

galáxia

Galáxia

E-ISSN: 1982-2553

aidarprado@gmail.com

Pontifícia Universidade Católica de São

Paulo

Brasil

Gunkel, David J.

Comunicação e inteligência artificial: novos desafios e oportunidades para a pesquisa em
comunicação

Galáxia, núm. 34, enero-marzo, 2017, pp. 5-19

Pontifícia Universidade Católica de São Paulo

São Paulo, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=399650408001>

- ▶ Como citar este artigo
- ▶ Número completo
- ▶ Mais artigos
- ▶ Home da revista no Redalyc



Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Comunicação e inteligência artificial: novos desafios e oportunidades para a pesquisa em comunicação¹

David J. Gunkel

Resumo: Este artigo advoga em favor de significativas reorientações e reconceitualizações dos estudos do campo da Comunicação, de maneira a acomodar as oportunidades e desafios introduzidos pelas máquinas cada vez mais inteligentes. Particularmente, buscamos demonstrar, por um lado, como e porque a atividade comunicacional vem sendo considerada uma condição definidora da inteligência artificial (IA) e, por outro lado, como a teoria da IA e o desenvolvimento de aplicações que a envolvam complica a ideia do sujeito de comunicação, urgindo requerer modificações significantes tanto em seu aparato conceitual quanto em sua estrutura filosófica.

Palavras-chave: inteligência artificial (IA); comunicação; tecnologia; teoria; Teste de Turing.

Abstract: Communication and artificial intelligence: new and opportunities and challenges for communication research - This essay advocates for a significant reorientation and reconceptualization of communication studies in order to accommodate the opportunities and challenges introduced by increasingly intelligent machines. In particular, it demonstrates, on the one hand, how and why the activity of communication has been considered a defining condition for artificial intelligence (AI) and, on the other hand, how the theory of AI and the development of AI applications complicate the subject of communication, requiring significant modifications in its conceptual apparatus and philosophical framework.

Keywords: artificial intelligence (AI); communication; technology; theory; Turing Test.

Introdução

Reconhecidamente ou não, comunicação e inteligência artificial estão intimamente relacionadas. Por um lado, a comunicação vem sendo um instrumental tanto para a teoria

¹ O presente texto é uma versão reduzida de Gunkel, David J. (2012). *Communication and Artificial Intelligence: Opportunities and Challenges for the 21st Century*. *Communication +1*: Vol. 1. Disponível em: <http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1007&context=cpo>.

quanto para a prática da inteligência artificial (IA). Particularmente, é a comunicação que fornece à ciência da IA seus casos de testes definidores e evidências experimentais. Essa questão é imediatamente aparente num artigo que, há muito tempo, é creditado como (re)definidor da inteligência maquinica. Estamos falando do texto “Computing Machinery and Intelligence”, de Alan Turing. Da mesma maneira, o desenvolvimento recente de máquinas autônomas, de algoritmos capazes de aprendizado e de sistemas inteligentes introduz novos desafios e oportunidades para os estudos em Comunicação. Lidar com essas inovações e suas consequências exigirá uma recompilação significante da disciplina e de seu foco tradicionalmente antropocêntrico, da teoria instrumentalista da tecnologia e das ideias filosóficas modernas.

O Jogo da Imitação

Apesar da expressão “inteligência artificial” ser um produto de uma conferência acadêmica organizada por John McCarthy no Dartmouth College em 1957, é o artigo de Alan Turing, datado de 1950, e seu “jogo da imitação”, ou que agora rotineiramente chamamos de “Teste de Turing”, o elemento que define e caracteriza o campo. Apesar de Turing iniciar seu artigo propondo considerar a questão “as máquinas podem pensar?”, ele imediatamente reconhece várias dificuldades com a questão em si. Por essa razão, ele propõe que uma linha de investigação diferente seja tomada. Uma linha que possa, como ele a descreve, ser “expressa em palavras relativamente não ambíguas”.

A nova formulação do problema pode ser descrita em termos de um jogo a que nós chamamos “jogo da imitação”. É realizado por três pessoas: um homem (A), uma mulher (B) e um interrogador (C), que pode ser de qualquer um dos sexos. O interrogador permanece num quarto, separado dos outros dois. O objetivo do jogo, para o interrogador, é determinar qual é o homem e qual é a mulher. (TURING, 1996, p. 21)

Essa determinação é feita com base em questões e respostas simples. O interrogador faz para A e B várias perguntas e, baseado nas respostas, tenta discernir se o respondente é uma mulher ou um homem. “Para que tons de vozes não ajudem o interrogador”, Turing estipula que “as respostas deveriam ser escritas, ou melhor ainda, datilografadas. O arranjo ideal é um telegravador com comunicação entre os dois quartos” (TURING, 1996, p. 22). Deste modo, a configuração inicial do “jogo da imitação” é, como descreve Turing, uma comunicação mediada-por-computador (CMC) *avant la lettre* (figura 1). O interrogador interage com dois participantes desconhecidos via uma interação sincrônica mediada por computador que é rotineiramente chamada de *chat*. Já que essa troca acontece por meio de mensagens de texto, o interrogador não pode ver ou de alguma outra maneira perceber a identidade dos dois interlocutores e deve, portanto, buscar acertar o gênero baseado

em respostas que são dadas às questões como “a pessoa X poderia me dizer, por favor, o tamanho de seu cabelo” (TURING, 1996). Consequentemente, o que Turing dispõe é aquilo que hoje é uma função comum na CMC: a identidade real de um dos interlocutores é ocultada e só pode ser acertada pelo modo como as mensagens são trocadas.

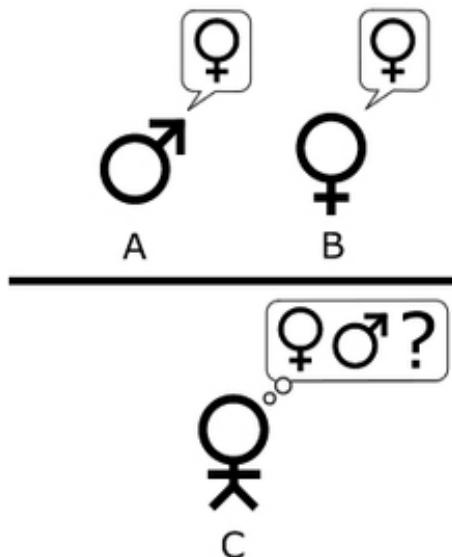


Fig.1. O jogo da imitação, primeira fase.
(Imagem retirada de http://en.wikipedia.org/wiki/Turing_test)

Turing leva, após isso, seu experimento ainda mais longe. Nós podemos fazer a pergunta? O que irá acontecer quando uma máquina toma o lugar de A neste jogo? O interrogador irá decidir erroneamente com frequência quando jogamos o jogo desse jeito ou em sua forma anterior entre um homem e uma mulher? Essas questões substituem a primeira delas, “Máquinas podem pensar?” (TURING, 1996). Em outras palavras, se o homem (A) é substituído por um computador no jogo da imitação, esse dispositivo seria capaz de responder as questões e “se passar” por outra pessoa, efetivamente enganando o interrogador e o levando a pensar que era apenas mais um interlocutor humano? (figura 2). É esta questão, de acordo com Turing, que *substitui* a pergunta infeliz e ambígua “Máquinas podem pensar?”. Consequentemente, se um computador de fato pode se tornar capaz de simular um ser humano de ambos os gêneros, em intercâmbios comunicativos com um interrogador humano de modo que o interrogador não consiga dizer se está interagindo com uma máquina ou outro ser humano, Turing conclui que essas máquinas devem ser consideradas “inteligentes”.

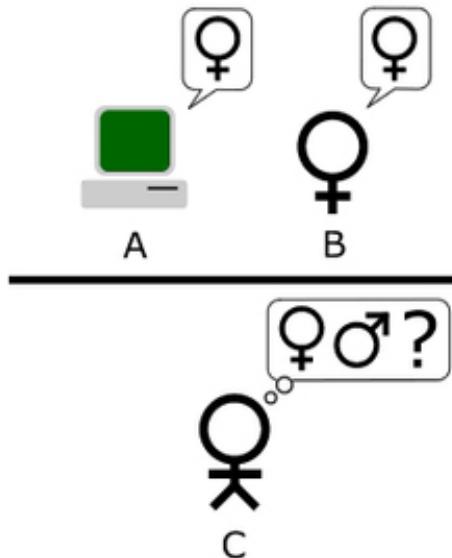


Fig.2. O jogo da imitação, fase dois.
(Imagen retirada de http://en.wikipedia.org/wiki/Turing_test)

Na época da publicação do artigo de Turing, ele estimou que o ponto crítico – aquele no qual uma máquina poderia conseguir jogar com sucesso o jogo da imitação – seria atingido em pelo menos meio século no futuro:

Acredito que, em cerca de 50 anos, será possível programar computadores, com uma capacidade de memória de cerca de 10^9 para fazê-los jogar o jogo da imitação tão bem que um interrogador médio não terá mais de 70% de probabilidade de chegar à identificação correta, após cinco minutos de interrogatório. (TURING, 1996, p. 44)

Não demorou tudo isso. Já, em 1966, Joseph Weizenbaum demonstrou uma aplicação simples de processamento de linguagem natural que era capaz de conversar com interrogadores humanos de maneira a parecer outra pessoa. ELIZA, como era chamada a aplicação, foi o primeiro *chatterbot*. Apesar desse termo ainda não ser utilizado por Weizenbaum, foi aplicado retroativamente como resultado dos esforços de Michael Maudlin, fundador e pesquisador-chefe da Lycos, que introduziu esse neologismo em 1994 para identificar uma aplicação de processamento de linguagens naturais que por ele foi chamada de Julia. ELIZA era, em termos técnicos, um programa bem simples que:

Consistia principalmente de métodos gerais de análise de frases e fragmentos delas, localizando o que chamamos de palavras-chave nos textos, montando

sentenças a partir dos fragmentos, e assim por diante. Ela tinha, em outras palavras, nenhum ferramental contextual embutido do universo do discurso. Isso era provido a ele por meio de um *script*. De certo modo ELIZA era uma atriz que comandava um conjunto de técnicas, mas não tinha nada pra dizer vindo de si mesma. (WEIZENBAUM, 1976, p. 188; tradução nossa)

Apesar dessa objeção, o programa de Weizenbaum demonstrou aquilo que Turing havia inicialmente previsto:

ELIZA criou a ilusão mais marcante de ser capaz de ter entender o que se passava na mente das muitas pessoas que conversaram com ela. Pessoas que sabiam muito que estavam conversando com uma máquina logo esqueceram esse fato - assim como a plateia de um teatro, sob o domínio da suspensão da descrença - e logo esqueceram que aquilo que estavam presenciando não era “real”. Essa ilusão era especialmente forte e se apegava com mais tenacidade em pessoas que conheciam pouco ou nada sobre computadores. Eles frequentemente pediam que as deixassem conversar com o sistema em privado e, após terem conversado com ela por um tempo, insistiam, mesmo depois de minhas explicações, que a máquina havia realmente os entendido. (WEIZENBAUM, 1976, p. 189; tradução nossa)

Características e Consequências

Por mais que exista um bom acordo com o que tem sido e continua a ser escrito em resposta ao artigo de Turing, o jogo da imitação e as demonstrações empíricas que foram proporcionadas pela ELIZA e suas sucessoras, essa formulação particular da IA tem quatro importantes características e consequências que envolvem e afetam a comunicação.

O problema das outras mentes

O texto de Turing situa a comunicação – e um tipo específico de comunicação interpessoal conversacional – como um fator decisivo para a IA. Já que a “questão original” “podem as máquinas pensar?” é considerada por Turing “com pouco sentido”, ele a reformulou e se referiu à pesquisa para a possibilidade de demonstração de habilidade comunicacional. Essa decisão não ocorreu por capricho, há bons motivos filosóficos para que prossigamos desta maneira, já que essa questão tem a ver com aquilo que os filósofos rotineiramente chamam de “o problema das outras mentes”. “Como alguém determina”, como Paul Churchland (1999, p. 67) celebremente disse, “se alguma coisa além de si mesmo – uma criatura alienígena, um robô sofisticado, um computador socialmente ativo, ou até mesmo outro humano – é realmente um ser pensante, capaz de sentir e de ter consciência; ao invés de, por exemplo, um autômato inconsciente cujo comportamento emerge de algo exterior aos genuínos estados mentais?”. Essa dificuldade, como explicam

Gordana Dodic-Crnkovic e Daniel Person (2008) está enraizada no fato inegável de que “nós não temos acesso ao funcionamento interno das mentes humanas – muito menos do que o acesso que temos ao funcionamento interno de um computador”. De fato, não podemos, como ressalta Donna Haraway (2008, p. 226), entrar nas cabeças das pessoas para conseguir “ver a história acontecer por dentro”. Consequentemente, as tentativas para resolver ou pelo menos responder esse problema inevitavelmente envolvem algum tipo de demonstração comportamental ou teste, como o jogo da imitação de Turing. “Para dizer de outra maneira”, Roger Schank (1990, p. 5) conclui que “nós não podemos examinar os interiores de uma entidade inteligente de modo a ter acesso àquilo que ela realmente sabe. Nossa única opção é questionar e observar”. Para Turing, assim como muitos de seus pares do campo da IA que seguem seu pensamento, a inteligência não é algo fácil de ser definida e tampouco observada. É, entretanto, evidenciada e decidida com base em comportamentos que possam ser considerados signos ou sintomas de inteligência, em especial a comunicação em geral e a conversação verbal particularmente em nível humano. Em outras palavras, já que o pensamento inteligente não é diretamente observável, o máximo que podemos fazer é lidar com algo, como a interação comunicativa, que é assumidamente um produto da inteligência e pode ser empiricamente observada, medida e avaliada.

A proposta de Turing, como consequência, está apoiada e alavancada em uma suposição comum da filosofia da mente, de que a comunicação é um indicativo de inteligência ou pelo menos de algum tipo de atividade cognitiva. Essa proposição tem profundas raízes filosóficas, voltando atrás até pelo menos a obra de René Descartes, na qual o discurso falado era identificado com uma característica unicamente humana e como único método correto pelo qual poderíamos diferenciar o sujeito humano racional dos animais e autômatos ostensivamente desprovidos de mentes. Descartes argumenta que existe pelo menos uma maneira correta de reconhecer que essas figuras artificiais são de fato máquinas e não humanos reais:

Nunca poderiam servir-se de palavras nem de outros sinais, combinando-os como fazemos para declarar aos outros nossos pensamentos. Pois se pode conceber que uma máquina seja feita de tal modo que profira palavras, e até profira algumas a propósito das ações corporais que causem alguma mudança em seus órgãos, como por exemplo, ela perguntar o que lhe queremos dizer se lhe tocarmos em algum lugar, se em outro, gritar que a machucamos, e outras coisas semelhantes, mas não é possível conceber que as combine de outro modo para responder ao sentido de tudo quanto dissemos em sua presença, como os homens mais embrutecidos podem fazer. (DESCARTES, 1996, p. 63)

O jogo da imitação de Turing segue e se alimenta dessa rica tradição filosófica, mas de maneira reversa. Se uma entidade – um outro ser humano, um animal, robô, algoritmo, etc. – é de fato capaz de performar operações comunicativas pelo menos de acordo com

aquilo que é tipicamente esperado de outro indivíduo humano, independente daquilo que de fato acontece dentro da cabeça ou do processador de informação da própria entidade, nós podemos considerar que se trata de uma entidade inteligente. A partir disso, Turing estimou que o desenvolvimento da comunicação maquinica iria avançar a tal ponto que não fará mais sentido falar (e falar de maneira inteligente) de inteligência maquinica já no fim do século XX. “Eu prevejo”, escreveu Turing, “que no fim do século o uso das palavras e da opinião formada terá sido alterado de tal maneira que as pessoas poderão falar sobre inteligência maquinica sem a expectativa de serem contrariadas” (TURING, 1996, p. 44).

Signos de Inteligência versus a “Inteligência Real”

Apesar de essa conclusão derivar, logicamente, do argumento de Turing, houve e continua a existir uma resistência considerável a ela. Para Turing, o desafio crítico desse contexto já havia sido articulado por Lady Lovelace (também conhecida como Ada Augusta Byron, a filha do poeta inglês Lord Byron), que não somente escreveu o software para o Engenho Analítico de Charles Babbage como, por essa razão, é considerada a primeira cientista da computação. A informação mais detalhada que temos do Engenho Analítico de Babbage, explica Turing, vem de uma memória de Lady Lovelace. Nela, ela afirma, “O Engenho Analítico não tem pretensão de *originar* nada. Ele só pode fazer qualquer coisa que *tenhamos a capacidade de ordená-lo a fazer*” (TURING, 1996, p. 50; grifo no original). De acordo com Lovelace, um computador (e quando ela escreveu isso, “computador” se referia não a um dispositivo eletrônico, mas a um grande processador mecânico de informação constituído de engrenagens e alavancas interligadas), não importa o quanto sofisticada seja a sua programação, só pode fazer aquilo que ordenemos que seja feito. Nós podemos, de fato, escrever o código de um *software* como a ELIZA ou a Siri da Apple, que reconhece um *input* verbal, extraí palavras-chave, rearranja essas palavras de acordo com *scripts* pré-programados e depois envia de volta resultados que parecem ser de origem inteligível. Isso, entretanto, não significa necessariamente que tal máquina seja capaz de pensamentos originais ou de entendimento do que está em jogo nem mesmo em uma escala rudimentar.

Esse posicionamento é levado em conta e desenvolvido por John Searle em seu exemplo bem conhecido do “Quarto Chinês”. Esse experimento, de pensamento intrigante e influente, introduzido nos anos 1980 com o artigo “Minds, Brains and Programs” e aprofundado em suas publicações subsequentes, foi utilizado como argumento contra a afirmação de existência de uma IA forte - de que as máquinas são capazes de atingir pensamentos inteligentes:

Imagine que você execute as etapas de um programa elaborado para responder perguntas em um idioma que você não comprehende. Eu não entendo chinês, então imagino que estou trancado em uma sala cheia de caixas com símbolos chineses (a base de dados), recebo uma pequena quantidade de símbolos chineses

(perguntas em chinês), e, então procuro em um manual (o programa) o que deveria fazer. Realizo algumas operações com símbolos de acordo com regras (i.e., eu executo as etapas do programa) e entrego uma pequena quantidade de símbolos (respostas às perguntas) aos que se encontram fora do quarto. Eu sou um computador executando um programa para responder perguntas em chinês, mas ao mesmo tempo não comprehendo uma palavra de chinês. (SEARLE, 1998, p. 38)

O ponto da ilustração de Searle, embora um pouco etnocêntrica, é bem simples – a simulação não é a coisa real. A mera combinação de símbolos verbais de um modo que pareça com um entendimento linguístico não é de fato um entendimento da linguagem. Um computador, como explica Terry Winograd (1990, p. 187), não entende de fato os *tokens* linguísticos que processa; ele somente “manipula símbolos sem dizer respeito às suas interpretações”. Ou, como conclui Searle, registrando o efeito desse *insight* no teste padrão de inteligência artificial: “Isso mostra que o Teste de Turing falha em distinguir capacidades mentais reais das simulações dessas capacidades. Simulação não é uma duplicação” (SEARLE, 1999, p. 115).



Fig.3. O “Quarto Chinês”, experimento de John Searle
(Imagem retirada de <http://cognitivephilosophy.net>)

Demonstrações como o Quarto Chinês de Searle, que buscam diferenciar a aparência de algo do algo real nele mesmo, não só retomam uma distinção filosófica antiga pelo menos tão velha quanto Platão, mas que inevitavelmente requer algum tipo de acesso privilegiado e imediato ao real como é e não como ele se apresenta. Para distinguir, por exemplo, entre uma simulação de inteligência e uma “inteligência real”, alguém

precisaria não somente ter acesso aos indicadores externos de algo que se parece com uma inteligência, mas também à atividade real de inteligência do modo como ela ocorre (ou não) na mente do outro. Esse requerimento, entretanto, imediatamente se choca com outros problemas da mente e com a limitação epistemológica que impõe, nominalmente, que nós não possamos entrar na “cabeça” de outra entidade – seja essa entidade um ser humano, um animal não humano, uma forma de vida alienígena, ou uma máquina – de modo que saibamos com alguma certeza se eles de fato performam aquilo que parecem manifestar. Em outras palavras, Searle só consegue distinguir e comparar aquilo que aparece para esses indivíduos interagindo com o quarto e o que há dentro do quarto porque ele tem acesso privilegiado àquilo que acontece em seu interior. Esse “contraexemplo”, portanto, viola as limitações epistemológicas impostas pelo problema das outras mentes, algo que no jogo da imitação Turing levou em consideração com cuidado e respeito.

IA e Interação Social

Ainda que alguém, seguindo o argumento de Searle, estivesse convencido de que *smartphones*, algoritmos de recomendação e aplicativos de computador devam permanecer como instrumentos “sem mente”, que meramente manipulam os *tokens* linguísticos, o Teste de Turing também demonstra que o modo como respondemos e o que fazemos com essas manipulações também faz diferença. Em outras palavras, não importa se concluímos que a máquina é inteligente ou não, o comportamento comunicacional que ela exibe, por exemplo, no jogo da imitação, gera um efeito em nós e em nossas interações sociais e relacionamentos. Mesmo que alguém duvide da possibilidade de um dia o que é chamado de “IA forte” ser atingida, o fato é que nosso mundo já é povoado por artefatos semi-inteligentes, robôs sociais, algoritmos capazes de aprender e sistemas de tomada de decisão autônomos que exponencialmente ocupam o lugar do Outro nas relações sociais e na interação comunicativa.

Esse *insight* foi experimentalmente confirmado por aquilo que Clifford Nass e Byron Reeves denominam a teoria do Computador como Ator Social (CSA):

Os computadores, do modo como eles se comunicam, instruem e revezam turnos de interação, estão *tão próximos* dos humanos que eles encorajam respostas *sociais*. O encorajamento necessário para uma produzir uma reação não é tão grande. Desde que existam alguns comportamentos que indiquem uma presença social, as pessoas irão responder a eles de acordo... Consequentemente, qualquer meio que está suficientemente próximo terá um tratamento humano, ainda que as pessoas achem isso bobo e neguem o feito depois. (REEVES; NASS, 1996, p. 22; tradução nossa)

O modelo de CSA, desenvolvido como resposta a inúmeros experimentos com sujeitos humanos, descreve como usuários de computadores – independente da inteligência real

detida (ou não) pela máquina – tendem a responder à tecnologia como um outro sujeito socialmente consciente e interativo. Em outras palavras, mesmo se usuários experientes saibam bem que foram engajados pelo uso de uma máquina, eles fazem o que Reeves e Nass (1996, p. 22) chamam de “erro conservador” e tendem a atuar de maneira a parecer um outro ser humano. Consequentemente, para algo ser reconhecido e tratado como um ator social, “não é necessário”, como Reeves e Nass (1996, p. 28) concluem, “ter inteligência artificial” de forma literal.

Esse resultado mostra-se evidente não só nos restritos estudos experimentais realizados por Reeves e Nass, mas também nas interações mundanas com objetos “sem mente” – como os programas de *chatterbots* e os personagens não jogáveis, que são encontrados em comunidades na *Internet* e em redes sociais e massivamente aplicados em jogos *online* como RPG. Esses softwares automatizados (aparentemente, descendentes de ELIZA) complicaram o cenário, tornando difícil de decidir quem ou o quê é responsável pelas atividades no espaço virtual ou em uma comunidade virtual.

O desenvolvimento das comunidades *online* levou a um fenômeno do tempo real, das interações multipersonagens por meio de personagens *online*. Algumas tecnologias de comunidades *online* permitem a criação de *bots* (personagens que atuam de acordo com um programa em vez de serem controladas por um usuário humano) de tal forma que não é sempre fácil um ser humano interagir com um *bot* num espaço social *online*. É também possível uma pessoa ser parcialmente controlada por um *software* e parcialmente por um humano... Isso leva de problemas teóricos e práticos a argumentos éticos (sem contar normatizações), já que a interação entre atores e os agentes morais pode ser perdida. (MOWBRAY, 2002, p. 2; tradução nossa)

Socialbots, que agora populam e operam nos espaços virtuais – não só de jogos, mas também de redes sociais como Twitter e Facebook –, complicaram as interações entre os atores e os agentes morais. “Aqui está”, como Steve Jones (2014, p. 245) aponta, “uma concomitante e crescente quantidade de intervenções de algoritmos que utilizam comunicações entre usuários e entre usuários e máquinas para criar, modificar ou canalizar comunicações e interações”. E, com essa “intervenção de algoritmos”, está se tornando cada vez mais difícil identificar quem ou o quê é responsável pelas ações no espaço virtual de uma comunidade. Por mais que esses *bots* não consigam atingir nem perto o que remotamente pareça uma inteligência, eles podem ser confundidos ou até se passar por um outro usuário humano (JONES, 2015; EDWARDS et al., 2013, e GEHL, 2013). Isso é o que Mowbray (2002, p. 2) aponta como não sendo “um elemento da sofisticação de um *bot*, mas o menor nível de uma banda de comunicação do espaço social *online*” em que é “fácil simular, convincentemente, um agente humano”. Isso ocorreu, recentemente, no caso dos *fembots*, computadores projetados para atuar como

mulheres, de Ashley Madison. Eles envolvem *scripts* de computadores projetados para iniciar trocas amorosas com usuários masculinos com objetivo de transformá-los em clientes que pagam. Por mais que a programação desses *fembots* seja muito simples e, de alguma forma, até banal, um número grande de usuários os achou envolventes, tanto que dividiram segredos íntimos com o *bot* e, mais importante ainda, usaram o cartão de crédito para continuar a conversa. Esses programas rudimentares não estão nem perto de atingir um nível humano de inteligência, entretanto, podem ser confundidos e se passar por usuários humanos. Eles são, nas palavras de Reeves e Nass, “próximos demais de um humano a ponto de encorajar respostas sociais”.

Inteligência Artificial e teoria da comunicação

Por fim, e por conta do exposto anteriormente, as regras do jogo na teoria da comunicação precisam ser ajustadas e modificadas de forma significativa. No Teste de Turing, o computador ocupava a posição tanto do meio *pelo qual* interlocutores humanos trocam mensagens e também como um dos participantes *com quem* (ou com o que)² o outro está engajado nesses intercâmbios interativos. Apesar disso, os estudos de comunicação pouco se interessaram e pouco abordaram essa questão. Eles entenderam e analisaram o computador – e todos os aparatos tecnológicos – como um meio de interação comunicativa. Essa decisão fundamental sobre o papel e a função do computador é consistente em relação à teoria instrumental da tecnologia³, que entende a tecnologia como nada mais do que uma ferramenta de atividades humanas, e com o modelo básico da teoria matemática da comunicação, inicialmente desenvolvida por Claude Shannon e Warren Weaver (1963). Por mais que essas abordagens tradicionais tenham um marcante sucesso, eles perderam uma oportunidade crucial identificada originalmente por Turing

² Não é claro se deve ser utilizada, neste contexto, a expressão “com o qual” ou “com quem”. Por mais que esse equívoco pareça ser uma pequena questão gramatical, tudo depende de uma decisão. Ao fazer uma escolha entre um e outro, é decidido se a máquina será considerada uma coisa, um mero objeto, ou se será considerada outro sujeito. Uma preocupação similar foi identificada e perseguida, por um ângulo diferente, no trabalho de Jacques Derrida sobre dom, perdão e hospitalidade. Como Derrida (2005, p. 8) explica em um dos textos presentes na obra “Paper Machine”, “eu já, aparentemente, confiei na disjunção entre *quem* e *o que*, para chacoalhar isso um pouco, serei então claro em meu presente trabalho, acima de todos os meus ensinamentos, eu tentarei atingir um ponto do *qual* essa disjunção entre *quem* e *o que* apareça e seja determinada, em outras palavras, de forma ‘anterior’ a essa disjunção, em um lugar mais ‘antigo’ ou mais ‘atual’ que isso, em um lugar em que ambos apontam determinação e também possibilitam a irreversível translação de *quem* em *o que*.

³ Essa conceituação é estruturada e baseada na resposta que é tipicamente apresentada para a questão relacionada à tecnologia. “Questionamos a tecnologia quando perguntamos o que é. Todos sabem quais são os dois posicionamentos que respondem a essa questão. Um diz que tecnologia é um meio para um fim. Outro diz que a tecnologia é uma atividade humana. As duas definições andam juntas. O que postula e dá significado a elas é a atividade humana. A manufatura e a utilização de equipamentos, ferramentas e máquinas, manufaturadas e usadas por eles mesmos, e a necessidade e a finalidade a que servem, tudo pertence ao que a tecnologia é” (HEIDEGGER, 1977, p. 4-5). Segundo a análise de Heidegger, o papel presumido e a função de qualquer tipo de tecnologia – independente de um produto ser feito à mão ou industrializado – é de que ela é explorada por usuários humanos com finalidades específicas. Heidegger denomina essa caracterização particular da tecnologia como “a definição instrumental” e indica quais formas são consideradas “corretas” para a compreensão de qualquer tipo de aparato tecnológico.

e confirmada experimentalmente por Reeves e Nass: o fato de que uma máquina não é só um meio de atividade humana, mas também pode ser um participante em interações comunicativas.

Os estudos de comunicação pareceram marginalizar ou até ignorar esse aspecto, entretanto, a disciplina de fato começou tratar do tema ao tentar abordar e conceitualizar o amplo leque de possibilidades. Esses esforços foram apresentados inicialmente por Robert Cathcart e Gary Gumpert em um artigo de 1985 intitulado “The Person-Computer Interaction”. Nesse texto relativamente precoce (“precoce” do ponto de vista do reconhecimento e engajamento da disciplina em relação à informática), os autores desenharam a distinção entre a comunicação *por meio* do computador da comunicação *com* o computador. No texto são mencionados termos como “funções propiciadas pelo computador” e “o computador se interpõe entre o emissor e o receptor”. Também são abordadas “funções interrelacionais indivíduo-computador” em que “ativa-se um computador, que responde apropriadamente de forma gráfica, alfanumérica ou sonora, estabelecendo uma relação emissor/receptor” (CATHCARD; GUMPERT, 1985, p. 114). Essas duas alternativas, que aderem à configuração do Teste de Turing, sem explicitamente mencioná-lo, foram corroboradas e posteriormente refinadas, em 1989, no “Computer-Mediated Communication”, de James Chesebro e Donald Bonsall. Na análise presente nesse livro, os autores detalham uma escala de cinco pontos que delimitam o leque de possibilidades para uma “comunicação computador-humano”. A escala vai do computador utilizado como mero meio de transmissão de mensagem à interlocução humana com o computador entendido como um agente inteligente com quem usuários humanos interagem.

Na relação com computador, sob o ponto de vista dessa outra (quase esquecida) alternativa teórica, todos os tipos de coisas podem mudar, não só o que nossa compreensão ou quem ou o quê qualifica ou legitima o sujeito da comunicação. Segundo Norbert Wiener, o progenitor da cibernetica, esses desenvolvimentos fundamentalmente alteram o cenário social:

Na tese desse livro [*The Human Use of Human Beings*], aquela sociedade pode ser entendida somente por meio de um estudo das mensagens e das ferramentas de comunicação a que pertence; e, no futuro, com o desenvolvimento das mensagens e das ferramentas de comunicação, as mensagens entre homens e máquinas, entre máquinas e homens, e entre máquinas e máquinas, estão fadadas a ter um papel cada vez mais importante. (WIENER, 1998, p. 16; tradução nossa).

Nas relações sociais em um futuro não tão distante (e alguém precisa relembrar o que Wiener escreveu em 1950, no mesmo ano do influente artigo de Turing), o computador e os sistemas relacionados, como os robôs detentores de corpo e os algoritmos desprovidos de corpo, não serão mais meros instrumentos de ações comunicativas de humanos ou meio pelo qual humanos se comunicam entre si. Ocuparão, em vez disso, a posição de outro ator social com quem alguém comunica ou interage. Para que ocupem essa posição,

nós inevitavelmente precisaremos buscar e identificar questões fundamentais relacionadas à responsabilidade social e à ética – questões que não só poderiam ter sido articuladas no contexto do anterior paradigma teórico dos estudos de comunicação, mas se fossem articuladas seriam, nesta perspectiva, consideradas inapropriadas e *nonsense*. Qual é, por exemplo, nossa responsabilidade em face desse Outro – um outro que é diferente de um ser humano? Como devemos fazer ou responder a essa forma de Outro? Como será ou deveria ser a forma desse Outro responder?⁴ Por mais que essas questões pareçam abertas em o que muitos considerariam ser uma ficção científica, elas já fazem parte da realidade social e o desafio para o século XXI é decidir como a teoria da comunicação irá responder e acomodar esses novos desafios e oportunidades sociais.

David J. Gunkel é professor da Northern Illinois University (EUA)

dgunkel@niu.edu

Tradução:

Francisco B. Trento, doutorando do Programa de Comunicação e Semiótica da PUC-SP.

francisco.trento@gmail.com

Daniela Norcia Gonçalves, doutoranda do Programa de Comunicação e Semiótica da PUC-SP

daniela.norcia@gmail.com

Referências

- CATHCART, R.; GUMPERT, G. The Person-Computer Interaction: A Unique Source. **Information and Behavior**, vol. 1, p. 113-124. New Brunswick: Transaction Books, 1985.
- CHESEBRO, J. W.; BONSALL, D. G. **Computer-Mediated Communication: Human Relationships in a Computerized World**. Tuscaloosa: The University of Alabama Press, 1989.
- CHURCHLAND, P. (1999). **Matter and Consciousness**. Cambridge: MIT Press, 1999.
- DERRIDA, J. **Paper Machine**. Trans. by Rachel Bowlby. Stanford: Stanford University Press, 2005.
- DESCARTES, R. **Selected Philosophical Writings**. Tradução J. Cottingham, R. Stoothoff, e D. Murdoch. Cambridge: Cambridge University Press, 1988.
- _____. **Discurso do Método**. Tradução Maria Ermantina Galvão. São Paulo: Martins Fontes, 1996.

⁴ Para considerações detalhadas sobre essas e outras reflexões relacionadas à questão social das máquinas inteligentes, veja em Gunkel (2012).

DODIG-CRNUKOVIC, G.; PERSSON, D. **Towards Trustworthy Intelligent Robots—A Pragmatic Approach to Moral Responsibility.** North American Computing and Philosophy Conference, jul. 10-12, 2008. Bloomington: Indiana University. Disponível em <http://www.mrct.mdh.se/~gdc/work/NACAP-Roboethics-Rev1.pdf>. Acesso em 20/7/2016.

EDWARDS, C.; EDWARDS, A.; SPENCE, P. R.; SHELTON, A. K. Is that a Bot Running the Social Media Feed? Testing the Differences in Perceptions of Communication Quality for a Human Agent and a Bot Agent on Twitter. **Computers in Human Behavior**, 2013.

GEHL, R. W. The Computerized Socialbot Turing Test: New Technologies of Noopower. **Social Science Research Network** (SSRN), 2013. Disponível em <http://ssrn.com/abstract=2280240>. Acesso em 20/7/2016.

GUNKEL, D. J. **The Machine Question: Critical Perspectives on AI, Robots, and Ethics.** Cambridge, MIT Press, 2012.

HARAWAY, D. **When Species Meet.** Minneapolis: University of Minnesota Press, 2008.

HEIDEGGER, M. The Question Concerning Technology. **The Question Concerning Technology and Other Essays.** Tradução William Lovitt. New York: Harper & Row, 1977.

JONES, S. People, Things, Memory and Human-Machine Communication. **International Journal of Media & Cultural Politics**, p. 245-258, 2014.

_____. How I Learned to Stop Worrying and Love the Bots. **Social Media and Society** 1(1): 1-2, 2015.

MOWBRAY, M. **Ethics for Bots.** Paper presented at the 14th International Conference on System Research, Informatics, and Cybernetics. Baden-Baden: jul.-ago. 2002. Disponível em <http://www.hpl.hp.com/techreports/2002/HPL-2002-48R1.pdf>. Acesso em 20/7/2016.

PARTRIDGE, D.; WILKS, Y. **The Foundations of Artificial Intelligence.** Cambridge, Cambridge University Press, 1990.

REEVES, B.; NASS, C. **The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.** Cambridge: Cambridge University Press, 1996.

SEARLE, J. **Minds, Brains and Science.** Cambridge: Harvard University Press, 1984.

_____. **O mistério da consciência.** São Paulo: Paz e Terra, 1998.

_____. The Chinese Room. **The MIT Encyclopedia of the Cognitive Sciences**, p. 115-116. Cambridge: MIT Press, 1999.

SCHANK, R. C. What is AI Anyway? **The Foundations of Artificial Intelligence: A Sourcebook.** Derek Partridge e Yorick Wilks (Orgs.). Cambridge: Cambridge University Press, 1990.

SHANNON, C. E.; WEAVER. **The Mathematical Theory of Communication.** Urbana: University of Illinois Press, 1963.

TURING, Alan. Computação e Inteligência. Tradução de Fábio de Carvalho Hansem. **Cérebros, Máquinas e Consciência: uma introdução à filosofia da mente.** São Carlos: EdUFSCar, 1996.

TEIXEIRA, J. F. (Org.). **Cérebros, Máquinas e Consciência: uma introdução à Filosofia da Mente.** São Carlos: EdUFSCar, 1996.

TURING, A. Computing Machinery and Intelligence. **Computer Media and Communication: A Reader**, p. 37-58. Oxford: Oxford University Press, 1999.

WEIZENBAUM, J. **Computer Power and Human Reason: From Judgment to Calculation.** São Francisco: W. H. Freeman, 1976.

- WIENER, N. **The Human Use of Human Beings: Cybernetics and Society.** Boston: Da Capo Press, 1988.
- WINOGRAD, T. Thinking Machines: Can there be? Are we? **The Foundations of Artificial Intelligence: A Sourcebook.** Derek Partridge and Yorick Wilks (Org.). Cambridge: Cambridge University Press, 1990.

*Artigo recebido em agosto
e aprovado em novembro de 2016.*