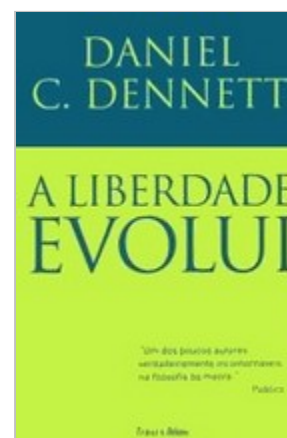


19 de Outubro de 2005 | Metafísica e lógica filosófica

A Liberdade Evolui, de Daniel C. Dennett

Tradução de Jorge Beleza

Lisboa: Temas e Debates, 2005, 365 pp.



A liberdade evolui. E o determinismo?

Foi finalmente editada em Portugal a tradução de *Freedom Evolves*, de Daniel Dennett, publicado no original em língua inglesa em 2003. A disponibilização entre nós de *A Liberdade Evolui* merece ser assinalada, que é o que pretendemos fazer com este texto.

A tese geral do livro é perfeitamente compreensível (e até esperada) dado o empenhamento do autor em sublinhar a importância da evolução natural na compreensão da forma como as coisas chegaram a ser o que são no nosso mundo:

"o livre-arbítrio é real, mas não é um aspecto preexistente da nossa existência, como a lei da gravidade. E também não é o que a tradição declara que é: um poder divino que dispensa a pessoa do tecido causal do mundo físico. Trata-se de uma criação da actividade e das crenças humanas que evoluiu, e é tão real quanto outras criações humanas (...)" (p. 28).

Temos, pois, que a tese central está toda no título do livro: a liberdade evolui. Uma parte importante da obra é dedicada a recensear várias tentativas para conceber como a evolução poderia ter produzido aquilo a que chamamos liberdade. Na verdade, não parece que Dennett esteja particularmente empenhado em fazer valer alguma dessas teorias em particular - estando, isso sim, mais interessado em mostrar que o padrão geral das explicações naturalistas é evolucionista também neste particular objecto.

Embora partilhemos essa ideia geral, esta obra de Dennett não se explica somente pelo que parece ser a sua tese central. É que, se essa tese é sobre a forma como a liberdade chegou a ser o que é, por via evolutiva, é preciso saber de que falamos quando falamos de liberdade. Que liberdade é essa que evolui? Neste livro Dennett procura construir uma resposta a essa questão que seja compatível com o determinismo. Essa opção terá, claro, um efeito decisivo sobre a própria noção de liberdade apresentada e defendida. Num certo sentido, o determinismo passa a ser, por essa via, a verdadeira questão desta obra. O que resultaria da eventual impugnação do argumento compatibilista, que sustenta a compatibilidade entre liberdade e determinismo? Resultaria uma outra pergunta, "virtual", sem resposta neste texto, assim formulada: *que liberdade_{não-determinista} é essa que evolui?*

Dennett adopta nas suas obras uma posição acerca da filosofia e das suas relações com as ciências que é reafirmada mais uma vez:

"as investigações filosóficas não são superiores ou mais importantes do que as investigações das ciências naturais, mas (...) são equiparadas a estas tentativas de procura da verdade (...)" (p. 29).

Esta é a forma como a tradução portuguesa fixa a declaração de naturalismo de Dennett. Talvez o original em inglês sublinhe mais a *parceria* entre a filosofia e as ciências do que a sua *equiparação* - mas, no que toca à filosofia, a higiene fundamental está lá: não lhe cabe uma autoridade particular,

competindo-lhe antes (o que nos parece muito mais exaltante) partilhar o esforço. Esta linha inspira a Dennett uma viagem constante ao mundo das ciências, onde colhe alguns dos materiais que fazem a beleza dos seus recursos argumentativos. Ilustração clara desse processo é a utilização recorrente de exemplos das ciências do artificial. Esse recurso, sempre fascinante, comporta, contudo, alguns riscos se deixarmos em insuficiente esclarecimento os pressupostos dos próprios dispositivos computacionais ou técnicos invocados. É assim que o recorrente investimento em exemplificação originada em ciências do artificial oculta, por vezes, tanto quanto mostra. Mais uma vez, neste livro, pode ser esse o caso.

Postas assim as coisas, o que nos propomos, neste texto, é (i) esclarecer de que fala Dennett quando fala de liberdade ou de livre-arbítrio - escolhendo para isso a via de (ii) fazer alguma luz sobre o significado dos exemplos que o autor mobiliza às ciências do artificial para ilustrar e apoiar os seus argumentos.

Dennett usa uma definição de determinismo (de Van Inwagen) e propõe-se defender três teses que contrariam o que considera serem três erros comuns em questões de liberdade e determinismo (p. 39). A definição é a seguinte: "a cada instante só há exactamente um futuro fisicamente possível". As três teses são: (1) o determinismo não implica inevitabilidade; (2) o indeterminismo não nos dá liberdade; (3) um mundo determinista não implica que as nossas opções sejam aparentes em vez de reais. O que procuraremos fazer, no caminho do nosso objectivo, é apreciar essas três teses e os argumentos que as servem ao longo da obra.

Chegaremos, desse modo, a um esclarecimento dos pressupostos do determinismo do autor, aos quais (para terminar) oporemos algumas notas dissonantes.

1. Determinismo

A definição de determinismo que Dennett toma para orientar o seu argumento é a seguinte: "a cada instante só há exactamente um futuro fisicamente possível" (p. 39). Temos de começar por considerar um pouco mais de perto essa tese determinista. Isso é necessário, porque o autor introduz certos dispositivos retóricos que tendem a descaracterizar, na aparência, a dita tese. O mais cândido desses dispositivos consiste em incluir o "demónio de Laplace" na secção dedicada às simplificações excessivas, quando nada de substantivo daí decorre para o argumento.

De facto, em certas passagens Dennett parece aperceber-se da ingenuidade do demónio de Laplace. É o caso da sua afirmação de que o determinismo é uma doutrina acerca da suficiência: se S_0 é uma descrição completa detalhada do estado do universo no momento t_0 e S_1 é uma descrição igualmente detalhada do estado do universo num momento posterior t_1 , S_0 é condição suficiente de S_1 em qualquer mundo físico. O determinismo não é uma doutrina acerca das condições necessárias para produzir S_1 . Por exemplo, de acordo com o determinismo, se S_0 descrever o estado do universo um segundo após o Big Bang, S_0 é causalmente suficiente para produzir o assassinio de J.F. Kennedy em 1963, mas não se pode dizer que S_0 causou tal acontecimento (pp. 99-100). Cabe notar, entretanto, que esta noção de determinismo não corresponde à tese determinista que Dennett anunciou ir guiar o seu argumento. Este "determinismo limitado à suficiência" diz apenas uma coisa simples: "o que aconteceu no mundo material podia ter acontecido no mundo material" - e, mesmo isso, só é dito retrospectivamente. Se entendermos útil usar a linguagem dos mundos possíveis, poderemos ainda acrescentar: "e o que aconteceu neste mundo podia ter acontecido em tais e tais mundos possíveis". É claro que Dennett pode estar aqui a visar outro tipo de alvo, por exemplo, aqueles que defendem que há intervenção divina na história ou que existem almas ontologicamente distintas da matéria e providas de eficácia causal sobre o

mundo físico. Não julgamos útil chamar essas hipóteses a qualquer tipo de explicação científica ou filosófica, mas também não nos parece que seja preciso ser determinista para tomar a decisão metodológica de descartar tais hipóteses. O que aqui interessa é que este "determinismo da suficiência" não é o determinismo da definição adoptada por Dennett. Nessa definição, se o determinismo é verdadeiro, o assassinio de J.F. Kennedy não podia ter deixado de acontecer, quando, onde e como aconteceu. É que se "a cada instante só há exactamente um futuro fisicamente possível" e se sempre assim foi (isto é, se o universo sempre foi determinista), então só há um percurso possível para este universo desde as condições do primeiro instante e as leis que então vigoravam. É isso que leva o matemático René Thom, num célebre manifesto determinista, a atacar as teses de Jacques Monod em *Le Hasard et la Nécessité* nos seguintes termos: "uma vez a Terra constituída nas condições climáticas e químicas da época, a vida e o pensamento deviam necessariamente nascer, mesmo em presença de "pequenas perturbações" desse ambiente" (Thom 1990:68).

Na verdade, Dennett dá ainda outros sinais de querer agilizar a sua tese determinista. Socorrendo-se de Whitehead ("a vasta independência causal de ocasiões contemporâneas é o que preserva a margem de manobra no Universo"), Dennett escreve mesmo que alguns eventos não têm qualquer causa (p. 100) - mas, infelizmente, a explicação que dá para isso é no mínimo obscura. Os exemplos fornecidos (pp. 100-103) assentam numa confusão grosseira entre a dificuldade de atribuir uma causa única a certos acontecimentos e o putativo carácter incausado desses mesmos acontecimentos. Para usar um dos seus exemplos, é certo que a I Guerra Mundial teve um conjunto complexo de causas, não tendo tido como causa única e separada o assassinio do Arquiduque Francisco Fernando, herdeiro da dupla coroa da Áustria e da Hungria, no dia 28 de Junho de 1914 em Sarajevo - mas isso não autoriza a conclusão de que I Guerra Mundial não teve qualquer causa.

O ponto de fuga de Dennett é o mesmo do demónio de Laplace: a ignorância. O autor afirma que "vivemos num mundo que é subjectivamente aberto" (p. 108), mas isso tem uma explicação: "Cada utilizador de informação finito tem um horizonte epistémico; não sabe tudo sobre o mundo que habita e esta ignorância não evitável assegura que tenha um futuro subjectivamente aberto" (p. 107). O nosso futuro está fechado, fixado, nós é que não sabemos como. A isso Dennett chama "suspense", mas estamos como um espectador a assistir a um filme de suspense: nada está em aberto, apenas o nosso conhecimento; a nossa sensação de liberdade é pura ignorância. Não vemos em que é que isso deva ser motivo de regozijo, mas concordamos com Dennett em não deixar de esclarecer cientificamente o que quer que seja por receio das consequências de compreender melhor quem somos e em que mundo estamos.

Onde a formulação de Dennett é mais sofisticada do que a do demónio de Laplace é ao focar-se na pluralidade das descrições possíveis de um mesmo estado do mundo (pp. 83-84). Dizer que o determinismo é verdadeiro é dizer que se quaisquer dois mundos partilham exactamente qualquer descrição dos seus estados no presente, esses dois mundos partilham todas as descrições dos seus estados subsequentes. Só que, mesmo com conhecimento perfeito e completo do estado do universo debaixo de uma determinada descrição, certos aspectos que podem de um modo ou de outro ser relevantes podem não ser acessíveis nessa descrição. De qualquer modo, o problema continua a ser o mesmo: o futuro só nos escapa a partir do presente porque não sabemos o suficiente acerca das leis físicas e do estado do mundo.

A tese determinista que Dennett propõe para orientar o seu argumento implica que o universo só podia ter seguido um percurso, exactamente aquele que seguiu, e nenhum outro. Nenhuma estratégia textual

lateral deve prejudicar a compreensão desse ponto. A tese da compatibilidade entre liberdade e determinismo (que passaremos a apreciar nas próximas secções) só pode ser devidamente apreciada se mantivermos clara a noção de determinismo que guia o argumento.

2. Compatibilismo

Como vimos, o compatibilismo de Dennett desdobra-se em três teses. Apreciemos cada uma delas à vez.

Determinismo e inevitabilidade. A primeira das teses que Dennett pretende defender é a seguinte: o determinismo não implica inevitabilidade. O argumento para essa tese é ilustrado com um dos dispositivos das ciências do artificial que aparecem recorrentemente nos argumentos do autor: o "jogo da vida".

O "jogo da vida", concebido por John Hoston Conway em 1970, é um autómato celular (AC), uma ferramenta típica da investigação em Vida Artificial (A melhor versão do Jogo da Vida (para ambiente Windows) que conheço pode ser obtida em linha [aqui](#)). Um AC é constituído por uma rede de células (um universo). Damos agora um exemplo de 3 linhas por 5 colunas. Cada célula pode estar activa ou inactiva. Neste exemplo, as células 22, 23 e 24 estão activas e todas as demais estão inactivas. Para o comportamento de cada célula interessa o seu próprio estado e o estado das suas vizinhas imediatas (na vertical, na horizontal e na diagonal).

11	12	13	14	15
21	22	23	24	25
31	32	33	34	35

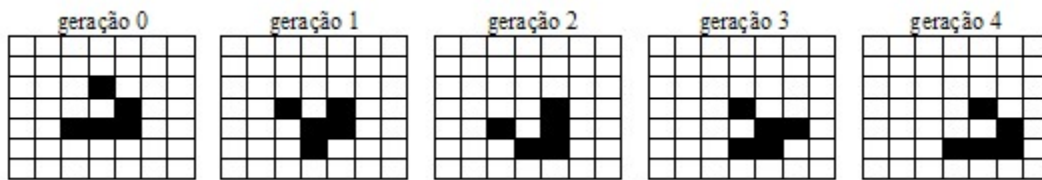
O "Vida" tem 3 regras: (i) uma célula inactiva que tenha como vizinhos 3 células activas, torna-se activa; (ii) uma célula activa que tenha como vizinhos 2 ou 3 células activas, mantém-se activa; (iii) em qualquer outro caso, uma célula torna-se ou permanece inactiva. Uma interpretação possível destas regras é, em coerência com a designação de "jogo da vida", a seguinte: para que haja nascimentos, tem de haver uma certa densidade populacional; as células não sobrevivem em isolamento extremo; uma densidade populacional demasiado forte não permite a sobrevivência. De qualquer modo, a interpretação não faz qualquer diferença ao jogo, embora possa fazer diferença à nossa maneira de olhar para ele.

O "Vida" funciona por etapas (gerações), sendo que em cada geração se aplicam todas as regras a cada célula. Calcula-se assim para cada célula o seu estado na geração seguinte e, quando esse cálculo está completo, procede-se à transição. Vamos aplicar esse procedimento ao exemplo apresentado acima e teremos a seguinte "geração 2" (em cada célula, agora, indica-se o número de vizinhos activos que ela tinha na geração 1):

1	2	3	2	1
1	1	2	1	1
1	2	3	2	1

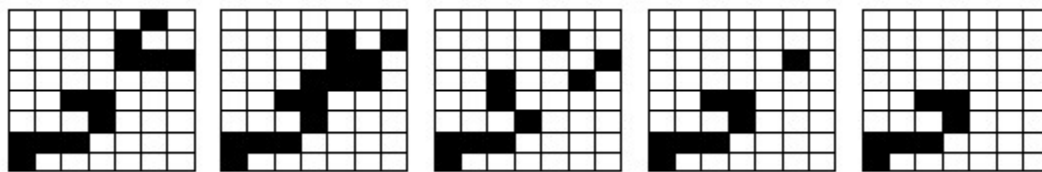
Tal como ilustrado neste caso, um AC tem três propriedades fundamentais: paralelismo: todas as células mudam de estado simultaneamente e independentemente umas das outras; localismo: a mudança de estado de uma célula depende apenas do seu estado inicial e da sua vizinhança imediata; homogeneidade: as leis são universais (comuns a todo o espaço do AC).

Como exemplo da dinâmica do "Vida", apresentam-se a seguir as primeiras gerações de um padrão, chamado "planador", que funciona precisamente segundo as regras anteriormente descritas.



O "planador" volta à forma original ao fim de quatro gerações, deslocando-se no "universo" - e assim continuará geração após geração.

Um outro tipo interessante de padrão é aquele em que uma configuração "invade" outra, o que pode ter várias consequências, uma delas sendo o desaparecimento de uma das configurações e a permanência de outra (aquela é "comida" por esta). É o caso no exemplo seguinte.



Vejamos então agora o uso que Dennett faz do jogo da vida. As regras do jogo e o estado inicial (quais células estão activas e quais estão inactivas) definem a física deste universo. Uma descrição física do último exemplo apresentado acima seria entediante: algo do género de uma longa lista em que, para cada geração, seria indicado o estado de cada célula. Se tivéssemos, em vez de um espaço de 7×8 células, um espaço de, digamos, 1000×1000 células, veríamos ainda melhor a inconveniência desse tipo de descrição. Podemos tornar a tarefa mais conveniente se adoptarmos a "postura do design": estudando os efeitos dinâmicos das regras "físicas" do jogo podemos projectar certas configurações celulares que "fazem" isto ou aquilo, que "se deslocam", "se transformam", que "comem" o que lhes aparece à frente, que "evitam" as configurações que "comem", que "persistem na sua identidade". E podemos descrever essas configurações com a linguagem própria dessa postura. Como escreve Dennett (p. 56):

"A física subjacente é a mesma para todas as configurações da *Vida*, mas algumas, em virtude nada mais do que da sua forma, possuem poderes que outras não possuem. Este é o facto fundamental do nível do *design*."

Se dermos mais um passo e passarmos da "postura do *design*" para a "postura intencional", podemos falar de configurações no *Vida* que "sabem" ou "acreditam" ou "querem" isto ou aquilo. Desse modo, aquelas configurações do *Vida* que se limitavam a exibir este ou aquele padrão de movimento são agora concebidas como sistemas intencionais ou agentes racionais: "Partimos simplesmente do princípio de que, independentemente da maneira como o fazem, fazem-no racionalmente - tiram as conclusões certas sobre o que fazer a seguir a partir da informação de que dispõem e de acordo com que querem" (p. 59). Dennett está a falar, convém não o esquecer, de configurações de quadradinhos acesos ou apagados no "jogo da *Vida*".

Dennett, apesar de reconhecer (p. 64) que não é possível especificar no *Vida* uma simulação suficientemente rica para comportar certos fenómenos complexos (como uma evolução com fim aberto, proporcionado pela possibilidade de mutação), apoia-se na demonstração matemática de que é possível implementar uma máquina de Turing universal num autómato celular deste género (pp. 60-61) para concluir que mesmo este "mundo determinista de brinquedo" tem todos os ingredientes necessários para permitir a evolução de "evitadores" (configurações "celulares" capazes de evitar o "dano" de serem "comidas" por outras). Então, conclui Dennett, se é possível neste mundo determinista que existam evitadores, isso é o suficiente para quebrar onexo entre determinismo e inevitabilidade: num mundo determinista, certos agentes podem evitar certas coisas, embora não possam evitar outras (p. 65).

Dennett resume assim o seu argumento (p. 70):

"Em alguns mundos deterministas há evitadores que evitam danos. Consequentemente, em alguns mundos deterministas algumas coisas são evitadas. Todas as coisas que são evitadas são evitáveis. Consequentemente, em alguns mundos deterministas nem tudo é inevitável. Consequentemente, o determinismo não implica inevitabilidade."

O que é central para compreender o real significado deste argumento é captar o que aqui quer dizer "evitar": não há nada em nenhuma estrutura do *Vida* que evite o que quer que seja (como nada é comido, ou se desloca, ou mantém uma identidade). No *Vida*, ao nível "físico", não há evitamento nenhum. Ao nível "físico" nem sequer há configurações, apenas "células". Não há nenhuma "lei física" (nenhuma regra) para configurações, só há regras para células. Não há nenhum "agente" ao nível "físico". Tudo depende, para nós, da ilusão de movimento provocada pelo padrão de "acender" e "apagar" de uns tantos pontos no ecrã do computador: é como aqueles anúncios de néon das lojas, que também criam a ilusão de movimento. Está tudo "no olho do observador", tem tudo a ver com as limitações do nosso processamento visual: as mesmas limitações que fazem que, quando vamos ao cinema, vejamos "imagens em movimento" em vez de uma sucessão de fotografias (que era o que veríamos se processássemos as imagens mais rapidamente e portanto as separássemos umas das outras). É tudo um artefacto da "postura intencional" inventada por Dennett. É-nos conveniente, a nós humanos intérpretes, falar de "evitamento" como um atalho de linguagem. Isso, aliás, ajuda a esclarecer o verdadeiro valor da "postura intencional", quando verificamos até que ponto ela permite este neo-animismo face a uma mera configuração de pontos pretos a piscar num ecrã de computador.

A primeira tese de Dennett (o determinismo não implica inevitabilidade) não obtém, no nosso entender, qualquer ponto de apoio nesta ilustração. As configurações no "jogo da *Vida*" não evitam nada: com a

dose certa de ingenuidade, parece-nos que "evitam", pela única e simples razão de que o domínio analítico das regras que impusemos ao projectar o jogo e ao desenhar certas configurações nesse universo não nos dá o domínio da dinâmica de um sistema a funcionar segundo essas regras (mas essa é uma limitação cognitiva nossa e não uma propriedade do sistema). Trata-se de mais um milagre da nossa ignorância. Para Dennett isso quer dizer que o ponto está no *design*. Connosco, humanos, a questão é basicamente a mesma. Nós e a lua somos feitos do mesmo material e obedecemos às mesmas leis físicas, mas a natureza da lua é fixa e a nossa não. Nós temos muitas maneiras de reagir às condições ambientais e a diferença está no *design*: nós "somos o produto de um processo de *design* competitivo a grande escala; a Lua não" (p. 108). Voltaremos a este ponto.

Indeterminismo e liberdade. A segunda das teses que Dennett se propõe defender é a seguinte: o indeterminismo não nos dá liberdade. Uma forma simples do argumento a favor desta tese é traduzida por outra das ferramentas ilustrativas a que Dennett recorre: o xadrez computacional. No essencial, o argumento corre como segue. (Para uma introdução breve aos principais conceitos e uma crítica às concepções acerca do xadrez computacional que prevalecem em certas correntes da investigação em Inteligência Artificial, cf. Silva 2004.)

Suponhamos que instalamos dois programas de xadrez (A e B) num computador e os pomos a jogar um torneio entre eles. Como de costume as partidas ficam registadas jogada a jogada e, quando posteriormente analisamos os registos, encontramos dois jogos exactamente iguais até à 12.^a jogada. No primeiro desses dois jogos, B (com as pretas) comete um erro crasso na 13.^a jogada e começa aí a sua derrota. No segundo desses dois jogos, A (agora com as pretas) faz uma 13.^a jogada completamente diferente e ganha. Analisando o funcionamento dos dois programas, verifica-se que B não chega à jogada "salvadora" porque as respectivas vantagens na dinâmica do jogo só se revelam muito longe na árvore de procura e B, com um mecanismo de busca menos eficiente do que o de A, abandonou esse ramo da árvore de procura antes de poder confirmar o seu interesse. O programa A, com um mecanismo de procura mais eficiente, chega a uma avaliação suficientemente esclarecedora da jogada "salvadora" dentro do tempo disponível, opta por essa jogada e ganha onde B perdera. Num certo sentido, B "podia ter feito aquela jogada": era uma jogada legal, sendo portanto uma "opção" para B. Contudo, continua Dennett, na realidade "não podia", porque não tinha "capacidade de análise" para lá chegar: só num mundo possível bastante diferente é que B podia ter feito aquela jogada (pp. 95-98). O design do programa A fizera a diferença: as maneiras que ele tinha de reagir às condições ambientais eram mais ricas do que era o caso com o programa B.

Para considerarmos, com este dispositivo, a tese agora em apreço, pergunta-se: o programa B poderia melhorar a sua condição se tivesse acesso a um processo físico indeterminista? Vejamos.

Os computadores incluem um gerador de números pseudo-aleatórios, que para muitas aplicações servem como se fossem números aleatórios. De cada vez que o gerador arranca (por exemplo, quando ligamos o computador) é criada uma longa sequência de números (sempre a mesma), aparentemente sem qualquer padrão. Se o computador, durante uma mesma sessão de trabalho, aceder várias vezes a essa sequência, retirará de cada vez segmentos diferentes da mesma e desse modo simula aleatoriedade. Suponhamos que o programa de xadrez, sempre que não consegue "decidir racionalmente" qual jogada fazer num determinado momento (terminou o tempo de avaliação e as duas jogadas que são as melhores candidatas a serem executadas na presente situação têm, com a análise que foi possível fazer, a mesma classificação), "decide" recorrendo ao gerador de números pseudo-aleatórios. Considere-se agora, por mor daqueles que dão muita importância ao indeterminismo físico, que o gerador de números pseudo-

aleatórios é substituído por um dispositivo genuinamente indeterminístico (com base em processos quânticos, por exemplo). Se o programa A ganhava sempre ao programa B, não vai deixar de acontecer o mesmo por se ter introduzido um mecanismo indeterminístico: o indeterminismo não acrescenta nada em termos de possibilidades realmente interessantes (p. 98).

De forma mais profunda, o que Dennett pretende mostrar (analisando as teses de Robert Kane, em *The Significance of Free Will*, de 1976) é que o indeterminismo físico (no sentido em que a mecânica quântica é indeterminista) não tem como ser um factor de liberdade humana. Dennett identifica nestes "libertistas" a ideia de que um agente humano só pode ser responsável por uma sua acção desde que a sua escolha dessa acção não tenha sido determinada pelo conjunto de condições físicas anteriores a essa escolha. Ora, Kane pretende garantir o livre-arbítrio introduzindo no mecanismo de tomada de decisão prática, algures entre a entrada e a saída de dados desse mecanismo, um processo aleatório. Dennett, e bem, considera que se isso tivesse alguma consequência (o que é duvidoso, como decorre do exemplo do xadrez computacional), ela só podia ser uma: passamos a ter decisões que, para situações idênticas, produzem resultados diferentes em momentos diferentes - o que é uma concepção assaz estranha de livre-arbítrio e de fundamento da responsabilidade (pp. 114, 126). Não poderíamos estar mais de acordo com Dennett neste ponto e consideramos que, com o seu argumento que aqui esquematizámos, consegue fazer valer a sua segunda tese: o indeterminismo não nos dá liberdade. De facto, podemos perguntar: quem quereria uma liberdade assente num mecanismo que pode ser simulado por um gerador de números pseudo-aleatórios? (Nada disto significa, contudo, que aceitemos por boas as razões apontadas por Dennett para desvalorizar a presença e o papel do acaso, do aleatório e do indeterminístico no mundo físico. A informação que Dennett carrega para esse ponto fica muito longe de fazer justiça a todas as dimensões do problema. Contudo, não sendo esse o ponto em análise, não cabe aqui aprofundar essa questão.)

Cabe assinalar, contudo, que Dennett pretende que a argumentação do seu capítulo 4, dedicado a mostrar as fraquezas dos "libertistas" que sigam as teses de Kane, faça outro trabalho: desmontar a tese do incompatibilismo. A tese do incompatibilismo é a seguinte: se o determinismo é verdadeiro, então não temos livre-arbítrio. A tese que Dennett pretende defender é o compatibilismo: o livre-arbítrio e o determinismo são compatíveis. Dennett mostrou bem, parece-nos, que pode haver indeterminismo e não haver qualquer forma de livre-arbítrio moralmente relevante. Em nosso entender, de modo nenhum demonstrou com este argumento que determinismo e livre-arbítrio possam conviver. Essa demonstração terá, eventualmente, de ser procurada noutra parte, porque é manifestamente abusivo considerar que o que chega para desmentir Kane chega para fundar a tese da compatibilidade entre determinismo e livre-arbítrio.

Opções aparentes, opções reais. A terceira tese de Dennett é que um mundo determinista não implica que as nossas opções sejam aparentes em vez de reais.

O que são "opções aparentes" e "opções reais"? Seja o caso de um homem em queda num poço de elevador. O homem "não sabe exactamente em que mundo possível de facto se encontra" mas sabe que está num "conjunto de mundos possíveis" em que, em qualquer deles, em breve inevitavelmente cairá no fundo do poço. Isso não implica inevitavelmente a sua morte: em alguns dos mundos possíveis (por exemplo, em todos aqueles mundos possíveis em que ele cai de cabeça) ele morrerá; mas em alguns outros mundos possíveis (por exemplo, alguns em que cai de pés e se enrola) pode sobreviver. O homem pode planear a sua acção para aproveitar a margem de manobra que lhe é dada por esse conhecimento que ele tem das possibilidades incrustadas na sua situação: pode planear racionalmente a sua acção para

tentar colocar-se num dos mundos possíveis em que sobrevive (pp. 103-104). Aprecie-se ou não este tipo de linguagem filosófica acerca de mundos possíveis, o ponto de Dennett é que por vezes podemos fazer alguma coisa para evitar certos danos, mesmo num mundo determinista: "uma coisa é inevitável para [ti] se [tu] não poderes fazer nada quanto a isso" (p. 74). Se temos os mecanismos apropriados para evitar tais danos, podemos fazer alguma coisa, temos opções e elas podem fazer toda a diferença para nós: o que interessa são as competências de tomada de decisão que o agente tem ou de que carece (p. 147). Nesse sentido, um organismo com um certo grau de competência tem, num mundo determinista, uma certa margem de manobra, opções reais, que são pontos de ramificação na nossa trajectória através da história (p. 117).

Organismos muito simples têm formas muito simples, automáticas e cegas, de reagir a certas características do ambiente que são particularmente importantes para evitar danos: meros interruptores do tipo "se topares com a condição C, faz A". Organismos mais complexos têm também "máquinas de escolhas", do tipo "se encontrares C, fazer A dará o resultado Z com a probabilidade p" (p. 177). Organismos muito simples têm raras ocasiões, durante toda a sua vida, de aplicar um desses mecanismos a uma ramificação de caminhos. Organismos muito complexos enfrentam mais ocasiões de aplicar esses mecanismos. Nos organismos muito simples, as opções disponíveis são muito limitadas; a complexidade vem a par com mais respostas possíveis às situações. Organismos simples são "sistemas intencionais balísticos", cujas metas são fixadas de uma vez por todas, sem reorientação possível; organismos mais complexos têm mais latitude para gerir metas (são "mísseis guiados") (pp. 166-169).

O problema de certas interpretações do determinismo é a concepção estreita, limitada, de "possibilidade": num mundo determinista pode haver aumento de competência e alargamento de oportunidades ao longo do tempo; aquilo que as pessoas podem fazer hoje é mais do que as pessoas podiam fazer no passado (pp. 307-308). Neste entendimento, há liberdade no sentido em que em certas situações pode dizer-se de uma pessoa que ela "podia ter feito de outra maneira". O que é que isso quer dizer? Simplesmente, que essa pessoa podia ter sido mais diligente a recolher informação pertinente para uma determinada decisão, que é informação sobre factores com um papel causal nas nossas acções (sejam factores externos ou internos ao agente) (p. 310).

Homero, no Canto XII da *Odisseia*, apresenta o episódio do canto das Sereias. As Sereias, na sua ilha, atraíam com um canto irresistível os marinheiros que navegavam ao largo, que assim se deixavam conduzir a uma armadilha mortal. Ulisses, avisado por Circe, sabendo que também ele e os seus companheiros não resistiriam à tentação, preparou-se para a ocasião explicando a situação à sua tripulação, tapando com cera os ouvidos dos seus marinheiros e ordenando-lhes que o amarrassem ao mastro do navio e que o prendessem ainda com mais cordas quando ele pedisse para o soltarem. Ulisses não expôs os seus companheiros à tentação e garantiu que ele próprio, concedendo-se a oportunidade de experimentar a situação, seria impedido nessa ocasião de tomar a má decisão que nesse momento haveria de querer tomar: aceder ao armadilhado convite das Sereias. Esta distribuição do processo de decisão revela uma competência sofisticada para, com antecipação, tornar evitável o que de outro modo (e para os que assim não procederam) era uma inevitabilidade (pp. 219-220).

Afinal, o problema da liberdade é uma questão de conhecimento: "Quanto mais sabemos, mais podemos fazer; quanto mais podemos fazer, mais obrigações enfrentamos" (p. 312). Devemos poder concluir, nesse caso, que, descontadas todas as variações irrelevantes (se não fosse possível descontar as variações irrelevantes entre situações, não seria viável em nenhum caso fazer regras e leis de aplicação geral), quaisquer duas pessoas que enfrentem situações similares agirão, se estiverem igualmente bem

informadas, substancialmente do mesmo modo nessa situação. Isso quer dizer que mil pessoas (ou um milhão de pessoas, ou...) que enfrentem uma situação equivalente na mesma posição, agirão do mesmo modo desde que estejam igualmente bem informadas. Dennett não nos indica como fugir a esta conclusão absurda.

A tese em apreço neste ponto, lembramos, é a seguinte: um mundo determinista não implica que as nossas opções sejam aparentes em vez de reais. Tendo esclarecido o que isto significa para Dennett, podemos desde já aceitar uma parte dessa tese: um mundo determinista não implica que não existam opções reais, no sentido em que há pontos de ramificação na trajectória de um sistema ao longo da sua história. Falta esclarecer, parece-nos, um aspecto essencial: em que medida podemos dizer que essas opções são "as nossas opções". Isto é: em que sentido participamos na escolha de qual dos caminhos seguimos em cada encruzilhada. Para isso interessa saber que liberdade é essa que aqui está em questão.

3. Liberdade

Como já vimos, para Dennett a liberdade é uma questão de *design*. Um agente tem mais ou menos liberdade consoante a sofisticação dos mecanismos de resposta ao ambiente com que tenha sido dotado. Certas configurações do autómato celular "jogo da Vida" foram projectadas de forma a deslocar-se no seu "universo", algumas foram projectadas para "comer" e outras para serem "comidas". O programa de xadrez computacional A foi projectado de forma a avaliar correctamente a posição que se lhe apresenta após a 12.^a jogada e assim consegue encontrar o caminho para a vitória, o programa B foi projectado de forma que o impede de encontrar esse caminho e acaba por não conseguir evitar uma derrota. O homem em queda no poço do elevador terá ou não um conjunto de mecanismos suficientemente sofisticados para elaborar uma resposta à sua difícil situação e disso pode depender a sua vida ou morte. De que decisões cada uma dessas "máquinas" ou "organismos" é capaz, consoante a forma como foi projectado, é que depende a competência para tomar o caminho certo nas encruzilhadas que o ambiente lhes apresenta. É nesse sentido que a liberdade - a "margem de manobra" - é uma questão de *design*.

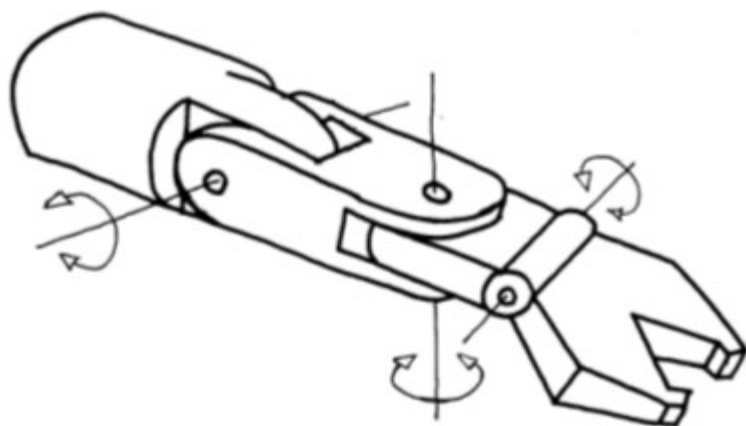
Numa perspectiva neodarwinista, o projectista dos organismos é a evolução natural. A evolução aumenta os graus de liberdade:

"as árvores podem "decidir" que chegou a Primavera e que está na altura de florir", mas essas opções são tão rudimentares que não são propriamente decisões. "Mas mesmo um comutador simples, que é ligado ou desligado por uma alteração do ambiente, marca um grau de liberdade, como dizem os engenheiros, e por isso é algo que precisa de ser controlado de uma maneira ou de outra. (...) Os comutadores (quer os que só alternam entre o estado ligado e desligado, quer os de escolha múltipla) podem ser ligados uns aos outros em série, em paralelo e em circuitos que combinam ambos os tipos de ligações. À medida que os circuitos proliferam, formando redes de comutação mais alargadas, os graus de liberdade multiplicam-se vertiginosamente e as questões de controlo tornam-se complexas e não lineares." (p. 176)

Os cérebros servem para recolher e tratar a informação que permite que o organismo se oriente nesse multidimensional espaço de possibilidades. Em alguns organismos o mecanismo das "máquinas de escolhas" torna-se particularmente sofisticado. Uma das formas mais sofisticadas de lidar com os sinais do ambiente, que parece só existir plenamente nos humanos, consiste em distinguir aparência e realidade. A cultura é, de forma mais geral, um patamar elevado de sofisticação desses mecanismos (p. 179).

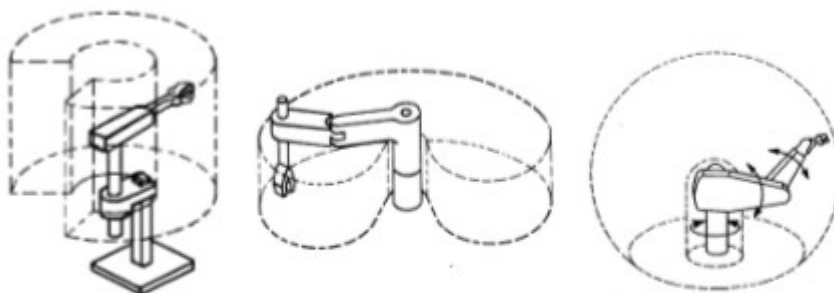
Para complementar a visão proporcionada pelo autómato celular e pelo programa de xadrez como instância de "decisores", convém também explicitar o que é, para os engenheiros invocados por Dennett, o conceito de "graus de liberdade". O conceito de "graus de liberdade" em engenharia compreende-se bem, por exemplo, em robótica. Tomemos uma caracterização básica a partir de (Couto 2000, capítulo 2).

Um determinado membro robótico é formado por vários braços, ligados entre si por juntas. Cada tipo de junta permite um certo movimento do segmento que lhe está ligado. O número de juntas e de braços de um membro robótico permite um certo número de possibilidade de movimentos combinados, confere um certo número de grau de liberdade (como se ilustra abaixo).



Um membro robótico com três graus de liberdade.
(Adaptação da figura 2.7 de Couto 2000)

Um membro robótico com uma determinada configuração permite que a sua extremidade (a "mão") alcance determinadas localizações no espaço circundante - mas não permite que alcance outras. Ao espaço desse modo acessível ao membro robótico pode chamar-se "volume de trabalho" (cf. figura abaixo para exemplos).



Volumes de trabalho de vários membros robóticos.
(Extraídos das figuras 2.11, 2.15 e 2.14 de Couto 2000)

Adquirido, de forma precisa, o conceito de "grau de liberdade" para um "corpo robótico" ("como dizem os engenheiros", segundo a expressão de Dennett), podemos associá-lo ao conceito de "programa de computador" (tal como usado por Dennett) e ficar assim com o par de dispositivos conceptuais básicos para compreender a "liberdade em determinismo" que nos é proposta: quais os danos que um

mecanismo/organismo é capaz de evitar no mundo é algo que depende do engenho posto pelo seu projectista no respectivo programa de controlo e dos graus de liberdade conferidos ao seu corpo (ou suporte físico) pelo mesmo projectista. É disso que depende a liberdade, porque a liberdade é um problema de *design*.

Aí reside um ponto essencial: a nossa liberdade é a margem de manobra que nos foi conferida pelo projectista (pela evolução). A liberdade é uma questão de competência, que depende do nosso programa e dos graus de liberdade do nosso corpo. Se o programa B de xadrez computacional tivesse sido dotado de um mecanismo mais eficiente de procura na árvore de jogadas, esse programa de computador teria mais liberdade. Dennett poderia ter acrescentado que o mero facto de correr o mesmo programa de xadrez em dois computadores com velocidades de processamento muito diferentes resultaria em duas máquinas de xadrez com competências muito diferentes - e, nesse caso, a máquina de xadrez (programa mais processador) com mais capacidade de cálculo seria a máquina com mais liberdade.

Se a máquina, em vez de ser o computador-com-programa, for eu ou o meu leitor, temos um problema: a nossa liberdade é a "competência" de fazer tudo aquilo que o nosso programa e os nossos mecanismos determinam que façamos em cada situação, até aos limites suportados pelo nosso *hardware*. Como Dennett explica longamente, o nosso programa será enormemente mais complexo do que o programa do computador e lidará com aspectos do mundo que deixam o computador indiferente. Contudo, o nosso programa não deixa de ser um programa, cuja "liberdade" é algo como uma "medida objectiva", que nos é exterior, da nossa competência para reagir de forma muito complexa a um ambiente complexo. Nós somos, nessa concepção, uma máquina de digerir informação e transformá-la em comportamento - tão livre como um computador. A diferença entre um humano que consideramos responsável e outro que acusamos de irresponsabilidade é, nestes termos, uma diferença de sofisticação, de precisão e de manutenção do mecanismo: o mecanismo do irresponsável está a funcionar mal, não fez toda a ginástica que consideramos normal num mecanismo humano e, portanto, fez um uso não óptimo da margem de manobra que consideramos padrão para um humano no nosso mundo. Essa é a liberdade do determinista, a liberdade definida pelo *design*.

Na liberdade do determinista, o que o agente faz numa certa situação depende, num certo sentido, de quem o agente é. Parece que aí poderá residir a responsabilidade moral. Mas "aquilo que o agente é" foi determinado por quem o projectou (pelo *design* do agente, pela evolução) e pelo ambiente. Em suma, o que o agente faz em cada instante é determinado pelo estado do universo no instante anterior e, cumulativamente, por todos os estados anteriores (de acordo com a definição de determinismo seguida por Dennett). Se o agente não ginasticou melhor o seu corpo, se não se preparou melhor para a situação, se não antecipou, se não treinou, se não tomou precauções - isso faz que o agente seja responsável pelas suas acções, porque podia ter feito de outra maneira? Não, porque o que o agente fez ou não fez, em cada instante dessa "fase preparatória", foi por sua vez determinado pelo estado do mundo no instante anterior (mecanismos internos e ambiente). A liberdade do agente na aceção do determinista é uma ilusão: só está no olho do observador. A liberdade é apenas mais um efeito da postura intencional. Nós pensamos que os outros são livres, porque não conhecemos os seus mecanismos de determinação - e pensamos que nós próprios somos livres, porque tão-pouco conhecemos os nossos próprios mecanismos de determinação.

Uma vez que Dennett descarta o modelo do "teatro cartesiano", o que o agente faz ou deixa de fazer é o resultado de um conjunto de processos internos centrados no sistema nervoso, também eles processos físicos igualmente sujeitos ao determinismo postulado: o estado do cérebro (ou do conjunto do sistema

nervoso central, ou mesmo de todo o sistema nervoso) num dado instante é estritamente determinado, em boa ordem, pela sucessão dos seus estados precedentes - até ao nascimento do organismo e, em bom determinismo, até qualquer ponto no passado do universo.

E isso é tudo, para o determinista: a "liberdade" reduz-se a uma forma conveniente de falar de certos sistemas. Aprofundemos, de seguida, a noção de liberdade que está aqui em causa, considerando o seu significado em sistemas multi-agentes.

4. Liberdade em sociedade

Um elemento central da estratégia argumentativa de Dennett em *A Liberdade Evolui* consiste em procurar substituir uma discussão em termos de livre-arbítrio por uma discussão das explicações possíveis para o aparecimento da liberdade no nosso planeta. O principal resultado negativo dessa estratégia é alimentar a ilusão de que um partidário de alguma forma de evolução natural tem de ser necessariamente determinista - o que não é o caso. De qualquer modo, é perfeitamente concebível que o autor nos ofereça uma tese ou um ponto de vista convincente sobre esse problema, a evolução da liberdade, mesmo que daí não resulte uma sustentação adequada para o ponto de vista determinista. Vale a pena, por isso, considerar o que nos diz Dennett numa parte substancial do seu texto, quando se dedica a desenhar uma perspectiva acerca da liberdade de seres sociais como os humanos.

Os termos em que Dennett coloca a questão do determinismo genético é um bom ponto de partida para esta abordagem, porque coloca em jogo o papel do ambiente. Criticando em particular Stephen Jay Gould e a sua oposição ao determinismo genético, Dennett é muito claro: o que não é determinado pelos nossos genes é determinado pelo nosso ambiente (não interessa muito onde se coloca o papel do acaso, porque isso não nos dá liberdade) e não há nenhuma boa razão para preferirmos o determinismo ambiental ao determinismo genético (p. 172). O ponto de partida do argumento é inescapável: ninguém pensa que possamos rever completamente o nosso legado genético (pelo menos por enquanto, não há vontade, educação ou aprendizagem que permitam a um humano do sexo masculino dar à luz). Alguns genes fixam de forma irrevogável certos aspectos do meu destino. Do lado do ambiente, como mostrou Jared Diamond (*Armas, Germes e Aço*), muito do que se passa nas tendências pesadas da evolução das comunidades humanas depende do seu ambiente: o nosso ambiente é em grande medida o que resulta das características genéticas dos animais e plantas que estão à nossa volta. No meio, estão as "máquinas de escolhas", que fazem o melhor que podem para evitar danos.

O que é curioso é que, sendo isto verdade, isto não é toda a verdade. Para a espécie humana, uma espécie organizada em sociedades densas e complexas, o "ambiente" é, em larga medida, constituído pelos nossos congéneres. Para o dizer brutalmente, somos hoje predadores e presas de outros humanos muito mais do que de qualquer outra espécie. Assim, os mecanismos de que resultam as razões da nossa acção são, em larga medida, mecanismos cuja funcionalidade está virada para a interacção social complexa. O problema da liberdade tem, pois, de ser pensado em termos de liberdade em sociedade.

Dennett, em perfeita coerência com a sua ideia de que a raiz da liberdade é o evitamento (como algumas configurações do "jogo da Vida" evitam ser comidas por outras), alinha com os inúmeros autores cujos esforços teóricos são profundamente marcados pela dificuldade em compreender o altruísmo. A ideia de proteger o interesse próprio por arranjos sociais que têm em conta o interesse comum com outros é vista como o que de mais próximo do altruísmo podemos conceber - mas sem ser realmente altruísmo, antes uma forma de pseudo-altruísmo ou, mais propriamente, de egoísmo prudente (pp. 207, 209).

"(...) os organismos podem vir a ser concebidos pela evolução para cooperar ou, mais precisamente, concebidos para se comportarem de modo a preferirem o bem-estar a longo prazo do grupo ao seu bem-estar individual imediato" (p. 210).

O argumento contra a cooperação e a favor do egoísmo é o seguinte: não se percebe como é que a cooperação pode ter evoluído naturalmente, porque "naturalmente" os altruístas seriam sempre suplantados pelos "oportunistas" no seio do grupo (pp. 210-211).

Parte importante de *A Liberdade Evolui* é utilizada para apresentar e discutir várias teorias, propostas e hipóteses acerca de como poderiam ter evoluído naturalmente (quer dizer, como poderiam ter uma base egoísta) as formas mais elevadas de moral tal como as reconhecemos entre os humanos. Não parece muito pertinente dar pormenorizada conta de todos esses recantos da exposição de Dennett. A facilidade com que se multiplicam teorias de base evolucionista parece-nos, em vez de um sinal de saúde desse programa de investigação, um indicador de risco de degenerescência prematura, de deriva ideológica, em que "vale tudo" por faltar uma compreensão mais aguda de que só uma ínfima parte do que poderia ter evoluído teve realmente lugar na evolução que efectivamente acabou por acontecer. Como Dennett afirma em *True Believers*, dizer que um sistema é como é porque a evolução assim o desenhou, é afirmar uma verdade que não explica nada: é preciso ser mais concreto acerca de como é que as coisas se passaram (ou se podiam ter passado) efectivamente no processo evolutivo (Dennett 1979: 33).

As concepções de Dennett acerca das formas superiores de liberdade em sociedade são pobres, na medida em que, exceptuando referências ocasionais a instituições e a acção política, se conformam a esquemas explicativos de exclusiva inspiração individualista. Sempre que se torna claro que é difícil compreender certas situações, comportamentos ou fenómenos sociais numa base ordinariamente individualista, Dennett volta à sua palavra de ordem (repescada de *Elbow Room*) segundo a qual quanto mais alargar o meu eu mais serei capaz de internalizar dimensões que de outro modo encararia como exteriores ou alheias a mim (pp. 137,227). No entender de Dennett, essa operação vai contra as concepções "pontilhistas" do eu, do eu como "ilha" ou como "enclave" protegido do resto do mundo (p. 138). Na verdade, pelo contrário, esse mecanismo do "alargamento do eu" torna-se necessário nesta concepção em que nunca se consegue ver na sociedade mais do que agregados de indivíduos em última instância egoístas. O "eu alargado" é indispensável para conferir alguma flexibilidade a este individualismo que, de outro modo, ficaria encurralado na sua própria estreiteza. Como não se quer admitir nada que não seja o egoísmo, mais ou menos "esclarecido", o que seja inexplicável pelo "egoísmo ordinário" reinterpreta-se como "egoísmo de um eu alargado": eu defendo o barco dos outros porque também navego nele e não me quero afundar.

Os instrumentos a que Dennett recorre para estimular a imaginação na direcção da sua particular concepção de liberdade em sociedade são, se bem lhes dermos atenção, esclarecedores. Focaremos a atenção em dois deles: os memes e as explicações em termos de teoria dos jogos.

Determinismo cultural: a máquina de memes

O leitor pode neste momento reproduzir, mesmo mentalmente, as quatro primeiras notas da *Quinta Sinfonia* de Beethoven? Não é extraordinário que essa "ideia" esteja tão espalhada entre os humanos, pelo menos de certas zonas do globo, de tal modo que provavelmente será reconhecida como "sinal musical" e reproduzida mesmo por pessoas que não sabem de onde ela vem exactamente? Outras ideias andam há muitos anos pelo mundo. Esses factos são interpretados de uma forma particular por uma

determinada teoria da cultura humana, a teoria dos memes. Dennett usa a teoria dos memes como entrada para a sua visão da liberdade em sociedade, razão pela qual é necessário clarificar o que está em causa por esse facto.

Richard Dawkins, em *O Gene Egoísta*, expõe a hoje em dia bem conhecida tese de que "nós, e todos os outros animais, somos máquinas criadas pelos nossos genes", "somos máquinas de sobrevivência - *robots* cegamente programados para preservar as moléculas egoístas conhecidas por "genes"", de tal modo que os nossos comportamentos resultam, pela evolução, não dos nossos próprios interesses, nem sequer dos interesses da espécie ou do grupo, mas dos "interesses" da unidade fundamental que é o gene. Bem vistas as coisas, o gene é apenas um caso particular de um "replicador", isto é, de uma molécula capaz de fazer cópias de si mesma. Esse replicador, em dado momento, descobriu uma forma de prolongar a sua existência: criar máquinas de sobrevivência, dentro das quais se instala e que controla a partir do interior. Nós somos uma etapa dessas máquinas de sobrevivência, *robots* desajeitados construídos e controlados por colónias de genes cuja preservação é a razão última da nossa existência (Dawkins 1976:29,23,41-42, 47,53-54).

No último capítulo dessa obra, Dawkins propõe uma forma de explicar uma especificidade da nossa espécie: a forma particular de cultura humana. A ideia é que tenha surgido recentemente uma nova estirpe de replicadores, os memes. Um meme é "uma unidade de transmissão cultural, ou uma unidade de imitação", uma estrutura que parasita os nossos cérebros, tomando-os como seus veículos, sendo que "tal como os genes se propagam no *pool* genético, saltando de corpo para corpo através dos espermatozoides ou dos óvulos, também os memes se propagam a si mesmos de cérebro para cérebro" (Dawkins 1976: 299-300). Exemplos de memes são: melodias, ideias, modas no vestuário, técnicas. O meme é "realizado fisicamente, milhões de vezes seguidas, como uma estrutura no sistema nervoso de homens individuais espalhados por todo o mundo". Um exemplo de meme é a ideia de Deus, a qual tem, no ambiente da cultura humana, "um elevado grau de sobrevivência ou de poder infectante". Como os cérebros e os seus corpos e outros veículos dos memes têm capacidades limitadas, há competição entre memes pelo controlo desses veículos - dos cérebros, mas também do tempo de antena nos meios de comunicação e do espaço nas bibliotecas, por exemplo (Dawkins 1976: 306). Os memes são uma produção dos genes, porque foram estes que deram cérebros às suas máquinas de sobrevivência. Tal como os genes, também os memes podem ser egoístas, isto é, "um traço cultural poderá ter evoluído da maneira como o fez simplesmente porque é *vantajoso para si próprio*" (Dawkins 1976:310), mas memes e genes podem combinar-se no sentido do reforço ou no sentido da oposição (este último caso pode ser exemplificado por um meme para o celibato).

Apesar de ter lançado uma tempestade intelectual centrada na ideia dos memes, Dawkins parece nunca ter fixado uma leitura coerentemente determinista dessa teoria, tendo salientado a imperfeição de um paralelismo entre genes e memes. Além das hesitações e reconsiderações que fará em obras posteriores dessa sua teoria, logo na sua versão original escreve, a fechar o livro (Dawkins 1976:311):

"Somos construídos como máquinas de genes e educados como máquinas de memes, mas temos o poder de nos revoltar contra os nossos criadores. Só nós, na Terra, temos o poder de nos rebelar contra a tirania dos replicadores egoístas."

Aparentemente, teríamos alguma capacidade para manobrar os nossos próprios memes, embora não se chegue a explicar como seria o correspondente mecanismo.

Dennett, por seu lado, não partilha dessas hesitações de Dawkins e vai levar tão longe quanto possível a "teoria social darwinista" (como Robert L. Trivers designa as teses de Dawkins em *O Gene Egoísta*). Tomemos as suas formulações sobre esse ponto em *A Ideia Perigosa de Darwin* (1995, principalmente capítulo XII).

Aí reconhece Dennett que a cultura humana nos diferencia como espécie e que ela pode contrariar resultados de processos genéticos anteriores (por exemplo, o aumento espectacular da estatura média dos humanos em poucas gerações é um efeito da cultura - alimentação e condições de vida - e não da genética), até porque as mudanças ambientais provocadas por inovações culturais podem mudar rapidamente a orientação das pressões da selecção genética. Contudo, cuidado com as ilusões, porque "a mente "independente" que luta para se proteger de memes estranhos e perigosos é um mito": nós somos os nossos genes e os nossos memes e o que faz a diferença no caso dos humanos é que os memes são capazes de trocar as voltas aos genes (Dennett 1995:364). A narrativa é esta: tal como nos estádios iniciais da nossa história evolutiva as células procarióticas foram invadidas por uma espécie de parasitas benéficos que, graças a um processo de simbiose mutuamente benéfico, originaram o aparecimento de células eucarióticas e dos organismos multicelulares - também o tipo de entidades que chamamos "pessoas" foi criado quando um certo tipo de animal (um macaco) foi infestado por um tipo específico de invasor, os memes (Dennett 1995: 338-339). Na verdade, continua Dennett, "eu não sou mais do que um sistema complexo de interacções entre o meu corpo e os memes que o infestam" (Dennett 1995: 365); "o que faz de alguém o que ele ou ela é são coligações de memes que governam - que desempenham papéis prolongados para determinar que decisões vão sendo tomadas" (Dennett 1995:367).

Precisamos dessa explicação, no entender de Dennett, porque, apesar da concepção corrente segundo a qual aceitamos uma ideia por ela ser verdadeira ou bela e rejeitamos ideias falsas ou feias, se dá muitas vezes o caso de aceitarmos ideias falsas ou feias enquanto rejeitamos ideias verdadeiras ou belas (Dennett 1995: 362-363). Seria bom esclarecer - entendemos nós - quem dita a verdade e a beleza de uma ideia, coisa que o autor não faz, como seria conveniente para revelar os pressupostos com que ataca o problema acabado de enunciar. Será que se trata tão-somente de afirmar que existem indivíduos que não se conformam a certos padrões de verdade e beleza? Ou será que se trata de afirmar que certos indivíduos aceitam ideias que consideram falsas ou feias? Se o problema está apenas em que nem todos os indivíduos partilham os mesmos cânones de verdade e beleza, talvez valesse a pena sugerir a Dennett que experimente visitar uma ou outra teoria da ideologia para resolver a dificuldade proposta.

O tipo de tática argumentativa de Dennett é, neste caso, iluminado pela estratégia da mono-explicação: "como a vida em si, e *todas* as restantes coisas maravilhosas, a cultura tem de ter uma origem darwinista" (Dennett 1995:339, ênfase nosso). Não há, pois, que ter receio de mostrar todas as faces do argumento. É o que Dennett faz, mais uma vez, com este exemplo: "Um académico não passa de um meio de uma biblioteca fazer outra biblioteca" (Dennett 1995: 344). Nós somos, bem vistas as coisas, marionetas: "O abrigo de que todos os memes dependem é a mente humana, mas esta é em si mesma um artefacto criado quando os memes reestruturam um cérebro humano de modo a tornarem-no um melhor habitat para memes" (Dennett 1995: 364). Na evolução cultural, o agente da transmissão não é o indivíduo humano: o agente da transmissão é o próprio meme (Dennett 1995: n.6 à p. 354). É essa a razão pela qual Dennett ridiculariza a ideia de que a evolução memética possa ser considerada lamarckiana.

Em *A Liberdade Evolui*, Dennett sugere que os memes podem abrir uma porta à ideia de que podemos fugir à nossa herança biológica, que esse ponto de vista nos abre um mundo de imaginação que de outro

modo nos estaria vedado (p. 192). Vista a história da ideia e a versão que dela dá Dennett, sabemos o que isso significa. Como essa é a porta de entrada para as teses do autor acerca da liberdade em sociedade, convém ter presente em que fundamentos repousam essas teses. No fundo, *A Liberdade Evolui* é um longo desenvolvimento de uma única frase de *A Ideia Perigosa de Darwin*: "Onde está a autonomia de que eu preciso para agir com livre-arbítrio? "Autonomia" não passa de uma palavra sofisticada para "autocontrole"" (Dennett 1995: 365).

Para vislumbrar o papel que desempenham as ciências do artificial nestes debates, interessa ainda esclarecer de outro modo o que são os memes. O meme é informação, informação independente do veículo, uma receita, um algoritmo. Na verdade, as diferenças entre capacidades humanas dependem de diferenças micro-estruturais nos cérebros humanos, a um nível que as neurociências (ainda) não captam. Essas diferenças são produzidas pelos memes (Dennett 1995:420). Os memes são pequenos programas, realizáveis em diferentes suportes (papel, vídeo, cérebro, computador) (Dennett 2003: 189). Isto é: a teoria dos memes é mais uma forma de nos aparentar aos computadores digitais, *hardware* com *software*, máquinas com programa incorporado. Essa concepção encaixa num pano de fundo no qual a transmissão cultural é (continuando a abusar de Shannon e ignorando os seus avisos) emissão e recepção de informação - o que se faz, por exemplo, ao "instalar energicamente tanta cultura quanto possível nos mais novos logo que estes a consigam absorver" (Dennett 2003:186). Somos, pois, objectos adequados, já não apenas para a engenharia genética, mas também para a engenharia memética (Dennett 2003:277).

Parece-nos difícil de perceber que vantagem pode ter esta teoria dos memes para um ponto de vista determinista genético-memético do tipo de Dennett. Talvez existam alguns ganhos ao nível da publicidade, um factor memético que Dennett valoriza. Contudo, esses ganhos dissipam-se rapidamente se tivermos o cuidado de ler com atenção tudo o que nos é dito e não descurarmos o que parecem ser pormenores. Certos autores tiveram mesmo oportunidade de mostrar o seu desagrado com a falta de consideração da memética por teorias anteriores que pretendiam fazer basicamente o mesmo trabalho. É o caso de Nick Rose, que preferia preservar as conquistas teóricas da sociobiologia e recorrer à memética só quando aquela fonte se mostrasse insuficiente. Rose exemplifica o problema ao analisar um argumento de Aaron Lynch. Lynch, na sua obra *Thought Contagion: How Belief Spreads through Society* (1996), em que defende uma versão da memética como "contágio social", escreve:

"Comer alimentos ricos em gorduras faz que as pessoas ganhem peso à medida que envelhecem. Assim, a percentagem de gordura corporal das mulheres está correlacionada com a sua idade. Mas os homens que preferem mulheres jovens podem ter carreiras reprodutivas mais longas, replicando os seus memes em mais filhos. Desse modo, nas sociedades modernas bem alimentadas, a preferência por parceiras magras, ao favorecer as mulheres jovens, tem um efeito replicador mais poderoso do que a preferência por parceiras gordas."

Ora, escreve Rose, basta substituir a palavra "memes" pela palavra "genes" para obtermos um argumento sociobiológico (Rose 1998, §4). Nem sempre a mesma substituição é tão fácil; talvez nem sempre seja possível; mas é útil compreender a proximidade das estratégias argumentativas.

Egoísmo para computadores

Quando trata de tentar compreender a liberdade em sociedade, Dennett socorre-se de novo de um instrumento popular para os praticantes das ciências do artificial: o dilema do prisioneiro, um dos modelos mais estudados em teoria dos jogos. A teoria dos jogos é uma tentativa altamente formalizada de representar e analisar situações em que jogadores (agentes que têm de tomar uma decisão) racionais

interagem entre si, tendo em conta a racionalidade dos seus oponentes (ou parceiros) quando definem a sua própria estratégia. O dilema do prisioneiro é usado para estudar o problema da cooperação, enunciado nos seguintes termos: como é que funciona a tensão entre o que é bom para o indivíduo no curto prazo e o que é bom para o grupo no longo prazo?

Uma apresentação clássica do "dilema do prisioneiro" é como segue (em inúmeras obras de exposição básica da teoria dos jogos aparece esta apresentação ou alguma equivalente; uma de fácil leitura é Davis 1970). Dois homens suspeitos de cometerem um crime grave em conjunto são presos e colocados em celas separadas. Interrogados, cada um deles pode confessar ou negar o crime. Se nenhum deles confessar, não haverá forma de provar o crime e os homens só serão condenados por um crime muito menos grave (um ano de prisão para cada um). A confissão confere o direito a um tratamento mais favorável, por constituir colaboração com a justiça (se ambos confessarem, cinco anos de prisão para cada um). Se apenas um deles confessar, o crime será considerado provado: o que confessa é libertado, o outro é condenado a uma pesada pena de 20 anos. A seguinte matriz traduz a situação.

		Suspeito 2	
		<i>Negar</i>	<i>Confessar</i>
Suspeito 1	<i>Negar</i>	(1 ano, 1 ano)	(20 anos, liberdade)
	<i>Confessar</i>	(liberdade, 20 anos)	(5 anos, 5 anos)

Do ponto de vista dos dois suspeitos (que são os jogadores), negar é cooperar com o outro, confessar é trair. Pode parecer que o melhor resultado para ambos resultaria da cooperação (ambos negam, 1 ano para cada um). Mas, cada um analisando a sua situação concluirá que, qualquer que seja a estratégia do outro, o melhor para si próprio é não cooperar (confessar). Vejamos o raciocínio do Suspeito 1: no caso do outro negar, se eu negar apanho 1 ano, se eu confessar vou em liberdade; no caso do outro confessar, se eu negar apanho 20 anos, se eu confessar apanho 5 anos. O raciocínio do Suspeito 2 dará o mesmo resultado.

O dilema do prisioneiro pode ser posto numa forma mais geral, como exemplificado na seguinte matriz:

		Jogador 2	
		<i>Cooperar</i>	<i>Desertar</i>
Jogador 1	<i>Cooperar</i>	(R, R)	(S, T)
	<i>Desertar</i>	(T, S)	(P, P)

Quanto aos resultados obtidos em cada caso, eles são os seguintes: R ("Recompensa" por cooperação mútua), P ("Punição" por deserção mútua), S (porque é "Sonso" aquele que coopera quando o outro é um desertor), T (resultado de ceder à "Tentação" de desertar quando o outro coopera). A estrutura de pagamentos típica do dilema do prisioneiro é a seguinte: $T > R > P > S$. Isto quer dizer várias coisas: o pior resultado possível para um jogador é o pagamento a uma vítima de deserção que não sabe agir em conformidade (S); o melhor resultado possível para um dos jogadores é o pagamento a um desertor individual (T); a estratégia dominante num único encontro é desertar, qualquer que seja a escolha do oponente; a melhor escolha individual para cada um dos jogadores (desertar) conduz ao pior resultado colectivo. É esta desigualdade nos pagamentos que gera a tensão entre o interesse colectivo e o interesse individual. Considera-se por vezes que uma segunda condição deve ser respeitada, a saber: $R > (S+T)/2$, isto é, a cooperação dá melhores resultados do que a alternância de mútuo acordo entre cooperação e deserção. É claro que a estipulação destas condições só faz pleno sentido à luz do "pressuposto da racionalidade", subjacente a este modelo como em geral a todos os modelos da teoria dos jogos: os

jogadores são plenamente racionais, cada um sabe que o outro é assim, ambos são dotados de capacidade de cálculo ilimitada. O estudo da cooperação no quadro da teoria dos jogos assume, geralmente, que o indivíduo racional é necessariamente um indivíduo egoísta (Axelrod 2000).

É este instrumento que ocorre a Dennett para considerar o problema da cooperação. Dennett, para quem "toda a circunstância na natureza em que algo de parecido com a cooperação surja, requer explicação", vira-se para este modelo explicativo: "é aqui que precisamos da teoria dos jogos, e do seu exemplo clássico, o Dilema do Prisioneiro" (p. 161). Tratemos, então, de compreender um pouco melhor o significado dessa opção. Para isso temos de mobilizar uma versão mais sofisticada do dilema do prisioneiro: o dilema do prisioneiro reiterado.

Como vimos, é a desigualdade $T > R > P > S$ nos pagamentos obtidos pelos jogadores que gera a tensão entre o interesse colectivo e o interesse individual nas situações que admitam como modelo o dilema do prisioneiro. É por isso que a solução do dilema do prisioneiro no quadro da teoria dos jogos só é possível na versão reiterada: quando os mesmo jogadores repetem as suas jogadas numa série de encontros sucessivos, a traição deixa de ser a estratégia dominante já que cada jogador sabe que os encontros se repetirão e que tem de ter em conta a reacção do oponente e isso despoleta um conluio tácito entre ambos (Macy 1998: §2.1). É em torno desta ideia que se desenvolve algum do trabalho mais conhecido de Axelrod.

Efectivamente, um dos trabalhos mais influentes na utilização de técnicas de simulação para tratar de fenómenos sociais é o de Robert Axelrod, que, nas suas obras *The Evolution of Cooperation* (1984) e *The Complexity of Cooperation* (1997), procura identificar as condições para a emergência da cooperação por mecanismos evolutivos, pela pura interacção de agentes individuais, procurando assim uma resposta para o problema da ordem social espontânea. O que Axelrod procurava era uma resposta para a questão: como é que o bem-estar do grupo pode prevalecer em situações em que os interesses individuais podem minar o interesse colectivo? O seu ponto de partida é que uma explicação dessas deve respeitar três condições gerais: primeiro, os agentes só adoptam estratégias cooperativas se elas têm sucesso no seu ambiente; segundo, essas estratégias cooperativas devem ser capazes de enfrentar com sucesso ambientes em que os participantes aprendem e se multiplicam; terceiro, populações de agentes cooperativos devem, uma vez estabelecidas, ser imunes à invasão de formas alternativas de comportamento. Qualquer estratégia cooperativa que não responda a estes critérios não poderia ter-se imposto num mundo dominado pela evolução. As suas experiências de simulação, com uma versão reiterada do dilema do prisioneiro, levaram-no à conclusão de que a estratégia (entre várias propostas por vários autores) que era a melhor candidata a explicar o sucesso evolutivo da cooperação era a estratégia de cooperação condicional conhecida pela designação de "Tit-For-Tat", em que o agente se comporta do seguinte modo: coopera na primeira ronda de qualquer jogo e a partir daí repete a jogada anterior do seu oponente, seja ela qual for (se o outro coopera, eu coopero; se o outro deserta, eu deserto).

Robert Hoffmann (2000) sintetiza o trabalho de vários autores que questionam a robustez dos resultados apresentados por Axelrod. O problema é o seguinte: será que as conclusões de Axelrod mostram características gerais da evolução da cooperação ou será que os seus resultados dependem de forma significativa dos valores escolhidos para os parâmetros das suas simulações? A dificuldade está em que, por vezes, embora seja relativamente fácil dar uma interpretação realista a um modelo de simulação na sua globalidade, pode ser inexequível dar qualquer interpretação realista para certas variações dos parâmetros.

Hoffmann identifica sete parâmetros que, de acordo com simulações realizadas por vários investigadores, ilustram este problema. Vejamos, abreviadamente, como os pressupostos que eles representam podem minar a robustez deste tipo de experiências.

1. *Representação do agente.* A evolução da cooperação é sensível aos pressupostos que dizem respeito à racionalidade do agente. Por exemplo, comportamentos envolvendo reciprocidade só são possíveis se o agente tem uma memória onde pode reter informação acerca das rondas anteriores, mas se a reciprocidade implica cálculos muito complexos isso terá custos de racionalidade que não estarão ao alcance de qualquer agente, sobretudo se quisermos que seja realista.
2. *População inicial.* Uma vez que o sucesso de uma estratégia depende do comportamento do opositor, a estrutura da população, só por si, pode ditar a sorte de certas estratégias. Por exemplo, as estratégias de cooperação condicional não dão resultado numa população excessivamente dominada por desertores. Note-se, entretanto, que os efeitos de pressão social estão ausentes deste tipo de simulação.
3. *Estrutura da população.* Experiências que abandonam a cláusula do anonimato da teoria dos jogos, usando a identidade dos jogadores para orientar a interação, mostram que isso pode ter efeitos sensíveis no sucesso relativo da cooperação. É o caso, por exemplo, em que os agentes podem recusar-se a interagir com um determinado oponente, devido à expectativa que formaram acerca do seu comportamento. Esta opção, só por si, pode promover a cooperação, porque os cooperadores podem recusar interagir com os desertores. De qualquer modo, a natureza do "dilema do prisioneiro reiterado" não admite o abandono da cláusula do anonimato. Outras experiências relativas à estrutura da população passam, por exemplo, por eliminar os que sofreram uma deserção a partir da ronda em que isso aconteceu, o que pode ser realista para certas condições de competição (casos de vida ou de morte) mas dificulta a evolução da cooperação. Outras experiências, também difíceis de enquadrar na teoria dos jogos, introduzem dimensões espaciais (interacções preferencialmente com os vizinhos e imitação dos mais bem sucedidos nas proximidades).
4. *Dinâmica da população.* A "dinâmica de replicação" usada por Axelrod para fazer variar a população de geração para geração, aumentando o peso relativo das estratégias que estavam a dar melhores resultados, impedia que uma estratégia que se tivesse extinto voltasse a emergir ou que surgisse qualquer estratégia nova. Técnicas de evolução que permitem outros cenários (por exemplo, permitindo que "mutações" apareçam na população a uma certa taxa) podem resultar em cenários diferentes, embora o próprio afinamento do parâmetro mutação suscite problemas de método (a mera modulação da taxa de mutação pode dar resultados muito diferentes, sem que seja claro como dar uma interpretação realista a essa variação). Um dos cenários que pode resultar de uma dinâmica mais flexível é a invasão de certas populações por estratégias aparentemente neutras (comportamentalmente idênticas) mas cuja diferença, quando tiver ocasião de se revelar, terá efeitos negativos irresistíveis. Seja o caso, por exemplo, de uma população largamente dominada pela estratégia RETALIAÇÃO PERMANENTE (coopera sempre até à primeira deserção do oponente, mas, a partir da primeira deserção deste, passa a desertar sempre, faça o outro o que fizer). Se esta população for invadida por estratégias COOPERA-SEMPRE (mesmo que o outro deserte), isso pode não se notar durante muitas gerações, enquanto houver um largo domínio de cooperadores na população (porque COOPERA-SEMPRE e RETALIAÇÃO PERMANENTE têm o mesmo comportamento face a oponentes cooperativos). Só que, se houver mais tarde uma invasão de estratégias DESERTA-SEMPRE (mesmo que o outro coopere), a população sucumbirá rapidamente a esta nova estratégia, porque COOPERA-SEMPRE não é capaz de se lhe opor.

5. *Variação dos pagamentos.* Experiências mostram que a variação dos pagamentos (por exemplo, variando o valor de T em relação a P) pode provocar evoluções diferentes (por exemplo, menor ou menor cooperação).
6. *Repetição.* A repetição (jogos de várias rondas) pode promover a cooperação, porque permite criar reputações. Contudo, a repetição associada a uma memória suficientemente poderosa pode permitir que os jogadores façam cálculos específicos acerca do fim do jogo (o jogador que deserta na última jogada nunca será penalizado por isso).
7. *Ruído.* O pressuposto de que os jogadores nunca cometem erros (seja na execução da sua própria estratégia, seja na percepção da estratégia dos outros) é pouco realista. O problema é que certas estratégias cooperativas são particularmente vulneráveis a esses erros. Por exemplo, num jogo entre duas estratégias TIT-FOR-TAT, um único erro despoletará uma série de deserções de ambos os lados. A capacidade para perdoar pode ajudar a remediar esse problema, mas tem de ser equilibrada para não cair em perfis de comportamento que permitam a exploração.

No entender de Hoffmann, este leque de demonstrações de falta de robustez dos resultados de Axelrod mostra o seguinte: primeiro, a evolução da cooperação é um fenómeno observável em circunstâncias muito diversas (ilustradas pela modulação dos vários parâmetros mencionados); segundo, a evolução pode conduzir ao sucesso da cooperação com vários tipos de estratégias de cooperação condicional e não apenas com a estratégia Tit-for-Tat; terceiro, a deserção também se pode impor como estratégia de sucesso nas situações modeladas pelo dilema dos prisioneiros reiterado. O que parece mais firme nos resultados de Axelrod é que a reciprocidade é um factor chave no sucesso das estratégias de cooperação.

O tipo de críticas ao trabalho de Axelrod recenseadas por Hoffmann merece, contudo, uma outra leitura, que implica levantar os olhos dos detalhes da simulação. Trata-se, em nosso entender, da necessidade de questionar um uso pouco esclarecido de modelos formais que carecem do realismo mínimo para serem autorizados como contribuições válidas para um esclarecimento de problemas da interacção social dos humanos. Como aponta Ken Binmore (1998), o dilema do prisioneiro repetido, tal como usado por Axelrod, modela uma situação de interacção entre dois estranhos, sendo que numa sociedade mais parecida com a sociedade dos humanos temos "jogos de múltiplos jogadores", num "jogo" que começou há muito tempo e em que a retaliação por comportamentos não cooperativos não é exercida apenas pelos envolvidos directamente, mas também pelos outros. E especifica:

"O jogo mais simples que parece capturar algo da intuição que os divulgadores erradamente aprenderam a etiquetar com a marca do Tit-for-Tat é um modelo de sobreposição de gerações no qual em qualquer altura estão vivos três jogadores. Ocasionalmente, um dos jogadores morre e é substituído imediatamente por um novo jogador. Em cada período, dois dos jogadores são escolhidos aleatoriamente para jogar o dilema dos prisioneiros, enquanto o terceiro jogador observa. Há muito tempo que de algum modo se estabeleceu um equilíbrio, que agora requer que cada jogador coopere sempre. Um jogador que não faça assim verificará que o seu oponente seguinte o punirá por deserção - qualquer que seja o oponente."

No fundo, o que Axelrod consegue fazer, com o seu uso do Dilema do Prisioneiro Reiterado, é ressuscitar um debate que já tinha começado nos alvares da teoria dos jogos, mas tinha sido abafado pelo domínio das interpretações formalistas. Destacamos apenas dois exemplos fornecidos por Alvin Roth (1995:8-13).

Quando, em 1950 (portanto, poucos anos após a publicação da obra de von Neumann e de Morgenstern que marca o nascimento da teoria dos jogos), Melvin Dresher e Merrill Flood realizaram na Rand

Corporation a primeira experiência da situação que viria a ser depois conhecida como o dilema do prisioneiro, os participantes eram remunerados pelo seu trabalho recebendo em dinheiro, no final do exercício, o resultado acumulado das suas jogadas ao longo de cem rondas. John Nash, o matemático e futuro prémio Nobel com uma das contribuições mais importantes para a teoria dos jogos, criticou a experiência por haver nela demasiada interação. O defeito da experiência consistia, segundo Nash, no facto de os jogadores, de facto, estarem num jogo de múltiplas jogadas e não numa sequência de jogos de uma única jogada cada - o resultado de estarem sempre a jogar contra os mesmos era que se criava uma reputação. A alternativa proposta por Nash era que os jogadores fossem sempre rodando e que nunca fosse possível a um jogador saber como tinha jogado nas rondas anteriores o seu actual oponente. Esse seria apenas um momento do esforço de depuração formalista da economia experimental envolvendo modelos da teoria dos jogos.

Autores houve que remaram num sentido diferente. Thomas Schelling realizou em 1957 experiências que mostravam que agentes em situações económicas são capazes de coordenação para resolver problemas, mesmo que isso exija recorrer a factores formalmente alheios à situação. Uma dessas experiências descreve-se de seguida. Um indivíduo entra num jogo em que, simultaneamente com dois outros jogadores, vai tentar ganhar uma certa quantia em dinheiro. Cada um dos três é designado por uma letra. A jogada, para cada um separadamente e sem comunicar com os demais, consiste em apresentar as letras que designam os jogadores (A, B, C) numa sequência qualquer. Se todos propuserem a mesma sequência, um prémio de montante x será distribuído por todos da seguinte maneira: para o jogador cuja letra apareça na primeira posição, para o jogador cuja letra apareça na segunda posição, para o jogador cuja letra apareça na terceira posição. Se nem todos propuserem a mesma sequência, ninguém recebe nada. Obviamente, cada um ganharia o máximo possível se a lista resultante tivesse à cabeça o seu próprio "nome". Contudo, dos 40 indivíduos que realizaram a experiência, 33 propuseram a sequência ABC. Dos 40, só 12 tinham a letra A. O que está aqui em causa, para Schelling, é que os jogadores encontram, fora da estrutura formal do problema, uma maneira de se coordenarem para alcançar um certo objectivo, enquanto a teoria dos jogos se empenha, pelo seu formalismo, em ignorar esses factores.

Uma polémica acerca das concepções implicadas neste debate sobre a cooperação, travada entre Michael Macy e Cristiano Castelfranchi em *The Journal of Artificial Societies and Social Simulation*, ajuda, em nosso entender, a esclarecer alguns pontos que dizem respeito a problemas mais gerais de interpretação das ciências do artificial.

O primeiro a tomar a palavra é Macy, que toma como objecto da sua crítica a ideia, de Axelrod, de que a cooperação é motivada pela "sombra do futuro": no dilema do prisioneiro reiterado a cooperação surge porque cada jogador consegue calcular que a sua deserção provocará, no futuro, a deserção do outro. Michael Macy (1998) vê, como outros, vários problemas nesse esquema e na teoria dos jogos em geral e, em consequência, expõe basicamente duas teses que pretendem perfilar-se para uma alternativa.

A primeira tese é que os humanos em interacção não são fundamentalmente calculadores racionais. O "jogador" que tem um modelo no dilema do prisioneiro reiterado é um matemático experimentado na própria teoria dos jogos, capaz de previsão estratégica baseada em informação completa e na captação perfeita da estrutura lógica de um problema bem definido. Já o "jogador" que encontramos na vida quotidiana, na generalidade das situações que a teoria dos jogos tem a pretensão de captar, não funciona assim. Nós não calculamos para concluir que é mais racional cooperar; nós aprendemos, por exposição repetida a certas situações, que certos comportamentos de interacção são mais apropriados ao nosso permanente convívio com outras pessoas. Isto é: em geral, nós não deliberamos racionalmente que

vamos cooperar, antes funcionamos segundo normas sociais, usos e costumes, convenções, rotinas, rituais, protocolos, regras morais, hábitos e heurísticas. Isso quer dizer que a ordem social não emerge da sombra do futuro, mas das lições do passado. Os insectos sociais (como as abelhas) estão geneticamente programados para seguirem regras conformes ao interesse do enxame, enquanto os humanos adquirem essas regras por repetição face à experiência. Se as nossas formas de interacção social dependessem usualmente de deliberação racional, as relações sociais seriam um emaranhado de respostas inapropriadas a reacções inesperadas por parte dos outros.

A segunda tese é em certa medida uma consequência da anterior: uma vez que os humanos em interacção são basicamente repetidores de padrões (e não seres em permanente deliberação racional), o que precisamos para explicar a cooperação é um modelo baseado na emergência. Isto é: a "cooperação emergente" explica como é que padrões de interacção simples e dependentes de informação acessível ao agente de forma puramente local são capazes de gerar soluções globais altamente complexas. Essa "cooperação emergente", assente numa teoria dos jogos evolucionista, só tem de assumir uma população de agentes adaptativos míopes regidos por regras - embora dê origem a sistemas dinâmicos não susceptíveis de compreensão analítica, devido à sua não linearidade e às suas propriedades estocásticas. Como extensão desta tese defende a teoria do "gene egoísta" e pretende que ela permite resolver o problema da explicação do altruísmo: se o altruísmo é sempre um sacrifício, como explicá-lo? Se for o "gene egoísta" a controlar o comportamento do seu portador, o altruísmo é explicado pelos conceitos de adaptação inclusiva e altruísmo de parentela: o altruísmo é um comportamento ditado pelo gene (e não pelo indivíduo portador); o sacrifício do indivíduo nada interessa ao gene de que ele é portador, desde que esse comportamento melhore as condições de reprodução do gene nos parentes desse portador. Esta explicação exemplifica, no entender de Macy, as vantagens de um modelo em que o "jogador" é subdeterminado, em vez de ser encarado como um agente intencionalmente deliberativo nas suas interacções sociais.

Cristiano Castelfranchi (1998) responde a Macy segundo as seguintes linhas.

Em primeiro lugar, a usual identificação entre "racionalidade" e "egoísmo" ocupa um lugar central na argumentação de Macy. Para essa perspectiva, qualquer cooperação é um sacrifício, uma jogada perigosa. Contudo, nada na teoria da decisão racional justifica essa identificação, porque essa teoria lida exclusivamente com a razão instrumental: essa teoria diz respeito aos meios para atingir os fins do agente, não diz respeito à forma de definir os fins e estes podem ser, por exemplo, o bem do grupo ou o bem de outro agente. O problema é que a teoria da decisão racional acaba por ser sistematicamente contaminada por certas visões da racionalidade económica, em que os fins do agente são entendidos de forma restrita como "lucro" (apesar de a "utilidade" não ter que ser, mesmo em termos económicos, reduzida ao lucro). Esta ideia da sociabilidade como um mal necessário, em que os agentes são fundamentalmente oponentes, é uma consequência indesejável de uma filosofia utilitarista.

Em segundo lugar, é inaceitável a ideia de que os agentes não podem ser deliberadamente cooperativos: eles são-no muitas vezes e de forma significativa. Os agentes são, em muitas circunstâncias, deliberadamente cooperativos, por exemplo por terem (ou por acreditarem que têm) objectivos comuns ou interesses comuns com outros agentes. É certo que há cooperação repetitiva (como diz Macy), mas também há cooperação deliberada, tal como há cooperação altruísta sem deixar de haver também cooperação egoísta. É preciso evitar as taxinomias simplistas: a "ordem social" não é sempre cooperativa, nem "boa".

Em terceiro lugar, Castelfranchi retira consequências importantes do facto de reconhecer parcialmente razão a Macy, quando este assinala que há aspectos emergentes da cooperação e que nem tudo é deliberado. É que, se os agentes podem planear, podem antecipar o futuro e isso tem consequências ao nível do seu comportamento, também é certo que mesmo assim não podem prever todos os efeitos globais e compostos das suas acções ao nível colectivo. É preciso reconciliar emergência e cognição, porque mesmo as acções plenamente intencionais dos agentes têm efeitos colectivos que eles não poderiam intencionar e, por essa via, há efeitos colectivos emergentes tanto dos aspectos intencionais como dos aspectos não intencionais da acção de agentes que interagem socialmente de forma deliberada.

Este debate entre Macy e Castelfranchi é muito útil para compreender o que o tipo de determinismo assumido por Dennett implica em termos de (dificuldade de) compreensão dos fenómenos sociais humanos. É que Dennett, embora pareça suspeitar de que estes modelos são demasiado simplistas, não consegue libertar-se do principal constrangimento que eles impõem. A sua adesão às teses do "gene egoísta" e da subdeterminação dos organismos individuais por "conspirações genéticas" impede-o de considerar soluções que não sejam "espontâneas" (isto é, que não sejam "naturais" de acordo com uma leitura reducionista do naturalismo). Em substância, o posicionamento de Castelfranchi tem o mérito de mostrar que a adopção exclusiva dessa tese é pouco razoável: a combinação desse tipo de mecanismos com mecanismos deliberativos seria susceptível de uma maior riqueza explicativa. Isso poderia passar, por exemplo, por dar maior atenção aos arranjos institucionais que contam com o papel de agentes livres e deliberativos em sociedade. Isso pede, contudo, um uso um pouco mais aberto e menos ortodoxo das ferramentas das ciências do artificial. Dennett, pelo seu lado, parece pouco interessado em procurar compreender as razões profundas da inadequação do "Dilema do Prisioneiro" à equação dos problemas da liberdade em sociedade. Esse facto traduz uma dificuldade de fundo, relativamente constante na obra de Dennett, que consiste numa certa credulidade e algum superficialismo no recurso a ferramentas das ciências do artificial. As intuições autorizadas pelo "dilema do prisioneiro" servem o determinismo de Dennett, mas à custa de adoptar como paradigma de dilema social uma situação que distorce fortemente o próprio problema em causa.

5. Resultados e perspectivas. Margem de manobra e tiros nos escuro: notas à metafísica do determinismo

Vejamos o que fizemos até agora nesta leitura de *A Liberdade Evolui*.

Na primeira secção verificámos que o autor introduz uma definição de determinismo para guiar o seu argumento, para vir depois a perturbar essa definição com variações que obscurecem o que está em causa.

Na segunda secção escrutinámos as três teses que fazem o compatibilismo de Dennett.

A primeira tese (o determinismo não implica inevitabilidade), ilustrada com um autómato celular artificial, é pesadamente dependente de uma versão muito instrumentalista da "postura intencional". A ilusão epistemológica embutida no "jogo da Vida" pareceu-nos insuficiente para apoiar esta primeira tese.

A segunda tese (o indeterminismo não nos dá liberdade) é, por via de uma crítica a autores libertistas e com ilustração no xadrez computacional, reduzida à tese de que, se um gerador de números pseudo-aleatórios fosse um mecanismo capaz de fundar a liberdade de um sistema, essa liberdade não valeria a

pena. Dennett tem sucesso a descartar essa noção degenerada de liberdade, mas isso não implica que tenha sucesso a mostrar que "o determinismo é compatível com a liberdade" - que era, isto sim, o que precisava para confortar as suas próprias teses. A segunda tese serve, pois, para desmontar uma particular concepção de liberdade, mas não contribui em nada para apoiar o compatibilismo proposto pelo próprio Dennett.

A análise da argumentação de Dennett a favor da terceira tese (um mundo determinista não implica que as nossas opções sejam aparentes em vez de reais) mostrou que essa tese é cindível em dois aspectos. O primeiro aspecto dessa tese (num mundo determinista pode haver pontos de ramificação na trajectória de um sistema ao longo da sua história) foi eficazmente defendido pelo Autor, que mostrou como a evolução de um sistema com um certo repertório comportamental é o resultado de uma interacção entre os mecanismos do próprio sistema e o ambiente, não resultando exclusivamente de factores externos. O segundo aspecto da terceira tese (as ramificações na trajectória de um sistema ao longo da sua história equivalem a liberdade do agente) implica, para ser avaliado, um esclarecimento da noção de liberdade: as opções "reais" são as "nossas" opções? O esclarecimento desse ponto é remetido para a secção seguinte - e esse esclarecimento é importante, porque a debilidade dos vários pontos de suporte do compatibilismo, revelada pelo escrutínio das três teses em que se desdobra, faz depender desta noção central a solidez de todo o argumento.

Na terceira secção analisámos o conteúdo da noção de liberdade nesta obra de Dennett, recorrendo designadamente ao esclarecimento da noção de "graus de liberdade". Concluimos que, no argumento de Dennett, a liberdade se reduz ao funcionamento escorreito do programa de controlo e dos mecanismos corporais que o projectista (a evolução natural) conferiu a um determinado sistema, de tal forma que esse sistema é competente para evitar certos danos no mundo. Trata-se, afinal, de uma interacção entre o nosso projectista e o nosso ambiente: uma espécie de "liberdade objectiva" que está apenas "no olho do observador".

Na quarta secção aprofundámos a noção de liberdade que nos é proposta, em direcção à "liberdade em sociedade". Tratámos, aqui, de tornar mais explícitos os pressupostos subjacentes à argumentação da obra, nomeadamente pela análise do uso que é feito de mais uma ferramenta típica das "ciências do artificial": o dilema do prisioneiro, em teoria dos jogos. Pensamos ter evidenciado a forma insuficientemente crítica como o Autor se faz valer dessas ferramentas, de uma forma que serve as suas intuições deterministas, mas sem ser capaz de ilustrar a verdadeira complexidade do debate que envolve essas "experiências de pensamento". Como mostrámos, a situação que Dennett apresenta como modelo de um dilema social está longe, no enquadramento proposto, de fazer jus à natureza do problema. A vantagem é que as exageradas simplificações que Dennett se autoriza iluminam os seus pressupostos - embora estejam longe de constituir um bom suporte para o seu determinismo e para a correspondente (empobrecida) noção de liberdade. No argumento de Dennett, o determinismo é o de sempre: a liberdade é que empobrece para servir o compatibilismo.

Para lá desta análise, que segue de perto o argumento de Dennett, concluiremos agora com algumas notas gerais que pretendem propor, de forma livre, algumas intuições metafísicas alternativas às do determinismo de *A Liberdade Evolui*.

A estratégia discursiva de Dennett, nesta obra como em outras, assenta num pressuposto implícito que se desdobra em duas alíneas: primeira, a da superioridade científica da sua proposta; segunda, a do carácter metafísico das teses opostas. Na verdade, segundo critérios que o próprio adopta num ou outro momento

da sua obra, o seu determinismo não é científico (na medida em que não é susceptível de invalidação por evidência empírica), mas metafísico.

Herbert Simon (1983), no quadro geral das suas teses acerca da racionalidade limitada (a racionalidade humana é fortemente cerceada pela situação e pelos poderes computacionais dos agentes), procura uma explicação para o facto de, mesmo assim, nos desvencilharmos de forma razoavelmente eficiente da maioria das circunstâncias correntes da nossa vida. Porque é que, apesar de ignorarmos inúmeros aspectos que nos poderiam afectar em consequência das nossas acções, os nossos comportamentos ainda assim servem razoavelmente a nossa sobrevivência e os nossos propósitos? Pelo menos em parte isso deve-se ao facto de certas carências que são constantes nos organismos (como a respiração) estarem a cargo de mecanismos fisiológicos que dispensam a nossa atenção - bem como ao facto de dispormos de mecanismos (como as emoções) que garantem aos problemas mais prementes a prioridade na sua resolução.

Simon aponta, além disso, uma razão geral para que seja possível vivermos, com certo êxito, segundo este modelo de racionalidade limitada, que nos permite compartimentar a nossas decisões: "*Vivemos no que se poderia chamar um mundo quase vazio* - um mundo no qual há milhões de variáveis que em princípio poderiam afectar cada uma das outras, mas que não o fazem na maior parte das vezes" (Simon 1983:30; ênfase nosso). Esta concepção do "mundo quase vazio" é largamente coincidente com a ideia de Whitehead, que (como vimos antes) Dennett menciona mas de que não tira proveito: "a ampla independência causal de ocasiões contemporâneas é que preserva a margem de manobra no seio do Universo".

O "mundo quase vazio" e o "mundo determinista" são duas metafísicas alternativas. Ora, para opor à visão do "mundo quase vazio" a visão do "mundo determinista" e dar ao determinismo uma formulação científica, seria necessário: primeiro, fornecer uma colecção de teorias deterministas locais, isto é, uma teoria determinista para qualquer classe de fenómenos considerada; segundo, fornecer uma teoria determinista global que, além de ligar todas as teorias deterministas locais, fosse capaz de excluir que qualquer aparente determinismo local resultasse de um mero efeito de construção engenhosa das fronteiras entre os diferentes domínios de fenómenos. Essa formulação científica de determinismo nunca foi fornecida: pressupô-la é uma preferência metafísica.

Mesmo que essa formulação científica do determinismo global chegasse alguma vez a ser fornecida, ela não poderia nunca ser estritamente provada ou falsificada apenas por "questões de facto". Isso resulta, no estado actual de conhecimento científico, de limitações inerentes à nossa posição no mundo físico. O passado acessível ao observador é representado por um cone espacio-temporal que, com origem no ponto que representa o presente (o aqui e agora) do observador, tem as fronteiras definidas pela velocidade finita da luz e pela limitação que ela impõe ao nosso acesso a informação sobre pontos distantes no espaço e no tempo. Assim sendo, o passado acessível ao observador não chega para obter a especificação das condições iniciais que seria requerida para uma previsão determinista - seria preciso, para tanto, o acesso ao passado integral, mas o passado integral extravasa o "cone de passado" do observador. Esse "ponto de vista do infinito", requerido pelo determinismo global, coloca-o fora do campo da investigação científica (Prigogine e Stengers 1990:264).

Uma outra consideração do papel e da condição dos sujeitos cognitivos no mundo leva-nos ao mesmo território. Mesmo na visão dennettiana, o sujeito cognitivo (em geral, qualquer agente) faz uma diferença no mundo: consoante a sua competência, consoante o que evita ou não evita de acordo com o alcance dos

seus mecanismos internos, faz com que em cada encruzilhada o curso dos acontecimentos siga por uma ou outra das ramificações possíveis. Ora, a diferença que o sujeito cognitivo faz no mundo é, em muitos casos, mediada, num sentido forte, pela actividade cognitiva do agente. Pelo menos desde o princípio de incerteza de Heisenberg que compreendemos a existência de fenómenos cujo conhecimento implica interacção física entre o sujeito e o objecto, de tal modo que o observador perturba o observado - mas não precisamos da microfísica para entender, noutra nível, que certos fenómenos sociais são perturbados por serem objecto de estudo. Assim, desde que a actividade cognitiva do agente não seja pura apreensão intelectual e envolva interacção física entre sujeito e objecto, a ignorância de qualquer sujeito de conhecimento é constitutiva do próprio universo. No plano da pergunta pelo que podemos saber, o demónio de Laplace não resiste a este obstáculo: a ignorância de cada um de nós perturba objectivamente o seu conhecimento (Amsterdamski 1990:230-231). Nesse plano, o determinismo está para lá do teste pela evidência empírica acessível a sujeitos cognitivos cuja interacção com o mundo não é puramente intelectual.

Para compreender os efeitos da metafísica determinista de Dennett na sua argumentação é útil considerar a sua noção de "factos históricos inertes". Factos históricos inertes são factos (do passado) que, no que toca às leis físicas, podiam ter sido de uma maneira ou de outra sem deixar qualquer efeito subsequente (p.84). Um exemplo de Dennett: num computador digital, as diferenças de voltagem que não afectam a distribuição dos eventos elementares pelas categorias 0 e 1 nunca poderão, passado o momento da sua ocorrência, fazer qualquer diferença no futuro (p.93). Vistas as coisas do lado do futuro: é muitas vezes impossível dizer quando é que um facto histórico até agora inerte virá a emergir como capaz de fazer a diferença para o que acontece a seguir. Aparentemente, estamos de novo no plano em que o determinismo se defende com a ignorância: os factos estão lá, apesar de nos faltar o conhecimento suficientemente detalhado para os tomar em devida conta.

Mais uma vez, teria sido preferível não deixar a imaginação ser condicionada por um artefacto tão raro no nosso mundo: o computador. Quando se perceberá que o fascínio do computador digital reside, não no facto de ele ser um modelo universal, mas antes na sua radical diferença em relação à generalidade das (outras) entidades naturais, em termos de constituição e funcionamento? Parece que podíamos reconhecer melhor o papel dos factos históricos inertes se alargássemos o seu viveiro e atendêssemos ao seu peso nas decisões dos agentes. Seja o seguinte caso. Estou desempregado há muito tempo e, não dispondo de quaisquer outros meios de sobrevivência além dos frutos do meu trabalho, caí na pobreza e já estou a passar fome. Estou a considerar duas opções: continuar durante alguns dias a procurar trabalho ou, se isso não der qualquer resultado, roubar para comer. É o caso (embora eu não saiba disso) que, se andar 100 passos para Norte encontro o indivíduo A, se andar 100 passos para Sul encontro o indivíduo B. O indivíduo A procura alguém com as minhas qualificações para empregar de imediato na sua empresa. O indivíduo B está a montar uma burla e procura recrutar alguém com as minhas qualificações para se juntar à sua quadrilha. Não tenho nenhuma razão particular para me dirigir para norte ou para sul, uma vez que já percorri todas as ruas deste bairro várias vezes durante o dia de hoje; nada indica que essa razão exista ao nível inconsciente. Assim, o facto de que acabo por me dirigir para Sul (ao encontro do burlão) não tem qualquer significado como "decisão" ou "escolha" relevante para a minha situação. Esta história resulta em que aderi ao esquema do burlão, aplicando-se neste caso o ditado "a ocasião faz o ladrão". Um determinista reducionista poderia raciocinar assim: nenhum dos passos dados pelo indivíduo em causa escapou às leis da física, à constituição do seu próprio corpo e à sua configuração mental; assim sendo, o resultado dos seus movimentos foi determinado por esses antecedentes. Ora, se o indivíduo A estivesse a Sul (em vez de estar a Norte) e o indivíduo B estivesse a Norte (em lugar de estar a Sul), os mesmos movimentos teriam dado um resultado completamente

diferente. Este exemplo mostra que só uma eliminação radical do acaso permite manter a coerência do determinismo: mas essa eliminação radical do acaso só pode ser metafísica, pelo menos no actual estado do conhecimento científico.

O facto é que qualquer "sequência de eventos banais" inclui inúmeros "factos inertes", isto é, que não serão tidos em conta por nenhum algoritmo razoável para lidar com a situação, mas que poderão condicionar o seu resultado. O que é "um algoritmo razoável para lidar com a situação": um algoritmo que, tendo obrigatoriamente em conta as limitações da capacidade de processamento do agente em causa, inclui o maior número possível de factores relevantes, mas exclui (além de possivelmente alguns factores relevantes) muitos factores que normalmente não são relevantes mas podem tornar-se "desviantes". Por isso, muitas consequências inadvertidas das acções de qualquer agente podem emergir, mais tarde, como condicionantes de decisões e de acções futuras desse mesmo agente. Se a acção dos agentes tem um papel no curso dos acontecimentos no mundo físico, o mundo físico só pode ser estritamente determinista após a operação metafísica da eliminação radical do acaso ao nível da acção dos agentes, incluindo a eliminação radical dos "factos inertes" admitidos por Dennett.

Num mundo determinista "denso" (por oposição a um "mundo quase vazio") deveria ser fácil para um agente "mudar o mundo": qualquer comportamento teria inúmeras consequências, identificáveis nas primeiras ondas de repercussão, com efeitos multiplicadores em inúmeras linhas causais atravessando o espaço e o tempo em todas as direcções. Na realidade, pelo contrário, inúmeros acontecimentos particulares são absorvidos pelo fluxo dos acontecimentos e não chegam a ter qualquer efeito, perdem-se, diluem-se nas interacções. Mudar o mundo é difícil e exige, em geral, doses massivas de insistência, mobilização, concertação, sustentação e direccionamento. Poucos acontecimentos particulares chegam a estar na crista da onda e a tornar-se relevantes - e poucos dos acontecimentos que chegaram a tornar-se relevantes chegaram à crista da onda por efeito do planeamento de qualquer agente individual agindo isoladamente.

Do mesmo modo, qualquer "decisão" isolada é o resultado de muitas decisões tomadas ao longo do tempo e que paulatinamente configuraram o espaço decisional. Muitas dessas decisões anteriores eram aspectos "inertes" da situação visada intencionalmente. O horizonte decisional que se nos abre em qualquer situação foi sendo configurado tanto pelos efeitos intencionados como pelos efeitos não intencionados de inúmeras decisões focadas em problemas muito díspares. Uma parte relevante da noção de responsabilidade pelos nossos actos está ligada às consequências aparentemente inertes de decisões e acções sem qualquer efeito prático imediato: as consequências eventuais de actos que parecem morrer neste momento mas podem emergir num momento futuro. (Veja-se o caso, em termos jurídicos, dos precedentes invocados para interpretar uma regra com uma lacuna.)

As preferências metafísicas de Dennett notam-se, por vezes, numa atenção insuficiente às ramificações das ferramentas que usa para estimular a intuição. Vejamos um exemplo.

Dennett argumenta muitos vezes, em muitos dos seus textos, recorrendo ao conceito de "máquina de Turing" (MdT). Alan Turing, no artigo em que introduziu esse dispositivo conceptual precursor dos modernos computadores digitais, distingue entre "máquinas automáticas" e "máquinas de escolhas". As máquinas automáticas são aquelas cuja operação é completamente determinada pela respectiva configuração. Já as máquinas de escolhas têm configurações ambíguas e, quando a máquina atinge uma dessas configurações, a continuação do processamento depende de algum operador externo que faça uma das escolhas possibilitadas por essa ambiguidade (Turing 1936:232). Aquilo a que mais habitualmente se

chama "máquina de Turing" é a máquina automática, o único caso analisado em detalhe no texto referido. Vale, contudo, a pena pensar um pouco mais no outro caso, na MdT não determinista. Uma forma particularmente interessante de MdT não determinista seria uma "MdT social": uma MdT que, em certas configurações, é influenciada pelo estado de outra MdT. Mesmo aqueles que insistem em considerar mais pertinente olhar para os humanos como máquinas (a partir dos conhecimentos e das metáforas permitidas pelas máquinas) do que como organismos animais (a partir dos conhecimentos e das metáforas permitidas pela biologia), terão de acrescentar alguma justificação para se restringirem sistematicamente à inspiração que encontram na MdT determinista, com exclusão das MdT não deterministas - entre as quais, sugerimos nós, se encontram como particularmente interessantes estas "MdT sociais".

Ora, uma tal MdT social poderia depender das entradas de dados provenientes de outras MdT por inúmeras razões: por exemplo, por aceitar como boa a informação fornecida por outros agentes, mesmo que não tenha maneira de verificar até que ponto ela é fiável (ou por preferir o risco de confiar ao custo de verificar); ou, ainda por exemplo, por ter adoptado o "programa" de se conformar às regras de certas instituições ou a certas rotinas que existem fora do seu mecanismo. É provável que uma abelha, no que toca à sua dança do pólen, não tenha meios de realizar essa consulta a outras abelhas ou a um código de conduta armazenado fora do seu organismo - mas também parece claro que os humanos usam intensamente, para determinar a sua acção, consultas a outros humanos e a códigos de conduta da mais diversa índole (sem que essa consulta tenha carácter algorítmico, nomeadamente devido a limitações cognitivas: caso contrário, em vez de tribunais penais só haveria códigos penais). Nesse sentido, pelo menos tomados como "máquinas" individuais, essas MdT são MdT sociais, uma subespécie das MdT não deterministas. O que faz com que para Dennett, como para muitos outros, só se pense em termos de MdT deterministas, é a preferência metafísica por indivíduos estritamente determinados.

Outra preferência metafísica do determinismo vai para as sociedades estritamente determinadas. Essa poderia ser uma saída para a questão suscitada no parágrafo anterior: mesmo que as "MdT sociais" não sejam determinadas ao nível individual, isso não exclui que, no seu conjunto, um sistema de tais máquinas não seja ainda determinista. Nesse cenário, as interferências que a máquina M espalha pelas máquinas que estão ao seu alcance são, todas elas, resultados das interferências que a máquina M sofreu, por seu turno, de inúmeras outras máquinas. Desse modo, a não determinação de cada máquina é apenas uma ilusão que se desfaz considerando a máquina social no seu conjunto.

Contudo, parece que a única forma de proteger este "determinismo social radical para MdT" é excluir qualquer grau de genuína deliberação por parte de cada uma das "MdT sociais" envolvidas. É isso, aliás, que faz o determinismo de Dennett, quando faz equivaler "liberdade" e "design": em cada estado do mundo, cada sistema faz o que o seu design lhe permite, dadas as suas competências. É isso que quer dizer a tese de que a liberdade é uma questão de conhecimento: "O conhecimento científico é o bom caminho - o único caminho - para a inevitabilidade" (p.286); "Quanto mais sabemos, mais podemos fazer; quanto mais podemos fazer, mais obrigações enfrentamos" (p.312). Na liberdade determinada pelo design, a genuína deliberação é uma aparência.

Cabe perguntar: se pudéssemos saber até à exaustão as consequências das nossas acções (digamos, até ao fim da nossa vida, ou da vida dos nossos filhos e netos), que liberdade teríamos para decidir? Nessas circunstâncias, qualquer decisão nossa seria determinada pela comparação entre os nossos objectivos (quaisquer que eles fossem) e as mais remotas consequências das diferentes linhas de acção ao nosso alcance. Mesmo essa tentativa poderia falhar, em caso de valores incomensuráveis (seria preferível

preservar o património de um filho ou a saúde de um inimigo?). O que proponho é que a nossa liberdade depende do facto de que temos de fazer certas opções que são, em certo sentido, não informadas - porque não resultam de um cálculo de variantes. Aderimos a normas sociais, a linhas de conduta, a regras morais, a princípios pessoais - por vezes apesar de nos assaltarem dúvidas acerca da sua adequação a todos os casos particulares que é suposto cobrirem - graças à feliz circunstância de ignorarmos o resultado do cálculo de variantes que nos daria acesso ao conhecimento exaustivo das consequências das nossas acções. Se tivéssemos esse conhecimento massivo das consequências das nossas acções, teríamos em cada decisão, para nosso governo, uma percepção extrema da mais pequena diferença entre altruísmo e egoísmo no nosso comportamento: suportaríamos isso?

De qualquer modo, não é essa a nossa condição. A nossa adesão a regras, códigos, princípios, valores - é uma aposta. No sentido pleno do termo: em certa medida, um tiro no escuro. Não podemos prever todas as consequências de aderir a determinadas instituições (costumes, convenções, ...), mas preferimos essa adesão - porque apostamos que, em geral, isso é correcto para nós e para os que partilham connosco as comunidades em que queremos viver. Essa aposta pode falhar. Aderimos a muitas instituições sem um cálculo prévio do resultado dessa opção, porque aderimos ao que podemos compreender dos princípios que as regem e porque temos alguma confiança de que estaremos melhor servidos por essa via, embora seja impossível saber que resultados concretos obteremos desse modo: quantas pessoas recorrem aos tribunais e, afinal, perdem a sua causa?

Um aspecto importante do argumento de Dennett é que ele integra a ideia de que, tal como mostrou a investigação em psicologia, estamos muito longe do modelo do "agente racional" em que certas teorias pretendiam enquadrar os humanos. Somos "imperfeitamente racionais", "movidos no espaço das razões por algo que não são razões conscientemente avaliadas" - por muito que isso custe à tendência dos filósofos para estabelecer uma moralidade ultra-racional, "livre" de disposições animais, paixões e emoções (pp. 279,280,281-282,291). As "razões" da nossa acção são, portanto, o produto de uma colecção de mecanismos que a espécie foi adquirindo ao longo da sua história evolutiva face aos constrangimentos mais salientes no ambiente a que foi fazendo face. Contudo, embora reconhecendo a nossa racionalidade limitada, o seu esquema determinista não dá qualquer explicação convincente para que mecanismos no mesmo estágio evolutivo tenham, para situações idênticas, comportamentos tão diferenciados. Esse é o caso dos humanos num dado momento histórico - mesmo que, se quisermos dar crédito à hipótese dos memes, tenhamos que restringir a situação a um contexto cultural comparável. Porque é que, se o design é que dita a competência e cada organismo faz o melhor que a sua competência lhe permita, os padrões de comportamento são tão díspares entre iguais? Porque (esta é a nossa resposta) a liberdade assenta tanto no conhecimento (nos termos defendidos por Dennett) como na ignorância.

Num sentido talvez escandaloso para os novos iluministas, num mundo em tempo real (com necessidade de agir atempadamente) e com ignorância (conhecimento e poder computacional limitados), a liberdade é possível. Condicional, mas possível. Porque gerimos uma "margem de manobra" limitada dando "tiros no escuro" orientados por princípios que escapam à racionalidade computacional.

Porfírio Silva

Bolseiro da Fundação para a Ciência e Tecnologia (SFRH / BD / 10127 / 2002)

Referências

(Amsterdamski 1990) AMSTERDAMSKI, Stefan, "Halte aux espoirs, silence aux accusations", in POMIAN, Krzysztof (org.), *La Querelle du Déterminisme*, Paris, Gallimard, 1990pp. 228-246

(Axelrod 2000) AXELROD, Robert, "On Six Advances in Cooperation Theory", in *Analyse & Kritik*, 22, pp. 130-151

(Binmore 1998) BINMORE, Ken, "[Review of "The Complexity of Cooperation" by Axelrod](#)", in *The Journal of Artificial Societies and Social Simulation*, 1(1)

(Castelfranchi 1998) CASTELFRANCHI, Cristiano, "[Though the Minds of the Agents](#)", in *The Journal of Artificial Societies and Social Simulation*, 1(1)

(Couto 2000) COUTO, Carlos, *Introdução à Robótica Industrial*, Lisboa, Universidade Aberta, 2000

(Davis 1970) DAVIS, Morton D., *Game Theory. A Nontechnical Introduction*, Nova York, Basic Books, 1970 (existe uma tradução em língua portuguesa: *Teoria dos Jogos: Uma Introdução Não-técnica*, São Paulo, Cultrix, s.d.)

(Dawkins 1976) DAWKINS, Richard, *The Selfish Gene* (todas as referências são para a tradução portuguesa, de Ana Paula de Oliveira, *O Gene Egoísta*, Lisboa, Gradiva, 1989)

(Dennett 1995) DENNETT, Daniel C., *Darwin's Dangerous Idea. Evolution and the Meanings of Life* (todas as referências são para a tradução portuguesa, *A Ideia Perigosa de Darwin*, Lisboa, Temas e Debates, 2001)

(Dennett 2003) DENNETT, Daniel C., *Freedom Evolves* (todas as referências são para a tradução portuguesa, *A Liberdade Evolui*, Lisboa, Temas e Debates, 2005)

(Hoffmann 2000) HOFFMANN, Robert, "[Twenty Years on: The Evolution of Cooperation Revisited](#)", in *The Journal of Artificial Societies and Social Simulation*, 3 (2)

(Macy 1998) MACY, Michael W., "[Social Order in Artificial Worlds](#)", in *The Journal of Artificial Societies and Social Simulation*, 1(1)

(Prigogine e Stengers 1990) PRIGOGINE, Ilya, e STENGERS, Isabelle, "La querelle du déterminisme, six ans après", in POMIAN, Krzysztof (org.), *La Querelle du Déterminisme*, Paris, Gallimard, 1990, pp. 247-265

(Rose 1998) ROSE, Nick, "[Controversies in Meme Theory](#)", in *Journal of Memetics - Evolutionary Models of Information Transmission*, 2

(Roth 1995) ROTH, Alvin E., "Introduction to Experimental Economics", in KAGEL, John H., e ROTH, Alvin E. (eds.), *Handbook of Experimental Economics*, Princeton, Princeton University Press, 1995, pp. 3-109

(Silva 2004) SILVA, Porfirio, "Por uma robótica institucionalista: um olhar sobre as novas metáforas da inteligência artificial", in *Trajectos*, 5, pp. 91-102

(Simon 1983) SIMON, Herbert A., *Reason in Human Affairs* (todas as referências são para a versão portuguesa, *A Razão nas Coisas Humanas*, Lisboa, Gradiva, 1989)

(Thom 1990) THOM, René, "Halte au hasard, silence au bruit", in *La Querelle du Déterminisme*, Paris, Gallimard, 1990, pp. 61-78

(Turing 1936) TURING, Alan M., "On Computable Numbers, with an Application to the *Entscheidungsproblem*", in *Proceedings of the London Mathematical Society*, Série 2, 42, pp. 230-265