



## **CAUSALIDADE, CONTRAFACTUAIS E LEIS**

*Causality, Counterfactuals and laws*

Tulio Roberto Xavier de Aguiar \*

**Resumo:** Examinando, neste artigo, a teoria contrafactual da causação de David Lewis, mostrando seu duplo aspecto como análise do conceito de causação e como investigação do que causação é na realidade. Eu comparo a teoria lewisiana com a teoria humiana da causação, enfatizando o papel diferente que as leis naturais têm em cada sistema e argumento que a teoria contrafactual tem maior sensibilidade para capturar os variados graus de contingência das relações causais.

**Palavras-chave:** Mundos possíveis. Superveniência. Graus de contingência. David Lewis. David Hume.

**Abstract:** In this article, I examine the counterfactual theory of causation of David Lewis, showing its dual aspect as analysis of the concept of causation and as research of what causation really is. I compare the Lewisian theory with Hume's theory of causation, emphasizing the different role natural laws play in each system and argue that the counterfactual theory is more sensitive to capture the varied degrees of contingency of causal relationships.

**Key-words:** Possible worlds. Supervenience. Degrees of contingency. David Lewis. David Hume.

---

\* Professor Associado do Departamento de Filosofia da UFMG. Artigo recebido em 12/09/2016 e aprovado para publicação em 18/07/2017.

## 1. Introdução

David Hume, ao chamar a atenção para os aspectos obscuros da noção de *causalidade*, impulsionou a demanda por uma análise reducionista da mesma. Desde então, os teóricos tentam reduzir a causalidade a itens supostamente mais básicos como leis, questões de fato, probabilidades, para citar alguns exemplos. Como é sabido, a própria teoria humiana opera uma redução da causalidade a regularidades entre eventos. A passagem a seguir é ilustrativa:

(...) nós podemos definir uma causa como sendo um objeto seguido por outro, tal que todos os objetos similares ao primeiro são seguidos por objetos similares ao segundo. Ou, em outras palavras, *se o primeiro não tivesse ocorrido, o segundo nunca teria existido* (HUME (1748): 146, grifo nosso).

A última sentença, entretanto, não é equivalente à teoria regularista oficial. Hume passa de causa como condição suficiente para causa como condição necessária em sentido modal, de uma ontologia atualista para o âmbito das possibilidades. Tal sentença aponta, ao que parece inadvertidamente, para uma teoria diferente – a teoria contrafactual da causalidade. Neste artigo, pretendo apresentar a teoria contrafactual, suas motivações, potencialidades e problemas. Particularmente, mostrarei que a conexão estabelecida pela teoria contrafactual entre causalidade e leis (regularidades) é mais adequada do que a que existe na teoria humiana. Em muitos pontos da exposição, seguirei de maneira um tanto livre a teoria como expressa por seu grande proponente David Lewis.

## 2. Um exemplo

Em nosso mundo observamos eventos que ocorrem sequencialmente: um evento e, depois, outro. Pensamos que algumas destas sequências são causais, outras vemos apenas como meras sequências. O que marca a diferença entre os dois tipos de sequências? Algum tipo de ligação física? Parece muito exigir que em todas as sequências causais haja ligação física. Pensemos nos casos de omissões: “a desatenção do pai causou o acidente da criança”. Não parece haver qualquer ligação física (ou processo) neste caso entre a causa e o efeito<sup>1</sup>. A ideia motivadora da teoria contrafactual é achar aquilo que está envolvido quando pensamos uma sequência como causal que não está presente nas sequências não-causais. Vejamos um exemplo.

---

<sup>1</sup> É pertinente notar que nós podemos fazer sentido da causalidade em um âmbito completamente ficcional. Por exemplo, em um contexto de histórias de magia, nós conseguimos entender uma sentença do tipo “as palavras de Merlin causaram a queda da muralha” e não supomos a necessidade de ligação física.

Imaginemos dois cenários envolvendo cada um dois relógios de pêndulo. No primeiro cenário, os dois relógios foram ajustados para que o relógio  $R_2$  badale logo após  $R_1$  a cada hora. No segundo cenário, há um mecanismo ligando os relógios  $R_4$  e  $R_3$  de maneira a que  $R_4$  badale após  $R_3$  a cada hora. Nosso julgamento intuitivo neste caso é que temos uma sequência causal apenas no segundo cenário. Se o nosso julgamento está correto, nós suporíamos que se  $R_3$  não tivesse badalado,  $R_4$  também teria permanecido em silêncio, ao passo que se  $R_1$  não tivesse badalado,  $R_2$  teria badalado assim mesmo. Os condicionais envolvidos nestes casos são chamados de condicionais contrafactuais e eles permitem capturar a diferença básica entre sequências causais e não-causais. Note que o condicional contrafactual envolvido no cenário 2 parece capturar o conteúdo do pensamento quando supomos estar diante de uma sequência causal mesmo que em um dado caso a atribuição causal seja errada. Além disso, mesmo se não conhecêssemos a descrição dos cenários acima, nós ainda suporíamos que o condicional contrafactual que aplicamos no cenário 2 (precisamente o “se o primeiro não tivesse ocorrido, o segundo nunca teria existido” de Hume) é a marca distintiva das sequências causais.

Outro ponto de grande importância a ser destacado é que nenhuma lei ou regularidade está *diretamente* envolvida em uma sequência causal, contrastando de maneira decisiva com a teoria Humiana. Na teoria contrafactual, o fato de um evento concreto (e não um evento tipo) ser a causa de outro não depende que estes instanciem alguma espécie de regularidade ou lei. O exemplo dos relógios acima ilustra o ponto, pois sua inteligibilidade não depende de regularidades para a construção do cenário. Mesmo que cada par de relógios batesse uma única vez, nós ainda conseguiríamos distinguir conceitualmente uma sequência causal de uma sequência não-causal. Nós saberíamos responder o que é uma sequência causal, embora talvez não soubéssemos dizer qual das sequências é causal. As regularidades ou leis, como veremos, entram em algum ponto da teoria, mas não estão implicadas diretamente na constituição de uma sequência causal.

### **3. Contrafactuais**

Até agora, utilizamos informalmente o termo “contrafactual” para denotar um condicional do tipo “se A tivesse sido o caso, B teria sido o caso” ou sua forma negativa, e mais útil na análise da causalidade, “se A não tivesse sido o caso, B não teria sido o caso”. São condicionais no modo subjuntivo e geralmente com antecedentes falsos como o termo “contrafactuais” indica. Os contrafactuais são absolutamente indispensáveis para o nosso pensamento ao considerarmos possibilidades e engajarmos em raciocínios hipotéticos. Eu penso com sentido que “se tivesse nascido

mais alto, teria sido um grande jogador de basquete”. A definição jocosa do termo “hipótese” também é ilustrativa: “hipótese é algo que não sabemos se é verdadeiro, mas supomos verdadeiro para ver como seria se fosse”. Daremos a seguir alguns esclarecimentos sobre o funcionamento dos condicionais contrafactuais.

Em primeiro lugar, o condicional contrafactual não funciona como o condicional material, pois obviamente ele não é automaticamente verdadeiro quando o antecedente é falso. Também não vale a regra do fortalecimento do antecedente que pode ser assim expressa:  $P \rightarrow Q \vdash P.R \rightarrow Q$ . Por exemplo, quando tomo como verdadeiro o contrafactual “se tivesse riscado o fósforo, ele teria acendido”, não infiro que “se tivesse riscado o fósforo e ele estivesse molhado, teria acendido”. Além disso, os contrafactuais não funcionam como condicionais estritos em que o antecedente acarreta logicamente o conseqüente. A questão que imediatamente se coloca então é a de quais seriam as condições de verdade dos contrafactuais. A linha de solução mais óbvia que se apresenta consiste em pensar que um contrafactual é verdadeiro quando o seu conseqüente é acarretado por seu antecedente mais alguma lei natural. O problema é que nenhum antecedente típico de um contrafactual acarreta sozinho, através de uma lei, o seu conseqüente. Embora nós tipicamente consideremos verdadeiro o contrafactual (f) “se este fósforo tivesse sido riscado, ele teria se acendido”, o nosso juízo mudaria se tivéssemos a informação que fósforo em questão estava molhado. Para continuarmos seguros que o contrafactual é verdadeiro, nós deveríamos assumir que tal condição não teria ocorrido se o antecedente do contrafactual *f* tivesse ocorrido. Estamos diante de um novo contrafactual cujo valor de verdade condiciona o valor de verdade do contrafactual original *f*. Assim, as condições de verdade de um dado contrafactual dependem regressivamente dos valores-de-verdade de outros contrafactuais, deixando-nos sem um critério estável para determinar a sua semântica (ver GOODMAN (1954)).

Para contornar essas dificuldades, David Lewis propôs o seguinte critério para a verdade dos contrafactuais:

A  $\gg$  C<sup>2</sup> é verdadeira se e somente se algum mundo, em que A e C são verdadeiras, é mais similar ao nosso mundo do que qualquer mundo, em que A é verdadeira e C é falsa.

Em outras palavras, um contrafactual é verdadeiro se ao tomarmos o antecedente verdadeiro e o conseqüente verdadeiro significar um menor afastamento do mundo atual do que tomar o antecedente verdadeiro e o conseqüente falso. Para tornar o critério acima operacional, precisamos falar de mundos possíveis e dar sentido à comparação de “distância” ou

<sup>2</sup> Lê-se A implica contrafactualmente C.

similaridade entre eles. Não é nosso objetivo aqui aprofundar a noção de mundo possível que tomaremos em um nível intuitivo. Ajuda a pensar que cada mundo possível reúne todas as maneiras que o nosso mundo poderia ser diferente ou uma recombinação de todas as coisas de nosso mundo, embora isso seja apenas uma aproximação heurística dos mundos possíveis<sup>3</sup>.

Quanto aos critérios de similaridade, nós temos que lidar com dois aspectos: as leis naturais e os fatos particulares. Em primeiro lugar, uma grande diferença nas leis naturais cria grande dissimilaridade entre os mundos. Em segundo lugar, grande dissimilaridade é criada se permitirmos diferença nos fatos particulares em uma extensa região do espaço e do tempo. Basicamente, quando fazemos um juízo sobre a similaridade entre dois mundos possíveis (incluindo o mundo atual) tentamos mobilizar ainda que de maneira difusa os dois elementos acima. Assim, conseguimos introduzir ordem nos mundos possíveis utilizando uma relação do tipo “o mundo  $w_1$  é mais similar ao mundo  $w_0$  do que o mundo  $w_2$ ”. Através deste ordenamento, conseguimos explicar a diferença entre sequências causais e não-causais.

#### **4. A Teoria**

Podemos reunir então os elementos necessários para precisar a análise contrafactual da causalidade. Diremos que “ $c$  causou  $e$ ” se e somente se  $c$  e  $e$  são eventos distintos e há uma cadeia de dependência contrafactual de  $e$  para  $c$  (possivelmente com apenas  $e$  e  $c$ ). No caso de uma cadeia causal sem intermediários, isso se resume ao contrafactual “se  $c$  não tivesse ocorrido,  $e$  não teria ocorrido”. Notar que a análise trata de eventos concretos e não eventos-tipo, não há a necessidade, como na teoria humiana padrão, que os eventos  $c$  e  $e$  sejam instâncias de alguma regularidade ou lei.

Precisamos entender também a exigência de que os relata da causalidade sejam eventos distintos. Lewis diz:

C e E devem ser eventos distintos – e distintos não apenas no sentido de não identidade, mas também no sentido de não-superposição e não-implicação (LEWIS (2000): 78).

Considere os seguintes contrafactuais (1) “se Sócrates não tivesse morrido, Xantipa não teria se tornado Viúva”, (2) “se hoje não fosse quarta-feira,

---

<sup>3</sup> Obviamente, a posição de Lewis, com o seu realismo dos mundos possíveis, é ontologicamente bem mais forte, pois admite como constituintes dos mundos possíveis particulares e propriedades que não existem em nosso mundo.

amanhã não seria quinta-feira” e (3) “se eu não tivesse lido o 1º capítulo deste livro, não teria lido o livro inteiro”. Nenhuma relação causal está realmente subjacente a estes contrafactuais. (1) talvez seja o caso mais difícil, pois parece superficialmente que poderíamos dizer que a morte de Sócrates causou a viuvez de Xantipa. Uma consideração, entretanto, nos dá motivos para dúvida. Mesmo separados no espaço, os eventos “morte de Sócrates” e “viuvez de Xantipa” ocorrem simultaneamente e é pelo menos um pouco estranho pensar em um mecanismo que possa agir de maneira instantânea. O que acontece neste caso é que “a morte de Sócrates” é o mesmo evento que “a morte do marido Xantipa” e este por definição acarreta a “viuvez de Xantipa”, violando o critério enunciado acima. (2) tem destino parecido, pois quarta-feira é o dia antes da quinta e, portanto, temos relação de implicação entre os *relata*. Finalmente, (3) apresenta uma relação de parte-todo e evidentemente viola o requisito de não-superposição dos *relata*.

A próxima qualificação necessária refere-se ao problema da assimetria causal e dos co-efeitos. Se a essência da causalidade consiste no contrafactual “se *c* não tivesse ocorrido, *e* não teria ocorrido”, então o que dizer do contrafactual “se *e* não tivesse ocorrido, *c* não teria ocorrido”? Se este contrafactual for verdadeiro também, a situação para a teoria contrafactual torna-se sombria, pois teríamos que um efeito pode causar a sua causa. Este tipo de contrafactual, chamado de retrocedente (*backtracking*), é interdito por David Lewis. O argumento é que o mundo mais próximo ao nosso, em que *c* e *e* ocorreram, é um mundo em que *c* ainda ocorreu e, por alguma razão, falhou em produzir *e*. Isto se deve ao fato de que o futuro sobredetermina o passado: cada evento passado tem vários determinantes no futuro chamados traços. A partir de um traço mais as leis da natureza, nós podemos remontar ao evento passado em questão. Por exemplo, alguém que dispara uma arma de fogo, introduz vários traços no futuro como vestígios de pólvora, os passarinhos que saíram voando assustados, traços na consciência do agente (eventualmente culpa e mudança de comportamento decorrente dela) e assim por diante. Seria muito difícil repor todo este estoque de traços em um mundo onde de fato não houve o disparo da arma e, assim, o mundo mais próximo ao nosso é um mundo onde a causa ocorreu, mas, por alguma razão, não produziu determinado efeito. Desta maneira, o contrafactual retrocedente “se *e* não tivesse ocorrido, *c* não teria ocorrido” é simplesmente falso. A tese de que o futuro sobredetermina o passado, segundo Lewis, é empírica, pois só vale para mundos que têm um grau de complexidade semelhante ao nosso, não valendo, por exemplo, para um mundo que consistisse de apenas dois átomos. Em tais mundos, parece-me, teríamos um tipo de relação causal degenerada.

O próximo problema com forte incidência na literatura sobre a causalidade é o dos co-efeitos. A queda da pressão atmosférica *p* tem como efeito a

queda do ponteiro do barômetro *b* e também a chegada de uma tempestade *t*. Para a teoria humiana, que reduz causalidade a meras regularidades, nós temos o incômodo resultado de que a queda do ponteiro do barômetro deve ser considerada a causa da chegada da tempestade já que ocorre antes e é regular com aquela. A interdição dos contrafactuais retrocedentes, na teoria contrafactual, permite também a solução do problema dos co-efeitos. Nós não afirmamos o contrafactual “se *b* não tivesse ocorrido, *t* não teria ocorrido” que suporia abolir a causa *p*, mas fixando esta, supomos que “mesmo que *b* não tivesse ocorrido, *t* ainda teria ocorrido como decorrência de *p*”.

## 5. Superveniência

Embora o conceito da *causalidade* e como nós pensamos sobre ela envolva de alguma forma os mundos possíveis, outra é a situação quando consideramos o que causalidade é no mundo. Neste último caso, Lewis defende o que ele mesmo chamou de tese da superveniência humiana: todos os fatos modais supervêm ao mosaico humiano. Tal mosaico é o conjunto das propriedades ocorrentes (não-modais) cuja instanciação em um ponto do espaço-tempo não necessita a instanciação de qualquer outra propriedade em outro ponto do mesmo espaço-tempo. Esta concepção é o corolário da doutrina humiana de que não existem conexões necessárias entre existências distintas. Assim, leis, relações causais, disposições e outros fatos modais supervêm ao mosaico humiano o que significa que se fossem diferentes, o mosaico seria diferente. No caso das leis, a concepção lewisiana é que elas seriam o melhor resumo das regularidades em um dado mundo – o sistema com a melhor combinação de simplicidade e força. Assim, as leis não governam o mundo, sendo apenas uma forma de organizar suas regularidades. O nosso mundo tem tal complexidade que apenas algumas regularidades pertencerão ao melhor sistema, deixando outras de fora. Não aprofundaremos aqui este ponto.

Como deve ter ficado claro pela discussão acima, as relações causais dependem das leis e fatos particulares de um dado mundo. A tese da superveniência, neste caso, diz precisamente que se as relações causais fossem diferentes, as leis e os fatos particulares seriam diferentes. Isso quer dizer que as relações causais em nosso mundo são determinadas em última análise pelo mosaico humiano. Desta forma, os mundos possíveis entram apenas para explicar o nosso conceito de causação e não o que a causação é no mundo. Armstrong explica bem o ponto:

As palavras de Lewis aqui são um pouco enganosas. É importante entender, como eu originalmente não fiz, e penso que muitos outros também não fizeram, que estes contrafactuais [que suportam as asserções causais] valem

apenas em virtude de feições do mundo em que as relações causais se dão. Como eu coloco, os *fazedores-de-verdade* para as verdades causais encontram-se apenas no mundo onde as relações causais são o caso. (Eu penso que isto segue diretamente da contingência da relação causal, uma contingência que Lewis não duvida). *Em sua teoria da causação* os mundos possíveis entram como mero expediente de cálculo (ARMSTRONG (2004): 445).

De fato, é uma confusão comum entre o que é para uma relação ser causal e quais relações são causais em um dado mundo, valendo a pena insistir no esclarecimento da questão. O primeiro ponto refere-se à busca pela análise correta do conceito de causação e tem a dependência contrafactual (com as qualificações acima) como resposta. Outra maneira de colocar o problema é lembrando que ele visa responder a pergunta “qual é a diferença no conteúdo do nosso pensamento quando distinguimos sequências causais de sequências não-causais?”. A explicação de Lewis para esta questão é, em parte, analítica: em todos os mundos possíveis, causação é um tipo de dependência contrafactual. Dizemos “em parte” porque a explicação do caráter assimétrico da relação causal tem, em Lewis, componentes empíricos como explicamos acima. Em mundos muito simples, por exemplo, com apenas dois átomos, a relação causal seria presumivelmente simétrica. O segundo ponto refere-se ao problema da identificação das relações causais em um dado mundo e sua resposta mobiliza apenas as feições do mundo em questão, notadamente suas leis e fatos particulares. Devemos ainda atentar para o fato de que ao procurarmos resposta para a questão de qual é o conceito de causalidade e do que a causalidade é no mundo, não devemos esperar necessariamente que a mesma teoria dê a melhor resposta para as duas questões. Poderíamos pensar, por exemplo, que a teoria contrafactual é a melhor para o conceito de causação, enquanto que alguma teoria física, postulando transferência de quantidades conservadas entre a causa e o efeito, é a melhor como descrição da realidade da causação. Não é claro que a mesma teoria possa ser otimizada para as duas tarefas<sup>4</sup>.

## **6. Causalidade e Graus de Contigência**

Segundo uma ampla tradição que remonta a Hume, as relações causais são contingentes e dependem do arranjo de qualidades ou dos fatos particulares em um dado mundo. Na interpretação costumeira da teoria humiana, quais relações causais existem dependem de quais regularida-

---

<sup>4</sup> David Kutach exemplifica o ponto com o exemplo da chave de fenda feita de cenoura. Podemos cortar uma cenoura no formato de uma chave de fenda tentando otimizá-la para apertar parafusos ao mesmo tempo em que é comestível. Tal chave de fenda teria a vantagem de ser mais leve e poderia ainda servir para matar a fome, mas dificilmente poderia substituir a chave de fenda tradicional que é mais especializada (KUTACH (2010): 52).

des (ou leis) se dão em nosso mundo e da situação dos fatos particulares e isto é contingente. Há uma crítica de longo termo a Hume, insistindo no fato de que parece excessivo exigir que para cada relação causal haja uma regularidade subtendida. Pensemos nas relações causais na história quando, por exemplo, tentamos explicar a eclosão de uma guerra. Parece extremamente irreal a expectativa de encontrarmos regularidades, ainda que muito complexas, subtendidas pelas asserções causais. Em um âmbito tão contingente quanto a história, não parece haver regularidades suficientes para cobrir todas as relações causais. Tal fato, entretanto, não significa que regularidades ou leis não desempenhem um papel importante no funcionamento da causalidade e é aqui precisamente que podemos comparar a teoria regularista de Hume com a teoria contrafactual.

O que inicialmente atrai alguns para a teoria humiana é que alguns dos seus exemplos paradigmáticos, como “este pão me alimentou”, têm uma ligação muito direta com regularidades que encontramos no mundo. Embora a relação causal que, neste caso, liga a ingestão do pão com a saciedade seja contingente, ela tem um alto grau de estabilidade, no sentido de que poucas situações desconectariam causa e efeito. Caso similar é dado por uma relação causal coberta pela regularidade “todo metal se expande quando aquecido” que também é bastante robusta. Contrastando fortemente com estes casos, temos o exemplo “a presença daquele automóvel naquela esquina e naquele momento causou o acidente” ou “o ter virado a minha cabeça naquele momento causou a visão daquela tragédia”. Estes casos exemplificam relações causais fortemente contingentes que dificilmente seriam assimiláveis a regularidades, sendo, contudo, bem capturados dentro da teoria contrafactual. No último caso, por exemplo, teríamos o contrafactual “se não tivesse virado a minha cabeça naquele momento, não teria visto aquela tragédia”, cuja verdade depende muito mais de fatos particulares ligados à situação no espaço e no tempo do que de regularidades. A teoria humiana não parece ter como acomodar este tipo de caso sem a suposição absurda de que todos os fatos particulares, como situação no espaço e no tempo, são cobertos por regularidades. Se há um consenso sobre regularidades e leis é que estas não devem mencionar particulares, inclusive pontos do espaço e do tempo.

O que está em jogo aqui é que as diversas relações causais em nosso mundo têm graus variáveis de contingência. Dentro do mosaico humiano que forma a base das relações causais, nós podemos falar sucessivamente em grau decrescente de contingência, de leis de alto nível, leis de baixo nível, regularidades simples e fatos particulares. As relações causais dependem de forma variável destes diversos níveis e explicar em que consiste exatamente a diferença entre eles é uma tarefa importante a ser realizada. A teoria contrafactual mostra uma sensibilidade interessante ao cobrir todo o espectro de contingência das relações causais.

## Bibliografia

- ARMSTRONG, D. (2004). "Going through the Open Door Again: Counterfactual versus Singularist Theories of Causation" in COLLINS, J. (ed.). *Causation and Counterfactuals*. Cambridge, The MIT Press, 2004.
- COLLINS, J. (ed.). *Causation and Counterfactuals*. Cambridge, The MIT Press, 2004.
- GOODMAN, N. (1954). *Fact, Fiction, and Forecast*. Cambridge, Harvard University Press, 1983.
- HUME, D. (1748). *An Enquiry Concerning Human Understanding*. Oxford, Oxford University Press, 1999 (edição crítica de Tom L. Beauchamp).
- HUME, D. (1748): *Investigação Acerca do Entendimento Humano* (col. "Os Pensadores). S. Paulo, ed. Nova Cultural, 1989.
- KUTACH, D. (2010). *Causation and its Basis in Fundamental Physics*. Draft.
- LEWIS, D. (1973). "Causation" in Sosa, E. & Tooley (eds.). *Causation*, Oxford, Oxford University Press, 1997.
- LEWIS, D. (1973). *Counterfactuals*. Malden, Blackwell Publishers.
- LEWIS, D. (1979). "Counterfactual Dependence and the Time's Arrow" in JACKSON, F. (ed.) *Conditionals*. Oxford, Oxford University Press, 1991.
- LEWIS, D. (2000). "Causation as Influence" in COLLINS, J. (ed.). *Causation and Counterfactuals*. Cambridge, The MIT Press, 2004.

Endereço do Autor:  
taguiar.bh67@gmail.com