

Por que a Inteligência artificial não faz justiça? - 2



Por **ARI MARCELO SOLON & ALAN BRAGANÇA WINTHER***

Os fundamentos da ciência da computação e da filosofia do direito mostram que a Inteligência Artificial é estruturalmente incapaz de realizar justiça, pois esta exige historicidade, interpretação contextual e uma “variável caótica” humana que transcende a mera racionalidade algorítmica

Atualmente a inteligência artificial tem ocupado cada vez mais espaço nas sociedades, seja como ferramenta para melhorar a eficiência do trabalhador ou em alguns casos mais perversos substituindo suas atividades fim e ocupando seu espaço de trabalho.

Discussões mais recentes começam a teorizar qual seria seu limite, pois seja do ponto de vista técnico ou ético em algum momento deve existir uma atividade que seja insubstituível para a máquina, devendo o ser humano continuar exercendo-a.

Este texto explora do ponto de vista técnico a substituição de inteligência artificial pelas decisões judiciais. Para tanto, ele inicia apresentando autores consagrados da teoria computacional, como John von Neumann e Alan Turing, considerados os pais do computador moderno.

Por não se tratar de um texto com viés matemático, as fórmulas algébricas que justificam suas teorias serão apenas citadas e remetidas aos textos originais dos mesmos autores, cabendo ao leitor, caso tenha interesse em uma comprovação pormenorizada, buscar essas fontes. O enfoque do texto é em suas conclusões e explicação, mesmo que breve, como essas teorias se aplicam no contexto atual da inteligência artificial.

Posteriormente, será tratado no texto a relação dessas teorias com o direito, buscando explicações, mais uma vez baseada em autores consagrados do gênero como Savigny e Hegel, o entendimento do sistema jurídico teórico e suas motivações. Para que então sejam postos em contraste os limites teóricos computacionais previstos nos primeiros tópicos com as necessidades de uma justiça saudável e com viés social.

Teoria da decisão

Conforme o andamento processual, chega a hora que o juiz precisa tomar uma decisão. Não se pretende adentrar nos tipos diferentes que elas podem assumir, como interlocutória ou de mérito, mas no ato do juízo de acatar uma das teses apresentadas pelas partes (ou uma híbrida) e tomar esse fato como verdade, deliberando sobre sua ocorrência no mundo material.

Sobre a decisão em si, pode-se adotar como embasamento filosófico a Teoria da Decisão, de von Neumann, conforme descrita em seu livro *Theory of Games and Economic Behavior*. Para tanto, são necessárias algumas considerações.

O modelo matemático da teoria dos jogos leva em consideração que sua aplicação se dará em um sistema econômico. Isso

a terra é redonda

significa que não é uma ciência exata, mas uma aplicação matemática que estuda os fenômenos para tentar prever o que ocorrerá em determinadas circunstâncias. As partes envolvidas na tomada de decisão também seriam agentes econômicos, ou seja, buscam a otimização racional para que façam a escolha.

Uma maneira simples de entender como a decisão racional é tomada é a partir do “problema da loteria”, que pode ser explicada resumidamente da seguinte forma: Tome-se como hipótese um jogo com 20% de chance de sucesso e que irá retornar o valor apostado em 4 vezes, deve-se participar desse jogo? A resposta simples é que a decisão racional diria para não participar do jogo, pois considerando X como o valor apostado, e as chances de vitória (20%) e derrota (80%) mutuamente excludentes, o resultado seria um retorno de $0,8X$, que é menor do que $1X$. Ou seja, para cada 1 real apostado, o retorno é de 80 centavos.

Mudando-se os parâmetros, pode-se tornar o jogo vantajoso ao participante. Por exemplo, se a chance de sucesso for 30% com o mesmo retorno de 4 vezes, para cada 1 real apostado espera-se receber 1,20. Rapidamente conclui-se que sendo P a probabilidade de vitória e A o coeficiente que multiplica o valor apostado, o jogo gera lucro ao apostador quando $P \times A > 1$ e deve ser evitado quando $P \times A < 1$.

Mesmo nesse modelo simplificado, com apenas duas variáveis, muitas vezes não é possível conhecer com exatidão os valores dos casos concretos. Ou seja, o ator racional da consideração inicial toma sua decisão influenciado por fatores que desconhece, valendo-se de sua experiência ou intuição. Como confirma von Neumann sobre o comportamento racional, *“the significant relationships are much more complicated than the popular and the ‘philosophical’ use of the word ‘rational’ indicates”*.

Retomando a decisão jurídica utilizada por juízes, pode-se traçar um paralelo. Eles conhecem o que foi apresentado no caso, mas não a verdade completa (o peso de cada variável) e buscam a otimização da justiça na sociedade, em comparação com os economistas que estudam o lucro. Porém por não terem acesso sequer a todas as variáveis do caso (muito menos conseguirem ponderar o valor delas) precisam de outras ferramentas além da racionalidade pura para chegarem a uma decisão.

Máquinas computacionais

Alan Turing ficou famoso como um dos pais da computação, pela criação da Enigma, o sistema que possibilitou aos aliados decifrar a criptografia usada pela Entente durante a Segunda Guerra. Seu funcionamento usa como teoria uma Máquina de Turing.

Esse modelo deu origem ao computador moderno. Ele é um dispositivo chamado de universal, pois permite que a partir de um único equipamento (*hardware*), seja criado um algoritmo (*software*) para realizar operações lógicas e resolver um problema concreto.

De maneira simplificada, a técnica dessa máquina consiste em uma fita – aqui definida como um dispositivo unidimensional fragmentado em n seções, cada uma capaz de armazenar um único símbolo – que possa ser lida e possui um significado. A partir desse par pré-configurado a máquina executa uma tarefa, que de forma resumida são: deslocar a fita para a direita, deslocar para a esquerda, imprimir um novo símbolo ou ler alguma seção específica da fita. Quando uma máquina é capaz de realizar essas operações ela é chamada de *“Turing complete”*.

Porém, mesmo com esse protótipo de computador moderno, existia o “problema da decisão”, ou seja, a possibilidade de prever se certo algoritmo conseguiria ser válido independentemente da sua entrada, ou existiria algum comando que o faria travar sem dar resultado ou continuar executando a operação de maneira infinita.

Alan Turing resolve tal questão em seu artigo *“On computable numbers, with an application to the Entscheidungsproblem”*, mostrando que o problema não pode ser resolvido.

Retornando ao mundo jurídico, tal teoria é importante para que seja considerado o limite do uso de inteligência artificial. Todo processo jurídico necessita de um fim, de preferência com uma análise material e uma decisão tomada. Porém, como mostrado por Turing, é impossível prever se determinado algoritmo pode ser calculável, o que ao ser utilizada uma inteligência artificial para o andamento processual pode resultar em erro ou infinitude do processo em questão.

Tal argumento pode parecer como uma ressalva que ocorreria apenas em casos mais complexos. Então, supondo que todos processos sejam calculáveis, surge outro problema de como a decisão é tomada. Para isso, cabe lembrar sobre o chamado Teste de Turing.

O teste, chamado pelo autor de “jogo da imitação” consiste em fazer uma máquina em, dialogando apenas através de texto, convencer um interlocutor que é um ser humano. Ele conclui que se uma máquina for capaz de se passar por um ser humano, então ela consegue pensar como um.

No mesmo artigo, Alan Turing apresenta algumas questões como que apesar dessa aparência de diálogo humano, ela nunca poderia pensar como um, trazendo argumentos desde o ponto de vista teológico até o de se comportar como um oráculo com percepção extra-sensorial.

Para que uma inteligência artificial consiga fazer justiça, é necessário o componente humano, como será visto posteriormente no presente artigo. Não basta imitar que seja um, mas é necessário realizar inferências muitas vezes inéditas, ou seja, sem base para que possa copiar.

Empatia da máquina

Em 1966 é publicado um artigo que apresenta o programa de computador chamado ELIZA, que por décadas foi considerado o mais próximo de resolver o Teste de Turing.

Para elaborar o algoritmo, seu criador se baseou na “*client-centered therapy*” de Carl Rogers, segundo a qual o terapeuta ao conversar com o paciente destaca palavras ou temas centrais e pede para que o paciente se aprofunde no assunto específico, assim o próprio encontraria uma resposta, ao invés de modelos anteriores que dependiam do terapeuta fazer a interpretação e chegar a uma conclusão.

A ELIZA realizava tarefas parecidas, ao receber o texto de um cliente (lembrando que o Teste de Turing originalmente é realizado por escrito) ela destaca as palavras centrais do tema e demonstra interesse no assunto, deixando que o interlocutor continue falando sobre o tema, até surgir uma ideia central diferente. Outra novidade é que ela armazenava esses temas (lembrando da Máquina de Turing, ela registra cada palavra chave em uma seção de sua fita de memória). Com isso, consegue fazer inferências, como no diálogo mais famoso dela, que consegue associar a relação dela com o pai com o pedido do namorado que ela faça terapia.

Esse experimento apesar de sucesso no campo da psicanálise – sucesso tamanho que seu criador decidiu desligar o programa –, ainda não resolvia o Teste de Turing, exatamente pela questão citada anteriormente, pois caso o interlocutor falasse algo sem ideia central, como um “*like what?*” ou a interlocução “*huh?*”, a resposta se tornava algo sem sentido.

Mesmo com essa limitação, o experimento demonstra como demonstrar empatia (no sentido de ouvir ativamente o interlocutor) é o que considerava a resposta mais parecida com o ser humano, uma imitação mais próxima do que se tivesse a capacidade de pensar como um.

A limitação também fica clara ao considerar que embora pareça um diálogo, apenas uma das partes (a humana) traz informações à conversa, sendo a outra apenas uma “parede que rebate as bolas de tênis”.

Considerando-se os experimentos mais recentes, a solução encontrada ainda não é suficiente. Por exemplo, o experimento

chamado “quarto chinês” de Searle. Nele é ensinada uma linguagem – no caso, chinês – para uma máquina. Ela então é colocada para se comunicar, da mesma forma que o “jogo da imitação” de Turing citado anteriormente, com seres humanos.

Embora ela consiga fingir ser um humano, a questão surge se ela realmente está pensando para gerar suas respostas ou apenas copiando aquilo que foi ensinada, em uma linguagem sem significado que apenas aparenta ser uma comunicação, mas não há diálogo, pois o conteúdo é ausente.

As inteligências artificiais mais modernas, considerando o estado da arte de quando esse artigo foi feito, seguem essa mesma teoria. Elas reconhecem pelo seu poder computacional como as frases e os textos são feitos, mas apenas os repete com palavras na ordem certa, sem que haja qualquer interpretação do que está sendo dito.

Para a comunicação ocorrer como entre seres humanos, é necessário um certo grau de entropia na conversa, pois não há certeza no que o outro irá responder.

No meio jurídico, essa incerteza é o que leva os processos aos tribunais. Se houvesse a resposta de antemão, não seria necessário o poder decisório do juiz e tampouco o trabalho de convencimento do advogado. O que cabe estudar, portanto, é como se dá a decisão do juiz, para ver se ela, considerando o que foi exposto até o momento, pode ser replicada por algum algoritmo de pensamento racional.

Diferença de justiça entre Themis e Dike

Ao utilizar o conceito de justiça, inicialmente é necessário explicar qual a interpretação que será dada para ela. Para tanto, Ari Solon demonstra existirem ao menos duas formas de se entender o que é a justiça, representadas na mitologia grega pelas deusas Themis e Dike.

Themis expressa a justiça aristocrática, aquela oriunda da expressão da vontade de Zeus, além de possuir uma origem relacionada aos oráculos (mais uma vez, dando voz ao divino). Por fim, é considerada como uma deusa da ética e como uma fonte de sabedoria. Pode-se entender portanto que a justiça de Themis é aquela baseada em leis, em ordenamentos superiores que os humanos devem seguir.

Já Dike carregava desde sua origem o conceito de juízo em sentido concreto, nas palavras de Solon, “a palavra *dike* também viria da linguagem processual de jogar a sorte para dirimir uma disputa”. Assim, Dike pode ser associada à justiça da decisão jurídica, a chance que os fracos (classes sociais menos favorecidas) têm de enfrentar os fortes (a aristocracia) e mesmo assim vencer, tratando-se de uma disputa não de forças, mas daquilo que é o justo contra o injusto.

A segunda interpretação de justiça, a de Dike, é a que será explorada nesse texto, pois alinha-se com o conceito de realismo jurídico (tema seguinte) assim como o objetivo para o qual a inteligência artificial foi proposta de ser usada, não para a criação de novas leis, mas para o auxílio (ou alguns dizem até a substituição) de juízes em decisões concretas.

Realismo jurídico (Savigny e Hegel)

O conceito de justiça como aplicação de decisões em situação completa (como explicado anteriormente, a vertente de Dike) é também defendida por outros autores na corrente chamada de realismo jurídico.

Hegel na *Fenomenologia do espírito* explica como a dialética ocorre a partir de um reconhecimento concreto (e não teórico) com a consciência universal. Porém esse reconhecimento da consciência universal seria inatingível, sendo necessário buscar a individual e cada pessoa chegar a sua própria conclusão acerca do tema concreto.

Ele também explica como através da história a consciência universal se modifica, o que pode ser comprovado pelas

diferentes atuações dos povos. Diferencia então três grandes períodos, terminando essa dialética com Napoleão, pois segundo o autor ao serem criados os direitos humanos universais, o conceito de divindade foi secularizado, e ao negar nessa “religião sem deus” foi alcançado o fim da história.

Posteriormente Savigny aprofunda o conceito de historicidade através de sua Escola histórica do direito. Segundo sua teoria, a codificação (como a feita por Napoleão) buscava congelar o direito e tudo que fora produzido até então, mas não abria espaço para novas criações.

E essas novas criações eram necessárias para a sociedade, pois o direito é algo dinâmico, as leis, mesmo escritas formalmente, recebem novas interpretações ao serem colocadas em casos concretos. E quem possui esse poder para novas interpretações são os juízes, através de suas decisões.

Dessa breve explanação sobre os dois autores, retomando o objeto desse texto, pode-se extrair que o direito é a aplicação das normas, e não apenas seu estudo teórico. Sendo assim, com as mudanças da sociedade ao longo do tempo, ocorre uma diferenciação de como as leis são aplicadas nos casos concretos, e essas alterações ocorrem a partir das decisões que os juízes tomam. Em outras palavras, confirma-se a tese do realismo jurídico, segundo a qual o direito surge nos tribunais.

Conclusão - máquina não faz justiça

Após essa longa apresentação de diversos autores e teses das mais distintas áreas do conhecimento, é possível conectar os pontos e chegar à conclusão almejada por esse texto.

De início, a justiça é formada pela decisão dos juízes, ela está nas cortes e não na letra fria da lei, como a interpretação de Dike demonstra. E ela muda com o tempo, conforme a teoria do realismo jurídico.

Do ponto de vista da programação racional, seria, portanto, necessária uma máquina com capacidade de se comportar como um juiz. Não basta a pura imitação, como a ELIZE era capaz, mas uma forma de pensar propriamente humana.

Neumann propôs, como demonstrado, um modelo computacional que replicasse o funcionamento cerebral. Ao invés de criar um *hardware* que realize um único algoritmo, a Máquina de Turing consegue a partir de uma única máquina fazer diversos algoritmos diferentes.

Mas como o problema da decisão (*Entscheidungsproblem*) demonstra, nem todo algoritmo consegue ser processado, em alguns casos ele pode ser infinito e, portanto, não gerar respostas.

Dito isso, existem duas maneiras que a inteligência artificial pode ser aplicada para gerar a justiça, e como será demonstrado, nenhuma das duas satisfaz.

A primeira é empregar uma máquina para resolver as questões em caso concreto. Ela consegue dar uma resposta, já que processa um argumento finito. Porém seu critério de decisão seria sempre racional, objetivando maximizar alguma variável (como Neumann demonstrou, na economia seria o lucro. Para o direito não é possível mensurar algo como “satisfação da sociedade”).

Dessa forma, buscava se basear em informações passadas, consulta a bancos de dados e fenômenos que se repetem. Assim, não haveria avanço no direito, toda a historicidade chegaria ao seu final e existiriam apenas máquinas que copiam o passado, sem se atentar aos avanços da sociedade.

A outra forma que uma inteligência artificial pode ser empregada é se ao invés de substituir um único juiz, ela seja usada como modelo jurídico e busque resolver todo o sistema, havendo uma variável caótica (explicada por Weizenbaum ao mostrar como é feito um diálogo).

Essa variável caótica poderia então alterar o sistema jurídico, mas acabaria por deslocá-lo cada vez mais da sociedade. Ele não mais caberia ao princípio de Dike de solucionar conflitos concretos, mas faria mutações no algoritmo de decisão apenas para alterá-lo, também não sendo uma solução cabível ao problema.

E essa situação talvez gere, como demonstrou Alan Turing, um algoritmo sem solução, já que precisaria de uma relação não-causal, trazendo problemas futuros para serem solucionados no presente, já que a historicidade é algo contínuo no tempo.

Assim, conclui-se que a inteligência artificial não é capaz de fazer justiça, tanto do ponto de vista material, pois suas decisões algorítmicas são baseadas em uma falsa racionalidade; como do ponto de vista técnico, pois as variáveis do direito na sociedade não podem ser calculadas como parâmetros desta hipotética máquina.

***Ari Marcelo Solon** é professor na Faculdade de Direito da USP. Autor, entre outros, livros, de *Caminhos da filosofia e da ciência do direito: conexão alemã no devir da justiça (Prisma)*. [<https://amzn.to/3Plq3jT>]

***Alan Bragança Winther** é mestrando na Faculdade de Direito da USP e engenheiro elétrico pela Escola Politécnica da USP.

Para ler o primeiro artigo desta série clique em <https://aterraeredonda.com.br/por-que-a-inteligencia-artificial-nao-faz-justica/>

Referências

CHRISTIAN, Brian. *The Most Human Human*. New York: Doubleday, 2011. p. 169-170.

HEGEL, Georg Wilhelm Friedrich. *Fenomenologia do Espírito (The Phenomenology of Spirit)*. Tradução de Arnold V. Miller; prefácio de J. N. Findlay. Oxford: Clarendon Press, 1977. xxxv, 595 p.

NEUMANN, John von; MORGENSTERN, Oskar. *Theory of Games and Economic Behavior*. 5. ed. Princeton: Princeton University Press, 1953.

ROGERS, Carl R. *Client-Centered Therapy: Its Current Practice, Implications, and Theory*. Boston: Houghton Mifflin, 1951. 560 p.

SAVIGNY, Friedrich Carl von. *Da vocação de nosso tempo para a legislação e a ciência do direito [Vom Beruf unsrer Zeit für Gesetzgebung und Rechtswissenschaft]*. 1814.

SEARLE, John R. Minds, brains, and programs. *Behavioral and Brain Sciences*, v. 3, n. 3, p. 417-457, 1980.

SOLON, Ari Marcelo. *Direito e Tradição: O Legado Grego, Romano e Bíblico*. Rio de Janeiro: Elsevier; Campus, 2009, p. 75-81.

TURING, Alan M. Computing machinery and intelligence. *Mind*, v. 59, n. 236, p. 433-460, 1950.

TURING, Alan M. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, v. 2, n. 42, p. 230-265, 1936.

WEIZENBAUM, Joseph. ELIZA — A computer program for the study of natural language communication between man and

a terra é redonda

machine. *Communications of the ACM*, v. 9, n. 1, p. 36-45, 1966.

A Terra é Redonda existe graças aos nossos leitores e apoiadores.

Ajude-nos a manter esta ideia.

CONTRIBUA

A Terra é Redonda