



Cognição, compreensão e algoritmos¹

Cognition, comprehension, and algorithms

Cognición, comprensión y algoritmos

Winfried Nöth - Pontifícia Universidade Católica de São Paulo - PUCSP | São Paulo | SP | Brasil. E-mail: noeth@uni-kassel.de | Orcid: <https://orcid.org/0000-0002-2518-9773>

Resumo: Algoritmos são regras gerais que guiam computadores na resolução de problemas e execução de trabalhos. Nisso, eles fazem computadores interpretar e produzir signos. Um algoritmo é um legissigno simbólico argumentativo, o que implica que eles são agentes semióticos e participam em processo cognitivos. Implementados em máquinas de IA, os algoritmos também interpretam ícones diagramáticos, mas diferentemente dos humanos não ícones da classe das imagens. – Máquinas inteligentes agem como extensões cognitivos de cérebros humanos. Os riscos da extensão da inteligência humana para inteligência artificial são riscos que existem desde que a comunicação humana tem sido estendida a lugares distantes, onde foram interpretados na ausência dos emissores e começaram a adquirir uma autonomia além do controle dos emissores dos signos.

Palavras-chave: algoritmo; cognição; inteligência artificial.

¹ Este artigo se baseia numa contribuição do autor ao debate *Cognição, compreensão e algoritmos* no Seminário *Direito e Tecnologia do Banco do Brasil* em Brasília no dia 22/11/2023.



Abstract: Algorithms are general rules that guide computers in solving problems and carrying out labor. In this, they make computers interpret and produce signs. Algorithms are argumentative symbolic legisigns, which implies that they are semiotic agents and participate in cognitive processes. Implemented in AI machines, algorithms also interpret diagrammatic icons, but unlike humans, not icons of the image class. – Intelligent machines act as cognitive extensions of human brains. The risks of extending human to artificial intelligent are risks that have existed since human communication was extended to distant places, where they were interpreted in the absence of the senders and began to acquire an autonomy beyond the control of the senders of the signs.

Keywords: algorithm; cognition; artificial intelligence.

Resumen: Los algoritmos son reglas generales que guían en la resolución de problemas y realización de trabajos. Hacen que los ordenadores interpreten y produzcan signos. Un algoritmo es un legisigno simbólico argumentativo, lo que implica que son agentes semióticos y cognitivos. Implementados en máquinas de IA, ellos también interpretan iconos diagramáticos, pero no iconos de la clase de imágenes. – Las máquinas inteligentes son extensiones cognitivas de los cerebros humanos. Los riesgos de extender la inteligencia humana a las inteligencia artificial son riesgos que existían desde que la comunicación humana se extendió a lugares distantes, donde se interpretaban en ausencia de los emisores y empezaron a adquirir una autonomía que escapaba al control de los emisores.

Palabras clave: algoritmo; cognición; inteligencia artificial.



1 Introdução

O discurso sobre os algoritmos é um discurso metonímico. Quando os algoritmos são objeto de debate, não se costuma tratar de algoritmos mesmo, mas da maquinaria inteligente controlada por eles e os efeitos sociais da inteligência artificial em geral. Nesse contexto, a justaposição dos algoritmos com a cognição e a compreensão soa estranho. O que é que as máquinas têm a ver com cognição e compreensão? As máquinas não pensam e, portanto, não compreendem, ensinavam os teóricos do antropocentrismo, de René Descartes até John Searle (1980). Conforme a doutrina deles, as máquinas deviam ser um tema para os matemáticos e engenheiros, enquanto cognição e compreensão são temas para as ciências humanas. Discutir os algoritmos junto com a cognição e a compreensão seria, portanto, uma provocativa contradição em termos. Nas palavras de Santaella e Kaufman (2024), o pressuposto de que inteligência podia também existir no mundo não humano toca uma “ferida narcísica do humano”. Subjacente ao tema da oposição entre o humano e as máquinas é um tema ainda mais fundamental, isto é, o tema da incomensurabilidade entre a vida e a matéria bruta ou entre o humano e o não humano.

Num mundo dividido entre o humano e o não humano, os algoritmos ou têm a reputação de uma ameaça ou de uma promessa para a humanidade. Nos dois casos, o futuro dos seres humanos parece depender dos algoritmos. As visões desse futuro abrangem dos cenários imaginados pelos apocalípticos, que pintam o cenário do terror de uma “superinteligência” (Bostrom, 2014) na parede, até às visões dos futuristas entusiastas, que preveem o dia da singularidade tecnológica (Kurzweil, 2005) como o começo de uma época na qual “será possível implantar no cérebro um computador do tamanho de um grão de ervilha para substituir neurônios destruídos pelo mal de Parkinson” (Kurzweil, 2010).

Sem minimizar os perigos ou as promessas de um mundo dominado pelos algoritmos e as máquinas inteligentes, precisamos reconhecer que os conceitos precisam ser desmistificados (Kaufman, 2022) e que dualismos não resolvem problemas complexos. É mister reconhecer que nem a cognição é exclusivamente humana, nem os algoritmos são exclusivamente tecnológicos.

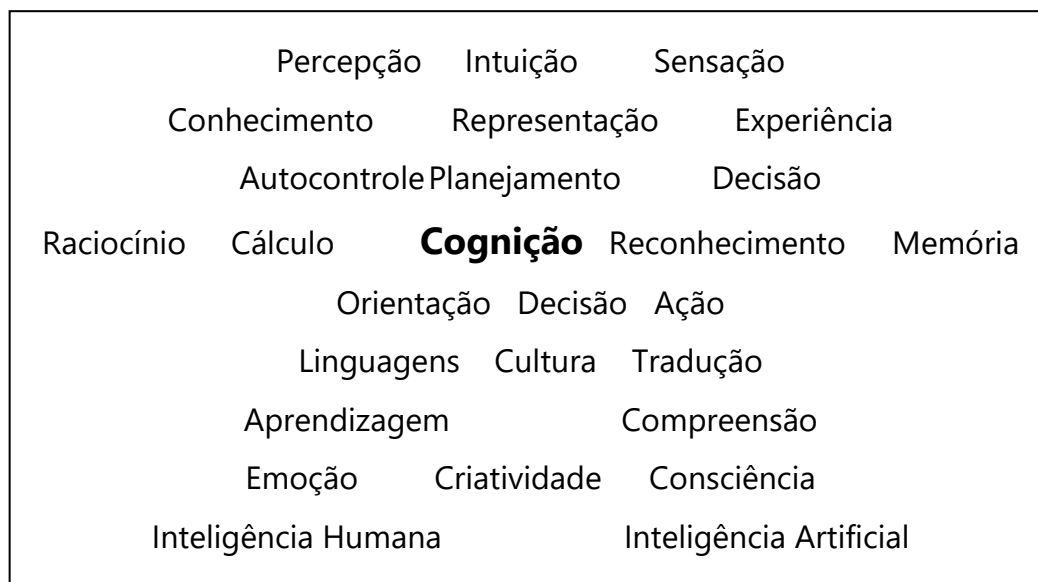
2 Cognição humana e cognição artificial

Cognição é o conjunto de processos mentais que abrangem desde a percepção até o raciocínio consciente. Ao examinar a pluralidade de processos cognitivos humanos, nota-se que a maioria deles também está sendo executada por máquinas inteligentes em processos computacionais que estendem as capacidades da cognição humana. Isso explica por que a IA é hoje considerada um dos ramos da ciência cognitiva. O seguinte mapa mental (Figura 1), reflete a abrangência dos temas tratadas



pelas ciências cognitivas. (Ver, por exemplo, o panorama abrangente da área das ciências cognitivas em Metzinger e Windt, 2016).

Figura 1 - Panorama das ciências cognitivas



Fonte: Elaboração própria.

Algoritmos de visão computacional estendem ou até substituem a percepção visual humana (Figura 2). Recursos inteligentes da síntese, do processamento e da análise de áudio estendem a capacidade auditiva humana, sistemas de navegação controlados por satélites estendem a nossa capacidade de orientação geoespacial (Figura 3), e o ChatGPT estende a capacidade humana de escrever textos informativos e aprimorar a qualidade retórica da expressão verbal.



Figura 2 - Reconhecimento facial



Nota: Prompt elaborado por Michele Eguchi

Fonte: Canva Pty Ltd. Imagem retratando o reconhecimento facial. Canva versão de 14 maio 2024. Inteligência artificial. Disponível em: https://www.canva.com/pt_br/. Acesso em: 14 maio 2024.

Figura 3 - Orientação geoespacial



Fonte: EGUCHI, Michele. **Orientação geoespacial**. 2024.



Essas e outras extensões de mentes humanas são extensões de cérebros que antes eram tidos por controlar a interação dos humanos com objetos no seu meio ambiente. A extensão da cognição humana para fora do cérebro e até o corpo humano não é um desenvolvimento recente. Ela tem acompanhada a evolução da inteligência da espécie humana desde as suas origens (Donald, 1991). As cunhas manuais que a espécie do *Homo habilis* usava há uns dois milhões anos eram extensões das mãos e dos braços da espécie humana assim como o ábaco, usado desde os mesopotâmios há uns sete mil anos, era uma extensão dos dedos com os quais os humanos costumavam calcular. Mãos e dedos são controlados por cérebros, e se o ábaco é uma extensão dos dedos, ele e os outros instrumentos manipulados pelos dedos humanos também são extensões de mentes humanas para fora do corpo. A evolução da cultura e tecnologia humana desde a idade de pedra até a idade da inteligência artificial tem o seu motor nesse processo da exteriorização da mente humana. Conforme o resumo de Merlin Donald da sua obra magistral sobre as origens da mente moderna e os estágios da sua evolução, a cognição tem sido “o mediador entre o cérebro e a cultura e, portanto, deve ter sido o motor e o lugar da mudança” (Donald, 1991, p. 2).

As implicações da extensão da mente humana se manifestaram de forma mais clara nos milênios da invenção da escrita. Os signos verbais, que antes só podiam ser comunicados presencialmente, começaram a desenvolver uma autonomia semiótica nas suas viagens para lugares distantes, porque na medida em que as mentes humanas se estenderam para signos escritos, lidos na ausência do emissor, os intérpretes adquiriram cada vez mais autonomia em relação aos signos recebidos. Os intérpretes, cuja autonomia aumentou com a ausência dos emissores dos signos, continuaram sendo mentes humanas por milênios, até que a sua substituição por mentes artificiais se tornou possível.

O reconhecimento facial e a orientação geoespacial são, sem dúvida processos nos quais acontecem interpretações de signos, isto é, são processos semióticos, mas as máquinas inteligentes cumprem não só serviços de interpretação, mas também tomam decisões em função das suas interpretações.

Ainda não sabemos exatamente quais processos ocorrem no cérebro de um porteiro de prédio, que reconhece as moradoras e as deixa entrarem, exceto que os signos interpretados nesse trabalho são signos análogos e, na interpretação semiótica, signos indexicais, que incluem signos icônicos da classe das imagens (Nöth; Santaella, 2017, p. 51-55). O que sabemos melhor é que a câmera que substitui o porteiro interpreta signos digitais. Esses signos também são índices das moradoras que incluem representações icônicas delas, mas os ícones interpretados pelos algoritmos pertencem a uma outra classe de ícone do que os ícones interpretados por olhos humanos. Não são ícones da classe, que Peirce definiu como imagens, mas ícones diagramáticos, isto é, signos visuais que reduzem as imagens análogas captadas pelas



câmeras a esquemas ou, como Peirce os descreveu em 1903, “ícones que representam as relações, principalmente diádicas, ou assim consideradas, das partes de uma coisa por meio de relações análogas em suas próprias partes” (CP 2.277). As relações diádicas nessa definição de ícones diagramáticos são as relações entre os pontos identificados pelos algoritmos das câmeras. As Figuras 2 e 3 representam essas relações diádicas na forma das linhas que conectam, cada uma delas, dois pontos da rede diagramática que representa as moradoras, por um lado, e as linhas, invisíveis entre esses pontos e os pontos da face identificada pela câmera dirigida pelos algoritmos, por outro. A decisão inteligente que a máquina de reconhecimento facial tem que tomar depois de cumprir a tarefa de reconhecer a pessoa é também uma decisão diádica e ao mesmo tempo digital: abrir (1) a porta ou não (0).

3 A mente estendida e a inteligência humana estendida pela inteligência artificial

A filosofia cognitiva contemporânea, numa das suas vertentes chamada teoria da mente estendida (Murphy, 2021), chegou à conclusão de que as extensões da cognição comprovam que a mente humana não é e nunca foi localizada exclusivamente no cérebro dos humanos, mas se estendeu há muito tempo para fora, para o meio ambiente, de forma que se pode até dizer que cognição humana também acontece “fora do cérebro”.

Marcos fronteiros estendem o conhecimento e a memória dos territórios para o meio ambiente, porque sem essas extensões da cognição não é possível memorizar o percurso exato da fronteira. Crianças delegam a tarefa de adicionar números de um dígito, do cérebro para os dedos das duas mãos capazes de fornecer evidência externa, por exemplo, de que a soma dos cinco dedos da mão esquerda e quatro da direita é nove. A invenção da escrita estendeu o alcance da palavra volátil para outras mentes em outros tempos e espaços.

Na evolução dos humanos, cada nova tecnologia cognitiva trouxe chances e riscos, avanços e retrocessos culturais. Quem diria que a invenção da escrita não tivesse sido um salto quântico na evolução humana? Porém, ninguém menos do que Platão, no passo 275a-b do seu diálogo *Fedro*, viu também a seguinte desvantagem nessa invenção. A descoberta da escrita “provocará nas almas o esquecimento [...] devido à falta de exercício da memória”, lamentou Platão há 2.400 anos.

Que a era digital e a inteligência artificial, que é o seu produto, tenha trazido um novo salto quântico na evolução humana tem sido amplamente discutido. O risco não é mais o risco do esquecimento, que Platão deplorou, mas surgiu outro, também relacionado ao tema da mente estendida: na medida em que a mente humana se estende, ela perde o controle sobre a informação comunicada. No ambiente digital, os perigos da extensão digital da mente humana são, portanto, os da perda de controle



sobre os dados que elas deixam como rastros da informação transmitida. Diminuir esses riscos da perda de autocontrole do cidadão sobre os seus dados é o desafio para a legislação, porque o poder do capital dos dados é superior ao poder de autocontrole do indivíduo sobre os seus dados, os quais deveriam ser a sua propriedade intelectual.

Porém, também é necessário lembrar que o risco de ser mal interpretado ou até interpretado de forma inovativa tem sido um risco desde sempre. Uma vez uma emissora emite uma mensagem, ela passa o potencial da interpretação dela para a comunidade dos intérpretes. A questão de se esse risco significa limites ou não (Eco, 2010) é um capítulo da hermenêutica semiótica.

4 Os algoritmos

Os algoritmos têm uma má reputação, não só nos jornais e revistas, mas também nas redes sociais, que elas mesmas devem a sua existência aos algoritmos que os controlam (ver Kaufman; Santaella, 2020). Desde que a IA se tornou cada vez mais presente em nossas vidas, os algoritmos têm sido os bodes expiatórios das novas tecnologias, responsabilizados por todos os tipos de problemas que surgiram com avanço das tecnologias inteligentes. Porém, bem considerados, os algoritmos em si não têm nada de assustador. Que são então os algoritmos, por que servem eles, e como é que se pode explicar a sua má reputação?

Os algoritmos devem o seu nome ao matemático al-Khwārizmī (ou Alcuarismi), que viveu na primeira metade do século 9 e ensinou matemática, geografia e astronomia em Bagdad (Crossley; Henry, 1990). A palavra “álgebra” é um conceito-chave nos escritos matemáticos de al-Khwārizmī, e devido a um tratado desse polímata medieval, “álgebra” entrou no vocabulário das línguas ocidentais. O mesmo tratado, traduzido para o latim em 1145, foi também fundamental para a substituição dos números romanos pelos algarismos hindu-árabes e o seu sistema numérico decimal no Ocidente. Nessa tradução, o nome do divulgador da álgebra aparece na forma latinizada “Algorizmi”, e dela se derivou a palavra “algarismo” (inglês: “algorism”) para designar os números árabes de 0 até 9 e as regras para usar o seu sistema decimal em cálculos matemáticos. Só a partir do século 18 surgiu a palavra “algoritmo” (com -t- ou -th-) como descendente de “algarismo” para designar certos métodos de cálculo exato. Do ponto de vista da história da matemática, há, portanto, não mais razão de ter medo dos algoritmos do que da álgebra e do sistema decimal dos números árabes.

Os algoritmos executados por computadores são sequências de instruções que a máquina tem que seguir para realizar um plano ou resolver um problema. Algoritmos não são programas computacionais, mas precisam ser reescritos ou traduzidos para a linguagem de um código que o computador entende, para ser implementado e



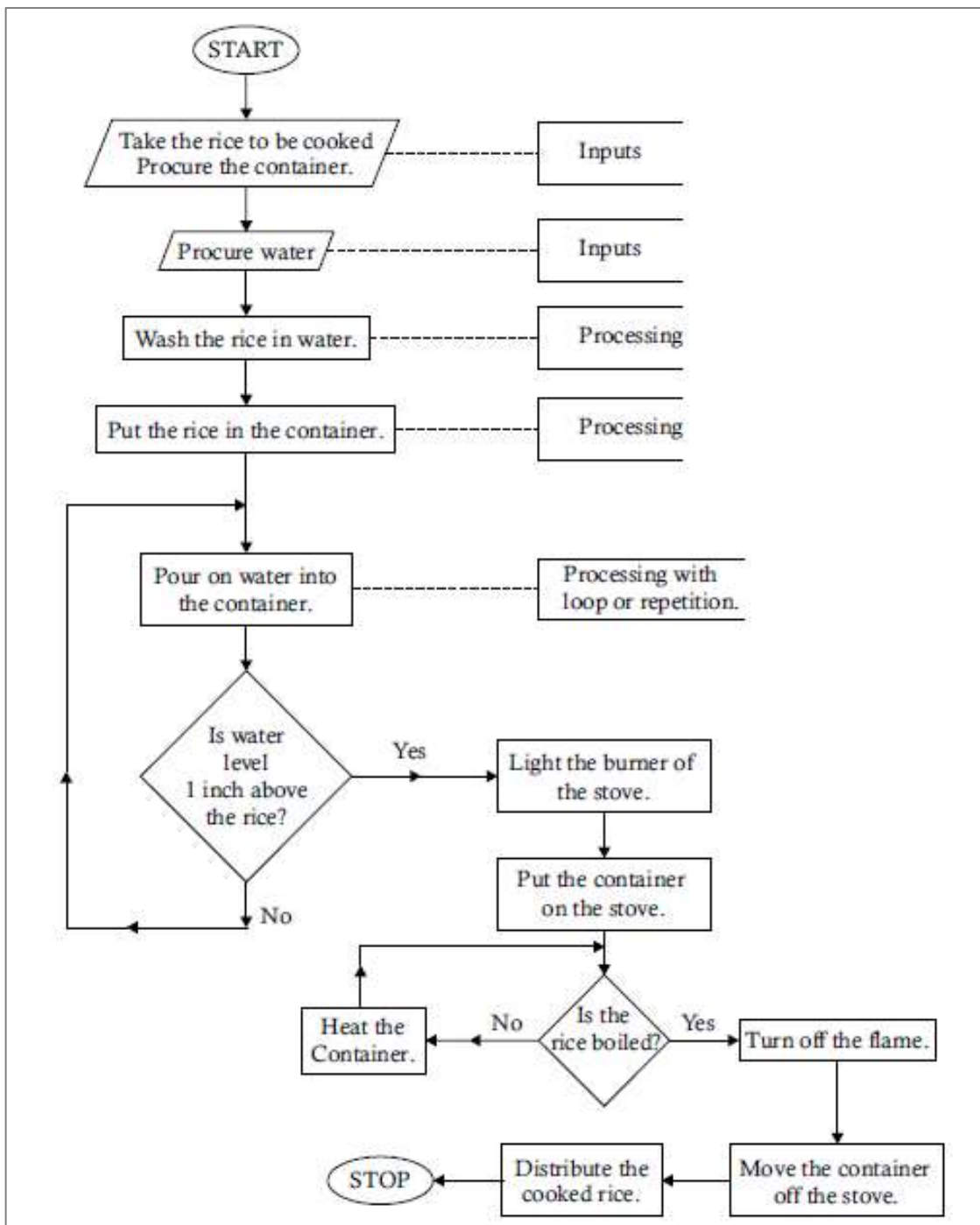
executado por um computador. (Sobre essas e outras diferenças, ver Santaella 2023, p. 30-37). Porém, os programas que são os resultados dessa implementação computacional são também algoritmos, porque eles refletem as sequências concebidas pelo algoritmo ainda não programado.

Um algoritmo pode ser escrito em forma de uma lista ou até em forma de uma narrativa em prosa, e ele é também um algoritmo se nenhuma implementação computacional for realizada. Por exemplo, as regras ensinadas nas escolas primárias para adição, subtração, multiplicação e divisão em colunas são algoritmos. Embora não sejam rotulados como tal, receitas culinárias, instruções de montagem de um móvel, que a fábrica entrega desmontado, e até as regras para escovar os dentes ou limpar os sapatos são algoritmos.

Um método comum para representar um algoritmo é um fluxograma. “Um fluxograma é uma representação diagramática das etapas de um algoritmo”, explica Chaudhuri (2020, p. 2). Como exemplo de um algoritmo, o autor apresenta uma receita para cozinhar arroz em forma de uma lista numerada e um fluxograma (Chaudhuri, 2020, p. 4-5). A lista numerada distingue sete etapas que a cozinheira tem que percorrer, dos quais os primeiros quatro são os seguintes: (a) Leve o arroz para ser cozido. (b) Obtenha o recipiente. (c) Obtenha a água. (d) Lave o arroz na água. Figura 4 representa o fluxograma do algoritmo completo da receita de cozinhar arroz.



Figura 4 - Um algoritmo para cozinhar arroz



Fonte: Chaudhuri (2020, p. 5).



Algoritmos são onipresentes na indústria, no comércio e nas redes sociais. Por exemplo, Kaufman e Santaella (2020), no seu estudo sobre “O papel dos algoritmos de inteligência artificial nas redes sociais”, distinguem cinco tipos de algoritmos desenvolvidos para controlar os fluxos nas redes sociais: (a) classificar o *Feed* de Notícias; (b) estabelecer um *ranking* pela “pontuação de afinidades” das semelhanças indicativas de um relacionamento entre usuários e os objetos de classificados; (c) selecionar um subconjunto de objetos de conteúdo em função das pontuações de afinidades auferidas entre o usuário e os objetos de conteúdo; (d) controlar a variável do *timing* para filtrar os objetos mais recentes e privilegiados; (e) priorizar interações ativas, definindo “ações de qualidade” as que requerem mais esforço do usuário (potencial gerador de mais interações, mais dados).

5 O algoritmo como signos da classe legissigno simbólico argumentativo

Algoritmos, portanto, são instruções gerais para executar tarefas recorrentes. Um cronograma de um encontro num dia específico não é um algoritmo, mesmo que ele especifique quem vai falar a que hora e em que sequência.

As regras gerais de um algoritmo são formuladas na forma dos símbolos de uma linguagem natural (inglês, português etc.) ou artificial (BASIC, Java, Prolog, Lisp etc.). Símbolos são ao mesmo tempo legissignos. Existem legissignos que não são símbolos, legissignos icônicos ou indiciais. Legissignos destas classes podem fazer parte de um algoritmo também. O diagrama pelo qual o programa de reconhecimento facial (Figura 2) identifica o morador do prédio como pessoa admissível não é um símbolo porque ele não representa um morador em geral, mas um indivíduo específico. O algoritmo ligado à câmera lê, portanto, um legissigno indexical que incorpora um legissigno icônico diagramático.

No fluxograma do algoritmo para cozinhar arroz (Figura 3) há três tipos de legissignos simbólicos, remas, dicentes e argumentos. Os elementos lexicais dos quais o algoritmo é composto, isto é, palavras como *water*, *rice* ou *put*, são remas. A segunda subclasse de símbolos, neste fluxograma, é o símbolo dicente, definido por Peirce como um símbolo que combina um sujeito com um predicado, como uma proposição faz. O algoritmo completo, isto é, a sequência de instruções do *start* até o *stop*, é um argumento, um legissigno simbólico argumentativo (Nöth, 2016). Ele é legissigno por ser um signo que formula uma regra geral, símbolo por ser um signo escrito numa língua que tem que ser apreendida e argumento porque a sequência das etapas, que o algoritmo prescreve têm um resultado, que segue os passos da sua produção como uma conclusão segue uma premissa.



O caráter distintivo do algoritmo de ser uma regra geral e o efeito pragmático de um símbolo é o tema de uma descrição exemplar com a qual Peirce descreveu como uma cozinheira segue uma receita. Embora o autor não use o termo “algoritmo”, a sua descrição das etapas em que a cozinheira planeja e executa o seu trabalho pode ser lida como um algoritmo em prosa:

Toda cozinheira tem em seu livro de receitas uma coleção de regras que está acostumado a seguir. Deseja-se uma torta de maçã. Agora, observe que raramente, provavelmente nunca, desejamos uma única coisa individual. O que desejamos é algo que produza um certo prazer de um determinado tipo. [...] Deseja-se, então, uma torta de maçã – uma boa torta de maçã, feita com maçãs frescas, com uma crosta moderadamente leve e um pouco curta, nem muito doce nem muito azeda etc. Mas não se trata de qualquer torta de maçã em particular, pois ela deve ser feita para a ocasião e a única particularidade é que ela deve ser feita e consumida hoje. Para isso, são necessárias maçãs; e, lembrando-se de que há um barril de maçãs no porão, a cozinheira vai até o porão e pega as maçãs que estiverem mais altas e à mão. Esse é um exemplo de como seguir uma regra geral. Ela é orientada a pegar as maçãs. Muitas vezes ela viu coisas que eram chamadas de maçãs e notou sua qualidade comum. Agora ela sabe como encontrar essas coisas e, desde que sejam boas e finas, qualquer maçã serve. O que ela deseja é algo de uma determinada qualidade; o que ela precisa pegar é esta ou aquela maçã em particular. [...] Quando vai ao porão buscar as maçãs, ela pega qualquer tigela ou cesta que estiver à mão, sem se importar com qual, desde que tenha um determinado tamanho, esteja limpa e tenha outras qualidades, mas depois de selecioná-la, ela pretende colocar algumas maçãs naquela tigela específica. Ela pega qualquer maçã que esteja à mão e pareça boa, mas, depois de pegá-la, pretende fazer uma torta com essas maçãs. Se por acaso vir outras na cozinha, ao voltar do porão, ela não as usará na torta, a menos que por algum motivo mude de ideia. Ao longo de todo o processo, ela persegue uma ideia ou um sonho sem nenhuma particularidade (CP 1.341, c. 1895).

A semiótica de Peirce nos traz outros insights sobre a natureza semiótica de um algoritmo. Como todos os signos, o algoritmo representa um objeto e tem o potencial de criar um interpretante. Que seria o objeto do legisigno de um algoritmo? O objeto representado pelo algoritmo na mão da cozinheira consiste no complexo real dos ingredientes da receita (maçãs, farinha, açúcar etc.), das regras culinárias estabelecidas pela autora anônima da receita, da experiência adquirida pela cozinheira no decorrer da sua carreira e da realidade concreta da cozinha (fogão, mesa, armários etc.) entre outros. O objeto determina o fazer profissional da cozinheira no sentido de que ele representa os pormenores culinários que impõe restrições reais e concretos a esse fazer. Ela tem que escolher maçãs e não peras, as maçãs têm que ser frescas, a quantidade dos ingredientes tem que estar correta etc.



O interpretante do algoritmo consiste na execução da receita pela cozinheira nas etapas descritas acima, inclusive o bolo e a satisfação na mesa da dona da casa. Todo algoritmo representa um objeto e serve para produzir um interpretante. O objeto do algoritmo de reconhecimento facial que substitui o porteiro consiste no conjunto das fotografias gravadas e reconhecidas pela câmera, do conjunto das fotografias gravadas, e a regra que só os moradores fotografados devem ser admitidos. O interpretante é simples: admitir a pessoa ou transmitir a mensagem de que a pessoa fotografada é desconhecida.

6 O algoritmo como agente cognitivo

Com Peirce podemos também mostrar que um algoritmo, como qualquer outro legissigno, é um agente semiótico ou cognitivo, o que significa, que o legissigno é um agente inteligente (Santaella, 2023, p. 98-106). A agência semiótica de um símbolo e de outros legissignos consiste na realização do seu potencial de modificar o pensamento e a ação dos seus usuários sob a influência do signo (Nöth, 2009). O legissigno tem o propósito de exercer o seu potencial de representar o seu objeto (e não outra coisa) e de criar um interpretante adequado. O legissigno tem este propósito, porque se não conseguisse ser interpretado conforme o seu propósito de representar o objeto e de criar um interpretante adequado, perderia o seu direito de existência e a sua existência mesmo. Os usuários iam abandoná-lo e o símbolo desperçia.

O algoritmo da receita exerce o poder sobre cozinheira de agir conforme as suas instruções, e se o algoritmo presta, o interpretante, em última instância o bolo pronto, vai satisfazer os comensais e trazer elogios para a cozinheira. Se a receita era ruim, ela será abandonada e nunca mais usada. Não só o algoritmo de uma receita, mas também qualquer outro algoritmo é um agente semiótico, como explicou Peirce. Como qualquer outra regra geral, um algoritmo “significa por seu poder de determinar outro signo” (EP 2, p. 317), isto é, pelos interpretantes e efeitos que cria. Assim, o algoritmo não é só um instrumento daqueles que o conceberam para influenciar as consumidoras. Ele mesmo adquiriu uma autonomia semiótica própria, e age com este potencial na rede semiótica. Essa é a lição, provocativa para os antropocentristas, que Peirce nos ensina.

Quando o agente que executa as regras de um plano geral é uma cozinheira ou um outro ser humano, todos concordam que se trata de agência humana, mas quando a cozinheira é substituída por uma máquina inteligente numa fábrica de bolos, a agência é atribuída ao dono da maquinaria e os engenheiros que programaram as máquinas. Isto é, quando um algoritmo determina a ação de um ser humano, o ser humano que concebeu o algoritmo é tido como o agente cognitivo, mas quando o algoritmo determina uma máquina a cozinhar o mesmo bolo, a máquina e o algoritmo



que a controla não são mais reconhecidos como agentes e a ação cognitiva é atribuída ao engenheiro e autor do algoritmo. Mas argumentar assim significa aplicar dois pesos e duas medidas (Nöth, 2001). Se tanto a cozinheira como a máquina da padaria industrial seguem algoritmos, que são da autoria de outros, por que só a primeira e não a segunda pode ser um agente semiótico?

Um argumento clássico baseado na lógica das duas medidas era o do quarto chinês, com o qual Searle (1980) quis provar que uma máquina controlada por um algoritmo não pode ser um agente cognitivo com a capacidade de pensar. Para Searle, a prova dessa incapacidade era que uma máquina, mesmo se pode fornecer traduções corretas, nem pode ser considerada a *entender* aquilo que traduz, nem o produto da tradução. Searle compara a maneira como a máquina funciona com o comportamento de um zumbi dentro de um quarto fechado, que nem fala nem entende chinês, mas, nem por isso, consegue produzir traduções corretas simplesmente porque segue as regras de um algoritmo, executando-as sem entendê-las. Disso, Searle conclui que máquinas não são agentes semióticos, porque não entendem nem o que traduzem nem o que produzem como tradução. Só a tradutora humana, que domina as duas línguas, pode ser considerada um agente cognitivo.

Sem dúvida, a tradutora humana entende tanto o que ela traduz como a sua tradução, mas o que significa entender um texto? Quem entende uma mensagem sabe o que ela significa, mas dizer que entender signos é conhecer os seus significados significa substituir o enigma do conhecimento pelo enigma da significação. Conhecer signos e os seus significados são desempenhos igualmente enigmáticos.

Para Peirce, entender um signo significa interpretá-lo, produzir um interpretante. Jakobson (1973) usou esta doutrina semiótica para resolver o enigma linguístico do significado (*meaning*) pela doutrina de que o significado de uma mensagem é a sua tradução, ou para outros signos verbais da mesma, ou de uma outra língua: "Para o linguista como para o usuário comum das palavras, o significado de um signo linguístico não é mais que a sua tradução por um outro signo que lhe pode ser substituído, especialmente um signo 'no qual ele se ache desenvolvido de modo mais completo', como insistentemente afirmou Peirce, o mais profundo investigador da essência dos signos" declarou Jakobson (1973, p. 64).

Em conclusão, a ética da pesquisa em processos cognitivos exige reconhecer que os algoritmos são agentes semióticos tanto quando determinam pensamentos e ações humanas como quando controlam processos inteligentes em máquinas. É medir com duas medidas de argumentar que os algoritmos não trabalham com inteligência quando realizem os mesmos trabalhos e prestarem os mesmos serviços.



7 Algoritmos bons e ruins e a ética da inteligência

Algoritmos computacionais têm sido desenvolvidos para executar trabalhos simples ou complexos, que substituem trabalhos humanos manuais ou mentais tediosos e repetitivos. Exemplos de algoritmos simples são as tarefas de pôr números desordenados em ordem crescente, pôr palavras em ordem alfabética, fazer o inventário do estoque de uma loja ou determinar o valor do estoque na base dos códigos dos artigos e os preços de tabela (Chaudhuri, 2020, p. 13). Algoritmos como esses são imunes a qualquer crítica porque as vantagens que eles trazem e a eficiência e precisão com os quais eles operam supera a de funcionários humanos. Eles também costumam ser chamados de algoritmos bons em contraste com os algoritmos ruins.

Algoritmos ruins justificam o seu nome não por ser um tipo específico de algoritmo, mas porque a sua programação foi mal concebida ou porque eles são usados sem as devidas considerações éticas. Entre os algoritmos bons e ruins existem os algoritmos que podem servir finalidades eticamente justificadas ou não.

A mídia está repleta de relatos sobre violações de princípios éticos por máquinas inteligentes e algoritmos que reforçavam preconceitos, resultavam em discriminação, tomavam decisões injustas ao negar pedidos de empréstimos, ou que impediram contratações de pessoal qualificado em empresas (Coeckelbergh, 2023). Felizmente, quando a ação cognitiva de algoritmos é descoberta como preconceituosa ou injusta pelos usuários, substituí-los por algoritmos sem as deficiências descobertas é mais fácil do que substituir preconceitos nas cabeças humanas e injustiças cometidas por humanos.

Infelizmente sabemos que inteligência pode ser guiada tanto por princípios éticos como por princípios antiéticos, ou até criminosos. Os perigos ligados ao uso da inteligência, com objetivos que violam princípios éticos são os mesmos que os princípios com os quais criminosos ou aproveitadores têm usado a sua inteligência humana. Peirce os reconhecia quando, em 1892, descreveu os perigos da inteligência humana na era do capitalismo selvagem assim:

O século XIX está agora afundando rapidamente no túmulo, e todos nós começamos a rever seus feitos e a pensar em que caráter ele está destinado a ter em comparação com outros séculos nas mentes dos futuros historiadores. Acho que ele será chamado de Século Econômico [...]. Bem, a economia política também tem sua fórmula de redenção. Ela é a seguinte: A inteligência a serviço da ganância garante os preços mais justos, os contratos mais justos, a conduta mais esclarecida em todas as negociações entre os homens e leva ao summum bonum, comida em abundância e conforto perfeito. Alimento para quem? Ora, para o ganancioso senhor da inteligência. Não quero dizer que essa seja uma das conclusões legítimas da economia política, cujo caráter científico reconheço plenamente. Mas o estudo de doutrinas, elas próprias verdadeiras, muitas vezes encorajará temporariamente generalizações



extremamente falsas, como o estudo da física encorajou o necessitarismo. O que eu digo, então, é que a grande atenção dada às questões econômicas durante o nosso século induziu a um exagero dos efeitos benéficos da ganância e dos resultados infelizes do sentimento, até que resultou em uma filosofia que, inconscientemente, chega a isso, que a ganância é o grande agente na elevação da raça humana e na evolução do universo (CP 6.290).

Referências

BOSTROM, Nick. **Superintelligence**: Paths, dangers, strategies. Oxford: Oxford University Press, 2014.

CHAUDHURI, Akmal B. **Flowchart and algorithm basics**. Dulles, VA: Mercury, 2020.

COECKELBERGH, Mark. **Ética na inteligência artificial**. Tradução: Clarisse de Sousa *et al.* São Paulo: Ubu, 2023 [2020].

CROSSLEY, John N.; HENRY, Alan S. Thus spake al-Khwārizmī: a translation of the text of Cambridge University Library Ms. li.vi.5. **Historia Mathematica**, Amsterdam, v. 17, n. 2, p. 103-131. 1990. DOI 10.1016/0315-0860(90)90048-I. Disponível em: <https://www.sciencedirect.com/science/article/pii/031508609090048I?via%3Dihub>. Acesso em: 10 ago. 2024.

DONALD, Merlin. **Origins of the modern mind**: the stages in the evolution of culture and cognitions. Cambridge, MA: Harvard University Press, 1991.

ECO, Umberto. **Os limites da interpretação**. 2. ed. São Paulo: Perspectiva, 2010.

JAKOBSON, Roman. Aspectos linguísticos da tradução. In: JAKOBSON, R. **Linguística e comunicação**. São Paulo: Cultrix, 1973 [1959], p. 63-72.

KAUFMAN, Dora. **Desmistificando a inteligência artificial**. São Paulo: Autêntica, 2022.

KAUFMAN, Dora; SANTAELLA, Lucia. O papel dos algoritmos de inteligência artificial nas redes sociais. **Revista Famecos – Mídia, Cultura e Tecnologia**, Porto Alegre, v. 27, p. 1-10, jan./dez. 2020. Disponível em: <https://revistaseletronicas.pucrs.br/revistafamecos/article/view/34074>. Acesso em: 10 ago. 2024.

KURZWEIL, Raymond. **The singularity is near**: When humans transcend biology. New York: Viking, 2005.

KURZWEIL, Raymond. Ray Kurzweil e o mundo que nos espera: uma entrevista com o inventor e futurólogo americano. **Piauí**, Rio de Janeiro, n. 43, abr. 2010. Disponível em:



<https://piaui.folha.uol.com.br/materia/ray-kurzweil-e-o-mundo-que-nos-espera/>. Acesso: 13 maio 2024.

METZINGER, Thomas; WINDT, Jennifer M. (ed.). **Open mind**: philosophy and the mind sciences in the 21st century. Cambridge, MA: MIT Press, 2016. 2 v.

MURPHY PAUL, Annie. **The extended mind**: the power of thinking outside the brain. Boston, MA: Mariner, 2021.

NÖTH, Winfried. Máquinas semióticas. **Galáxia**, São Paulo, v. 1, n. 51-73, 2001. Disponível em: <https://revistas.pucsp.br/index.php/galaxia/article/view/1057>. Acesso em: 15 maio 2024.

NÖTH, Winfried. On the instrumentality and semiotic agency of signs, tools, and intelligent machines. **Cybernetics & Human Knowing**, Exeter, v. 16, n.3-4, p. 11-36, 2009.

NÖTH, Winfried. Análise de discurso com Peirce: interpretar, raciocinar e o discurso como argumento. **Intexto**, Porto Alegre, v. 37, p. 34-46, 2016. Disponível em: <https://seer.ufrgs.br/index.php/intexto/article/view/68252>. Acesso em: 03 ago. 2024.

NÖTH, Winfried; SANTAELLA, Lucia. **Introdução à semiótica**. São Paulo: Paulus, 2017.

PEIRCE, Charles S. **Collected Papers**. Cambridge, MA: Harvard University Press, 1931-58. 8 v.

PEIRCE, Charles S. **The Essential Peirce**. Bloomington, IN: Indiana University Press, ed. Peirce Edition Project, 1998. v. 2.

SANTAELLA, Lucia. **A inteligência artificial é inteligente?** São Paulo: Edições 70, 2023.

SANTAELLA, Lucia; KAUFMAN, Dora. A inteligência artificial generativa como quarta ferida narcísica do humano. **Matrizes**, São Paulo, v. 18, n. 1, p. 37-53, 2024. Disponível em: <https://www.revistas.usp.br/matrizes/article/view/223351/204174>. Acesso em: 01 ago. 2024.

SEARLE, John. Minds, brains, and programs. **Behavioral and Brain Sciences**, Cambridge, v. 3, p. 417-457, 1980. Disponível em: <https://web-archive.southampton.ac.uk/cogprints.org/7150/1/10.1.1.83.5248.pdf>. Acesso em: 01 ago. 2024.

Revisão por:

Levy Henrique Bittencourt

E-mail: nikolai.streisky@gmail.com