

ELA: A QUESTÃO DA INTELIGÊNCIA ARTIFICIAL VISTA PELOS OLHOS DE UM CINEASTA

HER: THE QUESTION OF ARTIFICIAL INTELLIGENCE SEEN THROUGH THE EYES OF A FILMMAKER

Igor Assagra Rodrigues Barbosa ¹
Sergio Nojiri ²

Resumo

O filme “Ela” levanta questionamentos para o campo do direito. Pretende-se desenvolver um em especial, relacionado à inteligência artificial (IA): pode uma máquina pensar ou sentir como seres humanos? Quais as implicações para o jurídico? Alan Turing, Daniel Dennet e John Searle (entre outros) ofereceram argumentos pró e contra a possibilidade de uma IA. Nosso objetivo é modesto: traçar correlações entre o tema da IA e o relacionamento tal qual relatado no filme “Ela”, a partir de um levantamento bibliográfico dessas referências. Busca-se revelar aspectos filosóficos, jurídicos, científicos, principalmente, emocionais e sensíveis, que a linguagem da arte consegue expor.

Palavras-chave: Direito, Cinema, Inteligência artificial, Consciência

Abstract/Resumen/Résumé

The film “Her” raises questions for the law field. We intend to develop especially one related to artificial intelligence (AI): can a machine think or feel like human beings? What are the implications for the legal? Alan Turing, Daniel Dennett and John Searle (among others) have offered arguments for and against the possibility of an AI. Our goal is modest: draw correlations between the AI theme and the relationship portrayed in the film “Her”, from a reference literature review. The aim is to reveal philosophical, legal, scientific aspects, mainly, emotional and sensitive, that the language of art can expose.

Keywords/Palabras-claves/Mots-clés: Law, Film, Artificial intelligence, Consciousness

¹ Graduado em Direito pela UNESP/Franca; Mestrando em Direito pela Faculdade de Direito de Ribeirão Preto – FDRP/USP. igorassagra@usp.br.

² Professor Doutor da Faculdade de Direito de Ribeirão Preto – FDRP/USP. nojiri@usp.br

1. Introdução

O filme “Ela” (HER, Jonze, 2013) é uma espécie de ficção científica romântica com alguns toques cômicos, entremeada por outros dramáticos. É, ao mesmo tempo, um filme doce e melancólico. Sua interessante e criativa trama gira em torno de um escritor de cartas chamado Theodore Twombly (interpretado por Joaquin Phoenix) que está passando por um doloroso processo de divórcio de sua ex-esposa Catherine (Rooney Mara). O filme é ambientado em uma Los Angeles do futuro (ano 2025), com imensos prédios futuristas e deslumbrantes arranha-céus. Neste futuro, não tão distante, a relação das pessoas com computadores se tornou extremamente próxima, quase simbiótica. Em seu dia a dia, Theodore aparece sempre conectado a algum dispositivo computacional, seja ditando cartas (os teclados provavelmente foram extintos no futuro), ouvindo *e-mails* ou jogando vídeo games interativos.

Todo este ambiente tecnológico e futurístico serve como pano de fundo para um estranho e incomum caso de amor. Em um determinado momento do filme, Theodore resolve experimentar um novo sistema operacional, conectado a uma espécie de *smartphone*, com fone de ouvido sem fio. Por intermédio desses artefatos eletrônicos, ele conversa com uma voz feminina chamada Samantha (cuja voz é da atriz Scarlett Johansson). Ela, na realidade, é uma inteligência artificial extremamente sofisticada. No início, Samantha participa da vida de Theodore realizando apenas tarefas domésticas cotidianas como organizar a pasta de *e-mails*. No decorrer do filme, ela passa a auxiliá-lo em outras espécies de tarefas, mais complexas e íntimas. Com uma voz aconchegante e sensual, Samantha passa a desenvolver ações e reações cada vez mais “humanas”. A interação entre ambos parece inevitável. Samantha tem senso de humor, é inteligente (processa dados com rapidez inalcançável para os humanos) e foi programada para desenvolver “emoções”. Theodore, ainda carente e devastado pelo fracasso de seu casamento, termina por apaixonar-se por Samantha.

Surpreendente e original, a trama apresentada em “Ela” pode ser analisada sob os mais diversos aspectos. Se pensarmos no estado em que Theodore ficou após o fracasso de seu casamento e de como as pessoas vivenciam seu dia a dia no ano de 2025, tendemos a achar que a atual forma de interação por mídias eletrônicas (redes sociais do tipo *Facebook* e *Instagram*, plataformas de mensagens como *WhatsApp*, etc.) poderá nos levar, no futuro, a nos relacionarmos emocionalmente com sistemas operacionais. Neste sentido, explorar as diversas implicações que esse filme traz é relevante pois nos faz refletir sobre o tipo de sociedade que estamos nos transformando, tecnológica e solitária.

Apesar das inúmeras relevantes possibilidades discursivas que o filme proporciona, o foco deste trabalho será de explorar a questão da inteligência artificial, a partir do levantamento bibliográfico de autores que enfrentaram esse problema, bem como nas consequências, inclusive filosófico-jurídicas, decorrentes da possibilidade de uma entidade artificial desenvolver uma consciência.

2. O início da inteligência artificial: Turing

Uma das possíveis questões colocadas pelo filme “Ela” é a seguinte: computadores podem pensar, possuir mentes e ter consciência? Em outras palavras, é possível a criação de máquinas que pensem e atuem como seres humanos? Se a resposta for positiva, ela somente reforça a tese da Inteligência Artificial (IA), que se funda na afirmação de que, no fundo, o cérebro nada mais é que uma espécie de computador.

Pioneiro da tese da IA foi o matemático inglês Alan Turing (1912-1954). Uma das mais relevantes perguntas que Turing buscou responder foi a seguinte: Podem as máquinas pensar? Seu artigo *Computing Machinery and Intelligence*, publicado originalmente em 1950 na revista *Mind*, é largamente aceito como precursor da tese da IA.

Chamado inicialmente de “jogo da imitação”¹, o experimento que ficou posteriormente conhecido como “Teste de Turing” (TT) necessitava de três pessoas: (“A”) um homem, (“B”) uma mulher e (“C”) um interrogador, que pode ser de qualquer dos dois sexos. O objetivo do jogo, para o interrogador, é determinar, em relação aos outros dois, qual o homem e qual a mulher. O homem e a mulher são colocados em salas isoladas onde a única forma de comunicação é através de um *teletype* (uma espécie de *e-mail* da época). É permitido ao interrogador fazer perguntas a “A” e “B”, tais como: Será que “X” poderia me dizer qual o comprimento de seu cabelo? O objetivo do jogo, para “A”, é tentar induzir “C” a uma identificação errada, enquanto que, para “B”, é ajudar o interrogador. Nesse sentido, “B” pode utilizar a estratégia de dar apenas respostas verdadeiras, do tipo “eu sou a mulher, não dê ouvidos ao outro”, mas isso é inútil, porque o homem (“A”) pode dar respostas semelhantes. A questão formulada por Turing, no entanto, parte do seguinte pressuposto: o que acontecerá quando uma máquina ocupar o lugar de (“A”) neste jogo? Será que o interrogador decidirá erroneamente com a mesma frequência nas duas formas de jogar (com um homem e com uma máquina)? Esta questão substitui, segundo Turing, a pergunta original: “podem as máquinas pensar”? (TURING, 1973, p. 50-51).

¹ Este também foi o nome dado ao filme interpretado por Benedict Cumberbatch, baseado na biografia de Turing (*The Imitation Game*, TYLDUM, 2014).

Na visão de Turing, se o interrogador for incapaz de afirmar, com um índice de acerto superior a 50%, em qual quarto se encontra cada jogador, podemos concluir que a máquina em questão passou no TT. Um computador capaz de passar no TT, de acordo com Turing, pode ser declarado uma máquina pensante.

Inúmeras questões foram colocadas a partir do TT, muitas delas enfrentadas pelo próprio Turing. Se afirmou, por exemplo, que o pensamento é uma função apenas da alma humana, por ser uma dádiva divina. Turing respondeu que isso envolve uma contradição, uma vez que isso implicaria uma restrição à onipotência de Deus. Afirmou-se, ainda, que uma máquina nada cria de novo, uma vez que ela apenas realiza tarefas que lhe ordenamos e que os homens, pela sua imprevisibilidade, jamais agiriam da forma mecânica como fazem as máquinas. Turing refutou todas essas objeções, concluindo, ao final, ser possível a reprodução artificial da inteligência humana (TURING, 1973, p. 69-75).

A estratégia argumentativa de Turing se baseia no seguinte pressuposto: ao invés de tentar produzir um programa que simule a mente adulta, por que não tentar produzir um que simule a mente infantil? Turing tentou associar a noção de programa infantil com a de processo de aprendizagem. Ele acreditava que não podemos esperar encontrar uma boa máquina-criança logo na primeira tentativa, mas ensinar uma máquina para ver como ela aprende. Deve-se experimentar, como um processo de evolução (TURING, 1973, p. 77-78).

Ronaldo Martins e Mírian dos Santos (2013, p. 84), ao analisarem a posição de Turing, remetem ao comportamento cognitivo do ser humano. Afinal, ele pensa, aprende e cria, porque possuiu um programa biológico que o dotou com essas capacidades, juntamente com a capacidade de alteração pela interação com o ambiente. Este programa não será repetitivo e estereotipado, mas condicionado pelo ambiente. Ele terá que analisar seu próprio desempenho, diagnosticar suas falhas e providenciar mudanças para melhorar sua eficiência futura. Estes autores não possuem dúvidas que as máquinas podem “aprender”, ainda que não o façam de forma equivalente aos humanos. Especialmente no campo do processamento automático de línguas naturais, parece não haver dúvida sobre a possibilidade dos computadores adquirirem conhecimento linguístico por meio de inferência estatística. Isso é comprovado pelas competições entre sistemas promovidas pelo *National Institute of Standards and Technology (Open MT Evaluation, NIST)* (MARTINS; SANTOS, 2013, p. 87). Tome-se, ainda, como exemplo, o *Google Translate*, programa de tradução automática. Ele opera com uma memória “incremental”, cotidianamente atualizada de novas bases de conhecimento, novos glossários terminológicos, memórias de tradução e, especialmente, *feedback* de usuários. Da mesma forma que os *sites* de busca que utilizam parâmetros

dinâmicos como critérios, histórico de páginas visitadas, dados geoestacionários, entre outros, não há simplesmente uma repetição de atos por parte do computador. Os resultados apresentados por essas máquinas dependem, em grande medida, de seu processo de “aprendizado”, que muitas vezes pode ser individualizado e personalizado.

Todavia, conforme veremos a seguir, o problema da IA envolve vários outros aspectos. O mero desenvolvimento de algumas capacidades computacionais não resolve em definitivo a questão.

3. Um computador pode realmente pensar?

É possível reproduzir a consciência humana em máquinas? Esta é a pergunta que muitos filósofos e cientistas, de tempos em tempos, tentam responder. Iremos expor o posicionamento de Daniel C. Dennett, que responde afirmativamente a essa pergunta.

Dennett (1998, p. 120) acredita que a consciência é resultado da atividade de neurônios e módulos de processamento paralelamente distribuídos. A consciência, nesse sentido, não é uma entidade que se encontra em um local definido do cérebro. Para ele, a consciência é uma atividade neuronal que pode ser replicada em um sistema informático. Dessa maneira, afirmar que uma máquina composta de artefatos materiais é incompatível com a consciência, uma entidade imaterial, esconde um dualismo que tenta impedir o estudo científico do cérebro. Apesar da assustadora complexidade orgânica que envolve o cérebro, as coisas não devem ser assim, pensa Dennett.

O autor verifica que ao longo dos séculos, cada fenômeno inicialmente misterioso e sobrenatural sucumbiu a uma explicação científica. Para Thales, o protocientista pré-socrático, por exemplo, a magnetita possuía uma alma. Hoje sabemos exatamente que suas manifestações se dão por conta do magnetismo, um fenômeno bem conhecido. O próprio “milagre” da vida e de sua reprodução hoje são analisados pela biologia molecular. Por que, então, pergunta Dennett, o cérebro é o único objeto físico complexo no universo a ter uma interface com outra esfera de existência? Ao refutar o dualismo, afirma:

Os fenômenos da consciência são reconhecidamente deslumbrantes, mas suspeito que o dualismo jamais seria seriamente considerado se não houvesse uma forte corrente com vontade de proteger a mente da ciência, por supor que ela é composta de um material que é, em princípio, não investigável pelos métodos das ciências físicas (DENNET, 1998, p. 154, tradução nossa).²

² The phenomena of consciousness are an admittedly dazzling lot, but I suspect that dualism would never be seriously considered if there weren't such a strong undercurrent of desire to protect the mind from science, by supposing it composed of a stuff that is in principle uninvestigatable by the methods of the physical sciences.

É certo que a complexidade da bioquímica cerebral é enorme. Sem embargo, a ciência atual é capaz de imitar e suprir o funcionamento de alguns órgãos humanos. O coração artificial, por exemplo, funciona muito bem. Atualmente já se fala até na construção de um olho artificial. A questão proposta por Dennett é basicamente a seguinte: se é possível criar um coração e um olho, por que não um cérebro? Não há, pelo menos do ponto de vista lógico, uma objeção para a construção de módulos cerebrais através de um circuito eletrônico (DENNET, 1998, p. 159).

O que Dennett está querendo dizer é que a consciência e a inteligência nada mais são do que disparos neuronais. Dessa forma, os circuitos cerebrais podem ser replicados por meios informáticos, desde que atinjam os mesmos resultados. Isso significa que os estados mentais não passam de funções executadas pelo cérebro. Fenômenos como a intencionalidade não seriam atributos específicos do “espírito”, mas podem ser “rodados” em outros sistemas. Nesse sentido, quando um sistema “age como se fosse”, significa dizer que ele “está realmente agindo intencionalmente”, não havendo nada de misterioso por trás da consciência.

A teoria da consciência de Dennett é funcionalista. Para ele, estados mentais seriam funções executadas pelo cérebro humano, da mesma forma que o bombeamento de sangue é uma função do coração. O que importa, portanto, não é tanto o órgão, mas a sua função. Se não há problema na substituição de um coração orgânico por um mecânico, qual o problema de transportar funções mentais para as máquinas? E se substituíssemos pedaços do cérebro por peças de silício, metal ou algum outro material para salvaguardar uma determinada função? Para Dennett, considerando a complexidade da estrutura cerebral, o projeto de IA encontra obstáculos mais econômicos do que teóricos. (VIANA, 2013, p. 73-74).

Vale a pena destacar que, para Dennett (1998, p. 159), as experiências ditas subjetivas, como são os sentimentos de dor, prazer, felicidade, fome, etc. podem e devem ser objetos de estudo da ciência. Essas propriedades subjetivas são conhecidas pela palavra latina *qualia*.

Os *qualia*, segundo o filósofo David Chalmers (2010, p. 5-6), trazem consigo uma questão filosófica que, pela dificuldade em resolvê-la, ele a batizou de *Hard Problem*. Todos nós sabemos quando estamos falando sobre sensações, mas é notória a dificuldade que temos em defini-las. Quando, por exemplo, alguém experimenta uma fruta, digamos um abacaxi, e afirma que ele está doce, todos nós sabemos o que essa palavra significa, mas nunca saberemos ao certo se sua experiência subjetiva, de percepção de sabor, é exatamente a mesma que a minha. Talvez a sensação de “doce”, para essa pessoa, seja mais parecida com aquilo que eu posso considerar como amargo. Eis aí o *Hard Problem*.

Dennett se opõe a esse problema. Para ele, os *qualia* são fruto de uma mentalidade “obscurantista”. Postular qualidades internas especiais que são não apenas privadas e intrinsecamente valiosas, mas também que não podem ser confirmadas ou investigadas, para ele, é apenas obscurantismo (VIANA, 2013, p. 72).

Na realidade, Dennett pensa que o *Hard Problem* seria apenas um *Easy Problem*, a saber: encontrar os mecanismos neuronais que correspondam a estados mentais. Para além dos estímulos neuronais não existe a figura de um sujeito, ego ou *self*. Dennett nega, dessa forma, um espaço da subjetividade que concorra com a objetividade dos estímulos. O que ele pretende, de fato, é a elaboração de uma ciência da consciência que abandone a perspectiva em “primeira pessoa” com a finalidade de se obter uma pura objetividade ou neutralidade na pesquisa científica. Isso não quer dizer, no entanto, que a subjetividade não deva ser levada em consideração pela ciência, mas apenas que existem estados mentais privados e internos dos quais a consciência teria um acesso direto e infalível. Sua proposta, assim, é a da perspectiva em “terceira pessoa”, que seria a mais adequada para a ciência e a neurociência estudar a consciência. Daí que o método proposto não pode ser fenomenológico, que analisa a experiência subjetiva revelada à consciência, mas aquilo que Dennett chama de “heterofenomenológico”, capaz de explicar a consciência como um fenômeno produzido por *inputs* e *outputs* eletroquímicos. (VIANA, 2013, p. 73).

4. John Searle e o Argumento do Quarto Chinês

O experimento mental proposto pelo filósofo John R. Searle requer que imaginemos uma pessoa que não saiba chinês dentro de um quarto que contenha uma caneta e um pedaço de papel, juntamente com outros elementos (como um conjunto de regras) que possibilite a manipulação de palavras chinesas. O sujeito dentro do quarto, ao receber por uma abertura perguntas escritas em chinês, deverá respondê-las na mesma língua. Isso se torna possível pela utilização da lista de regras escritas em inglês, por exemplo, que ensina como correlacionar os caracteres em inglês e chinês. Nesse caso, oferecimento de respostas se dá apenas pela sua forma, ou seja, a pessoa adquire a capacidade de relacionar um conjunto de símbolos com outro sem sequer falar ou entender chinês (SEARLE, 2008, p.35).

O ponto defendido por Searle, a partir desse experimento mental, é que o mero seguimento de regras trata-se de uma operação no nível sintático, no qual o nível semântico, da compreensão, não faz parte. Isso quer dizer que, ainda que uma máquina passasse no TT,

ela não estaria “pensando” como seres humanos, uma vez que não teria utilizado a capacidade genuína da compreensão (semântica). Segundo o filósofo:

Na experiência mental, a pessoa na sala, no caso eu, é capaz de simular o comportamento de uma pessoa chinesa de uma maneira que satisfaça o teste de Turing porque, por exemplo, dá respostas corretas para as perguntas chinesas, ao mesmo tempo em que não entende uma única palavra em chinês. Por que não? Porque ela não tem nenhuma maneira de obter, a partir da manipulação dos símbolos, os seus significados, e se a pessoa no quarto não entende os significados dos símbolos em função da execução do programa, então nenhum outro computador, unicamente com base nisso, entenderia, porque nenhum computador, apenas em virtude de suas propriedades computacionais, possui algo mais que o homem no quarto chinês. Além disso, o quarto todo não possui nenhuma forma de conseguir significado a partir de símbolos. O sistema todo não possui um método que permita anexar a semântica à sintaxe dos símbolos do computador (SEARLE, 2008, p. 59, tradução nossa).³

O cerne do argumento do quarto chinês é de que a sintaxe do programa não é suficiente para garantir o conteúdo semântico (ou conteúdo mental ou significado) na mente de um falante chinês. Em resumo, Searle quer dizer que:

1. Programas são sintáticos;
2. As mentes possuem conteúdos semânticos;
3. A sintaxe, por si só, não é a mesma coisa, nem é suficiente para garantir o conteúdo semântico.

Para Searle, a atividade de pensar não consiste apenas em correlacionar símbolos de maneira correta. Há algo que vai além da manipulação de símbolos: a intencionalidade. O programa digital, nesta linha de raciocínio, não é intencional como a mente humana. Por quê?

Porque as manipulações formais de símbolos, por si só, não têm nenhuma intencionalidade; elas são completamente sem sentido; elas não são sequer manipulações de símbolo, uma vez que os símbolos não simbolizam nada. No jargão linguístico, eles possuem apenas sintaxe, mas não semântica. Essa intencionalidade que os computadores parecem possuir existe apenas nas mentes daqueles que os programam e daqueles que os utilizam, daqueles que transmitem na entrada e daqueles que interpretam a saída. (SEARLE, 1980, p. 422, tradução nossa).⁴

³ In the thought experiment the person in the room, namely me, is able to simulate the behavior of a Chinese person in a way that satisfies the Turing Test because, for example, he gives correct answers to Chinese questions, but all the same, he does not understand a word of Chinese. Why not? Because he has no way to get from the manipulation of the symbols to the meaning of the symbols, and if the person in the room does not understand the meanings of the symbols on the basis of implementing the program, then neither does any other computer solely on that basis, because no computer, just in virtue of its computational properties, has anything that the man in the Chinese Room does have. Furthermore, the whole room has no way of getting meaning from the symbols. The whole system has no way to attach semantics to the syntax of the computer symbols.

⁴ Because the formal symbol manipulations by themselves don't have any intentionality; they are quite meaningless; they aren't even symbol manipulations, since the symbols don't symbolize anything. In the linguistic jargon, they have only a syntax but not semantics. Such intentionality as computers appears to have is solely in the minds of those who program them and those who use them, those who send in the input and those whose interpret the output.

Ademais, há uma confusão a respeito da noção de “processamento de informação” (SEARLE, 1980, p. 423-424). Embora se afirme que o computador possa simular aspectos formais e processar informação da mesma forma que um cérebro, há, segundo Searle, uma ambiguidade na noção de informação. O sentido pelo qual as pessoas “processam informação”, quando refletem sobre problemas aritméticos ou quando leem ou respondem questões sobre história, não é o mesmo no qual um computador “processa informação”. Como o computador tem apenas sintaxe e não semântica, você pode perguntar a ele quanto é 2+2 e ele responderá 4, mas ele não tem a menor ideia do que é 4 e o que significa 4 ou que isso signifique alguma coisa. Esses símbolos, chamados de “primeira ordem”, não são interpretados. São apenas símbolos. Assim, em um primeiro sentido, o computador não processa informação, apenas manipula símbolos formais. Em um segundo sentido, apesar de o computador processar informação, é somente no sentido em que máquinas de somar, de escrever, estômagos e termostatos fazem: nós podemos descrevê-los como recebendo informação, transformando-a e produzindo informação como *output*.

Para Searle (1980, p. 424), portanto, a intencionalidade é um fenômeno biológico causalmente dependente da bioquímica específica de suas origens. Dessa forma, não é possível aceitar a tese de que a capacidade causal do cérebro, para produzir intencionalidade, possa consistir na instanciação de um programa de computador. Um programa de computador, para o autor, não é por si só suficiente para produzir intencionalidade.

E quanto ao argumento de Dennet, de que devemos utilizar um método científico que permita estudar os eventos mentais do ponto de vista de uma “terceira pessoa”?

Esse é o erro mais profundo e é a fonte da maioria dos outros cometidos por Dennet, segundo Searle (1998, p. 130). Ele chama atenção para a necessária distinção entre o sentido epistêmico e o ontológico de primeira e de terceira pessoa. Algumas afirmações podem ser conhecidas como verdadeiras ou falsas independentemente de quaisquer preconceitos ou posicionamentos por parte dos observadores. São objetivos no sentido epistêmico. Se afirmarmos, por exemplo, que Van Gogh morreu em *Auvers-sur-Oise*, na França, essa afirmação é epistemicamente objetiva, uma vez que independe de preconceitos ou preferências pessoais. Mas, se dissermos que Van Gogh foi mais talentoso que Renoir, essa afirmação passa a ser epistemicamente subjetiva, pois depende, ao menos em parte, de preferências pessoais. No entanto, há, ainda, um aspecto ontológico, para o qual a existência objetiva não depende de qualquer sujeito, como é o caso da existência de uma montanha. Outras são subjetivas, no sentido de que sua existência depende de ser sentida por um sujeito, como, por exemplo, uma dor de dente. O cerne da questão, afirma Searle, está em que, apesar

de a ciência buscar objetividade epistêmica, um conjunto de verdades livres de nossas preferências pessoais e preconceitos, a objetividade epistêmica do “método” não exige a objetividade ontológica do objeto. Conclui Searle:

É apenas um fato objetivo – no sentido epistêmico – que eu e pessoas como eu tenham dores. Mas o modo de existência destas dores é subjetivo – no sentido ontológico. Dennett tem uma definição de ciência que exclui a possibilidade de ela poder vir a investigar a subjetividade, e ele acredita que a objetividade de terceira pessoa da ciência o força a tal definição. Mas esse é um mau trocadilho com ‘objetividade’. O alvo da ciência é explicar sistematicamente o funcionamento do mundo. Uma parte do mundo consiste em fenômenos ontologicamente subjetivos. Se tivermos uma definição de ciência que nos proíba de investigar esta parte do mundo, é a definição que deve ser mudada, não o mundo (SEARLE, 1998, p. 132-133).

5. Máquinas e cérebros

As máquinas podem executar desde simples tarefas como um cálculo matemático (calculadoras) a atividades mais complexas como vencer um jogo de xadrez de um campeão mundial (*Deep Blue* versus Garry Kasparov). A cada dia, torna-se cada vez mais evidente a capacidade de execução, pelas máquinas, do maior número possível de funções mentais humanas. Ações que antes somente as pessoas podiam executar são atualmente realizadas por máquinas dos mais diversos tipos. Computadores participam de jogos, possuem percepção visual, controlam outras máquinas e, em certos casos, manipulam uma linguagem. Essa abordagem, puramente pragmática, é compatível com aquilo que cientistas e filósofos chamam de tese da Inteligência Artificial (IA) fraca. O seu contrário, a IA forte, parte do pressuposto de que um computador, desde que programado com um programa suficientemente sofisticado, pode possuir uma “mente” (MATTHEWS, 2007, p. 93).

O filme de Spike Jonze, nesse sentido, retrata uma realidade de IA forte.

Um computador com uma IA forte não somente seria capaz de responder que 3.456 vezes 2.815 é 9.728.640, mas que o aborto é moralmente certo ou errado por razões desta ou daquela natureza. E mais, quando responder a esta última pergunta, ele pode compreender os aspectos éticos e morais envolvidos no problema. Possuir uma mente, neste sentido, não é apenas ser capaz de responder corretamente certas perguntas, mas possuir um entendimento acerca delas.

A grande questão é a seguinte: por que o pensar pode ser somente algo executado por seres humanos e não, por exemplo, por um mecanismo construído? Ainda que partamos do pressuposto de que os atuais computadores não pensem como seres humanos, ou seja, seus atuais programas não possuam mentes, não seria possível, no futuro, o desenvolvimento de

tecnologias de uma nova geração (provavelmente de um tipo radicalmente diferente das atuais) que pudesse operar como uma mente humana?

Eric Matthews dá uma resposta interessante a essa pergunta. Ele parte da noção de pensamento como uso de uma linguagem. Este uso faz com que a linguagem obtenha seu significado a partir de conceitos partilhados pela totalidade de uma comunidade linguística. O uso das palavras, com significado, depende de certas regras que ditam o uso correto ou incorreto do termo em questão. Por essa razão, é impossível que exista uma linguagem puramente privada. Não posso, dessa forma, atribuir a certas palavras o sentido que “eu” quiser, sob o risco de não terem qualquer significado. Os significados das palavras são compartilhados. E o que isso tem a ver com o pensamento? Se o pensamento envolve considerar propostas com significado, então possuir um significado não pode ser simplesmente uma questão de algo que está se passando dentro do indivíduo privadamente. A compreensão das palavras exige interação com outros seres humanos do mundo exterior. Simplificando: é preciso aprender o que as palavras significam e o processo de aprendizado exige um uso que pode ser corrigido por outros (MATTHEWS, 2007, p. 102-103).

Neste ponto, Matthews concorda com Searle. As máquinas não podem pensar apenas porque seguem regras sintáticas (ordenar símbolos de formas particularmente aceitas). Elas não pensam porque são incapazes de semântica. E isso nada tem a ver com os materiais com os quais as máquinas são compostas, mas com a ausência de participação na sociedade, com a falta de atividades compartilhadas, seguidas de propósito. Elas não possuem um propósito próprio. Elas são, no sentido comum, mecanismos construídos pelos seres humanos para auxiliá-los no alcance de certos objetivos.

A grande diferença reside justamente aí: as pessoas, diferentemente das máquinas, possuem propósitos próprios. E esses propósitos, segundo Matthews, muitas vezes decorrem de necessidades biológicas inatas, da própria natureza, como um organismo biológico. Segundo o autor:

A maioria dos seres humanos, por exemplo, possui a intenção de fazer relacionamentos, como a amizade com outros seres humanos, que, por sua vez, estão obviamente relacionados à necessidade de se manter vivo ou de reproduzir a espécie: a forma particular assumida por esses relacionamentos parece ser, de fato, determinada bem mais pela cultura do que pela biologia. Não obstante os seres humanos podem ter tais propósitos, ou possuir uma cultura apenas porque são criaturas biológicas – seres humanos que agem de certas maneiras e podem, portanto, engajar-se em atividades em comum com outros membros da mesma espécie. Baseado nisso, seria possível argumentar que o pensar e a mente, em geral, podem pertencer apenas às criaturas biológicas, pois, somente em seu caso, conforme argumentado anteriormente, é possível existir propósitos intrínsecos à existência do ser em questão e, logo, conferir um significado baseado em objetos e em seu ambiente (MATTHEWS, 2007, p. 104).

O próprio Matthews, no entanto, reflete sobre a possibilidade de um dia construirmos máquinas modeladas no cérebro humano ou, ainda, no cérebro e no corpo humano. Máquinas com propósitos próprios que interagissem com outros seres humanos ou com outras máquinas. Parece, segundo Matthews, não haver impossibilidade lógica sobre a ideia de construir coisas vivas neste sentido. Mas, logo ele afasta essa ideia. Apoiado em Merleau-Ponty, afirma que a subjetividade é essencialmente incorporada a nós mesmos e que somos “subjetividade-em-um-corpo” ou “subjetividades incorporadas”:

Isso significa que, ser um sujeito ou possuir uma mente, não pode ser separado daquilo que significa ser um ser humano. No entanto, igualmente, ser um ser humano do nosso tipo significa possuir uma mente: a biologia humana deve ser compreendida, ao menos em parte, como intencional e subjetiva, pois nossos corpos são criaturas que podem relacionar-se ao mundo dessas formas. Eis porque o cérebro não é uma máquina. Não é um mecanismo pelo qual o ser humano alcança seu objetivo, e muito menos algo que leva o ser humano a agir de certas maneiras. Em vez disso, ele é uma parte essencial do ser humano e seus funcionamentos fazem parte da forma como o ser humano se relaciona com o mundo e atinge suas aspirações (MATTHEWS, 2007, p. 104-105).

6. IA e o direito

O aumento da capacidade dos computadores e a diminuição dos custos de armazenamento de dados, em conjunto com os significativos avanços no processo de aprendizagem das máquinas, torna a IA mais próxima de nossa realidade atual.

Por mais de 40 anos a velocidade dos transistores das unidades centrais de processamento (CPU's) tem dobrado a cada doze a dezoito meses. Este fato, que propiciou uma extraordinária inovação tecnológica, foi prevista por Gordon Moore, o fundador da Intel Corporation, que descreveu, em um conhecido artigo de 1965, aquilo que ficou conhecido como Lei de Moore. Naqueles primeiros anos, dobrar era aumentar de 5.000 a 10.000 hertz. Atualmente, o aumento da capacidade dos transistores é da ordem de bilhões (KATZ, 2013, p. 914).

Mas a velocidade da CPU não é única responsável pela rápida expansão da fronteira dessas possibilidades. Igualmente importante tem sido o rápido e consistente declínio no custo de armazenamento de dados. A lei de Kryder (o análogo de armazenamento para a Lei de Moore) sustenta que a diminuição dos custos de armazenamento de dados segue um padrão semelhante, se não superior, ao ritmo do aumento correspondente na capacidade do transistor. Para termos uma ideia disso, tomemos uma unidade muito comum de medição, familiar para muitos, que é o gigabyte. Um gigabyte pode armazenar 10^9 (1.000.000.000) bytes de informação. Em termos de conteúdo da informação, um gigabyte equivale a cerca de sete minutos de um vídeo em alta definição (HD) ou cerca de dezoito metros de livros em uma

típica prateleira. O preço de um gigabyte diminuiu muito rapidamente ao longo das últimas décadas. Em 1981, um gigabyte custava cerca de US\$ 300,000, em 1997, custava cerca de US\$ 100 e em 2011 cerca de US\$ 0,10. Em termos de escala, não se utiliza mais o gigabyte, mas o terabyte (10^{12} ou 1.000.000.000.000 bytes) e o petabyte (10^{15} ou 1.000.000.000.000.000 bytes). Um petabyte, por exemplo, equivale a mais de treze anos consecutivos de vídeo em HD e cinquenta petabytes é mais ou menos equivalente a toda produção escrita da humanidade registrada pela história, desde seu início, em todas as línguas (KATZ, 2013, p. 916-917).

Todavia, capacidade de armazenamento de dados e velocidade de processadores não são, por si só, suficientes a garantir uma “revolução” no desenvolvimento da Inteligência Artificial (AI). Temos ainda um longo caminho a percorrer, especialmente na compreensão do funcionamento do cérebro humano. Tomemos o exemplo do iCub. Ele é um robô humanoide criado na Universidade de Plymouth, na Inglaterra, e foi projetado para agir e aprender como uma criança humana. A ideia é fazer com que o robô se desenvolva tal qual uma criança – interagindo com o mundo, imitando e aprendendo através de exemplos (EAGLEMAN, 2015, p. 186). Todavia, será que os robôs poderiam aprender dessa maneira?

O iCub tem o tamanho de uma criança de 2 anos. Ele possui olhos, ouvidos e sensores de toque que o permitem interagir e aprender sobre o mundo. Se apresentamos um novo objeto ao iCub, ele o nomeia (“isto é uma bola vermelha”), seu programa computacional correlaciona a imagem visual do objeto com a correspondente qualificação verbal. Assim, da próxima vez que você apresentar uma bola vermelha e perguntar o que é isso, ele irá responder: uma bola vermelha.

Mas frequentemente as coisas não funcionam perfeitamente. Se você apresentar e nomear diversos objetos e pressionar o iCub a dizer o nome de todos eles, ele provavelmente irá errar muitas vezes e responderá, em um número grande de vezes, “eu não sei”.

O neurocientista David Eagleman passou um bom tempo com iCub. Segundo relata, quanto mais tempo ficava com ele, mais se convencia que não havia nenhuma mente por trás do programa de computador. Apesar de seus grandes olhos, voz amigável e movimentos semelhantes aos de uma criança, era evidente que iCub não possuía sentimentos ou emoções (EAGLEMAN, 2015, p. 187-188).

O físico Michio Kaku foi apresentado a um dos mais avançados robôs do mundo, chamado ASIMO, fabricado no Japão (onde 30% de todos os robôs industriais são fabricados) pela Honda Corporation. ASIMO, do tamanho de um jovem, pode andar, correr, subir escadas, falar diversas línguas e dançar. Kaku interagiu com o robô algumas vezes em

programas de televisão, tendo ficado muito impressionado com suas habilidades. Ao conhecer os criadores de ASIMO, perguntou: o quão inteligente é ASIMO comparado a um animal? Os criadores admitiram que ele possuía a inteligência de um besouro. O problema é que ASIMO era, no final das contas, um grande gravador. Ele possuía apenas uma lista modesta de funções autônomas. A grande maioria de suas falas ou movimentos tinham de ser cuidadosamente ensaiados anteriormente (KAKU, 2014, p. 219).

Os robôs podem ser muito bons em simular o futuro em apenas um campo específico, como em um jogo de xadrez, mas os cérebros não são como computadores digitais. Se você remover um único transistor de um chip Pentium o computador ficará totalmente comprometido, enquanto que o cérebro humano pode atuar razoavelmente bem mesmo que sua metade esteja faltando. Isso acontece porque o cérebro não é, de forma alguma, um computador digital, mas um sofisticado sistema neural. Diferente de um computador, que possui uma arquitetura fixa (input, output e processador), a rede de neurônios estão constantemente se reconectando e consolidando após o aprendizado de uma nova tarefa. O cérebro não possui um programa, um sistema operacional ou um processador central. Ao invés disso, as redes neurais possuem cem bilhões de neurônios disparando ao mesmo tempo, com a finalidade de alcançar um único objetivo: aprender (KAKU, 2014, p. 220).

Embora esses argumentos sejam convincentes, autores como Daniel Martin Katz (2013 p. 929) apostam em uma tecnologia que pode responder, de forma muito mais eficiente que os advogados humanos, perguntas como: Tenho chances de ganhar a ação? Qual a orientação jurídica mais favorável? Quanto vai custar este processo?

Katz menciona, por exemplo, a TyMetrix, uma empresa que desenvolve sistemas de tecnologia da informação (TI) projetados para melhorar o desempenho das operações internas, e fornecer soluções de dados que dão aos profissionais da área jurídica uma vantagem de informações em qualquer cenário. Entre as suas ofertas de produtos há uma plataforma de análise jurídica que fornece os gastos legais e dados de desempenho de todo um setor, que podem ser usados por clientes para determinar uma taxa aceitável a pagar por um determinado serviço jurídico (KATZ, 2013, p. 930-931).

Mas, provavelmente, o mais próximo que temos de uma IA, no direito, seja ROSS, um programa altamente sofisticado projetado sobre a base do supercomputador Watson, da IBM. Pergunte a ele sobre uma decisão judicial obscura de anos atrás e ele lhe responde instantaneamente com opiniões de como aquela decisão pode ser relevante para o caso que está sendo julgado no momento. E o mais interessante é que sua resposta é similar à de uma

pessoa, e não fragmentos com palavras chaves. ROSS irá “advogar” para o escritório americano Baker & Hostetler, especializado em casos de falência. Contudo, vários outros escritórios já estão interessados em Ross (LIBERATORE, 2016).

Isso significa que no futuro, nós, da área jurídica, seremos substituídos por computadores como ROSS? Por enquanto, tudo indica que não. Embora o armazenamento de dados e a velocidade dos processadores dos computadores possam atingir níveis impressionantes, há um elemento fundamental ausente nos robôs: a emoção. Neurocientistas, como Antônio Damásio, descobriram que quando a ligação entre o lobo pré-frontal (que governa o pensamento racional) e o centro das emoções (p. ex. o sistema límbico) encontra-se danificada, os pacientes são incapazes de realizar julgamentos de valor. Eles ficam paralisados diante de simples escolhas (que coisas comprar, quando marcar uma reunião, que cor de caneta usar) porque para eles tudo tem o mesmo valor. Isso significa que as emoções não são supérfluas, elas são absolutamente essenciais e sem elas um robô terá dificuldades em determinar o que é importante e o que não é. Assim, as emoções não são, portanto, periféricas para o progresso da IA, elas são centrais (KAKU, 2014, p. 230-231).

O problema com as emoções é que elas são, em algumas ocasiões, irracionais, enquanto que os programas computacionais são matematicamente precisos. Nossa consciência é influenciada por peculiaridades de nosso passado evolucionista que os robôs não possuem. Robôs podem não levar em consideração pistas sutis que os humanos utilizam quando se encontram, como, por exemplo, a linguagem corporal (KAKU, 2014, p. 231-232).

Ainda falta muito para descobrirmos como, de fato, o cérebro funciona e como reproduzir esse funcionamento em outra estrutura. Mas uma coisa é certa: estamos no início de algo que ainda não sabemos ao certo o que é. Vivemos um momento único da história, no qual a ciência do cérebro e a tecnologia evoluem juntas. O que virá desta interseção está prestes a mudar quem somos (EAGLEMAN, 2015, p. 201).

7. Conclusões

O filme “Ela”, apesar de também tratar do tema da IA, expande os limites discursivos de campos como a filosofia, a ciência e o direito. O filme é, antes de tudo, uma estória sobre relacionamento (humano?). A trama que envolve Theodore, um solitário escritor de cartas que se apaixona por um sistema operacional, à primeira vista, pode parecer tola e improvável. Mas ao assistir o filme, vamos percebendo que a carência afetiva de Theodore, preenchida por uma voz saída de um aparelho eletrônico, não é tão absurda ou distante

daquilo que a maioria das pessoas realiza cotidianamente, que é ficar conectada a um dispositivo eletrônico digitando palavras, vendo imagens, filmes ou simplesmente conversando. Aplicativos como *WhatsApp*, por exemplo, fazem com que as pessoas se curvem diante de seus aparelhos, com olhos fixos na tela, fazendo com que elas simplesmente se desliguem do mundo (real) ao seu redor para, mentalmente, se interligar em uma outra realidade, denominada virtual.

Ao longo do artigo buscamos esclarecer posicionamentos relevantes acerca da capacidade de máquinas dotadas de Inteligência Artificial (IA) desenvolverem uma consciência própria.

O “jogo da imitação” de Turing (posteriormente chamado de “Teste de Turing”) é tido como o marco inicial da controvérsia. A partir daí, inúmeras possibilidades foram aventadas. A da reprodução da consciência humana em máquinas, que poderiam ser plenamente replicados em sistemas informáticos, por exemplo, foi proposta por Daniel Dennett. John Searle, por sua vez, afirma que a consciência não pode ser pensada apenas num nível sintático, em razão de também possuir um nível semântico, que corresponde à capacidade de compreensão e entendimento.

No direito, propostas mais pragmáticas, que consideram, principalmente, as previsões judiciais quantitativas, utilizam programas como a *TyMetrix*, que buscam soluções para impasses nos quais as máquinas seriam capazes de apresentar, de forma mais eficiente do que os advogados humanos, reais chances de se ganhar uma ação, o melhor instrumento jurídico para a proteção de determinado direito e os custos totais de um processo. Outro surpreendente programa de IA extremamente desenvolvido para o cenário do direito é o programa *ROSS*. Você pode fazer perguntas em linguagem coloquial, como faria a um colega, e *Ross*, em seguida, pesquisa todo o corpo de leis e lhe devolve uma resposta com marcações tópicas da legislação, jurisprudência e outras fontes secundárias para a resolução do caso de forma rápida e eficiente.

Mas o impressionante avanço do desenvolvimento tecnológico não representa, por enquanto, uma verdadeira ameaça para os seres humanos que operam o direito, pois falta às máquinas um elemento indispensável: a emoção. Enquanto os robôs são programados para pensar de modo binário, matemático, as emoções surgem, em determinadas situações, de processos irracionais, difíceis de serem racionalmente explicados.

Tomemos o exemplo mostrado no filme “*Ela*”, que se passa em uma sociedade futura, tecnologicamente desenvolvida, em que Theodore, só e desiludido por conta de um amor fracassado, faz uso de um programa de inteligência artificial, *Samantha*, com o qual

acaba se envolvendo emocionalmente. Ele passa mais tempo interagindo com um sistema operacional do que com pessoas reais.

Esta descrição não parece estar muito longe de nossa realidade e o comportamento emocional de Theodore poderia levantar questionamentos acerca de sua sanidade mental ou de seu bom comportamento social.

Um programa de IA, por mais avançado que seja, dificilmente seria capaz de realizar uma adequada compreensão do comportamento de Theodore. Sofisticados algoritmos que proporcionam rápidas e numerosas previsões quantitativas dificilmente compreendem as sutilezas e a idiossincrasia das ações e emoções humanas. Um dos (muitos) diálogos inteligentes do filme serve como exemplo da complexidade do comportamento afetivo humano. A personagem Amy (Amy Adams), amiga de Theodore, em um certo momento, quer saber se ele está realmente se apaixonando por um sistema operacional:

Amy: Você está se apaixonando por ela?

Theodore: Sim. Isso faz de mim um cara esquisito?

Amy: Não...

Amy: Quer dizer, eu acho que todo mundo que se apaixona é esquisito. É uma espécie de insanidade socialmente aceitável (HER, Jonze, 2013, tradução nossa)⁵

Essa resposta, tão certa e que resume em tão poucas palavras o sentido da paixão, dificilmente seria dada por uma máquina, por mais rápida e inteligente que seja.

8. Referências

CHALMERS, David. Facing up to the problem of consciousness. In: **The Character of Consciousness**. Nova York: Oxford University Press, 2010.

DENNETT, Daniel C. **Brainchildren**: essays on designing minds. Massachusetts: The MIT Press, 1998.

EAGLEMAN, David. **The Brain**, Nova York: Pantheon Books, 2015.

HER. Direção: Spike Jonze. Fotografia: Hoyte Van Hoytema. Sony Pictures, 2013. DVD (126 min), NTSC, color.

KAKU, Michio. **The Future of The Mind**: the scientific quest to understand, enhance, and empower the mind. Nova York: Doubleday, 2014.

KATZ, Daniel M. Quantitative Legal Prediction – or – how I learned to stop worrying and start preparing for the data-driven future of the legal services industry. In: **Emory Law Journal**, Vol. 62, 2013. Disponível em:

⁵ Amy: Are you falling in love with her?

Theodore: Yeah. Does that make me a freak?

Amy: No...

Amy: I mean, I think that everyone who falls in love is a freak. It's a kind of socially acceptable insanity.

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2187752. Acesso em: 09 mar. 16.

LIBERATORE, Stacy. Your AI lawyer will see you now: IBM's ROSS becomes world's first artificially intelligent attorney. In: **MailOnline** - news, sport, celebrity, science and health stories, 2016. Disponível em: <http://www.dailymail.co.uk/sciencetech/article-3589795/Your-AI-lawyer-IBM-s-ROSS-world-s-artificially-intelligent-attorney.html>. Acesso em 21 mai. 16.

MARTINS, Ronaldo; SANTOS, Mírian. Podem as máquinas falar? In: DIAS, Cristiane. **Formas de mobilidade no espaço e-urbano: sentido e materialidade digital** [online]. Disponível em: http://www.labeurb.unicamp.br/livroEurbano/volumeII/arquivos/pdf/eurbanoVol2_RonaldoMartins_MiriandosSantos.pdf. Acesso em 15 jun.15.

MATTHEWS, Eric. **Mente: Conceitos-Chave em Filosofia**, tradução de Michelle Tsé, Porto Alegre: Artmed, 2007.

SEARLE, John R. Minds, brains, and programs. In: **The Behavioral and Brain Sciences**. vol. 3, p. 417-457, 1980.

_____. **O Mistério da Consciência**, tradução de André Yuji Pinheiro Uema e Vladimir Saflate. São Paulo: Paz e Terra, 1998.

_____. The Turing Test: fifty-five years later. In: **Philosophy in a New Century: Select Essays**. Cambridge: Cambridge University Press, 2008.

THE IMITATION GAME. Direção: Morten Tyldum. Fotografia: Óscar Faura. Diamond Films, 2014. DVD (115 min), NTSC, color.

TURING, Alan. Computadores e inteligência. In: EPSTEIN, Isaac (org.). **Cibernética e Comunicação**. São Paulo: Cultrix, 1973.

VIANA, Wellistony C. Técnica e inteligência artificial: o debate entre J. Searle e D. Dennett, **Pensando – Revista de Filosofia**, vol. 4, n. 7, 2013. Disponível em: <http://www.ojs.ufpi.br/index.php/pensando/article/viewFile/1326/1073>. Acesso em: 23 fev.15.