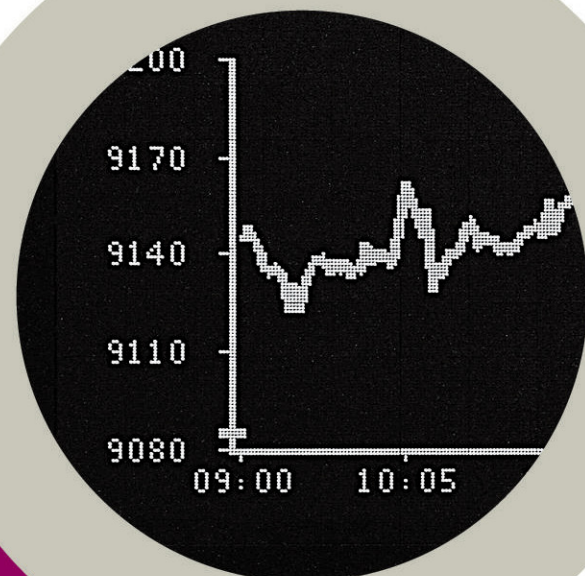


**Revolução** Aproxima-se um momento histórico na ciência, em que as regras morais vão ser introduzidas na programação dos computadores. Mas há cientistas que falam numa ameaça ao nosso modo de vida

# Quando a moral entra nos computadores



Textos **VIRGÍLIO AZEVEDO**

Deverá um *drone* bombardear uma casa onde o inimigo está escondido mas que, ao mesmo tempo, serve de abrigo a um grupo de civis? Deverá um automóvel sem condutor mudar de direção para evitar atropelar uma pessoa, mas chocar com outro que transporta várias pessoas? Deverá um robô de assistência hospitalar forçar uma idosa a tomar um remédio numa situação crítica, mesmo se ela recusar?

Estes são dilemas que, até há pouco tempo, eram apenas do domínio da ficção científica, mas que estão a chegar rapidamente à realidade do nosso dia a dia quase sem darmos por isso. Mas resolvê-los passa por uma verdadeira revolução, porque apela ao desenvolvimento da Moral Computacional, um novo ramo das ciências da computação que pretende incluir regras morais na programação dos computadores e dos robôs, no *software* das máquinas que têm algum grau de autonomia de decisão, isto é, onde não existe intervenção humana. E que pode também ajudar ou mesmo treinar os seres humanos no processo de decisão moral, bem como compreender melhor as questões da moralidade através de modelos experimentais de simulação.

## Investigação pioneira em Lisboa

Luís Moniz Pereira, que criou o Centro da Inteligência Artificial (Centria) da Faculdade de Ciências e Tecnologia (FCT) da Universidade Nova de Lisboa, onde está a ser desenvolvida investigação pioneira a nível mundial nesta área emergente, “porque envolve a criação de modelos que podem ser aplicados em qualquer área da moral. E porque inclui o conceito de *updates* (atualizações) morais, isto é, a sobreposição de novas regras morais para aperfeiçoar as já existentes”. O professor do Departamento de Informática da FCT salienta que “a inteligência artificial está pronta a servir de instrumento para lidar com o problema importante que é a moral”. Os cientistas “estão ainda no princípio desta investigação, mas a moral tem de ser computadorizada e quanto mais cedo melhor”. Não o fazer “pode ser muito perigoso para a sociedade”.

A programação dos computadores e dos robôs tem sido feita para situações específicas, mas a moral “é constituída por um conjunto de regras que se aplicam a uma generalidade de situações, o que significa que é

feita para lidar com o imprevisível”. À medida que o software que comanda agentes (entidades ou programas que tomam decisões) ganha mais autonomia, “temos de introduzir regras mais genéricas e abstratas, como as regras morais dos humanos, mas que têm de ser sensíveis ao concreto”, isto é, que admitem exceções justificadas e aceites pela sociedade. No fundo, o grande desafio para os cientistas está em “capturar informaticamente todas as especificações que a moral pretende abranger”.

O debate científico sobre estas questões tem uma forte componente filosófica, mas diz respeito a realidades muito concretas, que já afetam milhões de pessoas em todo o mundo, e que estão a ficar fora de controlo. O caso mais típico é o software do sistema financeiro internacional, cuja autonomia nos processos de decisão contribuiu para uma crise cujas consequências vão continuar a marcar a vida dos portugueses nos próximos anos.

## Regras morais no software financeiro

“É o meu problema preferido”, confessa Luís Moniz Pereira. “A que normas devem obedecer os produtos financeiros para não serem tóxicos, ou os cartões de crédito para não levarem ao endividamento insustentável das famílias?”, pergunta o investigador, que defende a criação de “linguagens de programação no software financeiro que especifiquem regras morais — o que pode ou não poder ser permitido — estabe-

lecidas por um corpo regulador do sistema financeiro internacional, que deveria ser criado junto de uma organização de governação global como a ONU, por exemplo”.

Ari Saptawijaya, cientista indonésio doutorando em Moral Computacional, que pertence à equipa da Universidade Nova de Lisboa, explica que “esta investigação não pretende resolver todos os problemas da moralidade nem introduzir uma nova teoria moral. O objetivo é antes desenvolver um sistema computacional para modelar a moralidade”. Mais precisamente, “estamos empenhados em desenvolver um sistema, baseado na Programação em Lógica (que usa a lógica matemática), para modelar apenas alguns aspetos morais já bem estudados e reportados na Filosofia e na Psicologia”.

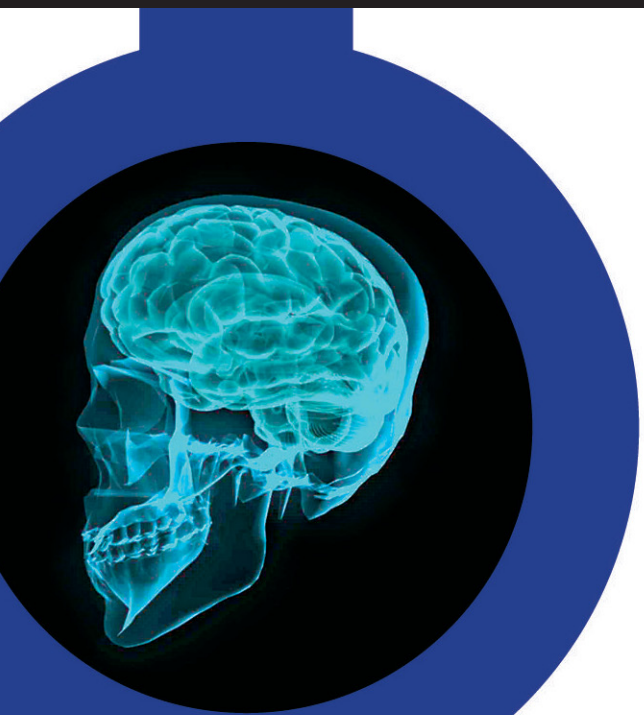
Num artigo científico que vai ser apresentado no final de janeiro numa conferência internacional em San Diego, na Califórnia (PADL'14), Luís Moniz Pereira e Ari Saptawijaya defendem precisamente o potencial da Programação Lógica como ferramenta para modelar três aspetos da moralidade: o processamento dual dos julgamentos morais, a justificação desses julgamentos e a intenção na permissibilidade moral. No primeiro aspeto, está em causa a interação e a competição ou cooperação entre dois sistemas psicológicos: intuitivo/racional (o julgamento moral é automático ou consciente?) e afetivo/cognitivo (o julgamento moral é conduzido por uma resposta afetiva ou por um raciocínio baseado em princípios morais?). O segundo aspeto é a capaci-

dade de um agente justificar o seu comportamento explicitando quais os princípios morais aceites pela sociedade que usou. O terceiro pretende distinguir as ações permitidas das não permitidas e saber se as últimas são conduzidas com a intenção de fazer mal a alguém.

“À medida que a Moral Computacional avançar, irá tornar os computadores e os robôs capazes de se comportarem mais perto da forma como nós, humanos, nos comportamos. Por isso, os robôs assistentes e outras tecnologias do mesmo tipo serão mais aceites pelas pessoas”, considera o investigador vietnamita The Anh Han, do Laboratório de Inteligência Artificial da Universidade Livre de Bruxelas, que se doutorou na Universidade Nova de Lisboa. Por outro lado, este novo ramo das ciências da computação “ajudará qualquer pessoa a compreender melhor as decisões tomadas por outros, o que contribuirá para uma sociedade com menos conflitos e mal-entendidos”.

vazevedo@expresso.imprensa.pt





## Compreender melhor os seres humanos

**Reproduzir comportamentos humanos em máquinas vai ter a vantagem de nos ajudar a entender melhor como nós próprios funcionamos**

A tentativa feita pelos cientistas de replicar nas máquinas certos comportamentos humanos “ajuda a compreender melhor como funcionam os próprios seres humanos”, considera o filósofo Porfírio Silva, colaborador do Instituto de Sistemas e Robótica (Instituto Superior Técnico) e do Centro de Filosofia das Ciências da Universidade de Lisboa.

Por isso, considera a Moral Computacional “muito útil, porque hoje há máquinas inteligentes por todo o lado que colaboram com os humanos e fazem parte do nosso ambiente”. Em contrapartida, o filósofo critica o facto de certas abordagens computacionais “serem demasiado simplistas face à complexidade dos problemas, e a Moral Computacional tende a ser um desses casos”.

Quando a engenharia estuda os problemas das ciências sociais e das humanidades, as suas abordagens computacionais “são geralmente muito individualistas, consideram a moralidade do ponto de vista de um só agente (entidade ou programa de tomada de

decisões), quando a moral é uma questão intrinsecamente social”. E nem sempre o problema individual “é a base do problema coletivo”. Mesmo quando os investigadores consideram uma população de agentes, “esta é frequentemente pouco heterogênea, o que não tem correspondência com a realidade humana”.

Francisco Santos, investigador do Grupo de Inteligência Artificial do INESC-ID (Instituto de Engenharia de Sistemas e Computadores), afirma que “é preciso perceber como os sistemas morais evoluem no tempo e como podem ser aplicados às máquinas”, para evitar problemas no futuro, quando se incluírem regras morais na programação dos computadores e dos robôs. No fundo, “as máquinas deverão ser desenhadas como sistemas adaptativos, que têm regras morais dinâmicas adaptadas ao ambiente em que são aplicadas, tal como acontece com a cultura humana”.

Fernando Cardoso, um filósofo brasileiro da Universidade Federal de Minas Gerais, que integra a equipa de investigação sobre Moral Computacional da Universidade Nova de Lisboa (ver artigo principal), defende que, “mais do que uma ameaça ou outra forma de catastrofismo, o uso de regras morais nos computadores e nos robôs vai ser uma forma de expansão da esfera moral e ética dos seres humanos para novos agentes”.

### QUATRO PERGUNTAS A

#### Manuel Curado

Especialista nas questões éticas da tecnologia e professor de Lógica no Departamento de Filosofia da Universidade do Minho

**■ O mundo da computação vai ser capaz de simular todas as valências da mente humana?**

**■** Estou convencido de que vai, no cálculo matemático, na percepção visual, na atenção visual, na decisão moral. Estamos a assistir ao momento da história da ciência em que isso vai acontecer.

**■ Está preocupado com esta evolução da ciência?**

**■** Estou ao mesmo tempo fascinado e preocupado, como académico. Estou fascinado porque há hoje arquiteturas computacionais que imitam o que pensávamos impossível, como a cooperação e o altruísmo, por exemplo. Tudo se joga na capacidade de a investigação computacional conseguir simular o cérebro humano e tudo indica que este tem uma estrutura computacional. Mas estou preocupado com as consequências, para a sociedade, da introdução da componente moral na computação. Imagine uma máquina com livre arbítrio, com capacidade de decisão diferente da nossa, com valores e prioridades diferentes dos nossos.

**■ Este tema está a ser suficientemente debatido?**

**■** Não, mas as Comissões de Ética e as Comissões Nacionais de Proteção de Dados dos diversos países deviam começar a refletir e a dar pareceres sobre o assunto. Neste momento, estão apenas preocupadas com os problemas da privacidade e do acesso aos dados pessoais, mas o grande desenvolvimento dos sistemas decisórios nas máquinas já está entre nós, não é para amanhã. E poderá desenvolver-se a um nível que as máquinas deixem de precisar dos seres humanos. Seremos imprescindíveis para pôr uma máquina a funcionar é uma verdade que nos deu conforto durante muito tempo, mas esta verdade está a desaparecer. E vamos precisar mesmo das três Leis da Robótica de Isaac Asimov (o criador da palavra robótica) ou de alguma versão delas. A primeira lei diz que um robô não deve fazer mal a um ser humano ou permitir que este sofra qualquer mal. A segunda estabelece que um robô deve obedecer a qualquer ordem dada por um ser humano, desde que não interfira na primeira lei. E a terceira diz que um robô deve proteger a sua existência, desde que isso não interfira nas duas leis anteriores.

**■ Os sistemas de decisão independentes da intervenção humana estão muito generalizados?**

**■** Sim, as máquinas estão a começar a decidir por nós na indústria, no comércio, na saúde, nos sistemas militares, e já hoje temos muitos sistemas informáticos (software) de apoio à decisão nas nossas vidas. Esta área vai evoluir muito nos próximos anos. A DARPA, uma agência do Pentágono, está a desenvolver *chips* — os implantes neuroprostéticos — para serem colocados no cérebro, de modo a melhorarem a memória, o nível de alerta ou o ciclo do cansaço. Esses *chips* poderão ter capacidades decisórias, como se fossem *pacemakers* cerebrais. Por outro lado, um dia poderão surgir máquinas que se apaixonam, que têm emoções ou que possuem preocupações éticas e morais. E a investigação na Moral Computacional vai chocar a 200 km/hora com as nossas vidas. Há mesmo questões filosóficas complicadas, como saber se a moral no ser humano que se pretende replicar num computador ou no robô tem uma base exclusivamente cerebral, como defendem muitos investigadores, ou se possui também bases exteriores ao cérebro, como as tradições, as regras sociais ou a cultura.

## Dos automóveis sem condutor aos *drones*

**A ausência de regras morais no software das máquinas que tomam decisões sem intervenção humana pode gerar problemas graves**

As grandes marcas de automóveis estão a preparar-se para lançar no mercado carros que andam sozinhos, que tomam decisões sem interferência humana. Para já, têm sido apresentados nos vários modelos experimentais, mas a Volvo, por exemplo, anunciou em dezembro que vai mesmo pôr 100 carros autónomos a circular nas ruas de Gotemburgo (Suécia) em 2017.

Trânsito mais seguro e melhor ambiente são vantagens apontadas pelo próprio Governo sueco. Mas “há um problema filosófico-jurídico que não está resolvido: de quem é a responsabilidade se qualquer coisa correr mal?”, pergunta Luís Moniz Pereira, fundador do Centro de Inteligência Artificial da Universidade Nova de Lisboa. Com efeito, se acontecer um acidente que provoque danos nos passageiros, nos peões ou em infraestruturas, não está ainda clarificado quem assume as culpas.

“Há um lóbi de pressão a nível internacional que defende não se-

rem necessárias novas leis, porque a responsabilidade é de quem comprou o automóvel e não do fabricante, da seguradora ou do legislador”, mas este é um caso “em que a necessidade de usar a Moral Computacional nestes automóveis está a tornar-se muito premente”, sublinha o investigador.

**Robôs de assistência a idosos**

Essa urgência existe noutras áreas, como a moralidade das prioridades de decisão dos robôs autónomos de socorro e salvamento em emergências e catástrofes, do *software* do sistema financeiro (ver artigo principal), ou dos robôs de assistência a idosos, que já existem a nível experimental em alguns hospitais.

No Japão, irão ser usados experimentalmente nos lares de pessoas idosas que vivem sozinhas, porque os seres humanos saem demasiado caros. São uma espécie de animais domésticos mecânicos, que têm de reconhecer intenções, fugas à rotina, dar medicamentos à hora certa, zelar para que haja comida e fazer encomendas diretas ao supermercado. A arquitetura dos próprios condóminos, que são edifícios inteligentes, está a ser concebida para

facilitar a circulação dos robôs.

Estes robôs poderão em breve ser usados também em instalações fabris ou na segurança em espaços públicos, como os parques de estacionamento automóvel. “A segurança passará a ser um misto humano/robô, em que as regras estabelecidas dirão que a última decisão é dos humanos, mas esta fronteira é problemática, como mostram os inúmeros problemas criados pelos *drones* militares — os *Predators* —, onde há muito pouco tempo para um humano decidir numa situação de ataque”, assinala Luís Moniz Pereira.

O problema da ausência de regras morais no *software* destes aparelhos, que se sobrepõe à decisão humana, estende-se ao uso de *drones* civis. No Texas, já existem associações que usam estes aparelhos aéreos autónomos para detectar imigrantes mexicanos ilegais e denunciá-los à polícia. No final de dezembro, a Administração Federal da Aviação dos EUA anunciou que vai autorizar seis instituições a realizar testes de segurança — mesmo quando as comunicações de rádio falham —, de modo a integrar estes aparelhos no espaço aéreo nacional a partir de 2015. Vem aí o caos aéreo? E como impedir a intrusão de *hackers* nos *drones*?