



INTELIGÊNCIA ARTIFICIAL FRACA E FORÇA BRUTA COMPUTACIONAL

Weak artificial intelligence and computational brute force

ALEXANDRE QUARESMA

Pontifícia Universidade Católica de São Paulo (PUC/SP), Brasil

KEY WORDS

Artificial Intelligence (AI)
Turing Universal Machine (MUT)
General artificial intelligence (IAG)
GPT-3
Cognition
Hard problem
Weak AI
Strong AI
Computer science
Computational brute force
Technology criticism

ABSTRACT

The purpose of this essay is to reflect critically on the current condition of our most modern artificial intelligence systems, their latent potentials, their technical possibilities, and also on their intrinsic structural and functional limitations. For this, we will take as an emblematic example of this time the new GPT-3 from OpenAI, which is a powerful specialist artificial intelligence system, therefore, a system designed and used according to the weak theory of AI, which is nothing more than a model of relatively complex artificial verbal language, structurally forged with 175 billion parameters and variables.

PALAVRAS-CHAVE

Inteligência artificial (IA)
Maquina Univeral de Turing (MUT)
Inteligência artificial geral (IAG)
GPT-3
Cognição
Hard problem
IA fraca
IA forte
Ciências computacionais
Força bruta computacional
Crítica da tecnologia

RESUMO

O objetivo desse ensaio é refletir criticamente sobre a condição atual de nossos mais modernos sistemas de inteligência artificial, seus potenciais latentes, suas possibilidades técnicas, e também sobre suas intrínsecas limitações estruturais e funcionais. Para tanto, tomaremos como exemplo emblemático desta feita o novo GPT-3 da OpenAI, que é um poderoso sistema de inteligência artificial especialista, logo, um sistema concebido e usado segundo a teoria fraca da IA, que nada mais é do que um modelo de linguagem verbal artificial relativamente complexo, forjado estruturalmente com 175 bilhões de parâmetros e variáveis.

Recebido: 14/01/2021

Aceite: 15/01/2021

1. Introdução

[Inteligência artificial é a] ciência que faz as máquinas realizarem coisas que os homens julgam inteligentes quando realizadas por eles.

Marvin Minsky apud Herbert Simon (1992, p. 223)

O estado da arte das Maquinas Universais de Turing e seus respectivos programas, em suas mais diferentes formas e aplicações – o que significa dizer, computadores, androides e robôs – melhora em potência e velocidade todos os dias, com processadores e sistemas mais rápidos e eficientes, com engenharia de software mais elaboradas e complexas, com hardwares mais robustos e poderosos, muitas vezes arranjados paralelamente, com algoritmos evolucionários, aprendizado profundo de máquina e redes neurais, todavia a lógica computacional e estrutural binária subjacente permanece absolutamente a mesma, o que significa dizer que, além de potenciais incríveis, essa lógica subjacente também possui suas limitações intrínsecas. Enquanto a IA forte¹ se debate contra o recalcitrante hard problem da consciência, prometendo mas nunca entregando a chamada capacidade cognitiva igual ou superior à humana, ou a IAG (inteligência artificial geral), a IA fraca – por seu turno – segue a passos largos, evoluindo com sistemas especialista mais e mais sofisticados a cada dia, elaborados segundo um número astronômico de parâmetros e diretrizes, rotinas e protocolos, e vão atuando nas mais diferentes frentes das sociedades atuais, como aliados poderosos. Um sistema especialista² pode

¹ “[A IA forte pretende] simular as funções superiores do cérebro”, informa-nos Changeux e Connes (1995, p. 169), enfim, “da inteligência humana, com auxílio de um computador. De algum modo, trata-se de substituir o cérebro do homem [ser humano] por uma máquina. Os êxitos da inteligência artificial são numerosos; robôs que pintam automóveis, computadores que comandam as viagens de naves espaciais até Marte, e além dele, sistemas especializados que informam sobre os mais recentes progressos da medicina etc. Todavia, a inteligência artificial não tem a ambição de compreender como o cérebro humano funciona, mas apenas “simular” algumas de suas funções. Esse procedimento é de um alcance bem limitado”.

² Massimo Negrotti (2012, p. 406) nos informa que “um sistema especialista [...] é um tipo de *software* que é capaz

sim superar o humano, e faz isso o tempo todo, pois suas diretrizes internas associadas a força bruta computacional³, podem gerar num átimo padrões de resposta ou memória que exigiriam o trabalho de centenas ou até milhares de seres humanos ao mesmo tempo. Tomemos um exemplo trivial: como comparar o trabalho de um algoritmo de inteligência artificial que opera com identificação facial, que examina um número astronômico de dados referentes às similaridades e diferenças, e apresenta seu resultado em poucos segundos? Quanto tempo um ser humano ou vários levariam para fazer esse mesmo trabalho? De modo que, é assim que os diversos sistemas especialistas – cada um na sua área – superam os seres humanos. Independentemente da concepção ou definição que se tenha de inteligência artificial (IA), e essas

de fornecer consultoria, em termos de explicação e previsão, para o usuário em um campo específico de conhecimento, como medicina, lei, ou outro qualquer. O sistema é capaz de fazer isso com uma taxa de sucesso aceitável graças à ‘doação’ de um perito humano, que decanta, como é seu conhecimento profissional em um banco de dados. Em seguida, o software, através de um conjunto de regras inferenciais e estatísticas incorporadas nele, torna-se capaz de entregar a sua consultoria como se fosse, dentro de certos limites, o próprio perito humano”. “O ponto-chave”, explica-nos Massimo Negrotti (2012, p. 406), “é que o que é modelado em um sistema especialista não é um cérebro humano, nem uma suposta mente, mas os resultados finais – conhecimento e regras – que os seres humanos obtiveram depois de ter trabalhado durante séculos sobre as melhores maneiras de raciocinar com os fatos dentro de um determinado domínio. É por isso que nenhum programa AI tem sido capaz ainda de propor algum problema novo, embora muitos desses programas sejam, sem dúvida, úteis no domínio de resolução de problemas”.

³ “Quando falamos de computação”, como discutimos em Quaresma (2020, p. 198-199), “estamos falando na verdade de uma máquina de calcular extraordinária, que de fato é capaz de fazer cálculos gigantescos em frações de segundo, manipulando enormes quantidades de memórias e dados, e isso permite a representação algorítmica de diversos tipos de dispositivos e funções diferentes, e com isso gerar e captar imagens, sons, desenhos, maquetes, projetos, modelos, objetos tridimensionais, e até mesmo alguns sistemas complexos, podem de fato ser representados dentro de um computador por meio de linguagem algorítmica. O que significa dizer que são traduzidos – no extremo – para sequências de zeros e uns infundáveis, e o que se passa basicamente no interior de um computador é força bruta computacional de cálculo, ou seja, de computação pesada para manipular e extrair informações desses mesmos dados e memórias”.

podem ser muito amplas e imprecisas⁴, é preciso dizer que inteligências artificiais funcionam em computadores, andróides e robôs, e que computadores, andróides e robôs são, sob todas as perspectivas e análises possíveis e pensáveis, são Máquinas Universais de Turing (MUT), ou, mais simplesmente, Turing Machines (TM). Uma Máquina Universal de Turing, que vai trabalhar com cálculos matemáticos, que por sua vez são expressões objetivas de algoritmos organizados deterministicamente em arranjos complexos numa fita virtual, não pode nem poderá lidar com sentido, significado, perspectiva de mundo, background cultural, nem com nada parecido que sequer o valha⁵. É exatamente por isso que ainda não estamos às voltas com computadores, andróides e robôs realmente inteligentes e conscientes no âmbito societal, caminhando autonomamente pelas ruas, avenidas e estradas do mundo.

Mas, por uma questão de método, comecemos pelos fundamentos basilares do nosso objeto de estudo e crítica nesse paper, ou seja, iniciemos afirmando que robôs, andróides e computadores operam segundo uma teoria, segundo uma lógica e segundo uma interpretação computacional, como já escrevemos doutra feita⁶. Segundo essa mesma interpretação, lógica e teórica, tais sistemas são concebidos e usados como

⁴ Como nos informa David Parnas (2017, p. 27), “notavelmente, aqueles que usam o termo ‘Inteligência artificial’ não definiram esse termo. Ouvi pela primeira vez o termo há mais de 50 anos e ainda não ouvi uma definição científica. Mesmo agora, alguns especialistas em IA dizem que definir IA é uma questão difícil (e importante) – na qual eles estão trabalhando. ‘Inteligência artificial’ continua a ser uma palavra da moda, uma palavra que muitos pensam entender, mas ninguém pode definir”.

⁵ “Numa situação em que o universo dos elementos candidatos à representação é constrangido e bem definido (por exemplo, quando um computador é programado ou quando uma experiência é conduzida a partir de um conjunto predefinido de estímulos visuais)”, informa-nos Francisco Varela (2017, p. 63), “a origem do sentido é clara. Cada elemento discreto, físico ou funcional, é concebido para *corresponder* a um objeto exterior (a sua referência), através de uma função que o observador pode facilmente fornecer. Fora destas obrigações, aos símbolos só resta sua forma: estão tão desprovidos de significado, como um grupo de *bits* cujo manual de informações desapareceu”.

⁶ *Inteligência artificial e os limites da computação*, A. Quaresma (2018).

Máquinas Universais de Turing. Máquinas Universais de Turing – por sua vez – são sistemas artificiais binários, quantitativos, deterministas, sintáticos e alopoiéticos *Ipso facto* – são limitados em relação ao que podem ou não fazer, ou seja, ao que é computável e ao que não é⁷. Nem tudo é computável, ou seja, nem tudo pode ser transformado em cálculo. A mente, a intencionalidade e a consciência – por exemplo, como fenômenos qualitativos e não quantitativos –, que estão presentes nas inúmeras formas vivas existentes, não são computáveis. O mesmo raciocínio é válido para o cérebro biológico, levando-nos – então –, obrigatoriamente, aos limites da computação. Os limites da computação nos remetem ao enorme abismo que ainda existe entre inteligência biológica consciente, e inteligência artificial inconsciente, em termos de performance e eficiência objetivamente inteligente no mundo. A teoria fraca da IA – ao contrário da forte, que pretende alcançar a consciência artificial, superando o *hard problem* –, resume-se à criação de sistemas cibernético-informacionais especialistas cada vez mais sofisticados e complexos, que são sistemas que, por definição e limitação funcional, fazem uma só coisa, que resolvem um só tipo de problema de cada vez, e isso é um sistema especialista. O mundo ao nosso redor está repleto desses sistemas especialistas, considerando que, de termostatos de geladeira, passando por aplicativos de smartphones, até pilotos automáticos de carros e aviões, o que temos são sistemas especialistas. Ao passo que a IA forte encontra forte impedimento no *hard problem* da consciência, como já discutimos em artigo progresso⁸, a IA fraca, por outro lado, tem se desenvolvido significativamente desde os primórdios desse campo nos anos 1950, em meados do século passado.

2. O GPT-3

Se pegamos um programa de IA e tentamos seguir a pista do que constitui a fonte de sua

⁷ *A falácia lúdica das três leis: Ensaio sobre inteligência artificial, sociedade e o difícil problema da consciência*, A. Quaresma (2020).

⁸ *Inteligências artificiais e o problema da consciência*, A. Quaresma (2019).

inteligência, veremos que esta desaparece dentro de um labirinto de subprocessos interconectados que, tomados separadamente, são completamente triviais.

Daniel Crevier (1996, p. 296)

Um exemplo extraordinário desse tipo de aperfeiçoamento e sofisticação é o GPT-3, da OpenAI. Trata-se de um poderoso sistema especialista, ou seja, um sistema que opera segundo a teoria e a lógica fraca da IA, sob a forma de um extraordinário modelo de linguagem artificial, e que é extraordinário justamente porque é forjado com – nada mais, nada menos – do que 175 bilhões de parâmetros, o que significa dizer que, a cada processamento de texto, a cada escolha de palavra, são consultadas ou podem ser consultadas 175 bilhões de hipóteses variáveis para tentar determinar a probabilidade da escolha da mais correta. Lembrando que um sistema especialista que opera binariamente em computadores, androides e robôs, é absolutamente dependente de seu programa, e das listas de parâmetros e variáveis armazenadas em seus programas, protocolos de funcionamento e rotinas de comandos, que ele pode utilizar em seus processos computacionais rotineiros para resolver um tipo específico de problema, no caso aqui, tentar articular textos em linguagem verbal, seja ela falada ou escrita. A primeira coisa que tem de ficar clara é que o GPT-3 não traz nenhuma inovação importante em termos de processos computacionais, pois ele simplesmente opera dentro dos parâmetros de qualquer sistema especialista, que é justamente a consulta de seus bancos e dados internos, buscando as possíveis melhores opções de palavras, dentre sua 175 bilhões de instruções e diretrizes. O que o GPT-3 tem de verdadeiramente inovador e até surpreendente, é justamente a força bruta computacional que ele utiliza na sua resolução de problemas computacionais⁹.

⁹ Quanto a isso, David Chalmers (2020) é enfático: “o GPT-3 não contém nenhuma nova tecnologia importante. É basicamente uma versão ampliada do GPT-2 do ano passado, que em si era uma versão ampliada de outros modelos de linguagem usando aprendizado profundo. Todas são enormes redes neurais artificiais treinadas em texto para prever qual será a próxima palavra em uma sequência. O GPT-3 é simplesmente maior: 100 vezes maior

Além disso, nada – é claro – nem parecido com consciência ou subjetividade pode ser encontrado no funcionamento ordinário do GPT-3. Mas, por outro lado, como ele abriga grandes porções de nossa própria inteligência humana, sob a forma de articulações de linguagem verbal, como extensões de nossa própria inteligência em seus sistemas, é possível identificar com certa clareza que o sistema acaba exibindo algum tipo de comportamento inteligente de algum modo semelhante ao humano, mas trata-se de algo que ainda se encontra muito longe de ser comparável com a própria inteligência humana que concebeu e produziu o GPT-3. E ainda que ele possa também criar programas de computador muito mais simples que ele mesmo¹⁰, não há um sujeito intencional e consciente subjacente ao sistema. Ao mesmo tempo que ele não pode ser considerado absolutamente desprovido de inteligência alguma, pois a inteligência que ele esbanja em seu funcionamento, é uma inteligência oriunda da inteligência biológica humana, de maneira que seria impossível não associar a articulação e produção de linguagens humanas a máquinas e sistemas que imitam o comportamento inteligente humano para os próprios humanos através também de linguagem humana. Há quem diga até que o GPT-3 seria, de algum modo ainda não comprovado nem

(98 camadas e 175 bilhões de parâmetros) e treinado em muito mais dados (CommonCrawl, um banco de dados que contém grande parte da Internet, junto com uma enorme biblioteca de livros e toda a Wikipedia)”. “Significativamente”, continua Chalmers (2020), “ele pode escrever seus próprios programas de computador com algumas instruções informais. Ele pode até projetar modelos de aprendizado de máquina. Felizmente, eles não são tão poderosos quanto o próprio GPT-3 (a singularidade ainda não está aqui)”.

¹⁰ “Todos sabem”, escreve Von Neumann (1966, p. 79), “que uma máquina-ferramenta é mais complicada que os elementos que podem ser feitos com ela, e que, em geral, um autômato A, que pode fazer um autômato B, deve conter uma descrição completa de B e as regras sobre como comportar-se enquanto efetua a síntese. Assim, dá-se uma impressão muito forte de que a complicação, a potencialidade produtiva de uma organização, é degenerativo, que uma organização que sintetiza algo é necessariamente mais complicada, de uma ordem superior, do que a organização que sintetiza”. Além disso, aponta Von Neumann (1966, p. 79), “apelar para o mundo orgânico, vivo, não nos ajuda muito, porque não entendemos o suficiente sobre como os organismos naturais funcionam”.

comprovável, consciente¹¹. Nós, por nosso turno, inclinamo-nos a discordar totalmente dessa hipótese, pois sabemos com alto grau de certeza e razoabilidade que a mera execução de um programa qualquer não é absolutamente suficiente para conferir uma mente intrinsecamente inteligente e consciente ao sistema computacional¹². Nesse ponto, ainda que tratemos objetivamente da IA fraca, logo, de sistemas especialistas, acabamos por esbarrar inelutavelmente na teoria da IA forte e seu inexorável *hard problem* da consciência, pois enquanto uma simulação de modelo de linguagem artificial, o GPT-3 se sai até que surpreendentemente bem em alguns casos, mas, por outro lado, como imitação da articulação de fala e escrita humana consciente, ele é significativamente deficiente, e o que está em jogo é apenas uma espécie de truque que induz ao equívoco de compreensão acerca do que está realmente acontecendo no interior do referido sistema¹³. Trata-se de uma lógica antiga, que remonta os primórdios da inteligência artificial, e

¹¹ Chalmers (2020) escreve *in verbis* que, “quanto à consciência, estou aberto à ideia de que um verme com 302 neurônios é consciente, então estou aberto à ideia de que o GPT-3 com 175 bilhões de parâmetros também é consciente”.

¹² Como aponta pertinentemente John Searle (2017, p. 52-53), “Nenhum programa de computador é, por si só, suficiente para dar uma mente a um sistema. Os programas, em suma, não são mentes e por si mesmos não chegam para ter mentes. Ora, esta é uma conclusão muito poderosa, porque significa que o projeto de tentar criar mentes unicamente mediante projetar programas está condenado, desde o início [grifos do autor].” “Alguma coisa só é um computador em relação a uma interpretação computacional” Searle (2010, p. 197). John Searle (2017, p. 42) adverte que “A razão por que nenhum programa de computador pode alguma vez ser uma mente é simplesmente porque um programa de computador é apenas sintático, e as mentes são mais do que sintáticas. As mentes são semânticas, no sentido de que possuem mais do que uma estrutura formal, têm um conteúdo”.

¹³ Shannon Vallor (2020) afirma que “a capacidade do GPT-3 de se deslumbrar com prosa e poesia que soa totalmente natural, até mesmo erudita ou lírica, é menos surpreendente. É um truque de salão que o GPT-2 já executou, embora o GPT-3 seja complementado com mais parâmetros [...] para aprimorar suas abstrações estilísticas e associações semânticas. Assim como sua bisavó ELIZA, ambos se beneficiam de nossa confiança em heurísticas simples para as habilidades cognitivas dos falantes, como ritmos de fala artísticos e sonoros”.

também de um sintoma de uma sociedade tecnológica que se mistura e se confunde com as suas próprias criações tecnológicas. Em linhas gerais, isso tudo aponta pelo menos dois aspectos relevantes: o fato de que talvez simplesmente (i) aumentar progressivamente a força bruta computacional não seja suficiente para pretensamente gerar a emergência de uma mente no interior do sistema. E que talvez (ii) tenhamos que trilhar outros caminhos teóricos e práticos para um dia poder alcançar a inteligência geral artificial, ou mais simplesmente IAG. Alcançar uma IAG, ou seja, um tipo de inteligência igual ou superior à humana, significa necessariamente também superar o *hard problem* da consciência, como igualmente alcançar um nível tecnológico de inteligência tão impressionantemente avassalador, que seria até mesmo perigoso para a própria humanidade que cria seus sistemas cibernético-informacionais¹⁴.

3. O recalcitrante e renitente *hard problem*

[Os fenômenos] são experimentados por um ‘eu’ que tem a experiência, e é nesse sentido que possuem uma ontologia de primeira pessoa.

John Searle (2010)

E é justamente aqui que o GPT-3 e os demais sistemas que simulam as capacidades biológicas humanas resvalam irreversivelmente no difícil problema da consciência. Grossíssimo modo, o problema de fundo e maior – em relação ao *hard problem* – repousa no fato inexoravelmente dado de que (i) a consciência é sempre subjetiva em relação ao sujeito que a experiencia, e que (ii) cada sujeito experimenta essa consciência de modo diferente, e que também (iii) ela é inacessível a outras consciências, já que nenhuma consciência pode experienciar a consciência de outra pessoa. Junte-se a isso o agravante fato de que (iiii) não sabermos como a

¹⁴ Nick Bostrom (2018, p. 139) sustenta que “se a IA de nível humano ainda não existir em virtude da falta de um determinado *insight* por parte dos programadores, então, quando a descoberta final ocorrer, a IA poderá saltar para um nível radicalmente superior ao humano sem ao menos passar pelos degraus intermediários”.

mente consciente funciona, e nem muito menos como representa-la por meio de linguagens formais conhecidas, e temos então, como resultado, o famigerado hard problem, ou – em bom português – o difícil problema da consciência e sua recalcitrante e renitente insolubilidade. Tudo isso dificulta a reprodução, emulação ou mesmo simulação de uma consciência artificial em sistemas cibernético-informacionais dotados com inteligência artificial (IA). De maneira que, o hard problem é justamente aquilo que a teoria forte da IA se propõe a solucionar, ou seja, decifrar o difícil problema da consciência e ainda conseguir instanciar uma inteligência artificial consciente igual ou superior em sistemas cibernético-informacionais. Porém, ser consciente não é um estado computacional, ou, dito de outra forma, o cérebro definitivamente não é um computador¹⁵, e, por isso mesmo, torna-se improvável que um sistema computacional atual possa ser considerado de fato consciente. Lembrando que os estados conscientes dizem respeito a agentes, a sujeitos, às subjetividades desses sujeitos¹⁶, e não sabemos sequer descrever o que se passa no interior dos cérebros das pessoas para poder quem sabe imitar ou reproduzir o fenômeno da consciência. Como disse muito bem o filósofo, existem coisas que os computadores não podem fazer¹⁷. Vinte, trinta, cinquenta anos depois disso, muito desse “não podem fazer” ou “não conseguem fazer” ainda continua vigente, principalmente porque as mais hegemônicas correntes das ciências cognitivas compreendem a consciência como algo matematizável, computável, traduzível e reproduzível num

¹⁵ Como nos informa Pentti Haikonen (2003, p. 03), “o cérebro definitivamente não é um computador. Pensar não é uma execução de cadeias de comandos programados. O cérebro também não é uma calculadora numérica. Nós não pensamos por números. Portanto, não deveria ser de se admirar que, além de alguns casos triviais, os ‘mesmos resultados’ não possam ser facilmente obtidos por algoritmos computacionais”.

¹⁶ “Os estados conscientes”, informa-nos John Searle (1998, p. 138), “existem apenas quando experimentados por um sujeito, o que significa que eles existem apenas a partir do ponto de vista da primeira pessoa desse sujeito”.

¹⁷ Referimo-nos a Hubert Dreyfus. Shannon Vallor (2020) aponta que “Dreyfus, em particular, entendeu que o obstáculo da IA não é o desempenho (contra toda leitura lamentável de Turing), mas a compreensão”.

computador, androide ou robô. No caso específico do GPT-3, que é um sistema especialista baseado na teoria fraca da IA, não há como sequer pensar em consciência no interior do sistema, pois não há e nem poderia haver agente consciente, não há sujeito intencional, e sem um sujeito ou agente que seja intencional e consciente, não poderá haver – ipso facto – compreensão¹⁸, mesmo porque, compreensão está bem além da capacidade do GPT-3¹⁹.

Observando que uma mente biológica – incluso a humana – não possui nem poderia possuir uma lista com 175 bilhões de itens e dados diretivos para consulta, simplesmente porque a mente não é um computador binário, e simplesmente também porque não é assim que tudo funciona na dinâmica intrínseca da mente biológica consciente, com suas sistêmicas oscilações²⁰. O GPT-3, nesse sentido, é uma caricatura sem talento ou mesmo arte do que

¹⁸ Daniel Crevier (1996, p. 50) nos lembra que, “para nós humanos, compreender consiste e ser capaz de colocar uma informação nova em relação com outros fatos que já conhecemos, e em fazê-lo de forma que possamos extrair conclusões que antes não tínhamos. Quanto mais rica for essa rede de conexões, mais profunda será nossa compreensão”.

¹⁹ Shannon Vallor (2020) aponta que “a compreensão está além do alcance da GPT-3 porque a compreensão não pode ocorrer em um comportamento isolado, não importa quão inteligente. A compreensão não é um ato, mas um trabalho. O trabalho é totalmente irrelevante para um modelo computacional que não tem história ou trajetória; uma ferramenta que simula infinitamente um novo significado a partir de um conjunto de dados não vinculado a seus esforços anteriores. Em contraste, a compreensão é um trabalho social para toda a vida. É um projeto sustentado que realizamos diariamente, à medida que construímos, reparamos e fortalecemos os laços de sentido em constante mudança que nos ancoram aos outros, coisas, tempos e lugares, que constituem um mundo”.

²⁰ Como aponta muito bem György Buzsáki (2006, p. 48), no funcionamento do cérebro donde emerge a consciência, “decisões coletivas globais são necessárias para a maioria das funções corticais. As decisões coletivas, no entanto, exigem ações cooperativas de áreas vizinhas e distantes. Acredita-se que tal cooperação flexível entre as assembleias de células locais e distantes esteja subjacente a quase todos os comportamentos cognitivos”. György Buzsáki (2006, p. VIII) acrescenta que “embora saibamos muito sobre os neurônios, os blocos de construção do cérebro, e termos amplo conhecimento sobre sua conectividade, ainda sabemos muito pouco como os módulos e sistemas de módulos funcionam juntos. É aqui que as oscilações oferecem seus serviços inestimáveis”.

seria ser um ser humano²¹. E a razão disso é muitíssimo simples: nenhum ato de fala ou escrita determina um significado fechado e completo de nada que de fato significa e representa numa comunicação trivial. O contexto, o tom de voz, a expressão facial, a gestalt corporal, a hierarquia social, o ambiente cultural, os valores, o momento, enfim, tudo isso está sempre impregnado nas variadas formas de comunicação verbal humana, e tudo isso completa – nesse sentido – o verbo. E mesmo que se quisesse informar ao computador, androide ou robô ‘tudo’ que pode ser comunicado, segundo as regras da sintaxe linguística, a semântica do significado, do sentido e da compreensão, não é passível de representação, logo, também não é passível de ser computada. Uma simples palavra – como “flor”, por exemplo – requer uma série de outras palavras agrupadas em sentenças para explicá-las²² – ou seja, uma flor é um ser vivo do reino vegetal, que faz fotossíntese, que produz oxigênio, composta basicamente por carbono e água, que depende de luz solar, por exemplo –, e essas sentenças – para serem explicadas e definidas – precisarão de novas palavras e sentenças, ou, na linguagem técnica da lógica, novos axiomas, e isso de fato

²¹ Amanda Askill (2020) é objetivamente clara, ao afirmar que, “embora sua capacidade de realizar novas tarefas com pouca informação seja impressionante, na maioria das tarefas o GPT-3 está longe do nível humano”.

²² “Mesmo que todas as ambiguidades genuínas tenham sido removidas da sentença”, informa-nos John Searle (2006, p. 261), “ainda ficamos com a vagueza e a incompletude. Palavras são inerentemente vagas, e descrições são sempre incompletas. Mas mais precisão e completude são acrescentadas ao entendimento pelo fato de os significados serem *suplementados* por um conjunto de expectativas habituais [que o autor chama de *background*]”. “Mesmo que os axiomas sejam escolhidos dentro da área do senso comum”, acrescenta John Von Neumann (1966, p. 77), “geralmente é muito difícil conseguir um acordo entre duas pessoas que fizeram isso independentemente. Por exemplo, na literatura de morfologia existem tantas notações quanto autores, e qualquer um que tenha usado uma anotação por algumas semanas acha que ela é mais ou menos superior a qualquer outra. Assim, embora a escolha das notações, dos elementos, seja extremamente importante e absolutamente básica para uma aplicação do método axiomático, essa escolha não é nem rigorosamente justificável nem humanamente inequivocamente justificável. Tudo o que se pode fazer é tentar submeter um sistema que resista a critérios de bom senso”.

não tem fim²³. Pois qualquer que seja o enunciado, sempre haverá necessidade de um outro enunciado ou muitos outros enunciados, para poder tentar explicar os precedentes, e isso é assim – também – por razões absolutamente matemáticas²⁴.

4. Derradeiras e provisórias considerações

Só um mundo predefinido pode ser representado.

Francisco Varela (2017, p. 73)

Nem tudo que existe é computável. E compreender depende não apenas da ação, mas também da necessidade ordinária de um ser vivo topologicamente situado num mundo não predeterminado, mas sim num mundo dinâmico e oscilante, que ele mesmo constrói para si representativamente enquanto o experiencia, e nisso consiste a percepção, e ser – nesse sentido ontológico – significa ser necessariamente um ser biológico. Quando tratamos sistemas cibernético-informacionais, ainda não temos algo que possa ser sequer comparado com um ser vivente, e nem estamos – acreditamos nós – próximos disso, em termos de estado atual da

²³ “O teorema da incompletude de Gödel [...], informa-nos Changeux e Connes (1995, p. 174), “afirma que, quaisquer que sejam os axiomas, em número finito ou dados de maneira recorrente, existem sempre questões às quais não podemos responder, que permanecem indecidíveis, e para as quais nos faltarão informações. Em outros termos, o teorema de Gödel especifica que é impossível tomar um número finito de axiomas de tal modo que toda questão seja decidível. O que não significa que não podemos analisar uma questão a partir do que sabemos, mas que o número de questões interessantes e novas que precisarão ser adicionadas à resposta é infinito”.

²⁴ “Qualquer que seja a complexidade de um sistema formal”, apontam Changeux e Connes (1995, p. 176-177), “sempre haverá um enunciado incidindo sobre os números inteiros positivos, que será ao mesmo tempo verdadeiro e indemonstrável no sistema formal. [...] As proposições verdadeiras que incidem sobre os números inteiros positivos não podem ser reduzidas, via inferência lógica, a um número finito de axiomas. Logo, a quantidade de informações contida no conjunto de todas essas proposições é infinita”.

arte em IA²⁵. O que torna o GPT-3 especial e, em certos sentidos até surpreendente, é que ele opera com o que os teóricos da computação chamam de força bruta computacional²⁶. O mais importante a ser frisado para a presente discussão, é que no caso do GPT-3 não existe

²⁵ Shannon Vallor (2020) pensa o mesmo, pois “a compreensão faz mais do que permitir que um agente inteligente surfe habilmente [como faz o GPT-3], de momento a momento, nas conexões causais e associativas que mantêm juntos um mundo de significado físico, social e moral. A compreensão diz ao agente como soldar novas conexões que se manterão, suportando o peso das intenções e objetivos por trás de nosso comportamento”. Shannon Vallor (2020) acrescenta que “sua improvisação instantânea [do GPT-3] não está ancorada em um mundo; em vez disso, está ancorada numa abstração baseada em dados de um tipo de comportamento isolado – o comportamento de escrever um ensaio sobre a consciência de IA”. “Em uma era em que o trabalho de compreensão de fazer sentido é suplantado como medida de inteligência pela capacidade de criar um aplicativo que reinventa outra coisa que já existe – onde agimos mais como GPT-3 a cada dia – não é uma surpresa que o campo de pesquisa de IA tenha perdido seu rumo. Mas o futuro da IA espera por aqueles que a escolherem”.

²⁶ Como escrevemos em Alexandre Quaresma (2020, p. 48), “quando podemos enumerar, mensurar e listar tudo que há no campo das possibilidades [como no jogo de xadrez, por exemplo], e em seguida introduzir isso no computador, de maneira que o sistema especialista, diante de suas finitas listas de possibilidades, possa agir de acordo com o que lhe foi programado, e, através de *força bruta computacional*, processar essas listas de possibilidades, de modo que, ao final, exiba comportamento que possa ser considerado inteligente por nós, ainda que a referida inteligência – se analisada friamente – não tenha absolutamente nada de artificial. O que vale dizer que, de fato há inteligência em máquina e sistemas, mas essa inteligência é – em todos os sentidos – absolutamente humana”. Também em Quaresma (2020, p. 48), apontamos que “sua qualidade primordial é poder efetuar esses cálculos matemáticos a velocidades inimagináveis, de modo que, quando um humano joga com um computador, o humano tenta antecipar jogadas – o que é a essência do jogo de xadrez –, imaginando-as por antecipação, enquanto que, o computador, avalia uma enorme coleção de variáveis numa velocidade extraordinariamente alta, no que os especialistas chamam de *força bruta computacional*”. Sobre isso, Henry Shevlin (2020) nos explica que, “o que torna o GPT-3 notavelmente diferente de qualquer modelo de linguagem anterior é a sua escala: seus 175 bilhões de parâmetros para 1,5 bilhão do GPT-2, seus 45 TB [terabytes] de dados de treinamento de texto em comparação com 40 GB [gigabytes] do GPT-2. O resultado desse aumento dramático em escala foi um aumento notável no desempenho em uma série de tarefas. O resultado é que conversar com o GPT-3 parece radicalmente diferente de se envolver com o GPT-2: ele acompanha as conversas, se adapta às críticas e até parece construir argumentos convincentes”.

nenhum elemento novo, no que se refere a problemas e soluções computacionais, já que o referido programa opera utilizando a já referenciada força bruta computacional com Máquinas Universais de Turing, e que os métodos, teorias, lógicas e sistemas continuam absolutamente os mesmos²⁷. O que parece indubitável, é que o progressivo surgimento de novas tecnologias de IA, mais e mais complexas e potentes a cada nova geração – como o GPT-3, por exemplo –, passa a exigir reflexões profundas e complexas, bem como acabam por suscitar questões que definitivamente não estavam presentes nas mentes dos engenheiros e projetistas destes mesmo sistemas artificiais. Cabe a sociedade que cria tais ou quais tecnologias, estar apta a administrar e tutelar a conduta e as consequências – previstas ou não – de tais sistemas cibernético-informacionais. É apenas a sociedade como um todo – e não apenas os tecnólogos, mas também os cientistas sociais e filósofos da tecnologia – que poderá criar as condições para que sujam novas sensibilidades sociais, e que valores importantes estejam presentes nas pautas de discussão e, principalmente, no momento das tomadas de decisão, acerca do que se vai ou não produzir em termos de tecnologia e objetos técnicos²⁸. Se tais

²⁷ “Conforme observado”, aponta Henry Shevlin (2020), “o GPT-3 não possui nenhum tipo de novo tipo de arquitetura revolucionária, e há um debate contínuo sobre se aumentos adicionais na escala resultarão em aumentos concomitantes no desempenho. E os tipos de saídas GPT-3 dramáticas que são amplamente compartilhadas online estão sujeitas a efeitos de seleção óbvios; interaja com o modelo você mesmo e você logo se deparará com incompreensões, erros e equívocos alienígenas.” “Embora os cientistas práticos da computação há muito tempo tenham ampliado o conceito de algoritmos para além da computação de funções”, sustentam Dina Goldin e Peter Wegner (2004, p. 07), “a ciência da computação teórica manteve a cosmovisão matemática. Apesar do trabalho teórico de complexidade avançada que se aventura fora dessa visão de mundo, como algoritmos on-line e distribuídos, jogos Arthur-Merlin e provas interativas, nosso tratamento da Teoria da Computação no nível de graduação não mudou. Os princípios matemáticos continuam a enquadrar a computação como baseada em funções e a delimitar nossa noção de problema computacional de acordo”.

²⁸ Henry Shevlin (2020): “Da mesma forma, o surgimento de tecnologias como GPT-3 torna gritante a necessidade de mais acadêmicos em humanidades adquirirem pelo menos conhecimento técnico rudimentar e compreensão para melhor lidar com o impacto de novas ferramentas sendo

sistemas vão ou não convencer as pessoas de que são entes vivos ou não, pouco importa, pois os seres humanos já travam relacionamentos afetivos até com robôs sexuais, por exemplo, com suas interfaces de fala ainda bastante rudimentares²⁹, sem que haja impeditivo algum, o que aponta para um contexto tecnopolítico novo, onde seres humanos podem de fato se relacionar afetivamente com objetos técnicos inanimados, como o GPT-3, por exemplo³⁰. Segundo alguns autores, tais relacionamentos seriam consequência direta de um acirramento de nossa comunicação impessoal cada vez mais mediada por objetos técnicos e sistemas dos mais variados tipos, onde, de certa forma, o aspecto humano desaparece por trás da própria tecnologia utilizada³¹.

produzidas por empresas como OpenAI, Microsoft e DeepMind”.

²⁹ Sharkey, Wynsberghe, Robbins e Hancock (2017, p. 3) explicam que “o sucesso das bonecas para a gratificação sexual estabeleceu um caminho claro para o papel da robótica no futuro do sexo. Bonecas sexuais foram oferecidas por várias empresas, algumas das quais passaram a adicionar capacidades robóticas para suas bonecas. O RealDoll (cuja empresa mãe é Abyss Creations) fornecem bonecas de tamanho humano desde 1996 e suas bonecas foram apresentadas na cultura popular, incluindo o filme Lars e a Real Girl, estrelado por Ryan Gosling. Eles oferecem bonecas masculinas e femininas, bem como a capacidade de bonecas transgênero de encomenda. O CandyGirl, com sede no Japão, também oferece bonecos de sexo realistas”.

³⁰ Henry Shevlin (2020) afirma – *in verbis* – que “parece uma questão de quando, e não se, as pessoas começarão a desenvolver sentimentos persistentes de identificação, afeição e até simpatia por essas teias bizantinas de parâmetros ponderados. Por mais equivocados que esses sentimentos possam ser, nós, como sociedade, teremos de determinar como lidar com eles. O que significará viver em um mundo em que as pessoas buscam amizades ou até casos de amor com esses simulacros cognitivos, talvez exigindo direitos para os sistemas em questão? Aqui, parece-me, há uma necessidade vital e urgente de os filósofos anteciparem, armarem-se e se prepararem para a onda de novas e estranhas interações homem-máquina que virão”.

³¹ Carlos Montemayor (2020) sustenta que “nossa comunicação online se desvinculou do cuidado da atenção conjunta síncrona. Parece que não encontramos um terreno comum e os preconceitos estão agravando a falta de comunicação. Devemos abordar esse problema como parte da estratégia geral para projetar máquinas inteligentes”.

Ademais, quando estudamos o meio interno do sistema, e quando seguimos minuciosamente seus protocolos e rotinas, percebemos com clareza que o GPT-3 não é e nem poderia ser uma mente, já que seus atos de fala ou escrita refletem apenas e tão somente a programação e os bancos de dados que foram pré-implantados no sistema. Todavia, é importante perceber também que, de uma maneira muito especial, ele também representa uma extensão inteligente de nossa própria mente inteligente, já que ele expressa ideias, conceitos, linguagens e até comportamentos tipicamente humanos. Por isso, há quem arrisque dizer que tal sistema não é apenas uma máquina, devido exatamente a esse tipo de extensão inteligente que estes sistemas instanciam³². A respeito disso, ouve-se muito falar sobre o teste de Turing, mas o que não se menciona é que o referido teste depende de um alto grau de subjetividade dos sujeitos envolvidos nele, para poder determinar ou não se o computador, androide ou robô, passa ou não no referido teste, para aí então ser considerado ou não consciente, e que a dimensão interna do sistema praticamente é esquecida ou renegada a um terceiro plano, diante de sua manifestação e comportamento externo³³. Mas, atenção: o computador, androide ou robô que expressar comportamento pretensamente consciente e programado não terá ipso facto uma consciência genuinamente subjetiva, intrínseca e intencional.

³² Regina Rini (2020), por exemplo, sustenta que o “GPT-3 não é uma mente, mas também não é inteiramente uma máquina. É outra coisa: uma representação estatisticamente abstrata do conteúdo de milhões de mentes, conforme expresso em seus escritos. Sua prosa jorra de um funil indutivo que absorve grandes quantidades de conversas humanas na Internet: postagens do Reddit, artigos da Wikipedia, notícias. Quando o GPT-3 fala, somos apenas nós falando, uma análise refratada dos caminhos semânticos mais prováveis trilhados pela expressão humana”. “Ao enviar uma consulta de texto para GPT-3”, informa-nos Regina Rini (2020), “você não está se comunicando com uma alma digital exclusiva. Mas você está chegando o mais perto que alguém chegou de literalmente falar com o *zeitgeist* [sinal dos tempos]”.

³³ Regina Rini (2020) aponta que “realmente não importa se essa fluidez digital prolixa poderia passar em um Teste de Turing irrestrito, porque nossas interações diárias com eles serão exatamente como nossas interações diárias com a maioria dos humanos online: breve, tarefa específica, transacional”.

A consciência e a intencionalidade não podem e nem devem ser medidas apenas com base no comportamento externo do robô, androide ou computador que construímos, e sim por uma análise lógica de seu meio interno, sua constituição e seu hipotético acoplamento estrutural com a realidade, com a cultura, com a sociedade e com a própria história bioevolutiva que os precederam³⁴. A IA fraca caracterizada por sistemas especialistas possuem diversas aplicações úteis, superando a capacidade humana de fazer a mesma atividade em diversos setores da sociedade, mas – não nos enganemos, pois – ainda há problemas a serem resolvidos. Sua lógica e utilidade está ligada à grande quantidade de dados e parâmetros que podemos oferecer ao sistema como base de trabalho (Big Data), e também à força bruta computacional para poder processá-los. A IA fraca e os sistemas especialistas não pretendem – como a IA forte faz – decifrar a consciência nem mesmo superá-la. Eles se propõem apenas a tratar um único problema de cada vez, com esmagadora eficiência e velocidade, e isso já muita coisa. Sistemas especialistas – os softwares de computadores, androides e robôs – manipulam símbolos cujos significados não podem compreender, pois não há um agente consciente e cognoscente em seu interior que possa constituir e efetuar essa mesma cognição. Independentemente do hard problem da consciência e da própria IA forte³⁵, os sistemas especialistas e a IA fraca continuam a avançar

significativamente, e a tendência é que tenhamos mais e mais deles em pontos chave das nossas sociedades. Por outro lado, a consciência – assim como a intencionalidade – está intimamente relacionada à biologia, à expressão dos seres vivos. A biologia tem uma história evolutiva pregressa única que – enquanto dinâmica – remonta o surgimento da primeira forma de vida no planeta há aproximadamente 4,6 bilhões de anos atrás. Ser consciente e intencional está ligado a comportamentos eficientes e funcionais de sobrevivência dos seres vivos em resposta a concorrentes, semelhantes e ao próprio meio. Ainda há um enorme abismo entre inteligência biológica consciente e inteligência artificial em termos de performance e eficiência objetivamente inteligente. Consciência e intencionalidade são e estão expressas – em graus e níveis distintos – nas inúmeras formas vivas existentes. Todas as formas vivas existentes têm em comum como regra a dinâmica da autopoiese. Computadores, androides e robôs dotados com IA são sistemas alopoiéticos, logo – ipso facto –, não são conscientes. Conclusivamente – e ainda que estejamos cômicos de que essa conclusão seja de fato relativamente provisória –, é sempre bom lembrar que, a única forma de consciência intencional conhecida até agora, é a biológica.

³⁴ Igor Aleksander (2014) aponta que “as máquinas de hoje não têm mente própria; sua assim chamada inteligência é alcançada através do sangue, suor e lágrimas, de exércitos de programadores humanos brilhantes. Os seres humanos escrevem as regras indispensáveis que fazem com que as máquinas reconheçam sons, respondam a padrões visuais, tomem o próximo passo no xadrez e até sugiram quais ações comprar no mercado de ações. Embora essas máquinas sejam limitadas às tarefas para as quais foram projetadas, um ser consciente tem outra coisa: um sistema complexo de estados internos instanciado através de seus mecanismos neurais”.

³⁵ Nick Bostrom (2018, p. 25) nos explica que “máquinas com inteligência geral comparável à dos humanos – ou seja, dotadas de bom senso e capacidade real de aprender, raciocinar e planejar a superação de desafios complexos de processamento de informação em uma vasta gama de domínios naturais e abstratos – têm sido esperadas desde a invenção dos computadores, na década de 1940. Naquele tempo, o advento de tais máquinas era frequentemente esperado para os vinte anos seguintes”.

Referências

- Aleksander, I.(2014). Machine consciousness: fact or fiction? In: *Footnote: showcasing research with the power to change our world*. Disponível em: <footnote.co/machine-consciousness-fact-or-fiction>. Acesso em: 17 ago. 2019.
- Askell, A. (2020). GPT-3: Towards Renaissance Models. In *Philosophers On GPT-3 (update with replies by GPT-3)*, DAILYNOUS: News Four and About the Philosopher Profession. Acesso em: 16/08/2020. Disponível em: <http://dailynous.com/2020/07/30/philosophers-gpt-3/>
- Bostrom, N. (2018). *Superinteligência: caminhos, perigos e estratégias para um novo mundo*. DarkSide Books.
- Buzsáki, G. (2006). *Rhythms of the brain*. Oxford University Press.
- Chalmers, D. (2020). GPT-3 and General Intelligence. In *Philosophers On GPT-3 (update with replies by GPT-3)*, DAILYNOUS: News Four and About the Philosopher Profession. Acesso em: 16/08/2020. Disponível em: <http://dailynous.com/2020/07/30/philosophers-gpt-3/>
- Changeux, J. y Connes, A. (1995). *Matéria e pensamento*. Editora UNESP.
- Crevier, D. (1996). *Inteligência artificial*. Acento Editorial.
- Shevlin, H. (2020). A Digital Remix of Humanity. In *Philosophers On GPT-3 (update with replies by GPT-3)*, DAILYNOUS: News Four and About the Philosopher Profession. Acesso em: 16/08/2020. Disponível em: <http://dailynous.com/2020/07/30/philosophers-gpt-3/>
- Goldin, D y Wegner, P. (2004). The origins of the Turing thesis myth. *Technical Report CS04-14*. Brown University. Disponível em: <ftp.cs.brown.edu/pub/techreports/04/cs04-14.pdf>. Acesso em: 17 ago. 2019.
- Haikonen, P. (2003). *The cognitive approach to conscious machines*. Imprint Academic.
- . (2007). *Robot brains circuits and systems for conscious machines*. Wiley.
- Herbert A, S. (1992). O computador-rei. In: *PESSIS-PASTERNAK, Guitta (org.). Do caos à inteligência artificial: quando os cientistas se interrogam*, pp. 223-232.
- Motemayor, C. (2020). Language and Intelligence. In *Philosophers On GPT-3 (update with replies by GPT-3)*, DAILYNOUS: News Four and About the Philosopher Profession. Acesso em: 16/08/2020. Disponível em: <http://dailynous.com/2020/07/30/philosophers-gpt-3/>
- Negrotti, M. (2012). From the natural brain to the artificial Mind. In: *SWAN, Liz. (org.), Origins of mind*, v. 8: *biosemiotics, Drodrecht: Springer Science+Business Media*, pp. 399-410.
- Parnas, D. (2017). The real risks of artificial intelligence. *Communications of ACM*, 60(10).
- Quaresma, A. (2020). *Inteligência artificial e bioevolução: Ensaio epistemológico sobre organismos e máquinas*. Dissertação de Mestrado pelo programa de pós-graduação em Tecnologias da Inteligência e Design Digital (TIDD). Pontifícia Universidade Católica de São Paulo (PUC/SP).
- . (2020a). A falácia lúdica das três leis: Ensaio sobre inteligência artificial, sociedade e o difícil problema da consciência. In *PAAKAT: Revista de Tecnología e Sociedad*, número 19(10).
- . (2019). Inteligências artificiais e o problema da consciência. In *PAAKAT: Revista de Tecnología e Sociedad*, 16(9).
- . (2018). Inteligências artificiais e os limites da computação. In *PAAKAT: Revista de Tecnología e Sociedad*, 15(8).
- Rini, R. (2020). The Digital Zeitgeist Ponders Our Obsolescence. In *Philosophers On GPT-3 (update with replies by GPT-3)*, DAILYNOUS: News Four and About the Philosopher Profession. Acesso em: 16/08/2020. Disponível em: <http://dailynous.com/2020/07/30/philosophers-gpt-3/>
- Searle, J. (2017). *Mente, cérebro e ciência*. Edições 70.
- . (2006). *A redescoberta da mente*. Martins Fontes.
- . (2010). *Consciência e linguagem*. WMF/Martins Fontes.
- . (1998). *O mistério da consciência*. Paz e terra.
- Sharkey, N, Wynsberghe, A.V., Robbins, S., & Honcoch, E. (2017). *Consultation report our sexual future with robots*. Disponível em: <responsible-robotics-myxf6pn3xr.netdna-ssl.com/wp-

content/uploads/2017/11/FRR-Consultation-Report-Our-Sexual-Future-with-robots-.pdf>.

Acesso em: 14 abr. 2019.

Vallor, S. (2020). GPT-3 and Missing Labor of Understanding. In *Philosophers On GPT-3 (update with replies by GPT-3)*, DAILYNOUS: News Four and About the Philosopher Profession. Acesso em: 16/08/2020.

Disponível em: <<http://dailynous.com/2020/07/30/philosophers-gpt-3/>>

Varela, F. (2017). *Conhecer: as ciências cognitivas, tendências e perspectivas*. Instituto Piaget.

Von neumann, J. (1996). *Theory of self-reproducing automata*. University of Illinois Press.