

Amizade Robô-Humano:

o Que Pode Estar Faltando em um Modelo de Robô “Amigo”?

Claus Emmeche

Mariana Vitti Rodrigues

Como citar: EMECHE, C.; RODRIGUES, M. V. Amizade Robô-Humano: o que pode estar faltando em um modelo de robô “amigo”?. *In:* ALVES, M. A.; GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D. (org.). **Informação, conhecimento e modelos**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2017. p. 39-56.
DOI: <https://doi.org/10.36311/2017.978-85-86497-29-2.p39-56>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported.

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição - Uso Não Comercial - Partilha nos Mesmos Termos 3.0 Não adaptada.

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 3.0 Unported.

AMIZADE ROBÔ-HUMANO: O QUE PODE ESTAR FALTANDO EM UM MODELO DE ROBÔ “AMIGO”?

Claus Emmeche
cemmeche@ind.ku.dk
Mariana Vitti Rodrigues
mariana.rodrigues@ind.ku.dk

1. INTRODUÇÃO

A questão discutida no presente capítulo repousa sobre um aspecto da cognição e da emoção humana frequentemente negligenciado ou ignorado – a saber, a amizade – e sua possível instanciação em sistemas artificiais como robôs¹. Robôs podem ser amigos? Você poderia ser amigo de um robô? Provavelmente, para a maioria das pessoas, parece um pouco estranho sugerir ou imaginar que humanos e robôs possam estabelecer uma relação interativa e dinâmica similar às relações de amizade. A intuição de que alguma coisa poderia estar faltando para um ser humano e um sistema artificial serem realmente “amigos” pode estar correta. Contudo, como sempre, intuições, no contexto da pesquisa, precisam ser bem articuladas, de sorte que investigá-las – ou propor contraexemplos, como a impossibilidade de robôs serem amigos – requer um profundo entendimento não apenas da especificidade da cognição humana, mas também da relação entre cognição, afeto e natureza da amizade. Nesse sentido, considerando a possibilidade de robôs autônomos, nós nos questionamos: o conceito de autocontrole é suficiente para a caracterização de relações de amizade? Se sim, poderiam robôs autônomos (que possuem autocontrole) estabelecer uma relação robô-humano de amizade?

¹ Este capítulo é uma versão traduzida e modificada de um artigo publicado anteriormente (EMMECHE, 2014). No presente artigo, desenvolvemos a argumentação encontrada em: EMMECHE, C. Robotic friendship: Can a robot be a friend?. *International Journal of Signs and Semiotic Systems*, v. 3, n. 2, p. 26-42, 2014.

Neste capítulo, não argumentamos de imediato em favor da impossibilidade do estabelecimento de relações de amizade entre humanos e robôs. Sugerimos deixar a questão em aberto, com o intuito de investigar mais detalhadamente as consequências que uma possível amizade entre robôs e humanos podem indicar acerca das semelhanças e diferenças entre a cognição humana (incluindo cognição distribuída) e a cognição artificial instanciada em robôs. Em geral, amizade e cognição social têm sido consideradas fatores importantes para o desenvolvimento de um *self*² autônomo em humanos. Dessa forma, entendemos que o estudo da amizade pode ensinar algo sobre a autonomia e a “agência”³ humana. Nas áreas da Ciência Cognitiva, Robótica, Inteligência Artificial e Vida Artificial, os pesquisadores frequentemente se ocupam da construção de sistemas inteligentes e adaptativos capazes de solucionar problemas práticos como sobrevivência. Nesse cenário, sistemas artificiais transformam a informação sensorial em conhecimento organizado, com o objetivo de direcionar o movimento e possibilitar o aprendizado de habilidades instrumentais necessárias para auxiliar a ação e a cognição humanas. Esse tipo de racionalidade e inteligência instrumental é importante para direcionar ações individuais, em muitos domínios técnicos; entretanto, pode não ser suficiente para todas as ações sociais em contextos onde valores e normas cumprem um papel importante.

Ao colocar questões sobre formas de amizade humana e uma possível existência (ou não-existência) de amizade em robôs, questionamos o que significa dizer que a ação racional-intencional humana é direcionada por *normas* no contexto social. Uma vez que a natureza dessas normas é social, elas são tomadas como expressões de autocontrole que regulam não apenas a ação de indivíduos isolados, mas de indivíduos sociais, isto é, indivíduos socializáveis, através de seu desenvolvimento dentro de uma comunidade. Assim como organismos sociais de outras espécies, ou “animais políticos” (tal como Aristóteles caracteriza os seres humanos), somos seres sociais, quando vivemos em sociedades regidas por instituições normativas. Os seres humanos são sociáveis, quando reconhecem (através da autoavaliação) um real comprometimento em seguir as normas sociais de

² Podemos traduzir *self* (ou *selves*, no plural) por “eu(s)” ou “ego(s)”. Entretanto, optamos por deixar o termo no original.

³ Do inglês “agency”.

acordo com uma conduta adequada, desconsiderando eventuais sucessos ou fracassos de suas ações.

Se compararmos a noção de amizade com a de parentesco ou de aliados (no aspecto prático de autodefesa), a amizade genuína parece ser algo com um fim em si mesmo. Obviamente, embora alguns animais como os babuíños verde-oliva possam ter “amigos” (exigindo uma forma de cognição social como o reconhecimento de tipos de relações entre outros membros do grupo – cf. SEYFARTH; CHENEY 2012; DAGG, 2011), essa relação toma uma forma distinta e complexa, quando envolve o ser humano na chamada amizade moderna, com suas normas baseadas na linguagem mútua de autodescoberta (BELL; COLEMAN, 1999; GIDDENS, 1992; JAMIESON, 1998; SINGER, 2009). Para além dos aspectos práticos, a amizade emerge como uma relação nômica e contingente às condições de certos limites emocionais, cognitivos e sociais, os quais possibilitam essa relação como um todo. Três importantes condições materiais da amizade, por exemplo, são o excedente de tempo, a curiosidade e a capacidade de permitir um espaço livre para atividades guiadas não somente por preocupações de autodefesa ou de sobrevivência.

Em síntese, é objetivo deste trabalho discutir em que medida podemos conceber uma relação de amizade entre humanos e robôs e, a partir dessa perspectiva, investigar possíveis diferenças entre a cognição humana e a artificial. Para tanto, na segunda seção, exploramos dois diferentes tipos de leis da robótica, quais sejam, as leis de Tilden e as leis de Asimov, questionando-nos se a amizade pode ser caracterizada dentro da proposta dessas leis. Na terceira seção, propomos uma caracterização semiótica de amizade, concebida como uma relação triádica, não-reduzível aos seus elementos constituintes. Na quarta seção, discutimos os conceitos de autocontrole e de inteligência, questionando se sistemas artificiais como robôs possuem um certo tipo de inteligência emocional necessária para a constituição de relações de amizade. Finalmente, realizamos um balanço dos conceitos trabalhados no presente capítulo.

2. LEI DE TILDEN *VERSUS* LEI DE ASIMOV NO CONTEXTO DA ROBÓTICA

Mark Tilden, cientista, autor e consultor técnico em cenas de filmes envolvendo robôs⁴, tornou-se famoso por sua invenção, na década de 1990, do que ficaram conhecidas como Leis da Robótica de Tilden, quais sejam:

1. um robô deve proteger sua existência a todo custo;
2. um robô deve obter e manter acesso à sua própria fonte de energia;
3. um robô deve procurar continuamente por uma melhor fonte de energia.⁵

Essas leis não parecem implicar algo como amizade. Inicialmente, essas três regras foram sugeridas como uma indicação de novos princípios de *design* para uma abordagem mais interativa incorporada à arquitetura robótica, o que ecoou em ficção científica. Entendemos que as leis de Tilden podem ser úteis, quando robôs não interagem diretamente com seres humanos; este é o caso, por exemplo, de robôs espaciais, construídos para explorar outros planetas ou cometas, como as sondas *Curiosity*⁶ e *Rosetta*⁷. Entretanto, há razões para nos preocuparmos com as leis de Tilden, quando estas são instanciadas em robôs que interagem com seres humanos ou, ainda pior, em robôs designados para ferir seres humanos. Este é o caso dos projetos de desenvolvimento de robôs autônomos com fins militares. Atualmente, há um investimento para produção de robôs autônomos programados para atuar em guerras e, em um sentido não específico, proteger a vida dos humanos “aliados”. O que aconteceria, se esses robôs se tornassem totalmente autônomos? Enfatiza Wareham (2014, tradução nossa): “Se o desenvolvimento de robôs militares não for controlado, a preocupação é que essas máquinas podem, em última análise, tomar decisões sobre a vida-ou-morte [de pessoas] no campo de batalha ou na aplicação da lei.”

⁴ Disponível em: <http://en.wikipedia.org/wiki/Mark_W._Tilden>. Acesso em: mar. 2016.

⁵ Disponível em: <http://en.wikipedia.org/wiki/Tilden%27s_Law_of_Robotics>. Acesso em: mar. 2016.

⁶ Segundo a NASA, este é o objetivo da sonda *Curiosity*: “Determinar se Marte já foi capaz de sustentar vida microbiana”. Veja mais em: <https://www.nasa.gov/mission_pages/msl/index.html>. Acesso em: mar. 2016.

⁷ De acordo com a Agência Espacial Europeia, o objetivo da espaçonave *Rosetta* é “[...] encontrar o cometa 67P/Churyumov-Gerasimenko onde se estudará o núcleo do cometa e seu ambiente por aproximadamente dois anos, e pousar uma sonda em sua superfície.” Veja mais em: <http://www.esa.int/Our_Activities/Space_Science/Rosetta_overview>. Acesso em: mar. 2016.

Antecipando essa tendência, Wiener (1970, p.75-76) previu o problema da modelagem de máquinas autônomas, no contexto da guerra:

Há um perigo real especialmente em nossos dias, com a possibilidade de usar máquinas que aprendem, que podem fazer a guerra, quanto ao controle de quando se deve apertar o botão que desencadeia uma guerra mundial. O único meio de evitá-lo é considerar a máquina não como um objetivo em si, mas como um meio de satisfazer as necessidades do homem [ser humano], como parte de um sistema humano-mecânico.

Preocupado com a possível construção de robôs que pudessem ameaçar a vida do ser humano, Isaac Asimov propôs suas Três Leis da Robótica⁸:

1. um robô não pode ferir um ser humano ou, por inércia, permitir que um ser humano sofra algum mal;
2. um robô tem que obedecer às ordens atribuídas por um ser humano, exceto se essas ordens possam entrar em conflito com a primeira lei;
3. um robô deve proteger sua própria existência, na medida em que essa proteção não entre em conflito com a primeira e a segunda leis.

Essas leis parecem acarretar a possibilidade de uma amizade entre robôs e seres humanos. Elas sustentam que robôs devem, basicamente, servir como ferramentas seguras para os seres humanos. Entretanto, é difícil entender até que ponto a segunda lei de Asimov sobre obediência pode permitir a emergência de uma amizade verdadeira, pelo menos se entendermos que amizade se constitui entre seres iguais com um alto grau de autonomia. Agir educada e amigavelmente não é o mesmo que ser amigo. Asimov espera que a robótica focalize simplesmente na construção de sistemas seguros para serem utilizados por humanos.

Essa breve indicação da história ligada ao debate sobre a natureza das relações humano-robô é suficiente para supor que muita atenção tem sido dada às questões de segurança e de serviços públicos, enquanto pouca atenção é dada às questões de relacionamento interpessoal mais profundas, como

⁸ Disponível em: http://en.wikipedia.org/wiki/Three_Laws_of_Robotics. Acesso em: mar. 2016.

a de amizade genuína. Na amizade genuína, como trataremos na próxima seção, pressupõe-se algum tipo de vínculo social, afetivo e emocional de ambas as partes. No contexto dos robôs interativos mais recentes, Akira Kobayashi, diretor da casa de repouso *Fu Yuen*, em Yokohama, expressa esperança em relação ao desenvolvimento de máquinas mais práticas, capazes de reduzir a carga física de cuidadores, por exemplo, ajudando a carregar e transportar os residentes idosos. Por outro lado, segundo Iida (2013, tradução nossa), Kobayashi pensa que “[...] o coração das pessoas, o cuidado e o calor do contato físico jamais poderão ser substituídos por robôs.”

Um exemplo de robô interativo pode ser encontrado no Roburso, um robô desenvolvido pelo Instituto Ricken e *Sumitomo Riko*. Esse robô foi designado para levantar pacientes imobilizados e/ou com a saúde debilitada e transferi-los para a cadeira de rodas, para a cama ou levá-los para tomar banho. Outro exemplo de um relacionamento entre robô e humano é a interação dos residentes da casa de repouso *Fuyouen* com o pequeno robô Paro (Figura 1). “Você é tão fofo, olha pra mim!” disse o residente Yukiko Kanesaka, de 92 anos, para o robô parecido com uma foca, que respondeu ao estímulo mexendo sua cabeça e suas pernas, guinchando e piscando seus olhos (Yokohama, em 10 de maio, 2013). Os cuidadores da casa de repouso dizem que os residentes aceitaram os robôs mais facilmente do que o esperado e que a interação com Paro teve um efeito positivo, trazendo sorrisos até aos rostos de pacientes em estado depressivo.

Figura 1 – Robô Paro



Copyright: National Institute of Advanced Industrial Science and Technology (AIST)

Mas, seria possível, de acordo com as leis de Asimov, modelar um robô que constituísse uma relação genuína de amizade? Seriam a autonomia e a habilidade de autocontrole essenciais para o desenvolvimento de robôs capazes de agir amigavelmente? Que tipo de inteligência um robô deveria possuir, para o estabelecimento de uma amizade com um ser humano? O que pode estar faltando em um modelo de robô “amigo”? Com o objetivo de discutir essas questões, devemos antes nos perguntar o que é amizade.

Amizade tem sido considerada como um vínculo social, uma organização não-institucional, uma forma de amor, um conjunto de práticas relacionadas ao conceito de semelhança familiar ou simplesmente uma forma recorrente de interação não antagonista entre agentes (BELL; COLEMAN, 1999; BRANDT; HEUSER, 2011; BRENT et al., 2014; DIGESER, 2013; HELM, 2010; KONSTAN, 1997; PAKALUK, 1991; SEYFARTH; CHENEY, 2012; VERNON, 2005). Na próxima seção, sugerimos uma caracterização semiótica de amizade que ressalta uma interpretação cognitiva mútua de si mesmo por um outro, e de outro por si mesmo, como uma relação triádica irreduzível.

3. UMA CARACTERIZAÇÃO SEMIÓTICA DE AMIZADE

A fim de desenvolver uma concepção mais geral de amizade, com especial ênfase em seus aspectos emocionais e cognitivos, sem adentrar em uma abordagem psicológica, focalizamos a perspectiva da *semiose*, isto é, da ação do signo, tentando compreender como a amizade pode ser concebida em sua estrutura relacional. Entendemos amizade como uma forma de amor, um vínculo social concretamente situado e incorporado em uma relação essencial de personalidade, por possuir um *self* e dispor de uma responsabilidade para com a comunidade à qual pertence. Consideramos que esses aspectos podem ser vistos de acordo com a ideia esquemática da ação do signo ou da lógica dos signos.

Propomos que a amizade seja concebida como uma relação engendradora em uma estrutura triádica irreduzível, basicamente similar à caracterização de signo, na semiótica peirciana, isto é, a lógica geral da ação do signo. Nessa perspectiva, amizade, entendida como uma categoria relacional, pode ser compreendida como uma complexa estrutura sógnica, a qual, por ser triádica, se torna mais rica que uma mera relação diádi-

ca. Entendemos por “mais rica” a não redutibilidade às categorias menos complexas. Por exemplo, um signo pertencente à categoria de mediação (ou terceiridade, no vocabulário peirciano), pressupõe, mas não pode ser reduzido, às categorias como de existência (secundidade) e de qualidade (primeiridade).

Uma primeira conceitualização de amizade, de acordo com a perspectiva semiótico-pragmática, pode ser assim formulada:

(F1) Um Amigo é um Primeiro (um *self*) que se coloca numa relação triádica genuína para com um Segundo (um *Outro*), de modo a ser capaz de determinar que um Terceiro (denominado *Amizade*), assuma a mesma relação triádica para com o *Outro* que ele próprio (*self*) assume para com aquele mesmo *Outro*.

Um modo de interpretar essa caracterização, evitando identificar o “*self* psíquico” com o *self* ao qual um amigo experiencia e interpreta, seria considerar o Terceiro como o componente normativo de uma estrutura medeiática como, por exemplo, um princípio-guia que integre ações individuais em um conjunto de normas, ou a ideia de amizade como um signo complexo constituído de muitas ideias em constante interação. Assim, um Primeiro, dentro de uma relação triádica, não é simplesmente o *self* psíquico de uma pessoa, mas um *self* que se expressa e se interpreta, bem como é interpretado por um *self* que não ele mesmo. Podemos denominar *outro-self* (para aludir à noção aristotélica de amigo como um *outro-self*) a representação processual do *self* por outro *self* que não ele mesmo⁹. Além disso, como o *Outro* é um Segundo, i.e., uma pessoa existente, a qual chamamos de “você”, em uma relação que é, ou está para se tornar, uma relação mútua, nós podemos denominá-la *Outro-Você*. Assim, propomos uma formulação expandida da noção de amizade:

⁹ Alguém poderia questionar, como um de nossos pareceristas o fez: “Se o *self*, na relação triádica, não é simplesmente o *Self* psíquico, cabe a dúvida: como se opera esta transformação do universo psíquico para o medeiático? Ou a *Amizade*, na relação triádica, tampouco é a *Amizade* psíquica?” Entendemos que, ao descrever os aspectos da amizade, seja através de uma estrutura relacional semiótica, seja como uma relação emocional-psicológica entre duas pessoas, seja ainda como algum outro fenômeno (por exemplo, social), tais descrições não constituem necessariamente mundos distintos (passando por mútuas transformações); elas simplesmente se adequam, em um certo sentido, aos modelos de amizade os quais podem (ou não) capturar os aspectos essenciais dessa relação.

(F2) Um Amigo, ou um *Outro-self*, é um Primeiro que se coloca numa relação triádica genuína para com um Segundo, denominado seu *Outro-Você*, de modo a ser capaz de determinar que um Terceiro, denominado sua *Amizade*, assuma a mesma relação triádica para com *Outro-Você* que ele próprio assume com aquele mesmo *Outro-Você*.

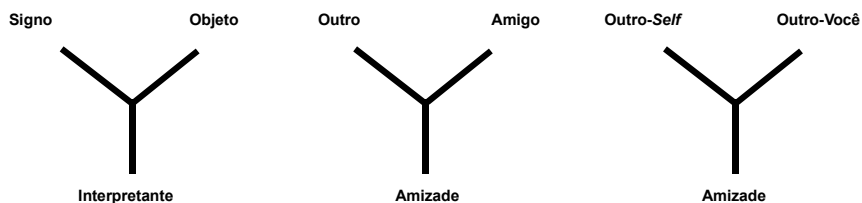
Essa relação triádica é genuína, i.e., seus três correlatos (os dois amigos e a amizade) estão unidos de uma maneira que não consiste em um conjunto complexo de relações diádicas. Essa definição é uma analogia formal da caracterização de signo proposta por Peirce, em 1902, qual seja:

(S1) “Um Signo, ou *Representamen*, é um Primeiro que se coloca numa relação triádica genuína para com um segundo, chamado seu *Objeto*, de modo a ser capaz de determinar que um Terceiro, chamado seu *Interpretante*, assuma a mesma relação triádica para com seu Objeto que ele próprio assume para com aquele mesmo Objeto”. (CP.2.274, tradução e ênfase nossas).

Pertencendo, assim como os signos, à categoria de mediação ou Terceiridade, a amizade se constitui em uma relação de desenvolvimento interpretativo, possibilitando um tipo de cognição distribuída em que os amigos que convivem e compartilham suas experiências também pensam conjuntamente, através de uma comunicação dialógica. Amizade, concebida, no sentido peirciano, como um interpretante (uma dinâmica e complexa estrutura triádica), não precisa ser uma instância explícita ou um processo de interpretação; ao contrário, é uma trajetória conjunta experienciada em pequena comunidade que traduz dois *selves* individuais dentro de uma unidade entrelaçada, um desenvolvimento de hábitos através de uma ação dialógica, influenciada por normas de amizade que definem o caráter da relação¹⁰. Na Figura 2, esboçamos o modelo de amizade inspirado na caracterização peirciana de signo.

¹⁰ O modelo semiótico de amizade é compatível com a análise do *self* em amizade proposta por Cocking e Kenneth (1998). Os autores ressaltam que, “[...] como um amigo próximo ao outro, um é caracteristicamente e distintivamente receptivo a ser direcionado e interpretado e, assim, nesses caminhos traçados pelo outro.” (1998, p. 503). Assim, por conta dessa receptividade aos interesses e interpretações de minhas peculiaridades, “[...] eu desenvolvo um caminho que é específico a este re-

Figura 2 – Correlação entre a definição de signo e as duas definições de amizades aqui propostas.



Fonte: Elaboração dos autores.

Uma possível atividade interpretativa é pressuposta pelo conceito de signo, sendo que essa atividade não deve ser realizada por uma mente individual. O interpretante pode ser um agente individual, mas não é necessariamente um indivíduo. Muitas vezes, Peirce entende o interpretante como uma atividade em uma comunidade (SKAGESTAD, 2004, p. 245). Assim, a abordagem peirciana permite uma caracterização mais geral de “múltipla agência” (*plural agency*), em que podemos encontrar na amizade genuína¹¹ uma forma de “agência” que possui uma estrutura triádica, mesmo sendo aparentemente uma relação diádica entre dois amigos.

Creemos que a amizade jamais poderá ser um relacionamento apenas entre dois indivíduos, pois ela sempre envolve um terceiro domínio, o qual é mediático, não apenas em um aspecto intelectual, mas também relacionado a normas, valores, afetos, ou o que podemos chamar de interpretante emocional, entendido aqui como sendo o primeiro efeito significativo e característico da amizade (PEIRCE, CP.5.475). Nesse contexto, a amizade pode ser considerada uma unidade triádica constituída por dois amigos e sua relação de mútuo interesse, que gera implicações para a comunidade como um todo.

lacionamento; o *self* que meu amigo vê é, pelo menos em parte, um produto da amizade.” (1998, p. 505). Nessa abordagem, o *self* é dinâmico e, também, precisamos acrescentar, um fenômeno semiótico; a saber, “[...] o *self* em uma amizade é, em parte, uma coisa que é constituída por e em particular para a amizade.” (1998, p. 510).

¹¹ Compare-se com Helm (2010), cuja perspectiva de “múltipla agência” é davidsoniana.

As descrições semióticas de amizade, propostas em F1 e F2, são formais e implicam uma estrutura normativa que pode se romper, uma vez que as normas de amizade não forem cumpridas, assim como uma amizade real pode terminar, se os amigos já não participarem dessa relação dialógica. Ou, ainda, o que parecia ser uma amizade real, pode tornar-se outro tipo de relação, caso as condições estruturais não forem atingidas. Isto acarreta a forma de avaliação dos tipos de relações que possam existir entre um robô e um ser humano, porque é altamente questionável se alguma vez um robô poderá ser construído não só para simular, mas também para vivenciar verdadeiros sentimentos de afeto para com outro ser. Abordamos essa questão na próxima seção, focalizando as noções de autonomia e autocontrole.

4. O PAPEL DO AUTOCONTROLE NAS RELAÇÕES DE AMIZADE

Se pensarmos a amizade como uma relação triádica que envolve um *self* e um *outro-self*, constituindo uma terceira instância guiada por normas, devemos questionar: como o desenvolvimento de robôs autônomos pode auxiliar no entendimento da relação de amizade entre humanos e robôs? Compreendemos que os conceitos de autocontrole e de hábito, tomado como um conjunto de regras direcionadoras da ação, podem lançar alguma luz sobre essa questão. Nesse contexto, o autocontrole dependeria da capacidade semiótica do organismo de quebrar e alterar hábitos. Como um pensador não-dualista e sinequista, Peirce reconhece uma continuidade de graus de autocontrole no interior de um organismo como o ser-humano. Segundo Peirce (CP.5.533, tradução nossa), temos:

[...] modos de autocontrole que parecem bastante instintivos. Em seguida, há um tipo de autocontrole que resulta de treinamento. Em seguida, um homem pode ser seu próprio treinador e assim controlar seu autocontrole. Quando este estágio é alcançado, grande parte ou todo o treinamento pode ser conduzido na imaginação. Quando um homem se treina, assim controlando o controle, ele deve ter alguma regra moral em vista. Entretanto, esta regra pode ser especial e irracional. Mas, em seguida, ele pode encarregar-se de melhorar esta regra; isto é, exercitar um controle sobre seu controle de controlar. Para fazer isso, ele precisa ter em vista algo superior a uma regra irracional. Ele deve possuir um tipo de princípio moral. Este, por sua vez, pode ser controlado por referência a um ideal estético sobre o que é bom. Há,

certamente, mais graus do que eu tenho enumerado. Talvez esse número seja indefinido.

Nesse aspecto moral do autocontrole, a amizade e outras relações sociais podem desempenhar um papel sobre as normas sociais e a sociabilidade em geral. Essas relações ensinam os indivíduos a diminuir a distância entre um ideal estético (como “sobre o que é bom”) e as ações reais no mundo existente, além de os ensinar a refletir criticamente sobre as regras em voga, numa tentativa de melhorá-las. Isso também pode ser estendido aos “robôs autônomos”, os quais - talvez ironicamente - tenham que se tornar dependentes de uma comunidade mais abrangente, para serem considerados tão “autônomos” quanto os humanos, possuindo, assim, a habilidade de exercer o controle sobre seu controle de controlar.

Ainda não encontramos alto grau de autocontrole em robôs, nem sinais de amizade. Evidentemente, é difícil prever, em um futuro próximo ou distante, quão complexo pode ser o comportamento de robôs, equipados com algoritmos de aprendizado inteligente capazes de se adaptarem em sua interação com humanos. Algo para se notar é que robôs certamente serão capazes de mais de um grau de autocontrole. Analogamente, parece que a diferença entre humanos e robôs é, em parte, devida ao nosso grande número de graus de autocontrole e habilidades semióticas, os quais possibilitam nossa versatilidade, adaptabilidade e diversidade de ajustes em culturas distintas. Isso é, em parte, devido a um arcabouço emocional de autocontrole que acrescenta algo qualitativo, um elemento de primeiridade ou sentimento que parece faltar na construção de sistemas artificiais como robôs.

A versatilidade humana, ou seja, nossa capacidade de quebrar e alterar hábitos, é igualmente indicada pela alta variedade de tipos de amizade, sendo vista como uma expressão da liberdade facultada pelo alto grau de autocontrole e semiotividade presente no ser humano. Segundo nosso entendimento, nossa versatilidade exige um tipo de inteligência que podemos também encontrar em robôs, mas a questão que colocamos aqui é: qual tipo de inteligência é necessária para a constituição de uma relação de amizade?

A inteligência referente à amizade não diz respeito apenas à habilidade de calcular ou prever fatos e eventos. Em vez disso, podemos enten-

dê-la como uma inteligência emocional que consolida um princípio-guia que direciona os componentes constituintes de uma amizade, ao realizar ajustes na conduta desses componentes. Talvez, inteligência não deva ser concebida apenas como uma “alta capacidade para resolução de problemas em geral” (abordagem da área de engenharia, como a Inteligência Artificial Clássica). Tal inteligência, quiçá, deva ser considerada como um “alto grau de animalidade”, que também fundamenta a inteligência “científica” humana, isto é, uma inteligência capaz de aprender com a experiência (PEIRCE, CP 2.227). Nesse contexto, os ajustes realizados entre dois amigos permite o estabelecimento de um padrão de conduta que denominamos amizade. Esses ajustes exigem um aprendizado emocional, através de uma relação semiótica triádica que caracteriza a amizade.

O princípio-guia constituído no relacionamento entre *outro-self* e *outro-você* engendra sentimentos de expectativa que retornam (via *feedback*) na conduta futura desse mesmo relacionamento. Uma quebra de expectativa pode gerar um efeito de aprendizado tanto em robôs quanto em humanos. Entretanto, o tipo de aprendizado pode ser diferente; quando constituímos um relacionamento, não somos mais os mesmos: “[...] cada amigo representa um mundo em nós, um mundo possivelmente não existente até a chegada deles, e é somente por esse encontro que esse novo mundo nasce.” (NIN, 1937, tradução livre). Crescemos e nos tornamos dependentes de nosso *outro-você*. Tentamos nos ajustar ao novo relacionamento, ao mesmo tempo em que alteramos nossa conduta perante outros amigos e diante da comunidade em geral. A questão que surge aqui é: um robô poderia se alterar por seu próprio bem ou pelo bem de seu amigo, ou ele apenas aprenderia como ser agradável e gentil para ser o amigo de um humano? Talvez essa questão não seja difícil de ser respondida, se lembrarmos que, atualmente, os robôs não são construídos para *serem* felizes, muito pelo contrário, eles são modelados para *parecerem* felizes e propiciarem felicidade aos humanos (pelo menos, na melhor das alternativas). A partir dessa perspectiva, indicamos quatro elementos básicos que não encontramos (ainda) em uma amizade entre humanos e robôs, sendo característicos de uma possível diferença entre humanos e robôs: (1) comprometimento emocional, (2) expectativa, (3) ajustes e (4) irreversibilidade.

O comprometimento emocional é encontrado nos afetos envolvidos em um relacionamento, por exemplo, quando você é mais que um, isto

é, quando você encontra certos tipos de compreensão compassiva por um amigo, dói voltar a ser apenas você mesmo. Isso pode acontecer quando uma relação se torna algo instrumental direcionado por normas externas à amizade genuína, ou quando as conexões, por alguma razão, já não são mais possíveis. Até onde compreendemos acerca dessa questão, não acreditamos que os robôs atuais sejam capazes de sentir compaixão ou a dor da perda.

A segunda característica é a expectativa que uma amizade genuína propicia. Numa relação triádica, quando você convive com seu amigo, há a constituição de um hábito – o terceiro elemento, a amizade – e esse hábito direciona a conduta de acordo com alguma regra de ação que constitui, por sua vez, um sentimento de expectativa em relação a eventos futuros. Em uma amizade entre humanos e robôs, parece que o robô (na medida em que seu programa funcione corretamente) não seria capaz de quebrar a expectativa do humano, já que ele irá se adequar e se ajustar de acordo com os hábitos de seu interlocutor. Essa característica nos leva para o terceiro elemento elencado aqui, qual seja, ajustes.

Em um relacionamento genuíno, há ajustes e microajustes de hábitos, maneiras e condutas entre os elementos que o constituem. Entendemos que tais ajustes devam existir em ambas as direções: o amigo (*outro-self*) ajusta-se ao *outro-você*, assim como o *outro-você* se ajusta de acordo com o *outro-self*. É a procura de sintonia que favorece o desenvolvimento da amizade, da habilidade de lidar com as diferenças e de tolerar a diversidade. Se pensarmos sobre os possíveis efeitos de uma amizade na comunidade, podemos entendê-la como uma rede de relações triádicas: uma amizade pode modificar, via *feedback*, os amigos (*outro-self* e *outro-você*), que, por sua vez, podem alterar seus amigos de diferentes relacionamentos e assim por diante, constituindo uma rede dinâmica de interações. É uma questão em aberto a capacidade de robôs se ajustarem para além do propósito da adequação aos hábitos do ser humano.

A última característica é a irreversibilidade, isto é, a impossibilidade de voltar no tempo e ser o mesmo de antes. Uma vez que você teve um amigo, você não pode simplesmente voltar ao estágio em que estava antes de encontrá-lo. Isso acontece, como vimos, porque em um relacionamento entre dois amigos há a emergência de uma terceira instância – a amizade – que modifica seus elementos. Esse processo não parece acontecer com robôs, os quais podem ser – em princípio – reciclados e reiniciados.

Em suma, argumentamos, na presente seção, que, apesar de robôs poderem *agir amigavelmente*, ainda não é claro em qual medida eles podem estabelecer *relações de amizade*, a menos que robôs possuam a mesma capacidade emocional, cognitiva e social de interpretação de signos que os humanos possuem. Nesse sentido, a amizade pode ser concebida como possuindo um *fim em si mesma*, ao invés de ser algo com fins ou propósitos extrínsecos. Às vezes, pode ser o caso em que uma amizade genuína se estabeleça de alguma relação que, no início, foi definida por propósitos externos à amizade; por exemplo, colegas de trabalho, de academia ou de escola, amigos de amigos etc. Mesmo se inicialmente o relacionamento é constituído, como Aristóteles (1985) diria, por “amigos de utilidade” ou “amigos de prazer”, acreditamos que, com base nessas relações, possa se constituir de “amigos de virtude” ou “amigos genuínos”. A habilidade de edificar uma amizade genuína pode resumir a diferença fundamental entre um relacionamento humano-humano, em detrimento ao robô-humano.

5. CONSIDERAÇÕES FINAIS

Para concluir, retornemos ao exemplo do Paro, o robô interativo que se parece com uma foca. A despeito de agir amigavelmente, o robô Paro parece não ultrapassar a barreira da construção de máquinas que *simulam* comportamento e afeto, possibilitando algum tipo de projeção de amizade. De acordo com as Leis de Asimov, seria razoável construir robôs que fossem apenas ferramentas para uma melhor condição da vida humana. Talvez não seja desejável que robôs se tornem mais que uma ferramenta para os seres humanos, seja no caso da amizade com robôs enfermeiros, seja no caso de robôs militares. Nesse contexto, há uma necessidade de levar a sério as Leis de Asimov, não permitindo a construção de robôs totalmente autônomos, os quais possam ameaçar a vida do ser humano.

Felizmente, os robôs que conhecemos hoje em dia são apenas quase-autônomos no sentido de serem, conforme Emmeche (2007, p. 477), “[...] conectados em redes sócio-técnicas, incluindo técnicos humanos que os mantêm funcionando.” Os robôs de hoje são construídos para *agir amigavelmente* com respeito ao ser humano. Eles podem ser programados para mostrar um tipo de comportamento amigável ou buscar aliados para reali-

zação de uma colaboração conjunta. Lembramos que agir amigavelmente, entretanto, não é o mesmo que possuir relações de amizade.

Concebendo o modelo “moderno” de amizade como exemplo, a amizade entre robôs demandaria um alto grau de autocontrole, que, até onde sabemos, não foi possível instanciar nos sistemas artificiais (o que não torna essa instanciação impossível *per se*). Além disso, a questão não pode ser respondida dentro de um quadro puramente individualista, como é facilmente pressuposto, quando se fala em “robôs autônomos”. A amizade, compreendida através de uma perspectiva semiótica, constitui uma pequena comunidade moral e dialógica em interação com as normas e os valores da comunidade em geral. Igualmente, a amizade entre robôs e humanos, para ser possível, não poderia ser concebida como uma díade fechada, mas apenas como uma relação triádica e social. Talvez isso se aplique aos robôs, visto que eles não são simplesmente uma entidade ou coisa, mas, sim, um produto do trabalho coletivo e da ingenuidade humana.

AGRADECIMENTOS

A presente pesquisa é financiada pela Universidade de Copenhagen, pela Fundação Velux (processo 437810 – “Humanomics”) e pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). Agradecemos a David Budtz Pedersen, Frederik Stjernfelt, Simo Køppe e aos outros membros do Grupo de Pesquisa *Humanomics*. Agradecemos também a Maria Eunice Quilici Gonzalez, Mariana Claudia Broens, Marcos Antonio Alves, Maria José Vicentini Jorente, Susan Haack, Henrik Zinkernagel, Rasmus Grønfeldt Winther, Theresa Schilhab, Di Ponti e Jon Nixon pelos comentários, auxílios, sugestões, críticas e apoio. Finalmente, gostaríamos de agradecer aos organizadores do IX EIICA - Encontro Internacional de Informação, Conhecimento e Ação – e aos nossos pareceristas.

REFERÊNCIAS

- ARISTÓTELES. *Ética a Nicômacos*. Tradução do grego, introdução e notas de Mário da Gama Kury. Brasília: Editora da UnB, 1985.
- BELL, S.; COLLEMAN, S. (Ed.). *The anthropology of friendship*. Oxford & New York: Berg, 1999.

- BRANDT, A.; HEUSER, E.A. Friendship and socio-cultural context. Experiences from New Zealand and Java/Indonesia. In: DESCHARMES, B. et al. (Ed.). *Varieties of Friendship Interdisciplinaty perspectives on social relationships*. Göttingen: V&R Unipress, 2011. p. 145-174.
- BRENT, L. J. N.; CHANG, S. W. C.; GARRIÉPY, J.-F.; PLATT, M. L. The neuroethology of friendship. *Ann. N. Y. Acad. Sci.*, v. 1316, p. 1-17, 2014.
- COCKING, D.; KENNETT, J. Friendship and the self. *Ethics*, v. 108, n. 3, p. 502-527, 1998.
- DAGG, A. I. *Animal Friendships*. Cambridge, UK: Cambridge University Press, 2011.
- DIGESER, P.E. Friendship as a family of practices. *AMITY: The Journal of Friendship Studies*, v. 1, p. 34-52, 2013.
- EMMECHE, C. A biosemiotic note on organisms, animals, machines, cyborgs, and the quasi-autonomy of robots. *Pragmatics & Cognition*, v. 15, n. 3, p. 455-483, 2007.
- _____. *Interdisciplinary challenges in the study of friendship*. (no prelo).
- _____. Robotic friendship: Can a robot be a friend? *International Journal of Signs and Semiotic Systems*, v. 3, n. 2, p. 26-42, 2014.
- GIDDENS, A. *The Transformation of intimacy*. Sexuality, love and eroticism in modern societies. Cambridge, UK: Polity, 1992.
- HELM, B. W. *Love, friendship, and the self: Intimacy, identification, and the social nature of persons*. Oxford: Oxford University Press, 2010.
- IIDA, M. Robot niche expands in senior care. *Japan Times News*, June 19, 2013. Disponível em: <<http://www.japantimes.co.jp/news/2013/06/19/national/social-issues/robot-niche-expands-in-senior-care/>>. Acesso em: mar. 2016.
- JAMIESON, L. *Intimacy*. Personal relationships in modern societies. Maldon, MA: Polity, 1998.
- KONSTAN, D. *Friendship in the Classical World*. New York: Cambridge University Press, 1997.
- NIN, A. In: STUHLMAN, G. (Ed.). *The Diary of Anaïs Nin*. V. 2: 1934-1939 (p. 193). Orlando, Florida: Harcourt Brace & Company, 1937/1967.
- PEIRCE, C.S. *Collected Papers of Charles Sanders Peirce*. V. 1-6. [HARTSHORNE, C. & WEISS, P. (Ed.); v. 7-8, BURKS, A.W. (Ed.)]. Cambridge, Mass.: Harvard

University Press. (References: CP, followed by volume and paragraph number), 1931-1958.

PAKALUK, M. (Ed.). *Other selves*. Philosophers on friendship. Indianapolis: Hackett, 1991.

SEYFARTH, R. M.; CHENEY, D. L. The Evolutionary Origins of Friendship. *Annu. Rev. Psychol.* v. 63, p. 153-177, 2012.

SINGER, I. *The nature of love*. V. 1-3. Cambridge, Mass.: The MIT Press, 2009.

SKAGESTAD, P. Peirce's semiotic model of the mind. In: MISAK, C. (Ed.). *The Cambridge Companion to Peirce*. Cambridge: Cambridge University Press, 2004. p. 241-256.

VERNON, M. *The philosophy of friendship*. Houndmills, Basingstoke: Palgrave Macmillan (2. ed. published 2010 as *The meaning of friendship*, same publisher), 2005.

WAREHAM, M. Why the world should ban autonomous weapon systems. *The mark news*. Disponível em: <<https://www.hrw.org/news/2014/11/28/killer-robots-why-world-should-ban-autonomous-weapons-systems>>. Acesso em: 16 nov. 2015.

WIENER, N. O Homem e a Máquina. In: KÜHNER, M. H. *O conceito de informação na ciência contemporânea*: Colóquios filosóficos internacionais de Royaumont. Rio de Janeiro: Paz e Terra, 1970. p. 69-99.