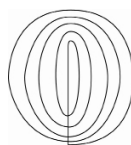


DILEMA DO PRISIONEIRO

EDIÇÃO DE 2021 do

COMPÊNDIO EM LINHA DE PROBLEMAS DE FILOSOFIA ANALÍTICA

2018-2021 FCT Project PTDC/ FER-FIL/28442/2017



Editado por
Ricardo Santos e Pedro Galvão

ISBN: 978-989-8553-22-5

Compêndio em Linha de Problemas de Filosofia Analítica
Copyright © 2021 do editor
Centro de Filosofia da Universidade de Lisboa
Alameda da Universidade, Campo Grande, 1600-214 Lisboa

Dilema do Prisioneiro
Copyright © 2021 do autor
Diogo Fernandes

DOI: <https://doi.org/10.51427/cfi.2021.0086>

Todos os direitos reservados

Resumo

O dilema do prisioneiro é um conhecido *puzzle* decisório, a análise do qual é feita, primeira e fundamentalmente, no contexto da Teoria dos Jogos. Torna-se, por isso, necessário apresentar uma tipologia dos jogos aí estudados, juntamente com as soluções propostas para os mesmos, e inserir o dilema na sua respectiva categoria. Cumprido esse objectivo, confrontar-se-á a solução que recomenda a escolha da acção dominante com outros argumentos favoráveis à escolha da acção dominada. Existindo uma versão do dilema que tem sido usada como ferramenta analítica para caracterizar certas situações de escolha social, aplicar-se-á a análise da mesma ao problema da justificação do poder político do Estado. Serão também analisadas as versões iteradas do dilema, dando-se especial atenção ao torneio de estratégias concebido por Robert Axelrod (1984), cujos resultados têm estimulado discussões importantes nos contextos das Ciências Sociais e da Biologia. Finalmente, comparar-se-á o Dilema do Prisioneiro com outro *puzzle* decisório, o Problema de Newcomb, apresentando-se aí o argumento favorável à tese de que os dois são estruturalmente idênticos.

Palavras-chave

Ponto de equilíbrio, ótimo de Pareto, dominação, estado de natureza, *Tit-for-Tat*.

Abstract

The prisoner's dilemma is a decision-theoretical puzzle the analysis of which is primarily and fundamentally made in the context of Game Theory. It will, therefore, be necessary to present a typology of games, and to place the dilemma in its respective category. Only after fulfilling this task, one can adequately confront the solution that recommends the choice of the dominant strategy with the arguments in favor of the dominated one. Since the multiplayer version of the dilemma has been used to characterize certain situations of social choice, that analysis will be used to assess the contractualist argument for the justification of the coercive power of the state. The iterated versions of the dilemma will also be considered, with an emphasis on the tournament of strategies designed by Robert Axelrod (1984), the results of which have promoted discussions in the Social Sciences and Biology. Finally, a comparison will be made between the Prisoner's Dilemma and Newcomb's problem, and the thesis that these two puzzles are structurally identical will be defended.

Keywords

Equilibria, Pareto-optimal, domination, state of nature, *Tif-for-Tat*.

Dilema do Prisioneiro

DOI: <https://doi.org/10.51427/cfi.2021.0086>

1 Introdução

Considere-se a seguinte situação:

Dois *gangsters* são presos por suspeita de assalto à mão armada e colocados em celas separadas sem possibilidade de comunicação. O procurador sabe que apenas poderá obter uma condenação caso um deles confesse o crime. Assim, ele oferece a cada um dos prisioneiros o seguinte acordo: 'Se confessares e o teu parceiro não confessar, ficas livre e ele leva a pena máxima. Se ambos confessarem, são os dois presos, mas nenhum de vós leva a pena máxima. Se nenhum confessar, trataremos de vos condenar a ambos por evasão fiscal'.

Esta situação consiste numa instância do *puzzle* decisório denominado 'dilema do prisioneiro' (DP). A questão que se impõe é a seguinte: 'O que faria em tal situação? Confessaria ou não confessaria?' Abaixo encontra-se uma tabela de decisão que representa o problema enfrentado por cada um dos prisioneiros, *A* e *B*, em que os valores numéricos correspondem aos anos de prisão que cada um receberá, consoante a combinação de estratégias seguidas:

		<i>B</i>	
		confessar	não-confessar
<i>A</i>	confessar	(3,3)	(0,5)
	não-confessar	(5,0)	(1,1)

A principal abordagem teórica a este problema tem sido feita no contexto da teoria matemática dos jogos, um empreendimento que tem unido economistas, matemáticos e filósofos. Existe, portanto, uma tipologia de jogos, analisados nesse contexto, da qual o DP faz parte. Portanto, para melhor se compreender o que está aqui em causa, é necessário percorrer parte dessa tipologia e comparar o DP com outros tipos de jogos.

Publicado pela primeira vez em 2021

Considere-se, por exemplo, o jogo conhecido por *chicken* (algo como 'medricas'): os dois jogadores encontram-se dentro de automóveis alinhados na mesma direcção, mas em sentidos contrários e a uma certa distância um do outro. Cada um acelera ferozmente contra o adversário e o primeiro a desviar o seu automóvel, evitando a colisão, passa a ser o *chicken*. Abaixo temos a tabela de decisão que representa este jogo:

		B	
		desviar	não-desviar
A	desviar	(2,2)	(1,4)
	não-desviar	(4,1)	(0,0)

Antes de qualquer abordagem técnica, uma simples análise da tabela mostra-nos várias coisas importantes. Primeiro, ambos os jogadores têm uma escala de preferências similar (com *payoffs* simétricos): ambos preferem seguir em frente e deixar o seu adversário desviar-se, o que corresponde à vitória no jogo (adquirir uma reputação de bravura); a preferência imediatamente inferior consiste em ambos desviarem simultaneamente os seus carros e evitarem a colisão; por último, resta a consequência desastrosa para ambos, a colisão frontal. Segundo, ambos os jogadores têm de ter em conta aquilo que pensam que o seu adversário irá fazer e, assim, tomar uma decisão que lhes seja vantajosa; chama-se isto escolher uma estratégia. Terceiro, duas das consequências possíveis, que correspondem à escolha de estratégias diferentes, possuem uma característica que se destaca: nenhum deles consegue sair-se melhor do que o adversário mudando unilateralmente de estratégia. Ou seja, se eu escolher a estratégia que consiste em desviar-me, e o meu adversário escolher a estratégia que consiste em seguir em frente, eu não consigo aumentar os meus ganhos mudando de estratégia, a menos que ele também o faça, e vice-versa. Aos pares de resultados que correspondem a estas duas consequências, nos cantos superior direito e inferior esquerdo da tabela, chamam-se *pontos de equilíbrio*. As questões que se colocam para ambos os jogadores são as seguintes: 'Haverá algo como uma solução para este jogo? Em caso de resposta afirmativa, em que poderá consistir tal solução?' Uma resposta adequada poderá ser a seguinte: escolher uma estratégia tal que, seja o que for que o adversário faça,

a escolha dessa estratégia permite alcançar sempre um melhor resultado do que a escolha da estratégia contrária. Neste exemplo não existe uma tal estratégia, e o problema parece resumir-se à questão de saber quem é que fica com o maior ganho, ou seja, qual dos pontos de equilíbrio corresponderá ao resultado do jogo.

A instância do DP acima apresentada partilha algumas características com o *chicken*. Primeiro, as escalas de preferências dos dois jogadores também são idênticas (com *payoffs* simétricos): ambos preferem a liberdade à pena menor, e esta à pena intermédia. Segundo, a decisão que cada um deles toma é também ela tomada de acordo com a expectativa relacionada com o comportamento do adversário. Terceiro, a análise da teoria pressupõe em ambos os casos que os agentes são racionais e possuem pleno conhecimento dos *payoffs* de todos os jogadores envolvidos, ou seja, um dos dados de partida da análise é que os jogadores são agentes racionais ideais.¹ Isto significa que a sua motivação é maximizar os seus ganhos e minimizar as suas perdas, e que, caso exista uma solução para o jogo, ambos optarão pelas estratégias que conjuntamente determinam essa solução. Não se trata aqui de adoptar um ponto de vista acerca da personalidade dos jogadores, classificando-os, por exemplo, como egoístas psicológicos, e, no caso de ser possível modelar situações reais através da análise formal, partir do princípio de que a racionalidade dos indivíduos se sobrepõe a quaisquer considerações morais. A teoria é completamente neutra a respeito deste tipo de considerações e o carácter estritamente competitivo, ou não, das situações em análise depende apenas da estrutura dos jogos em causa.

Apesar destas semelhanças, há um aspecto em que estes dois jogos divergem entre si. Se no *chicken* tínhamos dois pontos de equilíbrio, no DP temos apenas um. Se o jogador *A* confessar, duas coisas podem acontecer: 1) se o jogador *B* também confessar, ambos recebem três anos de prisão, 2) se o jogador *B* não confessar, o jogador *A* obtém a liberdade. Por outro lado, se o jogador *A* não confessar, duas coisas podem acontecer: 3) se o jogador *B* também não confessar, ambos recebem um ano de prisão, 4) se o jogador *B* confessar, o jogador *A* será condenado à pena máxima. Os factos são estes: se ambos jogam

¹ Para uma definição detalhada da condição do conhecimento comum, ver Lewis (1969: Cap. 2).

para o prémio máximo, ambos recebem apenas o terceiro resultado mais desejado; se um deles decidir jogar para o segundo prémio, duas coisas podem acontecer-lhe: ou ganha esse prémio ou expõe-se a ser explorado pela estratégia do adversário. Verifica-se, assim, que o único ponto de equilíbrio é o que corresponde à escolha mútua da estratégia que consiste em confessar, pois se algum deles alterar unilateralmente a sua decisão, alcançará um pior resultado. Assim, ao contrário do *chicken*, o DP parece ter uma solução: existe uma estratégia tal que, seja qual for a estratégia do adversário, alcança sempre melhores resultados do que a estratégia contrária. Uma tal estratégia é qualificada como 'dominante' e a estratégia contrária como 'dominada'. Sendo o princípio da dominação algo bem-establishado no contexto geral da teoria da decisão, é plausível argumentar que uma condição necessária para determinar a racionalidade de uma decisão no DP é a de não adoptar a estratégia dominada.² Logo, supondo a sua comum racionalidade, ambos os jogadores escolherão a estratégia dominante, vendo-se ambos condenados a permanecer três anos na prisão.

O dilema é gerado pelo seguinte aspecto da situação: se é racional escolher uma certa estratégia, esperar-se-ia que essa estratégia maximizasse os ganhos do agente. Contudo, a estratégia dominante não satisfaz esse objectivo, pois ambos os jogadores poderiam ter alcançado um resultado melhor caso ambos tivessem escolhido a estratégia dominada. Por outras palavras, a estratégia dominante não é uma estratégia *Pareto-ótima*. Para se compreender este conceito, considere-se uma situação em que desejamos distribuir uma certa quantidade de bens – 100 Euros – por um certo número de indivíduos – dois; neste caso, qualquer distribuição cujas parcelas somadas tenha valor 100 será uma estratégia *Pareto-ótima* (ou *eficiente*, se quisermos). Uma distribuição de 49 - 49 não será eficiente, pois um dos indivíduos pode ver a sua situação melhorada sem que o outro sofra qualquer prejuízo (uma distribuição 99 - 1 também será eficiente).

² Considere-se uma formulação um pouco mais precisa do *princípio da dominação*: entre duas acções x e y , é racional executar x em vez de y (ou que a estratégia x domina a estratégia y) caso duas condições sejam satisfeitas: 1) aconteça o que acontecer, fazer x nunca faz com que se fique pior do que fazendo y , 2) existe, pelo menos, uma consequência possível de fazer x que faz com que se fique melhor do que fazendo y .

No DP, a única combinação de estratégias que não é *Pareto-ótima* é quando ambos os jogadores escolhem a estratégia dominante; ou seja, a menos que um dos jogadores opte por não confessar, não existe qualquer outro resultado em que a situação de um deles melhore sem que a do outro piore. Assim, pode-se constatar que a combinação das estratégias dominadas obtém um resultado *Pareto-ótimo*; o mesmo é dizer que, nesse caso, não há lugar para qualquer *Pareto-improvement*, pois para que isso aconteça a situação de pelo menos um deles tem de poder melhorar em caso de alteração mútua ou unilateral de estratégias. Podemos, assim, clarificar de uma forma mais precisa a tragédia do dilema: o único ponto de equilíbrio deste jogo corresponde a um resultado ineficiente.

O DP assume, portanto, um aspecto distinto do *chicken*. Enquanto este possuía um carácter estritamente competitivo, o DP apresenta um aspecto cooperativo: se ambos os jogadores cooperarem obterão um resultado melhor para ambos do que aquele que será obtido caso ambos escolham a estratégia dominante. É por este motivo que as estratégias disponíveis são habitualmente designadas como 'não-cooperar/cooperar', ao invés de 'confessar/não-confessar', respectivamente.

2 Que tipo de jogo é o dilema do prisioneiro?

Ao procurar-se uma solução para o DP através da opção pela estratégia dominante, chegou-se à conclusão de que, afinal, essa 'solução' apresenta um carácter insatisfatório; daí que o argumento que lhe é favorável não seja tão persuasivo quanto à primeira vista poderia parecer. Mas, do ponto de vista da teoria dos jogos, qual é a razão técnica pela qual não podemos atribuir ao DP uma solução? Tal acontece porque o dilema não é um *jogo de soma-zero*, ou seja, os ganhos de um jogador não correspondem necessariamente a perdas equivalentes por parte do seu adversário. Num jogo de soma-zero, pelo contrário, se os ganhos e perdas dos jogadores (os seus *payoffs*) forem equivalentes a uma certa quantia de dinheiro, pode-se definir um jogo deste tipo como um jogo em que não se gera nem se perde qualquer riqueza. Um exemplo claro é o *poker*: começa-se com uma determinada quantia e chega-se ao final com a mesma quantia. A única diferença é que, no final, esta mudou de mãos ou encontra-se distribuída de maneira diferente. Uma outra maneira de qualificar os

jogos de soma-zero consiste em apontar que neles os jogadores não têm qualquer interesse em comum. Em suma, se estivermos perante um jogo de soma-zero, é realmente possível encontrar uma solução para o mesmo. Considere-se, para o efeito, a seguinte tabela de decisão para dois jogadores X e Y :³

		Y			
		I	II	III	IV
X	a	-3	17	-5	21
	b	7	9	5	7
	c	3	-7	1	13
	d	1	-19	3	11

Nesta tabela os valores devem ser interpretados da seguinte maneira: se X escolher a estratégia a e Y a estratégia II, X terá um ganho de 17. Se X escolher a estratégia a , mas Y escolher a estratégia III, X terá uma perda de 5. A análise da tabela mostra-nos que nenhuma das estratégias de X é dominada por qualquer estratégia de Y . Contudo, ao considerarmos as estratégias de Y , vemos que as suas estratégias I e III dominam as suas estratégias II e IV. Assim, podemos eliminar as colunas com as estratégias II e IV. Voltando a analisar a tabela, vemos agora que as estratégias a , c e d são dominadas pela estratégia b . Logo, resta a Y escolher a estratégia III, pois só desse modo conseguirá minimizar as suas perdas. Torna-se, assim, possível encontrar o ponto de equilíbrio deste jogo.

Esta maneira de encontrar pontos de equilíbrio (neste caso uma solução) para jogos de soma-zero pode ser generalizada: supondo que Y conhece as estratégias disponíveis de X , e vice-versa, Y irá escolher o valor mínimo de qualquer linha que X escolherá. Por seu lado, X irá escolher uma estratégia que obterá o máximo de entre estes valores mínimos. Este valor é chamado de *maximin*. Do mesmo modo, dado que X irá escolher o valor máximo de qualquer coluna que Y escolherá, então Y irá escolher uma estratégia que obterá o mínimo de entre estes valores máximos. Este valor é chamado de *minimax*. Quando os valores *maximin* e *minimax* coincidem, encontramos

³ Esta tabela é uma de entre várias apresentadas em Davis (1997: 12).

perante um ponto de equilíbrio e um respectivo par de estratégias de equilíbrio. O ganho (e a respectiva perda) associado a um ponto de equilíbrio corresponderá ao valor máximo na respectiva coluna e ao valor mínimo na respectiva linha (ou vice-versa). Quando é encontrado um ponto de equilíbrio num jogo de duas pessoas com soma-zero, considera-se encontrada a solução (ou o valor) desse jogo. Neste caso, o valor do jogo é 5.⁴

O argumento a favor da escolha da estratégia dominante num jogo de soma-zero com duas pessoas resulta, assim, claro. Primeiro, o jogador irá obter como resultado, pelo menos, o valor do jogo; segundo, o jogador irá impedir que o seu adversário ganhe mais do que o valor do jogo; terceiro, cada jogador, na suposição da sua racionalidade, estará motivado para minimizar as suas perdas e maximizar os seus ganhos, e, como tal, a jogar a estratégia dominante. Isto significa que o carácter satisfatório das soluções que a teoria dos jogos apresenta depende da natureza estritamente competitiva dos jogos de soma-zero. Apesar de se poder encontrar pontos e estratégias de equilíbrio para os jogos de soma não-zero, o teste do equilíbrio-*maximin/minimax* não se aplica a este tipo de jogos, resultando assim mais fraco o argumento em favor da estratégia dominante.

Será que de uma comparação entre estes dois tipos de jogos se pode extrair algum argumento favorável à estratégia cooperativa no DP? Parece difícil contestar a ideia de que a escolha de uma estratégia dominada só garante o sucesso se soubermos exactamente o que o nosso adversário irá fazer; ou seja, é importante não nos equivocarmos quanto ao *rationale* das vantagens de seguir a estratégia dominante nos jogos de soma-zero, pois tais estratégias não são boas apenas de um ponto de vista conservador ou avesso ao risco. A sua atractividade não resulta apenas da sua capacidade de evitar males maiores, mas também do facto de, independentemente da racionalidade ou irracionalidade do adversário, não parecerem existir estratégias mais

⁴ Uma apresentação informal do Teorema Minimax, segundo o qual pode ser atribuído um valor V a todos os jogos finitos de soma-zero, pode ser vista em Luce e Raiffa (1957: 71-73). Para uma demonstração detalhada, ver o Apêndice 2 da mesma obra. A primeira demonstração deste teorema foi a de John von Neumann em 1928. Com a publicação de *Theory of Games and Economic Behavior* (von Neumann e Morgenstern 1944), a área inter-disciplinar da teoria dos jogos ficou estabelecida.

atractivas. Já no DP, a irracionalidade do adversário (definida de acordo com os critérios dos que defendem a escolha da estratégia dominante) garantiria, pelo contrário, a obtenção imediata de um maior ganho, mesmo supondo, erradamente, que ele irá jogar de forma racional. Portanto, uma comparação entre os resultados da análise dos jogos de soma-zero e do DP só joga a favor da escolha da estratégia dominante.

Em suma, sempre que se tenha uma tabela de decisão com a estrutura da que se encontra abaixo, em que, sejam quais forem os valores dos *payoffs*, $x > y > z > 0$,

	A	B
A	(z, z)	(0, x)
B	(x, 0)	(y, y)

estaremos sempre na presença de um DP.⁵ É importante ser-se claro quanto a este resultado oferecido pela teoria dos jogos, pois qualquer tentativa de modelar uma situação real recorrendo à estrutura do DP encontra-se comprometida, dadas as suposições teóricas em causa, com a ideia de que o comportamento dos indivíduos numa tal situação é altamente previsível. No que diz respeito a aspectos relacionados

⁵ Esta tabela representa a forma mais simples de DP: simétrico, com dois jogadores e *payoffs* ordinais. Neste tipo de DP existe apenas um ponto de equilíbrio estrito: uma combinação de estratégias da qual um desvio unilateral piorará a situação para o jogador responsável pelo desvio. Enfraquecendo a condição acima, i.e., permitindo 'empates' no *ranking* de preferências dos jogadores (e.g., $x > y > z \geq 0$), teremos um ponto de equilíbrio num sentido mais fraco: nenhum jogador consegue melhorar a sua situação mudando unilateralmente de estratégia. Se se tratar de um jogo assimétrico, desde que o *ranking* de preferências se mantenha idêntico e a condição do *conhecimento comum* seja preservada, continuamos a ter um DP. Finalmente, se tivermos um jogo assimétrico com *payoffs* cardinais, a cooperação mútua poderá não ser *Pareto-ótima*, pois é possível que ambos obtenham melhores resultados adoptando uma estratégia mista: confessar com probabilidade p e não confessar com probabilidade $(1 - p)$ (ou vice-versa). Este tipo de jogos denominam-se PD's impuros, sendo que um PD puro consiste num PD em que a colaboração mútua é *Pareto-ótima*. Numa instância simétrica do DP, com *payoffs* cardinais, como a que foi apresentada em §1, a seguinte condição é suficiente para termos um DP puro: $y \geq \frac{1}{2}(x + 0)$. Para esta distinção, ver Kuhn e Moresi (1995) e Kuhn (2019).

com a abordagem epistemológica do DP no contexto das ciências sociais, este resultado pode apresentar, naturalmente, consequências práticas bastante importantes.

3 O dilema do prisioneiro e o contrato social

Considere-se o seguinte caso: existe uma crise de escassez de água doce no planeta Terra; se todos moderarem o seu consumo, todos poderão usufruir de uma quantidade moderada de água por um longo período. Do ponto de vista de cada um, se todos ou a maioria dos outros cooperarem, 'eu posso consumir livremente sem que ninguém saia prejudicado'. Contudo, 'se todos fizerem o mesmo que eu, os recursos esgotar-se-ão rapidamente e todos sairemos prejudicados'. A tabela abaixo representa, para cada agente, as escolhas à sua disposição, em que C consiste na acção de cooperar (poupar água), D na acção de desertar/não-cooperar (consumo imoderado), B ao benefício, T (taxa) ao custo, e n representa o número de jogadores:

	mais de n escolhem C	n ou menos escolhem C
D	B	0
C	$B+T$	T

Dado que, neste jogo, a deserção domina a cooperação, presume-se que a racionalidade dos agentes irá compeli-los a escolher essa alternativa e todos sairão prejudicados. Tal como no PN original, a cooperação de um número suficiente de agentes conduzirá a um resultado que todos preferirão quando comparado com a deserção geral. Por esta razão, é legítimo classificar este jogo como um DP com múltiplos jogadores.

O exemplo mais famoso de um jogo deste tipo é o da *caça ao veado*: vários caçadores podem tentar, isoladamente, caçar uma lebre, mas o esforço conjunto de todos é necessário e suficiente para caçar um veado. Se a parte do veado que cabe a cada um for melhor do que uma lebre, isso constituirá uma motivação para todos cooperarem entre si. Contudo, se 'todos ou a maioria dos restantes perseguirem o veado, eu sozinho conseguirei caçar uma quantidade de lebres que

é, para mim, melhor do que a parte do veado que me caberia se eu participasse na caçada'.⁶

Existe, contudo, uma diferença importante entre estes dois jogos, a qual pode ser identificada se tivermos em consideração a motivação psicológica para a deserção e as consequências da deserção para os restantes jogadores. Segundo Pettit (1986), os casos de PD's com múltiplos jogadores podem dividir-se em duas categorias fundamentais: problemas *free-rider* e problemas *foul-dealer*. Na primeira dessas categorias, em que não é necessário que todos cooperem para se obter o resultado desejado, o desertor cede àquilo que se pode designar como *free-rider temptation* (algo como 'a tentação de sair impune'), ou seja, ele tenta aproveitar-se da cooperação alheira para obter o benefício sem sofrer os custos. A contraparte negativa desta motivação é aquilo a que se pode chamar o medo de *being suckered* (algo como 'ser o otário'). Estes tipos de motivação aplicam-se ao primeiro exemplo.

Na verdade, a tabela acima consiste numa idealização de algumas situações reais bastante frequentes em que existe um número mínimo n necessário para que a cooperação obtenha sucesso; além disso, também não está a ser tido em conta que, caso as jogadas não sejam simultâneas, a 'minha' cooperação ou deserção pode aumentar ou diminuir a probabilidade de os outros virem a cooperar ou a desertar. A tabela abaixo representa um caso mais realista de DP com múltiplos jogadores, no qual está em causa a conservação de um recurso escasso:

	mais de n escolhem C	n escolhem C	menos de n escolhem C
D	B	0	0
C	$B + T$	$B + T$	T

Este caso é interessante porque, ao contrário do DP, não só a deserção já não é uma estratégia dominante, como existem agora dois resultados que constituem pontos de equilíbrio: a deserção de, pelo menos, $n - 1$ jogadores (um desvio unilateral resulta numa passagem de 0 para T), e a cooperação quando se verifica o estado do mundo que

⁶ Na verdade, este exemplo consiste numa generalização para múltiplos jogadores de um outro, com apenas dois, apresentado por Jean-Jacques Rousseau (1755).

corresponde à coluna do meio, aquilo que se designa como 'cooperação minimamente eficiente' (um desvio unilateral resulta numa passagem de $B+T$ para 0). Por este motivo, este jogo perde a sua faceta de dilema, embora a situação dos agentes não pareça menos trágica. Afinal, apesar de existirem equilíbrios *Pareto-ótimos*, não se sabe ao certo como é possível obter, por exemplo, uma cooperação minimamente eficiente, especialmente quando se trata de uma população bastante numerosa de agentes.⁷ Note-se, ainda, que num DP deste tipo, o *free-rider* não prejudica necessariamente a população de cooperadores.

O exemplo da caça ao veado, na suposição de que é necessária a cooperação de todos, constitui, de acordo com a classificação de Pettit, um problema *foul-dealer*. Neste caso, a motivação do desertor é diferente daquela que se verificava no exemplo anterior. Aqui o desertor já não é apenas um *free-rider* (ou oportunista), mas sim alguém cuja escolha de estratégia tem 'más intenções', ou seja, ele não só alcança o benefício desejado, como também prejudica necessariamente todos os outros.

Estes aspectos relacionados com a motivação dos agentes, tanto para cooperar como para desertar, permitem que se levantem questões relacionadas com as propriedades morais dos agentes, as quais podem conduzir a abordagens que interpretam o DP como um caso de conflito entre a racionalidade individual e a racionalidade colectiva; ou, se quisermos, entre a satisfação do interesse próprio e o contributo altruísta para o bem comum. Este tipo de abordagem é naturalmente acolhido por teorias normativas do tipo contratualista (Gauthier 1986). Contudo, a questão acerca do conflito entre estes dois tipos de racionalidade não tem necessariamente de ser entendida num

⁷ Para uma análise dos DP's de dois jogadores com jogadas assíncronas, ver Danielson (1992). No caso acima, se as jogadas forem assíncronas, a obtenção de qualquer um dos equilíbrios pode encontrar-se facilitada para um jogador que faz uma jogada tardia, dada a evidência que lhe é oferecida pelas jogadas anteriores. Se souber, ou tiver uma estimativa, do número total de jogadores envolvidos, ele poderá utilizar o princípio da maximização da utilidade esperada como regra de decisão.

sentido moral, podendo ser abordada apenas no contexto instrumental da satisfação das preferências de cada indivíduo.⁸

O tema do contratualismo entra em cena mediante a constatação do óbvio: para que todos ou a maioria dos agentes se encontrem motivados para cooperar em problemas como a conservação de recursos escassos ou a caça ao veado, todos exigirão uma garantia de que os outros também o farão. Uma das garantias viáveis neste tipo de situações consiste na adopção de um mecanismo que impeça, precisamente, que os desertores saiam impunes da sua não-cooperação, ou seja, uma arbitragem externa que dê garantias a todos os jogadores de que os acordos estabelecidos entre si, para vantagem mútua, serão, com elevado grau de probabilidade, cumpridos por todos. Foi essa a ideia central do argumento contratualista de Thomas Hobbes para justificar o exercício legítimo do poder coercivo do Estado.

De acordo com uma certa análise, o DP com múltiplos jogadores pode ser usado como caracterização analítica da situação do mundo denominada por Hobbes como 'estado de natureza' (Kavka 1983). A alegação de Hobbes tem uma longa história na tradição do pensamento ético-político, mais precisamente na constatação de que os desígnios da prudência, ou do interesse próprio, não se encontram em conflito com os da moral. Contudo, a moralidade não constitui, segundo ele, uma condição necessária para o surgimento das regras de cooperação mútua da sociedade política. Dados os pressupostos científicos que Hobbes aceitava, o seu argumento é adequado a uma abordagem que faz uso da teoria dos jogos. O *paradigma mecanicista* por ele adoptado, segundo o qual para compreender o funcionamento do todo é necessário separá-lo nas suas partes constituintes, é naturalmente compatível com uma abordagem à ciência social que trata, em primeiro lugar, de caracterizar os indivíduos que constituem as

⁸ Uma maneira de entender esse conflito consiste em supor que o nosso adversário no DP é uma parte temporal futura de nós próprios. Nesse caso, a estrutura do DP é preservada: o nosso eu actual tem de agir de acordo com a sua previsão acerca de como o nosso eu futuro irá agir. Uma situação que pode também ser modelada pelo DP verifica-se quando o nosso adversário é uma parte actual de nós próprios com a qual entramos em conflito: eu quero comer um doce após a refeição, mas o meu 'outro eu' não; talvez os meus 'eus' cooperem entre si e eu venha a comer uma peça de fruta. Para o tema da racionalidade intrapessoal, ver, por exemplo, Elster (1976).

unidades básicas desse mecanismo. A estrutura do argumento é simples: justificar e legitimar as vantagens alcançadas pelo poder do Estado, através de uma comparação com as desvantagens de uma situação na qual esse Estado não existe, famosamente caracterizada como um *estado de natureza*.

O recurso ao modelo da teoria dos jogos para justificar o argumento Hobbesiano pode suscitar a seguinte crítica: no contexto da teoria dos jogos não se pressupõe que os agentes são *egoístas psicológicos*, mas no contexto da teoria política de Hobbes verifica-se, de facto, que os agentes o são.⁹ É, todavia, arguível que esta é uma interpretação errada. A culpa é, sem dúvida, de várias considerações de Hobbes acerca dos elementos que caracterizariam a natureza humana (1999: 91):

Assinalo [...] como tendência geral de todos os homens um perpétuo e irrequieto desejo de poder e mais poder, que cessa apenas com a morte. E a causa disto nem sempre é que se espere um prazer mais intenso do que aquele que já se alcançou, ou que cada um não possa contentar-se com um poder moderado, mas o facto de não se poder garantir o poder e os meios para viver bem que actualmente se possuem sem adquirir mais ainda.

Contudo, a sua concomitante definição do poder dos indivíduos – os 'meios de que presentemente dispõe[m] para obter qualquer visível bem futuro' (*Ibid.* 83) – é favorável à interpretação de que o referido desejo de poder não implica necessariamente que os agentes sejam egoístas psicológicos. O único pressuposto motivacional é o de que as pessoas tentam maximizar um certo nível de bem-estar. Há um egoísmo racional predominante, relacionado com as condições do estado de natureza, que fazem com que este seja um estado de guerra de todos contra todos, mas não necessariamente um egoísmo psicológico.

Uma comparação com o contratualismo de Rawls (1971) pode ser elucidativa. No caso de Rawls, a posição original é a experiência de pensamento que, enquanto mecanismo justificatório, ocupa o lugar do estado de natureza em Hobbes. Ou seja, temos, primeiro, uma descrição das características dos indivíduos, a qual inclui igualmente a suposição da sua racionalidade, definida relativamente a um certo

⁹ Definir egoísmo psicológico é algo notoriamente difícil, mas penso que não é controverso afirmar que se trata de uma tese descritiva acerca da psicologia humana, segundo a qual toda a acção humana é motivada pelo interesse próprio (ver, por exemplo, Kavka 1983: 35-44).

conjunto de bens fundamentais; depois temos certas condições que definem as características da situação em que se encontram (o véu da ignorância); e, finalmente, um argumento que estipula qual o modo racional de agir sob essas condições: a adoção da regra *maximin*.¹⁰ Ou seja, apenas considerações de utilidade e vantagem individual são tidas em conta no argumento propriamente dito. Qualquer outro investimento moral já se encontra implícito nas condições da situação de escolha, as quais tentam modelar uma concepção de imparcialidade entre indivíduos com determinadas características morais, nomeadamente, uma concepção do bem e um sentido da justiça.

De acordo com a regra *maximin*, as opções disponíveis são avaliadas de acordo com as piores expectativas associadas a cada uma delas. Ou seja, supondo-se que se pode atribuir um valor de utilidade a cada uma das posições sociais que cada um irá ocupar, a regra diz-nos para compararmos as utilidades mínimas oferecidas por cada uma dessas situações e escolher a de maior valor. Como o próprio Rawls afirma, 'devemos escolher como se o nosso lugar na sociedade fosse escolhido pelo nosso pior inimigo' (1999: 132-33). Quando nos 'colocamos' na posição original é realmente como se estivéssemos a jogar um jogo de soma-zero contra a Natureza. Portanto, tanto no caso de Hobbes como no caso de Rawls, a motivação 'egoísta' das partes, ou a natureza prudencial das suas escolhas, é-lhes imposta pelas circunstâncias, não resultando de qualquer pressuposto acerca da natureza psicológica dos indivíduos em geral.

Qual é, então, no caso de Hobbes, a regra de decisão que é racional os agentes adoptarem para determinar o seu comportamento face às expectativas acerca do comportamento dos outros em estado de natureza? Para que o estado de natureza seja interpretado como um DP com muitos jogadores, tem de se mostrar que a não-cooperação é a estratégia dominante. A não-cooperação pode aí ser definida da seguinte maneira: fortalecer as próprias defesas, esperando que os outros também o façam, e atacar como forma de prevenção. Uma tabela de decisão para o estado de natureza seria algo do seguinte género:

10 Trata-se de uma regra utilizada para tomar decisões em condições de ignorância (ver, por exemplo, Resnik 1987, Cap.2), a qual não se encontra relacionada com a análise das soluções para os jogos de soma-zero.

	atacar	não-atacar
atacar	(1,1)	(5,0)
não-atacar	(0,5)	(3,3)

As condições objectivas que caracterizam o estado de natureza são conhecidas: não existe poder comum, verifica-se em geral uma igualdade de forças entre os indivíduos, existem preferências conflitantes devido à escassez de recursos, etc. Dadas estas condições, qual é a estratégia a seguir? A suposição que é legítimo atribuir a Hobbes é a de que não existem condições para estabelecer laços de confiança, pois basta existir uma minoria de egoístas psicológicos, que atacam sempre (seja qual for a estratégia dos adversários), para que haja rapidamente uma escalada e todos escolham a estratégia dominante (Kavka 1983: 109-125).¹¹ Estas condições são a base motivacional para criar uma autoridade comum que, para benefício de todos os agentes, se encarregue de punir aqueles que não cooperam.

Independentemente dos méritos do argumento de Hobbes, resalta da análise um aspecto que permite rever um dos dados com que o DP nos é normalmente apresentado: a impossibilidade de comunicação. Pode-se, assim, constatar que este elemento não tem qualquer influência na escolha dos agentes. A menos que haja uma maneira de garantir o cumprimento de promessas, o facto de os dois prisioneiros poderem ou não comunicar entre si é irrelevante. Afinal, o dilema não é um problema de pura coordenação, em que todas as partes envolvidas têm interesse em cooperar. Quando isso acontece, as convenções facilmente resolvem o problema (como no caso em que se tem de decidir em que lado da estrada todos devem conduzir). O DP é um jogo misto em que ambas as partes têm, por um lado, interesses em comum e, por outro, interesses individuais que conflituam entre si.

A existência de mecanismos de coacção eficazes é um assunto complexo (Schelling 1963). Se nalguns casos são relativamente fáceis de encontrar, noutras é muito difícil chegar-se a uma solução satisfatória. Por exemplo, o código de silêncio da máfia costuma ser suficientemente persuasivo para que ambos os prisioneiros não confessem, pois existe uma garantia quase certa de punição severa. Compreende-se

¹¹ Na verdade, as expressões que Kavka usa nas suas tabelas são *lie low* para 'cooperar' (não atacar) e *anticipate* para 'desertar' (atacar).

facilmente o trabalho a que uma organização criminosa se dá, e os elevados custos envolvidos, para castigar alguém que tenha quebrado esse código de silêncio, dado que a possibilidade de alguém sair impune destruiria imediatamente a eficácia dessa organização. Por outro lado, casos como os tratados internacionais de não-proliferação de armas nucleares ou de redução de emissões poluentes têm uma eficácia muito mais limitada.

4 O dilema iterado: *Tit-for-Tat* e a recompensa da cooperação

A análise do conceito de estado de natureza como um DP com múltiplos adversários pode abrir caminho para uma outra abordagem ao dilema que se tem vindo a mostrar profícua. Um aspecto que ressalta da análise em §3 consiste no facto de cada interacção não ser única, prolongando-se no tempo em reiteradas interacções. O problema que daí surge é designado como 'dilema do prisioneiro iterado' (DPI) e a sua análise tem resultado num grande número de evidência empírica, sendo que áreas de investigação como a Biologia e as ciências sociais beneficiam dos resultados dessa análise. Existem, contudo, vários tipos de DPI's; podemos ter DPI's finitos, i.e., com um número n finito de interacções, em que n é conhecimento comum; DPI's com um número infinito de interacções; e DPI's com um número n indefinido de interacções.¹²

Qual será, então, a melhor estratégia a adoptar se dois agentes se confrontarem num número finito de interacções? Podemos constatar que a deserção continua a ser a estratégia dominante e que a combinação de duas estratégias que consistem ambas em desertar em todas as interacções resulta num ponto de equilíbrio. Mas como podem os agentes perseverar nessa conduta se o resultado obtido através de um conjunto completo de cooperações é muito maior do que aquele que é obtido através de uma série constante de deserções? O argumento favorável ao emprego da estratégia dominante poderá depender da resposta à questão acerca do que se deve fazer na última interacção. Até aí as vantagens da cooperação são relativamente

12 Os DPI's infinitos são apenas uma hipótese teórica e a relevância da sua análise prende-se estritamente com a tentativa de criticar um certo de argumento em favor de estratégias a adoptar nos DPI's finitos (ver Kuhn 2019).

claras: incitar o outro jogador a cooperar e estabelecer uma relação de confiança mútua; para que a relação de confiança se estabeleça, é necessário que as duas partes demonstrem que são merecedoras de confiança; por exemplo, na primeira das interações, em que não existe ainda qualquer historial, parece não haver razões para qualquer um deles cooperar, pois o resultado directo consiste em deixar-se explorar. Por outro lado, cooperar de início pode ser encarado pelo outro como um convite a uma política de constante cooperação, cujo resultado é extremamente vantajoso para ambos. Mas, apesar de ser compensador tomar a iniciativa de cooperar, e fazê-lo daí em diante, o que será racional fazer na última interacção?

Através de um raciocínio designado como 'indução retroactiva' [*backwards induction*], cada um dos jogadores chega à conclusão de que a melhor estratégia a seguir em todas as interacções é a estratégia dominante: 'como ele terá a expectativa de que eu coopere na última jogada, ele não irá cooperar; logo, eu também não cooperarei. Mas, como ambos sabemos que não iremos cooperar na última jogada, pela mesma ordem de razões também não iremos cooperar na penúltima, e o mesmo acontecerá na jogada anterior a essa'. De uma forma mais geral, cada agente identifica qual é a decisão que é racional tomar na jogada n ; depois, com esta decisão em mente, determina qual é a decisão racional a tomar na jogada $n - 1$; com esta decisão em mente, determina qual é a decisão racional a tomar em $n - 2$; e assim sucessivamente até chegar à jogada inicial. É como se o agente enfrentasse em cada nova jogada um novo DPI que se estende por aí em diante; dessa maneira, o agente vai acrescentando, 'de trás para a frente', novas jogadas ao seu DPI, até chegar à jogada inicial.

Será que esta é uma conclusão inevitável? É claro que na vida real existem outros dados, para além do historial de interacções, de que os agentes dispõem para determinar o quão de confiança alguém poderá ser. Mas, nas condições do dilema, o argumento a favor da deserção continua a ser forte. Apesar disso, a maioria de nós terá a intuição de que na prática, principalmente em jogos bastante longos, não seria razoável adoptar a estratégia dominante. Considere-se, por exemplo, Selten (1978):

(...) a teoria indutiva é a que se encontra correcta de acordo com a teoria dos jogos [*game theoretically correct*], mas a teoria cooperativa parece ser um melhor guia para a acção [*better guide to practical behaviour*].¹³

De modo a resolver este conflito entre o que é racional e o que parece razoável fazer num DPI finito, uma alternativa consiste em jogar *estratégias condicionais*, nas quais a decisão acerca do que fazer numa determinada jogada depende daquilo que o adversário fez na jogada anterior ou numa série particular de jogadas anteriores, o que certamente vai ao encontro de uma concepção mais realista de racionalidade, sem que, com isso, seja necessário relaxar as restrições impostas aos jogadores num DPI típico. Mais, se esta for a alternativa adequada para lidar com um DPI finito, então ela será, provavelmente, a alternativa adequada para lidar com os DPI's indefinidos, em que o número de interacções é desconhecido de ambos os jogadores ou da totalidade dos indivíduos num determinado ambiente de jogadores.¹⁴

Nos DPI's indefinidos não existem bases para aplicar o raciocínio por indução retroactiva, embora se coloque também a questão de saber qual é a melhor estratégia a adoptar. O sucesso dessa estratégia irá depender das estratégias dos adversários, acerca das quais não temos condições para formar expectativas razoáveis.¹⁵ Sabemos apenas

13 Pettit e Sugden (1989) argumentaram no sentido de mostrar que o raciocínio por indução retroactiva é incoerente. Segundo eles, não há razões para acreditar que a minha crença inicial, acerca do comportamento racional do meu adversário na última jogada, permaneça intocada ao longo de todos os passos indutivos até à jogada inicial. Ou seja, o argumento só corre se se pressupuser, por exemplo, que mesmo que eu coopere na jogada inicial, o meu adversário acreditará que desde a última até à segunda jogada eu irei agir racionalmente. Contudo, o que o argumento tenta mostrar é precisamente que, no caso de se pressupor que agirei racionalmente na última jogada, então também o farei na primeira. Portanto, dada apenas a crença comum inicial na sua racionalidade, não se segue necessariamente daí que os jogadores tenham razões para acreditar que essa crença sobreviva ao longo do tempo, independentemente do que tenha acontecido no passado. Para uma resposta a esta objecção, ver Sobel (1993).

14 Podemos definir 'ambiente de jogadores' como o conjunto das estratégias adoptadas, numa determinada população, para enfrentar um DPI indefinido.

15 Na verdade, caso exista um limite superior para o número de jogadas, e se esse limite for conhecimento comum na população de jogadores, então é possível usar o argumento da indução retroactiva a favor da estratégia dominante

que, para cada ambiente de jogadores, existirá uma estratégia que amearhará mais pontos de utilidade. Nada nos garante, contudo, que uma estratégia vencedora num determinado ambiente consiga obter o mesmo sucesso num ambiente diferente. Por exemplo, num ambiente maioritariamente composto por estratégias que jamais respondem ao historial de interações, a deserção incondicional pode revelar-se adequada. Contudo, uma tal estratégia jamais conseguiria obter bons resultados num ambiente composto por um grande número de estratégias cooperantes. Numa analogia relevante, um indivíduo de uma certa espécie pode ter sucesso e prosperar num determinado ecossistema, mas, quando deslocado para outro, em que as aptidões para a sobrevivência são diferentes, esse sucesso pode transformar-se em falhanço e o indivíduo poderá acabar por não sobreviver. Ora, tal como existem indivíduos que conseguem prosperar em diferentes ecossistemas, existirão igualmente estratégias cujas características lhes permitirão prosperar numa grande variedade de ambientes.

Uma das abordagens ao DPI foi a experiência levada a cabo por Robert Axelrod (1981, 1984), em que este pediu a vários especialistas, oriundos de áreas científicas distintas, que desenhassem um programa de computador para competir num DPI. Cada programa confrontaria cada um dos outros num total de 200 jogadas, consistindo o objectivo em obter o maior número possível de pontos. Como as regras da competição preservavam os dados iniciais do problema, os resultados obtidos poderiam mostrar-se reveladores.¹⁶ Desse modo, foi constituído

(Carroll 1987). Dado que todos vivemos vidas finitas, constringidas pelas leis da Física, em última instância é sempre possível, de um modo mais ou menos razoável, atribuir um limite superior a cada DPI. Independentemente da plausibilidade desta ideia, um pressuposto menos forte é o de que existe uma probabilidade p , tal que em cada ronda existe essa probabilidade de o jogo continuar. Nesse caso, à medida que p se aproxima de 0, o PDI indefinido vai-se transformando num PDI finito (e o valor subjectivo da deserção aumenta), e à medida que p se aproxima de 1, o PDI indefinido vai-se transformando num PDI infinito (e o valor subjectivo da deserção diminui).

¹⁶ Duas cooperações seguidas eram sempre melhores do que a alternância entre cooperação e deserção — agora coopero eu, a seguir cooperas tu —, mas também, dada a cooperação do adversário, duas deserções seguidas eram sempre melhores do que duas cooperações. Continuava a verificar-se a condição ' $y \geq \frac{1}{2}(x + 0)$ ', mas y correspondia agora à cooperação mútua (3 pontos), x à vantagem do desertor (5 pontos) e z à deserção mútua (1 ponto).

um ambiente de jogadores em que os 'traços de carácter' de cada programa determinaram as suas condições de sucesso ou insucesso. Do ponto de vista da análise, as questões relevantes a extrair da experiência de Axelrod são as seguintes: 1) Pode realmente emergir cooperação num mundo de egoístas racionais? 2) Como pode começar de todo essa cooperação? 3) Que tipo de estratégias produzem maior sucesso (quais as suas características salientes)?

O programa vencedor, entre os catorze participantes, denominado *Tit-For-Tat* (algo como 'pagar na mesma moeda') foi apresentado por Anatol Rapoport, o qual já tinha anteriormente sugerido o emprego de estratégias condicionais nos DPI's (Rapoport *et al.* 1976). Uma característica formal do programa consistia no facto de ser o mais simples de todos. A sua estratégia consistia simplesmente no seguinte: cooperar sempre na primeira interacção e, a partir daí, fazer sempre na interacção seguinte o que o adversário tinha feito na anterior. Outros programas eram bastante mais complexos: poderiam introduzir uma jogada determinada pelo acaso após um certo número de jogadas, tentar sondar o adversário, reagir consoante este se mostrasse mais ou menos cooperativo, ou retaliar uma deserção com maior ou menor número de deserções, etc. Em suma, foram empregues algoritmos de alguma complexidade, mas nenhuma estratégia conseguiu bater a *Tit-For-Tat* (TFT).

Na avaliação dos resultados importa notar dois aspectos importantes. Primeiro, a característica saliente da TFT era a de cumprir sempre a sua parte no estabelecimento das bases de confiança de uma relação futura. Segundo, apesar de ter ganho o torneio com uma margem considerável, ela nunca venceu um único embate individual. Jogando contra si própria, a TFT teria empatado, pois ela e a sua versão gémea teriam começado por cooperar e tê-lo-iam continuado a fazer até à jogada número 200, acabando com 600 pontos cada. Teria empatado também contra uma estratégia que começasse por não cooperar e depois o fizesse sempre até ao fim. Em suma, num mundo ou ambiente de *soma não-zero*, os conceitos de sucesso e de vitória não dependem da capacidade de ganhar em cada uma das interacções. Num primeiro nível de raciocínio estratégico, é de certa maneira óbvia a vantagem de não cooperar; num segundo nível, em que se tem de lidar com possíveis retaliações, ainda é possível conceber-se estratégias que tentam impor-se a outras que sejam demasiado

complacentes. Contudo, num terceiro nível, os ecos das tentativas mútuas de exploração podem transformar-se numa série infindável de recriminações mútuas e, nesse caso, num ambiente em que existem algumas estratégias cooperantes que retiram grandes proveitos da interacção entre si, os ganhos destas podem compensar os prejuízos infligidos pelas estratégias mais agressivas.

Outras conclusões importantes foram as seguintes: face a uma estratégia que consiste em nunca cooperar, não há outra estratégia melhor senão também nunca cooperar. Uma tal estratégia nunca teria tido hipótese de obter sucesso, pois representa, afinal, o nível de raciocínio estratégico mais elementar num DPI; por outro lado, contra uma estratégia autista, que não reage de qualquer forma à interacção e se comporta como se estivesse a jogar sozinha, a melhor estratégia a adoptar consiste também em nunca cooperar. Isto resulta do facto de nunca se saber o que esperar de uma tal estratégia e, portanto, de não haver esperança de retirar qualquer proveito de uma interacção que nunca visa a cooperação. Talvez não seja exagerado retirar da análise dos resultados uma conclusão que vai ao encontro da tese, já considerada em §3, de que a prudência não se opõe à moralidade. De qualquer modo, é possível retirar algumas lições importantes desta experiência. Essas lições dizem respeito às características que tornaram a TFT numa estratégia de sucesso (Axelrod 1984: Parte II):

1. Simpatia [*niceness*]: a TFT cooperava sempre na primeira jogada.
2. Capacidade de perdoar [*forgiveness*]: castigava uma não-cooperação na jogada seguinte, mas voltava sempre a cooperar caso a outra também voltasse a fazê-lo.
3. Provocabilidade [*provocability*]: nunca ignorava as não-cooperações do adversário, retaliando sempre.
4. Clareza [*clarity*]: não era complexa ao ponto de ser inescrutável para o outro, nunca passando, assim, por autista.

Mas a característica talvez mais importante revelou-se depois de realizado um novo torneio, o qual testou seriamente as qualidades da TFT. Este contou com 62 participantes, incluindo duas estratégias que teriam ganho caso tivessem entrado no primeiro torneio, informação esta que foi divulgada, juntamente com os resultados, a todos aqueles que submeteram novas estratégias. A nova vitória da TFT mostrou,

mais uma vez, que o sucesso depende do ambiente em que se está inserido. Muitos dos novos participantes desenharam estratégias em que um dos principais objetivos era vencer essas duas estratégias referidas. O resultado realçou a característica que faltava a estas últimas, mas que veio a verificar-se que a TFT possuía: robustez. Assim, uma estratégia pode ser considerada robusta quando tem a capacidade de alcançar sucesso em vários ambientes diferentes. A matriz (63×63) que representou os resultados deste segundo torneio pôde ser utilizada para simular vários outros torneios, alterando-se o ambiente de uns para outros.¹⁷ Após Axelrod ter descoberto que havia seis tipos gerais de estratégias, entre as 63 participantes, passou então a realizar seis torneios hipotéticos em que predominavam, em cada um deles, um desses seis tipos. Mais uma vez, a TFT obteve um sucesso considerável: ganhou cinco desses torneios e no sexto ficou em segundo lugar.¹⁸

Um dos torneios mais interessantes foi o chamado 'torneio ecológico', também designado como 'DP evolucionário' (DPE), em que o ambiente de cada ronda sucessiva era determinado pelo sucesso de cada estratégia nas rondas imediatamente anteriores. Por outras palavras, o insucesso de uma estratégia numa ronda poderia conduzi-la à extinção, enquanto o seu sucesso seria promovido nas rondas seguintes com a possibilidade de 'reprodução'. Se, no início, algumas estratégias 'antipáticas' floresceram juntamente com as estratégias cooperativas, à medida que algumas estratégias com características mais débeis foram sendo extintas, o declínio destas arrastou consigo o das estratégias antipáticas, as quais encontravam cada vez menos

17 Em cada torneio participou sempre uma estratégia adicional, a *Random*, cujas jogadas eram sempre determinadas pelo puro acaso.

18 Após a experiência de Axelrod, têm vindo a ser apresentadas novas estratégias, com características parecidas com as da TFT, que obtiveram resultados semelhantes ou melhores em ambientes parecidos com os de Axelrod (ver, por exemplo, Nowak e Sigmund 1992; Beaufils *et al.* 1997). Mais recentemente, tem surgido evidência de que o sucesso da TFT dependeu em grande parte das condições iniciais estabelecidas por Axelrod; por exemplo, Rapoport *et al.* (2015) descobriram que dividindo a população de estratégias em sub-ambientes determinados ao acaso, cujos vencedores se confrontam numa ronda final, duas estratégias obtiveram maior sucesso do que a TFT. Para uma posição céptica mais geral acerca das vantagens da TFT, ver Binmore (2015).

possibilidades de explorar um adversário. Ou seja, num ambiente dinâmico como este, a capacidade de algumas estratégias de se darem bem quando interagem com outras mais fracas acaba por levá-las a derrotarem-se a si próprias no longo prazo. Portanto, uma população de antipáticos pode, com o tempo, vir a ser dominada por pequenos aglomerados iniciais de cooperantes simpáticos. Em suma, a cooperação pode, de facto, emergir num mundo de egoístas racionais; e, de maneira gratificadamente assimétrica, uma população em que predomina a simpatia e a cooperação resiste com robustez à tentativa de penetração e domínio por parte de estratégias antipáticas ou exploradoras, provando assim a sua estabilidade.

Este resultado da experiência gerou entusiasmo não apenas no campo das ciências sociais, mas também no da Biologia. Alguns autores tentaram mostrar de que modo o DP poderia ser aplicado, como ferramenta analítica, ao estudo do processo evolutivo e do modo como este favorece os comportamentos altruístas (Singer 1981, Kavka 1983).¹⁹ Axelrod sugere, com efeito, que as estratégias evolutivas de sucesso podem ser comparáveis às estratégias cooperativas no dilema do prisioneiro, tanto na preservação de certo tipo de genes numa determinada população, como no estabelecimento de relações simbióticas, as quais se mantêm enquanto pelo menos um dos elementos de uma relação sair beneficiado. Uma das conclusões é a de que a capacidade de reconhecer conscientemente outros indivíduos com os quais existe interacção não é sequer uma condição necessária para a cooperação, bastando para isso verificar-se uma interacção continuada entre dois organismos com mecanismos para o reconhecimento mútuo, como se verifica ao nível das bactérias. Ainda assim, há razões para acreditar que as condições impostas por Axelrod são demasiado irrealistas para justificar uma aplicação directa do PD, enquanto mecanismo explicativo, a situações reais de interacção bio-social. De certa forma, a artificialidade de alguma dessas condições é notória: 1) em condições que não se encontram livres de ruído, i.e., quando se inclui no modelo a possibilidade de erro, é expectável que os resultados

¹⁹ Note-se, contudo, que este tipo de análise é controversa e tem sido contestada (ver, por exemplo, Haji 1992). Sober (1988) tinha já defendido que o conceito de altruísmo usado pelos biólogos nos seus estudos evolutivos (Trivers 1971) era diferente daquele que é habitualmente usado em filosofia moral.

obtidos pela TFT sofram alterações consideráveis; 2) o DPE incluiu apenas as 63 estratégias que participaram no torneio inicial, em que se verificava um ambiente algo 'rarefeito', constituído apenas por especialistas, o que certamente não aconteceria em situações reais; 3) as únicas estratégias autorizadas a participar em cada uma das rondas foram apenas as que sobreviveram às rondas anteriores; em situações reais é de esperar que em cada ronda surjam estratégias mutantes que, por um lado, podem apresentar melhorias relativamente àquelas que se extinguiram e, por outro, contribuir para a sobrevivência de estratégias antipáticas, caso venham a apresentar características ingénuas ou demasiado simpáticas.²⁰

No que diz respeito à questão de saber qual a estratégia mais racional a empregar em DPI's finitos ou indefinidos, os resultados da experiência de Axelrod podem ser importantes. Mesmo num DP com apenas dois jogadores, o mínimo que se pode perder numa tentativa de cooperação não reciprocada é o ponto correspondente à 'punição do egoísmo'. Se, por outro lado, conseguirmos com isso alcançar algumas cooperações bem-sucedidas, é muito possível que os três pontos assim obtidos, correspondentes à 'recompensa da cooperação', venham no final a compensar uma série de 'teimosas' deserções, mesmo que o nosso adversário acabe por ganhar mais pontos do que nós. Se tivermos o espírito necessário para encarar esta 'derrota' como uma espécie de vitória, principalmente quando essa derrota não implica o desastre, podemos vir a sair beneficiados. É isso que acontece na vida real quando nenhum dos envolvidos tem uma fraqueza de espírito tal que não consiga suportar, uma vez por outra, 'a recompensa do otário'.

5 Soluções cooperativas para o dilema do prisioneiro

Da análise em §1 e §2 pode retirar-se uma conclusão que alguns poderão considerar insatisfatória: dois agentes 'irracionais' (segundo uma dada interpretação), que decidam cooperar, obterão um resultado melhor do que o resultado obtido por dois agentes racionais. Embora esta seja a raiz do dilema, ou o seu aspecto trágico, será que

²⁰ Para uma análise de torneios simulados que admitem a violação das características mencionadas, ver Nowak e Sigmund (1992).

a racionalidade tem necessariamente de ser garantia de obtenção do melhor resultado?

Davis (1977) sugere um exemplo que ilustra esta infelicidade aliada ao exercício da racionalidade. Suponha-se que um agente é forçado a escolher entre passar uma hora com uma cobra venenosa ou passar uma hora com um leão. A escolha racional parece ser a cobra, pois apesar de existir uma enorme probabilidade de sermos mordidos, podemos sempre contar com o antídoto, enquanto ser mordido pelo leão é garantia de morte certa. No entanto, o leão pode estar a dormir e escusávamos de passar pelas dores excruciantes associadas à mordida da cobra. Do mesmo modo, no DP, a opção pela escolha dominante parece equivaler à escolha da cobra: se o outro prisioneiro decidir não confessar, e nós fizermos o mesmo, escusamos de cumprir uma pena 'excruciante' de 5 anos. Mas o que transforma o DP num verdadeiro paradoxo, para a solução do qual existem dois argumentos contrários, é o pressuposto de que não estamos a lidar com um agente irracional ou imprevisível, como uma cobra ou um leão, mas com outro agente racional igual a nós.

Abandonando o argumento favorável à escolha dominante, somos aparentemente conduzidos a um raciocínio circular. Para sabermos qual é a alternativa racional, precisamos de estar seguros quanto às consequências de ambas as alternativas. Mas, para sabermos quais são essas consequências, precisamos de saber o que o outro irá fazer. O que o outro irá fazer depende daquela que para ele é a alternativa racional. Mas o que para ele é racional fazer depende de quais são as consequências das suas alternativas. Para ele saber quais são essas consequências, precisa de saber o que nós iremos fazer. Para sabermos o que iremos fazer... E assim por diante.

Haverá maneira de sair deste círculo? Afinal, sabe-se que ambos os prisioneiros são racionais e que, na suposição dessa racionalidade, ambos optarão pela mesma alternativa. Essa alternativa só pode ser aquela que maximiza a sua utilidade. Davis (1985: 53) apresenta uma versão do argumento favorável à cooperação, incorporando nas premissas 1 e 2 abaixo os princípios necessários para se fugir à circularidade:

1. Uma alternativa X encontra-se racionalmente prescrita para um agente y , se y sabe que existem apenas duas consequências possíveis m e n , tal que se y fizer X , a consequência é m , se y não fizer X a consequência é n , e m é melhor do que n .
2. Cada prisioneiro sabe que ambos optarão pela alternativa racionalmente prescrita.
3. Cada um sabe que uma alternativa se encontra racionalmente prescrita para um deles, apenas se também se encontrar racionalmente prescrita para o outro.
4. Cada um sabe que manterá o silêncio, apenas se o outro o fizer, e que confessará apenas se o outro o fizer.
5. Cada um sabe que, se o silêncio for racionalmente prescrito, a consequência será (1,1); e que, se confessar for racionalmente prescrito, a consequência será (5,5).
6. Cada um sabe que (1,1) e (5,5) são as únicas consequências possíveis.
7. Cada um sabe que ambos sabem que (1,1) é melhor do que (5,5).
8. Logo, manter o silêncio é a alternativa racionalmente prescrita para cada um.

A estratégia do argumento consiste em raciocinar a partir da suposição de que existe uma alternativa racionalmente prescrita, mesmo sem se saber ainda que alternativa é essa.²¹ Para que o argumento possa prosseguir é necessário pressupor-se não só que existe essa alternativa, mas que ambos sabem que ambos irão escolhê-la, i.e., que (2) é verdadeira. Ora, é difícil determinar se, fora do contexto idealizado da teoria, a crença na verdade de (2) é tão implausível quanto, à primeira vista, poderá parecer, implausibilidade essa que poderá roubar ao dilema alguma relevância prática. Com efeito, a definição das circunstâncias tem de ser suficientemente minuciosa para se poder determinar se estamos ou não perante um verdadeiro DP. Se incluirmos nos dados do problema, por exemplo, que não há honra

²¹ Suponha-se que não existe uma solução racional para o dilema. Nesse caso, deixa de haver qualquer segurança acerca do que o outro irá fazer. É como se estivéssemos novamente a jogar o dilema com uma cobra ou um leão. Como foi visto, nessa situação, a alternativa racional consiste em optar pela acção dominante, o que, por sua vez, contraria o pressuposto de que não existe uma solução racional.

entre bandidos', tal constitui uma circunstância relevante, pois se o código de conduta da mafia estiver em vigor, então ter-se-ia de atribuir diferentes *payoffs* às consequências. Nesse caso, a natureza do problema alterar-se-ia, deixando de se poder classificar como um DP. De qualquer modo, pressupondo que estamos a jogar contra uma espécie de réplica nossa – o que é consistente com os pressupostos da teoria dos jogos – o argumento de Davis poderá ser convincente.²²

Douglas Hofstadter (1983) apresentou considerações que podem ser interpretadas como uma versão do argumento de Davis. Inspirado pela experiência de Axelrod, Hofstadter decidiu convidar vinte amigos para jogarem entre si uma ronda do dilema, cada um deles contra cada um dos outros, amigos esses cujo brilhantismo não merecia contestação:

És muito inteligente [*bright*]. Os outros também são. São todos igualmente brilhantes [*bright*], diria eu (*Ibid.* 14).

Cada um dos jogadores deveria responder-lhe por carta apenas com uma letra: *C* para cooperar e *D* para desertar. Cada um deveria ainda oferecer uma breve justificação da sua resposta. Os *payoffs* eram os seguintes: 3\$ para a cooperação mútua, 1\$ para a deserção mútua, e, em caso de respostas diferentes, 0\$ para o cooperador e 5\$ para o desertor. A esperança de alguns jogadores era, obviamente, ser o único desertor, o que garantiria o prémio máximo de 95\$, contra os 54\$ (18×3) de cada um dos cooperadores.

Para Hofstadter, os resultados não só foram desapontantes, mas também contrários à sua expectativa: 14 desertores e 6 cooperadores. As razões apresentadas pelos primeiros podiam, quase na sua totalidade, ser reconduzidas à escolha da acção dominante. Por exemplo, Robert Axelrod, que defendeu uma solução cooperativa para os DPI's, desertou sem a mínima hesitação. Outro dos desertores respondeu simplesmente da seguinte maneira: 'Hofstadter, passa-me os meus 19\$!' Este valor corresponde ao prémio mínimo que cada um dos participantes receberia caso todos os vinte desertassem. Mas pior ainda, de acordo com a perspectiva de Hofstadter, a maioria dos cooperadores justificou a sua escolha pelas 'razões erradas', nomeadamente,

²² Talvez seja interessante mencionar que experiências levadas a cabo com pares de gémeos revelam que estes apresentam uma maior tendência para cooperar do que quando se trata de meros estranhos (Segal e Herschenberger 1999).

através do recurso a considerações de ordem moral.²³ Apesar destes resultados, Hofstadter defendeu a escolha cooperativa. Partindo de uma expectativa mútua de racionalidade, e pressupondo que existe de facto uma solução racional para o dilema, restava aos participantes, analisando os dados do problema, maximizarem os seus ganhos:

Um qualquer número de pensadores idealmente racionais, colocados perante a mesma situação, e experimentando de igual modo a agonia que acompanha este tipo de raciocínio, chegará inevitavelmente à mesma conclusão, desde que a justificação última da sua escolha resida apenas no exercício da sua racionalidade. Caso contrário, a racionalidade seria algo de subjectivo e não uma coisa objectiva, como é o caso da aritmética. [...] Se aceitarmos isto, teremos percorrido 90% do caminho. Tudo o que necessitamos agora de saber é o seguinte: 'Dado que todos vamos apresentar a mesma [solução], qual destas será a mais lógica? Ou seja, que mundo é melhor para o pensador racional *individual*: um mundo constituído apenas por cooperadores ou outro constituído apenas por desertores?' A resposta é imediata: 'Obtenho 57\$ se todos cooperarmos e obtenho 19\$ se todos desertarem. Eu prefiro claramente 57\$, logo este pensador racional particular prefere cooperar. Como eu sou um caso típico, *todos* os pensadores racionais devem preferir cooperar' (Hofstadter 1983: 22).

Segundo Hofstadter, os resultados do jogo deveram-se ao facto de os participantes não terem levado a sério a expectativa mútua de racionalidade, a qual foi sublinhada quando no convite ele mencionou o brilhantismo intelectual de todos os participantes. Levar a sério esse aspecto não é exactamente o mesmo que efectuar uma simulação mental do raciocínio de cada um dos outros participantes. Por exemplo, eu posso ponderar as razões favoráveis a cada uma das escolhas e considerar que durante um terço do tempo que levo a reflectir eu desejo cooperar e durante dois terços de tempo desejo desertar. Se me encarar como um participante típico, concludo então que dois terços dos meus pares vão desertar e, efectuando um cálculo simples, concludo que, para maximizar os meus ganhos, devo desertar. Contudo, independentemente do número de cooperadores que estimamos que venha a haver, este tipo de raciocínio levará sempre à mesma conclusão. Isto significa que uma simulação deste tipo não é

23 Dan Dennett, um dos cooperadores, justificou a sua escolha da seguinte maneira: 'Prefiro ser o indivíduo que comprou a ponte de Brooklin do que aquele que a vendeu. Do mesmo modo, sentir-me-ia melhor a gastar 3\$ ganhos através da cooperação do que 10\$ ganhos desertando' (*Ibid.* 27). Tal como afirmou outro dos cooperadores, o psiquiatra Charles Brenner, Dennett provavelmente não quis ver o seu nome incluído numa lista de desertores publicada numa revista de grande circulação como a *Scientific American*.

adequada para caracterizar o pensamento estratégico. Para que isso aconteça, é necessário compreender que, para além de idealmente racionais, todos os outros jogadores são também *super-rationais*: cada pensador racional deve ter em conta que cada um dos outros pensadores racionais deve ter em conta que cada um dos outros pensadores racionais deve ter em conta... e assim *ad infinitum*. Portanto, se levarmos a sério a ideia de que o nosso adversário no DP é uma espécie de réplica nossa, então é possível concluir que se deve cooperar, escapando-se assim ao mecanismo recursivo descrito.

6 Será o dilema do prisioneiro um Problema de Newcomb?

Um outro *puzzle* decisório bastante discutido na literatura sobre racionalidade prática é o Problema de Newcomb (PN) (Nozick 1969).²⁴ Cada um dos *puzzles* contém um aspecto essencial à sua natureza paradoxal, e esses aspectos, embora distintos, parecem estar relacionados entre si. Tal relação levou Lewis (1979) a defender que o DP não é mais do que uma versão do PN. No que respeita à tentativa de encontrar uma solução para o DP, se se conseguir mostrar que o DP e o PN são o mesmo problema, então a defesa de uma determinada solução para o PN será uma defesa da solução equivalente para o DP.

Com efeito, parece existir uma relação próxima entre o pressuposto da racionalidade mútua dos agentes no DP e a elevada taxa de sucesso do previsor no PN. Diríamos que estes dois aspectos são o espelho um do outro: 'tal como eu espero que no DP o meu adversário aja da mesma maneira que eu (ver o argumento pró-cooperação em

²⁴ Para uma discussão detalhada deste problema, ver (neste compêndio) Fernandes (2020). A seguinte é uma das possíveis formulações: temos perante nós duas caixas: uma delas é transparente e contém 1,000 Euros; a outra é opaca e contém 1,000,000 Euros ou nada. É-nos dado a escolha entre ficarmos com as duas caixas ou apenas com a caixa opaca. O conteúdo da caixa opaca é determinado da seguinte maneira: existe um previsor cujas previsões têm uma elevada taxa de sucesso. Ele não só previu correctamente todas as nossas decisões passadas, como tem vindo a prever correctamente as decisões de outros indivíduos mais ou menos semelhantes a nós. Se o previsor previu que iremos escolher as duas caixas, ele deixou a caixa opaca vazia. Se o previsor previu que iremos escolher apenas a caixa opaca, ele colocou 1,000,000 Euros na caixa opaca. O que devemos fazer para satisfazer a nossa preferência, sendo que esta consiste em adquirir a maior quantia possível?

§5), espero também no PN que, seja qual for a minha escolha, o previsor a preveja correctamente'.

Se isto for suficiente para tornar o DP e o PN semelhantes, então a suposta implausibilidade da premissa dois em §5 (que cada prisioneiro sabe que ambos optarão pela alternativa racionalmente prescrita) não retira qualquer importância teórico-prática ao DP; ou seja, dada a relevância epistemológica inegável do PN (ver Fernandes 2020, §2), as idealizações da teoria dos jogos não roubam qualquer relevância ao DP.

Um ponto em comum entre estes dois problemas é que em ambos se verifica um apelo controverso à dominação. Pressupondo que os prisioneiros são uma espécie de réplica um do outro, podemos redefinir o DP como confronto entre esse apelo à dominação e o princípio da maximização da utilidade condicional esperada (Nozick 1969, Jeffrey 1983). Colocando-nos na perspectiva de um dos jogadores, e considerando que as utilidades de cada consequência são lineares com os anos de prisão, temos que

$$\begin{aligned} \text{UCE (confessar)} &= u(5,5) \times pr((5,5)|\text{eu confesso}) + \\ &+ u(0,10) \times pr((0,10)|\text{eu confesso}). \end{aligned}$$

Como $pr((0,10)|\text{eu confesso}) = 0$, pois ambos faremos a mesma coisa, e como, pelo mesmo motivo, $pr((5,5)|\text{eu confesso}) = 1$, a utilidade de confessar é 5. Por outro lado,

$$\begin{aligned} \text{UCE (cooperar)} &= u(10,0) \times pr((10,0)|\text{eu coopero}) + \\ &+ u(1,1) \times pr((1,1) | \text{eu coopero}). \end{aligned}$$

Como $pr((10,0)|\text{eu coopero}) = 0$, pois ambos faremos a mesma coisa, e como, pelo mesmo motivo, $pr((1,1)|\text{eu coopero}) = 1$, a utilidade de cooperar é 1. Como a utilidade é aqui inversamente proporcional ao tempo de prisão, segue-se que a utilidade de cooperar é maior do que a utilidade de confessar.

Comparando as tabelas de decisão dos dois problemas podemos constatar, desde logo, que ambos têm a mesma estrutura. Sejam quais forem os *payoffs*, desde que $x > y > z > 0$, estaremos sempre na presença de um DP. Na matriz do PN, sabendo-se que a utilidade dessas consequências é inversa à do dilema, tal que $0 > z > y > x$, temos o seguinte:

	caixa opaca vazia	1 milhão na caixa opaca
duas caixas	1000 = z	1,001,000 = 0
caixa opaca	0 = x	1,000,000 = y

Um outro aspecto comum aos dois *puzzles* consiste na independência causal que se verifica entre os estados do mundo (colunas) e as acções disponíveis (linhas). No DP, a minha decisão não tem qualquer influência causal sobre a decisão do meu oponente; no PN, a minha decisão não tem qualquer influência causal sobre o conteúdo da caixa opaca.

Finalmente, ambos partilham um terceiro aspecto: os estados do mundo não são evidencialmente independentes das acções. A minha decisão de confessar é um sinal muito forte de que o meu oponente irá confessar, o mesmo acontecendo com a minha decisão de cooperar. No PN, a minha decisão de escolher as duas caixas é um sinal muito forte de que o previsor deixou a caixa opaca vazia, enquanto a minha decisão de escolher a caixa opaca é um sinal muito forte de que o previsor colocou um milhão na caixa opaca. Este terceiro aspecto é específico daqueles problemas que Sobel (1985) caracteriza como PN's de Previsão Certa [*near-certainty*] e DP's de expectativa comum [*near-certainty*], nos quais, respectivamente, a taxa de sucesso do previsor é quase um e a probabilidade de o outro prisioneiro agir como eu é quase um. Temos, assim, justificada a relação de simetria entre a elevada taxa de sucesso do previsor e o pressuposto da racionalidade comum dos prisioneiros. Estes três factores são responsáveis por tornar os dois problemas estruturalmente idênticos.

Consideremos uma nova matriz, comum aos dois problemas:

	E ₁	E ₂
A ₁	A ₁ ∧ E ₁	A ₁ ∧ E ₂
A ₂	A ₂ ∧ E ₁	A ₂ ∧ E ₂

Dada a dependência evidencial dos estados relativamente às acções, temos, para ambos os *puzzles*, as seguintes igualdades:

$$pr(E_1|A_1) \cong 1 \cong pr(E_2|A_2),$$

$$pr(E_2|A_1) \cong 0 \cong pr(E_1|A_2),$$

das quais se seguem:

$$pr(A_1 \wedge E_1) \cong 1 \cong pr(A_2 \wedge E_2),$$

$$pr(A_1 \wedge E_2) \cong 0 \cong pr(A_2 \wedge E_1).$$

Sendo os dois problemas estruturalmente idênticos, a acção que é racionalmente prescrita num deles tem de corresponder à acção equivalente no outro. Por exemplo, a acção de cooperar, que é racionalmente prescrita de acordo com o argumento em §5, equivale à acção de escolher a caixa opaca. Como afirma Sobel, cada DP de Expectativa Comum é um PN de Previsão Certa.

David Lewis (1979: 236-7), argumentando a favor da identidade entre os dois problemas, coloca a questão em termos que fazem realçar essa simetria:

1. São-me oferecidos 1,000 Euros e outros 1,000 ao outro prisioneiro; é pegar ou largar.
2. Talvez venha também a receber 1,000,000, mas venha isso a acontecer ou não, tal é causalmente independente do que eu farei.
3. Receberei o milhão extra se, e somente se, o outro recusar os seus mil.

O DP e o PN são semelhantes quanto aos pontos 1 e 2. O PN difere apenas quanto ao ponto 3, que deve ser reformulado, segundo Lewis, da seguinte maneira:

- 3'. Receberei o milhão extra se, e somente se, for previsto que eu recusei os mil.
- 3''. Receberei o milhão extra se, e somente se, um qualquer processo preditivo (que pode decorrer antes, durante ou depois da minha escolha) tem como resultado uma previsão de que não escolho ficar com os mil.

Pode, contudo, alegar-se que esta maneira de colocar o problema sofre de enviesamento (Fernandes 2019), adequando-se apenas a quem defende a acção dominante, que é o caso de Lewis. Tal contenção pode ser defendida se aceitarmos que em 3' e 3'' se encontra implícita

a ideia de que 'eu confesso e espero para ver' (transpondo para o PN, 'eu escolho as duas caixas e espero que o previsor erre'). A simetria entre os dois problemas pode também ser apresentada nos seguintes termos:

- a. São-me oferecidos 1,000 Euros (3 anos de prisão) e outros 1000 ao outro prisioneiro; é pegar ou largar (confesso ou não confesso).
- b. Posso aceitar os mil, ou recusá-los e ganhar um milhão (1 ano de prisão); mas, consiga ou não ganhar o milhão, tal é causalmente independente do que eu farei (aceitar ou não aceitar os mil).
- c. Ganharei o milhão (1 ano de prisão) se, e somente se, o outro recusar os seus mil (o outro não confessa).
- c'. Ganharei o milhão (1 ano de prisão) se, e somente se, for previsto que recusei os mil.

Um aspecto importante da caracterização de Lewis é que parece excluir do cenário duas consequências da acção de cooperar: ganhar 1,000,000 Euros, respectivamente um ano de prisão (extremamente improvável) ou 0, respectivamente 5 anos de prisão (extremamente improvável). Nenhuma referência a estas consequências é feita na caracterização do problema. Para um defensor da deserção (e da escolha das duas caixas), a segunda parte do ponto *b* é claramente decisiva para a sua tomada de posição. Para um defensor da cooperação (e da escolha da caixa opaca), apesar da crença inabalável em *b*, dois novos pontos da maior importância devem ser incluídos na caracterização:

- d. Se eu recusasse os 1000, então (probabilisticamente) o outro recusaria os 1000; e se eu aceitasse os mil, então (probabilisticamente) o outro aceitaria os mil.
- d'. A probabilidade de o outro prever que eu não aceitaria os mil, caso eu não os aceitasse, é quase 1.

Se os pontos 2 e *b* realçam a independência causal entre nossas acções e as acções do nosso adversário (ou as previsões), o que favorece os defensores da deserção/escolha das duas caixas, *d* e *d'* realçam a dependência evidencial entre as mesmas, o que favorece os defensores da cooperação/escolha da caixa opaca.

Com o intuito de contestar esta identidade entre os dois *puzzles*, poder-se-á apontar o seguinte: temos um PN se, e somente se, a probabilidade de o previsor acertar for superior a 0.5005 (Fernandes 2020: §2); contudo, um DP em que a probabilidade com que cada prisioneiro prevê a ação do outro é baixa deixa de apresentar um aspecto paradoxal; nestes casos, como se viu em §5, o argumento em defesa da deserção é muito forte (o leão pode estar acordado). Mas isto não derrota a análise, pois um DP deste último tipo não é um PN; como foi visto, nem todos os DP's são PN's, apenas o são os DP de Expectativa Comum. Isto significa que se removermos do DP a cláusula da dependência evidencial, deixamos de ter uma identidade estrutural entre os dois *puzzles*.

Para quem acredita que a cláusula da expectativa mútua de racionalidade é implausível, pertencendo apenas ao domínio ideal da teoria dos jogos, a importância prática do DP e do PN vê-se algo diminuída. Talvez venha daí a relutância de Jeffrey (1983: 25) em ter em conta o PN para efeitos de revisão do seu modelo de escolha racional:

Encaro o Problema de Newcomb como um dilema do prisioneiro para pilotos de naves-espaciais: um sucessor secular, num cenário de ficção científica, dos problemas da predestinação que ocuparam pensadores como Jonathan Edwards.

Aceitando-se o argumento a favor da identidade estrutural entre os dois *puzzles* ficamos em mãos com muitos mais DP's e PN's do que aqueles que à partida esperaríamos. Não só existem casos em que parece legítimo supor a comum determinação dos agentes em satisfazer as suas preferências, sem que encontremos motivos para desconsiderar a sua racionalidade (em situações que podem ser análogas à do estado de natureza de Hobbes) – um exemplo comum é o das estratégias de armamento nuclear durante a Guerra Fria – como foram ainda apresentadas versões plausíveis do PN – no contexto da teoria económica (Broome 1989) – as quais não envolvem os cenários irrealistas do PN tradicional.

Diogo Fernandes
Grupo LanCog, Centro de Filosofia da Universidade de Lisboa
diogo.fernandes@campus.ul.pt

Referências

- Axelrod, Robert. 1981. The Emergence of Cooperation Among Egoists. *The American Political Science Review* 75: 306-18.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Beaufils, Bruno, Jean-Paul Delahaye e Philippe Mathieu. 1997. Our Meeting With Gradual: A Good Strategy For The Iterated Prisoner's Dilemma. *Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems*. MIT Press, 202-12.
- Binmore, Kenneth. 1994. *Playing Fair: Game Theory and the Social Contract 1*. Cambridge, MA: MIT Press.
- Binmore, Kenneth. 2015. Why All the Fuss: The Many Aspects of the Prisoner's Dilemma. In Martin Peterson (ed.), *The Prisoner's Dilemma*. Cambridge: Cambridge University Press.
- Broome, John. 1989. An Economic Newcomb Problem. *Analysis* 49: 220-22.
- Carroll, J. W. 1987. Indefinite Terminating Points and the Iterated Prisoner's Dilemma. *Theory and Decision* 22: 247-256.
- Danielson, Peter. 1992. *Artificial Morality: Virtual Robots for Virtual Games*. London: Routledge.
- Davis, Lawrence. 1977. Prisoner's, Paradox and Rationality. *American Philosophical Quarterly* 14: 319-27.
- Davis, Lawrence. 1985. Is the Symmetry Argument Valid? In Richmond Cambell e Lanning Snowden (eds.), *Paradoxes of Rationality and Cooperation: Prisoner's Dilemma and Newcomb's Problem*. Vancouver: University of British Columbia Press, 255-62.
- Davis, Morton D. 1997. *Game Theory, A Nontechnical Introduction*. Mineola, New York: Dover Publications.
- Elster, Jon. 1977. Ulysses and the Sirens: A Theory of Imperfect Rationality. *Social Science Information* 16 (5): 469-526.
- Fernandes, Diogo. 2020. O Problema de Newcomb. In Ricardo Santos e Pedro (eds.), *Compêndio em Linha de Problemas de Filosofia Analítica* Lisboa: Centro de Filosofia da Universidade de Lisboa.
- Gauthier, David. 1986. *Moral's by Agreement*. Oxford: Clarendon Press.
- Haji, Ishtiyaque. 1992. Evolution, altruism, and the prisoner's dilemma. *Biology and Philosophy* 7 (2): 161-75.
- Hofstadter, Douglas R. 1983. Metamagical Themas: The calculus of cooperation is tested through a lottery. *Scientific American, Inc.*, Setembro.
- Hobbes, Thomas. [1651]/1999. *Leviatã*. Trad. de João Paulo Monteiro e Maria Beatriz Nizza da Silva. Lisboa: Imprensa Nacional da Casa da Moeda.
- Jeffrey, Richard. [1964]/(1983). *The Logic of Decision*, 2nd ed. Chicago: University of Chicago Press.
- Kavka, Gregory. 1983. Hobbesean Moral and Political Theory. Princeton: Princeton University Press.
- Kuhn, Steven. 2019. Prisoner's Dilemma. In *The Stanford Encyclopedia of Philosophy* (Edição 2019 de Inverno), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2019/entries/prisoner-dilemma/>>.
- Kuhn, Steven e Serge Moresi. 1995. *Pure and Utilitarian Prisoner's Dilemmas*. *Economics and Philosophy* 11: 123-133.
- Lewis, David. 1969. *Convention. A Philosophical Study*. Cambridge, Massachusetts: Harvard University Press.
- Lewis, David. 1979. Prisoners' Dilemma is a Newcomb Problem. *Philosophy and Public Affairs* 8 (3): 235-40.
- Luce, Duncan R. e Howard Raiffa. 1957. *Games and Decisions*. New York: John Wiley and Sons.

- Nowak, Martin e Karl Sigmund. 1992. Tit for Tat in Heterogeneous Populations. *Nature* 355 (1 de Janeiro): 250-53.
- Nowak, Martin e Karl Sigmund. 1994. A Strategy of Win-Stay, Lose-Shift that Outperforms Tit-for-Tat in the Prisoner's Dilemma Game. *Nature* 364 (1 de Julho): 56-8.
- Nozick, Robert. [1969]/(1985). Newcomb's Problem and Two Principles of Choice. In Richmond Campbell e Lanning Sowden (eds.), *Paradoxes of Rationality and Cooperation: Prisoner's Dilemma and Newcomb's Problem*. Vancouver: The University of British Columbia Press, 107-33.
- Peterson, Martin (ed.). 2015. *The Prisoner's Dilemma*. Cambridge: Cambridge University Press.
- Pettit, Phillip. 1986. Free Riding and Foul Dealing. *Journal of Philosophy* 83: 361-79.
- Pettit, Phillip e Robert Sugden. 1989. The Backward Induction Paradox. *Journal of Philosophy* 86: 169-182.
- Rapoport, Anatol, Melvin J. Guyer e David G. Gordon. 1976. 2×2 Games. Ann Arbor: University of Michigan Press.
- Rapoport, Ammon, D. A. Seale e A. M. Colman. 2015. Is Tit-for Tat the Answer? On the Conclusions Drawn from Axelrod's Tournaments. *PLoS ONE* 10 (7): e0134128.
- Rawls, John. [1971]/1999. *A Theory of Justice*, revised edition. Oxford University Press.
- Resnik, Michael D. 1987. *Choices – An Introduction to Decision Theory*. University of Minnesota Press.
- Rousseau, Jean-Jacques. [1755]/1997. Discourse on the Origin and the Foundations of Inequality Among Men or Second Discourse. In Victor Gourevitch (trad. e ed.), *Rousseau: The Discourses and Other Early Political Writings*. Cambridge University Press, 116-93.
- Schelling, Thomas C. 1963. *The Strategy of Conflict*. New York: Oxford University Press.
- Segal, Nancy e Scott Hershberger. 1999. Cooperation and Competition Between Twins: Findings from a Prisoner's Dilemma Game. *Evolution and Human Behavior* 20: 29-51.
- Selten, Reinhard. 1978. The Chain-Store Paradox. *Theory and Decision* 9: 127-159.
- Sobel, John H. 1985. Not Every Prisoners' Dilemma is a Newcomb Problem. In Richmond Campbell e Lanning Sowden (eds.), *Paradoxes of Rationality and Cooperation, Prisoner's Dilemma and Newcomb's Problem*. Vancouver: The University of British Columbia Press, 263-74.
- Sobel, John H. 1993. Backward-induction arguments: A paradox regained. *Philosophy of Science* 60 (1):114-133.
- Sober, Elliott. 1988. What Is Evolutionary Altruism? *Canadian Journal of Philosophy*, Supp. 14: 75-99.
- Trivers, Robert L. 1971. The Evolution of Reciprocal Altruism. *Quarterly Review of Biology* 46: 35-57.
- von Neumann, John e Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.