OPEN ACCESS

COGNITIO
Revista de Filosofia
Centro de Estudos de Pragmatismo
São Paulo, v. 26, n. 1, p. 1-18, jan.-dez. 2025
e-ISSN: 2316-5278

https://doi.org/10.23925/2316-5278.2025v26i1:e70004

TRADUÇÃO

Naturalizando a lógica: como o conhecimento de mecanismos reforça a inferência indutiva[1]

Paul Thagard

Tradução
Marcos Rodrigues da Silva*
mrs.marcos@uel.br

Gabriel Chiarotti Sardi**
gabrielchi@hotmail.com

Resumo: Este artigo naturaliza a inferência indutiva, indicando como o conhecimento científico de mecanismos reais proporciona grandes benefícios para essa forma de inferência. Apresento a ideia de que o conhecimento sobre mecanismos contribui para a generalização, para a inferência da melhor explicação, para a inferência causal e para o raciocínio probabilístico. Partindo da ideia de que alguns A são B, uma generalização de que todos A são B se torna mais plausível quando um mecanismo conecta A e B. A inferência da melhor explicação é fortalecida quando as explicações empregam mecanismos e quando as hipóteses explicativas são elas próprias explicadas por meio de mecanismos. As inferências causais na explicação médica, no raciocínio contrafactual e por meio da analogia também se beneficiam de conexões por meio de mecanismos, os quais também auxiliam em problemas relativos à interpretação, disponibilidade e cálculo de probabilidades.

Palavras-chave: Explicação científica. Inferência indutiva. Mecanismos.

1 Introdução

Uma velha piada filosófica (atribuída a Morris Cohen, ainda que não haja certeza quanto a Cohen ter de fato feito a piada) diz que os manuais de lógica se dividem em duas partes: na primeira metade eles tratam de lógica dedutiva, e as falácias são desmascaradas; na segunda metade, sobre lógica indutiva, aparecem raciocínios falaciosos. Ainda que essa piada seja muito dura com a inferência indutiva, pois este tipo de inferência é indispensável na ciência e na vida cotidiana, ela assinala a diferença entre dedução e indução, e esta última introduz incertezas de um modo inevitável. Este artigo defende que uma apreciação de mecanismos pode ajudar substancialmente para reduzir os problemas que estão presentes em raciocínios indutivos.

A filosofia naturalista fez progressos substanciais na integração da epistemologia, da metafísica e da ética com ciências como física, psicologia e neurociência (Thagard, 2019). No entanto, a lógica pode ser compreendida como localizada além do alcance do naturalismo, tendo em vista que ela fornece ideais normativos acerca de como as pessoas deveriam raciocinar, e não descrições de como as mentes realmente

Recebido em: 18/01/2025.
Aprovado em: 28/04/2025.
Publicado em: 18/09/2025.

* Universidade Estadual de Londrina.
** Universidade de São Paulo.

1 **Nota dos tradutores**: texto publicado originalmente na *Philosophies* (v. 6, n. 52, 2021). Os tradutores agradecem ao filósofo Paul Thagard e aos editores da *Philosophies* pela disponibilização dos direitos de tradução deste artigo.

funcionam. No entanto, ocorreram avanços na psicologia da dedução (Johnson-Laird; Byrne, 1991; Rips, 1994) e foram propostos modelos computacionais e psicológicos de inferência indutiva (Holland *et al.*, 1986; Thagard, 2012). A inteligência artificial se desenvolveu com aplicações para a aprendizagem intensa e para outros tipos de inferência indutiva (Goodfellow; Bengio; Courville, 2016; Thagard, 2021).

Este artigo explora uma maneira diferente e adequada de naturalizar a inferência indutiva não apenas a psicologizando, mas também mostrando como o conhecimento científico de mecanismos reais fornece grandes benefícios para ela. Não estou argumentando que o conhecimento de mecanismos seja essencial para a indução, mas apenas que alguns tipos de inferência indutiva se tornam mais substanciais com tal conhecimento. Especificamente, mostro como o conhecimento sobre mecanismos contribui para a generalização, para a inferência da melhor explicação, para a inferência causal e para o raciocínio probabilístico. Até mesmo a dedução pode ser influenciada pelo conhecimento sobre mecanismos reais.

Uma busca pelo termo "mecanismo" no Google Scholar gerou, apenas no ano de 2020, mais de 7 milhões de respostas em vários campos das ciências naturais e sociais e mais de 200.000 menções. A filosofia da ciência atual tem investigado com profundidade a noção de mecanismos, os quais podem ser entendidos como combinações de partes conectadas cujas interações produzem mudanças regulares (Thagard, 2019; Bechtel, 2008; Craver; Darden, 2013; Craver; Tabery, 2015; Glennan, 2017). No entanto, essas investigações negligenciaram a contribuição que o entendimento dos mecanismos fornece para os problemas interligados da descrição e da justificação de vários tipos de inferência indutiva. A indução pode se efetivar sem mecanismos, mas ela é mais compreensível e confiável quando as inferências são apoiadas por informações sobre mecanismos.

A contribuição dos mecanismos para uma boa inferência indutiva é importante tanto por razões práticas quanto por razões teóricas. O mundo está repleto de desinformação sobre questões como saúde, clima e política. Uma forma de distinguir informação de desinformação se dá por meio da observação de como elas diferem em sua base inferencial. Ao passo que a informação surge de inferências indutivas sólidas baseadas em mecanismos, a desinformação baseia-se frequentemente em inferências que ignoram mecanismos ou dependem de mecanismos que são extremamente inconfiáveis. Assim, uma tarefa importante da lógica indutiva é discriminar mecanismos confiáveis de mecanismos não confiáveis.

2 Mecanismos

Dois dos problemas mais preocupantes que a humanidade enfrenta são as alterações climáticas e as epidemias virais. Há uma exigência de inferências indutivas para se responder a esses problemas; por exemplo: a atividade industrial humana que produz gases com efeito estufa conduz a um aquecimento global irreversível? Novas doenças virais, como a COVID-19, podem ser controladas? Felizmente, como mostra a Tabela 1, foi alcançado um conhecimento substancial sobre os mecanismos relevantes. Ela mostra as diversas terminologias usadas por diferentes filósofos na discussão dos mecanismos.

Tabela 1: Mecanismos relevantes para o aquecimento global e epidemias virais. Na linha horizontal superior, os parênteses mostram as várias terminologias usadas em discussões filosóficas de mecanismos. As outras linhas horizontais abaixo descrevem operações climáticas (Ranney; Clark, 2016) e mecanismos virais (Dimitrov, 2004).

	Combinação (Totalidade, Sistema, Estrutura)	Partes (Entidades, Componentes)	Interações (Atividades, Operações)	Mudanças	Resultados (Comportamentos, Funções, Fenômenos)
Aquecimento Global	Sistema Solar incluindo a Terra	Sol, radiação solar, atmosfera da Terra, superfície da Terra, moléculas de efeito estufa	A Terra absorve raios solares. A Terra emite luz infravermelha. Gases de efeito estufa absorvem luz, retendo energia. A energia aquece a Terra	A Terra esquenta	A temperatura da Terra aumenta permanentemente. Temperaturas extremas e inundações são cada vez mais comuns
Epidemia Viral	População Humana	Corpos, células, vírus	Vírus infectam as células e se reproduzem. Vírus se espalham para outros corpos	As infecções se espalham entre os corpos	Ocorrem epidemias e pandemias

Quando mecanismos são levados em conta, obtemos um auxílio para a solução de dois problemas interligados ao problema da indução: descrição e justificação. O problema de descrição é o de caracterizar como usualmente fazemos inferências indutivas. De um ponto de vista lógico, a forma é a de um conjunto de regras que são aplicadas às premissas e em seguida se infere uma conclusão. A ciência cognitiva pode estender esse esquema: pode incluir uma variedade de representações mentais – tais como representações pictóricas e representações de percepções de movimento –, além de procedimentos computacionais diferentes das regras lógicas de inferência. O problema da justificação é o de se os procedimentos indutivos são legitimados pela sua produção de conclusões confiáveis e úteis. Argumentarei que a inclusão de mecanismos para a descrição de vários tipos de inferência indutiva torna tais inferências mais úteis e justificáveis.

3 Generalização indutiva

A forma mais simples e familiar de indução é a generalização de alguns para todos; por exemplo: a partir da observação de que alguns porcos-formigueiros estão comendo formigas na terra, conclui-se que todos os porcos-formigueiros comem formigas na terra. John Stuart Mill percebeu que algumas dessas inferências são mais plausíveis do que outras:

> Por que, em alguns casos, um único exemplo é suficiente para uma indução completa, enquanto em várias outras situações – mesmo quando não haja uma exceção –

> avançamos tão pouco no sentido de estabelecer uma proposição universal? Quem puder responder essa questão conhece mais da filosofia da lógica do que o mais sábio dos antigos, e terá resolvido o problema da indução. (Mill, 1970, p. 206).

Por exemplo: algumas instâncias de porcos-formigueiros podem ser suficientes para nos convencer que todos os porcos-formigueiros comem formigas, ao passo que precisaríamos de muito mais casos para ter certeza de que eles comem amendoim.

O termo "plausível" deve ser entendido do seguinte modo: uma afirmação é plausível se ela é mais coerente com a evidência disponível do que afirmações contrárias, embora o grau de coerência possa não ser suficiente para que a afirmação seja definitivamente aceita (Connell; Keane, 2006; Thagard, 1989). A coerência pode ser avaliada computacionalmente pelo fato de maximizar o cumprimento de exigências relativas a como as hipóteses explicam as evidências e outras hipóteses, e pelo fato de maximizar o cumprimento de exigências relativas à concorrência entre hipóteses incompatíveis.

Na década de 1980, a pesquisa em psicologia e filosofia forneceu uma resposta à pergunta de Mill, a partir da noção de variabilidade (Nisbett *et al.*, 1983; Thagard; Nisbett, 1982). Alguns estudos mostraram que pessoas que ouvem falar sobre um metal que não existe rapidamente inferem, a partir de poucos casos, que uma alta porcentagem deste suposto metal queima com uma chama azul; por outro lado, pessoas que ouvem falar sobre um pássaro que não existe relutam em inferir que uma alta porcentagem deste suposto pássaro é azul. Uma explicação plausível para essa diferença é que as pessoas estão cientes de que os metais têm pouca variabilidade em suas propriedades de combustão e cientes de que pássaros têm muito mais variabilidade em suas cores. Quando a generalização indutiva trata de tipos de coisas e propriedades que têm pouca variabilidade, então mesmo uma única instância ou algumas delas podem ser suficientes.

Essa análise provavelmente captura diferenças psicológicas, mas uma resposta mais profunda emerge quando se consideram os mecanismos. Na década de 1780, Antoine Lavoisier identificou o mecanismo da combustão como sendo do seguinte tipo: a combinação de materiais com oxigênio para produzir calor e luz. Posteriormente, o mecanismo foi aprofundado para explicar como os átomos de elementos como o carbono interagem com átomos de oxigênio para produzir calor: moléculas se movem rapidamente e a luz é resultante da emissão de fótons. Assim, um novo elemento ainda não descoberto também teria sua combustão acompanhada de uma chama azul.

Em contraste, no caso de pássaros ainda não descobertos, os mecanismos fornecem muito menos garantias sobre a plausibilidade da generalização indutiva. O principal mecanismo relevante para inferir a cor dos animais vem da genética, por meio da suposição de que os descendentes herdam a cor de seus pais. Muitas vezes, no entanto (como no caso de gatos, papagaios e humanos), os genes não produzem cores uniformes, e sim diversas cores de cabelo e pele. Os genes têm variantes chamadas alelos, e alelos diferentes podem produzir variações na cor do cabelo. Em humanos, por exemplo, a maioria dos ruivos tem uma mutação no gene do receptor de melanocortina 1, mutação essa que afeta o cabelo e a pele. Assim, na ausência de amplo conhecimento sobre a genética de pássaros ainda não descobertos, inferências sobre tais supostos pássaros seriam pouco confiáveis.

A noção de mecanismo também esclarece os paradoxos que afetaram as tentativas, nas décadas de 1940 e 1950, de estabelecer relatos puramente sintáticos de generalização indutiva. Teorias da confirmação propuseram que o suporte indutivo para generalizações da forma $(x)(Fx \rightarrow Gx)$ vinha de instâncias como Fa e Ga. Por exemplo: observar um corvo preto confirma a hipótese de que todos os corvos são pretos. No entanto, Carl Hempel notou que $(x)\ (Fx \rightarrow Gx)$ é logicamente equivalente a $(x)\ (\sim Gx \rightarrow \sim Fx)$, e isso sugere que um corvo preto também confirma a estranha hipótese de que todas as coisas não-pretas são não-corvos (Hempel, 1965). É igualmente estranho que um sapato branco confirma a afirmação de que todos os corvos são pretos.

A estranheza desaparece quando uma hipótese conecta um tipo a uma propriedade, como no caso de "todos os corvos são pretos", conexão esta que torna a hipótese muito mais plausível. Os genes

dos corvos não dispõem de alelos para cores além de preto, e os raros corvos brancos surgem devido ao albinismo por meio de mutações que produzem melanina. Portanto, os mecanismos conhecidos da genética conduzem à confirmação, quando se observa um corvo preto, de que todos os corvos são pretos. Ao contrário, nenhum mecanismo conecta coisas não-pretas com coisas não-corvos, então não há razão para levar esta hipótese a sério, apesar de sua equivalência sintática com "todos os corvos são pretos". Uma das razões do fracasso do positivismo lógico como filosofia da ciência é a de que o raciocínio científico vai além da sintaxe e deve, em vez disso, considerar a constituição física do mundo em termos de mecanismos.

Uma outra forma de lidar com o paradoxo dos corvos, proposta por Nelson Goodman, é chamada de "novo enigma da indução" (Goodman, 1965). A observação de esmeraldas verdes parece apoiar a generalização de que todas as esmeraldas são verdes, mas também apoiam a generalização de que todas as esmeraldas são verduis – "verdul" denota esmeraldas verdes se elas são observadas antes do tempo t e simultaneamente verde e azul (verduis) se observadas após o tempo t. Felizmente, mecanismos fornecem um contraste valioso entre as generalizações de que todas as esmeraldas são verdes e de que todas as esmeraldas são verduis. Podemos ler no *The Gem Encyclopedia*:[2] "As esmeraldas são formadas quando o cromo, o vanádio e o ferro estão presentes nas gemas verdes (berilo). A presença variável destes três elementos é que dá cor às esmeraldas. O cromo e o vanádio produzem uma cor verde intensa, enquanto o ferro dá à pedra uma tonalidade azulada".

As esmeraldas são formadas quando o cromo, o vanádio e o ferro estão presentes no mineral berilo. É a variação destes três elementos que confere à esmeralda as possibilidades de suas cores. O cromo e o vanádio produzem uma cor verde forte, já o ferro dá à pedra uma tonalidade azulada.

A cor das esmeraldas resulta primeiro da forma como os seus elementos constituintes (suas partes) interagem com a luz para refletir a luz em uma frequência específica (cerca de 550 nanômetros), e, em segundo lugar, de como essa frequência de luz estimula os receptores da retina a enviar sinais para o cérebro, sinais estes que são interpretados como verdes. Levando em consideração esses dois mecanismos, o tempo não é relevante para tornar as esmeraldas azuis e, portanto, o conhecimento anterior sobre os mecanismos é muito mais útil do que a sintaxe pura (expressa, por exemplo, no paradoxo dos corvos de Hempel) para compreender generalizações indutivas.

A generalização indutiva atribui uma propriedade a um tipo, e muitos filósofos têm reconhecido que a noção de tipos naturais sustenta melhor um raciocínio indutivo do que por meio de coleções artificiais (Bird; Tobin, 2017). Tipos naturais são às vezes identificados como essências metafísicas – sendo essas as mesmas em todos os mundos possíveis; porém, essa identificação não tem utilidade para a filosofia da ciência. Uma boa descrição de tipos naturais foi desenvolvida por Richard Boyd, que propôs que as espécies biológicas são grupos de propriedades unidas por propriedades subjacentes que são homeostáticas: uma série estável de propriedades é mantida porque as anomalias de tais propriedades têm pouca chance de persistir (Boyd, 1991). Assim, devemos pensar que a indução a partir de tipos naturais é melhor sustentada como resultando de mecanismos subjacentes.

A generalização indutiva não requer um conhecimento dos mecanismos; por vezes, mesmo que não tenhamos conhecimento de um mecanismo que conecte A e B, podemos ter um bom número de exemplos de A e de B para alcançar a conclusão de que todos A são B. Por exemplo: há séculos se sabia que as cascas da planta do salgueiro reduzem a dor; em 1897 a casca foi isolada e com isso se descobriu a aspirina, e seu mecanismo bioquímico foi desvelado em 1971. No entanto, o conhecimento de mecanismos é altamente útil para compreender as contribuições da variabilidade e dos tipos naturais para a inferência indutiva e para a compreensão do fracasso da abordagem puramente sintática da teoria da confirmação.

2 https://www.gia.edu/seeing-green

4 Inferência da melhor explicação

Quando se toma "indução" em um sentido estrito, estamos a falar apenas da generalização de alguns para todos; mas, em um sentido mais amplo, "indução" é qualquer inferência que difere da dedução – uma vez que a dedução, ao contrário da indução, garante certeza. Há diversas formas de indução, como a analogia e a inferência estatística; uma outra forma é conhecida como "inferência da melhor explicação" (Harman, 1973; Lipton, 2004; Thagard, 1978). Essa expressão não era empregada até a década de 1960, mas a inferência de hipóteses explicativas já era sugerida por filósofos do século XIX, tais como William Whewell e Charles Peirce; além disso, os astrônomos da Renascença e mesmo Aristóteles podem ser vistos como precursores da inferência da melhor explicação. Peirce introduziu o termo "abdução" para a geração e aceitação de hipóteses explicativas, e muitos trabalhos recentes em filosofia e inteligência artificial analisam variedades de inferência abdutivas (Thagard, 2012; Josephson; Josephson, 1994; Magnani, 2009; Peirce, 1958).

A forma básica da inferência da melhor explicação (doravante "IBE") é a seguinte:

A evidência *E* demanda explicação;
A hipótese *H* fornece uma explicação melhor de *E* do que explicações alternativas disponíveis;
Portanto, se infere *H*.

IBE é comum na vida cotidiana; por exemplo: quando as pessoas atribuem estados mentais a outras pessoas e quando a mecânica identifica as causas de panes em automóveis. Também é comum na justiça (quando os jurados concluem que um acusado é culpado) e na medicina (quando os médicos concluem que um paciente tem uma doença que explica os sintomas de tal paciente).

Assim como ocorre com generalizações, o conhecimento dos mecanismos não é essencial para a inferência da melhor explicação, mas os mecanismos ajudam a reduzir o risco lógico inerente a IBE. A forma menos rigorosa de IBE tem sido rejeitada por se apesentar na forma de um argumento inválido:

Se *A* então *B*;
B;
Portanto *A*.

Esta forma de inferência é obviamente fraca porque pode haver muitas outras razões para *B* que não são *A*. Uma maneira de tornar IBE mais rigorosa seria exigir uma conexão causal (*A* causa *B*), mas ainda poderia haver outras causas que precisariam ser levadas em consideração. Para que um raciocínio seja uma forma de IBE, alguma avaliação comparativa de alternativas tem de ter ocorrido.

Defensores de IBE são, em geral, vagos acerca do que é uma explicação; eles caracterizam a noção de explicação do seguinte modo: enquadrar algo intrigante em um padrão familiar. Padrões úteis vão desde a narrativa solta, frequente na vida cotidiana, até a dedução exata encontrada em campos matemáticos como a física. Em biologia, medicina, ciência cognitiva e outras disciplinas, a explicação é muitas vezes uma descrição do mecanismo causal; por exemplo: quando a gripe é explicada pela infecção de células por vírus.

As explicações por meio de mecanismos fortalecem IBE. Em primeiro lugar, mecanismos fornecem uma conexão muito mais forte entre hipóteses e evidências do que meras relações causais (se-então) ou causas abstratas. Ao se afirmar que os sintomas de febre, tosse e dores de um paciente são o resultado da gripe, tal afirmação pode ser reforçada por muitos detalhes causais, inclusive o de que um vírus conhecido infectou o sistema respiratório do paciente, causando reações corporais específicas. Quando um mecanismo é conhecido, temos boas razões para considerar que uma hipótese que explica os sintomas é uma hipótese séria; já a hipótese de que o paciente está possuído por demônios, por não explicitar os mecanismos, é fantasiosa. Uma avaliação por meio de IBE é comparativa (uma hipótese é

mais explicativa do que outras hipóteses alternativas), mas empregar explicações que explicitam seus mecanismos estabelece uma medida elevada que deverá ser seguida por todos aqueles que propõem hipóteses alternativas: tais hipóteses alternativas também devem ser capazes de explicitar os mecanismos que conectam a hipótese à evidência.

Em segundo lugar, os mecanismos são importantes para IBE porque as hipóteses são avaliadas não apenas pelo quanto explicam, mas também pelo fato de que elas próprias podem ser explicadas (Thagard, 1989; Harman, 1986). Por exemplo: no direito, a hipótese de que um acusado é culpado de assassinar uma vítima tem de explicar muitos aspectos da cena do crime, tais como as impressões digitais do acusado na arma do crime. No entanto, a hipótese de que um acusado é culpado também recebe apoio quando se fornece um motivo pelo qual o assassino matou a vítima (por exemplo: por ciúme). Tais explicações jurídicas baseiam-se em um conhecimento de senso comum, o qual, por sua vez, se baseia em mecanismos psicológicos não muito rigorosos, como crenças e desejos (por exemplo: a crença de que a vítima havia provocado um sentimento de vingança no acusado).

Na ciência e na medicina, explicações (de um nível superior) de hipóteses muitas vezes empregam mecanismos mais profundos apoiados por evidências substanciais. Por exemplo, a teoria da evolução de Darwin obteve suporte por sua capacidade de explicar muitas observações, tais como as distribuições de espécies; mas a teoria da evolução foi por sua vez explicada pelos mecanismos de seleção natural e de transmissão genética de características herdadas. O ceticismo acerca da verdade das teorias científicas é inspirado pela observação de que muitas hipóteses científicas se revelaram falsas; por exemplo: a teoria do éter de Aristóteles e a teoria da substância química da tradição de pesquisa do flogisto (Laudan, 1981). No entanto, essa tese pessimista pode ser enfraquecida quando se nota que todas as teorias científicas aceitas e que foram aprofundadas por meio de explicações a partir de mecanismos e a partir de evidências adicionais resistiram a teorias alternativas (Thagard, 2007).

E com isso obtemos uma formatação de uma IBE forte:

A evidência *E* demanda explicação;

A hipótese *H* fornece explicações por meio de mecanismos, e *H* é melhor do que as hipóteses alternativas e seus mecanismos são melhores do que outros mecanismos;

Além disso, os mecanismos fundamentais de *H* são explicados por mecanismos mais fundamentais;

Portanto, se infere *H*.

A aplicação da formatação acima não elimina completamente a incerteza da inferência (pois IBE é uma forma de raciocínio indutivo), mas ajuda a reduzir a aparente arbitrariedade de IBE. Usar mecanismos ajuda a superar o problema identificado por Bas van Fraassen de que a melhor explicação pode ser apenas a melhor explicação de um conjunto defeituoso de explicações (Van Fraassen, 1980). Se uma hipótese fornece um mecanismo (M1)[3] dos fenômenos que constituem uma evidência para a hipótese, se (M1), por sua vez, for explicado por mecanismos subjacentes (M2) que explicam por que as partes e suas interações se comportam da forma como se comportam, e se M1 e M2 forem avaliados em relação a explicações alternativas, então teremos bases sólidas para aceitar tal hipótese.

Sustentar IBE por meio da utilização de mecanismos pode parecer circular: os mecanismos se justificariam pela própria IBE. Contudo, de um ponto de vista naturalista, o objetivo não é fornecer uma justificativa *a priori* da inferência indutiva, mas sim o de identificar como a indução funciona quando ela funciona bem. IBE e inferência indutiva não satisfazem o ideal dedutivo de inferências que partem de axiomas indubitáveis para teoremas; ao invés, IBE e inferência indutiva requerem um ideal alternativo baseado na coerência geral entre uma série interligada de hipóteses e evidências. As noções filosóficas de coerência explicativa eram inicialmente vagas, mas já podemos entender a coerência mecanicamente

3 **Nota dos tradutores**: as siglas (M1) e (M2) não constam na publicação original de Thagard, mas optamos por incluí-las para facilitar a compreensão do texto do autor.

– ou seja: como um processo computacional realizado por redes neurais (Thagard, 2000). A existência de mecanismos (psicologicamente e neurologicamente plausíveis) acerca de como IBE reúne as evidências e as hipóteses apoiam a conclusão de que IBE é um bom relato de boa parte da indução humana. É claro que precisamos considerar hipóteses alternativas a IBE, e uma alternativa interessante a IBE é a inferência bayesiana, conforme será discutido em seguida.

5 Causalidade e contrafactuais

Uma das aplicações mais importantes de IBE ocorre em afirmações causais; por exemplo: a doença COVID-19 é causada pelo novo coronavírus; o aquecimento global é causado pelo aumento da emissão humana de gases de efeito estufa. Tais afirmações vão além de generalizações indutivas (de que todos A são B e, portanto, A causa B). Essas afirmações são de importância prática, pois sugerem que podemos lidar com efeitos indesejáveis, tais como doenças e aquecimento global, e assim poderíamos atuar em suas causas. Os mecanismos não explicam a causalidade porque pressupõem noções causais de que partes e interações produzem, geram, ou são responsáveis pelas mudanças.

Uma análise da inferência causal depende de quais causas são consideradas. Os céticos que afirmam que a causalidade é uma ideia falsa e não científica não têm necessidade de avaliar inferências causais, ainda que não consigam explicar por que relatos causais são onipresentes na ciência, como mostram as milhões de menções do *Google Scholar* ao termo "causa". David Hume afirmou que a causalidade era apenas uma conjunção constante (Hume, 1888), o que faria com que a inferência causal não passasse de uma generalização indutiva; contudo, a distinção entre correlação e causalidade é geralmente reconhecida. Teorias probabilísticas de causalidade que procuram causas onde P (efeito | causa) > P (efeito) também tentam fazer inferência causal baseada em dados, mas têm problemas com causas não observáveis, como partículas subatômicas. Teorias que tratam de intervenção na causalidade enfatizam como as causas podem ser inferidas através da identificação de intervenções que alteram os seus efeitos, mas elas têm problemas quando lidam com relações causais em outras partes da galáxia que estão além da intervenção humana.

Prefiro uma descrição ecumênica da causalidade, descrição essa que deixa de lado a definição e, ao invés, busca identificar exemplos-padrão e características típicas de causalidade, observando seu papel explicativo (Thagard, 2019). Exemplos-padrão de relações causais incluem empurrões, puxões, movimentos, colisões, ações e doenças. As características típicas de causalidade (menos rigorosas do que condições necessárias e suficientes) são as seguintes: ordenamento temporal, com causas antes dos efeitos; padrões sensório-motores como chutar uma bola; regularidades expressas por regras gerais; condições estatísticas; redes causais de influência. A causalidade explica por que os eventos acontecem e por que as intervenções funcionam.

A causalidade, nessa perspectiva, se traduz em uma inferência da melhor explicação – pois leva em consideração uma série de evidências sobre padrões temporais, correlações, probabilidades e intervenções. O conhecimento dos mecanismos não é essencial para IBE, mas é de muito auxílio em casos nos quais as interações das partes conectam uma suposta causa com um efeito; por exemplo: a alegação de que a causa da COVID-19 é a infecção pelo novo coronavírus (SARS-CoV-2) não é apenas correlacional, pois há amplo conhecimento de como o vírus infecta células e perturba órgãos como pulmões e vasos sanguíneos.

Pesquisadores médicos têm dedicado muita atenção à análise das considerações para inferir as causas das doenças, inclusive a força das associações empíricas e o conhecimento anterior (Hill, 1965; Hennekens; Buring, 1987; Dammann, 2020). Todas essas considerações se prestam à análise por meio de modelos computacionais baseados na coerência explicativa (Dammann; Poston; Thagard, 2019). A identificação de mecanismos é apenas uma das considerações necessárias para reconhecer a coerência

de uma afirmação causal, mas fornece um aval importante para afirmações como a de que fumar causa câncer. Esta hipótese foi aceita na década de 1960, antes que se soubesse sobre como a fumaça do cigarro perturba o crescimento normal das células, mas tornou-se mais forte graças à compreensão de como os produtos químicos são cancerígenos para as células pulmonares. A hipótese de que o vírus Zika causa defeitos neurais em bebês, fundamentada inicialmente em correlações entre infecções por Zika e defeitos congênitos, tornou-se mais plausível quando se baseou também na compreensão de como o vírus infecta neurônios e produz crescimento anormal. Em 2021, quando foi identificado um mecanismo pelo qual vacinas baseadas em adenovírus causam coagulação do sangue, houve um justificado receio quanto à existência de coágulos sanguíneos no cérebro de pessoas que tomaram dois tipos de vacinas para COVID-19.

Os mecanismos são relevantes para decidir se um fator *C* é a causa de um evento E nas quatro seguintes situações (Thagard, 1999):

(1) Existe um mecanismo conhecido pelo qual *C* produz *E*.
(2) Existe um mecanismo plausível pelo qual *C* produz *E*.
(3) Não existe nenhum mecanismo conhecido pelo qual *C* produz *E*.
(4) Não existe nenhum mecanismo plausível pelo qual *C* produz *E*.

A quarta situação é problemática para a inferência indutiva porque sugere que a ligação entre *C* e *E* não pode ser observada sem que se abandone a ciência bem estabelecida. Muitas alegações paranormais – como possessão demoníaca, percepção extrassensorial e a telecinesia – são incompatíveis com a física baseada em evidências.

Os contrafactuais fornecem um dos domínios mais problemáticos da inferência causal. Como deveríamos avaliar afirmações como a de que, se o novo coronavírus não tivesse se espalhado para um mercado úmido em Wuhan, então a pandemia de COVID-19 nunca teria acontecido, ou a de que se a revolução industrial não tivesse ocorrido, então não haveria aquecimento global? Tratamentos lógicos habituais de contrafactuais usando mundos possíveis conectados por relações de similaridade são matematicamente elegantes, mas cientificamente inúteis.

O pesquisador de Inteligência Artificial Judah Pearl desenvolveu uma descrição muito mais plausível de contrafactuais, descrição essa baseada em relações causais (Pearl, 2011). Quando se analisa causas em termos de contrafactuais, se procede do seguinte modo: se a causa não tivesse ocorrido, então o efeito não teria ocorrido. Em vez disso, Pearl sugere que as afirmações contrafactuais, sendo verdadeiras ou não, ainda seriam plausíveis ou implausíveis dependendo das relações causais do mundo. Para avaliar causalmente os contrafactuais, podemos trabalhar com uma rede causal e ajustar algumas das causas para ver o que acontece, seja excluindo uma causa, seja alterando a força de sua conexão com um efeito. Métodos computacionais para tais ajustes estão disponíveis por meio de redes bayesianas ou redes de coerência explicativa.

O conhecimento causal nem sempre depende de mecanismos, mas os mecanismos realçam a inferência causal e também podem contribuir para julgamentos contrafactuais mais plausíveis. Geralmente, para avaliar uma afirmação contrafactual de que, se o evento 1 não tivesse acontecido, então o evento 2 não teria ocorrido, ajuda a fazer as seguintes perguntas: existe um mecanismo que conecta o evento 1 e o evento 2? Existem outros mecanismos que podem produzir o evento 2 sem o evento 1? Como um exemplo, considere a afirmação contrafactual de que se Donald Trump não tivesse sido infectado com o novo coronavírus, então ele não teria contraído a COVID-19. Os mecanismos pelos quais o vírus produz a doença são bem conhecidos, e nenhum outro mecanismo produz a COVID-19, e assim o contrafactual acima é plausível.

Os mecanismos também ajudam no que diz respeito a outro tipo instável de inferência indutiva que se beneficia das relações causais: a analogia. Na sua forma menos rigorosa, a inferência analógica

apenas atesta que duas coisas ou eventos são semelhantes em alguns aspectos e permite a inferência de que serão semelhantes em outro aspecto. Por exemplo: Montreal é semelhante a Toronto por ser uma grande cidade canadense; Toronto tem metrô, então provavelmente Montreal também tem.

Dedre Gentner percebeu que as analogias são muito mais úteis quando se baseiam em relações causais sistemáticas (Gentner, 1983). Se você conhece o contexto político das cidades canadenses e como este contexto se manifesta nos níveis nacional, provincial e municipal, então você pode construir uma história causal sobre como o processo de decisão que gerou um metrô em Toronto provavelmente irá gerar um metrô em Montreal. Tais analogias causais ficam ainda mais fortes quando há uma correspondência entre os mecanismos; por exemplo: construir um metrô em Toronto e construir um metrô em Montreal. Outro exemplo seria o de que uma das razões para pensar que o vírus Zika causa defeitos congênitos é a semelhança com o mecanismo pelo qual o sarampo causa defeitos congênitos (Dammann; Poston e Thagard, 2019).

Os mecanismos não são obrigatórios para as inferências analógica, contrafactual e causal em geral. No entanto, eles ajudam a reduzir a incerteza das inferências indutivas.

6 Probabilidade

Embora a teoria da probabilidade só tenha sido inventada no século XVIII, muitos filósofos assumem que a inferência indutiva deve estar baseada na noção de probabilidade (Howson e Urbach, 1989; Olsson, 2005). Visto que o conjunto de inferências qualitativas discutidas até agora (generalização, IBE, causal, contrafactual, analógica) não se reduz ao raciocínio probabilístico, considero implausível a ideia de que a inferência indutiva deve estar baseada na noção de probabilidade. Seja como for, as probabilidades são indispensáveis para muitos tipos de inferências estatísticas; por exemplo: para estimar a eficácia das vacinas na prevenção da COVID-19, em que os dados são usados para estimar P (infecção | vacinação).

No núcleo da inferência probabilística está o teorema de Bayes, que diz que a probabilidade de uma hipótese dada à evidência depende da probabilidade anterior da hipótese multiplicada pela probabilidade da evidência dada a hipótese, tudo dividido pela probabilidade da evidência. Em símbolos: P (H | E) = P(H) X P(E | H)/P(E). Como um teorema do cálculo de probabilidade, esse resultado é objetivo, mas, ao aplicá-lo a casos reais de inferência indutiva, enfrentamos problemas relativos à interpretação, disponibilidade e cálculo de probabilidades. Considerações sobre mecanismos ajudam com todos esses três problemas.

A sintaxe da teoria da probabilidade é incontroversa devido à axiomatização de Kolmogorov, mas ainda há disputas no que diz respeito à sua semântica (Hájek, 2019; Hájek; Hitchcock, 2016). As probabilidades devem ser interpretadas como frequências, graus de crença, relações lógicas ou propensões? A interpretação de que são frequências parece mais consistente com as práticas estatísticas, mas tem dificuldade em estabelecer o que significa uma probabilidade de frequência de longo prazo e em aplicar esta noção para a probabilidade de eventos únicos. Bayesianos assumem que as probabilidades são graus de crença, mas enfrentam problemas sobre como tais crenças subjetivas podem objetivamente descrever o mundo; além disso, há descobertas experimentais que não podem ser explicadas pela noção de probabilidade (Kahnema; Tversky, 2000). As tentativas de descrever probabilidades como relações lógicas têm enfrentado o problema de como considerações abstratas de lógica e evidência podem gerar probabilidades que, ao mesmo tempo, satisfaçam aos axiomas da teoria das probabilidades e sirvam como um guia prático para a vida.

Devido a essas questões, penso que a interpretação mais plausível da probabilidade é a teoria da propensão: as probabilidades são tendências de situações físicas para gerar frequências relativas de longo prazo; por exemplo: P (*infecção* | *vacinação*) = x é uma propriedade objetiva do mundo por meio da qual as interações entre pessoas, vírus e vacinas têm uma disposição para produzir, a longo prazo,

uma proporção x de pessoas vacinadas que foram infectadas. No entanto, essa interpretação ignora em grande medida o que são realmente propensões, tendências ou disposições.

O que significa dizer que o vidro tem tendência a quebrar se for atingido? A fragilidade não é apenas uma questão de relações lógicas, tais como "se o vidro for atingido, então ele quebrará", ou contrafactuais, como "se o vidro tivesse sido atingido, ele teria quebrado". Em vez disso, podemos descobrir os mecanismos pelos quais o vidro é formado (por exemplo: descobrir que moléculas não muito bem-organizadas geram fissuras microscópicas, riscos ou impurezas e com isso se tornam fracas e quebram quando o vidro é atingido), e, partir disso, explicar a sua fragilidade (Khun, 2015). Da mesma forma, os mecanismos da infecção viral, do contágio, da vacinação e da imunidade explicam a disposição das pessoas de serem protegidas por vacinas. Os mecanismos lançam luz na interpretação da probabilidade como propensão e apontam para uma nova interpretação da probabilidade por meio de mecanismos (Thagard, 2019; Abrams, 2012).

Karl Popper introduziu a interpretação das probabilidades como propensões a fim de resolver problemas enfrentados pela interpretação de frequência para aplicações a eventos singulares. Ele afirmou que "as propensões podem ser explicadas como possibilidades (ou como medidas ou 'pesos' de possibilidades) que são dotadas de tendências ou disposições para se realizarem, e que são aceitas como responsáveis pelas frequências estatísticas que irão ocorrer em longas sequências de repetições de um experimento" (Popper, 1959, p. 30).

Assim como ocorre com as forças, as propensões apontam para propriedades disposicionais não observáveis do mundo físico.

Popper, no entanto, não elucidou a natureza dessas possibilidades, tendências ou disposições e não deixou claro como elas explicariam as frequências. Essas lacunas deixadas por Popper são preenchidas caso se entendam as propensões como mecanismos que geram e explicam frequências. As propensões são disposições para gerar frequências que resultam das conexões e interações entre as partes de mecanismos subjacentes; por exemplo: a probabilidade de dois dados jogados darem 12 é 1/36 devido às interações dos dados com seu ambiente e assim, a longo prazo, 12 ocorrerá em uma proporção aproximada de 1/36.

A interpretação da propensão por meio de mecanismos funciona bem para análises de probabilidades estatísticas, mas não se aplica aos tipos de inferência indutiva considerados neste artigo. A generalização indutiva e a inferência da melhor explicação não geram conclusões com probabilidades determináveis, pois nenhuma propensão conhecida gera conclusões como a de que todos os corvos são pretos e que as espécies evoluíram por seleção natural. Deste modo, as probabilidades são irrelevantes para inferências não estatísticas (Thagard, 2019).

O segundo problema com as abordagens bayesianas para a inferência indutiva é que as probabilidades relevantes geralmente não estão disponíveis, seja como frequências, graus de crença ou propensões. Os bayesianos geralmente apresentam apenas exemplos simplificados, mas, se trabalhassem com exemplos que tivessem dezenas ou centenas de eventos ou proposições, perceberiam que o cálculo bayesiano exige um grande número de probabilidades condicionais (Thagard, 2004). Prestar atenção aos mecanismos ajuda a restringir a identificação de probabilidades que importam em um contexto inferencial específico; por exemplo: compreender os mecanismos da infecção, contágio, vacinação e imunidade deixa claro que muitos fatores estranhos – como a possessão demoníaca – podem ser ignorados.

Os mecanismos também ajudam no terceiro problema com abordagens bayesianas à inferência indutiva: a inferência probabilística tem se revelado computacionalmente intratável, pois a quantidade de cálculo aumenta exponencialmente à medida que aumenta o número de variáveis (Kwisthout; Wareham; Van Rooij, 2011). O cérebro humano tem milhares ou milhões de crenças e grandes bancos de dados de um computador podem ter centenas ou milhares de variáveis inter-relacionadas. Foram desenvolvidas redes bayesianas que eliminam as redes potencialmente problemáticas, por meio da introdução de um operador (chamado de *DO operator*) que torna as redes restritas a conexões causalmente plausíveis,

mas a semântica desse operador não é bem especificada (Pearl; Mackenzie, 2018). Compreender os mecanismos subjacentes em uma situação reduz em grande medida as relações causais que fornecem conexões plausíveis entre variáveis em uma rede bayesiana, reduzindo assim o número de probabilidades que precisa ser computado.

Assim, o conhecimento dos mecanismos auxilia em três problemas da abordagem bayesiana à inferência indutiva: interpretação, disponibilidade e cálculo. Esse auxílio não é suficiente para defender a teoria da probabilidade como uma abordagem geral à inferência indutiva, uma vez que isso exigiria análises probabilísticas de todos os outros tipos de indução que não foram aqui discutidos. No entanto, o uso legítimo de probabilidades em muitos casos importantes de raciocínios empregados na vida real é aprimorado pela incorporação do conhecimento sobre mecanismos.

7 A avaliação de mecanismos

Apresentei a contribuição da informação sobre mecanismos para vários tipos de inferência indutiva, mas não tratei do problema da avaliação da qualidade dos mecanismos. Num exemplo extremo, alguém poderia afirmar que a possessão demoníaca é o mecanismo responsável pela COVID-19: as partes são demônios, organismos e almas, as interações são demônios invadindo organismos conectados às almas, e o produto disso são organismos infectados e almas que sofrem. Felizmente, a filosofia da ciência pode avaliar o que torna alguns mecanismos muito mais explicativos do que outros.

Carl Craver e Lindley Darden identificam três vícios que podem ocorrer por meio de representações de mecanismos: superficialidade, incompletude e incorreção (Craver; Darden, 2013, cap. 6). Mecanismos superficiais meramente redescrevem o fenômeno a ser explicado, sem fornecer qualquer estrutura interna, como na piada de Molière de que remédios para dormir fazem as pessoas dormir porque têm uma virtude dormitiva. Mecanismos superficiais não podem ser empregados para uma inferência da melhor explicação. Meu exemplo do demônio não é superficial, pois, pelo menos, tenta dizer algo sobre demônios infectando organismos e almas para produzir sintomas.

Mecanismos incompletos fornecem apenas esboços de mecanismos, excluindo partes e interações que precisariam ser levadas em consideração. A incompletude é por vezes inevitável devido à falta de conhecimento; por exemplo: quando Darwin não conseguiu explicar a herança de características pela prole. No entanto, o objetivo geral da ciência é usar os mecanismos como esquemas que fornecem detalhes sobre partes, conexões, interações e resultados causais. Quando avaliam teorias concorrentes, os cientistas podem comparar o grau de completude dos mecanismos empregados em suas explicações. Meu exemplo do demônio é bastante incompleto porque não diz nada sobre como os demônios conseguem infectar organismos e nem como as alterações nestes organismos causam sofrimento mental.

Um mecanismo é incorreto quando não consegue descrever com precisão as partes, conexões e interações, ou, usando a terminologia de Craver e Darden, suas entidades, organização e atividades. Um esquema deve explicar como um mecanismo realmente funciona e não apenas como poderia funcionar. No início de uma investigação, os cientistas podem legitimamente especular sobre como um mecanismo poderia funcionar, mas deve haver evidências acumuladas caso se deseje sugerir que o mecanismo é pelo menos plausível à luz do conhecimento anterior e, em última análise, é plausível para sugerir que o mecanismo explica como o mundo funciona. Podemos deixar de lado o mecanismo de um demônio, pois a ciência não encontrou nenhuma evidência da existência de demônios e almas, e também não encontrou um papel causal para demônios e almas quanto a infecções.

Em contextos científicos, a corretude pode ser uma questão de grau; por exemplo: quando há boas evidências disponíveis para a existência de partes dos mecanismos, mas as interações propostas ainda não foram estabelecidas empiricamente. Quando há competição entre diferentes teorias acerca de mecanismos, podemos compará-las em relação ao seu grau de corretude, bem como pelo seu grau de completude.

Uma forma ainda mais forte de verificar que um mecanismo pode estar incorreto foi acima mencionada: em relação à inferência causal. Um mecanismo que postula partes e interações incompatíveis com a ciência consolidada nem sequer pode ser julgado como fornecendo uma explicação possível; por exemplo: os demônios têm capacidades mágicas, como tomar posse de almas – mas isto é incompatível com a física e a psicologia científicas.

Assim, os mecanismos propostos podem ser avaliados a partir de sua superficialidade, completude e corretude. Podemos avaliar as evidências das partes, conexões e interações hipotéticas, e também se as interações são possíveis, plausíveis ou realmente causam o resultado que deve ser explicado.

Em resumo: podemos julgar um mecanismo como sendo forte, fraco, defeituoso ou prejudicial. Um mecanismo forte é aquele com boas evidências de que suas partes, conexões e interações realmente produzem o resultado a ser explicado. Um mecanismo fraco é aquele que não é superficial, mas ao qual faltam detalhes importantes sobre as partes, conexões, interações e sobre sua eficácia na produção do resultado a ser explicado. Mecanismos fracos não devem ser descartados, porque podem ser o melhor que pode ser feito em uma situação, tal como no início das investigações sobre conexões entre tabagismo e câncer e entre casca de salgueiro e alívio da dor.

Mecanismos podem ser rotulados como defeituosos quando temos boas razões para duvidar da existência de suas partes ou interações, ou para duvidar de uma produção alegadamente causal. As dúvidas sobre a existência de partes podem vir de três direções. Primeiro, se esforços consideráveis não foram suficientes para encontrar evidências quanto às partes, então temos razões para acreditar que tais partes não existem. O clichê de que a ausência de evidência não é evidência de ausência não se aplica quando tentativas consideráveis de encontrar evidências não conseguiram fornecer razões para acreditar em sua existência. A inexistência de evidências sobre demônios, unicórnios e deuses justifica a crença em sua inexistência.

Além disso, hipóteses sobre partes e interações podem ser rejeitadas quando as teorias que as propõem forem substituídas por outras que forneceram explicações melhores; por exemplo: a teoria do flogisto, que dominou as explicações químicas da combustão durante a maior parte do século XVIII, foi substituída pela teoria do oxigênio de Lavoisier, que propôs diferentes partes e interações. Portanto, temos boas razões para duvidar da existência do flogisto e de suas interações com materiais inflamáveis. Posteriormente, o oxigênio foi isolado da água e de outros gases e átomos de oxigênio e foi inclusive fotografado através de microscópios eletrônicos, e então a evidência das partes e interações do mecanismo do oxigênio se tornou progressivamente mais forte.

Por fim, a existência de partes e interações pode ser posta em dúvida quando, conforme já sugerido, seu funcionamento é inconsistente com os princípios científicos estabelecidos. A medicina homeopática se tornou popular no início do século XIX por meio da sugestão de que quantidades de substâncias que tivessem alguma semelhança com sintomas de doenças poderiam ser usadas para curar a doença. Este mecanismo é deficiente para explicar as doenças devido à implausibilidade de afirmações causais baseadas em quantidades e semelhanças diminutas.

Alguns mecanismos propostos não são apenas defeituosos, mas também prejudiciais, na medida em que suas aplicações são perigosas aos seres humanos, ameaçando a sua integridade física ou psicológica. O alegado mecanismo homeopático é tóxico porque as pessoas que usam tratamentos ineficazes para problemas médicos graves podem não obter tratamentos baseados em evidências que de fato funcionam. No início do século XIX acreditava-se que a homeopatia era provavelmente melhor do que tratamentos usuais devido a crenças sobre desequilíbrios de humor, os quais também são deficientes como mecanismos, pois apelavam a tratamentos de restauração do equilíbrio, como sangria e purificação do corpo, o que piorava os pacientes.

Em resumo: os mecanismos contribuem efetivamente para a inferência indutiva se forem fortes ou, ainda que fracos, são mais fortes do que as alternativas disponíveis. Mecanismos defeituosos e perigosos bloqueiam a eficácia epistêmica e a prática da inferência indutiva.

8 Conclusões

A importância das contribuições dos mecanismos para a inferência indutiva se estende para além da filosofia. Alguns psicólogos perceberam a importância dos mecanismos para o pensamento e a aprendizagem das pessoas (Johson; Ahn, 2015; Keil; Lockhart, 2021). Um melhor entendimento de como os mecanismos são mentalmente representados e processados deve contribuir para uma análise mais aprofundada das vantagens dos mecanismos causais para o conhecimento relativamente superficial das associações entre eventos observáveis. Da mesma forma, a inteligência artificial teve grande sucesso através do método associativo de aprendizagem profunda, mas a inteligência de nível humano exigirá que os computadores compreendam a causalidade com base em mecanismos (Goodfellow; Bengio; Courville, 2016; Thagard, 2021). Tanto as pessoas quanto os computadores podem aprender melhor se forem usados mecanismos que oferecem contribuições para a inferência indutiva.

O significado social do papel dos mecanismos da inferência indutiva vem da necessidade de diferenciar desinformação de informação. Lidar com as mudanças climáticas e com a COVID-19 gerou muitas evidências informativas, mas as controvérsias também geraram muitos casos de desinformação, como alegações de que as alterações climáticas são uma mistificação e de que a COVID-19 poderia ser tratada com alvejantes (Thagard – *no prelo*). Separar informação de desinformação requer a identificação de bons padrões de inferência indutiva: os que levam à informação e os que levam à desinformação. Observar o uso de mecanismos para a indução justificada é uma das contribuições para esta separação, como vimos na Tabela 1.

Pode parecer ridículo que os mecanismos também possam ser relevantes para a inferência dedutiva, mas considere um argumento devido a Gilbert Harman (Harman, 1986). Suponha que você acredite que todos os porcos-formigueiros são cinzentos e que o animal do zoológico é um porco-formigueiro. Você deveria, portanto, inferir dedutivamente que o animal é cinza? Se você perceber que o animal é realmente marrom, você pode querer considerar, em vez disso, que não é um porco-formigueiro ou que você estava errado em acreditar que todos os porcos-formigueiros são cinzentos. Em geral, não se pode inferir, a partir das premissas de um argumento dedutivo, que a conclusão do argumento é verdadeira, porque você pode precisar questionar algumas das premissas ou mesmo se preocupar com a validade de um tipo particular de argumento dedutivo, tal como o silogismo disjuntivo. O argumento de Harman sugere que toda inferência dedutiva é, na verdade, indutiva, de modo que os mecanismos são potencialmente relevantes. Você pode saber, por exemplo, que mutações em genes de cor são comuns em animais semelhantes ao porco-formigueiro e, portanto, há mais motivos para duvidar de sua inferência dedutiva.

Deixo claro que a inferência indutiva nem sempre depende de mecanismos. No entanto, quando o conhecimento dos mecanismos está disponível, muitas vezes ele pode ser valioso para tornar a inferência indutiva mais confiável. Eu argumentei a favor da relevância de informação baseada em mecanismos para a generalização, para a inferência da melhor explicação, para o raciocínio causal e para o raciocínio baseado em probabilidades. Pensar por meio de mecanismos torna as pessoas mais inteligentes e ajuda a naturalizar a lógica.

Referências

ABRAMS, Marshall. Mechanistic probability. *Synthese*, v. 187, p. 343–375, 2012. https://doi.org/10.1007/s11229-010-9830-3.

BECHTEL, William. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience.* New York: Routledge, 2008.

BIRD, Alexander; TOBIN, Emma. Natural Kinds. In *Stanford Encyclopedia of Philosophy*. Stanford: Stanford University, 2017.

BOYD, Richard. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philos. Studies Int. J. Philos. Anal. Trad*., v. 61, p. 127–148, 1991. https://doi.org/10.1007/BF00385837.

CONNELL, Louise; Keane, Mark T. A model of plausibility. *Cogn. Sci*., v. 30, p. 95–120, 2006. https://doi.org/10.1207/s15516709cog0000_53.

CRAVER, Carl F.; DARDEN, Lindley. *In Search of Mechanisms:* Discoveries across the Life Sciences. Chicago: University of Chicago Press, 2013.

CRAVER, Carl F.; TABERY, James. Mechanisms in Science. In: *Stanford Encyclopedia of Philosophy*. Stanford: Stanford University, 2015.

DAMMANN, Olaf. *Etiological Explanations*. Boca Raton: CRC Press, 2020.

DAMMANN, Olaf; POSTON, Ted; THAGARD, Paul. How do Medical Researchers Make Causal Inferences? In: MCCAIN, K; KAMPOURAKIS, K. (eds.). *What Is Scientific Knowledge? A*n Introduction to Contemporary Epistemology of Science. New York: Routledge; p. 33–51, 2019.

DIMITROV, Dimiter S. Virus entry: Molecular mechanisms and biomedical applications. *Nat. Rev. Microbiol.*, v. 2, p. 109–122, 2004. https://doi.org/10.1038/nrmicro817.

GENTNER, Dedre. Structure-mapping: A theoretical framework for analogy. *Cogn. Sci*., v. 7, p. 155–170, 1983. https://doi.org/10.1207/s15516709cog0702_3.

GLENNAN, Stuart. *The New Mechanical Philosophy*. Oxford: Oxford University Press, 2017.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. *Deep Learning*. Cambridge: MIT Press, 2016.

GOODMAN, Nelson. *Fact, Fiction and Forecast* - 2nd ed. Indianapolis: Bobbs-Merrill, 1965.

HÁJEK, Alan. Interpretations of Probability. In: *Stanford Encyclopedia of Philosophy.* Stanford: Stanford University, 2019.

HÁJEK, Alan; Hitchcock, Christopher. *The Oxford Handbook of Probability and Philosophy.* Oxford: Oxford University Press, 2016.

HARMAN, Gilbert. *Change in View: Principles of Reasoning*. Cambridge: MIT Press/Bradford Books, 1986.

HARMAN, Gilbert. *Thought*. Princeton: Princeton University Press, 1973.

HEMPEL, Carl G. *Aspects of Scientific Explanation.* New York: The Free Press, 1965.

HENNEKENS, Charles H.; BURING, Julie E. *Epidemiology in Medicine.* Boston: Little, Brown and Co., 1987.

HILL, Austin B. The environment and disease: Association or causation? *Proc. R. Soc. Med*., v. 58, p. 295–300, 1965. DOI: https://doi.org/10.1177/0035915765058005.

HOLLAND, John H.; HOLYOAK, Keith J.; NISBETT, Richard E.; THAGARD, Paul R. *Induction: Processes of Inference, Learning, and Discovery*. Cambridge: MIT Press, 1986.

HOWSON, Colin; URBACH, Peter. *Scientific Reasoning: The Bayesian Tradition*. Lasalle: Open Court, 1989.

HUME, David. *A Treatise of Human Nature*. Selby-Bigge, L.A., Ed. Oxford: Clarendon Press, 1888.

JOHNSON, Samuel G.; AHN, Woo-Kyoung. Causal networks or causal Islands? The representation of mechanisms and the transitivity of causal judgment. *Cogn. Sci*., v. 39, p. 1468–1503, 2015. https://doi.org/10.1111/cogs.12213.

JOHNSON-LAIRD, Phillip N.; BYRNE, Ruth. M. *Deduction*. Hillsdale: Lawrence Erlbaum Associates, 1991.

JOSEPHSON, John R.; JOSEPHSON, Susan G. (Eds.) *Abductive Inference:* Computation, Philosophy, Technology. Cambridge: Cambridge University Press, 1994.

KAHNEMAN, Daniel; TVERSKY, Amos. (Eds.) *Choices, Values, and Frames.* Cambridge: Cambridge University Press, 2000.

KEIL, Frank; LOCKHART, Kristi. Beyond cause: The development of clockwork cognition. *Curr. Dir. Psychol. Sci*., v. 30, p. 167–173, 2021. https://doi.org/10.1177/0963721421992341.

KUHN, Jen. *Part 2: Why does glass break? Corning Museum of Glass*. 2015. Disponível em: https://blog.cmog.org/2015/06/03/part-2-why-does-glass-break/. Acesso em: 18/06/2021.

KWISTHOUT, Johan; WAREHAN, Todd; VAN ROOIJ, Iris. Bayesian intractability is not an ailment that approximation can cure. *Cogn. Sci.*, v. 35, p. 779–784, 2011. https://doi.org/10.1111/j.1551-6709.2011.01182.x.

LAUDAN, Larry. A confutation of convergent realism. *Philos. Sci*., v. 48, p. 19–49, 1981. DOI: https://doi.org/10.1086/288975.

LIPTON, Peter. *Inference to the Best Explanation* - 2nd ed. London: Routledge, 2004.

MAGNANI, Lorenzo. *Abductive Cognition:* The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning. Berlin: Springer, 2009.

MILL, John S. *A System of Logic* - 8th ed. London: Longman, 1970.

NISBETT, Richard E.; KRANTZ, David; JEPSON, Christopher; KUNDA, Ziva. The use of statistical heuristics in everyday inductive reasoning. *Psychol. Rev*., v. 90, p. 339–363, 1983. https://psycnet.apa.org/doi/10.1037/0033-295X.90.4.339.

OLSSON, Erik. *Against Coherence: Truth, Probability, and Justification.* Oxford: Oxford University Press, 2005.

PEARL, Judea. The algorithmization of counterfactuals. *Ann. Math. Artif. Intell*., v. 61, p. 29–39, 2011. https://doi.org/10.1007/s10472-011-9247-9.

PEARL, Judea; MACKENZIE, Dana. *The Book of Why:* The New Science of Cause and Effect. New York: Basic Books, 2018.

PEIRCE, Charles S. *Collected Papers*. Hartshorne, W.P., Burks, A., (eds.). Cambridge: Harvard University Press, 1958.

POPPER, Karl R. The propensity interpretation of probability. *Br. J. Philos. Sci*., v. 10, p. 25–42, 1959. https://doi.org/10.1093/bjps/X.37.25.

RANNEY, Michael A.; CLARK, Dav. Climate change conceptual change: Scientific information can transform attitudes. *Top. Cogn. Sci*., v. 8, p. 49–75, 2016. https://doi.org/10.1111/tops.12187.

RIPS, Lance J. *The Psychology of Proof:* Deductive Reasoning in Human Thinking. Cambridge: MIT Press, 1994.

THAGARD, Paul. *Bots and Beasts:* What Makes Machines, Animals, and People Smart? Cambridge: MIT Press, 2021.

THAGARD, Paul. Causal inference in legal decision making: Explanatory coherence vs. Bayesian networks. *Appl. Artif. Intell*., v. 18, p. 231–249, 2004. https://doi.org/10.1080/08839510490279861.

THAGARD, Paul. *Coherence in Thought and Action*. Cambridge: MIT Press, 2000.

THAGARD, Paul. Coherence, truth, and the development of scientific knowledge. *Philos. Sci.*, v. 74, p. 28–47, 2007. https://doi.org/10.1086/520941.

THAGARD, Paul. Explanatory coherence. *Behav. Brain Sci.*, v. 12, p. 435–467, 1989. https://doi.org/10.1017/S0140525X00057046.

THAGARD, Paul. *How Scientists Explain Disease*. Princeton: Princeton University Press, 1999.

THAGARD, Paul. Mechanisms of misinformation: Getting COVID-19 wrong and right. *No prelo*.

THAGARD, Paul. *Natural Philosophy: From Social Brains to Knowledge, Reality, Morality, and Beauty*. Oxford: Oxford University Press, 2019.

THAGARD, Paul. The best explanation: Criteria for theory choice. *J. Philos*., v. 75, p. 76–92, 1978. https://doi.org/10.2307/2025686.

THAGARD, Paul. *The Cognitive Science of Science: Explanation, Discovery, and Conceptual Change.* Cambridge: MIT Press, 2012.

THAGARD, Paul; NISBETT, Richard E. Variability and confirmation. *Philos. Studies*, v. 42, p. 379–394, 1982. DOI: https://doi.org/10.1007/BF00714369.

VAN FRAASEN, Bas. *The Scientific Image*. Oxford: Clarendon Press, 1980.

OPEN ACCESS

COGNITIO

Revista de Filosofia
Centro de Estudos de Pragmatismo

São Paulo, v. 26, n. 1, p. 1-18, jan.-dez. 2025
e-ISSN: 2316-5278

https://doi.org/10.23925/2316-5278.2025v26i1:e70004