As Teorias de Revisão de Crenças e os Condicionais Contra-factuais

Diego Pinheiro FERNANDES* Wagner SANZ

12 de Junho de 2011

Palavras-Chave: revisão de crenças, condicionais contra-factuais, teste de Ramsey

1 Introdução

O estudo do processo de revisão de crenças, além da perspectiva de aplicação de seus resultados na computação, mostra-se altamente relevante para a compreensão conceitual de um tipo de inferência investigada pelo menos há seis décadas: a inferência de condicionais contra-factuais. O problema desse tipo de inferência, que apresenta tanta resistência à propostas de solução, é que em um argumento em defesa de um condicional contra-factual, isto é, em um argumento que utiliza hipóteses contra-factuais, as noções tradicionais de cogência e validade não funcionam. Nesse tipo de argumento não é suficiente que as premissas sejam verdadeiras e que haja consequência dedutiva das premissas para a conclusão.

Os critérios de cogência argumentativa já não podem se resumir a aceitar simplesmente premissas verdadeiras. É necessário que, além disso, as premissas sejam selecionadas de acordo com critérios de compatibilidade e relevância informativa. Por exemplo, seja que (B) Verdi é italiano e (C) Bizet é francês. Suponha (H) que eles são compatriotas ($\neg(B\&C)$). Agora, temos um argumento A_1 que parte de H e B, e conclui (D) Bizet é italiano, A_1 tem uma premissa verdadeira e claramente apresenta consequência lógica de H e B para D. Também temos outro argumento A_2 que parte de H e C e conclui (E) Verdi é francês. A_2 também tem uma premissa verdadeira e consequência lógica de H e C para E. A_1 e A_2 são conflitantes, como decidir qual é correto, se algum dos dois é?

O problema apresentado logo acima pode ser aproximado ao problema de revisar um conjunto de crenças da seguinte maneira. O caso realmente problemático com o qual as teorias de revisão de crenças tem que lidar é quando se deseja adicionar uma sentença H ao conjunto de sentenças K que representa o estado de crenças e essa sentença contradiz as sentenças já presentes em K. Para que H possa ser adicionada, K deve ser revisado de modo que a adição de H não resulte em um conjunto inconsistente. Esse processo de revisão se constitui de remover sentenças de K que estejam em conflito com H e, além disso, remover o mínimo de informações de K em tal processo. O problema está em que raramente esse processo é unívoco. Como no exemplo acima, suponhamos que B e C esteja em nosso estado de crenças K. Desejamos adicionar $\neg(B\&C)$ e manter a consistência. Há pelo menos duas formas de fazer isso, a primeira removendo C de K, e a segunda removendo B de K. Pela primeira forma, o argumento A_1 acima é asserível a partir de K, e pela segunda, o argumento A_2 é asserível a partir de K.

*Aluno do Mestrado em Filosofia da Faculdade de Filosofia da Universidade Federal de Goiás e bolsista da CAPES.

Assim, o problema da aceitabilidade de premissas em um argumento contra-factual é bastante similar ao problema de selecionar as sentenças que devem ser removidas de um conjunto K que representa um estado de crenças, quando se deseja adicionar uma sentença H a K que seja incompatível com as sentenças de K.

O objetivo deste texto é apresentar a teoria de revisão de crenças que ficou conhecida como "modelo AGM", e mostrar que as operações de mudanças de crenças são incompatíveis com um critério de adequação adotado por várias abordagens de condicionais contra-factuais, o teste de Ramsey.

2 O Modelo AGM

A teoria de revisão de crenças desenvolvida por ALchourrón, Gardenfors e Makinson em um artigo de 1985[1] se converteu na teoria de revisão de crenças modelo, que posteriormente foi chamado de "o modelo AGM de revisão de crenças". Esse modelo é caracterizado pela forma em que são representados os estados de crenças por um lado (a estática de crenças), e por outro, pelos critérios que delineiam as possíveis mudanças nos estados de crenças (a dinâmica de crenças).

Com respeito a estática de crenças, o modelo AGM tem três postulados elementares: (1) é postulado que uma crença sempre pode ser representada por uma sentença, (2) que um estado de crenças pode ser representado por um conjunto de sentenças e, (3) este conjunto é fechado sob consequência lógica. É consequência de (1) que toda crença tem que ter um conteúdo proposicional. Por (2), um agente tem a crença em A quando $A \in K$, em que K é o conjunto que representa o estado de crenças do agente, e tem a crença em $\neg A$ quando $\neg A \in K$. Se nem A nem $\neg A$ pertencem a K, se diz que o agente não tem opinião sobre A. Por (3), se um agente tem a crença em A, então ele está comprometido a ter como crença todas as consequências lógicas de A. Dado que uma sentença A tem infinitas consequências, o estado de crenças de um agente será representado por um conjunto infinito.

Com relação à dinâmica de crenças, por consequência de como é representada a parte estática, há três possíveis modificações em um conjunto de crenças: a contração, a expansão e a revisão. A expansão, representada por "+", é a simples adição de uma crença A a um conjunto K, em que o conjunto resultante K_A^+ é fechado por consequência lógica. Isto significa que na operação de expansão não é assegurada a consistência de K_A^+. A contração, representada por "–", é a eliminação de uma crença A de um conjunto K. Como K é fechado sob consequência lógica, isto deve ser efetuado de modo que o conjunto resultante K_A^- continue sendo fechado sob consequência lógica. Isto significa que o conjunto K_A^-, resultante da contração de K por A, não pode seguir contendo A implicitamente, como consequência das demais sentenças dele. A operação de revisão, representada por "*", é a adição de uma crença a um conjunto de crenças K, em que se deve assegurar que o conjunto resultante K_A^* seja consistente.

2.1 A Operação de Revisão

A operação de revisão é caracterizada de forma geral no modelo AGM pelos seguintes postulados.

Critério 1*: K_A^* é um conjunto de crenças para qualquer K e A.

O segundo critério assegura que o sistema nunca recusará a revisão por uma sentença A:

Critério 2*: $A \in K_A^*$.

[1][Alchourrón, Gärdenfors e Makinson 1985]

O terceiro garante que a revisão de K por A será sempre um subconjunto da expansão de K por A. Quando A for inconsistente com K, isso se dá trivialmente, e quando A não for inconsistente, $K_A^* = K_A^+$.

Critério 3*: $K_A^* \subseteq K_A^+$.

O próximo assegura a inclusão oposta, dado que $A \notin K$:

Critério 4*: $A \notin K \;\Rightarrow K_A^+ \subseteq K_A^*$.

A revisão sempre garante a consistência do conjunto resultante, exceto quando o seu argumento é ele mesmo inconsistente:

Critério 5*: $K_A^* = K_\perp \;\Leftrightarrow \vdash \neg A$.

O critério 6* diz que sentenças logicamente equivalentes produzem o mesmo resultado:

Critério 6*: $\vdash A \leftrightarrow B \;\Leftrightarrow\; K_A^* = K_B^*$.

Os critérios seguintes foram chamados "postulados adicionais" para a revisão, eles incluem a revisão por conjunções. A ideia é que revisar K por $A \wedge B$ tem que ser um subconjunto de revisar K por A e depois adicionar B (isso será o caso trivialmente se $\neg B \in K$).

Critério 7*: $K_{A\wedge B}^* \subseteq (K_A^*)_B^+$.

E o último afirma a inclusão oposta quando $\neg B \notin K_A^*$:

Critério 8*: $\neg B \notin K_A^* \;\Rightarrow\; (K_A^*)_B^+ \subseteq K_{A\wedge B}^*$

3 Revisão de Crenças e os Condicionais

Gärdenfors começou sua pesquisa na teoria de revisão de crenças com o propósito de investigar a inferência de condicionais e ser capaz de oferecer critérios de aceitabilidade razoáveis para condicionais. Em seu livro "Knowledge in Flux" ele é enfático: "A hipótese inicial deste capítulo é que as sentenças condicionais em várias formas são sobre mudanças de estados de crenças".[2]

Na abordagem em questão é utilizado um critério de aceitabilidade para condicionais já bastante frequente na literatura, conhecido como "teste de Ramsey". O critério é como se segue:

(RT) Aceite uma sentença da forma 'Se A, então C' em um estado de crença K se e somente se a mudança mínima necessária para aceitar A requer que C seja aceito.[3]

[2] The starting hypothesis of this chapter is that conditional sentences in various forms are about changes of states of belief. [Gärdenfors 1988, p. 146]

[3] *ibid.*

3.1 A incompatibilidade de RT com 4*

Há um problema quando se tenta juntar o critério RT com os operadores de revisão para obter uma análise dos contra-factuais. O postulado 4* é um postulado que, dado seu apelo intuitivo, deveria ser válido na lógica dos contra-factuais. Se o postulado na sua força total não pudesse ser aceito, pelo menos uma consequência dele, aqui denominada "critério de preservação"(P), deveria ser aceita:

> (P) Se uma proposição B é aceita em um dado estado de crenças K e A é consistente com as crenças K, então B é ainda aceito na mudança mínima de K necessária para aceitar A.

O critério P tem seu fundamento no critério de economia informacional e é fortemente endorsado na tradição bayesiana. Como Gärdenfors diz "[...] informação não é gratuita e nos não queremos remover crenças desnecessariamente"[4]. Assim, tanto o teste de Ramsey quanto o critério de preservação são de grande interesse na análise da dinâmica de crenças. Gärdenfors mostra, contudo, que eles não podem ser combinados sem trivialidade.

Para que seja possível mostrar os resultados de trivialidade da combinação do teste de Ramsey e do critério de preservação, faremos uma lista das assunções necessárias para a prova. Relembramos o teste de Ramsey:

> (RT) $A > C \ \in K$ sse $C \in K^*_A$.

Observamos que RT depende da seguinte assunção:

> (0) Conjuntos de crenças incluem proposições que contém o conectivo condicional $>$.

Embora essa assunção pareça bastante razoável, há pelo menos um autor[5], que colocou nela a culpa dos resultados de trivialidade.

Para que seja obtido um resultado ainda mais forte, não será utilizado na prova o teste de Ramsey propriamente dito, mas uma consequência dele denominada "critério de monotonicidade":

> (M) para todos conjuntos de crenças K e K' e todas as proposições A, se $K \subseteq K'$, então $K^*_A \subseteq K'^*_A$.

Outro requisito para prova é o critério de preservação (apresentado logo acima), que será formalizado assim:

> (P) Se $\neg A \notin K$ e $B \in K$, então $B \in K^*_A$.

Serão utilizadas três propriedades da operação de revisão, a primeira assere que um conjunto de crenças K revisado pela sentença A contém A:

> (I) $A \in K^*_A$.

A segunda propriedade da operação de revisão diz que o conjunto K revisado pela sentença A só será inconsistente no caso limite de A ser uma contradição:

> (II) Se $K \neq K_\perp$ e $K^*_A = K_\perp$, então $\vdash \neg A$.

E a terceira propriedade diz que um conjunto revisado é logicamente fechado:

> (III) $K^*_A = Cn(K^*_A)$

[4] "[...] information is not gratuituous and we do not want to give up beliefs unnecessarily." [Gärdenfors 1986, p. 82]

[5] Isaac Levi [Gärdenfors 1988, p. 164]

Finalmente, faremos uso de mais duas propriedades da operação de expansão. Pela operação de expansão do conjunto K pela sentença A se entende a união de K e A fechada por consequência lógica, em termos formais, $K_A^+ = Cn(K \cup \{A\})$.

(IV) $(K_A^+)_B^+ = K_{A\&B}^+$
(V) $K_{A\vee B}^+ \subseteq K_A^+$

Por sistema de revisão de crenças não-trivial se entende um sistema $< \mathfrak{K}, * >$ em que há pelo menos três proposições A, B e C tais que, tomadas de par em par, são disjuntas, *i.e.* $\vdash \neg(A\&B)$, $\vdash \neg(A\&C)$ e $\vdash \neg(B\&C)$, e que há um conjunto de crenças $K \in \mathfrak{K}$ que é consistente com as três A, B e C, *i.e.*, $\neg A \notin K$, $\neg B \notin K$ e $\neg C \notin K$.

Podemos enunciar agora o teorema de trivialização:

> *Teorema*: Não há sistema de revisão de crenças não-trivial que satisfaça todas as condições (1), (2), (M) e (P).[6]

4 Considerações Finais

As operações para revisão no modelo AGM, portanto, não podem ser compatibilizadas com o teste de Ramsey. À primeira vista diríamos que (RT) é o "culpado" pelos resultados de trivialização, por implicar o critério de monotonicidade (M). Isaac Levi, contudo, discorda[7]. Ele põe a "culpa" na assunção (0). O propósito da presente investigação é descobrir a causa dos resultados de trivialização e procurar propor alternativas plausíveis.

Referências

[Alchourrón, Gärdenfors e Makinson 1985]ALCHOURRÓN, C.; GÄRDENFORS, P.; MAKINSON, D. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, v. 50, n. 2, p. 510–530, 1985.

[Gärdenfors 1986]GÄRDENFORS, P. Belief revisions and the ramsey test for conditionals. *The Philosophical Reveiw*, v. 95, n. 1, p. 81–93, 1986.

[Gärdenfors 1988]GÄRDENFORS, P. *Knowledge in Flux*. [S.l.]: Mit Press, 1988.

6

Segue a prova. Assuma, para redução ao absurdo, que

(H) algum sistema $< \mathfrak{K}, * >$ não-trivial satisfaça (1), (2), (M) e (P).

Segue de (H) que (K_A^+, K_B^+, K_C^+, $K_{A\vee B}^+$, $K_{A\vee C}^+$, $K_{C\vee B}^+$) $\neq K_\perp$, Assim, $(K_A^*)_{B\vee C}^* \neq K_\perp$ e $B \vee C \in (K_A^*)_{B\vee C}^*$ por (I). Agora temos que $\neg B \notin (K_A^*)_{B\vee C}^* \;\vee\; \neg C \notin (K_A^*)_{B\vee C}^*$, visto que $\{B \vee C, \neg B, \neg C\} \vdash \perp$. Suponha que $\neg C \notin (K_A^*)_{B\vee C}^*$, então por (V) $K_{A\vee B}^+ \subseteq K_A^+$ e visto que $\neg A \notin K$, por (P) $K \subseteq K_A^*$. Por (I) temos $A \in K_A^*$, consequentemente $K_A^+ \subseteq K_A^*$ e $K_{A\vee B}^+ \subseteq K_A^*$. Disso, por (M) segue-se $(K_{A\vee B}^+)_{B\vee C}^* \subseteq (K_A^*)_{B\vee C}^*$. Assim, $\neg C \notin (K_{A\vee B}^+)_{B\vee C}^*$.

Suponha agora que $\neg(B \vee C) \in K_{A\vee B}^+$, então $(A \vee B) \to \neg(B \vee C) \in Cn(K)$ e $\neg B \wedge \neg(A \vee C) \in Cn(K)$, o que implica que $\neg B \in Cn(K)$, o que é absurdo. Portanto, $\neg(B \vee C) \notin K_{A\vee B}^+$. Disso se segue, por (P) que $K_{A\vee B}^+ \subseteq (K_{A\vee B}^+)_{B\vee C}^*$ e, por (I) que $B \vee C \in (K_{A\vee B}^+)_{B\vee C}^*$. Por fechamento temos $Cn((K \cup \{A \vee B\}) \cup \{B \vee C\}) \subseteq (K_{A\vee B}^+)_{B\vee C}^*$. Visto que $A \wedge C$ é inconsistente, $Cn(K \cup \{A \vee B, B \vee C\}) = Cn(K \cup \{B\})$. Assim, $K_B^+ \subseteq (K_{A\vee B}^+)_{B\vee C}^*$ e $B \in (K_{A\vee B}^+)_{B\vee C}^*$. Por (H) $\vdash B \to \neg C$, então, por fechamento, $\neg C \in (K_A^*)_{B\vee C}^*$, o que contradiz a hipótese feita anteriormente.

A demostração de que $\neg B \notin (K_A^*)_{B\vee C}^*$ também leva a um absurdo é similar.

[7] [Gärdenfors 1988, p. 164]