

UMA RESPOSTA FUNCIONALISTA AO ARGUMENTO DO QUARTO CHINÊS DE SEARLE

A FUNCTIONALIST RESPONSE TO SEARLES'S CHINESE ROOM ARGUMENT

André Sathler Guimarães

Coordenação de Pós-Graduação do Centro de Formação, Treinamento e Aperfeiçoamento da Câmara dos Deputados, Brasília – Brasil
andre.sathler@gmail.com

Resumo: O Argumento do Quarto Chinês de John Searle permanece como um desafio aos pesquisadores da área de Inteligência Artificial (IA), tendo sido objeto de respostas e contra-respostas ao longo dos últimos anos. O presente artigo busca apresentar uma nova possibilidade de resposta para o Argumento do Quarto Chinês, a partir de uma perspectiva funcionalista, característica dos primórdios da investigação sobre IA. Trata-se de abordagem filosófico-especulativa, construída a partir de argumentação e contra-argumentação. Da análise do cerne do argumento e dos contra-argumentos até aqui apresentados, foram extraídos elementos indicativos de que Searle operou uma mudança de campo da discussão, tornando impossível uma resolução de seu desafio nos termos apresentados. O artigo conclui apresentando uma possibilidade de resposta ao Argumento do Quarto Chinês, que sinaliza para sua irrelevância do ponto de vista da pesquisa em IA.

Palavras-chave: Filosofia da Mente. Argumento do Quarto Chinês. Inteligência Artificial.

Abstract: *John Searle's Chinese Room Argument remains a challenge for researchers in Artificial Intelligence (AI); it has been the subject of responses and counter-responses during the last years. This paper aims to present a new possibility of answer to the Chinese Room Argument, from a functionalist perspective, a characteristic of early AI researches. It is the philosophic-speculative approach, built on argumentation and counter-argumentation. The analysis of the argument's and counter-argument's core provided elements indicating that Searle has made a change in the field of discussion, making it impossible to solve his challenge in the terms presented. The paper is concluded with a possible response to the Chinese Room Argument, which points out the irrelevance of this issue to AI research.*

Keywords: *Philosophy of Mind. Chinese Room Argument. Artificial Intelligence.*

Introdução

John Searle ganhou notoriedade por seus constantes ataques aos projetos filosófico e científico da Inteligência Artificial (IA). Sua obra mais conhecida nesse campo foi o artigo “*Minds, Brains and Programs*”, no qual apresentou um forte argumento contra a possibilidade do pensamento nas máquinas, que veio a ficar conhecido como o Argumento do Quarto do Chinês – AQC¹. A premissa do AQC é a impossibilidade de se atribuir capacidade de compreensão à máquina dada a sua incapacidade de transcender ao nível das operações sintáticas.

Ao longo do tempo, foram desenvolvidas e apresentadas várias tentativas de resposta ao AQC, cujo foco, de forma geral, esteve na postulação da capacidade de compreensão da máquina. Entende-se que essa linha de argumentação foi perseguida justamente por buscar atacar a premissa central do AQC. O presente artigo pretende argumentar pela irrelevância da questão para o campo de pesquisas da IA, por consistir, na verdade, em um falso dilema.

¹ As linhas gerais do AQC serão apresentadas com base na versão “*Mentes, Cérebros e Programas*”, traduzida por Cléa Regina de Oliveira Ribeiro, publicada em Teixeira (1996).

1. O Argumento do Quarto Chinês

O AQC foi desenvolvido por Searle na forma de um *Gedankenexperiment*: uma pessoa, trancada em um quarto, recebe um calhamaço de papel com um texto em chinês. Essa pessoa desconhece completamente o chinês, a ponto de não conseguir discerni-lo de rabiscos. Um segundo calhamaço de papel é entregue a essa pessoa, com texto em sua língua nativa, contendo instruções para correlacionar esse segundo texto com o primeiro. Ou seja, a pessoa adquire a possibilidade de relacionar um conjunto de símbolos com outro, apenas pela sua forma. Um terceiro calhamaço é entregue à pessoa, com símbolos em chinês e mais instruções em língua nativa²,

as quais me possibilitarão correlacionar elementos deste terceiro maço com os dois primeiros; estas regras me instruem como relacionar determinados símbolos em chinês com certos tipos de configuração e os devolver como resposta a determinadas configurações dadas no terceiro calhamaço (SEARLE, 1996, p. 66).

Com o tempo, a pessoa adquire experiência em seguir as instruções e manipular os símbolos em chinês. Do ponto de vista de um observador externo, as respostas (*outputs*) que a pessoa que está dentro do quarto emite, apenas manipulando formalmente os símbolos (pela sua forma, seguindo as instruções), tornam-se indiscerníveis das respostas de um falante de chinês. Para Searle, a pessoa dentro do quarto obtém “respostas manipulando símbolos formais em chinês, sem significação. No que diz respeito ao chinês [essa pessoa se comportou] como um computador; executando operações computacionais com base em elementos formalmente especificados” (SEARLE, 1996, p. 69). Searle prossegue afirmando que “por mais que se coloque no computador princípios formais isto não será suficiente para a compreensão, uma vez que um ser humano será capaz de seguir tais princípios formais sem compreender nada (*Ibidem*). A essência do AQC, na visão de Searle, é que se o homem no quarto não compreende chinês apenas com base na implementação de um programa apropriado para compreensão do chinês, então nenhum computador digital, com a mesma base, poderia compreender chinês, porque nenhum computador tem algo que o ser humano não tenha.

Searle pretendia, em última instância, responder de forma definitiva ao teste de Turing. O pioneiro da Inteligência Artificial (IA) esboçou seu famoso teste, segundo um qual um computador seria inteligente caso conseguisse passar por humano em uma interação às cegas com outros seres humanos. A linha de fundo do argumento de Searle é que meramente seguir regras é operar em um nível sintático, no qual não existe o nível semântico, que seria a compreensão. Programas computacionais podem vir a conversar em linguagem natural, porém não conseguirão desenvolver a capacidade de compreender o que estão falando. Ou seja, ainda que fossem capazes de passar no teste de Turing, sustentando uma interação dialógica plausível (baseada na linguagem natural) com outros seres humanos, os computadores não seriam inteligentes ou teriam capacidade genuína de compreensão. Compreensão, para Searle, é a posse simultânea de estados mentais (intencionais) e a condição de verdade desses estados (validade, sucesso). Assumindo-se essa acepção, compreender assemelha-se a pensar.

² Searle utiliza da linguagem em primeira pessoa em seu *Gedankenexperiment*, o que faz com que a língua nativa seja o inglês. Ao nos referirmos ao experimento, utilizaremos a linguagem de terceira pessoa, o que o tornar mais abrangente, sem invalidar suas premissas. A linguagem em primeira pessoa é mantida nas citações originais, em respeito ao autor.

Portanto, o AQC consiste, na realidade, em um argumento contra a possibilidade de uma inteligência artificial “real”³.

O AQC pode ser contextualizado a partir de alguns antecedentes históricos, cujos argumentos, em essência, defendem que o comportamento exterior não pode ser assumido como referência para a existência de pensamento inteligente interior. O funcionamento interior, com características físicas (sintática), não é suficiente para explicar o comportamento exterior (semântico). Aristóteles já apontava para a questão, ao afirmar a impossibilidade de discernir entre um autômato que apresentasse comportamento idêntico a um ser humano e um ser humano, sem que fosse exposto o seu interior. Leibniz no chamado argumento “Leibniz Mill” defende que uma percepção depende de *bases mecânicas inexplicáveis*, as quais, quando analisadas em si, não possibilitam a explicação do fenômeno perceptivo. Alan Turing, antes de formular seu famoso teste, propôs que um humano que operasse uma máquina de jogar xadrez apenas a partir de instruções para gerar os movimentos, não se distinguiria de um computador e não precisaria conhecer os fundamentos do jogo de xadrez – apenas seguir as instruções. Pode-se afirmar que a essência desses argumentos – que viria a estar presente também no AQC – é a diferença entre *aparência de compreensão* e *compreensão genuína*.

Em termos de lógica formal, o AQC pode ser apresentado da seguinte forma:

- (1) Caso a IA⁴ seja verdadeira, então existe um programa para chinês de modo que se qualquer sistema computacional execute o sistema, o sistema então venha a compreender chinês.
- (2) Eu posso executar um programa para chinês sem vir a compreender chinês.
- (3) Então a IA é falsa.

O AQC seria a sustentação da segunda premissa.

2. Respostas ao AQC

O artigo de Searle provocou muita controvérsia, com réplicas e trélicas. Uma primeira linha de argumentação contrária foi a *resposta dos sistemas de Berkeley*⁵. O fundamento dessa réplica está na noção de que embora o indivíduo dentro do quarto não compreenda chinês, ele é uma parte de um sistema maior que compreende [quarto + calhamaços + banco de dados]. Essa crítica concede ao AQC o fato de que inexistente compreensão no indivíduo que está no quarto, porém aponta para o fato de que a compreensão acontece em alguma outra instância, ou entidade diferente.

Uma segunda linha de argumentação foi a *resposta do robô*, de Yale. Postula-se aqui que um computador, dentro de um robô, que não apenas operasse símbolos formais, mas fizesse coisas como andar, mover-se, comer, beber, etc. teria compreensão nos termos postulados por Searle. Essa crítica concede ao AQC o fato de que apenas executar um programa de linguagem natural, conforme descrito no Argumento, não criaria compreensão,

³ Compreendida essa afirmação como a possibilidade de uma inteligência semelhante à humana ou antropomórfica.

⁴ Inteligência Artificial no sentido forte, que assume como projeto a construção de softwares que teriam a capacidade de igualar o comportamento humano inteligente. Posteriormente, essa linha inicial de pesquisa veio a ser chamada de GOF AI – *Good Old-fashioned Artificial Intelligence*. A sigla foi criada em 1981 pelo filósofo J. Haugeland e significa, em uma tradução literal, a *boa e velha inteligência artificial*.

⁵ Algumas das contra-argumentações são tratadas pelo próprio Searle em seu artigo. As denominações dos argumentos são as mesmas empregadas por aquele autor.

porém um programa de computador *corporificado*, interagindo com o mundo físico por meio de sensores e motores, teria compreensão.

Uma terceira linha de argumentação, oriunda de Berkeley e do MIT, defende que poderia haver uma simulação do cérebro – um programa que simule a sequência efetiva da atividade dos neurônios nas sinapses do cérebro de um falante nativo de chinês, quando este entende histórias e dá respostas a elas. O quarto argumento contrário é o da *combinação* (Berkeley e Stanford), que reúne as três linhas de contra-argumentação anteriores. Um quinto contra-argumento é a objeção das *outras mentes* (Yale), que defende que é impossível saber se uma pessoa compreende chinês apenas pelo que ela expressa discursivamente. Para esse argumento, caso seja impossível atribuir compreensão a um computador com base no comportamento exibido no AQC, seria identicamente impossível atribuir compreensão a um ser humano na mesma situação.

3. A irrelevância do AQC

Searle pretende provar que o computador não compreende, uma vez que apenas manipula símbolos formais, os quais, em última instância, nem deveriam ser considerados símbolos, porque não simbolizam nada e são tomados em si mesmos (pelas suas formas)⁶. As diversas tentativas de resposta se concentraram em provar que o computador pode compreender, nos termos colocados por Searle, o que aproxima-se muito de provar que o computador poderia pensar. Colocado nesses termos o argumento de Searle, uma contra-prova definitiva somente poderia ser encontrada quando se obtivesse um computador que pensasse da mesma forma que um ser humano.

Porém, mesmo essa suposta prova definitiva teria dificuldades de ratificação, pois levaria à mesma situação em que se fica diante dos autômatos. Não há como saber se *o outro* é humano, ou um marciano revestido com um disfarce de humano, se os comportamentos forem indiscerníveis (o problema das *outras mentes*).

Nossa resposta ao AQC é: aceite-se que um computador não pode compreender / pensar. Porém, assumo-se como consequência lógica que o ser humano dentro do quarto também não compreende. Ou seja: no âmbito dos limites bizarros do *Gedankenexperiment* de Searle, tanto um computador quanto um ser humano não teriam a faculdade de compreensão. Nossa resposta segue a linha dos pioneiros da IA: reconhecer a existência do problema, mas devolvê-lo como um problema igualmente comum ao gênero humano.

No contexto dessa discussão, há os que chamam a atenção para o fato de que está em questão a possibilidade de existência de uma inteligência artificial e não a possibilidade de um ser humano artificial. Ser inteligente e ser humano são coisas distintas e não faz sentido postular que uma máquina, para ser inteligente, precise ter necessidades sexuais, fome, pulso, emoções, ou, ainda, um corpo antropomórfico. Esse problema deriva parcialmente da tradição cultural, fortemente enraizada, de considerar que o que distingue o *Homo sapiens* das demais espécies é a sua capacidade de pensar. Para Hofstadter (2001, p. 186),

talvez estejamos inconscientemente assoberbados com um chauvinismo semelhante com respeito à inteligência e, em consequência, com respeito ao significado. Em nosso chauvinismo, consideraríamos ‘inteligente’ qualquer

⁶ Talvez para enfatizar esse ponto Searle tenha escolhido o chinês, uma linguagem morfêmica, ao invés de uma linguagem fonêmica, embora a eventual substituição não afete o argumento (os fonemas, tomados isoladamente, não significam nada).

ser com um cérebro suficientemente parecido com o nosso e recusar-nos a reconhecer como inteligentes outros tipos de objetos.

O que está em jogo é a definição de quais seriam os predicados que estamos dispostos a atribuir às máquinas, sem que essa atribuição resulte em uma ontologia ingênua e eticamente inerte. Segundo Teixeira (2000, p. 125),

a noção de uma inteligência artificial como realização das tarefas inteligentes, ou seja, a possibilidade de replicação mecânica de segmentos da atividade mental humana por dispositivos que não têm a mesma arquitetura nem a mesma composição biológica e físico-química do cérebro foi a grande motivação para o aparecimento das teorias funcionalistas.

A aceitação da equivalência dos funcionalistas entre o *ser* e o *fazer* dos funcionalistas obstrui a argumentação de que algo que se comporta conscientemente não seja consciente⁷. Adversamente, ainda que se repila a equivalência *ser/fazer*, há que se reconhecer seus isorresultados. Nos dizeres de Wittgenstein (2001, p. 167),

que haja uma regra geral por meio da qual a música pode extrair a sinfonia da partitura, uma por meio da qual se pode derivar a sinfonia dos sulcos do disco e, segundo a primeira regra, derivar novamente a partitura, é precisamente nisso que consiste a semelhança interna dessas configurações, que parecem tão completamente diferentes.

Como seres humanos, podemos aprender a imitar a Máquina de Turing. Logo, por definição, somos, no mínimo, Máquinas de Turing. Esse parece ter sido o exercício minimalista de Searle – reduzir um ser humano, isolado em um quarto, a uma Máquina de Turing. Contudo, pelos princípios básicos da lógica, se A é B, B é A. Turing se perguntava: “pode alguém fazer uma máquina que tenha sentimentos, como você e eu?”. Ao que respondia: “eu devo nunca saber, não mais do que eu nunca deverei ter certeza de que você sente como eu sinto” (Turing, 2004, p. 569). Turing lançava uma linha de argumentação que se tornaria clássica como defesa da IA. Diante de argumentos que assumam premissas do tipo *o computador não pode fazer isso*, contra-argumenta-se *concordo, mas o homem também não*.

Por exemplo: contra a possibilidade de uma intencionalidade de primeira pessoa aplicada a máquinas computacionais usualmente se levanta o Teorema de Gödel, referindo-se à Proposição VI do referido autor: “*Proposition VI: To every ω -consistent recursive class c of formulae there correspond recursive class-signs r , such that neither v Gen r nor Neg (v Gen r) belongs to Flg (c) (where v is the free variable of r)*” (GÖDEL, 1992, p. 57). O próprio Gödel, explicando o que buscou provar, afirma que se trata do fato de que problemas relativamente simples, na teoria dos números ordinais inteiros, não podem ser decididos a partir de seus axiomas. Portanto existem proposições que não podem ser provadas ou refutadas dentro do sistema. A prova gödeliana, portanto, não se relaciona diretamente à

⁷ Teixeira revela sua preocupação com a aceitação integral da perspectiva funcionalista: “esse salto corresponderia também a alguma quintessência que, segundo Descartes, ficaria faltando na forma de um autômato, pois, na medida em que ser consciente não seria uma propriedade física, a replicação física integral de um cérebro não implicaria, necessariamente, na replicação do caráter consciente dos estados mentais que esse autômato poderia vir a ter.” (TEIXEIRA, 2000, p. 77). Teixeira parece se referir à afirmação cartesiana sobre os autômatos: “primeiro, eles não podem jamais usar palavras ou outros sinais construídos, como nós usamos para declarar nossos pensamentos aos outros (...) segundo, enquanto eles podem fazer muitas coisas tão bem quanto qualquer um de nós ou até melhor, eles vão infalivelmente falhar em outras, revelando que eles não agem com base em conhecimento mas apenas com base na disposição de seus órgãos.” (DESCARTES, 2000, p. 20).

questão de se os computadores poderão ou não pensar. Sua relevância consiste no fato de apontar para a existência de limites nos sistemas formais (inclusive para a lógica formal).

Como a programação de computadores opera basicamente com a lógica booleana, a prova de Gödel coloca limites intransponíveis para aquilo que um computador poderá fazer, inclusive quanto à sua eventual *capacidade cognitiva*. Porém, a réplica dos fundadores da IA, como Turing, consiste em afirmar a irrelevância desse elemento, uma vez que também existiriam limites para a capacidade cognitiva humana⁸:

a resposta mais simples a esse argumento é a de que, embora esteja estabelecido que há limitações aos poderes de qualquer máquina específica, enunciou-se apenas, sem qualquer espécie de prova, que nenhuma limitação desse tipo se aplica ao intelecto humano (TURING, 1996, p. 38).

Sobre esse ponto, vale recuperar também uma argumentação de Hofstadter (2001, p. 21):

ocorre que nenhum método algoritmo pode dizer como aplicar o método de Gödel a todos os tipos possíveis de sistemas formais. E, a menos que se tenha inclinações algo místicas, tem-se de concluir, portanto, que qualquer ser humano simplesmente alcançará os limites de sua própria capacidade de gödelização em algum ponto.

A metáfora computacional altera a questão chave da pesquisa no campo da IA. Não se trata mais de perguntar se um computador (Máquina de Turing universal⁹) poderá simular os processos de funcionamento cerebral. Mas sim de se saber se todos os processos conscientes humanos são algoritmizáveis, ou seja, redutíveis a alguma espécie de procedimento computacional.

Na perspectiva de responder afirmativamente, pode-se considerar a afirmação de von Neumann de que “a observação mais imediata no que concerne ao sistema nervoso é a de que o seu funcionamento é, *prima facie*, digital” (VON NEUMANN, 2005, p. 74). Para von Neumann, apesar do impulso gerado pelo/no neurônio abranger aspectos variados, de ordem elétrica, química e mecânica, o processo de sua geração é invariável ou idêntico sob todas as condições, representando uma resposta unitária e essencialmente reproduzível para uma variedade imensa de estímulos.

⁸ A discussão em torno desse assunto terminou por ficar mais conhecida pelo seu nome em inglês – *the bounding problem* – que, literalmente, seria traduzido como um problema de limites, ou um problema de fronteiras. Aqueles que defendem a existência desses limites partem da configuração característica do humano, em seus aspectos ontológicos, biológicos e morfológicos, explicitando que esse conjunto teria limites aos quais não conseguiria ultrapassar. Afina-se com essa idéia Noam Chomsky, quando diz que “quanto à questão do alcance cognitivo, se os humanos são parte do mundo natural e não seres sobrenaturais, então a inteligência humana tem seu escopo e seus limites determinados pelo *design* inicial.” (CHOMSKY, 2005). Ou seja, certos fatos não estarão ao alcance do sistema cognitivo peculiar ao ser humano. Isso se repetiria com outros animais, em outros níveis. Chomsky se refere, também, à prova experimental de que ratos são incapazes de atravessar labirintos com propriedades numéricas para afirmar que isso ocorre pela ausência nesses roedores de conceitos apropriados. Haveria, portanto, ainda segundo Chomsky, “mistérios para ratos”, bem como “mistérios para humanos.”

⁹ Para Dennett (1991, p. 212), a Máquina de Turing universal é uma idealização brilhante e uma simplificação de um fenômeno hiperracional – um matemático realizando um cálculo rigoroso.

Entretanto, o próprio Turing, argumenta que apesar do pensamento ser claramente uma atividade motora das células cerebrais, a resposta a essa questão não estava clara: “os processos da máquina são mosaicos de partes muito simples e padronizadas, mas os designs podem ter grande complexidade, e não está óbvio onde estão os limites quanto aos padrões de pensamento que eles podem imitar.” (TURING, 2004, p. 500).

A redução algorítmica, com seus requisitos de precisão, sempre traz o dilema de como prever formas de lidar com as exceções e o imprevisto. Até o momento, os defensores da IA insistem que o obstáculo é de natureza tecnológica, e não filosófica. Conforme Teixeira, “a idéia do conhecimento como representação parece estar na raiz das dificuldades tecnológicas aparentes envolvidas na construção desses sistemas: explosão combinatorial, rigidez de estrutura e assim por diante.” (TEIXEIRA, 2004, p. 40).

Um algoritmo efetivamente computável precisa atender aos seguintes requisitos:

- 1 – ser finito (em tempo e extensão);
- 2 – ser completamente explícito e livre de ambigüidades;
- 3 – ser infalível;
- 4 – poder ser realizado por um *idiot savant*.

Mesmo que teoricamente, um ser humano paciente e metuculoso deve ser capaz, sem a ajuda de qualquer instrumento e sem a necessidade de *insights*, de chegar ao fim do procedimento, com o uso apenas de uma quantidade potencialmente ilimitada de papel, tinta e tempo. A Máquina de Turing universal é capaz de realizar qualquer algoritmo, mas isso não se deve à uma impressionante capacidade proto-cognitiva. Ela o faz por meio de funções recursivas, ou seja, uma função f que seja definida em termos da aplicação repetida de um número de funções simples aos seus próprios valores, com a especificação de uma fórmula recursiva e uma cláusula base. Nessa linha, pode-se recuperar o pensamento de Fodor (1996, p. 39),

se, como muitos de nós supõem, mentes são essencialmente dispositivos de manipulação de símbolos, deve ser útil pensar em mentes pelo modelo da Máquina de Turing, uma vez que as Máquinas de Turing são tão genéricas quanto qualquer dispositivo manipulador de símbolos possa ser.

A generalidade decorre de sua simplicidade. Máquinas de Turing são sistemas computacionais fechados, cujas computações são determinadas somente pelo estado atual da máquina, a configuração da fita e o programa. McCulloch e Pitts (2000) mostraram que uma rede neural pode calcular qualquer número¹⁰ que possa ser calculado por uma Máquina de Turing. Essa demonstração foi importante porque uniu um modelo do funcionamento neuronal humano à teoria dos autômatos. Ao provar que as operações de uma rede neural e uma Máquina de Turing formalmente convergem, McCulloch e Pitts confirmaram seu *insight* de que “cérebros não secretam pensamentos como o fígado secreta a bile, mas eles computam o pensamento da forma como os computadores eletrônicos calculam números” (McCULLOCH-PITTS, 2000, p. 351).

A aproximação com o funcionamento cerebral passa pela noção de que todos os aspectos do pensamento podem ser vistos como descrições de nível alto de um sistema que, em um nível baixo, é governado por regras simples e formais. A ruptura decisiva reside na adoção da idéia do computador como um sistema simbólico e não como um dispositivo causal-físico ordinário, como outras máquinas. Pois há uma noção fortemente enraizada de

¹⁰ No contexto da lógica booleana, isso equivale a dizer qualquer proposição.

que o elemento distintivo da espécie humana é sua capacidade de processar símbolos. Aceitas ambas as premissas, resta estabelecido o isodinamismo entre mentes e máquinas computacionais.

Considerações finais

Searle assume, logo no início de seu trabalho, que faria um ataque a IA no sentido forte¹¹ - GOFAI. Posteriormente, com base no seu *Gedankenexperiment*, concentra o argumento no fato de que o computador não pode pensar. As réplicas já conhecidas são de cunho defensivo, postulando a possibilidade do pensamento na máquina. Provar que a máquina pode pensar é apresentar uma máquina que pense. Dessa forma, solucionar o AQC, nos termos colocados por Searle, assemelha-se a provar que Deus existe – apresentando-se Deus. Defendemos que Searle confundiu propositalmente os níveis da argumentação e lançou um despiste com seu bem elaborado experimento mental. A resposta ao seu artigo, considerado como uma objeção à GOFAI, deve ser: *e daí?*

Um ser humano submetido às condições de seu *Gedankenexperiment* se sai igual ou pior¹² que um computador. O argumento de Searle é irrelevante para o campo de pesquisa em IA. De fato, as pesquisas continuam avançando, à revelia da insolubilidade do AQC.

Vale abordar aqui a expectativa de que um ser inteligente deverá ser igual ao homem. Parafraseando Wittgenstein, mesmo se um dia os computadores vierem a pensar, nós não seremos capazes de compreender os seus pensamentos. Turing também estava atento a esse fato, dizendo que “haveria muito a fazer na tentativa de entender o que as máquinas estivessem tentando dizer.” (TURING, 1945/2004, p. 475). Computadores e cérebros são instâncias materiais radicalmente diferentes (inorgânico/orgânico) e, portanto, sempre haverá uma diferença qualitativa nas formas de movimento da matéria que ocorrem em um e no outro.

A instanciação material da Máquina de Turing (conceito abstrato) traria a possibilidade, teórica, de construtos identitários singulares a máquinas digitais. A existência dessa possibilidade deixa aberto o caminho para que as máquinas digitais se tornem autônomas e inteligentes, embora isso não signifique que venham a se tornar humanas. É provável que se essa máquina um dia vier a pensar, ao tentarmos ler esses pensamentos, estejamos como que diante do leão de Wittgenstein.

Referências bibliográficas

CHOMSKY, Noam. *Novos horizontes no estudo da linguagem e da mente*. São Paulo: Editora UNESP, 2005.

DENNETT, Daniel C. *Consciousness explained*. New York: Back Bay Books, 1991.

DESCARTES, René. *Discurso sobre o método: para bem dirigir a própria razão e procurar a verdade nas ciências*. Curitiba: Hemus, 1614/2000.

FODOR, Jerry A. *The modularity of mind*. Cambridge, USA: A Bradford Book, 1996.

¹¹ Em seu início, a IA mesclava a abordagem da então incipiente ciência cognitiva com a ciência da computação e tinha como propósito a criação de modelos computacionais para a compreensão da cognição humana..

¹² Considerando que um computador digital teria mais velocidade para tratamento das instruções sintático-formais recebidas.

GÖDEL, Kurt. *On formally undecidable propositions of Principia Mathematica and related systems*. New York: Dover Publications Inc., 1992.

HOFSTADTER, Douglas R. *Gödel, Escher, Bach: um entrelaçamento de gênios brilhantes*. Brasília: Editora Universidade de Brasília; São Paulo: Imprensa Oficial do Estado, 2001.

McCULLOCH, Warren S.; PITTS, Walter. "A logical calculus of the ideas immanent in nervous activity". In: CUMMINS, Robert; CUMMINS, Denise Dellarosa (Org.). *Minds, Brains and Computers: the foundations of cognitive science, an anthology*. Malden, USA: Blackwell Publishers Inc., 2000, 19 p.

TEIXEIRA, João de Fernandes. *Mente cérebro e cognição*. Petrópolis, RJ: Vozes, 2000.

_____. *Filosofia e ciência cognitiva*. Petrópolis, RJ: Vozes, 2004.

TURING, Alan. "Computação e inteligência". In: TEIXEIRA, João de Fernandes. *Cérebros, máquinas e consciência: uma introdução à filosofia da mente*. São Carlos, SP: EDUFSCar, 1996, 16 p.

TURING, Alan. "On computable numbers, with an application to the Entscheidungsproblem" (1936). In: COPELAND, B. Jack (Org.). *The essential Turing: the ideas that gave birth to the computer age*. Oxford: Clarendon Press, 2004. 32p.

_____. "Intelligent Machinery, a Heretical Theory" (1945). In: COPELAND, B. Jack. (Org.). *The essential Turing: the ideas that gave birth to the computer age*. Oxford: Clarendon Press, 2004, 34 p.

_____. "Lecture on the Automatic Computing Engine" (1947). In: COPELAND, B. Jack. (Org.). *The essential Turing: the ideas that gave birth to the computer age*. Oxford: Clarendon Press, 2004, 28 p.

TURING, Alan. "Intelligent Machinery: a report" (1948). In: COPELAND, B. Jack. (Org.). *The essential Turing: the ideas that gave birth to the computer age*. Oxford: Clarendon Press, 2004, 27 p.

_____. "Can automatic calculating machines be said to think?" (1952). In: COPELAND, B. Jack. (Org.). *The essential Turing: the ideas that gave birth to the computer age*. Oxford: Clarendon Press, 2004, 17 p.

_____. "Chess" (1953). In: COPELAND, B. Jack. (Org.). *The essential Turing: the ideas that gave birth to the computer age*. Oxford: Clarendon Press, 2004, 18 p.

VON NEUMANN, John. *O computador e o cérebro*. Lisboa: Relógio d'Água Editores, 2005.

WITTGENSTEIN, Ludwig. *Tractatus Lógico-Philosophicus*. São Paulo: Editora da Universidade de São Paulo, 2001.