

# A epistemologia bayesiana e o problema do desacordo

## *Bayesian epistemology and peer disagreement*

**José Carlos Camillo Castro Neto**

Mestrando em Filosofia pela UFMT

josecarloscamillo@gmail.com

**Resumo:** O problema do desacordo entre pares foi levantado recentemente nos debates filosóficos. Se duas pessoas, com igual capacidade epistêmica, têm em mãos o mesmo conjunto de evidências, mas, quando vão apontar suas conclusões entre si acerca do mesmo assunto, descobrem que discordam, que atitude eles teriam o direito de tomar? Várias respostas foram dadas para essa pergunta desde então e elas podem ser agrupadas em Conformistas e Não-Conformistas. Contudo, ainda mais recentemente, utilizaram-se pressupostos e cálculos da teoria bayesiana para lidar com a questão. Foram muitas as tentativas realizadas a partir da teoria bayesiana, mas, analisá-las exaustivamente seria impossível. Por isso, este trabalho analisará duas abordagens feitas a partir da teoria bayesiana ao problema do desacordo entre pares. Uma delas defenderá a proposta Conformista e a outra, a Não-Conformista. Ao fim, será evidenciado que a abordagem que tem resultados intuitivos e uso fidedigno das restrições da teoria bayesiana é a proposta Não-Conformista.

**Palavras-chave:** Desacordo entre Pares; Epistemologia Bayesiana; Teoria Justificacionista.

**Abstract:** *The question about peer disagreement was raised up recently in the philosophical realm. If two people, with identical epistemic abilities, hold the same set of evidences, but present different conclusions about the same issue, and one present his own conclusion to the other in order to both of them realize that they disagree about that issue, what epistemic attitude are they entitled to take? Lots of answers came up since the question was made and all of them can be regrouped in two major positions: Conformist and Nonconformist. More recently, philosophers tried to answer this very question by the Bayesian theory. Once again, lots of proposals were made and the aim of this paper is to analyze two of them, one that holds the Conformist view and one that attacks it. At the conclusion, it will be evident that, from the two approaches analyzed that the one more internally coherent and with the Bayesian theory is the Nonconformist one.*

**Keywords:** *Peer Disagreement; Bayesian Epistemology; Justificationist View.*

O assunto do desacordo é muito recente. Embora algumas propostas sejam um pouco mais antigas, o debate cresceu de fato com a proposta de Feldman (2006). Trabalhando com filosofia da religião, ele lidava muito com desacordo e decidiu estudá-

lo de um aspecto epistemológico. Sendo muitos de seus alunos adeptos de religiões diferentes, eles tinham uma atitude de tolerância para com os que seguiam outras religiões. O que chamou mais a atenção de Feldman é que, a partir de um mesmo conjunto de evidências, as crenças de seus alunos diferiam muito. Mesmo assim, os alunos consideravam uns aos outros como sendo tão inteligentes e capazes quanto consideravam a si mesmos.

Ruivo (2016) apresentou o problema do desacordo através da seguinte forma: levando em conta que (1) você, com base na evidência E formou a crença que P de forma independente e que (2) um par epistêmico seu, também de modo independente, chega à crença que não-P com base na mesma evidência E; a pergunta que podemos fazer é “[...] qual atitude racional você está intitulado a ter com respeito a P?” (RUIVO, 2016, p. 51).

Antes de apresentar as principais respostas a essas perguntas, é importante definir o que é um par epistêmico. Lackey (2010) define um par epistêmico como um sujeito que preencha os seguintes critérios: 1) igualdade evidencial – ambos os sujeitos em desacordo devem dispor das mesmas evidências; 2) igualdade cognitiva – ambos os sujeitos devem ser igualmente competentes, inteligentes e razoáveis; 3) abertura completa (*full disclosure*) – ambos os sujeitos devem compartilhar as razões para sua crença. Além disso, Feldman (2006) já havia estabelecido que ambos os sujeitos devem se reconhecer como pares epistêmicos.

O objetivo deste artigo é analisar duas propostas que utilizam o método da epistemologia formal ou epistemologia bayesiana para responder à questão do desacordo. As duas propostas foram escolhidas por apresentarem duas visões diferentes, defendendo as duas principais teorias sobre o desacordo: a conformista e a não-conformista. Contudo, não é objetivo deste artigo defender o cálculo bayesiano nem a proposta dos graus de crença a ele acoplado por muitos estudiosos. Assim, analiso as conclusões das duas propostas e concluo com algumas considerações para o uso da epistemologia bayesiana na procura de uma solução para o problema do desacordo.

Para responder a isso, primeiro vou apresentar algumas teorias que tentam responder à questão do desacordo que não utilizam a teoria bayesiana. Depois apresento como funciona de modo rápido a abordagem bayesiana e como o desacordo entre pares passa a ser visto nessa abordagem. Por fim, analiso duas propostas: a de Shogenji (2007), mencionada por Christensen (2007) e outros que defendem teorias conformistas, e a de Lang (2014), que defende uma teoria não-conformista.

## 1. PRINCIPAIS TEORIAS SOBRE O DESACORDO

As principais teorias sobre o assunto se dividem em conformistas e não-conformistas (ou *teorias da permanência*)<sup>1</sup>. Além dessas, há a TET, de Kelly (2005; 2010), que também pode se encaixar como não conformista (ver RUIVO, 2016), e a Teoria Justificacionista, de Lackey (2010), que passeia entre as duas.

As teorias conformistas se apoiam no princípio chamado de *Perspectiva do Peso Equivalente* (PPE). De acordo com essa perspectiva, em um caso de desacordo entre

---

1 Frances e Matheson (2018) chamam as teorias não-conformistas de *steadfast view*, que preferi traduzir por teoria da permanência.

pares, deve-se dar o mesmo peso para ambas as posições, o que deveria levar ambos a suspender o juízo. Frances e Matheson (2018) apontam que essa perspectiva é uma combinação de três teses: 1) tese do derrotador – saber que um par discorda de você acerca de P lhe dá razões para você acreditar que sua crença está errada sobre P; 2) tese do peso equivalente – ambas as posições têm peso igual; 3) independência – as razões da crença do seu par acerca de P devem ser diferentes do próprio desacordo.

Os defensores dessa perspectiva utilizam principalmente três linhas argumentativas. A primeira se baseia em examinar casos centrais. O principal é o caso da conta no restaurante, apresentado por Christensen (2007, p. 193). Imagine o seguinte caso: você e mais cinco amigos vão a um restaurante, e, na hora de pedir a conta, decidem que todos pagarão valores iguais, independentemente de quanto comeram. Você vê o valor da conta e, mentalmente, chega ao resultado de que cada um pagará R\$ 43,00 enquanto sua amiga, que você reconhece também ser boa em contas, chega ao resultado de R\$ 45,00. Ora, imediatamente você suspende a crença em seu resultado e volta a fazer novas contas para tentar descobrir o resultado correto.

Neste caso, temos uma defesa da tese do derrotador, afinal, saber que seu par chegou a conclusões diferentes o fez repensar suas próprias conclusões. Ele também apoia a tese do peso equivalente, afinal, você considerou ambas as possibilidades como tendo a mesma probabilidade, suspendeu a crença e teve que refazer os cálculos. Por fim, ele também se encaixa na tese da independência, já que ambos chegaram a resultados diferentes independentemente de saberem o resultado do outro.

Uma segunda linha argumentativa utiliza-se do valor do testemunho como fonte de conhecimento/justificação. Como muito tem sido debatido sobre o valor do testemunho e as propostas normalmente o apresentam como confiável, então o testemunho de outra pessoa serve como evidência acerca de P que fará você repensar sua crença sobre essa proposição.

Os conformistas ainda utilizam uma terceira linha argumentativa que é utilizar analogias. Frances e Matheson (2018) apresentam a analogia do termômetro, em que o aparelho pega um pouco de informações do ambiente e chega a um resultado. Caso nova informação chegue ao termômetro, o resultado será alterado. Da mesma forma, a crença é formada através de uma quantidade de informação. O testemunho de outra pessoa é uma informação nova que alterará a crença final.

Contudo, essa posição apresenta alguns problemas. O primeiro é uma inconsistência, já que o conformismo leva à conclusão de que não existe desacordo entre pares. Isso porque, toda vez que alguém fica sabendo que outro discorda sobre o mesmo assunto, há suspensão da crença. Esse desdobramento é problemático porque é como se apresentar como uma resposta a um problema que não existe e parece evidente, pelo exemplo de Feldman mencionado no começo deste texto, quanto pelo que vivemos no dia a dia que existe o desacordo.

Além disso, o conformismo leva a uma falha argumentativa já que, como muitos não concordam com essa perspectiva, em casos de desacordo acerca do próprio problema do desacordo, a atitude pregada por essa teoria seria a de suspensão da crença, e não a afirmação da teoria conformista. Ou seja, a proposta conformista seria incapaz de resolver uma situação de desacordo acerca do próprio problema que ela tenta responder.

Christensen (2007) tenta responder a isso dizendo que apenas no caso do método não deve haver suspensão da crença. Isso leva a uma teoria conformista que não se conforma tanto assim. Há uma arbitrariedade aqui na utilização do conformismo. Ainda contra a PPE, Kelly (2013) mostra que ela se apoia na *Tese da Unicidade*. A Tese da Unicidade diz que, para um corpo de evidência E e a proposição P, E justifica, no máximo, uma atitude doxástica em relação a P. Em outras palavras, só uma crença pode ser justificada a partir de um mesmo grupo de evidências. Utilizando os alunos de Feldman como exemplo, cada um deles tinha sua crença, mas todos partiam das mesmas evidências. Segundo a Tese da Unicidade, apenas um deles, no máximo, teria crença justificada enquanto os outros não estariam justificados.

Kelly (2013) apresenta a tese de que a evidência é permissiva, especialmente em casos de pouca evidência. Ou seja, para uma evidência E há mais de uma atitude doxástica justificável em relação a P. Ele utiliza o seguinte caso para demonstrar isso: suponha que seis meses antes das eleições presidenciais dos EUA você possua evidência de que o candidato democrata irá vencer. Com base na mesma evidência, é, portanto, irracional afirmar que o republicano vencerá. Contudo, discutindo com um companheiro que disponha das mesmas evidências e que também concorda que o democrata será eleito você chegue ao seguinte desacordo: você acredita que ele vencerá com 60% dos votos, enquanto seu amigo acredita que a vitória será com 67% dos votos. Ora, Kelly (2011) argumenta que a evidência disponível permite essa discordância<sup>2</sup>. Neste caso, a teoria conformista estaria equivocada.

Com base nessas e outras críticas à teoria conformista, as teorias não-conformistas argumentam que a atitude a se tomar de modo justificado diante de um desacordo é manter a crença. Para defender isso de maneira mais eficaz, Kelly (2010) propõe a TET. Essa teoria leva em conta a distinção entre evidências de Primeira Ordem (EPO) e evidências de Ordem Superior (EOS) (FRANCES; MATHESON, 2018). As EPOs para P são diretamente pertencentes a P, enquanto as de Ordem Superior para P são evidências acerca da evidência de outrem para P. Kelly (2010) argumenta que EOSs, quando somadas às de Primeira Ordem, não devem ser consideradas como de igual valor.

Desse modo, utilizando o mesmo caso da conta do restaurante, caso sua amiga chegue à conclusão de que o valor a ser pago é R\$ 450,00 não há motivo para você mudar sua crença de que é R\$ 43,00, pois a diferença absurda, neste caso, fica evidente.

Ruivo (2016) afirma que considerar alguém como par epistêmico é uma generalização indutiva que não se impõe a um caso individual. Desse modo, a amiga continua sendo par por continuar sendo igualmente competente, mas o que aconteceu foi que, naquela ocasião, ela falhou, já que, mesmo que competente, todo ser humano é falível.

Além dessas teorias, Lackey (2010) apresenta a Teoria Justificacionista. Ela começa criticando uma visão que permeia as duas teorias acerca do desacordo. Ela chama essa visão de Uniformidade. A Uniformidade defende que a resposta ao desacordo deve ser

---

2 Essa crítica de Kelly tem aspectos que merecem ser considerados também. O desacordo em relação à porcentagem não implica, na teoria conformista, na suspensão da crença na vitória do candidato, mas a suspensão da crença na porcentagem específica com a qual o candidato será eleito. Pode-se contra-argumentar que a vitória numa eleição depende ao menos de uma porcentagem mínima. Contudo, a crença na proposição “o candidato terá uma porcentagem de votos que está acima da porcentagem mínima para eleição” não é, necessariamente afetada pela suspensão da crença na proposição “o candidato terá uma porcentagem de votos específica de 67%”.

a mesma para todas as ocasiões. Para defender a crítica à Uniformidade, ela também critica a paridade epistêmica e aponta dois tipos de desacordo.

Conforme já exposto no início deste artigo, a paridade epistêmica deve satisfazer os critérios de igualdade evidencial, igualdade cognitiva e abertura total (*full disclosure*). Lackey (2011) critica especialmente este último. A abertura total foi sugerida por Feldman (2006) como critério para que de fato haja uma paridade. Contudo, que os pares revelem todas as razões para a posição que tomam não apenas parece muito demandante, mas também irrealizável. Por isso, casos em que os pares cumprem os critérios são chamados por Lackey (2011) de *desacordo idealizado* enquanto os desacordos que ocorrem com frequência no dia a dia não passam de *desacordos ordinários*.

Não obstante, podemos pensar que, se a satisfação dos critérios de paridade é muito demandante, diferenciar um desacordo idealizado de um ordinário também o seria. Mas, Lackey (2011) coloca o resultado como critério de diferenciação entre ambos. O desacordo idealizado é aquele em que, após tomarem conhecimento da posição contrária do par e da argumentação deste, ambos permanecem com a posição que têm sem deixar de considerar a posição do outro como racional. Desse modo, ela também está indo contra a Tese da Unicidade, afirmando ser possível que um mesmo corpo de evidência permita atitudes doxásticas diversas acerca de uma mesma proposição.

Porém, Lackey (2011) desenvolve um pouco mais sua tese ao apresentar que tipo de evidência seria permissiva ao suportar mais de uma atitude doxástica. Para ela, a distinção se dá pela fonte da justificação da evidência. Fontes que ela considera como mais confiáveis como percepção e memória espacial não seriam permissivas. Já fontes menos confiáveis como o cálculo mental do caso do restaurante permitiria mais de uma atitude doxástica.

Assim, com a Teste Justificacionista, a resposta varia de conformista a não-conformista dependendo da situação. E a situação se define principalmente, para Lackey (2011), pela fonte da justificação da EPO.

A seguir, apresentarei de modo resumido a epistemologia bayesiana e como ela pode se relacionar com o problema do desacordo entre pares.

## **2. EPISTEMOLOGIA BAYESIANA**

O teorema de Bayes para probabilidade tem sido usado pela epistemologia formal para estudar casos como o do desacordo. A epistemologia bayesiana, grosso modo, defende que crença aparece em graus. Quando não se utilizam graus em epistemologia, apenas três atitudes doxásticas são previstas: crença que  $p$ , crença que não- $p$  e suspensão da crença. O caso do desacordo entre pares evidencia a necessidade do uso de graus especialmente na proposta conformista. Conforme vimos, quando duas pessoas em situação de paridade têm atitudes doxásticas diferentes acerca da mesma proposição (digamos que um crê que  $p$  e outro crê que não- $p$ ) o que ambos deveriam fazer seria suspender a crença (posição intermediária entre as duas atitudes). Contudo, quando há um desacordo na seguinte situação: um crê que  $p$  e outro já está em suspensão de crença em relação a  $p$ , que atitude deveriam tomar? A suspensão da crença em ambos não é coerente com a visão conformista já que um dos dois mantém sua posição, por outro lado, ambos passarem a crer que não- $p$  não parece uma atitude

intuitiva. Por isso, parece haver necessidade de estados intermediários entre a crença que  $p$ , a suspensão da crença e a crença que não- $p$ . Em outras palavras, parece haver a necessidade de graus de crença.

Além disso, intuitivamente percebemos que temos maior convicção em relação a algumas proposições do que em relação a outras. Creio que há um computador em minha frente enquanto escrevo este texto com maior convicção do que creio que o homem foi à lua. Creio em ambos, mas se tivesse de apostar, teria mais vontade de apostar na verdade do computador em minha frente do que na ida do homem à lua. Tais níveis de convicção são chamados de graus de crença.

Esses graus seguiriam, conforme a epistemologia formal, princípios de probabilidade. Por isso, se adota o teorema bayesiano de probabilidade. A base desse teorema é a expressão da probabilidade condicional, que considera a probabilidade da hipótese  $H$  em caso da evidência  $E$  é igual à probabilidade de  $H$  e  $E$  serem o caso dividido pela probabilidade de  $E$  ser o caso. Assim se expressa esse teorema:

$$P(H|E) = (P(H\&E)) / (P(E))$$

Começamos a construir outra equação de probabilidade invertendo-a para se descobrir a probabilidade de  $E$  dado que  $H$ :

$$P(E|H) = (P(E\&H)) / (P(H))$$

Já que  $P(H\&E)$  e  $P(E\&H)$  são os mesmos (a ordem dos fatores não altera o produto), podemos isolá-lo na segunda equação e substituí-lo na primeira:

$$P(E|H).P(H) = P(E\&H)$$

$$P(H|E) = (P(H).P(E|H)) / (P(E))$$

Esta última fórmula será a mais utilizada em cálculos na epistemologia bayesiana.

Existem algumas restrições exigidas para que o teorema bayesiano possa ser usado para o estudo de graus de crença. A primeira é o probabilismo ou axioma de Kolmogorov, que apresenta três restrições para que uma função seja probabilística (BRADLEY, 2015): as proposições  $X$  e  $Y$  num conjunto de crenças  $B$  formam uma função probabilística se e somente se essa respeita as seguintes restrições:

1. Não-negatividade:  $P(X) \in \mathbb{R}$ ,  $P(X) \geq 0$  para todo  $X \in B$ ;
2. Normalização:  $P(T) = 1$  para toda tautologia  $T \in B$ ;
3. Aditividade finita:  $P(X \cup Y) = P(X) + P(Y)$ .

Podemos defender cada uma das restrições. A 1) (não-negatividade) implica em que nenhuma probabilidade pode ser negativa. Isso provém de uma convenção para ajudar na análise de probabilidades. A probabilidade de que não- $p$  deve ser 1 menos a probabilidade de que  $p$ , indicando a restrição 3). Sendo 1 a convicção total, o maior grau de crença, então uma tautologia ( $p$  ou não- $p$ ) deve receber grau de crença máximo, o que evidencia a restrição 2).

A segunda restrição é a condicionalização. De acordo com ela, o grau de crença racional acerca de uma hipótese  $H$  ao saber que a evidência  $E$  é o caso é a probabilidade condicional de  $H$  dado que  $E$ . Sendo  $P_E H$  o grau de crença em  $H$  depois de saber que  $E$  é o caso, temos que

$$P_E H = P(H|E) = P(H \& E) / P(E) = (P(H) \cdot P(E|H)) / (P(E))$$

À luz da epistemologia bayesiana, pode-se dizer (adaptado de LANG, 2014, p. 45) que os sujeitos  $A$  e  $B$  discordam acerca da hipótese  $H$  se e somente se

1.  $P_A(H) = x$ ;
2.  $P_B(H) = y$ , onde  $x \neq y$ ;
3.  $A$  e  $B$  estão conscientes de que têm atitudes doxásticas diferentes acerca de  $H$ .

Deste modo, um desacordo não ocorre apenas quando dois sujeitos têm atitudes doxásticas opostas, mas quando têm atitudes doxásticas diferentes. Utilizando o exemplo de Kelly (2013) acerca da atitude doxástica em relação ao resultado das eleições, atitudes doxásticas opostas, num modo simples, ocorre quando um sujeito  $A$  crê que o candidato  $X$  vencerá as eleições e  $B$  crê que o mesmo candidato irá perdê-las. Atitudes doxásticas diferentes ocorrem quando os sujeitos  $A$  e  $B$  creem que o candidato  $X$  será eleito, mas os dois sujeitos têm diferentes graus de crença em relação a essa proposição.

### 3. BREVE ANÁLISE DE DUAS PROPOSTAS

Tendo em vista os conceitos apresentados, analisaremos duas propostas a seguir. A de Shogenji (2007) e a de Lang (2014). A primeira tenta apontar uma abordagem conformista a partir da epistemologia bayesiana e a segunda se propõe a analisar o desacordo como EOS e esta como um derrotador-solapador. A partir desses pressupostos, Lang (2014) apresenta algumas conclusões que serão expostas a seguir.

#### 3.1. A VISÃO CONFORMISTA DE SHOGENJI

Shogenji (2007) considera a visão conformista mais coerente e tenta adequá-la à epistemologia bayesiana, mas esbarra em alguns problemas. Para ele, o grau de crença final ( $C^*$ ) dos sujeitos 1 e 2 acerca de uma proposição  $p$  é dado pela soma dos graus de crença iniciais iniciais de 1 ( $C_1$ ) e 2 ( $C_2$ ) em que cada parcela da soma é multiplicada por um peso  $q$  que é proporcional à qualificação epistêmica dos sujeitos. Desse modo, se os sujeitos são pares epistêmicos, tem-se que  $q_1 = q_2$ . Seguindo a aditividade infinita, tem-se que  $q_1 + q_2 = 1$ , conclui-se que

$$C^*(p) = q_1 C_1(p) + q_2 C_2(p)$$

$$C^*(p) = 1/2 C_1(p) + 1/2 C_2(p)$$

Acontece que essa proposta não se adéqua à condicionalização. A tentativa de Shogenji (2007) de resolver isso é deixar de lado a média aritmética para alcançar o

grau de crença final e tentar outra forma de encontrar uma média. Ele, então, tenta com média geométrica, que se calcula assim:  $(\prod_{i=1}^n x_i)^{(1/n)}$ . Contudo, essa média não respeita a restrição da aditividade finita do probabilismo bayesiano.

Essa é uma situação muito problemática, pois ou a proposta não se adéqua à condicionalização ou desrespeita a restrição da aditividade finita. Shogenji chama essa situação de duplo vínculo (*double bind*). O modo que ele encontra para fugir desse duplo vínculo é utilizar uma asserção relaxada de uma asserção livre de contexto (*context-free assumption*), ou seja, fazer com que o resultado não dependa de  $C_1(p)$  e  $C_2(p)$ . Neste caso, o grau de crença final  $C^*$  é produtivo do grau de crença do sujeito 1 em  $p$  ( $C_1(p)$ ) elevado a  $q_1$  dividido pela somatória do produtivo do grau de crença do sujeito 1 em não- $p$  ( $C_1(\neg p)$ ) elevado a  $q_1$ .

$$C^*(p) = \prod_{i=1}^n C_1(p)^{(q_1)} / (\sum_{j=1}^m (\prod_{i=1}^n C_1(\neg p)^{(q_{-1}^j))})$$

O problema é que tal grau de crença viola o Princípio da Unanimidade, ou seja,  $C^*(p)$  não é uma agregação de  $C_1(p)$  e  $C_2(p)$ . Shogenji (2007) então propõe que se respeite a unanimidade apenas para casos em que os graus de confiança sejam básicos e não derivados de outros.

Existe outro problema com essa proposta. Os resultados que ela apresenta são contra intuitivos. Se, por exemplo,  $C_1$  for 0,3 e  $C_2$  for 0,7,  $C^*$  fica perto de 0,15. Parece complicado aceitar essa proposta por ser tão contraintuitiva, pois, ao se deparar com um par com grau de crença maior que o meu em uma proposição, a posição que essa proposta sugere para que eu esteja tomando uma atitude racional é diminuir meu grau de crença e não aumentá-lo. Por isso, autores posteriores como Christensen (2007) a consideram como inválida. Contudo, Shogenji (2007) defende-se afirmando que a intuição se equivoca algumas vezes e pode ter se equivocado aqui. Além disso, podemos mencionar que, conforme Ruivo (2016), o problema do desacordo deve apontar aquilo que os pares devem (estão no direito de) fazer e não o que acontece intuitivamente.

### 3.2. EVIDÊNCIA DE SEGUNDA ORDEM COMO DERROTADOR-SOLAPADOR EM LANG

Lang (2014), por outro lado, procura não ter nenhuma teoria de resposta em mente, mas começa com alguns pressupostos. O pressuposto com o qual ele trabalhará e defenderá é que o desacordo pode ser resolvido quando se descobre o que fazer com a EOS e ele considera essa evidência como um derrotador-solapador. Lang (2014) distingue dois tipos de derrotadores: o derrotador-refutador e o derrotador-solapador. Por derrotador, ele entende  $D$  como tal se e somente se:

1.  $R$  é uma razão epistêmica para o sujeito  $S$  acreditar que  $p$ .
2.  $(RAD) \Rightarrow$  sem razões para acreditar que  $p$ .

Um derrotador-refutador não apenas anula a razão  $R$  para crer que  $p$  como também é uma razão  $R_1$  para crer que não- $p$ . Um derrotador-solapador apenas anula a razão

R para crer que p. Lang (2014) utiliza como base para expor cada tipo de derrotador algum acontecimento na nave Enterprise, de Star Trek.

No caso do derrotador-refutador, a tripulação do Enterprise está em um batalha. Para se protegerem, ela utiliza um escudo que tampa a visão do que está fora da nave, mas os instrumentos de Spock permitem analisar as coisas que estão no lado externo. Os instrumentos de Spock passam a proposição  $p_{\text{destruída}}$  (a nave inimiga foi destruída) e Spock, sendo plenamente racional, forma a crença  $B_{\text{destruída}}$ . Contudo, ao baixarem os escudos, toda a tripulação pode ver que a nave inimiga não foi destruída. Nesse caso, não há mais razão para manter a crença  $B_{\text{destruída}}$  e há razão suficiente para acreditar em não- $p_{\text{destruída}}$ .

Aqui, vamos trabalhar inicialmente com a probabilidade de dois termos. Primeiro, a probabilidade de  $p_{\text{destruída}}$  e a probabilidade do relato de  $p_{\text{destruída}}$  ( $\text{Rel}_{\text{Pdestruída}}$ ) e a probabilidade de  $p_{\text{destruída}}$  dado que  $\text{Rel}_{\text{Pdestruída}}$ . Utilizando a fórmula da epistemologia bayesiana, temos a fórmula 1.

$$P(P_{\text{destruída}} | \text{Rel}_{\text{Pdestruída}}) = (P(\text{Rel}_{\text{Pdestruída}} | P_{\text{destruída}}) \cdot P(P_{\text{destruída}})) / (P(\text{Rel}_{\text{Pdestruída}}))$$

O primeiro termo ( $P(P_{\text{destruída}} | \text{Rel}_{\text{Pdestruída}})$ ) indica a confiabilidade do relato porque evidencia a probabilidade do relato apontar que a nave foi destruída caso esse seja o caso. Vamos chamar o parâmetro de confiabilidade de r. Para que r indique, de fato, uma confiabilidade racional, Lang (2014) divide esse parâmetro em dois: a razão (v) verdadeiro positivo e a razão (f) falso positivo. (v) é a probabilidade do relato avisar que algo é o caso quando, de fato, o for. (f) é a probabilidade do relato avisar que algo é o caso, quando não o for ( $P(\text{Rel}_p | \neg P)$ ). A razão de confiabilidade geral de um relato pode ser dada, então, pela razão  $r = 1 - f/v$ .

O segundo termo ( $P(P_{\text{destruída}})$ ) é a probabilidade prioritária, porque é a probabilidade de que  $p_{\text{destruída}}$  independente do relato. Ela pode influenciar consideravelmente o resultado final do grau de crença de Spock no relato, mesmo que este seja muito confiável, porque, por exemplo, caso Spock nem tenha acionado as armas contra a nave inimiga, ele vai saber que a é praticamente zero e isso, tecnicamente, zeraria qualquer crença num relato sobre a destruição da nave.

O último termo é o denominador ( $P(\text{Rel}_{\text{Pdestruída}})$ ). De novo, temos um fator de confiabilidade do relato, porque ele pode apontar que p quando esse for o caso e quando não for o caso. Assim, Lang (2014), propõe expandir o denominador para abranger ambos os casos (verdadeiro positivo e falso positivo). Para isso, é necessário multiplicar o verdadeiro positivo com a probabilidade de p e multiplicar o falso negativo com a probabilidade de  $\neg p$ , para que a fórmula continue em equilíbrio.

$$P(\text{Rel}_{\text{Pdestruída}}) = P(\text{Rel}_{\text{Pdestruída}} | P_{\text{destruída}}) \cdot P(P_{\text{destruída}}) + P(\text{Rel}_{\text{Pdestruída}} | \neg P_{\text{destruída}}) \cdot P(\neg P_{\text{destruída}})$$

Conforme já vimos, podemos chamar a probabilidade  $P(\text{Rel}_{\text{Pdestruída}} | P_{\text{destruída}})$  de v e a  $P(\text{Rel}_{\text{Pdestruída}} | \neg P_{\text{destruída}})$  de f. Assim,

$$P(\text{Rel}_{\text{Pdestruída}}) = v \cdot P(P_{\text{destruída}}) + f \cdot P(\neg P_{\text{destruída}})$$

Agora, podemos voltar à fórmula 1) e fazer a substituição dos termos.

$$P(P_{\text{destruída}} | \text{Rel}_{\text{Pdestruída}}) = v \cdot P(P_{\text{destruída}}) | v \cdot P(P_{\text{destruída}}) + f \cdot P(\neg P_{\text{destruída}})$$

Podemos cortar o v de cima e de baixo acrescentando a fração no segundo termo na soma do denominador:

$$\begin{aligned} & v \cdot P(P_{\text{destruída}}) | v \cdot P(P_{\text{destruída}}) + f \cdot v | v \cdot P(\neg P_{\text{destruída}}) = \\ & v \cdot P(P_{\text{destruída}}) | v \cdot P(P_{\text{destruída}}) + v \cdot f | v \cdot P(\neg P_{\text{destruída}}) = \\ & v \cdot P(P_{\text{destruída}}) | v \cdot P(P_{\text{destruída}}) + f | v \cdot P(\neg P_{\text{destruída}}) \end{aligned}$$

Desse modo, obtemos a fórmula da fonte 1 (os instrumentos de Spock, no caso):

$$P(P_{\text{destruída}}) | P(P_{\text{destruída}}) + f | v \cdot P(\neg P_{\text{destruída}})$$

A fórmula da fonte 2 (os olhos, no caso) se desenvolve de forma semelhante, mas como ela afirma que não-p, então é só inverter os valores de “não” e obteremos:

$$P(P_{\text{destruída}}) | P(P_{\text{destruída}}) + v | f \cdot P(\neg P_{\text{destruída}})$$

Para analisar o efeito que a fonte 2 tem sobre o grau de crença que a fonte 1 havia proporcionado acerca de p, basta utilizar na fórmula da fonte 2, como  $P(P_{\text{destruída}})$ , o resultado da fórmula da fonte 1. Voltando ao exemplo inicial de Spock, como a primeira fonte é os instrumentos de Spock, supondo que eles tenham  $v=0,9$  e  $f=0,05$  e, como Spock não sabe nada acerca da nave inimiga a  $P(P_{\text{destruída}})$ , o resultado da fonte 1 é  $P^1=0,9474$ .

A fonte 2 são os olhos de Spock, tradicionalmente mais confiável que os aparelhos. Quando ele vê que a nave inimiga não foi destruída, ela vai contra o que os aparelhos dizem. Supondo que, dos olhos, o  $v = 0,99$  e o  $f = 0,01$  e a  $P(P_{\text{destruída}})$ , que é o resultado da fonte 1, o grau de crença final (após o relato dos instrumentos e após ver com os próprios olhos) de Spock em  $p_{\text{destruída}} = 0,1539$ . Em outras palavras, a fonte 2 basicamente derrotou a fonte 1 em um nível que Spock ficou muito tendente a crer que  $\neg P_{\text{destruída}}$ , o contrário do que a fonte 1 dizia. Isso é um exemplo de derrotador-refutador.

No caso de um derrotador-solapador, vamos retomar o exemplo da Enterprise em batalha. Mais uma vez, a tripulação utiliza um escudo que tampa a visão, mas os instrumentos de Spock permitem analisar o que está do lado de fora. Os instrumentos afirmam que  $p_{\text{destruída}}$  e, então, Spock passa a sustentar a crença  $B_{\text{destruída}}$ . Contudo, o tenente Montgomery avisa Spock que a região em que estão tem muitas erupções solares e essas erupções atrapalham o funcionamento dos instrumentos de Spock, tornando-os fontes não confiáveis. Neste caso, Spock deixa de crer em  $p_{\text{destruída}}$ , mas não tem razão para crer em  $p_{\text{destruída}}$ . Vamos ver como funciona esse tipo de derrotador-solapador na epistemologia bayesiana.

Como a fonte 1 é idêntica à fonte 1 do caso do derrotador-refutador, simplesmente vamos reutilizar sua fórmula. Mas a fórmula da fonte 2 deve mudar porque não é mais

sobre a proposição  $p_{\text{destruída}}$ , mas sobre o relato dos instrumentos  $\text{Rel}_{\text{Pdestruída}}$  ou seja, é uma informação sobre  $v$  e  $f$  da fonte 1.

Antes de continuar, Lang estabelece o que seria uma fonte não-confiável. Uma fonte totalmente confiável seria aquela que a  $P(\text{Rel}_p|P) = 1$  e a  $P(\text{Rel}_p|\neg P) = 0$ . Ou seja, a probabilidade de que o relato aponte que  $p$  quando  $p$  for o caso é 1 (todos os casos) e a probabilidade de que o relato aponte que  $p$  quando  $p$  não for o caso é 0 (nenhum caso). Um caso de não-confiabilidade é um caso  $a = P(\text{Rel}_p|P) = P(\text{Rel}_p|\neg P)$ . Ou seja, não-confiável é a fonte que emite relatos de que  $p$  quando  $p$  for e não for o caso em igual probabilidade.

Então, como a fonte 2 dá um relato de não confiabilidade sobre a fonte 1, temos a probabilidade de confiabilidade da fonte 1 (instrumentos) dado o relato de sua não confiabilidade provido pela fonte 2 ( $P(\text{Conf}_{ins})|\text{Rel}_{\neg\text{Confins}}$ ), que seria igual à fórmula da fonte 2, feitas as devidas substituições:

$$P(\text{Conf}_{ins}) = P(\text{Conf}_{ins}) = \frac{v}{v+f} \cdot P(\neg\text{Conf}_{ins})$$

É importante destacar que, nesta fórmula, o  $v$  e o  $f$  são do tenente Montgomery (a fonte 2), pois são eles que irão informar a não confiabilidade dos instrumentos. Então, para descobrirmos a  $P(\text{Conf}_{ins})$  inicial, basta lembrarmos que o denominador da fórmula da fonte 1 era parâmetro de confiabilidade. Num caso de um relato de não-confiabilidade, temos como  $v$  e  $f$  as seguintes fórmulas:

$$\begin{aligned} v &= P(\text{Conf}) + a \cdot (1 - P(\text{Conf})) \\ f &= a \cdot (1 - P(\text{Conf})) \end{aligned}$$

Substituindo o  $f$  no  $v$ , temos:

$$\begin{aligned} v &= P(\text{Conf}) + f \\ v - f &= P(\text{Conf}) \\ P(\text{Conf}) &= v - f \end{aligned}$$

Desse modo, se utilizarmos os mesmos valores iniciais que usamos no caso do derrotador-refutador para o  $v$  e para o  $f$  (0,9 e 0,05, respectivamente), temos que a inicial  $P(\text{Conf}) = 0,85$ . Suponhamos que os valores da fonte 2 (tenente Montgomery, no caso), sejam  $v = 0,85$  e  $f = 0,1$ , temos que a  $P(\text{Conf}_{ins})|\text{Rel}_{\neg\text{Confins}} = 0,4$ .

Como o valor da  $\text{Conf}_{ins}$  mudou de 0,85 iniciais para 0,4 após o relato, surge uma nova  $P(P_{\text{destruída}})$ . Para calculá-la, com base na mesma fórmula da fonte 1, mas com os novos valores de confiabilidade, é preciso descobrir os novos  $v$  e  $f$  da fonte 1. Para isso, vamos utilizar as mesmas fórmulas do  $v$  e  $f$  no caso de não-confiabilidade descritas logo acima. Com o novo valor de  $P(\text{Conf}_{ins})$ ,  $v = 0,58$  e  $f = 0,18$ . Assim, a nova  $P_2(P_{\text{destruída}}) = 0,76$ . A probabilidade inicial era 0,9474 e ficou 0,76 após o relato de não-confiabilidade. Esse valor não se parece com o de um derrotador, que, mesmo sendo apenas solapador, tiraria a razão para crer que  $p$ .

Por isso, Lang (2014) propõe que um derrotador-solapador não funcione como um caso de não-confiabilidade, mas como um caso de anti-confiabilidade. Ele define

anti-confiabilidade como um caso em que  $P(\text{Rel}_p|P) = 0$  e a  $P(\text{Rel}_p|P) = 1$ . Ou seja, todas as vezes em que o relato disser que p, não-p será o caso. Numa linguagem mais simples, uma fonte anticonfiável seria aquela que diz o oposto da realidade em todas as situações. E já que  $P(\text{Rel}_p|P) = 0$ , então,  $v = P(\text{Conf}^+)$ , onde  $P(\text{Conf}^+)$  é a probabilidade  $P(\text{Rel}_p|P)$  e  $f = P(\text{Conf})$ , onde  $P(\text{Conf})$  é  $P(\text{Rel}_p|P)$ . Utilizando essa informação de v e f, temos os resultados de que  $v = 0,5143$  e  $f = 0,309$ .

Com esses dados da anti-confiabilidade, a  $P_2(P_{\text{destruída}}) = 0,6246$ , o que é muito mais intuitivo. Lang (2014) repete os mesmos cálculos para mostrar que uma EOS funciona exatamente da mesma forma. Por isso, ele passa a tratar de um tipo de EOS que funcionaria de forma diferente, que seriam relatos probabilísticos. O exemplo que ele dá deriva de Christensen (2010) *apud* Lang (2014).

Dr. McCoy pede para que Spock participe de um experimento. Enquanto Spock bebe a água que lhe é oferecida, McCoy explica o experimento. Funcionaria da seguinte forma: Spock tomaria uma droga. De acordo com resultados de uma pesquisa anterior com essa droga, 80% das pessoas erravam as respostas de lógica que lhe eram feitas, sem perceberem que estavam errando. Sendo um sujeito muito racional e com grande capacidade lógica, Spock fica muito interessado em participar. McCoy, então, propõe que ele comece fazendo uma pergunta antes de aplicar a droga em Spock, para efeito de comparação. A primeira pergunta foi: "Suponha que todos os búfalos sejam ferozes e Ferdinando não é um búfalo feroz. Qual das seguintes alternativas procede das premissas expostas?". As alternativas eram a) Ferdinando é feroz; b) Ferdinando não é feroz; c) Ferdinando é um búfalo; d) Ferdinando não é um búfalo. Spock responde confiante a letra b). McCoy sorri e lhe diz que a droga já estava na água que Spock bebeu.

Nesse caso, há um relato de anticonfiabilidade. A questão aqui é diferente porque o relato não afirma que Spock é 80% anticonfiável, mas que 80% das pessoas que tomam a droga se tornam anticonfiáveis. Assim, o que acontece é que McCoy está 80% confiante de que a droga funcionará. Vamos chamar essa probabilidade que a fonte (McCoy) atribui a si mesma de  $\tilde{n}$ . A fonte está dizendo que ou p é o caso, com a probabilidade  $\tilde{n}$  (neste caso, McCoy está dizendo que a droga funcionou, sendo  $\tilde{n} = 0,8$ ) ou que p é o caso, com a probabilidade  $1 - \tilde{n}$  (a probabilidade de a droga não ter funcionado é  $1 - \tilde{n} = 0,2$ ).

Utilizando as fórmulas das fontes 1 e 2, podemos calcular a variação da probabilidade de que p (Spock está certo) quando se recebe um relato de que q é verdadeiro (a droga funcionou) com a probabilidade  $\tilde{n}$  de q e de q, com a probabilidade  $1 - \tilde{n}$ . A primeira variação, vamos chamar de  $\Delta^+$ , e a segunda de  $\Delta^-$ .

$$\begin{aligned}\Delta^+ &= (P^0(p))|(P^0(p) + f|v.P^0(\neg p)) - P^0(p) \\ \Delta^- &= (P^0(p))|(P^0(p) + v|f.P^0(\neg p)) - P^0(p)\end{aligned}$$

Considerando que o que está dentro do parêntese de  $\Delta^+$  é igual a  $P(p|\text{Rel}_p)$  e o que está dentro do parêntese de  $\Delta^-$  é igual a  $P(p|\text{Rel}_{\neg p})$  temos:

1.  $P(p|\text{Rel}_p) = P^{1+}(p) = P^0(p) + \Delta^+$
2.  $P(p|\text{Rel}_{\neg p}) = P^{1-}(p) = P^0(p) + \Delta^-$

Lembrando que ambos são influenciados pela probabilidade  $\tilde{n}$  ( $\Delta^+$  é influenciado por  $\tilde{n}$  e  $\Delta^-$  por  $1 - \tilde{n}$ ), Lang (2014) propõe que

$$P(p|Rel_p) = P^1(p) = P^0(p) + \Delta^+ \cdot \tilde{n} + \Delta^- \cdot (1 - \tilde{n})$$

Essa proposta faz todo sentido porque, quando a fonte está absolutamente certa acerca de sua informação,  $\tilde{n} = 1$ , o que nos leva à equação (1) acima. E quando a fonte está absolutamente certa que sua informação não procede,  $\tilde{n} = 0$ , o que nos leva à equação (2).

No caso de Spock e a droga que ele tomou na água, McCoy está dizendo que Spock está sob influência da (sendo alterado pela) droga com a probabilidade  $\tilde{n} = 0,8$ . Nesse caso, o  $\Delta^- = 0$  (o relato de que a droga não fez efeito não alteraria o grau de crença na resposta de Spock). Considerando o  $v_{Spock} = 0,8$  e o  $f_{Spock} = 0,2$  e o  $v_{McCoy} = 0,8$  e o  $f_{McCoy} = 0,2$  (vamos supor que sejam iguais), lembrando que nas fórmulas da fonte 2 (que é a que estamos considerando agora), o  $v$  e o  $f$  utilizados são os da fonte 2, no caso o  $v_{McCoy}$  e o  $f_{McCoy}$ . Utilizando a fórmula da fonte 1, chegamos a 0,8 como probabilidade inicial de Spock acertar a questão. Com a informação de que Spock já havia tomado a droga, refazendo os cálculos, levando em conta agora o  $\tilde{n}$ , essa probabilidade cai para 0,56.

Ou seja, de fato a EOS serviu como derrotador-solapador neste caso e é muito parecido com isso que funciona um caso de desacordo entre pares (ou mesmo um caso de desacordo que não seja entre pares). O  $\tilde{n}$  seria o grau de crença que a própria fonte tem sobre a informação que está dando. Quando há o desacordo, é intuitivo que o  $\tilde{n}$  mude, como aconteceu com Spock e McCoy, caso façamos os cálculos.

Utilizando esses cálculos, os resultados mostram que em caso de desacordo entre pares (num caso de pares, o  $v$  e o  $f$  devem ser considerados os mesmos para ambos os indivíduos já que ambos são igualmente confiáveis), o que tem grau de crença maior a respeito de  $p$  diminui seu grau de crença e o que tem grau de crença menor a respeito de  $p$  aumenta seu grau de crença. Isso se aproxima, de acordo com Lang (2014), da TET.

Como exemplo, tomemos um caso em que um paciente apresente numa radiografia uma mancha no pulmão. O Dr. Dolittle tem grau de confiança 0.6 de que seja câncer enquanto Dr. House tem grau de confiança 0.3 do mesmo. Dr. Dolittle e Dr. House consideram-se ambos pares epistêmicos com  $v = 0.9$  e  $f = 0.2$  cada um. Após saberem do grau de crença um do outro, o grau de crença final de cada será, respectivamente, 0.442 e 0.404.

Esta abordagem parece muito mais intuitiva além de ter a vantagem de não quebrar nenhuma restrição ou princípio da probabilidade bayesiana. Como mencionado, ela se aproxima muito da TET. Além disso, pode-se considerar que, atribuindo  $v$  e  $f$  como o mesmo valor para ambos os sujeitos, ela também se aproxima da PPE, servindo, na verdade, de um mediador.

Contudo, a questão que pode ser levantada é se, de fato, todo caso de desacordo entre pares seria como um derrotador-solapador (funcionando especificamente como um relato de anticonfiabilidade<sup>3</sup> e, mais especificamente, como uma EOS do tipo relato

---

3 Como vimos acima, o relato de não-confiabilidade não é suficiente para se encaixar na definição de derrotador-solapador.

probabilístico). Embora os resultados sejam bem intuitivos, haveria necessidade de se enquadrar o desacordo entre pares a essa definição tão exigente. No entanto, poderia ser contra-argumentado quanto a essa questão que apenas a construção das fórmulas se baseia numa definição muito demandante. Neste caso, mesmo que se utilize a definição inicial que demos para o desacordo numa proposta bayesiana, os cálculos se encaixam bem e dão resultados intuitivos. E, de fato, é o que acontece.

Além do mais, apesar de Lang (2014) afirmar que sua proposta se adéqua muito à TET, os resultados dela se encaixam de modo mais adequado à Teoria Justificacionista de Lackey (2010). Isso porque há casos em que, de fato, os pares terão razão de entrar num acordo e casos em que isso não ocorrerá. Isso dependerá do grau de confiabilidade das evidências, conforme a própria Lackey (2010) propôs. Na proposta de Lang (2014), o grau de confiabilidade das evidências seria o ñ de cada um dos membros do par.

#### 4. CONSIDERAÇÕES FINAIS

Este artigo analisou duas propostas da epistemologia bayesiana em casos de desacordo entre pares. O desacordo entre pares se dá quando dois indivíduos epistemicamente habilitados de modo igual têm atitudes doxásticas diferentes acerca de uma mesma proposição. Na epistemologia bayesiana, isso implica em atribuir graus de crença diferentes a uma mesma proposição.

A proposta de Shogenji (2007) de utilizar a epistemologia bayesiana para solucionar o problema do desacordo falha inicialmente no duplo vínculo (ou viola a condicionalização ou viola a normalização), e, quando ele consegue resolver os dois, os resultados são contra-intuitivos. A proposta de Lang (2014) se mantém fiel às restrições do teorema bayesiano e serve de evidência para a TET, e, mais especificamente, para a Teoria Justificacionista. Contudo, ela pode falhar por exigir uma definição de desacordo entre pares que seja muito demandante, embora tal definição demandante não seja necessária para a realização dos cálculos.

Assim, ao comparar as duas propostas, a que respeita as exigências da teoria bayesiana e cujos resultados são mais intuitivos é a que aponta para a Teoria Justificacionista. Mas as propostas são inúmeras e a abordagem bayesiana ao problema do desacordo está apenas começando.

#### REFERÊNCIAS BIBLIOGRÁFICAS

BRADLEY, D. *A critical introduction to formal epistemology*. New York: Bloomsbury, 2015.

CHRISTENSEN, D. Epistemology of disagreement: the good news. *Philosophical Review*, vol. 116, pp. 187-217, 2007.

FRANCES, B.; MATHESON, J. *Disagreement*. 2018. Available on: <<https://plato.stanford.edu/entries/disagreement/>>. Accessed on: 18/11/2018.

FELDMAN, R. Epistemological puzzles about disagreement. In: HETHERINGTON, S. (Ed.). *Epistemology futures*. Oxford: Oxford University Press, 2006.

KELLY, T. Peer disagreement and Higher Order Evidence. In: FELDMAN, R.; WARFIELD, T. (Eds.). *Disagreement*. New York: Oxford University Press, 2010.

\_\_\_\_\_. Evidence can be permissive. In: STEUP, M.; TURRI, J.; SOSA, E. (Eds.). *Contemporary debates in Epistemology*. Vol. 1. New York: Blackwell, 2013.

LACKEY, J. A justificationist view about disagreement's epistemic significance. In: HADDOCK, A.; MILLAR, A.; PRITCHARD, D. (Eds.). *Social Epistemology*. Oxford: Oxford University Press, 2010.

LANG, P. *Bayesian epistemology of disagreement*. Dissertation (Master in Philosophy). Universität Wien, Viena, 2014.

RUIVO, J. L. Algumas razões em defesa da posição não-conformista no desacordo entre pares. In: ROSA, L.; NEIVA, A. (Orgs.). *Proceedings of the Brazilian Group on Epistemology*. Porto Alegre: Editora Fi, 2016.

SHOGENJI, T. *A conundrum in Bayesian Epistemology of disagreement*. 2007 (Manuscrito não publicado). Available on: [www.fitelson.org/few/few\\_07/shogenji.pdf](http://www.fitelson.org/few/few_07/shogenji.pdf). Accessed on: 07/11/2018.

**Recebido em:** 24/Jun/2019 - **Aceito em:** 27/Fev/2020.