

## Inteligência Artificial: Um novo centro de poder

António Raimundo

Thomas Hobbes, em 1652, expressou a opinião que o “pensamento humano resulta da manipulação de símbolos” e afirmou mesmo que “quando um homem raciocina, não faz mais do que conceber uma soma total, pela adição de parcelas, ou conceber um resto, pela subtração de uma soma a outra”. Numa expressão mais elementar e simplificando o complexo, o cérebro humano é um processador de símbolos. Ora um computador é, no essencial, um mecanismo que processa símbolos, mas isso não o faz inteligente, mas simplesmente mais rápido. Em 1936, Alan Turing, matemático, demonstrou que qualquer computador com suficiente memória, que manipule símbolos, consegue fazer cálculos e Alonzo Church descreveu um mecanismo de cálculo que pode ser executado por uma máquina que processe símbolos. O conjunto destes dois princípios base, conhecido como tese Church-Turing, permite induzir que qualquer resultado que possa ser efetivamente calculado pode ser processado por uma máquina de Turing<sup>1</sup>. A partir desta tese, não podemos concluir que um sistema automático, um computador, possa ser considerado inteligente, mas, e tão só, que pode processar modelos matemáticos que simulam (emulam) funções consideradas como constituintes da inteligência humana – aprender, reconhecer, simular, etc. As técnicas utilizadas na emulação destas funções são suportadas em modelos, modelos matemáti-

cos, que simulam o funcionamento de neurónios e sinapses, em processos de complexidade crescente.

### “Passada a primeira fase – motores de regras e redes neuronais –, a IA entrou na fase de tratamento massivo de dados como fulcro da aprendizagem(...)”

O funcionamento destes modelos, por mais complexos que sejam, procura simular o funcionamento do cérebro humano, mas é, ainda, substancialmente diferentes daquele. António Damásio expressa de forma simples a situação atual ao mencionar que “tentaram emular aquilo que consideravam ser o mais essencial e útil – chamam-lhe de inteligência simples – e omitiram a outra componente da inteligência, os sentimentos, porque o universo dos afetos, como manifestação histórica de inteligência, é crucial para o aparecimento e desenvolvimento da criatividade<sup>2</sup>.

A investigação em Inteligência Artificial (IA) tem já várias décadas e centrou-se inicialmente nos Estado Unidos da América e na União Europeia. O termo Inteligência Artificial reporta-se a 1956 e ao “Dartmouth Summer Research Project on Artificial Intelligence”. A China, no presente, é um ator

determinante com bastante capacidade instalada (humana e técnica) e que avança a ritmos elevados. Estima-se que, em 2025, a maioria dos artigos de primeiro nível científico, publicados e citados, serão de produção chinesa<sup>3</sup>. Os primeiros produtos de IA foram chamados de “motores de regras”. Na prática, os algoritmos mais não eram que transposições (codificação) de regras de comportamento/conhecimento que os especialistas disponibilizavam. Foi assim que foram programados os algoritmos que jogam xadrez e ganham aos melhores especialistas humanos<sup>4</sup>. Nestes casos, os algoritmos foram programados de modo a transmitir-lhes as regras e todas as jogadas possíveis a partir de cada situação de jogo.

A evolução seguinte da IA foram as chamadas “redes neuronais”, onde apenas são transmitidas aos algoritmos as regras de funcionamento. As redes neuronais possuem a “inteligência” para, com base nas regras do sistema, “pensarem os movimentos”, avaliarem possíveis respostas e decidirem para cada situação, as jogadas seguintes.

Passada a primeira fase – motores de regras e redes neuronais –, a IA entrou na fase de tratamento massivo de dados como fulcro da aprendizagem. Assim, e em função do nível de “profundidade tecnológica”, existem quatro tipologias de soluções que começaram a integrar produtos concretos<sup>5</sup>:

- IA como motor de recomendações:

prende-se com a capacidade de os algoritmos apreenderem com as nossas preferências e apresentarem conteúdos escolhidos em função dessas preferências. Existem implementações de sistemas que geram “fakenews” ou mesmo “deepfakes” (manipulação de imagens), que deturpam notícias, denigrem personalidades e destroem imagens de pessoas e empresas;

- IA como algoritmo funcional: prende-se com a utilização de algoritmos no contexto empresarial. São exemplos os “brokers de alta frequência” que analisam os comportamentos da bolsa e desencadeiam de forma automática ordens de compra ou venda em funções de parâmetros predefinidos;

- IA com capacidade de interação: prende-se com a capacidade de os sistemas reconhecerem padrões – fotografias, rostos humanos ou sons. Esta tecnologia começa a fazer a ponte entre o mundo real e o mundo digital. A identificação de pessoas e o processamento da voz fazem parte deste tipo de funcionalidades. Uma evolução particularmente relevante prende-se com a melhoria da capacidade de visão e as chamadas redes neuronais convolucionais (RNC);

- IA global: reside na integração das vagas tecnológicas anteriores, ou seja, na produção de sistemas que olham, veem, “sentem” e interagem com o mundo físico. Os primeiros sistemas integram robôs, mas com capacidades ainda muito limitadas.

Um campo particular de investigação está centrado na capacidade de tratamento da linguagem natural assim como de imagens. Neste contexto,

os algoritmos têm de ter a capacidade de “memorizar” passos anteriores. Esta função algorítmica é chamada de “Long-Short Term Memory” – LSTM e baseia-se numa função que pondera a cada passo o dado de entrada e o resultado do passo anterior. Esta função de realimentação permite as chamadas RNN – Recurrent Neuronal Network e os “transformers”. Ambos são modelos de aprendizagem avançada e baseiam-se em algoritmos matemáticos que, como sucintamente mencionado, realimentam funções com resultados de iterações anteriores. Estas funções foram, no essencial, desenvolvidas pela Google, em 2017, no contexto do projeto “Google Brain”<sup>6</sup>. Uma das aplicações desta tecnologia mais em “moda” é o ChatGPT que responde a questões abertas e consegue mesmo produzir um texto sobre quase qualquer tema.

A forma simplificada como fizemos a “história” da IA não pode transmitir uma ideia de facilidade. Muito investimento e muitas horas de investigação permitem o estado atual da tecnologia no reconhecimento da caligrafia, no reconhecimento e legendagem de imagens, no reconhecimento da fala ou no reconhecimento facial, para apenas referir alguns domínios de investigação. Tecnologias tão complexas como a aprendizagem profunda por reforço (DRL), a otimização bayesiana de hiperparâmetros, as grelhas LSTM (memórias de curto e longo prazo), as redes de memória, os autocodificadores variacionais, a representação vetorial da palavra nas frases, as redes generativas adversas, os modelos generativos baseados na atenção, para além de abordagens diversas da pro-

gramação probabilística, são alguns dos domínios de investigação que constroem os sistemas inteligentes do futuro<sup>7</sup>.

Uma das aplicações dos sistemas de reconhecimento facial é utilizada pelas camaras e sistemas de segurança, quer aeroportuárias quer urbanas. Todas as grandes metrópoles possuem sistemas sofisticados suportados em algoritmos de reconhecimento facial que cruzam, em tempo real, informação de qualquer eventual “suspeito” com toda a informação relevante que possa existir. Como qualquer tecnologia de utilização dual, pode ser utilizada para benefício dos cidadãos, aumentando a segurança pública, ou para funções políticas, como o que acontece na China no controle dos Uigures<sup>8</sup>.

**“No domínio militar ninguém quer dar o primeiro passo nos sistemas 100% autónomos, mas a IA domina hoje cerca de 80% das transações financeiras de Wall Street(...)”**

Na aplicação ao domínio dos armamentos, importa separar duas realidades. Por um lado, as armas assistidas e suportadas em algoritmos de IA, o que lhes confere maior precisão, mais eficiência, maior capacidade para evitar sistemas de contramedidas e que com intervenção ou sem intervenção humana direta, desempenham missões com mais precisão e eficácia; por

outro, as armas equipadas com sistemas de IA completamente autônomas da supervisão e decisão humana, ou seja, com capacidade autônoma de decisão. Nos sistemas autônomos, o suporte à decisão está detalhadamente especificado, mas, em última análise, o julgamento tem sempre de depender do contexto específico, o que é muito difícil de expressar numa função matemática sempre que duas situações críticas se apresentam em alternativa. Um algoritmo só conhece as instruções que foram codificadas, não possuindo funções que permitam expressar considerações éticas ou dúvidas morais<sup>9</sup>.

Em 2020, a “IA resolveu um enigma da biologia – dobramento das proteínas – com 50 anos, ultrapassou humanos no reconhecimento de discursos e objetos, simulou humanos na fala, teve melhor aproveitamento em provas académicas, foi mais justo na aplicação da justiça, mais preciso no diagnóstico médico e mais eficiente a gerir ‘drones’ usados na agricultura, em expedições a lugares impossíveis para humanos, e em tecnologia militar”<sup>10</sup>.

Todos os sistemas automáticos geram informação, que pode ser apreendida por algoritmos que assim podem realimentar e otimizar o seu funcionamento. Os sistemas industriais e todos os sistemas técnicos, quer integrem uma cidade, uma fábrica, ou sistemas críticos, como os nucleares, podem ser objeto de maior ou menor automatização, mas esse grau de implementação, na sociedade civil ou no militar, apenas depende do nível de confiança, porque a sofisticação e a emulação continuarão a evoluir.

O controle de cidadãos, a partir de sistemas de reconhecimento facial, que integram dados com sistemas de pagamento, redes sociais e sistemas de mailing, constroem uma sociedade sem privacidade ou liberdade. Sistemas de controlo de armas associadas a supervisão por satélite que identificam movimentações e equipamentos e guiam sistemas autônomos, podem mudar o curso da guerra e os conceitos de segurança. Sabotagem informática de sistemas vitais, como energia, saúde ou telecomunicações, podem bloquear um país ou uma região. Autômatos que analisam exames médicos, realizam atos cirúrgicos, administram a lei, ensinam e realizam funções administrativas e sociais, podem alterar profundamente a estrutura do emprego e da sociedade.

De acordo com a teoria da singularidade e dado o crescimento exponencial da computação, a IA autodirigida poderá também crescer exponencialmente e ganhar superinteligência<sup>11</sup>. Estamos longe de qualquer destes objetivos pese embora as aplicações, em cada um dos domínios, se sucederem a ritmo crescente. Porém, subsistem problemas de dimensão e qualidade dos dados para, nomeadamente, suportar a aprendizagem automática necessária para a análise de exames médicos ou o reconhecimento facial, o que provoca erros e viés de análise.

No domínio militar ninguém quer dar o primeiro passo nos sistemas 100% autônomos, mas a IA domina hoje cerca de 80% das transações financeiras de Wall Street (a partir de autômatos cada dia mais sofisticados e que pesquisam automaticamente informa-

ções sobre empresas, pessoas e negócios), Londres tem mais câmaras de vigilância do que Beijing e a Huawei vendeu sistemas de reconhecimento facial a mais de cinquenta países e quer a China quer os EUA trabalham em sistemas que aumentam a capacidade de identificação de objetos e pessoas quer a partir de satélites quer de “drones” por forma a tornar os veículos e sistemas autônomos mais poderosos e eficientes<sup>12</sup>.

Uma máquina superinteligente superará o ser humano mais dotado e terá capacidade para criar máquinas ainda mais inteligentes. Talvez possamos imaginar que esse sistema venha a possuir uma unidade de comando externa (sistema de segurança) que proteja a Humanidade da sua capacidade de decisão (I. J. Good)<sup>13</sup>. Mas a conjugação entre estes diferentes campos de aplicação, permite antever um poder tecnológico que poderá bloquear qualquer modelo de sociedade (em particular as democráticas) e criar, por si, a capacidade autônoma para decidir<sup>14</sup>.

## Referências

- <sup>1</sup> Oliveira, Arlindo. (2019). Inteligência Artificial. Fundação Manuel dos Santos.
- <sup>2</sup> Damásio, António. (2020). Sentir e Saber. Temas e Debates.
- <sup>3</sup> Lee, Kai-Fu. (2023). Inteligência Artificial 2041. Relógio D'Água.
- <sup>4</sup> O'Reilly, Tim. (2017). Como será o Futuro e porque depende de Nós. D. Quixote.
- <sup>5</sup> Lee, Kai-Fu. (2018). As Superpotencias da Inteligência Artificial. Relógio de Água.
- <sup>6</sup> Kaku, Michio. (1997). Visões. Bizancio.
- <sup>7</sup> Bostrom, Nick. (2014). Superinteligência. Relógio D'Água
- <sup>8</sup> Ford, Martin. (2022). O Futuro da Inteligência Artificial. Bertrand.
- <sup>9</sup> Bostrom, Nick. (2014).
- <sup>10</sup> Lee, Kai-Fu. (2023).
- <sup>11</sup> Idem.
- <sup>12</sup> Ford, Martin. (2022).
- <sup>13</sup> Citado por Martin Ford em O Futuro da Inteligência Artificial, p.295.
- <sup>14</sup> Ford, Martin. (2022).