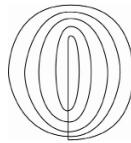


ARGUMENTO DA CONSEQUÊNCIA

EDIÇÃO DE 2022 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 © 2022 do editor
Centro de Filosofia da Universidade de Lisboa
Alameda da Universidade, Campo Grande, 1600-214 Lisboa

Argumento da Consequência
Copyright © 2022 do autor
Luís Veríssimo

DOI: <https://doi.org/10.51427/cfi.2021.0095>

Todos os direitos reservados

Resumo

O Argumento da Consequência é um dos mais discutidos argumentos no que diz respeito à controvérsia em torno do problema da compatibilidade entre livre-arbítrio e determinismo. Este argumento tem sido o principal cavalo de batalha dos incompatibilistas e nenhum compatibilista pode defender satisfatoriamente a sua posição sem tentar dar-lhe uma resposta. Neste artigo, apresenta-se uma visão geral do argumento, bem como algumas das críticas que os compatibilistas lhe têm dirigido.

Palavras-chave

Incompatibilismo, argumento da consequência, livre-arbítrio, determinismo.

Abstract

The Consequence Argument is one of the most discussed arguments in the dispute between free will and determinism. This argument has been one of the incompatibilists main defenses against compatibilism and no sound version of compatibilism can be suitably sustained without a proper reply to it. This article presents the general outline of the argument, as well as some of the main objections compatibilists have pointed out over the years.

Keywords

Incompatibilism, consequence argument, free will, determinism.

Argumento da Consequência

DOI: <https://doi.org/10.51427/cfi.2021.0095>

1 Introdução

O Argumento da Consequência é um dos mais discutidos argumentos no que diz respeito à controvérsia em torno do problema da compatibilidade entre livre-arbítrio e determinismo. Este problema pode ser explicitamente formulado da seguinte forma: 'O livre-arbítrio é compatível com o determinismo?'

A expressão 'livre-arbítrio' é aqui entendida como sinónimo de 'liberdade da vontade'. Assim, dizer que temos livre-arbítrio, ou uma vontade livre, é o mesmo que dizer que podemos controlar algumas das coisas que acontecem.

O termo 'determinismo', por sua vez, é utilizado para designar a tese de que tudo o que acontece é consequência necessária do passado e das leis da natureza. O determinismo diz-nos, então, que vivemos num universo com determinadas leis naturais que regem as relações entre tudo aquilo que existe, não sendo, por isso, possível que as mesmas causas tenham efeitos diferentes, ou seja, se dois mundos deterministas tiverem exatamente as mesmas leis da natureza, então se forem idênticos em algum momento do tempo, serão igualmente idênticos em todos os outros.

Aparentemente, existe uma tensão entre a crença de que temos livre-arbítrio e a crença no determinismo. Afinal de contas, se tudo o que acontece é uma consequência necessária do passado e das leis da natureza, então, dado que as nossas ações também são acontecimentos, também elas serão uma consequência necessária desses dois fatores. Mas poderão as nossas ações depender de nós, se são a consequência necessária de fatores sobre os quais, aparentemente, não temos qualquer tipo de controlo?

Existem duas respostas possíveis para este problema: a primeira, que dá pelo nome de 'incompatibilismo', sustenta que o livre-arbítrio não é compatível com o determinismo; a segunda, designada

'compatibilismo', sustenta precisamente o contrário, ou seja, que o livre-arbítrio é compatível com o determinismo.

O Argumento da Consequência tem sido o principal cavalo de batalha dos incompatibilistas e nenhum compatibilista pode defender satisfatoriamente a sua posição sem tentar dar-lhe uma resposta.

2 O argumento da consequência

A formulação clássica do Argumento da Consequência é-nos apresentada por Peter van Inwagen (1983: 56), nos seguintes termos:¹

Se o determinismo é verdadeiro, então as nossas ações são consequências das leis da natureza e de acontecimentos que ocorreram num passado remoto. Mas tanto as leis da natureza como aquilo que aconteceu antes de termos nascido não dependem de nós. Logo, as consequências destas coisas (incluindo os atos que realizamos agora) não dependem de nós. (van Inwagen 1983: 56)

Este argumento pode ser explicitamente formulado conforme se segue:

- (1) Se o determinismo é verdadeiro, então as nossas ações (tal como tudo o que acontece) são a consequência das leis da natureza e de acontecimentos que ocorreram num passado remoto.
- (2) Não somos capazes de alterar as leis da natureza nem os acontecimentos que ocorreram num passado remoto.
- (3) Se as nossas ações são a consequência das leis da natureza e de acontecimentos que ocorreram num passado remoto e não somos capazes de alterar as leis da natureza, nem os acontecimentos que ocorreram num passado remoto, então não temos possibilidades alternativas.
- (4) Se não temos possibilidades alternativas, então não temos livre-arbítrio.
- (5) Logo, se o determinismo é verdadeiro, então não temos livre-arbítrio.

A conclusão do argumento segue-se validamente das suas premissas,

¹ Esta formulação tornou-se canónica, embora, em 1966, Carl Ginet tenha avançado uma proposta semelhante (Cf. Ginet: 1966).

portanto, para poderem consistentemente rejeitar a tese incompatibilista, os compatibilistas terão forçosamente de contestar alguma das suas premissas. Mas qual?

A premissa (1) limita-se a apresentar uma definição de determinismo ou, mais precisamente, uma aplicação do determinismo ao caso concreto das nossas ações.

Deste modo, os compatibilistas têm ao seu dispor apenas três possibilidades:

1. Negar a premissa (2);
2. Negar a premissa (3);
3. Negar a premissa (4).

A premissa (2) compromete-nos com a ideia, à partida, bastante plausível, de que nem o passado remoto, nem as leis da natureza dependem de nós, ou seja, não está ao nosso alcance mudar o que aconteceu num passado remoto (antes de termos nascido), nem violar as leis da natureza. Parece trivial dizer que não podemos fazer nada para alterar o que aconteceu antes de termos existido, nem temos poderes que nos permitam violar as leis da natureza.

Contudo, há quem defenda que esta ideia está longe de ser assim tão óbvia. Alguns autores, como J. K. Campbell (2007), vão tentar mostrar que essa premissa é falsa, pois não é necessariamente verdade que o passado remoto não depende de ninguém. Outros autores, como David Lewis (1981), vão defender que, de acordo com uma determinada leitura, a premissa é falsa, pois ainda que não possamos violar as leis da natureza, podemos fazer coisas que implicariam a ocorrência de uma violação das mesmas.

2.1 *A objeção do passado inexistente*

Em primeiro lugar, iremos debruçar-nos, sobre a estratégia seguida por Campbell. No seu artigo 'Free will and the necessity of the past' (2007), Campbell pretende precisamente mostrar que, ainda que a proposição 'O passado remoto não depende de nós' possa ser *contingentemente verdadeira* (i.e., embora possa ser verdadeira *no mundo atual*), esse não é *necessariamente* o caso (i.e., ela não é verdadeira em todos os mundos possíveis).

De facto, a existência de um passado remoto é um facto contingente acerca do mundo atual. Num mundo possível onde não exista um passado remoto, ou seja, num mundo possível onde exista alguém (um agente) desde sempre (ou desde que esse mundo começou a existir), a ideia de que o passado não depende de ninguém parece não ser de todo evidente. Assim, se concebermos um mundo possível onde o determinismo é verdadeiro, mas no qual algum ser humano existe desde sempre, podemos perfeitamente considerar que a primeira escolha desse ser humano é livre – apesar de este se encontrar num mundo determinista.

Isto significa que o Argumento da Consequência não pode aspirar a mais do que mostrar que o livre-arbítrio não pode coexistir com o determinismo, *no mundo atual*, mas não que os dois são *necessariamente* incompatíveis entre si (i.e., incompatíveis *em todos os mundos possíveis*).

Atendendo ao facto de apelarem a experiências mentais que envolvem um mundo determinista sem um passado remoto, este tipo de objeções tornou-se conhecido como 'Objeções do Passado Inexistente'.²

2.2 *O compatibilismo dos milagres locais*

Passemos agora à objeção de David Lewis. No artigo 'Are we free to break the laws?' (1981), Lewis defende uma perspetiva designada 'Compatibilismo de Milagres Locais'.³ Lewis começa por conceder que caso o determinismo seja verdadeiro, então tudo o que acontece (incluindo as nossas ações) é uma consequência necessária do passado e das leis da natureza. Neste sentido, se estava determinado que alguém, por exemplo, Lewis, não levantaria o braço, só podemos dizer que Lewis podia ter levantado o braço, no caso de ser verdade que ele poderia ter feito algo que tornaria falsa

- (i) uma proposição acerca do passado remoto, ou
- (ii) uma proposição correspondente a uma lei da natureza.

Contudo, uma vez que não há obviamente nada que alguma vez alguém pudesse ter feito que tornasse falsa uma proposição acerca do passado

² 'No Past Objections' no original em inglês.

³ 'Local Miracle Compatibilism', no original em inglês. Esta designação parece dever-se a John Martin Fischer (cf. Fischer 1994, Capítulo 4).

remoto, segue-se do facto de que Lewis podia ter levantado o braço (embora estivesse determinado que não o faria) que ele poderia ter feito algo que tornaria falsa uma proposição correspondente a uma lei da natureza.

Ainda assim, em vez de encarar isto como uma redução ao absurdo da perspectiva compatibilista, Lewis vai procurar mostrar que, embora seja verdade que, nessas circunstâncias, o facto de ele levantar o braço exigir a ocorrência de um milagre e ninguém ser capaz de realizar milagres, ele podia ter levantado o seu braço, uma vez que o facto de uma pessoa poder realizar uma ação que exige a ocorrência de um milagre, não implica que esta pessoa seja capaz de fazer milagres.

As leis da natureza são, por definição, (pelo menos) regularidades sem exceção. Assim, se alguém é capaz de fazer algo que aparentemente constituiu uma violação das leis da natureza, então é simplesmente porque aquilo que se pensava ser uma lei da natureza afinal não o era.⁴ Por exemplo, não podemos correr mais depressa do que a velocidade da luz, nem construir uma máquina capaz de lançar pedras a essa velocidade. Mas se assim é, como podemos afirmar que a proposição 'As leis da natureza não dependem de nós' é falsa?

Lewis começa por sustentar que essa proposição admite duas leituras – uma delas corresponde a uma tese forte, a outra corresponde a uma tese fraca. Supondo que *L* abrevia a conjunção de todas as leis da natureza (sejam elas quais forem), estas duas leituras dizem respetivamente o seguinte:

- Tese forte: *L* é verdadeira e não há qualquer agente *X* e ação *A*, tal que *X* possa fazer *A* e *A* seja ou cause um evento que torna *L* falsa – isto é, ninguém tem a capacidade de violar uma lei da natureza.*X*
- Tese fraca: *L* é verdadeira e não há qualquer agente *X* e ação *A*, tal que *X* pode fazer *A* e, se *X* fizesse *A*, *L* seria falsa – isto

⁴ A definição aqui em causa incorpora a expressão 'pelo menos' entre parêntesis, pois existem diferentes formas de entender as leis da natureza e, de acordo com algumas delas, as leis da natureza são mais do que simples regularidades sem exceção. Contudo, uma vez que mesmo de acordo com essas perspectivas, embora nem todas as regularidades sem exceção sejam leis da natureza, estas serão seguramente regularidades sem exceção (ainda que de um determinado tipo). Assim sendo, esta definição mais lata dá conta de ambas essas possibilidades.

é, ninguém tem a capacidade de fazer algo tal que, se o tivesse feito, uma lei da natureza teria sido falsa.

De maneira a especificar um valor de verdade para as duas leituras no que respeita ao exemplo em causa (a saber, o ato de Lewis levantar o braço, quando estava determinado que ele não o faria), bem como o que se segue de cada uma dessas leituras, Lewis sugere que consideremos três exemplos de capacidades.

Exemplo 1

[S]ou capaz de partir uma janela. Para começar: sou capaz de fazer algo tal que, se eu o tivesse feito, uma janela teria sido partida. Mas há mais. Sou capaz de fazer algo tal que, se eu o tivesse feito, esse ato causaria um evento de partir janelas. (Lewis 1981: 115)

Exemplo 2

[S]ou capaz de quebrar uma promessa. Para começar: sou capaz de fazer algo tal que, se eu o tivesse feito, uma promessa teria sido quebrada. Mas há mais. Sou capaz de fazer algo tal que, se eu o tivesse feito, esse ato causaria um evento de quebrar promessas. (Lewis 1981: 115)

Exemplo 3

Em seguida, consideremos o que seria uma capacidade verdadeiramente incrível de violar uma lei – capacidade essa que eu não poderia plausivelmente reclamar. Suponhamos que eu era capaz de atirar uma pedra com muita muita força. E suponhamos que se eu o fizesse, a pedra acabaria por se deslocar mais rápido do que a velocidade da luz, um evento contrário às leis. Para começar: eu seria capaz de fazer algo tal que, se eu o tivesse feito, uma lei teria sido violada. Mas há mais. Eu seria capaz de fazer algo tal que, se eu o tivesse feito, esse ato causaria um evento de violar leis da natureza. (Lewis 1981: 115)

A lição dos três exemplos anteriores é, um pouco adiante, apresentada por Lewis nos seguintes termos:

Se nenhum ato meu causou ou consistiu num evento de partir janelas, de quebrar promessas ou de violar leis da natureza, então penso que não se pode dizer que eu tenha partido uma janela, quebrado uma promessa, ou violado uma lei. Assim sendo, só sou

capaz de partir uma janela, quebrar uma promessa ou violar uma lei da natureza, se sou capaz de fazer algo tal que, se eu o tivesse feito, esse ato teria causado ou seria (em si mesmo) um evento de partir janelas, de quebrar promessas ou de violar leis da natureza. (Lewis 1981: 116)

Os exemplos levam então Lewis a concluir que a leitura forte tornaria a proposição 'As leis da natureza não dependem de nós' verdadeira (e, conseqüente, tornaria verdadeira a premissa (2) do argumento da consequência, visto que se trata da conjunção desta ideia com a ideia de que o passado remoto não depende de nós). De facto, não há nada que possamos fazer que cause ou seja (em si mesmo) uma violação das leis da natureza (leitura forte).

Contudo, uma vez que isso não implica que não podemos fazer algo que, caso o tivéssemos feito, uma lei da natureza teria sido falsa (leitura fraca), se entendêssemos essa proposição numa aceção forte, então a premissa (3) do Argumento da Consequência seria falsa e, por conseguinte, o argumento não seria sólido. Ainda que seja verdade que não controlamos o que aconteceu num passado remoto e que não podemos fazer nada que cause ou seja (em si mesmo) uma violação das leis da natureza (leitura forte), isso não significa que não temos possibilidades alternativas, pois mesmo nessas circunstâncias há coisas que poderíamos ter feito que implicariam que uma lei da natureza teria sido falsa (leitura fraca).

Por outro lado, se interpretássemos a proposição 'As leis da natureza não dependem de nós' num sentido fraco, Lewis sustenta que ela seria falsa – e, por esse motivo, o argumento também não seria sólido, pois a premissa (2) seria falsa.

Ora, para entender exatamente por que razão a leitura fraca torna a proposição em causa falsa, temos de ser capazes de determinar o valor de verdade de uma condicional contrafactual como

- se Lewis tivesse levantado o braço, então uma lei da natureza teria sido falsa.

Saber como determinar as condições de verdade das condicionais contrafactuais é um tópico particularmente controverso. De qualquer modo, o próprio Lewis (1979: 465) é o responsável por uma célebre

proposta quanto a este tópicó que achamos conveniente apresentar e que pode ser sintetizada conforme se segue:

Uma contrafactual como 'Se A tivesse sido o caso, então B teria sido o caso'⁵ é (não-vacuamente) verdadeira num mundo possível W se, e só se, B é verdadeira no(s) mundo(s) possível(eis) mais próximo(s) de W em que A é verdadeira (ou, dito de outra forma, no(s) mundo(s)- A mais próximo(s) de W). Por sua vez, o(s) mundo(s) possível(eis) mais próximo(s) de um dado mundo W é (ou são, caso haja um empate entre dois ou mais mundos) o(s) mundo(s) mais semelhante(s) a W .

Para avaliar a semelhança entre mundos possíveis, a proposta de Lewis especifica ainda dois fatores. O primeiro é a correspondência entre mundos no que diz respeito às leis da natureza. O segundo fator é a correspondência entre mundos no que diz respeito a questões de facto particulares. Para Lewis, estes dois fatores encontram-se mutuamente contrabalançados de acordo com a seguinte hierarquia:

1. O mais importante é evitar grandes, prolíferas e variadas violações das leis da natureza.
2. A segunda coisa mais importante é maximizar a região espaço-temporal ao longo da qual prevalece uma correspondência perfeita entre questões de facto particulares.
3. A terceira coisa mais importante é evitar violações das leis da natureza, ainda que pequenas, localizadas e simples.
4. É de pequena ou nenhuma importância assegurar a semelhança aproximada entre questões de facto particulares concretas, ainda que nos digam imenso respeito. (Lewis 1979: 472)

De modo a ilustrar a forma como os dois fatores se articulam entre si para determinar a semelhança entre mundos, consideremos o seguinte exemplo retirado do artigo 'Local Miracle Compatibilism' (2003), de Helen Beebe:

[S]uponhamos que eu, efetivamente, permaneço sentada à minha

⁵ Uma condicional contrafactual pode ser representada simbolicamente ($A \Box \rightarrow B$), para a distinguir da implicação (ou condicional) material, ($A \rightarrow B$), e de uma implicação necessária, $\Box(A \rightarrow B)$.

secretária na próxima meia-hora. Qual é o mundo mais próximo no qual chego ao *pub* daqui a dez minutos (evento E)? Num mundo- E , W_1 , eu permaneço sentada à minha secretária por, digamos, 9 minutos e 59.9 segundos, de maneira que todo o passado até esse momento é exatamente como no mundo atual, e depois eu espontaneamente desapareço, reaparecendo à porta do *pub* um décimo de segundo mais tarde. Num outro mundo- E , W_2 , algo se passa comigo ao fim de, digamos, 5 minutos – sendo todo o passado até esse momento (anterior) exatamente como no mundo atual – que faz com que eu comece a caminhar de forma perfeitamente normal até ao *pub*, chegando lá exatamente 5 minutos depois. Vamos assumir que aquilo que me acontece passados 5 minutos em W_2 é que eu adquiro subitamente um forte desejo por cerveja. (Beebe 2003: 262)

Ora, assumindo que o determinismo é verdadeiro, tanto W_1 como W_2 contêm milagres por relação com o mundo atual. Isto é assim porque, uma vez que em ambos os mundos existem instantes no passado remoto integralmente idênticos aos do mundo atual, caso as suas leis da natureza fossem exatamente as mesmas, o evento E não poderia nunca vir a ocorrer.

Mas como é que podemos determinar qual destes dois mundos se encontra mais próximo do atual recorrendo à hierarquia lewisiana? Bem, podemos constatar que embora em W_1 a correspondência perfeita de questões de facto particulares seja preservada durante mais uns minutos do que em W_2 , o milagre que ocorre em W_1 parece ser muito maior do que aquele que ocorre em W_2 . Em W_2 , o único milagre exigido é que o agente adquira um forte desejo por cerveja – o que pode simplesmente vir a ocorrer devido à ativação de uns neurónios extra no seu cérebro; ao passo que em W_1 é necessário que um grande número de moléculas desapareça de um sítio e que se reagrupe, exatamente com a mesma configuração, noutra.

Assim sendo, ainda que a correspondência entre questões de facto particulares seja ligeiramente maior em W_1 do que em W_2 , este último é mais próximo do mundo atual, pois contrariamente ao que acontece em W_1 , em W_2 apenas é exigida a ocorrência de um pequeno milagre local e não a ocorrência de uma enorme diferença no que diz respeito às leis da natureza.

Deste modo, ao procurar determinar o valor de verdade de uma contrafactual como 'Se Lewis tivesse levantado o braço, o mosquito não o teria picado' podemos pensar duas coisas. A primeira coisa é que há mundos nos quais Lewis levantou o braço porque diferem do atual no que diz respeito a questões de facto particulares. Podemos conceber, por exemplo, mundos nos quais Lewis avistou um conhecido que estava a passar perto do local onde ele se encontrava e, por isso, levantou o braço para o cumprimentar, ou para chamar a sua atenção, mundos em que Lewis levantou o braço porque estava há muito tempo na mesma posição e começou a sentir uma certa dormência, etc. No entanto, uma vez que nenhuma dessas possibilidades ocorre no mundo atual, podemos também considerar que, à partida, nenhum desses mundos se encontra entre os mais próximos do nosso.

A segunda coisa que podemos pensar é que há mundos nos quais Lewis levanta o braço porque diferem do atual no que diz respeito às leis da natureza. Podemos imaginar, por exemplo, que a lei da gravidade não existe, que existem leis que fazem com que os objetos tenham tendência para se repelir mutuamente, uma espécie de anti-gravidade, que faz com que o braço de Lewis se tivesse levantado devido à proximidade a que se encontrava da secretária, etc. No entanto, podemos considerar, uma vez mais, que, à partida, nenhum desses mundos se encontra entre os mais próximos do atual, pois se as suas leis divergissem assim tanto das nossas, seria de esperar que se registassem diferenças significativas entre os dois mundos.

Contudo, conforme foi notado anteriormente, também parece seguir-se do determinismo o seguinte:

- se dois mundos deterministas coincidirem em termos de leis da natureza, então se forem idênticos em algum momento do tempo, serão igualmente idênticos em todos os outros.

Ora, se assim fosse, parece que afinal de contas Lewis não poderia ter levantado o braço há momentos (e, conseqüentemente, evitado a picadela de mosquito). Mas, aparentemente, isso serviria apenas para reforçar a tese incompatibilista. De maneira que, para evitar este resultado, um compatibilista – como o próprio Lewis – tem de se mostrar disposto a admitir como mais próximos do atual, mundos que são indiscerníveis do mesmo até (aproximadamente) ao momento

em que ele levanta o braço, mas que, por este motivo, acabam por divergir do atual em relação às leis da natureza.

Suponhamos, por exemplo, que W_2 é um mundo factualmente indiscernível do nosso (i.e., suponhamos que ele é muito próximo do mundo atual, W_a) até ao momento em que Lewis está prestes a ser picado pelo mosquito.

Imaginemos, agora, que nesse mundo, devido a um milagre de divergência – o qual constitui uma violação das leis da natureza do mundo atual, L_a , mas não uma violação das leis, L_2 , desse mundo, pois isso seria uma contradição nos termos – Lewis levantou o braço há um minuto evitando a picadela de mosquito. Isto significa que, no mundo atual, Lewis tinha a capacidade de levantar o braço há um minuto.

No entanto, segundo Lewis, isso não implica que ele tem a capacidade de violar as leis da natureza (num sentido forte) – isto é, que ele tem a capacidade de fazer algo tal que cause ou seja, em si mesmo, uma violação das leis. Implica apenas que ele tem a capacidade (num sentido fraco) de fazer algo tal que, caso o tivesse feito, uma lei da natureza teria sido falsa. E ele não tem essa capacidade num sentido forte porque

- i) a sua ação não pode ser ela mesma o milagre de divergência, pois, uma vez que essa ação se encontra ausente da cadeia de acontecimentos do mundo atual, para que esta pudesse vir a ocorrer, um milagre de divergência teria de já ter ocorrido previamente, e
- ii) de acordo com a noção de causalidade de Lewis, a sua ação também não pode causar ela própria o milagre de divergência.

Para compreender melhor o que está em jogo em *ii*, torna-se crucial dizer algo mais sobre a noção lewisiana de causalidade. A noção lewisiana de causalidade pode ser assim apresentada:

Um evento, digamos C , causa outro evento, digamos E , se e só se existe uma cadeia causal que leva de C a E .

Uma cadeia causal é uma sequência de eventos, digamos $\langle \dots, C, D, E, \dots \rangle$, tal que D depende causalmente de C , E depende causalmente de D , e assim sucessivamente.

Um evento, digamos E , depende causalmente de outro, digamos

D, se e só se, se *D* acontecesse, *E* aconteceria, e se *D* não acontecesse, *E* não aconteceria.

Podemos agora compreender por que razão a ação de Lewis não poderia causar o milagre de divergência. De acordo com a anterior análise da causalidade, não há qualquer ação que Lewis poderia ter realizado que pudesse causar um determinado milagre de divergência. À primeira vista, pode certamente parecer que existe uma relação de dependência causal entre o seu ato de levantar o braço e a ocorrência de um determinado milagre, pois parece que (A) se Lewis tivesse levantado o braço, o milagre de divergência M_1 teria ocorrido, e (B) se Lewis não tivesse levantado o braço, o milagre de divergência M_1 não teria ocorrido.

No entanto, conforme Lewis faz notar, (A) é falsa. De facto, é verdade que, se Lewis tivesse levantado o braço, um milagre ou outro teria de ter ocorrido, mas nada nos garante que seja o milagre M_1 , e não um outro milagre qualquer, digamos M_2 .

Assim, uma vez que a ação de Lewis levantar o braço pode vir a ocorrer, sem que o milagre M_1 tenha ocorrido, a contrafactual expressa em (A) é falsa (i.e., há pelo menos um mundo próximo do atual no qual a antecedente é verdadeira e a consequente falsa) e, por conseguinte, é falso que essa ação causou o milagre de divergência M_1 — a ação pode ter ocorrido graças a um outro milagre de divergência qualquer, digamos M_2 .

Fica assim claro por que razão Lewis sustenta que a proposição 'As leis da natureza não dependem de nós' será falsa, se for entendida num sentido fraco. Isso acontece porque, uma vez que podemos considerar que existem mundos possíveis próximos do nosso onde fazemos algo tal que, caso o tivéssemos feito, uma lei da natureza teria sido falsa, não podemos dizer, *neste sentido*, que 'ninguém tem nem nunca teve o poder de tornar falsa a conjunção das leis da natureza'.

De igual modo, fica também claro por que razão Lewis considera que, se essa proposição for entendida num sentido forte, a premissa (2) é, de facto, verdadeira. Conforme também acabamos de ver, ainda que seja verdade que há mundos possíveis próximos do nosso onde, por exemplo, fazemos algo tal que, caso o tivéssemos feito uma lei da natureza teria sido falsa, a nossa ação não pode causar, nem ser ela mesma, um evento de violação das leis. Ainda assim, segundo Lewis,

daqui não se segue que ‘se o determinismo for verdadeiro, não temos a capacidade de agir de outro modo’, pois não está em nada do que se disse implicado que não existem mundos possíveis próximos do nosso onde fazemos algo tal que, caso o tivéssemos feito, uma lei da natureza teria sido falsa (leitura fraca).

É deste modo que Lewis conclui que, qualquer que seja a interpretação que queiramos atribuir à proposição ‘As leis da natureza não dependem de nós’, o Argumento da Consequência não é sólido: numa das interpretações – a chamada ‘leitura forte’ – a premissa (2) é verdadeira, mas a premissa (3) é falsa; na outra interpretação – a chamada ‘leitura fraca’ – a premissa (3) é verdadeira, mas a premissa (2) é falsa.

Apesar da força, aparentemente, devastadora deste ataque ao argumento modal da consequência, a sua estratégia não está isenta de problemas. Para seguir este debate veja-se, por exemplo, Helen Beebe (2003), Shane Oakley (2006) e Peter Graham (2008). O primeiro destes artigos constitui uma crítica à perspectiva defendida por Lewis, os dois últimos são respostas a essa mesma crítica.

2.3 O princípio de transferência de incapacidade

Passemos agora a analisar a premissa (3) do Argumento da Consequência. Esta premissa estabelece que se as nossas ações são consequência de coisas que não podemos controlar, então, quando agimos, não temos verdadeiramente possibilidades alternativas ao nosso dispor, uma vez que nada do que acontece resulta de coisas que possamos controlar. Esta ideia parece ter subjacente o seguinte princípio:

*Princípio de transferência de incapacidade*⁶

Para todo e qualquer agente S , e para quaisquer proposições P e Q , se P é verdadeira e S não tem a capacidade de mudar isso, e se Q é uma consequência de P e S não tem a capacidade de mudar isso, então Q é verdadeira e S não tem a capacidade de mudar isso.

⁶ Esta designação foi retirada de Ned Markosian (2010), mas este princípio também tem sido designado por ‘Regra Condicional’ [*Conditional Rule*] (ver van Inwagen 2008) e, de certa forma, corresponde àquilo que na versão modal do argumento da consequência se designa usualmente por ‘Regra β ’.

Para ilustrar o que está aqui em causa, van Inwagen (1983) sugere que imaginemos o seguinte exemplo:

Suponhamos que p representa a proposição 'O Sol explode no ano 3.000' e q representa a proposição 'A vida na Terra acaba no ano 3.000'. Se p é verdadeira e ninguém tem a capacidade de alterar esse facto, isto é, se é verdade que o Sol explode no ano 3.000 e ninguém tem a capacidade de alterar esse facto, e além disso é verdade que q é uma consequência de p e ninguém tem capacidade de alterar esse facto, isto é, o fim da vida na Terra no ano 3.000 é uma consequência do facto de o Sol explodir no ano 3.000 e ninguém tem a capacidade de alterar esse facto, então parece seguir-se que q é verdadeira e ninguém tem a capacidade de alterar esse facto, ou seja, a vida na Terra acaba no ano 3.000 e ninguém tem a capacidade de mudar isso.

Parece inquestionável que se algo não depende de mim e se não depende de mim o facto de isso ter certas consequências, então essas consequências também não dependem de mim em nenhum sentido relevante, isto é, não há nada que eu pudesse ter feito para evitar que estas viessem a acontecer. Com efeito, este princípio parece ser intuitivamente válido, mas o facto de isso se verificar no que diz respeito a casos particulares, não é suficiente para concluirmos que isso se verifica em todos os casos concebíveis. Será que não existem contraexemplos relevantes que permitam rejeitar este princípio e, consequentemente, rejeitar o Argumento da Consequência?

Uma estratégia frequentemente utilizada pelos compatibilistas para atacar o Argumento da Consequência passa precisamente por atacar diretamente a premissa (3) do argumento, procurando contra-exemplos para o princípio que lhe subjaz.

Para refutar a premissa (3) do Argumento da Consequência, o compatibilista terá a seu cargo a tarefa de mostrar que apesar de as nossas ações serem a consequência das leis da natureza e de acontecimentos que ocorreram num passado remoto e de não sermos capazes de alterar as leis da natureza nem os acontecimentos que ocorreram num passado, é ainda assim possível que tenhamos possibilidades alternativas.

Para defender esta ideia, alguns compatibilistas sugerem que a ideia das possibilidades alternativas seja entendida de uma forma

condicional. De acordo com esta interpretação, dizer que 'podíamos ter agido de outra forma' é para ser entendido de forma condicional, isto é, deve ser entendido como se aquilo que estamos a afirmar fosse o seguinte: 'teríamos agido de outra forma, se tivéssemos escolhido fazê-lo'.

Esta estratégia da análise condicional está associada àquilo que se tem designado 'compatibilismo clássico', defendido por autores como Thomas Hobbes, John Locke, David Hume, John Stuart Mill, G. E. Moore, A. J. Ayer, entre outros.

A análise condicional permite ao compatibilista sustentar que o livre-arbítrio não é incompatível com o determinismo, pois embora seja verdade que só temos livre-arbítrio se pudermos agir de modo diferente daquele que agimos – isto é, se tivermos possibilidades alternativas – é falso que num universo determinista não existem tais possibilidades. Isto porque, de acordo com esta análise condicional, dizer que temos possibilidades alternativas é o mesmo que dizer que caso tivéssemos decidido fazer outra coisa, teríamos efetivamente feito outra coisa. Ora, isso pode ser verdadeiro, mesmo num mundo onde tudo (incluindo a nossa vontade, as nossas crenças e os nossos desejos, etc.) está determinado, pois mesmo nessas circunstâncias continuaria a ser verdade que, por vezes, teríamos agido de modo diferente daquele que agimos, se tivéssemos decidido fazê-lo, isto é, se as nossas crenças e desejos tivessem sido diferentes daquilo que efetivamente são.

Esta perspetiva passa a encarar o livre-arbítrio como ausência de coação: somos livres na medida em que não somos forçados (por fatores externos) a fazer aquilo que não queremos, nem impedidos (por fatores externos) de fazer aquilo que pretendemos fazer (ainda que aquilo que pretendemos seja inteiramente determinado).

Assim, o argumento subjacente à estratégia da análise condicional pode ser reconstruído conforme se segue.

- (1) Poderíamos ter agido de modo diferente daquele pelo qual agimos se, e só se, tivéssemos agido de um modo diferente, caso tivéssemos desejado fazê-lo. [Análise Condicional]
- (2) Ainda que as nossas ações sejam a consequência das leis da natureza e de acontecimentos que ocorreram num passado remoto e não sejamos capazes de alterar as leis da natureza

nem os acontecimentos que ocorreram num passado remoto, é verdade que, por vezes, teríamos agido de modo diferente, se assim o tivéssemos desejado.

- (3) Logo, é falso que se as nossas ações são a consequência das leis da natureza e de acontecimentos que ocorreram num passado remoto e não somos capazes de alterar as leis da natureza nem os acontecimentos que ocorreram num passado remoto, então não temos possibilidades alternativas (i.e., nunca poderíamos ter agido de modo diferente).

Muitos autores consideram que esta estratégia não é procedente. Roderick M. Chisholm (1964: 720–729), por exemplo, sustenta que a análise condicional é altamente implausível, pois compromete-nos com a ideia de que o significado profundo de uma frase como

- a) 'S podia ter agido de outro modo' (*analysandum*)

se deixa captar adequadamente por uma frase como

- b) 'Se S tivesse escolhido agir de outro modo, então ele teria agido de outro modo' (*analysans*).

Ora, de acordo com Chisholm, a menos que

- c) 'S podia ter escolhido agir de outro modo'

seja verdadeira, é possível que *b* seja verdadeira e *a* seja falsa. Isto significa que, contrariamente ao que é sugerido pela estratégia da análise condicional, estas duas afirmações, *a* e *b*, não têm o mesmo significado.

No artigo 'Cans without ifs' (1968), Keith Lehrer ilustra este tipo de possibilidade, recorrendo ao seguinte exemplo:

1. Suponhamos que me ofereciam um frasco com vários doces, incluindo gomas vermelhas; e que, uma vez que tenho uma aversão patológica a gomas vermelhas (imaginemos, por exemplo, que me fazem lembrar gotas de sangue), decidi não tirar nenhum doce.
2. Ora, apesar de ser logicamente possível supor que, se eu tivesse escolhido tirar uma goma vermelha, eu tê-lo-ia feito, a verdade é que, dada a minha aversão, eu não poderia ter escolhido fazê-lo.

3. Isto significa que eu não tenho, efetivamente, a capacidade de tirar uma goma vermelha (isto é, a proposição que representa o *analysandum* é falsa), mas, se eu tivesse escolhido tirar uma, eu tê-lo-ia feito (isto é, a proposição que representa o *analysans* é verdadeira).

Assim, um incompatibilista pode sempre rejeitar a análise condicional, sustentando que é uma trivialidade sem qualquer relevância metafísica afirmar que teríamos agido de modo diferente daquele pelo qual agimos, se assim o tivéssemos decidido (i.e., desde que tivéssemos crenças e desejos diferentes daqueles que efetivamente temos). Para esses autores, o problema do livre-arbítrio consiste precisamente em tentar perceber se num mundo determinista é possível que a nossa vontade seja livre, ou seja, se aquilo que decidimos, ou não, fazer é algo que de certa forma podemos controlar, que depende de nós em algum sentido relevante, ou se não passa de mais uma consequência do passado e das leis da natureza que escapa completamente ao nosso controle.

Ora, de acordo com a imagem determinista do mundo, em cada instante existe apenas um estado de coisas possível, pois cada estado de coisas é consequência necessária dos estados do mundo que o antecederam e das leis da natureza. Assim sendo, a história do mundo é como um comboio que viaja numa linha sem bifurcações. O que significa que, num mundo determinista, não faz sentido dizer que poderíamos ter decidido agir de outra forma ou ter desejos diferentes daqueles que efetivamente temos, pois estes são a consequência pelo menos da nossa história pessoal até ao momento e das leis da natureza.

2.4 O princípio das possibilidades alternativas

Resta-nos então voltar a nossa atenção para a premissa (4) do Argumento da Consequência. Como vimos acima, a premissa (4) expressa a ideia que sem possibilidades alternativas não há livre-arbítrio. Esta ideia ficou conhecida na literatura relevante como Princípio das Possibilidades Alternativas (*PPA*). Segundo este princípio:

Princípio das Possibilidades Alternativas

Um determinado agente *S*, faz uma dada ação *A* livremente só se *S* podia ter feito outra coisa que não *A*.

A ideia subjacente a este princípio também parece bastante intuitiva. Aparentemente, se não temos quaisquer possibilidades alternativas ao nosso dispor, então não faz sentido dizer que agimos de livre vontade, pelo menos não em qualquer sentido relevante para a questão da responsabilidade moral.

O Princípio das Possibilidades Alternativas foi aceite de forma praticamente consensual até à publicação do artigo 'Responsabilidade Moral e o Princípio das Possibilidades Alternativas' (1969), de Harry Frankfurt. Nesse artigo, Frankfurt inventou uma experiência mental que constitui um contraexemplo a este princípio. Este tipo de experiência mental ficou conhecido como 'casos de Frankfurt'.

Num típico caso de Frankfurt, existe uma circunstância, *C*, tal que:

1. Um agente *A* toma uma determinada decisão *D*.
2. Se *A* não decidir *D* por si mesmo, *C* entra em ação e força *A* a decidir *D*.
3. *C* em nada contribui para que *A* decida *D*.

Por exemplo, suponhamos que Black é o chefe de uma poderosa organização criminosa e que Jones é um dos mais eficientes assassinos dessa organização. Black quer matar o Presidente e sabe que Jones é a pessoa certa para o trabalho. No entanto, existem rumores de que Jones quer abandonar a profissão, razão pela qual o seu compromisso para com a organização é incerto.

Nesse momento, Black recorda-se de uma das invenções mais recentes dos cientistas da sua organização – o 'neuroscópio'. O 'neuroscópio' é um aparelho que, uma vez introduzido no cérebro de uma pessoa, permite vigiar e controlar os seus estados cerebrais. Black apercebe-se de que o 'neuroscópio' lhe permite resolver dois problemas de uma só vez.

Se implantar secretamente o 'neuroscópio' no cérebro de Jones, ficará a saber se este continua a ser um dos seus fiéis assassinos – pode acontecer que ele decida, por si mesmo, matar o Presidente, sem que seja necessária a intervenção do 'neuroscópio' – ao mesmo tempo que se certifica de que Jones cumpre a sua missão e mata

o Presidente – porque se o 'neuroscópio' detetar qualquer indício de que ele não o vai fazer, entra em ação e força-o a decidir nesse sentido.

Imaginemos, agora, que Jones decide pelos seus próprios motivos matar o Presidente. Nesse caso temos uma situação em que:

1. Jones toma, por si mesmo, a decisão de matar o Presidente.
2. Se Jones não decidisse, por si mesmo, matar o Presidente, o 'neuroscópio' entraria em ação e forçaria Jones a tomar essa decisão.
3. A presença do 'neuroscópio' no cérebro de Jones em nada contribui para a sua decisão de matar o Presidente.

A conclusão que Frankfurt extrai deste tipo de casos é a seguinte: há situações nas quais, apesar de não termos possibilidades alternativas, temos livre-arbítrio. Isto corresponde à negação do princípio das possibilidades alternativas, pois afirma a sua antecedente ('não temos possibilidades alternativas') e nega a sua consequente ('temos livre-arbítrio').

Assim, os compatibilistas contemporâneos concluem que aquilo que é fundamental para que possamos considerar que temos livre-arbítrio não é o facto de termos possibilidades alternativas, mas sim o facto de as nossas ações terem origem na nossa vontade.

A argumentação destes pensadores pode ser reconstruída conforme se segue:

- (1) Temos livre-arbítrio, se, e só se, algumas das coisas que acontecem dependem fundamentalmente da nossa vontade.
- (2) Ainda que o determinismo seja verdadeiro e não existam possibilidades alternativas, algumas das coisas que acontecem dependem fundamentalmente da nossa vontade (tal como demonstram os casos de Frankfurt).
- (3) Logo, ainda que o determinismo seja verdadeiro e não existam possibilidades alternativas, temos livre-arbítrio.

Uma objeção frequentemente utilizada contra os casos de Frankfurt surge sob a forma de um dilema – habitualmente designado 'dilema WKG' devido ao nome dos três autores que avançaram este tipo de

objeção – David Widerker (1995), Robert Kane (1996) e Carl Ginet (1996). Estes começam por chamar a atenção para o seguinte:

- uma vez que pretendem demonstrar a falsidade de *PPA*, os casos de Frankfurt não podem pressupor que o agente se encontra numa situação em que age livremente, apesar de viver num mundo determinista.

De facto, isso condenaria toda esta estratégia ao fracasso, pois implicaria incorrer na falácia da petição de princípio, isto é, estar-se-ia a assumir à partida aquilo mesmo que se pretende provar.

Assim sendo, o defensor da estratégia de Frankfurt vê-se forçado a supor que o agente se encontra num mundo indeterminista. Contudo, num mundo indeterminista, a decisão do agente não pode ser conhecida de antemão, pois, uma vez que a sua decisão não decorre necessariamente do passado e das leis da natureza, a única forma de o elemento interventor saber ao certo qual é a decisão do agente é depois desta ter sido tomada. No entanto, nesse momento, será tarde de mais para atuar sobre ela, pelo que, se temos a intuição de que esta foi uma decisão genuinamente livre, isso pode simplesmente dever-se ao facto de considerarmos que ela ocorreu num cenário em que o agente tinha possibilidades alternativas ao seu dispor. Por sua vez, qualquer intervenção póstuma implicaria que o agente foi vítima de uma interferência externa na sua vontade e que, por conseguinte, não agiu livremente.

Deste modo, para que o elemento interventor possa funcionar devidamente, isto é, de forma meramente contrafactual, não há como evitar assumir a existência de alguma espécie de sinal prévio que indique que o agente vai tomar, por si mesmo, a decisão que era suposto tomar. É, precisamente, neste ponto que o dilema aparece.

- (1) Numa situação frankfurtiana (ou caso de Frankfurt) existe sempre um sinal prévio que indica que o agente vai tomar, por si mesmo, a decisão que era suposto tomar e esse sinal prévio é suficiente nas circunstâncias para a tomada de decisão em causa ou não é.
- (2) Se for suficiente, então Frankfurt está a pressupor que o determinismo é verdadeiro, pois dada a ocorrência do sinal prévio há apenas um curso de acontecimentos possível. Contudo,

como dissemos anteriormente, nesse caso, a experiência mental incorre numa falácia de petição de princípio – isto é, pressupõe aquilo que pretende provar – e, por conseguinte, não é capaz de mostrar que pode haver livre-arbítrio sem haver possibilidades alternativas, ou seja, não consegue provar que o *PPA* é falso.

- (3) Se o sinal não for suficiente, isso significa que o agente dispõe de possibilidades alternativas, pois a ativação do sinal prévio é, ainda assim, compatível com diferentes possibilidades de decisão.
- (4) Mas, nesse caso, a situação frankfurtiana também não é capaz de mostrar que pode haver livre-arbítrio sem haver possibilidades alternativas, ou seja, não consegue provar que o *PPA* é falso.
- (5) Logo, uma situação frankfurtiana não é capaz de mostrar que pode haver livre-arbítrio sem haver possibilidades alternativas, ou seja, não consegue provar que o *PPA* é falso.

Do lado dos defensores da estratégia de Frankfurt tem havido um esforço por conceber casos que não pressupõem a existência de nenhum tipo de sinal prévio. Este tipo de casos pode envolver um interventor contrafactual omnisciente, ou um viajante no tempo, ou um dispositivo que atua preventivamente, isto é, que atua independentemente da decisão do agente, mas cuja eficácia é nula no caso deste tomar a decisão por si mesmo. Esta alternativa é explorada por Alfred Mele e David Robb, no artigo 'Rescuing Frankfurt-Style Cases' (1998). Nesse artigo pode ler-se o seguinte:

O nosso cenário apresenta um agente, Bob, que habita um mundo onde o determinismo é falso. Isto não significa que não existem eventos deterministicamente causados no mundo de Bob. Por exemplo, uma certa partícula radioativa decaiu na cidade de Bob ao meio-dia, e esse evento não foi deterministicamente causado; no entanto, essa partícula tinha sido recrutada como um gatilho aleatório para uma bomba, e assim que a partícula decaísse a explosão da bomba seria deterministicamente causada.

Em t_1 , Black inicia um certo processo determinístico P no cérebro de Bob com o intuito de fazer com que este decida, em t_2

(digamos, uma hora mais tarde) roubar o carro de Ann. O processo, que é monitorizado a partir da consciência de Bob, irá culminar deterministicamente no facto de Bob decidir, em t_2 , roubar o carro de Ann, a não ser que ele decida por si mesmo roubá-lo, em t_2 ... O processo não é de forma alguma sensível a qualquer espécie de 'sinal' daquilo que Bob irá decidir. Acontece simplesmente que, em t_2 , Bob decide por si mesmo roubar o carro, com base no seu próprio processo indeterminista de deliberação [...], e a sua decisão não tem nenhuma causa determinística. Mas se este não tivesse acabado de tomar por si mesmo a decisão de roubar o carro, P teria culminado, em t_2 , na sua decisão de o roubar. (Mele e Robb 1998: 101-102)

A coerência deste cenário pode ser posta em causa. Afinal de contas, como é que Bob pode decidir por si mesmo, em t_2 , roubar o carro de Ann, se existe um certo processo determinístico implantado no seu cérebro com o intuito de o forçar a tomar essa decisão?

Para responder a esta questão, Mele e Robb sugerem que se compare o funcionamento do processo P , com o funcionamento de uma determinada máquina M , que descrevem nos termos que se seguem:

A máquina [M] produz *widgets* artísticos de diferentes formas e cores. A cor dos *widgets* produzidos é determinada pela cor de um porta-bolas (*pb*) que atinge o recetor da máquina num determinado intervalo de tempo. A máquina M está rodeada por vários disparadores de *pbs* automáticos, cada um dos quais contendo *pbs* de várias cores. O aspeto relevante do design mecânico de M , para os nossos propósitos, é relativamente simples. Em primeiro lugar, [...] se um *pb* da cor x atinge o recetor de M e esta ainda não se encontrar no processo de produzir um *widget*, então M inicia de imediato um processo projetado para resultar na produção de um *widget* de cor x . Em segundo lugar, e porque ocasionalmente dois ou mais *pbs* atingem o recetor em simultâneo, o artista projetou a sua máquina de tal forma que de cada vez que isso aconteça (quando M não está ocupada a produzir nenhum *widget*) M inicia de imediato um processo projetado para resultar na produção de um *widget* da cor do *pb* mais à direita. Nenhum outro impacto simultâneo no recetor de M desempenha qualquer papel na ativação de M . (Mele e Robb 1998: 103)

Assim sendo, se, numa situação frankfurtiana, o dispositivo contrafactual fosse programado de modo análogo à máquina *M*, se a decisão de roubar o carro não tivesse sido iniciada espontaneamente, o próprio mecanismo teria eficazmente assegurado que essa decisão era tomada. Acontece simplesmente que, uma vez que o agente toma a decisão por si mesmo, o dispositivo não chega a desempenhar qualquer espécie de papel causal relevante para esse processo de tomada de decisão. Deste modo, os defensores da estratégia de Frankfurt podem afirmar, sem correr o risco de cometer uma petição de princípio, que o agente não tinha efetivamente possibilidades alternativas ao seu dispor, mas, apesar disso, tomou a decisão de roubar o carro por sua própria iniciativa – e, por conseguinte, parece ser legítimo responsabilizá-lo pela mesma.

Contudo, ainda que a estratégia do dilema não seja bem-sucedida, alguns críticos de Frankfurt insistem que existem razões adicionais para rejeitarmos os casos de Frankfurt.

De acordo com tais críticos,⁷ os casos de Frankfurt não funcionam como contraexemplos ao *PPA*, pois não constituem efetivamente situações em que o agente age livremente, apesar de não ter possibilidades alternativas, pois mesmo num caso de Frankfurt existem possibilidades alternativas, a saber, o agente pode sempre:

- (A) tomar por si próprio uma decisão; ou
- (B) ser forçado por uma circunstância contrafactual *C* a tomar essa decisão, o que significa que esta é uma decisão diferente daquela que o agente teria espontaneamente tomado.

Ora, os críticos de Frankfurt defendem que é justamente a existência dessas possibilidades (ainda que escassas) que explica a nossa intuição de que, num típico caso de Frankfurt, o agente é, apesar de tudo, dotado de livre-arbítrio e, conseqüentemente, moralmente responsável pelas suas ações.

A estratégia ficou conhecida como 'estratégia da centelha de

⁷ David Widerker e Michael McKenna, por exemplo, escrevem na introdução da sua coletânea *Moral Responsibility and Alternative Possibilities*: 'Estritamente falando, os exemplos de Frankfurt não eliminam todas as possibilidades alternativas, uma vez que existem efetivamente, incorporadas nos exemplos, centelhas de liberdade' (Widerker & McKenna 2003: 7).

liberdade',⁸ graças a John Martin Fischer que, na obra *The Metaphysics of Free Will* (1994) introduz a expressão no debate sobre o problema do livre-arbítrio. Fischer afirma o seguinte:

Os Casos de Frankfurt parecem, à partida, não envolver qualquer tipo de possibilidades alternativas. Mas, à luz de uma análise mais minuciosa, podemos aperceber-nos de que, embora não envolvam possibilidades alternativas num sentido normal, envolvem ainda assim algum tipo de possibilidades alternativas. Isto significa que, embora os interventores contrafactuais eliminem a maioria das possibilidades alternativas, pode defender-se que eles não as eliminam na totalidade: mesmo num caso tipo-Frankfurt, parece haver uma 'centelha de liberdade'. Consequentemente, existe espaço para se argumentar que estas possibilidades alternativas (as centelhas de Liberdade) têm de estar presentes, mesmo nos casos tipo-Frankfurt, para que exista responsabilidade moral. (Fischer 1994: 134)

Assim, embora uma e a mesma ação venha a ocorrer seja qual for a decisão tomada pelo agente, este teria ao seu dispor um leque (ainda que muito estreito) de possibilidades alternativas, a saber:

- (i) tomar a decisão, por si mesmo, ou
- (ii) não tomar a decisão por si mesmo e ser forçado a tomar essa decisão por um fator externo.

Embora Fischer considere que estas alternativas não são suficientemente robustas para alicerçar a responsabilidade moral, outros julgam que a diferença entre agir e ser coagido é uma diferença relevante para os nossos juízos intuitivos sobre responsabilidade moral e, por conseguinte, pensam que a estratégia das centelhas de liberdade representa um sério desafio para os casos de Frankfurt.

3 O argumento modal da consequência

Uma das versões atualmente mais discutidas do argumento da consequência é a sua versão modal. Nesta formulação, van Inwagen⁹ recorre

⁸ 'Flicker-of-freedom strategy', no original em Inglês.

⁹ Cf. van Inwagen (1983: 93-95).

à Lógica Modal Proposicional para mostrar que partindo do determinismo como premissa seremos conduzidos à conclusão de que ninguém tem, nem nunca teve, genuinamente controlo algum sobre as suas ações.

Para este efeito, van Inwagen começa por acrescentar os operadores habituais em lógica modal – o operador de necessidade (normalmente representado pelo símbolo ' \Box ') e o operador de possibilidade (normalmente representado pelo símbolo ' \Diamond ') – e um terceiro operador, representado pela letra ' N ', que corresponde à expressão 'não depende de nós que ...'.¹⁰ Assim, dada uma proposição P , temos o seguinte sistema de abreviaturas:

- $\Box P$ representa: Necessariamente, P .
- $\Diamond P$ representa: Possivelmente, P .
- NP representa: P (é o caso) e ninguém tem nem nunca teve qualquer escolha acerca disso.

Adicionalmente vamos também considerar o seguinte:

- \vdash representa: A relação de consequência lógica entre fórmulas, quando colocado entre elas; ou a ocorrência de uma fórmula tautológica (i.e., verdadeira em todas as interpretações possíveis dessa fórmula), quando é prefixado à mesma, sem qualquer fórmula à sua esquerda.

De seguida, van Inwagen estabelece que o operador N se rege pelas seguintes regras de inferência:

- Regra (α): $\Box P \vdash NP$

Ou seja, se P é uma proposição necessária (como, por exemplo, ' $2 + 2 = 4$ '), então ninguém tem, ou alguma vez teve, qualquer escolha acerca de P .

- Regra (β): $N(P \rightarrow Q), NP \vdash NQ$

Ou seja, no caso de não depender de nós que P implique Q , e não depender de nós que P , então Q não depende de nós.

¹⁰ Como veremos adiante, numa versão posterior deste argumento, van Inwagen substituiu esta expressão pela seguinte: 'É uma verdade humanamente inalterável que...'. (van Inwagen: 2015: 19)

Para percebermos o que realmente vai estar em jogo no argumento modal da consequência, vamos agora traduzir as ideias fundamentais do argumento para a linguagem formal da Lógica Modal Proposicional acrescida do operador N . Introduzimos, para esse efeito, a seguinte lista de abreviaturas:

- L representa: a conjunção das leis da natureza;
 H representa: uma proposição verdadeira que descreve integralmente um dado estado de coisas no passado remoto;
 P representa: uma qualquer proposição acerca de uma ação, tal como 'eu levantei o meu braço'.

De acordo com o estabelecido, podem, ainda, notar-se duas coisas. A primeira é que P funciona efetivamente como uma variável proposicional que pode ser substituída por qualquer proposição acerca de uma ação, eventualmente diferente de 'eu levantei o meu braço', e a segunda é que ' NP ' abrevia a seguinte proposição:

- Eu levantei o meu braço e ninguém tem nem nunca teve qualquer escolha acerca de eu ter levantado o meu braço.

Isto basta para ficarmos em condições de formalizar as várias perspetivas em confronto, no que diz respeito ao problema da compatibilidade.

- Libertismo $\stackrel{\text{def}}{=} \neg NP$

Ou seja, pelo menos um agente tem, ou teve, o poder de tornar falsa pelo menos uma proposição (a proposição em causa pode ser tanto acerca das ações desse agente, como acerca das suas escolhas).

- Determinismo $\stackrel{\text{def}}{=} \Box((H \wedge L) \rightarrow P)$

Ou seja, necessariamente, a conjunção das leis da natureza juntamente com uma proposição verdadeira que descreve integralmente um dado estado de coisas num passado remoto implicam P (implicam, por exemplo, que eu levantei o meu braço).

- Incompatibilismo $\stackrel{\text{def}}{=} \Box((H \wedge L) \rightarrow P) \rightarrow NP$

Ou seja, caso o determinismo seja verdadeiro, não há livre-arbítrio.

- Compatibilismo $\stackrel{\text{def}}{=} \neg(\Box((H \wedge L) \rightarrow P) \rightarrow NP)$

O compatibilismo, tal como era expectável, corresponde neste momento, à negação do incompatibilismo, ou seja, à tese de que o determinismo é compatível com a existência de livre-arbítrio.¹¹

Com base no anterior conjunto de definições, o Argumento Modal da Consequência pode ser traduzido na linguagem formal da Lógica Modal Proposicional do seguinte modo:

- (1) $\Box((H \wedge L) \rightarrow P)$ [premissa-suposição, def. de determinismo]
- (2) NH [premissa, o passado remoto não depende de nós]
- (3) NL [premissa, as leis da natureza não dependem de nós]
- (4) $\therefore \Box((H \wedge L) \rightarrow P) \rightarrow NP$ [conclusão, tese do incompatibilismo]

O argumento revela-se, assim, uma prova condicional – o que significa que, se partimos da suposição que o determinismo é verdadeiro, somos conduzidos à conclusão de que ninguém tem, ou alguma vez teve, qualquer escolha acerca de qualquer proposição verdadeira, inclusive proposições sobre ações aparentemente livres dos seres humanos.

A premissa (1) corresponde à definição de determinismo, ou seja, à ideia de que a conjunção das leis da natureza com uma proposição verdadeira acerca do passado remoto implica, necessariamente, P (implica, de acordo com o nosso exemplo, que eu levantei o meu braço). Ela diz, por outras palavras, o seguinte:

- tudo o que acontece (incluindo as nossas ações) é uma consequência necessária do passado e das leis da natureza.

Na premissa (2) afirma-se que ninguém tem, ou alguma vez teve, qualquer escolha acerca do passado remoto ou, de outra forma, que

¹¹ David Lewis propõe uma conceção diferente de compatibilismo no artigo 'Are we free to break the laws?' (Cf. Lewis 1981: 113). De acordo com a sua perspectiva, a tese compatibilista pode ser formulada conforme se segue: $\Diamond((\Box(H \wedge L) \rightarrow P) \wedge \neg NP)$. No entanto, uma vez que, ao longo do artigo, Lewis vai assumir a perspectiva – que designa 'determinismo moderado' – coincidente com a negação da tese incompatibilista defendida por van Inwagen, podemos continuar a formular a tese compatibilista nos termos apresentados. Até porque de outra forma essa tese não seria inconsistente com o incompatibilismo, tal como foi apresentado, i.e., como uma tese acerca do mundo atual e não como uma tese acerca de todos os mundos possíveis.

aquilo que aconteceu antes de existirem seres humanos não depende de nós.

Na premissa (3) sustenta-se que ninguém tem, ou alguma vez teve, qualquer escolha acerca das leis da natureza, isto é, afirma-se que as leis da natureza não dependem de nós.

A partir destas premissas tem-se a conclusão (4), de acordo com a qual se o determinismo é verdadeiro, então P (é o caso) e ninguém tem, ou alguma vez teve, qualquer escolha acerca de P . Ora, uma vez que P pode ser substituída por qualquer proposição acerca de uma ação, caso o determinismo seja verdadeiro, então ninguém tem, ou alguma vez teve, qualquer escolha em relação às suas próprias ações. O que também significa, claro, que ninguém poderia ter feito algo diferente do que de facto fez ou ter escolhido (ou, ainda, decidido) agir de forma diferente – e, por conseguinte, caso o determinismo seja verdadeiro, não temos livre-arbítrio.

Para demonstrar a validade do argumento, van Inwagen apresenta a seguinte derivação:

- | | | |
|-----|---|---|
| (1) | $\Box((H \wedge L) \rightarrow P)$ | [premissa-suposição, def. de determinismo] |
| (2) | NH | [premissa, o passado remoto não depende de nós] |
| (3) | NL | [premissa, as leis da natureza não dependem de nós] |
| (4) | $\Box(H \rightarrow (L \rightarrow P))$ | [1, regra de exportação para a lógica modal] |
| (5) | $N(H \rightarrow (L \rightarrow P))$ | [4, regra (α)] |
| (6) | $N(L \rightarrow P)$ | [2 e 5, regra (β)] |
| (7) | NP | [3 e 6, regra (β)] |
| (8) | $\Box((H \wedge L) \rightarrow P) \rightarrow NP$ | [1-7, introdução da condicional] |

Mas será este argumento procedente?

Claro que o compatibilista não pode aceitar a conclusão na linha (8), pois isso seria precisamente a negação da sua tese. Assim, o compatibilista terá de defender, pelo menos uma, das seguintes ideias:

- i. NH é falsa.
- ii. NL é falsa.
- iii. A regra de exportação para a lógica modal é inválida.
- iv. A regra (α) é inválida.
- v. A regra (β) é inválida.

Não nos iremos debruçar sobre *i* e *ii*, pois, uma vez que estas correspondem sensivelmente à *objeção do passado inexistente*, de J. K. Campbell (*i*) e ao compatibilismo de milagres locais, de D. Lewis (*ii*), já foram exploradas anteriormente. Assim, passaremos em seguida a explorar os pontos *iii*, *iv* e *v*.

3.1 O ataque à regra de exportação para a lógica modal

A regra de exportação para a lógica modal estabelece que podemos inferir $\Box(P \rightarrow (Q \rightarrow R))$ a partir de $\Box((P \wedge Q) \rightarrow R)$, ou seja, a regra pode ser formulada do seguinte modo:

$$\Box((P \wedge Q) \rightarrow R) \vdash \Box(P \rightarrow (Q \rightarrow R))$$

Ora, esta regra é, aparentemente, incontroversa para quem quer que aceite o sistema K , que é o mais fraco dos sistemas de lógica modal (no sentido em que K é o sistema de lógica modal com o menor número de axiomas). Isto acontece porque, uma vez que em K podemos colocar o operador ' \Box ' antes de qualquer tautologia da Lógica Proposicional Clássica (LP) e visto que ' $((P \wedge Q) \rightarrow R) \rightarrow (P \rightarrow (Q \rightarrow R))$ ' é uma tautologia em LP , podemos assumir que $\Box(((P \wedge Q) \rightarrow R) \rightarrow (P \rightarrow (Q \rightarrow R)))$ é uma fórmula válida em K ; e, assim sendo, podemos concluir que a *regra da exportação para a lógica modal* é válida em K utilizando o *axioma da distribuição da necessidade* (também designado *Axioma K*, precisamente por ser o axioma que caracteriza este sistema), segundo o qual $(\Box(P \rightarrow Q) \rightarrow (\Box P \rightarrow \Box Q))$. Por substituição teríamos a seguinte fórmula:

$$\Box(((P \wedge Q) \rightarrow R) \rightarrow (P \rightarrow (Q \rightarrow R))) \rightarrow \Box(((P \wedge Q) \rightarrow R) \rightarrow \Box(P \rightarrow (Q \rightarrow R)))$$

Deste modo, rejeitar a *regra de exportação para a lógica modal* implicaria rejeitar o mais fraco dos sistemas de lógica modal, o que equivaleria a rejeitar a própria lógica modal como um todo. Não estou a presumir que isso seja necessariamente um erro, pois embora se considere que a lógica modal tem uma comprovada utilidade no tratamento de vários problemas metafísicos, também é sabido que nem todos os filósofos lhe reconhecem esse mérito (Cf. Quine, 1960).

No entanto, penso que esta crítica é pouco satisfatória, pois consiste basicamente em dizer que o argumento modal da consequência não é bom porque é modal. O que implica pressupor que a ideia de que

a lógica modal como um todo está errada é mais plausível do que o próprio argumento modal da consequência. Além disso, uma vez que os critérios para avaliar a validade de um argumento são, geralmente, inerentes ao próprio sistema formal utilizado na sua estruturação, isso significa que os critérios para avaliar um argumento modal são inerentes à própria lógica modal, e que apenas a própria lógica modal nos fornece as ferramentas necessárias para atacar a estrutura formal do argumento.

3.2 O ataque à regra (α)

A regra (α) sustenta que a seguinte inferência é válida:

$$\Box P \vdash NP$$

Por exemplo, $\Box(2 + 2 = 4) \vdash N(2 + 2 = 4)$, o que significa que se ' $2 + 2 = 4$ ' é necessariamente verdadeira, então ninguém tem ou alguma vez teve o poder de a tornar falsa. Negar ' $N(2 + 2 = 4)$ ' seria alegar que a verdade de ' $2 + 2 = 4$ ' depende, de alguma forma, de nós. Ora, à partida, isso parece descabido, daí que a regra (α) seja intuitivamente válida, mas nem todos estão dispostos a aceitá-la. Em primeiro lugar, importa salientar que, dado que o sistema K não se caracteriza pela relação de reflexividade (i.e., em K não é verdade que qualquer mundo possível acede a si mesmo), a regra (α) não é válida em K . Considerando que a inferência ' $\Box P \vdash P$ ' não é válida em K , podemos imaginar uma circunstância em que ' $\Box P$ ' é verdadeira e ' P ' falsa e, portanto, ' $\Box P \vdash NP$ ' seria uma inferência inválida, visto que, nessa circunstância, ' $\Box P$ ' seria verdadeira e ' NP ' seria falsa, pois esta última abrevia ' P ' e ninguém tem, nem nunca teve, qualquer escolha acerca de ' P ' e, neste caso, a sua primeira conjunta seria falsa, pelo que a conjunção no seu todo também o seria.

Contudo, parece mais plausível aceitar a ideia de que o sistema K é incompleto do que negar a regra (α). Na verdade, a rejeição da relação de reflexividade, associada ao sistema K , parece bastante contraintuitiva, pois uma vez que ' P é necessariamente verdadeira' num mundo possível, M_1 , significa que ' P é verdadeira em todos os mundos possíveis acessíveis a partir de M_1 ', e visto que, em K , nada nos garante que M_1 aceda a si próprio, é possível que em M_1 ambas as proposições que se seguem sejam verdadeiras: ' P é necessariamente

verdadeira' e '*P* é falsa'; o que intuitivamente parece inadequado. Assim, parece mais razoável aceitar que qualquer mundo acede a si mesmo do que aceitar que '*P* é necessariamente verdadeira' e '*P* é falsa' podem ser simultaneamente verdadeiras, o que significa que talvez *K* não seja o melhor sistema de lógica modal para lidar com questões metafísicas como aquelas que estão em jogo no problema do livre-arbítrio.

Uma outra forma de atacar (α) é alegar que existe um Deus todo-poderoso que criou o universo e tudo quanto existe, incluindo as verdades necessárias e, portanto, ainda que seja verdade que, por exemplo, ' $\Box(2 + 2 = 4)$ ', é falso que ninguém tem nem nunca teve escolha sobre o assunto. Contudo, muitos filósofos (e.g., Leibniz, Plantinga, entre outros) defendem que nem um Deus onipotente, caso existisse, conseguiria alterar o valor de verdade de proposições necessárias. Além disso, o problema que temos em mãos diz respeito ao livre-arbítrio humano, não à liberdade de uma hipotética divindade. O que queremos saber é se, efetivamente, nós, os seres humanos, somos (ou não) dotados de livre-arbítrio. Portanto, o facto de haver um Deus todo-poderoso que criou o universo e tudo quanto existe, incluindo as verdades necessárias, é irrelevante para a inferência aqui em causa. O operador '*N*' aplica-se aos seres humanos e, uma vez que é esse o escopo do argumento modal da consequência, isso é tudo quanto precisamos para preservar a validade do mesmo.

Contudo, apesar do seu apelo intuitivo, alguns autores têm vindo a desferir ataques mais sérios à validade de (α). No artigo 'Able to do the impossible' (2016), Jack Spencer desenvolve um engenhoso tipo de contraexemplos a esta regra. Nos casos descritos por Spencer, um dado agente *G* tem a capacidade de realizar uma determinada ação Φ , ainda que seja metafisicamente impossível *G* fazer Φ . Por exemplo,

Suponhamos que o determinismo é verdadeiro. Seja *H* a especificação completa das condições iniciais do universo. Seja *L* a especificação completa das leis da natureza deterministas. Seja (*H* \wedge *L*) a sua conjunção. Suponhamos que *G* não acredita, nunca acreditou e nunca acreditará que (*H* \wedge *L*). *G* nunca leu um livro ou ouviu um programa de rádio sobre as condições iniciais do universo; *G* faltou as aulas porque estava com gripe no dia em que a sua professora de Física deu as condições iniciais e as leis da natureza

nas aulas, e a professora nunca se deu ao trabalho de rever essa matéria. Podemos dizer que $(H \wedge L)$ é um conhecimento muito comum na comunidade de G , que se espera que os alunos do 12.º saibam que $(H \wedge L)$, que muitos dos colegas de turma de G sabem que $(H \wedge L)$, e que G é uma das alunas mais brilhantes da sua turma. A proposição $(H \wedge L)$ não está para além das capacidades cognitivas de G , quer em termos de extensão, quer em termos de complexidade, e não existem obstáculos especiais que impedem G de formar essa crença. (Spencer 2016: 468)

Num caso como o descrito, dado que todos os mundos possíveis em que $(H \wedge L)$ é verdadeira são mundos deterministas idênticos ao mundo de G , são igualmente mundos possíveis onde G desconhece a verdade dessa proposição, logo, não existe qualquer mundo possível no qual G sabe que $(H \wedge L)$, o que significa que a proposição ' G não sabe que $(H \wedge L)$ ' – doravante, K – é necessariamente verdadeira.

Contudo, embora seja o caso que $\Box K$, parece certo, pela forma como o caso é descrito, que G tem a capacidade de saber $(H \wedge L)$ – isto é, $\neg NK$ –, ainda que nunca chegue a exercê-la em nenhum mundo possível. Ora, isso significaria que a verdade de NK não se segue necessariamente de $\Box K$ e, por conseguinte, seria suficiente para mostrar a invalidade de (α) .

Uma estratégia semelhante é seguida por Merluzzi e Lampert, no artigo 'Counterfactuals, counteractuals, and free choice' (2021). Ainda que o alvo dos autores nesse artigo seja uma determinada versão da regra (β) – chamada (β_2) ¹² –, a sua objecção consiste em tentar mostrar que essa regra implica (α) e, por conseguinte, não é válida, dado que existem contraexemplos a (α) . Os autores recorrem ao operador da actualidade – @ – bem como a descrições rigidificadas – como por exemplo 'o actual cantor principal dos *Beatles*' – para construir os seus contraexemplos a (α) . Suponhamos que numa fase inicial não estava ainda muito bem definido quem seria o cantor principal dos *Beatles*, imaginemos, por exemplo, que tanto Paul como John eram candidatos igualmente promissores para o lugar. Contudo, John acabou por se chegar à frente, assumindo ele esse papel. Isto faria com que a proposição 'O actual cantor principal dos *Beatles* é John'

¹² Ver adiante.

fosse verdadeira. Ora, ainda que haja mundos possíveis nos quais o cantor principal dos *Beatles* não é John, a verdade é que isso aconteceu no mundo actual e, uma vez que a expressão utilizada se refere ao mundo actual, seja qual for o mundo possível que tomemos como referência é verdade que, tal como as coisas decorreram no mundo actual, o principal cantor dos *Beatles* neste mundo é, de facto, John. Isto significa que a proposição 'O actual cantor principal dos *Beatles* é John' é necessariamente verdadeira. Contudo, parece que não faz sentido dizer que ninguém tem nem nunca teve qualquer tipo de escolha acerca do assunto, pois John poderia não se ter chegado à frente, deixando assim o lugar à disposição de Paul, ou Paul poderia ter tomado a iniciativa mais cedo, etc. Assim sendo, teríamos um contraexemplo a (α), pois teríamos uma situação possível na qual embora seja verdade que $\Box P$, é falso que NP .

Apesar disso, como iremos ver adiante, existem versões alternativas do argumento da consequência que não recorrem a (α), nem a regras de inferência que impliquem (α), por isso o desafio levantado por este argumento ao compatibilismo subsiste independentemente da validade desta regra de inferência.

3.3 O ataque à regra (β)

A regra (β) sustenta que:

$$N(P \rightarrow Q), NP \vdash NQ$$

van Inwagen (1983: 98) recorre ao seguinte exemplo para ilustrar a plausibilidade desta regra:

- (1) Se o sol explodir no ano 3.000, então acaba a vida na terra no ano 3000 e ninguém tem, ou alguma vez teve, qualquer escolha acerca de 'se o sol explodir no ano 3000, então acaba a vida na terra no ano 3.000'. [$N(P \rightarrow Q)$]
- (2) O sol vai explodir no ano 3.000 e ninguém tem, ou alguma vez teve, qualquer escolha acerca de 'o sol vai explodir no ano 3.000'. [NP]
- (3) Logo, a vida na terra acaba no ano 3.000 e ninguém tem, ou alguma vez teve, qualquer escolha acerca de 'a vida na terra acaba no ano 3.000'. [$\therefore NQ$]

Este argumento é uma instância da regra (β) e, aparentemente, é válido. Para mostrar a invalidade do argumento teríamos de defender que é possível que as premissas sejam verdadeiras e a conclusão falsa. Contudo, tal tentativa parece fracassar, pois ao negarmos a conclusão estamos também a negar alguma das premissas. Se 'a vida na terra acaba no ano 3.000' depende de nós, então ou é falso que não depende de nós que caso o sol venha a explodir no ano 3.000, a vida acaba na terra; ou é falso que o facto de o sol explodir no ano 3.000 não depende de nós. Mas, o facto de isso se verificar neste caso não é suficiente para demonstrar a validade da regra, pois basta que seja possível encontrar um contraexemplo – ou seja, uma situação em que NQ é falsa, apesar de $N(P \rightarrow Q)$ e NP serem verdadeiras – para podermos legitimamente pôr em causa a validade da regra. Mas haverá tal contraexemplo?

3.4 O contraexemplo da análise condicional

Como vimos anteriormente, alguns compatibilistas sugerem que a ideia de possibilidades alternativas seja entendida de uma forma condicional. De acordo com esta interpretação, dizer 'podíamos ter agido de outra forma' deve ser entendido de forma condicional, isto é, deve ser entendido como se aquilo que estamos a afirmar fosse o seguinte: 'podíamos ter agido de outra forma, se e só se, caso tivéssemos escolhido agido de outra forma, teríamos agido de outra forma'.

Ora, no seu artigo 'Compatibilism and the argument from unavoidability' (1987), o compatibilista Thomas Flint recorre à análise condicional para mostrar que é possível encontrar um contraexemplo à regra (β) no interior do próprio *Argumento da Consequência*.

Segundo Flint, a análise condicional mostra que a forma como os incompatibilistas interpretam ' NP ' e ' $\neg NP$ ' deixa de fora algo de uma importância fundamental. O incompatibilista interpreta ' $\neg NP$ ' do seguinte modo:

$\neg NP (I)$: Um agente, S , tem, ou teve, a capacidade de tornar pelo menos uma proposição, P , falsa.

Isto significa que ' $\neg NP (I)$ ' implica que S tem ao seu dispor possibilidades alternativas. Ao contrário de ' N ', que significa ' P é o caso e ninguém tem, nem nunca teve, a capacidade de tornar falsa a proposição

P' , ou seja, relativamente a P ninguém tem, nem nunca teve, possibilidades alternativas.

No entanto, com recurso à análise condicional, um compatibilista como Flint pode interpretar ' $\neg NP$ ' do seguinte modo:

$\neg NP$ (C): Um agente teria a capacidade de tornar pelo menos uma proposição, P , falsa, *se tivesse crenças e desejos diferentes daqueles que efetivamente tem.*

O que implica que, para esse compatibilista, ' NP ' passaria a significar:

NP (C): Ninguém teria o poder de tornar falsa a proposição P , *mesmo que tivesse crenças e desejos diferentes daqueles que efetivamente tem.*

Desta forma, a análise condicional permite ao compatibilista encontrar um contraexemplo à regra (β), no interior do próprio Argumento modal da consequência, pois este estabelece que ' NP ' se segue de ' NL ' e ' $N(L \rightarrow P)$ ' – respetivamente, as linhas 3 e 6 da derivação anteriormente apresentada do argumento. No entanto, através da análise condicional somos conduzidos a uma situação em que essas premissas são verdadeiras, mas é falso que NP , porque nessas circunstâncias mesmo que tivéssemos crenças e desejos diferentes, não conseguiríamos alterar tais coisas como as leis da natureza, ' L ', ou a implicação ' $L \rightarrow P$ ', mas a conclusão ' NP ' seria falsa, pois, de acordo com a análise condicional, se tivéssemos crenças e desejos diferentes, teríamos o poder de tornar falsa a proposição P . Portanto, se a análise condicional fornece uma interpretação correta de ' $\neg NP$ ', então a regra (β) autoriza-nos a inferir uma ideia falsa a partir de premissas verdadeiras, o que implica que a regra (β) é inválida.

Contudo, é legítimo perguntar se o incompatibilista se vê forçado a aceitar este contraexemplo e, por conseguinte, a abandonar a regra (β). Afinal de contas, este pode sempre defender que a regra (β) é mais plausível e intuitiva do que a análise condicional, pois, como vimos anteriormente, a análise condicional não está isenta de problemas.

Aparentemente, mesmo entre os compatibilistas não há consenso acerca da análise condicional. Michael Fara, por exemplo, no seu artigo de 2008, 'Masked Abilities and Compatibilism', defende que a análise condicional não fornece uma interpretação correta do conceito de 'capacidade' envolvido na controvérsia em torno do livre-arbítrio,

pois não fornece uma explicação adequada das situações em que uma capacidade se encontra mascarada.

Fara sugere que a noção de 'capacidade' deve ser entendida de modo análogo à noção de 'disposição', de forma que a primeira permita acomodar a ideia de 'capacidade mascarada', tal como a segunda acomoda a de 'disposição mascarada'. Considere-se, a título de exemplo, um determinado pedaço de porcelana que tem a disposição para se partir. Se este for revestido com um material protetor que o impeça de partir, não dizemos que a porcelana perdeu a disposição para partir. Acontece simplesmente que essa disposição foi mascarada pelo facto de este ter sido revestido com um material protetor.

Analogamente, Fara acredita que existem situações em que podemos dizer que uma capacidade se encontra mascarada, descrevendo a noção de 'capacidade mascarada' nos seguintes termos:

A capacidade de um agente, *S*, para fazer *A* numa circunstância, *C*, é uma capacidade mascarada se, e só se,

- (i) *S* tenta fazer *A*
- (ii) na circunstância *C*;
- (iii) *S* tem a capacidade de fazer *A* enquanto tenta fazê-lo; contudo,
- (iv) *S* não é bem-sucedido a fazer *A*. (Fara 2008: 848)

Suponhamos que Cristiano Ronaldo é um excelente jogador de futebol e que tipicamente é um bom marcador de golos. Isto faz com que a proposição 'O Cristiano Ronaldo tem a capacidade de marcar golos' seja verdadeira. No entanto, sempre que recebe um mau passe na grande área, Cristiano Ronaldo é simplesmente incapaz de realizar a mesma proeza que, noutras circunstâncias, parece não ter qualquer dificuldade em realizar.

Imaginemos agora, que Cristiano Ronaldo recebe um mau passe em plena grande área, mas que, ainda assim, arrisca um remate à baliza. Apesar de não ser bem-sucedido, não podemos dizer que este perdeu a sua capacidade de marcar golos; acontece simplesmente que, nestas circunstâncias, essa capacidade se encontra mascarada. Ou seja, temos uma situação em que:

- (i) Cristiano Ronaldo tenta marcar golo
- (ii) na sequência de um mau passe na grande área.

- (iii) Cristiano Ronaldo não perdeu a capacidade de marcar golos; porém
- (iv) não é bem-sucedido.

Assim, de acordo com esta sugestão, o conceito de 'capacidade' passaria a ser interpretado de forma disposicional (e não condicional), ou seja:

DISP: Um agente, *S*, tem a capacidade para fazer *A* nas circunstâncias *C* se, e só se, está disposto a fazer *A*, quando tenta fazê-lo nessas circunstâncias.

Esta forma de entender o conceito de 'capacidade' é, intuitivamente, melhor do que a análise condicional, que se limita a estabelecer que:

COND: Um agente, *S*, tem a capacidade de fazer *A*, desde que *S* pudesse fazer *A*, se assim o decidisse.

Desde logo porque, conforme foi demonstrado através do exemplo, existem situações em que uma capacidade pode estar mascarada, ou seja, pode ser verdade que um determinado agente, *S*, tem a capacidade de fazer *A*, embora seja falso que este faria *A*, se assim o decidisse.

Assim, se, para escapar à conclusão do argumento modal da consequência, o compatibilista pretende atacar a regra (β), tem de desenvolver uma estratégia alternativa que não dependa da análise condicional. Fara acredita que a sua análise disposicional consiste justamente nessa alternativa, pois permite igualmente construir um contraexemplo à regra beta. Para o demonstrar, Fara sugere que nos debrucemos sobre o caso que se segue:

Suponhamos, por exemplo, que um golfista experiente está perante um *putt* particularmente fácil. Sendo experiente, o golfista tem a capacidade de meter o *putt*. Mas imaginemos que, neste caso, essa capacidade se encontra mascarada por uma súbita rajada de vento. O golfista tenta meter o *putt* mas falha. Substituindo, na regra (β), *P* por 'Ocorre uma súbita rajada de vento' e *Q* por 'O golfista falha', temos um contraexemplo a esta regra. Ninguém é capaz de fazer com que seja o caso que não ocorre uma súbita rajada de vento, e ninguém é capaz de fazer com que seja o caso que ocorreu uma súbita rajada de vento, mas o golfista não falhou – podemos supor que o golfista tinha experiência apenas em *putts*

sem vento. Ainda assim, o golfista é capaz de fazer com que seja o caso que ele não falha – ele tem a capacidade de meter o *putt* (Fara 2008: 862)

Ou seja, Fara pensa que a regra (β) é falsa devido à possibilidade de existirem capacidades mascaradas. Na sua opinião, casos como o apresentado mostram que é possível haver circunstâncias em que existe algo que ninguém pode controlar, a saber, as condições que mascaram uma dada capacidade, C , de um agente, A (no exemplo apresentado, ninguém pode controlar a ocorrência de uma súbita rajada de vento, que mascara a capacidade que o golfista experiente tem de meter o *putt*) [NP], ao mesmo tempo que ninguém controla o facto de essas circunstâncias serem suficientes para impedir que esse agente, A , manifeste essa capacidade, C , (no exemplo apresentado, ninguém pode controlar o facto de a rajada de vento ser suficiente para o golfista falhar o *putt*) [$N(P \rightarrow Q)$], embora seja falso que ninguém tem a capacidade em causa (no exemplo apresentado, é falso que ninguém é capaz de tornar falsa a proposição 'O golfista falha o *putt*'. Afinal de contas, tal como é dito no exemplo, trata-se de um golfista experiente e, por conseguinte, com capacidade para meter aquilo tipo de *putt*, embora, nas circunstâncias em causa, essa capacidade se encontrasse mascarada) [$\neg NQ$].

Contudo, esta estratégia de Fara também não parece ser inteiramente satisfatória. Se acrescentarmos que o facto de ter ocorrido uma rajada de vento faz parte das circunstâncias em que o golfista se encontrava, então não podemos dizer que este tinha a capacidade de meter o *putt*. Aliás, o próprio Fara parece reconhecer isso quando afirma que 'podemos supor que o golfista tinha experiência apenas em *putts* sem vento'. Analogamente, se incluirmos o facto de que estamos situados num mundo determinista às circunstâncias em que habitualmente agimos, nunca poderemos dizer que tínhamos capacidade para agir de outro modo, i.e., para agir de um modo diferente daquele que estava predeterminado, pelo passado e pelas leis da natureza, que agiríamos.

Deste modo, podemos considerar que a estratégia da análise disposicional de Fara falha ao vindicar o compatibilismo, porque, uma vez que num mundo determinista, nunca podemos dizer que estamos dispostos a ser bem-sucedidos ao agir de outro modo quando tenta-

mos fazê-lo, numa circunstância em que estava predeterminado que não o faríamos, ou seja, nunca podemos dizer que temos a capacidade para agir de outro modo.

Dada a suposição de que o determinismo é verdadeiro, façamos nós o que fizermos, estaremos a fazer aquilo que estava predeterminado (pelo passado e pelas leis da natureza) e, por conseguinte, não podemos dizer que estamos dispostos a ser bem-sucedidos a agir de um modo diferente daquele que estava predeterminado. Nessa circunstância, sempre que somos bem-sucedidos a fazer seja o que for, é porque essa forma de agir é precisamente aquela que estava predeterminada, o que significa que não estaremos a ser bem-sucedidos a agir de outro modo.

Ao contrário do que acontece com capacidades como cozinhar, saltar à corda ou andar de bicicleta, que, mesmo assumindo que o determinismo é verdadeiro, podem, ou não, vir a manifestar-se consoante o que aconteceu no passado e as leis da natureza, a capacidade de agir de um modo diferente daquele que estava predeterminado é uma capacidade que nunca pode vir a manifestar-se num mundo determinista.

No caso descrito por Fara, dadas as circunstâncias em que se encontrava, havia várias capacidades, neste segundo sentido, que o golfista podia ter manifestado, mas se considerarmos que o determinismo é verdadeiro, seremos forçados a concluir que o passado e as leis da natureza implicam necessariamente se alguma delas se viria a manifestar ou não e, em caso afirmativo, qual delas seria exatamente.

Assim sendo, mesmo de acordo com a análise disposicional proposta por Fara, agir de um modo diferente daquele que estava predeterminado é uma capacidade que nenhum agente situado num mundo determinista pode aspirar possuir; portanto, à semelhança do que acontece com a análise condicional, esta linha de defesa do compatibilismo também parece estar condenada ao fracasso. Mas isto não significa que os compatibilistas tenham esgotado todas as alternativas disponíveis para tentar refutar o argumento modal da consequência.

Uma outra alternativa bastante explorada pelos compatibilistas consiste em mostrar que a regra (β) permite concluir que o livre-arbítrio não só é compatível, como pressupõe o determinismo. Peter van Inwagen rotulou este tipo de estratégia argumentativa de

'Argumento da *Mind*', por ter sido sistematicamente confrontado com ela nas páginas dessa revista. Vejamos mais detalhadamente qual é a estrutura subjacente a este tipo de argumentação.

3.5 *O argumento da Mind*

O argumento da *Mind* é um argumento compatibilista que visa mostrar que o livre-arbítrio não só é compatível com o determinismo, como implica que este é verdadeiro. De acordo com van Inwagen:

O Argumento da *Mind* desenvolve-se identificando o indeterminismo com o acaso e argumentando que um ato que ocorre por acaso, se é que se pode chamar ato a um evento que ocorre por acaso, não pode estar sob o controlo do seu alegado agente e, por conseguinte, não pode ser realizado livremente. Os proponentes [deste argumento] concluem, portanto, que o livre-arbítrio não só é compatível como implica o determinismo. (van Inwagen 1983: 16)

A ideia principal do argumento é reduzir o Argumento modal da consequência ao absurdo mostrando que aceitar a validade da regra (β) conduz a resultados inaceitáveis para a maioria dos incompatibilistas. Para compreender o argumento temos de começar por imaginar um mundo onde os eventos relevantes para a formação de ações livres são indeterminados – as ações são causadas, mas não determinadas, por estados anteriores da mente (tais como, por exemplo, crenças e desejos) do agente. Num tal mundo indeterminista, ninguém tem escolha sobre se um estado da mente terá um resultado particular ou não. De igual modo, o estado da mente de uma pessoa, sendo o resultado final de eventos anteriores sobre os quais ela não tem escolha, também é algo sobre o qual ela não tem qualquer tipo de controlo. Portanto, uma vez que nesse mundo ninguém teria escolha sobre qualquer coisa que seja relevante para produzir as suas ações, aparentemente, ninguém teria escolha sobre essas próprias ações. O argumento pode ser explicitado conforme se segue:

Suponha-se que estamos num mundo indeterminista e que as ações são causadas, mas não determinadas, por conjuntos particulares de

crenças e desejos de agentes. Assim, de acordo com o seguinte sistema de abreviaturas:

- CD : um particular conjunto crença-desejo de um dado agente.
- P : uma ação exclusivamente produzida por CD (considere-se que CD é a única coisa causalmente relevante para a ocorrência de P).

Dada a suposição de que vivemos num mundo indeterminista com as características descritas, P é uma consequência indeterminada de CD e ninguém tem, ou alguma vez teve, escolha sobre se P se segue (ou não) de CD , ou seja:

- $\neg \Box (((H \wedge L) \rightarrow P) \rightarrow N(CD \rightarrow P))$

De igual modo, uma vez que, nesse tipo de mundo, CD passaria a ser encarado como uma consequência indeterminada de factos anteriores sobre os quais o agente não tem controlo, parece que ninguém tem, ou alguma vez teve, escolha sobre ter CD . O que significa que:

- $\neg \Box (((H \wedge L) \rightarrow P) \rightarrow NCD)$

Dadas estas premissas, podemos construir o argumento conforme se segue:

- | | | |
|-----|--|-------------------------------------|
| (1) | $\neg \Box ((H \wedge L) \rightarrow P)$ | [suposição da negação determinismo] |
| (2) | $\neg \Box ((H \wedge L) \rightarrow P) \rightarrow N(CD \rightarrow P)$ | [premissa] |
| (3) | $\neg \Box ((H \wedge L) \rightarrow P) \rightarrow NCD$ | [premissa] |
| (4) | NCD | [1 e 3, <i>Modus Ponens</i>] |
| (5) | $N(CD \rightarrow P)$ | [1 e 2, <i>Modus Ponens</i>] |
| (6) | NP | [4 e 5, regra (β)] |
| (7) | $\neg \Box ((H \wedge L) \rightarrow P) \rightarrow NP$ | [1-6, introdução da condicional] |

As premissas 2 e 3 representam as consequências de um agente estar situado num mundo indeterminista (1). O que significa que se supusermos que o indeterminismo é verdadeiro, então não temos livre-arbítrio. Isto significa que se a regra (β) for válida, então o livre-arbítrio não é compatível com o indeterminismo.

Contudo, a menos que seja complementado com a uma premissa adicional como:

(0.1) O livre-arbítrio tem de ser compatível com o determinismo ou com o indeterminismo.

este argumento não é, por si só, capaz de estabelecer a verdade do compatibilismo. Isto porque um incompatibilista que não aceite o dilema presente em (0.1) pode optar por abraçar as consequências que resultam de aceitar a regra (β) e concluir que, uma vez que o livre-arbítrio não é compatível nem com o determinismo (tal como pretende demonstrar o argumento modal da consequência), nem com o indeterminismo (tal como é sugerido pelo argumento da *Mind*), então o livre-arbítrio é simplesmente impossível, ou seja, pode optar por subscrever o impossibilismo.

Assim, sem essa premissa adicional (altamente disputável), esta objeção limita-se a exercer pressão sobre os incompatibilistas que não estejam dispostos a negar a existência de livre-arbítrio (i.e., sobre os libertistas), pois deixa-os entalados entre duas alternativas que eles rejeitam: abandonar a crença no livre-arbítrio, ou abrir mão da regra (β) e, conseqüentemente, deitar por terra o argumento modal da consequência que suporta a sua crença no incompatibilismo.

Por conseguinte, é legítimo perguntar se não haverá uma forma mais eficaz de atacar a regra (β). Haverá algum contraexemplo que seja capaz de estabelecer, de uma vez por todas, a invalidade desta regra?

Num artigo intitulado 'A Reconsideration of an Argument Against Compatibilism' (1996), Thomas McKay e David Johnson desenvolveram um contraexemplo à regra (β) que não parece padecer deste tipo de limitações.

3.6 O contraexemplo de McKay e Johnson

O contraexemplo de McKay e Johnson não é um ataque direto à regra (β). Em vez disso estes autores decidiram mostrar que (β) implica que podemos aplicar ao operador N o Princípio da Aglomeração (*PAGL*). O *PAGL* estabelece o seguinte:

PAGL: Diz-se que um dado operador, χ , está sujeito ao princípio da aglomeração quando de χP e χQ se segue que $\chi(P \wedge Q)$.
Ou seja: $\chi P, \chi Q \vdash \chi(P \wedge Q)$.

Assim, da aplicação do *PAGL* ao operador *N* resulta o seguinte:

PAGL (N): De *P* (é o caso) e ninguém tem, ou alguma vez teve, qualquer escolha acerca de *P* e *Q* (é o caso) e ninguém tem, ou alguma vez teve, qualquer escolha acerca de *Q*, segue-se que $P \wedge Q$ e ninguém tem, ou alguma vez teve, qualquer escolha acerca de *P* e *Q*. [$NP, NQ \vdash N(P \wedge Q)$]

Mas McKay e Johnson mostram que não é possível inferir $N(P \wedge Q)$ a partir de *NP* e *NQ*, e, por conseguinte, mostram (por *Modus Tollens*) que a regra (β) é inválida. O exemplo de McKay e Johnson baseia-se no seguinte cenário:

Imaginemos que temos uma moeda que poderia ser lançada ao ar em condições normais e que não está viciada. Considerando a seguinte lista de abreviaturas:

- *P*: a moeda não sai cara,
- *Q*: a moeda não sai coroa,

o contraexemplo pode ser formulado assim:

- (1) *NP*
- (2) *NQ*
- (3) $N(P \wedge Q)$

Ora, dadas as circunstâncias descritas no exemplo, '*NP*' é uma premissa verdadeira, pois ao lançar a moeda ao ar ninguém tem, nem nunca teve, qualquer escolha acerca de a moeda não sair cara, e o mesmo acontece com '*NQ*', pois ao lançar a moeda ao ar ninguém tem, nem nunca teve, qualquer escolha acerca de a moeda não sair coroa. Contudo, ' $N(P \wedge Q)$ ' é uma conclusão falsa, pois afirma que ninguém tem, nem nunca teve, qualquer escolha acerca de a moeda não sair cara nem coroa; ou seja, não depende de nós alterar o valor de verdade da proposição ' $(P \wedge Q)$ ' e, no entanto, ao lançarmos a moeda ao ar, nós podemos alterar o valor de verdade dessa proposição, visto que ao fazer o lançamento a moeda vai sair cara ou coroa, ' $(\neg P \vee \neg Q)$ '. Ora, como a negação de ' $(P \wedge Q)$ ' é logicamente equivalente a ' $(\neg P \vee \neg Q)$ ', então ao lançar a moeda ao ar temos a capacidade de tornar essa proposição falsa. O que mostra que a aplicação do *PAGL* ao operador '*N*' não é válida.

Em seguida, McKay e Johnson fazem notar que a regra (β) implica que esse princípio se aplica ao operador 'N', conforme se demonstra em seguida:

- | | |
|--|-------------------------------------|
| (1) NP | [premissa] |
| (2) NQ | [premissa] |
| (3) $\Box(P \rightarrow (Q \rightarrow (P \wedge Q)))$ | [necessidade de uma verdade lógica] |
| (4) $N(P \rightarrow (Q \rightarrow (P \wedge Q)))$ | [3, regra (α)] |
| (5) $N(Q \rightarrow (P \wedge Q))$ | [1 e 4, regra (β)] |
| (6) $N(P \wedge Q)$ | [2 e 5, regra (β)] |

Este contraexemplo revela-se eficaz na refutação da regra (β), o que significa que o incompatibilista terá de abandonar ou reformular o Argumento modal da consequência.

3.7 Resposta ao contraexemplo de McKay e Johnson

No artigo 'Some Thoughts on An Essay on Free Will' (2015), van Inwagen reconhece a pertinência deste ataque à regra (β). De acordo com van Inwagen, a possibilidade de um contraexemplo como aquele que foi apresentado por McKay e Johnson escapou-lhe porque supôs, erradamente, que 'a única maneira de não termos ter escolha acerca do valor de verdade de uma proposição é esse valor estar, de certa forma, tão rigidamente "fixado" que ninguém seria capaz de o modificar' (van Inwagen 2015: 19), e prossegue dizendo que não viu que 'existe outra maneira de não termos escolha acerca do valor de verdade de uma proposição: o facto de esse valor de verdade ser meramente fruto do acaso' (van Inwagen 2015: 19).

Assim, para bloquear este tipo de contraexemplos, van Inwagen sugere que a regra (β) seja substituída por uma versão revista da mesma. Esta revisão passa por uma interpretação diferente do operador 'N'. Em vez de 'Não depende de nós que...', 'N' passa a ser entendido do seguinte modo: 'É uma verdade humanamente inalterável que...'¹³. Ou seja:

$NP \leftrightarrow P$ e não existe ninguém que pudesse ter feito algo tal que se essa

¹³ Esta formulação já tinha sido utilizada por van Inwagen em 2008, na obra *Metaphysics. The big questions*.

pessoa o tivesse feito isso poderia ter tido como consequência $\neg P$.

Assim, de acordo com esta sugestão, a nova regra (β) revista estabelece o seguinte:

Regra (β) revista:

Se é uma verdade humanamente inalterável que P e é uma verdade humanamente inalterável que P implica Q , então podemos concluir que é uma verdade humanamente inalterável que Q .

Ora, esta nova versão de (β) acaba por suportar inferências como a que se segue:

- (1) É uma verdade humanamente inalterável que P .
- (2) É uma verdade humanamente inalterável que P implica Q .
- (3) Logo, é uma verdade humanamente inalterável que Q .

Segundo van Inwagen, esta nova versão de (β) resiste ao contraexemplo em causa, pois, de acordo com esta interpretação, as suas premissas não seriam verdadeiras. Na sua opinião, 'A moeda não sai cara' não é uma verdade humanamente inalterável (e o mesmo se aplica, claro, à proposição 'A moeda não sai coroa'). Isto porque qualquer pessoa com a capacidade de atirar uma moeda ao ar poderia ter feito algo tal que se o tivesse feito isso poderia ter tido como consequência a falsidade da proposição 'A moeda não sai cara' (assim como da proposição 'A moeda não sai coroa').

3.8 β_2 entra em campo

Outra forma de superar o problema levantado por McKay e Johnson, foi explorada por David Widerker (1987), Alicia Finch e Ted A. Warfield (1998), que propuseram a substituição da regra (β) pela regra (β_2), de acordo com a qual:

Regra (β_2): $\Box(P \rightarrow Q), NP \vdash NQ$

Ou seja, não se tem escolha sobre as consequências lógicas daquelas verdades acerca das quais não se tem escolha.

Conforme se pode perceber, esta nova versão da regra (β) não implica o princípio de aglomeração relativamente ao operador 'N':

- (1) NP [premissa]
- (2) NQ [premissa]
- (3) $\Box(P \rightarrow (Q \rightarrow (P \wedge Q)))$ [necessidade de uma verdade lógica]
- (4) $N(P \rightarrow (Q \rightarrow (P \wedge Q)))$ [1 e 3, regra (β_2)]
- (5) ...

E esta nova versão da regra (β) também não permite a formulação anterior do argumento da *Mind*:

- (1) $\neg\Box((H \wedge L) \rightarrow P)$ [suposição da negação determinismo]
- (2) $\neg\Box((H \wedge L) \rightarrow P) \rightarrow N(CD \rightarrow P)$ [premissa]
- (3) $\neg\Box((H \wedge L) \rightarrow P) \rightarrow NCD$ [premissa]
- (4) NCD [1 e 3, *Modus Ponens*]
- (5) $N(CD \rightarrow P)$ [1 e 2, *Modus Ponens*]
- (6) ...

Mas esta nova versão da regra (β) permite-nos fazer uma nova derivação da conclusão do argumento modal da consequência:

- (1) $\Box((H \wedge L) \rightarrow P)$ [premissa-suposição, def. de determinismo]
- (2) $N(H \wedge L)$ [premissa, passado e leis da natureza não dependem de nós]
- (3) NP [1 e 2, (β_2)]
- (4) $N((H \wedge L) \rightarrow P) \rightarrow NP$ [1-3, introdução da condicional]

Contudo, parecem existir boas razões para nos preocuparmos com estas novas versões de (β). Em 2001, no artigo 'The Consequence Argument and the Mind Argument', Dana Nelkin propôs uma nova formulação do argumento da *Mind* que é válida com a regra (β_2). A estrutura central do argumento mantém-se inalterada:

- (1) Se não tivermos escolha sobre as coisas que constituem tudo o que é relevante para produzir as nossas ações, então não temos escolha sobre as nossas próprias ações.
- (2) Se o indeterminismo for verdadeiro, então não temos escolha sobre as coisas que constituem tudo o que é relevante para produzir as nossas ações.
- (3) Logo, se o indeterminismo for verdadeiro, não temos escolha sobre as nossas próprias ações.

Ora, a primeira premissa deixa-se captar por (β_2) , que alega que se não tivermos escolha sobre um estado total de coisas que implica logicamente um segundo estado de coisas, então não temos escolha sobre o segundo estado de coisas. A segunda premissa expressa as consequências de vivermos num mundo indeterminista, tal como foi anteriormente descrito. É plausível que, num mundo indeterminista, uma pessoa não tenha escolha sobre a conjunção da proposição que especifica o seu estado mental e a proposição de que o seu estado mental causa as ações que atualmente ela tem. De maneira que o argumento da Mind pode ser reformulado do seguinte modo:

- (5) $\neg \Box ((H \wedge L) \rightarrow P)$ [suposição da negação determinismo]
- (6) $\neg \Box ((H \wedge L) \rightarrow P) \rightarrow N(CD \wedge (CD \rightarrow P))$ [premissa]
- (7) $N(CD \wedge (CD \rightarrow P))$ [1 e 2, *Modus Ponens*]
- (8) $\Box (CD \wedge (CD \rightarrow P))$ [premissa, verdade lógica]
- (9) NP [2 e 3, regra (β_2)]
- (10) $\neg \Box ((H \wedge L) \rightarrow P) \rightarrow NP$ [1-5, regra introdução da condicional]

Isto significa que um libertista, como van Inwagen, continua a ter problemas, mesmo com a nova regra (β_2) , pois, aparentemente, quer o mundo seja determinista, quer seja indeterminista, não parece haver espaço para a conceção libertista de livre-arbítrio.

Além disso, Alex Blum (2000, 2003) mostrou que ambas as versões de (β) implicam o fatalismo. Schnieder (2008) rejeita a crítica de Blum por considerar que esta se baseia num princípio inválido, mas formula o seu próprio contraexemplo a (β) . Marco Hausmann (2018) argumentou, com base em Blum e Schnieder, que o argumento da consequência está entalado num dilema que conduz a resultados inaceitáveis em ambas as suas bifurcações.

Assim, nada no que foi dito até aqui pode garantir que existe uma versão de (β) completamente imune a qualquer tipo de contraexemplo, por isso, a menos que possamos fornecer uma prova formal da sua validade, a nossa confiança nessa regra não pode ser estabelecida de forma segura e não passa de um mero apelo às nossas intuições, o que, conforme se pode constatar em relação à confiança depositada por van Inwagen na regra (β) original, é muitíssimo insatisfatório. Foi precisamente por esse motivo, que, no artigo 'Incompatibilism Proved' (2013), Alexander Pruss se propôs desenvolver uma prova formal de (β_2) ,

estabelecendo assim, de uma vez por todas, a validade do argumento da consequência. Vejamos, em seguida, em que consiste essa prova.

3.9 Uma prova formal de β_2

Para provar a validade de (β_2) , Pruss começa por oferecer uma definição mais detalhada do operador 'N':

$$NP \stackrel{\text{def}}{=} P \wedge \neg \exists x \exists \alpha (\text{Can}(x, \alpha) \wedge (\text{Does}(x, \alpha) \Box \rightarrow \neg P))$$

Ou seja, P e não existe nenhum x (variável que pode ser substituída por qualquer ser humano) tal que existe um α (variável que pode ser substituída por qualquer ação) tal que x é capaz de fazer α e caso x tivesse feito α , P teria sido falsa.

Em seguida, serve-se da semântica de David Lewis para contrafactuais para justificar a aceitação do seguinte princípio (ao qual dá o nome de *weakening*):

$$\textit{Weakening: } P \Box \rightarrow Q, \Box(Q \rightarrow R) \vdash Q \Box \rightarrow R$$

Ou seja, de se P tivesse sido o caso, Q teria igualmente sido o caso e Q implica necessariamente R segue-se validamente que se Q tivesse sido o caso, teria igualmente sido o caso que R .

A justificação apresentada por Pruss é a seguinte: se P for impossível, então $P \Box \rightarrow Q$ é trivialmente verdadeira. Se P for possível, então $P \Box \rightarrow Q$ é verdadeira se, e só se, existe um mundo M_1 no qual P e Q são ambas verdadeiras que é mais próximo do mundo atual do que qualquer outro mundo no qual P é verdadeira e Q falsa. Seja M_1 esse mundo. Uma vez que Q implica necessariamente R , R será igualmente verdadeira em M_1 . Seja M_2 um qualquer mundo em que P e $\neg R$ são verdadeiras. Ora, visto que Q implica necessariamente R , este seria um mundo no qual P e $\neg Q$ são verdadeiras, mas isso faria de M_2 um mundo mais afastado do mundo atual do que M_1 . Portanto, $Q \Box \rightarrow R$ segue-se validamente de $P \Box \rightarrow Q$ e $\Box(Q \rightarrow R)$.

Conjugando a definição de 'N' apresentada com o princípio *weakening*, Pruss pode formular a seguinte demonstração de (β_2) :

- | | |
|--|------------|
| (1) NP | [premissa] |
| (2) $\Box(P \rightarrow Q)$ | [premissa] |
| (3) $P \wedge \neg \exists x \exists \alpha (\text{Can}(x, \alpha) \wedge (\text{Does}(x, \alpha) \Box \rightarrow \neg P))$ | [1, df.] |

- (4) $\neg\exists x \exists\alpha (\text{Can}(x, a) \wedge (\text{Does}(x, \alpha) \Box\rightarrow \neg P))$ [3, E \wedge]
(5) $\forall x \forall\alpha \neg(\text{Can}(x, a) \wedge (\text{Does}(x, \alpha) \Box\rightarrow \neg P))$ [4, De Morgan]
(6) $\neg(\text{Can}(b, a) \wedge (\text{Does}(b, a) \Box\rightarrow \neg P))$ [de 5, E \forall]
(7) $\text{Can}(b, a) \rightarrow \neg(\text{Does}(b, a) \Box\rightarrow \neg P)$ [6, Equivalência]
(8) $\Box((P \rightarrow Q) \rightarrow (\neg Q \rightarrow \neg P))$ [nec taut]
(9) $\Box(P \rightarrow Q) \rightarrow \Box(\neg Q \rightarrow \neg P)$ [8, distr.]
(10) $\Box(\neg Q \rightarrow \neg P)$ [2,9 E \rightarrow]
(11) $\text{Can}(b, a)$ [sup]
(12) $\text{Does}(b, a) \Box\rightarrow \neg Q$ [sup]
(13) $\text{Does}(b, a) \Box\rightarrow \neg P$ [10, 12 *weakening*]
(14) $\neg(\text{Does}(b, a) \Box\rightarrow \neg P)$ [7, 11 E \rightarrow]
(15) $\neg(\text{Does}(b, a) \Box\rightarrow \neg Q)$ [12-14, *reductio*]
(16) $\text{Can}(b, a) \rightarrow \neg(\text{Does}(b, a) \Box\rightarrow \neg Q)$ [11-15, I \rightarrow]
(17) $\neg(\text{Can}(b, a) \wedge (\text{Does}(b, a) \Box\rightarrow \neg Q))$ [16, Equivalência]
(18) $\forall x \forall\alpha \neg(\text{Can}(x, \alpha) \wedge (\text{Does}(x, \alpha) \Box\rightarrow \neg Q))$ [17, I \forall]
(19) $\neg\exists x \exists\alpha (\text{Can}(x, \alpha) \wedge (\text{Does}(x, \alpha) \Box\rightarrow \neg Q))$ [18, *De Morgan*]
(20) $P \rightarrow Q$ [2, T]
(21) P [3, E \wedge]
(22) Q [20, 21 E \rightarrow]
(23) $Q \wedge \neg\exists x \exists\alpha (\text{Can}(x, a) \wedge (\text{Does}(x, \alpha) \Box\rightarrow \neg Q))$ [19, 22 I \wedge]
(24) NQ [23, df.]

Com esta prova da validade de (β_2) , Pruss está convencido de que o debate em torno da validade do argumento da consequência está resolvido e de que podemos, por isso, voltar as nossas atenções para a sua solidez. Contudo, os problemas subsistem, pois, como vimos acima, Merluzzi e Lampert (2021), defendem que (β_2) implica (α) e existem contraexemplos a (α) . Já vimos anteriormente em que consistem os contraexemplos de Merluzzi e lampert a (α) , resta mostrar de que forma estes estabelecem que (β_2) implica (α) .

Segundo os autores, segue-se de (β_2) que seja qual for o enunciado necessariamente verdadeiro que se tenha em mente, ninguém alguma vez teve qualquer tipo de escolha acerca da sua verdade. Tome-se, a título de exemplo, uma qualquer verdade necessária, digamos $\Box\psi$.

Se ψ é uma verdade necessária, então é necessariamente implicada por qualquer proposição. Assim, tomando uma dada proposição φ , ψ é necessariamente implicada por φ , ou seja, $\Box(\varphi \rightarrow \psi)$. Mas se assumirmos que φ é uma proposição acerca da qual ninguém tem ou alguma vez teve qualquer tipo de escolha, ou seja, $N\varphi$, segue-se, por (β_2) , que $N\psi$, ou seja, que ψ e ninguém tem nem nunca teve qualquer escolha acerca disso. Isto significa que seja qual for a verdade necessária de que partimos, (β_2) permite-nos concluir que ninguém tem nem nunca teve qualquer escolha acerca da mesma, que é precisamente aquilo que era estabelecido por (α) .

No mesmo artigo, os autores propõem uma nova versão de (β) , à qual chamaram de (β_3) , que permite uma nova formulação do argumento da consequência. Contudo, a posição dos autores é a de que contraexemplos análogos aos formulados para (β_2) poderiam igualmente ser levantados a propósito desta nova regra, bem como de outras similares, pondo assim em causa a validade de qualquer *princípio de transferência de incapacidade* que possa ser utilizado em novas versões do argumento da consequência.

4 Conclusão

O argumento da consequência é absolutamente incontornável no que toca à discussão do problema do livre-arbítrio e do determinismo, e tem animado o debate em torno do mesmo. Haverá uma forma satisfatória de estabelecer a sua solidez? Como vimos, grande parte da discussão atual gravita em torno do *princípio de transferência de incapacidade*, mais concretamente, da sua contraparte modal: a regra (β) , nas suas diferentes formulações. Alguns autores procuraram demonstrar que esta enfrenta sérios contraexemplos ou tem implicações inaceitáveis, outros tentam melhorar a sua formulação para responder a esses problemas, e há ainda quem tente oferecer algum tipo de demonstração formal da mesma, apenas para ver essa suposta prova ser desafiada por alguém põe em causa alguns dos pressupostos em que esta assenta.

Luis Verissimo
luisverissimo@campus.ul.pt

Referências

- Beebe, Helen. 2003. Local Miracle Compatibilism. *Noûs* 37(2): 258–277.
- Blum, Alex. 2000. ‘N’. *Analysis* 60 (3): 284–286.
- Blum, Alex. 2003. The Core of the Consequence Argument. *Dialectica* 57 (4): 423–429.
- Campbell, Joseph K. 2007. Free will and the necessity of the past. *Analysis* 67(2): 105–111.
- Chisholm, Roderick M. 1964. Human Freedom and the Self. In T. Crane; K. Farkas (Eds.) *Metaphysics: A Guide and Anthology*. Oxford: Oxford University Press, 2004: 720–729.
- Fara, Michael. 2008. Masked Abilities and Compatibilism. *Mind* 117(468): 843–865.
- Finch, Alicia e Ted A Warfield. 1998. The Mind argument and libertarianism. *Mind* 107(427): 515–528.
- Fischer, John Martin. 1994. *The Metaphysics of Free Will*. Oxford: Blackwell.
- Flint, Thomas P. 1987. Compatibilism and the argument from unavoidability. *Journal of Philosophy* 84 (August), 423–40.
- Frankfurt, Harry. 1969. Alternate Possibilities and Moral Responsibility. *The Journal of Philosophy* 66(23), 829–839.
- Ginet, Carl. 1966. Might We Have No Choice? In K. Lehrer (Ed.), *Freedom and Determinism*. New York: Prometheus Books, 1976, 87–104.
- Ginet, Carl. 1996. In Defense of the Principle of Alternative Possibilities: Why I Don’t Find Frankfurt’s Argument Convincing. *Philosophical Perspectives* 10, 403–417.
- Graham, Peter A. 2008. A Defense of Local Miracle Compatibilism. *Philosophical Studies* 140(1), 65–82.
- Hausmann, Marco. 2018. The consequence argument ungrounded. *Synthese* 195 (11): 4931–4950.
- Kane, Robert. 1996. *The Significance of Free Will*. Oxford: Oxford University Press.
- Lehrer, Keith. 1968. Cans without Ifs. *Analysis* 29(1), 29–32.
- Lewis, David. 1979. Counterfactual Dependence and Time’s Arrow. *Noûs* 13(4), 455–476.
- Lewis, David. 1981. Are We Free to Break the Laws? *Theoria* 47(3), 113–121.
- McKay, Thomas; Johnson, David. 1996. A Reconsideration of an Argument against Compatibilism. *Philosophical Topics* 24(2), 113–122.
- Mele, Alfred e David Robb. 1998. Rescuing Frankfurt-Style Cases. *Philosophical Review* 107(1), 97–112.
- Nelkin, Dana. 2001. The Consequence argument and the Mind argument. *Analysis* 61(270), 107–115.
- Oakley, Shane. 2006. Defending Lewis’s Local Miracle Compatibilism. *Philosophical Studies* 130(2), 337–349.
- Pruss, Alexander. 2013. Incompatibilism Proved. *Canadian Journal of Philosophy* 43 (4): 430 – 437.
- Quine, Willard van O. 1960. *Word and Object*. Cambridge: M.I.T. Press.
- Schnieder, Benjamin. 2008. On What We Can Ensure. *Synthese* 162 (1): 101–115.
- van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.
- van Inwagen, Peter. 2008. *Metaphysics: The Big Questions*. Oxford: Blackwell.
- van Inwagen, Peter. 2015. Some Thoughts on An Essay on Free Will. *The Harvard Review of Philosophy* 22, 16–30.
- Widerker, David. 1987. On an Argument for Incompatibilism. *Analysis* 47, 37–41.
- Widerker, David. 1995. Libertarianism and Frankfurt’s Attack on the Principle of Alternative Possibilities. *Philosophical Review* 104(2), 247–261.
- Widerker, David e Michael McKenna (Eds). 2003. *Moral Responsibility and Alternative Possibilities*. Aldershot: Ashgate Press.