

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA  
MESTRADO EM FILOSOFIA**

**O PROBLEMA DO LIVRE-ARBÍTRIO E DO DETERMINISMO:  
UMA DEFESA DO INCOMPATIBILISMO**

PEDRO MERLUSSI

**FLORIANÓPOLIS  
2013**



Pedro Merlussi

O PROBLEMA DO LIVRE-ARBÍTRIO E DO DETERMINISMO: UMA  
DEFESA DO INCOMPATIBILISMO

Dissertação submetida ao Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Catarina para a obtenção do grau de Mestre em Filosofia.

Orientador.: Professor Doutor Alexandre Meyer Luz

**Florianópolis**  
**2013**



## RESUMO

Nesta dissertação discuto alguns poucos problemas relacionados ao livre-arbítrio. Entretanto, o problema principal com o qual lido é o problema da compatibilidade entre o determinismo e o livre-arbítrio, o qual chamarei de “o problema do livre-arbítrio e do determinismo”. Ele pode ser formulado intuitivamente da seguinte maneira: será que a verdade do determinismo exclui a existência do livre-arbítrio? O incompatibilista pensa que, se o determinismo for verdadeiro, também será verdadeiro que não há livre-arbítrio. Acredito que o incompatibilista está certo, de modo que avanço um argumento a favor do incompatibilismo, o Argumento da Consequência. Este é o principal argumento a favor do incompatibilismo, mas ele enfrenta objeções importantes. Assim, esta dissertação tem três partes. Na primeira parte, discuto alguns problemas relacionados ao livre-arbítrio, tal como ofereço uma formulação do problema do livre-arbítrio e do determinismo. Depois disso, na segunda parte, formulo o Argumento da Consequência. Finalmente, na terceira parte, lido com suas principais objeções.

## ABSTRACT

In this dissertation I discuss several problems related to free will. However, the main problem that I deal with is the problem of compatibility between determinism and free will, which I will call “the problem of free will and determinism”. It can be formulated intuitively as follows: does the truth of determinism rule out the existence of free will? The incompatibilist thinks that if determinism is true, it is also true that there is no free will. I think the incompatibilist is right, and I put forward an argument for incompatibilism, which is the Consequence Argument. This is the main argument for incompatibilism, but it faces important objections. So this dissertation has three parts. In the first one, I discuss some problems related to free will, and I also offer a formulation of the problem of free will and determinism. After that, in the second part, I formulate the Consequence Argument. Finally, the third part deals with its main objections.



## SUMÁRIO

<b>INTRODUÇÃO.....</b>	<b>11</b>
<b>1 – OS PROBLEMAS DO LIVRE-ARBÍTRIO.....</b>	<b>21</b>
1.1 – OS PROBLEMAS DO LIVRE-ARBÍTRIO.....	21
1.2 – O PROBLEMA DO LIVRE-ARBÍTRIO E DO DETERMINISMO.....	30
1.3 – O DILEMA DE HUME.....	43
<b>2 – O ARGUMENTO DA CONSEQUÊNCIA.....</b>	<b>49</b>
2.1 – O ARGUMENTO DA CONSEQUÊNCIA.....	50
2.2 – AS REGRAS ALFA E BETA.....	61
2.3 – O PRIMEIRO ARGUMENTO FORMAL.....	67
2.4 – SERIA A OBJEÇÃO DE WESTPHAL BEM-SUCEDIDA?.....	72
<b>3- AS OBJEÇÕES AO ARGUMENTO DA CONSEQUÊNCIA.....</b>	<b>76</b>
3.1 – TIPOS DE OBJEÇÕES.....	76
3.2 – A REGRA BETA E AS RÉPLICAS COMPATIBILISTAS.....	77
3.3 – A OBJEÇÃO DE WIDERCKER.....	80
3.4 – A OBJEÇÃO DE MCKAY E JOHNSON.....	81
3.5 - A OBJEÇÃO DE MICHAEL FARA.....	86
3.6 – A OBJEÇÃO DE DAVID LEWIS.....	87
3.7 – COMPATIBILISMO HUMIANO.....	92
3.8 – A OBJEÇÃO DE SCOTT SEHON.....	95
3.9 – A OBJEÇÃO DE TED WARFIELD.....	96



<b>CONSIDERAÇÕES FINAIS.....</b>	<b>98</b>
<b>REFERÊNCIAS.....</b>	<b>99</b>

# INTRODUÇÃO

## O problema tradicional

Antes de tudo, gostaria de apresentar duas passagens relacionadas ao livre-arbítrio. Não são passagens de autores que serão aqui discutidos. Porém, elas resumem claramente as posições que tipicamente se assume acerca do livre-arbítrio nos debates informais:

Quando uma pessoa diz: “Não sou livre, e no entanto levanto e deixo cair minha mão”, todos compreendem que esta resposta ilógica é a prova irrefutável da liberdade. (TOLSTOI 1973, segundo epílogo, Cap. VIII).

Não é o menor dos encantos o fato de que uma teoria possa ser refutável; é precisamente por isso que ela atrai as mentes mais sutis. Parece-me que a teoria mil vezes refutada do "livre arbítrio" deve sua sobrevivência apenas a essa qualidade, posto que sempre vemos surgir alguém disposto a refutá-la ainda (NIETZSCHE 1966, Cap. 1, §18).

Como já tinha adiantado, as citações acima expressam concepções tradicionais sobre as questões suscitadas pelo livre-arbítrio. Nas palavras de Nietzsche, muito embora ele não lide diretamente com o problema, a tese do livre-arbítrio fora amiúde refutada. Na citação de Tolstói, em contrapartida, há precisamente a tese oposta, e algo muitíssimo mais forte: simplesmente levantar meu braço é uma prova irrefutável da minha liberdade. Mas o que há de interessante, em todo o caso, é que ambas as citações chamam a nossa atenção para o que fora tradicionalmente considerado como o problema do livre-arbítrio. O que seria, então, este problema tradicional?

Intuitivamente, pode-se dizer que muitas coisas não dependem de nós. Não depende de nós se os prótons têm de ser milhões de vezes menores que o pingo de tinta neste “i”. E também não depende de nós como o mundo era antes de nosso nascimento. Por outro lado, parece que outras coisas estão sob nosso controle e que, portanto, dependem de nós. Coube a mim, por exemplo, escolher cursar filosofia, assim como despendei algum tempo para escrever esta introdução. A ideia segundo a qual temos controle sobre como agimos parece ser compartilhada por todos nós. À primeira vista, estamos propensos a

pensar que temos livre-arbítrio; parece que realizamos decisões livremente e que essas decisões tipicamente fazem uma diferença no nosso futuro.

Por outro lado, há quem coloque em causa esta ideia intuitiva. O pensador alemão Friedrich Nietzsche, como vimos acima, parece pensar que ela é obviamente falsa. Uma pessoa pode colocá-la em causa por vários motivos. Por exemplo, alguém poderia pensar que a própria noção de livre-arbítrio é de alguma maneira incoerente. Poder-se-ia também defender, baseando-se exclusivamente em crenças religiosas, que nosso arbítrio não é livre. Cientistas poderiam adicionalmente dizer que não temos livre-arbítrio porque somos determinados pelas leis da física. Assim como um copo de vidro estaria determinado a se estilhaçar caso arremessado violentamente ao chão, estaríamos determinados a realizar certas ações.

Em resumo, o debate tradicional gira em torno do problema de saber se temos livre-arbítrio. Pré-teoricamente este parece ser o caso. Em contrapartida, pode-se colocar esta tese em causa, tal como diversos cientistas, filósofos e religiosos repetidamente o fizeram. Este é um dos problemas mais vivos e estimulantes ao longo da história do pensamento humano.

Como sugere Pink (2004, p. 2), pode-se esclarecer a noção de livre-arbítrio pelas duas palavras que formam o seu nome: “liberdade” e “arbítrio”. Quando Aristóteles, em sua obra *Ética a Nicômaco*, discutiu as ações e o nosso controle sobre elas, ele curiosamente não utilizara a palavra grega “ελευθερία”, que pode ser traduzida como “liberdade”, para descrever nosso controle sobre as ações. Ao invés, o filósofo grego utilizara “ἐφ’ ἡμῶν” para se referir à capacidade que temos para controlar nossas ações, o que literalmente significa aquilo que depende de nós ou depende de nós. “ελευθερία”, explica Pink, era um termo usado apenas em discussões políticas para se referir à liberdade política. Foi apenas após Aristóteles que os filósofos gregos começaram a utilizar “liberdade” para se referir à nossa capacidade de controlar ações. É por isso, conclui Pink, que há uma ambiguidade na palavra “liberdade”: ela pode ser usada tanto para nos referirmos à liberdade no sentido político, quanto à liberdade no sentido de que controlamos nossas ações. Porém, é apenas neste último sentido que a palavra “liberdade” é relevante para esta discussão. “Arbítrio”, por outro lado, foi usado de várias maneiras. Mas o sentido mais importante a ser aqui destacado é o poder que temos para tomar decisões. Temos o poder não apenas de realizar ações, mas também de decidir se vamos realizar tais ações. Quando falamos de liberdade do arbítrio, referimo-nos a esse poder psicológico que temos para realizar decisões. Uma vez caracterizada essa noção, surge uma pergunta: teríamos motivos para colocar essa noção em causa?

### **Algumas questões suscitadas pelo livre-arbítrio**

Como afirmei anteriormente, as discussões suscitadas pelo livre-arbítrio afetam, assim como são afetadas, tanto pela filosofia quanto pela religião e a ciência. Pelo lado da religião, temos uma ilustração do problema descrito pelo clássico poema *Paradise Lost* de John Milton. Ali, o autor descreve um vivo debate entre anjos no qual eles discutem como alguns deles poderiam ter pecado em virtude do livre-arbítrio, uma vez que Deus os criara assaz inteligentes. E se eles eram inteligentes, por que teriam pecado? E se fora Deus quem os criara daquela maneira, por que seriam eles próprios os responsáveis pelos seus pecados? Não seria antes Deus? Afinal, Deus – por ser presciente – já sabia de antemão que eles cometeriam pecados. Milton descreveu o caráter enigmático desses problemas como uma questão nada confortável para nós humanos (cf. KANE, 2005, p.1).

Pelo lado da ciência, inúmeros problemas são também suscitados, muitos dos quais parecem colocar em causa que tenhamos livre-arbítrio. Um experimento realizado no centro Bernstein de Neurociência Computacional em Berlim parece colocar em xeque uma concepção pré-teórica que possuímos do conceito de livre-arbítrio. No experimento, uma tela exibia uma sequência aleatória de letras, e voluntários deveriam escolher uma letra e apertar um botão quando ela surgisse. Entretanto, ao monitorar os cérebros dos voluntários via ressonância magnética, os cientistas mostraram que dez segundos antes de os voluntários resolverem apertar o botão, sinais elétricos correspondentes a essa decisão apareciam nos córtices frontopolar e medial, as regiões do cérebro que controlam a tomada de decisões; (acredite ou não, isso acontecia 10 segundos antes mesmo!). A ideia do experimento é a de que, mesmo que sejamos responsáveis por nossas escolhas, elas não são conscientes.

Pelo lado da filosofia, inúmeros argumentos foram apresentados para mostrar que o conceito de livre-arbítrio é de alguma maneira incoerente. Grosso modo, esses argumentos são tipicamente conhecidos como argumentos fatalistas, como o argumento da batalha naval e a falácia fatalista. Eis um exemplo. Considere a seguinte proposição: amanhã haverá uma batalha naval ou não. Ora, uma vez que essa proposição é uma verdade lógica, cuja forma é a do terceiro excluído, pode-se plausivelmente inferir que ela também é uma verdade necessária. Portanto, necessariamente, haverá uma batalha naval ou não. O defensor desse argumento infere daí que necessariamente haverá uma batalha naval ou necessariamente não haverá uma batalha naval. Em outras palavras, quer ocorra uma batalha naval, quer não ocorra, é necessário que esse

seja o caso. E é claro que o caso da batalha naval foi apenas um exemplo aleatório. Necessariamente, eu poderia escolher escrever outra introdução ou não. Seja como for, é necessário que eu tenha escolhido escrever outra introdução, ou é necessário que eu não tenha escolhido, o que significa que não fui realmente livre.

Como veremos mais à frente, penso que há muita confusão nas questões levantadas acima, principalmente com relação ao argumento da batalha naval. No primeiro capítulo procurarei desfazer muitas dessas confusões, assim como mostrar por que o argumento da batalha naval é falacioso. Mas, deixando isso de lado por enquanto, alguém poderia se perguntar: se o livre-arbítrio é tão problemático, por que deveríamos continuar a pensar que temos livre-arbítrio? Não poderíamos antes rejeitar essa ideia obscura que apresenta tantos problemas?

### **Livre-arbítrio e responsabilidade moral**

Muitas pessoas de fato fazem isso. Especialmente motivadas por alguns experimentos científicos, concluem que o livre-arbítrio é uma ilusão, e que a neurociência solucionara este problema definitivamente. Contudo, se isso for verdade, ou seja, se realmente não houver livre-arbítrio, um desastre parece ocorrer; é porque, se não houver livre-arbítrio, não há responsabilidade moral! Em outras palavras, é muito plausível assumir que o livre-arbítrio é uma condição necessária para a responsabilidade moral, o que significa que a seguinte condicional tem um elevado grau de plausibilidade: se somos moralmente responsáveis, então temos livre-arbítrio. Por contraposição, se não temos livre-arbítrio, então não somos moralmente responsáveis. Sider ilustra vividamente o que significaria não ter livre-arbítrio:

Suponha que você tenha sido sequestrado e forçado a cometer uma série de assassinatos terríveis. Os sequestrados fizeram você atirar na primeira vítima ao forçar o seu dedo a pressionar o gatilho de uma arma, também o hipnotizaram de modo a fazer com que você envenenasse uma segunda vítima, e depois jogaram você de um avião, causando a morte de um terceiro. Milagrosamente, você sobrevive à queda do avião. A situação deixa você atordoado, aliviado pela dolorosa experiência ter chegado ao fim. Mas então, para a sua surpresa, você é apreendido pela polícia, que o algema, acusando-o de assassinato. Os pais das vítimas gritam obscenidades a você ao mesmo tempo que a polícia o

leva, humilhado.

Estarão os pais e polícia sendo justos ao culpá-lo pelas mortes? Claro que não, pois você tem uma desculpa irrefutável: você não agiu de acordo com o seu livre-arbítrio. Você não podia evitar o que tinha feito; não podia ter feito de outro modo. E apenas aqueles que agem livremente são moralmente responsáveis (SIDER & CONEE 2005, p.112)

Precisamente por isto rejeitar o livre-arbítrio é tão constrangedor: a renúncia ao livre-arbítrio parece também ser uma renúncia à moralidade, pois apenas aqueles que agem livremente são passíveis de culpa ou de atribuição de mérito. Um assassino não seria moralmente responsável por um crime sádico que cometeu. Aquele aluno que estudou horas e horas para ser bem-sucedido em um teste não teria mérito algum por isso. Em segundo lugar, renunciar nossa liberdade significaria abandonar nossos planos para o futuro. Afinal de contas, por que faríamos planos se não somos livres para mudar o que acontecerá? Sem livre-arbítrio somos como meras máquinas, altamente complexas, porém incapazes de controlar nossos destinos. Por que valeria a pena viver uma vida desse tipo? A rejeição do livre-arbítrio parece também apontar para certa resposta em relação ao problema do sentido da vida: sem livre-arbítrio, nossa vida parece não ter sentido algum.

Em suma, o mais desejável seria solucionar esses quebra-cabeças mencionados acima. A rejeição do livre-arbítrio é *prima facie* demasiado implausível. Infelizmente, há outra tese – argumentavelmente tão plausível quanto a do livre-arbítrio – que intuitivamente ameaça nossa liberdade: o determinismo.

### **Mais problemas para o livre-arbítrio: o determinismo**

O que é o determinismo? Grosso modo, o determinismo é a tese de que o passado mais as leis da natureza determinam, a cada instante, um futuro único. A tese determinista parece violar a ideia segundo a qual temos livre-arbítrio porque afirma que o passado, mais as leis da natureza, determinam, por exemplo, minha escolha de escrever esta introdução. E um determinista radical diria que eu não poderia ter escolhido não escrever esta introdução em virtude

do passado e das leis da natureza. Não poderia escolher não escrever esta introdução do mesmo modo que, dadas as leis da natureza mais o passado, a caneta que acabo de soltar não poderia não cair. A ideia intuitiva do determinismo é mais ou menos a seguinte: nossas escolhas são parte do mundo natural. Portanto, assim como a caneta não pode deixar de cair após soltá-la (em virtude do passado e da lei da gravidade), nossas escolhas são de tal modo que não poderiam ser diferentes (em virtude do passado e das leis da natureza). Portanto, o determinismo *parece* negar a proposição *poderíamos escolher agir de outro modo*, e assim parece negar que temos livre-arbítrio.

Mas seria o determinismo realmente plausível? Há quem pense que sim. Ted Sider, por exemplo, endossa essa tese:

Nossa crença no determinismo é razoável porque vimos a ciência ser bem-sucedida, várias e várias vezes, em sua busca pela causa subjacente das coisas. As inovações tecnológicas devem sua existência à ciência: arranha-céus, vacinas, naves espaciais, a Internet. A ciência parece explicar tudo aquilo que observamos: a mudança das estações, o movimento dos planetas, o funcionamento interno de plantas e animais. Dado este registro de êxitos, esperamos razoavelmente que a marcha do progresso científico continue; esperamos que a ciência eventualmente irá descobrir as causas de tudo (SIDER & CONEE 2005, p.114).

Quando dizemos que estamos determinados a escolher tal e tal, o que queremos dizer é que não poderíamos não escolher tal e tal. O que o determinismo parece mais intuitivamente implicar é que não temos cursos alternativos de escolha ou de ação. Ou, dito de outra maneira, o determinismo parece implicar que não poderíamos escolher ou agir diferentemente daquilo que escolhemos ou agimos. Numa palavra: a tese determinista viola o princípio de possibilidades alternativas. Segundo este princípio, ao menos segundo certa formulação, “se uma pessoa não poderia ter decidido de outro modo, então essa pessoa não é responsável por aquilo que ela faz” (BROOK & STANTON, 2002, p.137). Tendo em conta o princípio de possibilidades alternativas, e se o determinismo acarretar que não poderíamos decidir agir de outro modo, seguir-se-ia que não somos responsáveis pelas decisões que fazemos ou pelo que vamos fazer. Não poderíamos repreender um assassino porque ele não poderia ter decidido de outro modo. Não poderíamos recompensar uma pessoa caso realizasse uma ação virtuosa porque ela não poderia ter decidido (ou agido) de

outro modo. Em resumo, o determinismo levanta os mesmos problemas que aqueles quebra-cabeças mencionados anteriormente.

O que temos de notar, contudo, é que há aqui um problema distinto. O primeiro problema mencionado, que chamei de o Problema Tradicional, é precisamente o de saber se temos livre-arbítrio. O determinismo levanta outro problema. Aqui, queremos saber quais são as relações conceituais entre o determinismo e o livre-arbítrio. Queremos saber se a tese determinista implica a negação da tese do livre-arbítrio. Este é o que irei chamar de o Problema da Compatibilidade. É o problema de saber quais as relações conceituais entre o determinismo e o livre-arbítrio.

### **O Problema da Compatibilidade**

Resumidamente, o problema da compatibilidade surge de certa tensão existente entre duas de nossas crenças mais razoáveis: a crença de que temos livre-arbítrio e a crença de que o determinismo é verdadeiro. Como vimos, o determinismo é a tese de que o passado mais as leis da natureza determinam um futuro único. Por exemplo, se o determinismo for verdadeiro, dado o passado mais as leis da natureza, a caneta que acabei de soltar está determinada a cair. Quanto ao livre-arbítrio, é difícil dizer algo consensual, mas comumente se aceita que uma condição necessária para ter livre-arbítrio é poder se decidir de outro modo. Por exemplo, se o leitor tem livre-arbítrio, parece que poderia ter decidido não ler esta introdução. E, como vimos, diríamos que à primeira vista o leitor tem livre-arbítrio, assim como qualquer pessoa. Por outro lado, parece que também temos razões para pensar que o determinismo é verdadeiro. Vimos as razões apresentadas por Sider, muito embora eu não concorde com elas, pelas razões que apresentarei no primeiro capítulo<sup>1</sup>. Além disso, poder-se-ia dizer que, intuitivamente, as leis da natureza e o passado determinam que as coisas aconteçam apenas desse modo. E se assim o for, aparentemente se segue que nossas decisões são também determinadas, e que por isso não poderiam ser diferentes. Assim voltamos à tensão inicial: as crenças de que temos livre-arbítrio e de que o determinismo é verdadeiro são plausíveis, mas parecem incompatíveis. Mas será que de fato o são? Alguns pensam que sim, outros pensam que não. Os primeiros, conhecidos como incompatibilistas, têm a seu favor a intuição de que essas crenças não são compatíveis. Já os últimos, conhecidos como compatibilistas, defendem a conclusão mais desejável; tanto

---

<sup>1</sup> Sider confunde o determinismo com o Princípio de Causação Universal. Como procurarei mostrar no capítulo seguinte, as teses não são equivalentes.



melhor se conseguirmos compatibilizar as duas crenças que, consideradas separadamente, são muito plausíveis. Motivações para ambas as posições, bons argumentos para cada lado e muita controvérsia. Eis aqui um dos principais problemas da filosofia.

Quem propôs a divisão do problema do livre-arbítrio entre o Problema da Compatibilidade e o Problema Tradicional foi Peter van Inwagen. O primeiro é este que foi exposto até aqui. Ou seja, é o problema de saber se o livre-arbítrio e o determinismo são compatíveis. Para colocar em termos mais precisos, trata-se do problema de saber se a condicional “se o determinismo é verdadeiro, então não temos livre-arbítrio” é verdadeira. Enquanto o incompatibilista sustenta que essa condicional é verdadeira, o compatibilista defenderá que é falsa, ou ao menos implausível. Em contrapartida, o Problema Tradicional não é o de saber se aquela condicional é verdadeira, mas o de saber se a afirmação “temos livre-arbítrio” o é. Uma das maneiras de lidar com o problema tradicional, adicionalmente, é a de saber se a tese determinista é verdadeira. Não sigo esta perspectiva, contudo, pois parece-me que este último problema é empírico. Muito terei a dizer sobre o determinismo no último capítulo, em todo o caso, de modo que pouparemos esse assunto por ora.

Peter van Inwagen argumentou que a existência do Problema Tradicional depende de uma solução correta para o Problema da Compatibilidade, de modo que este último é, na verdade, mais importante que o primeiro. Como ele afirma:

Se o determinismo e o livre-arbítrio são compatíveis e, *a fortiori*, se do livre-arbítrio se *deriva* o determinismo, então não há o problema tradicional, não mais do que o problema de saber como minhas frases podem ser compostas simultaneamente de palavras em português e de letras em romano. (VAN INWAGEN 1983: 2)

Embora a razão apresentada por van Inwagen seja um tanto obscura, penso que a conclusão que ele procura nos persuadir a aceitar seja verdadeira (embora eu a sustente por outras premissas, claro). Suponhamos que se consiga provar que o determinismo seja verdadeiro. Disso não se segue que não temos livre-arbítrio. Para se seguir, teríamos de mostrar que o determinismo e o livre-arbítrio são incompatíveis. Note que um dos argumentos clássicos mais famosos a favor da tese de que não temos livre-arbítrio pressupõe o incompatibilismo sem oferecer razões. Trata-se do dilema de Hume, como veremos no próximo capítulo. Uma das premissas desse argumento é a

seguinte: se o determinismo for verdadeiro, então não temos livre-arbítrio. Ora, essa é a proposição que está em jogo no Problema da Compatibilidade. Para o dilema de Hume ser um argumento cogente, temos de saber (ou pelo menos crer justificadamente) nessa proposição. O mais interessante, desse modo, seria solucionar o Problema da Compatibilidade. Suponha, em contrapartida, que se consiga provar que temos livre-arbítrio. Disso também não se segue que a tese determinista seja falsa. Para se seguir, novamente, seria preciso sustentar o incompatibilismo. Portanto, em suma, uma resposta ao Problema Tradicional dependerá de uma resposta ao Problema da Compatibilidade.

Ademais, é também argumentável que o Problema Tradicional, entendido sob certa visão, não é um problema filosófico. Como afirmei acima, se pressupusermos que o problema tradicional é o de saber se o determinismo é verdadeiro, teremos de lidar com o seguinte quebra-cabeça: o determinismo é uma questão empírica; como podemos, mediante mera análise conceitual, solucionar um problema de natureza empírica? Uma concepção meta-filosófica aparentemente razoável é a de que os problemas filosóficos são *a priori* por natureza. O Problema da Compatibilidade se enquadra nessa concepção. O mesmo já não é tão claro com relação ao Problema Tradicional.

Nesta dissertação, apresento uma solução para o Problema da Compatibilidade. Não é uma solução original, no entanto. O que irei fazer é apresentar e defender um argumento popular a favor do incompatibilismo, a tese de acordo com a qual o determinismo e o livre-arbítrio são incompatíveis. Vários argumentos a favor do incompatibilismo foram apresentados ao longo da história da filosofia, como o Argumento do Jardim das Passagens Forquilhadas, o Argumento da Cadeia Causal e os argumentos do desígnio e da manipulação (cf. VIHVELIN 2011). No entanto, penso que o único argumento cogente a favor do incompatibilismo seja este:

Se o determinismo for verdadeiro, então nossas ações são a consequência das leis da natureza e dos eventos no passado remoto. Mas não depende de nós o que se passou antes de nascermos, e nem depende de nós o que as leis da natureza são. Portanto, as consequências dessas coisas (incluindo nossas ações presentes) não dependem de nós (VAN INWAGEN, 1983, p.56).

Este é o famoso Argumento da Consequência. Após formular o Problema da Compatibilidade, no primeiro capítulo, dedico um capítulo inteiro para oferecer uma formulação mais precisa desse argumento. É claro que,

como qualquer argumento bem-conhecido em filosofia, o Argumento da Consequência enfrenta sérias objeções. O último capítulo lida justamente com elas.

### **Visão geral**

O problema com o qual irei lidar principalmente nesta dissertação é o Problema da Compatibilidade. Como procurei mostrar, esse problema é distinto daquele de saber se temos livre-arbítrio, ou o de saber se o determinismo é verdadeiro. Isso não quer dizer que irei negligenciar o Problema Tradicional ao longo de toda esta dissertação; na verdade, dediquei algumas páginas do primeiro capítulo para lidar com algumas questões suscitadas pelo problema tradicional. No entanto, o tópico principal é efetivamente o Problema da Compatibilidade. Como eu disse, meu objetivo principal é o de defender o Argumento da Consequência, que é um argumento apresentado por Peter van Inwagen. Esta dissertação foi, então, dividida em três partes principais. Na primeira delas, procuro desfazer diversas confusões relacionadas ao livre-arbítrio e formular o Problema da Compatibilidade com toda a precisão que sou capaz de oferecer. Depois disso, na parte seguinte, formulo o argumento de van Inwagen, esclarecendo conceitos como o de proposição, leis da natureza, etc. A terceira parte lida com as principais objeções ao argumento.

A partir do primeiro capítulo, lançarei mão de certas ferramentas da lógica proposicional modal. Assumirei, de agora em diante, os axiomas de S5, presumivelmente o melhor candidato para a lógica da necessidade metafísica.

## 1. OS PROBLEMAS DO LIVRE-ARBÍTRIO

Nesta dissertação, apresento razões para convencer o leitor de que o Argumento da Consequência é um bom argumento a favor do incompatibilismo. Mas a formulação dessa tese envolve termos técnicos, como “incompatibilismo”, “compatibilismo” e “Argumento da Consequência”. Por exemplo, para formular a tese incompatibilista, é preciso formular o problema da compatibilidade entre o determinismo e o livre-arbítrio. Obviamente, os conceitos de determinismo e livre-arbítrio não são precisos e merecem algum esclarecimento. O objetivo deste capítulo é o de esclarecer esses conceitos, formular o problema da incompatibilidade entre o determinismo e o livre-arbítrio, as teses em disputa e, finalmente, indicar qual tese pretendo defender. Como já adiantei, o determinismo e o livre-arbítrio me parecem incompatíveis e o argumento por mim defendido a favor dessa incompatibilidade é o Argumento da Consequência. Esse argumento não foi primeiramente apresentado por mim e, portanto, não é original. Porém, seguindo a máxima de que é preferível estar certo a ser original, será aqui apresentado para defender a tese que acredito ser verdadeira. De qualquer forma, muitas respostas às principais objeções a este argumento são originais.

Na primeira seção deste capítulo, argumento que a expressão “o problema do livre-arbítrio” é enganadora e deve ser, tanto quanto possível, evitada. Como procurarei mostrar, há vários problemas relacionados à liberdade do arbítrio. A segunda seção consiste numa formulação do Problema da Compatibilidade e foi dividida em duas partes. Na primeira, defino o conceito de determinismo. Na segunda, esclareço o conceito de livre-arbítrio. E é só depois disso que formulo o problema com o qual me ocuparei com maior intensidade nesta dissertação, o problema da compatibilidade entre o determinismo e o livre-arbítrio. Finalmente, na última parte deste capítulo, apresento razões para mostrar que um famoso argumento a favor do impossibilismo – a tese de acordo com a qual o livre-arbítrio é impossível – é falacioso.

### 1.1 Os problemas do livre-arbítrio

Em seu romance *Guerra e Paz*, Tolstoi nos disse que o assim chamado “problema do livre-arbítrio” tem ocupado as melhores inteligências humanas desde há muito tempo. De fato, a discussão filosófica sobre o livre-arbítrio fez

parte da produção intelectual de Guilherme de Ockham (*Predestination, Foreknowledge, and Future Contingents*), Tomás de Aquino (*Summa Theologica*, Ia, q. 14, art 13), David Hume (*A Treatise of Human Nature* 2.3.1-2, *Enquiry concerning Human Understanding* sec. 8), Thomas Hobbes<sup>2</sup>, Immanuel Kant (*Crítica da Razão Pura*, terceira antinomia), entre muitos outros. Não menos certo é que este problema até hoje não foi consensualmente solucionado. Muitos filósofos estão propensos a dizer, em virtude desses resultados, que o comumente denominado “problema do livre-arbítrio” é um mistério (VAN INWAGEN 2000 e MCGINN 1993). Alguns, com inspiração fortemente kantiana, talvez queiram inferir daí que ele sequer pode ser solucionado. A despeito de tudo isso, a filosofia fez progressos notáveis em relação a uma melhor compreensão do que se tem chamado, ao longo da história, de “o problema do livre-arbítrio”.

No parágrafo acima, como se deve ter notado, não me referi ao problema do livre-arbítrio, mas apenas ao que se tem chamado “o problema do livre-arbítrio”. Espero deixar suficientemente claro por que fiz isto. Suponha que o problema do livre-arbítrio seja o problema de saber se temos livre-arbítrio. A afirmação “o problema do livre-arbítrio é o de saber se temos livre-arbítrio”, segundo a teoria das descrições definidas de Bertrand Russell (1905), equivale logicamente à afirmação de que “há um, e apenas um, problema do livre-arbítrio que é o de saber se temos livre-arbítrio”. Ora, o que ocorre é que não há propriamente *um único* problema do livre-arbítrio, mas, ao invés, vários problemas relacionados ao livre-arbítrio. Considere, por exemplo, o problema de saber se o livre-arbítrio é compatível com o determinismo. Este problema é independente daquele de saber se temos livre-arbítrio. Digamos que se solucione o primeiro problema, e que sua solução seja um argumento cogente contra a existência do livre-arbítrio. Isto, por si só, não solucionará o problema de saber se a existência do livre-arbítrio é compatível com o determinismo. Mesmo que nosso arbítrio não seja livre, permanece em aberto o problema de saber se as teses de que há livre-arbítrio e de que o determinismo é verdadeiro poderiam ser compatíveis. Assim, a proposição expressa pela frase “há um único problema sobre o livre-arbítrio” é falsa. Por isso seria mais adequado utilizar outra expressão: “os problemas do livre-arbítrio”, que é justamente o título desta seção.

---

<sup>2</sup> A posição de Hobbes acerca do livre-arbítrio está disponível em tradução para a língua portuguesa aqui:

<http://www.revistafundamento.ufop.br/Volumel/n2/vol1n2-2.pdf>.

A despeito do que afirmei acima, não quero com isso dar a entender que grande parte dos filósofos, como Hobbes, Ayer, entre muitíssimos outros, diziam que “há *apenas um* problema do livre-arbítrio”. Eu nada disse sobre atos de asserção e não estou cometendo o mesmo engano, creio, que Strawson cometeu com Russell em uma de suas famosas objeções à teoria das descrições (STRAWSON 1950, p.330). Embora a frase “o problema do livre-arbítrio é tal e tal” implique logicamente a frase de que “há apenas um problema do livre-arbítrio”, disso não se segue que quem profere a primeira também profere a segunda. Em outras palavras, aquilo que é logicamente implicado por uma frase não tem de ser asserido por quem a profere. Por exemplo, suponha que eu diga que “não há o problema do livre-arbítrio”. Embora essa frase implique logicamente que “não há o problema do livre-arbítrio ou a lua é feita de queijo”, não estarei asserindo que não há o problema do livre-arbítrio ou a lua é feita de queijo. Portanto, é bem plausível assumir que a tradição filosófica não tenha dito a falsidade *óbvia* de acordo com a qual há apenas um problema do livre-arbítrio.

Claro, a tradição filosófica não disse essa falsidade *óbvia*, mas disse, mesmo assim, uma falsidade: não há *o* problema do livre-arbítrio. Por isso devo evitar, de agora em diante, usar a expressão “*o* problema do livre-arbítrio”. Permita-me apresentar um rápido argumento a favor disso. Se a proposição expressa pela frase “há *o* problema do livre-arbítrio” é falsa, então não é conhecida (Princípio da Factividade do Conhecimento). Se não é conhecida, não deve ser asserida (Tese da Asserção de Conhecimento<sup>3</sup>). Logo, se a proposição expressa pela frase “há o problema do livre-arbítrio” é falsa, então não deve ser asserida.

Voltemos, no entanto, aos *problemas* do livre-arbítrio. Há outras questões antigas, e igualmente honráveis, relacionadas à liberdade do arbítrio. Considere, por exemplo, o problema da presciência divina. A concepção tradicional de Deus teísta assume que Ela instancia<sup>4</sup> as seguintes propriedades essenciais<sup>5</sup>: suma bondade, onipotência, auto-existência, distância e

---

<sup>3</sup> Esta tese é quase que ortodoxa na literatura corrente; é defendida – embora em versões sutilmente diferentes – por Williamson (2000) e De Rose (2002).

<sup>4</sup> De agora em diante, vou utilizar a expressão “instanciar uma propriedade”, ao invés de “ter uma propriedade”, pois o verbo “instanciar” é neutro em relação a certa resposta ao Problema dos Universais.

<sup>5</sup> A ortodoxia sobre o conceito de propriedade essencial afirma que algo é uma propriedade essencial de um particular sse é instanciada por ele em todos os mundos nos quais esse particular existe. Por exemplo a propriedade de ser um ser humano é essencial a Sócrates porque ela a instancia em todos os mundos possíveis em que existe.

independência do mundo e, a mais importante para os nossos propósitos, a onisciência. À primeira vista, se Deus é onisciente, então Ela tudo sabe, incluindo todo o futuro, o que presumivelmente inclui nossas ações. Intuitivamente, as seguintes teses parecem incompatíveis: (1) Deus é onisciente e (2) há livre-arbítrio. Mas será que (1) e (2) são realmente incompatíveis? Eis um breve argumento para pensar que sim (WESTPHAL 2011, p.246):

Deus criou o mundo em seis dias e, à sexta-feira, já sabia de tudo. É um desafio compreender claramente a expressão “sabia de tudo”. O conhecimento é comumente distinguido entre três tipos: proposicional (*e.g.*, saber que  $7 + 5 = 12$ ), prático (*e.g.*, saber andar de bicicleta) e por contato (*e.g.*, conhecer Florianópolis visitando-a). Admitamos, para fins de argumentação, que o conhecimento proposicional faça parte da extensão do conceito de onisciência. Algum esclarecimento adicional será oferecido mais à frente ao conceito de proposição, mas, para já, estipularei apenas que proposições são portadores primários de valor-de-verdade. Desse modo, se Deus é onisciente, Ela conhecia, na sexta-feira, todas as proposições verdadeiras, incluindo  $F$ , uma proposição verdadeira sobre o que o Bruno iria fazer na segunda-feira (digamos, levantar sua mão). Disso podemos inferir que  $BgF$ , em que  $BgF$  é a proposição de indexação temporal de que na sexta-feira Deus acreditara que  $F$ . Suponha, contudo, que o Bruno tenha livre-arbítrio e que poderia levantar sua mão, ou seja, poderia tornar  $F$  falsa. Desse modo, Deus acreditaria numa falsidade ( $BgF \ \& \ \sim F$ ). Portanto, se o Bruno tem livre-arbítrio, então Deus não é onisciente; afinal, a tese de que Deus é infalível ( $BgP \supset P$ ) seria falsa<sup>6</sup>. Ora, mas se Deus é onisciente, e acredita que o Bruno levantará sua mão na segunda-feira, então o Bruno não é capaz de agir de modo a tornar  $F$  falsa. E assim enfrentamos um dilema: ou o Bruno tem livre-arbítrio e pode agir de modo a tornar  $F$  falsa, caso em que Deus não é onisciente, ou Deus é onisciente, caso em que o Bruno não tem livre-arbítrio<sup>7</sup>.

O problema é que a apresentação acima é um tanto informal. Não é de todo claro qual a forma do argumento que parece expressar o paradoxo acima.

O termo “mundo possível” é um termo técnico, mas oferecerei uma breve caracterização desse problema mais à frente.


<sup>6</sup> Estou assumindo que a tese da onisciência acarreta a tese da infalibilidade divina.

<sup>7</sup> Este dilema, não obstante apresentado informalmente e de modo intuitivo, recebeu inúmeras respostas ao longo da história da filosofia. Linda Zagzebski (2011) elenca as respostas principais ao problema, como as respostas Aristotélica, Agostiniana, Ockhamista, Boeciana e Molinista.

Uma maneira seria a seguinte (assumindo que “ $\Box$ ” representa a necessidade lógica, WESTPHAL 2011, p.247):

1.  $\Box (BgP \supset P)$                       premissa, infalibilidade divina
2.  $BgF$                                       premissa
3.  $\Box F$                                       de 1 e 2.

Ora, 3 seria a conclusão de que, necessariamente, o Bruno levantou a mão na segunda-feira, o que é logicamente equivalente à proposição de que o Bruno não poderia não levantar a mão na segunda-feira. Assim, o raciocínio acima presumivelmente expressaria o paradoxo informalmente apresentado anteriormente. Entretanto, o problema óbvio desta formulação é o de que a inferência de 3, por meio de 1 e 2, é inválida. Para mostrar isso, considere um sistema de árvores de lógica modal proposicional que seja uma extensão da lógica proposicional clássica (considerando que uma teoria lógica L2 é uma extensão de uma teoria lógica L1 sse os teoremas de L1 são também teoremas de L2). O sistema resultante permitirá mostrar que a inferência acima não é lícita:

- |   |                           |       |            |
|---|---------------------------|-------|------------|
| 1.  | 0: $\Box (BgP \supset P)$ |       |            |
| 2.  | 0: $BgP$                  |       |            |
| 3.  | 0: $\sim \Box P$          |       |            |
| 4.  | 0: $\Diamond \sim P$      | (3)   |            |
| 5.  | 1: $0 \rightsquigarrow 1$ | (4)   |            |
| 6.  | 1: $\sim P$               | (4,5) |            |
| 7.  | 1: $BgP \supset P$        | (1,5) |            |
|  |                           |       |            |
| 8.  | 1: $\sim BgP(7)$          | 9.    | 1: $P$ (7) |
- x

Como se pode observar, a inferência é inválida, pois um dos ramos da árvore não fecha, a saber, 8. Para a inferência ser lícita nessa lógica, teríamos



de ter  $\Box BgP$  como premissa. Entretanto, parece-me difícil encontrar boas razões a seu favor.

Obviamente não é meu intuito oferecer uma solução para o problema da compatibilidade entre a presciência divina e o livre-arbítrio. Apenas mostrei que certa formulação de um tradicional argumento a favor da incompatibilidade é inválido. A propósito, trata-se de uma falácia análoga à famosa falácia fatalista (que mencionarei brevemente no capítulo seguinte). Em todo o caso, a pergunta ainda permanece: será a onisciência divina compatível com o livre-arbítrio? Alguns dizem que sim, outros dizem que não. Independentemente disso, este problema está relacionado ao livre-arbítrio. E o problema da presciência divina apenas é um problema à medida em que parece colocar em causa que agentes como nós tenham livre-arbítrio.

Outra questão igualmente honrável é o que irei chamar de “o problema do fatalismo lógico”. O fatalismo lógico é a tese de que é uma verdade lógica ou conceitual que ninguém é capaz de agir de modo diferente daquilo que agiu (van Inwagen 1983, cap. 2). Isso é controverso, mas podemos admitir, para fins de formulação do problema, que uma condição necessária para haver livre-arbítrio é ser capaz de agir de modo diferente daquilo que se agiu. Um argumento tradicional famoso a favor do fatalismo lógico, por exemplo, é o argumento da batalha naval. Se a tese do fatalismo lógico for verdadeira, então não somos capazes de agir de modo diferente daquilo que agimos. E se não somos capazes de agir de modo diferente daquilo que agimos, então não temos livre-arbítrio. Logo, se o fatalismo lógico for verdadeiro, então não temos livre-arbítrio.

Quando mencionei o problema da onisciência, mostrei por que uma formulação de um argumento tradicional a favor de uma resposta a esse problema é inválida. Aqui, como disse, o argumento tradicional a favor do fatalismo lógico é o da batalha naval. Espero mostrar que uma das formulações desse argumento é inválida. Vejamos:

1. Haverá uma batalha naval ou não.
2. Necessariamente, haverá uma batalha naval ou não.
3. Logo, ou necessariamente haverá uma batalha naval, ou necessariamente não haverá uma batalha naval.

Intuitivamente, a conclusão, como podemos observar, é a de que, ocorra ou não uma batalha naval, é necessário que esse seja o caso. Seja  $P$  a

proposição expressa pela frase “haverá uma batalha naval” e admitindo que “ $\square$ ” representa a necessidade lógica, podemos formalizar o argumento acima da seguinte maneira:

- |                                    |                 |
|------------------------------------|-----------------|
| 1. $P \vee \sim P$                 | premissa        |
| 2. $\square (P \vee \sim P)$       | 1, IK $\square$ |
| 3. $\square P \vee \square \sim P$ | 2               |

Uma vez que “ $\square$ ” representa a necessidade lógica, a conclusão é a de que é logicamente necessário que haverá uma batalha naval ou logicamente necessário que não haverá uma batalha naval.  $P$  é uma variável proposicional que pode ser substituída por qualquer proposição, inclusive qualquer proposição acerca de uma ação. Assim, seja  $P$  a proposição de que o Bruno levantará a mão, a conclusão é a de que é logicamente necessário que o Bruno levantará a mão ou logicamente necessário que não levantará a mão.

1. É logicamente necessário que: se é logicamente necessário que o Bruno levantará a mão, então ele não é capaz de não levantar a mão.

2. É logicamente necessário que ele levantará a mão.

3. É logicamente necessário que ele não é capaz de não levantar a mão.

O problema, claro, é que a primeira inferência é inválida, o que é fácil de ver recorrendo novamente a uma árvore lógica:

- |     |   |        |
|-----|---|--------|
| 1.  | 0: $\square (P \vee \sim P)$              |        |
| 2.  | 0: $\sim (\square P \vee \square \sim P)$ |        |
| 3.  | 0: $\sim \square P$                       | (2)    |
| 4.  | 0: $\sim \square \sim P$                  | (2)    |
| 5.  | 0: $\diamond \sim P$                      | (3)    |
| 6.  | 0: $\diamond P$                           | (4)    |
| 7.  | 0: $0 \rightsquigarrow 1$                 | (5)    |
| 8.  | 1: $\sim P$                               | (5, 7) |
| 9.  | 0: $0 \rightsquigarrow 2$                 | (6)    |
| 10. | 2: $P$                                    | (6, 9) |
| 11. | 1: $P \vee \sim P$                        | (1, 7) |





Novamente, não pretendo oferecer uma solução para o problema do fatalismo lógico, mas apenas mencionar que uma formulação desse clássico argumento é inválida, quer numa lógica modal que seja uma extensão da lógica clássica, quer na lógica intuicionista. (Para um interessante ensaio sobre o fatalismo, ver VAN INWAGEN 1983, cap. 2; para outras referências ao problema, veja o já clássico TAYLOR 1962).

Poderia prolongar a apresentação de problemas relacionados ao livre-arbítrio, mas esses acima mencionados bastam para o que eu tinha afirmado nos parágrafos iniciais desta seção. Não há somente um problema relacionado ao livre-arbítrio, mas vários. Apresentei os problemas da presciência divina e do fatalismo lógico de maneira breve porque esses problemas não serão discutidos nesta dissertação.

Finalmente, é importante notar que, de agora em diante, discutirei alguns problemas *filosóficos* relacionados ao conceito de livre-arbítrio. Não vou discutir os problemas *científicos* a ele relacionados. Há muitas notícias, principalmente entre os jornais, de que os cientistas estariam próximos de resolver o problema do livre-arbítrio. A pergunta apropriada, neste caso, seria a seguinte: qual problema eles irão resolver? Muito provavelmente não é o problema de saber se o livre-arbítrio é compatível com o determinismo. Este é um problema de saber quais as relações conceituais entre a tese determinista e a tese de que há livre-arbítrio, e não é um problema que compete aos cientistas resolverem. É claro que, para apresentar um argumento cogente a favor disso, eu teria de prolongar com discussões meta-filosóficas a respeito da natureza de um problema filosófico. Não quero me envolver nessa discussão. O que preciso é apenas o seguinte: se alguém pensa que uma ciência empírica pode resolver um problema de análise conceitual, essa pessoa tem o ônus de apresentar um argumento cogente a favor dessa posição. Se *prima facie* não há tal argumento, então *prima facie* o problema da compatibilidade entre o determinismo e o livre-arbítrio é um problema que compete apenas aos filósofos solucionarem. Portanto, *prima facie* há alguns problemas relacionados ao livre-arbítrio que não competem aos cientistas solucionarem. Assim, em suma, é importante distinguir os problemas filosóficos dos problemas científicos relacionados ao livre-arbítrio. Considerando nossa compreensão pré-teórica desses problemas, proponho-me aqui apenas a discutir os problemas filosóficos relacionados ao livre-arbítrio.

## 1.2 O problema do livre-arbítrio e do determinismo

O problema com o qual me ocuparei com maior interesse é o problema do livre-arbítrio e do determinismo, ou, para abreviar, o Problema da Compatibilidade. Grosso modo, este é o problema de saber se o determinismo e o livre-arbítrio são compatíveis. Chamemos de “Compatibilismo” à tese de que o livre-arbítrio e o determinismo são compatíveis; ou, em outras palavras, à tese segundo a qual a verdade do determinismo não exclui a existência do livre-arbítrio. Chamemos de “Incompatibilismo” à tese de que o livre-arbítrio e o determinismo são incompatíveis; ou, em outras palavras, à tese segundo a qual a verdade do determinismo exclui a existência do livre-arbítrio. Os proponentes da primeira tese são conhecidos como compatibilistas, enquanto que os proponentes da segunda tese são conhecidos como incompatibilistas. Em resumo, para colocar o problema em outras palavras (VIHVELIN 2011): o problema filosófico do livre-arbítrio e do determinismo é o de saber quem está certo, o compatibilista ou o incompatibilista. Sem dúvida, esta é uma caracterização demasiado informal do problema. Peço ao leitor paciência. Oferecerei uma caracterização mais precisa nesta seção. Mas, para que isso seja feito, preciso oferecer alguns breves comentários sobre os conceitos de determinismo e livre-arbítrio, respectivamente.

### **a) O determinismo**

O primeiro aspecto a ser notado é o de que o termo “determinismo” é ambíguo. Não sei o que esse termo significa em outras disciplinas, como a física ou a psicologia. O que posso dizer é que, em filosofia, ele é usado para se referir a teses diferentes. O determinismo teológico é a tese de que Deus é onisciente; o que, como vimos, quer dizer que conhece todas as proposições verdadeiras, incluindo proposições verdadeiras sobre nossas ações futuras. O problema do determinismo teológico e do livre-arbítrio é o problema da onisciência divina caracterizado na seção anterior. O determinismo lógico é a tese de que o princípio da bivalência (a saber, o princípio de que, para qualquer proposição, essa proposição é verdadeira ou falsa) vale para todas as proposições, incluindo proposições sobre nossas ações futuras. O problema do determinismo lógico e do livre-arbítrio é o de saber se ambos são compatíveis. O determinismo psicológico é a tese de que as ações de uma pessoa são determinadas por motivos mais fortes. Infelizmente não sei qual é o problema filosófico relacionado ao determinismo psicológico, muito provavelmente porque não sei o que significa “determinado por um motivo mais forte”. Seja como for, a concepção de determinismo usada ao longo de toda esta dissertação deve ser distinguida dessas concepções de determinismo.

Chamemos de “determinismo nomológico” à tese determinista que será discutida de agora em diante. Intuitivamente, é a tese de que o passado e as leis da natureza *determinam* um único futuro. Em outras palavras, é a tese de que, dado o passado, há exatamente um futuro possível. O termo “possível” talvez seja mais ambíguo do que “determinismo” e por isso merece alguma explicação. O sentido empregado de “possibilidade” nessa caracterização é o de possibilidade nomológica:

$\Phi$  é nomologicamente possível com relação às leis da natureza<sup>8</sup>  $L$  se, e só se, é logicamente consistente com  $L$ .

Notemos que há *diversos futuros possíveis* em outros sentidos de possibilidade. Vejamos, por exemplo, a possibilidade epistêmica:

$\Phi$  é epistemicamente possível para um sujeito  $S$  se, e só se,  $\Phi$  não é excluído pelo que  $S$  sabe (VAIDYA 2007).

Considerando tudo o que sei, é possível – no sentido de ser epistemicamente possível – que ocorra um eclipse solar esta tarde, assim como é possível que não ocorra. Isso também é possível com relação à possibilidade conceitual, pois

$\Phi$  é conceitualmente possível se, e só se,  $\Phi$  não é excluído pelo conjunto de verdades conceituais (VAIDYA 2007).

Mas note que, do fato de haver muitos futuros possíveis (nos sentidos epistêmico ou conceitual de possibilidade), isso não constitui uma refutação à tese determinista. O que a tese determinista quer dizer é que, dado o passado, há um único futuro *nomologicamente* possível. E um futuro nomologicamente possível é um futuro logicamente consistente com as leis da natureza. A tese determinista é falsa se, e só se, dado o passado, há mais de um futuro nomologicamente possível. O determinismo é falso se, e só se, o indeterminismo é verdadeiro.

Essa caracterização tem de ser precisada. Para um futuro ser logicamente consistente com as leis da natureza, ele tem de ser entendido como uma proposição, não como um evento, um estado de coisas ou qualquer coisa

---

<sup>8</sup> Não farei distinção entre os termos “leis físicas” e “leis da natureza”.

que não seja um portador-de-verdade. É argumentável que seja um erro categorial dizer que um estado de coisas  $E$  é logicamente consistente a um estado de coisas  $F^9$ . No entanto, não é um erro categorial dizer que uma proposição  $P$  é logicamente consistente com uma proposição  $Q$ . Como se pode observar, uma vez que a tese determinista precisa da noção de que o futuro seja logicamente consistente com as leis da natureza, estipularei que, ontologicamente falando, leis da natureza são proposições.

Além disso, na caracterização acima do termo “determinismo”, o termo “passado” foi invocado. Ora, o que significa a expressão “dado o passado, há exatamente um futuro possível”? Disse que a possibilidade é para ser entendida como a possibilidade nomológica e que o futuro tem de ser de alguma forma entendido como uma proposição. O mesmo ocorre com o passado. Nessa caracterização de determinismo, o passado será entendido como uma proposição verdadeira sobre o estado total do mundo em qualquer tempo  $t$ ; o futuro será entendido como uma proposição verdadeira sobre o mundo em qualquer tempo posterior a  $t$ . Em virtude dessas considerações, o determinismo pode agora ser definido:

Determinismo = *def.* uma proposição verdadeira sobre o estado total do mundo em qualquer tempo  $t$ , e uma proposição que seja a conjunção de todas as leis da natureza, implicam estritamente uma proposição verdadeira sobre estado do mundo em qualquer tempo posterior a  $t$ .

Há conceitos no *definiens* que não foram adequadamente explicados; por exemplo, os conceitos de leis da natureza, proposição e estado total do mundo. Isto, contudo, não será necessário para este capítulo. Irei oferecer uma caracterização mais precisa desses conceitos quando for apresentar o Argumento da Consequência, no capítulo seguinte.

Antes de prosseguir para a noção de livre-arbítrio, será importante distinguir a concepção de determinismo que mencionei acima de concepções diferentes assumidas por outros filósofos. Em *The Open Universe: an argument for indeterminism*, Karl Popper discute a tese que chamara de “determinismo científico”, a saber, a tese de que

---

<sup>9</sup> É argumentável porque podemos definir “consistência lógica” em termos da noção de verdade; por exemplo,  $A$  é logicamente consistente com  $B$  sse é possível que  $A$  e  $B$  sejam verdadeiras. É comum considerar que estados de coisas não são portadores-de-verdade.

a estrutura do mundo é tal que *qualquer evento pode ser racionalmente previsto, com qualquer grau desejado de previsão, se nos é dada uma descrição suficientemente precisa dos eventos passados, junto com todas as leis da natureza* (POPPER, 1988, pp.1-2).

Embora a definição do conceito de determinismo de Popper recorra aos conceitos de eventos passados e leis da natureza, ela se difere de uma maneira substancial daquela que apresentei acima: na concepção de Popper, o determinismo é uma tese sobre *previsibilidade*. Eu nada disse sobre previsibilidade na definição que apresentei. Se há uma noção legítima de “determinismo” que pode ser entendida como uma tese sobre previsibilidade, é algo que permanece controverso. Kadri Vihvelin (2011, §1), por exemplo, afirma que “o determinismo não é uma tese sobre previsibilidade”. E ainda asseve que “a Teoria do Caos nos diz que alguns sistemas determinísticos são muito difíceis de prever. A mecânica quântica nos diz, pelo menos de acordo com algumas interpretações, que o comportamento de sistemas probabilísticos é, pelo menos em algum aspecto, fácil de prever” (2011, §1). Não vou argumentar que o conceito determinismo que estou a discutir é o correto; apenas estipulo que o que chamo de “determinismo” pode ser legitimamente assim chamado. De fato, a discussão sobre o Argumento da Consequência se pauta nessa concepção de determinismo, e acredito que isso seja o bastante para me justificar a assumir essa concepção. E nessa concepção, é claro, o determinismo não é uma tese sobre previsibilidade<sup>10</sup>.

Em todo o caso, tanto eu quanto Popper concordamos que o determinismo não deve ser confundido com uma tese sobre causação. De fato, a bibliografia antiga sobre o Problema da Compatibilidade (representada pelos Grandes Filósofos Mortos, como Espinosa, Hobbes, Hume, Kant, Mill) assumia que o determinismo era equivalente ao Princípio de Causação Universal:

Todo evento (mudança, acontecimento, ou qualquer *relata* causal) tem uma causa.

---

<sup>10</sup> A concepção assumida nesta dissertação está mais próxima do que Popper chamou de “determinismo metafísico”, a qual não é uma tese sobre previsibilidade.



Como já distingui o determinismo de uma tese sobre previsibilidade, será importante distingui-lo dessa tese sobre causalção.

Em primeiro lugar, para mostrar que o determinismo e o Princípio de Causação Universal não são teses equivalentes, suponha que a doutrina da causalção *imane*nte (ou a doutrina da causalção do *agente*, como dizem alguns) seja verdadeira. A doutrina da causalção imane<sup>11</sup>nte é muito bem ilustrada na seguinte passagem de Chisholm (1964, p.62):

Quando agimos, cada um de nós é um primeiro motor imóvel. Ao fazer o que fazemos, causamos certos eventos que acontecem, e nada – ou ninguém – nos causa a causar esses eventos que acontecem

Suponhamos, então, que alguma mudança tenha ocorrido num agente e esse mesmo agente tenha sido o responsável por essa mudança. Admitamos que a mudança nesse caso não seja uma mera mudança de Cambridge<sup>12</sup>. Segundo a doutrina da causalção imane<sup>11</sup>nte, o agente é o único responsável por determinar sua mudança, e esta não é determinada pelas leis da natureza e o passado. A ideia intuitiva de que as leis da natureza e o passado não determinam a mudança do agente, segundo a definição de determinismo apresentada acima, é equivalente à afirmação de que o passado e as leis da natureza não implicam estritamente uma proposição verdadeira acerca de um instante do tempo posterior ao do passado, digamos, uma proposição acerca da mudança desse agente. Portanto, se a doutrina da causalção imane<sup>11</sup>nte é verdadeira, o determinismo é falso. No entanto, a doutrina da causalção

---

<sup>11</sup> Será informativo distinguir a causalidade imane<sup>11</sup>nte da transeunte. Nas palavras de Zimmerman (1997, p.433): “A causalção imane<sup>11</sup>nte foi tipicamente contrastada com a ‘causalção transeunte’ do seguinte modo. Na causalidade transeunte, diz Johnson ‘a ocorrência da causa e a ocorrência do efeito referem-se a contínuos diferentes, enquanto que na causalção imane<sup>11</sup>nte a ocorrência da causa e a ocorrência do efeito são atribuídas ao mesmo contínuo’.

<sup>12</sup> O termo “mudança de Cambridge” foi utilizado por Peter Geach (1969) para se referir aos filósofos de Cambridge, como Russell, os quais pensavam que um particular sofre mudanças sse ganha ou perde propriedades. Ora, como notou Geach, essa concepção de mudança é implausível. Admita que eu tenha dado um passo agora mesmo. Digamos que ganhei, assim, a propriedade de estar 2cm mais longe de um átomo de carbono numa galáxia muito, muito distante. Não parece plausível admitir que eu tenha mudado.

imane é totalmente compatível com Princípio de Causação Universal. O que o princípio afirma é que, se algo é um evento, então tem uma causa. A mudança que ocorreu no agente é um evento que tem uma causa, a saber, o próprio agente. Já o agente propriamente dito não é causado por nada, mas isso não falsifica o Princípio de Causação Universal, pois o agente não é um evento; como o antecedente é falso, o Princípio de Causação Universal é verdadeiro.

Enfim, para evitarmos esse tipo de objeção à equivalência entre o determinismo e o Princípio de Causação Universal, precisaríamos da seguinte premissa (VAN INWAGEN 1983, p.4):

(1) Se um evento tem uma causa, então sua causa tem de ser um evento e nunca uma substância ou contínuo, tal como uma pessoa.

Mesmo que 1 seja verdadeira, precisaríamos de premissas adicionais para mostrar que as teses são equivalentes. Considere o seguinte: “suponha que todo o evento seja causado por algum evento anterior ou eventos anteriores e que essas causas não *necessitam*, mas meramente *produzem*, seus efeitos” (van Inwagen 1983: 4). Essa suposição é compatível com o Princípio de Causação Universal. No entanto, se ela for verdadeira, o passado não *determina* um único futuro. Se o passado não determina um futuro único, então o determinismo é falso. Assim, seria possível que o Princípio de Causação Universal fosse verdadeiro e o determinismo fosse falso. Portanto, precisaríamos da seguinte premissa adicional (VAN INWAGEN 1983, p.4):

(2) Se um evento A for a causa de um evento B, então segue-se que, dado que A aconteceu e dado as leis da natureza, A “causalmente necessitou” B, que B não poderia ter falhado em acontecer.

Finalmente, mesmo que (1) e (2) sejam verdadeiras, para mostrar que o Princípio de Causação Universal acarreta o determinismo, precisaríamos da seguinte premissa (van Inwagen 1983: 4):

(3) Toda cadeia de causas que não tem um primeiro membro é tal que, para todo tempo  $t$ , algum evento na cadeia acontece num momento anterior a  $t$ .

O determinismo pode ser falso, se (3) for falsa, mesmo que o Princípio de Causação Universal e (1) e (2) sejam verdadeiras. Como não consegui ser mais claro, citarei novamente van Inwagen (1983: 5): “pode haver dois

instantes de tempo,  $t_1$  e  $t_2$ , tal que (i) um certo evento  $A$  acontece em  $t_2$  (ii)  $A$  é o último membro de uma cadeia infinita de causas, e (iii) todo membro desta cadeia ocorre depois de  $t_1$ .” Isto foi notado por Lukasiewicz (1967) em “On Determinism”. Um exemplo dado por van Inwagen talvez torne o ponto mais claro (nota 7 do primeiro capítulo): digamos que  $t_1$  seja meia-noite e  $t_2$  1:00h da manhã; suponha que a causa de  $A$  ocorreu 12:30h, a causa da causa 12:15h, a causa da causa da causa 12:07:30, e assim *ad infinitum*.

Em suma, para que o Princípio de Causação Universal acarrete o determinismo, precisamos das seguintes premissas:

(1) Se um evento tem uma causa, então sua causa tem de ser um evento e nunca uma substância ou contínuo, tal como uma pessoa.

(2) Se um evento  $A$  for a causa de um evento  $B$ , então segue-se que, dado que  $A$  aconteceu e dado as leis da natureza,  $A$  “causalmente necessitou”  $B$ , que  $B$  não poderia ter falhado em acontecer.

(3) Toda cadeia de causas que não tem um primeiro membro é tal que, para todo tempo  $t$ , algum evento na cadeia acontece num momento anterior a  $t$ .

Como não vou discutir a plausibilidade dessas premissas, não vou assumir que o Princípio de Causação Universal acarreta o determinismo. Agora note que, mesmo que essas três premissas fossem verdadeiras, mostraríamos *apenas* que o Princípio de Causação Universal acarreta o determinismo. Não mostraríamos o acarretamento converso. Ou seja, não mostraríamos que o determinismo acarreta o Princípio de Causação Universal. Seria interessante ver com cuidado se há este acarretamento. Contudo, se isso é assim, vai depender de qual é a teoria correta da causação (VIHVELIN, 2011, §1); essa discussão sobre causação, porém, será evitada tanto quanto possível, como veremos. Por ora, é o bastante para o determinismo. Procurarei doravante oferecer alguns esclarecimentos sobre o conceito de livre-arbítrio e, depois, caracterizar com mais precisão as teses em disputa acerca do Problema da Compatibilidade.

## **b) O livre-arbítrio**

Há pouquíssimo consenso sobre o que seja o livre-arbítrio. O que há de mais consensual diz respeito a uma compreensão intuitiva desta noção, a saber, que algumas coisas estão sob nosso controle, enquanto outras não (PINK, 2007). Por exemplo, ter nascido como membro da espécie *homo*

*sapiens*, brasileiro, no estado de São Paulo, etc., é algo que não estava sob o meu controle. O que se passou há muito tempo atrás, em galáxias muito, muito distantes, também é algo que intuitivamente está fora do meu controle. Em contrapartida, tipicamente se assume que minhas ações presentes e futuras estão sob meu controle. Ter livre-arbítrio, intuitivamente, é a ideia de que temos controle, pelo menos em alguns casos, acerca de como agimos. Assim, pré-teoricamente, o livre-arbítrio tem a ver com o nosso poder para realizar certas ações. Agora, se quisermos definir “livre-arbítrio”, será importante oferecer alguns esclarecimentos sobre “capacidade” e “poder”. Começarei por “capacidade”.

À primeira vista, capacidades são parecidas com disposições. Tipicamente pensamos que as disposições são aquelas propriedades atribuídas por predicados como “é frágil” e “é solúvel” (CARNAP 1936, GOODMAN 1954). Por exemplo, uma xícara de porcelana tem a propriedade disposicional de se quebrar, enquanto o sal tem a propriedade disposicional de se dissolver em água. Capacidades são semelhantes a disposições no sentido de que ambas podem não se manifestar. Uma xícara de porcelana tem a propriedade disposicional de se quebrar mesmo que não se quebre, um bom jogador de futebol tem a capacidade de fazer um belo gol mesmo que não manifeste essa capacidade.

Embora semelhantes no aspecto apresentado acima, capacidades e disposições se distinguem em outro: é uma condição necessária para a capacidade, mas não para a disposição, ser uma propriedade de agentes. Isso significa que a seguinte condicional é verdadeira para “capacidade”: se algo é uma capacidade, então é uma propriedade de um agente. Se substituirmos o termo “capacidade” por “disposição”, veremos que a condicional não será verdadeira para este último: a propriedade de ser solúvel não é instanciada por agentes, caso em que a antecedente é verdadeira (a propriedade de ser solúvel é disposicional) e a consequente, falsa (essa propriedade não é instanciada por quaisquer agentes).

Nesse sentido, “capacidade” é mais parecido com “poder”. Tanto o poder quanto a capacidade podem não se manifestar. Tanto o poder quanto a capacidade são ambas propriedades de um agente. Eles só diferem num aspecto: as capacidades estão relacionadas a ações no sentido de que um agente é capaz de realizar algo se, e só se, ele é capaz de realizar uma ação. Contudo, é conceitualmente possível que um agente tenha o poder de fazer algo que não seja uma ação. Nesse sentido, eu tenho o poder, mas não a capacidade, de entender diversas frases da língua portuguesa (desde que os textos não sejam traduções de Deleuze e Heidegger). Meu poder para fazer algo será uma

capacidade apenas na medida em que envolver uma relação a uma ação (MAIER 2010 §1.2, VAN INWAGEN 1983, 8-13).

Além disso, será importante notar que ser capaz de realizar uma ação não é equivalente a ser fisicamente possível realizar essa ação. Suponha que eu esteja dentro de um quarto e que sua única saída, digamos uma porta, esteja trancada e que a chave para abri-la se encontre nas mãos de uma pessoa que esteja fora do quarto. É fisicamente possível que eu saia do quarto, já que é totalmente consistente com as leis da natureza que a porta seja aberta. No entanto, intuitivamente não sou capaz de sair do quarto, pois não está em meu poder realizar esta ação.

Resumindo: nesta taxonomia, o conceito de capacidade aplica-se a ações realizadas por agentes: sou capaz, por exemplo, de realizar as ações de levantar minha mão, arremessar uma pedra e escrever esta dissertação (ou assim espero!). E ser capaz de realizar uma ação não pode ser entendido como ser fisicamente possível realizar essa ação. Acredito que seja tudo o que eu possa falar, de maneira mais ou menos incontroversa, sobre “capacidade”. Infelizmente não sei definir esse termo. O que eu tinha pensado inicialmente era entender “capacidade” como um operador de formação de frases, que operasse sobre afirmações acerca de ações para gerar afirmações mais complexas. Parecia-me plausível admitir que um sistema correto para esse operador seria o sistema de lógica modal alética T (CHELLAS 1980), que conteria, entre outros, os seguintes esquemas de axiomas:

$$(T\Diamond): \vdash A \rightarrow \Diamond A$$

$$(K\Diamond): \vdash \Diamond (A \vee B) \rightarrow (\Diamond A \vee \Diamond B)$$

Mas penso que Anthony Kenny (1975) tenha argumentado persuasivamente a favor de que  $(T\Diamond)$  e  $(K\Diamond)$  não devem ser esquemas de axiomas se entendermos o diamante como o operador de capacidade. Apresento dois contraexemplos inspirados no artigo original de Kenny.

Contraexemplo a  $(T\Diamond)$ : Suponha que eu nunca tenha manejado uma arma de fogo em toda a minha vida (o que é bem verdade) e, por alguma coincidência, fui convidado a acertar um alvo; digamos, uma pequena marca numa lata. Por coincidência, acerto o alvo bem na mosca. No entanto, embora eu tenha realizado essa ação, não se segue que eu seja capaz de realizá-la; tive apenas sorte.

Contraexemplo a (K $\diamond$ ): Esse contraexemplo é completamente inspirado no artigo de McKay & Johnson (1996) do qual terei muito a falar no terceiro capítulo. Suponha que eu tenha uma moeda de dois lados, em condições normais de uso. Sou capaz de realizar a seguinte ação: fazê-la virar cara ou coroa. Mas não sou capaz de fazê-la virar cara, já que isso se deve ao acaso, nem sou capaz de fazê-la virar coroa, pelo mesmo motivo anterior.

Esses resultados não parecem decisivos. Embora uma lógica modal para a ação e a capacidade tenha seus inconvenientes (por exemplo, foi considerada “muito abstrata” pela literatura), há desenvolvimentos interessantes. É o caso, por exemplo, do sistema de Hilbert de ação e capacidade desenvolvido por Dag Elgesem (1997). E não obstante tenha sido mostrado que esse sistema é incompleto com relação à semântica pretendida, conforme se vê no artigo de Governatori & Rotolo (2005), os próprios autores sugerem que a completude pode ser recuperada. Em todo o caso, não entrarei nessa discussão, pois, para a defesa do Argumento da Consequência que pretendo desenvolver, as explicações que ofereci sobre “capacidade” são suficientes<sup>13</sup>.

Agora algum esclarecimento sobre “poder”. Como já disse acima, a diferença principal entre “poder” e “capacidade” é que este último refere-se apenas a ações, enquanto o primeiro não. Assim como “capacidade”, “poder” tem de ser cuidadosamente distinguido de “possibilidade física”. O exemplo que apresentei anteriormente para distinguir “capacidade” de “possibilidade física” pode ser igualmente usado aqui. Além disso, “poder” não significa “permissibilidade moral”. No sentido que irei aqui empregar, um criminoso tem o poder de torturar crianças apenas por diversão, embora isso não seja moralmente permissível.

Essa breve caracterização dos termos “poder” e “capacidade” é útil para traçar a clássica distinção entre liberdade de escolha e liberdade de ação. Intuitivamente, a distinção é a de que posso não ser livre, por exemplo, para sair do meu quarto (não tenho liberdade de ação), embora seja livre para escolher ficar no quarto (liberdade de escolha). Agora, na presente taxonomia, a distinção será apresentada da seguinte maneira. Se “livre-arbítrio” for entendido no sentido de liberdade de ação, isso significa que o termo “capacidade” estará presente no *definiens*, pois “capacidade” é um termo que se aplica a ações realizadas por agentes. Por outro lado, se “livre-arbítrio” for

---

<sup>13</sup> Não me esqueci da famosa análise condicional do conceito de capacidade. Ela será apresentada e refutada, pelo menos em sua versão clássica, no último capítulo desta dissertação.

entendido no sentido de liberdade de escolha, isso significa que “poder” estará no *definiens*, já que esse termo não se aplica apenas a ações; aplica-se também a escolhas, que são estados mentais.

Em todo o caso, o “livre-arbítrio” será entendido em termos de “poder”, já que “capacidade” foi caracterizado como “poder” mais “ação”. Tal como não sou capaz de definir “capacidade”, não sou capaz de definir “poder”. Portanto, não sou capaz de definir “livre-arbítrio”. Na seção seguinte, apesar de tudo, ofereci uma caracterização mais precisa da tese do livre-arbítrio; acredito que qualquer definição bem-sucedida de “livre-arbítrio” terá de acomodar essa caracterização. Assim, minha incapacidade de defini-lo não será um problema para a defesa do Argumento da Consequência que pretendo desenvolver.

### **c) As teses do livre-arbítrio e do determinismo e o Problema da Compatibilidade**

A tese do livre-arbítrio é a de que pelo menos um agente (ou, se achar melhor, uma criatura não-divina) tem livre-arbítrio. Neste sentido, *o livre-arbítrio é uma tese sobre agentes*, enquanto que *o determinismo é uma tese sobre proposições*. O Problema da Compatibilidade é o problema de saber quais são as relações conceituais entre essas duas teses. E eis que surge uma pergunta óbvia: como vamos investigar as relações conceituais entre uma tese sobre agentes e outra sobre proposições? Peter van Inwagen (1983, p.66) oferece uma solução bastante plausível: o livre-arbítrio terá de ser entendido como uma tese sobre agentes e proposições.

Assim, Peter van Inwagen desenvolve, como ele mesmo diz, “uma maneira de descrever nossos poderes para agir – e, ao agir, modificar o mundo – como poderes sobre valores-de-verdade de proposições” (*ibid*). A ideia intuitiva disso é muito simples. Tome-se as proposições expressas pelas seguintes frases verdadeiras:

- (a) Nenhum objeto viaja mais depressa que a luz
- (b)  $7 + 5 = 12$
- (c) A xícara de café em cima da minha mesa nunca se quebrou
- (d) Ninguém leu a *Crítica da Razão Pura* inteira em voz alta

A diferença que há entre essas proposições, como afirma van Inwagen (*ibid*), é que elas poderiam ser descritas de várias maneiras. Por exemplo,

ninguém tem, nem nunca teve, o poder de mudar o valor de verdade das proposições expressas pelas frases (a) e (b), mas presumivelmente há agentes que têm o poder de mudar o valor de verdade das proposições expressas pelas frases (c) e (d). Por exemplo, presumivelmente posso agir e modificar o mundo de modo a quebrar a xícara de café que está em cima da minha mesa. Usei nesse caso a expressão “presumivelmente” porque o problema de saber se há realmente agentes que têm o poder de mudar o valor de verdade das proposições expressas pelas frases (c) e (d) é algo que depende da tese do livre-arbítrio ser verdadeira. Assim, se houver uma distinção entre (a) e (b) e (c) e (d), há livre-arbítrio; do contrário, não (ou seja, a distinção será ilusória).

Para traduzir a tese do livre-arbítrio para a linguagem de uma lógica modal, é preciso apresentar o operador modal “N”, que é intuitivamente caracterizado do seguinte modo:

“NP” abrevia “*P* e ninguém tem, nem nunca teve, qualquer escolha sobre se *P*”.

A expressão “ter escolha sobre se *P*” pode ser entendida como “ter o poder de mudar o valor-de-verdade de *P*” ou, simplesmente, “ter o poder de tornar *P* falsa”. Seja *P* uma variável proposicional que pode ser substituída por qualquer proposição, a tese do livre-arbítrio pode ser formulada da seguinte maneira:

Tese do livre-arbítrio = def.  $\sim NP$

Ou seja, coloquialmente, é a tese de que não é o caso que: *P* e ninguém tem, nem nunca teve, qualquer escolha sobre se *P*. Essa é uma tese mínima sobre o livre-arbítrio: é a tese de que pelo menos um agente teve o poder de tornar ao menos uma proposição falsa; essa proposição pode ser tanto acerca das ações desse agente quanto acerca de suas escolhas.

Agora será possível apresentar o Problema da Compatibilidade de maneira mais precisa. Para fazer isso, será preciso apenas traduzir a definição do determinismo para a mesma linguagem. Seja “L” a abreviação de uma frase que expressa uma proposição que é a conjunção de todas as leis da natureza e “P0” a abreviação de uma frase que expressa uma proposição verdadeira sobre o estado total do mundo em algum tempo, num passado distante, antes de quaisquer agentes existirem, o determinismo é a seguinte tese:

Determinismo:  $\Box ((P_0 \ \& \ L) \supset P)$ , em que “ $\Box$ ” é a necessidade alética, “ $\supset$ ” a



condicional material e “P” uma variável proposicional que pode ser substituída por qualquer proposição acerca de uma ação.

Finalmente, podemos definir tanto o incompatibilismo quanto o compatibilismo. O incompatibilismo é a tese de que o determinismo implica que não há livre-arbítrio. O compatibilismo é a negação do incompatibilismo.

Incompatibilismo:  $\square ((P_0 \ \& \ L) \supset P) \supset NP$

Compatibilismo:  $\sim (\square ((P_0 \ \& \ L) \supset P) \supset NP)$

O Problema da Compatibilidade é o problema de saber qual tese é verdadeira, a saber, o compatibilismo ou o incompatibilismo.

Considerando as definições acima, é fácil definir as teses em disputa do Problema Tradicional, ou seja, o problema de saber se há livre-arbítrio, ou de saber se o determinismo é verdadeiro. O determinista é um incompatibilista que aceita a antecedente da condicional. O libertista é um incompatibilista que pensa que  $NP$  é falsa e, portanto, que o determinismo é falso.

Essa é a maneira que vejo o problema. Ela pode ser colocada em disputa. Uma das objeções que é feita por Warfield (2000) ao Argumento da Consequência, por exemplo, pressupõe que a formulação que fiz do problema é malsucedida. Irei responder essa objeção no último capítulo desta dissertação. Por ora, acredito que a formulação que fiz é pelo menos intuitivamente plausível.

Os capítulos seguintes desta dissertação se ocuparão com o Problema da Compatibilidade, mas não com o Problema Tradicional, tal com caracterizado na introdução. Uma das resposta ao Problema Tradicional é conhecida como impossibilismo, a saber, a tese de que o livre-arbítrio é impossível. Na última seção deste capítulo, como eu já havia adiantado, procuro refutar um famoso argumento a favor do impossibilismo: o Dilema de Hume.

### 1.3 O Dilema de Hume

O Dilema de Hume é um argumento aparentemente muito persuasivo a favor da tese de acordo com a qual a tese do livre-arbítrio é, de alguma maneira, incoerente. Embora o argumento tenha recebido este nome, o filósofo

David Hume, tanto quanto sei, não o defendeu. Ademais, o argumento já havia sido apresentado anteriormente por outros filósofos, como Leibniz e Tomás de Aquino. Seja como for, a discussão em torno de saber quem defendeu esse argumento é filosoficamente irrelevante, embora talvez historicamente interessante. A tese de que o livre-arbítrio é, de alguma maneira, incoerente, é imprecisa. Essa tese informal pode ser formulada em termos mais precisos da seguinte maneira: a tese do livre-arbítrio é uma impossibilidade conceitual ou metafísica. Chamemos essa tese de Impossibilismo.

É sempre importante lembrar, entretanto, que “possível” é um termo vago. Tal como formulado, há dois sentidos de possibilidade do Impossibilismo que irei aqui discutir. Portanto, para já, sinto-me compelido a definir as possibilidades e necessidades metafísica e conceitual. A possibilidade conceitual já fora definida na seção 1.2. Mas, apenas para não esquecermos, apresento a definição novamente:

$\Phi$  é conceitualmente possível se, e só se,  $\Phi$  não é excluída pelo conjunto de verdades conceituais

$\Phi$  é conceitualmente necessária se, e só se,  $\neg\Phi$  é excluída pelo conjunto de todas as verdades conceituais (VAIDYA 2007).

Até aqui, quando defini os tipos de modalidades, recorri às definições de Anand Vaidya. Como agora preciso definir a modalidade metafísica, senti-me num primeiro momento inclinado a oferecer da mesma maneira as definições de Vaidya, que são as seguintes:

$\Phi$  é metafisicamente possível se, e só se,  $\Phi$  é verdadeira em algum mundo metafisicamente possível.

$\Phi$  é metafisicamente necessário se, e só se,  $\Phi$  é verdadeira em todos os mundos metafisicamente possíveis (VAIDYA 2007).

Contudo, essas definições me parecem malsucedidas, pois não há consenso quanto a uma definição bem-sucedida de “mundo metafisicamente possível”. Penso que uma definição melhor da possibilidade e necessidade metafísicas seja esta:

$\Phi$  é metafisicamente possível se, e só se,  $\Phi$  é possível.

$\Phi$  é metafisicamente necessário se, e só se,  $\Phi$  é necessário.

Poder-se-ia objetar que esta definição é pouco informativa. Bem, entenderei nesta dissertação os conceitos de possibilidade e necessidade como primitivos, de modo que não podem ser adequadamente analisados; por exemplo, se a análise tripartite de conhecimento fosse bem-sucedida, o conceito de conhecimento poderia ser eliminado pelos conceitos de crença, verdade e justificação. Não penso que o mesmo possa ser feito com relação aos conceitos de possibilidade e necessidade. Mas é claro que essa tese tem de ser sustentada com argumentos. Infelizmente não posso apresentá-los, pois isso envolveria uma longa discussão em metafísica modal que desviaria os propósitos desta dissertação. Apenas penso ser mais plausível entender a modalidade metafísica como a modalidade *simpliciter*; metafisicamente possível é aquilo que é possível, metafisicamente necessário é aquilo que é necessário. E os conceitos de necessidade e possibilidade, por sua vez, são interdefiníveis: possível é aquilo cuja negação não é necessária, necessário é aquilo cuja negação não é possível.

Uma vez caracterizado o Impossibilismo, o Dilema de Hume pode ser apresentado em sua forma intuitiva deste modo:

(Dilema de Hume): Ou o determinismo é verdadeiro, ou não. Se for verdadeiro, então não há livre-arbítrio. Se não for verdadeiro, então as nossas ações são aleatórias e, portanto, não há livre-arbítrio. Seja como for, portanto, não há livre-arbítrio.

Considerando as definições da tese determinista e do livre-arbítrio, podemos formular o Dilema de Hume da seguinte maneira:

1.  $\Box ((P_0 \ \& \ L) \supset P) \vee \sim \Box ((P_0 \ \& \ L) \supset P)$
2.  $\Box ((P_0 \ \& \ L) \supset P) \supset \text{NP}$
3.  $\sim \Box ((P_0 \ \& \ L) \supset P) \supset \text{NP}$

## 4. NP

A conclusão, 4, segue-se das premissas 1, 2 e 3, pela regra de eliminação da disjunção da lógica clássica. O argumento, portanto, é válido. Teremos agora de ver se ele é um *bom* argumento. Um argumento bom não pode ser simplesmente um argumento sólido, ou seja, um argumento que, além de válido, tem premissas verdadeiras. Considere, por exemplo, o seguinte argumento:

1. Se  $1 + 1 = 2$ , então  $1 + 1 = 2$
2.  $1 + 1 = 2$
3. Logo,  $1 + 1 = 2$

Pela definição de “solidez” dada acima, o argumento supracitado é sólido. Porém, não é bom. Trata-se da famosa falácia *petitio principii*, pois a conclusão do argumento já está contida na premissa. Usarei o termo “cogência” para me referir aos argumentos bons. Um argumento cogente é, intuitivamente, um argumento sólido e racionalmente persuasivo (ou seja, se o argumento é sólido e o agente reconhece sua solidez, ele está racionalmente obrigado a aceitar sua conclusão). Não é preciso definir “racionalidade epistêmica” para entendermos essa noção de cogência: qualquer definição desse termo terá de acomodar a intuição de que um agente não está racionalmente obrigado a aceitar um argumento que seja falacioso, como a petição de princípio. Se o argumento acima, apesar de sólido, não é racionalmente persuasivo, então não é cogente. O que procurarei mostrar é que o Dilema de Hume não é um argumento cogente exatamente nessa caracterização.

Não penso que primeira premissa apresente quaisquer problemas para o argumento. Ela me parece perfeitamente plausível porque é uma instância do princípio do terceiro excluído da lógica clássica; rejeitar essa premissa para refutar o argumento é um estratégia muito geral não deveria ser tomada. A rejeição dessa premissa nos leva à rejeição da maneira como tipicamente argumentamos. Uma vez que não estou disposto a rejeitar isso, não estou disposto a rejeitar a primeira premissa.

A segunda premissa já é bem problemática. Talvez haja a intuição inicial de que, se o determinismo é verdadeiro, então não há livre-arbítrio. Eu

realmente não sei se esta intuição esteja de acordo com o senso comum. Porém, seja como for, a segunda premissa nada mais é do que a tese *incompatibilista*, tal como definido na seção anterior. Embora eu pense que essa premissa seja verdadeira, como procurarei argumentar a favor dela no capítulo seguinte, ela não pode ser meramente pressuposta, sem quaisquer argumentos adicionais. Diversos filósofos pensam que esta premissa é falsa; e, de fato, por um bom tempo foi assumida como obviamente falsa. São os compatibilistas, obviamente, que têm razões independentes para rejeitar a segunda premissa. Uma vez que não há consenso acerca do Problema da Compatibilidade, e ninguém sabe quem está certo, se é o compatibilista ou incompatibilista, o compatibilista poderá muito bem rejeitar racionalmente a conclusão do Dilema de Hume. Ora, dada a definição de argumento cogente, é fácil ver que, por assumir sem argumentos adicionais a premissa 2, o dilema de Hume não é um argumento cogente.

É interessante notar que o Dilema de Hume é um argumento comumente apresentado nas discussões sobre o Problema Tradicional, tal como caracterizado na introdução desta dissertação. No entanto, sua segunda premissa depende de uma resposta correta para o Problema da Compatibilidade. Nesse aspecto, o Problema da Compatibilidade é mais fundamental do que o Problema Tradicional, pois, na maior parte dos casos, uma solução correta para este último depende de uma solução correta para o primeiro. Já adiantei, em todo o caso, que defenderei o incompatibilismo e, portanto, eu teria de aceitar a segunda premissa do dilema. Se aceito que o argumento é válido, bem como as premissas 1 e 2, e mesmo assim penso que o argumento não é bom, tem de haver algo de errado com a terceira premissa.

Na apresentação intuitiva do Dilema de Hume, há um entimema a favor da terceira premissa, a saber, este: “se não for verdadeiro, então as nossas ações são aleatórias e, portanto, não há livre-arbítrio”. É preciso tornar esse argumento explícito, pois a terceira premissa é, à primeira vista, implausível.

Por contraposição, a terceira premissa é a tese de que o livre-arbítrio implica o determinismo. Um argumento a favor dessa tese pode ser apresentado brevemente da seguinte maneira. Suponha que um agente, digamos o Rafael, tenha arremessado uma pedra na janela de uma casa, mas que tenhamos descoberto que não havia *qualquer tipo de causa* para sua ação de arremessar aquela pedra. Se não houve qualquer tipo de causa para ação do Rafael, então teríamos de dizer que sua ação fora meramente casual, arbitrária, que nada tinha a ver efetivamente com o Rafael. Ora, é nesse sentido que o livre-arbítrio implica o determinismo, pois uma condição necessária para o

Rafael ter realmente arremessado aquela pedra é que sua ação tenha uma causa, ou seja, que ele próprio tenha causado o arremesso da pedra.

Este argumento é mau: o que ele mostra, na melhor das hipóteses, é que o livre-arbítrio *implica* o Princípio de Causação Universal. No entanto, como mostrei anteriormente, o Princípio de Causação Universal não é equivalente à tese determinista. O argumento, mesmo que seja sólido, é falacioso, pois é uma *ignoratio elenchi*; afinal, sua conclusão é irrelevante para mostrar a verdade da terceira premissa. Uma vez que não há boas razões para aceitar a verdade da terceira premissa do Dilema de Hume, não há razões para considerar o dilema como um argumento cogente.

E por falar em *ignoratio elenchi*, note que não só o argumento a favor da terceira premissa é uma falácia desse tipo, mas o próprio Dilema de Hume. Este dilema fora formulado como um argumento a favor da tese de acordo com a qual o livre-arbítrio é impossível (seja no sentido de possibilidade conceitual, seja no sentido de possibilidade metafísica). Ora, a conclusão do argumento é apenas a de que não há livre-arbítrio. Se não houver problemas em falar de mundos possíveis, a conclusão do argumento é apenas a de que não há livre-arbítrio no mundo atual; o argumento não estabelece a conclusão de que, em todos os mundos metafísica ou conceitualmente possíveis, não há livre-arbítrio. Portanto, mesmo que as três premissas do argumento sejam verdadeiras, o Dilema de Hume é uma *ignoratio elenchi* e, portanto, não é um argumento cogente a favor do impossibilismo.

Assim termino minha exposição inicial. Defini as teses do livre-arbítrio e do determinismo e procurei desfazer algumas confusões comuns que há nas discussões sobre os problemas do livre-arbítrio. Acredito que isso permitiu mostrar que um argumento clássico, e aparentemente muito persuasivo, a favor da tese de que a tese do livre-arbítrio é incoerente não é bom. No capítulo seguinte, não irei mais discutir o Problema Tradicional e passarei a discutir diretamente o Problema da Compatibilidade, ao defender o Argumento da Consequência.

## 2. O ARGUMENTO DA CONSEQUÊNCIA

O Problema da Compatibilidade é um dos problemas que mais tem recebido atenção nos debates atuais sobre as questões relacionadas ao livre-arbítrio. E isso se deve principalmente à discussão em torno do mais importante argumento a favor do incompatibilismo, o Argumento da Consequência. Desde sua principal formulação, a qual se deve a Peter van Inwagen (1974), o problema fora tratado por autores como Michael Slote (1982), David Lewis (1981), McKay & Johnson (1996), Crisp & Warfield (2000), Helen Beebe & Alfred Mele (2003), Michael Fara (2008), Scott Sehon (2011), Jonathan Westphal (2012), entre muitos outros. Como procurei mostrar na introdução, uma resposta apropriada para o Problema Tradicional depende de uma resposta apropriada para o Problema da Compatibilidade. Formulei este último problema no primeiro capítulo e já antecipei que defendo o incompatibilismo. E se defendo o incompatibilismo, o mais natural a se esperar é um argumento a favor dessa posição. O argumento que acredito ser cogente a favor do incompatibilismo é, como seria natural esperar, o Argumento da Consequência. Eis o argumento em sua forma intuitiva:

Se o determinismo for verdadeiro, então nossas ações são a consequência das leis da natureza e dos eventos no passado remoto. Mas não depende de nós o que se passou antes de nascermos, e nem depende de nós o que as leis da natureza são. Portanto, as consequências dessas coisas (incluindo nossas ações presentes) não dependem de nós (van Inwagen, 1983: 56).

Entretanto, tal como apresentado acima, o argumento não está adequadamente formulado. Como qualquer formulação intuitiva, esta apresenta o problema de ser consideravelmente imprecisa. Neste capítulo apresentarei duas formulações mais precisas desse argumento intuitivo, as quais na verdade são formulações de Peter van Inwagen, em seu livro *An Essay on Free Will* (1983). Gostaria de já adiantar, contudo, que defenderei formulações diferentes daquelas de van Inwagen. Penso que minhas formulações, embora preservem o caráter intuitivo do argumento acima apresentado, são imunes a algumas objeções persuasivas às formulações originais de van Inwagen. Vou deixar a apresentação das minhas próprias formulações para o capítulo 3, infelizmente, para não perder a dialética da argumentação. O capítulo 3 lida com as principais objeções ao Argumento da Consequência; minhas formulações são

tentativas de consertar algumas das falhas que os argumentos de van Inwagen apresentam.

Este capítulo, assim como o anterior, está dividido em três partes. A primeira parte é a formulação modal do Argumento da Consequência, o que incluirá certo esclarecimento sobre os conceitos de proposição, leis da natureza e estado total do mundo. Depois disso, apresento algumas razões para mostrar a plausibilidade das regras de inferência utilizadas no argumento e, principalmente, sobre a regra beta. Como vimos no primeiro capítulo, alguns argumentos incompatibilistas repousam em falácias modais óbvias. “Incompatibilismo” aqui é para ser entendido em sentido amplo; por exemplo, o argumento que apresentei a favor da incompatibilidade entre a onisciência e o livre-arbítrio é para ser entendido como um argumento a favor do incompatibilismo em sentido amplo. Isso não fora apresentado por acaso, claro. Historicamente, os filósofos pensavam que os argumentos a favor do incompatibilismo repousavam em falácias modais óbvias. A discussão sobre a regra beta será importante para mostrar que o Argumento da Consequência não repousa numa falácia modal *óbvia*. Irei comparar, por exemplo, a regra beta com a famosa falácia fatalista.

Finalmente, a última seção deste capítulo consiste na apresentação de outra formulação do Argumento da Consequência. Trata-se, aliás, da formulação original que van Inwagen propusera no já clássico “The incompatibility of determinism and free will”. Ela será importante porque discutirei uma recente objeção a esse argumento apresentada por Jonathan Westphal (2012). Por que então apresentar inicialmente a versão modal? Primeiro porque a versão modal é a mais discutida na literatura recente. Em segundo lugar, ela não sofre do problema levantado por Westphal com relação às duas primeiras formulações.

## 2.1 O Argumento da Consequência

Antes de formular o argumento de maneira mais precisa, precisaremos retornar às definições de determinismo e livre-arbítrio. A definição que ofereci no capítulo anterior de determinismo era a seguinte:

Determinismo = *def.* uma proposição verdadeira sobre o estado total do mundo em qualquer tempo  $t$ , e uma proposição que seja a conjunção de todas as leis da natureza, implicam estritamente uma proposição verdadeira sobre o mundo em qualquer tempo posterior a  $t$ .



Como já tinha dito anteriormente, os conceitos de proposição, leis da natureza e estado total do mundo não foram devidamente explicados no capítulo anterior. O objetivo desta seção é o de esclarecer esses conceitos. Vamos começar com o conceito de proposição.

Para início de conversa, tenho infelizmente de dizer que não sei oferecer uma boa definição para nenhum desses conceitos. O mesmo se aplica a proposição, naturalmente. Alguns filósofos, mais notavelmente Robert Stalnaker (1976), oferecem uma análise do conceito de proposição em termos de mundos possíveis. Grosso modo, como afirma Hoffman, esses filósofos “sustentam que toda a proposição é idêntica ao conjunto de mundos possíveis em que ela é verdadeira” (HOFFMAN, 2010, p.449). Tome-se como exemplo a proposição expressa pela frase “Aristóteles é um filósofo grego”. Segundo essa concepção, aquela proposição é idêntica ao conjunto de mundos possíveis em que Aristóteles é um filósofo grego. Essa é uma visão promissora, mas dizer que ela é imune a problemas é um exagero.

Além disso, afirmei no primeiro capítulo que não sei definir o conceito de mundo possível. Não quero com isso dizer que eu veja algum problema em usá-lo intuitivamente, pois a noção de mundo possível tem um aspecto intuitivo que parece iluminar as noções modais de necessidade e possibilidade. Porém, eu gostaria aqui de oferecer uma definição rigorosa do conceito de proposição; e uma compreensão intuitiva do conceito de mundo possível não seria o bastante para esse propósito.

O que usualmente se assume com relação ao termo “proposição” está descrito nas palavras de McGrath (2012):

O termo “proposição” tem um amplo uso na filosofia contemporânea. Ele é usado para referir a algumas (ou todas) das seguintes coisas: os portadores primários de valor de verdade, os objetos de crença e outras “atitudes proposicionais” (i.e., o que é acreditado, duvidado, etc.), os referentes das cláusulas-quê, e os significados das frases.

Embora não seja incomum usar o termo “proposição” para se referir aos portadores primários de valor de verdade, essa posição não é incontroversa. Muitos filósofos a colocam em causa, assim como consideram como problemático, por exemplo, estipular a existência de entidades como proposições para explicar o significado de frases. No entanto, para os

propósitos dessa dissertação, não dependerei de características especiais de alguma explicação particular das proposições, muito menos lidar com os problemas relacionados à natureza da proposição; por exemplo, proposições realmente existem? Se existem, elas possuem localização espaço-temporal? Esses problemas são estimulantes, mas não são importantes aqui. Tudo o que preciso admitir é que as proposições tenham suas características comuns. E por “característica comum” eu quero dizer que as proposições instanciam as seguintes propriedades (VAN INWAGEN 1983, p.58):

(1) Para todo o modo possível que o mundo poderia ser, corresponde pelo menos uma proposição, uma proposição que é necessariamente tal que ela é verdadeira se, e só se, o mundo é daquele modo.

(2) Toda proposição é verdadeira ou falsa.

(3) A conjunção de uma proposição verdadeira e uma proposição falsa é uma proposição falsa.

(4) Proposições obedecem à lei da contraposição no que diz respeito ao acarretamento. Isto é, para todo  $x$  e para todo  $y$ , se  $x$  e  $y$  são proposições, e se é impossível para  $x$  ser verdadeira e  $y$  falsa, então é impossível para a negação de  $y$  ser verdadeira e a negação de  $x$  ser falsa.

Admitindo isso, o leitor pode pensá-las como quiser. Por exemplo, pode pensá-las como funções de mundos possíveis em valores de verdade, admitir a Análise Relacional da Atribuição de Atitudes (McGrath 2012, §3.1), etc.

Passemos agora à noção de leis da natureza. Ainda é uma questão em aberto saber o que é uma lei da natureza. Porém, o que se pressupõe nesta dissertação não nos compromete substancialmente com uma resposta ao problema. O objetivo aqui é apenas mostrar, para fins de argumentação, algumas características que o conceito de leis da natureza tem de satisfazer.

Como já tinha afirmado no capítulo anterior, leis da natureza, ontologicamente falando, são proposições. Algumas proposições instanciam a propriedade de ser uma lei da natureza, outras não. Intuitivamente, é fácil de oferecer alguns exemplos de proposições que *não* são leis da natureza. Por exemplo, proposições como aquelas expressas pelas frases “Sócrates bebeu a cicuta”, “Aristóteles é o autor da *Ética a Nicômaco*”, etc., *não* são leis da natureza. É muito mais difícil, entretanto, oferecer exemplos de proposições que são leis da natureza.

Os filósofos tipicamente aceitam que uma lei da natureza não é uma generalização accidental; supõe-se, adicionalmente, que tal generalização tem de ser suficientemente abrangente. Assim, por exemplo, a proposição expressa pela frase “Todas as notas de cinco reais são azuis” não é uma lei da natureza; embora essa proposição seja uma generalização verdadeira, ela é simplesmente accidental. Ademais, leis da natureza têm de ser regularidades abrangentes, o que não ocorre quando levamos em conta a proposição expressa pela frase “Todas as notas de cinco reais são azuis”.

Embora aceite-se que uma lei da natureza não seja uma generalização accidental, ainda é controverso se as leis da natureza são necessariamente verdadeiras. Por exemplo, admitindo que a proposição expressa pela frase “Nenhum objeto viaja mais rápido que a luz” seja uma lei da natureza, é argumentável (cf. CARROLL, 2010, §8) que, por ser concebível que um objeto não viaje mais depressa do que a luz, essa proposição é apenas contingentemente verdadeira. Esse argumento merece ser explicitado. Considere o seguinte raciocínio:

1. Se  $P$  é concebível, então  $P$  é possível.
2. Se a proposição expressa por “algum objeto viaja mais depressa do que a luz” é concebível, então é possível. (Instância do esquema 1).
3. A proposição expressa por “algum objeto viaja mais depressa do que a luz” é concebível.
4. A proposição expressa por “algum objeto viaja mais depressa do que a luz” é possível. De 3 e 4, por *modus ponens*.

Agora vejamos outro raciocínio:

1. A proposição expressa por “Nenhum objeto viaja mais depressa do que a luz” é verdadeira.
2. A proposição expressa por “Nenhum objeto viaja mais depressa do que a luz” é possível. (De 1, T<sup>14</sup>).

Finalmente, considere a seguinte definição de contingência:

$P$  é contingente sse  $P$  é possível e  $\sim P$  é possível.

---

<sup>14</sup> Estou assumindo a seguinte regra de inferência:  $P \vdash \diamond P$ .

Admitamos que  $P$  é a proposição expressa por “algum objeto viaja mais depressa do que a luz”. Considerando os dois raciocínios anteriores, nomeadamente, a conjunção das conclusões dos dois raciocínios, podemos inferir que a proposição expressa por “algum objeto viaja mais depressa do que a luz” é contingente.

Ora, tenho sérias dúvidas de que este seja um argumento cogente para nos persuadir a aceitar que as leis da natureza não são necessariamente verdadeiras. Tenho sérias dúvidas de que seja um argumento cogente porque tenho sérias dúvidas quanto ao esquema 1 do primeiro raciocínio. *Prima facie*, a concebilidade não implica a possibilidade. Por exemplo, se é concebível que a água não seja duas moléculas de hidrogênio e uma de oxigênio, não se segue, como muitos essencialistas parecem pensar, que a água possivelmente não seja duas moléculas de oxigênio e uma de hidrogênio. É concebível que Sócrates seja um cartão de crédito VISA, mas não é possível que Sócrates seja um cartão de crédito VISA.

É claro que esses contraexemplos podem ser respondidos (veja, em especial, Yablo (1998), Geirsson (2005), Chalmers (2002)). Pode-se, por exemplo, simplesmente enfraquecer a tese. Admitindo que  $C(P)$  significa “é concebível que  $P$ ”,  $\Diamond P$  “é possível que  $P$ ” e  $I(X/Y)$  “ $X$  é um indício para  $Y$ ”, considere o seguinte. Ao invés de defender

$$1) C(P) \supset \Diamond P,$$

defender-se-ia

$$2) I(C(P)/\Diamond P).$$

Em resumo, ao invés de defender que algumas leis da natureza são contingentes, defender-se-ia a tese mais fraca de acordo com a qual *temos indícios* para acreditar que algumas leis da natureza são contingentes. Em todo o caso, não gostaria de prolongar a discussão, pois isso desviaria sobremaneira o foco desta dissertação. O problema em causa é o de saber se a concebilidade implica a possibilidade ou, pelo menos, se é fonte de justificação para crenças modais. Este último tópico é realmente estimulante, mas pertence à epistemologia modal. Humianos acerca de leis da natureza defendem justamente essa tese mais fraca. Por exemplo, alguns humianos têm interesses

tanto epistêmicos quanto ontológicos sobre conexões necessárias. No primeiro caso, defendem que, uma vez que conexões necessárias não podem ser observadas, a existência de tais conexões é inconsistente com uma epistemologia empirista. No segundo caso, humianos defendem que não há quaisquer razões para postular entidades exóticas, tais como a existência de conexões necessárias (BEEBEE & MELE 2003, p.202). Em resumo, tipicamente quem tem uma perspectiva humiana acerca de leis da natureza defenderá que não temos indícios para sustentar que as leis da natureza são necessárias.

Seja como for, a discussão acima não é importante porque, tal como entendi o conceito de leis da natureza, está explícito que não irei pressupor que elas sejam necessariamente verdadeiras; portanto, não pressuponho uma concepção “necessarista” de leis da natureza. Um exemplo dessa concepção pode muito bem ser ilustrada recorrendo ao trabalho de Armstrong (1983: 11-2). De acordo com Armstrong, leis da natureza são constituídas por uma relação contingente de necessitação entre universais. Ou seja, é uma lei que todos os  $Fs$  são  $Gs$  se, e somente se, ocorre a relação de necessitação contingente  $N$  entre os universais  $F$  e  $G$  (essa relação é contingente no sentido de que é uma questão contingente que  $N$  seja instanciada). Em oposição à concepção de Armstrong, temos, naturalmente, uma perspectiva humiana acerca de leis da natureza, a mais plausível das quais provavelmente seja a chamada Perspectiva de Ramsey-Lewis (RAMSEY 1978, LEWIS 1973, p.73), a qual evita alguns contraexemplos à visão humiana. O leitor verá no próximo capítulo que será importante deixar esse debate em aberto, pois algumas objeções ao Argumento da Consequência repousam numa visão humiana acerca das leis da natureza.

Ademais, vou admitir – muito plausivelmente, suponho – que o conceito de leis da natureza *não* é epistêmico. Uma proposição é uma lei da natureza independentemente de a conhecermos ou de termos uma crença justificada sobre ela. Pode ser que as proposições que pensamos ser leis da natureza de fato não o sejam, e que também não venhamos a descobrir qualquer lei da natureza, ou a ter uma crença justificada na sua existência. O conceito de lei da natureza, nesse sentido, é semelhante ao conceito de verdade. Uma proposição é verdadeira independentemente de a conhecer ou de termos justificção para crer nela.

Alguém poderia ser tentado a dizer que as leis da natureza suportam contrafactuais. Estou totalmente de acordo que suportar contrafactuais seja uma condição necessária para ser uma lei da natureza. Porém, “suportar contrafactuais” é condição suficiente para uma proposição ser uma lei da

natureza? Parece que não. Van Inwagen apresenta um contraexemplo bem-sucedido:

Considere a proposição “Os cães morrem se expostos ao vírus V” [sic.]. A afirmação de que essa proposição suporta as suas contrafactuais é, penso, equivalente à afirmação de que “Todo cão é tal que, se fosse exposto ao vírus V, morreria” é *verdadeira*. Suponhamos que essa última proposição *seja* verdadeira, e que a quantificação seja entendida como sendo sobre todos os cães, passados, presentes e futuros. A sua verdade, parece-me, é bem consistente com o fato de que criadores de cães *poderiam* (mas não *irão*) instituir um programa de reprodução seletiva que *produziria* um tipo de cão que é imune ao vírus V. Mas se os criadores de cães *poderiam* fazer isso, então claramente “Cães morrem se expostos ao vírus V” não é uma lei da natureza, visto que nesse caso a verdade da contrafactual universalmente quantificada correspondente depende de uma circunstância accidental: se os criadores instituíssem um certo programa de reprodução seletiva que eles são totalmente capazes de instituir, então “Todo cão é tal que, se fosse exposto ao vírus V, morreria” poderia ser falsa. Assim, uma proposição pode “suportar as suas contrafactuais” e, ainda assim, não ser uma lei da natureza (VAN INWAGEN 1974, pp.187-8).

Para fins de argumentação, estipularei que não faz parte da extensão do conceito de leis da natureza leis psicológicas que incluam leis sobre o comportamento voluntário dos agentes racionais (cf. VAN INWAGEN, 2004, pp. 696-7 e 1983, p.64). Essa restrição serve para impedir que o determinismo seja trivialmente incompatível com o livre-arbítrio: é que se houver leis sobre o comportamento dos agentes, então esses agentes têm de se comportar do modo como essas leis descrevem os seus comportamentos. Finalmente, convém apenas lembrar que, para a formulação do argumento, “L” será entendida como a abreviação de uma frase que expressa uma proposição que é a conjunção de todas as leis da natureza.

Agora o conceito de estado total do mundo num passado distante. Tudo o que será preciso aqui é apenas uma compreensão intuitiva. Esta ideia intuitiva é a de que há uma proposição que descreve o estado total do mundo num passado distante, antes da existência de quaisquer agentes; por exemplo,

podemos pensar numa proposição que descreva o estado total do mundo no início do universo. Não é decisivo explicar o conceito de estado total do mundo porque o argumento de van Inwagen que irei apresentar é independente de seu conteúdo. O que precisamos ter em mente é apenas a seguinte restrição: o conceito de estado total do mundo tem de ser tal que, dado o que o mundo é num certo estado e em certo instante, nada se segue desse estado em qualquer outro instante. Por exemplo, o conceito de estado não pode permitir a cláusula “e, neste instante, o mundo é tal que a mão de alguém será levantada 10 segundos depois deste instante” (cf. VAN INWAGEN, 2004, p.696 e 1983, p.58). Esta restrição serve para não permitir que o conceito *estado do mundo* possa ser definido de tal modo que o determinismo seja trivialmente verdadeiro, pois poderíamos acrescentar uma cláusula contendo informações sobre o futuro que dirão sempre o que iremos escolher amanhã. Finalmente, lembremos que “ $P_0$ ” será entendida como a abreviação de uma frase que expressa uma proposição verdadeira sobre o estado total do mundo em algum tempo, num passado distante, antes de quaisquer agentes existirem. Em suma, temos as seguintes definições:

- “L” é a abreviação de uma frase que expressa uma proposição que é a conjunção de todas as leis da natureza.
- “ $P_0$ ” é a abreviação de uma frase que expressa uma proposição verdadeira sobre o estado total do mundo em algum tempo, num passado distante, antes de quaisquer agentes existirem.

“L” e “ $P_0$ ” são entendidas como abreviações de frases que expressam proposições, e não como nomes, por uma razão simples. Se fossem entendidas como nomes, uma frase como “ $P_0$  & L” seria agramatical, pelo mesmo motivo que a frase “Pedro e João” é agramatical. E, como devemos nos lembrar, precisamos de “ $P_0$  & L” para a formulação do determinismo:

Determinismo:  $\square ((P_0 \ \& \ L) \supset P)$

Lembremos que, enquanto o determinismo é uma tese sobre proposições, o livre-arbítrio é uma tese sobre agentes; e para investigar as relações conceituais entre essas duas teses, precisamos considerar o livre-arbítrio como uma tese sobre agentes e proposições (cf. VAN INWAGEN, 1983, p.66). Como vimos anteriormente, o operador modal “N” ajuda a esclarecer as relações conceituais entre essas duas teses. Este operador fora caracterizado da seguinte maneira: “NP” abrevia “P, e ninguém tem, nem

nunca teve, qualquer escolha sobre se  $P$ ". Portanto, caso se mostre que do determinismo se segue que " $\mathbf{N}P$ ", estabelece-se a tese incompatibilista.

A expressão "ter escolha sobre se  $P$ ", como mencionei no capítulo anterior, será entendida como "ter o poder de tornar  $P$  falsa". Um dado agente tem escolha sobre se  $P$  se, e só se, pode tornar  $P$  falsa. Por exemplo, seja  $P$  a proposição de que dois e dois são quatro, é verdade que  $\mathbf{N}P$ , pois não podemos tornar essa proposição falsa. Por outro lado, seja  $Q$  a proposição de que levantei minha mão, é *prima facie* falso que  $\mathbf{N}Q$ , pois supostamente poderia tornar  $Q$  falsa ao não levantar minha mão.

Precisamos de uma definição mais precisa de " $S$  pode tornar  $P$  falsa". Van Inwagen nos oferece a seguinte definição:

Está dentro do poder de  $S$  organizar ou modificar os objetos concretos que constituem seu ambiente de tal modo que não é possível, em sentido lógico amplo, que ele organize ou modifique aqueles objetos daquele modo e o passado tenha sido exatamente como de fato foi e  $P$  seja verdadeira (VAN INWAGEN 1983, p.68).

Nesta definição, "poder" é para ser entendido tal como sugeri no capítulo anterior. "Possível em sentido amplamente lógico" é para ser entendido como "possibilidade conceitual", tal como defini no capítulo anterior. "Proposição" fora brevemente caracterizado nos parágrafos acima. Os outros termos, como "objeto concreto", tomarei como primitivos.

Seria ótimo se tivéssemos uma lógica completa para o operador " $\mathbf{N}$ ". Infelizmente ainda não a temos. Contudo, isso não será um problema para a apresentação do argumento; o que podemos fazer é estipular, tal como van Inwagen faz, que as duas regras de inferência seguintes são válidas (VAN INWAGEN, 1983, p.94):

$$(\alpha) \Box P \vdash \mathbf{N}P$$

$$(\beta) \mathbf{N}(P \supset Q), \mathbf{N}P \vdash \mathbf{N}Q$$

Intuitivamente, a regra  $(\alpha)$  diz que, no caso de  $P$  ser uma proposição necessariamente verdadeira, então podemos concluir que  $P$  e ninguém tem, nem nunca teve, qualquer escolha sobre se  $P$ . A regra  $(\beta)$  é apenas a regra de fechamento para o operador " $\mathbf{N}$ ". Parece plausível acreditar que essas regras



sejam válidas. Essa plausibilidade, contudo, pode ser posta em causa, de modo que, após a apresentação do argumento, mostrarei algumas das razões a favor da crença de que essas regras sejam válidas.

Considerando toda a discussão até aqui, podemos finalmente oferecer uma formulação mais precisa do Argumento da Consequência. Em sua forma intuitiva, é importante lembrar, o argumento é o seguinte:

Se o determinismo for verdadeiro, então nossas ações são a consequência das leis da natureza e dos eventos no passado remoto. Mas não depende de nós o que se passou antes de nascermos, e nem depende de nós o que as leis da natureza são. Portanto, as consequências dessas coisas (incluindo nossas ações presentes) não dependem de nós (VAN INWAGEN, 1983, p.56).

Há duas premissas que estão explícitas nesta versão intuitiva do argumento, a saber, a de que não depende de nós o que as leis da natureza são e a de que não depende de nós o que é o passado. De modo mais preciso, as premissas são as seguintes:  $NL$  e  $NP_0$ . O argumento pode, enfim, ser apresentado de modo mais preciso.

Suponhamos que o determinismo seja verdadeiro, isto é, suponhamos que necessariamente, se “ $L$ ” e “ $P_0$ ” são verdadeiras, então “ $P$ ” também o é:

$$1. \Box ((P_0 \& L) \supset P)$$

Da primeira premissa, infere-se pela regra de exportação da lógica modal alética

$$2. \Box (P_0 \supset (L \supset P))$$

Aplicando a regra ( $\alpha$ ) em 2, temos

$$3. \mathbf{N} (P_0 \supset (L \supset P))$$

Agora introduzimos a premissa de que não temos escolha sobre o passado. A ideia intuitiva a essa premissa, como vimos, é a de que não temos o poder de tornar falsa uma proposição sobre o estado total do mundo num

passado distante. Em outras palavras, não temos escolha sobre o que se passou antes de existirmos.

4.  $NP_0$

e de 3 e 4 infere-se, pela regra ( $\beta$ ) que

5.  $N(L \supset P)$ .

Novamente acrescentamos mais uma premissa, a saber, a premissa segundo a qual não temos escolha sobre as leis da natureza. A ideia intuitiva dessa premissa é a de que não temos o poder de tornar falsa uma proposição que seja a conjunção de todas as leis da natureza.

6.  $NL$

Inferimos agora de 5 e 6, pela regra ( $\beta$ )

7.  $NP$

Logo, se 1 é verdadeira (isto é, se o determinismo é verdadeiro), então 7 é verdadeira, pela regra de introdução da condicional material.

8.  $\square ((P_0 \ \& \ L) \supset P) \supset NP$

Ora, 8 é precisamente a tese incompatibilista definida no capítulo 1.

Penso que este argumento, ao menos à primeira vista, coloca o compatibilista em dificuldades. Para o defensor do compatibilismo mostrar que o argumento não é cogente, terá de mostrar que pelo menos uma de nossas crenças nestas proposições não é plausível:

- 1) A regra de exportação para a lógica modal alética é válida<sup>15</sup>
- 2) A regra ( $\alpha$ ) é válida
- 3) A regra ( $\beta$ ) é válida

---

<sup>15</sup> A rigor, não há tal coisa como a regra de exportação para a lógica modal alética. Há a regra de exportação para a lógica clássica. O que fiz foi estipular que a expressão “regra de exportação para a lógica modal alética” refere-se à inferência de 2 por meio de 1 no argumento.

- 4) NL
- 5) NP<sub>0</sub>

Objecções a 1 e a 2 não são comuns, mas ao menos são concebíveis; por exemplo, filósofos que têm alguma razão independente para rejeitar a lógica modal alética podem rejeitar 1; um defensor da teoria convencionalista da necessidade poderia objetar 2, ou algum cartesiano poderia conceder que Deus tem escolha sobre se  $P$ , apesar de  $P$  ser necessariamente verdadeira. Com relação a 3, há diversas objeções, e algumas delas foram apresentadas por Michael Slote (1982), McKay e Johnson (1996), e Fara (2008). David Lewis (1983) e Beebe (2003) objetaram 4. Foley (1979) e Campbell (2007), por exemplo, objetaram 5. As objeções a 1 e a 2 serão consideradas aqui brevemente.

### 1.2 As regras ( $\alpha$ ) e ( $\beta$ )

Vejamos novamente o argumento:

- |    |                                       |                          |
|----|---------------------------------------|--------------------------|
| 1. | $\square ((P_0 \& L) \supset P)$      | determinismo             |
| 2. | $\square (P_0 \supset (L \supset P))$ | de 1                     |
| 3. | $N (P_0 \supset (L \supset P))$       | 2, regra ( $\alpha$ )    |
| 4. | NP <sub>0</sub>                       | premissa                 |
| 5. | $N (L \supset P)$                     | 3 e 4, regra ( $\beta$ ) |
| 6. | NL                                    | premissa                 |
| 7. | NP                                    | 5 e 6, regra ( $\beta$ ) |

Há três regras de inferência utilizadas explicitamente no argumento acima. A regra de exportação da lógica modal alética e as regras ( $\alpha$ ) e ( $\beta$ ). Nesta seção, discutirei a plausibilidade dessas regras. As premissas 4 e 6 serão defendidas no capítulo seguinte.

A premissa 2 se segue de 1 por meio da regra de exportação da lógica modal alética. Os filósofos que tiverem alguma razão independente para rejeitar a lógica modal alética, como o filósofo americano W. O. Quine (1960) com certeza tinha, podem rejeitar a inferência de 2 por meio de 1. Entretanto, não vejo essa manobra como promissora. O argumento que tenho contra essa réplica repousa na noção de cogência usada na última seção do primeiro capítulo. Qualquer argumento que pressuponha a rejeição da lógica modal

alética não será cogente, pois, mesmo que esse argumento seja sólido, ele não será racionalmente persuasivo. Para ser racionalmente persuasivo, teríamos de ter um argumento sólido racionalmente persuasivo para abandonar a lógica modal alética. Como não temos esse argumento, não temos razão para rejeitar a inferência de 2 por meio de 1.

Já a premissa 3 se segue de 2 através da regra ( $\alpha$ ). Esta regra é muito plausível. Anteriormente, eu (2012a) havia apresentado o seguinte argumento a favor de ( $\alpha$ ) no intuito de excluir contraexemplos (admitindo que as condicionais das premissas são condicionais estritas):

1. Se  $P$  é necessariamente verdadeira, então  $P$  não pode ser falsa.
2. Se  $P$  não pode ser falsa, então ninguém pode tornar  $P$  falsa.
3. Se ninguém pode tornar  $P$  falsa, então ninguém tem escolha sobre se  $P$ .
4. Logo, se  $P$  é necessariamente verdadeira, então ninguém tem escolha sobre se  $P$ .

O problema desse argumento é que ele é uma petição de princípio. A premissa 2 é a conclusão que quero estabelecer, a saber, se  $P$  não pode ser falsa (o que é equivalente a “ $P$  é necessariamente verdadeira”), então ninguém pode tornar  $P$  falsa. Parece-me plausível que qualquer pessoa que tenha alguma razão para recusar 4, recusará também 2. E talvez, na verdade, exista algum contraexemplo a 2. Considere o seguinte.

Por contraposição, o que a premissa afirma é que, se alguém pode tornar  $P$  falsa, então  $P$  é possivelmente falsa; ou, se quisermos usar a linguagem dos mundos possíveis, se alguém pode tornar  $P$  falsa, então há um mundo possível no qual  $P$  é falsa. Mas talvez seja plausível pensar que essa condicional não seja verdadeira. Considere, por exemplo, a seguinte condicional:

C1: Se um copo tem a disposição de se quebrar, então é possível que ele se quebre.

Pode-se argumentar que C1 é falsa. Mesmo que não exista um mundo possível no qual o copo se quebra, ele ainda assim não perde a disposição de se quebrar. No capítulo anterior, assumi que disposições podem não ser efetivadas. Ora, o mesmo parece ocorrer com o poder:

C2. Se alguém pode tornar  $P$  falsa, então é possível que  $P$  seja falsa.

Suponha, por exemplo, que “ $P$ ” seja a abreviação de uma frase que expresse a proposição de que ninguém nunca leu a *Crítica da Razão Pura* toda em voz alta. Essa proposição será possivelmente falsa sse possivelmente alguém ler a *Crítica* toda em voz alta; admitindo (plausivelmente, suponho) que não exista um único mundo possível no qual alguém tenha paciência de ler a *Crítica* toda em voz alta, mesmo assim é verdade que alguém pode lê-la toda e, portanto, tornar  $P$  falsa. Ora, se há razões para pensar que 2 seja falsa, há razões para pensar que 4 seja falsa. Logo, há razões para pensar que a regra ( $\alpha$ ) seja inválida.

Penso que o problema dessa objeção é o de não levar em conta qual o tipo de modalidade está em causa. A necessidade de “ $\Box P$ ” é a necessidade lógica. Teríamos de mostrar que a seguinte condicional seja falsa:

C2\*. Se alguém pode tornar  $P$  falsa, então é logicamente possível que  $P$  seja falsa.

O erro do contraexemplo é o de admitir que a proposição expressa pela frase “Ninguém nunca leu a *Crítica da Razão Pura* toda em voz alta” não é uma possibilidade lógica; é claro que é uma possibilidade lógica, pois, considerando a definição de modalidade lógica apresentada anteriormente, essa proposição é perfeitamente consistente com os axiomas de qualquer sistema. Portanto, o contraexemplo não é bem-sucedido.

Outro contraexemplo que alguém poderia imaginar à regra ( $\alpha$ ) seria o seguinte: suponha que a inferência “ $\Box P \vdash P$ ” seja inválida (e que, portanto, o sistema T não seja correto para o raciocínio modal), ou seja, que exista uma circunstância na qual “ $\Box P$ ” é verdadeira e “ $P$ ” falsa. Portanto, “ $\Box P \vdash NP$ ” também é inválida, uma vez que “ $\Box P$ ” é verdadeira e “ $NP$ ” é falsa, já que esta última abrevia “ $P$  e ninguém tem, nem nunca teve, qualquer escolha sobre se  $P$ ” e sua primeira conjunta seria falsa. O problema dessa objeção é o mesmo problema que eu já havia apresentado anteriormente à objeção da inferência de 2 por meio de 1: um argumento que pressuponha que T não seja um sistema correto para o raciocínio modal não será cogente. Entretanto, essa objeção tem um aspecto esclarecedor: uma lógica completa para “N” terá de admitir um sistema pelo menos mais forte do que K.

Essas foram as melhores objeções que pude imaginar à regra ( $\alpha$ ), mas me pareceram malsucedidas. Vou admitir, de agora em diante, que a regra ( $\alpha$ ) é válida até que se prove o contrário. Essa suposição é perfeitamente plausível, tanto que não há qualquer artigo influente objetando o Argumento da Consequência em virtude da rejeição de ( $\alpha$ ).

Passemos agora à regra ( $\beta$ ). Como disse anteriormente, alguns argumentos incompatibilistas repousam em falácias modais óbvias. Afirmar também que “incompatibilismo” aqui é para ser entendido em sentido amplo, ou seja, no sentido de que o argumento que apresentei a favor da incompatibilidade entre a onisciência e o livre-arbítrio é um argumento a favor do incompatibilismo. Ademais, afirmar que esses argumentos não foram apresentados por acaso. Pelo menos até o trabalho de van Inwagen, a esmagadora maioria dos filósofos assumia o compatibilismo, ao mesmo tempo em que negligenciavam os argumentos a favor do incompatibilismo. Isto é cristalino na seguinte passagem de Donald Davidson.

Não estarei diretamente interessado nisso [nos argumentos a favor da incompatibilidade entre a liberdade e a determinação causal], uma vez que sei muito bem que esses argumentos são, na melhor das hipóteses, apenas superficialmente plausíveis. Hobbes, Locke, Hume, Moore, Schlick, Ayer, Stevenson, e uma série de outros tem feito o que pode ser feito, ou o que deveria ter sido feito, para remover as confusões que podem aparentemente tornar o determinismo oposto à liberdade (DAVIDSON, 1973, p.139).

Pelo menos até o trabalho de Peter van Inwagen, tipicamente se pensava que os argumentos a favor do incompatibilismo repousavam em alguma falácia modal óbvia, ou que eles confundiam determinação e compulsão (objeção de Ayer), leis descritivas com leis prescritivas, etc. Com relação à acusação de os argumentos incompatibilistas repousarem em falácias modais óbvias, não é de todo claro. Por exemplo, a regra ( $\beta$ ) do argumento de van Inwagen não deve ser confundida com a seguinte forma argumentativa inválida:

1.  $\square (P \supset P)$
2. P

3. Logo,  $\Box P$ 

Um dos argumentos mais famosos a favor do fatalismo metafísico tem a forma argumentativa supracitada. Trata-se da famosa “falácia fatalista”. Eis um exemplo:

1. Necessariamente (se o João decidiu levantar sua mão, então o João decidiu levantar sua mão)
2. O João decidiu levantar sua mão
3. Logo, necessariamente, o João decidiu levantar sua mão.

Uma de nossas intuições com relação ao livre-arbítrio é que, embora tenhamos decidido agir de um modo, poderíamos decidir agir de outro. Segundo este argumento, porém, o João não poderia decidir não levantar sua mão. Mas é claro que o nome já diz tudo: “falácia fatalista”. Uma simples árvore lógica mostra por que a falácia fatalista faz jus ao nome.

1. 0:  $\Box (P \supset P)$

2. 0: P

3. 0:  $\sim\Box P$

4. 0:  $\Diamond \sim P$  (3)

5. 1:  $0 \rightsquigarrow 1$  (4)

6. 1:  $\sim P$  (4,5)

7. 1:  $P \supset P$  (1,5)



8. 1:  $\sim P$  (7)

9. 1: P (7)

x

A regra ( $\beta$ ) do argumento de van Inwagen parece estabelecer a diferença entre o Argumento Modal da Consequência e a falácia fatalista. Considere a seguinte forma argumentativa válida.

1.  $\Box(P \supset Q)$
2.  $\Box P$
3. Logo,  $\Box Q$

Novamente recorrerei a uma árvore lógica para demonstrar que a forma argumentativa supracitada é válida.

1. 0:  $\Box(P \supset Q)$
2. 0:  $\Box P$
3. 0:  $\sim\Box Q$
4. 0:  $\Diamond\sim Q$  (3)
5. 0 $\sim$ 1 (4)
6. 1:  $\sim Q$  (4,5)
7. 1: P (2,5)
8. 1:  $P \supset Q$  (1,5)



- |  |  |
|--|--|
| <ol style="list-style-type: none"> <li>9. 1: <math>\sim P(8)</math></li> </ol> | <ol style="list-style-type: none"> <li>10. 1: Q (8)</li> </ol> |
| x  | x  |

Intuitivamente, a regra ( $\beta$ ) está mais próxima dessa forma argumentativa do que a falácia fatalista. Claro, o operador modal da “não escolha” é, de fato, diferente do operador modal da necessidade. Portanto, do fato de a regra funcionar para o operador modal da necessidade não se segue que ela também funcione para o operador da não escolha. Não obstante, o operador da não escolha parece ser similar o bastante ao operador da necessidade para pensarmos que a regra ( $\beta$ ) seja uma regra de inferência válida. Se o Argumento Modal da Consequência repousa em uma regra falaciosa, isso ao menos não é de todo óbvio. Como veremos no próximo capítulo, há ótimas razões para pensar que a regra ( $\beta$ ) realmente não é válida. Mas, como veremos, os argumentos são bastante sofisticados, e de modo algum mostram que o equívoco dessa formulação seja evidente.



Qualquer leitor atento de van Inwagen poderia se perguntar por que optei por apresentar inicialmente a versão modal do Argumento da Consequência, ao invés das duas primeiras versões, já que a versão modal é a última das formulações apresentadas por van Inwagen em *An Essay on Free Will*. Optei por isso por uma razão muito simples. A formulação modal é a melhor maneira de tornar o Argumento Modal mais preciso. Isso é controverso, pois van Inwagen afirma que as três formulações são maneiras distintas de apresentar o mesmo argumento:

Denominei este capítulo de “Três Argumentos a favor do Incompatibilismo”, mas o princípio de individuação para os argumentos é muito claro; Eu poderia denomina-los “Um argumento a favor do incompatibilismo feito de três maneiras (VAN INWAGEN, 1963, p.56).

Ao contrário do que van Inwagen afirma, contudo, há quem pense que a formulação modal é, sim, diferente do Primeiro Argumento Formal, assim como do Segundo Argumento Formal. Em um artigo recente, Jonathan Westphal (2012) argumentou que van Inwagen comete uma falácia de escopo modal óbvia em sua defesa da segunda premissa do Primeiro Argumento Formal; e defendeu também que o problema apresenta-se com igual força ao Segundo Argumento Formal. Como veremos, concordo com a objeção de Westphal; van Inwagen realmente comete uma falácia modal. Entretanto, penso que há uma maneira natural de contornar essa dificuldade e defender a segunda premissa do Primeiro Argumento Formal.

### 1.3 O Primeiro Argumento Formal

Antes de apresentar o Primeiro Argumento Formal, temos de notar que, na formulação modal, “ $P_0$ ” e “ $L$ ” foram usadas como abreviações de frases que expressam proposições, o que significa que uma frase como “ $P_0 \supset L$ ” é gramatical, mas também que “ $P_0$  é verdadeira” é agramatical, pelo mesmo motivo que “Todos os homens são mortais é verdadeira” é agramatical. No Primeiro Argumento Formal, “ $P_0$ ” e “ $L$ ” são usados como nomes para proposições, o que significa que uma frase como “ $P_0 \supset L$ ” é agramatical, pelo mesmo motivo que “Se Pedro, então João” é agramatical, mas “ $L$  é verdadeira” é gramatical (cf. VAN INWAGEN, 1983, p.56).

Nesta formulação, van Inwagen imagina um exemplo no qual um juiz não levanta a sua mão para absolver um prisioneiro; no país do juiz, a ação de

levantar a mão é um ato de clemência. Conseqüentemente, o prisioneiro é sentenciado à morte. O que van Inwagen argumenta é que, se o determinismo for verdadeiro, então o juiz não poderia ter levantado a sua mão, e por isso o prisioneiro seria sentenciado à morte. Como afirma Westphal, “nada poderia ilustrar melhor o comprometimento de van Inwagen ao que ele toma como sendo as rígidas implicações do determinismo do que este sombrio exemplo” (WESTPHAL, 2012, p.36).

Note-se que este é um exemplo escolhido aleatoriamente. Um argumento paralelo pode ser facilmente desenvolvido para o caso de qualquer agente e qualquer ação não realizada. Assim, precisamente porque este argumento não dependerá de quaisquer características peculiares ao caso imaginado, acredita van Inwagen, a incompatibilidade entre o livre-arbítrio e o determinismo será estabelecida (VAN INWAGEN, 1983, p.68).

De maneira um tanto confusa, van Inwagen usa, no Primeiro Argumento Formal, a letra maiúscula J para a constante individual que denota o juiz, bem como T para denotar certo momento do tempo (digo de maneira confusa porque seria muito mais natural usar *j* e *t*, respectivamente):

No Primeiro Argumento Formal e em meu comentário subsequente acerca dele, usarei “ $T_0$ ” para denotar algum instante do tempo arbitrariamente escolhido anterior ao nascimento de J, “ $P_0$ ” para denotar a proposição que expressa o estado do mundo em  $T_0$ , “P” para denotar a proposição que expressa o estado do mundo em T, e “L” para denotar a conjunção de uma única proposição acerca de todas as leis da natureza. Todos esses símbolos são considerados como “designadores rígidos” (VAN INWAGEN, 1983, p.70).

Vejamos agora o Primeiro Argumento Formal:

- (1) Se o determinismo é verdadeiro, então a conjunção de  $P_0$  e L acarreta P.
- (2) Não é possível que J tenha levantado sua mão em T e P seja verdadeira.
- (3) Se (2) é verdadeira, então se J pudesse ter levantado sua mão em T, J poderia ter tornado P falsa.

(4) Se J pudesse ter tornado P falsa, e se a conjunção de  $P_0$  e L acarreta P, então J poderia ter tornado a conjunção de  $P_0$  e L falsa.

(5) Se J pudesse ter tornado a conjunção de  $P_0$  e L falsa, então J poderia ter tornado L falsa.

(6) J não poderia ter tornado L falsa.

(7) Se o determinismo é verdadeiro, J não poderia ter levantado sua mão em T.

Após a apresentação do argumento, van Inwagen prossegue defendendo suas premissas. O que será relevante para os nossos propósitos é a defesa que o autor faz da segunda premissa. Segundo Jonathan Westphal, o argumento de van Inwagen a favor de (2) é falacioso, análogo às falácias modais mencionadas ao longo de toda esta dissertação. Vejamos o argumento que van Inwagen avança a favor de (2). (Cito aqui os argumentos que ele apresentou no artigo original de 1974, bem como no livro de 1983):

Premissa (2). O símbolo “P” é o nosso nome para a proposição que expressa o estado em que o mundo estava de fato em T, um tempo no qual a mão de J não fora levantada. Portanto, *é impossível* [minha ênfase] que P seja verdadeira se a mão de J não fora levantada em T, ou, de fato, se em qualquer caso as coisas fossem, em T, diferentes do modo que elas efetivamente foram (van Inwagen 1983: 70)<sup>16</sup>.

---

<sup>16</sup> O leitor pode comparar com a defesa que van Inwagen fez dessa premissa no artigo original de 1964: “(2) Se J tivesse levantado a sua mão em T, então o mundo estaria em T em um estado diferente daquele que de fato estava. (Conferir a nossa segunda condição do conteúdo do “estado do mundo”). E, portanto, se J tivesse levantado a sua mão em T, alguma contrária de P exprimiria o estado do mundo em T. Deve-se enfatizar que “P” não *significa* “a proposição que exprime o estado do mundo em T”. Antes, “P” *denota* a proposição que exprime o estado do mundo em T. Na terminologia de Kripke, “P” está sendo usada como um *designador rígido*, enquanto “a proposição que exprime o estado do mundo em T” é forçosamente não-rígida” (VAN INWAGEN 1974, p.192).

Qual o argumento de van Inwagen? A sugestão de Westphal é a seguinte. (Em virtude da confusão de van Inwagen mencionada anteriormente, Westphal usa  $J$  para representar a proposição de que o juiz levantou sua mão, e  $j$  para a constante individual que denota o juiz):

O argumento de van Inwagen parece ser o seguinte. O juiz não levantou sua mão em  $t$ , e por isso  $\sim J$  é verdadeira. Mas como van Inwagen está usando os nomes das proposições,  $P$  é verdadeira, e  $P$  implica  $\sim J$ . Assim, de  $P \ \& \ J$  podemos derivar uma contradição, e  $P \ \& \ J$  expressa um estado de coisas impossível. Consequentemente, se  $P$  é verdadeira, então o estado de coisas representado por  $J$  é um estado de coisas impossível. Mas  $P$  é verdadeira. Assim, o estado de coisas representado por  $J$  é impossível (WESTPHAL, 2012, p.38).

Repare, contudo, na conclusão: o estado de coisas representado por  $J$  é *impossível*, se  $P$  é verdadeira; veja bem, é *impossível*. Em outras palavras, é impossível que a proposição de que  $J$  levantou a mão seja verdadeira, se  $P$  é verdadeira. Como escreve Westphal:

A conclusão que van Inwagen está de fato legitimado a extrair é o que chamarei de (2'):  $j$  não levantou sua mão e  $P$  é verdadeira. O ponto é que disso não se segue que "Se (2) é verdadeira, então se  $J$  pudesse ter levantado sua mão em  $T$ ,  $J$  poderia ter tornado  $P$  falsa."

O que há de errado com o que van Inwagen de fato escreveu encontra-se na proposição de que se  $p$  é verdadeira, então  $\sim p$  não pode ser verdadeira. Uma coisa é dizer que (i) é uma verdade necessária que, se  $Tp$ , então  $T\sim p$ . Outra coisa totalmente diferente é dizer que (ii) se  $Tp$ , então é uma verdade necessária que  $F\sim p$ . Pois (i) é verdadeira e (ii) é falsa. E é quase que certo que (ii) não se segue de (i) (WESTPHAL, 2012, p.38).

Para piorar as coisas, o argumento de van Inwagen não é a favor de

(2)  $\sim \diamond (P \ \& \ J)$ ,

mas a favor de

(C)  $P \supset \sim \diamond J$ .

(2) nada mais é do que a segunda premissa que van Inwagen presumivelmente tem de justificar. Porém, van Inwagen apresenta um argumento a favor de (C). Um argumento a favor de (C) não é um argumento a favor da segunda premissa; (C) não é equivalente a (2), pois as condições de verdade de (2) e (C) são obviamente diferentes. Mas como alguém poderia confundir as condições de verdade dessas duas proposições? Considere as mesmas sem os operadores modais:

(i)  $\sim(P \& J)$ .

(ii)  $P \supset \sim J$ .

Sem os operadores modais, as proposições são de fato equivalentes. Com os operadores modais, não. Van Inwagen parece ter cometido um erro elementar ao negligenciar o fato de que há operadores modais operando nessas proposições e gerando proposições mais complexas.

Retornando ao argumento a favor de (C), van Inwagen parece pensar que a seguinte condicional é verdadeira: se  $P$  é verdadeira, então  $\sim P$  não pode ser verdadeira. Ou seja, se  $P$  é verdadeira, então  $\sim P$  tem de ser falsa. Van Inwagen já parece assumir (2), ou seja,  $\sim \diamond (P \& J)$ , e ao considerar a verdade de  $P$ , infere que se  $P$  é verdadeira, não é possível que  $J$  seja verdadeira. Em outras palavras, van Inwagen pensa que o esquema  $(\Box \sim (P \& \sim P) \& P) \supset (P \supset \Box \sim \sim P)$  é verdadeiro. Porém, é claro que esse esquema não é verdadeiro. A propósito, isso nada mais é do que a forma proposicional daquela típica falácia de escopo modal mencionada ao longo desta dissertação.

Antes de prosseguir para a minha resposta à objeção de Westphal, vejamos como a mesma se apresenta com igual força ao Segundo Argumento Formal de van Inwagen. De acordo com Westphal, van Inwagen precisa oferecer um argumento a favor da seguinte premissa: “Não é possível que eu tenha visitado Arcturo em  $T$  e  $P$  seja verdadeira”, em que  $P$  inclui a proposição de que eu não visitei Arcturo. Na verdade, van Inwagen não apresenta qualquer argumento a favor dessa premissa. Porém, se ele fosse apresentar algum, sugere Westphal, o argumento seria este (WESTPHAL, 2012, p.40):

O símbolo “P” é o nosso nome para a proposição que expressa o estado em que o mundo estava de fato em  $t$ , um tempo em que não visitei Arcturo ( $\sim A$ ). Portanto, é *impossível* que P seja verdadeira se eu não visitei Arcturo (A), ou se as coisas fossem de fato de algum modo diferente em  $t$  do modo que elas efetivamente foram

Esse argumento, assim como o anterior, sugere que de  $P \supset \sim A$  se segue que  $A \supset \sim \diamond P$ . Mas é claro que não se segue. Portanto, tanto o primeiro argumento, quanto o segundo, apresentam o problema levantado por Westphal. Na seção seguinte vou esboçar uma resposta a esta objeção. A discussão a seguir se concentrará sobre o Primeiro Argumento Formal. No entanto, ficará evidente como minha resposta é apropriada também à objeção ao Segundo Argumento Formal.

#### 1.4. Seria a objeção de Westphal bem-sucedida?

Em virtude da invalidade do argumento de van Inwagen a favor de (2), Westphal conclui que “todo o ‘Primeiro Argumento Formal’ não é sólido” (WESTPHAL, 2012, 36). Veja bem, *todo argumento não é sólido*. Concordo com ele que o argumento de van Inwagen a favor de (2) seja falacioso. Entretanto, não concordo que, em virtude apenas disso, tenhamos razões para pensar que o argumento não seja sólido.

Primeiramente, Westphal parece cometer um erro elementar. Do fato do argumento a favor da segunda premissa ser inválido, não se segue que “todo o Primeiro Argumento Formal não seja sólido”. Um argumento é sólido sse é válido e tem premissas verdadeiras. Para mostrar que o Primeiro Argumento Formal não é sólido, Westphal teria de mostrar que a segunda premissa é falsa. Ele não mostra isso. Ele apenas mostra que *um* argumento a favor da segunda premissa não é válido. E do fato de mostrar que um argumento a favor de uma proposição não é válido, não se segue que essa proposição seja falsa. Por exemplo, posso apresentar um argumento inválido a favor da proposição de que nenhum solteiro é casado. Obviamente, disso não se segue que a proposição expressa por “nenhum solteiro é casado” seja falsa. Agora suponha que eu tenha um argumento cuja segunda premissa seja a de que nenhum solteiro é casado. Melhor, suponha que eu apresente o seguinte argumento:

- (1) João é solteiro
- (2) Nenhum solteiro é casado

(3) Logo, o João não é casado.

Admita que o João seja realmente solteiro. Ora, este é um argumento sólido. Além de válido, suas premissas são verdadeiras. Só porque apresentei um argumento inválido a favor de (2) não se segue que meu argumento não seja sólido.

Talvez o que Westphal realmente queira dizer é que, uma vez que o argumento apresentado a favor de (2) não é bom, não temos razões para acreditar nessa premissa. Se não temos razões para acreditar em ao menos uma premissa, o argumento não é racionalmente persuasivo. Se o argumento não é racionalmente persuasivo, então não é cogente. Logo, se não temos razões para acreditar em ao menos uma das premissas do Primeiro Argumento Formal, então esse argumento não é cogente. Em termos mais precisos, o que Westphal queria dizer, portanto, é que o Primeiro Argumento Formal não é um argumento cogente.

Van Inwagen apresentou um argumento inválido a favor de (2). Não obstante, disso não se segue que não tenhamos justificção para acreditar em (2). Em primeiro lugar, eu diria que a segunda premissa é uma verdade analítica:

(2) Não é possível que J tenha levantado sua mão em T e P seja verdadeira.

“P”, como diz van Inwagen, é “o nosso nome para a proposição que expressa o estado em que o mundo estava de fato em T, um tempo no qual a mão de J não fora levantada” (1983, p.70). O que a premissa (2) diz é apenas o seguinte:

Não é possível que a proposição de que J não levantou sua mão em T seja verdadeira e o estado de coisas de que ele levantou sua mão em T se verifique.

Analogamente, é impossível que a proposição de que a neve é branca seja verdadeira e o estado de coisas de que a neve não é branca se verifique. Essa parece-me uma verdade analítica. E por “verdade analítica”, não quero dizer que a frase é verdadeira em virtude de seu significado. Estou aqui assumindo a distinção traçada por Paul Boghossian (1997) entre uma noção metafísica e uma noção epistêmica de analiticidade:

*Analicidade metafísica:* Uma frase é analítica se, e somente se, é verdadeira em virtude do seu significado e da sintaxe.

*Analicidade epistemológica:* Uma frase é analítica se, e somente se, a mera compreensão do seu significado e da sua sintaxe for suficiente para nos justificar a tomá-la como verdadeira.

Irei aqui assumir – sem argumentos – que a noção epistêmica de analiticidade é mais plausível que a metafísica. Como aquela premissa do argumento de van Inwagen é uma verdade analítica, temos justificação para acreditar nela. Logo, não se segue que o Primeiro Argumento Formal não seja cogente.

Retornemos ao argumento que eu apresentara anteriormente. Seria cogente aquele argumento a favor da conclusão de que o João não é casado? Acredito que sim. Suponha que eu conheça pessoalmente o João e que ele tenha me dito que ele seja um solteirão; assim, tenho justificação para acreditar na primeira premissa. Quanto à segunda, eu havia apresentado um argumento inválido a favor dela. Porém, esse argumento era, na verdade, inútil. Eu já estava justificado em acreditar na premissa de que nenhum solteiro é casado simplesmente porque eu compreendia o significado e a sintaxe da frase que exprimia aquela proposição.

O mesmo ocorre com o argumento de van Inwagen. O que ele pode alegar é que simplesmente estamos justificados em acreditar na segunda premissa porque ela é uma verdade analítica. Claro, compreender uma verdade analítica, por definição, nos justifica a tomá-la como verdadeira. Portanto, mesmo que o argumento de van Inwagen a favor de (2) seja uma falácia modal, não se segue que não tenhamos justificação para acreditar em (2). Temos justificação para acreditar em (2) precisamente porque (2) é uma verdade analítica.

Assim, se essa resposta esta correta, as duas primeiras formulações de van Inwagen do Argumento da Consequência não precisam ser descartadas. Isso não significa, contudo, que o Argumento da Consequência esteja imune a críticas. No capítulo seguinte, vamos lidar com algumas das principais objeções ao argumento.



### 3. OBJEÇÕES AO ARGUMENTO DA CONSEQUÊNCIA

No primeiro capítulo discuti questões mais gerais relacionadas ao livre-arbítrio, muito embora elas tenham sido importantes para formular o Problema da Compatibilidade. O segundo capítulo consistiu na apresentação do argumento mais importante a favor do incompatibilismo, tal como definido no primeiro capítulo. E eis que finalmente chegamos à parte crucial: as principais objeções à tese que venho até aqui defendendo. Como qualquer argumento filosófico, o Argumento da Consequência enfrenta sérias críticas.

Este capítulo está dividido em três partes principais. Dividi as objeções ao argumento quanto à sua validade, solidez e cogência. Por exemplo, com relação à sua validade, discutirei três contraexemplos à regra Beta. O primeiro desses contraexemplos foi apresentado por David Widerker (1987), em seu artigo “On an Argument for Incompatibilism”. Após oferecer uma réplica apropriada para este contraexemplo, a qual se deve a Timothy O’Connor, lidarei com mais dois contraexemplos bastante persuasivos: o primeiro deles é provavelmente a crítica mais influente ao Argumento da Consequência; trata-se da objeção de McKay & Johnson exposta no célebre artigo “A reconsideration of an argument against incompatibilism”. A segunda é a objeção de Michael Fara em seu artigo “Masked Abilities and Compatibilism”.

No que diz respeito à sua solidez, irei considerar apenas as objeções à tese de que não podemos violar as leis da natureza. Na verdade, são objeções inspiradas na visão humiana acerca de leis da natureza. A primeira delas, consideravelmente famosa, foi formulada por David Lewis em seu artigo “Are we Free to Break the Laws?”. A segunda apresenta de maneira mais precisa como a visão humiana acerca das leis da natureza motiva-nos a rejeitar a ideia de acordo com a qual não podemos violar as leis da natureza.

Por fim, a última parte lida com objeções que atacam a cogência do argumento. Segundo essas críticas, o Argumento da Consequência é uma *ignoratio elenchi*, pois ele seria irrelevante para demonstrar a incompatibilidade entre o determinismo e o livre-arbítrio; aqui, apresentarei as objeções de Ted Warfield (2000) e Scott Sehon (2011).

#### 3.1 Tipos de objeções

De agora em diante, discutiremos a versão modal:

1.	$\square ((P_0 \& L) \supset P)$	determinismo
2.	$\square (P_0 \supset (L \supset P))$	de 1
3.	$\mathbf{N} (P_0 \supset (L \supset P))$	2, regra ( $\alpha$ )
4.	$\mathbf{NP}_0$	premissa
5.	$\mathbf{N} (L \supset P)$	3 e 4, regra ( $\beta$ )
6.	$\mathbf{NL}$	premissa
7.	$\mathbf{NP}$	5 e 6, regra ( $\beta$ )

Quando formulei o Argumento da Consequência no capítulo anterior, afirmei que seus adversários terão de mostrar que ao menos uma de nossas crenças nestas proposições não é plausível:

- 1) A regra de exportação para a lógica modal alética é válida
- 2) A regra ( $\alpha$ ) é válida
- 3) A regra ( $\beta$ ) é válida
- 4)  $\mathbf{NL}$
- 5)  $\mathbf{NP}_0$

Poder-se-ia acrescentar ainda mais uma proposição, a saber,

- 6) O argumento não é uma *ignoratio elenchi*.

Como veremos, algumas objeções admitem a solidez do argumento ao mesmo tempo em que negam que sua conclusão estabeleça a incompatibilidade entre o determinismo e o livre-arbítrio.

Já tratei de objeções comuns à plausibilidade das regras de inferência no capítulo anterior. Penso, aliás, que as objeções à plausibilidade de 1 e 2 foram adequadamente respondidas naquele capítulo. Portanto, deste capítulo irei lidar apenas com as réplicas compatibilistas à plausibilidade de 3, 4, 5 e 6. Começaremos com os contraexemplos à regra beta.

### 3.2 A regra beta e as réplicas compatibilistas

Van Inwagen reconheceu que o ponto fraco do poderoso Argumento da Consequência é a regra beta. A defesa que ele ofereceu a favor dessa regra fora a seguinte:

Devo confessar que minha crença na validade de beta tem apenas duas fontes, uma incomunicável e outra inconclusiva. A primeira fonte é aquela que os filósofos satisfatoriamente chamam de “intuição”... A última fonte é o fato de que não sou capaz de pensar nenhuma instância de beta que tenha, ou poderia possivelmente ter, premissas verdadeiras e conclusão falsa (VAN INWAGEN 1983, pp.97-8).

Peter van Inwagen tem razão num aspecto: a regra beta é – sim – intuitivamente válida. Um indício a favor disso é que, pelo menos até o artigo de David Widerker, os adversários do Argumento da Consequência apenas se contentaram em apontar a possibilidade de contraexemplos à validade de beta; é o caso, por exemplo, do artigo “Selective Necessity and the Free Will Problem” de Michael Slote (1983). Entretanto, como irei lidar com três contraexemplos diretos à validade de beta, deixarei o artigo de Slote de lado.

A primeira reação compatibilista a esse poderoso argumento foi argumentar que ele é inválido porque a explicação compatibilista da capacidade de agir de outro modo é correta. Em virtude da plausibilidade das premissas e da regra beta, se o argumento é inválido, beta teria de ser inválida. Como afirma Vihvelin:

Por exemplo, se “S é capaz de fazer X” significa “se S escolhesse fazer X, S faria X”, então as premissas do argumento são verdadeiras (pois, mesmo que se S escolhesse mudar as leis ou o passado, ela não seria bem-sucedida), mas a conclusão é falsa (dado que o determinismo é consistente com a verdade de condicionais como “se S escolhesse levantar sua mão, ela faria isso” (VIHVELIN, 2011, §5)

Kadri Vihvelin está a apresentar acima a famosa análise condicional. A análise pode ser apresentada mais precisamente da seguinte maneira:

**Análise Condicional:** Um agente tem a capacidade de fazer *A* nas circunstâncias *C* se, e somente se, ela faria *A* se ela escolhesse, nas circunstâncias *C*, fazer *A*.

Vihvelin tem razão ao afirmar que o Argumento da Consequência é inválido se a famosa análise condicional é bem-sucedida. No entanto, a réplica incompatibilista óbvia a esse tipo de objeção consiste em considerar que a regra beta é mais plausível que a análise condicional. Em primeiro lugar, não há qualquer consenso acerca de qual a análise correta do conceito de capacidade mesmo entre os compatibilistas. Por exemplo, compatibilistas como Michael Fara (2008) não aceitam que a análise condicional seja uma análise bem-sucedida do conceito de capacidade. Em segundo lugar, poder-se-ia simplesmente levantar as dificuldades que a análise condicional levanta.

Os contraexemplos mais proeminentes que existem na literatura dizem respeito à *suficiência* da análise. Em seu artigo “Cans without ifs”, Keith Lehrer nos apresentou o seguinte contraexemplo. Suponha que me ofereçam um pote com vários doces e que nesse pote estejam algumas jujubas vermelhas. Como tenho uma aversão patológica a esse doce (digamos, jujubas vermelhas me lembram gotas de sangue), escolhi não pegar um único docinho... Como afirma Lehrer, “é logicamente consistente supor que, se eu tivesse escolhido pegar a jujuba, eu a teria pegado, mas ao não escolher, sou inteiramente incapaz de tocá-la” (LEHRER, 1968, p.32). Ou seja, eu não tenho a capacidade de pegar uma jujuba vermelha (a proposição que representa o *analysandum* é falsa), mas, se eu tivesse escolhido pegar uma, eu a teria pegado (a proposição que representa o *analysans* é verdadeira).

Contraexemplos à *necessidade* da análise não têm sido tão comuns, embora também existam. Consideremos uma versão de um famoso exemplo apresentado por Austin (1958). Suponha que um habilidoso jogador de futebol, digamos, Ronaldinho Gaúcho, erre uma fácil cobrança de pênalti. Uma vez que Ronaldinho escolheu cobrar o pênalti e errou, é falso que ele teria acertado o pênalti se ele tivesse escolhido cobrá-lo. Apesar disso, Ronaldinho ainda tem a capacidade de acertar o pênalti.

Em suma, a tradicional análise condicional do conceito de capacidade apresenta fortes objeções. Obviamente, não quero com isso dizer que não existam versões mais plausíveis dessa análise. Como veremos mais à frente, há análises mais recentes do conceito de capacidade inspiradas na análise condicional; uma delas é precisamente a análise de Michael Fara (2008), a qual apresentará problemas à regra beta de van Inwagen. No entanto, por ora, o argumento apresentado por Vihvelin não é cogente, dado que, como vimos até aqui, a análise condicional é menos plausível que a regra beta.

Uma maneira mais interessante de se argumentar contra a validade de beta é apresentar contraexemplos que não pressuponham a compatibilidade

entre o determinismo e o livre-arbítrio. Este, na verdade, foi o desafio lançado por van Inwagen. Seguindo Crisp & Warfield (2000), irei me referir aos contraexemplos propostos como o Primeiro Desiderato:

PRIMEIRO DESIDERATO: Os contraexemplos propostos à regra beta não devem pressupor a compatibilidade entre o determinismo e o livre-arbítrio.

De agora em diante, os contraexemplos que irei apresentar satisfarão essa condição. O primeiro deles foi apresentado por David Widerker (1987).

### 3.3 O contraexemplo de Widerker

Eis o contraexemplo de Widerker:

Suponha que ao destruir um pedaço de rádio  $r$  antes de  $t_9$ , Sam impeça a emissão de uma partícula subatômica de  $r$  em  $t_9$ . Suponha, além disso, que este seja o único modo pelo qual Sam possa se assegurar que  $r$  não emitirá radiação em  $t_9$ . Finalmente, suponha que Sam seja o único ser senciente que exista ou tenha existido. Seja ‘R’ e ‘S’

R:  $r$  emite em  $t_9$  uma partícula subatômica

S: Sam destrói  $r$  antes de  $t_9$ . (WIDERKER, 1987, p.38).

Esses são os ingredientes adequados para arquitetar um contraexemplo à regra beta:

(P1)  $N\sim R$

(P2)  $N(\sim R \supset S)$

(C)  $NS$

(P1) é verdadeira, pois  $r$  não emite uma partícula em  $t_9$  e Sam não pode assegurar que esse seja o caso. (P2) também é verdadeira; Sam não pode torná-la falsa porque ele teria de tornar a condicional “ $\sim R \supset S$ ” falsa. Mas se Sam tornasse  $S$  falsa,  $\sim R$  não seria verdadeira, pois a única maneira de  $r$  não emitir uma partícula subatômica em  $t_9$  era por meio de sua destruição (a qual

cabia apenas a Sam). Entretanto, a conclusão é falsa: embora S seja verdadeira, é totalmente consistente com o exemplo que Sam seja livre para não destruir r.

Uma das vantagens dessa objeção ao Argumento da Consequência é que Widerker apresenta um contraexemplo direito à regra beta, ao invés de meramente indicar a possibilidade de existir um contraexemplo. Além disso, o contraexemplo de Widerker satisfaz o primeiro desiderato. Agora, teria Widerker apresentado um contraexemplo bem-sucedido, forte o bastante para refutar o Argumento da Consequência?

Penso que não. Na verdade, como Timothy O'Connor corretamente notou (1993), esse contraexemplo dá apenas uma vitória oca aos adversários do Argumento da Consequência. Por quê? Lembremos novamente da regra beta:

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

Segundo O'Connor, no exemplo de Widerker “a proposição correspondente a  $Q$  torna-se verdadeira em um tempo *anterior* ao tempo no qual a proposição correspondente a  $P$  se torna verdadeira” (O'Connor 1993: 209). O problema é que esta manobra é imprópria para persuadir o defensor do incompatibilismo. Qualquer incompatibilista estará primariamente interessado na incompatibilidade entre o determinismo e o livre-arbítrio *do passado ao futuro*. A versão revisada da regra beta proposta por O'Connor é mais do que suficiente para bloquear o contraexemplo de Widerker. O'Connor na verdade apresenta um princípio, que é formulado da seguinte maneira:

BETA':  $\Box (NP \ \& \ N(P \supset Q) \supset NQ)$ , para todo  $P$  e para todo  $Q$  tal que  $Q$  se torna verdadeira depois de  $P$ .

Embora o contraexemplo de Widerker seja decisivo à regra beta tal como formulada por Peter van Inwagen, ele é completamente ineficaz para refutar o princípio reformulado proposto por O'Connor. Todas as inferências relevantes que precisamos de fazer por meio de beta são legitimadas pelo princípio revisto de O'Connor. A resposta de O'Connor tem um *insight* importante, mas ela infelizmente não vai a fundo na defesa da regra beta. Como veremos, ela é ineficaz para lidar com os contraexemplos que apresentarei a seguir.

### 3.4 O contraexemplo de McKay e Johnson

Se a regra beta não é válida – e a aqui veremos boas razões para pensar que não – é algo que não é imediatamente óbvio. O contraexemplo de Widerker foi engenhoso – embora adequadamente respondido por O'Connor. Ademais, o caso de que a regra beta não é imediatamente inválida pode ser compreendido mais claramente quando consideramos o contraexemplo de McKay & Johnson, o que, na minha visão, é o mais persuasivo a favor da invalidade de beta. Ao invés de apresentarem um contraexemplo direto à regra ( $\beta$ ), McKay & Johnson apresentaram um contraexemplo à regra de aglomeração para o operador “N”. E argumentaram que, se aglomeração é inválido, beta também o é, como mostrarei a seguir. Assim, eles não apresentaram um contraexemplo direto, mas um bom argumento para nos persuadir a favor da invalidade de beta. Consideremos a regra de aglomeração.

Aglomeração:  $NP \ \& \ NQ \vdash N(P \ \& \ Q)$

O primeiro ponto a ser destacado é que aglomeração é inválido. Para ver isso, considere uma moeda, de dois lados, em condições normais de uso. Suponha que ninguém tenha lançado essa moeda, mas que poderia lançá-la.

$P$ : a moeda não virou cara.

$Q$ : a moeda não virou coroa.

Considere a proposição expressa pela frase “a moeda não virou cara”. Vimos que “ $NP$ ” abrevia “ $P$  e ninguém tem, nem nunca teve, qualquer escolha sobre se  $P$ ”; e podemos entender “ter escolha sobre se  $P$ ” como “poderia tornar  $P$  falsa”. Seja  $P$  a proposição de que a moeda não virou cara, é verdade que ninguém poderia torná-la falsa, pois ninguém tem a capacidade de, ao arremessar a moeda, fazer com que ela vire cara. E o mesmo acontece com a proposição de que a moeda não virou coroa. Seja  $Q$  a proposição de que a moeda não virou coroa, é verdade que ninguém poderia torná-la falsa, pois ninguém tem a capacidade de, ao arremessar a moeda, fazer com que ela vire coroa. Portanto, “ $NP \ \& \ NQ$ ” é verdadeira.

No entanto, “ $N(P \ \& \ Q)$ ” é falsa. É fácil ver que alguém poderia tornar a proposição expressa pela frase “a moeda não virou cara e não virou coroa” falsa. Poderia torná-la falsa precisamente porque poderia arremessar a moeda e, nesse caso, ela viraria cara ou cora. Ora, a proposição expressa pela frase “a moeda virou cara ou coroa” é a contraditória da proposição expressa por “a moeda não virou cara e não virou coroa”. Portanto, “ $N(P \ \& \ Q)$ ” é falsa. Logo,

a regra de aglomeração para o operador “**N**” é inválida. Considere, finalmente, esta derivação:

- |   |                                |
|---|--------------------------------|
| 1. $NP$                                       | premissa                       |
| 2. $NQ$                                       | premissa                       |
| 3. $\square (P \supset (Q \supset (P \& Q)))$ | premissa                       |
| 4. $N(P \supset (Q \supset (P \& Q)))$        | de 3 e da regra ( $\alpha$ )   |
| 5. $N(Q \supset (P \& Q))$                    | de 1, 4 e da regra ( $\beta$ ) |
| 6. $N(P \& Q)$                                | de 2, 5 e da regra ( $\beta$ ) |

As premissas são verdadeiras. Como vimos, “ $NP$ ” é verdadeira porque a moeda não virou cara e ninguém tem, nem nunca teve, a capacidade de, ao lançar a moeda, fazê-la virar cara. Se “ $NP$ ” é verdadeira, “ $NQ$ ” também o é. 3 é praticamente indisputável. 4 se segue de 3, pela regra ( $\alpha$ ). Por meio da regra ( $\beta$ ) chegamos facilmente a 6, que, como vimos, é falsa. Portanto, o que permitiu, no raciocínio acima, chegar de premissas verdadeiras a uma conclusão falsa foi justamente a regra ( $\beta$ ). Logo, a regra ( $\beta$ ) é inválida.

Irei assumir – sem argumentos – que o argumento de McKay & Johnson é imune à versão revista do princípio beta proposta por O’Connor. O leitor interessado nesse problema achará informativa a discussão de Crisp & Warfield (2000) sobre esse ponto. A questão interessante a ser notada é que há uma característica distintiva no contraexemplo de McKay & Johnson que os faz pensar que o contraexemplo deles é bem-sucedido em persuadir o defensor do Argumento da Consequência. Segundo os autores, no exemplo da partícula de Widerker ocorre um processo essencialmente indeterminístico. E se esse é o caso, o contraexemplo de Widerker está em sérios apuros, de acordo com McKay & Johnson. “Assumir que o mundo é indeterminístico é um modo problemático de argumentar contra beta porque essa regra é usada para extrair as consequências do determinismo” (MCKAY & JOHNSON, 1996, p.118). Os autores sugerem que, se todo contraexemplo a beta tiver de ser indeterminístico, então uma versão revisada da regra seria suficiente para manter o argumento de van Inwagen. A versão revisada da regra beta é o que eles chamam de delta:

**(D):**  $\square ((P_0 \& L) \supset P), NP, N(P \supset Q) \vdash NQ$

McKay & Johnson estão obviamente certos em notar que não há quaisquer contraexemplos indeterminísticos à regra delta. Crisp & Warfield afirmam que a sugestão de McKay & Johnson é um segundo desiderato que os



contraexemplos à regra beta deveriam satisfazer para ameaçar o Argumento da Consequência (CRISP & WARFIELD, 2000, p.180):

SEGUNDO DESIDERATO: Os contraexemplos propostos à regra beta não devem pressupor a verdade do indeterminismo.

Entretanto, admitindo o primeiro desiderato, assim como o segundo, surge um problema para o contraexemplo de McKay & Johnson. Os autores afirmam que o contraexemplo deles é efetivo não só à regra beta, mas também à regra delta. Contudo, se levarmos em conta ambos os desiderata, o contraexemplo não será bem-sucedido. A réplica que apresento foi oferecida por Crisp & Warfield (2000). Antes de apresentá-la, vejamos novamente o primeiro desiderato:

PRIMEIRO DESIDERATO: Os contraexemplos propostos à regra beta não devem pressupor a compatibilidade entre o determinismo e o livre-arbítrio.

A ideia de Crisp e Warfield é a de que o contraexemplo da moeda não se adequa bem a ambos os desiderata. Segundo eles, McKay e Johnson enfrentam um dilema, que é o seguinte. O exemplo ocorre num mundo determinista ou não.

Suponha que não. Portanto, o caso ocorre num mundo indeterminista (a definição do conceito de indeterminismo foi apresentada no primeiro capítulo, sendo apenas a negação do determinismo). Assim, se o caso ocorre num mundo indeterminista, o contraexemplo pode ser bloqueado tal como o contraexemplo de Widerker poderia ser bloqueado, a saber, usando a regra delta; esta, aliás, foi a sugestão de McKay e Johnson.

Suponha que o mundo seja determinista. Se esse é o caso, o contraexemplo viola o primeiro desiderato. O cenário introduzido por McKay e Johnson depende da suposição de que a moeda não foi arremessada, mas poderia ser. Ora, como o mundo em questão é determinista, isso é o mesmo que assumir que eu poderia ter agido diferente daquilo que agi num mundo determinista. Agir diferente daquilo que se agiu, mesmo que o mundo em questão seja determinista, é nada mais do que o compatibilismo. E se não é nada mais do que o compatibilismo, o contraexemplo de McKay e Johnson viola o primeiro desiderato.

Em suma, mesmo que o contraexemplo seja bem-sucedido em mostrar que a regra beta seja inválida, ele não é persuasivo, dado que não satisfaz os desiderata.

Além da réplica de Crisp e Warfield que apresentei acima, há adicionalmente duas respostas principais ao contraexemplo de McKay e Johnson. Em “Free Will Remains a Mystery”, van Inwagen concedeu a invalidade da regra ( $\beta$ ) e propôs substituir o operador “N” por “N\*”, em que “N\*P” diz “P e ninguém pode, nem nunca poderia, fazer alguma coisa tal que, se fizesse, P poderia ser falsa”. Neste caso, segundo van Inwagen, o princípio de aglomeração é válido para “N\*” e o contraexemplo não funcionaria. Se a resposta de van Inwagen estiver correta, a objeção ao princípio de aglomeração para refutar a regra ( $\beta$ ) não se aplica a “N\*”.

Finch e Warfield (1998) também propuseram uma nova forma de tornar o Argumento Modal da Consequência imune à objeção supracitada, substituindo a regra ( $\beta$ ) pela regra ( $\beta$ )2.

$$(\beta)2: NP, \Box (P \supset Q) \vdash NQ$$

Esta substituição nos daria o seguinte argumento:

- |    |   |                    |
|----|---|--------------------|
| 1. | $\Box((P_0 \ \& \ L) \supset P)$            | suposição          |
| 2. | $N(P_0 \ \& \ L)$                           | premissa           |
| 3. | NP  | 1, 2 ( $\beta$ ) 2 |
| 4. | $\Box((P_0 \ \& \ L) \supset P) \supset NP$ | 1-3, $\supset$     |

Ao substituir a regra ( $\beta$ ) por ( $\beta$ )2, observa-se que esta última não nos permite derivar aglomeração:

- |    |  |                            |
|----|--|----------------------------|
| 1. | NP   | premissa                   |
| 2. | NQ   | premissa                   |
| 3. | $\Box(P \supset (Q \supset (P \ \& \ Q)))$ | premissa                   |
| 4. | $N(Q \supset (P \ \& \ Q))$                | de 1 e 3 por ( $\beta$ ) 2 |

A partir de 4, só seria possível derivar  $N(P \ \& \ Q)$  caso admitíssemos a regra ( $\beta$ ) original. No entanto, na reformulação de Finch e Warfield, o Argumento Modal da Consequência não precisa da regra ( $\beta$ ) original. E como a

regra ( $\beta$ )<sup>2</sup> não nos permite derivar aglomeração, a nova formulação é imune à objeção de McKay e Johnson.

De agora em diante, será importante ter em mente tanto as respostas de Crisp e Warfield (2000) quanto a de (1998) Finch e Warfield. Acredito que essas duas respostas sejam bem-sucedidas em bloquear o contraexemplo apresentado por Michael Fara, como mostrarei a seguir.

### 3.6. O contraexemplo de Michael Fara

Ao contrário do exemplo de McKay e Johnson, Fara apresenta um contraexemplo direto à regra ( $\beta$ ). O que Fara propõe é empregar a noção de máscara – muito comum com relação a disposições – a capacidades. Por exemplo, estamos propensos a acreditar que uma xícara de porcelana tem a propriedade disposicional de se quebrar. Mas podemos simplesmente mascarar a disposição da xícara se quebrar ao revesti-la, digamos, com algum material protetor, um isopor, etc. Se fizermos isso, a xícara não perde a disposição de se quebrar, ocorre apenas que sua disposição foi mascarada. Fara (2008) argumenta que o mesmo acontece com relação ao conceito de capacidade. Algumas capacidades, assim como algumas disposições, são mascaradas. A definição proposta por Fara do conceito de capacidade mascarada é a seguinte:

Uma capacidade de um agente para fazer  $A$  nas circunstâncias  $C$  é mascarada sse (FARA, 2008, p.848)

- (i) o agente tenta fazer  $A$
- (ii) as circunstâncias  $C$  ocorrem
- (iii) o agente retém sua capacidade enquanto tenta fazer  $A$ ; porém
- (iv) o agente não é bem-sucedido em fazer  $A$ .

Irei assumir sem argumentos que a análise de Fara é bem-sucedida; e que ele está certo ao empregar a noção de máscara com relação a disposições. Motivado por isso, Fara oferece a seguinte análise do conceito de capacidade:

**A análise disposicional:** um agente tem a capacidade para fazer  $A$  nas circunstâncias  $C$  se, e somente se, ela tem a disposição para fazer  $A$ , quando, nas circunstâncias  $C$ , ela tenta fazer  $A$ .

A análise de Fara é motivada pela tradicional análise condicional do conceito de capacidade, que discutimos no início deste capítulo. Mas deve ser óbvio ao leitor que as análises não são equivalentes. Um dos motivos pelos quais Fara rejeita a análise condicional é o fato dela não acomodar casos nos quais as capacidades são mascaradas. Podemos deixar isso de lado, em todo o caso, pois o que será relevante aqui é entender como a análise de Fara será saliente para formular um contraexemplo à regra beta.

Considere então o seguinte exemplo. Suponha que o João seja um bom golfista, e que tipicamente consegue acertar a bolinha de golfe no buraco correto. Assim, o João tem a capacidade de acertar a bolinha de golfe no buraco. No entanto, se estiver ventando muito, o João é simplesmente incapaz de realizar a mesma proeza que sempre realiza num dia sem muito vento. Suponhamos que o João vá jogar golfe justamente num dia em que há muito vento. O João fracassará em acertar a bolinha de golfe no buraco correto. No entanto, ele não perde a sua capacidade de acertá-la; apenas ocorre que a sua capacidade foi mascarada:

- (i) o João tenta acertar a bolinha de golfe no buraco
- (ii) o João está num campo de golfe
- (iii) o João retém sua capacidade enquanto tenta acertar a bolinha de golfe no buraco; porém
- (iv) o João não é bem-sucedido em acertar a bolinha de golfe no buraco.

Seja  $P$  a proposição de que está ventando muito e  $Q$  a proposição de que a bolinha de golfe cai no buraco, temos o seguinte contraexemplo. “ $\mathbf{N}P$ ” é verdadeira precisamente porque venta muito e ninguém tem, nem nunca teve, qualquer escolha sobre se está ventando muito. “ $\mathbf{N}(P \supset Q)$ ” é verdadeira porque, se venta muito, então a bolinha de golfe não cai no buraco, e ninguém tem, nem nunca teve, qualquer escolha sobre a proposição expressa pela frase “se venta muito, então a bolinha de golfe não cai no buraco”. Mas “ $\mathbf{N}Q$ ” é falsa. O João tem a capacidade de tornar a proposição de que a bolinha de golfe não cai no buraco, pois ele tem a capacidade de acertar a bolinha de golfe no buraco; o que ocorre é apenas que sua capacidade foi mascarada. Logo, a regra ( $\beta$ ) é inválida.

Seria o contraexemplo de Fara bem-sucedido? Penso que não. Em primeiro lugar, veremos que o contraexemplo não é bem-sucedido para refutar a regra beta 2.

( $\beta$ )2:  $NP, \Box(P \supset Q) \vdash NQ$

A razão é bem simples. A proposição de que está ventando muito não *acarreta* a proposição de que a bolinha cai no buraco. Veja que a condicional da regra beta 2 é a condicional estrita; é totalmente plausível pensar que há um mundo possível no qual venta muito e, no entanto, o João é bem-sucedido em acertar a bolinha no buraco. Fara teria de apresentar um contraexemplo no qual a condicional não fosse a condicional estrita, o que ele não faz. Claro, o contraexemplo é bem-sucedido em refutar a regra beta original, mas parece não funcionar quando aplicado à reformulação de Finch e Warfield.

Em segundo lugar, e mais importante, o contraexemplo não é bem-sucedido em satisfazer os dois desiderata descritos por Crisp e Warfield. O mesmo argumento apresentado contra McKay e Johnson pode ser formulado aqui. O exemplo de Fara ocorre num mundo determinista ou não.

Suponha que não. Se não ocorre num mundo determinista, então ocorre num mundo indeterminista. Portanto, o exemplo falha em satisfazer o segundo desiderato. E podemos bloquear o contraexemplo simplesmente ao utilizar a regra delta, ao invés da regra beta.

Suponha que sim. Logo, o exemplo ocorre num mundo determinista. Contudo, o cenário introduzido por Fara depende da suposição de que a capacidade do agente de fazer com que a bolinha caia no buraco seja mascarada. Ora, mas pela própria definição do conceito de capacidade mascarada, o agente tem de reter sua capacidade, embora não seja bem-sucedido em realizar a ação. Assim, temos um cenário de um mundo determinista no qual o agente retém a sua capacidade, o que equivale a assumir o compatibilismo. Assim, o contraexemplo viola o primeiro desiderato.

Em suma, embora a análise do conceito de capacidade proposta por Michael Fara seja iluminante, precisamente por permitir que capacidades sejam mascaradas, ela falha em ser usada para formular um contraexemplo bem-sucedido às novas formulações do Argumento da Consequência. Acredito que as duas formulações sejam suficientes para lidar com as dificuldades originais relacionadas à validade do argumento:

(1) Formulação de Finch e Warfield:

1.  $\Box((P_0 \ \& \ L) \supset P)$  suposição

- |    |   |                    |
|----|---|--------------------|
| 2. | $N(P_0 \& L)$                           | premissa           |
| 3. | $NP$                                    | 1, 2 ( $\beta$ ) 2 |
| 4. | $\Box((P_0 \& L) \supset P) \supset NP$ | 1-3, $\Box$        |

(2) Formulação minha inspirada em McKay, Johnson, Crisp e Warfield

( $\alpha$ )  $\Box P \vdash NP$

(D),  $\Box((P_0 \& L) \supset P), N(P \supset Q), NP \vdash NQ$

- |    |   |                               |
|----|---|-------------------------------|
| 1. | $\Box((P_0 \& L) \supset P)$            | premissa                      |
| 2. | $\Box(P_0 \supset (L \supset P))$       | 1, exportação                 |
| 3. | $N(P_0 \supset (L \supset P))$          | 2, pela regra ( $\alpha$ )    |
| 4. | $NP_0$                                  | premissa                      |
| 5. | $N(L \supset P)$                        | <b>1, 3, 4 pela regra (D)</b> |
| 6. | $NL$                                    | premissa                      |
| 7. | $NP$                                    | <b>1, 5, 6 pela regra (D)</b> |
| 8. | $\Box((P_0 \& L) \supset P) \supset NP$ | 1-7 $\Box$                    |

Acredito que a discussão até aqui tenha dado conta das objeções quanto à validade do argumento. A seção seguinte discute sua solidez.

Primeiramente, irei reconstruir o argumento de David Lewis, tal como apresentado pelo mesmo. O problema é que a objeção de David Lewis é muito imprecisa. Então, para motivar sua posição, apresentarei a objeção de Beebee e Mele (2002). Como veremos, a motivação que os autores apresentam, embora seja um problema para os incompatibilistas, mina a principal resposta traçada por Lewis. Terminarei a seção após apresentar as razões pelas quais o argumento de Beebee e Mele não é cogente.

### 3.6 Somos livres para violar leis da natureza? A objeção de David Lewis

A resposta de David Lewis ao Argumento da Consequência é elegante, não obstante controversa: certas ações nossas podem violar uma lei da natureza. Lewis argumenta que um compatibilista não tem de aceitar a premissa segundo a qual não podemos tornar uma lei da natureza falsa se fizermos uma distinção entre uma tese forte e outra fraca do compatibilismo. Para tornar a objeção de Lewis mais clara, apresentarei o Argumento da Consequência reconstruído pelo próprio.

Suponhamos que o determinismo seja verdadeiro, e que eu tenha levantado minha mão. Um compatibilista terá de afirmar que esta é uma ação livremente realizada, ao mesmo tempo em que o determinismo é verdadeiro. Se minha ação foi livre, então poderia ter agido de outro modo; eu poderia, por exemplo, não ter levantado minhas mãos. No entanto, se o determinismo é verdadeiro, “ $P_0$ ” e “ $L$ ” conjuntamente determinam minha ação, e conjuntamente contradizem a proposição segundo a qual não levantei minhas mãos. Mas se não tivesse levantado minhas mãos, teríamos uma de três coisas: *ou* (a) contradições teriam de ser verdadeiras, *ou* (b) “ $P_0$ ” não seria verdadeira, *ou* (c) “ $L$ ” não seria verdadeira. Então, se eu pudesse não ter levantado as mãos, realizaria algo realmente impressionante: ou tornaria contradições verdadeiras, ou teria a capacidade de mudar o passado, ou a capacidade de mudar ou violar as leis da natureza. Como estas três opções são absurdas, segue-se, por *reductio ad impossibile*, que eu não poderia ter levantado minhas mãos.

Em virtude desse argumento, um compatibilista estaria obrigado a aceitar a seguinte consequência: se eu tivesse agido de outro modo (por exemplo, ao não levantar minhas mãos), então alguma lei da natureza seria violada. Eis o que diz Lewis: “‘Isto é dizer’, parafraseando meu oponente, ‘que você é capaz de violar uma lei da natureza. E com tão pouco esforço! Um poder realmente maravilhoso. Você também pode entortar colheres?’” (Lewis 1983: 292). Não obstante, a paráfrase do oponente feita por Lewis não está completamente correta. Isso porque há uma distinção a ser feita entre uma tese forte e uma tese fraca. A tese forte sem dúvida tem uma consequência inacreditável, semelhante à paráfrase de Lewis de seu oponente. Entretanto, um compatibilista não tem de aceitar a tese forte. Ele pode aceitar simplesmente uma tese fraca. Lewis traça uma distinção entre duas contrafactuais.

(C1) Se eu não tivesse levantado minha mão,  $L$  não seria verdadeira.

(C2) Se eu não tivesse levantado minha mão, isso seria ou causaria um evento que acarretaria não- $L$ .

Um compatibilista que apenas aceita a tese fraca aceita a primeira contrafactual (C1), mas não precisa aceitar C2. Lewis então faz uma distinção entre duas afirmações acerca de capacidades:

(Ca1) Eu tenho a capacidade de fazer algo tal que, se eu fizesse, *L* não seria verdadeira.

(Ca2) Eu tenho a capacidade de fazer algo tal que, se eu fizesse, minha ação seria ou causaria um evento que acarretaria não-*L*.

Como um compatibilista que aceita C1, Lewis está comprometido com a verdade da primeira afirmação de capacidade (Ca1). Porém, não está comprometido com a verdade de (Ca2). E, segundo Lewis, é (Ca2), não (Ca1), que descreve uma capacidade incrível. Uma coisa é dizer que, se eu não tivesse levantado minhas mãos, seria a mesma coisa que arremessar uma pedra na janela e minha pedra causar o evento da janela ter sido quebrada. Outra coisa é dizer que, se eu não tivesse levantado minhas mãos, eu violaria uma lei da natureza. E a mesma ideia que Lewis apresenta a favor da ideia de que temos escolha sobre as leis da natureza pode ser aplicada com relação ao passado (veja Vihvelin 2011).

Para você ver como o argumento de Lewis ataca o Argumento da Consequência, precisamos considerar novamente as premissas 5 e 6 do Primeiro Argumento Formal:

(5) Se *J* pudesse ter tornado a conjunção de  $P_0$  e *L* falsa, então *J* poderia ter tornado *L* falsa.

(6) *J* não poderia ter tornado *L* falsa.

Para Lewis, a resposta apropriada ao Argumento da Consequência depende de qual é a noção de “ser capaz de tornar *P* falsa” empregada por van Inwagen. Se van Inwagen assume o sentido fraco, então Lewis rejeita a premissa 6: podemos tornar *L* falsa, pois, como ele diz, isso não significa dizer que nossa ação causou o evento que acarreta não-*L*, mas apenas que uma lei estaria violada de antemão. Por outro lado, se o sentido empregado por van Inwagen é o sentido forte, então Lewis rejeita 5, pois tudo o que se segue da antecedente da condicional da premissa 5 é que *j* poderia ter tornado *L* falsa no sentido fraco; não se segue da antecedente que ele poderia ter tornado *L* falsa no sentido forte.



O que Lewis faz é uma distinção entre um poder *causal* sobre as leis da natureza (sentido forte) e um poder *contrafactual* sobre tais leis (sentido fraco). O problema é que surge uma tensão inicial aqui, como foi notado por Laura Ekstrom (1998): como Lewis pode defender tal distinção ao mesmo tempo em que defende uma explicação contrafactual da causação? Parece que, se um agente tem o primeiro tipo de capacidade, a capacidade contrafactual, isso significa que também tem a capacidade causal.

Lewis tem uma resposta a essa dificuldade que repousa em sua explicação dos eventos. Entretanto, essa saída enfrenta sérias dificuldades, como argumentou Ekstrom (1998, pp.336-40). Não entrarei em detalhes porque discutirei uma objeção que tornará a ideia lewisiana mais precisa: como veremos, a concepção humiana acerca de leis da natureza – que Lewis aceita – irá minar sua distinção entre a tese forte e a tese fraca.

### 3.7 Compatibilismo humiano

Vejamos a seguinte a passagem:

Considere Fred, que está indeciso em *t* sobre se ele irá ou não comer torradas no café da manhã no dia seguinte, e suponha que Fred de fato *irá* comer torradas amanhã de manhã. É logicamente impossível que este fato ocorra e que Fred não tome o café da manhã. Mas essa impossibilidade não impõe qualquer limite à sua capacidade para não tomar o café da manhã; e não acarreta que não depende de Fred se ele irá ou não evitar o café da manhã, ou que ele não poderia agir de outro. Fred *não irá* agir de outro modo, mas isso não acarreta que ele *não pode* agir de outro modo (BEEBEE & MELE 2002, p. 207).

Beebee e Mele estão chamando atenção para o fato de que, mesmo que seja um fato acerca do futuro, por exemplo, que o Fred irá tomar café da manhã, a proposição que descreve esse fato *não acarreta* a proposição de que Fred não é capaz de agir de outro; apenas acarreta a proposição de que Fred não irá tomar o café da manhã. Mas, é claro, dizer que Fred não irá tomar o café da manhã é algo completamente diferente de dizer que Fred não era capaz de não tomar o café da manhã.

O que foi dito acima diz respeito acerca de fatos sobre o futuro. Mas assumamos que o humianismo acerca das leis da natureza seja verdadeiro e que o determinismo seja verdadeiro. Para um humiano, o argumento que se aplica acima, *mutatis mutandis*, aplica-se aqui:

- (A) Considere Fred, que está indeciso em  $t$  sobre se ele irá ou não comer torradas no café da manhã no dia seguinte, e suponha que Fred de fato *irá* comer torradas amanhã de manhã. Não porque é um fato acerca do futuro que Fred irá comer torradas, mas porque o determinismo acarreta que Fred irá comer torradas. É logicamente impossível que esse fato ocorra e que Fred não tome o café da manhã. Mas essa impossibilidade não impõe qualquer limite à sua capacidade para não tomar o café da manhã; e não acarreta que não depende de Fred se ele irá ou não evitar o café da manhã, ou que ele não poderia agir de outro modo. Fred *não irá* agir de outro modo, mas isso não acarreta que ele *não pode* agir de outro modo.

Se o determinismo é verdadeiro, o passado mais as leis da natureza acarretam que Fred irá tomar café da manhã no dia seguinte. Porém, disso não se segue que ele não seja capaz de não tomar o café da manhã; apenas se segue que Fred irá tomar o café da manhã. Tal como fatos sobre o futuro não impõe restrições à capacidade de Fred, o determinismo analogamente não impõe.

Nós podemos apresentar o mesmo argumento de outra maneira. Quando dizemos que alguém é capaz de agir de outro, podemos nos perguntar se há quaisquer mundos possíveis num modo particular de mundos possíveis no qual esse agente *age de outro modo*. A sugestão do defensor do Argumento da Consequência seria a seguinte: os mundos nos quais o agente age de outro modo têm de ser mundos com o mesmo passado que o nosso. Além disso, seria plausível dizer que a classe deveria incluir também mundos que têm as mesmas leis da natureza que o nosso. Assim, o defensor do Argumento da Consequência concluiria que o agente em questão não poderia ter agido de outro modo. No entanto, Beebe e Male afirmam que essa suposição adicional, a qual diz respeito às leis da natureza, é totalmente desmotivada para um humiano. Precisamente porque as leis da natureza são regularidades, mesmo que o passado seja o mesmo, o mundo poderia ter seguido sua história de diferentes maneiras. Retornando ao exemplo de Fred, um humiano diria que Fred poderia, sim, não ter tomado café da manhã, pois as leis da natureza não *garantem* que ele irá tomar o café da manhã.

Embora seja mais plausível, o compatibilismo humano esboçado por Beebe e Mele também minam a distinção traçada por Lewis entre as teses forte e fraca. Como eles dizem:

Em contrapartida, alguém poderia questionar que a perspectiva humana esboçada acima tem como consequência não apenas que Fred é capaz de tomar café da manhã no dia seguinte na medida em que fazer isso é consistente com o seu mundo passado (e as leis da lógica), mas também que ele é capaz, neste mesmo sentido, de mover seu braço mais depressa do que a luz [...] Este é um questionamento legítimo, porque a perspectiva humana tem de fato essas consequências. A conclusão disso é que a motivação que oferecemos para a tese fraca de Lewis simultaneamente mina sua negação da tese forte (BEEBEE & MALE, 2002, p.212).

A posição de Lewis, também conhecida como compatibilismo de milagre local (FISCHER, 1994, p.69), está assim em uma posição delicada. Beebe e Male esboçaram o compatibilismo humano para oferecer suporte à tese fraca. Mas isso também destrói uma das plausibilidades do compatibilismo de milagre local, a saber, a de que nunca temos capacidades fortes, no sentido de viajar mais depressa do que a luz. O problema acima, claro, é um problema para o compatibilismo de milagre local e um problema, portanto, para Lewis. Não é meu objetivo aqui defendê-lo. E o problema que o compatibilismo de Lewis enfrenta não é um problema para o compatibilismo humano que tem sido esboçado até aqui. Uma vez que o compatibilismo de milagre local não é um problema para a tese que venho defendendo, resta-me apenas lidar com o compatibilismo humano sustentado por Beebe e Male.

Em primeiro lugar, o compatibilismo humano depende da suposição de que as leis da natureza são, grosso modo, regularidades. Assim, dada a definição de cogência que apresentei anteriormente, o argumento esboçado por Beebe e Male não é cogente, dado que não é racionalmente persuasivo; afinal, depende de uma suposição altamente controversa acerca do que são leis da natureza. Na melhor das hipóteses, a perspectiva endossada por Beebe e Male os justifica a recusar a aceitação da premissa de acordo com a qual as leis da natureza não cabem a nós, mas não pode ser usado como um argumento persuasivo a favor da rejeição dessa premissa.

Segundo, e mais importante, penso que o argumento por eles apresentado tem uma conclusão contra-intuitiva. Permita-me apresentar o argumento novamente:

- (A) Considere Fred, que está indeciso em  $t$  sobre se ele irá ou não comer torradas no café da manhã no dia seguinte, e suponha que Fred de fato *irá* comer torradas amanhã de manhã. Não porque é um fato acerca do futuro que Fred irá comer torradas, mas porque o determinismo acarreta que Fred irá comer torradas. É logicamente impossível que esse fato ocorra e que Fred não tome o café da manhã. Mas essa impossibilidade não impõe qualquer limite à sua capacidade para não tomar o café da manhã; e não acarreta que não depende de Fred se ele irá ou não evitar o café da manhã, ou que ele não poderia agir de outro modo. Fred *não irá* agir de outro modo, mas isso não acarreta que ele *não pode* agir de outro modo.

Vejamos agora o mesmo argumento, mas num caso diferente:

- (A') Considere Fred, que está indeciso em  $t$  sobre se ele irá ou não torturar crianças por pura diversão no dia seguinte, e suponha que Fred de fato *irá* torturar essas crianças amanhã. O determinismo acarreta que Fred irá torturar as crianças, embora ele não tenha decidido fazer isso. É logicamente impossível que esse fato ocorra e que Fred não torture as crianças. Mas essa impossibilidade não impõe qualquer limite à sua capacidade para não torturar as crianças; e não acarreta que não depende de Fred se ele irá ou não torturá-las por pura diversão, ou que ele não poderia agir de outro modo. Fred *não irá* agir de outro modo, mas isso não acarreta que ele *não pode* agir de outro modo.

Minha intuição aqui é a de que, precisamente porque é um fato que Fred *irá* torturar crianças (apesar de não ter decidido fazer isso), ele não é moralmente responsável por torturar as crianças. O argumento que posso apresentar a favor disso é o de que Fred não era capaz de agir de outro modo, dado que era um fato que ele iria torturar crianças por pura diversão, mesmo não tendo decidido fazer coisa alguma. Na perspectiva de Beebe e Male, Fred poderia ser moralmente responsável por suas ações, apesar de tudo isso, pois, segundo eles, ele tem a capacidade de agir de outro modo, apesar de ser um fato que *ele não irá agir de outro modo*.

Certamente estou apelando a intuições, mas isso não é um problema. O que procurei fazer foi apenas desmotivar o compatibilismo humiano. Parece-me que essa tese tem consequências contra-intuitivas, como mostrei acima, e que, portanto, não temos razões para aceitar.

Assim encerro a seção relacionada às objeções que atacam a solidez do Argumento da Consequência. As seções seguintes lidam com duas objeções que atacam a cogência do argumento. A primeira delas é apresentada por Scott Sehon (2011) e a segunda por Ted Warfield (2000).

### 3.8 A objeção de Sehon ao argumento da consequência

A objeção de Scott Sehon à versão modal do Argumento da Consequência resume-se basicamente a um modo de ler a caixa em “ $\Box(P_0 \& L \supset P)$ ” (a concepção de determinismo até aqui empregada), a qual deveria supostamente expressar uma concepção razoável de determinismo. Sehon lê a caixa como a necessidade lógica. A leitura feita pelo filósofo é a de que “ $\Box$ ” representa a necessidade lógica. De acordo com Sehon, esta concepção de determinismo acarreta que um Deus Intervencionista (doravante IG) é logicamente impossível.

Segundo a concepção usual de um IG, Deus tem a capacidade de intervir na ordem natural e, por exemplo, milagrosamente transformar água em vinho (cf. SEHON, 2011, p.31). Isto quer dizer que, necessariamente, se um IG existe, então  $\Diamond(P_0 \& L \& \sim P)$ . Sehon argumenta que, se aceitarmos que é logicamente possível que um IG exista, então teremos de aceitar que o determinismo, tal como entendido por van Inwagen, é falso. Seu argumento pode ser apresentado do seguinte modo:

- |    |   |                      |
|----|---|----------------------|
| 1. | $\Box(IG \supset \Diamond(P_0 \& L \& \sim P))$ | Premissa             |
| 2. | $\Diamond IG$                                   | Premissa             |
| 3. | $\Diamond(P_0 \& L \& \sim P)$                  | de 1&2, assumindo S4 |
| 4. | $\sim \Box(P_0 \& L \supset P)$                 | de 3                 |

Assim, se assumimos a ideia, inicialmente plausível, de que um IG é logicamente possível, bem como o sistema S4 na lógica modal alética (o que é perfeitamente aceitável), teremos um sério problema para a concepção de determinismo tal como empregada no Argumento da Consequência.

Contudo, é importante notar que, no argumento acima, estamos lendo as caixas e os diamantes como a necessidade e possibilidade lógicas. Ora, mas e se a concepção de determinismo fosse entendida aqui não como envolvendo a necessidade lógica (como Sehon faz), mas, ao invés, a metafísica? Neste caso, a conclusão do argumento de Sehon não seria a negação do determinismo, pois

$$(L) \sim \Box_l(P_0 \& L \supset P)$$

e

$$(M) \Box_m(P_0 \& L \supset P)$$

obviamente não são contraditórias (em que “ $\Box_m$ ” representa a necessidade metafísica e “ $\Box_l$ ” representa a necessidade lógica). Portanto, ao contrário do que afirma Sehon, a possibilidade lógica de um IG é, sim, compatível com a maneira de van Inwagen entender o determinismo. A única coisa que precisamos especificar é que “ $\Box$ ” representa a necessidade metafísica.

*Primeira réplica:* Uma objeção que se poderá fazer é a de que (L) e (M) são, afinal de contas, contraditórias, argumentando da seguinte maneira. Assuma a tese de que a possibilidade lógica implica a possibilidade metafísica. Assim, conseguiríamos mostrar que (L) e (M) são contraditórias:

- |    |   |          |
|----|---|----------|
| 1. | $\Diamond_l(P_0 \& L \& \sim P)$  | premissa |
| 2. | $\Diamond_l(P_0 \& L \& \sim P) \supset \Diamond_m(P_0 \& L \& \sim P)$ | premissa |
| 3. | $\Diamond_m(P_0 \& L \& \sim P)$  | de 1 & 2 |
| 4. | $\sim \Box_m(P_0 \& L \supset P)$                                       | de 3     |

(4) é a contraditória de (M). Portanto, ler a caixa de ‘ $\Box(P_0 \& L \supset P)$ ’ como a necessidade metafísica não salvaria o Argumento da Consequência da objeção de Sehon.

A réplica a esta objeção é a de que (2) é implausível porque é incompatível, sem razão para tal, com perspectivas filosóficas independentes como, por exemplo, o essencialismo. Esta é a tese de que existem possibilidades lógicas – por exemplo, Sócrates ser filho de pais diferentes – que não são possibilidades metafísicas (assumindo a tese da necessidade da origem). Deste modo, um essencialista rejeita (2). E uma vez que não vemos razão para Sehon querer que seu argumento seja incompatível com o essencialismo, não vemos razão para aceitar essa objeção.

*Segunda réplica:* Talvez seja possível reconstruir o raciocínio de Sehon da seguinte maneira:

1.  $\Box_1(IG \supset \Diamond_m(P_0 \& L \& \sim P))$  premissa
2.  $\Diamond_1 IG$  premissa
3.  $\Diamond_1 \Diamond_m(P_0 \& L \& \sim P)$  de 1 e 2
4.  $\Diamond_m(P_0 \& L \& \sim P)$  de 3

A ideia aqui é a de que a afirmação de que é logicamente possível que um IG acarrete a possibilidade metafísica da falsidade do determinismo (premissa (1)). Isto é muito plausível. Se um IG efetivamente existe, então ele tem a capacidade, digamos, para transformar água em vinho. Entretanto, isso significa que é metafisicamente possível que a água seja transformada em vinho, violando portanto o determinismo.

O problema é o passo (4). S4 não ajuda aqui porque estamos lidando com dois tipos diferentes de possibilidade. Alguém poderia estipular que o movimento é aceitável até que se prove o contrário. Pois bem, considere isto. Parece logicamente possível que seja metafisicamente possível que Wittgenstein seja uma raquete de tênis. Por quê? Porque tudo o que é preciso para uma dada proposição  $P$  ser logicamente possível é que  $P$  não seja uma contradição. Ora, não é uma contradição que seja metafisicamente possível Wittgenstein ser uma raquete de tênis. Filósofos essencialistas, no entanto, não aceitariam que é metafisicamente possível que Wittgenstein seja uma raquete de tênis. Portanto, novamente, o essencialismo é incompatível com essa versão revisada do argumento.

Minha conclusão é a de que essa réplica não é plausível precisamente porque é incompatível com o essencialismo, e parece não haver qualquer razão independente para nós rejeitarmos o essencialismo apenas para aceitar o argumento de Sehon.

*Terceira réplica:* Assuma que o determinismo deva ser entendido de modo a envolver a necessidade metafísica, ao invés da necessidade lógica. Desse modo, enfrentamos algo parecido com o que Sehon chama de o

Problema da Inescrutabilidade Epistemológica. Considerando o determinismo de modo a envolver a necessidade lógica, o Problema da Inescrutabilidade Epistemológica é este: como sabemos, ou pelo menos temos crença justificada, de que o determinismo é verdadeiro? Presumivelmente, o determinismo deve ser entendido como uma tese empírica. Entretanto, é difícil de ver como indícios empíricos poderiam decidir em favor da necessidade lógica.

Na concepção de determinismo que venho aqui defendendo, esse problema não existe, pois estou entendendo o determinismo de modo que ele deve ser entendido como envolvendo a necessidade metafísica, ao invés da lógica. Portanto, a principal preocupação de Sehon não se aplica aqui. A pesquisa empírica é tão relevante para saber se o determinismo é verdadeiro, em meu entendimento, como é relevante para saber se água é essencialmente H<sub>2</sub>O. Se há um problema aqui é apenas o problema de explicar a epistemologia da modalidade metafísica. E um problema geral dessa natureza não deveria contar como um ponto forte a favor de qualquer concepção de determinismo.

Pois bem, seria a possibilidade lógica de um IG compatível com o estilo de van Inwagen de entender o determinismo? Minha sugestão é a de que sim. Um ponto diferente, mais substancial, é o de saber se a concepção metafísica do conceito de determinismo é adequada. Acredito que seja. Mas essa é outra questão. Meu objetivo aqui foi apenas mostrar que não há um modo implausível de entender o determinismo que seja imune à interessante objeção de Sehon.

### 3.9. A objeção de Ted Warfield

A objeção de Sehon ao Argumento da Consequência é, basicamente, a de que o argumento é uma *ignoratio elenchi*. Dado que a concepção de determinismo do argumento é implausível, e desse modo não é uma concepção genuína de determinismo, o argumento falha em estabelecer a incompatibilidade entre o determinismo e o livre-arbítrio. Em outras palavras, sua conclusão é irrelevante para o problema. A objeção de Ted Warfield é muito similar neste aspecto. Em seu artigo “Causal Determinism and Human Freedom Are Incompatible: A New Argument for Incompatibilism”, podemos extrair uma objeção a favor da tese de que o Argumento da Consequência é uma *ignoratio elenchi*. A conclusão do argumento de van Inwagen, vale lembrar, era a seguinte:



$$(C) \square ((P_0 \ \& \ L) \supset P) \supset NP$$

Ou seja, se o determinismo é verdadeiro, então não há livre-arbítrio.

A objeção de Warfield é simples. (C) na verdade não é o incompatibilismo. Como podemos notar, “ $\supset$ ” representa a condicional material. Para Warfield, o incompatibilismo genuíno deveria ser representado pela condicional estrita. Em outras palavras, esta deveria ser a tese incompatibilista:

$$(I) \square ((P_0 \ \& \ L) \supset P) \neg NP$$

Uma vez que o Argumento da Consequência não estabelece (I), mas apenas (C), ele não estabelece a incompatibilidade entre o determinismo e o livre-arbítrio.

Tanto quanto vejo, Warfield tem toda a razão em chamar atenção para o fato de que (I) é sem dúvida mais forte que (C). Mas não vejo motivação para não considerar (C) como uma tese incompatibilista. Podemos simplesmente traçar uma distinção entre duas teses incompatibilistas: uma tese fraca, defendida até aqui, e uma tese forte, que seria precisamente (I). Tal como defini o incompatibilismo, é evidente que o argumento que venho aqui defendendo não é uma *ignoratio elenchi*. E há motivação para aceitar essa concepção mais fraca de incompatibilismo. Warfield chamou sem dúvida atenção para o fato de que essas versões do Argumento da Consequência não estabelecem a tese mais forte. No entanto, isso não mostra que (C) não seja uma versão plausível do incompatibilismo.

Em resumo, se as respostas às principais objeções ao Argumento da Consequência são bem-sucedidas, então há boas razões para pensar que o incompatibilismo é verdadeiro. Como se observou, as réplicas às principais críticas me levou a defender versões diferentes do argumento. No entanto, penso que, não obstante modificado, o argumento é ainda assim bastante persuasivo.

## CONSIDERAÇÕES FINAIS

O primeiro capítulo consistiu principalmente na formulação do problema da compatibilidade entre o determinismo e o livre-arbítrio. Defendi que é preciso distinguir cuidadosamente o problema tradicional, que é o de saber se temos livre-arbítrio, ou o de saber se o determinismo é verdadeiro, do problema de saber quais são relações conceituais entre as teses do determinismo e do livre-arbítrio. As discussões do primeiro capítulo também me permitiram argumentar que o famoso Dilema de Hume não é um argumento cogente a favor do impossibilismo.

Considerando a caracterização do problema no primeiro capítulo, apresentei o Argumento da Consequência em duas formulações principais oferecidas por van Inwagen, oferecendo respostas às objeções mais intuitivas. Com relação ao Primeiro Argumento Formal, argumentei que a objeção de Westphal (2012) não é forte o bastante para mostrar que o argumento não é sólido.

Finalmente, no terceiro capítulo, lidei com algumas das principais objeções ao argumento. Dividi as objeções com relação à validade, solidez e cogência do Argumento da Consequência. A versão final do argumento que acabei por defender repousa numa regra sutilmente diferente daquela oferecida por Finch e Warfield:

$$(\beta)2_m: NP, \Box_m(P \supset Q) \vdash NQ$$

A diferença, é claro, está no índice da caixa. A versão final do argumento por mim defendida é a seguinte:

- |    |   |                    |
|----|---|--------------------|
| 1. | $\Box_m((P_0 \ \& \ L) \supset P)$            | suposição          |
| 2. | $N(P_0 \ \& \ L)$                             | premissa           |
| 3. | $NP$  | 1, 2 ( $\beta$ ) 2 |
| 4. | $\Box_m((P_0 \ \& \ L) \supset P) \supset NP$ | 1-3, $\Box$        |

Penso que esta versão é tão intuitivamente plausível quanto a primeira versão de van Inwagen, além de consertar muitas das falhas do argumento original. Assim, creio que o Argumento da Consequência seja um bom argumento a favor do incompatibilismo; ao menos por enquanto.

## REFERÊNCIAS

- ARMSTRONG, D. (1983). *What Is a Law of Nature?*, Cambridge: Cambridge University Press.
- BEEBEE, H. (2003). “Local Miracle Compatibilism”. *Noûs*, 37: 258–277.
- BEEBEE, H. & Mele (2002). “Humean Compatibilism”, *Mind*, 111(442): 201–23.
- BOGHOSIAN, P. (1997). “Analyticity,” in B. Hale and C. Wright (eds.), *A Companion to the Philosophy of Language*, Oxford: Blackwell.
- BROOK, A. & STANTON, R. (2002). *Knowledge and Mind: a philosophical introduction*. Cambridge: The MIT Press.
- CAMPBELL, J. (2005). “Compatibilist Alternatives”, *Canadian Journal of Philosophy*, 35: 387–406.
- CARNAP, R. (1936). “Testability and Meaning”, *Philosophy of Science* 3: 419–471 and 4: 1–40.
- CARROLL, J. (2010). Laws of Nature. In: *The Stanford Encyclopaedia of Philosophy*.
- CHALMERS, D. (2002). “Does Conceivability Entail Possibility”, in J. Hawthorne & T. Gendler (eds.), *Conceivability and Possibility*, Oxford: Oxford University Press, pp. 145–200.
- CHELLAS, B. (1980). *Modal Logic: An Introduction*, Cambridge: Cambridge University Press.
- CHISHOLM, R. (1964). “Human Freedom and the Self”, *The Lindley Lectures*, Department of Philosophy, University of Kansas. Reprinted in Watson 2003.
- CRISP, T. M. & WARFIELD, T. (2000). “The Irrelevance of Indeterministic Counterexamples to Principle Beta.” *Philosophy and Phenomenological Research* 61, No. 1 pp.173-84.
- DAVIDSON, D. (1973). "Freedom to Act" in Ted Honderich, ed., *Essays on Freedom of Action*, London.
- DE ROSE, K. (2002). “Assertion, Knowledge and Context”, *The Philosophical Review*, 111(2): 167–203.

- ELGESEM, D. (1997). The modal logic of agency, *Nordic J. Philos. Logic* 2(2), pp.1–46.
- FARA, M. (2008), “Masked Abilities and Compatibilism”. *Mind*, 117 (468): 843–865.
- FINCH, A. & WARFIELD, T. (1998). “The Mind Argument and Libertarianism”. *Mind* 107, No 427, pp.515-28.
- FOLEY, R. (1979). “Compatibilism and Control over the Past”, *Analysis*, 39: 70–74.
- GEACH, P. (1969) *God and the Soul*. South Bend, IN: St. Augustine’s Press, 2001.
- GEIRSSON, H. (2005). “Conceivability and Defeasible Modal Justification”, *Philosophical Studies*, 122 (3): 279–304.
- GOODMAN, N. (1954). *Fact, Fiction and Forecast*, Cambridge, Mass.: Harvard University Press.
- GOVERNATORI, G. & ROTOLO, A. (2005). “On the Axiomatisation of Elgesem’s Logic of Agency and Ability.” *Journal of Philosophical Logic*, 34: pp. 403-31.
- HOFFMAN, A. (2012). “Are propositions set of possible worlds?”. *Analysis*, 72, number 3: pp. 449-455.
- KANE, R. (2005). *A Contemporary Introduction to Free Will*. New York: Oxford University Press.
- KENNY, A. (1975). *Will, Freedom, and Power*, Oxford: Blackwell.
- ŁUKASIEWICZ, J. (1967). “On Determinism”, in *Polish Logic*, ed. Storrs MacCall. Oxford: Oxford University Press.
- LEWIS, D. (1973). *Counterfactuals*, Cambridge: Harvard University Press.
- LEWIS, D. (1983). “Are We Free to Break the Laws?” *Theoria* 47 (1981): pp. 113–21.
- MAIER, J. (2010). “Abilities”. In: *The Stanford Encyclopaedia of Philosophy*.
- MERLUSSI, P. (2012). “O Argumento Modal da Consequência contra o Determinismo Moderado”. *Fundamento: revista de pesquisa em filosofia*. 1, no. 4: pp.61-79. Disponível em: <http://www.revistafundamento.ufop.br/Volume1/n3/vol1n3-4.pdf>.

- MERLUSSI, P. (2012). “Em defesa do incompatibilismo”. In: *Crítica* (London). Disponível em: <http://criticanarede.com/inwagen.html>.
- MCGINN, C. (1993). *Problems in Philosophy*. Oxford: Blackwell Publishers.
- PINK, Thomas (2004). *Free Will: A Very Short Introduction*. Oxford: Oxford University Press.
- MCGRATH, M. (2012). Propositions. In: *The Stanford Encyclopaedia of Philosophy*.
- MCKAY, T. & JOHNSON, D. (1996). “A Reconsideration of an Argument against Compatibilism”, *Philosophical Topics*, 24: 113–22.
- PLANTINGA, A. (1983). “On existentialism. In *Philosophical Studies*”, 1–20. vol. 44. Reprinted in his *Essays in the Metaphysics of Modality*, (2003), ed. M. Davidson, 158–75. Oxford: Oxford University Press.
- POPPER, K. (1988). *The Open Universe: an argument for indeterminism*. New York: Routledge.
- QUINE, W. O. (1960). *Word and Object*, Cambridge, Mass.: M.I.T. Press.
- RAMSEY, F. (1978). *Foundations*, London: Routledge and Kegan Paul.
- TAYLOR, R. 1962, “Fatalism”, *Philosophical Review*, 71: 56–66.
- RUSSELL, Bertrand (1905). "On Denoting", *Mind* 14, pp. 479–493.
- SIDER, T. & CONEE, E. (2005). *Riddles of Existence: A Guided Tour of Metaphysics*. New York: Oxford University Press.
- SLOTE, M. (1982). “Selective Necessity and the Free Will Problem”, *Journal of Philosophy*, 79: 5–24.
- STALNAKER, R. (1976). “Possible worlds”. In *Nou’s*, 65–75. vol. 10. Revised version in his *Inquiry*, (1984) 43–58. Cambridge, MA: The MIT Press and in his *Ways a World Might Be*, (2003) 25–39. Oxford: Oxford University Press.
- STALNAKER, R. (2010). Merely possible propositions. In *Modality: Metaphysics, Logic, and Epistemology*, eds. B. Hale and A. Hoffmann, 21–34. Oxford: Oxford University Press.
- STRAWSON, P. F. (1950). “On Referring”. *Mind*, 59: 320–344. New York: Routledge.

- TAYLOR, R. (1962). "Fatalism", *Philosophical Review*, 71: 56–66.
- van Inwagen, Peter, (1974). "The Incompatibility of Free Will and Determinism", *Philosophical Studies*, 25: 185–99.
- VAN INWAGEN, P. (1983). *An Essay on Free Will*, Oxford: Clarendon Press.
- VAN INWAGEN, P. (2000). "Free Will Remains a Mystery", in Tomberlin 2000, 1–19. Reprinted in Kane 2002B.
- VAIDYA, A. (2007). "Epistemology of Modality". In: The Stanford Encyclopaedia of Philosophy.
- VIHVELIN, K. (2011). "Arguments for Incompatibilism". In: The Stanford Encyclopaedia of Philosophy.
- WARFIELD, T. A. (2000). "Causal Determinism and Human Freedom are Incompatible: A New Argument for Incompatibilism", in Tomberlin: pp. 167–80.
- WESTPHAL, J. (2011). "The compatibility of divine foreknowledge and freewill". *Analysis* 71: 246–52.
- WESTPHAL, J. (2012). "Is there a modal fallacy in van Inwagen's 'First Formal Argument'?" *Analysis* 72(1): 36-41.
- WIDERKER, D., (1987). "On an Argument for Incompatibilism", *Analysis*, 47: 37–41.
- WILLIAMSON, T. (2000). *Knowledge and its Limits*, Oxford: Oxford University Press.
- WILLIAMSON, T. (2002). "Necessary existents". In *Logic, Thought, and Language*, ed. A. O'Hear, 233–51. Cambridge: Cambridge University Press.
- YABLO, S. (1993). "Is Conceivability a Guide to Possibility?". *Philosophy and Phenomenological Research*, 53: 1–42.
- ZAGZEBSKI, L. (2011). "Foreknowledge and Free Will". In: The Stanford Encyclopaedia of Philosophy.
- ZIMMERMAN, D. (1997). "Immanent Causation". *Nous* 31, pp 433-71.