



A IA VAI A JULGAMENTO: AS IAS PODEM SER REALMENTE INTELIGENTES? *

AI goes on trial: can AIs really be intelligent?

Daniel De Luca Silveira de Noronha **

Bruno Dornelas ***

Resumo: Alan Turing previu que o uso das palavras se modificaria a tal ponto que hoje poder-se-ia falar sem constrangimento a respeito de máquinas que pensam. Assim, a questão “As máquinas podem pensar?” seria considerada insignificante. No entanto, na literatura atual, chega-se a afirmar que a aplicação não metafórica de predicados como *inteligente* e *pensar* a máquinas e a sistemas é sempre ilegítima ou mesmo absurda, pois carece de sentido. Em contrapartida, tenta-se, não apenas justificar a aplicação desses predicados, mas provar a existência da própria IA. O objetivo geral deste artigo, portanto, é lançar alguma luz sobre essa “Babel Tecnológica”, recorrendo a uma consciência linguística mais plena. Assim, emprega-se a metodologia filosófica de Garth Hallett a fim de dissolver o entendimento padrão da questão. Dessa forma, abre-se espaço para a predição adequada no contexto da IA e da robótica. Em seguida, por intermédio do ‘Princípio de Semelhança Relativa’ (PSR) e da ‘Norma de Correspondência Linguística’ (NCL) hallettianas, que aquilo que pode ser rotulado como IA não apenas pode receber o veredito “inteligente”, mas também os vereditos “não-inteligente” e “não-decيدido”. A adoção dessa posição permite contribuir para

* Artigo recebido em 02.10.2025 e aprovado para publicação em 20.11.2025.

** Doutor em Filosofia pela UFMG (2013). Professor assistente na Faculdade Jesuíta de Filosofia e Teologia (FAJE). ORCID: <https://orcid.org/0000-0002-4882-5023>. E-mail: deluca.11@gmail.com.

*** Mestre em Filosofia e doutorando em Filosofia pela FAJE. ORCID: <https://orcid.org/0009-0008-2736-2657>. E-mail: brunoluizdornelas@outlook.com.

o esclarecimento de questões disputadas na IA, que abrigam sérias implicações jurídicas e morais, como as de 'autoria' e 'liberdade'.

Palavras-chave: Inteligência Artificial. Filosofia da Linguagem. Garth Hallett. Princípio de Semelhança Relativa. Norma de Correspondência Linguística.

Abstract: Alan Turing predicted that the use of words would change to such an extent that today one could speak without embarrassment about machines that think. Thus, the question "Can machines think?" would be considered insignificant. However, in current literature, it is even stated that the non-metaphorical application of predicates such as intelligent and thinking to machines and systems is always illegitimate or even absurd, as it lacks meaning. In contrast, there are attempts not only to justify the application of these predicates, but also to prove the existence of AI itself. The general objective of this paper, therefore, is to shed some light on this "Technological Babel" by resorting to a fuller linguistic awareness. Thus, Garth Hallett's philosophical methodology is employed in order to dissolve the standard understanding of the issue. This, in turn, opens up space for adequate predication in the context of AI and robotics. Next, through the Hallettian 'Principle of Relative Similarity' (PRS) and the 'Norm of Linguistic Correspondence' (NLC), it is concluded that what can be labeled as AI can not only receive the verdict "intelligent", but also the verdicts "not-intelligent" and "undecided". Adopting this position helps elucidate disputed issues in AI, which carry serious legal and moral implications, such as those of 'authorship' and 'freedom'.

Keywords: Artificial Intelligence. Philosophy of Language. Garth Hallett. Principle of Relative Similarity. Linguistic Correspondence Norm.

Introdução

Em seu artigo seminal *Computing Machinery and Intelligence*, Alan Turing previu que o uso das palavras se modificaria a tal ponto que hoje poder-se-ia falar sem constrangimento a respeito de máquinas que pensam. Assim, a questão "As máquinas podem pensar?" seria considerada insignificante.¹ Normalmente, uma resposta a essa questão exigiria primeiro uma definição do significado de *máquina* e *pensamento*. Porém, segundo Turing, se as definições refletem o uso costumeiro das palavras, e se o significado de uma palavra deve ser examinado por meio de seu uso comum, então o significado deverá ser buscado, perigosamente, na pesquisa estatística.² Por essa razão, ele propõe substituir a procura por uma definição do sig-

¹ TURING, A. *Computing Machinery and Intelligence*, in *Mind*, New Series, v. 59, n. 236, Oct. 1950, p. 433-460.

² TURING, *Computing Machinery and Intelligence*, p. 433.

nificado pelo *Jogo da Imitação* – mais conhecido como *Teste de Turing*. Esse teste propôs uma definição operacional de *pensamento* e *inteligência* e foi o responsável, em alguma medida, pelo desenvolvimento do campo da IA.³ Contudo, além de sua validade ser considerada controversa, a crescente percepção de quão difícil seria para uma máquina obter sucesso nele levou a uma mudança na discussão em torno da IA. Assim, apesar de a previsão de Turing ter se confirmado no âmbito da cultura, na literatura acadêmica atual, chega-se a afirmar que a aplicação não metafórica de predicados como *inteligente* e *pensar* a máquinas e a sistemas é sempre ilegítima ou mesmo absurda, pois carece de sentido. Em contrapartida, busca-se não apenas justificar a aplicação desses predicados, mas também provar a existência da própria IA.

O objetivo geral deste artigo, portanto, é lançar alguma luz sobre essa “Babel Tecnológica”, recorrendo a uma consciência linguística mais plena. Para esse fim, emprega-se a metodologia filosófica e o arcabouço conceitual desenvolvidos por Garth Hallett, com o objetivo de dissolver o entendimento padrão da questão. Dessa forma, abre-se espaço para a predicação adequada no contexto da IA e da robótica. Como uma espécie de matéria-prima a servir de pano de fundo para a discussão, é utilizada amplamente a obra *Neuroscience and Philosophy: Brain, Mind, and Language* de 2007.⁴ Ela apresenta os debates de Max Bennett e Peter Hacker contra John Searle e também contra Daniel Dennett, sobre uma vasta gama de tópicos entrelaçados à IA.⁵ O desenvolvimento do artigo concentrar-se-á, sobretudo, na dialética entre Hacker e Dennett, apresentando, em primeiro lugar, a posição do último, bem como algumas objeções centrais a ela. Em segundo lugar, será examinada a posição do primeiro, juntamente com algumas objeções centrais a ela. Por fim, será empregada a metodologia filosófica de Hallett, a fim de apontar uma solução sintética para as respectivas questões. Podemos aplicar legitimamente – não metaforicamente – termos como *inteligente* e *pensar* a máquinas e a sistemas? A ‘Autoridade da Língua’ pode nos fornecer a base comum (como *medium*) de superação da “Babel Tecnológica”?

³ FRENCH, R. The Turing Test: The First 50 Years, in *Trends in Cognitive Sciences*, v. 4, n. 3, 2000, p. 115-122, aqui p. 116-117.

⁴ BENNETT, M.; DENNETT, D.; HACKER, P.; SEARLE, J. *Neuroscience and Philosophy: Brain, Mind, and Language*. New York: Columbia University, 2007.

⁵ Essa obra se torna especialmente relevante para nossa discussão devido à clareza e à precisão com que os debatedores são obrigados a apresentar suas posições e pressuposições, em virtude das demandas impostas por seus opositores. A obra conta ainda com o valioso comentário de Daniel Robinson, que a qualifica como “instrutivamente dialética”. BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 174.

1. Um terno apertado demais: a semelhança essencial

1.1 Negando a inteligência: a *Forma de Vida*

Peter Hacker é conhecido por trafegar pela rodovia pavimentada por Wittgenstein – ou como diria Dennett, por seguir fielmente a doutrina de São Ludwig. Assim, sua contribuição para a ciência consiste essencialmente no esclarecimento das estruturas conceituais sobre as quais a ciência trabalha.⁶ A tarefa da filosofia é, sob este prisma, eminentemente conceitual. Ela se preocupa com nossa forma de representação e não com verdades empíricas. Por conseguinte, a filosofia situa-se em um registro lógico distinto da ciência. As “verdades conceituais” reveladas pela investigação filosófica são constitutivas das condições de sentido que, por sua vez, são pré-condição da verdade e da descrição científica correta. Como determinantes do sentido, essas verdades delineiam o espaço lógico que abriga os fatos, bem como traçam a linha que separa o sentido do sem sentido (absurdo). Como verdade conceitual, a proposição “O vermelho é mais escuro que o rosa” assume um papel implicitamente normativo (gramatical) e não descritivo. Tais declarações constituem parcialmente o significado de suas expressões constituintes, atribuindo propriedades ou relações internas a elas. Portanto, são essas relações lógicas de implicação e pressuposição, compatibilidade e exclusão, que caracterizam o uso dos conceitos e devem ser observadas a fim de evitar combinações de palavras que resultem em absurdo.⁷ Fica evidente a aderência de Hacker à máxima: “a significação de uma palavra é seu uso na linguagem”. Esse breve resumo esboça o quadro a partir do qual Hacker aborda questões relativas à IA. Contudo, ele se guia especialmente por uma passagem-chave das *Investigações*: “somente de um ser humano e do que *se assemelha* (*se comporta como*) a um ser humano vivo é que se pode dizer: ele tem sensações [pensamentos; inteligência]; ele vê, é cego; ouve, é surdo; é consciente ou inconsciente”.⁸

Essa passagem das *Investigações* elabora as condições necessárias e suficientes de aplicação para predicados psicológicos. Em várias oportunidades, Hacker expõe sua interpretação desse trecho. Todavia, a demanda por maior clareza e precisão em sua tréplica no debate com Searle torna a Nota 18, encontrada entre as páginas 207 e 208 de *Neuroscience*, particularmente elucidativa. De acordo com essa nota, a aplicação de tais predicados observa tanto bases evidenciais indutivas quanto não indutivas (criteriais). O principal elemento que assegura a aplicação dos termos psicológicos é o critério comportamental, havendo, pois, uma conexão interna entre

⁶ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 127.

⁷ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 128-130.

⁸ WITTGENSTEIN, L. *Investigações Filosóficas*. São Paulo: Abril Cultural, 1975, p. 104.

conceito e comportamento. Esse critério é, contudo, não indutivo porque sua natureza é lógica. Assim, apesar de esses critérios não esgotarem o significado dos termos psicológicos, eles são parcialmente constitutivos dos seus significados. Portanto, as evidências indutivas têm de pressupor essa base lógica que as fundamenta. É ela que determina como e para onde se deve olhar indutivamente; a saber, para comportamentos sofisticados que exibem capacidades complexas em circunstâncias apropriadas. Como Hacker observa, isso não constitui behaviorismo como se poderia supor, pois os critérios diferem das condições de verdade, assim como as condições de verdade das declarações pelas quais se atribuem predicados psicológicos são distintas de sua verdade. Um animal, por exemplo, pode estar com dor sem demonstrar comportamento naturalmente equivalente, assim como pode não estar com dor e, ainda assim, exibir comportamento típico de dor. Depois de apresentar a base teórica por meio da qual Hacker aborda a questão da aplicabilidade de “inteligência” e “pensamento” a IAs e robôs, torna-se possível considerar mais profundamente sua posição.

Hacker acredita com Wittgenstein, que a questão “É possível para uma máquina pensar?” é semelhante à questão “O número 3 tem uma cor?”,⁹ e que “a frase ‘Uma máquina pensa (percebe, deseja)’ parece de alguma forma sem sentido”.¹⁰ Isto é, não apenas a resposta positiva à questão parece destituída de sentido, mas o mesmo se aplica a uma resposta negativa, uma vez que, como se sugere, a própria questão da qual ambas decorrem carece de sentido. Mas por qual razão tanto as respostas quanto a questão, se consideradas em sua literalidade, formam uma combinação absurda de palavras?¹¹

⁹ WITTGENSTEIN, L. *The Blue and Brown Books*. Oxford: Basil Blackwell, 1958, p. 47. Embora o fenômeno sinestésico grafema-cor associe cores a números, o número em si mesmo não possui cor, ao contrário de um objeto físico.

¹⁰ No *Blue Book*, Wittgenstein afirma que a questão “Uma máquina pode pensar?” é análoga a “Uma máquina pode ter dor de dente?” e conclui que o uso de *pode* na sentença “Uma máquina não pode ter dor de dente”, denota uma impossibilidade lógica, e não experiencial – como nunca ter visto uma máquina que sentisse dor de dente. Em seguida, ele chama a atenção para a questão central que será apresentada adiante: “Qual é a relação entre o pensamento (ou dor de dente) e o sujeito que pensa, tem dor de dente, etc.?” (WITTGENSTEIN, *The Blue and Brown Books*, p. 16). Segundo Oswald Hanfling, em tais passagens Wittgenstein emprega ‘máquina’ no sentido costumeiro, isto é, entendido como um dispositivo que não se comporta de modo semelhante ao ser humano. Por outro lado, uma criatura artificial que satisfizesse tais condições não poderia ser considerada uma mera máquina, mas sim uma “pessoa artificial”. HANFLING, O. *Wittgenstein and the Human Form of Life*. New York: Routledge, 2002, p. 148.

¹¹ Entre as várias opções consideradas por Hacker para explicar a fala comum dos termos psicológicos aplicados aos computadores – usos metonímicos, analógicos, metafóricos, figurativos e secundários (HACKER, P. *Men, Minds and Machines*. In: HACKER, P. *Wittgenstein, Meaning and Mind*. Cambridge, MA: Blackwell, 1990, p. 89 e ss.; cf. BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 152 e ss.) – ele observa que os engenheiros aplicam o vocabulário psicológico com naturalidade, de forma não figurativa, sem precisar recorrer a paráfrases literais, por tratarem essas expressões como *metáforas mortas*. Por outro lado, ele não elucida ou oferece indícios de que essa explicação possa ser aplicada ao restante da fala comum e atual sobre IA.

O ponto principal, para Hacker, é que IAs e robôs não possuem formas de vida; e um vocabulário psicológico e seu uso só são possíveis no fluxo da vida humana. Cada expressão ou termo psicológico constitui um nó numa trama conceitual complexa da qual não podem se afastar; não podem, pois, ser usados uma única vez. O padrão de uso de cada termo se caracteriza por manter determinada posição e desempenhar, a partir dela, uma função na vida humana, cuja possibilidade é garantida pela série de conexões lógico-conceituais existentes entre conceitos relacionados. Portanto, “[o] conceito de dor está ligado não apenas ao comportamento característico da dor [...], mas também à pena e comiseração, medo e ansiedade, crueldade e misericórdia. Está entrelaçado com atitudes especiais em relação aos doentes (tanto cuidado quanto desdém), com *conceitos* de saúde e bem-estar e, portanto, com o do bem de um ser”.¹² O comportamento de dor está entrelaçado com “comportamentos”, “atitudes especiais” e com outros “conceitos” na trama complexa da vida humana. Todos esses elementos, por seu turno, se entrelaçam ainda mais, formando uma trama muito mais densa. Essa característica conceitual constitui um problema para o discurso da IA, na medida em que seu emprego de termos psicológicos problemáticos se encontra abstraído de todo um conjunto de conceitos fundamentais que estabelece as condições para os critérios de aplicação de tais termos. Por consequência, seu uso não tem sentido. Segundo Hacker, critérios para aplicação de conceitos como “opinar”, “imaginar”, “supor”, pressupõem a aplicação “de desejo e vontade, motivo e intenção, satisfação e decepção, etc.”.¹³ Os primeiros sempre pressupõem os últimos.

Portanto, conforme Hacker, uma IA ou robô não pode exemplificar os critérios comportamentais (inseparáveis das circunstâncias da vida) necessários para indicar pensamento.¹⁴ Todavia, como o próprio Hacker nota, o “Teste de Turing” aparentemente contornaria essas objeções gramaticais, contanto que um computador fosse capaz de exibir respostas indistinguíveis daquelas que um ser humano poderia dar. Logo, ele se comportaria como um ser humano inteligente. Porém, ao invés de exibirem comportamento, os computadores estariam exibindo apenas o produto do comportamento humano. Em sua visão, os computadores nem sequer calculam de fato.¹⁵

¹² HACKER, Men, Minds and Machines, p. 98.

¹³ HACKER, Men, Minds and Machines, p. 99.

¹⁴ É interessante notar que Hacker está alinhado, a contragosto, a muito do que se tem dito sobre a construção de uma IA forte. Tal alinhamento pode ser observado ao final de seu artigo de 1990: “Se, por alguma razão estranha e perversa, desejássemos criar artificialmente uma coisa pensante, em oposição a um dispositivo que nos poupe o trabalho de pensar, teríamos que começar, por assim dizer, com a animalidade, não com a racionalidade. Desejo e sofrimento são as raízes do pensamento, não a computação mecânica” (HACKER, Men, Minds and Machines, p. 111).

¹⁵ Pelas razões que se observam abaixo, essa conclusão vale não somente para a IA generativa, mas também para IA clássica, cujo programa internaliza regras matemáticas tradicionais e não opera por meio de estatística preditiva.

Dizer que computadores calculam literalmente é semelhante a dizer que ábacos e calculadoras mecânicas também o fazem. Isso porque as sentenças “Máquinas não podem seguir regras” e “Não se pode seguir uma regra mecanicamente” representam proposições gramaticais.¹⁶ Isso se explica tanto pela constatação de que computadores modernos não são diferentes, em princípio, de outras máquinas, quanto pelo fato de que estas não podem seguir regras, mas apenas operar (passar por processos) em conformidade com sua regularidade. O seguimento de uma regra, por sua vez, só pode ocorrer num contexto comportamental normativo, isto é, o nexos das transformações dos signos matemáticos deve ser lógico-normativo e não causal.¹⁷ O ato de seguir uma regra só pode ocorrer no interior de práticas complexas, em que estão conectadas diversas “atividades normativas reais e potenciais [...] de justificar, perceber erros e corrigi-los com referência à regra relevante, criticar desvios da regra, e, se solicitado, explicar uma ação como estando em conformidade com a regra ou ensinar a outros o que conta como seguir a regra”.¹⁸ Assim, afirmar que computadores calculam não constitui um mero erro empírico, mas um erro lógico-conceitual, visto que o discurso literal segundo o qual máquinas seguem regras sequer faz sentido.

1.2 A tese do dilema: a excessiva rigidez do significado

Apresentada a posição de Hacker, deve-se agora ser observado que a densidade da trama conceitual – fundamentada na forma de vida humana e estabelecida como critério para aplicação de termos psicológicos – restringe severamente o uso desses termos. Toda a discussão conceitual hackeriana sugere que os aspectos da *semelhança* e do *comportar-se como*, presentes na passagem-chave que orienta o uso destas expressões, recebem

¹⁶ HACKER, Men, Minds and Machines, p. 106.

¹⁷ Pode-se dizer que, enquanto Hacker nega que o computador calcule, porque não pensa, Hauser sustenta que uma calculadora de bolso pensa, justamente porque calcula. Cf. HAUSER, L. The Sense of Thinking? Reply to Rapaport’s Comments, in *Minds and Machines*, v. 3, n. 1, 1993, p. 21-29; HAUSER, L. Why Isn’t My Pocket Calculator a Thinking Thing?, in *Minds and Machines*, v. 3, n. 1, 1993, p. 3-10. Tais conclusões opostas constituem duas faces da mesma moeda, cuja inscrição poderia ser: ‘Calcular é pensar!’. Assim, só pode haver duas alternativas excludentes: ou a calculadora não calcula, ou então ela pensa. Como a análise abaixo sugere, nenhum deles considera efetivamente a variedade daquilo que se poderia chamar *calcular*.

¹⁸ HACKER, Men, Minds and Machines, p. 106-107. Como observa Glock, uma objeção pode ser levantada ao assumir que computadores suficientemente complexos podem ser programados para exibir tal comportamento. No entanto, ele concorda com Hacker ao sustentar que, para Wittgenstein, seguir regras só é possível para aquele que persegue objetivos, tem vontade e possa se interessar (GLOCK, H-J. *A Wittgenstein Dictionary*. Cambridge, MA: Wiley-Blackwell, 1996, p. 158-159). Se admitirmos, com Hacker, que mesmo que o computador possa reproduzir tal comportamento ainda assim não haveria comportamento de fato, então, poder-se-ia perguntar se a posição não estaria sob a égide de um essencialismo que sub-repticiamente age através do próprio conceito de *comportamento*.

uma interpretação excessivamente rígida. Por conseguinte, a semelhança de comportamento, estabelecida como critério de aplicação dos termos psicológicos no § 281 das *Investigações*, é considerada de maneira demasiadamente estreita.

No entanto, essa não parece ser a forma mais adequada de interpretar o § 281. Essa passagem, ao contrário, supõe que a *semelhança* exigida em relação ao critério de comportamento é estabelecida dentro de limites amplos.¹⁹ Tal amplitude é exemplificada logo à frente no § 284a:

Olhe para uma pedra e imagine que ela tem sensações. — Alguém diz a si mesmo: como alguém poderia sequer ter a ideia de atribuir uma *sensação* a uma *coisa*? Poderia muito bem atribuí-la a um número! — E agora olhe para uma mosca se contorcendo e imediatamente essas dificuldades desaparecem; a dor parece conseguir firmar-se ali, onde antes tudo era, por assim dizer, demasiado liso para ela.²⁰

Para exemplificar a aplicação do conceito de *dor*, Wittgenstein não precisou saltar do inanimado para um ser muito superior. O comportamento de uma mosca foi suficientemente semelhante para satisfazer o critério conceitual. Embora não especifique em que medida, Hacker afirma que o conceito de *dor* começa a escorregar para criaturas inferiores, aquelas que não são humanas ou mamíferos, e termina no inanimado.²¹ Ele não traça, contudo, nenhum limite entre o inanimado e os animais inferiores. No caso da mosca não se observa – ao menos em situações normais – feições de dor, gemidos de dor, sofrimento em seu olhar etc., mas, ainda assim a dificuldade de aplicação desaparece em sua ação de se contorcer. Todavia, imagine uma mosca se contorcendo e inquieta. Seja como for, esse comportamento guarda apenas uma semelhança bem distante com o comportamento de dor do ser humano. Ao que parece, a explicação hackeriana conflita com essa aceitação. Talvez até sua própria aceitação, mesmo que escorregadia, da aplicação de dor a seres inferiores seja conflitante, pois aqui a rede conceitual da forma de vida humana, base da semelhança, parece perder grande parte de sua força como fundamento criterial. Contudo, outra análise pode tornar a situação mais clara.

Concordamos com o Professor Dennett que muitas das crenças de uma criança são crenças em um sentido atenuado. A compreensão de uma menina sobre o conceito de médico pode ser defeituosa, mas ela dirá corretamente “Papai é médico” e responderá à pergunta “Onde está o médico?” dizendo “Lá dentro” (apontando para o consultório do papai). Assim, pode-se dizer que ela acredita, em um sentido atenuado, que seu pai é médico. Ela satisfaz, em seu comportamento verbal e dêitico, *alguns* dos critérios normais para

¹⁹ HALLETT, G. *A Companion to Wittgenstein's "Philosophical Investigations"*. Ithaca/London: Cornell University Press, 1977, p. 356.

²⁰ WITTGENSTEIN, *Investigações Filosóficas*, p. 105.

²¹ HACKER, *Men, Minds and Machines*, p. 99.

acreditar que seu pai é médico (mas também satisfaz alguns dos critérios para a falta dessa crença) [...] em seu comportamento verbal e dêitico, a criança pode manifestar crença *rudimentar*.²²

Este excerto é retirado da tréplica de Hacker em sua discussão com Dennett – inclusive o exemplo foi apresentado por este último contra Hacker – a respeito da questão de se o cérebro pode ou não realmente possuir crenças.²³ Dada a raridade de um exemplo “cinzento” como esse discutido por Hacker, tal caso mostra-se particularmente útil. Pode-se interpretar “crença em sentido atenuado” como “crença rudimentar”. Então, como essa crença rudimentar se caracteriza? Em primeiro lugar, é dito que a criança possui um conceito defeituoso, daí segue-se que ela acredita que seu pai é médico em um sentido atenuado. Entretanto, um conceito defeituoso deveria produzir uma crença defeituosa ou errônea não atenuada. É o conceito que é atenuado, não a crença. O conceito pode ser estreito, pouco informado, logo, diminuído. Mas a crença é diminuída? Ao ser perguntada se acredita realmente que seu pai é médico, se não possui alguma desconfiança quanto a essa afirmação, ou ainda se o médico ao qual se refere é mesmo seu pai, ela deveria, em virtude de sua crença atenuada, responder igualmente de forma atenuada – talvez acentuando sua hesitação ou sua desconfiança. Em suma, deveria dizer que não sabe se realmente acredita em tais fatos. E o mesmo deveria valer para o conceito de *pai*. Serão, então, todas as suas crenças atenuadas, assim como todos os seus conceitos defeituosos?

Se tal crer é defeituoso, então deve haver um crer exato, genuíno. Mas onde exatamente está fixado o limite? Será necessário apenas o conhecimento biológico, ou também o sociológico, psicológico e religioso, para empregar de maneira não defeituosa o conceito de *pai*. Mas tais conhecimentos devem ser adquiridos mínima ou completamente? Deve-se supor que a maior parte dos homens possui um conhecimento defeituoso do conceito de *matemática*? Como observa Hallett, uma criança que possui um conceito domina uma forma de vida, um jogo de linguagem. Ela não aprende os fatos sobre a constituição da água, mas sim o que é chamado de água depois de se familiarizar, em uma base mínima, com suas propriedades sensíveis. Ela aprende os usos mais básicos do conceito de *água*. Porém, o significado aprendido não constitui um conceito defeituoso de segunda categoria, mas possui um *uso mínimo* na língua. Os conceitos crescem em graus para pessoas e povos, “a ontogenia espelhando a filogenia”.²⁴ Assim,

²² BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 141.

²³ Para dificuldades relativas à atribuição de um tipo atenuado de crença ao cérebro, veja o Epílogo de D. Robinson, em BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 188.

²⁴ HALLETT, G. The Theoretical Content of Language, in *Gregorianum*. Roma: Pontificia Universita Gregoriana, v. 54, n. 2, 1973, p. 330.

os nativos não possuem um conceito defeituoso ou de segunda categoria de água e médico, apenas por não conhecerem a temperatura de fervura da água ou as técnicas da medicina moderna. Eles possuem um *significado mínimo* adequado para a forma de vida que partilham.

Em segundo lugar, se a criança verifica alguns critérios e contradiz outros, como decidir sobre a aplicação correta no caso – isto é, entre crença e não crença? De quais traços de comportamento pode-se abdicar e de quais não – onde traçar o limite, sob essa perspectiva? Não se pode fugir da determinação do que seria considerado um *comportamento suficientemente semelhante* para a aplicação correta do conceito de *crer*, simplesmente atribuindo-lhe defeito.²⁵ Parafraseando-se livremente a própria avaliação de Hacker sobre a postura intencional de Dennett, quando este fala que algumas partes do cérebro *meio que (sort of) acreditam, meio que decidem etc.*²⁶, pode-se concluir: “Hacker parece comprometido com a ideia de que uma criança crê defeituosamente que seu pai é um médico”. Tudo isso, presumivelmente, deve explicar defeituosamente o que os filósofos da linguagem pretendem explicar. Mas, se os explananda são pseudocrenças, protocrenças ou semicrenças (como Hacker sugere), na melhor das hipóteses, apenas fazem sentido defeituosamente e presumivelmente são defeituosamente verdadeiras. E como alguém poderia fazer inferências válidas a partir de tais premissas permanece mais do que apenas defeituosamente obscuro. Defeituosamente acreditar, pseudocrer, protocrer ou demicrer em algo fornece a uma criança uma razão para agir? Ou apenas uma razão defeituosa? – para uma ação igualmente defeituosa?²⁷

A análise de Hacker parece cometer uma simplificação essencialista com relação à significação das palavras.²⁸ Nas palavras de Hallett: “[s]e os significados das palavras fossem itens indivisíveis e invariáveis que tivessem

²⁵ Como mencionaremos abaixo (item 4.1), casos limítrofes genuínos não encontram apoio no uso, pois não podem satisfazer as condições de verdade e inteligibilidade dos termos. No entanto, a crença de uma criança não constitui um desses casos.

²⁶ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 87-88.

²⁷ Cf. BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 140. Michael Dummett distingue duas intuições a respeito da noção de *precisão*. A primeira é caracterizada precisamente pelas seguintes palavras: “temos uma simpatia intuitiva pela visão de Frege de que uma expressão vaga é aquela que, nessa medida, é *defeituosa em sentido*” (DUMMETT, M. *Frege's Philosophy of Language*. 2nd ed. Cambridge, MA: Harvard University Press, 1981, p. 647; ênfase adicionada). A intuição contrária sustenta que “a vagueza é uma característica indispensável da nossa linguagem” (DUMMETT, *Frege's Philosophy of Language*, p. 647). A primeira é defendida por Frege, enquanto a segunda é defendida por Wittgenstein. No entanto, o trecho de Hacker sobre o conceito defeituoso sugere que, pelo menos no contexto de nossa discussão, ele pode ser colocado ao lado de Frege contra Wittgenstein.

²⁸ Em comentário à discussão entre Hacker e Dennett a respeito dos pares conceituais ‘sentido’ e ‘absurdo’, ‘verdade’ e ‘falsidade’, Robinson observa: “Tomadas com talvez demasiada sobriedade, as classificações de Hacker aqui são ternos apertados demais para o uso diário” (BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 185). Não cabe aqui examinar o mérito do comentário de Robinson. Pretende-se apenas salientar que, se as

que ser conhecidos (aprendidos, lembrados, ensinados, etc.) por completo ou não fossem conhecidos de forma alguma, então o argumento poderia se sustentar”.²⁹ Esse trecho destina-se a combater o argumento de Waismann, que se opõe à generalidade da fórmula segundo a qual o significado de uma palavra é seu uso. No entanto, o raciocínio de Hallett se aplica, com igual ou maior êxito, à análise de Hacker. Waismann indaga se, “ao aprender o que é um número, é preciso também aprender o uso de frases como ‘um número de pessoas’, ‘um pequeno número’, ‘um número redondo’, etc.”.³⁰ As afirmações de Hacker sugerem uma resposta positiva a essa questão. A essa questão, segue-se outra: “Seria correto dizer que, se uma criança não está familiarizada com tais expressões, ela não sabe o que é um número?”.³¹ Mais uma vez, sugere-se, Hacker teria de responder afirmativamente. No entanto, como o próprio Waismann afirma corretamente em sua crítica, a adição de uma nova expressão, como por exemplo “Tempo é dinheiro”, não altera o significado da palavra *tempo*. A resposta de Waismann a ambas as questões é, portanto, negativa – única que parece realmente defensável. Em suma, somente em uma visão essencialista um conceito precisa ser adquirido por inteiro, ou seja, levando consigo toda a rede conceitual da qual faz parte e que constitui a condição para seu critério de aplicação.

Se a maior parte dos conceitos, em vez de adotar um padrão de precisão ou de essencialismo, seguem um padrão de semelhança, vale a pena ter em mente essa conseqüente e ampla variedade e flexibilidade resultantes. Considerando-se as crenças de um ser humano adulto, de uma criança, de uma pessoa com deficiência intelectual grave, de um cão, de animais com comportamentos menos complexos e, por fim, de um robô ou IA, percebe-se que não há uma *única* rede conceitual *comum* pressuposta em todas as aplicações do conceito. Observam-se parentescos, com algumas características comuns sendo mantidas, outras perdidas e novas surgindo. Diferentes funções são exercidas e novas habilidades são requeridas. Visto que as semelhanças podem se cruzar e se envolver de múltiplas formas, não haveria, a princípio, um limite pré-estabelecido por regras que impeça novas extensões de um conceito. E, se Hacker acompanha Wittgenstein aqui,

classificações são muito apertadas para o uso diário, então se tornam particulares ou enganosas, pois os termos adquirem sua classificação, isto é, sua significação e seu sentido, nos diversos contextos de usos diários. Compare-se o que foi dito acima a essa observação de Robinson. Hacker, contudo, não se vê como o gramático cômico a que proíbe novas extensões e exclama: “você não pode dizer isso!”, como o qualifica Dennett (*Ibid.*, p. 89). De sua própria perspectiva, limita-se a apontar incoerências conceituais nos cruzamentos dos sentidos. Desse modo, segundo ele, “[n]ão há nada de errado em falar sobre o pé de uma montanha – desde que não se pergunte se ele tem um sapato. Não há nada de errado em falar de comida não saudável – desde que não se pergunte quando ela recuperará sua saúde” (*Ibid.*, p. 154).

²⁹ HALLETT, G. *Essentialism: A Wittgensteinian Critique*. Albany: State University of New York Press, 1991, p. 59.

³⁰ WAISMANN, F. *Analytic-Synthetic II*, in *Analysis*, v. 11, n. 2, 1950, p. 29.

³¹ WAISMANN, *Analytic-Synthetic II*, p. 29.

terá de reconhecer que qualquer limite que tracemos para um conceito nunca coincidirá com seu uso real, pois ele não possui limites nítidos, mas apenas contornos imprecisos. A situação indica que não há, logicamente, uma circunscrição clara de um conceito,³² nem regras nítidas separando um uso possível de um impossível. Assim, também não há também uso exato para palavras, pois a linguagem permite a invenção de novos usos para elas, desde que concordem, em maior ou menor grau, com seu uso comum. As regras são instituídas e modificadas com o jogo em andamento. Portanto, a flexibilidade e variabilidade que a trama conceitual entrelaçada com o conceito de *crer* – assim como aquelas ligadas aos conceitos de *pensar*, *calcular* e outros –, pode assumir, e de fato assume, é muito maior do que a concedida por Hacker. Afinal, “não há um critério preciso para separar o significativo do sem sentido”.³³

2. Um terno frouxo demais: a mera semelhança

2.1 Afirmando a inteligência: a *Postura Intencional*

Enquanto a contribuição de Hacker para a IA e a neurociência baseia-se essencialmente no esclarecimento de estruturas conceituais, a de Daniel Dennett apoia-se nos poderes explicativos e preditivos do que ele denominou *postura intencional*. A inspiração para a postura intencional veio, segundo ele, de suas conversas com os pesquisadores da ciência da computação e da cognição.³⁴ Esses pesquisadores falavam naturalmente, sem qualquer autorreflexão, a respeito de computadores e cérebros pensando, querendo, decidindo, entre outros. Refletindo sobre quais regras estavam em jogo nesse discurso, Dennett concluiu que a postura intencional apreende ao menos parte das regras nele empregadas. Desde seu nascimento, ela tem sido empregada para caracterizar sistemas complexos, como computadores, e subsistemas, como partes do cérebro humano. Ele sintetiza a postura intencional da seguinte forma.

A ideia é que, quando projetamos um sistema complexo (ou fazemos engenharia reversa de um sistema biológico, como uma pessoa ou o cérebro de uma pessoa), podemos progredir ao decompor toda a pessoa maravilhosa

³² WITTGENSTEIN, *The Blue and Brown Books*, p. 25, p. 19. Cf. WITTGENSTEIN, *Investigações Filosóficas*, p. 52-57.

³³ HALLETT, G. *Wittgenstein's Definition of Meaning as Use*. New York: Fordham University Press, 1967, p. 109. Cf. WITTGENSTEIN, *The Blue and Brown Books*, p. 27; WITTGENSTEIN, *Investigações Filosóficas*, p. 58.

³⁴ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 86-87.

em subpessoas, por assim dizer: sistemas que agem como agentes, que possuem *parte* das capacidades de uma pessoa. Em seguida, esses homúnculos podem ser decompostos ainda mais em agentes ainda mais simples e menos parecidos com uma pessoa, e assim por diante — uma regressão *finita*, não infinita, que atinge seu limite quando chegamos a agentes tão estúpidos que podem ser substituídos por uma máquina.³⁵

Segundo Dennett, há pelo menos três posturas importantes que podem ser adotadas para a explicação e previsão.³⁶ A primeira é a *postura de design*, como central a noção de função (relativa à teleologia e ao propósito). Grosso modo, pode-se explicar e prever o funcionamento de um computador pelo funcionamento de suas partes, a partir de diferentes níveis de abstração, ao se focar em elementos funcionais superiores ou inferiores. O essencial nessa postura “é que fazemos previsões unicamente a partir do conhecimento ou de suposições sobre o design funcional do sistema, independentemente da constituição física”.³⁷ A segunda é a *postura física*. Nela, a explicação e a previsão se baseiam em nosso conhecimento das leis da natureza aplicado ao estado físico do *explanandum*. No entanto, segundo Dennett, dada a complexidade dos computadores que jogam xadrez, essas duas posturas são incapazes de fornecer uma explicação adequada de sua atividade. Assim, a melhor forma de um homem derrotar a máquina no xadrez seria prever suas jogadas em termos de “raciocínio”, “escolha” e correlatos, em suma, atribuindo-lhe racionalidade. Encarar o computador sob essa perspectiva é vê-lo como um sistema intencional, com base na suposição de que ele possui informações e é orientado por objetivos. E, se a noção de “posse de informações” é tão intencional e epistêmica quanto a de “crença”, pode-se afirmar que o computador possui crença. Do mesmo modo, pode-se dizer que o computador possui desejo, pois essa noção é tão intencional quanto a de “objetivo”.

Contudo, a partir da postura intencional, não se deve confirmar nem infirmar que os computadores possuem de fato crenças e desejos, pois a postura diz respeito apenas às explicações e previsões que podem ser obtidas por meio dela. Além disso, a adoção da postura é de natureza pragmática. Ela não é propriamente errada ou certa, mas compreendida como pragmaticamente necessária à previsão lógico-normativa (não empírica) da psicologia de todo e qualquer sistema intencional. Ela funciona quando nenhuma outra estratégia parece funcionar.³⁸ Dennett observa

³⁵ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 88.

³⁶ DENNETT, D. Intentional Systems, in *The Journal of Philosophy*, v. 68, n. 4, 1971, p. 87-106, p. 87.

³⁷ DENNETT, Intentional Systems, p. 88-89.

³⁸ Conforme Hacker, visto que a postura intencional carece de precisão empírica e poder explicativo, ela oferece apenas uma aparência de explicação e, assim, contribui para impedir o progresso científico (BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 160).

que a intencionalidade pressupõe a racionalidade. A racionalidade, por sua vez, fundamenta-se em processos adaptativos e evolutivos. Embora Dennett reconheça que a capacidade de crença – assim como as outras capacidades atribuídas aos sistemas pela postura intencional – possui pré-condições lógicas de aplicação, intrinsecamente ligadas às pré-condições lógicas do processo evolutivo, ele afirma que é necessário separar a postura intencional desses critérios e fazê-la caminhar por conta própria. Isso é feito, segundo ele, supondo a racionalidade do sistema, mesmo quando o desempenho racional está aquém do ideal.

2.2 A antítese do dilema: a demasiada permissividade do significado

A predicação proporcionada pela postura intencional se baseia em critérios pragmáticos, mas não somente neles. Segundo Dennett, “[a]s capacidades de crença e de comunicação têm seus próprios pré-requisitos [...] não faz sentido atribuir crenças a um sistema, a menos que as crenças atribuídas sejam, em geral, *apropriadas ao ambiente* e que o sistema responda *apropriadamente* a essas crenças”.³⁹ Assim como na abordagem hackeriana, o critério para a correção da predicação é comportamental – fundado na adequação contextual de respostas comportamentais.⁴⁰ No entanto, as posições de Dennett e Hacker quanto à predicação da IA e robôs são diametralmente opostas. Mas o que explicaria tal divergência? Um retorno à passagem-chave das *Investigações* é capaz de elucidar essa questão. Nas palavras de Dennett,

[é] aqui que Hacker e eu nos separamos. Fico feliz em citar eu mesmo essa passagem de Wittgenstein; na verdade, considero que estou *estendendo* a posição de Wittgenstein: vejo que robôs e computadores que jogam xadrez e, sim, cérebros e suas partes *de fato* “se assemelham a um ser humano vivo (por se comportarem como um ser humano)” – e essa semelhança é suficiente para justificar um uso ajustado do vocabulário psicológico para caracterizar esse comportamento. Hacker não vê isso, e ele e Bennett chamam todas as instâncias desse tipo de uso de “incoerentes”, insistindo repetidamente que elas “não fazem sentido”. Agora, quem está certo?⁴¹

Aqui se vê o ponto-chave de toda a diferença: semelhança, ou melhor, *semelhança suficiente*. Enquanto a explicação de Hacker exige semelhança muito próxima para o uso correto das expressões, a explicação de Dennett

³⁹ DENNETT, *Intentional Systems*, p. 101 e ss.; ênfase adicionada.

⁴⁰ Assim como Hacker, seguindo Wittgenstein, Dennett abstrai das questões de composição e constituição internas, permanecendo, desse modo, apenas com o comportamento, interpretado sob uma perspectiva intencional. Entretanto, diferentemente de Wittgenstein e Hacker, ele também se afasta, em consonância com essa abordagem, da exigência da existência de pensamento e consciência literais (DENNETT, *Intentional Systems*, p. 100).

⁴¹ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 78.

aceita uma semelhança bastante frouxa como suficiente para assegurar um uso adequado. Ele explica as predicções exemplificadas em sua passagem como inovações linguísticas produzidas pela extensão do uso comum dos termos psicológicos. Esse duplo movimento interconectado é justificado em virtude da surpreendente semelhança entre os comportamentos do homem e os dos sistemas e subsistemas.⁴² Em contrapartida, os esclarecimentos conceituais de Hacker qualificam tais predicções como incoerências, as quais incorreriam em flagrante desrespeito às regras de uso.⁴³ Antes que se possa responder à questão dennettiana sobre quem está correto – se ele ou Hacker –, é necessário o exame de alguns pontos essenciais.

Ao mesmo tempo em que Dennett se vê estendendo a aplicação do § 281, ele rejeita, de modo geral, o método de análise linguística de Wittgenstein-Hacker, juntamente com seus pressupostos. Em sua perspectiva, o exame do uso das palavras constitui uma investigação empírica (não *a priori*), semelhante a uma antropologia linguística. Na realização desse trabalho, segundo Dennett, o filósofo não contaria com verdades conceituais, mas principalmente com suas *intuições* gramaticais (*autoantropologia*), e, em menor grau, com a intuição daqueles que estão à sua volta.⁴⁴ Embora ele reconheça valor nas intuições e suas implicações lógicas, é preciso considerar seriamente quanto reconhecimento deve ser dado a elas como guia para a verdade, isto é, como elemento normativo do discurso.⁴⁵ No entanto, se para ele a análise do uso das expressões é insuficiente para guiar a predição e, dado que a postura, enquanto tal, não pode fornecer as regras para a correção da predição, como Dennett pode justificar as predicções que defende?

Em primeiro lugar, como mencionado anteriormente, Dennett afirma que a extensão da intencionalidade é finita. E ao se perguntar até que ponto

⁴² BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, 86. Como justificção adicional, ele indica que a literatura científica adotou o modo de fala da postura intencional. Por essa razão, ele destaca que a estratégia da postura intencional não é inédita, pois cientistas já vinham tomando emprestado o vocabulário intencional ao explicar eventos subpessoais. Para Robinson, “[a] validação de Dennett para sua escolha de locuções estranhas citando a frequência com que as utiliza, é – para adaptar um exemplo de Wittgenstein – ‘como se alguém comprasse várias cópias do jornal da manhã para assegurar a si mesmo de que o que dizia era verdade’ (*Ibid.*, p. 190; cf. WITTGENSTEIN, *Investigações Filosóficas*, p. 102).

⁴³ Por outro lado, Dennett não vê a questão de forma ingênua. Ele concede que há problemas conceituais sérios no modo de fala de cientistas cognitivos, mas sustenta que tais problemas geralmente podem ser resolvidos voltando-se para os modelos e suas explicações, pois, ao final, são os modelos que explicitam o que o termo problemático quer dizer. E, quando os cientistas se deixam enganar por interpretações exageradas de seus próprios modelos, elas devem ser corrigidas – como de fato já foram em alguns casos. No entanto, segundo Dennett, a abordagem de Hacker mostra-se de pouca utilidade nessa área (cf. BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 92-93; p. 204, Nota 19).

⁴⁴ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 81.

⁴⁵ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 202, Nota 12; p. 81-82.

a extensão ainda é válida para a atribuição de intencionalidade real, sua resposta é: “Não pergunte”.⁴⁶ No entanto, embora não ofereça um critério adequado para a extensão, ele supõe que as “razões para considerar um neurônio individual (ou um termostato) como um sistema intencional são pouco impressionantes, mas *não nulas*”.⁴⁷ Em seguida, como se estivesse se antecipando a possíveis objeções, adverte que “a segurança de nossas atribuições intencionais nos níveis mais altos não depende de identificarmos um nível mais baixo de intencionalidade real”. Duas questões se impõem aqui.

Primeira: a adoção da postura intencional é um requisito necessário à aplicação correta dos termos psicológicos? Segunda: ela é suficiente para a aplicação correta desses termos? Além disso, a postura intencional oferece a melhor explicação disponível? Deve-se levar em consideração que o sucesso na extensão correta dos termos em um nível mais alto pode ser acidental ou pressupor outros fatores dos quais não se tomou consciência. Por exemplo, ao adotar como uma espécie de limitador a ideia de que “agentes tão estúpidos que podem ser substituídos por uma máquina”, pode-se incorrer em certas dificuldades. Não está claro se agentes bastante inteligentes não poderiam, *em alguma medida*, ser substituídos por uma máquina estúpida. Em segundo lugar, agentes muito estúpidos ainda podem possuir parte da destreza de uma pessoa, principalmente quando não há nenhum critério determinando o grau de destreza que contaria como “tão estúpidos”, “não tão estúpidos” ou “nada estúpidos”. Finalmente, mesmo a ação de uma máquina (por mais simples que seja, lembre-se do termostato) pode ainda apresentar semelhança com a ação de uma pessoa em algum aspecto importante. Com efeito, virtualmente, qualquer máquina possui, em alguma medida, parte da destreza de um sistema complexo ou pessoa.

Mais uma vez, contudo, o ponto crucial está na semelhança suficiente. Interpretando-se as razões mencionadas por Dennett em termos de seme-

⁴⁶ DENNETT, D. Self-portrait. In: GUTTENPLAN, S. (Ed.). *A Companion to the Philosophy of Mind*. Cambridge: Blackwell, 1994, p. 236-244, *apud* BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 88.

⁴⁷ DENNETT, Self-portrait, *apud* BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 88; ênfase adicionada. Em um dado momento da discussão, Dennett afirma que a postura intencional permite ‘licença poética’ a fim de facilitar a tarefa explicativa da ciência (BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 89). Em outro momento, ele diz estar falando metonimicamente (*Ibid.*, p. 95). Sobre a licença poética, Hacker assevera que ela é válida para a poesia, mas não contribui para explicação ou precisão empírica (*Ibid.*, p. 160), podendo até mesmo impedir o progresso científico (*Ibid.*). Não se trata aqui de avaliar se a conclusão de Hacker sobre a licença poética decorre de sua costumeira rigidez linguística ou se a licença poética dificulta ou favorece a ciência. Mas apontar que, mesmo na licença poética, cujas expressões podem não ser facilmente rotuladas como verdadeiras ou falsas, ainda assim há a escolha das expressões, que podem ou não ser adequadas. Isto é, mesmo na licença poética, as expressões continuam funcionando segundo a norma NCL, levando em conta expressões rivais e concorrentes, como se verá no item 4.1.

lhança, pode-se dizer que as extensões defendidas por ele se baseiam em mera semelhança. No entanto, mera semelhança não é suficiente para uma justificação adequada. Pois “[e]stendendo nossos conceitos ‘naturalmente’, de semelhança em semelhança, poderíamos nos sentir no direito de dizer que um livreto é um livro, depois que um folheto também é, depois um panfleto, depois uma única página, uma única frase, uma única palavra”.⁴⁸ A incapacidade de fornecer, ou mesmo de se guiar por, um critério para a semelhança suficiente pode gerar abuso linguístico e conceitual.⁴⁹ Supondo que seja razoável adotar a postura intencional na relação com a IA (já que as pessoas assim o fazem em sua fala cotidiana), ainda seria razoável aplicá-la a termostatos, dado que não vemos um discurso semelhante empregado no dia a dia?⁵⁰

Retornando às justificações de Dennett para suas predicções, em segundo lugar, ele nota que os filósofos têm julgado comumente os desvios de sentido das expressões com base em argumentos que se fundamentam no que as pessoas diriam em determinadas situações, sem, contudo, jamais terem articulado as regras para o uso de nenhuma expressão.⁵¹ Como a própria discordância entre Dennett e Hacker sobre a predicação correta indica, o § 281 apontaria uma direção – possivelmente a correta –, mas, por si só, não é capaz de estabelecer plenamente as regras para a verdade e a inteligibilidade das expressões psicológicas. É necessária uma regra geral, como observa Dennett. Em sua réplica a Hacker, a respeito da natureza da filosofia, Searle nota que a análise filosófica dos atos de fala se torna mais poderosa à medida que ela conta com uma teoria geral da linguagem e dos atos de fala. Pode-se, talvez corretamente, suspeitar que a questão levantada por Dennett (“quem está certo?”) exige justamente uma teoria geral como condição para que se possa fornecer uma resposta adequada.

Em resumo, identificam-se dois complicadores: 1) a discussão tem falhado tanto em fornecer quanto em se apoiar em alguma regra geral válida para

⁴⁸ HALLETT, G. *Language and Truth*. New Haven: Yale Press University, 1988, p. 182.

⁴⁹ Segundo Robinson, “Dennett insiste em encaixar inúmeros e variados fatos em um *recipiente conceitual muito elástico* para ter forma e muito fino para o peso dos problemas reais” (BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 189; ênfase adicionada).

⁵⁰ John MacCarthy, que cunhou o termo ‘inteligência artificial’ na famosa conferência de 1956 em Dartmouth College, atribui crença *literal* aos termostatos (PAGELS, H.; DREYFUS, H.; MCCARTHY, J.; MINSKY, M.; PAPERT, S.; SEARLE, J. Panel Discussion, in *Annals of the New York Academy of Sciences*, v. 426, 1984, p. 138-160, p. 152). Trata-se de uma atribuição ainda mais forte do que a de Dennett.

⁵¹ É por essa mesma razão que Dennett se recusa a aceitar a reclamação de Hacker, segundo a qual haveria incoerência nos usos dos termos técnicos empregados por neurocientistas, em virtude do cruzamento conceitual entre o novo e o antigo uso. Dennett, uma vez que não foi apresentada nenhuma regra para orientar o cruzamento, tratar-se-ia de uma petição de princípio. Ele chega a afirmar de modo contundente que Hacker está apenas inventando tal coisa, já que não tem a menor noção de quais seriam as regras que governariam o uso dessas expressões (BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 85; cf. p. 186).

o uso das expressões em questão; 2) a falha dupla observada em (1) gera um dilema. Parece que se é obrigado a escolher *arbitrariamente* entre duas alternativas antitéticas: uma conduz a uma rigidez excessiva (Hacker) e a outra a uma permissividade excessiva (Dennett). Ambas as alternativas apresentam um caráter não realista em relação ao uso efetivo das expressões. Os complicadores, por sua vez, impõem uma demanda de natureza normativa. Torna-se, portanto, imprescindível o estabelecimento de uma norma geral para o uso das expressões que seja capaz de levar em conta tanto a natureza quanto o grau da semelhança exigidos – dimensões já implícitas no célebre § 281 das *Investigações*.

3. Um terno feito sob medida: a semelhança relativa

3.1 Síntese da inteligência: o Princípio de Semelhança Relativa (PSR) e a Norma de Correspondência Linguística (NCL)

Quase quarenta anos antes de Dennett, Clement Mundle denominou “linguística legislativa” a prática de *afirmar falsamente* o que *é* e o que *não é* dito, bem como o que *pode* e o que *não pode* ser dito, sem, contudo, oferecer quaisquer regras ou princípios gerais a partir dos quais tais afirmações, que sustentam suas teses, pudessem ser derivadas.⁵² Hallett reconheceu a validade da queixa a respeito da prática linguística dogmática. Com a elevação do conceito de *verdade* a lugar central da investigação filosófica, Hallett pôde articular o tipo apropriado de regra ou princípio de que a filosofia linguística necessitava para sua própria sobrevivência. Influenciado por Wittgenstein, ele decide seguir a recomendação no *Brown Book*, iniciando sua análise do conceito de *verdade* a partir da consideração de que as palavras devem sua vida às manipulações de signos que ocorrem num determinado *sistema linguístico*.⁵³ O raciocínio que Hallett engendra a partir desse eixo é reconstruído silogisticamente em Dornelas:

⁵² MUNDLE, C. *A Critique of Linguistic Philosophy*. Oxford: Clarendon, 1970, p. 16. Mundle qualifica aqueles que aderem a essa prática como “Policiais da Linguagem”, e a própria prática como falaciosa – intitulada por ele de “Falácia do Policiamento da Linguagem” – pois, à maneira da Falácia Naturalista, ela deriva, de forma ilegítima, um *dever-ser* (o que pode e o que não pode ser dito) a partir de um *é* (o que é e o que não é dito) (MUNDLE, *A Critique of Linguistic Philosophy*, p. 15; p. 114).

⁵³ WITTGENSTEIN, *The Blue and Brown Books*, p. 42. Cf. HALLETT, G. *Linguistic Philosophy: The Central Story*. Albany: State University of New York Press, 2008, p. 106 e ss.; HALLETT, G. *Theology within the Bounds of Language: A Methodological Tour*. New York: State University of New York Press, 2011, p. 30; HALLETT, *Language and Truth*, p. 205, Nota 13.

Argumento (1):

P1: O acordo na língua determina as coisas que possuem qualidade veritativa (IF § 241); P2: os enunciados são coisas que possuem qualidade veritativa; C: os enunciados (enquanto coisas que possuem qualidade veritativa) são determinados pelo acordo na língua.

Argumento (2):

P1: a verdade dos enunciados depende dos significados de seus elementos; P2: os significados de seus elementos dependem dos usos estabelecidos na língua como um todo; C: a verdade dos enunciados depende dos usos estabelecidos na língua como um todo.⁵⁴

A análise hallettiana torna patente não só a importância do uso estabelecido para a verdade dos enunciados e para o significado das palavras e expressões, mas também que o uso estabelecido só pode ser corretamente compreendido por referência ao sistema da língua como totalidade. Uma norma ou princípio geral dependerá, portanto, da elucidação da natureza do uso. Hallett revela as principais características do uso ao sistematizar a notável fórmula que Wittgenstein avançou para fins terapêuticos no § 43 das *Investigações*: “a significação de uma palavra é seu uso na linguagem”. Ele apresenta seis traços principais, dos quais dois são decisivos neste contexto: (i) *Flexibilidade*: a extensão das palavras não está fechada por nenhuma fronteira, pois nenhum limite intrínseco foi traçado. No entanto, nem tudo conta, por exemplo, como jogo; (ii) *Limitabilidade*: uma palavra pode ser estendida apenas até que esbarre em outra palavra. Toda palavra de uma língua é delimitada por seus termos concorrentes no espaço do mapa conceitual.⁵⁵

Apoiado na descrição do uso, Hallett chega à “correspondência linguística”⁵⁶ como critério da verdade. A fórmula que torna um enunciado (proposição, declaração, etc.) verdadeiro pode ser expressa nos seguintes termos: “para um enunciado ser verdadeiro *o emprego na ocasião deve estar em consonância com o uso estabelecido*”.⁵⁷ Certamente não se poderia afirmar que Dennett segue implicitamente uma fórmula semelhante, mas talvez se possa dizer isso de Hacker. Aqui surge uma questão central não discutida por Hacker:

⁵⁴ DORNELAS, B. *Uma (dis)solução do problema da linguagem religiosa a partir do pensamento de Garth Hallett*. 167 p. Dissertação (Mestrado em Filosofia) — Faculdade Jesuíta de Filosofia e Teologia, Belo Horizonte, 2018, p. 84-85.

⁵⁵ WITTGENSTEIN, *Investigações Filosóficas*, p. 43. Cf. *Ibid.*, p. 53.

⁵⁶ Para uma explicação da posição linguística da verdade defendida por Hallett, cf. DORNELAS, *Uma (dis)solução do problema da linguagem religiosa a partir do pensamento de Garth Hallett*, item 3.3.

⁵⁷ DORNELAS, *Uma (dis)solução do problema da linguagem religiosa a partir do pensamento de Garth Hallett*, p. 88. Por questões de espaço e objetivo, as características essenciais da fórmula não serão completamente detalhadas; uma análise mais plena pode ser consultada em *Ibid.*, p. 88 e ss.

“Quão próxima a correspondência deve ser para que uma afirmação seja verdadeira?”.⁵⁸ Como mencionado anteriormente, as extensões de Dennett pretendem se apoiar em semelhança suficiente, mas, na falta de um critério apropriado, acabam baseando-se apenas em mera semelhança. Esse modo de proceder, como observa Hallett, não é capaz de estabelecer por que o enunciado “O gato está no tapete” pode ser verdadeiro, enquanto “O guaxinim está no carpete” não, já que tanto gatos se assemelham a guaxinins quanto tapetes se assemelham a carpetes. Considerando a língua como um sistema autônomo e completo, Hallett pôde avançar um princípio capaz de satisfazer a exigência de um critério adequado de semelhança, chamado de “Princípio de Semelhança Relativa” (PSR):

Uma sentença é verdadeira se e somente se seu uso dos termos se assemelha mais de perto aos usos estabelecidos dos termos (sejam eles padronizados ou estipulados) do que se assemelharia qualquer substituição por uma expressão rival e incompatível.⁵⁹

Se qualquer tipo de semelhança e qualquer grau de similaridade fossem justificados, poderia ser considerado verdadeiro afirmar, de um mesmo ente, tanto a frase “Ali está um gato” quanto a frase “Ali está um guaxinim”. Contudo, o PSR evita tal problema, ao indicar o grau e a natureza da similaridade, sem perder de vista o modo como esses elementos de fato funcionam no discurso. A natureza da similaridade é *linguística* e, portanto, o estabelecimento de seu grau como *maior proximidade ao uso estabelecido na língua* deve observar, ainda que implicitamente, as expressões rivais e incompatíveis no âmbito da língua como um todo. Desse modo, o PSR encerra um critério comparativo em vez de absoluto. Enquanto norma implícita da verdade, o PSR pode dar lugar a uma norma explícita da predicação, a “Norma de Correspondência Linguística” (NCL):

Nosso uso de termos deve se assemelhar mais aos seus usos estabelecidos do que se assemelharia a substituição de qualquer expressão rival e incompatível.⁶⁰

É pela observação da NCL, ou de norma equivalente, que se torna inteligível chamar um determinado ente de gato e não de guaxinim – e vice-versa –, pois a maior proximidade com o uso estabelecido na língua assim determina. Dito de outro modo, poder-se-ia estender o termo *árvore* para nos referir ao que chamamos arbusto, dada a incrível semelhança, não fosse a existência do termo concorrente *arbusto*. Neste ponto, contudo, impõe-se uma característica central da teoria. Como proceder nas situações em que

⁵⁸ HALLETT, *Language and Truth*, p. 91.

⁵⁹ HALLETT, *Linguistic Philosophy*, p. 117.

⁶⁰ HALLETT, *Theology within the Bounds of Language*, p. 41.

os itens reais se situam entre categorias como *árvore* e *arbusto*, *adolescente* e *adulto*, ou *chuva* e *granizo*?

Embora na maioria das vezes, as realidades possam ser claramente identificadas como uma instância de um conceito ou de outro, existem casos nos quais a língua é incapaz de diferenciar instâncias de não-instâncias. Ou seja, casos intermediários são comuns para muitos conceitos. Em uma zona cinzenta autêntica, a língua não fornecerá suporte nem para a afirmação nem para a negação da inclusão de uma realidade em um determinado conceito, em razão da fronteira conceitual, que é caracteristicamente vaga, difusa e irregular. Via de regra, discussões que giram em torno de casos limítrofes erram ao assumir que sempre existe uma resposta. No debate entre Searle e Hacker acerca da consciência de animais inferiores, o primeiro traça o limite na posse de um sistema nervoso, e o segundo, no comportamento exibido. Um agarra o padrão comportamental e o outro o padrão biológico. O comportamento ausente representa a dessemelhança, e o sistema nervoso presente, a semelhança. Porém, como a autoridade da língua não pende para um lado ou para o outro, a disputa se torna meramente verbal. Sem ela, não adianta investigar os animais inferiores, pois eles não podem nos dizer por si mesmos se constituem instâncias autênticas de consciência, nem onde a consciência termina ou começa. É o próprio conceito de *consciência* que possui uma zona vaga e imprecisa; ele inclui claramente Searle e Hacker, e exclui claramente árvores e pedras, porém, seu uso não é favorável nem contrário à inclusão de certos animais inferiores.⁶¹ E quando nenhuma resposta, negativa ou positiva, se aproxima mais de perto do uso estabelecido do termo, o PSR não se abstém, mas declara o caso 'não-decidido', isto é, ele rejeita ambas as alternativas, pois nenhuma delas satisfaz as condições do princípio.⁶² Com o esboço desse quadro teórico, pode-se agora voltar o foco para o discurso da IA.

Hacker assevera que, além do sentido costumeiro das expressões, temos usos: (1) derivados; (2) analógicos; (3) homônimos; e (4) figurativos ou metafóricos.⁶³ Ao propor essa divisão, seu modelo aproxima-se da clássica tipologia tripartite de Aristóteles: (a) unívoco; (b) analógico; e (c) equívoco. Nesse esquema, o uso costumeiro se aproxima de (a), os usos (1) e (2) correspondem a (b), e (3) a (c). Neste ponto, Hallett se afasta radicalmente de Hacker, pois o PSR, ao reconhecer a complexidade, a flexibilidade e a diversidade do uso, não se restringe a uma classe de palavras, mas se estende a toda a língua, dissolvendo, portanto, o esquema tripartite. Desse modo, as distinções entre univocidade, analogicidade e equivocidade, de um lado, e entre figuratividade e literalidade, de outro, entram em colapso. Em seu lugar, o PSR estabelece a *analogia* como a

⁶¹ HALLETT, G. *Logic for the Labyrinth*. Boston: University Press of America, 1984, p. 166-167.

⁶² HALLETT, *Theology within the Bounds of Language*, p. 61.

⁶³ BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 149.

configuração padrão da língua.⁶⁴ Hallett explica melhor esse ponto considerando um exemplo direto da IA.

As pessoas jogam “xadrez” com tabuleiros e peças, ou por correspondência, ou em suas cabeças, ou com computadores; e agora os computadores podem jogar xadrez com computadores. Notando apenas os dois extremos neste contínuo – “xadrez” aplicado ao jogo de pessoas, com tabuleiros e peças, por um lado, e ao jogo dos computadores, sem tabuleiros ou peças, por outro – poderíamos chamar a primeira aplicação de “unívoca” e a outra de “analógica”. Mas, em nenhum ponto em tal alongamento de um termo, uma linha divisória nítida indica onde a “univocidade” cessa e a “analogia” começa.⁶⁵

Desde a perda do movimento das peças sobre o tabuleiro prescrita pela regra, até a ausência de pessoas, peças, tabuleiros e comunicação orais, é a semelhança mais próxima que garante a atribuição do termo *xadrez* a todas essas atividades, inclusive àquela situada no extremo do continuum. O PSR explicita a natureza do continuum como um continuum analógico. O mesmo vale para o discurso religioso. Quando se diz “Deus ama os homens”, o verbo *amar* não é empregado em dois sentidos exatamente iguais, nem em dois sentidos distintos, mas empregado de forma semelhante – cobrindo tanto a similaridade quanto a diferença. A palavra *amar* se aproxima mais do suposto sentimento divino do que *ódio*, *indiferença*, *mera simpatia*, ou outras. Do mesmo modo, o termo *xadrez* se aproxima mais daquilo que comumente chamamos xadrez do que de jogos como truco ou paciência, e o verbo *jogar* se aproxima mais da atividade desempenhada pelo computador do que *assistir*, *comentar*, entre outros. A lógica que busca impedir o discurso da IA – por exemplo, sobre “computadores jogando xadrez” – provavelmente alcançaria o discurso religioso. No entanto, o discurso religioso se baseia numa prática consolidada, com o uso de suas expressões ancorado na língua, e fora da academia, raramente, ou quase nunca, esse discurso é considerado puro absurdo. Ao aceitar esse último

⁶⁴ Dreyfus afirma que Wittgenstein vai mais longe do que geralmente se reconhece. Segundo ele, Wittgenstein propõe uma espécie de ‘reconhecimento da similaridade’, que não se distingue claramente do reconhecimento da semelhança. Dreyfus esclarece a ideia da seguinte forma: “cada uma das características que ele menciona na análise da semelhança de família [...] não é idêntica em quaisquer dois membros da família, mas, em troca, consiste numa rede de similaridades entrecruzadas. [...] Assim, nenhum par de membros da família precisa ter *quaisquer* características idênticas para que todos compartilhem uma semelhança de família. A semelhança é a noção fundamental na análise de Wittgenstein e não pode ser reduzida [...] a uma lista ou disjunção de características idênticas e determinadas. [...] Aqueles que são capazes de reconhecer um membro de uma ‘família’ não precisam ser capazes de identificar *quaisquer* traços comuns exatamente similares sequer a dois membros, nem há razão para supor que tais traços existam”. Ele conclui que “De fato, a formalização da semelhança de família em função de traços exatamente semelhantes eliminaria um tipo de abertura para novos casos, a qual constitui a característica mais impressionante desta forma de reconhecimento” (DREYFUS, H. *What Computers Can't Do: A Critique of Artificial Reason*. 1st ed. New York: Harper & Row, 1972, p. 39).

⁶⁵ HALLETT, *Theology within the Bounds of Language*, p. 69-70.

tipo de discurso, por que não se deveria aceitar o primeiro? No caso do termostato, os termos *detectar*, *medir* e *regular* se aproximam mais de perto do uso estabelecido do que, por exemplo, sua substituição por *pressentir*, *adivinhar* e *desregular*. Por essa razão, *pensar* e *acreditar* não se assemelham mais de perto às atividades do termostato a fim de autorizar uma predição legítima; para descrever suas atividades, os termos adequados são justamente *detectar*, *medir* e *regular*. Do mesmo modo que se diz que um relógio marca as horas, mas não que acredita ou pensa ser meio-dia, diz-se que o termostato detecta a temperatura, mas não que pensa, duvida ou acredita que, em um determinado momento, faz 28°C no quarto. E quanto a toda a rede conceitual que apoia a predição “joga xadrez”?⁶⁶

O erro, neste caso, consiste em não notar que a rede conceitual que sustenta a predição é ela própria constituída de termos e conceitos que não possuem limites estritamente definidos e que podem ser estendidos até que eles esbarrem nos demais termos e conceitos que compõem o mapa da língua. Assim, quando se diz que um computador joga xadrez, tanto a palavra *xadrez* quanto a palavra *joga* têm seu sentido estendido; e, se jogar é uma atividade, então *atividade* também é tomada analogicamente e, portanto, pode ser estendida. O mesmo vale para todos os nós conceituais que compõem a rede. Mas à medida que uma palavra é estendida, muita coisa se perde e muita coisa se ganha, a teia conceitual que a sustenta se modifica tanto qualitativa quanto numericamente.⁶⁷ Enquanto uma língua é essencial para o pensar humano, não o é para o pensar do cão. Contudo, o que é ou não essencial não está – nem é decidido – de antemão; a extensão pode vir depois, desde que respeite o sistema linguístico como um todo, à luz de um princípio como o PSR.⁶⁸ Considere, a seguir, o exemplo elucidativo fornecido por Waismann.

⁶⁶ Destacando os complexos fatores disposicionais e culturais envolvidos até mesmo na prática de brincar, Robinson declara que o “Deep Blue só ‘joga’ xadrez no mesmo sentido em que um micro-ondas ‘cozinha’ sopa, embora sua programação seja incomparavelmente mais complexa” (BENNETT, DENNETT, HACKER, SEARLE, *Neuroscience and Philosophy*, p. 184). A gramática de superfície das expressões *máquina* e *dispositivo eletrônico* pode ocultar importantes distinções lógicas. Em primeiro lugar, pode-se destacar o fato de o micro-ondas não realiza nenhuma “jogada” que não tenha sido previamente determinada por seus fabricantes. Em segundo lugar, é correto dizer “Hoje vou cozinhar no micro-ondas em vez de cozinhar no fogão” (não por acaso o micro-ondas é um forno). Mas seria igualmente cabível dizer: “Hoje vou jogar xadrez no Deep Blue em vez de jogar no tabuleiro”? Parece plausível dizer: “Joguei xadrez contra, ou em parceria com, o Deep Blue”. Porém, seria coerente dizer, no mesmo sentido: “Cozinhei com, ou contra, o micro-ondas” – talvez numa competição? Por fim, poderia o micro-ondas vencer uma competição de culinária do mesmo modo que o Deep Blue poderia sair vitorioso de uma competição de xadrez?

⁶⁷ Como explica Wittgenstein, quando tentamos definir *jogo* e percebemos elementos comuns de divertimento e habilidade, convém notar a diferença entre o que chamamos habilidade no xadrez e no tênis (WITTGENSTEIN, *Investigações Filosóficas*, p. 52).

⁶⁸ Segundo Wittgenstein, o conceito de *pensar* pertence à linguagem comum e apresenta um contorno eminentemente irregular, pois se refere a fenômenos muito distintos e serve a múltiplas manifestações. Seu aspecto é tão “terrivelmente ramificado” que Wittgenstein sugere que poderia haver mais de uma palavra para as diversas coisas às quais chamamos *pensar* (WITTGENSTEIN, L. Zettel. ANSCOMBE, G.; WRIGHT, G. (Ed.). Oxford: Blackwell, 1967, p. 21-23).

Não devemos esquecer que nossa ideia de um número vem, em primeiro lugar, dos números cardinais. Se prosseguimos destes para os números inteiros, racionais e reais, e finalmente para os complexos e hipercomplexos, a cada etapa desta série perdemos certas propriedades associadas ao ‘número’, tal como inicialmente o concebemos. Os números complexos, por exemplo, deixam de ser comparáveis quanto à magnitude; os números hipercomplexos (como os quaternions) deixam de obedecer a certas leis formais de cálculo (como a lei comutativa da multiplicação); e o mesmo se aplica aos números cardinais transfinitos de Georg Cantor. Em suma, aquilo que chamamos de ‘números’ torna-se progressivamente menos semelhante aos números cardinais – que são, evidentemente, nossos protótipos – até que no final essa semelhança desaparece por completo.⁶⁹

O exemplo apresenta uma forma de extensão tão complexa e radical que parece não apenas falsificar o método de Hacker, mas também desafiar a norma de extensão do PSR. Porém, à medida que observamos que cada nova extensão de “número” está de acordo com o PSR – semelhança relativa –, essa impressão se esvai.⁷⁰ Esse trecho tem importância central, ao mostrar que nem sempre o paradigma inicial precisa guardar semelhança relevante com uma extensão que se afastou o suficiente dele. Além disso, novas extensões podem se aproximar mais de outras extensões intermediárias do que do paradigma inicial, seguindo em múltiplas direções. Pense no cálculo realizado pelo ser humano com papel e caneta, no cálculo de cabeça, no cálculo feito por uma calculadora, no cálculo produzido por uma IA generativa e, por fim, no cálculo enigmático dos *savants*.

Portanto, dizer que a IA sabe ou acredita em algo significa, ao menos implicitamente, que a relação da IA com esse algo se aproxima mais dos termos *acredita* ou *sabe* do que de expressões como *ignora*, *não sabe*, *duvida*. A partir do reconhecimento desse fato, pode-se aceitar a existência de um continuum da inteligência. Colocando-se o ser humano adulto típico no meio desse continuum, pode-se situar, de um lado, em uma das extremidades, e a uma distância imensurável, Deus; do outro lado, bem próximo ao ser humano, ou no mesmo local, o “ser humano artificial”; um pouco mais distante, a IA forte;⁷¹ ainda mais distante, a pessoa com deficiência intelectual grave; em seguida, a criança pequena, o cão, outros animais inferiores, até alcançar, na extremidade oposta, a IA fraca.⁷²

⁶⁹ WAISMANN, F. *The Principles of Linguistic Philosophy*. New York: St. Martin’s Press, 1965, p. 185.

⁷⁰ HALLETT, *Language and Truth*, p. 182.

⁷¹ Para evitar uma confusão comum, faz-se mister explicitar que os termos ‘IA forte’ e ‘IA fraca’ são empregados, respectivamente, como sinônimos de ‘*artificial general intelligence*’ (AGI) e ‘*artificial narrow intelligence*’ (ANI). Essa distinção, no entanto, não segue o sentido original de Searle, que cunhou os termos para diferenciar a duplicação da mente de uma mera simulação dela. Contudo, a discussão em torno da distinção searleana não é o foco deste estudo e exigiria uma investigação separada.

⁷² Esse continuum serve meramente a propósitos ilustrativos e não pretende representar fielmente a realidade. Além disso, vários de seus elementos poderiam figurar em posições diferentes ou em mais de uma ao mesmo tempo.

3.2 O Veredito da inteligência: para além da permissividade do ‘Sim’ e da rigidez do ‘Não’

Esta seção retoma, mais propriamente, as questões principais. Trata-se de saber agora se podemos aplicar legitimamente – e não metaforicamente – termos como *inteligente* e *pensar* a máquinas e a sistemas. Qual é, então, o veredito do PSR e da NCL, por exemplo, para *inteligente*?

O veredito compreende quatro possíveis julgamentos:

1. *Inteligente*. Se um artefato⁷³ rotulado como IA satisfaz as condições linguísticas e factuais de aplicação do termo “inteligente” conforme previstas pelo PSR e pela NCL, ele pode ser verdadeira e significativamente chamado de inteligente.
2. *Não inteligente*. Se um artefato rotulado como IA não satisfaz as condições linguísticas e factuais de aplicação do termo, conforme previstas pelo PSR e pela NCL, ele não pode ser verdadeira e significativamente chamado de inteligente.
3. *Não decidido*. Se um artefato rotulado como IA constitui uma instância limítrofe – isto é, quando nem a predicação *inteligente* nem a predicação rival *não inteligente* se aproximam mais do uso estabelecido da expressão, nenhuma delas satisfaz a condição do PSR e da NCL – então ele não pode ser verdadeira e significativamente chamado de inteligente, mas também não pode ser chamado verdadeira e significativamente de não inteligente.
4. *Inteligente num sentido secundário*.⁷⁴ Se um artefato rotulado como IA é chamado de inteligente em conformidade com o fenômeno das extensões verbais que vão da causa para o efeito – nesse esquema, fala-se tanto de perguntas inteligentes e ações prudentes quanto de pessoas inteligentes e prudentes (que fazem essas perguntas e realizam essas ações); assim, pode-se falar igualmente de escovas inteligentes e dos sistemas de IA que as produziram –, então ele pode ser verdadeira e significativamente chamado de inteligente, pois, no dia a dia, dificilmente haverá algum mal-entendido.

Com relação à segunda questão, o arcabouço conceitual e metodológico de Hallett, aplicado em conjunto com a classificação sugerida, lança sua luz sobre a análise da legitimidade do discurso da IA. Desse modo, ilumina-se a “Babel Tecnológica” e, conseqüentemente, indica a observância da “autoridade da língua” como meio de superação.

⁷³ *Artefato* deve ser entendido em sentido amplo, abrangendo sistemas de IA, robôs e afins.

⁷⁴ Embora esse veredito pudesse, em princípio, ser integrado ao primeiro, é vantajoso distingui-lo dos demais no contexto teórico, a fim de preservar clareza e precisão.

Considerações finais

Ao final da discussão, conclui-se, sobretudo, que as questões no campo da IA podem beneficiar-se amplamente se levarem em consideração tanto a “autoridade da língua”, em geral, quanto a orientação proporcionada pelo PSR e pela NCL.

A análise mostrou essencialmente que os paradigmas representados por Hacker e Dennett falham por não disporem de uma teoria geral da verdade e do significado adequada, mediante a qual se pudesse orientar e avaliar legitimamente o uso dos termos problemáticos no discurso da IA. A privação de tal teoria corresponde à privação da consciência linguística necessária à discussão no campo da IA. A consciência da “semelhança relativa”, explicitada pelo PSR, evitaria tanto uma posição excessivamente rígida, como a adotada por Hacker, quanto uma posição demasiadamente permissiva, como a adotada por Dennett. Nenhuma das duas posições parece condizer com a forma pela qual a língua é efetivamente utilizada por falantes competentes. O PSR mostrou-se apto ao exibir um equilíbrio coerente entre o “nada pode” do primeiro paradigma e o “tudo pode” do segundo. Além disso, diferentemente das outras posições analisadas, o PSR captura razoavelmente a complexidade da língua, ao explicitar, por exemplo, tanto seu padrão analógico quanto as fronteiras difusas pertinentes aos casos limítrofes.

Por intermédio do “Princípio de Semelhança Relativa” e da “Norma de Correspondência Linguística”, concluiu-se que aquilo que pode ser rotulado verdadeira e inteligivelmente como IA pode receber não apenas o veredito “inteligente”, mas também os vereditos “não-inteligente” e “não-decidido”, bem como o veredito adicional “inteligente num sentido secundário”. Adotar a posição hallettiana permite contribuir para o esclarecimento de questões disputadas na IA que carregam sérias implicações jurídicas e morais, tais como as relativas à “autoria” e à “liberdade”. Sem adentrar propriamente a questão da autoria, não seria irrazoável supor, por exemplo, que em certos casos envolvendo empresas, engenheiros, sistemas generativos e usuários, o veredito correto seja “não-decidido”. Assim, a atribuição de autoria deverá observar tal situação e ser decidida, em última instância, por meio do debate jurídico-político.

Finalmente, a análise sugeriu que, embora Turing rejeite a orientação dos usos estabelecidos dos termos, seu teste se mostra plausível justamente por pressupor, de maneira implícita, uma norma como o PSR.⁷⁵ Visto que

⁷⁵ TURING, Computing Machinery and Intelligence, p. 433. Para compreender corretamente esse ponto, é preciso destacar que o teste de Turing é interpretado como um *experimento mental*, cujo objetivo é demonstrar que expressões como ‘pensamento’ e ‘inteligência’ poderiam ser atribuídas a computadores sob certas condições – não como a descrição de um *experimento empírico* a ser conduzido, destinado a verificar factualmente se certos computadores são capazes de pensar e de exibir inteligência.

o Jogo da Imitação, sabiamente, tenta emular comportamento humano da maneira mais estrita possível, até torná-lo indistinguível do da máquina, ele não apenas satisfaz as condições do PSR, mas também garante semelhança bem próxima, em vez de muito distante. Portanto, o comportamento exibido por um computador que passe no teste de Turing se aproximaria mais de perto do que chamamos inteligência e pensamento, o que garante condição suficiente para atribuí-los à máquina. Ademais, pode-se sugerir que o teste não recebeu a devida aceitação justamente por não dispor, ou não explicitar, a consciência linguística necessária à sua validação. Por outro lado, com a consciência linguística apropriada, o teste parece tornar-se obsoleto.

Referências

BENNETT, M.; DENNETT, D.; HACKER, P.; SEARLE, J. *Neuroscience and Philosophy: Brain, Mind, and Language*. New York: Columbia University, 2007.

DENNETT, D. Intentional Systems, in *The Journal of Philosophy*, v. 68, n. 4, Feb. 1971, p. 87-106.

DENNETT, D. Self-portrait. In: GUTTENPLAN, S. (Ed.). *A Companion to the Philosophy of Mind*. Cambridge: Blackwell, 1994, p. 236-244.

DORNELAS, B. *Uma (dis)solução do problema da linguagem religiosa a partir do pensamento de Garth Hallett*. 167 p. Dissertação (Mestrado em Filosofia) — Faculdade Jesuíta de Filosofia e Teologia, Belo Horizonte, 2018.

DREYFUS, H. *What Computers Can't Do: A Critique of Artificial Reason*. 1st ed. New York: Harper & Row, 1972.

DUMMETT, M. *Frege's Philosophy of Language*. 2nd ed. Cambridge, MA: Harvard University Press, 1981.

FRENCH, R. The Turing Test: The First 50 Years, in *Trends in Cognitive Sciences*, v. 4, n. 3, 2000, p. 115-122.

GLOCK, H-J. *A Wittgenstein Dictionary*. Cambridge, MA: Wiley-Blackwell, 1996.

HACKER, P. Men, Minds and Machines. In: HACKER, P. *Wittgenstein, Meaning and Mind*. Cambridge, MA: Blackwell, 1990, p. 89-111.

HALLETT, G. *A Companion to Wittgenstein's "Philosophical Investigations"*. Ithaca/London: Cornell University Press, 1977.

HALLETT, G. *Essentialism: A Wittgensteinian Critique*. Albany: State University of New York Press, 1991.

HALLETT, G. *Language and Truth*. New Haven: Yale Press University, 1988.

HALLETT, G. *Linguistic Philosophy: The Central Story*. Albany: State University of New York Press, 2008.

HALLETT, G. *Logic for the Labyrinth*. Boston: University Press of America, 1984.

- HALLETT, G. *Theology within the Bounds of Language: A Methodological Tour*. New York: State University of New York Press, 2011.
- HALLETT, G. The Theoretical Content of Language, in *Gregorianum*. Roma: Pontificia Università Gregoriana, v. 54, n. 2, 1973, p. 307-337.
- HALLETT, G. *Wittgenstein's Definition of Meaning as Use*. New York: Fordham University Press, 1967.
- HANFLING, O. *Wittgenstein and the Human Form of Life*. New York: Routledge, 2002.
- HAUSER, L. The Sense of Thinking? Reply to Rapaport's Comments, in *Minds and Machines*, v. 3, n. 1, p. 21-29, 1993.
- HAUSER, L. Why Isn't My Pocket Calculator a Thinking Thing? *Minds and Machines*, v. 3, n. 1, 1993, p. 3-10.
- MUNDLE, C. *A Critique of Linguistic Philosophy*. Oxford: Clarendon, 1970.
- PAGELS, H.; DREYFUS, H.; MCCARTHY, J.; MINSKY, M.; PAPERT, S.; SEARLE, J. Panel Discussion. *Annals of the New York Academy of Sciences*, v. 426, p. 138-160, 1984.
- TURING, A. Computing Machinery and Intelligence., in *Mind*, New Series, v. 59, n. 236, Oct. 1950, p. 433-460.
- WAISMANN, F. Analytic-Synthetic II, in *Analysis*, v. 11, n. 2, Dec. 1950, p. 25-38.
- WAISMANN, F. *The Principles of Linguistic Philosophy*. New York: St. Martin's Press, 1965.
- WITTGENSTEIN, L. *The Blue and Brown Books*. Oxford: Basil Blackwell, 1958.
- WITTGENSTEIN, L. *Investigações Filosóficas*. São Paulo: Abril Cultural, 1975.
- WITTGENSTEIN, L. *Zettel*. ANSCOMBE, G.; WRIGHT, G. (Ed.). Oxford: Blackwell, 1967.