

A INTELIGÊNCIA ARTIFICIAL E PROCESSOS DE DECISÃO¹

José Júlio Alferes²

Foi-me pedido pelos organizadores deste encontro para falar sobre “Inteligência Artificial” e seu uso em sistemas de decisão e apoio à decisão, com vista a iniciar um debate sobre novos problemas éticos que tal uso pode levantar.

Embora o convite muito me tenha honrado, a tarefa não se adivinhava nada fácil. A começar pelo facto de, apesar de se centrar na Inteligência Artificial, o meu trabalho de investigação não estar directamente ligado às suas aplicações a processos de decisão. Com efeito, a minha actividade de investigação recai essencialmente no estudo e concepção de linguagens de programação e representação de conhecimento apropriadas para a Inteligência Artificial. Por outras palavras, o meu trabalho visa construir “ferramentas” que permitem depois construir programas para, entre outras coisas, sistemas de decisão e apoio à decisão. E não, a construção desses programas.

Assim sendo, optei nesta intervenção por vos falar um pouco sobre o que é Inteligência Artificial, e o que se tem conseguido fazer usando-a para a construção de sistemas automáticos de decisão e apoio à decisão, em especial no campo médico, deixando a porta aberta para debatermos depois se, e que, novos problemas éticos estes sistemas podem levantar.

1 A Inteligência Artificial

Como disciplina científica, a Inteligência Artificial estuda os princípios do comportamento inteligente. Pretende compreender as entidades inteligentes mas, ao contrário da Psicologia ou da Filosofia que também se interessam pela “inteligência”, visa também o *design* e construção de sistemas eficientes que exibam comportamentos inteligentes.

Neste último aspecto é uma “ciência do artificial”, semelhante a uma engenharia: nas abordagens que faz à “inteligência” tem um cariz bastante pragmático, preocupa-se com compromissos necessários de custo/benefício e ocupa-se com a totalidade dos aspectos relevantes à construção dos sistemas “inteligentes”. Tal como uma engenharia.

Embora não seja consensual qual a data do aparecimento da Inteligência Artificial como disciplina científica, é para mim claro que os objectivos e problemas hoje a ela associados apareceram quase simultaneamente com o aparecimento dos computadores modernos, no início dos anos 50.

Os computadores são máquinas que processam símbolos de forma automatizada e eficiente. Mas, talvez mais importante, fazem-no com grande generalidade. Para isso contribui o facto de haver uma distinção clara entre a forma de processar os símbolos (programa, ou *software*) e a máquina que suporta esse processamento (*hardware*). É essa distinção que permite que um mesmo computador possa servir para levar a cabo tarefas tão diferentes. Ora, assumindo que na actividade de raciocínio se processam símbolos e que pode

¹ Reunião do Núcleo de Évora do Centro de Estudos de Bioética, Évora 10/99

² Professor de Informática da Universidade de Évora

haver independência entre o processamento desses símbolos e as operações orgânicas que o suportam, porque não usar estas novas máquinas genéricas (computadores) para processar símbolos por forma a exibirem comportamentos inteligentes? Foi esta questão que se colocou logo desde cedo, e que deu origem ao aparecimento da Inteligência Artificial.

Mas afinal o que é que a Inteligência Artificial trouxe de substancialmente novo, relativamente às outras áreas do saber que antes já se preocupavam com o estudo da inteligência?

Em primeiro lugar, possibilita a construção, em computador, de modelos funcionais de inteligência. Modelos esses com comportamento observável e repetível. Permite-se assim, à semelhança das outras ciências, a formulação de hipóteses e o teste dessas hipóteses de forma objectiva (experimentação). Com isto, podem-se experimentar novas formas de raciocínio e inteligência.

Além disso, a Inteligência Artificial ambiciona criar novas entidades inteligentes. Embora seja por agora apenas uma ambição, não sendo ainda possível a criação completa de tais entidades, os passos dados nesse sentido ameaçam levar à dissociação entre a noção de inteligência e o Homem. Pretende-se criar uma inteligência fora do Homem e diferente da inteligência humana. Pretendem-se automatizar, em máquinas, faculdades mentais que antes eram exclusivo da espécie humana.

Embora em Inteligência Artificial se construam modelos de comportamento inteligente, não é sua preocupação imitar o comportamento inteligente dos humanos. Aspectos como comportamentos neuróticos, patológicos e erróneos, necessariamente ligados ao estudo do comportamento inteligente de humanos, não são objecto de estudo na Inteligência Artificial. Muito menos se preocupa em imitar os processos internos que levam os seres humanos à exibição de comportamentos inteligentes. Se os sistemas construídos exibirem o comportamento inteligente desejado, pouco importa saber se o fazem da mesma forma que os seres humanos. A este respeito relembro o que há pouco disse sobre o carácter engenheirístico/pragmático da Inteligência Artificial. Obviamente, sendo desenvolvida por humanos, inspira-se certamente na “inteligência humana”. Mas, repito, sem se ver obrigada a imitar.

Apesar dos objectivos da Inteligência Artificial serem tão ambiciosos como o de criar entidades inteligentes, não vislumbro que possam vir a ser atingidos na totalidade proximamente, se é que alguma vez o serão. Afinal a inteligência é algo de bastante complexo! Não será tão cedo que se verão máquinas com potencialidades de inteligência genéricas e abrangentes, semelhantes às de um ser humano.

O que sobra então, agora, para a Inteligência Artificial? Que “formas de inteligência” é que já se conseguem em máquinas? Essencialmente constroem-se sistemas que necessitam de formas de inteligência bastante específica, e nos quais a inteligência possa resultar de um compromisso entre o que é necessário e o que é possível obter de forma eficiente. Casos em que tal é já hoje possível são: sistemas de aprendizagem de padrões e de regras de formação de padrões a partir de grandes quantidades de dados que permitam, por exemplo, descobrir padrões de compras feitas em grandes superfícies comerciais, ou padrões de comportamento de mercados bolsistas a partir dos dados relativos a todas as transações, ou ainda descobrir regras de formação de compostos químicos a partir de grandes quantidades de informação relativas a experiências feitas³; sistemas de apoio à decisão em domínios específicos como,

³ A este respeito, refira-se um sucesso recente desta área da Inteligência Artificial, ao conseguir, exclusivamente de forma automática, obter resultados novos e importantes acerca de predição de

por exemplo, diagnóstico médico (destes falarei mais detalhadamente daqui a pouco), atribuição de crédito em instituições bancárias, etc; sistemas de reconhecimento de voz e imagem; *interfaces* homem-máquina em língua natural; sistemas para planeamento de tarefas; sistemas de jogos, capazes de vencer ao campeão mundial de xadrez; sistemas de diagnóstico de máquinas ou redes de máquinas complexas; sistemas de representação de conhecimento relativo a assuntos específicos, que permitem por exemplo guardar as nossas preferências relativamente a música, livros, ou programas de televisão e, tendo-as em conta, aconselhar-nos sobre que programas e canais de televisão ver em cada momento, e sugerir-nos ou alertar-nos para livros ou discos que possamos estar interessados em adquirir.

Este último é um bom exemplo para explicar o que quero dizer, neste âmbito, com compromissos custo/benefício. O que é que cada um de vós preferia? Um sistema que, conhecendo as vossas preferências, rapidamente vos sugerisse uma sequência de programas de televisão a ver, sequência essa que na maior parte das vezes fosse do vosso agrado, embora por vezes se “enganasse”. Ou um outro, que acertasse sempre nas vossas preferências, mas que demorasse vários anos a dar uma resposta. Certamente o primeiro. O que aqui está em causa é um compromisso entre o custo (em tempo) associado a uma solução, e a qualidade dessa solução.

Apesar de já hoje se conseguir que máquinas façam raciocínios considerados tão complexos como os associados à elaboração de um diagnóstico médico, está longe de se conseguir que estas façam coisas que qualquer um de nós classificaria como banais ou triviais. Tudo o que tenha a ver com o chamado “conhecimento ou raciocínio de senso comum”, como por exemplo o entendimento de uma banal conversa de café, está longe de se conseguir implementar em computador. Tal como tudo o que tenha a ver, por exemplo, com “imaginação” ou “sensibilidade”.

Um computador consegue lembrar-se sempre dos mais pequenos detalhes, e fá-lo de forma bastante eficiente. Consegue fazer cálculos matemáticos complexos e lidar com muitos dados simultaneamente, certamente melhor que qualquer ser humano. Por outro, lado está longe de ter as capacidades de processamento paralelo existentes no cérebro humano, ou de conseguir lidar, de forma eficiente, com informação incompleta e de formular hipóteses. É isto, em boa parte, que justifica a sua aptidão para formas de inteligência muito específica e que envolva bastante cálculo, e a sua inaptidão para modelizar uma inteligência mais genérica e abrangente como a “associada ao conhecimento de senso comum”.

Assim sendo, a Inteligência Artificial não deverá ser vista como algo que vem para substituir os humanos nas suas actividades inteligentes. Deverá antes ser encarada como um instrumento que permite potenciar a forma de pensar humana com as potencialidades que a máquina lhe acrescenta. Da mesma forma que hoje não passa pela cabeça de um engenheiro, na elaboração dum projecto, fazer todos os cálculos de forma manual, e o leva necessariamente a usar uma máquina para os fazer, no futuro não passará pela cabeça de um decisor tomar uma decisão sem recorrer a uma máquina que o ajude, desenvolvendo raciocínios detalhados sobre a questão em causa. Haverá, na minha opinião, uma simbiose homem-máquina que permita ao homem produzir “melhores raciocínios” e de forma mais eficiente.

2 Aplicações de Inteligência Artificial em Medicina

Existem actualmente várias aplicações da Inteligência Artificial na área da Medicina. Entre as mais conhecidas, e talvez de maior interesse para esta reunião, estão os sistemas de apoio ao diagnóstico médico.

Tratam-se de sistemas periciais de regras⁴ que, dada toda a informação relevante sobre o paciente, conseguem elaborar um diagnóstico médico. Além disso, este tipo de aplicações pode ainda fornecer explicações pormenorizadas sobre a forma como chegou ao diagnóstico.

Embora seja possível construir sistemas destes que exibem uma grande fiabilidade nos diagnósticos que produzem, na prática a sua utilização generalizada defronta-se com uma série de problemas. Logo a começar pela forma de passar à máquina “toda a informação relevante sobre o paciente”. Um médico experiente consegue, rapidamente, recolher uma série de dados relevantes para a formulação do diagnóstico. Consegue-o olhando para o doente, conversando com ele. Para que a máquina possa elaborar o diagnóstico, o médico terá que lhe passar toda essa informação relevante que conseguiu recolher. E esse processo, que em muito depende do *interface* homem-máquina, é nos actuais sistemas um processo moroso e por vezes enfadonho. De tal forma que é lícito perguntar porque é que um médico se há-de socorrer de uma máquina para a elaboração de um diagnóstico quando, por vezes, só actividade de lhe fornecer todos os dados relevantes demora mais tempo do que a elaboração pelo médico do diagnóstico? Por forma a tornar estes sistemas aplicáveis em geral, muito há ainda a fazer, não tanto em Inteligência Artificial, mas, essencialmente, na área de desenvolvimento de bons *interfaces* entre o homem e a máquina.

No entanto, em certas áreas específicas da Medicina, e para certo tipo de diagnósticos específicos, este problema não se coloca com tanta acuidade. Em particular, nos casos em que grande parte dos dados relevantes para a formulação do diagnóstico são já obtidos de forma automática, por máquinas. Por exemplo, em cardiologia, na formulação de diagnósticos com base em exames de electrocardiograma. Nestas áreas, existem já vários sistemas a funcionar na prática, tendo-se vindo a obter bons resultados no que respeita à fiabilidade dos diagnósticos assim produzidos (da ordem dos 90%, comparando com os diagnósticos elaborados pelos melhores especialistas).

Outro dos problemas com que se defronta a utilização destes sistemas automáticos de diagnóstico é o da confiança dos médicos nos diagnósticos produzidos. Logo a começar pela confiança dos médicos nos peritos que forneceram as regras ao sistema e, consequentemente, no sistema de regras obtido. Poder-se-ia pensar que esse problema seria minimizado pelo facto dos sistemas fornecerem explicações sobre a forma como chegaram ao diagnóstico. Mas não se revela prático, no dia-a-dia, estar a recorrer a essas explicações pormenorizadas para, antes de aceitar ou não o diagnóstico fornecido, o médico avaliar da qualidade do mesmo. Mais uma vez, porque é que um médico se há-de socorrer de uma máquina para a elaboração de um diagnóstico quando, por vezes, a actividade de avaliação do diagnóstico obtido demora mais tempo do que a elaboração pelo médico do diagnóstico? De notar que isto não significa que não seja útil o fornecimento de explicações por parte destes sistemas. Este tipo de sistemas com explicações pode revelar-se bastante útil para uso no ensino.

⁴ I.e. sistemas constituídos por uma série de regras e que simulam o comportamento de um perito na matéria em causa

Uma outra aplicação de técnicas de Inteligência Artificial em Medicina, recente e provavelmente menos polémica, é a de automatização de protocolos clínicos. Como saberão certamente melhor que eu, estes protocolos resultam de consensos pelo que, ao contrário do que acontecia com os conjuntos de regras dos sistemas de diagnóstico, têm uma maior aceitação na comunidade. Embora não produzam diagnósticos, os sistemas automáticos criados a partir desses protocolos poderão ser um bom instrumento de ajuda ao médico, guiando, monitorando e alertando-o no decurso do processo de diagnóstico, prognóstico e prescrição. A existência destes sistemas permite pedir ao médico os dados adequados no momento certo e, recolhidos os dados, estruturar a informação clínica num formato que possibilite a sua consulta posterior, de forma automatizada. Esta última é uma vantagem que, só por si, justificaria a existência e uso deste tipo de sistemas.

À primeira vista, a automatização destes protocolos seria um problema de Informática tradicional, que não necessitaria de nada de especial vindo da Inteligência Artificial. Trata-se simplesmente da automatização de processos com recolha e estruturação de informação. E, com efeito, assim seria não fosse o facto da descrição desses processos ser por vezes vaga e aparentemente ambigua, de por vezes se basear em informação incompleta, de necessitar de formular hipóteses, hipóteses essas que agora se assumem e mais tarde se podem vir a revelar falsas. Tudo questões com as quais a Informática tradicional não costuma lidar. Tudo problemas típicos de qualquer sistema de Inteligência Artificial. Daí que, na construção destes sistemas de automatização de protocolos clínicos, se utilizem várias técnicas desenvolvidas na Inteligência Artificial.

Em todos os casos, quer os sistemas de diagnóstico quer os de automatização de protocolos clínicos apenas sugerem um diagnóstico ou procedimento, cabendo sempre ao médico a decisão de o aceitar ou não. Tratam-se de sistemas de apoio à decisão e não de sistemas de decisão. Esta limitação não se deve a questões técnicas da Inteligência Artificial. Deve-se apenas a questões éticas. Mas deverá ser sempre assim? Deixar-se-á sempre ao médico a decisão e a aceitação do diagnóstico? Ou haverá situações em que se possa “abdicar” de tal?

Para ilustrar melhor o problema que levanto com estas minhas questões, tomemos como exemplo o sistema *TraumAID*⁵, que tem vindo a ser desenvolvido e aplicado desde 1984 nos Estados Unidos, na “University of Pennsylvania” e na “Allegheny University of the Health Sciences”.

O *TraumAID* é um sistema concebido para assistir os médicos no diagnóstico e tratamento de urgências por traumas resultantes de tiros (ou outros acidentes com perfuração) no peito ou abdomen. Nestes situações de emergência, é importante saber quando é mais apropriado prosseguir com o diagnóstico, e quanto é mais apropriado iniciar a terapia. Por vezes o médico tem de iniciar a terapia sem antes completar o diagnóstico sobre a extensão das lesões causadas. Outras vezes, acções terapêuticas podem ajudar no diagnóstico do paciente. O sistema *TraumAID*, recolhendo toda a informação conhecida, vai propondo, em cada momento, acções (de diagnóstico ou terapia), produzindo assim um plano para tratamento do paciente. A escolha da acção a desencadear em cada momento pode ser baseada no tempo que esta demora a ser efectuado, localização (por vezes, certas acções só podem ser levadas a cabo em locais específicos, e.g. blocos operatorios), custo e outros factores.

Acontece que o uso deste programa tem vindo a mostrar que os planos por ele concebidos são em regra melhores do que aqueles produzidos pelos médicos nas mesmas

⁵ Mais informação sobre este sistema poderá ser obtida a partir de:

<http://www.cis.upenn.edu/~traumaid/>

situações⁶. Ora, tratando-se de situações de emergência, bem específicas, em que é importante o tempo de resposta e a rapidez do raciocínio, e havendo sistemas automáticos que conseguem planejar o que fazer, melhor e mais rapidamente que os médicos, porquê deixar aos médicos essa actividade de planeamento, como se faz actualmente? Porque não deixar a máquina actuar? Esta é sem dúvida uma questão que, com os avanços na Inteligência Artificial, surgirá cada vez mais e que levantará vários problemas éticos.

⁶ Numa comparação cega em casos reais, três especialistas consideraram os planos propostos pelo *TraumAiD* melhores do que os elaborados pelos médicos, com um rácio de 64:17 (com 16 empates).