

# **O estatuto ético do algoritmo na decisão clínica**

## **Uma reflexão prospetiva**

**Paulo César Pereira Rodrigues de Roboredo**

Dissertação para obtenção do Grau de Mestre em  
**Gestão de unidades de saúde**  
(2<sup>o</sup> ciclo de estudos)

Orientador: Prof. Doutor António Campelo Amaral

**Junho de 2020**



# **Dedicatória**

A todos quantos tornaram esta dissertação possível, em especial à minha companheira.



# **Agradecimentos**

Ao Professor Doutor António Amaral que me auxiliou na germinação das ideias e durante todo o processo de desenvolvimento deste trabalho nas inúmeras reuniões.



# Índice

Lista de Acrónimos	x
Resumo	xii
<i>Abstract</i>	xiv
<i>Preâmbulo</i>	xvi
Introdução. O algoritmo, a ética e a gestão de unidades de saúde.	1
Parte I :O estado da Arte	4
1- Fundamentação teórica	4
1.1- Abordagem ao tema	4
1.2- Contexto histórico	6
2- Ética e máquinas	8
2.1- Responsabilização e transparência	8
2.2- Humano dispensável	9
2.3- Máquinas éticas	11
3- Limitando as máquinas rebeldes	13
3.1- Leis da robótica	13
3.2- <i>ETHICS GUIDELINES FOR TRUSTWORTHY AI</i>	15
4- Limitações atuais	17
5- Considerações finais da primeira parte	22
Parte II: Ética e Moral	23
1- A necessidade da dilucidação dos termos	23
1.1- Algoritmo	23
1.2- Decidir e escolher	24

1.3- Diferenciação entre ética e moral, contexto e evolução	26
2- A ética aplicada ao contexto clínico	30
2.1- Contexto histórico e o ping-pong entre o paternalismo e autonomia	31
2.2- Mais além da ética, a bioética	33
2.3- Os três níveis de juízo médico	35
2.4- A clínica e a ética	37
Parte III: Deliberar e decidir	39
1- Conceito de deliberação	39
2- Pode uma máquina deliberar?	39
3- Dissecação da deliberação	41
3.1- Dilema e clínica, contexto histórico	41
3.2- Esquematização/perspetivas da deliberação	43
3.3- Robotização e deliberação	46
4- A deliberação como ponte para a decisão razoável	49
4.1- A ética da demora	50
4.2- Paradoxo da ponderação em contexto crítico de aceleração temporal	51
4.3- No tempo da ética aristotélica à ética da compressão temporal acelerada	55
5- A máquina como integrante da ética clínica	56
Conclusão. Decisão do algoritmo e algoritmo na decisão. Em que ficamos?	57
1- A máquina como parte da decisão ética	58
2- A máquina no cerne da decisão ética	58
Bibliografia	61



# Lista de Acrónimos

AAA	<i>American Automobile Association</i>
AI	<i>Artificial intelligence</i>
ADAS	<i>Advanced driver-assistance systems</i>
ADN	<i>Ácido desoxirribonucleico</i>
IA	<i>Inteligência Artificial</i>
IAS	<i>Inteligências Artificiais</i>
BSOD	<i>Blue screen of death</i>
<i>Captcha</i>	<i>Completely Automated Public Turing test to tell Computers and Humans Apart</i>
EPIC	<i>Electronic Privacy Information Center</i>
PDCA	<i>Plan-Do-Check-Act</i>
ROM	<i>Read only memory</i>



# Resumo

O processo de decisão está inerente à prática médica e é a base do diagnóstico e tratamento, as questões de moralidade dos atos médicos surgem logo desde o aparecimento da medicina. As decisões entre as várias opções são muitas vezes fontes de dilemas morais ou éticos e para a resolução destes, a bioética veio contribuir de forma ímpar para definir os trâmites daquilo que se pode hoje considerar como uma decisão mais correta.

Na equipa de saúde esta função decisória está na maioria das vezes a cargo do médico, cuja formação e experiência o sobrecarrega com o ónus de coordenação do processo de tratamento e intervenção, porém este é tanto mais difícil quanto maior é a urgência. Nestas condições o tempo disponível para contemplar as opções é escasso e torna-se imperiosa uma escolha, uma decisão mais ou menos ponderada.

A gestão na área da saúde está sempre rodeada de um contexto de incertezas, em que se discutem reais benefícios e acesso às novas tecnologias em saúde. Assim o contexto de decisão no ato individual tem implicações maiores em termos de gestão de saúde. O fazer ou não fazer, a decisão de internar ou dar alta, de tratar ou paliar, reanimar ou não reanimar deveriam sempre assentar num pressuposto que qualquer deliberação em (bio) ética clínica resulta do princípio de que o respeito pelo ser humano é indispensável para o agir e decidir corretamente.

Apesar de a ética no geral estar a descartar-se da individualidade para um pensamento mais sistémico, no que concerne à parte da ética organizacional esta está ainda muito virada para a culpabilização individual. Esta dissertação de mestrado compromete-se a apresentar uma reflexão teórica sobre ponderações ou abordagens correlacionadas ao processo decisório em (bio) ética clínica e sobre o papel das máquinas ou algoritmos de decisão automatizada. Desta reflexão teórica poderão ser depois assentadas bases para uma mais ética gestão em saúde.

## Palavras-chave

**Ética; algoritmo; decisão crítica; inteligência artificial.**



# Abstract

The decision-making process is inherent to medical practice and is the basis of diagnosis and treatment, questions of morality of medical acts arise immediately since the appearance of medicine. The decisions between the various options are often sources of moral or ethical dilemmas and for the resolution of these, bioethics has come to contribute in a unique way to define the procedures of what can be considered today as a more correct decision.

In the health team, this decision-making function is mostly in the responsibility of the doctor, whose training and experience overload him with the burden of coordinating the treatment and intervention process, but this is all the more difficult the greater the urgency. In these conditions, the time available to contemplate the options is scarce and a choice, a more or less considered decision, is imperative.

Health management is always surrounded by a context of uncertainty, in which real benefits and access to new health technologies are discussed. Thus, the decision context in the individual act has major implications in terms of health management. Doing or not doing, the decision to hospitalize or discharge, to treat or palliate, revive or let go should always be based on the assumption that any deliberation in clinical (bio) ethics results from the principle that respect for human beings is indispensable to act and decide correctly.

Although ethics in general is parting from individuality to a more systemic thinking, in regard to the part of organizational ethics this is still very much turned to individual blame. This master's project undertakes to present a theoretical reflection on considerations or approaches related to the decision process in clinical (bio) ethics and on the role of machines or automated decision algorithms. From this theoretical reflection, bases can be laid for a more ethical health management.

# Keywords

**Ethics;algorithm;critical decision;artificial intelligence.**



# ***Preâmbulo***

## **Um diálogo prospetivo**

Como preâmbulo da realização da tese de mestrado e no contexto da discussão prévia com o orientador, achei premente deixar um ponto de situação do conhecimento/opinião da temática antes do aprofundamento da dissertação para elaborar posteriormente um termo comparativo após a pesquisa.

### **Poderá um algoritmo alguma vez ser responsável pela decisão ética num caso limítrofe/crítico?**

Talvez. Um algoritmo remete para a ideia de uma sequência de procedimentos para o atingir de um fim. Uma operação realizada por uma máquina de maneira precisa, mecânica com um resultado correto. Será uma instrução de transformação de um dado *input* para um *output*. A existência destas restrições quanto à qualidade de *inputs* que pode processar representa como que paredes que delimitam o âmbito de atuação e tornarão a máquina incapaz de atuar quando se transbordam estes limites, portanto com fraca adaptabilidade a diferentes situações.

Uma máquina de calcular simples devolve um “*error*” se em vez de um número usássemos uma letra ou se o resultado da conta for maior que os espaços disponíveis para o apresentar. No entanto dentro dos cálculos mais simples, ou seja aqueles para os quais foi programada, todas as respostas fornecidas pela máquina são corretas e fiáveis uma vez que são mecânicas.

### **Poderá esta calculadora resolver um problema matemático simples por si só?**

Obviamente não. A modo de exemplo, um problema matemático escrito numa folha de papel jamais será resolvido pela calculadora sendo essencial a existência de um “intérprete” humano que atue de transdutor transformando a informação visual que lê, em parâmetros percebidos pela máquina e depois novamente na tradução das respostas da máquina para o papel.

### **Atualmente já existem algoritmos e IA (inteligência artificial) capazes de resolver o problema matemático simples descrito numa folha de papel?**

Já. Uma pergunta simples formulada numa folha de papel poderá já ser decifrada por um programa de IA que traduza automaticamente a linguagem escrita de modo a poderem ser processados os cálculos e ser debitada uma resposta por um outro algoritmo, até mesmo em papel, recorrendo então a um algoritmo adicional. Isto já é possível descurando claro, a imensa falibilidade ainda existente na tradução de linguagem escrita por parte das máquinas, imensamente pior que a tradução por parte de um humano pela ausência de contexto (embora estas diferenças se esbatam por exemplo se for um humano a traduzir a linguagem escrita num idioma que não conhece e os resultados seriam mais similares).

### **E se o problema matemático fosse apresentado sob a forma de imagem? Já há IA para o resolver?**

Já também, relativamente. Aqui reside todo o dilema atual do *takeover* preconizado pelas inteligências artificiais (IAs). O incremento de computação das últimas décadas permite que haja algum grau de interpretação de imagens, o que deu aso a novidades como o reconhecimento facial ou mesmo a condução autónoma, recorrendo a uma variedade imensa de algoritmos diferentes.

Estas novidades não aparentam tanto uma perspectiva de futuro mas de presente, estão já a ser realizadas multitudes de experiências que demonstram como a autonomia e automatização são realidades mas ainda de uma forma cruda pouco acessível ao público em geral com verdadeira capacidade de independência do humano. Mesmo no veículo automático ainda é o humano que necessita introduzir o destino, decidir sobre paragens, cruzamentos e rotundas ou tratar da manutenção do veículo sendo esta multitude de algoritmos destinada a funcionar em circuito fechado limitando-se apenas à tarefa de condução, da qual há ainda sérias dúvidas na possibilidade de excluir desta a intervenção humana direta (veículos sem pedais ou volante). De notar que a discrepância entre a interpretação humana e a de IA das mesmas imagens, é ainda o meio mais comumente usado para discernir entre humano e máquina e tal é usado por exemplo nos “*captchas*” ou “*Completely Automated Public Turing test to tell Computers and Humans Apart*”, portanto a IA não “vê” como um humano.

Regressando ao assunto, podemos dizer que uma série de algoritmos alinhados poderia realizar um problema matemático submetido sob a forma de uma imagem com contexto e seguidamente elaborar uma resposta de forma mais ou menos satisfatória com ajuda de intervenção humana para corrigir alguns erros.

## **Aumentando a complexidade, poderia um algoritmo decidir sobre uma decisão limite crítica num paciente de forma ética?**

Obviamente que não. Não existe hipótese qualquer de um algoritmo tomar este tipo de decisão de contextos complexos por si só. Como vimos anteriormente, à medida que se aumenta a complexidade do problema aumentam as necessidades computacionais de forma exponencial e num contexto de vida real estas necessidades tenderiam ao infinito pois o algoritmo não admite situações não previstas. Impossível portanto. O funcionamento algorítmico funciona em circuito fechado como já vimos anteriormente, sem ambiguidade, e jamais conseguirão manter-se em circuito fechado todas as variáveis que possam entrar numa decisão muito complexa.

Ao surgir uma situação inesperada o algoritmo debitará simplesmente aquilo que faz atualmente um computador quando algo inesperado acontece, sai do programa que estava a correr e no ecrã aparece um erro de “*divide by zero*”, havendo que reiniciar o sistema ou o programa. Em modo “*debug*” o algoritmo mais avançado pergunta ao humano como proceder quando o inesperado acontece.

## **Mas poderíamos contornar estas necessidades de computação infinita? E porque não prescindir da necessidade de colocar todos os dados de todas as situações possíveis no algoritmo e simplesmente agrupássemos as diversas situações em grupos *quasi-homogéneos*? Será que não poderíamos fazer a abreviação das decisões, a discriminação de casos conforme padrões?**

Claro, e assim conseguiríamos manter o circuito fechado e de certa maneira impedir o imprevisível, com recurso a um número limitado de grupos poderíamos mais facilmente abarcar todos os casos possíveis. Mas então surgiria uma outra variável com a qual as máquinas não conseguem lidar, a possibilidade de erro.

Agrupou-se um caso “excepcional” num grupo, no qual cumpria critérios para estar e a decisão foi errada, surgem então implicações do tipo de responsabilidade que jamais se poderiam atribuir atualmente a uma máquina (o algoritmo tomou a decisão errada, que fazemos? Cadeia com ele? Na lógica da máquina ela procedeu corretamente, para o que estava programada). A responsabilização do algoritmo ou da máquina necessitaria que entrássemos em domínio de máquina consciente, responsável, o que ainda está longe de sequer ser vislumbrado (e pode uma decisão ética partir de algo sem responsabilidade? Decisão ética *vs* programação ética).

A outra hipótese seria o desviar da responsabilidade para a entidade ou para a empresa responsável pela programação da máquina. Neste caso seria dúbia a responsabilidade criminal e a solução seria contar com os erros de decisão, admitir que eles acontecem e preparar-se para indemnizar todos os casos excepcionais. Há aqui ética? Por outro lado existe possibilidade de estarmos a exigir a uma máquina, mais ética do que a que exigiríamos a um humano no mesmo dilema e numa situação de emergência, em que é necessária a atuação imediata, a definição de certo ou errado é mais esbatida. Uma máquina aí poderia talvez ter acesso a mais dados tomando uma decisão que, apesar de não ser formada a partir de uma base ética pode estar mais correta estatisticamente e então o somatório de múltiplas decisões deste tipo por parte das máquinas poderia ser mais ético do que o somatório das decisões tomadas pelos humanos na falta de dados. O desenvolvimento deste texto com a comparação máquina/humano iria indubitavelmente cair num dilema básico da teoria da evolução, a consequência da eliminação da variabilidade que por exemplo podemos resumir na seguinte pergunta: Se todos os carros fossem automatizados salvaríamos vidas na estrada, é um facto, devemos então proibir a condução não automatizada. Isto é ético? Numa outra questão onde talvez resida a solução intermédia podemos perguntar: Um algoritmo terá sempre que ser executado por uma máquina? Porque não um humano? Algoritmo -> *guideline* para situações limite/críticas -> filosofia sobre a tomada de decisões, múltiplos autores.

Portanto, antes da elaboração desta dissertação de mestrado a minha posição sobre os algoritmos na decisão é que estes serão extremamente úteis para automatizar decisões simples, fechadas que proporcionam tempo para deliberar na resolução do “inesperado”. Não há ao momento o vislumbre de ser possível uma tomada de decisão “humana” por parte de uma máquina sem intervenção humana pois estamos sempre a supor a hipótese de uma decisão errada, errar é humano e lidar com o erro talvez só possa ser humano.



# Introdução

## O algoritmo, a ética e a gestão de unidades de saúde

Seja qual for o nível de cuidados de saúde, primário, secundário ou terciário o processo de decisão está inerente à prática médica e é a base do diagnóstico e tratamento, as questões de moralidade ou não dos atos médicos surgem logo desde o aparecimento da medicina. As decisões entre as várias opções são muitas vezes fontes de dilemas morais ou éticos e para a resolução destes, a bioética veio contribuir de forma ímpar para definir os trâmites daquilo que se pode hoje considerar como uma decisão mais correta.

A resolução de um dilema não costuma ter uma resposta direta, tal resulta da própria definição de dilema que implica que os possíveis caminhos ou respostas ao mesmo tenham um valor hierarquicamente similar e de tal resulta o impasse de escolha. A resolução destes impasses requer tempo, ponderação, deliberação para mais profunda análise das várias opções nas suas vertentes mais profundas usando lógica e razão.

Nos pilares da base de atuação e previamente a esta está o conceito de “decisão”, seja qual for a profissão em análise, o poder de decisão é maior quanto mais alta seja a hierarquia num grupo de trabalho. Na equipa de saúde esta função decisória está na maioria das vezes a cargo do médico, cuja formação e experiência o sobrecarrega com o do ónus de coordenação do processo de tratamento e intervenção, porém este é tanto mais difícil quanto maior é a urgência. Nestas condições o tempo disponível para contemplar as opções é escasso e torna-se imperiosa uma escolha, uma decisão mais ou menos ponderada.

*“A ética organizacional foi descrita como o próximo passo na evolução da bioética que se focou primariamente em questões éticas no cuidado direto ao paciente”<sup>1</sup>*

A gestão na área da saúde está sempre rodeada de um contexto de incertezas, em que se discutem reais benefícios e acesso às novas tecnologias em saúde. Assim o contexto de decisão no ato individual tem implicações maiores em termos de gestão de saúde. O fazer ou não fazer, a decisão de internar ou dar alta, de tratar ou paliar, reanimar ou não reanimar deveriam sempre assentar num pressuposto que qualquer deliberação em (bio)ética clínica resulta do princípio de que o respeito pelo ser humano é indispensável

---

<sup>1</sup> Singer, 2008, p. 243

para o agir correto. Referindo-se à ética organizacional Peter Singer diz que a ética é aquela cujas delineações respondem a princípios que considera universais de humanidade, benefício recíproco, confiança, gratidão, serviço, regulação (Singer, 2008, p. 341).

Apesar de a ética no geral estar a descartar-se da individualidade para um pensamento mais sistémico, no que concerne à parte da ética organizacional esta está ainda muito virada para a culpabilização individual.

“Infelizmente a ética clínica não embarcou na tendência de pensamento por sistema (*systems thinking*) nos cuidados de saúde”<sup>2</sup>.

Assim urge trabalhos que possam impulsionar a evolução neste sentido e como tal, o estudo sobre *guidelines*, protocolos ou algoritmos de ação automatizada são um grande passo nesta direção pois a sua ênfase é a atuação e a responsabilidade ao nível de um sistema e não de um indivíduo.

Esta dissertação de mestrado compromete-se a apresentar uma reflexão teórica sobre ponderações ou abordagens correlacionadas ao processo decisório em (bio)ética clínica e sobre o papel das máquinas ou algoritmos de decisão automatizada. Desta reflexão teórica poderão ser depois assentadas bases para uma mais ética gestão em saúde.

Tendo em conta o atual contexto de evolução tecnológica impõe-se a realização de trabalhos deste tipo para avaliar as contingências da substituição da ação humana por decisão maquinizada.

Este trabalho de mestrado pretende servir de base analítica para a tomada de decisão possível e não como um guia normativo, trata-se de uma investigação com propósito descritivo e interpretativo.

O contexto decisionário não se aplica somente à doença, há muitas outras situações com indefinição moral onde existe dilema de decisão em contexto de urgência até porque a definição de doença é ambígua em si pois presume-se sempre uma escassez de informação acerca do sujeito ou condições sobre as quais se vai decidir. Se na eventualidade de um trauma podemos claramente definir um antes e um pós em termos de gravidade, este período de tempo compõe um *degradê* de gravidade no caso de estarmos em presença de uma doença crónica que evolui a brotes, nunca sabendo qual será o ponto de irreversibilidade pois tal necessita uma delimitação com base numa história clínica detalhada e exames atualizados, raramente acessíveis na decisão crítica mas com os quais

---

<sup>2</sup> Singer, 2008, p. 314

esta terá obrigatoriamente de lidar. Nestas situações, a opção de decisão está muitas vezes baseada em regras de decisão pessoal, fundamentos pouco sólidos como a opinião, a fraca experiência do médico ou a intuição e, como em qualquer opinião, estas percepções são alvo de muita variabilidade até dentro do mesmo indivíduo em períodos de tempo diferentes.

Sabendo então desta variabilidade de respostas ante a incerteza que apartam uma estatística de resultados de uma atuação perfeita, podemos socorrer-nos de protocolos, *guidelines* para atuação protocolada em situações específicas, regras ou instruções que ditam como atuar em determinada situação. Ora este modo de atuar pode em parte prescindir do profissional quando se encontra dentro do seu âmbito. Numa altura de rápidos avanços tecnológicos em que somos assoberbados por máquinas que se prontificam a automatizar toda a nossa vida (piloto automático, carro automático, robôs de cozinha, procura de parceiro automatizado com base em perfis...) poderá o ato clínico da decisão crítica também ser alvo de automatização? Provavelmente sim, o presente trabalho disserta sobre a viabilidade e riscos dessa automatização no que concerne à ética da ação.

A tese a desenvolver promete refletir sobre este tema, sobre o papel dos algoritmos, da automação e inteligência artificial na capacidade de decisão humana e na automatização desta. No fundo, sobre o impacto da automação e automatização na redefinição das noções do que consideramos humano e ético sabendo que a maneira como afeta tais conceitos tem implicações tão profundas como as conceções de livre arbítrio.

# Parte I- O estado da arte

## 1 - Fundamentação teórica

### 1.1 – Abordagem ao tema

Teoriza-se que a chegada dos antibióticos possa ter como consequência um impacto negativo na resistência humana às infecções pois possibilitou que se deixassem sobreviver indivíduos menos resistentes, uma vez que deixam de estar à mercê da seleção natural aqueles mais predispostos e mais vulneráveis às doenças infecciosas, estes puderam sobreviver e assim perpetuar os seus genes defeituosos e menos resistentes.

O valor global do aparecimento deste milagre da medicina será então negativo e a humanidade irá mais cedo ou mais tarde pagar o alto preço da perda de resistência natural às infecções. Mas será mesmo assim? Passados quase cem anos da descoberta da penicilina parecemos ainda longe de pagar uma primeira prestação dessa dívida e a descoberta não parece ter trazido nada além de benefícios incontáveis.

Mesmo o advento da resistência microbiana no seu pico teórico não aparenta trazer mais consequências possíveis além da inutilidade do antimicrobiano e um regresso ao ponto de partida somente.

De uma maneira paralela o aparecimento dos computadores, dos algoritmos, das redes sociais fazem pensar que a substituição do intelecto humano no comando das operações possa levar à perda de capacidades sociais e intelectuais.

A ideia que o facilitismo providenciado pela automatização leva a um enfraquecimento das capacidades surge em oposição a um futuro de ficção científica em que as máquinas ganham consciência e se apoderam do mundo através da força, estaríamos nesta hipótese alternativa então a caminhar para um futuro que podemos denominar de “geração torpe”, de perda de capacidades (Bauerlein, 2009, p. 26). Os indícios à nossa volta parecem ser mais sugestivos deste tipo de futuro, por todo o lado se observam pais a fornecer um telemóvel ou um *tablet* como meio de calar a sua prole reclamante que vidrada em vídeos

de *youtube* desprovidos de conteúdo aparenta zombificar. Que futuro terá esta geração de crianças?<sup>3</sup>

À medida que as necessidades do humano forem sendo satisfeitas por robôs, a custo zero, a tendência seria para uma apatia por parte dos humanos, uma perda de capacidades e potencial, desvalorização das capacidades física e mentais, da criatividade, contrastada com uma maior competência da cognição artificial, um "suicídio evolutivo" com o ganhar de força progressivo das máquinas cuja disseminação ganharia a forma de desenvolvimento de uma vida própria e autônoma, uma mente própria.<sup>4</sup>

Independentemente do que o futuro nos reservar, é inegável que as implicações da inteligência artificial serão incontornáveis, e tal, é alvo do crescente interesse da filosofia já desde o início da computação, mais concretamente nos anos cinquenta. Moor diz-nos que muitas das violações de código ético como roubar, matar, difamar, existem hoje também no mundo digital e como tal as noções de ética e moral são aqui aplicáveis (Moor, 1985, pp. 266-275). A mesma opinião é espelhada em Mitcham.

*"Ethics in relation to computing and information technology (IT) has raised the issue of whether there are new ethical questions to be answered, or just new versions of old questions"*<sup>5</sup>.

Desde os anos sessenta que os modelos computacionais têm vindo a trazer nova luz a problemas clássicos da filosofia, Brey e Soraker, racionalizam mesmo que a computação é responsável por uma revolução filosófica trazida por tecnologia e que o assunto da inteligência artificial é o tópico que mais atenção capta na filosofia (Brey & Soraker, 2009, p. 28). No contexto deste trabalho o maior interesse será o âmbito da ética.

Sem alargar nas definições da mesma, uma vez que tal será abordado ao pormenor em capítulos posteriores, fica uma noção geral que jamais um comportamento será ético, seja qual for a perspetiva ou noção de ética, se este contribuir significativamente para uma

---

<sup>3</sup> O dilema da inquietude das crianças é um tema recorrente ao passar de cada geração e que há sempre a tentação de no limite da paciência seguir a via mais fácil e fornecer algo para acalmar a criança pensando apenas no imediato sem nos darmos conta da recorrência das situações. Para não ficar a ideia de um valor pejorativo relativamente ao entretenimento digital, deixo a lembrança do extenso uso da ritalina®, usada a partir dos anos 60 para os défices de atenção diagnosticados e presumidos, sendo que classifico o *tablet* da mesma maneira que a que o metilfenidato (ritalina®) é classificado legalmente nos Estados Unidos como substância controlada. Trata-se de uma substância com reconhecido valor médico com um alto potencial para o abuso.

<sup>4</sup> Tal é espelhado no filme de 2008 da Disney WALL-E (Bork, 2015, p. 6).

<sup>5</sup> Mitcham, 2005, p. 89 .

diminuição das capacidades humanas ao ponto da irrelevância ou para fazer caminho rumo a um evento apocalíptico de extinção em que as máquinas tomam o poder à força.<sup>6</sup>

## 1.2 - Contexto histórico

A inteligência artificial surgiu na década de cinquenta como método para o estudo da inteligência humana através da simulação (Brey & Søraker, 2009, pp. 28-30).

Uma vez que os processos de computação são conhecidos mas não os processos cerebrais, foi teorizado que tentando simular comportamento/inteligência humana com computadores se poderia compreender os métodos pelos quais funciona o cérebro, a inteligência. A explosão de conhecimento potencial com o desenvolvimento da tecnologia fazia prever que seria possível rapidamente simular a inteligência de um adulto humano. Em 1965, o investigador Herbert Simon previa que isto seria possível já em 1985 (Simon, 1965, p. 34). Trinta anos depois ainda aparentamos estar longe de tal objetivo.

Searle, delimita a fronteira do que chama de *Strong AI*, esta corrente estipula que um programa bem organizado pode corretamente simular um estado mental ou cognitivo, ao ponto de ganhar consciência (Searle, 1980, pp. 417-424). Este pensamento deriva pouco depois nas teorias computacionalistas (Pylyshyn, 1984, pp. 49-59); (Shapiro, 1995, pp. 467-487) com grande divulgação na década de setenta e que defendem que o cérebro é uma máquina digital e que há uma correspondência entre cognição e computação ao ponto de serem a mesma coisa.

*“...we are classified along with rocks , atoms , and galaxies for the purpose of revealing how we move in response to physical forces ...these are merely ways of classifying individuals for the purpose of discovering some of their operating principles”<sup>7</sup>*

Atualmente estas correntes têm vindo a desvanecer. Apesar de a simulação ser ainda um pedestal de ciência cognitiva, há uma tendência para a inteligência artificial ser usada de uma forma mais prática, uma ciência aplicada em que os programas são usados simplesmente para tarefas que possam requerer inteligência, esta abordagem é denominada *Weak AI* (Brey & Søraker, 2009, p. 30).

---

<sup>6</sup> Ao estilo *Skynet* em “O Exterminador” 1984.

<sup>7</sup> Pylyshyn, 1984, p. xii

Independente destas duas correntes, um facto é que ao longo das últimas décadas temos vindo a assistir a uma substituição de humanos por máquinas nas funções que tradicionalmente eram realizadas apenas por inteligência humana e o ritmo desta substituição não parece estar a abrandar.

As máquinas, os programas, as IAs têm já um grande impacto na evolução humana, por mecanismos indiretos afetam já as populações tendo imensa influência em questões de vida ou de morte embora maioritariamente de forma indireta. Tomemos pois como exemplo o alegado uso da inteligência artificial que disseminava notícias falsas, criava contas em redes sociais para influenciar os resultados das eleições, para escolher um decisor. Deste sujeito cuja validação como decisor é uma forma de fraude democrática, vão depender decisões sobre política de imensa importância, embargos, sanções, declarações de guerra, portanto questões de vida ou morte.

Aqui convém fazer a distinção entre autonomia e automação/automatização sendo que a primeira se refere à capacidade de prescindir de ajuda, trata o sistema pelo qual os mecanismos controlam o seu próprio funcionamento, são autónomos, regem-se a eles mesmos (de notar que autónomo aplicado às máquinas tem um sentido menos abrangente que o conceito de autonomia aplicada à pessoa, muito por utilização em diferente contexto, mais restrito no caso da máquina). Já automatização ou o diminutivo automação refere-se ao facto de o homem deixar de realizar as tarefas realizadas agora por máquinas, tarefas automatizadas. Ambos os conceitos estão intimamente relacionados com a inteligência artificial e os algoritmos, embora expressem diferentes facetas destes fenómenos (Mitcham, 2005, p. 146).

Podemos apenas especular o que acontecerá quando as máquinas começarem a tomar decisões com implicações diretas na vida/morte da população e é compreensível que será neste domínio que surgem as principais questões éticas relativamente à IA. Já esteve mais longe o dia em que a máquina consegue raciocinar, racionalizar<sup>8</sup>. Quando o raciocínio, aquele que é o pilar da definição de humanidade for realizado melhor por uma máquina do que por um humano, que nos resta depois? Que resta para decidir? E sendo as decisões consistentemente delegadas para as IAs, implicaria isso uma redução de autonomia humana?

---

<sup>8</sup> 1. Fazer uso da razão para depreender, julgar ou conhecer. 2. Procurar e alegar razões para convencer outrem. 3. Discorrer. "Raciocinar", in Dicionário Priberam da Língua Portuguesa [em linha], 2008-2020, <https://dicionario.priberam.org/raciocinar> [consultado em 30-04-2020].

## 2 – Ética e máquinas

### 2.1 – Responsabilização e transparência

Um dos grandes problemas do uso da IA para tomar decisões pelos humanos é que tal leva a uma erosão da noção de responsabilidade que é o conceito mais importante quer na ética quer na lei. A máquina errou e decidiu premir o gatilho, quem se responsabiliza? O programador? O usuário que ligou o aparelho aquela manhã esperando uma decisão acertada? O fabricante? E porque não a máquina em si? Certamente esta última maneira de ver as coisas tem cabimento nas teorias de Strong AI e derivadas. Segundo Sullins, se for possível a IA ter uma noção de responsabilidade então deveria ter direitos e responsabilidades independentemente de ser pessoa ou não (Sullins, 2006, pp. 24-30). Tal visão implicaria falarmos então da IA como agente moral, como uma fonte de bem ou mal. Aqui poderia dar-se uma desresponsabilização do *designer* e isto é usado por Moor para defender que os computadores não deveriam poder tomar decisões sobre princípios básicos e valores (Moor, 1979, pp. 224-229).

O problema é mais vasto, hoje não se trata somente de um único *designer* a programar um algoritmo ou máquina, na maioria dos casos são múltiplos, centenas, a trabalhar cada um na sua parte da escrita de um programa, muitas vezes alheios às minuciosidades do trabalho dos restantes colegas ao qual integram o seu próprio trabalho e avaliam posteriormente o resultado mediante experiências de campo (versões *alfa*, *beta* e finais), com erros, muitos erros que vão sendo limados à medida que vão sendo reportados, os chamados *bugs*, tão mais frequentes quanto mais complexo o programa e mais diversos os seus cenários de aplicação. Se o cenário de aplicação for o mundo, com certeza podemos esperar um sem-fim de *bugs*.

Um algoritmo, tratando-se de um procedimento preciso, não ambíguo, mecânico, eficiente e correto e esta definição, especialmente pela parte da ausência de ambiguidade remete para uma fraca adaptabilidade a situações. Um algoritmo portanto com esta definição funciona sempre em circuito fechado dentro do âmbito para o qual foi programado e somente nesse âmbito.

Tomemos então a modo de exemplo um algoritmo simples como o existente nas máquinas de calcular mais simples, as não científicas que permitem cálculos básicos. Todas as respostas fornecidas pela máquina são corretas mas esta só funciona dentro daquilo para a qual foi programada, se introduzirmos um caractere que a calculadora não tem previsto

aparece um erro no ecrã, o aparelho bloqueia e terá de ser reiniciado, o mesmo sucede com uma operação simples de tentar dividir um número por zero.

A maioria dos sistemas de IA são opacos, os critérios que derivam em escolhas e decisões são de difícil controlo pelo usuário e muitas vezes pelos próprios *designers*, especialmente quando construídos através de sistemas conexionistas ou de arquitetura evolucionista sem regras formais (e grande complexidade). Logo seria impossível garantir um funcionamento sem falhas, daí que a criação de AI para fins bélicos seria inerentemente não ético.

Então que dizer quanto à aplicação da IA para fins médicos, para decisões de vida ou de morte? Aqui o objetivo inerente seria o de salvar vidas mas a questão de fundo é similar, se pode uma máquina tomar uma decisão de implicações de vida/morte direta de forma ética e uma vez que não seria possível garantir o correto funcionamento, sem falhas, supomos que mais cedo ou mais tarde um erro iria ser cometido.

No hipotético pináculo da simulação, uma máquina poderia apresentar comportamentos sobreponíveis aos humanos inclusivamente aqueles aos quais adjudicamos características que fazem das pessoas, pessoas e no entanto ser constituída de maneira completamente diferente do humano, uma arquitetura da máquina muito díspar da arquitetura cerebral mas com comportamentos iguais.

Ganharia então a máquina sapiência e seria merecedora de estatuto de pessoa? Poderíamos adjudicar a este ser um *status* moral? Decerto estamos de acordo que se o seu comportamento for sobreponível ao de um humano, sendo este um ser sapiente, com capacidade de raciocínio, seria então a máquina capaz de responsabilização como uma pessoa, mesmo apesar de ser dúbia a existência de consciência ou experiências conscientes. Se se comporta tal como uma pessoa então talvez devesse ter estatuto e responsabilidade de pessoa. Estaríamos então talvez a atribuir estatuto de pessoa a uma casca vazia (sem alma?). Chalmers denomina este problema como “problema zombie” e a resposta a estas questões metafísicas demorará ainda a ter uma resposta (Chalmers, 1996, p. 95).

## **2.2 - Humano dispensável**

A evolução acelerada do progresso tecnológico parece fazer prever que não tarda as máquinas vão assumir o controlo, o preconizado *takeover* pelas IAs. Temos já métodos de simulação de discurso muito similar à de um humano, capacidade visual igualmente

similar que aliadas a bases de dados imensas fazem da máquina já um memorizador muito mais eficaz que qualquer humano. A máquina está sem dúvida a simular melhor o humano.

Resta ainda alguma coisa para evoluir nos termos da interpretação das coisas, do discurso, das imagens e baseada nestas discrepâncias ainda é notória a diferença entre o comportamento humano e algorítmico pelo que ainda podemos verificar falsidades e manipulação. Remeto novamente para o exemplo dos “*captchas*” que nos mostram que a IA não “vê” como um humano. Até quando?

Estas novidades tecnológicas proporcionadas por IA autónoma apesar de não estarem ainda implementadas (ainda não há realmente carros autónomos à venda)<sup>9</sup>, parecem estar já ali ao virar da esquina havendo inclusivamente uma multitude de experiências com sucesso no terreno. Em Setembro, 2019 o Washington Post (Shin, 2019) publicou um simulador onde se detalham os desafios atuais da condução autónoma. O nível cinco de condução autónoma, em que não há necessidade de intervenção humana estará ainda a décadas de distância. Atualmente encontramos-nos nos níveis dois e três em que um humano consegue ver quatro vezes mais longe que a tecnologia atual durante o dia, e esta torna-se totalmente inútil com mau tempo.

Portanto, todas estas experiências têm umas coisas em comum: o facto de estarmos a trabalhar com estas máquinas em circuito fechado, não haver variáveis desconhecidas e o facto de todas elas necessitarem de uma forma ou outra de intervenção humana.

Aquilo que atualmente chamamos de condução autónoma ainda é muito dependente do humano no que respeita às idiossincrasias do caminho. Aliás se ainda não é possível um operador de um comboio de metro ou de um intercidades automático, sendo estas situações em que o caminho é só um e a velocidade está limitada por defeito, como se pode esperar um veículo autónomo numa estrada na via pública?

Ao que parece, não é ainda possível criar um circuito fechado suficientemente abrangente para automatizar um veículo que circule num carril com uma rota diária fixa. Podemos presumir então que o advento dos carros autónomos verdadeiros (veículos sem pedais ou volante) está ainda longe. Há que referir por aqui que aquilo que achamos bom ou

---

<sup>9</sup> Uma publicação da AAA (*American Automobile Association*) de 10/2019 (*American Automobile association*, 2019) comunica uma serie de testes feitos aos sistemas de travagem automáticos atualmente disponíveis, (ADAS de *Advanced driver-assistance systems*) e revelam que em condições óptimas o sistema falha e atropela um adulto a atravessar a estrada durante o dia a 30Km/h 60% das vezes e uma criança 89% das vezes. Esta estatística foi ainda pior à noite ou condições de mau tempo sendo estes sistemas simplesmente classificados de ineficazes.

conveniente numa IA é muitas vezes fruto de estratégias de *marketing* somente. Talvez a percepção que temos de os algoritmos estarem já disponíveis e prontos a tomar conta das nossas vidas seja apenas uma coisa que às empresas tecnológicas convém para nos venderem produtos inacabados (Bory, 2019, p. 635).

Também se pode propor uma visão menos catastrofista em tudo isto, já aconteceu antes, múltiplas vezes na história. Destaco a revolução industrial, em especial na agricultura, a lembrar que os milhares de pessoas necessários a uma ceifa foram substituídos por um ou dois tratores, que os trabalhos nas minas de carvão também sofreram quebras de empregabilidade e no entanto o mundo não acabou. Todos sabemos de profissões que já deixaram de existir. Talvez esta vista catastrofista seja a simples expressão da falta de visão por quem a conjura. Ainda somos do tempo em que se dizia que substituir as pessoas por máquinas nos controlos das portagens nas auto-estradas era imoral, que devíamos abdicar da IA somente para manter gente empregada, assim ia ficar um humano a cobrar dinheiro aos condutores, a obrigá-los a parar com custos de tempo, ambientais, económicos, somente para justificar um ordenado.

O humano ao contrário da máquina é mais adaptável e lá acabará por encontrar sítio no mercado de trabalho.

### **2.3 - Máquinas éticas**

Moor, discute se a disciplina de ética das máquinas deverá existir uma vez que a ética é expressão de emoções e as máquinas não têm emoção. Fala sobre as considerações do que se pode entender como ética das máquinas porque nem todos os valores são éticos (Moor, 2006, pp. 18-21).

Como exemplo tomemos a durabilidade e utilidade, são valores muitas vezes atribuídos às máquinas mas que não têm em conta valores éticos. Se o chocolate pode e deve ser ético (aqui atribuível às condicionantes de produção e fabrico) então as máquinas também deveriam ser éticas e não só em termos de produção mas também de uso e avaliadas como tal. Um robô que substitua um trabalhador escravo do Bangladesh a fabricar roupa permitiu a libertação desse trabalhador do trabalho escravo, no entanto pouco fez pelas condições que são a propensão à escravização desse trabalhador no mesmo país. Tomando um exemplo mais local, um computador permite-nos socializar através de redes sociais muito mais facilmente do que era possível há uns anos, melhorando a nossa comunicação, mas também predispõe ao uso não ético dos nossos dados.

Debruça-se então sobre duas formas de ética das máquinas, como uma forma de julgamento ético (com base em aprendizagem, mais humana) e em termos de programação ética.

Pode então um computador operar eticamente por ser internamente ético de alguma maneira?

Para tornar uma máquina ou algoritmo mais éticos podemos por um lado delimitar o âmbito de aplicação dessa máquina de modo a que não possa atuar de forma não ética, no entanto isso diminui grandemente os contextos possíveis de atuação, uma rede fechada de comunicação permite alguma segurança na privacidade dos dados mas também limita as possibilidades de comunicação fora da rede.

Podemos também atuar noutra vertente, através de programação ética que promova comportamento ético, assim, um aviso de alerta surge num computador de um banco quando existe uma transferência não habitual e insta o usuário a confirmar ou averiguar os trâmites da transação, o computador funciona assim como um agente ético implícito.

Moor, indaga também sobre a possibilidade de um computador atuar eticamente de forma explícita, capaz de “raciocínio ético” e da mesma maneira como pode jogar xadrez poderia atuar como agente ético e justificar as suas ações (Moor, 2006, pp. 18-21).

A própria dificuldade de delimitação do que é a ética atualmente torna a sua implementação e definição no âmbito técnico muito difícil. Moor crê também que não está previsto que haja máquinas capazes de ser agentes éticos completos para breve mas tal será possível se limitarmos a ética a um contexto restrito. Argumenta também que mais difícil do que definir numa máquina o ser ou não ser ético, será dar à máquina a noção de senso comum, imprescindível para as noções de comportamento ético.

Joseph Weizenbaum escreveu uma extensa literatura sobre a ética e a programação e sugere que as questões éticas não devem ser colocadas depois do programa feito mas sim antes da sua génese, que uma vez que não conseguimos fazer máquinas sábias então não deveríamos deixar aos computadores tarefas que necessitem sabedoria (Weizenbaum, 1976, p. 227). Com efeito,

*“Ethics is not a question that should be raised by AI, but it should be the very foundation of AI; The justification for building a system in the first place should be an ethical question.”<sup>10</sup>*

Esta visão traz consigo encargos no desenho de um programa, não pode ser simplesmente idealizado para uma função mas terá que ser construído passo a passo de forma a manter um grau de transparência de funcionamento que permita fiscalização.

Surgem então numerosos entraves à ideia de desenhar um algoritmo mais simples e deixar a máquina aprender por ela, sem controlo absoluto sobre os porquês de uma decisão de modo a permitir superintendência. Por outro lado, o imperativo ético antes da conceção da máquina e a construção passo a passo faria com que as decisões da máquina pudessem ser compreensíveis, mecanizadas e a ética das suas resoluções estaria assegurada num pilar de previsibilidade, sem possibilidade de “livre arbítrio” por parte da máquina. O excesso de previsibilidade acabaria um pouco com a noção de máquina como agente moral, fonte de bem ou mal, com a noção de consciência da mesma. Por outro lado, se fosse já possível esquematizar todos os passos de resolução de um dilema ético complexo não teria qualquer sentido sequer a discussão deste trabalho, tal será abordado em capítulos posteriores.

### **3- Limitando as máquinas rebeldes**

#### **3.1 As leis da robótica**

Em 1942 Isaac Asimov nas suas obras de ficção científica elabora os preâmbulos das conhecidas seis leis da robótica as quais eram a base da programação dos robôs e estipulam o seguinte, segundo a ordem cronológica em que foram surgindo (Asimov, 2004, p. 37).

Lei Um: Um robô não deve causar mal a um humano ou, através de inação, permitir que um humano seja magoado.

Lei dois: Um robô deve obedecer às ordens dadas por humanos excepto quando essas ordens entrarem em conflito com a primeira lei.

---

<sup>10</sup> Brey & Søraker, 2009, p. 41

Lei três: Um robô deve proteger a sua própria existência enquanto esta não entrar em conflito com a primeira e segunda lei.

Lei zero: Um robô não deve fazer mal à humanidade ou por inação permitir que a humanidade seja contundida.

Lei quatro: Um robô deve ser capaz de se identificar ao público (identificação simétrica).

Lei cinco: Um robô deve ser capaz de explicar ao público o seu processo de tomada de decisão (transparência algorítmica).

As leis um, dois e três são as primeiras a surgir nas obras de Asimov. Estas são incorporadas em todos os robôs e o seu objetivo é funcionar como mecanismo de segurança não podendo ser ultrapassadas.

O enredo das obras gira em torno dos conflitos entre as leis. Apesar de aparentarem ser óbvias e de interpretação direta, as ambiguidades na concretização de conceitos como o de “não fazer mal” permitem muitas vezes resultados inesperados consoante a interpretação dos termos. Por exemplo, imagine-se um ponto em que um robô tentando obedecer às leis pode aperceber-se que os humanos fazem mal uns aos outros e são eles mesmo a maior ameaça às suas próprias existências, os robôs podem então chegar à conclusão que estarão mais protegidos se enjaulados, tal é descrito no filme *I, Robot* de 2004, inspirado na obra de Asimov.

Nas suas obras mais tardias foi adicionada uma outra lei, a quarta mas que foi denominada de lei zero como forma de manter a precedência sobre as outras leis. Esta lei zero permite já separar os conceitos de humano e humanidade mas acaba por trazer este último termo também carregado de ambiguidade por ser uma abstração.

Assim como as passagens da bíblia dependem das várias interpretações, Moor (Moor, 1985, pp. 266-275) menciona que a aplicação à letra é fonte de fraca previsibilidade dando como exemplo um robô a viajar pelo mundo para tentar evitar o dano em todos os humanos.

Outros autores tentaram adicionar outras leis para tentar resolver os problemas de conflitos e fiscalização, destaco Marc Rotenberg (Rotenberg, 2016), presidente da *Electronic Privacy Information Center* (EPIC) responsável por inúmeras ações de proteção ao consumidor e envolvida em processos contra *Uber*, *Snapchat*, *WhatsApp*, entre outros, que adiciona as leis quatro e cinco (Heiser, 2018).

### **3.2- ETHICS GUIDELINES FOR TRUSTWORTHY AI**

Em 8 de Abril de 2019, foi tornado público pela Comissão Europeia um documento denominado *ETHICS GUIDELINES FOR TRUSTWORTHY AI* (European Commission, 2019) com vista à promoção de uma IA confiável. Este está dividido em três capítulos.

No primeiro capítulo, chamado de as Fundações de IA confiável, são referidas três premissas para uma IA confiável, três componentes que devem ser cumpridos durante todo o ciclo de vida do sistema, são eles:

-Ser cumpridor da lei, cumprindo com as leis e regulamentos aplicáveis.

-Ser ético, com aderência a princípios e valores.

-Ser robusto, de uma perspectiva técnica e social, porque mesmo com boas intenções há possibilidade de dano não intencional.

Estes componentes isoladamente atuam como fator necessário mas insuficiente e idealmente trabalham em sintonia com sobreposição de funções.

Descartando a parte legal (provavelmente por necessitar adaptação a cada contexto) as *guidelines* têm ênfase na eticidade e robustez apenas. Estes complementam-se, a base serão os direitos humanos fundamentais, europeus e internacionais. Estão condensados em quatro imperativos ou princípios éticos, não hierarquizados, que podemos comparar de forma análoga às leis da robótica de Asimov, são eles o respeito pela autonomia humana, a prevenção do mal, a justiça, a explicabilidade ou transparência.

Os direitos fundamentais do homem estão associados na Europa a força de lei pelo que colaboram na adequação à primeira premissa (cumpridor da lei) e segunda (ético).

O documento advoga a mitigação de tensões entre os princípios, especialmente em sectores vulneráveis da população, os benefícios devem ultrapassar os riscos individuais (exemplo da vigilância para prevenção de crime) mas refere não haver uma fórmula fixa para lidar com estas tensões e conflitos e incumbe a sociedade de proceder para alinhar os princípios e evitar tensões entre eles.

No segundo capítulo, intitulado de realização de IA confiável, aborda-se o desenvolvimento e implementação de IA mediante o cumprimento de sete requerimentos de confiabilidade, supervisão humana, decisão humana (autonomia humana como questão central), robusteza técnica e segurança, privacidade e controlo de dados,

transparência, diversidade, não discriminação e justiça, bem-estar social e ambiental, responsabilização/prestação e contas, prevenir responsabilização das ações da IA antes e depois do seu desenvolvimento (isto obviamente entra em conflito com a noção de imprevisibilidade).

Tal como nas obras de Asimov, estas leis e requerimento de IA estão carregadas de subjetividade na noção de não fazer o mal, de justiça, de explicabilidade e de autonomia humana podendo já antever inúmeros conflitos entre estas leis. Uma vez que não há fórmula fixa para lidar com os choques entre os requerimentos e princípios, sugerem uma extensa documentação de conflitos, promoção da pesquisa, transparência com as partes interessadas. Sugerem que se deve presumir a imprevisibilidade e contínua correção e avaliação durante todo o ciclo de vida da IA através de um ciclo análogo PDCA (*plan-do-check-act*) -> uso -> análise -> redesenho -> desenvolvimento -> Uso... Promoção de um ambiente de verdade nas capacidades, limitações e requisitos facilitando o seguimento e auditabilidade pelas partes interessadas.

No terceiro capítulo intitulado de avaliação de IA confiável é sugerida a avaliação da IA programada através de uma extensa panóplia de perguntas do qual fornecem um teste piloto para avaliação da confiabilidade da IA, é feito finca-pé na necessidade de avaliação contínua e melhoria de resultados com envolvimento de todas as partes interessadas durante todo o ciclo de vida do sistema.

Em suma, nas *guidelines* está bem frisado que o documento serve apenas de guia e não tem força de lei servindo apenas como molde, para promover a segurança e a necessidade de estar sempre com “o pé atrás” no que respeita às IAs, tal é o potencial perigo representado.

As regras que sugerem, com o envolvimento de todos, a contínua avaliação, parecem ser regras genéricas de segurança ao ponto de poderem ser aplicadas no âmbito da exploração de qualquer sector em que há ainda muito para descobrir. A IA aparenta estar já à distância de todos e com os moldes de IA já disponíveis em *opensource* ao alcance de qualquer programador, estas regras vão criar um entrave de logística e burocracia que poderá deslocar a capacidade de fazer IA apenas para alguns, para os que tiverem recursos para lidar com a fiscalização e manutenção.

A facilidade com que aparenta hoje ser possível criar uma IA vai levar certamente à disseminação de várias IAs caseiras, sem respeito por *guidelines* e a contenção da expansão destas terá de levar à aplicação das *guidelines* com força de lei, mediante

penalizações e fiscalização assim como se faz de maneira análoga para o controlo da comercialização de tabaco caseiro e drogas leves e todos sabemos como resultou a tentativa de imposição da Lei Seca, da proibição de um artigo que estava ao alcance de produção por qualquer um.

Estas *guidelines* deixam também claro que apesar da aplicação na europa, almejam fomentar a pesquisa e reflexão de um molde ético de sistemas de IA a um nível mais global. Está bem claro que adoptam como fundo geral as leis da união europeia, a democracia e a cartilha dos direitos fundamentais do homem e como é de conhecimento geral, estas não formam parte dos princípios e valores de todo o mundo, podemos predizer que a aplicabilidade noutros contextos será difícil.

Sem entrar em juízos de valor sobre este “puxar a brasa à minha sardinha” é óbvio que este almejar de divulgação a nível mundial procura disseminar os valores atuais europeus.

È de salientar que o elevar deste tipo de valores ao que consideramos ao momento o pináculo do bem-estar, resulta provavelmente do facto de serem adotados por países cujos níveis de desenvolvimento são superiores, onde a qualidade de vida é objetivamente melhor e assim tem sido durante centenas de anos. Por uma questão de simples lógica é fácil atribuir a adoção destes padrões de comportamento, destes valores, como estando relacionados com tal desenvolvimento do país, ainda que seja difícil concluir relação de causalidade, ou seja, se os valores são respeitados pelo facto de haver qualidade de vida ou se contribuem como de forma direta à formação da base do que consideramos qualidade de vida. Saliento então que no momento atual há países como a China que apesar de estarem completamente fora do âmbito de aplicação destes valores aparentam estar a caminhar para um nível de qualidade de vida comparável ao dos tradicionais países desenvolvidos e se, com o passar das décadas se verificar que é possível o bem-estar sem adesão a estes valores, então a relevância dos mesmos e as características daquilo que consideramos moral ou ético hoje podem vir a modificar-se. Daqui se conclui que apesar de a forma como são elaboradas as *guidelines*, aparentarem ser apenas guias de elaboração de IA ética para todos, na prática acabam por não ser.

#### **4- Limitações atuais**

Retomando o assunto do *Preâmbulo*, é muito discutível se existe alguma hipótese de um algoritmo tomar uma decisão de contextos complexos em que estejam em causa vida/morte por si só, de forma ética. O aumento da complexidade e da necessidade

computacional cada vez que se dá um passo para a implementação de um algoritmo na vida real parecem oferecer ainda sérios obstáculos à emancipação das máquinas. Se considerarmos o algoritmo como sempre funcionando em circuito fechado então o cenário parece muito distante mesmo.

O circuito fechado, a atuação sem ambiguidade, são incompatíveis com o processamento de todas as variáveis que possam entrar numa decisão muito complexa. O exemplo do *Preâmbulo* do erro de “*divide by zero*” ao surgir uma situação inesperada, que obriga a desligar o programa e a forçar o reinício está ainda patente. Quem experienciou os sistemas operativos *Windows* dos últimos trinta anos sabe que era assim mas também se dá conta que agora não é qualquer situação inesperada a forçar um reinício de sistema. Quer seja por ter aumentado o âmbito do circuito fechado com abarcamento de mais variáveis quer seja pela inclusão de algoritmos com redundância, a verdade é que há menos caixas no ambiente de trabalho a pedir *input* ao utilizador sobre como proceder, “*Repeat? Ignore? Cancel?*”.

Com recurso a discriminação de casos e à criação de grupos quasi-homogêneos poderemos manter em circuito fechado o funcionamento do algoritmo, dando à sua atuação esta capacidade de “redundância” sem necessitar a colocação de todos os casos individualmente e portanto sem sobrecarregar memória e sem requerer computação infinita. O problema é que esta redundância dá aso a más interpretações além de manter os mesmos erros quando surge uma variável “*out of the box*”.

Não somente sobre forma de contornar o erro em termos de continuidade programática, há que ter em conta que sempre existirão exceções e erros e tal leva a incerteza sobre outras dimensões do âmbito do erro, como é o caso da responsabilidade. Se for a máquina a tomar a decisão quem é o responsável pela mesma?

Recordando o exemplo do *Preâmbulo* “o algoritmo tomou a decisão errada, que fazemos? Cadeia com ele? Na lógica da máquina ela procedeu corretamente, para o que estava programada.”

A elaboração de máquinas com base em teorias conexionistas ou evolutivas apresentam em parte resposta para este problema, um erro é uma oportunidade de eliminação de ligações em detrimento na hierarquia de outras ligações que melhor se coadunem à situação, e ao longo de muitas tentativas e erros poderíamos ter uma máquina mais afinada e replicá-la. Os custos óbvios seriam pois de transparência e previsibilidade de resposta, a transparência é seriamente comprometida neste tipo de sistemas conexionistas

uma vez que a máquina evolui ou aprende por ela, sendo difícil discernir sobre o que está na base da aprendizagem e prever o comportamento. A máquina não é um humano, não tem consciência, não tem *status* moral e portanto a responsabilização da mesma não é uma possibilidade.

*“While it is fairly consensual that present-day AI systems lack moral status, it is unclear exactly what attributes ground moral status”<sup>11</sup>.*

Pode uma decisão ética partir de algo sem responsabilidade? (Decisão ética vs programação ética). As *guidelines* europeias também têm em conta estes “erros” de uma forma análoga, em sistemas não conexionistas, portanto mais transparentes, preconizam processos de melhoria contínua para tentar a minimização do erro, uma outra forma de “aprender com os erros”. No entanto ambas hipóteses apesar de permitirem evolução frente a um determinado contexto presumem que o “zero erros” nunca seria alcançado. Há que se ter em conta que a evolução não pára e os contextos alteram-se, os erros irão sempre existir, não são preveníveis uma vez que só se modifica o programa depois de acontecerem, trazem então com eles a necessidade de responsabilidade e responsabilização mantidas à qual oferecem pouca resposta.

Ao embater com o joelho numa mesa a culpabilização da mesa tem pouco sentido. A solução consiste no desviar da responsabilidade para a entidade ou para a empresa responsável pela colocação ou construção da mesa ou mesmo na atitude negligente da vítima. No que concerne à atuação mais ativa que no caso da mesa, ou seja, uma decisão de um algoritmo, as dúvidas sobre responsabilidade judicial mantêm-se e a solução seria talvez a apresentada no *Preâmbulo*: “contar com os erros de decisão como expectáveis, admitir que eles acontecem e preparar-se para indemnizar todos os casos excepcionais”, o que seria possível e viável se no preço de aquisição do algoritmo este custo estivesse contemplado. Será esta atitude ética?<sup>12</sup>

O *Preâmbulo* levanta também a possibilidade de haver uma “discriminação” relativamente à ética das máquinas sendo estas penalizadas por decisões individuais quando há a possibilidade de no global o benefício do somatório das decisões das mesmas ser superior ao do caso em que as mesmas decisões eram tomadas por humanos. Se a definição de certo ou errado é mais esbatida no caso da intervenção emergencial ou situações de aceleração temporal em que prima a escassez de tempo e em que o fator

---

<sup>11</sup> Bostrom & Eliezer, 2014, p. 6

<sup>12</sup> Imaginar uma máquina de cirurgias ao apêndice com 10% de mortalidade mas cujo custo de aquisição fosse suficiente para permitir cobrir os custos das indemnizações e ainda ter lucro.

“sorte” está muito colado à decisão humana, talvez possamos deixar a cargo das máquinas somente estas decisões em que esta tem acesso a mais dados, atuação mais rápida e baseada em protocolos.

Relativamente à questão do dilema básico da teoria da evolução, a consequência da eliminação da variabilidade referido no *Preâmbulo* que, por exemplo podemos resumir na seguinte pergunta:

“Se todos os carros fossem automatizados salvaríamos vidas na estrada, é um facto, conclui-se então que devemos proibir a condução não automatizada. Isto é ético?”<sup>13</sup>

Retomando o dado anteriormente em que as máquinas podem estatisticamente atuar de forma global de forma mais benéfica que os humanos nas mesmas situações podemos fazer um ponto intermédio. Apesar de aparentar ser um facto que no exemplo das situações de aceleração temporal a máquina pode atuar melhor, a realidade é que haverá sempre situações onde a atuação humana é preferível. Será talvez na discriminação sobre qual o tipo de decisões que cada um deve tomar que esteja a solução intermédia. É necessário para isto que haja uma hipótese de escolha sobre quem toma a decisão. No exemplo da condução autónoma podemos manter a variabilidade de comportamentos se permitirmos voluntariamente a intervenção humana na condução, devidamente responsabilizada e incentivada e não como é atualmente em que é relegada a situações de emergência cuja atuação está já coagida por horas de condução autónoma monótona que desviam a atenção da estrada por parte do humano.

O texto todo até aqui discute sobre a possibilidade de o programa cometer erros por si, mas as *guidelines* europeias alertam para outro tipo de variáveis a ter em conta que apesar de mais facilmente evitáveis em teoria, acabam por ser elas a manchar mais a reputação dos algoritmos. Falamos pois dos erros feitos na altura da programação, intencionais ou simplesmente como reflexos dos vieses a que os programadores estão sujeitos. Um exemplo recente será o algoritmo para deturpar os valores de emissões no escândalo da *Volkswagen* que ao que parece estende-se já a várias outras marcas em múltiplos contextos, mas claramente intencionais.

Char, num artigo saído no *New England Journal of Medicine* fala dos problemas éticos relacionados com o espelhar dos erros comumente cometidos por humanos nas decisões médicas e dos resultantes do desalinhamento entre as intenções dos programadores (lucro) e dos utentes (saúde) (Char & Shah, 2018, pp. 981-983). Por outro lado os

---

<sup>13</sup> *Preâmbulo* vide supra, pp. xix

algoritmos com aprendizagem com base em padrões estatísticos tendem a ser altamente discriminatórios, seria o caso ao avaliar o tratamento a alguém a quem foi diagnosticado um tumor de pâncreas com base em estatísticas de mortalidade de 95% aos cinco anos sem ter em conta a raridade mas ainda assim possível existência de variantes mais benignas.

A disponibilização de serviços de saúde é muitas vezes subtilmente discriminatória e de difícil antecipação pelo que prevenir que um algoritmo atue com base em estatística sem aquisição deste perfil discriminatório se torna muito difícil.

Como exemplo podemos verificar que os serviços de saúde são algo discriminatórios no que concerne ao seu acesso, os centros hospitalares estão num determinado local por alguma razão logística ou qualquer que seja, mas será sempre um facto que nas grandes cidades os habitantes têm melhor acesso a estes centros nem que seja por estarem mais perto destes. Ora então para a formação dos dados do hospital vão entrar desproporcionalmente mais os habitantes que vivem mais perto, dando lugar depois na elaboração de estatística um viés de discriminação relativamente à população mais distante que é muito menos representada. Estes vieses são muito complicados de eliminar.

A qualidade da atuação do algoritmo vai estar muito ligada à qualidade dos dados sobre os quais extrapola decisões e sabemos que na prática clínica na base de uma decisão há que ter sempre em conta tempo, lugar e idiosincrasias do utente, ou seja, o contexto, difícil de refletir na estatística e muito difícil de avaliar por um algoritmo (Obermeyer, 2019, pp. 447-453).

Ainda na temática da responsabilização, também como noutros contextos, o evoluir das máquinas tem tendência a criar mais afastamento entre prestador e utente, neste caso médico e utente, também com grandes implicações de responsabilidade. Na prática de medicina clínica esta relação tem sido essencial e a introdução de computadores faz com que muitas vezes o utente seja “visto” apenas pelos resultados dos exames, sem o tradicional exame clínico, com óbvios prejuízos em termos de responsabilização a que assistimos diariamente (a culpa não foi minha, foi da máquina). Mais além assiste-se à transformação da relação médico-utente numa relação entre utente e serviço de saúde originando uma ainda maior diluição de responsabilidade (a culpa não é minha, é do sistema).

## 5- Considerações finais da primeira parte

Analisando o contexto histórico e fazendo uma previsão com base na evolução ao momento atual parece que o único consenso é que o uso dos algoritmos na decisão é que estes serão indispensáveis para automatizar decisões simples, fechadas que proporcionam tempo para deliberar na variável *out of the box* que é geralmente posto a cargo de um humano ou para lidar com problemas cujas violações éticas são problemas de gravidade mais solucionáveis<sup>14</sup> (privacidade na disponibilização de registos médicos eletrónicos) mas que são hoje a principal fonte de problemas éticos da digitalização com a disseminação da *big data* sob a forma de uma globalização digital.

Não há ao momento a possibilidade de uma tomada de decisão “humana” por parte de uma máquina e menos ainda de forma ética, se não tivermos em conta algum tipo de intervenção humana. O algoritmo com o seu contexto de sistema fechado ajudará à tomada da decisão, enriquecerá a mesma com o seu acesso a informação privilegiado e capacidade de processamento ímpar, no entanto não fará parte do cerne da decisão ética e não substituirá a decisão de forma autónoma. Pensar de outro modo levará a uma diferente interpretação sobre os conceitos de algoritmo, ética e de decisão ética que se abordam no capítulo seguinte.

---

<sup>14</sup> Comparado com a panóplia de possibilidades de violações éticas possíveis ao colocar o controlo aos comandos de um computador, a privacidade dos registos médicos desencadeia problemas éticos com os quais se está a conviver atualmente e apesar da frequência das violações deste tipo, não se considera sequer a hipótese de estes registos não estarem em suporte informático.

## Parte II: Ética e moral

### 1- A necessidade da dilucidação dos termos

Apesar de já se poder tirar uma conclusão preliminar com base na análise histórica do texto introdutório, o aprofundar da delimitação do algoritmo na decisão ética necessita uma sondagem mais exaustiva das vertentes em questão.

A necessidade do apuramento e dissecação dos termos em jogo é fundamental para a questão em estudo, para a sua problematização. Em primeiro lugar há que ter uma definição mais clara do que é um algoritmo. Seguidamente damos conta que, no que concerne aos algoritmos no processo de decisão há dois campos a ter em conta.

O algoritmo como parte da decisão.

O algoritmo é a decisão, decide.

Para saber o papel do algoritmo na decisão ética teremos de examinar estes dois termos mais profundamente de forma individual; decisão e ética.

#### 1.1 – Algoritmo

Para dilucidar este termo, opta-se pela tradução de um provedor de informação sobre linguagem conceituado americano, o *Merriam-Webster*, escolhido tendo em conta o contexto de surgimento de inovações tecnológicas neste continente e porque muitas das palavras utilizadas no contexto de informática têm uma conjuntura anglo-saxónica.

Segundo este provedor, algoritmo está definido como: “Um procedimento para solucionar um problema matemático. Um processo passo a passo para atingir um fim”. “O conjunto de regras seguidas por uma máquina para atingir um fim desejado” (Merriam-Webster, 2020)

De um ponto de vista genérico dentro da dualidade *hardware/software* fica a impressão que o algoritmo será um programa, portanto *software* processado por *hardware*, ou seja, fisicamente transposto por um meio tangível para atingir um fim ou um resultado.

Por exemplo, na elaboração de uma *pizza*, as instruções para a realizar será o algoritmo que é lido por *hardware* e processado cozinhando e juntando os ingredientes sequencialmente até ao resultado final, a *pizza*. No entanto a distinção entre *software* e *hardware* não é tão linear assim e não podemos simplesmente classifica-los recorrendo somente à tangibilidade, funcionalidade ou maleabilidade de cada um, sendo que há uma equivalência lógica entre ambos. “Qualquer operação realizada por *software* pode ser *imbuída* diretamente no *hardware*... Qualquer instrução executada por *hardware* pode ser simulada em *software*” (Mitcham, 2005, pp. 896-898)

O *hardware* muda de estado conforme vai processando o algoritmo (a impressora imprime) e o *software* (as instruções) permanece estático. Tal não é uma regra pois na abrangência da evolução computacional atual e de importância para esta tese temos de ter em conta que os programas de *software* podem ser dotados de capacidade de auto-modificação e não são somente uma sequência de repetições, podem portanto ser adaptáveis, plásticos e esta definição aparta-se um pouco da rigidez considerada aquando a elaboração do *Preâmbulo*. Assim “um meio para atingir um fim” não tem de estar vinculado a uma regra estática como o algoritmo presente numa máquina de calcular simples. Numa máquina que aprenda o algoritmo pode funcionar como uma sequência de iniciação que se vai modificando, no término do seu exercício será já uma função muito mais complexa.

## **1.2- Decidir e escolher**

No que concerne à decisão há que incidir sobre a sua diferenciação com a palavra escolha. Apesar de se usarem frequentemente como sinónimos, expressam coisas diferentes apesar de serem interdependentes. Os termos não podem ser alvo de tradução direta a partir do inglês *decide* e *choose* uma vez que significam conceitos não totalmente sobreponíveis.

No Moraes escolher está como sinónimo de *eleger*, separar o bom do mau (Bluteau, 1789). O dicionário da Porto Editora adiciona como sinónimos preferir e marcar. Aí há indicação de proveniência do termo a partir do latim *excolligere* que significa «recolher; obter» (Porto Editora, 2020).

Analisando o significado no dicionário Priberam, para o termo “escolher” temos como significado, o mesmo que: fazer escolha de, preferir, estremar, separar, marcar, optar (Priberam, 2020).

A notar que não figura nos exemplos a palavra decidir como sinónimo de escolher.

Já para o verbo “decidir” o dicionário Moraes tem como significado o mesmo que resolver, julgar, sentenciar, determinar. O dicionário da Porto Editora acrescenta deliberar e o Priberam soma ainda: ser causa imediata, concluir, fechar, ultimar (contratos, negócios, etc.), emitir opinião ou voto; dispor.

Ou seja, em “escolher” temos a noção de ser uma coisa imediata, com poucas contemplações enquanto os sinónimos de decisão apontam mais para o final de um processo. Esta *nuance* de definição de escolha e decisão aparenta ter-se mantido desde o século XVIII.

Tal como na lei, uma decisão deve ser munida de um princípio subjacente que a legitime e não somente uma escolha. Um juiz decide, não escolhe uma pena de entre as possibilidades para o delito em causa, não escolhe por convicção moral. A decisão está relacionada com o juízo, presume um compromisso. O ato de decidir implica a justificação dos motivos da decisão. Decidir não deve simplesmente ser uma manifestação de uma escolha entre várias hipóteses possíveis, é um ato ponderado, uma razão para a escolha, uma fundamentação (Streck, 2013, p. 107). Podemos falar da decisão como uma escolha com deliberação (conceito este abordado em capítulo posterior). Logo uma boa decisão implica boas escolhas, mas para escolher não é preciso decidir.

Assim, uma máquina escolhe ou decide dependendo do partido que tomemos sobre a possibilidade de máquina ser capaz de algo similar a um juízo (Bory, 2019, pp. 627-642). As hipóteses de tal foram já discutidas no ponto 4 do capítulo anterior que neste texto foi deixado em aberto o campo teórico da possibilidade de raciocínio por parte da máquina e assim validar a hipótese de uma máquina poder “ser a decisão” no futuro. Isto claro, ainda que tomemos a máquina como ser decisor, tal pouco nos diz sobre os aspetos éticos dessa decisão.

O consenso parece mais claro sobre o papel das máquinas como parte da decisão, nem que seja para contexto. Nos tempos atuais é difícil tomar decisões de grande âmbito (políticas ou de planeamento) sem auxílio informático para pelo menos acelerar análises de dados, como otimizador de decisão. O uso das máquinas como auxiliares de decisão não parece criar obstáculo à eticidade da decisão, salvo nos casos não tão raros de existência de vieses no algoritmo usado mas tal poderia existir igualmente se os dados fossem tratados por humanos e não por máquinas. Estes também têm idiossincrasias, vieses próprios ou simplesmente diferentes crenças ou formas de educação.

O busílis da questão está no facto de que talvez não seja possível simplesmente separar os termos decisão e ética como sendo distintos. Como nos diz Köhler (Köhler, 2015, p. 10) na sua psicologia da forma, da “*Gestalt*”, “o todo não é igual à soma das partes”.

Atribuindo as diferenças anteriormente explanadas para escolha e decisão, o termo “escolha ética” terá algum sentido? Tal necessita elaboração sobre o termo de “ética” e outra palavra muitas vezes usada como sinónimo, moral.

### **1.3- Diferenciação entre ética e moral, contexto e evolução**

As palavras “ética” e “moral” usam-se com frequência como sinónimo mas podemos denotar uma pequena diferenciação conforme a aplicação do termo seja usada para se referir áquilo que é considerado bom ou ao obrigatório. Usam-se para designar a moral vivenciada, a vida moral e também para designar a moral pensada, intelectualizada ou cientificamente aceiteada, ao que se chama filosofia moral ou Ética.

A ética como filosofia moral pertence ao campo da axiologia da filosofia que estuda o valor das coisas (Internet Encyclopedia of Philosophy "Ethics" , 2020). Divide-se em três áreas, são elas a meta-ética que investiga a origem dos princípios e o seu significado, a ética normativa que concerne à parte prática ou aplicativa da ética que determina o curso de ação moral com padrões morais e por fim a ética aplicada que examina assuntos controversos do dia-a-dia definindo o que é proibido ou permitido. Singer defende que a “*Global Ethics*” ou uma ética de cariz mais universal promove valores aceites universalmente como o respeito pela vida humana, direitos humanos, equidade, liberdade, democracia, sustentabilidade, solidariedade e portanto éticos serão os comportamentos que melhor contentarem estes parâmetros (Singer, 2008, p. 341).

Este basear num sistema de referência, ainda que este possa variar consoante o tempo e localização geográfica, promove as bases para considerar a ética uma reflexão crítica sobre a conduta humana.

*“From the perspective of science, technology, and ethics, ethics itself—that is, critical reflection on human conduct—may be viewed as a science, as a technology, and as providing multidimensional independent perspectives on science and technology”<sup>15</sup>.*

---

<sup>15</sup> Mitcham, 2005, p. 700

Moral usa-se para fazer referência ao conjunto de normas e critérios pelos quais se rege um indivíduo ou grupo nos seus atos, assim como para avaliar os mesmos, para aprovar ou desaprovar (moral/imoral). Assim podemos referir-nos ao decálogo como uma moral para os cristãos e à deontologia médica e seu conjunto de normas como uma moral da profissão médica, que regula as condutas do médico com a sociedade em geral e com o âmbito profissional. Ética, como sinónimo de filosofia Moral seria a parte da filosofia que orienta, que trata de dizer como devemos comportar-nos fundamentando racionalmente os critérios pelos quais pretende reger.

*“Aristotle draws a clear distinction between moral and intellectual virtues. The former are acquired by habituation and produce right action in changing circumstances. The latter are acquired by learning and are oriented toward an understanding of the nature of things”<sup>16</sup>.*

A ética como parte da filosofia diferencia-se do conhecimento moral da vida quotidiana pela tentativa de ser um saber preciso, sistemático e racionalizado que não se finda na mera descrição mas levando uma vertente prática, não diz como são as condutas mas como devem ser. Afirma-se assim a ética não como mero saber teórico mas teórico-prático sendo que o seu objeto é a “vida moral”, a reflexão sobre a moralidade (relação intrínseca com a praxis moral).

*“Ethics is not just a science but a practice, a technique for self- and social improvement. Insofar as this is the case, ethics provides guidelines for development of character and counsel for political organization and rule. Ethics leads to politics, meaning not just political action but political philosophy”<sup>17</sup>.*

Aristóteles diz, com razão que na ética não refletimos para saber o que é a virtude (saber teórico) mas sim para ser virtuosos (praxis).

*“The body of the Ethics is accordingly devoted to a treatment of virtues such as bravery, temperance, generosity, and justice”<sup>18</sup>.*

No entanto esta perspectiva assente em valores fundamentais não é bem o conceito de ética aristotélica. A herança aristotélica caracteriza a ética de acordo com o fim (telos), com o objetivo final, adoptando uma perspectiva Teleológica em que uma atuação é boa somente

---

<sup>16</sup> Mitcham, 2005, p. 108

<sup>17</sup> Ibid., p. 701, citando ética a nicomaco.

<sup>18</sup> Ibid., p. 107.

se o resultado da atuação for bom<sup>19</sup>, independente do método do ato em si. Em contrapartida, Kant (Kant, 1996, p. 117) define a moralidade com o carácter obrigatório da norma, ênfase no dever, na obrigatoriedade para agir de determinado modo que é o quanto basta para valorar o ato em bom ou mau, independentemente do resultado final do mesmo. Falamos de uma perspectiva deontológica. Podemos resumir as duas visões dizendo que a Teleológica tem uma ênfase sobre «aquilo-em-vista-do-qual-se-age» e a outra, a deontológica, sobre «o-que-deve-ser-feito». A ligação entre as duas é necessária para a compreensão da articulação entre o alvo finalístico e as normas necessárias para o atingir, sendo a base da modelação decisionária entre a escolha premeditada e a deliberação (Amaral, 2019, pp. 121-132).

Esta associação já é de certa forma preconizada por Aristóteles na medida em que divide as virtudes em éticas e dianoéticas. As virtudes éticas ou de carácter derivam do hábito e as dianoéticas (ou da inteligência), derivam da aprendizagem exigindo experiência e tempo (Silveira, 2000, p. 10).

Ética procede do grego clássico (que provém do sânscrito que não vem ao caso). O substantivo “*êthos*” originariamente significava toca, nicho, covil, lugar ou num âmbito mais alargado, moradia. Está em relação com refúgio, local que oferece segurança. Homero assim o descreve, sendo o *êthos* o local onde coabitam humanos e animais em relação uns com os outros, uma relação simbiótica. *Êthos* está estreitamente relacionado com “costume” em provável associação dos costumes aos locais de habitação. Pensar que é fácil associar diferentes costumes consoante as regiões geográficas onde habitam povos diferentes. Posteriormente “*éthos*” passou a significar “caráter, forma de ser” e este muitas vezes varia conforme o local de habitação. Diferentes regiões e diferentes povos têm diferentes costumes.

Aristóteles, um leitor de Homero, pega em *êthos* e transforma a palavra em conceito e dá-lhe uma ênfase filosófica falando do *êthos* interior, do habitat interior ou carácter, o lugar onde moram os atos humanos. Altera assim o significado conceptual da palavra. Pergunta-se quais os tijolos, a matéria com que se constrói essa moradia interior, com que se edifica o carácter. Surge então a palavra “*éthos*” referindo-se a esta matéria, sendo esta os atos, os costumes. *Éthos* mantém proximidade com as relações do habitat tal como *êthos* pois o carácter permite o relacionamento com os outros. O carácter permite o relacionamento numa base de confiança fundada na expectativa que podemos traduzir em “eu sou o meu

---

<sup>19</sup> A perspectiva teleológica almejava fazer o máximo de bem ou optava pelo resultado em que a soma dos bens individualmente fosse maior, muitas vezes priorizando o bem comum em desprezo do individual com as contradições que isso implica (Santos, 2012, pp. 62-66).

carácter”. *Éthos* faz então referência ao modo de ser e está estreitamente relacionado com o “costume”, com o modo habitual de agir. A palavra “*éthos*” ou “costume” derivou e está relacionada intrinsecamente com “*êthos*”, significa “carácter” e da mesma maneira o carácter ou forma de ser das pessoas procede dos costumes, dos hábitos ou formas de fazer as coisas.

Podemos então expor que as virtudes éticas em certa medida pela nossa “forma de ser” podem ser recebidas, aperfeiçoando-se pelo hábito, do costume. As virtudes éticas surgem então através de comportamentos treinados a partir de hábitos cuja repetição induz a sua formação. Por outro lado, através de um fazer aprendendo induz-se o exercício de virtudes dianoéticas, aprende-se a agir, fazendo. Recorre-se então a uma espécie de educação cívica que promove o desenvolvimento da ação desde o hábito virtuoso arquitetando a finalização da eticidade.

A palavra moral que se corresponde com o vocábulo latino “*mos*” inicialmente significava “costume”. Pelas mãos de Cícero passa a significar “carácter” ou modo de ser com a necessidade de traduzir ao latim o vocábulo grego original “*êthos*” acabando por fundir na forma “*éthos*” a significação de hábito, costume e preceito que o termo moral ainda conserva atualmente, perdendo-se assim a anterior diferenciação entre “*éthos*” e “*êthos*” simplesmente por enviesamento de tradução. Moral, refere-se a boas práticas, a comportamentos, recursos. Refere-se à parte material, a costumes (Amaral, 2019, p. 126).

Esta definição de ética e moral não é consensual, mas a discussão já existia no tempo de Aristóteles<sup>20</sup>.

Segundo Aristóteles, a distinção entre hábito e caráter adquire relevância praxiológica pela sua interdependência. Tal significa que a virtude advém de uma capacidade decisória eletiva, da avaliação e ponderação e não somente de um mecanismo repetitivo de aquisição de hábitos. O benefício da capacidade decisória irá então além do interesse individual, irá teleologicamente, em vista do fim, à procura do interesse comum. Esta procura do interesse comum vincula estreitamente a ética à política como o veículo para ser arquitetada.

Decidir bem, decidir eticamente, implica conformidade com disposições legais e demais restrições “morais” necessitando educação. A educação articula então o objetivo virtuoso com o hábito, na conceção do indivíduo prudente, que delibera e elege com razoabilidade e sabedoria. A virtude moral, induzida por hábito ou treino tem como termo, um processo

---

<sup>20</sup> A palavra ética contém em si dois caminhos, foi o Aristóteles o primeiro a ter isso em conta, já tem uma história de dupla interpretação. ARIST., EN, II, 1103a 14-26

de metamorfose ética que se ajusta aos restringimentos do caso concreto personalizando a atuação e levando o sujeito a refletir em “porque é que se age como age”, tal culminando na tomada de decisão. A decisão é então a modelação do universal ao particular, passa-se de uma ética mais geral à particularidade do caso, o decisor modela a forma à matéria, a decisão depende do contexto, (ARIST., EN, VI, 7, 1141b 8-23).

Com estas definições aplicadas ao tema atual, sendo a moral derivada do hábito, do costume, podemos dizer que a moral será o conjunto de instruções para determinar o sentido de uma ação, o comportamento de um programa fechado, escolhas deontológicas e opções técnicas.

A ética será então referida aos moldes do carácter, individual que se valora em ético ou não ético conforme a adesão de um conjunto de normas de carácter mais universal e racionalizado, ponderado e refletido num contexto e não somente a aplicação de uma tabela de comportamentos, esta aplicação seria do domínio da moralidade.

Apesar destas definições de ética e moral aparentarem dar ao ato de moralizar algo uma aparência de exercício intelectual, podemos por outro lado pensar que esta intelectualização será talvez mais uma tentativa de dar sentido a “instintos” que existem como consequência evolutiva de uma seleção natural de comportamentos ao longo da história da evolução humana. No entanto, tal conceito transborda o âmbito deste texto.

## **2- A ética aplicada ao contexto clínico**

Os termos ética médica ou ética clínica são a aplicação de um conceito de ética geral ao contexto específico que é a profissão médica, aqui interessa a sua aplicação clínica. A medicina e a filosofia têm evoluções paralelas na medida em que uma trata da saúde corporal e a outra da saúde da alma e são inúmeras as analogias e as sobreposições entre os dilemas clássicos da medicina e da filosofia, sobretudo interessando a parte ética (Carrick, 2001, p. 21). A ética e a medicina são conceitos que estão intimamente relacionados com os avanços das ciências da vida e da saúde e atuam como um filtro de relação entre o que é considerado bom ou mau na atividade clínica onde há um nível ético e um nível técnico. A ética clínica concerne à maneira como se relacionam os médicos com os seus utentes enquanto pessoas, processos decorrentes dessa atividade profissional seja

ela a parte clínica ou de investigação e com os problemas morais/éticos que possam surgir decorrentes do exercício profissional.

## **2.1- Contexto histórico e o *ping-pong* entre o paternalismo e autonomia**

A relação médico-doente é uma entidade de difícil conceptualização e definição visto a constante metamorfose que sofreu ao longo da história ao sabor das correntes filosóficas e do pensamento dominante de cada época e local (Silva, 2016, p. 3).

De um modo geral as correntes dominantes são o reflexo da diferença de força/conhecimento entre médico e paciente e a relação médico-paciente vigente é o reflexo desta assimetria que varia muito não só com o tempo mas com o local geográfico onde se desenrola não havendo uma harmonia global, adquirindo evolução com relação a diferentes tempos e locais.

Podemos dizer que a relação iniciou-se provavelmente ligada à tradição mágica em que nem o médico nem o paciente têm controlo vigente sobre a doença ou sobre o processo terapêutico atuando o médico/terapeuta/sacerdote apenas como instrumento para as forças divinas atuarem no paciente, sujeito passivo, especialmente no que concerne a patologias psiquiátricas e outras para as quais não havia um componente físico claro e observável (Silva, 2016, p. 4). Já na Grécia antiga é introduzido o raciocínio e o conhecimento em detrimento da mística religiosa, os objetivos terapêuticos deslocam-se no sentido de obter satisfação do utente e como tal, este ganha direito a opinião e a participação ativa no processo terapêutico. Surgem as condutas éticas cujo pináculo verte no aparecimento do juramento de Hipócrates (Carrick, 2001, p. 14).

Durante a época da Europa medieval esta evolução sofre algum revés, novamente com o terapeuta a situar-se num patamar superior frente a doente que passa novamente a sujeito passivo com reintrodução da temática religiosa e retrocesso do pensamento científico. Apesar disto há que destacar que durante a idade média surgem os hospitais, ficando a cargo de ordens religiosas com muito ênfase na solidariedade e beneficência, mais virados para as populações mais pobres havendo como consequência do mesmo grande interesse no bem-estar do paciente, é nesta época que surge o valor da caridade (Bynum & Porter, 1993, p. 56). O facto da relação médico-doente estar presidida por um forte cunho religioso permite-nos dizer que esta relação está provida de um “*éthos*” de serviço com

muitas obrigações que são sobreponíveis ao juramento hipocrático com uma ênfase de caridade adicionada ao sentido de missão.

Assim se o enfermo estava triste era adicionado vinho à refeição e fazia-se o possível para o abstrair do processo morboso com números de teatro, preconizava-se o alívio do sofrimento e a felicidade deste era um sinal de bom resultado terapêutico ou pelo menos que os clérigos estavam a desempenhar bem a sua missão de servitude.

Surgem as necessidades de mudança de roupa dos leitos como progresso higiénico.

Apesar da ausência dos meios técnicos e da crueza com que realizavam procedimentos cirúrgicos ou médicos muitas vezes ineficazes e com algum impacto no agravar de doenças (Por exemplo as sangrias terapêuticas e as desidratações para curar as diarreias) os clínicos da altura dispunham de habilidade no controlo do efeito placebo e na arte do consolo.

A partir do renascimento há ressurgimento do pensamento científico. A escassez de clínicos levava a que apenas classes sociais abastadas pudessem aceder a cuidados de saúde. Dá-se também nesta fase uma preponderância do utente sobre o médico pois uma vez que o doente paga, a competição dos clínicos era sobre quem conseguia agradar mais ao utente, independentemente da patologia no geral, o objetivo principal era a eliminação de sintomas (Bynum & Porter, 1993, p. 60). O ressurgimento do pensamento científico acaba por contagiar os hospitais que gradualmente acabam por sair da alçada religiosa mas herdando o realce da necessidade de proporcionar cuidados aos mais pobres, aqui justificados por imperativo humanista. Progressivamente os hospitais religiosos transitaram para jurisdição secular.

A partir do século dezoito a nova leva de conhecimentos científicos acarreta a que a doença seja encarada como um processo e não somente como um sintoma. O tratamento assenta num plano terapêutico em que o médico examina o utente e trata-o. Este sistema encontra-se em linha com o despotismo iluminado vigente em sociedades da altura. Assim as classes mais baixas são persuadidas a acarretar planos de saúde de índole mercantilista através da “polícia médica”. Submetendo-se às indicações dos governos esta população mais fragilizada tinha acesso a saúde. Gera-se novamente uma assimetria adoptando o doente de novo uma atitude passiva em que como ser ignorante ao processo terá de se sujeitar ao tratamento sem contestação, como implica este modelo paternalista da época.

O século vinte implicou uma evolução sem par na história da medicina, os avanços da ciência aliados à disseminação do conhecimento e à pronta disponibilidade deste graças

aos avanços informáticos elevou o padrão de conhecimento do utente a um patamar que só a sua disponibilidade pode limitar. Também os avanços na área da psicanálise ajudam novamente a colocar o doente no centro do processo de tratamento como parte participativa iniciando o processo hoje conhecido como “medicina centrada no doente” (Silva, 2016, p. 9). Novamente o doente assume preponderância no processo morboso e terapêutico mas de uma forma distinta da anterior, agora é visto como “pessoa”, como entidade afeta por uma doença, com capacidade de autonomia de pensamento e vontade.

No final da segunda guerra mundial o tribunal de Nuremberga (Bynum & Porter, 1993, p. 1554) põe a descoberto as atrocidades perpetradas por médicos com a desculpa de progresso científico ou de ordens superiores o que desencadeia um compromisso posterior de não mais descurar aspetos éticos da atividade médica. O cume vem com a “Declaração de Genebra” em 1948 que atualiza o juramento hipocrático estabelecendo princípios éticos fundamentais sobre os quais se deveria basear a atividade médica em todo o mundo.

Mais recentemente há uma nova leva sem precedentes de progresso científico e este evoluir demasiado rápido da difusão de informação fez pender a balança da relação médico paciente para o extremo contrário do paternalismo. Nos últimos cinquenta anos descambou-se numa mentalidade em que se somaram inúmeros processos judiciais contra médicos, inicialmente contra atitudes de paternalismo e mais posteriormente buscando deliberadamente o erro com vista ao dinheiro indemnizatório. Isto visto sob uma perspectiva principialista<sup>21</sup> é o resultado da preponderância do princípio da autonomia sobre os demais, o utente faz do médico o que quer e os seus serviços são mais para cumprir os seus desejos do que para tratar doenças.

## **2.2- Mais além da ética, a bioética**

Nos Estados Unidos, não ao acaso mas porque desde o início do século XX que são parte da maior representação de desenvolvimento técnico e científico a nível mundial surgem no Hospital Hopkins encontros informais dirigidos por André Hellegger com o objetivo de debater questões relacionadas com os avanços tecnológicos e científicos. Este desenvolvimento é mais destacado que outros sobre o mesmo tema em universidades de Nova Iorque e Wisconsin, contemporâneos. Estes encontros são em parte motivados pelo surgimento de preocupações com os direitos dos utentes, na moda pelo impacto e

---

<sup>21</sup> Corrente da bioética que se baseia na formulação de quatro princípios fundamentais, de beneficência, não maleficência, justiça e autonomia. Sem hierarquia e com igual valor. (Beauchamp & Childress, 1994, p. 100)

interesse como tendência difundida pela comunicação social e contrastam opiniões científicas com filosóficas iniciando assim o estudo interdisciplinar científico/humanista mesmo antes de ser nomeado de bioética (Garrett, Jotterand, & Ralston, 2013, p. 44).

Helleger estava convicto que o futuro apresentaria como grandes problemas o dilema da distribuição de recursos pela consciência da finitude destes em contraste com o crescimento da população a nível mundial e o conseqüente aumento de consumo. Propunha para a sua solução o diálogo entre a economia da saúde e a demografia com vista ao planeamento. Dentro desta dicotomia consumo/natalidade surgia o problema da deterioração ambiental.

O termo bioética surge pela primeira vez a partir de um texto de Fritz Jahr num texto sobre um “imperativo bioético” sobre o uso de animais e plantas na investigação em 1926 (Sass, 2007, pp. 279-295). Posteriormente em 1970, Ransselaer Potter (Lolas, 2008, pp. 119-123) pela junção de “Bio” e “ethos” descrevia a comunicação entre as ciências e as humanidades como a construção de uma “ponte para o futuro”.

O ser humano é então encarado como ser ativo e passivo numa tripla vertente, biológica, fisiológica e cultural (paradigma bio-psico-social) com vista à sobrevivência da humanidade num mundo em mudança. Roga nos seus textos a evolução dos comportamentos humanos para abandonar as tendências inatas de buscar vantagens a curto prazo em prol do planeamento através da interiorização da noção da necessidade de visão de futuro como requisitos mínimos para se falar de sobrevivência aceitável contraposta a sobrevivência miserável (Potter, 1970, pp. 127-153).

As perspetivas de Helleger e Potter acabam por ser semelhantes, uma Bioética Global que engloba os problemas recentes da denominada globalização.

A corrente da bioética acaba por semear a ideia da pessoa autónoma que conclui acabando por dar aso à maior mudança na relação médico-doente dos últimos cinquenta anos que é o reconhecimento do utente como agente moral autónomo e a capacidade de recusar um tratamento se este não se encaixa com o seu perfil de valores.

Em contracorrente surge mais recentemente um movimento denominado de “Ética das virtudes médicas” que veicula a ideia do “médico como pessoa” em que este é visto como alguém com direitos e autonomia e não somente um instrumento de serviços. Atualmente almeja-se que a relação médico-doente seja a derivada de uma relação que tenha em conta o contexto único social e cultural próprio dos intervenientes, baseando o progresso

terapêutico numa questão de consensos entre médico e doente, partilhando o poder, a responsabilidade e o respeito.

### **2.3- Os três níveis de juízo médico**

Paul Ricoeur (Ricoeur, 2010, pp. 1-22) diferencia três níveis de juízo em jogo entre médico e paciente, são eles o Pacto de confiança, o Contrato médico e o Não-dito dos códigos. Na relação medico-paciente as questões éticas surgem quando há intervenções deliberadas no processo de vida humana e não humana e estas surgem tanto no âmbito clínico como de investigação embora Ricoeur faça a distinção de níveis mais aplicada à parte clínica.

O primeiro, o Pacto de confiança é também é chamado de prudencial (do grego *Phronesis*): a faculdade de julgar é decorrente do ensino e exercício e de natureza mais ou menos intuitiva surgindo em consequência da relação entre dois intervenientes individualizados, o médico e o paciente. Quer isto dizer que na relação médico-paciente existe uma assimetria criada no início da relação. Por um lado está o médico, aquele que sabe o que fazer, por outro lado o doente, aquele que sofre. O sofrimento é pois uma característica única de individualidade e a maneira como se sente ou se vive. Tal como o prazer, é um filtrado de toda uma vida de experiências singulares daquele indivíduo. O sofrimento é também a motivação fundamental da medicina tendo como foco a saúde física e mental. Na base dos juízos prudenciais está o desejo de finalizar com o sofrimento, neste objetivo assenta a estrutura do ato médico. Subjaz a este ato um pacto de confiança, de compromisso entre as duas partes em que ambos vão colaborar com o mesmo objetivo, sendo este o processo de tratamento da doença. Esta colaboração esbate a assimetria inicial prometendo empenho das duas partes, do médico com a função de “seguir” o doente e do paciente em se “conduzir” como agente do próprio tratamento. Esta promessa implícita é constitutiva do estatuto prudencial do juízo moral subentendido no “ato de linguagem” da promessa.

O segundo nível ou deontológico, deriva da aplicação de normas instituídas como sejam os códigos de deontologia médica à relação médico-paciente de forma individualizada. Esta generalização do pacto de confiança na relação médico-doente implica a sua transformação em norma, o que constitui o movimento deontológico do juízo. Assim, os preceitos da prudencialidade são revestidos deontologicamente. A confidencialidade passa a ser regulada pela justiça sob a forma de segredo médico e o pacto de cuidados passa a ser

refletido sob a forma de relação contratual, de obrigação refletida por exemplo nos códigos de deontologia médica.

Os códigos e a deontologia têm ainda outra função tão ou mais importante, a de sanar possíveis conflitos, de os arbitrar formulando compromissos decorrentes da experiência dos debates nos diferentes níveis do corpo médico, opinião pública e poder político.

Os principais conflitos da prática médica surgem em duas frentes, a primeira é relativa à zona de fronteira entre a prática clínica e a investigação, a segunda, na frente da saúde pública. Relativamente à bifurcação clínica/investigação há que considerar que ambas são interdependentes mas têm objetivos diferentes. A clínica serve para melhorar os cuidados enquanto a investigação serve para fazer avançar a ciência. A ciência não avança sem a clínica e esta é melhor quanto mais avançada a ciência. O doente é ao mesmo tempo um ser pessoal e um objeto de investigação. Se ninguém duvida que a ciência melhora a clínica na globalidade, o mesmo não se pode dizer a nível mais individual. Aquele que é alvo de experimentação, de investigação, nem sempre sofre procedimentos compatíveis com a melhor clínica disponível e provada. A sujeição do doente a este desnível entre o cuidado provado e o experimental necessita sempre o seu “consentimento informado”. De notar que esta imposição deontológica de consentimento protege ambas as partes, por um lado o doente fica protegido dos abusos de poder por parte do médico e este fica também protegido contra os abusos judiciais em processos de má prática em casos de simples fracasso terapêutico.

Relativamente à frente da saúde pública, fica a noção que estes conflitos não são tão resolúveis com recurso a códigos de deontologia e que estes muitas vezes conservam ambiguidade para dar aso a escolhas conforme o contexto. A verdadeira resolução destes problemas está a cargo da bioética, dos comités e do sistema judicial. Como exemplo destaca-se o facto de o segredo médico não ser absoluto podendo ser anulado no caso de risco de saúde pública (veja-se o exemplo atual da declaração do estado de emergência ante a pandemia de coronavírus que permite inclusivamente suspensão de direitos constitucionais adquiridos), tal exceção está por determinar pela ética legal que varia conforme o contexto e a legislação aplicável de acordo com a sociedade em questão.

Ricoeur coloca num terceiro nível a parte da bioética correspondente à legitimação dos juízos dos níveis anteriores através da reflexão. Refere-se ao não dito dos códigos ou aquilo que não figura nestes. Está em jogo a noção de saúde, pública ou privada que não se pode dissociar da forma como uma sociedade encara a vida, a morte, a identidade, o ser ou não ser... Aqui a deontologia entra no campo da antropologia filosófica que não consegue

fugir do pluralismo das sociedades democráticas. Este pluralismo implica que tais conflitos e as suas decisões sejam legitimados de forma razoável apenas, consenso por comparação (Pegoraro, 2002, pp. 75-107).

## **2.4- A clínica e a ética**

Diego Garcia diz-nos que a clínica e a ética compartilham o mesmo método de origem hipocrática. Esta modificação teve origem em Aristóteles que provavelmente pelas suas raízes de contato com a medicina ao descrever a lógica de raciocínio prático. Este tipo de raciocínio é contraposto ao intitulado raciocínio teórico, típico da matemática, que assenta na “demonstração” e na “certeza” pois conhecidas as variáveis, o desenrolar do raciocínio dá-se sobre uma estrutura mais rígida, concreta, cujos resultados serão objetivos (Garcia, 2001, pp. 18-23).

Na clínica e na ética, ao contrário do que acontece com a matemática, os resultados e as variáveis não são concretos, têm matizes ou *nuances* e jogam com o desconhecido. Tal conjugação impossibilita esta rigidez de um resultado concreto e assim inviabilizam o uso do raciocínio teórico, não há demonstrações ou certezas. A lógica da clínica é a própria da incerteza, da probabilidade onde a certeza é impossível e como tal pretende-se que as decisões sejam não certas mas “razoáveis” e são passíveis de retificação. Ao invés do teórico, o raciocínio prático assenta em duas condições básicas, são elas a “prudência” e a “deliberação” que não são mais que uma transposição da “demonstração” e da “certeza” através de um filtro de probabilidade, de incerteza. Com a probabilidade em jogo, a certeza não é certa, é provavelmente a mais certa, a mais razoável. A decisão torna-se tão mais razoável quanto menor a incerteza. O procedimento de decisão assenta nesta redução de incerteza através da “ponderação” dos fatores intervenientes, das variáveis. Este processo é o chamado de “deliberação” e à decisão “razoável” (aquela feita após deliberação) foi chamada “prudente”.

Assim podemos atribuir toda a teoria de deliberação, a prudência, o raciocínio, a toma de decisões na incerteza de uma mesma maneira à ética e à clínica pois ambas compartilham o método do raciocínio prático (Garcia, 2001, pp. 18-23).

Um problema ético consiste num conflito de valores e este conflito também existe na clínica. Aqui a diminuição da incerteza é muito baseada na identificação dos vários valores em conflito e para tal usa-se a história clínica. Para avaliar a eticidade de uma deliberação há sempre que contrastar os princípios em jogo e as suas consequências previsíveis.

A ética de Aristóteles aparta-se do extremismo dos teleologistas e dos deontologistas navegando num meio-termo. Se por um lado os principialistas outorgam aos princípios um carácter absoluto independentemente das consequências dos mesmos, os consequencialistas importam-se somente com as consequências dos atos dando menos importância aos métodos para lá chegar. Ao dissecar os métodos destas duas correntes vamos acabar por verificar que ambas são assentes em valores rígidos, sem plasticidade, valores absolutos e como tal acabam por se aproximar mais às práticas do raciocínio teórico pois a existência de algo absoluto baseia-se sempre na ausência de incerteza.

Como já vimos anteriormente a clínica e a ética estão cheias de incerteza por os valores e as variáveis em jogo nunca serem totalmente conhecidos, daí é relativamente fácil encontrar conflitos reais que tombem estes modos analíticos e teóricos de ver a vida. Tome-se como exemplo os frequentes dilemas e conflitos de valores entre os princípios de autonomia e beneficência que de tão frequentes na clínica, agem como um obstáculo intransponível à rigidez do método. Já relativamente aos teleologistas e consequencialistas temos de concordar que o melhor resultado e mais ético é aquele em que haverá maior benefício. Já esta definição de maior benefício depende muito da perspetiva e do contexto.

Sendo os deontologistas, apologistas da ética do dever, do modo (importa que algo seja feito segundo as regras apenas, independentemente das consequências) e os teleologistas apologistas da ética de fins em que o resultado deve ser aquele em que haja maior benefício global, menos interessa o caminho para lá chegar. Podemos dizer que Aristóteles era apologista de uma ética de intenções, em que um ato é ético se foi bem estruturado, bem-intencionado e deliberado ainda que as consequências venham a ser nefastas ou que se tenham de saltar regras para lá chegar. O próprio Kant definiu que o único valor absolutamente mau é a má vontade (Kant, 1996, p. 117). Não se quer com isto dizer que a ética de intenções passa por cima das regras e dos objetivos, longe disso. Os princípios e os resultados continuam a ser os pilares orientadores da decisão, este método apenas nos diz que tem de haver exceções sendo que exceções são isso mesmo, eventos excepcionais que devem ser devidamente justificados com o contexto. Se a exceção se torna regra então passaria a ser um princípio, uma norma e fazer a exceção por princípio acaba por trair o propósito inicial da deliberação que é tornar a decisão como única, aplicada somente àquele contexto e não generalizável.

# Parte III: Deliberar e decidir

## 1- Conceito de deliberação

A palavra deliberar vem do latim *deliberare*, que significa pensar seriamente sobre um assunto ou tomar uma decisão.

Em latim, *deliberare* vem da junção de duas palavras: *de* e *librare*. *De* é uma preposição com o sentido de “originalmente de” ou “a partir de”. *Librare* significa pesar na balança, equilibrar duas coisas ou refletir sobre um assunto. Assim, *deliberare* tem o sentido literal de “tomar uma decisão a partir da reflexão sobre as opções” (Dicionário etimológico, 2020).

*Librare* vem da raiz *libra*, que era uma unidade de medida de peso romana. A balança para pesar também era chamada de *libra*. As balanças tinham dois pratos e funcionavam pelo equilíbrio de dois pesos. Quando duas coisas tinham o mesmo peso, os pratos ficavam equilibrados, mas se um objeto fosse mais pesado, seu prato afundaria.

A ideia de deliberar vem dessa imagem de pesar duas coisas na balança. A pessoa que delibera “pesa” as opções, analisando a informação que tem, para ver se existe uma opção que é melhor que a outra. Assim, toma uma decisão mais informada. Por exemplo, se alguém quer comprar um carro mas está indeciso entre dois diferentes, vai deliberar, pensando nas vantagens e desvantagens de cada carro para decidir qual é o melhor carro para si.

## 2- Pode uma máquina deliberar?

A resposta a esta pergunta é fundamental para saber se a máquina ou algoritmo são capazes de entrar no processo decisório como figura central, se são capazes de decidir e não somente auxiliar a decisão.

Estando a ética de definição aristotélica ligada ao dever de deliberação para evitar cair em formalismos e decisões “programadas”, poderia uma máquina que se presume intimamente vinculada a uma formalidade de programação decidir de forma ética? Poderia uma máquina ser capaz de ponderação?

Mais uma vez, para a resposta afirmativa neste dilema estaríamos novamente a atribuir características humanas a uma máquina. De certo modo a formalidade de programação pode já não ser um entrave uma vez que há máquinas cujo desenvolvimento tem características de autonomia, aprendem pelo meio, pelas experiências e o resultado acaba por ser uma amálgama cujos processos de resposta são difíceis de discernir e prever (Moor, 1979, p. 218). De uma forma análoga um humano também é “programado” através da educação, das experiências. Tomando o exemplo de um jogo de xadrez, a máquina que aprende não se limita a seguir um algoritmo, evolui, modifica a maneira como responde ao estímulo. Analisa um contexto, visualiza o tabuleiro e integra esta imagem com a aprendizagem anterior e “decide” uma jogada. Esta visão de Moor não é consensual e podemos arguir que a máquina é sempre regida por um algoritmo mas se considerarmos que este é modificável com a aprendizagem, então “o algoritmo” é mais uma sequência de iniciação apenas, termina como algo diferente consoante o contexto aprendido. De uma maneira similar um computador liga-se e a primeira coisa que faz é correr o algoritmo memorizado em ROM (*read only memory*) que apesar do nome, o facto é que esta memória é programável ainda que por vezes só eletronicamente. Sabendo como foi educado um humano, podemos presumir as suas respostas ante uma determinada situação mas perante a complexidade das interações da personalidade com as experiências será sempre muito difícil prever com certeza uma resposta pois os processos íntimos que levam a ela são indecifráveis. Uma coisa é certa, não podemos simplesmente dizer que a resposta é somente dependente do ADN desse humano que segue estático a vida toda, do seu algoritmo de iniciação. De uma maneira análoga isto acontece também com estas máquinas que aprendem. Não se deve também cair no erro de atribuir uma espécie de livre arbítrio somente fundado na nossa incapacidade de prever comportamentos<sup>22</sup>. Essa previsibilidade das máquinas também deveria ser claramente um limite à criação, à criatividade pois um algoritmo fechado, pré-programado jamais poderia fazer algo de novo. Tal não é bem assim<sup>23</sup>.

A máquina também não tem as respostas diretas para um dilema, um impasse obriga-a muitas vezes a recorrer ao utilizador procurando *input* ou a uma outra forma pré-programada de resolução para não surgir um erro que finde o processo em “*stop*” (surge o exemplo do “*blue screen of death*”<sup>24</sup> do *Windows*). Mas recorrendo a este último método

---

<sup>22</sup> Se sabemos como um humano vai responder dizemos que é fruto de ser coagido, no caso de uma máquina dizemos que são as limitações de programação fechada.

<sup>23</sup> Recordando os duelos máquina/humano de xadrez e de Go que terminaram com a vitória das máquinas, foram atribuídas à “invenção” de jogadas novas, não conhecidas sendo que a possível intervenção humana nessa criação seja especulatória (Bory, 2019, p. 639)

<sup>24</sup> BSOD, do inglês *Blue screen of death* vem o ecrã azul da morte que é uma tela que surge nos sistemas *Windows* no caso de ocorrência de erro grave e irrecuperável de sistema deixando como única alternativa o reinício do sistema forçado e consoante perda de dados de trabalho em curso.

pré-programado (ou pré experienciado no caso de máquinas que “aprendem”) podemos dizer que a máquina “delibera” à sua maneira e não somente calcula, computa.

Escavando mais a equiparação homem-máquina, ante a situação de dilema ético, o humano “dá-se conta” da situação, pondera nela, falamos de uma capacidade de compreensão da situação, de consciência, que ao momento no que concerne às máquinas pouco passa de um tema teórico, alvo de obras de ficção. Não existe definição legal de consciência e a presunção de uma consciência artificial levaria a outros dilemas éticos tais como a definição dos direitos da máquina como entidade consciente, como ser. Contudo, esta consciência do problema não parece ser um requisito para um ato se considerar uma decisão<sup>25</sup>.

### **3- Dissecação da deliberação**

Apesar do exposto anteriormente nos sugerir que é impossível fazer um algoritmo universal para a resolução de problemas de forma ética, a capacidade de deliberação é passível de ser sistematizada ou esquematizada.

#### **3.1- Dilema e clínica, contexto histórico**

O dilema, o impasse tanto na vida como na clínica é uma constante. Podemos dizer que a necessidade de decidir faz parte da raiz do aparecimento da medicina pois estamos a lidar com uma disciplina com ampla vertente prática. O desenrolar da atividade clínica é fonte de dilemas morais e éticos de forma constante mas aquilo que consideramos como a resolução mais correta está muito dependente da evolução das mentalidades e dos valores do contexto, a bioética de maneira recente obrigou à implicação de fatores externos à dualidade “paciente – médico” com a introdução das questões de impacto ambiental (da natureza) na equação.

Como está na própria definição do dilema ou impasse, em jogo estão valores com peso similar. A resolução destes impasses requiere tempo, ponderação, deliberação, usando lógica e razão para a maturação das decisões. Requer a análise de minuciosidade dos valores já considerados ou contemplação de outros valores novos que possam fazer

---

<sup>25</sup> Numa prateleira do supermercado decidimos por uma marca de manteiga após análises de preços, experiência prévia, aspeto da embalagem, a maioria das vezes de forma automática ou pouco consciente mas tal não justifica relegar esta decisão a mera escolha.

sobressair a diferença de peso entre as opções em jogo e deste modo se poder decidir melhor.

Antes do aparecimento do exercício da livre racionalidade na Grécia Antiga a fonte dos valores para a tomada de decisão era a tradição mítico-religiosa. Esta oferecia uma “razão” para uma escolha de determinada forma, quer fosse para apaziguar um Deus, para escolher conforme as escolhas de um ídolo, existia já uma dicotomia entre bem e mal. Atualmente a esfera de valores para uma tomada de decisão é muito mais abrangente, sejam a experiência, a moral, a religião, a lei, entre outras que conforme o contexto se tornam mais relevantes a uma tomada de decisão considerada ética.

Sócrates introduz a conceção de ética baseada na racionalidade que distingue aquilo que é uma opinião individualizada ou de grupos de um conceito mais universal criado através do reconhecimento de valores e de metodologia. A introdução desta dimensão prática da razão oferece a delimitação das opções possíveis. Nestas opções há que distinguir patamares, variáveis de importância ou valores diferentes. Acontece que o valor dos valores é inconstante, habitualmente a rotina e o hábito trazem com eles um automatismo, sendo que a moralidade é vivida de acordo com a correspondência às obrigações habituais (Cabral, Cardoso, & Pena, 2012). Mesmo dentro do principialismo existe o termo *prima facie*<sup>26</sup> que nos indica a preponderância de alguns valores sobre outros, a existência de prioridades consoante as circunstâncias (Beauchamp & Childress, 1994, pp. 99-102).

Na situação limite há sempre uma insuficiência de uns valores sobre outros não só pelo seu valor abstrato mas porque levam o decisor a questionar tal valor e o que ele representa em termos de opção por uma “melhor escolha”. O valor de qualquer valor é medido consoante a circunstância em que a estamos a medir e tal não pode ser universalizado ou sujeito a uma simplificação algorítmica. Ou pode?

Na tentativa de organizar a hierarquização dos valores surge a possibilidade da colocação destes em esquema através de uma árvore de decisão. Com esta solução há um grupo de interrogações iniciais (um tronco) que são desdobradas em várias possibilidades e por sua vez ramificadas noutras escolhas conforme a adequação ao caso, desembocando finalmente numa das várias soluções finais (os ramos terminais). Isto permite a exclusão progressiva de opções com vista ao afunilamento das hipóteses de decisão, trilha-se assim

---

<sup>26</sup> O termo *prima facie* foi pela primeira vez usado por David Ross (Ross, 1930, pp. 19-36) na década de 30 adjudicando o termo às obrigações que se devem cumprir salvo em conflito com outras obrigações hierarquicamente iguais ou superiores. Ross categorizada os deveres como: 1. Deveres para com os outros devido a atos próprios prévios, são eles os de Fidelidade, Reparação e Gratidão. 2. Deveres para com os outros não baseados em acções prévias. São eles os de beneficência, não maleficência e justiça. 3. Deveres para consigo mesmo em relação com o desenvolvimento do potencial próprio e melhoria constante.

um caminho único pela árvore que se inicia num tronco comum e desemboca num ramo único, terminal, numa solução.

Segundo Diego Garcia, na ética clínica as árvores de decisão usam-se vulgarmente de forma simplificada dos quais se destacam pioneiros como Howard Brody (Bellamy, 2003, p. 51) com recurso a métodos analíticos de árvores de decisão ou David C. Thomasma (Thomasma, 1978, pp. 33-36). Esta forma de orientar a decisão de um procedimento médico não é mais do que uma maneira algo analógica de representar um algoritmo, instruções de forma mecânica, traduzida num papel e executada por um humano.

Mais recentemente Beauchamp e Childress criam, baseados em Ross, dois patamares de deveres, os *prima facie* e os *atual duties* (Garcia, 1998, p. 15). Acreditavam ter encontrado um processo universal e incontestável de resolução de conflitos éticos mas a prática veio demonstrar que tal não se verificava. Os princípios da bioética são *prima facie* e como tal não poderiam ter exceções, mas quando aplicados ao mundo real por vezes entram em conflito e há que escolher um para tomar primazia sobre os restantes. As exceções faziam-se sempre no nível dos *atual duties* e não dos princípios.

Da descrição de Beauchamp e Childress não fica patente a forma de hierarquizar mas é sugerido pelo texto que deveriam ser priorizados de acordo com as consequências, similar a uma ética de fim, teleológica. Para tentar resolver este conflito, Diego Garcia hierarquiza os princípios em dois patamares. Um primeiro patamar que corresponde aos princípios de não maleficência e de justiça que se encontram hierarquicamente acima dos princípios de autonomia e beneficência, num segundo patamar. Os primeiros implicam todos por igual independentemente da opinião ou vontade e correspondem a uma ética de mínimos, em relação com os deveres para com os demais. Os segundos são negociáveis dependendo do contexto da decisão e uma imposição dos mesmos seria imoral, são de índole mais privada e correspondem a uma ética de máximos (Garcia, 1998, p. 16).

### **3.2 Esquematização/perspetivas da deliberação**

Carlise Nora procura numa revisão de literatura focar o tema da deliberação ética como solução para analisar os problemas éticos de forma sistematizada e contextualizada. Mais uma vez a prudência revela-se como pilar da decisão ética e consiste na capacidade de avaliar alternativas envolvidas no conflito ético de forma razoável chegando à solução mais prudente e não somente a mais correta (Nora, 2015, pp. 114-123).

O conceito de deliberação mais premente no estudo é o proposto por Diego Garcia que oferece uma estruturação para a descoberta da solução óptima ou menos prejudicial, assentada numa crítica aos princípios (Garcia, 2001, pp. 18-23).

Este autor descreve que na altura de tomar uma decisão bioética há que ter em conta três níveis (Bioeticawiki contributors):

Nível um: Sistema de referência (ontológico): Passa pela universalização da ação moral para manter em primeira linha os deveres para com os demais. Podemos desdobrar em:

1. Premissa ontológica: o homem é pessoa, tem dignidade e não preço.
2. Premissa ética: enquanto pessoas, todos os homens são iguais e merecem igual respeito.

Nível dois: Delineamentos morais (deontológico). Ao contrário dos conteúdos da moral que são sujeitos a evolução, o conteúdo formal da ética não está dependente de evolução por parte da razão. Trata de conjuntos de conteúdos que a razão elabora partindo de um sistema de referência ideal, são hipóteses, suposições, esboços ou rascunhos.

1. A não maleficência e a justiça correspondem diretamente à não discriminação biológica descrita no sistema de referência e como tal estes princípios adquire um carácter obrigatório típico de uma ética de mínimos, do dever, referem-se ao bem comum, aos direitos.
2. A autonomia e a beneficência são relativos ao indivíduo e ao contexto deste, são negociáveis e dependentes do próprio sistema de valores, típicos de uma ética de máximos. Estamos no campo da Moral.

Nível três: Experiência moral ou experiência racional (teleológico).

Os delineamentos têm de ser contrastados com o sistema de referência, mas também com a experiência. Corresponde ao teste com a realidade. Sabemos que os delineamentos morais nunca se adequam perfeitamente ao sistema de referência que se caracteriza pela sua adequação formal à realidade. O critério da hierarquia não serve para exceções mas somente para ordenar valores. Assim, com base na experiência podemos justificar infringir os patamares escolhendo um nível inferior, serão caso da urgência.

Verificação moral (justificação): Seguidamente há que fazer uma verificação moral ou uma justificação para verificar a solidez da decisão a tomar. Contrasta-se o caso com a regra,

verifica-se a possibilidade de justificar uma exceção à regra no caso concreto e por fim contrasta-se a decisão tomada com o sistema de referência.

Ficaremos então prontos para tomar uma decisão.

Com esta explicação Diego Garcia pensa conseguir uma universalidade da ética num sentido formal e não material através de uma fundamentação racional e sem cair na falácia naturalista ou num racionalismo material.

Aplicando estes conceitos à clínica obtemos o processo deliberativo de Diego Garcia que adaptado por Nora se formula em 10 passos (Nora, 2015, pp. 114-123):

1) apresentação do caso pela pessoa responsável por tomar a decisão; 2) esclarecimento dos factos do caso; 3) identificação dos problemas éticos; 4) identificação do problema ético fundamental; 5) determinação dos valores em conflito; 6) identificação dos cursos extremos de ação; 7) identificação dos cursos intermédios; 8) análise do curso de ação óptimo; 9) decisão final; 10) aplicação das provas de consistência que consistem em:

Prova de legalidade. “Isso é uma decisão legal?”

Publicidade. “Você estaria preparado para defender publicamente o curso de ação escolhido?”

Teste de tempo. “Com mais algumas horas ou dias chegaria à mesma decisão?”

A decisão considerada mais prudente é a que responde positivamente às três questões.

### **Método Nijmegen**

Este método de deliberação é aplicável especificamente ao entorno clínico. Envolve um eticista profissional que sirva tanto como crítico como facilitador. Baseia-se num processo estruturado que tenta ter em consideração todos os aspetos do caso clínico. Diferencia-se em quatro passos:

- 1- O problema moral, que consiste no cerne do problema implícito.
- 2- O inventário dos fatos do caso: diagnóstico, tratamento, valores do utente e outras dimensões relevantes.
- 3- Avaliação do peso relativo das normas e valores em jogo que incluem questões de justiça, consentimento e responsabilidade profissional.
- 4- Elaboração de decisão com razões justificáveis e execução.

O processo permite uma deliberação de base multidisciplinar e em equipa motivando a decisão com base na racionalidade e no consenso (Kazeem, 2014, pp. 73-79).

### **“Four Box Method”**

Do inglês traduzido como método das quatro caixas, trata de um método simples cujo cerne consiste numa forma de apresentação dos dados disponíveis de modo a poder ser visualizado o problema como um todo. Isto permite verificar saliências morais e focar no que realmente interessa. Consiste na elaboração de uma tabela, quatro quadrículas ou espaços distintos que consistem em:

Caixa um: Indicações médicas: Prognóstico, problemas médicos atuais, tudo o relativo ao processo de doença.

Caixa dois: Preferências do Utente: Esta dedica-se aos desejos do paciente, as vontades e objetivos suas ou representadas por um familiar.

Caixa três: Qualidade de Vida: descrição sobre a capacidade do utente de poder usufruir de qualidade de vida no seu contexto, a maneira como a sua condição atual lhe permite experienciar uma determinada decisão e as perspetivas ou possibilidades de melhoria.

Caixa quatro: Contexto: Referências de aspeto religioso e cultural, especificidades familiares.

Seguidamente depois desta organização visual do contexto a informação está pronta para ser contrastada com o método principialista tentando encontrar uma maneira de encaixar os aspetos descritos nas caixas com os princípios de beneficência, não maleficência, justiça e autonomia (Jonsen, Sieger, & Winslade, 2005, pp. 1-9).

### **3.3- Robotização e deliberação**

A questão é que além do método de Diego Garcia, mais desenvolvido aqui por ser o mais representativo nos estudos considerados é que há outros, dos quais estão dois exemplos mas todos eles com complexidade pouco passível de colocação num algoritmo ou numa máquina. Ainda assim, como já referido antes, estas podem ajudar muito nas tomadas de decisão por permitem acelerar muitos destes pontos. Será possível colocar num algoritmo, numa máquina um destes métodos de deliberação assim como se pode traduzir em papel um algoritmo sob a forma de uma árvore de decisão?

Analisando as várias fases do processo de Diego Garcia desde o sistema de referência às premissas morais, podemos encontrar analogias com as leis da robótica de Asimov previamente descritas. Deixando de lado a complexidade da tradução informática de tais leis e a aplicação destas num sistema robótico, numa base de assunção que seria possível (assim como presumimos que as leis de Asimov também seriam aplicáveis) é impossível não ceder à tentação de pensar se tal não seria meio caminho para a criação da máquina como agente ético e moral.

O maior problema seria que ao traduzir os conceitos para normas estaríamos novamente a entrar em absolutismos de ideias, a uma normatização de algo que deve ser relativo e que o próprio autor critica noutros textos (para os quais remeto ao ponto 2 desta parte) e aqui admite mesmo que não será válido para exceções.

Como exemplo mais óbvio tomemos o sistema de referência “O homem é pessoa, tem dignidade e não preço, portanto o “homem não tem preço”.

Apesar de ser bonito dizer que a vida humana, que o homem não tem preço a verdade é que tem. E se por algum motivo estivermos a ler a frase de uma perspectiva condicionada pelo negativismo inculcado no termo depreciativo de “preço” podemos alterar esta palavra para o sinónimo de “valor”, assim podemos dizer que “a vida humana não tem preço mas tem valor”. A existência de um “valor” da vida humana permite logo interrogarmo-nos sobre as características desse valor, se é muito ou pouco valor. Certamente concordamos que o valor da vida humana não é infinito e mesmo que assim o considerássemos tal não invalidaria o termo comparativo com outros valores. De uma mesma maneira, na matemática infinito ( $\infty$ ) não é igual a infinito e uma coisa pode ser mais infinita que outra (Internet Encyclopedia of Philosophy, 2020).

Podemos dizer que entre dois números, por exemplo entre 0 e 1 existe uma infinidade de valores intermédios, um infinito ( $\infty$ ). Da mesma maneira entre 0 e 2 existe também uma infinidade de valores intermédios, um infinito ( $\infty$ ). Mas na realidade entre 0 e 2 existe o dobro da infinidade de valores que há entre 0 e 1 sendo que para cada número intermédio entre 0 e 1 existem dois números entre 0 e 2. Portanto apesar de  $2 \times \infty = \infty$  mas  $\infty \neq 2\infty$ . Falamos de algo similar a estar a usar os mesmos termos para nos referirmos a coisas diferentes. Atentemos nos exemplos para ilustrar esta equação:

-Uma vida humana não tem preço.

-A obra de *Da Vinci, Gioconda* não tem preço.

-> Perder a *Gioconda* é equivalente a perder uma vida humana.

-A vida humana não tem preço.

-A liberdade não tem preço.

-> A liberdade não justifica a perda de uma vida.

Ainda que seja difícil associar a vida humana a um valor concreto podemos falar no seu preço/valor em comparação.

-Uma vida humana vale menos que muitas vidas humanas.

Mesmo esta última afirmação não é certa, depende apenas da perspectiva pela qual a encaramos havendo exemplos de exceção em toda a parte, assim podemos estar errados se a aplicarmos num caso concreto como: a vida de uma criança vale menos que duas vidas de velhinhos em fim de vida. Portanto a vida tem um preço, e tem mesmo um preço monetário. Uma vida está associada a um custo e custo esse que sim, está associado a um valor monetário concreto e tal é refletido nos baremos e nas tabelas indenizatórias de avaliação funcional especificamente para perdas de braços, pernas, visão, funcionalidade potencial ou mesmo perda de vida humana que tem um valor consoante seja jovem, criança, adulto ou velho. Este valor não pode ser interpretado como o valor que representa todas as dimensões do que está a ser valorado, é apenas uma tentativa de restituição após uma causalidade e encara um valor dentro de um determinado contexto. Assim a perda de um filho por uma mãe apesar de irreparável (diferente de impagável), tem um valor monetário fixo, relativamente baixo dependendo da perspectiva e todos sabemos que a perda de um filho não pode ser compensada mesmo que dessemos a *Gioconda* a essa mãe, obra essa que não tem preço. No fundo, paga-se o que se pode pagar, o que se pode traduzir num valor monetário. O resto remedeia-se.

Este tema da tentativa de normativizar a incerteza foi abordado anteriormente na parte da clínica e da ética, pelo olhar crítico do mesmo autor, Diego Garcia. Aqui é usado como exemplo para enfatizar a ideia da dificuldade de algoritmização da ética pela incapacidade de a reduzir a termos fixos e absolutos, em suma, pela incapacidade de colocar um algoritmo a trabalhar na ambiguidade.

Estamos então a trabalhar um pouco dentro dos dilemas daquilo que se considera o pluralismo ético.

*“Most simply ethical pluralism holds that the values or goods legitimately pursued by human beings are plural, incompatible, and incommensurable”<sup>27</sup>.*

Há muitos valores que simplesmente não se podem reduzir ou ser descritos com um sistema absoluto de valores que se sobrepõem a tudo o resto, não podem ser hierarquizados, haverá sempre conflito entre eles e este conflito é dificilmente sanável por uma máquina.

Tal não implica menorizar o papel dos algoritmos e das máquinas na ética ao ponto de os excluir de uma decisão ética. No entanto a verdade é que esta minoração parece implicar a exclusão do algoritmo do cerne da decisão propriamente dita, não pode deliberar, não pode haver eticidade na escolha. No entanto ainda resta o ramo em que a máquina é usada de forma acessória na decisão, como parte dela. Fazendo o trabalho correspondente à parte moral da decisão dando uma vertente mais virada para incutir boas práticas e que tecnicamente possa permitir escolhas corretas, em suma, um otimizador moral.

#### **4- A deliberação como ponte para a decisão razoável**

A definição de decisão ética aristotélica implica que a decisão ética é a mais razoável e chega-se a tal através da ponderação, da deliberação. A análise dos textos anteriores orientam para a conclusão de que uma decisão não pode ser ética se não for ponderada, que ao atuarmos por reflexo não estamos a exercer ética porque não estamos a ter em conta o contexto próprio e isto é independente do resultado final ser benéfico ou não para o sujeito em questão uma vez que a decisão foi razoável, foi ponderada e dentro do contexto era a mais adequada e a intenção era realizar o bem.

Ora como as máquinas não podem deliberar, estão fora do cerne da decisão, não podem decidir.

E esta definição será para a ética clínica também ou somente para a ética de uma forma geral?

---

<sup>27</sup> Mitcham, 2005, p. 690

Esta associação entre a razoabilidade e a deliberação como moldes para a base ética é o argumento principal pelo qual se excluem as máquinas da decisão ética. Estas não podem deliberar.

Diz-nos que a decisão é tão mais ética, tão mais razoável quanto mais ou melhor ponderada seja. Isto claro, sem cair no extremismo de uma ponderação *ad aeternum* que adie a decisão indefinidamente sendo que, o “mais ou melhor ponderada” está associado ao número e qualidade de variáveis de contexto a ter em conta na ponderação e não somente o tempo que se pondera. O problema é que mesmo não caindo no extremismo do *ad aeternum*, o tempo de ponderação não pode ser considerado inócuo.

Esta vinculação absolutista da razoabilidade com a ponderação torna incongruente o conceito de ética aristotélica numa ética de intenções, na intenção de causar o bem na prática da ética clínica. Se não podemos ponderar então não podemos ter decisões éticas.

A partir de aqui surgem dúvidas sobre a verdadeira definição do conceito de ética na clínica e na continuidade do uso deste termo no texto uma vez que a estou a desvincular do dever de deliberação em prol da razoabilidade e da boa intenção<sup>28</sup>.

#### **4.1- A ética da demora**

Os dois últimos métodos exemplificados no apartado anterior, o Nijmegen e o *Four boxes* apesar de serem esquematizados e organizados, claramente necessitam um interlocutor humano capaz de entranhar o contexto da situação, que faça uma comunicação adequada com os intervenientes mas saltam rapidamente à ideia as inúmeras possibilidades que os algoritmos e a informática em geral podem fazer para a disponibilização, processamento e apresentação da informação. A esquematização do método de deliberação não tem como objetivo somente alterar o tipo de decisão a tomar, mas de a agilizar.

Aqui temos um pretexto para introduzir uma outra vertente da ética que está muitas vezes negligenciada na discussão da ética das máquinas e da definição de ética que é indissociável do contexto da clínica. Falamos do tempo.

---

<sup>28</sup> Há que recordar que Aristóteles não tem em conta o tempo, é filosófico, não fala em tempo cronometrado. O dilema exposto resulta apenas da tentativa de aplicação da filosofia à prática clínica.

No contexto clínico a ética está relacionada não somente com as características da decisão mas com o tempo que esta leva a tomar. Não podemos esquecer que as doenças não fazem pausas enquanto se espera por uma sentença. A progressão e atividade da doença, ou seja, o atraso na orientação decisória tem custos éticos consideráveis. Não é ético manter o utente a sofrer à espera de uma decisão.

Então qual o mais importante numa decisão ética? A razoabilidade ou a ponderação? É que num cenário onde o tempo de ponderação é escasso não podemos deixar de tomar decisões razoáveis, decisões estas que apesar de pouco ponderadas são mais éticas que a inação ou indecisão para aumentar os dados sobre o contexto do caso. Isto criaria um limbo, uma neutralidade que escuda a inação através de uma ausência de má intenção para ponderar enquanto o utente se deteriora e se perde a oportunidade de melhor servir os deveres de justiça, respeito e confiança para com o utente, que são a base da valoração ética. Falamos da necessidade de fazer uma “escolha” bem-intencionada. Não poderá esta “escolha” servir o campo da ética?

Há que salientar que a intervenção das máquinas nestes contextos não é meramente acessória pois nestes contextos clínicos qualquer ajuda que permita encurtar tempo de decisão permite poupar no sofrimento (princípios de beneficência e não maleficência) e limitar a pressão deste fator na autonomia do utente e além do mais, cortar em custos globais associados a demoras, adequando-se assim ao princípio de justiça.

#### **4.2- Paradoxo da ponderação em contexto crítico de aceleração temporal**

Na clínica a ponderação e o tempo são indissociáveis, coisa que não sucede ao considerar o plano filosófico somente. Apesar de muitas das decisões da prática clínica permitirem extensa avaliação e negociação, a demora num tratamento ou decisão em que há restrições de tempo ainda que não voluntária, acaba por infringir todos os princípios da bioética, ao ponto que especialmente em caso de urgência pode ser ético não ponderar ou pelo menos ponderar não ponderar.

Esta analogia parece colidir de frente com a noção de Aristóteles que define a ética de um comportamento sempre com base na ponderação mas o exemplo da urgência levanta questões sobre a quantidade ou qualidade da ponderação que não se podem ignorar. Teremos de considerar as decisões sob pressão temporal como não éticas? Não deveria a

ética estar presente aquando a tomada de decisões em contexto de pressão temporal? De que forma?

Não somente na urgência, o tempo adquire extrema importância no caso de patologias complexas e em casos de cuidados de fim de vida ou doença crónica e nos quais há extremos e “exceções” para os quais a ponderação passa a não coincidir com a decisão mais razoável. Quando falamos da prática médica, da clínica estamos sempre a falar em sofrimento sob alguma forma. Na valoração ética de uma decisão tem sempre de se equacionar o sofrimento com o tempo que lhe é proporcional.

Tomemos o exemplo do utente com doença grave que necessita uma reunião multidisciplinar para deliberação do curso a seguir e cujo tempo útil de realização da mesma talvez não seja o mais benéfico para o utente pois a sua doença está a progredir, aí talvez uma decisão menos ponderada possa ser a mais razoável. Durante este tempo o utente está a sofrer, física ou psicologicamente.

Surge então na clínica uma hierarquia que associa o tempo, a razoabilidade e a ponderação mas de uma forma assimétrica.

No plano filosófico, maior/melhor ponderação implica maior razoabilidade, de uma forma matemática isto traduz-se em:

Traduzindo numa equação:  $ax = by$

Sendo que “x” corresponde à deliberação e “y” corresponde à razoabilidade. Desta equação destaca-se a existência de proporcionalidade, o aumento do valor/qualidade da deliberação leva a um incremento da razoabilidade.

Relacionando pois a incerteza com qualquer uma destas variáveis (porque são proporcionais), a incerteza assume uma proporcionalidade inversa quer com a qualidade da deliberação quer com a razoabilidade da decisão:

Traduzindo numa equação:  $\frac{1}{x} = y$

Sendo que “x” corresponde à abrangência da deliberação (ou à razoabilidade visto serem proporcionais) e “y” corresponde à incerteza,

Assim:  $\frac{1}{\text{Deliberação}} = \text{Incerteza}$  ou  $\frac{1}{\text{Razoabilidade}} = \text{Incerteza}$

Aumentando o valor da deliberação vamos diminuindo o valor da incerteza sendo que o valor de “y” (incerteza) quando “x” (deliberação ou razoabilidade) tende para infinito é Zero. Ou seja, a melhor deliberação possível é aquela que considera todas as perspectivas possíveis, com incerteza zero e uma razoabilidade que se chama agora decisão correta, típica de uma decisão de máquina de circuito fechado.

O problema é que na tradução da filosofia para a clínica, o tempo limita a ponderação/deliberação deixando esta de ser tão proporcional à razoabilidade podendo uma decisão ser tanto mais razoável quanto menos deliberada. O objetivo da terapêutica sempre assenta na diminuição do sofrimento e este é diretamente proporcional ao tempo, assim como a qualidade da deliberação. Quanto mais deliberamos sobre um assunto melhor é a qualidade da deliberação, mais pontos de vista temos em conta, mas maior é o sofrimento do utente durante o tempo que se pondera.

$$\frac{\textit{Tempo}}{\textit{Deliberação}} = \textit{Razoabilidade}$$

Assim para manter um determinado valor de razoabilidade na decisão, sob pressão do tempo, teremos de encurtar o valor da deliberação para minimizar o sofrimento do utente. O mesmo é dizer que para um determinado tempo disponível há que variar a deliberação ou a razoabilidade para ter uma decisão aceitável pois o tempo é função da razoabilidade pela deliberação.

$$\textit{Tempo} = \textit{Razoabilidade} \times \textit{Deliberação}$$

Pior, aparta-se a proporcionalidade direta entre deliberação e razoabilidade pois uma decisão mais razoável pode ser aquela com maior incerteza. Como exemplo, um utente com um diagnóstico potencialmente mortal se não tratado, com dúvidas se uma lesão será resultado de uma infeção ou de um tumor. Para diminuir a incerteza necessitamos um exame que pode demorar duas semanas, tempo este que pode ser fatal no caso de se tratar de infeção. Então a decisão mais razoável será a de tratar empiricamente a infeção (mesmo que submetendo o utente ao risco de iatrogenia) mesmo com maior incerteza de diagnóstico. A decisão tomada é então aquela que não permitiu obter todos os dados para uma melhor deliberação do tratamento. Filosoficamente realizar-se-iam os exames e pouparíamos a necessidade de submeter o utente a tratamentos potencialmente desnecessários.

E assim:  $\frac{1}{\textit{Deliberação}} = \textit{Incerteza}$  é diferente de  $\frac{1}{\textit{Razoabilidade}} = \textit{Incerteza}$

As consequências da introdução do tempo como variável, ou seja, da aplicação da definição de ética filosófica à clínica é que deixa de ser razoável uma decisão muito ponderada e bem-intencionada (ou não mal intencionada) se o tempo para a tomada de decisão levar expetavelmente à progressão de uma doença de forma desmesurada que anule os benefícios da melhor decisão possível com a ponderação.

Traduzindo para um outro caso de clínica, podemos exemplificar que um utente de aparente idade avançada mas comorbilidades e estado prévio desconhecido entra na sala de emergência, apresenta alterações do estado de consciência graves por uma situação potencialmente facilmente reversível, no caso uma hipoglicémia (açúcar baixo no sangue). Nestes casos o dilema coloca-se quando ponderamos se será ético a reanimação ou a reversão de um estado que poderá ter já causado danos cerebrais a um utente já com pouca expectativa de vida e tornar o final da mesma muito sofrida ao utente. A verdade é que não sabemos o contexto do utente ou a qualidade de vida que aferia nem os danos já causados pelo tempo desconhecido da duração da hipoglicémia. A hipoglicémia trata-se de uma situação com gravidade e dá direito quando muito a equacionar opções alguns segundos, entre os quais podemos comprovar a existência ou não da mesma medindo o açúcar no sangue. Esta situação dá-nos algum tempo para um grau de deliberação que permita uma avaliação rápida com base em pouco fatores:

-Confirma-se a hipoglicémia?

-Apresenta algum indício que possa sugerir um estado do terminal por evolução de doença irreversível?

-Está em sofrimento?

E pouco mais do que isto, qualquer demora mais a procurar contexto estará a arrastar o utente para uma fase de maior irrecuperabilidade e de maior dano residual ao reverter a condição de hipoglicemia, ou seja dispomos de um tempo de ponderação do caso extremamente reduzido. Tal será ainda pior no caso de uma paragem cardio-respiratória ou seja na verdadeira urgência que implica atuação imediata. O tempo nunca é uma coisa inexistente, não é um zero absoluto mas é quase. Na questão da paragem cardio-respiratória temos somente o tempo que demora a verificação da ocorrência da paragem, o tempo que demora a verificar a ausência de circulação ou de contração cardíaca. Aí podemos especular, olhar bem para o utente em questão avaliar a idade, a presença de sinais de recuperabilidade ou que possam impelir a obstinação terapêutica (por exemplo o aspeto de ser um utente acamado anquilosado, com lesões de decúbito e provavelmente

sem vida de relação) mas pouco mais, e a dúvida é algo muito frequente que não se aclara com uma observação de meros segundos.

Em suma, qualquer ponderação acima do muito básico pode implicar a que haja danos irreversíveis neurológicos que por si inviabilizam qualquer tentativa de obter um comportamento ético na decisão. Então o razoável será consoante essa quase ausência de deliberação decidir logo e não ir consultar o processo ou a família para saber se não estaremos a infringir o princípio de não maleficência ou de justiça através da obstinação terapêutica. Ou seja, deve haver uma proporcionalidade entre o tempo disponível e o tempo de deliberação sendo que quando o tempo é zero e não podemos deliberar temos de fazer escolhas e a decisão de fazer uma escolha não pode simplesmente ser considerado não ético ou fora do âmbito da ética.

#### **4.3- No tempo da ética aristotélica à ética da compressão temporal acelerada**

Para não cair novamente em absolutismo de ideias, há que ter em conta que a ponderação como base para uma decisão ética na clínica não está condenada pela introdução do tempo na equação quando este é escasso. Há alguns álibis que esta teoria pode alegar para continuar vigente quando aplicada à clínica desde que com recurso a algum exercício de plástica.

Da dissecção da ética aristotélica ficamos com a impressão que a ponderação é essencial e é proporcional à razoabilidade independentemente do fator tempo entrar em equação.

Esta afirmação também pode ser verdade se não estivermos limitados a considerar ponderação somente à deliberação que sucede após o acontecimento que desencadeou o dilema e a decisão a tomar. Para isso podemos socorrer-nos da casuística, da previsibilidade. E quem melhor para manter uma casuística sempre à mão que o suporte tecnológico?

Assim apesar de vincular a ponderação e a ética às especificidades de um caso concreto, a realidade é que na clínica um caso é concreto mas não é totalmente único e dissociável de outros casos semelhantes. Apesar da especificidade inerente ao caso há certamente alguns pontos em comum com outros casos semelhantes ocorridos anteriormente, sobre os quais alguém já ponderou e elaborou normas ou guias de atuação. Já alguém levou todo o tempo que quis a ponderar o modo de atuação mais razoável para um caso que viria a suceder

posteriormente deixando apenas a interrogar a necessidade de deliberação sobre os aspetos mais únicos do caso por parte do médico e doente que o testemunham.

Assim podemos de novo encaixar a atitude menos ponderada novamente na definição aristotélica de ponderação e razoabilidade proporcionais uma vez que o tempo está menos em conta quando se está a equacionar modos de atuação para casos futuros.

## **5- A máquina como integrante da ética clínica**

Ante os limites das capacidades humanas reconhecidos e a existência desta variável tempo como impactante na qualidade da decisão, as máquinas são capazes de atuar nesta variável de extrema importância ética, o tempo. Uma máquina com a sua capacidade de computação permite a realização de milhões de cálculos de variáveis a partir de bases de dados de uma maneira impossível para a mais sagaz das mentes humanas o que por si só não implica que seja capaz de decidir.

Moor acha que sim, mesmo desconsiderando a possibilidade de a máquina ser um análogo de decisão humana, descreve que “há situações em que a atividade de um computador pode ser entendida como uma análise complexa de informação resultando na seleção de um caminho a tomar”, ou seja, uma decisão e não somente uma escolha (Moor, 1979, p. 217).

De qualquer maneira é de consenso que os recursos a protocolos, a algoritmos, a suporte informático, registos e casuística vieram estender e dilatar a qualidade da deliberação possível por um humano no tempo que é limitado. Podemos dizer que não podendo tomar uma decisão sem humanos, as máquinas em contextos específicos clínicos em que o tempo é figura de importância (e que são muitos) permitem que uma decisão seja mais ética do que sem elas, podendo levar ao extremo de dizer que recusar os algoritmos e as máquinas no tratamento dos utentes não é ético porque não é ético recusar algo que melhora a qualidade da deliberação segundo definição aristotélica.

## Conclusão

### Decisão do algoritmo e algoritmo na decisão.

#### Em que ficamos?

Uma decisão ética é aquela que adere a princípios mais globais e “universais” fora do estrito foro individual e terá então características de respeito, justiça, confiança e responsabilidade (entre outros)<sup>29</sup>, que serão os correspondentes ao sistema de referência proposto por Diego Garcia<sup>30</sup>. Estes permitem classificar uma atitude como boa ou má fora do simples campo da moralidade, além disso a decisão ética foi pensada ou deliberada e não mero alvo de uma escolha. Estas características estão presentes de uma forma mais relevante ainda quando falamos da ética e da clínica, da relação entre médico e paciente e da indissociação dos conceitos de ética e clínica na forma como surgiram e evoluíram. Ao processo de adequação de uma decisão a estas características ou a estes “princípios mais globais” chama-se prudência (do grego *phronesis*) que na decisão é expressa sob a forma de deliberação ponderada (Amaral, 2019, p. 149). Durante esta ponderação pesam-se alternativas e eliminam-se aquelas não éticas ou que menos representam estes valores, dando-se primazia à alternativa melhor ou menos má, portanto, à mais ética. Este tipo de decisão demarca-se dos aspetos puramente morais de um código mais fixo de carácter mais individual que definem um ato como moral ou imoral simplesmente, que são mais dependentes de uma escolha e não de uma verdadeira decisão.

Decisão define-se em função de uma circunstância, do contexto e não de um código.

Relativamente ao núcleo do problema em questão, sobre o papel dos algoritmos na decisão, há que recordar que este papel se pode apresentar em duas medidas. Sendo o algoritmo como mera parte da decisão ou sendo o algoritmo o cerne da decisão.

---

<sup>29</sup> Diferenciação entre ética e moral, contexto e evolução, *vide supra* p. 26

<sup>30</sup> Esquematização/perspetivas da deliberação, *vide supra*, p. 44

## **1- A máquina como parte da decisão ética**

Relativamente ao algoritmo como parte da decisão, aparenta ser óbvio que deve na maioria dos casos formar parte da mesma, num grau mais ou menos importante desta, conforme ajude a expandir ou delimitar o contexto das variáveis intervenientes nas opções. Se a análise for feita estritamente relativamente à origem das palavras verificamos que no sentido aristotélico da palavra ética original está a vinculação à ponderação, e aí o algoritmo tem um papel mais acessório, fornece o código, é a parte moral (Moral = conjunto de instruções para determinar o comportamento de um programa fechado, escolhas deontológicas e opções técnicas) e não a parte ética pois a deliberação e o cerne da decisão terão sempre de partir do usuário humano, o único com o “éthos” que forma o seu carácter capaz de orientar a decisão ao contexto. O algoritmo e as máquinas com o seu acesso a capacidade de processamento de dados muito superior à humana, são o suporte ideal para o auxílio do decisor na vertente mais técnica, no código, na parte moral.

Neste contexto, as máquinas assumem um papel de suma importância. As melhorias que aportam à qualidade da decisão são vitais e dificilmente alcançáveis por outros meios pelo que a recusa da entrada das mesmas no processo de decisão será até eticamente questionável.

## **2- A máquina no cerne da decisão ética**

Relativamente ao outro braço da questão central, sobre a possibilidade da máquina poder ser a parte mais importante do processo decisório, pela tomada de uma verdadeira decisão ética, o tema é mais complicado. O autor apresenta duas soluções para este cenário ser possível. A primeira no campo mais da teoria e de uma perspectiva mais futura implica a evolução da máquina de um simples sistema fechado para algo mais próximo ao humano, para uma entidade de sistema aberto capaz de aprendizagem, da tomada de consciência do problema, capaz de uma abrangência que permita a adaptação de uma decisão a um contexto não programado previamente, uma máquina capaz de lidar com a ambiguidade, com o inesperado, que esteja dotada de uma capacidade de podermos falar nela em termos de individualidade e responsabilidade. A criação de máquina deste género poderia existir tanto como máquina de simulação perfeita da mente e comportamento humano como por uma evolução de tipo convergente através de algoritmos e computadores com desenvolvimento de tipo conexionista, autónomo e como tal de difícil previsibilidade na interpretação própria destes conceitos de respeito, justiça e confiabilidade que serão a

base da valoração entre ético e não ético. Com estas características quasi-humanas então não haveria dúvidas que uma decisão ética poderia ser tomada por este ente, se um paciente é desconhecido então mais facilmente a “máquina” cria um contexto mais amplo por análises de bases de dados que o médico, especialmente na emergência.

Num patamar menos consensual ainda mas ao alcance das máquinas atuais, partilho parte da visão de Moor sobre considerar que as máquinas conseguem decidir. Já se essa decisão é ética ou não, trata-se apenas de uma questão da comparação sobre se essa decisão é mais ou menos ética que a realizada por um humano e tudo isso será uma contenda de contexto apenas. Assim podemos considerar hipoteticamente que uma máquina poderia decidir mais eticamente um tratamento para uma infecção a um habitante de Lisboa do que um aborígine (tome-se como alguém completamente fora do contexto), simplesmente por a máquina ser capaz de se adaptar melhor ao contexto do doente e assim servir melhor os interesses deste.

A outra hipótese discutida para a máquina poder ser o cerne da decisão ética aproveita-se também um pouco do malabarismo de contextos. Consiste no alargamento da definição deste cerne pela aplicação específica da decisão na clínica. O algoritmo é o cerne da decisão e indispensável a esta, mas não sozinho.

Todas as decisões críticas devem ser éticas, independentemente de ser emergentes ou não e como a clínica tem a variável tempo em jogo (que é diretamente proporcional à qualidade da deliberação) uma decisão deve poder ser ética ainda que não seja deliberada porque não existe tempo para tal. Esta situação cria obviamente um paradoxo porque se a ponderação é parte da definição de decisão, de ética então não pode haver uma decisão ética não pensada. A junção dos termos “escolha ética” não deveria ter sentido mas pode ter (Singer, 2008, p. 59). No caso de não haver contatos relacionais de um paciente, a máquina ou um algoritmo podem talvez ajustar-se melhor ao contexto por análise de uma base de dados, se essa análise permitir uma melhor adequação aos desejos do paciente que a decisão em tempo útil que o médico teria então poderíamos deixar a decisão/escolha para o algoritmo.

Não passa pela cabeça de ninguém que numa sala de emergência o efetuar a escolha de reanimar alguém ou não em paragem cardio-respiratória possa não ser ético simplesmente porque não houve tempo suficiente para ponderar. Não se pode pensar que a ética está excluída das salas de emergência. A solução para o paradoxo está na introdução da ponderação na equação de maneira não convencional, admitir que se faz um escolha no momento de reanimar ou não e se segue simplesmente um algoritmo ou

protocolo. Algoritmo ou protocolo esse que foi elaborado por alguém que ponderou e equacionou modos de atuação futuros baseados em casos que seriam de características similares aos da situação (porque na medicina todos os casos são únicos mas nenhum caso é realmente único), deixando apenas nuances para diferenças de contexto na globalidade da decisão. Assim se entrarmos na equação com essa pré-ponderação ou ponderação na elaboração do algoritmo podemos transformar a escolha em decisão alargando o cerne do que consideramos decisão para incluir o algoritmo que acerta com ele essa pré-ponderação, pois sem ele falamos somente de uma simples escolha e tal não é ético.

Tomemos o exemplo: O utente entra na sala de emergência em paragem cardio-respiratória, não há tempo a perder a procurar na história do utente para saber de contexto, para ponderar reanimar ou não, nesse momento entra em jogo o algoritmo de reanimação que com base nas informações a que consegue aceder decide em milissegundos se ativa os aparelhos para iniciar a reanimação ou se esta é claramente infrutífera no salvamento do utente. Em caso positivo os aparelhos começam a fazer compressões torácicas e a administrar fármacos, algoritmo esse em que alguém pensou e ponderou com boas intenções achando que seria mais ético ante a incerteza insistir na reanimação com base numa informação de contexto similar à disponível.

Acabamos assim por obrigatoriamente incluir no cerne da decisão o autor do algoritmo, que decidiu e ponderou na elaboração deste, o algoritmo e o médico na sala de emergência, todos indispensáveis para tornar a escolha numa decisão sendo que ao médico na sala de emergência cabe também ser parte deste algoritmo como supervisor, com uma análise do caso durante o procedimento do computador e à medida que vai ganhando informações responde a uma questão de forma contínua: “continua-se ou suspende-se o algoritmo de tratamento?”

## Bibliografia

- Amaral, A. (2019). Da (im)pertinente diferenciação entre ética e moral na filosofia de Aristóteles. In *PRAKTON. Discursividades da Acção em Aristóteles* (pp. 121-132). Covilhã: UBI: Editora LabCom.IFP.
- Amaral, A. (2019). Formas de vida e escolha deliberada na ética teleológica de Aristóteles. In *PRAKTON Discursividades da Acção em Aristóteles* (p. 149). Covilhã: UBI: Editora LabCom.IFP.
- American Automobile association. (October de 2019). *Automatic Emergency Braking With Pedestrian Detection*. Obtido de <https://www.aaa.com/AAA/common/aar/files/Research-Report-Pedestrian-Detection.pdf>
- Aristóteles, *Ética a Nicomaco*, trad. A. Caeiro. Lisboa: Quetzal Editores, 2004.
- Asimov, I. (2004). Runaround. In *I, robot* (p. 37). New York: Bantam Books.
- Bauerlein, M. (2009). In *The Dumbest Generation: How the Digital Age Stupefies Young Americans and Jeopardizes Our Future* (p. 26). New York: Jeremy P. Tarcher/Penguin.
- Beauchamp, T. L., & Childress, J. F. (1994). In *Principles of Medical Ethics* (pp. 99-102). New York: Oxford University Press.
- Bellamy, R. F. (2003). Theories of medical ethics: The philosophical structure. In *Military medical Ethics* (pp. 23-54). Office of the Surgeon General, department of the USA army. Obtido de <https://www.cs.amedd.army.mil/borden/bookDetail.aspx?ID=106cb6b1-3327-47f3-a350-2fb57beb928&pageTitle=Military%20Medical%20Ethics%20Vol.%201>
- Bioeticawiki contributors. (2020). *Principialismo, revisión de Diego Garcia*. (Bioeticawiki, Editor) Obtido 4 15, 2020, de Bioeticawiki.com: [https://www.bioeticawiki.com/Principialismo\\_\(Revisi%C3%B3n\\_de\\_Diego\\_Gracia\)](https://www.bioeticawiki.com/Principialismo_(Revisi%C3%B3n_de_Diego_Gracia))
- Bluteau, R. (1789). *Dicionário da Língua Portuguesa*. (A. M. Silva, Ed.) Lisboa: Officina de Simão Thaddeo Ferreira. Obtido de <http://purl.pt/29264/3/>

- Bork, J. (2015). "Critical Programming: Toward a Philosophy of Computing". *Electronic Theses and Dissertations*. 1356, University of Central Florida, pp. 4-7.
- Bory, P. (2019). Deep new: The shifting narratives of artificial intelligence from Deep Blue to AlphaGo. *The International Journal of Research into New Media Technologies*, 25(4), 627-642.
- Bostrom, N., & Eliezer, Y. (2014). Forthcoming. "The Ethics of Artificial Intelligence.". In K. Frankish, & W. Ramsey (Eds.), *Cambridge Handbook of Artificial Intelligence* (p. 6). New York: Cambridge University Press.
- Brey, P., & Søraker, J. (2009). Philosophy of Computing and Information Technology' Philosophy of Technology and Engineering Sciences. In A. Meijers (Ed.), *Handbook for Philosophy of Science* (Vol. 14, pp. 28-30, 41). Elsevier.
- Bynum, W. F., & Porter, R. (1993). In *Companion Encyclopedia of the History of Medicine* (pp. 56-62, 1554). London: Routledge.
- Cabral, I. V., Cardoso, T. S., & Pena, R. C. (2012). O racionalismo socrático. *Contribuciones a las Ciencias Sociales*. Obtido de [www.eumed.net/rev/cccss/22/](http://www.eumed.net/rev/cccss/22/)
- Carrick, P. (2001). In *Medical ethics in the ancient world* (pp. 14-21). Washington D.C.: Georgetown University Press.
- Chalmers, D. J. (1996). In *The Conscious Mind: In Search of a Fundamental Theory*. *Philosophy of Mind Series*. (p. 95). New York: Oxford University Press.
- Char, D. S., & Shah, N. H. (2018). Implementing Machine Learning in Health Care — Addressing Ethical Challenges. *New England Journal of Medicine*, 378(11), 981-983.
- Dicionário etimológico. (2020, 4 15). *Origem da palavra deliberar*. Obtido de Dicionário etimológico: <https://www.dicionarioetimologico.com.br/deliberar/>
- European Commission. (2019). ETHICS GUIDELINES FOR TRUSTWORTHY AI. European Commission.
- Garcia, D. (1998). In *Bioética Clínica* (pp. 15-16). Bogotá, Colombia: Editorial el Buho LTDA.

- García, D. (2001). La deliberación moral: el método de la ética clínica. *Med Clin*, 117, pp. 18-23.
- Garrett, J., Jotterand, F., & Ralston, D. (2013). The Development of Bioethics in the United States. *Philosophy and Medicine*, 115, 42-50.
- Heiser, J. (2018, 03 12). Artificial Stupidity: How are Algorithms Changing the Way Art is Seen? *Frieze*(194). Obtido 4 25, 2020, de <https://frieze.com/article/how-are-algorithms-changing-way-art-seen?language=de>
- Internet Encyclopedia of Philosophy "Ethics"* . (2020, 4 30). Obtido de <https://www.iep.utm.edu/ethics/>
- Internet Encyclopedia of Philosophy. (2020, 4 30). "*Zeno's Paradoxes, a Paradoxes of motion*". Obtido de <https://www.iep.utm.edu/zeno-par/>
- Jonsen, A., Siegel, M., & Winslade, W. (2005). In *Clinical Ethics: A Practical Approach to Ethical Decisions in Clinical Medicine*. (8th ed., pp. 1-9). New York: McGraw-Hills. Obtido de [https://depts.washington.edu/lend/pdfs/4\\_box\\_model.revised\\_Fitzpatrick\\_2017.pdf](https://depts.washington.edu/lend/pdfs/4_box_model.revised_Fitzpatrick_2017.pdf)
- Kant, M. (1996). In *Fundamentación de la metafísica de las costumbres* (p. 117). Barcelona: Ariel.
- Kazeem, F. A. (2014). The Nijmegen Method of Case Deliberation and Clinical Decision in a Multicultural Society. *Bangladesh Journal of Bioethics*, 5(2), 73-79. doi:<https://doi.org/10.3329/bioethics.v5i2.19618>
- Köhler, W. (2015). In *The Task of Gestalt Psychology* (p. 10). Princeton: Princeton University press.
- Lolas, F. (2008). "Bioethics and animal research: A personal perspective and a note on the contribution of Fritz Jahr". *Biological Research (Santiago)*, 41(1), pp. 119-123. doi:10.4067/S0716-97602008000100013.
- Merriam-Webster. (2020, April). *Algorithm*. In *Merriam-Webster.com dictionary*. Obtido de <https://www.merriam-webster.com/dictionary/algorithm>

- Mitcham, C. (2005). In *Encyclopedia of Science, Technology, and Ethics*. (Vols. 1-4, pp. 89, 107, 146, 690, 700-710, 896-898). Thomson Gale.
- Moor, J. H. (1979). Are There Decisions Computers Should Never Make? *Nature and System*, 1, pp. 217-229.
- Moor, J. H. (1985). What is Computer Ethics? *Metaphilosophy*, 16, 266-275.
- Moor, J. H. (2006). The Nature, Importance, and Difficulty of Machine Ethics. *IEEE Computer Society*, pp. 18-21.
- Nora, C. R. (2015). Deliberação ética em saúde: revisão integrativa da literatura. *Revista Bioética*, 23(1), pp. 114-123. doi:10.1590/1983-80422015231052
- Obermeyer, Z. e. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366, pp. 447-453.
- Pegoraro, O. A. (2002). In *Ética e Bioética: da subsistência à existência* (pp. 75-107). Petrópolis: Editora Vozes.
- Porto Editora. (2020, 04 20). *Dicionário infopédia da Língua Portuguesa*. (Porto Editora) Obtido de <https://www.infopedia.pt/dicionarios/lingua-portuguesa/>
- Potter, V. R. (1970). Bioethics, the Science of Survival. *Perspectives in Biology and Medicine*, 14(1), pp. 127-153. doi:10.1353/pbm.1970.0015
- Priberam. (2020, 16 2). "escolher"; "decidir"; in *Dicionário Priberam da Língua Portuguesa*. Obtido de <https://dicionario.priberam.org/escolher>
- Pylyshyn, Z. W. (1984). In *Computation and Cognition, Toward a Foundation for Cognitive Science* (pp. xii, 49-59). MIT Press.
- Ricoeur, P. (2010). Os três níveis de juízo médico. (J. m. rosa, Ed.) *Colecção: textos clássicos LusoSofia*, 1-22.
- Ross, W. (1930). In *The right and the good* (pp. 19-36). Oxford: Clarendon.
- Rotenberg, M. (2016, October 19). Privacy in the Modern Age: The Search for Solutions. *EPIC 38th ICDPPC Marrakech, Morocco*.
- Santos, J. M. (2012). In *Introdução à ética* (pp. 62-66). Lisboa: Documenta.

- Sass, H.-M. (2007). Fritz Jahr's 1927 concept of bioethics. *Kennedy Inst Ethics J*, 17(4), pp. 279-295.
- Searle, J. (1980). Minds, Brains and Programs. . *Behavioural and Brain Sciences*, 3(3), pp. 417-424.
- Shapiro, S. C. (1995). Computationalism. *Minds and Machines*, 5(4), 467-487.
- Shin, Y. (2019). How does an autonomous car work? Not so great. *The Washington Post, US*. Obtido 9 2019, de <https://www.washingtonpost.com/graphics/2019/business/how-does-an-autonomous-car-work/>
- Silva, A. (2016). A Relação Médico Doente: uma revisão da literatura. *Faculdade de Medicina Da universidade de Coimbra, Coimbra*, pp. 3-10.
- Silveira, D. (2000). As Virtudes em Aristóteles. *Revista de Ciências Humanas*, 1(1), 2-10.
- Simon, H. A. (1965). In *The Shape of Automation for Men and Management* (p. 34). Harper & Row.
- Singer, P. A. (2008). In *The Cambridge Textbook of Bioethics* (pp. 58-59, 243, 314, 341). Cambridge: Cambridge University Press.
- Streck, L. L. (2013). In *O que é isto – decido conforme minha consciência?* (pp. 106-110). Porto Alegre: Livraria do Advogado Editora Ltda.
- Sullins, J. P. (2006). When is a Robot a Moral Agent? *International Review of Information Ethics*, 6, pp. 24-30.
- Thomasma, D. E. (1978). Training in Medical Ethics: An Ethical Workup. *Forum in Medicine*, 1, pp. 33-36.
- Weizenbaum, J. (1976). In *Computer Power and Human Reason: From Judgment to Calculation* (p. 227). New York: W. H. Freeman and company.