo, se estende também a seres humanos e, em geral, a todos os tipos de entes que se pode encontrar na vida. As atribuições de inteligência a autorias ficam, nesse sentido, interditas em geral – não porque factualmente falsas, mas, antes, pela incompatibilidade em princípio entre a textualidade manifesta no fenômeno da compreensão e qualquer tipo de natureza de efeito ou produto, condição sem a qual não há sentido em se falar em autoria.

De fato, nesse sentido de *texto*, lidamos com textos variados, e um conceito de *humanidade* se estrutura e se mostra nessa lida. Novos textos e textualidades correspondentes têm se manifestado segundo perspectivas componentes do feixe de preconceitos mais ou menos coeso a que se se pode referir como o cânone da tradição ocidental desde que a há, e transformado, de acordo com suas próprias características, os contornos desse conceito de humanidade que vivenciamos. Computadores – antes de nos enredarmos em perguntas que nos limitam de saída a compreender essa noção como uma designação funcional possivelmente análoga à de *autor* ou *intérprete* – têm sido, no mínimo desde a consolidação de algo como a caricatura da seção 3 num âmbito conversacional socialmente relevante, precisamente um desses novos tipos de expressão textual; assim como diferentes comunidades, espécies de animais, livros, pedras, florestas equatoriais, motores a vapor, etc. em geral, o são.

Assim, a pergunta “pode um computador apresentar inteligência?” é sempre carente de sentido se não emerge de uma situação textual determinada em todas as suas condições particulares, a partir das quais respostas são, ao

menos em alguns casos, possíveis. Um “computador” como texto – assim como um ser humano como texto – não é uma categoria ontológica *a priori*, mas um acontecimento interpretativo sempre contingente e particular; assim, a pergunta não consiste em um questionamento imposto previamente a respeito de uma noção supostamente disponível de “computadores” em geral, mas, antes, se apresenta e permite com ela uma lida apenas nos contextos específicos que a fazem emergir como uma pergunta *genuína* – o que Gadamer chamaria de uma pergunta *radical* –, e não como um mero dispositivo auxiliar à confirmação de eventuais preconceitos prévios. Esses contextos são aqueles em que *os sentidos das próprias noções envolvidas na pergunta estão em jogo* – condição sem a qual uma resposta, qualquer que seja, não passaria de uma dedução de consequências de definições adotadas previamente em mera virtude de preconceitos em vigor, nos moldes da resposta negativa da seção 3. Em outros termos: quem pergunta “pode um computador apresentar inteligência?” sem estar nisso em jogo o que entende por computador, por inteligência e/ou pela relação entre essas noções, não se pergunta efetivamente coisa alguma: apenas oferece a si o ensejo de *calcular*, a partir de juízos e normas previamente adotadas, a atribuição de valor afirmativo ou negativo em que consistirá sua “resposta”. Se alguém efetivamente se pergunta tal coisa, portanto, fá-lo a partir de circunstâncias históricas que tiram o trânsito conceitual no campo semântico dos computadores ou da inteligência dos seus trilhos, por assim dizer – esse interrogador genuíno, portanto, genuinamente *carece* de compreensão de algo do que significa aquilo pelo que pergunta.

A pergunta então se volta, uma vez mais, para os casos materiais que a trouxeram à tona em contextos históricos particulares. Fenômenos como o Deep Blue ou o ChatGPT colocaram ou têm colocado em jogo o significado de computador, de inteligência ou da relação entre essas noções? Se sim, como e até que ponto? Que interesses econômicos, políticos e sociais têm motivado respostas afirmativas ou negativas à pergunta acerca da apresentação de inteligência por esses dispositivos computacionais em particular? E quanto a casos fictícios, de HAL-9000 ao robô David, protagonista do filme de 2001 *I.A. - Inteligência Artificial*, de Spielberg – que tipo de questionamento ensejam, ou ainda, em que tipo de questionamento consistem?

A inteligência apresentada pelo Deep Blue talvez seja creditável a seus desenvolvedores e programadores, além de a enxadristas cujas partidas serviram de tijolos de construção do seu programa – e certamente não lhe permitiu feitos relativamente banais esperados de seres humanos. Já a inteligência insinuada no ChatGPT, talvez, seja aquela das inúmeras pessoas

envolvidas no desenvolvimento das tecnologias teóricas e práticas ligadas aos LLMs (*Large Language Models*), e também daquela manifesta na quantidade enorme de dados textuais que o “alimentam” – além, vale mencionar, da de vendedores e publicitários empenhados em alimentar fantasias “utópicas” e naturalizar fantasias distópicas a respeito da própria noção de humanidade. HAL-9000 aterrorizou e aterroriza, e David despertou e desperta a compaixão de enormes audiências. Presumivelmente, novos fenômenos de computação ocorrerão, e novos ensejos à pergunta se darão. Quão radicais serão os questionamentos ensejados, e que contornos emprestarão a nossa compreensão e a nossa lida com noções como computador, inteligência e humanidade não são – ao menos ainda não – questões que emergem da situação histórica em que nos encontramos.

Referências

- DE CASTRO ALVES, T. Towards an evaluation of the normalisation thesis on identity of proofs: the case of the Church-Turing thesis as touchstone. *Manuscrito – Rev. Int. Fil.* Campinas, v. 43, n. 3, p. 114-163, Jul.-Sep. 2020.
- DE LAS CASAS, B. Tratado de Indias y El Doctor Sepúlveda, In: *Fuentes para la Historia Colonial de Venezuela*. Caracas, 1962.
- DESCARTES, R. *Discurso do método*. In: Coleção Os pensadores, vol. XV., Trad. J. Guinsburg e Bento Prado Jr.. São Paulo: Abril Cultural, 1973, p. 33-80.
- DESCARTES, R. *Meditações concernentes à primeira filosofia*. In: Coleção Os pensadores, vol. XV., Trad. J. Guinsburg e Bento Prado Jr. São Paulo: Abril Cultural, 1973. p. 80-222.
- DE SOUZA, C. S. *Reflexões sobre Semiótica, Linguagem, Ontologia e Computação*. EMAPS-Notas #02. Rio de Janeiro, RJ - Brasil: SERG, Departamento de Informática, PUC-Rio, 2023.
- FLEMING, M. Gadamer's Conversation Does the Other Have a Say?. In: Code, L. *Feminist Interpretations of Gadamer*. University Park: Pennsylvania State University Press, 2003.
- FIUMARA, G. The Development of Hermeneutic Prospects. In: Code, L. *Feminist Interpretations of Gadamer*. University Park: Pennsylvania State University Press, 2003.

- GADAMER, H.-G. *Verdade e Método*. Tradução de Flávio Paulo Meurer. Petrópolis, Vozes; Bragança Paulista, SP: Editora Universitária São Francisco, 1997.
- GADAMER, H.-G. *Verdade e método II*. Tradução de Enio Paulo Giachini; revisão da tradução de Márcia Sá Cavalcante Schuback. Petrópolis, Vozes; Bragança Paulista, SP: Editora Universitária São Francisco, 2002.
- LEIBNIZ, G. *Monadologia*. In: Coleção Os Pensadores, v. XIX, Trad. M. S. Chauí Berlinck. São Paulo: Abril Cultural, 1974, p. 61-73.
- OPPY, G. e DOWE, D. The Turing Test, *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.) Disponível em <https://plato.stanford.edu/archives/win2021/entries/turing-test/> 2021. Acesso em 10 mai. 2024.
- PALMEIRO, T. *Informática: representações sem exterioridade*. Manuscrito não publicado, 2023.
- SEARLE, J. Minds, Brains and Programs, *Behavioral and Brain Sciences*, 3: 417–57, 1980.
- SEARLE, J. The Chinese Room, in R.A. Wilson and F. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA: MIT Press, 1999.
- SEARLE, J. Why Dualism (and Materialism) Fail to Account for Consciousness, in Richard E. Lee (ed.), *Questioning Nineteenth Century Assumptions about Knowledge* (III: Dualism), New York: SUNY Press, 2010.)
- SEPÚLVEDA, J. G. *Democrates Alter de Justis Belli Causis Apud Indios*. Edição bilingue latim-espanhol: *Tratado Sobre las Justas Causas de la Guerra contra los Indios*. México, Fondo de Cultura Económica, 1987.
- TURING, Alan M. On Computable Numbers, with an Application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society*, second series, 42(1): 230–265, 1936.
- TURING, A. Computing Machinery and Intelligence, *Mind*, 59(236): 433–460, 1950.
- TURING, A. *Programmers' Handbook for Manchester Electronic Computer Mark II*, Computing Machine Laboratory, University of Manchester, 1950.
- WITTGENSTEIN, L. *Investigações filosóficas*. In: Coleção Os Pensadores, v. XLVI. Trad. J. C. Bruni. São Paulo: Abril Cultural, 1975.