UNIVERSIDADE FEDERAL DE GOIÁS

FACULDADE DE FILOSOFIA

MESTRADO EM FILOSOFIA

As Teorias de Revisão de Crenças e os Condicionais Contrafactuais

Diego Pinheiro Fernandes

Goiânia
2012

UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE FILOSOFIA
MESTRADO EM FILOSOFIA

AUTORIZAÇÃO PARA PUBLICAÇÃO DE DISSERTAÇÃO EM FORMATO ELETRÔNICO

Na qualidade de titular dos direitos de autor, **AUTORIZO** a Faculdade de Filosofia da Universidade Federal de Goiás – UFG a reproduzir, inclusive em outro formato ou mídia e através de armazenamento permanente ou temporário, bem como a publicar na rede mundial de computadores (*Internet*) e na biblioteca virtual da UFG, entendendo-se os termos "reproduzir" e "publicar" conforme definições dos incisos VI e I, respectivamente, do artigo 5º da Lei nº 9610/98 de 10/02/1998, a obra abaixo especificada, sem que me seja devido pagamento a título de direitos autorais, desde que a reprodução e/ou publicação tenham a finalidade exclusiva de uso por quem a consulta, e a título de divulgação da produção acadêmica gerada pela Universidade, a partir desta data.

Título: As Teorias de Revisão de Crenças e os Condicionais Contrafactuais

Autor(a): Diego Pinheiro Fernandes

Goiânia, 04 de outubro de 2012.

Diego Pinheiro Fernandes – Autor

Wagner de Campos Sanz – Orientador(a)

DIEGO PINHEIRO FERNANDES

As Teorias de Revisão de Crenças e os Condicionais Contrafactuais

Dissertação apresentada ao Programa de Pós Graduação em Filosofia da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Filosofia.

Linha de Pesquisa: Filosofia

Orientador(a): Wagner de Campos Sanz

Goiânia
2012

Dados Internacionais de Catalogação na Publicação (CIP)
GPT/BC/UFG

	Fernandes, Diego P.
F363t	As Teorias de Revisão de Crenças e os Condicionais Contrafactuais [manuscrito] / Diego Pinheiro Fernandes. – 2012.
	112f.
	Orientador: Prof(a) Wagner de Campos Sanz
	Dissertação (Mestrado) – Universidade Federal de Goiás, Faculdade de Filosofia, 2012.
	Bibliografia.
	Índice Remissivo.
	1. Revisão de Crenças 2. Condicionais Contrafactuais I. Título.
	CDU: 37.015.3:101.2

DIEGO PINHEIRO FERNANDES

As Teorias de Revisão de Crenças e os Condicionais Contrafactuais

Dissertação apresentada ao Programa de Pós Graduação em Filosofia da Universidade Federal de Goiás como requisito parcial para obtenção do título de Mestre em Filosofia, aprovado em 04 de outubro de 2012, pela Banca Examinadora constituída pelos professores abaixo assinados.

Prof(a) Wagner de Campos Sanz
Presidente da Banca

Prof(a) Frank Thomas Sautter
Faculdade de Filosofia – UFSM

Prof(a) Alexandre Costa Leite
Faculdade de Filosofia – UNB

à Daniela, Mariana e à memória de meu pai

Agradecimentos

Este trabalho somente foi possível graças à bolsa de pesquisa concedida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Agradeço ao meu orientador Wagner Sanz por todas as boas sugestões, conselhos e pela paciência em ler e reler as numerosas versões deste trabalho. Agradeço ao Hermógenes e ao Bruno pelas discussões de grande proveito para mim durante as aulas do mestrado. Também agradeço ao Hermógenes pelo desenvolvimento da classe LaTeX utilizada neste texto. Gostaria de agradecer à Daniela e à Mariana o enorme apoio em todo o período de investigação e à Renata pelo carinho e pelas incansáveis leituras e correções deste trabalho.

Resumo

Fernandes, Diego P.. **As Teorias de Revisão de Crenças e os Condicionais Contrafactuais**. Goiânia, 2012. 112 páginas. Dissertação de Mestrado. Faculdade de Filosofia, Universidade Federal de Goiás.

Este trabalho tem como propósito investigar o processo de inferência dos condicionais contrafactuais por meio das teorias de revisão de crenças, como o modelo AGM, a dinâmica de bases de crenças, o Sistema de Manutenção das Razões (RMS) desenvolvido por Doyle e abordagens afins como a de Goodman e a de Rescher. Será analisado em pormenor o modelo AGM e mostraremos que a abordagem de contrafactuais do modelo AGM não é adequada. Também mostraremos que a abordagem de contrafactuais por um modelo mais restrito que o modelo AGM —a dinâmica de bases de crenças— também não é adequada. A partir do Sistema de Manutenção das Razões, tomando algumas ideias de Goodman, Parry e Rescher, desenvolveremos o esboço de uma variante que consegue lidar de forma adequada com os contra-exemplos para as outras teorias de contrafactuais.

Palavras-chave: Revisão de crenças, condicionais contrafactuais, modelo AGM, dinâmica de bases de crenças, RMS.

Abstract

Fernandes, Diego P.. **The Belief Revision Theories and the Counterfactual Conditionals**. Goiânia, 2012. 112 pages. Master's Dissertation. Faculdade de Filosofia, Universidade Federal de Goiás.

This work has the intention to investigate the inferential process of counterfactual conditionals by means of the belief revision theories, such as the AGM model, the belief base dynamics, the Reason Maintenance System (RMS) developed by Doyle and like approaches such as Goodman's and Rescher's. The AGM model will be analysed in detail, and we will show that the AGM model approach to the counterfactuals is inadequate. We also will show that the approach to the counterfactuals based on a stricter model —the belief base dynamics— still is not adequate. We will take the Reason Maintenance System and some ideas from Goodman, Parry and Rescher, and develop a sketch of a RMS derivative system that can deal well with the counterexamples presented against the other approaches to counterfactuals.

Keywords: Belief revision, counterfactual conditionals, AGM model, belief base dynamics, RMS.

Sumário

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 contrafactuais . Condicionais contrafactuais são condicionais do tipo "se a tivesse sido o caso, então c teria sido o caso", "se a fosse o caso, então c seria o caso" e outras variantes que tenham em comum o fato de que quem as afirma acredita na falsidade de ambos a e c.

O problema desse tipo de condicional, que apresenta tanta resistência à propostas de solução, é que em um argumento em defesa dele, *e.g.*, em um argumento que utiliza hipóteses contrafactuais, as noções tradicionais de cogência e validade não funcionam. Ou seja, 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

(v) Verdi é italiano e
(b) Bizet é francês.

Suponha que

(c) eles são compatriotas.

Agora, temos um argumento A_1 que parte de c e v, e conclui

(b') Bizet é italiano.

A_1 tem uma premissa verdadeira e claramente apresenta consequência lógica de c e v para b'. A partir de A_1 seria inferível o condicional "se Bizet e Verdi fossem compatriotas, Bizet seria italiano".

Também temos outro argumento A_2 que parte de c e b e conclui

(v') Verdi é francês.

A_2 também tem uma premissa verdadeira e consequência lógica de c e b para v'. A partir de A_2 seria inferível o condicional "se Bizet e Verdi fossem compatriotas, Verdi seria francês". Os condicionais sustentados por A_1 e A_2 são conflitantes e intuitivamente falsos. Mas, já que ambos têm argumentos com premissas verdadeiras e consequência lógica das premissas para a conclusão, em que reside sua falsidade? A falsidade de ambos somente pode residir no fato que o tipo de premissa utilizado em ambos argumentos A_1 e A_2 não é admissível em argumentos com hipóteses contrafactuais. Assim, premissas verdadeiras e consequência lógica das premissas para a conclusão não são suficientes para um argumento com hipóteses contrafactuais ser cogente. Portanto, de maneira geral, o estudo dos condicionais contrafactuais consiste na investigação de condições de admissibilidade para premissas em argumentos com hipóteses contrafactuais.

As primeiras abordagens dos condicionais contrafactuais, como as de Chisholm [Chi46] e Goodman [Goo47], procuravam resolver a questão das condições de verdade de um contrafactual sob a seguinte assunção: um contrafactual "se a fosse o caso, então c seria o caso" é verdadeiro se e somente se o antecedente a, juntamente com algum conjunto relevante de premissas S, implica o consequente c. Assim, toda a questão se resumia em buscar oferecer as condições sob as quais um conjunto S de premissas que suporta um contrafactual é relevante. Por exemplo, condições de relevância usualmente adotadas são as de que S contenha somente sentenças verdadeiras, de que S seja compatível com o antecedente a, de que S e a implique o consequente c, etc. Goodman, no decorrer de seu artigo propõe condições de relevância para S e mostra os eventuais problemas com tais condições, até chegar em uma condição razoavelmente complexa. Não obstante a complexidade dessa condição, Goodman mostra, através de um contra-exemplo, que ela também é falha. Ele afirma então que a única condição de relevância para premissas que ele conseguia elaborar para se livrar do contra-exemplo era a seguinte:

> um condicional "se a fosse o caso, então c seria o caso" é verdadeiro se e somente se existe um S tal que a e S são co-sustentáveis e tal que de a, S segue-se c,

em que "co-sustentável" significa que não é o caso que S não seria verdadeiro se a fosse verdadeiro.[i] Naturalmente, a utilização do critério de co-sustentabilidade para estabelecer a verdade de um contrafactual "se a fosse o caso, então c seria o caso" requer que a verdade de outro contrafactual "não é verdade que S não seria verdadeiro se a fosse verdadeiro" seja estabelecida e isso leva a um regresso infinito. Esse problema ficou conhecido como o "problema da co-sustentabilidade" e é com ele que a abordagem de Goodman ficou sem saída (o problema será exposto na seção 4.2 abaixo).

Posteriormente, Rescher [Res64] propõe uma nova abordagem, embora continue seguindo o caminho de Goodman de tentar definir as condições relevantes para a inferência de um contrafactual. Ele propõe que o problema dos condicionais contrafactuais seja abordado da perspectiva mais ampla do problema de como adicionar uma crença a a um estado de crenças K, quando a é incompatível com K, em que "estado de crenças" significa a totalidade das crenças de algum agente em um dado instante. Esse problema é denominado por Rescher de "o problema das hipóteses crença-conflitantes" [Res64, pág. 26]. Rescher argumenta que o problema

[i] Goodman argumenta [Goo47, pág. 120] que as duplas negações nesse critério não podem ser eliminadas.

de estabelecer a verdade de um contrafactual "se a fosse o caso, então c seria o caso" (que representaremos daqui em diante por $a > c$), é o mesmo problema de **revisar** um estado de crenças, isto é, o problema de estabelecer qual é o melhor método de resolver o conflito causado pela adição da hipótese contrafactual (ou mais exatamente, hipótese crença-conflitante) a um estado de crenças. No exemplo sobre Bizet e Verdi acima, em um estado de crenças K que contenha, entre outras crenças, $b, v, \neg c$. Para adicionar c a K deve-se assegurar que o estado resultante permaneça consistente. Uma forma de revisar K seria remover b e $\neg c$ de K (correlata ao argumento A_1 acima). Outra seria remover v e $\neg c$ de K (correlata ao argumento A_2 acima). Cada uma dessas formas de revisar K por c corresponde a escolha de qual condicional seria razoável inferir em tal circunstância.

Rescher propõe alguns critérios para escolher dentre as diversas formas de revisar um estado de crenças. Esses critérios estabelecem que no estado de crenças a ser revisado determinados tipos de crenças têm mais prioridade de retenção do que outras. No entanto, somente pelo fato de considerar os contrafactuais sobre essa perspectiva mais ampla do problema de revisar um estado de crenças, o problema da co-sustentabilidade de Goodman não desaparece. Rescher, não obstante, faz algumas considerações com o intuito de dissolver o problema (ver seção 4.3 abaixo).

Outra abordagem dos condicionais pelo processo de revisão de crenças já tinha sido sugerida por Ramsey, que formulou um critério de adequação que ficou conhecido por "teste de Ramsey". O critério é o seguinte [Ram50, pág. 247, nota 1]:

> [...] se duas pessoas estão argumentando sobre "se p, então q" e estão ambas em dúvida quanto a p, elas estão adicionando p hipoteticamente a seus estoques de conhecimentos e argumentando nessa base acerca de q.

Com a grande popularização da semântica de mundos possíveis proposta por Kripke para a lógica modal, não tardou surgir análises da noção de condicional contrafactual utilizando esse tipo de semântica. Stalnaker propõe uma abordagem dos contrafactuais, tomando o teste de Ramsey acima e reformulando-o em termos de mundos possíveis [Sta68, pág. 102]:

> [...] considere um mundo possível no qual a é verdadeira, e que de outro modo difira minimamente do mundo atual. [O condicional] "Se a, então b" é verdadeiro (falso) somente no caso em que B é verdadeira (falsa) nesse mundo possível".[1]

Seguindo essa abordagem de Stalnaker, Lewis [Lew73] construiu um sistema bastante intrincado que ficou conhecido como a "lógica dos contrafactuais". Nesse sistema *grosso modo* um condicional $a > c$ é verdadeiro (no mundo atual) se e somente se em todos os mundos mais similares ao mundo atual em que a é verdadeira, também c é verdadeira. O sistema de Lewis desde então passou a ser a referência de abordagem de contrafactuais.

Não obstante a sua popularidade e complexidade, o sistema de Lewis tem pelo menos dois problemas graves. O primeiro é que a noção de mundo possível é extremamente obscura e o segundo é que a noção de similaridade entre mundos possíveis consegue sê-lo ainda mais. Motivado pelo caráter intragável de tais noções, Gärdenfors tenta fazer uma abordagem sistemática dos condicionais em termos de estados de crenças, na linha de investigação proposta por Rescher. No entanto, ele toma como critério de adequação para a asseribilidade de contrafactuais a seguinte modificação do teste de Ramsey formulado por Stalnaker [Gär78, pág. 381]:

> Aceite $a > b$ em um estado de crenças P se e somente se a mudança mínima de P necessária para aceitar a requer que b seja aceito.[2]

Nessa abordagem de Gärdenfors as noções de mundo possível e similaridade são substituídas pelas noções mais palatáveis de estados de crença e mudança mínima. No sistema de Gärdenfors são propostos critérios para regular as mudanças de estados de crenças de modo que sejam mudanças mínimas. Ele propõe axiomas para uma lógica de contrafactuais que são correlatos a cada critério para mudança em estados de crenças e prova que seu sistema é equivalente em termos expressivos, à lógica "oficial" de Lewis.

Posteriormente, Gärdenfors alia seus esforços de investigação na teoria de revisão de crenças com Alchourrón e Makinson, que vinham estudando as mudanças que ocorriam em códigos legais, provocadas por derrogações e emendas. Desse trabalho conjunto surgiu o artigo seminal de 1985 [AGM85]. A partir das iniciais dos autores, a teoria de revisão de crenças contida nesse artigo ficou conhecida como modelo AGM. O modelo AGM é considerado o marco conceitual para então recém fundada teoria de revisão de crenças.

Dada a grande quantidade de resultados na teoria de revisão de crenças obtidos pelo modelo AGM, entre eles com ênfase a lógica de contrafactuais de Gärdenfors (que usa basicamente os mesmos critérios para a revisão que o modelo AGM), o modelo AGM tem um papel destaque na presente investigação.

No primeiro capítulo serão investigados pormenorizadamente os critérios empregados pelo modelo AGM, as críticas que receberam na literatura e suas contra-críticas.

No segundo capítulo será apresentada a abordagem dos condicionais contrafactuais de Gärdenfors. Posteriormente será exposto o resultado de Gärdenfors de que um postulado relativamente elementar que rege uma operação de mudança de crenças é incompatível com o teste de Ramsey.

No terceiro capítulo investigaremos um resultado de Tennant acerca da força dos postulados do modelo AGM e as consequências desse resultado para a análise dos condicionais contrafactuais que adota o modelo AGM como estrutura subjacente. Mostraremos que a lógica de Gärdenfors admite condicionais intuitivamente inaceitáveis. Em seguida, mostraremos que outra abordagem dos condicionais contrafactuais, tendo como teoria subjacente uma teoria de revisão de crenças muito mais estrita que o modelo AGM (a revisão por base de crenças), embora evite o resultado de Tennant, também admite condicionais implausíveis.

No quarto capítulo retomaremos, por assim dizer, aos primórdios da análise dos contrafactuais. Examinaremos a abordagem de Goodman, os problemas com os quais Goodman se deparou e algumas críticas e sugestões que sua abordagem recebeu na literatura. Em seguida será investigada a sugestão de Rescher de uma abordagem mais geral da questão dos condicionais contrafactuais em termos de revisão de crenças; a sua proposta de dissolução do problema da co-sustentabilidade e a resposta de Goodman. Finalmente, tomaremos algumas boas ideias de Goodman, Rescher, Parry, entre outros, e as adaptaremos em uma teoria de revisão de crenças denominada Sistema de Manutenção das Razões (RMS ou TMS) de Doyle [Doy79]. Mostraremos que o esboço de sistema resultante dessas adaptações, que denominamos RMS-C ("C" para "contrafactuais"), consegue elucidar alguns problemas que outras abordagens de contrafactuais não conseguiram.

Notação

As letras minúsculas $a, b, c, \ldots$ estão por sentenças e as letras maiúsculas $A, B, C, \ldots$ por conjuntos de sentenças. A lógica subjacente às teorias de revisão de crenças tratadas neste texto é **lógica proposicional clássica**. Os teóricos do modelo AGM não explicitam[ii] se utilizam nos seus resultados a noção de consequência lógica (semântica) ou a noção de consequência dedutiva (sintática). Adotaremos a mesma posição. A indiscriminação da relação de consequência não será prejudicial porque neste texto só será considerada a lógica proposicional. A relação de consequência lógica é representada por $\vdash$ e o operador de consequência lógica é representado por Cn. Quanto a Cn, é suposto apenas que ele satisfaz três propriedades: $K \subseteq Cn(K)$, $Cn(Cn(K)) \subseteq Cn(K)$ e se $K \subseteq J$, então $Cn(K) \subseteq Cn(J)$. Um conjunto K é **fechado sob** Cn (*i.e.*, $Cn(K) = K$) quando para todo φ tal que $K \vdash \varphi$, $\varphi \in K$. Por **teoria** entende-se um conjunto fechado sob Cn.

Os símbolos $\neg$, $\vee$, $\wedge$, $\rightarrow$, $>$ significam **negação, disjunção, conjunção, implicação material** e **implicação contrafactual**, respectivamente. As expressões "$\alpha \Rightarrow \beta$" e "$\alpha \Leftrightarrow \beta$" são abreviações de "se α, então β" e "α se e somente se β", respectivamente. No contexto de critérios, regras e definições, em geral utilizaremos o símbolo "&" para a conjunção "e".

De modo a simplificar e possibilitar a investigação, o conteúdo de uma crença de determinado agente será entendido como uma sentença declarativa. O estado de crenças de um agente será representado por um conjunto, seja de sentenças declarativas, seja de sentenças declarativas e suas justificativas (formadas a partir de sentenças declarativas). As operações de mudança de crenças são de três tipos: expansão, contração e revisão. A expansão de K por a é representada por K_a^+ e consiste na adição da crença designada pela sentença a ao estado de crenças designado por K. A contração de K por a é representada por K_a^- e consiste na remoção da sentença a de K. A revisão de K por a é representada por K_a^* e tem por objeto a adição de a a K de forma que o conjunto resultante K_a^* seja consistente. $K^\perp$ representa o conjunto de todas as sentenças da linguagem. O conjunto dos subconjuntos maximais por inclusão de K e que não implicam uma sentença a é representado por $K \perp a$. Deste modo, para que um conjunto Y pertença a $K \perp a$,

[ii] Ver [AGM85, pág. 511].

quando $a \in Cn(K)$, deve ocorrer que $a \notin Cn(Y)$ e para qualquer Y' tal que $Y \subset Y' \subseteq K$, segue-se que $a \in Cn(Y')$. Exemplo: seja $K = Cn(r, r \to p, p \to q)$, $K \perp q$ terá três elementos: $\{r, p \to q\}$, $\{p \to q, r \to p\}$ e $\{r, r \to p\}$. Os elementos de $K \perp a$ para K e a quaisquer serão chamados de **conjuntos resto**. O conjunto potência de K será representado por $\mathcal{P}(K)$.

Quanto à formatação, serão utilizadas no texto dois tipos de notas: de fim de capítulo e de rodapé. As notas de fim de capítulo somente conterão as citações originais em inglês e serão referidas pelos numerais arábicos. As notas de rodapé utilizarão os numerais romanos minúsculos.

Notas

[1]Consider a possible world in which A is true, and which otherwise differs minimally from the actual world. "If A, then B"is true (false) just in case B is true (false) in that possible world.

[2]Accept $a > b$ in a state if belief P if and only if the minimal change of P needed to accept A also requires accepting B.

Capítulo 1

O Modelo AGM de Revisão de Crenças

Serão analisados em detalhe os critérios empregados no modelo AGM, algumas críticas que esses receberam na literatura e algumas respostas dos autores do modelo AGM a essas críticas. Também analisaremos as diversas operações de mudanças de crenças que são definidas a partir de tais critérios.

1.1 Introdução

Gärdenfors, no fim da década de 1970, desenvolvia uma série de critérios de adequação para as operações de mudanças de crenças, isto é, as operações de adicionar ou remover crenças de estados de crenças [Gär78], mas até então não tinha ainda desenvolvido nenhuma construção explícita destas operações[i]. Alchourrón e Makinson, por outro lado, pesquisavam exatamente construções explícitas para essas operações [AM82]. Gärdenfors eventualmente tomou conhecimento que ele, Makinson e Alchourrón pesquisavam um mesmo tema sob diferentes perspectivas. A partir de então começaram a trabalhar juntos, publicando um artigo em 1985 [AGM85] que passou a ser a pedra fundamental das teorias de revisão de crenças. Nesse artigo eles mostram que as construções explícitas desenvolvidas por Alchourrón e Makinson se adequam perfeitamente aos critérios de Gärdenfors, e que todas essas operações caem sob um tipo de operação geral que se denominou "revisão (contração) de intersecção parcial". Os critérios e resultados presentes neste artigo caracterizaram o que ficou conhecido por "modelo AGM".

O modelo AGM consiste em uma série de critérios ou postulados acerca do que são crenças e o que representa um estado de crenças, assim como critérios que regulam as mudanças incidentes sobre os estados de crenças. Um estado de crenças é a totalidade das crenças de um agente em determinado instante. No modelo AGM, os estados de crenças são representados por conjuntos de sentenças fechados sob consequência lógica e serão designados por **conjuntos de crenças**. Se uma sentença a pertence a um conjunto de crenças K de certo agente, então dizemos que o agente **acredita** em a. Se $\neg a$ pertence a K, então diz-se que o agente **acredita**

[i]Exceto a mudança de adicionar uma crença a um estado de crenças consistente com ela.

em $\neg a$ e, se nem a nem $\neg a$ estão presentes em K, dizemos que o agente **ignora** a. O conjunto de crenças K de um agente pode mudar de três maneiras, com respeito a uma sentença a: (i) o agente pode perder a crença em a, isto é, remover a de K (em símbolos: K_a^-), esta mudança é denominada **contração**; (ii) o agente pode passar a acreditar em a, isto é, adicionar a a K (em símbolos: K_a^+), esta mudança é denominada **expansão**; (iii) o agente pode perder a crença em a e passar a acreditar em $\neg a$, isto é, adicionar $\neg a$ a K, quando a pertencia anteriormente a K (em símbolos: $K_{\neg a}^*$). Essa mudança é denominada **revisão**. No modelo AGM, um critério de adequação fundamental para as mudanças de crenças é a noção de mudança mínima, da qual falaremos adiante.

1.2 Do Critério de Mudança Mínima

Pode-se dizer que em geral as informações de que dispomos em nosso estado de crenças são úteis e a que obtenção de tais informações demanda alguma medida de esforço. Desse modo, é razoável requerer que, ao se efetuar uma mudança em um conjunto de crenças, deve-se efetuar uma mudança mínima, ou seja, que um mínimo de informação seja descartada para que dada sentença seja removida, no caso da contração, ou adicionada, no caso da revisão. Tem havido, não obstante, grande dificuldade em captar formalmente esse requisito intuitivo sobre as mudanças de crenças. Como medir a perda de informação? Uma resposta imediata seria: pela quantidade de informação perdida. Assim, diminuir a quantidade de informação perdida significaria minimizar a mutilação do estado de crenças que sofre a mudança. Neste caso, por exemplo, pode-se propor o critério:

se $p \in K$ & $p \notin K_a^-$, então $K_a^- \cup \{p\} \vdash a$.

Isto é, que seja mantida a maior quantidade de sentenças de K em K_a^-. Contudo, esse critério é problemático se os estados de crenças são representados por teorias, pois, se K é uma teoria e $a \in K$, então $a \vee b \in K$, para qualquer $b \notin K$. Agora, utilizando o critério proposto acima, se $K_a^- \cup \{a \vee b\} \nvdash a$, então $a \vee b \in K_a^-$. Ou seja, uma sentença que dependia exclusivamente de a em K (*i.e.*, $a \vee b$) é mantida mesmo quando a já não está mais disponível. Contudo, deveria ser levado em consideração o fato de que a permanência de $a \vee b$ em K_a^-, embora logicamente possível, não é razoável, já que a é **única** justificativa[ii] em K de $a \vee b$, $a \vee b$ será em K_a^- uma informação desprovida de valor, espúria.

Tennant oferece [Ten6b, pág. 491] uma formulação um pouco mais explícita dessa ideia: (i) K_a^- contém tanto de K quanto é prudentemente possível; (ii) seja Θ o conjunto de justificativas de uma determinada sentença $p \in K$ (isto é, cada $x \in \Theta$ é suficiente para justificar p), então $p \in K_a^-$ somente se algum $x \in \Theta$ pertence a K_a^-. Em outras palavras, todas as crenças de K_a^- tem pelo menos uma das justificativas suficientes que tinham em K. Assim, essa noção de mensuração

[ii] Por "justificativa" se entende, *grosso modo*, uma relação inferencial, de modo que a justifica b, quando de a se segue b.

da perda de informação leva em consideração não a quantidade de informação, mas sim o valor informativo.

Makinson, por outro lado, defende a mensuração da perda de informação por meio da **quantidade** de informação, dizendo que o problema não reside nessa noção mas sim em como são representados os estados de crenças no modelo AGM. Segundo ele, se os estados de crenças são representados como uma base finita de crenças, medir a perda de informação pela quantidade de informação perdida é um procedimento particularmente interessante. Diz Makinson [Mak87, pág. 392]: "[...] parece ao autor que a perda de conteúdo, mensurada por inclusão, é um tipo particularmente interessante de perda para considerar no contexto de contração de teorias."[3] Assim, os resultados contra-intuitivos que surgem da tentativa de mensuração da perda de informação por meios quantitativos surgem da aplicação das operações de mudança de crenças sobre o tipo errado de argumento: conjuntos fechados sob consequência lógica. As operações de contração e revisão deveriam ser aplicadas sobre bases finitas, como diz ele (*ibid.*): "[...] Contração na vida real, nós sugerimos, é aplicada a tais bases [conjuntos finitos], e não a conjuntos fechados sobre consequência."[4]

A mensuração da perda de informação adotada pelo modelo AGM é essa de quantidade de informação descartada. Procederemos agora à análise das operações de mudanças de crenças no modelo AGM, que tentam capturar de alguma forma essa noção de mudança mínima.

1.3 A Operação de Expansão

A operação de expansão (+) no modelo AGM é, por assim dizer, descompromissada com a consistência do conjunto resultante. Esta característica será apontada como causadora de resultados negativos na seção 2.4 abaixo. Os critérios para a expansão são expostos abaixo.

Os critérios para a expansão O primeiro critério assegura que o resultado da operação de expansão de K por a (K_a^+) é sempre um conjunto de crenças:

($^+$**-fecho**) K_a^+ é um conjunto de crenças (*i.e.*, fechado sob consequência lógica).

O critério $^+$**-sucesso** assegura que o sistema nunca rejeitará uma expansão por uma sentença qualquer. Assim,

($^+$**-sucesso**) $a \in K_a^+$.

O critério $^+$**-inclusão** garante que K seja um subconjunto de K_a^+. O que ocorrerá mesmo se a for inconsistente com K:

($^+$**-inclusão**) $K \subseteq K_a^+$.

Se a pertence a K, a mudança mínima para adicionar a a K é não alterar K, portanto

($^+$**-vacuidade**) $a \in K \Rightarrow K^+_a = K$.

Se $K \subseteq J$, então, mesmo que $\neg a \in J$, o resultado da expansão K^+_a tem que estar incluso na expansão J^+_a (que nesse caso seria igual à $K^\perp$, isto é, o conjunto de todas as sentenças da linguagem[iii]). Isso significa que a operação de expansão é monotônica:[iv]

($^+$**-monotonicidade**) $K \subseteq J \Rightarrow K^+_a \subseteq J^+_a$.

O último critério para expansão tenta captar a ideia de mudança mínima para a operação de expansão. A expansão de K por a deve ser uma operação mínima para adicionar a a K:

($^+$**-minimalidade**) K^+_a é o menor conjunto de crenças que contém K e a.

1.4 A Operação de Revisão

A operação de revisão ($*$) é uma operação de expansão segura: para **qualquer** sentença a, a revisão deve garantir que o conjunto resultante K^*_a seja consistente, exceto quando a é uma contradição, em que $K^*_a = K^\perp$. Um fator digno de nota da operação de revisão no modelo AGM é que ela não é monotônica: se $K \subseteq J$, então não é necessário que $K^*_a \subseteq J^*_a$, quando $\neg a$ ocorre em J mas não ocorre em K. Para que a possa ser adicionada consistentemente à J é necessário que $\neg a$ seja removida. A remoção de $\neg a$ em J pode remover sentenças que pertencem a K, por isso não está garantido que $K^*_a \subseteq J^*_a$. Abaixo (na seção 2.4) exporemos a crítica de Tennant à não-monotonicidade da operação de revisão no modelo AGM.

O problema da operação de revisão é que a retirada de uma sentença a quando pretendemos incluir $\neg a$ pode em geral ser feita de várias maneiras. E se faz necessário o desenvolvimento de algum método de escolha entre essas diversas maneiras. Um exemplo simples é: seja $K = Cn(b, b \to a)$ ($Cn(K)$ está para o conjunto das consequências lógicas de K). Ao revisar K por $\neg a$ há duas possíveis soluções: $K^*_{\neg a} = Cn(b, \neg a)$ ou $K^*_{\neg a} = Cn(b \to a, \neg a)$, cada um desses é um subconjunto maximal de K que não implica a (isto é, um elemento de $K \perp a$). Isso significa que a revisão, tal como regulada pelos postulados, não é uma função, pois seu valor não depende unicamente de K e de a. No modelo AGM é **assumido** que para cada conjunto de crenças K e sentença a, há uma única revisão que é **a** mudança mínima de K para acrescentar a. Gärdenfors diz [Gär88, pág. 54] que embora os postulados não determinem univocamente uma função de revisão, com a adição de certos fatores de natureza epistemológica, é possível obter uma função de revisão, isto é, uma operação cujo resultado é determinado unicamente pelos seus argumentos.

[iii]Isto se dá porque na lógica subjacente (clássica) é válido o princípio *ex contraditione quodlibet*: $a, \neg a \vdash b$, para qualquer b.

[iv]Uma operação é **monotônica** quando a relação de inclusão entre conjuntos é preservada quando esses conjuntos são submetidos à operação em questão.

Os critérios para revisão Os postulados propostos por Gärdenfors para a revisão são os seguintes. O primeiro, tal como ($^+$**-fecho**), assegura que o resultado da revisão seja um conjunto de crenças.

(***-fecho**) K_a^* é um conjunto de crenças.

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

(***-sucesso**) $a \in K_a^*$.

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

(***-inclusão**) $K_a^* \subseteq K_a^+$.

O quarto assegura a inclusão oposta, dado que $\neg a \notin K$:

(***-vacuidade**) $\neg a \notin K \Rightarrow K_a^+ \subseteq K_a^*$.

O quinto critério garante a consistência do conjunto resultante de uma revisão, exceto quando o argumento é ele mesmo inconsistente:

(***-consistência**) $K_a^* = K^\perp \Leftrightarrow \vdash \neg a$.

O sexto assegura que sentenças logicamente equivalentes produzem o mesmo resultado:

(***-extensionalidade**) $\vdash a \leftrightarrow b \Leftrightarrow K_a^* = K_b^*$.

Os critérios a seguir foram chamados "postulados suplementares" para a revisão. Eles tentam captar como deve se comportar 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 $\neg b \in K$):

(***-suplementar**$_1$) $K_{a \wedge b}^* \subseteq (K_a^*)_b^+$.

E o último afirma a inclusão oposta quando $\neg b \notin K_a^*$:

(***-suplementar**$_2$) $\neg b \notin K_a^* \Rightarrow (K_a^*)_b^+ \subseteq K_{a \wedge b}^*$.

1.5 A Operação de Contração

Há uma operação de contração quando uma crença é retirada de um conjunto e nenhuma nova crença é adicionada. O problema fundamental da operação de contração é que dado um conjunto K e uma sentença a, retirar a de K significa normalmente remover não só a mas também as sentenças que implicam a. Novamente, o problema é agravado porque em geral há várias maneiras distintas de remover uma dada sentença a de um conjunto K, cada uma delas removendo o mínimo de informação possível.

Qualquer uma dessas maneiras corresponde a escolha de um certo elemento Y de $\mathcal{P}(K)$, que contenha um máximo da informação contida em K, sem que a seja consequência lógica de Y. Como na operação de revisão, aqui também é suposto [Gär88, pág. 61] que a operação de contração é uma função, ou seja, o resultado de uma contração é sempre único. E, de forma análoga à operação de revisão, somente os postulados para contração não serão suficientes para obter tal unicidade.

Os critérios para contração O primeiro critério assegura que o resultado da contração ($^-$) seja uma teoria:

(**–fecho**) K_a^- é um conjunto de crenças.

O segundo critério assegura que novas crenças não sejam adicionadas em uma operação de revisão, isto é, que a operação de contração é efetivamente contração e não uma operação de revisão. Para tal, é requerido que o resultado da contração seja subconjunto do conjunto inicial:

(**–inclusão**) $K_a^- \subseteq K$.

O terceiro afirma que se $a \notin K$, não há motivos para retirar nada de K:

(**–vacuidade**) $a \notin K \;\Rightarrow\; K_a^- = K$.

O quarto assegura que quando a não é um teorema, a operação de contração deve ser bem sucedida:

(**–sucesso**) $\nvdash a \;\Rightarrow\; a \notin K_a^-$.

O quinto critério é o que tenta aplicar o critério de mudança mínima mais fortemente. Segundo ele, em uma contração de K por a, é retirado de K o estritamente necessário, de forma que, ao expandir K_a^- por a, temos o conjunto K novamente. Esse critério foi amplamente criticado na literatura. Adiante (seção 2.7) será exposta a discussão em torno dele.

(**–recuperação**) $a \in K \;\Rightarrow\; K \subseteq (K_a^-)_a^+$.

O sexto critério assegura que o resultado de contrair sentenças logicamente equivalentes de um mesmo conjunto resulte em conjuntos iguais:

(**–extensionalidade**) $\vdash a \leftrightarrow b \;\Rightarrow\; K_a^- = K_b^-$.

Os postulados seguintes foram denominados "postulados suplementares", propostos com a intenção de regular o comportamento da contração por conjunções. De acordo com **–suplementar**$_1$, o que há de comum em K_a^- e K_b^- deve ser um subconjunto da remoção de $a \wedge b$ de K:

(**–suplementar**$_1$) $K_a^- \cap K_b^- \;\subseteq\; K_{a\wedge b}^-$.

De acordo com **–suplementar**$_2$, quando é necessário contrair $a \wedge b$ de K, a operação pode ser levada a cabo retirando ou a, ou b, ou ambos de K. Se for possível, o critério de mudança mínima requer que a e b não sejam ambas removidas nessa operação; que seja removida, por exemplo, a que tem menor importância epistêmica:

(**–suplementar**$_2$) $a \notin K_{a\wedge b}^- \;\Rightarrow\; K_{a\wedge b}^- \subseteq K_a^-$.

De agora em diante, utilizaremos **–fecho/–suplementar**$_2$ e ***-fecho/*-suplementar**$_2$ para nos referirmos à **coleção completa** dos postulados de contração e revisão AGM, respectivamente.

1.5.1 As Identidades de Levi e Harper

A operação de revisão de um conjunto K por a pode ser entendida intuitivamente como composta de duas fases: a primeira consiste na retirada em K de todas as sentenças incompatíveis com a, isso redunda em contrair de K a sentença $\neg a$. A segunda fase, assegurada a inexistência de sentenças incompatíveis com a, resume-se à expansão do conjunto $K_{\neg a}^-$ por a. Essa equivalência da operação de revisão com as operações de contração e expansão ficou conhecida como **identidade de Levi**, em termos formais:

(**Identidade de Levi**) $K_a^* = (K_{\neg a}^-)_a^+$.

Embora a operação de contração possa parecer mais simples que a de revisão, a contração pode ser definida por meio desta, pois a contração de K por a contém tudo aquilo que for comum a ambos K e $K_{\neg a}^*$. Essa equivalência é denominada **identidade de Harper** e é expressada do seguinte modo:

(**Identidade de Harper**) $K_a^- = K \cap K_{\neg a}^*$.

1.6 A Construção de Funções de Contração e Revisão

É razoável, em uma primeira análise do que constituiria uma função de contração para um conjunto K e sentença a, sugerir uma função de seleção γ cujo domínio seja $K \perp a$, isto é, $\{X \mid X \subseteq K, X \nvdash a \;\&\; X \subset Y \subseteq K \;\Rightarrow\; Y \vdash a\}$, e que retorne um único X tal que $X \in K \perp a$. Um exemplo: seja $K = \{p \to q, p, q\}$ e $K \perp q = \{\{p \to q\}, \{p\}\}$. Assim, γ efetuaria uma seleção sobre os subconjuntos maximais de K em questão: $\{p \to q\}$ e $\{p\}$, escolhendo o "melhor" elemento de acordo com alguma ordem de preferência e retornando esse subconjunto de K como a contração de K por q.

Segundo Makinson [Mak85, pág. 349], essa escolha deveria ser feita com base em um ordenamento dos elementos de $K \perp a$ a partir de noções pragmáticas e epistêmicas, como simplicidade, conveniência ou valor informativo. O resultado de $\gamma(K \perp a)$ seria então o "melhor" elemento, de acordo com alguma medida, de $K \perp a$. Essa função de contração foi denominada *maxichoice* e será designada neste texto por $(^{-mc})$ e a revisão, definida por meio de $(^{-mc})$ será denotada por $(^{*mc})$.

Contudo, Alchourrón e Makinson mostraram [AM82] que $(^{-mc})$ tem consequências problemáticas quando a operação de revisão é definida por meio de $(^{-mc})$ pela identidade de Levi $(K_a^{*mc} = (K_{\neg a}^{-mc})_a^+)$. Se a contração é *maxichoice*, então para todo conjunto de crenças K e uma sentença a qualquer, K_a^{*mc} é completo, ou seja, para toda fórmula p da linguagem ou p ou $\neg p$ $\in K_a^{*mc}$.

Isso conduz à investigação acerca do resultado de uma operação de contração cuja função de seleção γ selecionasse não um, mas todos os elementos de $K \perp a$, ou seja, se $\gamma(K \perp a)$ fosse igual à $K \perp a$. A função de contração resultante seria então definida por $\cap(K \perp a)$, isto é, aquele conjunto de crenças que seria comum aos conjuntos maximais de K que não implicam a. Essa operação foi denominada **contração de intersecção total** e será designada neste texto por $(^{-fm})$ ("fm" para *full meet*) e a revisão correspondente será denominada $(^{*fm})$.

Alchourrón e Makinson mostraram (*ibid.*) que, se uma operação de contração é de intersecção total, então $K_{\neg a}^{-fm} = K \cap Cn(a)$, ou seja, só restará em $K_{\neg a}^{-fm}$ as sentenças de K que são consequências de a. Portanto, a revisão $K_a^{*fm} = (K_{\neg a}^{-fm})_a^+ = Cn(a)$.

Por um lado, a operação de revisão (contração) *maxichoice*, formada pela escolha de apenas um conjunto resto[v] resulta em conjuntos demasiado "grandes". Por outro, a operação de revisão por intersecção total, formada pela intersecção de todos conjuntos resto, resulta em conjuntos demasiado "pequenos". Isso levou os teóricos do modelo AGM [Gär88, pág. 80] a considerarem que elas configurariam os limites superior e inferior de todas as operações de contração e revisão possíveis (o que se mostrará não ser o caso adiante na seção 3). Assim, a sequência de investigação natural passou a ser a de estudar as funções de contração e revisão obtidas através de seleções sobre **alguns** elementos resto.

1.6.1 A Operação de Contração de Intersecção Parcial

Alchourrón, Gärdenfors e Makinson desenvolveram em [AGM85] uma operação geral de contração, chamada "contração de intersecção parcial" (**partial meet contraction**). Para um conjunto K e uma sentença a, essa contração é definida pela seguinte fórmula: $K_a^- = \cap\gamma(K \perp a)$.

[v] *i.e.* um X tal que $X \in K \perp a$ para algum K e a.

As operações de intersecção total e *maxichoice* são dois sub-casos da operação de interseção parcial, no primeiro, $\gamma(K\perp a) = (K\perp a)$ e no segundo, $\gamma(K\perp a)$ é igual um único X tal que $X \in K\perp a$.

A função de seleção γ A função de seleção γ para um conjunto K é dita **relacional** se existe alguma relação $\sqsubseteq$ sobre $\mathcal{P}(K)$, tal que para todo $a \notin Cn(\varnothing)$, $\sqsubseteq$ define $\gamma(K\perp a)$ da seguinte forma[vi] :

(Def γ) $\quad \gamma(K\perp a) = \{K' \in K\perp a \mid K'' \sqsubseteq K' \text{ para todos } K'' \in K\perp a\}$.

Se o ordenamento $\sqsubseteq$ é transitivo, então a função γ é chamada de função "transitivamente relacional" e a contração definida pela função γ é chamada "função de contração de intersecção parcial transitivamente relacional".

A função γ pode selecionar, por exemplo, os elementos de $K\perp a$ mais arraigados epistemicamente. Para que seja possível uma seleção dos elementos de $K\perp a$ é necessário que haja um ordenamento entre eles. Gärdenfors diz [Gär88, pág. 81]: "Esse ordenamento dos subconjuntos maximais deveria ser independente de qual sentença nós estamos retirando [...]", mas não esclarece a necessidade de tal requisito. Aparentemente, a razão da imposição desse requerimento reside na necessidade de ter uma noção prévia à qualquer mudança de quais são os conjuntos de sentenças que são mais úteis na deliberação ou investigação, para que essa escolha dos conjuntos mais úteis seja consistente. Vejamos um exemplo (baseado em [Han99, pág. 81]). Seja o conjunto:

$$K = \{p \vee q, p \rightarrow q, q \rightarrow p\},$$

em que p e q são logicamente independentes. Os subconjuntos de K são:

$$B_1 = \{p \vee q, p \rightarrow q\}$$
$$B_2 = \{p \vee q, q \rightarrow p\}$$
$$B_3 = \{p \rightarrow q, q \rightarrow p\}.$$

Agora, temos os seguintes conjuntos-resto:

$$K\perp(p \leftrightarrow q) = \{B_1, B_2\}$$
$$K\perp(p \wedge q) = \{B_1, B_2, B_3\}$$

Suponha agora que a função de seleção γ faça as seguintes seleções, com base em algum ordenamento $\sqsubseteq$ de importância epistêmica:

(1) $\gamma(K\perp(p \leftrightarrow q)) = \{B_1\}$
(2) $\gamma(K\perp(p \wedge q)) = \{B_2\}$

[vi] A parte estrita de $\sqsubseteq$ é definida da maneira usual: $K \sqsubset J =_{def} K \sqsubseteq J$ e $J \not\sqsubseteq K$.

Para que seja notada a necessidade de que o ordenamento dos conjuntos de K seja feito de maneira independente das mudanças incidentes sobre K, perceba que no caso (1) acima, o ordenamento $\sqsubseteq$ em $\mathcal{P}(K)$ colocou $B_2 \sqsubset B_1$, isto é, B_1 é epistemicamente mais útil em K que B_2. Mas em (2), o ordenamento $\sqsubseteq$ colocou $B_1 \sqsubset B_2$, ou seja, B_2 é epistemicamente mais útil em K que B_1. Ou seja, a função γ é inconsistente! E isso se dá porque γ não faz sua seleção com base em um ordenamento $\sqsubseteq$ das sentenças de K que seja independente das possíveis mudanças que K possa sofrer.

Suponha que seja feito um ordenamento dessa maneira nos conjuntos B acima. Esse ordenamento tem que estabelecer de antemão que conjuntos devem ser retidos em detrimento de outros, por exemplo, se $p \to q$ e $q \to p$ são leis bem estabelecidas, sendo $q \to p$ a mais importante e $p \vee q$ apenas uma contingência, então, o ordenamento será $B_1 \sqsubset B_2 \sqsubset B_3$. A partir desse ordenamento, não mais é possível obter inconsistências nos resultados da função γ.

Será apresentado em seguida o teorema de representação do modelo AGM, em que uma operação de contração é obtida pela fórmula $K_a^- = \cap\gamma(K\perp a)$ se e somente se satisfaz os postulados AGM para contração.

O teorema de representação Um dos resultados centrais de [AGM-1985] é o teorema que diz que uma função é uma operação de contração se e somente se satisfazem os postulados para contração (**–fecho, –inclusão, –vacuidade, –sucesso, –recuperação, –extensionalidade, –suplementar$_1$, –suplementar$_2$**). Pela Identidade de Levi, o mesmo vale para as funções de revisão) (*ibid.*, pág. 514):

> **OBSERVAÇÃO 2.5** Seja $(-)$ uma função definida para conjuntos A de proposições e proposições x. Para toda teoria A, $(-)$ é uma contração de intersecção parcial sobre A sse $(-)$ satisfaz os postulados de Gärdenfors [**–fecho, –inclusão, –vacuidade, –sucesso, –recuperação e –extensionalidade**] para contração sobre A.[vii] [5]

e (*ibid.*, pág. 520):

> **COROLÁRIO 4.5** Seja A qualquer teoria, e $(-)$ uma função de contração parcial sobre A determinada por uma função de seleção γ. Então $(-)$ é transitivamente relacional sse satisfaz ambos **–suplementar$_1$** e **–suplementar$_2$**.[viii] [6]

Estes resultados fazem com que os esforços dos teóricos do modelo AGM sejam direcionados à construção explícita de funções de contração e revisão por intersecção parcial. Dizem eles (*ibid.*, pág. 520):

> Visto que esta coleção de postulados pode ser motivada de forma independente (cf. [Gär78]), há uma forte razão para focar nas funções de intersecção parcial transitivamente relacionais como uma representação ideal do processo intuitivo de contração.[7]

[vii] A notação foi adaptada ao nosso texto.
[viii] A notação foi adaptada ao nosso texto.

1.6.2 Contração por Arraigamento Epistêmico

Mesmo considerando que todas as crenças que constituem um estado de crenças tenham o mesmo grau de credibilidade, há certos tipos de crenças que, quando derrogadas, causam mais mudanças ou danos ao estado de crenças que outras. Por exemplo, retirar por um lado de um conjunto de crenças a sentença "uma maçã foi solta no ar e caiu" e retirar por outro "todos os corpos se atraem mutuamente", resulta em conjuntos de sentenças muito diferentes de um ponto de vista epistemológico. Se consideramos um estado de crenças como uma representação da realidade, então, a retirada de uma crença acerca da queda de uma maçã altera muito pouco o estado de crenças; enquanto que a retirada da lei da gravitação universal altera enormemente a organização estrutural do estado de crenças. Isso se dá de tal forma que agora faz sentido adicionar a este último caso, sem que haja inconsistência, a crença de que a maçã foi solta no ar e subiu.

Quanto mais útil for uma determinada informação na organização de nosso conhecimento, no planejamento de ações futuras e nas investigações científicas, mais resistência haverá em derrogá-la, e mais dano na estrutura de nosso conhecimento a derrogação dela acarretará. Dizemos, como no exemplo acima, que a lei da gravitação universal está **epistemicamente mais arraigada** do que a crença de que alguma maçã foi solta e caiu.

Analisemos o seguinte caso. Suponha que entre as crenças de João estão as seguintes:

(a) este pincel é de plástico,
(b) objetos de plástico não conduzem eletricidade,
(c) este pincel não conduz eletricidade,
(d) objetos de metal conduzem eletricidade.

Suponha que João descubra que (e) o pincel é de metal. Frente a essa nova informação, a certamente deve ser removida das crenças de João, mas além disso, a adição de e faz com que João deva se decidir quanto à remoção de c ou d. Isso significa que João deve escolher entre remover uma informação geral (d) e uma informação particular (c). Considerações pragmáticas o levarão a remover c em detrimento de d, entre elas, que d é uma informação mais útil para os afazeres do dia-a-dia do que c e que a remoção de c é uma forma mais econômica de reestruturação por provocar menos distúrbios no seu conjunto de crenças. É importante observar que a escolha de João em remover c e não d não é feita sob a alegação que d é mais correta ou tem mais evidência que c, pois geralmente é exatamente o oposto que ocorre. Em geral as crenças mais arraigadas epistemicamente, como leis naturais, são menos fundamentadas do que observações empíricas. Assim, João remove c porque d é mais arraigada epistemologicamente.

Quando são observados conflitos entre as crenças, é razoável re-estabelecer a consistência de forma que dentre as crenças envolvidas no conflito sejam eliminadas aquelas que são epistemicamente menos arraigadas. Da mesma maneira, quando é necessário remover alguma crença de um conjunto, de forma a bloquear a derivação de outra crença, procuramos remover aquelas que são menos arraigadas.

Um fator importante a notar é que as relações de importância ou arraigamento epistêmico não são fixas, mas variam de acordo com o estado de crenças. Por exemplo, para alguém que saiba que objetos de plástico, quando submetidos a uma diferença de potencial elétrico adequada, conduzem eletricidade, a sentença b acima seria minimamente arraigada em seu estado de crenças.

Visto que essas considerações sobre arraigamento epistêmico tomam parte nas mudanças de crenças que fazemos intuitivamente, Gärdenfors se propôs a dar um tratamento formal do conceito de arraigamento epistêmico. Todavia, segundo Gärdenfors, esse tratamento formal não pretende representar aspectos **quantitativos** da relação de arraigamento epistêmico, fazendo alguma espécie de mensuração do valor epistêmico de uma sentença em relação à outra. O que é pretendido é um tratamento formal dos aspectos **qualitativos** da relação de arraigamento epistêmico, na tentativa de capturar formalmente algumas características desse conceito. Em um primeiro momento [Gär84], as únicas características que ele assume para esse ordenamento é que seja transitivo e conectado. Posteriormente Gärdenfors e Makinson [GM88] tentaram captar mais características do ordenamento epistêmico. Apresentaremos a primeira abordagem de Gärdenfors, na qual são detalhadas algumas propriedades do arraigamento epistêmico e os requerimentos que o conjunto de crenças sobre o qual o ordenamento se dá deve cumprir. Em seguida examinaremos os postulados para o arraigamento epistêmico elaborados posteriormente.

A primeira abordagem

A operação de contração por intersecção parcial (seção 1.6.1) depende de um ordenamento dos conjuntos maximais de $K \perp a$ para todos $a \in K$. Gärdenfors propõe uma abordagem diferente no artigo em que apresenta a operação de contração por arraigamento epistêmico [Gär84]. Ele propõe um ordenamento de arraigamento epistêmico não dos conjuntos maximais em um $K \perp a$, mas sim das sentenças em K, que "abrangem" os elementos de um $X \in K \perp a$, *e.g*, a sentença $(a \wedge b) \wedge c$ abrange o conjunto $\{a, b, c\}$. A diferença entre as duas abordagens não é clara, vejamos os detalhes.

Para que seja possível obter para cada elemento de $K \perp a$ uma sentença que o abranja, é preciso que K seja um conjunto **fechado por consequência lógica finito**; em outras palavras, é necessário que a linguagem contenha somente um número finito de sentenças atômicas. Isto permite que os elementos de um conjunto de crenças K possam ser organizados em um número finito de classes de equivalência, no caso, classes cujos elementos são logicamente equivalentes As classes de equivalência têm infinitos elementos, visto que o conjunto é fechado sob consequência lógica, não obstante, havendo apenas um número finito de sentenças atômicas, há um número finito de classes de equivalência. Cada classe é representada por um elemento dela. Seja $Cn(K) = K$ e X um elemento do conjunto $K \perp a$, para alguma sentença $a \in K$. Visto que $Cn(K) = K$, segue-se que $Cn(X) = X$. Não obstante, X pode ser denotado pelos elementos representativos $x_1, \ldots, x_n$ de suas classes de equivalência, que são finitas. Assim, é possível fazer referência a X através da **sentença-abrangente** S_X:

> $S_X : x_1 \wedge, \ldots, \wedge x_n$, tal que $x_1, \ldots, x_n$ são elementos representativos das classes de equivalência do conjunto X.

A relação de arraigamento epistêmico $a \sqsubseteq b$ significa que b **é pelo menos tão epistemicamente arraigado quanto** a, em outros termos, **a disposição em reter** b **é pelo menos igual à disposição de reter** a, ainda outra leitura: **a disposição em remover** a **é maior ou igual a disposição em remover** b. A relação $a \sqsubset b$ é definida da maneira usual, *i.e.* $a \sqsubseteq b$ e $b \not\sqsubseteq a$, e quando $a \sqsubseteq b$ e $b \sqsubseteq a$, se diz $a \equiv b$.

Nesta abordagem, a relação $\sqsubseteq$ é definida sobre cada sentença-abrangente S_X, para todos os $X \in K \bot a$. Seja M_a^K:

$$M_a^K = \{X \mid S_{X'} \sqsubseteq S_X \text{ para todos } X' \text{ em } K \bot a\}$$

Em outros termos, M_a^K é o conjunto dos elementos topo sob o ordenamento de arraigamento epistêmico $\sqsubseteq$ das sentenças-abrangentes dos conjuntos em $K \bot a$.

A contração de K por a é definida então como:

$$K_a^- : \{p \mid p \in \cap M_a^K\}$$

K_a^- é formada pela intersecção dos elementos de M_a^K, isto é, os melhores elementos X de $K \bot a$ de acordo com o ordenamento $\sqsubseteq$ das sentenças-abrangentes S_X.

A operação de contração definida dessa maneira **não vale para conjuntos infinitos** (*i.e.* conjuntos que não podem ser organizados em um número finito de classes de equivalência), visto que para tais conjuntos não é possível obter uma sentença-abrangente que os identifique. A operação de contração definida por Alchourrón e Makinson [1982], e a operação de intersecção parcial (seção 1.6 acima), por sua vez, são baseadas em um ordenamento dos **subconjuntos** de um conjunto K fechado sob consequência lógica, ao invés de suas sentenças abrangentes. Gärdenfors enumera dois problemas de tais abordagens [Gär84, pág. 147]:

> 1. Primeiramente, ordenar todos os subconjuntos maximais de K de acordo com sua importância epistêmica requer habilidades cognitivas muito maiores que meramente ordenar as sentenças de K de acordo com sua importância epistêmica.
> 2. Em segundo lugar, Alchourrón e Makinson assumem que K_a^- seja **o** elemento maximal sobre um ordenamento dos conjuntos $K \bot a$ e, como eles mostraram, esse conjunto conterá muitas sentenças que intuitivamente não deveriam estar inclusas na contração de K por a.[8]

Quanto à primeira dificuldade, note-se que, se um conjunto K é logicamente fechado, ele tem uma quantidade infinita enumerável de elementos. Assim, a cardinalidade de seu conjunto potência $\mathcal{P}(K)$ é $2^{\aleph_0}$. Assumir a existência de um ordenamento de importância epistêmica de um conjunto de tal cardinalidade é um tanto despropositado.

Com relação à segunda dificuldade, visto que um $X \in K \perp a$ "tem sentenças demais"[ix], é assumido tacitamente na proposta da função de contração por intersecção parcial que a intersecção dos "melhores" elementos de $K \perp a$ eliminaria as sentenças que "super-populam" os elementos de $K \perp a$. Não obstante, tal assunção é digna de questionamento e tomada por Tennant como uma das causas do seu teorema de degeneração do modelo AGM, a ser apresentado adiante (seção 3).

A segunda abordagem

A abordagem de contração por arraigamento epistêmico proposta posteriormente por Gärdenfors e Makinson [GM88] consiste na definição, através de postulados, de um ordenamento de arraigamento epistêmico **entre as sentenças** de um conjunto de crenças K, ao invés de ordenar os subconjuntos de K. Os postulados propostos são de natureza **qualitativa**, isto é, não são utilizadas noções de natureza **quantitativa** para medir a relação de importância epistêmica. Os postulados seguem abaixo.

O critério fundamental para uma relação $\sqsubseteq$ de arraigamento epistêmico é que ela seja transitiva:

(**EE1**) Se $a \sqsubseteq b$ e $b \sqsubseteq c$, então $a \sqsubseteq c$.

O segundo postulado relaciona derivabilidade e arraigamento epistêmico. Segundo ele, sentenças logicamente mais fortes são mais fracas epistemicamente (trataremos deste postulado pormenorizadamente abaixo):

(**EE2**) Se $a \vdash b$, então $a \sqsubseteq b$.

O terceiro postulado regula a relação de arraigamento epistêmico no caso das conjunções. Se o propósito é retirar $a \wedge b$ de um conjunto K, isso pode ser feito retirando apenas a ou apenas b. Isso significa que a perda informativa de remover $a \wedge b$ é no mínimo a perda de remover a ou de remover b.

(**EE3**) Para todos a e b em K, ou $a \sqsubseteq a \wedge b$ ou $b \sqsubseteq a \wedge b$.

Os dois postulados restantes são para os casos-limite da relação de arraigamento epistêmico. **EE4** é para o caso em que uma sentença a não está no conjunto K em consideração, nesse caso, a é um elemento mínimo sob $\sqsubseteq$:

(**EE4**) Se $K \neq K^{\perp}$, então $a \notin K$ sse $a \sqsubseteq b$ para todo b.

Se a é mais ou tão arraigada epistemicamente do que qualquer outra sentença, então a é logicamente válida:

(**EE5**) Se $b \sqsubseteq a$ para todo b, então $\vdash a$.

[ix] Para todo $b \in K$, $a \to b \in X \mid X \in K \perp a$

Para que **EE5** seja válido em relação a um K, K deve ser logicamente fechado. De outro modo, se uma sentença a logicamente válida não pertencer a K, para qualquer $b \in K$, por **EE4**, $a \sqsubseteq b$. Assim, os postulados para arraigamento epistêmico acima são relativos apenas a conjuntos de crenças.

De agora em diante, nos referiremos à lista de postulados **EE1, EE2, EE3, EE4, EE5** abreviadamente por **EE1/EE5**.

Os postulados que Gärdenfors e Makinson propõem para a relação de arraigamento epistêmico têm como propósito captar formalmente a noção intuitiva de valor informativo. Dizem eles (*ibid.*, pág. 87):

> O critério fundamental para determinar o arraigamento epistêmico de uma sentença é quão útil ela é na investigação e deliberação.[9]

Se os postulados para arraigamento epistêmico pretendem capturar formalmente alguns aspectos da noção de importância epistêmica, como parece ser o caso pela citação acima, então **EE2** é contra-intuitivo. Intuitivamente, se uma sentença implica outra, então o valor informativo dela é maior ou igual do que o da outra. De acordo com Gärdenfors [Gär88, pág. 89–90], esse postulado "[..] parece ser tudo o que pode ser postulado para um tratamento **qualitativo** do arraigamento epistêmico."[10] Gärdenfors justifica **EE2** dizendo que ele foi adotado em função do critério de mudança mínima. Se uma sentença b é deduzível a partir de a e é necessário remover b, deve-se remover a também, porque de outro modo b permaneceria como consequência de a. Assim, a disposição em remover a é maior ou igual a disposição em remover b ou seja $a \sqsubseteq b$. Agora, se a deve ser removida, não é necessário que b também o seja, neste caso, não é necessário que $b \sqsubseteq a$. De acordo com o critério de mudança mínima, se $a \sqsubseteq b$, então, na medida do possível, b será preservada, quando a for removida.

Zenker aponta [Zen09] que a proposta de elucidação do conceito de arraigamento epistêmico feita pelos teóricos do modelo AGM acaba separando duas noções intimamente conectadas: **retratilidade comparativa** e **importância epistêmica**. Na citação acima, Gärdenfors claramente indica que pretende formalizar a noção de **importância epistêmica**. No entanto, o postulado **EE2** não pode ser usado para captar essa noção, visto que intuitivamente concebemos que determinada proposição é ou mais informativa ou igualmente informativa que suas consequências, como é o caso da lei da gravitação universal com respeito as suas instâncias. Esse postulado só pode ser considerado razoável se tomado como captando a noção de **retratilidade comparativa**: se há fortes motivos para remover a instância de uma lei, então haverá fortes motivos para remover a lei. Isto é, uma lei é mais ou igualmente retrátil que suas instâncias.

Nessa abordagem da noção de arraigamento epistêmico, é possível que uma lei seja mais retrátil ou menos imune que suas instâncias e que, não obstante, seja epistemicamente mais importante. Isso é contra-intuitivo na medida que em geral consideramos que o caráter da importância informativa de uma sentença em um estado de crenças K está associado à alguma medida de imunidade dessa sentença em K. **EE2** falha em obedecer um critério de adequação bem razoável para arraigamento epistêmico, como diz Zenker (*ibid.*, pág. 14):

> [...] qualquer ordenamento de arraigamento epistêmico deve espelhar nossa menor inclinação em duvidar da lei comparado à aceitar que alguma de suas instâncias seja uma exceção ou que seja não-cooperativa, em uma palavra: anômala. [11]

Apesar disso, Zenker mostra (*ibid.*, pág. 15) o seguinte. Seja um conjunto de crenças K formado pela representação formal de uma lei: $\forall x(Px \rightarrow Qx)$ e a sentença Pa[x] obtida como resultado de uma observação empírica. Assim, Qa será consequência de K. Se for descoberto que $\neg Qa$, isto é, que K deva ser revisado por $\neg Qa$, há que se fazer uma escolha entre remover Pa ou remover $Pa \rightarrow Qa$, e essa última implicaria remover $\forall x(Px \rightarrow Qx)$. Se o ordenamento de arraigamento epistêmico inicial colocar $Pa \sqsubset \forall x(Px \rightarrow Qx)$, então Pa será removida. Se o ordenamento inicial colocar $\forall x(Px \rightarrow Qx) \sqsubset Pa$, então $\forall x(Px \rightarrow Qx)$ será removida. E se $Pa \equiv \forall x(Px \rightarrow Qx)$, ambas serão removidas.

Segue-se que, na abordagem de arraigamento epistêmico delineada pelos postulados **EE1/EE5**, para que uma lei como $\forall x(Px \rightarrow Qx)$ seja mantida em detrimento de sentenças particulares como Pa, basta fazer com que $\forall x(Px \rightarrow Qx)$ seja mais arraigada que Pa; não é necessário que $\forall x(Px \rightarrow Qx)$ seja mais arraigada que sua instância $Pa \rightarrow Qa$. Isto significa que os postulados **EE1/EE5** não interferem nas relações de arraigamento epistêmico entre sentenças logicamente independentes, como Pa e $Pa \rightarrow Qa$.

Dos postulados **EE2** e **EE3** segue-se para quaisquer a e b, que ou $a \equiv a \wedge b$ ou $b \equiv a \wedge b$.[xi] Isso significa que uma conjunção é no máximo tão arraigada quanto cada um de seus componentes, ferindo a noção intuitiva de que há conjuntos de informações cujo valor informativo de sua conjunção é maior que o valor informativo de suas partes. Um exemplo simples é o seguinte: (a) o carro tem combustível e (b) o carro está em perfeitas condições mecânicas e elétricas. Evidentemente, o valor informativo de $a \wedge b$ é maior de que o valor informativo de a ou b tomados em isolado, no sentido em que $a \wedge b$ é mais útil para tomarmos decisões e fazer planos que qualquer dos a e b tomados isoladamente. Portanto, essa característica dos postulados os tornam bastante inadequados para modelar o que normalmente entendemos por arraigamento epistêmico e valor informativo. Adiante vamos mostrar outra característica implausível de **EE1/EE5**.

Seja **conectividade** a seguinte propriedade:

(**conectividade**) Para todos a e b em K, ou $a \sqsubseteq b$ ou $b \sqsubseteq a$.

Gardenfors [Gär88, pág. 90] obtém o seguinte resultado:

Teorema 1.6.1. *Os postulados* ***EE1/EE3*** *implicam* ***conectividade***.

[x]O modelo AGM está definido sobre uma lógica proposicional, não obstante, o fato de este exemplo utilizar lógica de predicados é irrelevante.

[xi]Devemos essa observação ao Prof. Frank Sautter.

Abaixo apresentamos a prova, tomando os postulados **EE1/EE3** como regras de inferência.

$$\frac{a \sqsubseteq b \quad b \sqsubseteq c}{a \sqsubseteq c}\,\scriptstyle{EE1}\textstyle\;,\quad \frac{\begin{matrix}[a]\\ b\end{matrix}}{a \sqsubseteq b}\,\scriptstyle{EE2}\textstyle\;,\quad \frac{}{(a \sqsubseteq a \wedge b) \vee (b \sqsubseteq a \wedge b)}\,\scriptstyle{EE3}\textstyle\;.$$

Demonstração. Por **EE2** temos que $a \wedge b \sqsubseteq a$ e $a \wedge b \sqsubseteq b$. Dado que (por **EE3**) $(a \sqsubseteq a \wedge b) \vee (b \sqsubseteq a \wedge b)$, suponha que $a \sqsubseteq a \wedge b$, visto que $a \wedge b \sqsubseteq b$, logo (por **EE1**), $a \sqsubseteq b$ e, consequentemente, $b \sqsubseteq a \vee a \sqsubseteq b$. Suponha que $b \sqsubseteq a \wedge b$, visto que $a \wedge b \sqsubseteq a$, portanto (por **EE1**), $b \sqsubseteq a$ e, consequentemente, $b \sqsubseteq a \vee a \sqsubseteq b$. Assim, segue-se de $(a \sqsubseteq a \wedge b) \vee (b \sqsubseteq a \wedge b)$ que $b \sqsubseteq a \vee a \sqsubseteq b$. □

O que resulta disso é que, dadas duas crenças a e b, deve estar determinado se o valor epistêmico de a é maior ou igual ao de b, ou vice-versa. Mas é razoável pensar em duas crenças cujo valor epistêmico não seja comparável. Por exemplo, até que se prove o contrário, não faz sentido comparar o valor epistêmico da lei da gravitação universal com a lei de Ohm[xii]. **EE1/EE3** conduzem a resultados contra-intuitivos. Logo abaixo trataremos mais pormenorizadamente da conectividade em um ordenamento de arraigamento epistêmico.

A operação de contração por arraigamento epistêmico

Gärdenfors aponta duas maneiras de relacionar os postulados **EE1/EE5** com uma operação de contração. Uma delas é assumir uma operação de contração como dada e, a partir dela, extrair o ordenamento $\sqsubseteq$. A outra é assumir um ordenamento de arraigamento epistêmico e a partir dele definir a operação de contração. Evidentemente, o caminho mais construtivo é o último, porque a relação de ordenamento epistêmico parece ter um caráter mais primitivo do que a operação de contração. Gärdenfors está de acordo com esse ponto [Gär88, pág. 88]: "Eu tomo a noção de arraigamento epistêmico como sendo mais fundamental que a noção de uma função de contração ou revisão".[12]

O ordenamento de arraigamento epistêmico pode ser definido a partir de uma operação de contração. Essa definição é obtida através da observação de que se uma sentença b não pertence a $K^-_{a\wedge b}$ é porque ela ou não pertence a K ou ela foi retirada em detrimento de a, em ambos os casos $b \sqsubseteq a$.

$$(C_{\sqsubseteq}) \qquad b \sqsubseteq a \Leftrightarrow b \notin K^-_{a\wedge b}$$

O caso que importa de um ponto de vista construtivo é o segundo, em que uma operação de contração é definida a partir de um ordenamento epistêmico. Gärdenfors e Makinson definem essa operação de contração da seguinte maneira:

$$(C_-) \qquad b \in K^-_a \Leftrightarrow b \in K \;\&\; (a \sqsubset a \vee b \text{ ou } \vdash a).$$

[xii] A corrente elétrica é igual à razão da voltagem pela resistividade.

Se $b \in K$, é claro que b está em K_a^- quando $\vdash a$, visto que $K_a^- = K$. O problema consiste em entender a relação da pertinência de b em K_a^- quando $a \sqsubset a \vee b$. Os autores não oferecem uma explicação intuitiva e independente de C_-.[xiii] A justificativa dada para C_- é que ele "simplesmente funciona" [GM88, pág. 89] e isso é corroborado pelo seguinte resultado:

> se um ordenamento $\sqsubseteq$ satisfaz **EE1/EE5**, então a função de contração que é determinada por C_- satisfaz os critérios **–fecho/–suplementar**$_2$ e também $C_\sqsubseteq$ (*ibid.*, pág. 90).

Possivelmente esse é o motivo para a asserção de Gärdenfors [Gär88, pág. 91] de que **EE1/EE5** podem ser considerados um conjunto "completo" de postulados para o ordenamento de arraigamento epistêmico.

Da não conectividade do ordenamento

Quando se trata de um ordenamento entre conjuntos de sentenças ao invés de um ordenamento de sentenças, é mostrado em [AGM85] que a imposição de conectividade em uma relação ordenadora adiciona muito pouco ou nada à relação, como dizem Alchourrón *et al.* (*ibid.*, pág. 523) "Talvez surpreendemente, acontece que no caso infinito ela [a conectividade] adiciona muito pouco e, no caso finito, absolutamente nada." [13]

Não obstante, que a relação de arraigamento epistêmico delineada pelos postulados **EE1/EE5** seja conectada faz com que ela seja bastante implausível como uma formalização da noção intuitiva de arraigamento epistêmico ou valor informativo. Sabemos intuitivamente que há sentenças cujos valores informativos respectivos não são comparáveis, por estarem em âmbitos do conhecimento completamente distintos, como é o caso, citado acima, da lei de Ohm e da lei da gravitação universal. Tennant argumenta em favor da implausibilidade do postulado **EE3** (para todos a e b em K, ou $a \sqsubseteq a \wedge b$ ou $b \sqsubseteq a \wedge b$), dizendo que não é razoável ter que decidir acerca do arraigamento epistêmico de duas crenças tomadas arbitrariamente [Ten97, pág. 580]:

> Hansson e Rott notam esse fato (na p. 363), mas falham em extrair a conclusão óbvia que a própria desconectividade da relação de arraigamento intuitiva, em geral, torna a condição escolhida por eles (**EE3**) extremamente implausível como parte de uma caracterização formal da noção intuitiva.[14]

A mesma crítica foi feita por Lindström e Rabinowicz em [LR91, pág. 8], em que é apresentada uma abordagem da relação de arraigamento epistêmico em que não vale conectividade. Essa abordagem toma os postulados de Gärdenfors-Makinson e apenas substitui o terceiro (**EE3**) pelo seguinte:

$(EE3')$ se $a \sqsubseteq b$ e $a \sqsubseteq c$, então $a \sqsubseteq b \wedge c$.

[xiii] A explicação oferecida envolve o operador $C_\sqsubseteq$.

Agora, de **EE1**, **EE2**, **EE3'**, **EE4** e **EE5** não se segue que $\sqsubseteq$ seja conectada. Na abordagem de Gärdenfors-Makinson $a \wedge b$ é tão arraigado quanto o elemento menos arraigado de $\{a,b\}$; enquanto que na abordagem de Lindström-Rabinowicz a e b podem ser incomparáveis com relação ao arraigamento epistêmico, caso em que $a \wedge b$ é estritamente menos arraigado que ambos a e b (*ibid.*).

Os postulados de Lindström e Rabinowicz já não definem apenas ordenamentos conectados. Isso porque os ordenamentos não-conectados satisfazendo os postulados podem ser representados por intersecções de ordenamentos conectados. Isto é, alguns ordenamentos conectados têm como intersecção o mesmo ordenamento não-conectado. Dizem eles (*ibid.*, pág. 13):

> [...] a representação do arraigamento epistêmico em termos de conjuntos de relações conectadas é mais **fina** ou **discriminadora** do que a representação em termos de relações não-conectadas. Mas talvez tal representação seja discriminadora **demais**. [15]

Uma consequência importante da possibilidade de não-conectividade da relação de arraigamento epistêmico de Lindström e Rabinowicz é que não é possível em todos os casos, dado um conjunto K e uma sentença a, definir uma **única** revisão de K por a. Se a relação $\sqsubseteq$ é conectada, a revisão de K por a nesse sistema é o conjunto de todos os b tal que $\neg a \sqsubset b$, mas, se a relação $\sqsubseteq$ não é conectada, essa definição será imprópria, pois haverá algum b que seja incomparável com $\neg a$. Isso significa que a operação de revisão nesse sistema **não é funcional** (*ibid.*, pág. 15).

1.7 O Critério de Recuperação e a Noção de Mudança Mínima

O critério de recuperação (ou simplesmente **–recuperação**) foi proposto na tentativa de captar a ideia de que a operação de contração de uma sentença a de um conjunto de crenças K deveria ser o menos incisiva possível, de forma a retirar o estritamente necessário para que $K \nvdash a$. A motivação para **–recuperação** na contração de K por a é que seja retirado o estritamente necessário, de modo que se a for adicionado novamente, tudo o que se seguia de K, continue seguindo-se de $(K_a^-)_a^+$.

Que **–recuperação** valha para a contração *maxichoice* (seção 1.6 acima), *i.e.* aquela que, para um $Cn(K) = K$ e uma sentença a, retorna um elemento de $K \perp a$, é mostrado facilmente. Para tanto, basta mostrar que, para cada $b \in K$, $a \to b$ pertence a qualquer $X \in K \perp a$, de modo que $b \in X_a^+$.

Teorema 1.7.1. *Se (-) é uma operação de contração* maxichoice*, então (-) satisfaz* ***–recuperação****.*

Abaixo são dadas em forma de regras de dedução natural as definições de fecho e conjunto resto. Abreviamos "$(x_1 \wedge \ldots \wedge x_n) \to y \mid x_i \in X$" por "$X \to y$".

$$\frac{b \in K}{(a \to b) \in K}{\scriptstyle fecho_1}\ ,\quad \frac{X \to a}{a \in X}{\scriptstyle fecho_2}\ ,\quad \frac{a \in X}{X \notin K \perp a}{\scriptstyle K\perp a_1}\ ,\quad \frac{X \in K \perp a}{a \notin X}{\scriptstyle K\perp a_2}\ ,$$

$$\frac{b \in K \quad X \in K\perp a \quad b \notin X}{(X \wedge b) \to a}{\scriptstyle K\perp a_3}\ ,\quad \frac{Cn(K) = K \quad X \in K\perp a}{Cn(X) = X}{\scriptstyle K\perp a_4}\ \text{e}\ \ \frac{(a \wedge b) \to c}{a \to (b \to c)}{\scriptstyle \to\wedge}\ .$$

Demonstração. Seja que $Cn(K) = K$, $b \in K$, e que $X \in K\perp a$. Logo, $a \to b \in K$. Suponha que $a \to b \notin X$, então (por $k\perp a_3$) segue-se que $X \wedge (a \to b) \to a$ e, (por $\to\wedge$) $X \to ((a \to b) \to a)$. Dado que $\vdash ((a \to b) \to a) \to a$, segue-se que $X \to (((a \to b) \to a) \to a)$. Portanto, $X \to a$ e, (por $K\perp a_4$) $a \in X$ e, (por $K\perp a_1$) $X \notin K\perp a$, o que contraria a hipótese. Portanto $a \to b \in X$. □

O resultado contra-intuitivo é que o critério de recuperação também vale para a contração de intersecção total $(-^{fm})$ (seção 1.6 acima), recordando que $K_b^{-fm} = K \cap Cn(\neg b)$, isto é, a contração $(-^{fm})$ de K por b em que só permanecem aqueles elementos de K que são consequências de $\neg b$; todos os eventuais elementos de K não relacionados com b são removidos conjuntamente com b, ferindo de modo direto o critério de mudança mínima. Para mostrar que **–recuperação** vale para $(-^{fm})$ há que se mostrar para um $a \in K$, que $a \in Cn((K \cap Cn(\neg b)) \cup \{b\})$.

Teorema 1.7.2. *Se (-) é uma operação de intersecção total, então (-) satisfaz* ***–recuperação****.*

Demonstração. Suponha que $a \in K$ e $Cn(K) = K$, então $b \to a \in K$. Dado que $\vdash \neg b \to (b \to a)$, então (por $fecho_2$) $b \to a \in Cn(\neg b)$. Portanto, $b \to a \in Cn(\neg b) \cap K$ e $(Cn(\neg b) \cap K) \cup \{b\} \to a$. Assim, $a \in Cn((Cn(\neg b) \cap K) \cup \{b\})$. □

O critério de recuperação foi adotado na tentativa de assegurar que tão pouco seja removido na contração de K por a que, se a for adicionado novamente, é obtido tudo o que havia antes. O problema é que, em casos como a contração de intersecção total $(-^{fm})$ acima, a "segurança" de que crenças não sejam removidas desnecessariamente redunda em que restem apenas as consequências de $\neg a$ em K_a^{-fm}. Isto significa que o critério de recuperação desempenha papel praticamente nulo para assegurar que a operação de contração realize mudanças mínimas.

1.7.1 Contra-exemplos ao Critério de Recuperação

O critério **–recuperação** foi bastante criticado na literatura[xiv], no sentido de que são raros os casos em que a adição de uma crença a um conjunto de crenças do qual ela tinha sido contraída faz com que o conjunto volte a ser como era antes. Seguem dois contra-exemplos típicos, um abstrato e um concreto.

1º contra-exemplo Seja K o conjunto de crenças obtido pelo fecho lógico de $\{a\}$, *i.e.* $K = Cn(a)$. Para remover $b \to a$ de K deve-se remover a **e também é razoável que se remova todas as dependências de** a, de forma que $K^-_{b\to a} = Cn()$. Agora a adição de $b \to a$ à $K^-_{b\to a}$ resulta em um conjunto que não mais implica a.

[xiv]Por exemplo, em [Ten94] e [Han06].

2º contra-exemplo Suponha que João acredite em (p) "Maria tem uma menina", e em (q) "Maria tem um menino". Representemos a crença de que Maria tenha filhos por "$p \vee q$". Deste modo, João acredita em $p \wedge q$ e, consequentemente, em $p \vee q$. Depois, por algum motivo, João abandona a crença de que Maria tenha filhos ($p \vee q$), para tanto é necessário abandonar ambas as crenças p e q e **tudo aquilo que dependa delas**, $p \wedge q$, por exemplo. Posteriormente, João volta a acreditar que Maria tenha filhos ($p \vee q$), mas agora ele não pode inferir daí nem p nem q. Ou seja, o critério da recuperação falharia neste caso.

A resposta de Makinson

Makinson argumenta [Mak97] que os contra-exemplos só são válidos se é suposto que os conjuntos de crenças não são fechados sob consequência lógica ou que são fechados por consequência lógica mas possuem algum tipo de estrutura de justificação entre as crenças.

1º caso De acordo com Makinson, no modelo AGM a ação em negrito do 1º contra-exemplo não pode ser levada a cabo por que nos conjuntos de crenças não são distinguidas relações de justificação e dependência. $K = Cn(a)$, logo, $\neg b \to a \in K$ e, dado que $\neg b \to a$ não implica $b \to a$, segue-se que $K^{-}_{b\to a}$ contém $\neg b \to a$. Agora, de $(K^{-}_{b\to a})^{+}_{b\to a}$, segue-se a, posto que $\neg b \to a, b \to a \vdash a$.

2º caso O contra-argumento para este caso segue as mesmas linhas do anterior, no sentido de que deve-se considerar que o conjunto K que representa o estado de crenças de João é logicamente fechado e que não sejam distinguidas relações de dependência em K. Visto que $p \wedge q \vdash p \leftrightarrow q$, então $p \leftrightarrow q \in K$. Dado que $p \leftrightarrow q \nvdash p \vee q$, não há porquê removê-lo para que K não implique $p \vee q$, mesmo que saibamos intuitivamente que **nenhuma das justificativas que $p \leftrightarrow q$ tinha em K, persiste em $K^{-}_{p\vee q}$**. De acordo com isso, $K^{-}_{p\vee q}$ será igual à $Cn(p \leftrightarrow q)$. Agora se $p \vee q$ for acrescentado a $K^{-}_{p\vee q}$, o conjunto resultante $(K^{-}_{p\vee q})^{+}_{p\vee q}$ conterá tanto p quanto q, tal como o conjunto original K.

Os dois contra-exemplos para o critério de recuperação são baseados no seguinte critério proposto por Fuhrmann [Fuh91, pág. 184] denominado "condição de filtragem":

(Condição de Filtragem)

> Quando uma crença y é retirada de modo a formar a contração das crenças de alguém por x, então a contração por x não deveria conter nenhum item que é acreditado "simplesmente porque" y era acreditado.[16]

A condição de filtragem é bastante plausível. Se ela é imposta, **–recuperação** não é um critério válido. Contudo, a relação de dependência inferencial referida pela expressão "simplesmente porque" não tem lugar no modelo AGM, visto que os estados de crenças nesse modelo são completamente desprovidos de quaisquer relações de dependência ou justificação. Makinson defende que **–recuperação** vale para teorias fechadas sob consequência lógica e desprovidas de relações de justificação. Ele adiciona [Mak97, pág. 479] : "A adequação [de **–recuperação**] simplesmente serve para enfatizar que conjuntos fechados nus estão distantes de ser uma representação adequada de conjuntos de crenças no mundo real. Estes deveriam ser bases ou conjuntos fechados equipados com estruturas justificativas ou alguma outra".[17]

1.7.2 Makinson e o Critério de Recuperação

Makinson aponta [Mak87, pág. 385] quatro peculiaridades do critério de recuperação: uma intuitiva e três formais. A intuitiva é que **–recuperação** é o único dos seis critérios para contração que é passível de questionamento **sob a interpretação pretendida** da operação de contração. Segundo ele, embora **–recuperação** seja razoável, não é evidentemente correto como os demais. As peculiaridades formais do critério são:

1. **–recuperação** é o único critério que faz com que uma determinada operação de contração não seja "contração de intersecção parcial"[xv] se o operador de consequência lógica não incluir a lógica proposicional clássica.

2. Somente os postulados **–fecho** e **–recuperação** requerem que os estados de crenças sejam representados por conjuntos de sentenças logicamente fechados.

3. A operação de contração pode ser definida no modelo AGM através da revisão pela seguinte identidade, chamada **identidade de Harper**: $K_a^- = K \cap K_{\neg a}^*$. O curioso é que os postulados para revisão não incluem o postulado de recuperação.

Por outro lado, são apresentadas por Makinson (*ibid.*, pág. 387) certas características "interessantes" do critério do ponto de vista formal:

1. **–recuperação** faz com que um conjunto de crenças K, mesmo sofrendo mudanças severas quando da remoção de uma sentença a, continue contendo informação o suficiente para recuperar as informações removidas, caso a seja adicionado.

2. **–recuperação** é **extremamente** útil, senão fundamental, para o modelo AGM na demonstração do teorema de representação que prova que toda operação de contração que obedece **–fecho/–suplementar$_2$** é uma contração de intersecção parcial.

3. Outro tipo de contração (que não será abordada neste texto), denominada "contração segura" e considerada bastante atrativa por Makinson, também satisfaz **–recuperação**.

Contudo, um resultado surpreendente (*ibid.*, pág. 389) é:

> dada uma operação de contração $(^-)$ que não satisfaz recuperação, há uma operação $(^{-'})$ que satisfaz recuperação, a qual retém o maior número de sentenças do

[xv] A operação de intersecção parcial é a seguinte: $K_a^- = \cap\gamma(K \perp a)$, em que γ é uma função de seleção definida sobre os elementos de $K \perp a$.

conjunto original e que é "revisão-equivalente" a aquela, isto é, ambas $(^-)$ e $(^{-'})$ definem o mesmo operador de revisão, por meio da identidade de Levi.

Desta maneira, Makinson conclui que quando se está interessado na operação de revisão, o critério de recuperação é, além de útil, inofensivo. Neste caso, recuperação não é só útil para estabelecer certos resultados para a operação de revisão e contração, simplificando as provas, mas também não tem como resultado nada além do que outro operador de contração teria.

1.7.3 Isaac Levi: A Condição de Filtragem e o Critério de Recuperação

Levi defende que a condição de filtragem[xvi] não é um bom critério para regular as mudanças incidentes sobre um estado de crenças. Ele argumenta que não cabe apelo às relações de justificação para a decisão de se determinada crença deve ser removida ou não. O que deveria ser considerado neste caso são critérios que estejam relacionados à noção de **valor informativo**, como capacidade explicativa, poder preditivo, etc. Diz Levi [Lev03, pág. 214]:

> Argumentos a favor e contra a contração envolvem a perda de coerência explicativa, capacidade preditiva, simplicidade e, mais amplamente, valor informativo da teoria provocados pela remoção de um item do conjunto de crenças. Estas considerações não envolvem uma "condição de filtragem".[18]

Recordaremos (de forma simplificada) o segundo contra-exemplo ao critério de recuperação acima, por conveniência. João acredita que Maria tenha dois filhos, um menino e uma menina ($p \wedge q$). Depois, João abandona a crença de que Maria tenha filhos ($p \vee q$). Posteriormente, João volta a acreditar que Maria tenha filhos ($p \vee q$), mas agora ele não pode inferir daí nem p nem q. O contra-argumento de Makinson vai no sentido de que no modelo AGM a sentença $p \leftrightarrow q$ seria mantida após a contração por $p \vee q$, e ela permitiria a recuperação tanto de p quanto de q, quando $p \vee q$ for re-adicionada.[xvii]

O argumento de Levi adaptado ao exemplo acima, assere que o fato de que $p \leftrightarrow q$ seja consequência de $p \wedge q$ é irrelevante para deliberar se ela deve ou não ser removida na contração de $p \vee q$. Segundo ele, se $p \leftrightarrow q$ deve legitimamente ser removida, então sua remoção não provoca mais perda de valor informativo que a remoção de $p \vee q$.

1.7.4 Contração em Teorias sem o Critério de Recuperação

Embora o critério de recuperação seja válido para teorias (conjuntos fechados sob consequência lógica) e seja ademais "inofensivo" quando o interesse está na operação de revisão, seu caráter contra-intuitivo levou à busca de substitutos para ele. Surpreendentemente, todos

[xvi] Relembrando: (**condição de filtragem**) Quando uma crença y é retirada de modo a formar a contração das crenças de alguém por x, então a contração por x não deveria conter nenhum item que é acreditado "simplesmente porque" y era acreditado

[xvii] Lembre-se que $p \vee q,\ p \leftrightarrow q \vdash p \wedge q$

os substitutos que pareciam razoáveis revelaram-se equivalentes à recuperação na presença dos outros postulados para contração.

Um critério de adequação bastante razoável que tenta assegurar que a contração efetue mudanças mínimas foi proposto em [Han91, pág. 253]:

> [...] crenças que de nenhuma forma contribuem para o fato de que K implica p não deveriam ser excluídas de K_p^-. [19]

Uma formalização bastante direta desse critério de adequação é a seguinte:

$$q \in K \;\&\; q \notin K_p^- \;\Rightarrow\; p \in Cn(K_p^- \cup \{q\})$$

O problema com o critério acima é que, se adotado juntamente com os demais postulados para contração, exceto **–recuperação**, segue-se que toda contração de K por p é um elemento de $K \perp p$. Pela própria definição do conjunto $K \perp p$,[xviii] é simples ver, para todo $q \in K$ e $X \in K \perp p$, que $p \to q \in X$. Disso se segue que **–recuperação** é satisfeito como consequência desse critério e dos demais postulados para contração. Portanto o critério acima não é um candidato para substituir **–recuperação**.

Outro critério, denominado **–relevância**, foi também analisado (*ibid.*, pág. 253):

(**–relevância**) $\quad q \in K \;\&\; q \notin K_p^- \;\Rightarrow\; \exists A(K_p^- \subseteq A \subset K \;\&\; (p \notin Cn(A)) \;\&\; (p \in Cn(A \cup \{q\})))$.

A ideia que o critério tenta formalizar é que, se algum elemento q pertencente a K não pertence a K_p^-, isso se deve ao fato de que q é membro de algum subconjunto de K que implica p e, ademais, q é relevante para que isso ocorra.

O critério **–relevância** pode ser ainda enfraquecido retirando a condição de que $K_p^- \subseteq A$, assim obtemos outro critério, denominado **–retenção-de-núcleo**:

(**–retenção-de-núcleo**) $\quad x \in K \;\&\; x \notin K_p^- \;\Rightarrow\; \exists A(A \subset K \;\&\; (p \notin Cn(A)) \;\&\; (p \in Cn(A \cup \{x\})))$.

Hansson sugere (*ibid.*, pág. 254) que **–retenção-de-núcleo** é o critério mais fraco que se pode utilizar para formalizar a ideia de que, se uma sentença q é excluída de um conjunto K de modo que p seja removida, então q toma alguma parte no fato de que K implique p.

Não obstante, é demonstrado (*ibid.*, pág. 258) que **–retenção-de-núcleo**, quando aplicado a teorias, implica **–recuperação**. Isto é feito mostrando que qualquer operação aplicada a uma teoria e que satisfaça os critérios para contração (menos **–recuperação**) e **–retenção-de-núcleo** é uma operação de intersecção parcial. Mas toda operação de intersecção parcial sobre teorias satisfaz **recuperação**.[xix] Assim, um critério plausível e fraco como **–retenção-de-núcleo**, combinado com os demais critérios aparentemente inofensivos, implica o critério de recuperação. Isso faz com que diversos autores, como Makinson e Hansson[20], considerem-no como

[xviii] Isto é, o conjunto dos $X \subseteq K$, tal que $X \nvdash p$, e para qualquer $Y \subseteq K$ tal que $X \subset Y$, então $Y \vdash p$

[xix] Observação 2.3 em [AGM85]

uma “propriedade emergente” das operações de contração sobre conjuntos logicamente fechados. Hansson diz [1991, p. 255] (note-se que ele fala de operadores no âmbito de conjuntos fechados sob consequência lógica):

> Visto que não parece sensato que uma operação de contração viole retenção de núcleo ou qualquer dos (**–fecho, –sucesso, –inclusão, –vacuidade** e **–extensionalidade**), um operador razoável de contração sem a propriedade de recuperação não parece ser possível.[xx] [21]

Hansson analisa o seguinte postulado (**EC**) em substituição ao postulado **–fecho** (K_a^- é um conjunto de crenças (fechado sob consequência lógica)) do modelo AGM:

(**EC**) $\quad K \cap Cn(K_a^-) \subseteq K_a^-$

Uma operação $(^-)$ sobre o conjunto de crenças K satisfaz **EC**, **–inclusão, –vacuidade, –sucesso, –extensionalidade** e **–relevância** se e somente se $(^-)$ é uma operação de intersecção parcial. Ademais, quando $(^-)$ é aplicada sobre conjuntos finitos ou bases de crenças, $(^-)$ não satisfaz **–recuperação** (*ibid.*, pág. 256).

Por outro lado, se **–relevância** for substituído por **–retenção-de-núcleo**, outra operação de contração é obtida: a **retirada conservativa**. Um operador $(^-)$ é uma **retirada conservativa** se e somente se $(^-)$ satisfaz **–inclusão, –vacuidade, –sucesso, –extensionalidade e –retenção-de-núcleo**. Hansson mostra que neste caso toda operação de intersecção parcial é uma retirada conservativa, mas que o inverso nem sempre ocorre. Como no caso seguinte: $K = \{x,y,z\}$ e $K^-_{x\wedge(y\vee z)} = \{y\}$. Essa contração é uma retirada conservativa, mas não é uma operação de intersecção parcial, porque $K\perp(x\wedge(y\vee z)) = \{\{x\},\{y,z\}\}$, e {y} não é a intersecção de nenhum subconjunto desse conjunto. Esse tipo de contração, portanto, não retém tudo o que poderia reter do conjunto inicial sem incompatibilidade lógica e, portanto, claramente viola o postulado de recuperação. Adiante (seção 3) outras considerações sobre os postulados AGM serão feitas.

No próximo capítulo será apresentada a abordagem dos condicionais contrafactuais pelo modelo AGM de revisão de crenças.

Notas

[3][...] it seems to the author that loss of content, measured by inclusion, is a particularly interesting kind of loss to consider in the context of theory contraction.

[4] [...] Real-life contraction, we suggest, is applied to such bases, and not to sets already closed under consequence.

[5] OBSERVATION 2.5. Let $-$ be a function defined for sets A of propositions and propositions x. For every theory A, $-$ is a partial meet contraction operation over A iff $-$ satisfies the Gärdenfors postulates $(-1)-(-6)$ for contraction over A.

[6]COROLLARY 4.5. Let A be any theory, and $-$ a partial meet contraction function over A determined by a selection function γ. Then $-$ is transitively relational over A iff satisfies both (-7) and (-8).

[xx]A notação foi adaptada ao nosso trabalho.

[7]Since this collection of postulates can be independently motivated (cf. [Gär78]), there is strong reason to focus on transitively relational partial meet contraction functions as an ideal representation of the intuitive process of contraction.

[8] Firstly, ordering all the maximal subsets of K according to their epistemic importance requires much stronger cognitive abilities than merely ordering the sentences of K according to their epistemic importance. Secondly, Alchourrón and Makinson assume that K_a^- be an element of $K\perp a$; and as they have themselves shown (the first result mentioned earlier), if K is a belief set, K_a^- will then contain many sentences that, intuitively, should not be included in the contraction of K with respect to A.

[9]The fundamental criterion for determining the epistemic entrenchment of a sentence is how useful it is in inquiry and deliberation.

[10][...](**EE2**) seems to be all that can be postulated for a **qualitative** account of epistemic entrenchment."

[11][...] any ordering of epistemic entrenchment must mirror our lower inclination to doubt the law compared to accepting that one of its instances may be exception-ridden or otherwise uncooperative, for short: anomalous. (*ibid.*, pág. 13)

[12]"I take the notion of epistemic entrenchment to be more fundamental than the notion of a contraction or a revision function.

[13]"Perhaps surprisingly, it turns out that in the infinite case it adds very little, and in the finite case nothing at all."

[14]Hansson and Rott note this fact (at p. 363), but fail to draw the obvious conclusion that the very disconnectedness of the intuitive entrenchment relation, in general, makes their chosen condition (**EE3**) extremely implausible as part of a formal characterization of the intuitive notion.

[15][...] the representation of epistemic entrenchment in terms of sets of connected relations is more **finegrained** or **discriminating** than the representation in terms of non-connected relations. But perhaps such a representation would be too discriminating.

[16]When a belief y is withdrawn in order to form the contraction of one's beliefs by x, then the contraction by x should not contain any items that are believed "just because" y was believed.

[17]"The apropriateness merely serves to emphasize that naked closed sets are far from being an adequate representation of belief sets in the real world. These should be bases or closed sets equipped with justificational and perhaps other structure."

[18]Arguments for and against contraction concern the loss in explanatory coherence, predictive capacity, simplicity and, more generally, informational value of the theory incurred by giving up an item from the belief set. These considerations do not invoke a "filtering condition".

[19][...] beliefs that do not in any way contribute to the fact that K implies p should not be excluded from K_p^-.

[20][Mak97], [Han91]

[21]Since it does not seem sensible for a contraction operation to violate core-retainment or any of $(^-1)$–$(^-)5$, a reasonable contraction operator without the recovery property does not seem to be possible. (A ordem dos postulados foi mudada de modo a se adequar a numeração do presente texto)

Capítulo 2

O Modelo AGM e as Sentenças Condicionais

*Apresentaremos a lógica para contrafactuais de Gärdenfors e a correlação de cada axioma com os postulados para revisão do modelo AGM. Posteriormente apresentaremos a prova de Gärdenfors (teorema de impossibilidade) de que o conhecido critério de adequação para contrafactuais –o teste de Ramsey **RT**– não é compatível com o postulado ***-vacuidade** ($\neg a \notin K \Rightarrow K_a^+ \subseteq K_a^*$) para revisão. Por fim analisaremos as tentativas de barrar o teorema de impossibilidade e suas supostas "causas".*

2.1 Introdução

Gärdenfors começou suas investigações sobre revisão de crenças com o propósito de analisar a inferência de condicionais e ser capaz de oferecer critérios de aceitabilidade razoáveis para os mesmos. No capítulo sobre condicionais de seu livro "Conhecimento em Fluxo" ele é enfático [Gär88, pág. 146]:

> A hipótese inicial deste capítulo é que sentenças condicionais em várias formas são sobre mudanças de estados de crenças.[22]

Na abordagem em questão é utilizado um critério de aceitabilidade para condicionais contrafactuais bastante conhecido na literatura, o "teste de Ramsey" (**RT**). Segundo **RT**, um condicional é aceito em um estado de crenças quando o seu consequente pertence ao estado de crenças revisado (minimamente) pelo antecedente.

Gärdenfors [Gär86] mostrou que **RT** é incompatível com o critério ***-vacuidade** ($\neg a \notin K \Rightarrow K_a^+ \subseteq K_a^*$). ***-vacuidade** é um dos critérios propostos no sentido de assegurar que as operações de revisão incidentes nos estados de crenças efetuassem mudanças mínimas. Segundo ***-vacuidade** se a é consistente com K, a mudança mínima para adicionar a a K é simplesmente uma expansão.

Portanto, para fazer uma abordagem dos condicionais contrafactuais em termos de revisão de crenças, um dos critérios, **RT** ou ***-vacuidade**, deve ser abandonado. Gärdenfors desenvolveu [Gär78] uma lógica para contrafactuais que prescinde de ***-vacuidade**, e mostrou que essa lógica (que tem como substrato estados de crenças) é equivalente à lógica de contrafactuais "oficial" de Lewis [Lew73] (que tem como substrato mundos possíveis). O teste de Ramsey, sustenta Gärdenfors, parece uma forma razoável de análise lógica dos condicionais. Contudo, [Gär86, pág. 88]:

> [...] é questionável se as revisões de crenças usadas nesta análise realmente são **mínimas**, visto que elas não satisfazem [***-vacuidade**].[23]

Em seguida falaremos sobre o teste de Ramsey, apresentaremos a lógica para contrafactuais de Gärdenfors, o resultado de incompatibilidade de **RT** com ***-vacuidade** e, por fim, discutiremos alguns conceitos envolvidos no teorema de incompatibilidade.

2.1.1 O Teste de Ramsey

O termo "teste de Ramsey"[i] aparece na literatura tendo pelo menos três significados distintos, dois derivados e um original. O original vem de uma nota de rodapé em um trecho em que Ramsey faz considerações acerca da implicação material; na nota Ramsey diz [Ram50, pág. 247, nota 1]:

> (1) [...] se duas pessoas estão argumentando sobre "se p, então q" e estão ambas em dúvida quanto a p, elas estão adicionando p hipoteticamente a seus estoques de conhecimentos e argumentando nessa base acerca de q.

Chisholm é o primeiro a mencionar (1) em seu artigo pioneiro sobre os condicionais contrafactuais [Chi46, pág. 298], embora ele não o adote como critério em sua análise. A proposta de (1) só passa a ser amplamente divulgada a partir de sua reformulação por Stalnaker [Sta68]. Com o propósito de fazer (1) valer também para condicionais contrafactuais, Stalnaker propõe duas reformulações, uma epistêmica (2) e uma ontológica (3). Somente essa última apresenta a noção de diferença ou mudança mínima com relação a mundos possíveis. Posteriormente, mesmo as análises de contrafactuais de cunho epistemológico apresentam o "teste de Ramsey" contendo a ideia de mudança mínima não mais com relação a mundos possíveis mas com relação a estados de crenças.

Stalnaker sustenta que na sua forma original (1) o teste de Ramsey só abarcava aqueles condicionais "abertos", *i.e.*, aqueles que têm o valor de verdade de sua proposição antecedente desconhecido, e propõe uma nova forma que abrangeria também aqueles condicionais que tem a proposição antecedente falsa. O critério (2) formulado por Stalnaker para conjuntos de crenças é [Sta68, pág. 102]:

---

[i]Termo que parece ter sido criado por Harper segundo [LR95, pág.148]

> (2) [...] Assim se avalia um condicional: Primeiro adicione o antecedente (hipoteticamente) a seu estoque de crenças; segundo, faça todos os ajustes necessários para manter a consistência (sem modificar a crença hipotética no antecedente); finalmente, considere então se o consequente é ou não verdadeiro. [24]

Stalnaker propõe uma divisão do problema tradicional dos contrafactuais em dois aspectos: o problema lógico e o problema pragmático. A resolução do problema lógico envolve a explicitação de critérios formais para a verdade de contrafactuais. Para a resolução do problema pragmático cabem considerações de natureza epistemológica, como de plausibilidade e valor informativo. Visto que Stalnaker estava em busca de condições de verdade e não de asseribilidade, ele propõe que as considerações sobre estoques de crenças sejam substituídas por considerações acerca de mundos possíveis pois, segundo ele, um mundo possível é o análogo ontológico de um estoque de crenças. Então, (2) se transforma no critério (3) de Stalnaker para mundos possíveis. Ele afirma (*ibid.*) que, para avaliar um condicional "se a, então b":

> (3) [...] considere um mundo possível no qual a é verdadeira, e que de outro modo difira minimamente do mundo atual. [O condicional] "Se a, então b" é verdadeiro (falso) somente no caso em que B é verdadeira (falsa) nesse mundo possível".[25]

Gärdenfors toma o critério (3) de Stalnaker e o formula em termos de estados de crenças ao invés de mundos possíveis [Gär88, pág. 146]:

> (4) Aceite uma sentença da forma 'Se a, então c' em um estado de crenças K se e somente se a mudança mínima necessária para aceitar a requer que c seja aceito.[26]

Representando a frase "a implica contrafactualmente c" por $a > c$, o critério (4) é captado formalmente por Gärdenfors da seguinte maneira (*ibid.*,p. 148]:

(RT) $\quad a > c \in K \;\Leftrightarrow\; c \in K^*_a$

Essa formulação contém uma assunção tácita importante da abordagem de Gärdenfors e que será criticada posteriormente como a causadora de certos resultados negativos. A assunção é que sentenças do tipo $a > c$ pertencem a linguagem objeto e que podem ser elementos do conjunto de crenças. Analisaremos essa assunção posteriormente na seção 2.4.3.

Em seguida será exposta a lógica de contrafactuais de Gärdenfors, na qual ele propõe alguns axiomas e regras de inferência e mostra que esses axiomas correspondem mais ou menos aos postulados para operação de revisão.

2.2 A Lógica dos Contrafactuais de Gärdenfors

Seja $\mathcal{L}$ uma linguagem do cálculo proposicional tal que sua lógica subjacente contenha *modus ponens*, seja consistente e completa. Seja $\mathcal{L}'$ a extensão mínima de $\mathcal{L}$ fechada sob o conectivo > de implicação contrafactual (a implicação material será representada por "→"). Um sistema de revisão de crenças é definido da seguinte maneira. Seja $\mathcal{K}$ um conjunto de estados de crenças e $(*)$ uma operação de revisão de $K \times \mathcal{L}'$ para $\mathcal{K}$. É assumido que a operação $(*)$ satisfaz **RT** e que o conjunto $\mathcal{K}$ é fechado sob expansão, isto é, se $K \in \mathcal{K}$, então $K_a^+ \in \mathcal{K}$ para todo $a \in \mathcal{L}'$.

A definição de validade neste sistema é como se segue (*ibid.*):

(Validade)

> Uma fórmula a de $\mathcal{L}'$ é **satisfazível em um sistema de revisão de crenças** $< \mathcal{K}, * >$ sse existe algum $K \in \mathcal{K}$ tal que $K \neq K^{\perp}$ e $a \in K$. Uma fórmula a é **válida em um sistema** $< \mathcal{K}, * >$ sse $\neg a$ não é satisfazível no sistema. Uma fórmula a é (logicamente) **válida** sse a é válida em todos os sistemas de revisão de crenças.[27]

Gärdenfors propõe (*ibid.*, pág. 149) um sistema dedutivo **CM** que consiste em um conjunto de axiomas básicos (A_1)–(A_3), um conjunto de axiomas suplementares (A_4)–(A_{10}) e duas regras de inferência. Os axiomas (A_4)–(A_{10}) baseiam-se nos postulados para operação de revisão. É mostrado para cada um desses (A_i), que um sistema satisfaz (A_i) sse satisfaz o critério para revisão correspondente.

Os axiomas básicos são:

(A_1)	Todas tautologias vero-funcionais;
(A_2)	$(a > b) \wedge (a > c) \rightarrow (a > b \wedge c)$;
(A_3)	$a > \top$.

As regras de inferência são:

(DR_1)	*modus ponens*;
(DR_2)	Se $b \rightarrow c$ é um teorema, então $(a > b) \rightarrow (a > c)$ é um teorema.

Seguem os demais axiomas que são baseados nos critérios para revisão e no teste de Ramsey **RT**. Antes de apresentá-los, recordamos os critérios para revisão e **RT**:

(***-fecho**)	K_a^* é um conjunto de crenças;
(***-sucesso**)	$a \in K_a^*$;
(***-inclusão**)	$K_a^* \subseteq K_a^+$;
(***-vacuidade**)	$\neg a \notin K \Rightarrow K_a^+ \subseteq K_a^*$;
(***-consistência**)	$K_a^* = K^{\perp} \Leftrightarrow \vdash \neg a$;
(***-extensionalidade**)	$\vdash a \leftrightarrow b \Leftrightarrow K_a^* = K_b^*$;
(***-suplementar**$_1$)	$K_{a \wedge b}^* \subseteq (K_a^*)_b^+$;
(***-suplementar**$_2$)	$\neg b \notin K_a^* \Rightarrow (K_a^*)_b^+ \subseteq K_{a \wedge b}^*$;
(**RT**)	$a > c \in K \Leftrightarrow c \in K_a^*$.

O axioma que corresponde ao critério ***-sucesso**, facilmente obtido através de **RT**, é o seguinte:

$(A_4) \qquad a > a.$

O critério ***-inclusão** é garantido pelo seguinte axioma, que diz que de um contrafactual se segue a implicação material correspondente:

$(A_5) \qquad (a > b) \to (a \to b).$

Como será mostrado adiante, ***-vacuidade** é incompatível com **RT**. Portanto, ***-vacuidade** não pode ser formulado completamente. O axioma A_6 deve ser formulado de acordo com uma versão mais fraca deste critério:

(***-vacuidade**$_f$) $\qquad a \in K \;\&\; K \neq K^{\perp} \;\Rightarrow\; K \subseteq K^*_a.$

Gärdenfors propõe que ***-vacuidade**$_f$ seja captado pelo axioma 6:

$(A_6) \qquad a \wedge b \;\to\; (a > b).$

O caráter contra-intuitivo desse axioma é facilmente detectado através de um exemplo. Suponha que João tenha um carro e uma moto. Intuitivamente não consideramos que isso seja razão para afirmar que, se João tivesse um carro, ele teria uma moto.

Uma maneira direta de representar ***-consistência** não foi encontrada. A forma mais próxima, de acordo com Gärdenfors é a seguinte:

$(A_7) \qquad (a > \neg a) \;\to\; (b > \neg a).$

A ideia por trás de (A_7), utilizando **RT**, é que, se $\neg a \in K^*_a$, então $\neg a \in K^*_b$ para um b qualquer e isso significa que $\neg a$ é um teorema.

O critério ***-extensionalidade** corresponde ao seguinte esquema:

$(A_8) \qquad (a > b) \wedge (b > a) \;\to\; ((a > c) \to (b > c)).$

Por **RT**, observe que, se $b \in K^*_a$ e $a \in K^*_b$ e algum c está em K^*_a, então $c \in K^*_b$, *i.e.* $K^*_a = K^*_b$

Quanto ao critério ***-suplementar**$_1$, seu axioma correspondente é:

$(A_9) \qquad (a > c) \wedge (b > c) \;\to\; (a \vee b > c).$

Utilizando **RT**, se $K^*_a \vdash c$ e $K^*_b \vdash c$, de uma forma análoga à regra de eliminação da disjunção em sistemas de dedução natural[ii], $K^*_{a\vee b} \vdash c$. O critério ***-vacuidade** é um caso especial do critério ***-suplementar**$_2$ e, como ***-vacuidade** não pode ser formulado completamente, ***-suplementar**$_2$

[ii] A regra de eliminação da disjunção é a seguinte:

$$\frac{a \vee b \qquad \begin{array}{c}[a],\Gamma\\ c\end{array} \qquad \begin{array}{c}[b],\Theta\\ c\end{array}}{c}$$

também não pode. A ideia de Gärdenfors é que b pode ser adicionado consistentemente a K_a^* quando "b **poderia** ser aceito [em K] se a fosse aceito [em K]" [Gär78, pág. 392]. Ele diz que há duas maneiras de capturar formalmente essa ultima oração, ou como $\neg(a > \neg b) \in K$ ou $\neg b \notin K_a^*$, embora, como Gärdenfors reconhece (*ibid.*), a última não implique a primeira. De $\neg b \notin K_a^*$, por **RT**, só é possível concluir que $(a > \neg b) \notin K$. Para que fosse possível concluir ademais que $\neg(a > \neg b) \in K$ seria necessário supor que K fosse completo com relação aos condicionais contrafactuais, o que não seria nada razoável.

Gärdenfors escolhe (*ibid.*) formular a ideia de que b **poderia** ser aceito em K, se a fosse aceito, como "$\neg(a > \neg b) \in K$". Usando essa ideia, o seguinte critério foi proposto em lugar do ***-suplementar**$_2$ ($\neg b \notin K_a^* \;\Rightarrow\; (K_a^*)_b^+ \subseteq K_{a\wedge b}^*$):

$$(\textbf{*-suplementar}_{2f}) \qquad \neg(a > \neg b) \in K \;\Rightarrow\; (K_a^*)_b^+ \subseteq K_{a\wedge b}^*.$$

O axioma que corresponde a ***-suplementar**$_{2f}$ é:

$$(A_{10}) \qquad (a > b) \wedge \neg(a > \neg c) \;\rightarrow (a \wedge c > b).$$

Não obstante, essa escolha é problemática por dois motivos: (i) não foram dadas condições suficientes para a inferência da negação de um contrafactual, (ii) há contra-exemplos intuitivos para o axioma resultante dessa formulação.

Quanto a (i), observe-se que o teste de Ramsey só fornece uma condição suficiente e necessária para a **asserção** de contrafactuais, a saber, que um contrafactual pode ser asserido por um agente se e somente se seu consequente pertence ao estado de crenças do agente revisado pelo antecedente. Isto não diz nada acerca de quando a negação de um contrafactual pode ser asserida.

Goodman faz uma análise bastante interessante do significado dos contrafactuais e de suas negações [Goo47]. Para ele, um condicional contrafactual como $(a > c)$ "se o fósforo tivesse sido riscado, ele teria acendido", assere uma conexão entre o antecedente a "o fósforo foi riscado" e o consequente c "o fósforo acendeu", essa asserção deve ser feita a partir de um conjunto S apropriado de informações adicionais. Com relação à negação de contrafactuais, Goodman sustenta que a sentença $(\neg(a > c))$ "**é falso** que se o fósforo tivesse sido riscado, ele teria acendido" é equivalente à sentença (**A**) "mesmo que o fósforo tivesse sido riscado, ele ainda assim não teria acendido". Goodman denomina a **negação** de um condicional contrafactual por "condicional **semifactual**". A nomenclatura faz referência ao fato de que quem assere um semifactual, como a sentença **A**, acredita na verdade do consequente, isto é, acredita que o fósforo não acendeu, ao contrário de quem assere o contrafactual $a > c$, que acredita na falsidade de ambos a e c. Ademais, ao contrário dos contrafactuais, a asserção de um semifactual manifesta a crença na **desconexão** entre antecedente e consequente. Quem assere **A** está se comprometendo a defender que o riscamento do fósforo não permite concluir ou não implica o seu acendimento. De acordo com essas considerações, a partir do fato de que $\neg c \notin K_a^*$, não é possível asserir que o semi-factual $\neg(a > \neg c)$ está em K, outras condições teriam que ser cumpridas para que essa asserção pudesse ser feita. Um critério para a asserção de semifactuais poderia ser: se $\neg c \in K^*_a$ e $\neg c \in K^*_{\neg a}$, então $\neg(a > \neg c) \in K$.

Quanto a (ii), seja K o conjunto de crenças formado a partir das sentenças

($\neg a$) não há combustível;
($\neg b$) o carro não funciona;
($\neg c$) o motor não está com defeito.

É razoável dizer, a partir de K que $b \in K^*_a$ e, portanto, que $a > b \in K$, em outros termos, afirmar que, se houvesse combustível, o carro funcionaria. Do mesmo modo, é plausível dizer que é falso que, se houvesse combustível, o motor não estaria com defeito ($\neg(a > \neg c)$). Neste caso, usando o critério proposto logo acima para a asserção de semi-factuais, é admissível dizer que $\neg(a > \neg c) \in K$, visto que aceitamos que $\neg c \in K^*_{\neg a}$ e $\neg c \in K^*_a$. Contudo, a partir de K não é aceitável dizer que se houvesse combustível e o motor estivesse com defeito, o carro funcionaria ($a \wedge c > b$), em outros termos, $(a \wedge c > b) \notin K$.

Outros contra-exemplos podem facilmente ser encontrados, por exemplo, tomando $\neg a$, $\neg b$ e $\neg c$ por "eu não sou rico", "eu não sou feliz" e "eu não tenho câncer", respectivamente. Pode-se aceitar os condicionais "se eu fosse rico, seria feliz" ($a > b$), "é falso que se eu fosse rico, eu não teria câncer" ($\neg(a > \neg c)$), mas não é possível aceitar o condicional "se eu fosse rico e tivesse câncer, eu seria feliz" ($a \wedge c > b$).

Em suma, os axiomas e regras de inferência para a lógica de contrafactuais **CM** de Gardenfors são:

(A_1) Todas tautologias vero-funcionais;
(A_2) $(a > b) \wedge (a > c) \rightarrow (a > b \wedge c)$;
(A_3) $a > \top$;
(A_4) $a > a$;
(A_5) $(a > b) \rightarrow (a \rightarrow b)$;
(A_6) $a \wedge b \rightarrow (a > b)$;
(A_7) $(a > \neg a) \rightarrow (b > \neg a)$;
(A_8) $(a > b) \wedge (b > a) \rightarrow ((a > c) \rightarrow (b > c))$;
(A_9) $(a > c) \wedge (b > c) \rightarrow (a \vee b > c)$;
(A_{10}) $(a > b) \wedge \neg(a > \neg c) \rightarrow (a \wedge c > b)$;
(DR_1) *modus ponens*;
(DR_2) Se $b \rightarrow c$ é um teorema, então $(a > b) \rightarrow (a > c)$ é um teorema.

A partir do sistema **CM** acima Gärdenfors consegue obter um resultado muito interessante. Ele consegue mostrar que **CM** é equivalente à conhecida lógica para contrafactuais "oficial" ou **VC** de Lewis [Lew73, pág. 130]. Gärdenfors obtém que [Gär78, pág. 151]:

> Uma fórmula a é um teorema em **VC** sse ela é derivável a partir de **CM** juntamente com os esquemas de axioma (A_4)–(A_{10}).[28]

VC pode ser resumido nas seguintes regras e axiomas [Lew73, pág. 123]. Só serão dadas duas definições intuitivas para os operadores. Caso o leitor tenha interesse nas demais, consultar (*ibid.*, págs. 118-119).

Regras:
(1) *Modus Ponens*,
(2) Regra para Possibilidade Comparativa: para qualquer $n \geq 1$,

$$\frac{\vdash \varphi \rightarrow (\psi_1 \vee \ldots \vee \psi_n)}{\vdash (\psi_1 \preccurlyeq \varphi) \vee \ldots \vee (\psi_n \preccurlyeq \varphi)}$$

Axiomas:
(1) Tautologias vero-funcionais,
(2) Definições dos operadores,
(3) Transitividade: $((\varphi \preccurlyeq \psi) \wedge (\psi \preccurlyeq \chi)) \rightarrow (\varphi \preccurlyeq \chi)$,
(4) Conectividade: $(\varphi \preccurlyeq \psi) \vee (\psi \preccurlyeq \varphi)$,
(5) Centralização (**C**): $\diamondsuit\varphi \rightarrow \varphi$.

A relação $\varphi \preccurlyeq \psi$ significa que φ é ou mais ou igualmente próxima ao mundo atual quanto ψ. A sentença "$\diamondsuit\varphi$" é verdadeira em um mundo i se e somente se φ é verdadeira no "melhor" mundo (de acordo com $\preccurlyeq$) em relação a i. O axioma C diz, portanto, que o melhor mundo (de acordo com $\preccurlyeq$) é o mundo atual [Lew73, pág. 101].

O teorema de Gärdenfors acima é extremamente interessante, pois mostra que a lógica **CM**, que tem como conceito básico o conceito de estado de crenças, consegue demonstrar os mesmos teoremas que a lógica **VC**, cujas noções fundamentais baseiam-se no conceito complexo de mundo possível. Ademais, a lógica **VC** possui onze constantes lógicas além das constantes do cálculo proposicional[iii]. E ainda uma regra de inferência para lidar com a noção obscura de possibilidade comparativa.

Ao adotar a abordagem de Gärdenfors em detrimento da abordagem de Lewis, livramo-nos dos problemas que a noção de similaridade entre mundos possíveis acarreta, e ficamos com a noção menos problemática de mudança mínima entre estados de crenças, além disso, com um aparato formal muito mais simples. Essa escolha, do ponto de vista conceitual, apresenta um ganho considerável.

2.3 A Incompatibilidade de RT com *-vacuidade

A abordagem de revisão de crenças para os condicionais contrafactuais tem um ponto problemático. O postulado ***-vacuidade** ($\neg a \notin K \Rightarrow K_a^+ \subseteq K_a^*$) é um dos postulados que tentam fazer com que as revisões sobre conjuntos de crenças apenas efetuem mudanças mínimas. ***-vacuidade** é um postulado que, dado seu apelo intuitivo, deveria ser válido na lógica **CM**. Caso

[iii] Ver [Lew73, pág. 118 e ss]

o postulado na sua força total não possa ser incorporado, pelo menos uma consequência dele, aqui denominada **preservação**, deveria poder ser incorporada:

(**preservação**)

Se uma sentença b pertence a um dado estado de crenças K e a é consistente com as crenças K, então b é ainda aceita na mudança mínima de K necessária para incluir a.

O critério de preservação tem seu fundamento ainda mais fortemente no critério de mudança mínima. Assim, tanto o teste de Ramsey quanto **preservação** são de grande interesse na análise da dinâmica de crenças. Gärdenfors mostra, contudo, que eles são incompatíveis.

Para mostrar a impossibilidade de combinar **RT** com **preservação**, faremos uma lista das assunções necessárias para a prova. Relembramos o teste de Ramsey, tal como formalizado por Gärdenfors:

(**RT**) $\quad a > c \in K \Leftrightarrow c \in K_a^*.$

Observamos que **RT** depende da seguinte assunção:

($\varnothing$) Conjuntos de crenças incluem sentenças que contém o conectivo condicional “>”.

Embora essa assunção pareça razoável, há pelo menos um autor[iv] que colocou nela a culpa dos resultados de impossibilidade. Para que seja obtido um resultado ainda mais forte, não será utilizado na prova o critério **RT**, mas uma consequência dele denominada “critério de monotonicidade” (***-monotonicidade**):

(***-monotonicidade**) $\quad K \subseteq K' \Rightarrow K_a^* \subseteq K_a'^*.$

Teorema 2.3.1. ***RT*** *implica* ****-monotonicidade****.*

Demonstração. Suponha que $b \in K_a^*$ e $K \subseteq K'$. Segue-se (por **RT**) que $a > b \in K$, logo $a > b \in K'$. Portanto, (por **RT**), $b \in K_a'^*$. Assim, $K_a^* \subseteq K_a'^*$. □

Outro requisito para prova é o critério de preservação, que é formalizado assim:

(***-preservação**) $\quad \neg a \notin K \ \& \ b \in K \Rightarrow b \in K_a^*.$

Serão utilizadas três postulados da operação de revisão, ***-fecho**, ***-sucesso** e ***-consistência**

(***-fecho**) K_a^* é um conjunto de crenças para quaisquer K e a;
(***-sucesso**) $a \in K_a^*$;
(***-consistência**) $K_a^* = K^\perp \Leftrightarrow \vdash \neg a$.

[iv] Isaac Levi, citado por [Gär88, pág. 164]

Outro pressuposto[v], apontado como importante por Hansson [Han92b, pág. 524] e Morreau [Mor92, p. 49] é o fecho por expansão do sistema de crenças:

(**fecho-por-+**) $K \in \mathcal{K} \Rightarrow K_a^+ \in \mathcal{K}$, para toda sentença a.

Finalmente, faremos uso de mais duas propriedades da operação de expansão (+). Lembrando que $K_a^+ = Cn(K \cup \{a\})$

($^+$**A**) $(K_a^+)_b^+ = K_{a\&b}^+$;
($^+$**B**) $K_{a\vee b}^+ \subseteq K_a^+$.

Por **sistema de revisão de crenças não-trivial** entende-se um sistema $<\mathcal{K}, *>$ em que há pelo menos três sentenças a, b e c tais que, tomadas de par em par, são incompatíveis, *i.e.* $\vdash \neg(a\&b)$, $\vdash \neg(a\&c)$ e $\vdash \neg(b\&c)$, e que há um conjunto de crenças $K \in \mathcal{K}$ que é consistente com as três sentenças, *i.e.*, $\neg a \notin K$, $\neg b \notin K$ e $\neg c \notin K$.

Que a assunção de não-trivialidade seja inócua é facilmente perceptível por um exemplo (retirado de [Ten08, pág. 411]): seja K um conjunto de crenças que não contenha nenhuma informação acerca da cor de determinada bola, e seja (a) "a bola é verde", (b) "a bola é azul" e (c) "a bola é vermelha". As três são incompatíveis entre si, embora, tomadas individualmente, sejam compatíveis com K.

O teorema da impossibilidade de Gärdenfors [Gär88, pág. 168] é o seguinte:

Teorema 2.3.2. (Teorema da impossibilidade) *Não há sistema de revisão de crenças não-trivial que satisfaça todas as condições* ****-fecho**, ***-sucesso**, ***-consistência**, $^+$**A**, $^+$**B**, ***-monotonicidade*** *e* ****-preservação***.

Demonstração. Assuma, para redução ao absurdo, que (H) algum sistema $<\mathcal{K}, *>$ não-trivial que satisfaça todas as condições ***-fecho**, ***-sucesso**, ***-consistência**, $^+$**A**, $^+$**B**, ***-monotonicidade** e ***-preservação**.

Segue-se pela não-trivialidade que K_a^+, K_b^+, K_c^+, $K_{a\vee b}^+$, $K_{a\vee c}^+$ e $K_{c\vee b}^+$ não são inconsistentes. Assim, $(K_a^*)_{b\vee c}^* \neq K^\perp$ e, por ***-sucesso**, $b \vee c \in (K_a^*)_{b\vee c}^*$. Visto que $\{b \vee c, \neg b, \neg c\} \vdash \perp$, segue-se que $\neg b \notin (K_a^*)_{b\vee c}^* \vee \neg c \notin (K_a^*)_{b\vee c}^*$.

Suponha que $\neg c \notin (K_a^*)_{b\vee c}^*$. Por $^+$**B**, $K_{a\vee b}^+ \subseteq K_a^+$. Visto que $\neg a \notin K$, por ***-preservação**, segue-se que $K \subseteq K_a^*$. Por ***-sucesso**, temos $a \in K_a^*$. Dado que $K_a^+ = K \cup \{a\}$, logo $K_a^+ \subseteq K_a^*$ e $K_{a\vee b}^+ \subseteq K_a^*$. Disso, por ***-monotonicidade**, segue-se $(K_{a\vee b}^+)_{b\vee c}^* \subseteq (K_a^*)_{b\vee c}^*$. Logo, $\neg c \notin (K_{a\vee b}^+)_{b\vee c}^*$.

Suponha agora que $\neg(b \vee c) \in K_{a\vee b}^+$, então $(a \vee b) \rightarrow \neg(b \vee c) \in Cn(K)$ ou, equivalentemente (visto que $\vdash \neg(a \wedge c)$), $\neg b \wedge \neg(a \wedge c) \in Cn(K)$, o que implica que $\neg b \in Cn(K)$, contrariando a hipótese de não-trivialidade. Portanto, $\neg(b \vee c) \notin K_{a\vee b}^+$. Disso se segue, por ***-preservação**, que $K_{a\vee b}^+ \subseteq (K_{a\vee b}^+)_{b\vee c}^*$ e, por ***-sucesso**, que $b \vee c \in (K_{a\vee b}^+)_{b\vee c}^*$. Por ***-fecho**, 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, segue-se que $Cn(K \cup \{a \vee b, b \vee$

[v] em [Gär86, p. 85]

$c\}) = Cn(K \cup \{b\})$. Assim, $K^+_b \subseteq (K^+_{a\vee b})^*_{b\vee c}$ e $b \in (K^+_{a\vee b})^*_{b\vee c}$. Pela assunção de não-trivialidade, $\vdash\ b \to \neg c$. Logo, por ***-fecho**, $\neg c \in (K^*_{a\vee b})^*_{b\vee c}$, o que contradiz a conclusão obtida anteriormente.

A demostração de que $\neg b \notin (K^*_a)^*_{b\vee c}$ também leva a um absurdo é similar. □

2.3.1 Tentativas para Evitar o Teorema da Impossibilidade

Foi apontado por Rott [Gär88, pág. 160] que o teorema de impossibilidade é consequência da seguinte consequência de **RT** ($a > c\ \in K\ \Leftrightarrow\ c \in K^*_a$):

($RT_{trivial}$) $\qquad a \wedge c \in K \Rightarrow a > c\ \in K.$

Essa consequência é implausível em geral porque a partir de que João tenha uma bicicleta e que João tenha uma moto, não segue-se que, se João tivesse uma bicicleta ele teria uma moto. Isto é, do fato de que duas sentenças pertençam a um estado de crenças não segue que haja uma relação de implicação contrafactual entre elas.

Gärdenfors propõe então os seguintes critérios que não implicam $RT_{trivial}$:

(**RT**$_{f1}$) $\qquad a > c\ \in K \Leftrightarrow c \in K^*_a\ \&\ c \notin K,$

(**RT**$_{f2}$) $\qquad a > c\ \in K \Leftrightarrow c \in K^*_a\ \&\ c \notin K^*_{\neg a},$

(**RT**$_{f3}$) $\qquad a > c\ \in K \Leftrightarrow c \in (K^-_c)^*_a,$

(**RT**$_{f4}$) $\qquad a \vee c\ \notin K \Rightarrow (a > c\ \in K \Leftrightarrow c \in K^*_a).$

O critério **RT**$_{f1}$ tenta captar a ideia de um contrafactual $a > c$ é válido apenas quando $c \in K^*_a$ é consequência de que $a \in K^*_a$. **RT**$_{f2}$ tenta formalizar a noção de que $a > c$ é válido apenas quando a é relevante para inferir c a partir de K^*_a, visto que $c \notin K^*_{\neg a}$. A ideia por trás do critério **RT**$_{f3}$ é bastante similar àquela do critério **RT**$_{f1}$. Por fim, o critério **RT**$_{f4}$ procura captar a noção de que $a > c$ é válido quando nem a nem c estão em K e a adição do antecedente a a K faz com que seja possível inferir c.

São demonstráveis[vi] as seguintes implicações:

- **RT**$_{f1}$ ⇒ **RT**$_{f4}$,
- **RT**$_{f2}$ & ***-inclusão** ⇒ **RT**$_{f4}$,[vii]
- **RT**$_{f3}$ & **–vacuidade** ⇒ **RT**$_{f4}$.[viii]

Seja ***-monotonicidade**$_f$ o seguinte critério:

(***-monotonicidade**$_f$) $\qquad (K \subseteq K', a \vee c\ \notin K'\ \&\ c \in K^*_a) \Rightarrow c \in K'^*_a.$

Gärdenfors demonstra (*ibid.*, lema 7.13) que **RT**$_{f4}$ ⇒ ***-monotonicidade**$_f$ e, além disso, que ***-monotonicidade**$_f$ é incompatível com o critério ***-preservação**.[ix] Assim, mesmo ado-

[vi] As demonstrações se encontram em [Gär88, pág. 241, lemma 7.12]
[vii] (*-inclusão) $K^*_a \subseteq K^+_a$.
[viii] (–vacuidade) $a \notin K\ \Rightarrow\ K^-_a = K$.
[ix] (*-preservação) $\neg a \notin K$ e $b \in K\ \Rightarrow\ b \in K^*_a$

tando algum dos critérios $\mathbf{RT}_{f1}$–$\mathbf{RT}_{f3}$ em lugar de **RT**, obtém-se a seguinte variação do teorema da impossibilidade:

> Não há sistema de revisão de crenças não-trivial que satisfaça todas as condições ***-fecho, *-sucesso, *-consistência**, $^{+}$**A**, $^{+}$**B**, ***-monotonicidade**$_f$ e ***-preservação**.

Gärdenfors aponta (*ibid.*, pág. 162) que não conseguiu obter resultados de trivialidade para o seguinte princípio:

$$(\mathbf{RT}_{f5}) \qquad \neg a \in K \Rightarrow (a > c \in K \Leftrightarrow c \in K^{*}_{a}).$$

A razão para tal está no fato de que $\mathbf{RT}_{f5}$ assume que $\neg a \in K$ fazendo que a condição do critério ***-preservação** ($\neg a \notin K$) não seja cumprida. Gärdenfors, não obstante, afirma que $\mathbf{RT}_{f5}$ é muito fraco para capturar o conteúdo intuitivo de **RT**, pois $\mathbf{RT}_{f5}$ não abarcaria os chamados "condicionais abertos", isto é, aqueles condicionais $a > c$, tais que nem a nem $\neg a$ pertencem a K.[x] Substituir **RT** por $\mathbf{RT}_{f5}$, segundo ele, não seria uma forma promissora de escapar dos resultados de trivialidade.

A posição de Gärdenfors é questionável. Nas análises conceituais é razoável e talvez até recomendável discriminar as nuances de cada conceito e tratá-las em separado. Neste sentido, discriminar os condicionais contrafactuais dos condicionais abertos e dar-lhes uma abordagem em separado pode não só facilitar a tarefa como evitar que apareçam problemas alheios aos conceitos em foco, como foi o caso da incompatibilidade de **RT** com ***-preservação**. Quando um condicional contrafactual $a > c$ é asserido com base em um conjunto de crenças K, há normalmente a crença na verdade de $\neg a$, isto é, $\neg a \in K$. É exatamente por isso que tais condicionais são denominados "contrafactuais". Levando isso em conta, $\mathbf{RT}_{f5}$ poderia servir de critério de aceitabilidade de contrafactuais em termos de revisão de crenças. Contudo, como veremos adiante (seção 3), esse sistema de revisão de crenças não pode ser tão irrestrito como o modelo AGM.

As operações para revisão no modelo AGM, portanto, não podem ser compatibilizadas com o teste de Ramsey. As propriedades ($^{+}$**A**) e ($^{+}$**B**) são simples consequências do fato de que os conjuntos de crenças são fechados sob consequência lógica. Como os critérios ***-fecho, *-sucesso, *-consistência** são mais elementares que ***-preservação**, é razoável dizer que a incompatibilidade se dá particularmente entre **RT** e ***-preservação**, de modo que um dos dois deve ser rejeitado. Gärdenfors atribui [Gär88, pág. 159] a "culpa" à **RT**:

> Minha posição presente é que **RT** é o culpado. Uma forma de argumentar em favor disso é, acredito, considerar ***-monotonicidade** e suas consequências para os sistemas de revisão de crenças. [29]

Gärdenfors argumenta (*ibid.*) contra ***-monotonicidade** com o seguinte exemplo. Seja K o conjunto das crenças de Vitória. Suponha que ela acredite que (a) tem sangue tipo O, que (b)

[x] Esses condicionais são geralmente expressados utilizando o verbo do antecedente no modo indicativo presente, como "se chove, então faz frio".

João seja seu pai e, ademais, suponha que Vitória ignore o tipo sanguíneo de seu pai. Suponha que Vitória aprenda que (c) uma pessoa com sangue tipo AB não pode ter filhos com sangue do tipo O. Seja J seu novo estado de crenças. Como c é consistente com K, J=K_c^+ e $K \subset J$. Suponha que Vitória descubra que (d) seu pai tem sangue tipo AB. Agora é razoável que $b \notin J_d^*$. Contudo é razoável que b permaneça em K_d^*, o que significa que o critério de monotonicidade da revisão falha nesse caso.

2.4 Defendendo o Teste de Ramsey

2.4.1 Neil Tennant e o Teorema de Impossibilidade de Gärdenfors

Tennant argumenta que o problema está no critério ***-vacuidade** (que implica ***-preservação**). O argumento dele é como se segue.

O modelo AGM tem como propósito circunscrever o âmbito daquilo que poderia ser chamado de uma "teoria de revisão de crenças de agentes idealmente racionais", isto é, agentes que estão cientes de todas as consequências de suas crenças e que não sustentam crenças incompatíveis entre si. As verdades e falsidades lógicas não podem efetuar mudanças no estado de crenças de um agente racionalmente ideal, visto que este não pode deixar de acreditar em verdades lógicas e tampouco pode passar a acreditar em falsidades lógicas.[xi] Portanto, as únicas sentenças que podem provocar mudanças nos estados de crenças são as sentenças contingentes. Com respeito a uma sentença contingente a e um conjunto K consistente, há três possíveis situações:

1. a não está em K e é consistente com K,
2. a está em K,
3. a não está em K e é inconsistente com K.

Tennant sustenta que uma teoria de revisão de crenças deveria separar os âmbitos de atuação e lidar com os casos principais em cada âmbito. Assim, cada situação (1–3) deveria corresponder exclusivamente a um tipo de mudança de crença. Em (1) a única mudança de crença que pode ocorrer é de que a seja adicionada a K, essa mudança seria denominada **expansão** (propriamente dita). Em (2) só se pode mudar K com relação a a, retirando a de K, mudança chamada de **contração** (propriamente dita). Em (3) a mudança que pode incindir sobre K consiste na adição de a a K de forma que o conjunto resultante permaneça consistente e que essa seja uma mudança mínima para fazê-lo, essa mudança é denominada **revisão** (propriamente dita). Portanto, cada operador de mudança **só deve estar definido para a situação que lhe é própria**, isto é, a expansão só está definida para a situação (1), contração para (2) e revisão para (3).

[xi]Deste modo, tal como formulado no modelo AGM, um agente idealmente racional não pode mudar de opinião sobre a lógica que rege suas inferências.

Segundo Tennant, o teorema de impossibilidade de Gärdenfors torna-se demonstrável somente quando são mesclados os âmbitos de atuação (2) e (3) das funções de contração e revisão. De acordo com essa abordagem, vários postulados do modelo AGM perdem o valor: o postulado $^+$**-vacuidade** ($a \in K \;\Rightarrow\; K_a^+ = K$), ***-vacuidade** ($\neg a \notin K \;\Rightarrow K_a^+ \subseteq K_a^*$) e o postulado **–vacuidade** ($a \notin K \;\Rightarrow\; K_a^- = K$). Ele diz [Ten08, pág. 414]: "[...] se φ é consistente com K, então não há realmente nada, intuitivamente, como uma revisão de K com respeito a φ. Porque nada em K necessitaria ser descartado ao passar a comprometer-se com φ". [30]

Ademais, Tennant alega que [Ten08, pág. 416]

> [...] em [Ten6b] prova-se um teorema que afirma a existência de funções de revisão que satisfazem ambos os postulados AGM básicos para revisão (propriamente reformulados) e o princípio da monotonicidade para revisões (propriamente restringidos ao caso em que $K, \varphi \vdash \bot$). [31]

Não obstante, no referido artigo, Tennant oferece apenas duas construções pouco interessantes, por assim dizer, de revisão monotônica. Na primeira é definida uma função de revisão $\rho(K,a)$ em que a é consistente e K é um conjunto de crenças consistente que implica $\neg a$. $\rho(K,a) = Cn(a)$ para todo K consistente e incompatível com a. Assim, se $K \subseteq K'$, trivialmente $\rho(K,a) \subseteq \rho(K',a)$. No modelo AGM (seção 1.5 acima) uma função equivalente a $\rho()$ é definível, ela é a função de revisão de intersecção total (^{*fm}) (*i.e.* $K_a^{*fm} = (K_{\neg a}^{-fm})_a^+$). Para qualquer K, K_a^{*fm}=$Cn(a)$. Assim, se $K \subseteq K'$, então $K_a^{*fm} \subseteq K_a'^{*fm}$, ou seja, (^{*fm}) é monotônica.

A segunda construção que mostra a existência de funções de revisão monotônicas utiliza indução transfinita sobre teorias consistentes e sentenças consistentes incompatíveis com tais teorias [Ten6b, pág. 503]. Nenhum caso concreto desse tipo de função de revisão é mostrado. O aparato utilizado para provar a existência de tais funções tampouco tem apelo intuitivo de modo a ter alguma valia para a elucidação da questão da existência de funções de revisão monotônicas.

Em suma, a crítica de Tennant com relação à separação dos "papéis" de cada função de mudança de crenças é relevante. Se a crítica é aceita, então não vale o postulado ***-vacuidade** para revisão. Gärdenfors enuncia seu teorema de impossiblidade, como a prova de incompatibilidade entre **RT** e ***-vacuidade** . Diz Gärdenfors [Gär88, pág. 159] "O teorema e seu corolário mostram que o teste de Ramsey **RT** e a condição de preservação [***-preservação**, ou, equivalentemente, ***-vacuidade**] não podem ambos serem critérios racionais para revisões de crenças".[32] Se ***-vacuidade** é rejeitado a princípio como um postulado para revisão, naturalmente a prova não funciona. O que não significa que não seja possível construir outra prova da impossibilidade de se combinar ***-monotonicidade** com os outros postulados para revisão do modelo AGM. Para que a existência de tal prova fosse realmente descartada, um contra-exemplo de função de revisão não-trivial monotônica deveria ser apresentada.

2.4.2 Sven O. Hansson Sobre o Teste de Ramsey

O teorema de impossibilidade de Gärdenfors depende das seguintes premissas:

($\varnothing$)	Conjuntos de crenças incluem sentenças que contém o conectivo condicional ">";
(***-monotonicidade**)	$K \subseteq K' \Rightarrow K^*_a \subseteq K'^*_a$;
(***-preservação**)	$\neg a \notin K$ & $b \in K \;\Rightarrow\; b \in K^*_a$;
(***-fecho**)	K^*_a é um conjunto de crenças para quaisquer K e a;
(***-sucesso**)	$a \in K^*_a$;
(***-consistência**)	$K^*_a = K^\perp \;\Leftrightarrow\; \vdash \neg a$;
(**fecho-por-+**)	$K \in \mathcal{K} \Rightarrow K^+_a \in \mathcal{K}$, para todo a;[xii]
($^+$**A**)	$(K^+_a)^+_b = K^+_{a\&b}$;
($^+$**B**)	$K^+_{a\vee b} \subseteq K^+_a$.

Hansson defende que o problema reside no critério **fecho-por-+**, seu argumento será exposto em seguida. As expressões $\mathcal{L}_0$, $\mathcal{L}_1$ e s são definidas da seguinte maneira:

- Seja $\mathcal{L}_0$ a linguagem formada a partir das sentenças da linguagem proposicional $\mathcal{L}$ que não contém o conectivo contrafactual ">";

- Seja $\mathcal{L}_1$ a linguagem composta das sentenças de $\mathcal{L}_0$ e de composições destas pelo conectivo ">" até o grau 1, isto é, para todas as sentenças $a > b$ de $\mathcal{L}_1$, a e $b \in \mathcal{L}_0$;

- Seja $s(K)$ o conjunto das sentenças **suportadas** ou **justificadas** a partir de K.

Hansson defende que algumas propriedades elementares válidas para estados de crenças formados a partir de $\mathcal{L}_0$ já não são válidas para aqueles estados de crenças que contenham sentenças de $\mathcal{L}_1$, isto é, que contenham condicionais contrafactuais. Um exemplo é a relação de inclusão. Para um conjunto das sentenças sem o conectivo contrafactual $>$ suportadas por K, *i.e.* $s(K) \cap \mathcal{L}_0$, é trivial encontrar um subconjunto próprio K' de K, tal que $s(K') \cap \mathcal{L}_0 \subset s(K) \cap \mathcal{L}_0$. Para tanto, basta efetuar uma contração em K. Para o caso de encontrar K'' tal que $K \subset K''$ e $s(K) \cap \mathcal{L}_0 \subset s(K'') \cap \mathcal{L}_0$, basta efetuar uma expansão em K.

No caso de conjuntos $s(K) \cap \mathcal{L}_1$, isto já não é tão simples, nem está garantido que seja possível [Han92b, pág. 525]. Por exemplo: alguém que acredita ser jovem e ter câncer, quando deixa de acreditar que tem câncer, poderia justificar a crença de que, se fosse rico, seria plenamente feliz. Se essa mesma pessoa passasse a acreditar que dinheiro não traz felicidade, a crença de que, se fosse rico, seria plenamente feliz, deixaria de estar justificada.

Em resumo, retirar sentenças de $\mathcal{L}_0$ de um conjunto K em $\mathcal{L}_1$, pode fazer com que novas crenças em $\mathcal{L}_1$ sejam justificadas, *i.e.*, são adicionadas novas crenças em $s(K)$. No exemplo acima, a retirada da crença de que se está com câncer (sentença de $\mathcal{L}_0$), justifica uma nova crença "se fosse rico, seria plenamente feliz" (sentença de $\mathcal{L}_1$). Já a adição de crenças em $\mathcal{L}_0$ a K, pode fazer com que crenças em $\mathcal{L}_1$ de $s(K)$ sejam removidas, por terem perdido sua justificação. No exemplo acima, a adição da crença de que dinheiro não traz felicidade (sentença em $\mathcal{L}_0$) faz com que seja abandonada a crença "se fosse rico, seria plenamente feliz" (sentença em $\mathcal{L}_1$).

Para que a prova do teorema da impossibilidade de Gärdenfors possa ser levada adiante, é necessário que haja três conjuntos de crenças em $\mathcal{L}_1$, K, K' e K'', tais que $K' \subset K$, $K'' \subset K$, $K' \nsubseteq K''$ e $K'' \nsubseteq K'$. Hansson mostra que não existem tais conjuntos K, K' e K'' que cumpram certas propriedades básicas. As propriedades que os estados de crenças em $\mathcal{L}_1$ deveriam cumprir são (*ibid.*, pág. 528):

(**FL**) Se $K \subseteq s(J)$ & $a \in Cn(K)$, então $a \in s(J)$ (fecho lógico);

(**CN**) Se $K \nvdash \bot$ & $a \in s(K)$, então $\neg a \notin s(K)$ (consistência por negação);

(**CC**) Seja $K \nvdash \bot$ & $\nvdash \neg a$. Se $a > b \ \in s(K)$ e $a > c \ \in s(K)$, então $b \wedge c$ é consistente (consistência condicional);

(**PC**) Para quaisquer a e b, se $a \in s(K)$, $b \notin s(K)$ & $\neg b \notin s(K)$, então $b > a \wedge b \ \in s(K)$ (preservação condicional).

As propriedades **FL**, **CN** e **CC** são aceitáveis mais ou menos diretamente como consequências dos conceitos de sentença suportada por um estado de crenças e de consistência de um estado de crenças. Por outro lado, a razoabilidade da propriedade **PC** não é tão clara. Hansson oferece o seguinte exemplo em seu favor (*ibid.*, pág. 529). Suponha que João tenha a crença que (p) Estocolmo é a capital da Suécia e que não tenha nenhuma opinião acerca de (q) Estocolmo é a maior cidade da Suécia. Assim, João acredita que, ($q > p \wedge q$) se Estocolmo for a maior cidade da Suécia, então ela será tanto a capital da Suécia quanto a maior cidade da Suécia.

Hansson mostra (*ibid.*) que para quaisquer estados de crenças K, K' e K'', que satisfaçam **FL**, **CN**, **CC** e **PC**, se $K' \cap \mathcal{L}_1 \subseteq K \cap \mathcal{L}_1$ e $K'' \cap \mathcal{L}_1 \subseteq K \cap \mathcal{L}_1$, então ou $K' \cap \mathcal{L}_1 \subseteq K'' \cap \mathcal{L}_1$ ou $K'' \cap \mathcal{L}_1 \subseteq K' \cap \mathcal{L}_1$. Isto é, a condição para a prova do teorema da impossibilidade não é cumprida. Como os requerimentos **FL, CN, CC** e **PC** sobre os estados de crenças em $\mathcal{L}_1$ são razoáveis, há que se reconsiderar os pressupostos iniciais da prova.

Dentre todos os pressupostos, Hansson coloca a "culpa" na condição **fecho-por-+**, dizendo que embora ela seja válida para conjuntos de crenças em $\mathcal{L}_0$, não vale para conjuntos em $\mathcal{L}_1$.

2.4.3 Isaac Levi e a Assunção $\varnothing$

Levi defende (*apud* [Han92b, pág. 524]) que o teorema da impossibilidade é "causado" pela assunção $\varnothing$:

($\varnothing$) Conjuntos de crenças incluem sentenças que contém o conectivo condicional ">".

Segundo ele, condicionais contrafactuais não são objetos de crença propriamente ditos e não podem pertencer a um estado de crenças. As crenças pertencentes a um estado de crenças seriam aquelas que um agente acredita serem verdadeiras. Para Levi, os contrafactuais carecem de valor de verdade. Eles seriam uma espécie de crença de segunda ordem, isto é, crenças acerca de como os estados de crenças seriam modificados, caso houvesse a crença no antecedente. Assim, o teste de Ramsey poderia ser formulado assim: $a > b$ é aceitável em K se $b \in K^*_a$. Isso bloquearia a derivação do critério ***-monotonicidade** porque $a > b$ não pertenceria a K.

Um problema dessa abordagem, de acordo com Gärdenfors [Gär86, pág. 90] é que ela não consegue lidar com condicionais iterados, do tipo "se *a*, então (se *b*, então *c*)" ou "se (se *a*, então *b*), então (se *c*, então *d*)". Ele diz (*ibid.*): "Que tipo de crença seria uma proposição do tipo seguinte nessa perspectiva "Se este vaso quebra quando jogado no chão, então ele quebra quando jogado na parede"?.[33] No caso de condicionais iterados do primeiro tipo citado, parece ser possível reduzi-los a condicionais não-iterados do tipo "se *a* e *b*, então *c*". Por exemplo, o condicional iterado "se o carro funcionar normalmente por um tempo, então, se ele apagar, é porque acabou o combustível", poderia ser reduzido a "se o carro funcionar normalmente por um tempo e apagar, então acabou o combustível". No entanto, não parece ser possível encontrar um substituto não iterado razoável para condicionais do segundo tipo.

No próximo capítulo investigaremos a força dos postulados AGM para revisão, isto é, a capacidade dos postulados de delimitar as operações de revisão permissíveis sobre um estado de crenças. Aplicaremos as consequências dessas investigações à lógica para contrafactuais de Gärdenfors, apresentada acima.

Notas

[22]The starting hypothesis of this chapter is that conditional sentences in various forms are about changes of states of belief.

[23]Substituímos (P) por ***-vacuidade**. Gärdenfors diz que "[...] elas não satisfazem (P)", (P) é uma consequência de ***-vacuidade** e será apresentado abaixo.
However, it is questionable whether the belief revisions used in this analysis really are **minimal** since they do not satisfy (P)

[24]First, add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent); finally, consider whether or not the consequent is then true.

[25]Consider a possible world in which A is true, and which otherwise differs minimally from the actual world. "If A, then B"is true (false) just in case B is true (false) in that possible world.

[26]Accept a sentence of the form 'If *A*, then *C*' in a state of belief *K* if and only if the minimal change of *K* needed to accept *A* also requires accepting *C*.

[27]A formula A in $\mathcal{L}'$ is **satisfiable in a belief revision system** $< \mathcal{K}, * >$ if there is some $K \in \mathcal{K}$ such that $K \neq K^{\perp}$ and $A \in K$. A formula A is **valid in a system** $< \mathcal{K}, * >$ iff $\neg A$ is not satisfiable in the system. A formula A is (logically) **valid** iff A is valid in all belief revision systems.

[28]A formula A is a theorem in **VC** iff it is derivable from **CM** together with the axiom schemata A_4–A_{10}.

[29]My present position is that **RT** is the culprit. One way to argue for this is, I believe, to consider (M) and its consequences for belief revision systems.

[30][...] if φ is consistent with K, then there really is no such thing, intuitively, as the revision of K with respect to φ. For nothing in K would need to be given up upon undertaking a commitment to φ.

[31][...] in [Ten6b] a theorem is proved to the effect that there are revisions functions satisfying both the basic AGM-postulates for revision (properly re-formulated) and the Principle of Monotonicity for Revisions (properly

restricted to the case where $K, \varphi \vdash \perp$).

[32]"The theorem and its corollary show that the Ramsey Test **RT** and the preservation condition (P) [or, equivalently, (*4)] cannot both be rational criteria for belief revisions"[Foi feita uma adaptação ao presente texto da nomenclatura dos critérios **RT** e (P).]

[33]What kind of belief would a proposition like the following be on this view? "If this vase breaks if dropped on the floor, then it breaks if thrown against the wall."

Capítulo 3

Da Força dos Postulados do Modelo AGM

Neste capítulo apresentaremos um resultado de Tennant, acerca da força dos postulados do modelo AGM, isto é, da capacidade de separação entre operações de revisão que intuitivamente são consideradas boas daquelas que são reconhecidas intuitivamente como ruins. Serão apresentadas algumas réplicas dos teóricos do modelo AGM sobre a questão. Analisaremos as consequências desse resultado para a análise dos condicionais contrafactuais em termos do modelo AGM e mostraremos que a lógica de contrafactuais de Gärdenfors admite condicionais intuitivamente falsos. Investigaremos as "causas" do resultado de Tennant. Por fim apresentaremos uma abordagem de contrafactuais baseada em um sistema de revisão de crenças mais restrito que o modelo AGM, denominado "revisão por bases de crenças". Nesse sistema não é demonstrável o resultado de Tennant. Mostraremos, não obstante, que essa última abordagem também admite contrafactuais intuitivamente falsos.

3.1 Introdução

Alchourrón e Makinson publicaram [AM82] dois resultados de degeneração das funções de contração e revisão. O primeiro resultado mostra que a revisão *maxichoice*[i] ($*^{mc}$), quando aplicada a um conjunto K e a uma sentença a, resulta em um conjunto completo[ii] se $\neg a \in K$. O segundo resultado mostra que a revisão de intersecção total[iii] ($*^{fm}$) aplicada a um conjunto K e a uma sentença a, é igual a $Cn(a)$, isto é, todas as sentenças que eventualmente estejam em K são removidas para que a seja incluída.

Os teóricos do modelo AGM passaram a investigar as propriedades de interseções de subconjuntos de $K\perp a$, já que nem a escolha de um elemento de $K\perp a$ (*maxichoice*), nem a escolha de todos elementos de $K\perp a$ (intersecção total), deu resultados favoráveis. Se fosse possível definir alguma operação de mudança razoável, ela teria que ser definida "no meio-termo" da função *maxichoice* e da intersecção total.

---

[i] $K_a^{*mc} = (K_{\neg a}^{-mc})_a^+$, em que $K_a^{-mc} = \gamma(K\perp a)$, tal que $\gamma(K\perp a) = X \in K\perp a$.

[ii] Por "completo" significamos que para toda fórmula p da linguagem, ou p ou $\neg p \in K_a^{*mc}$.

[iii] $K_a^{*fm} = (K_{\neg a}^{-fm})_a^+$, em que $K_a^{-fm} = \cap\gamma(K\perp a)$, tal que $\gamma(K\perp a) = K\perp a$.

A partir dos resultados referentes às funções de revisão e contração *maxichoice* e de intersecção total, Alchourrón, Gärdenfors e Makinson começaram a desenvolver as chamadas funções de intersecção parcial (seção 1.6.1 acima). As funções de interseção parcial de contração (as funções de revisão são definidas por meio das funções de contração) pressupõem uma função de seleção γ que toma uma classe dos "melhores elementos" de um $K\perp a$. O resultado da contração é então a intersecção destes "melhores elementos" [AGM85, pág. 511].

As investigações acerca das funções de intersecção parcial culminam no teorema de representação (seção 1.6.1 acima) [Gär88, pág. 82]:

> Seja K um conjunto de crenças qualquer e seja $(^-)$ uma função de contração definida sobre K. Então $(^-)$ é uma função de intersecção parcial transitivamente relacional **se e somente se** $(^-)$ satisfaz **–fecho, –inclusão, –vacuidade, –sucesso, –recuperação, –extensionalidade, –suplementar$_1$ e –suplementar$_2$**.[iv] [34]

Gärdenfors diz (*ibid.*) que esse teorema dá uma forte razão para acreditar que tais funções de seleção sobre $K\perp a$ representariam o processo intuitivo de contração. Diz ele:

> Pelo fato da coleção **–fecho, –inclusão, –vacuidade, –sucesso, –recuperação, –extensionalidade, –suplementar$_1$ e –suplementar$_2$** ter sido motivada independentemente no capítulo 3, este teorema nos dá uma forte razão para focar em funções de intersecção parcial transitivamente relacionais como uma representação ideal do processo intuitivo de contração.[v] [35]

Tennant atribui (na seção 3.3 adiante, essa atribuição será questionada) aos teóricos do modelo AGM a crença de que uma função intermediária entre *maxichoice* e intersecção total representaria o processo de revisão intuitivo, e sustenta que tal crença não estaria nem um pouco justificada. Diz ele [Ten6a, pág. 6]:

> A ideia parecia ser, portanto, que ao evitar os dois extremos e ficar no meio-termo, os postulados agora permitiriam apenas revisões racionalmente admissíveis [...] parecia ser um artigo de fé que revisões de intersecção parcial de K com respeito a a, que fossem construídas nesse espaço intermediário por meio de subconjuntos próprios não-vazios de $(K\perp\neg a)$, evitariam os dois extremos e seus respectivos tipos de degeneração. [36]

Tennant mostra que mesmo no caso de seleções de subconjuntos próprios não unitários de um $K\perp a$ é possível construir infinitas funções de revisão e contração que respeitam os postulados do modelo AGM e mesmo assim são funções degeneradas tais como *maxichoice* e intersecção total, o que indica uma fraqueza bem grande nos postulados para contração e revisão do modelo AGM. Serão apresentados agora o resultados de Tennant e, posteriormente, discutiremos acerca da força dos postulados AGM, da chamada "degeneração do modelo AGM", das "raízes" de tal degeneração e das consequências para a análise de Gärdenfors para os contrafactuais.

[iv] A notação de Gärdenfors foi adaptada ao nosso texto.
[v] A notação de Gärdenfors foi adaptada ao nosso texto.

3.2 O Teorema de Degeneração de Tennant

Tennant utiliza o artigo [AGM85] como referência para os postulados do modelo AGM. Neste artigo alguns postulados para operação de revisão são diferentes dos apresentados acima, embora sejam equivalentes a aqueles. Listamos todos os postulados, por conveniência:

(***-fecho**)	K_a^* é um conjunto de crenças;
(***-sucesso**)	$a \in K_a^*$;
(***-vacuidade**$_1$)	$K, a \nvdash \bot \;\Rightarrow\; K_a^* = K_a^+$;
(***-consistência**$_1$)	$\nvdash \neg a \;\Rightarrow\; K_a^* \nvdash \bot$;
(***-extensionalidade**)	$a \leftrightarrow b \;\Rightarrow\; K_a^* = K_b^*$;
(***-ident-harper**)	$K_a^* \cap K \;=\; K_{\neg a}^-$;
(***-suplementar**$_1$)	$K_{a\wedge b}^* \subseteq (K_a^*)_b^+$;
(***-suplementar**$_2$)	$\neg b \notin K_a^* \;\Rightarrow\; (K_a^*)_b^+ \subseteq K_{a\wedge b}^*$.

Para obter os resultados de degeneração, Tennant define o operador $(\cdot)$ da seguinte maneira [Ten6a, pág. 11]. Abreviamos $Cn(\{a\} \cup \Gamma)$ por $Cn(a,\Gamma)$, em que a é uma sentença e Γ um conjunto de sentenças. Para qualquer a e Γ

$$a \cdot \Gamma =_{df} \left\{ \begin{array}{l} a,\Gamma \nvdash \bot \;\Rightarrow\; a\cdot\Gamma = Cn(a,\Gamma) \\ a,\Gamma \vdash \bot \;\Rightarrow\; a\cdot\Gamma = Cn(a) \end{array} \right\}.$$

Ou seja, se $\{a\} \cup \Gamma$ é um conjunto inconsistente, $a \cdot \Gamma$ é o fecho lógico de a, se não, $a \cdot \Gamma$ é igual ao fecho lógico de $\{a\} \cup \Gamma$. A partir da definição do operador $(\cdot)$, Tennant lista seis observações derivadas de sua definição:

(**Observação 1**)	$a \cdot \Gamma$ é uma teoria;
(**Observação 2**)	$a \cdot \Gamma \vdash a$;
(**Observação 3**)	$a \cdot \; = Cn(a)$;
(**Observação 4**)	$a \nvdash \bot \Rightarrow a \cdot \Gamma \nvdash \bot$;
(**Observação 5**)	$a \leftrightarrow b \Rightarrow a \cdot \Gamma = b \cdot \Gamma$;
(**Observação 6**)	$K \cap (a \cdot \Gamma), \neg a \vdash K$.

Posteriormente são utilizados os seguintes lemas (ver prova em (*ibid.*)):

(**L1**)	$a \wedge b \cdot \Gamma \;\subseteq\; Cn(a\cdot\Gamma, b)$,
(**L2**)	Se $a \cdot \Gamma$, $b \nvdash \bot$, então $a \wedge b \cdot \Gamma \;=\; Cn(a\cdot\Gamma, b)$.

Finalmente é possível provar o resultado central do artigo (*ibid.*, pág. 14):

Teorema 3.2.1. (Teorema de Degeneração) *Seja* Γ *um conjunto consistente de sentenças. Para qualquer teoria consistente* K *e qualquer sentença contingente* a*, tal que* $K \vdash \neg a$*, seja*

$$K_a^* =_{df} a \cdot \Gamma.$$

Então (*) é uma função de revisão que satisfaz os postulados-AGM para a revisão.

Demonstração. Pelas Observações 1 e 2, ***-fecho** e ***-sucesso** são satisfeitos, respectivamente. Visto que só está sendo levado em consideração o caso principal da operação de revisão, isto é, quando $K \nvdash \bot$, $a \nvdash \bot$ e $K, a \vdash \bot$, ***-vacuidade** se torna irrelevante.
Pela Observação 4, ***-consistência** é satisfeito, visto que K_a^* é consistente.
Pela Observação 5, ***-extensionalidade** é satisfeito: se $a \leftrightarrow b$, então $K_a^* = K_b^*$.
Se $K_{\neg a}^- =_{def} K \cap (a \cdot \Gamma)$, então, pela Observação 6, $K_{\neg a}^-$ satisfaz **–recuperação**. Que $K_{\neg a}^-$ definida desta maneira satisfaz os demais postulados para contração é facilmente verificável.
Pelos lemas **L1** e **L2**, os postulados ***-suplementar**$_1$ e ***-suplementar**$_2$ são satisfeitos, respectivamente. □

Gärdenfors [Gär88, pág.58] afirma que "Os postulados que foram introduzidos são motivados pela interpretação de K_a^* como a mudança mínima de K necessária para aceitar a. As consequências dos postulados que foram estabelecidas também suportam essa interpretação."[37] Uma das consequências razoáveis à qual Gärdenfors se refere (*ibid.*) é a seguinte. Se $b \in K_a^*$, então a mudança mínima de K para adicionar b ou é menor ou igual à mudança mínima em K para adicionar a. Do mesmo modo, se $a \in K_b^*$, a mudança mínima de K para adicionar a ou é menor ou igual à mudança mínima de K para adicionar b. Pela assunção de que há uma única mudança mínima de K para adicionar a, e uma única mudança mínima para adicionar b, segue-se que K_a^* e K_b^* devem ser idênticos.

Teorema 3.2.2. $K_a^* = K_b^* \iff b \in K_a^* \ \& \ a \in K_b^*$

Demonstração. A implicação da esquerda para direita segue-se diretamente do postulado ***-sucesso**. Para prová-la da direita para esquerda, suponha que $b \in K_a^*$ & $a \in K_b^*$. Por ***-consistência**, segue-se que $\neg b \notin K_a^*$ e, por ***-suplementar**$_2$, segue-se que $(K_a^*)_b^+ \subseteq K_{a \wedge b}^*$. Logo, por ***-suplementar**$_1$, $(K_a^*)_b^+ = K_{a \wedge b}^*$. De forma similar obtemos que $(K_b^*)_a^+ = K_{a \wedge b}^*$. Visto que $b \in K_a^*$, $K_a^* = (K_a^*)_b^+$. Do mesmo modo, $K_b^* = (K_b^*)_a^+$. Portanto, $K_a^* = K_b^*$. □

Apesar do resultado acima, que mostra que os postulados para revisão capturam um aspecto interessante da noção de mudança mínima, eles ainda deixam a desejar no que diz respeito a uma caracterização mais definida dessa noção, visto que, no teorema 3.2.1, Γ é um conjunto consistente qualquer de sentenças, e pode não ter absolutamente nada em comum com o conjunto de crenças inicial K. Isso significa que a operação de revisão (*), tal como definida no teorema 3.2.1, permite mutilações e inflações arbitrárias de K. A esperança de que as funções que satisfizessem os postulados de revisão acima representariam uma idealização do processo intuitivo de revisão não parece mais ter lugar, pois os postulados são demasiado fracos para captar a noção de mudança mínima.

Da mesma forma, a noção de contração de intersecção parcial também acaba por ser pouco restritiva. Vejamos um exemplo [Osherson, *apud* [Ten6a, pág. 19]]. Seja K um conjunto que refuta a. Seja J qualquer teoria que implica a. Por $a \to J$, representaremos $\{a \to b \,|J \vdash b\}$. Dado que $K \vdash \neg a$, então $a \to J \subseteq K$. Seja Γ o conjunto dos $X \in K \perp a$, tal que $a \to J \subseteq X$. Visto que Γ é um subconjunto de $K \perp a$, $Cn(\cap\Gamma, a)$ é uma revisão por intersecção parcial de K por a e, ademais, $Cn(\cap\Gamma, a) = J$, o que demonstra que $K_a^* = J$. (ver prova detalhada em (*ibid.*, pág. 20)).

A noção intuitiva de que a revisão de um conjunto K pela sentença a deveria efetuar uma mudança mínima em K necessária para acomodar a de maneira consistente não é alcançada pelos postulados-AGM. De acordo com Tennant (*ibid.*, pág. 7], pelo fato de estes terem mostrado-se demasiado fracos, os resultados de Alchourrón, Gärdenfors e Makinson de 1985 só asseguram que

> [...] uma noção de revisão delineada de forma totalmente inadequada definida implicitamente pelos postulados escolhidos é equiparada por uma noção de revisão delineada de forma totalmente inadequada pelas intersecções parciais.[38]

3.3 Os Contra-Argumentos dos Teóricos do Modelo AGM

Gärdenfors, após apresentar os postulados para revisão [Gär88, pág. 58], analisa mais alguns critérios, entre eles o seguinte:

(***-permanência**) $\qquad b \in K \;\Rightarrow\; b \notin K_a^*$ ou $\neg b \in K_a^*$

Em outras palavras, b só não pertence a K_a^* porque K_a^* implica $\neg b$. Esse critério é razoável no sentido de que a remoção de determinada sentença b de K só está justificada se b contradiz as crenças em K_a^*. Todavia, é demonstrável (ver prova em (*ibid.*, pág. 59)) que a adição de ***-permanência** aos critérios para revisão, faz com que para qualquer conjunto K e sentença a, que K_a^* seja um conjunto completo, o que é implausível.

Depois de considerar outros critérios, Gärdenfors afirma: “[...] Eu argumento que os postulados [citados acima] em um certo sentido cobrem as propriedades **lógicas** do processo de revisão, de modo que não são necessários axiomas extras.”[39] Os postulados não caracterizam completamente uma função de revisão, contudo, Gärdenfors sustenta (*ibid.*): “Eu acredito que seria um erro esperar que apenas propriedades lógicas são suficientes para caracterizar um processo de revisão.”[40] Então, o que se desprende é que, de acordo com Gärdenfors, os postulados AGM para revisão (e contração) foram motivados pelo critério de mudança mínima e, embora não identifique somente as funções de revisão (e contração) **razoáveis**, eles circunscrevem o âmbito no qual essas funções **razoáveis** podem ser encontradas. Contudo, Gärdenfors sustenta, para encontrar as funções de revisão e contração plausíveis, já não podem ser usados critérios de natureza lógica.

Um ponto que merece nota é que os postulados adicionais como ***-permanência** (logo acima) ou a “condição de filtragem” (seção 1.7.1) a princípio poderiam ser aplicados a conjuntos que não são fechados sob consequência lógica, sem os resultados contra-intuitivos que têm no modelo AGM. Isso significaria que o problema não é o de que não se pode mais formular critérios de caráter lógico para a teoria de revisão de crenças, mas sim o de que não se pode fazê-lo considerando estados de crenças como conjuntos fechados sob consequência lógica.

Makinson (comunicação particular, 2012) diz que os teóricos do modelo AGM nunca alegaram a impossibilidade de que funções de revisão como aquela definida por Tennant (acima) fossem construídas obedecendo os postulados do modelo AGM. De acordo com ele, nunca houve a pretensão de que os postulados de revisão e contração fossem **suficientes** para que qualquer função que os obedeça fosse considerada "razoável", mas que os postulados deveriam ser tomados como condições **necessárias** que tais funções devem satisfazer, quando definidas sobre conjuntos logicamente fechados e sem referência a nenhuma base a partir da qual eles podem ser derivados.

3.4 Os Condicionais Contrafactuais e o Resultado de Tennant

Tennant mostrou que os postulados AGM para revisão admitem como resultado da revisão de um conjunto K por uma sentença a, conjuntos muito distantes daqueles que seriam esperados intuitivamente. Cabe investigar se tal resultado afeta a abordagem oferecida por Gärdenfors para os condicionais contrafactuais. Mostraremos que a resposta para essa questão é afirmativa.

Os axiomas para a lógica dos contrafactuais de Gärdenfors que apresentamos anteriormente (seção 2.2) são:

(A_1)	Todas tautologias vero-funcionais;
(A_2)	$(a > b) \wedge (a > c) \rightarrow (a > b \wedge c)$;
(A_3)	$a > \top$;
(A_4)	$a > a$;
(A_5)	$(a > b) \rightarrow (a \rightarrow b)$;
(A_6)	$(a \wedge b) \rightarrow (a > b)$;
(A_7)	$(a > \neg a) \rightarrow (b > \neg a)$;
(A_8)	$(a > b) \wedge (b > a) \rightarrow ((a > c) \rightarrow (b > c))$;
(A_9)	$(a > c) \wedge (b > c) \rightarrow (a \vee b > c)$;
(A_{10})	$(a > b) \wedge \neg(a > \neg c) \rightarrow (a \wedge c > b)$.

Os axiomas A_{4-10} foram elaborados a partir do teste de Ramsey

(RT) $\quad a > c \in K$ sse $c \in K^*_a$,

e a partir dos seguintes postulados para revisão de crenças (respectivamente):

(***-fecho**)	K^*_a é um conjunto de crenças para quaisquer K e a;
(***-sucesso**)	$a \in K^*_a$;
(***-inclusão**)	$K^*_a \subseteq K^+_a$;
(***-vacuidade**$_f$)	$a \in K$ & $K \neq K^\perp \;\Rightarrow\; K \subseteq K^*_a$;
(***-consistência**)	$K^*_a = K^\perp \;\Leftrightarrow\; \vdash \neg a$;
(***-extensionalidade**)	$\vdash a \leftrightarrow b \;\Leftrightarrow\; K^*_a = K^*_b$;
(***-suplementar**$_1$)	$K^*_{a\wedge b} \subseteq (K^*_a)^+_b$;
(***-suplementar**$_{2f}$)	$\neg(a > \neg b) \in K \;\Rightarrow\; (K^*_a)^+_b \subseteq K^*_{a\wedge b}$

Devido à incompatibilidade de ***-vacuidade** e ***-suplementar**$_2$ com **RT**, eles foram substituídos por consequências deles, mais fracas. A lista de postulados enfraquecida acima será referida por ***-fecho/*-suplementar**$_{2f}$, enquanto a lista dos postulados AGM padrão para revisão será referida por ***-fecho/*-suplementar**$_2$.

Mostramos em seguida que ***-fecho/*-suplementar**$_2$ implica ***-fecho/*-suplementar**$_{2f}$. Para tanto provamos, a partir da assunção de que K é consistente, que ***-vacuidade**$_f$ segue-se de ***-vacuidade** e que ***-suplementar**$_{2f}$ segue-se de ***-suplementar**$_2$.

Lema 3.4.1. ***(*-vacuidade $\Rightarrow$ *-vacuidade$_f$)***
$(\neg a \notin K \;\Rightarrow\; K^+_a \subseteq K^*_a) \Rightarrow (a \in K \;\&\; K \neq K^\perp \;\Rightarrow\; K \subseteq K^*_a)$

Demonstração. Suponha que $a \in K$ e $K \neq K^\perp$. Logo $\neg a \notin K$ e, (por ***-vacuidade**), $K^+_a \subseteq K^*_a$. Dado que $a \in K$, segue-se que (por $^+$**-vacuidade**), que $K = K^+_a$. Portanto, $K \subseteq K^*_a$. □

Lema 3.4.2. ***(*-suplementar$_2$ $\Rightarrow$ *-suplementar$_{2f}$)***
$(\neg b \notin K^*_a \;\Rightarrow\; (K^*_a)^+_b \subseteq K^*_{a\wedge b}) \Rightarrow (\neg(a > \neg b) \in K \;\Rightarrow\; (K^*_a)^+_b \subseteq K^*_{a\wedge b})$

Demonstração. Suponha que $\neg(a > \neg b) \in K$ e $K \neq K^\perp$. Logo $(a > \neg b) \notin K$ e, (por **RT**) $\neg b \notin K^*_a$. Consequentemente, (por ***-suplementar**$_2$), $(K^*_a)^+_b \subseteq K^*_{a\wedge b}$. □

Lema 3.4.3. ****-fecho/*-suplementar$_2$ $\Rightarrow$ *-fecho/*-suplementar$_{2f}$.***

Demonstração. Diretamente a partir dos lemas 3.4.1 e 3.4.2. □

Como vimos acima (seção 2.2), Gärdenfors prova (demonstrações em [Gär88, pág. 149]) que os axiomas (A_1)–(A_{10}) são válidos em um sistema de revisão de crenças $(\mathcal{K},{}^*)$, em que $\mathcal{K}$ é um conjunto de conjuntos de crenças fechado sob expansão,[vi] se e somente se o sistema $(\mathcal{K},{}^*)$ satisfaz os postulados ***-fecho/*-suplementar**$_{2f}$. Em seguida, aplicaremos um argumento de Quine contra esta abordagem de contrafacfuais de Gärdenfors.

Quine propôs [Qui82] um par de contrafactuais com o propósito de argumentar contra a possibilidade de formulação de uma teoria de contrafactuais razoável. O argumento em linhas gerais é que, embora cada um dos contrafactuais seja falso, eles são consequência de inferências R_1 e R_2 que são corretas do ponto de vista usual, isto é, que preservam a verdade e têm

[vi] Isto é, se $K \in \mathcal{K}$, então $K^+_a \in \mathcal{K}$, para todo a

premissas verdadeiras. As inferências são as seguintes:

$$\frac{\text{Bizet e Verdi são compatriotas, } \Gamma}{\text{Bizet é italiano}}\,{\scriptstyle (R_1)} \quad \text{e} \quad \frac{\text{Bizet e Verdi são compatriotas, } \Theta}{\text{Verdi é francês}}\,{\scriptstyle (R_2)}\ .$$

A inferência R_1, a partir de

(c) "Bizet e Verdi são compatriotas"

e das premissas Γ, que incluem

(v) "Verdi é italiano",

conclui

(b') "Bizet é italiano"

e sustenta o contrafactual

($c > b'$) "se Bizet e Verdi fossem compatriotas, Bizet seria italiano".

A inferência R_2, a partir de c, e das premissas Θ, que incluem

(b) "Bizet é francês",

conclui

(v') "Verdi é francês"

e sustenta o contrafactual

($c > v'$) "se Bizet e Verdi fossem compatriotas, Verdi seria francês".

Ambos $c > b'$ e $c > v'$ são claramente falsos, portanto, uma teoria de contrafactuais que os valida não pode ser uma boa teoria. Mostramos em seguida que isto ocorre com a teoria de contrafactuais de Gärdenfors.

Teorema 3.4.4. *A abordagem de Gärdenfors para os contrafactuais valida ambos contrafactuais ($c > b'$) "se Bizet e Verdi fossem compatriotas, Bizet seria italiano" e ($c > v'$) "se Bizet e Verdi fossem compatriotas, Verdi seria francês"*

Demonstração. Seja K um conjunto de crenças consistente que contenha b, v e $(\neg c)$ "Bizet e Verdi não são compatriotas". O que se pretende é revisar o conjunto K por c. Seja que $J = Cn(b', v', c)$. Dado que J não contém nenhuma informação no sentido em que se Bizet é italiano e Verdi é francês, eles não são compatriotas, J é consistente. De acordo com o teorema 3.2.1, existe uma função (*) que satisfaz ***-fecho/*-suplementar**$_2$ e, pelo lema 3.4.3, que satisfaz ***-fecho/*-suplementar**$_{2f}$, tal que $K_c^* = J$. Desse modo, $b' \in K_c^*$ e $v' \in K_c^*$, o que, por **RT**, implica que $c > b' \in K$ e $c > v' \in K$. □

3.5 As “Causas” do Teorema de Tennant

Como mostraram o resultados de Tennant apresentados na seção anterior, os postulados para as mudanças de crenças do modelo AGM são demasiado amplos. Segundo Tennant [Ten6b, pág. 522], os postulados regem muito pouco as mudanças sobre estados de crenças. Um exemplo claro é o postulado ***-inclusão** ($K_a^* \subseteq K_a^+$). Ora, no caso em que realmente deve haver revisão, ou seja, quando a é incompatível com K, K_a^+ é igual a toda a linguagem $\mathcal{L}$ na qual os operadores estão definidos (pelo princípio de *ex contraditione quodlibet*[vii]), então o que o postulado diz para o caso principal, onde deve haver revisão de K, que K_a^* é um conjunto de $\mathcal{L}$! Ou seja, ***-inclusão** é praticamente inócuo.

Usando funções de contração que satisfazem os oito postulados do modelo AGM (dentre eles, merece ênfase, **–recuperação**) é possível construir funções de revisão através da identidade de Levi que resultem conjuntos tão alheios ao conjunto original quanto se queira, satisfazendo unicamente a restrição de que contenham a sentença pela qual o conjunto inicial foi revisado.

Makinson havia dito [Mak87, pág. 388] que, ao menos quando se trata da operação de revisão, “o postulado de recuperação é inofensivo”[41]. Tennant redargúi [Ten6b, pág. 251] “[...] à luz dos resultados anteriores, se poderia dizer que o postulado de recuperação é inútil [...]”[42]. Tennant analisa outro critério que aparentemente cumpriria a função que os postulados AGM para contração e revisão deixam de cumprir, a saber, de assegurar que as operações de mudança de crenças efetuem mudanças mínimas. O critério é o seguinte

> (***-conservatividade**) Tudo o que se segue de K_a^* deve ser consequência das sentenças de K que permanecem em K_a^*, mais a. Em outros termos, que se $K_a^* \vdash p$, então $K_a^* \cap K, a \vdash p$.

O critério ***-conservatividade**, a princípio, não permitiria que sentenças completamente alheias a K e a pudessem estar presentes em K_a^*. Tennant mostra, não obstante, que esse critério já é consequência dos postulados básicos para revisão. Exporemos o seu argumento abaixo.

Tennant propõe uma abordagem relacional das operações de mudança de crenças do modelo AGM que, ao contrário destas, não se busca definir **a** contração de K por b, nem **a** revisão de K por b. O que se pretende definir nessa abordagem é quando J pode ser considerada como **uma** contração de K por b; e quando J pode ser considerada como **uma** revisão de K por b. Outra diferença importante entre a abordagem de Tennant e o modelo AGM (mencionada acima, seção 2.4.1) é que ele adota o que chama de “análise do caso principal”, em que a operação de expansão de K por b, de contração de K por b e de revisão de K por b só estão definidas quando $b \notin K$, $b \in K$ e $\neg b \in K$, respectivamente.

A prova de que ***-conservatividade** é consequência dos postulados da teoria relacional de revisão de crenças de Tennant é simples: suponha que J é uma revisão de K por a e que $J \vdash p$.

[vii] De $a \wedge \neg a$ segue-se qualquer b

Visto que J é uma revisão de K por a, segue-se que $K \vdash \neg a$ e que $K \vdash a \rightarrow p$. Se $J \vdash p$, então $J \vdash a \rightarrow p$; logo $a \rightarrow p \in J \cap K$, o que implica que $J \cap K, a \vdash p$. Portanto, se J é uma revisão de K por a e $J \vdash p$, então, $J \cap K, a \vdash p$. A mesma prova pode ser levada a cabo no modelo AGM se levado em conta que só se obtém K_a^* quando $\neg a \in K$.

Tennant mostra [Ten6b, pág. 514] que o teorema de degeneração (teorema 3.2.1 acima) também vale para a sua reformulação relacional dos postulados AGM. O que se desprende é que a adição de ***-conservatividade** não alteraria em nada o teorema de degeneração, visto que ele é obtido através dos postulados iniciais e ***-conservatividade** é consequência destes. Ocorre que intuitivamente ***-conservatividade** parece captar o requerimento bastante razoável de que tudo que seja derivável da revisão de um conjunto K por a seja consequência de algumas sentenças de K mais a. Por que então mesmo com a presença de ***-conservatividade** ainda é derivável o teorema de degeneração? Qual é a sua "causa"?

Possível solução: das relações de justificação nos estados de crenças

Tennant oferece (*ibid.*, pág. 524) uma explicação razoável da "causa" do teorema de degeneração, que é como se segue. Seja J um conjunto consistente qualquer que seja uma revisão de K por q: para tanto, basta que $q \in J$. Seja p uma proposição qualquer, tal que $p \in J$ e $p \notin K$. Visto que as revisões só estão definidas para $K \nvdash \perp$, $q \nvdash \perp$ e $K, q \vdash \perp$, então $\neg q \in K$ e, consequentemente, $q \rightarrow p \in K$. Mas como $p \in J$, então $q \rightarrow p \in J$. Assim, $q \rightarrow p \in K \cap J$. Logo, $K \cap J, q \vdash p$.

A causa do problema é que $q \rightarrow p$ está em K pelo fato de que $\neg q$ está em K, a justificativa de $q \rightarrow p$ em K é $\neg q$. Por outro lado, $q \rightarrow p$ está em J como consequência de p, assim, p é a justificativa de $q \rightarrow p$ em J. Portanto, embora $q \rightarrow p$ esteja em $K \cap J$, a justificativa que $q \rightarrow p$ tem em K é completamente diferente da justificativa que $q \rightarrow p$ tem em J, e isso infringe o senso comum que as crenças de K que permanecem em J deveriam ter pelo menos alguma das justificativas suficientes que tinham em K.

O que deve ser mudado no modelo AGM parece não se resumir à adição ou remoção de postulados. É necessária uma nova abordagem quanto à representação dos estados de crenças, tal que estes tenham uma estrutura suficientemente rica para poder representar a relação de justificação entre as crenças. Somente assim poderíamos dizer que uma crença é comum a dois estados de crenças diferentes: quando possui pelo menos alguma justificativa suficiente comum a ambos estados.

Seguramente, em uma abordagem dos estados de crenças que apresente justificativas, critérios como ***-conservatividade** seriam muito mais eficazes na regulação da operação de revisão. Adiante apresentaremos uma teoria de revisão de crenças que dá um passo nessa direção, denominada "dinâmica de bases de crenças". Nessa teoria os estados de crenças são representados por um conjunto finito de sentenças –a base de crenças– a partir das quais as suas consequências são justificadas.

3.6 A Dinâmica de Bases de Crenças e os Contrafactuais

O teorema de degeneração não vale para estados de crenças não fechados sob consequência lógica, aqui denominados **bases de crenças**. A abordagem de estados de crenças como bases de crenças já havia sido proposta em [AM82] mas, devido à grande popularidade do artigo de Alchourrón *et al.* de 1985, que tomava os estados de crenças como teorias, a ideia de Alchourrón e Makinson de 1982 só foi levada a sério alguns anos depois [Han92a].

Hansson defende (*ibid.*) que a representação de estados de crenças por bases de crenças captura melhor as nossas intuições acerca de como deveriam ser efetuadas as mudanças sobre um estado de crenças. As mudanças de crenças deveriam, de acordo com essa abordagem, incidir sobre as crenças atualmente sustentadas por um agente e não sobre as consequências em geral dessas crenças, isto é, sobre uma base B e não sobre $Cn(B)$.

Quando as mudanças de crenças incidem sobre bases de crenças, a operação de revisão ganha uma nuance que não possuía no modelo AGM: há dois tipos de revisão em bases de crenças, a revisão **interna** e **externa**. A operação de revisão interna é aquela que é definível pela identidade de Levi, a saber, $B^*_a = (B^-_{\neg a})^+_a$. Na revisão interna, para adicionar a a B, primeiro assegura-se que B não implica $\neg a$ e depois a é adicionada a B. Neste caso, após o primeiro estágio da revisão, a é **indiferente** em B, isto é, nem a nem $\neg a$ pertence a B. A revisão externa segue o caminho inverso, para revisar B por a, primeiro adiciona-se a e depois $\neg a$ é contraída. Esse tipo de revisão não está disponível no modelo AGM, em que $Cn(B) = B$, visto que, se $\neg a \in B$, então $B \cup \{a\} = B^\perp$, por *ex contraditione quodlibet*. Na revisão externa, pode haver um estágio em que a base de crenças é inconsistente. Hansson defende [Han99, pág. 207] que os dois tipos de revisão são independentes um do outro, mostrando casos em que cabe somente a revisão interna e outros em que cabe unicamente a revisão externa.

Cabe à revisão interna casos em que no processo de revisão da base B por a, aparentemente nem a nem $\neg a$ pertencem a B. Já a revisão externa parece ser mais próxima ao processo intuitivo real quando, diz Hansson (*ibid.*, pág. 205),

> [...] é óbvio que a informação nova deve ser aceita, mas menos óbvio quais crenças prévias ela [a operação de revisão] deveria retirar [...].[43]

Hansson propõe (*ibid.*, pág. 200) alguns postulados para as operações de revisão interna e externa. Os postulados são apresentados abaixo:

(***-sucesso**)	$a \in B^*_a$;
(***-inclusão**)	$B^*_a \subseteq B^+_a$;
(***-consistência**)	$B^*_a = B^\perp \Leftrightarrow\ \vdash \neg a$;
(***-relevância**)	$b \in B \ \&\ b \notin B^*_a \Rightarrow \exists B'(B^*_a \subseteq B' \subseteq B \cup \{b\}, B' \nvdash \perp \ \&\ B', \{b\} \vdash \perp)$;
(***-uniformidade**)	$\forall B' \subset B(B', \{a\} \vdash \perp \Leftrightarrow B', \{b\} \vdash \perp) \Rightarrow B \cap B^*_a = B \cap B^*_b$;
(***-uniformidade$_f$**)	$\forall B' \subset B((a, b \in B) \ \&\ (B', \{a\} \vdash \perp \Leftrightarrow B', \{b\} \vdash \perp)) \Rightarrow B \cap B^*_a = B \cap B^*_b$;
(***-pré-expansão**)	$(B^*_a)^+_a = B^*_a$.

Nos referiremos ao conjunto destes postulados por ***-sucesso/*-pré-expansão**.

Uma operação de intersecção parcial sobre bases de crenças é definida de forma análoga àquela definida para teorias. Isto é, (*) é uma operação de revisão por intersecção parcial sobre B se existe alguma função de seleção γ, definida sobre $\mathcal{P}(B)$, tal que para algum a, $B^*_a = \cap\gamma(B\perp a)$. Hansson obtém os seguintes resultados (*ibid.*, pág. 208)

> O operador (*) é um operador de intersecção parcial **interna** para uma base de crenças B se e somente se ele satisfaz ***-sucesso, *-inclusão, *-consistência, *-relevância, *-uniformidade**;
> O operador (*) é um operador de intersecção parcial **externa** para uma base de crenças B se e somente se ele satisfaz ***-sucesso, *-inclusão, *-consistência, *-relevância, *-uniformidade**$_f$ e ***-pré-expansão**.[viii] [44]

Hansson propõe o seguinte sistema de revisão sobre base de crenças para condicionais [Han92b, pág. 533], em que as revisões incidentes sobre as bases de crenças estão definidas apenas para as sentenças da linguagem proposicional $\mathcal{L}$ (que não contém o conectivo >). Hansson utiliza na definição desse sistema a operação de revisão por intersecção parcial *i.e.* $B^*_a = \cap\gamma(B\perp\neg a) \cup \{a\}$. Diz ele:

> Um **sistema de revisão de crenças** (em uma dada linguagem $\mathcal{L}$ e relativo a uma dada lógica) é um par $< \mathcal{B}, f >$, em que $\mathcal{B}$ é um conjunto de bases de crenças e f é uma função que, para cada $B \in \mathcal{B}$, associa um operador de revisão $(*^b)$ tal que para todas as sentenças indicativas p da linguagem, $B^{*b}_p \in \mathcal{B}$. [ix] [45]

O critério de aceitabilidade para condicionais oferecido por Hansson (*ibid.*, pág. 534) é uma formulação do teste de Ramsey, em que $s(B)$ significa o conjunto das sentenças suportadas/justificadas por B:

> (**RT**$_1$) Seja p e q sentenças indicativas e B um conjunto de sentenças indicativas. Então $p > q \in s(B)$ sse $B^*_p \vdash q$, em que (*) é o operador de revisão único que é associado a B no modelo de revisão de crenças. [x] [46]

Dado que Hansson provou (enunciado acima) que toda operação em bases de crenças é de revisão por intersecção parcial se e somente se satisfaz uma das combinações dos postulados ***-sucesso/*-pré-expansão**. O sistema de Hansson será examinado somente pelo ponto de vista dos postulados, por considerarmos que a abordagem dos postulados tem um teor mais regulativo e intuitivo.

[viii] Ênfase nossa. A notação de Hansson foi adaptada ao presente texto.

[ix] A notação de Hansson foi adaptada ao presente texto.

[x] O critério original de Hansson contém a frase "$p > q \in s(B)$ sse $q \in B^*_p$", no lugar de "$p > q \in s(B)$ sse $B^*_p \vdash q$". É possível perceber que o critério original apresenta uma falha importante, pois requer que $q \in B^*_p$ para que $p > q \in s(B)$. Mas se $B^*_p = \cap\gamma(B\perp\neg p) \cup \{p\}$, q só pertencerá a B^*_p se $q \in B$, visto que B^*_p não é fechado sob consequência lógica e não contém as consequências da adição de p. Assim, o critério (**RT**$_1$) só validaria condicionais semifactuais, isto é, aqueles cujo consequente pertence ao estado de crenças original (ver seção 2.2 acima). Para que seja o critério se aplique a condicionais contrafactuais, o critério deve ser formulado assim: $p > q \in s(B)$ sse $B^*_p \vdash q$. Utilizaremos nas considerações abaixo o critério com essa modificação.

A pesar de os postulados serem mais restritivos e intuitivos que aqueles do modelo AGM, ainda é possível construir uma função de revisão (*) que obedeça os postulados para revisão em base de crenças, tanto externa quando interna, e permita inferir por $\mathbf{RT}_1$ acima, ambos condicionais $c > b'$ e $c > v'$ que já havíamos utilizado para examinar a lógica dos contrafactuais de Gärdenfors e que são ambos intuitivamente insustentáveis. Recordaremos o contra-exemplo.

$(\neg c)$	Bizet e Verdi não são compatriotas,
(b)	Bizet é francês,
(v)	Verdi é italiano,
$(\neg b')$	Bizet não é italiano,
$(\neg v')$	Verdi não é francês.

As seguintes implicações (entre outras) são intuitivamente ou materialmente corretas:

$(b \to \neg b')$	se Bizet é francês, ele não é italiano,
$(v \to \neg v')$	se Verdi é italiano, ele não é francês,
$(b \wedge v \to \neg c)$	se Bizet é francês e Verdi é italiano, eles não são compatriotas,
$(b' \wedge v' \to \neg c)$	se Bizet é italiano e Verdi é francês, eles não são compatriotas,
$(b \wedge c \to v')$	se Bizet é francês e Bizet e Verdi são compatriotas, Verdi é francês,
$(v \wedge c \to b')$	se Verdi é italiano e Bizet e Verdi são compatriotas, Bizet é italiano.

Considere a seguinte base B:

$$B = \{\neg c, b, v,\ b \to \neg b',\ v \to \neg v',\ b \wedge v \to \neg c,\ b' \wedge v' \to \neg c,\ b \wedge c \to v',\ v \wedge c \to b'\}.$$

Seja $\circledast$ a função de revisão para a base B acima definida da seguinte maneira. Seja x uma sentença da linguagem proposicional $\mathcal{L}$ e seja c a sentença "Bizet e Verdi são compatriotas". Então ou $x = c$ ou $x \neq c$.
Quando $x = c$

$$B_x^{\circledast} = \{c, b, v,\ b \wedge c \to v',\ v \wedge c \to b'\}.$$

Quando $x \neq c$, há dois casos:

(i) $B \nvdash \neg x$ ou

(ii) $B \vdash \neg x$.

No caso (i), seja $B_x^{\circledast} = B \cup \{x\}$.
Para o caso (ii), definimos a seguinte operação $\downarrow$: $\mathcal{P}(\mathcal{L}) \times \mathcal{L} \longrightarrow \mathcal{P}(\mathcal{P}(\mathcal{L}))$:

Definição 3.6.1. $B \downarrow \neg x = \{B' \subseteq B \mid B' \vdash \neg x \ \&\ \forall B''((B'' \subseteq B' \ \&\ B'' \vdash \neg x) \Rightarrow B'' = B')\}$.

Em outros termos, $B \downarrow \neg x$ é o conjunto dos subconjuntos mínimos por inclusão de B que implicam $\neg x$.

Seja a operação $e : \mathcal{P}(\mathcal{P}(\mathcal{L})) \longrightarrow \mathcal{P}(\mathcal{L})$:

Definição 3.6.2. $e(X)$ escolhe de maneira uniforme um elemento p, por meio de algum ordenamento, de cada subconjunto Y de X. Uma escolha é **uniforme** no caso em que se p é um elemento escolhido de $Y \in X$ e se $p \in Z$ tal que $Z \in X$, então p é um elemento escolhido de Z.

Intuitivamente, a operação e toma os subconjuntos mínimos de B que implicam $\neg x$ $(B \downarrow \neg x)$ e escolhe uniformemente um elemento y de cada um; assim, quando cada y é removido de B, o conjunto resultante $B - e(B \downarrow \neg x)$ não implica $\neg x$. Por exemplo, seja $J = \{p, q, r, r \rightarrow p, q \rightarrow p\}$. Logo $J \downarrow p = \{\{p\}, \{r, r \rightarrow p\}, \{q, q \rightarrow p\}\}$ e $e(J \downarrow p)$ poderia ser $\{p, r, q\}$.

Agora definimos a revisão de B por x como

Definição 3.6.3. $B_x^{\circledast} = (B - e(B \downarrow \neg x)) \cup \{x\}$.

Lema 3.6.4. $\forall B' \subset B(B', \{a\} \vdash \bot \Leftrightarrow B', \{b\} \vdash \bot) \Rightarrow B \downarrow \neg a = B \downarrow \neg b$.

Demonstração. Suponha que $\forall B' \subset B(B', \{a\} \vdash \bot \Leftrightarrow B', \{b\} \vdash \bot)$ e que $Y \in B \downarrow \neg a$. Logo, $Y \vdash \neg a$ e, pela hipótese do lema, segue-se que $Y \vdash \neg b$. Suponha que $Y \notin B \downarrow \neg b$. Então, existe algum $Y' \subset Y$, tal que $Y' \vdash \neg b$. Mas pela hipótese do lema, $Y' \vdash \neg a$, o que contradiz a hipótese de que $Y \in B \downarrow \neg a$. Portanto, $Y \in B \downarrow \neg b$. A prova da direita para esquerda é análoga. Portanto, $B \downarrow \neg a = B \downarrow \neg b$. □

Lema 3.6.5. $(\forall B' \subset B(B', \{a\} \vdash \bot \Leftrightarrow B', \{b\} \vdash \bot) \& (B \cap B_a^{\circledast} \nvdash \neg b)) \Rightarrow B \cap B_a^{\circledast} \subseteq B_b^{\circledast}$.

Demonstração. Para provar o lema 3.6.5, assuma que $\forall B' \subset B(B', \{a\} \vdash \bot \Leftrightarrow B', \{b\} \vdash \bot)$ e que $(B \cap B_a^{\circledast} \nvdash \neg b)$. Deve ser demonstrado que $\nexists X \subseteq B \cap B_a^{\circledast}(X \subseteq e(B \downarrow \neg b))$. Para tanto, suponha que exista um X tal que $X \subseteq (B \cap B_a^{\circledast})$ e que $X \subseteq e(B \downarrow \neg b)$. Pelo lema 3.6.4, $B \downarrow \neg a = B \downarrow \neg b$. Pela definição do operador e, temos que $e(B \downarrow \neg a) = e(B \downarrow \neg b)$. Portanto, $X \subseteq e(B \downarrow \neg a)$, o que implica que $X \nsubseteq (B \cap ((B - e(B \downarrow \neg a)) \cup \{a\})$, contrariando a hipótese de que $X \subseteq (B \cap B_a^{\circledast})$. Logo, $\nexists X \subseteq B \cap B_a^{\circledast}(X \subseteq e(B \downarrow \neg b))$. □

Teorema 3.6.6. *A operação* $(\circledast)$ *satisfaz* ***-sucesso, *-inclusão, *-consistência, *-relevância, *-uniformidade*** $_f$ *e* ****-pré-expansão***.

Demonstração. $x \in B_x^{\circledast}$ e $B_x^{\circledast} \subseteq B \cup \{x\}$, portanto, ***-sucesso** e ***-inclusão** são satisfeitos.

Quanto a ***-consistência**, B é consistente, suponha que $B_x^{\circledast}$ é inconsistente. Então $B_x^{\circledast} \vdash \neg x$, e $B - e(B \downarrow \neg x) \cup \{x\} \vdash \neg x$, logo $B - e(B \downarrow \neg x) \vdash \neg x$. Então existe algum Y tal que $Y \subseteq B - e(B \downarrow \neg x)$ e que $Y \vdash \neg x$. Seja Y' um subconjunto mínimo de Y que implica $\neg x$. Então para algum $z \in Y'$, $z \in e(B \downarrow \neg x)$. Então, $Y' \nsubseteq B - e(B \downarrow \neg x)$ e, consequentemente, $Y \nsubseteq B - e(B \downarrow \neg x)$ o que contradiz o que concluímos anteriormente.

Para ***-relevância** seja $y \in B$ e $y \notin B_x^{\circledast}$. Então $y \notin B - e(B \downarrow \neg x)$, o que implica que y é um elemento de um subconjunto mínimo de B que implica $\neg x$ e está presente em $e(B \downarrow \neg x)$. Portanto, $B_x^{\circledast} \cup \{y\} \vdash x \wedge \neg x$.

Para ***-uniformidade**, suponha que para todos B' tais que $B' \subseteq B$, ocorre que $B' \cup \{a\} \vdash \bot \Leftrightarrow B' \cup \{b\} \vdash \bot$. Seja que $x \in B \cap B_a^{\circledast}$, devemos provar que $x \in B \cap B_b^{\circledast}$. Há dois casos: (i) $B \nvdash \neg a$ e (ii)

$B \vdash \neg a$. Caso (i), visto que $B \nvdash \neg a$, então $B^{\circledast}_a = B \cup \{a\}$ e $B \cap B^{\circledast}_a = B$. Assim, $\nexists B' \subseteq B(B' \cup \{a\} \vdash \bot)$. Pela hipótese, $\nexists B' \subseteq B(B' \cup \{b\} \vdash \bot)$. Logo, $B \nvdash \neg b$ e $B \cap B^{\circledast}_b = B$. Visto que $x \in B$, $x \in B \cap B^{\circledast}_b$. Caso (ii): $B \vdash \neg a$. Suponha que $x \in B \cap B^{\circledast}_a$, devemos provar que $x \in B \cap B^{\circledast}_b$. Para tanto, suponha que $x \notin B \cap B^{\circledast}_b$. Então, $x \notin B^{\circledast}_b$. Suponha que $B \cap B^{\circledast}_a \cup \{b\} \vdash \bot$, então pela hipótese de ***-uniformidade**, $B \cap B^{\circledast}_a \cup \{a\} \vdash \bot$, logo $B \cap B^{\circledast}_a \vdash \neg a$ e $B^{\circledast}_a \vdash \neg a$, o que é absurdo pois ($\circledast$) satisfaz ***-sucesso** e ***-consistência**. Portanto, $B \cap B^{\circledast}_a \cup \{b\} \nvdash \bot$. Disso se segue que $B \cap B^{\circledast}_a \nvdash \neg b$ e, pelo lema 3.6.5 segue-se que $B \cap B^{\circledast}_a \subseteq B^{\circledast}_b$. Dado que $x \in B \cap B^{\circledast}_a$, obtemos então que $x \in B^{\circledast}_b$, o que contradiz a hipótese anterior. Portanto, $x \in B \cap B^{\circledast}_b$. A prova de que se $x \in B \cap B^{\circledast}_b$, então $x \in B \cap B^{\circledast}_a$, é análoga.

***-uniformidade**$_f$ é consequência direta de ***-uniformidade**. Que ($\circledast$) satisfaça aquela segue-se diretamente da prova de que ($\circledast$) satisfaz ***-uniformidade** .

Quanto a ***-pré-expansão**, observe que pela definição de ($\circledast$), se $x \in (B^{\circledast}_a)^+_a$, então ou $x \in B$ e x não pertence a $e(B \downarrow \neg a)$ ou $x = a$. Dado que ($\circledast$) obedece ***-sucesso, *-inclusão, *-consistência**, então $x \in B^{\circledast}_a$. Se $x \in B^{\circledast}_a$, dado que ($\circledast$) obedece ***-sucesso, *-inclusão, *-consistência**, $x \in (B^{\circledast}_a)^+_a$.

A função ($\circledast$) quando aplicada a B e c, claramente também obedece os postulados ***-sucesso, *-inclusão, *-consistência**.

No caso de ***-relevância**, tome B' como $B^{\circledast}_c$; para cada um dos $x \in \{\neg c, b \to \neg b', v \to \neg v', b \wedge v \to \neg c, b' \wedge v' \to \neg c\}$, temos que $x \in B$ e $B^{\circledast}_c \cup \{x\} \vdash \bot$.

Que ($\circledast$) para os argumentos B e c satisfaça ***-uniformidade**, segue-se da prova geral para quaisquer bases de crenças X e sentenças x. A satisfação de $B^{\circledast}_c$ de ***-pré-expansão** segue-se imediatamente se verificarmos que $\{c, b, v, b \wedge c \to v', v \wedge c \to b'\} \cup \{c\} = \{c, b, v, b \wedge c \to v', v \wedge c \to b'\}$. □

Corolário 3.6.7. *O sistema de revisão de crenças* $< \mathcal{B}, f >$ *de Hansson valida os contrafactuais implausíveis* ($c > v'$) *"se Bizet e Verdi fossem compatriotas, Verdi seria francês" e* ($c > b'$) *"se Bizet e Verdi fossem compatriotas, Bizet seria italiano".*

Demonstração. Seja a base $B = \{\neg c, b, v, b \to \neg b', v \to \neg v', b \wedge v \to \neg c, b' \wedge v' \to \neg c, b \wedge c \to v', v \wedge c \to b'\}$. Seja ($\circledast$) o operador descrito acima. Pelo teorema 3.6.6, ($\circledast$) satisfaz os postulados ***-sucesso/*-pré-expansão**. Portanto, visto que $B^{\circledast}_c = \{c, b, v, b \wedge c \to v', v \wedge c \to b'\}$, segue-se que $B^{\circledast}_c \vdash v'$ e $B^{\circledast}_c \vdash b'$. Consequentemente, por $\mathbf{RT}_1$, $c > v' \in s(B)$ e $c > b' \in s(B)$. □

A revisão em base de crenças apresenta um grande avanço em relação ao modelo AGM, por exemplo, barrando o teorema de degeneração (teorema 3.2.1). A inclusão do postulado de relevância (***-relevância**) e a eliminação do postulado de fecho (***-fecho**) do modelo AGM são fundamentais na restrição das operações de revisão que obedeçam os postulados. Não obstante, ainda é possível construir funções de revisão que obedeçam os postulados ***-sucesso/*-pré-expansão** e que ainda permitam obter conjuntos revisados que, através de $\mathbf{RT}_1$, justifiquem condicionais contrafactuais que não são plausíveis.

Um possível contra-argumento consistiria em dizer que a função ($\circledast$) construída acima foi "mal intencionada", que o aparato formal não foi pensado para evitar tais funções. Não obstante, se Hansson oferece [Han92b, pág. 534], condições de asseribilidade de contrafactuais em termos de $\mathbf{RT}_1$, e dos postulados ***-sucesso/*-pré-expansão**, ou deveriam ser explicitados os casos em que elas não se aplicam, ou deveria estar claro que as condições de asseribilidade dependem no fim de que alguma função razoável de revisão deva ser escolhida especialmente, para cada base de crenças. Naturalmente, esta última opção faz com que os postulados sejam de pouca ou nenhuma valia para a confecção de um critério de asseribilidade razoável para contrafactuais. Tendo em mãos um aparato formal como o oferecido por Hansson, ao invés de utilizá-lo para guiar nossas inferências contrafactuais, antes utilizaremos as nossas intuições pré-formais sobre contrafactuais para corrigir o aparato formal.

Acreditamos, não obstante, que uma mescla de algumas boas ideias dos pioneiros das investigações sobre contrafactuais, como Goodman, Parry e Rescher, com alguns conceitos mais recentes, provenientes da revisão por base de crenças e do sistema de manutenção de razões, nos permitirá avançar um pouco nestas investigações. No próximo e último capítulo, apresentaremos rapidamente a forma natural e intuitiva da abordagem inicial de Nelson Goodman e as críticas e sugestões de Parry e Rescher. Finalmente, tentaremos aplicar algumas dessas ideias à teoria de revisão de crenças.

Notas

[34]Let K be any belief set, and let $(^{-})$ be a contraction function defined over K. Then $(^{-})$ is a transitively relational partial meet contraction function **if and only if** $(^{-})$ satisfies $(^{-}1$–$(^{-}8)$.

[35]Because the collection $(^{-}1)$–$(^{-8})$ has been indepently motivated in chapter 3, this theorem gives us a strong reason to focus on transitively relational partial meet contraction functions as an ideal representation of the intuitive process of contraction.

[36]The thought seemed to be, therefore, that by avoiding the two extremes and sticking to the middle range, the postulates would now permit only rationally admissible revitions to take place [...] it appeared to be an article of faith that partial meet revisions of K with respect to A that were constructed in this middle range by means of non-empty, proper, subsets of $(K\bot\neg A)$ would avoid the two extremes and their respective kinds of degeneracy

[37]"The postulates that have been introduced are motivated by the interpretation of K_a^* as the minimal change of K needed to accept a. The consequences of the postulates that have been established also support this interpretation".

[38][...] a totally inadequately constrained notion of revision implicitly defined by the chosen postulates is matched by a totally inadequately constrained notion of revision constructed by means of partial meets.

[39][...] I argue that the postulates $(^*1)$–$(^*8)$ in a certain sense cover the **logical** properties of the revision process, so that no further axioms are necessary."

[40]"I believe it would be a mistake to expect that only logical properties are sufficient to characterize a revision process."

[41][...] but also **harmless** as far as revision is concerned.

[42]"[...] in light of the results proved above, [one might add] that RECOVERY is **toothless** [...]

[43][...] it is obvious that the new information must be accepted, but less obvious which previous belief(s) it should push out [...].

[44]The operator * is an operator of internal partial meet revision for a belief base A if and only if it satisfies consistency, inclusion, relevance, success and uniformity; The operator * is an operator external partial meet revision for a belief base A if and only if it satisfies consistency, inclusion, relevance, success and weak uniformity and pre-expansion.

[45]A **belief-revision system** (in a given language and relative to a given logic) is a pair $< \mathcal{B}, f >$, where $\mathcal{B}$ is a set of belief bases and is a function that to each $K \in \mathcal{B}$ assigns a revision operator $*_K$ such that for all indicative sentences p of the language, $K *_k p \in \mathcal{B}$.

[46]Let p and q be indicative sentences and K a set of indicative sentences. Then $p > q \in s(K)$ iff $q \in K * p$, where * is the unique revision operator that is assigned to K in the belief-revision model.

Capítulo 4

Contrafactuais: De Nelson Goodman até as Teorias de Revisão de Crenças

Neste capítulo serão expostas inicialmente as abordagens de contrafactuais dos pioneiros dessas investigações, Goodman e Rescher. Mostraremos os benefícios e algumas falhas dessas abordagens. Por fim, alguns aspectos dessas abordagens serão adaptados para o Sistema de Manutenção de Razões, de Doyle.

4.1 Introdução

A abordagem de Nelson Goodman [Goo47] acerca dos condicionais contrafactuais consiste em propor critérios de admissibilidade para as premissas de um argumento que defende um condicional contrafactual. No artigo de 1947, Goodman paulatinamente vai propondo critérios de admissibilidade e mostrando os problemas da adoção de tais critérios. Em determinado ponto do artigo, Goodman expõe a sua maior dificuldade com relação à sua análise: o problema da co-sustentabilidade (*cotenability*). Esse problema faz com que os critérios de admissibilidade para as premissas, que vinham sendo propostos por ele para a inferência de contrafactuais, não conseguissem discriminar um condicional contra-intuitivo de um razoável. O único critério que Goodman encontrou para discriminar a inferência de um condicional intuitivamente verdadeiro de um condicional falso, o critério de co-sustentabilidade, já envolve o conceito de condicional contrafactual, de modo que a análise de Goodman torna-se circular. Nicholas Rescher [Res64] analisa a abordagem de Goodman para os contrafactuais do ponto de vista mais geral do problema da revisão de crenças e propõe uma solução para o problema da co-sustentabilidade de Goodman. Argumentaremos que a abordagem de Rescher, embora apresente um avanço sobre a abordagem de Goodman, Gärdenfors, Lewis e Hansson, não consegue resolver a questão completamente. Por fim, mostraremos que uma modificação da proposta de Rescher a partir de algumas ideias de Parry [Par57] pode ser encaixada de maneira natural no sistema de manutenção de razões de Doyle [Doy79].

4.2 A Abordagem de Goodman para os Contrafactuais

Goodman em [Goo47] propõe uma abordagem de condicionais contrafactuais que pode ser resumida da seguinte maneira: um contrafactual $a > c$ é verdadeiro se e somente se o seu antecedente a em conjunção com um conjunto relevante S de sentenças verdadeiras, implicam (relativamente às leis[i]), o consequente c. Na busca de estabelecer critérios para caracterizar o conceito de conjunto relevante, Goodman vai mostrando os problemas decorrentes de sucessivos critérios até chegar no seguinte (*ibid.*, pág. 119) (em que a e c estão para antecedente e consequente, respectivamente):

> (**CR**) Nossa regra então diz que um contrafactual é verdadeiro se e somente se [α] existe algum conjunto S de sentenças verdadeiras tal que S seja compatível com c e com $\neg c$, e tal que a & S seja auto-compatível e leve por lei a c; [β] enquanto não existe um conjunto S' compatível com c e $\neg c$, tal que a & S' seja auto-compatível e leve por lei a $\neg c$.[47]

O conjunto S é auto-compatível se e somente se S não é inconsistente relativamente às leis. A cláusula β foi adicionada pelo seguinte motivo. Seja que

($\neg j$) Jones não está na Carolina,
(c) Carolina é igual a Carolina do Norte mais a Carolina do Sul.

Se for tomada como S, as sentenças verdadeiras c e

($\neg n$) Jones não está na Carolina do Norte,

de acordo com a cláusula α é possível asserir o contrafactual

($j > s$) se Jones estivesse na Carolina, ele estaria na Carolina do Sul.

Por outro lado, tomando como S, c e a sentença também verdadeira

($\neg s$) Jones não está na Carolina do Sul,

de acordo com a cláusula α também é possível asserir o contrafactual

($j > n$) se Jones estivesse na Carolina, ele estaria na Carolina do Norte.

Assim, dois condicionais intuitivamente incompatíveis ($j > s$) e ($j > n$), estão de acordo com a cláusula α. Por isso, Goodman propõe a cláusula β de forma a eliminar ambos. A ideia de Goodman seguramente era de que não pode haver dois bons argumentos para contrafactuais incompatíveis. O problema é que a cláusula β ao rejeitar a razoabilidade de condicionais como "Se Jones estivesse na Carolina, estaria na Carolina do Sul", acaba desqualificando condicionais sustentados por bons argumentos [Cha03, pág. 393-394]. É fácil encontrar um situação em que é perfeitamente razoável asserir este condicional:

seja o conjunto S

[i]Tomamos "leis" no sentido amplo do termo.

Sempre que Jones está na Carolina, está hospedado na casa de sua mãe,
A mãe de Jones mora na Carolina do Sul.

A união de j e S forma um conjunto auto-compatível e dele segue-se o consequente (s) "Jones está na Carolina do Sul". A existência de outro conjunto de premissas S' (*i.e.*, $S' = \{c, \neg s\}$) que tomado com a mesma hipótese leva a n, não deveria afetar a razoabilidade dessa inferência.

Não obstante, mesmo não validando alguns contrafactuais que tomamos como razoáveis, se o critério **CR somente** validasse condicionais razoáveis, ele poderia ser aceito como um critério demasiado restrito, mas correto.

Infelizmente, Goodman dá um contra-exemplo mostrando que isso não é o caso. Posteriormente ele afirma [Goo47, pág. 121] que não consegue ver saída senão adicionando um critério que envolva o conceito de contrafactual. O contra-exemplo é o seguinte.

Exemplo 1

Para um dado fósforo **b**, seria razoável afirmar

($r > a$) Se o fósforo **b** tivesse sido riscado, ele teria acendido,

com base nas premissas S que cumprem o critério **CR** acima

(b) o fósforo **b** é bem feito,
(s) o fósforo **b** está seco,
(o) há oxigênio suficiente,

e na lei

(l_1) todo fósforo que é riscado, está seco, é bem feito e está em um ambiente com suficiente oxigênio, acende.

Contudo, o seguinte conjunto S' de premissas também cumpre **CR**:

(b) o fósforo **b** é bem feito,
(o) há oxigênio suficiente,
($\neg a$) o fósforo **b** não acendeu,

e por meio dele e l_1 poderia ser inferido um condicional inaceitável:

($r > \neg s$) Se o fósforo **b** tivesse sido riscado, ele não teria estado seco.

Isso ocorre, porque em S' está contida ($\neg a$) "o fósforo **b** não acendeu". Uma proposição que, nas palavras de Goodman, embora seja compatível com (r) "o fósforo **b** foi riscado", não seria verdadeira se r fosse verdadeira. Assim, $\neg a$ não é **co-sustentável** com r. Para que o condicional $r > \neg s$ pudesse ser eliminado, a única forma seria acrescentar o critério de que S seja co-sustentável com r. Mas, para que seja estabelecida a co-sustentabilidade entre S e r é preciso saber se não é verdade que S seria falsa se r fosse verdadeira. Ou seja, é preciso determinar a verdade de outro contrafactual.

O problema da co-sustentabilidade Alguns comentários surgiram na literatura sobre o problema da co-sustentabilidade, como [Par57] e [Coo57]. Eles argumentam que o problema da co-sustentabilidade surge da ausência de consideração das relações temporais usuais entre antecedente e consequente. Segundo eles, explicitada tal relação temporal, desapareceria o problema. Todavia, Goodman oferece um contra-argumento (exemplo 2, abaixo) em resposta a Cooley, no qual não ocorre nenhuma variação temporal entre o antecedente e o consequente do contrafactual, em que o problema da co-sustentabilidade permanece.

4.3 Rescher: Uma Abordagem Mais Geral da Questão

Rescher adota uma análise mais geral da questão dos contrafactuais subsumindo-a ao problema das hipóteses crença-conflitantes, ou seja, o problema de adicionar a um conjunto de crenças uma crença conflitante com as demais e re-estabelecer a consistência de uma forma racional. Ele diz [Res64, pág. 25]: "Um condicional contrafactual é, efetivamente, nada mais que um condicional que extrai uma consequência a partir de um antecedente que é na verdade uma hipótese crença-conflitante".[48]

Frente a uma hipótese crença-conflitante, Rescher observa que, normalmente, há várias maneiras de re-estabelecer a consistência do conjunto resultante e que usando apenas um aparato formal, aparentemente não é possível efetuar uma escolha razoável dentre elas. A hipótese crença-conflitante é ambígua no sentido em que não contém nela própria as instruções de como devem ser efetuadas as mudanças do conjunto de crenças frente à sua adição [Res61, pág. 186]. Para ilustrar, analisemos a resposta de Goodman à Cooley a partir da perspectiva de Rescher sobre a questão.

Cooley atribui o problema que põe fim a abordagem de Goodman dos contrafactuais –o problema da co-sustentabilidade– à falta de explicitação das relações temporais entre o antecedente e o consequente do condicional [Coo57]. Goodman oferece [Goo57, pág. 532] o seguinte argumento em resposta à Cooley.

Exemplo 2

Seja a hipótese crença-conflitante

(t) A temperatura de **b** era 650ºC no tempo **t**.

E seja S o conjunto de informações sobre a questão:

($\neg t$) A temperatura de **b** não era 650ºC no tempo **t**,
(f) **b** era um parafuso de ferro em **t**,
(p) **b** era preto em **t**,
($\neg v$) **b** não era vermelho em **t**,
($l2$) Todo objeto de ferro, se está a 650ºC em **t**, é vermelho em **t**.

A assunção da hipótese crença-conflitante t faz com que seja possível reestruturar o conjunto S, mantendo algumas informações iniciais e a consistência, de pelo menos três formas:

Alternativa 1	**Alternativa 2**	**Alternativa 3**
remover: $\neg t$, p, $\neg v$	**remover**: $\neg t$, f	**remover**: $\neg t$, $l2$
reter: f, $l2$	**reter**: p, $\neg v$, $l2$	**reter**: f, p, $\neg v$

• Da alternativa 1 e t segue-se v e é asserível o condicional

($t > v$) se a temperatura de **b** tivesse sido 650ºC no tempo **t**, então **b** teria estado vermelho em **t**.

• Da alternativa 2 e t segue-se $\neg f$, e inferimos

($t > \neg f$) se a temperatura de **b** tivesse sido 650ºC no tempo **t**, então **b** não teria sido de ferro em **t**.

• Da alternativa 3 e t segue-se $\neg l2$, e é consequência

($t > \neg l2$) se a temperatura de **b** tivesse sido 650ºC no tempo **t**, então nem todo objeto de ferro aquecido a 650º é vermelho.

Assim, o problema de estabelecer qual dos condicionais $t > v$, $t > \neg f$ e $t > \neg l2$ acima estão justificados é o mesmo que estabelecer qual escolha entre as alternativas 1 a 3 é a correta. Rescher diz [Res64, pág.27]:

> [...] a solução do problema dos condicionais contrafactuais reside em fazer e suportar uma distinção, dentro do grupo de alternativas logicamente elegíveis, entre formas "naturais" e "não-naturais" de efetuar uma reconciliação entre uma hipótese crença-conflitante por um lado, e por outro o conjunto total das crenças residuais que continuam a ser coletivamente inconsistentes com ela. [49]

4.3.1 O Princípio de Retenção

Segundo Rescher, fazer uma escolha unívoca entre as várias alternativas de restruturação das crenças frente a uma hipótese crença-conflitante consiste em estabelecer uma ordem prévia de preferência entre as crenças. Essa ordem seria estipulada por um princípio de retenção de crenças estabelecendo que determinados tipos de crenças são mais facilmente derrogáveis que outros.

Rescher faz uma distinção dos condicionais contrafactuais em dois tipos: os nomológicos e os puramente hipotéticos. Os nomológicos, como sugere o nome, são aqueles cuja inferência envolve algum tipo de lei, por exemplo, os condicionais do exemplo 2 acima. Os puramente hipotéticos são aqueles que não pressupõem uma lei, por exemplo: "se Bizet e Verdi fossem compatriotas, Bizet seria italiano".

Em ambos os tipos de condicionais ocorre a ambiguidade da suposição crença-conflitante, contudo, no caso dos contrafactuais nomológicos, Rescher argumenta que é perceptível que atribuição de razoabilidade ou não razoabilidade a um contrafactual é refletida pela priorização ou não priorização das leis sobre os fatos singulares. Vejamos um exemplo de contrafactual nomológico em que isso ocorre:

Exemplo 3

Seja a hipótese crença-conflitante

(m) O pedaço de manteiga foi aquecido a 100ºC.

E seja S o conjunto de informações sobre a questão:

($\neg m$) O pedaço de manteiga não foi aquecido 100ºC,
($\neg d$) O pedaço de manteiga não derreteu,
($l3$) Toda manteiga aquecida a 100ºC, derrete.

A assunção da hipótese crença-conflitante m faz com que seja possível reestruturar o conjunto S, mantendo um número máximo de informações iniciais e a consistência, de duas formas:

Alternativa 1	Alternativa 2
remover: $\neg m$, $\neg d$	**remover**: $\neg m$, $l3$
reter: $l3$	**reter**: $\neg d$

• Da alternativa 1, e de m, segue-se d e
($m > d$) Se o pedaço de manteiga tivesse sido aquecido a 100ºC, ele teria derretido.

• Da alternativa 2, e de m, segue-se $\neg l3$ e
($m > \neg l3$) Se o pedaço de manteiga tivesse sido aquecido a 100ºC, nem toda manteiga aquecida a 100ºC derrete.

O condicional $m > d$ correspondente à alternativa 1 soa mais “natural” visto tratar-se de um contrafactual nomológico, as leis que tomam parte na sua inferência devem ter prioridade sobre os fatos. Neste caso, deve-se adaptar todo o conjunto de informações disponível em função da lei em questão, e em função da hipótese crença-conflitante que, por princípio, não deve ser derrogada nesse caso. Neste exemplo, aceitar como verdadeiro o condicional contra-intuitivo, ($m > \neg l3$) corresponderia a uma priorização de um fato singular ($\neg d$) em detrimento da lei geral ($l3$).

4.3.2 O Problema da Co-sustentabilidade

Cabe observar que no exemplo 1 acima os dois conjuntos de premissas disponíveis fazem uso da lei l_1 em questão na inferência dos condicionais. E também no exemplo 2 os conjuntos de premissas correspondentes às alternativas 1 e 2 contêm a lei l_2. O que significa que, nesses casos, na reestruturação das informações frente à hipótese crença-conflitante, as leis não foram derrogadas em detrimento de fatos particulares. Desse modo, embora a lei não seja derrogada nos exemplos 1 e 2, subsistem alternativas e, portanto, permanece o problema da escolha. Portanto, a explicação de Rescher da irrazoabilidade do condicional $m > \neg l_3$ acima não é adequada para explicar a irrazoabilidade dos condicionais anteriores $r > \neg s$ (exemplo 1) e $t > \neg f$ (exemplo 2). Ou seja, com o princípio de retenção e rejeição, tal como apresentado na seção anterior, o problema da co-sustentabilidade permanece, pois não temos instrumentos para explicar por que os condicionais $r > a$ e $t > v$ dos exemplos 1 e 2 são razoáveis e por que seus respectivos opostos $r > \neg s$ e $t > \neg f$ não o são.

A proposta de Rescher

Rescher adiciona então mais um critério [Res64, pág. 66]:

> Têm prioridade sobre as outras informações aquelas que garantem a aplicabilidade da lei em questão.

Assim, no caso do contra-exemplo de Goodman (exemplo 1 acima), a informação s (o fósforo estava seco) tem prioridade sobre $\neg a$ (o fósforo não acendeu); bem como, no exemplo 2, f (**b** era um parafuso de ferro em **t**) tem prioridade sobre $\neg v$ (**b** não era vermelho em **t**).

Ao fazer com que as leis tenham preferência de retenção sobre as informações particulares, e que as informações particulares que garantem a aplicabilidade da lei tenham preferência sobre as outras, Rescher parece resolver a questão.

No contra-exemplo de Goodman, respeitando esses critérios de preferência, obtemos a ordem $\neg a \equiv \neg r \sqsubset s \equiv o \equiv b \sqsubset l_1 \sqsubset r$, onde $\equiv$ e $\sqsubset$ significam "igualmente preferível" e "menos preferível", respectivamente. Assim, s tem preferência sobre $\neg a$, e l_1 tem preferência sobre s. O único conjunto S de premissas verdadeiras e compatíveis com a hipótese contrafactual que satisfaz essa ordem de preferência é $\{l_1, o, s, b\}$, aquele que permite a inferência do condicional (mais razoável) $r > a$.

No exemplo 2, obtém-se a ordem $p \equiv \neg v \equiv \neg t \sqsubset f \sqsubset l_2 \sqsubset t$. Da mesma maneira, a única alternativa que fornece um conjunto de premissas obedecendo esse ordenamento preferencial é a alternativa 1, que também é a que permite a inferência do condicional plausível $t > v$.

Goodman: Mais Problemas

Goodman, em correspondência com Rescher, oferece um contra-argumento ao último critério proposto por Rescher [Res61, nota 13, p. 193].

Goodman observa que qualquer lei l_x, formulada como

(l_x) O que satisfaz as condições b_1, b_2, ..., b_n, deve exibir a característica c,

também pode ser reformulada em n versões diferentes, logicamente equivalentes entre si:

(l_x') O que satisfaz as condições $\neg c$, b_2, ..., b_n, deve exibir a característica $\neg b_1$.

Assim, a lei l_2 do exemplo 2, representada esquematicamente de forma simplificada (sem quantificação),

(l_2) $t\&f \Rightarrow v$,

pode ser reformulada em

(l_2') $t\&\neg v \Rightarrow \neg f$.

Se esse argumento procede, o último critério de Rescher, relacionado à priorização das informações que permitem a aplicabilidade da lei, já não funciona. Tomemos o caso do exemplo 2. A alternativa 2 foi excluída pelo último critério de Rescher por reter $\neg v$ em detrimento de f, não priorizando as informações que permitem a aplicabilidade da lei l_2. Mas se l_2 é equivalente a l_2', então agora $\neg v$ é uma informação que permite a aplicabilidade da lei l_2', então ela deve ser priorizada sobre f. Sumarizando, f deve ser priorizada sobre $\neg v$, por um lado, e por outro, $\neg v$ deve ser priorizada sobre f. Portanto, o argumento de Goodman defende que o problema da co-sustentabilidade permanece, mesmo com os critérios de priorização de informações formulados por Rescher.

A resposta de Rescher

Rescher, em resposta à Goodman, sustenta que a inter-deduzibilidade entre as leis l_2 e l_2' não justifica a sua inter-substituição dentro do âmbito da lógica indutiva, ou seja, que nesse âmbito, não só o conteúdo das leis deve ser levado em consideração, mas também sua forma. Isso por que, embora as leis l_2 e l_2' e suas equivalentes sejam rechaçadas por uma mesma instância, elas são confirmadas por instâncias diferentes.

Vejamos o argumento mais pormenorizadamente. Tanto l_2 ($(t \,\&\, f) \Rightarrow v$), quanto l_2' ($(t \,\&\, \neg v) \Rightarrow \neg f$), são rechaçadas quando se descobre uma instância que satisfaz suas condições do antecedente, mas que não exibe a característica do consequente. Assim, se ocorre t e f, mas ocorre $\neg v$, a lei l_2 é refutada. Por esses mesmos fatos, a lei l_2' também é refutada. Mas as sentenças l_2 e l_2' são confirmadas por instâncias muito diferentes. Rescher oferece o seguinte exemplo. Seja a sentença

l: "Tudo que é x, é y",

e sua equivalente dedutiva

l': "Tudo que é não-y, é não-x".

Suponha que de um total de 100 objetos obtém-se a seguinte tabela da divisão dos objetos em x, não-x, y e não-y.

	y	não-y
x	7	0
não-x	13	80

Ou seja, dos 100 objetos, apenas 7 são relevantes para confirmar l apesar de que 80 instancias confirmem l'. Nenhuma instância rejeita l e l'. Nesse exemplo l' é várias vezes mais suportada que l, o que mostra uma grande diferença entre ambas. Se posteriormente forem descobertas mais 10 instâncias de x que são y, então l será confirmada por 17 casos e o *status* de confirmação de l' não será alterado.

Assim, embora as sentenças l_x e suas reformulações l_x' sejam logicamente equivalentes, elas não são intercambiáveis quando se toma l_x como uma lei estabelecida por indução e sujeita a confirmação e rejeição.

Aplicando essas considerações ao caso dos contrafactuais nomológicos, Rescher conclui [Res61, pág. 193, nota 13] que não é admissível utilizar uma sentença geral, dedutivamente equivalente a uma outra que se toma como lei, para refutar um dado contrafactual, nos moldes dos exemplos acima.

Portanto, ao restabelecer a consistência frente a hipótese crença-conflitante do contrafactual em questão, deve-se levar em conta a formulação específica da lei que garantirá a inferência do contrafactual. Assim, se uma lei l tem as condições a, b e c, então deve-se tomar esse fato como uma base para dar prioridade, dentre as informações particulares, para as que dizem respeito a a, b e c. Diz Rescher [Res64, pág. 84]:

> Nós usamos o modo de formulação da sentença da lei (e assim um elemento acima do "conteúdo assertivo" da lei) como um guia para estabelecer a categoria modal relativa [isto é, as relações de preferência de retenção] das sentenças-fato relevantes para o contexto da hipótese contrafactual e necessária para sua análise.[50]

Das leis operantes

O princípio de retenção e rejeição para hipóteses crença-conflitantes relativas a contrafactuais nomológicos proposto finalmente por Rescher é o seguinte. Os tipos de informações estão enumerados de acordo com a prioridade de retenção [Res64, pág. 34-5].

1. A hipótese crença-conflitante,

2. As leis operantes na questão,

3. Condições para a aplicabilidade das leis operantes na questão,

4. Outras questões de fato.

Assim, em caso de um conflito causado pela hipótese crença-conflitante, as primeiras informações a serem descartadas são questões de fato que não são condições de aplicabilidade das leis operantes na questão.

Como foi apresentado na última seção, Rescher não aceita que duas generalizações logicamente equivalentes (do tipo $t\&f \Rightarrow v$ e $t\&\neg v \Rightarrow \neg f$) sejam ambas tomadas como a formulação de uma mesma lei. Assim, conclui Rescher (*ibid.*, pág. 34, n. 1], se um contrafactual C_i foi estabelecido a partir da lei L_i, seguindo o princípio de retenção e rejeição apresentado acima, não é lícito estabelecer um contrafactual oposto a C_i com base em uma sentença dedutivamente equivalente a l_i, mesmo que o princípio de retenção e rejeição também seja seguido.

Suponhamos, relembrando o exemplo 1, que se tome como lei operante na questão a seguinte sentença:

(l_1') Todo fósforo que é riscado, está em um ambiente com suficiente oxigênio, é bem feito e não acende, não está seco.

Obedecendo o princípio de retenção e rejeição acima, e utilizando as premissas do exemplo 1, é justificável o seguinte contrafactual:

($r > \neg s$) se o fósforo tivesse sido riscado, ele não teria estado seco.

Agora, de acordo com as considerações de Rescher, ninguém pode apresentar outro argumento constituído da lei

(l_1) todo fósforo que é riscado, está seco e está em um ambiente com suficiente oxigênio, acende;

e das mesmas premissas, utilizando o mesmo princípio de retenção e rejeição, para justificar o contrafactual oposto:

($r > a$) se o fósforo tivesse sido riscado, ele teria acendido.

Visto que l_1' foi tomada como lei operante na questão, e a sentença l_1, embora dedutivamente equivalente, não representaria a mesma lei.

Essas considerações são suficientes para mostrar que a proposta de Rescher não consegue resolver a questão. Queremos que o condicional $r > a$ seja justificável, enquanto o $r > \neg s$ não seja. Ora, mas se for tomada como lei operante l_1', o condicional justificável é condicional $r > \neg s$ e não $r > a$.

A questão agora se desloca para o caráter de uma lei operante em um âmbito em que ocorre hipóteses crença-conflitantes. Na abordagem de Rescher, a escolha da lei operante não pode ser livre, sob pena de permanecer enredado com o problema da co-sustentabilidade de Goodman. Mas como estabelecer que determinada sentença geral representa efetivamente uma lei, que justifica os condicionais que queremos que sejam justificados, e que todas as sentenças dedutivamente equivalentes a ela não podem ser operantes em nenhuma argumentação que contém hipóteses crença-conflitantes? As respostas que surgem a esta questão não estão nada longe de serem simplesmente respostas *ad hoc*.

4.4 Adotando Ideias de Goodman, Parry e Rescher para Revisão de Crenças

O princípio da retenção de Rescher é uma boa tentativa de resolução do problema da co-sustentabilidade. Proporemos algumas nuances extras a ele e introduziremos algumas noções para a análise do conceito de condição relevante que foram influenciadas pela abordagem dos contrafactuais de Parry e Oswaldo Chateaubriand.

Rescher propõe o ordenamento de preferência hipótese $\sqsupset$ lei operante $\sqsupset$ condições de aplicabilidade da lei $\sqsupset$ outras questões. Mostramos na seção anterior que o problema da co-sustentabilidade transmite-se à escolha da lei operante na questão. Uma questão importante no estudo do conceito de condição relevante para inferência de contrafactuais, apontada primeiramente por Parry e depois enfatizada por Chateaubriand[ii], é a noção de que **as premissas utilizadas nas inferências contrafactuais devem ser independentes da hipótese contrafactual**. Especificamente sobre o exemplo de Jones na Carolina, Parry diz [Par57, pág. 90]:

> Agora, quando começamos com "Se Jones estivesse na Carolina..." nós estamos abstraindo de sua localização geográfica atual, e colocando ele em uma localização presumivelmente não atual, enquanto deixando outras coisas tanto quanto possível da mesma forma. É claro que as "outras coisas" que podem ser incluídas nas condições relevantes, *S*, **não podem incluir nada acerca da sua localização atual** no tempo do antecedente.[51]

Ou seja, a informação que consideramos espúria ($\neg n$) "Jones não está na Carolina do Norte", não é independente da proposição que queremos revisar ($\neg c$) "Jones não está na Carolina". Se, em outra circunstância, asserimos $\neg n$ e alguém nos questiona por que $\neg n$, podemos perfeitamente dizer: $\neg n$ porque $\neg c$. Agora, como diz Parry, se estamos revisando a localização geográfica de Jones (expressada por $\neg c$), para extrair alguma consequência não podemos utilizar como premissa uma informação que **dependa** daquilo que estamos revisando. Neste exemplo, as informações $\neg n$ e $\neg s$ **não podem ser consideradas relevantes** para extrair conclusões acerca de c.

Tomemos o exemplo 1 acima sobre o fósforo. O condicional contra-intuitivo ($r > \neg s$) "se o fósforo tivesse sido riscado, ele não teria estado seco" faz uso das seguintes premissas (o) "há oxigênio suficiente", (b) "o fósforo é bem feito" e ($\neg a$) "o fósforo não acendeu". Suponha que alguém esteja em posse de um fósforo que satisfaça essas premissas e que ($\neg r$) não tenha sido riscado. Se essa pessoa assere o, podemos questionar por que o, e uma possível resposta seria a de que o ambiente está ventilado. Se ela assere b, quando questionada, pode responder que o fósforo passou pelo controle de qualidade. Todavia, ainda que seja uma asserção um tanto estranha e enganadora, se a pessoa diz que $\neg a$, ou algo mais razoável como "o fósforo não está aceso", a justificativa dela vai depender de $\neg r$, por exemplo, que o fósforo nunca saiu da caixa ou mesmo que o fósforo não foi riscado.

[ii] Conferência do VII Simpósio Principia, Florianópolis

Analisemos agora o exemplo sobre o parafuso de ferro (exemplo 2 acima), que parece oferecer certa complicação à análise que vinhamos fazendo. Recordemos as informações disponíveis.

($\neg t$) A temperatura de **b** não era 650ºC no tempo **t**,
(f) **b** era um parafuso de ferro em **t**,
(p) **b** estava preto em **t**,
($\neg v$) **b** não estava vermelho em **t**,
($l2$) Todo objeto de ferro, se está a 650ºC em **t**, está vermelho em **t**.

Suponha que alguém afirme que ($\neg v$) o parafuso **b** não estava vermelho em **t**. Se questionada sobre essa asserção, ela pode responder "$\neg v$ porque (p) **b** estava preto em t". Quando questionada acerca de p, diria "p porque estou vendo". Assim, pareceria que o conhecimento de que p é independente do conhecimento de que $\neg t$, e a discriminação dos conjuntos razoáveis dos não-razoáveis iria água abaixo pela distinção entre independência e dependência da negação da hipótese contrafactual.

Contudo, imaginemos que alguém afirme ($t > \neg f$) "se a temperatura de **b** tivesse sido 650ºC no tempo **t**, então **b** não teria sido de ferro em **t**", e justificasse tal asserção dizendo "porque sei que o parafuso era preto em **t** e que $l2$ é válida". Essa suposta justificativa haveria de ser recusada pois supor que (t) o parafuso foi aquecido, e reconhecer a validade de $l2$, implica rechaçar p. Outro modo de ver a dependência de p com $\neg t$ é o seguinte. Alguém pode justificar o conhecimento acerca da cor de um objeto de ferro sem a necessidade de uma experiência visual. Reconhecendo a validade de $l2$, a forma de fazê-lo é medir a temperatura do objeto. Assim, há pelo menos uma circunstância em que a cor do objeto de ferro depende razoavelmente da sua temperatura. Portanto, é plausível dizer que, embora p acima possa, por assim dizer, depender de si próprio, p também é dependente de $\neg t$ e $l2$. Repare que o mesmo não poderia ser dito acerca da relação entre $\neg t$ e f, pois, como uma regra geral, alterações na temperatura não alteram a qualidade de ser ferro.

Propomos uma modificação no princípio de retenção de Rescher na seguinte enumeração em ordem decrescente de prioridade de retenção:

Princípio de Retenção

1. hipótese ou crença a ser revisada;
2. definições dos termos
3. leis gerais e regras de inferência
4. questões de fato independentes (relativamente às definições, leis gerais e regras de inferência) da hipótese
5. outras questões de fato

Na seção seguinte o princípio de retenção acima será aplicado na teoria de revisão de crenças formulada por Doyle [Doy79] e denominada Sistema de Manutenção das Razões ou RMS.

4.4.1 O Sistema de Manutenção das Razões

As bases de crenças oferecem uma estrutura de justificação precária baseada somente em consequência lógica. A representação dos estados de crenças por bases de crenças não é adequada para captar noções de dependência e independência, requeridas pelo conceito de condição relevante. RMS [Doy79] é um sistema de revisão de crenças bastante rico, que consegue representar vários tipos de relações de justificação entre as crenças como justificações monotônicas e não-monotônicas. Essa riqueza estrutural é particularmente propícia a representar as noções de dependência e independência entre as crenças. Nesta seção apresentaremos o sistema de manutenção das razões, tal como formulado originalmente por Doyle [Doy79]. Posteriormente serão propostas algumas alterações de modo a incorporar o princípio de retenção na operação de revisão do RMS.

No sistema RMS os **estados de crenças** são compostos de dois tipos de estruturas: os **nós** e suas **justificativas**. Um nó representa uma sentença, uma justificativa para um nó um conjunto de sentenças e uma **crença** é representada por um nó devidamente justificado. Um estado de crenças B no sistema RMS é um conjunto de triplas ordenadas do tipo $< n, \mathcal{P}, \mathcal{A} >$, em que n é um nó e $\mathcal{P}$ e $\mathcal{A}$ são as justificativas em B para n. $\mathcal{P}$ está para os nós cuja **presença** no estado de crenças justifica o nó n. $\mathcal{A}$ está para os nós cuja **ausência** no estado de crenças justifica o nó n.

Um nó n juntamente com suas justificativas ($< n, \mathcal{P}, \mathcal{A} >$) será designado como uma **entrada**. Um nó n pode estar justificado em um estado de crenças B por meio de várias entradas, assim como uma sentença pode ser justificada de diversas maneiras. Os nós são de dois tipos, as **assunções** ou **nós fundamentais** e os **nós derivados**. Os nós fundamentais são justificados por si próprios e os nós derivados são justificados a partir dos nós fundamentais. Em um estado de crenças B um nó n tem uma **justificativa válida** quando (i) existe pelo menos uma entrada para n do tipo $< n, \{\}, \{\} >$, caso em que diz-se que n é uma assunção ou nó fundamental; ou quando (ii) existe em B pelo menos uma entrada para n do tipo $< n, \mathcal{P}, \mathcal{A} >$, em que ou existe pelo menos um $a \in \mathcal{P}$ tal que a tem uma justificativa válida, ou todos os nós $b \in \mathcal{A}$ não têm justificativas válidas; nesse caso dizemos que n é um nó derivado. Em outros termos, os nós derivados podem ser justificados validamente em B por meio da presença de outro nó justificado validamente em B e também pelo fato de outro nó não ter uma justificativa válida em B.

Uma **crença** é representada por um nó com uma justificativa válida. Caso a um nó n não esteja associado em B a nenhuma justificativa $\mathcal{P}$ ou $\mathcal{A}$, válida ou não, então $< n, -, - > \in B$. Por exemplo, seja o estado de crenças $B = \{< n, \{m\}, \{o\} >, < m, \{\}, \{\} >, < o, -, - >\}$, em que o nó ($m$) "a pluma tem mais resistência ao atrito do ar" é um nó fundamental; o nó (n) "a pluma cai mais devagar que uma bola de ferro", é justificado pela presença de m e pela ausência de (o) "a pluma e a bola de ferro estão no vácuo"; e o não tem justificativa, portanto não é acreditado.

Quando ao menos uma entrada para n apresente uma justificativa válida em B, dizemos que n está *dentro* (com relação a B), na situação oposta, quando nenhuma entrada para n apresente uma justificativa válida, dizemos que n está *fora* (com relação a B). Assim, um nó derivado m está *dentro* de B quando existe uma entrada para m tal que ao menos um elemento de sua lista $\mathcal{P}$ de presentes está *dentro* ou quando todos nós da sua lista $\mathcal{A}$ de ausentes está *fora*. Um nó n está *fora* em B em três circunstâncias: (i) quando todos os nós de sua lista $\mathcal{P}$ estão *fora*, (ii) quando algum nó de sua lista $\mathcal{A}$ está *dentro* e (iii) quando a única entrada para n em B é nula, *i.e.*, do tipo $< n, -, - >$.

Vejamos um exemplo[iii]:

	Justificação		
Nó	Lista dos Presentes	Lista dos Ausentes	Status
(n_1) Oscar não é culpado de difamação	n_2, n_6	n_3	*dentro*
(n_2) O acusado deve ter o benefício de dúvida	{ }	{ }	*dentro*
(n_3) Oscar chamou a rainha de meretriz	n_4 & n_5	{ }	*fora*
(n_4) Pode ser assumido que o relatório da testemunha é correto	{ }	{ }	*dentro*
(n_5) A testemunha disse que ouviu Oscar chamar a rainha de meretriz	–	–	*fora*
(n_6) A testemunha é uma mentirosa compulsiva	–	–	*fora*

Nesse exemplo, os nós n_2 e n_4 são nós fundamentais e estão *dentro*. Os nós n_5 e n_6 estão *fora* porque não têm justificativas associadas a eles. O nó n_1 está *dentro* porque o n_2 está dentro e n_3 está fora (n_1 poderia estar *dentro* também se n_6 estivesse *dentro* e n_3 continuasse *fora*). O nó n_3 está *fora* porque os nós n_4 e n_5 da sua lista de presentes não estão ambos *dentro*.

Para expandir o estado de crenças B pelo nó n (B_n^+), basta adicionar a B uma entrada $< n, \mathcal{P}, \mathcal{A} >$ que apresente uma justificativa válida para n em B, obtendo $B \cup < n, \mathcal{P}, \mathcal{A} >$. Para contrair n de B (B_n^-), deve-se invalidar todas as justificativas para n em B. Para revisar B por n (B_n^*), deve-se expandir B por n e assegurar que o sistema continue consistente.

Na próxima seção apresentaremos os prospectos de um sistema, denominado RMS-C, resultante de algumas modificações do sistema RMS exposto acima.

4.4.2 O Sistema RMS-C

Propomos, à guisa de elucidação das inferências contrafactuais, algumas modificações estruturais no sistema RMS. Denominaremos o (esboço de) sistema resultante de RMS-C. Está fora do escopo deste texto o desenvolvimento de um sistema de revisão de crenças completo, portanto, não há a presunção de que tal esboço de sistema seja tão bem desenvolvido e estruturado como o sistema RMS de Doyle.

[iii] Exemplo retirado de [Gär88, pág. 34]

Substituiremos a noção de nó fundamental ou assunção do RMS pelas noções mais flexíveis de **nó-fato** e **nó-fato-independente**. Os nós derivados do RMS terão no sistema RMS-C outro significado. Tanto os nós-fato quanto os nós-fato-independentes podem pertencer a um estado de crenças como nós-fundamentais, isto é, nós com as justificativas $<\{\},\{\}>$. Mas também podem pertencer a um estado de crenças por meio de alguma justificativa $\mathcal{P}$ ou $\mathcal{A}$ (como os nós derivados do RMS). A diferença entre os nós-fato e os nós-fato-independentes é que se houver necessidade de resolver algum conflito entre nós de ambas categorias, os nós-fato-independentes serão mantidos sempre que possível. O RMS-C é constituído ademais de nós de tipo especial representando as **regras de inferência** e as **definições**. Aos nós deste tipo somente podem ser associadas justificativas do tipo $<\{\},\{\}>$ ou do tipo $<-,->$, caso em que estão *dentro* (de algum estado de crenças) ou *fora* (de algum estado de crenças), respectivamente. As regras de inferência representam as leis e passos inferenciais válidos ou inválidos (caso estiverem $dentro$ ou $fora$, respectivamente) de um estado de crenças. As definições representam basicamente o significado de certos nós em termos de outros nós.

Deste modo, um estado de crenças do sistema RMS-C é representado por um conjunto de quádruplas ordenadas $<n,\mathcal{P},\mathcal{A},t>$, em que n é o nome do nó, $\mathcal{P}$ e $\mathcal{A}$ as justificativas presentes e ausentes, respectivamente; e em que t está para o tipo de nó —nó-fato ($\mathcal{F}$), nó-fato-independente ($\mathcal{I}$), regra ($\mathcal{R}$), definição ($\mathcal{D}$).

Quando alguma entrada $<n,\mathcal{P},\mathcal{A},t>$ que apresente uma justificativa válida em B para n é adicionada a B (B_n^+), ou quando todas as justificativas para algum nó n são invalidadas em B (B_n^-), pode ocorrer que alguns nós que não eram inferíveis a partir dos nós presentes em B, passem a sê-lo a partir de B_n^+ ou B_n^-, isto é, passa a haver “nós implícitos”. Esses são os **nós-derivados** e só existem enquanto está ocorrendo uma mudança no estado de crenças a que pertencem. No decorrer da mudança de crenças eles são adicionados a B_n^+ ou B_n^- como nós (explícitos). Assim que os nós-derivados são inferíveis, eles são adicionados explicitamente a B_n^+ ou B_n^- tendo como justificativa o conjunto de informações das quais é derivado e sendo do nível do nó de menor nível da sua justificativa. Por exemplo, seja $B = \{<o,\{\},\{\},\mathcal{F}>,<o\&r \Rightarrow d,\{\},\{\},\mathcal{R}\}$; se a B for adicionado um nó-fato-independente $<r,\{\},\{\},\mathcal{I}>$, então será inferível um nó-derivado d. Assim, o nó d tem “$o\&r\&(o\&r \Rightarrow d)$” como justificativa $\mathcal{P}$ e o elemento de menor nível da sua justificativa é o. Portanto, d é do mesmo nível que o, que é nó-fato. O nó-derivado d então é adicionado explicitamente com nível de nó-fato a B pela entrada $<d,\{o\&r\&(o\&r \Rightarrow d)\},\{\},\mathcal{F}>$.

Quando ocorrer uma inconsistência, a ordem de prioridade de retenção será: definições, regras, nós-fato-independentes, nós-fato.

Revisão em RMS-C O procedimento de revisão dos estados de crenças desse sistema é descrito a seguir. Para revisar o estado de crenças B pelo nó n, deve-se expandir B por uma entrada com justificativa válida em B, do tipo $<n,\mathcal{P},\mathcal{A},t>$, em que se $<n,\mathcal{P}',\mathcal{A}',t'> \in B$, então t deve ser igual a t'. Isto é, as entradas para dado nó n em B, devem ser consistentes com relação à prioridade de retenção de n em B anteriormente estabelecida. Marcar o *status* de n como *rev*. Remover todos os nós do tipo $\neg n$ que eventualmente estiverem *dentro*, respeitando a ordem de prioridade. Se algum nó m foi removido, verificar o *status* dos nós que contém m nas suas listas $\mathcal{P}$ e $\mathcal{A}$. Computar os eventuais nós-derivados e acrescentá-los em seus respectivos níveis, de acordo com o procedimento descrito acima, e marcar o *status* dos nós-derivados recém-adicionados como *rev*; se houver algum conflito entre os nós que estão *dentro* e os nós *rev*, resolver o conflito dando prioridade máxima de retenção aos nós com *status rev*. Se algum nó m for removido neste estágio, verificar o *status* dos nós que contém m nas suas listas $\mathcal{P}$ e $\mathcal{A}$ e resolver os conflitos que surgirem. Removidos todos os conflitos, mudar os *status* dos nós modificados de *rev* para *dentro*.

Exemplos Ilustraremos o procedimento de revisão com alguns exemplos. Tomemos o exemplo 2 acima sob o RMS-C, adicionando a ele a informação

($term$) a temperatura do parafuso foi medida pelo termômetro,

e denominaremos o estado de crenças resultante de B. De acordo com a proposta acima, B será representado da seguinte maneira:

nós-fato ($\mathcal{F}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$\neg t$	$term$	$\{\}$	$dentro$
p	$\neg t, \{\}$	$\{\}$	$dentro$
$\neg v$	$p\&(dentro(p) \Rightarrow dentro(\neg v))$	$\{\}$	$dentro$

nós-fato-independentes ($\mathcal{I}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
f	$\{\}$	$\{\}$	$dentro$
$term$	$\{\}$	$\{\}$	$dentro$

Regras ($\mathcal{R}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$dentro(f\&t) \Rightarrow dentro(v)$	$\{\}$	$\{\}$	$dentro$

Definições ($\mathcal{D}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$dentro(p) \Rightarrow dentro(\neg v)$	$\{\}$	$\{\}$	$dentro$

Suponhamos agora que desejamos fazer com que o nó t fique *dentro* em B. Para tanto, adicionamos $< t, \{\}, \{\}, \mathcal{F} >$ a B, isto é, t será adicionado aos nós-fato como nó fundamental. O *status* de t será provisoriamente *rev*. Depois de adicionar $< t, \{\}, \{\}, \mathcal{F} >$ a B, entrada que justifica um nó incompatível com alguns nós em B, deve-se proceder à restauração da consistência. Estando um nó com *status rev*, como é o caso de t, todos os nós incompatíveis com ele devem ser removidos, como é o caso de $\neg t$. Para que $\neg t$ seja removido, $term$ deve sê-lo. Como $term$ foi introduzido como um nó fundamental, para removê-lo, retiramos as justificativas associadas a ele. Procede-se agora à revisão do *status* de cada nó que depende de $\neg t$ ou $term$, isto é, que os menciona em suas listas $\mathcal{P}$ e $\mathcal{A}$. O nó p perdeu uma de suas justificativas ($\neg t$) mas ainda está *dentro*, visto que também foi introduzido como nó fundamental. Contudo, a adição de t permite, conjuntamente com f que o nó-derivado v fique *dentro*, e que o nó-derivado $\neg p$ fique *dentro*. O nó v tem como justificativa $f\&t\&(dentro(f\&t) \Rightarrow dentro(v))$ e o nó $\neg p$

tem como justificativa $v\&(dentro(v) \Rightarrow dentro(\neg p))$. Visto que o menor nível dos nós da justificativa de v é de nó-fato, a categoria de v será de nó-fato, consequentemente, a categoria de $\neg p$ será também de nó-fato. Com esses dados (justificativa e categoria), entradas para v e $\neg p$ são adicionados explicitamente ao estado de crenças com o *status rev*. Agora deve-se novamente invalidar todas as justificativas para os nós de *status dentro* que são incompatíveis com os nós de *status rev*, como é o caso de p e $\neg v$. Deve-se novamente restabelecer a consistência eliminando as sentenças que estão no menor nível, portanto, o nó-fato p fica *fora* e com isso, $\neg v$ perde a justificativa e fica *fora*. Feito tal procedimento, deve-se verificar se os nós removidos não fizeram que outros nós também perdessem a justificativa. Finalmente, os nós com *status rev* passam para o *status dentro*. O estado de crenças B_t^* resultante então é:

nós-fato ($\mathcal{F}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$\neg t$	$term$	$\{\}$	$fora$
p	$\neg t$	$\{\}$	$fora$
$\neg v$	$p\&(dentro(p) \Rightarrow dentro(\neg v))$	$\{\}$	$fora$
t	$\{\}$	$\{\}$	$dentro$
v	$f\&t\&(dentro(f\&t) \Rightarrow dentro(v))$	$\{\}$	$dentro$
$\neg p$	$v\&(dentro(p) \Rightarrow dentro(\neg v))$	$\{\}$	$dentro$

nós-fato-independentes ($\mathcal{I}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
f	$\{\}$	$\{\}$	$dentro$
$term$	–	–	$fora$

Regras ($\mathcal{R}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$dentro(f\&t) \Rightarrow dentro(v)$	$\{\}$	$\{\}$	$dentro$

Definições ($\mathcal{D}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$dentro(p) \Rightarrow dentro(\neg v)$	$\{\}$	$\{\}$	$dentro$

Observe-se que o nó v está dentro em B_t^* e é inferível o condicional razoável ($t > v$) "se a temperatura de **b** tivesse sido 650ºC no tempo **t**, então **b** teria estado vermelho em **t**"; enquanto que $\neg f$ não está dentro de B_t^*, e não é inferível o condicional ($t > \neg f$) "se a temperatura de **b** tivesse sido 650ºC no tempo **t**, então **b** não teria sido de ferro em **t**".

Será analisado em seguida o exemplo de Quine acerca de Bizet e Verdi. Sejam as informações disponíveis abaixo:

$(\neg c)$	Bizet e Verdi não são compatriotas,
(b)	Bizet é francês,
(v)	Verdi é italiano,
$(\neg b')$	Bizet não é italiano,
$(\neg v')$	Verdi não é francês,
$(b \to \neg b')$	se Bizet é francês, ele não é italiano,
$(v \to \neg v')$	se Verdi é italiano, ele não é francês,
$(b \wedge v \to \neg c)$	se Bizet é francês e Verdi é italiano, eles não são compatriotas,
$(b' \wedge v' \to \neg c)$	se Bizet é italiano e Verdi é francês, eles não são compatriotas,
$(b \wedge c \to v')$	se Bizet é francês e Bizet e Verdi são compatriotas, Verdi é francês,
$(v \wedge c \to b')$	se Verdi é italiano e Bizet e Verdi são compatriotas, Bizet é italiano.

O estado de crenças B formado pelas informações acima é:

nós-fato ($\mathcal{F}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$\neg c$	$b\&v\ \&\ dentro(b\&v) \Rightarrow dentro(\neg c)$	$\{\}$	$dentro$
$\neg b$	$b\ \&\ dentro(b) \Rightarrow dentro(\neg b')$	$\{\}$	$dentro$
$\neg v'$	$v\ \&\ dentro(v) \Rightarrow dentro(\neg v')$	$\{\}$	$dentro$

nós-fato-independentes ($\mathcal{I}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
b	$\{\}$	$\{\}$	$dentro$
v	$\{\}$	$\{\}$	$dentro$

Regras ($\mathcal{R}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$dentro(b\&v) \Rightarrow dentro(\neg c)$	$\{\}$	$\{\}$	$dentro$
$dentro(b'\&v') \Rightarrow dentro(\neg c)$	$\{\}$	$\{\}$	$dentro$
$dentro(b\&c) \Rightarrow dentro(v')$	$\{\}$	$\{\}$	$dentro$
$dentro(v\&c) \Rightarrow dentro(b')$	$\{\}$	$\{\}$	$dentro$

Definições ($\mathcal{D}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$dentro(b) \Rightarrow dentro(\neg b')$	$\{\}$	$\{\}$	$dentro$
$dentro(v) \Rightarrow dentro(\neg v')$	$\{\}$	$\{\}$	$dentro$

Para revisar o estado de crenças B acima por c, adicionamos c a B como nó fundamental do tipo nó-fato com *status* rev. Procedemos à restauração da consistência. $\neg c$ deve ser removido e com ele, alguma de suas justificativas. Como $b\&v$ tem menos prioridade de retenção que $dentro(b\&v) \Rightarrow dentro(\neg c)$, $b\&v$ será removido. Para remover $b\&v$ seria necessário remover apenas um de $\{b, v\}$. Contudo, visto que ambos estão na mesma categoria de prioridade, não há motivos para escolher um em detrimento do outro, é necessário remover ambos b e v, sob pena de permitir que inferências contra-intuitivas, como as apresentadas anteriormente (seções 3.4 e 3.6), possam ser feitas. Dado que b e v estão $fora$, $\neg b'$ e $\neg v'$ perderam suas justificativas também passam a estar $fora$. Muda-se o *status* do nó c para $dentro$.

O estado B^*_c é o seguinte:

nós-fato ($\mathcal{F}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$\neg c$	$b\&v\ \&\ dentro(b\&v) \Rightarrow dentro(\neg c)$	$\{\}$	$fora$
$\neg b$	$b\ \&\ dentro(b) \Rightarrow dentro(\neg b')$	$\{\}$	$fora$
$\neg v'$	$v\ \&\ dentro(v) \Rightarrow dentro(\neg v')$	$\{\}$	$fora$
c	$\{\}$	$\{\}$	$dentro$

nós-fato-independentes ($\mathcal{I}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
b	$-$	$-$	$fora$
v	$-$	$-$	$fora$

Regras ($\mathcal{R}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$dentro(b\&v) \Rightarrow dentro(\neg c)$	$\{\}$	$\{\}$	$dentro$
$dentro(b'\&v') \Rightarrow dentro(\neg c)$	$\{\}$	$\{\}$	$dentro$
$dentro(b\&c) \Rightarrow dentro(v')$	$\{\}$	$\{\}$	$dentro$
$dentro(v\&c) \Rightarrow dentro(b')$	$\{\}$	$\{\}$	$dentro$

Definições ($\mathcal{D}$)

Nó	$\mathcal{P}$	$\mathcal{A}$	Status
$dentro(b) \Rightarrow dentro(\neg b')$	$\{\}$	$\{\}$	$dentro$
$dentro(v) \Rightarrow dentro(\neg v')$	$\{\}$	$\{\}$	$dentro$

Note-se que nem $v' \in B^*_c$ nem $b' \in B^*_c$, o que implica que nenhum dos condicionais contra-intuitivos ($c > v'$) "se Bizet e Verdi fossem compatriotas, Verdi seria Francês" e ($c > b'$) "se Bizet e Verdi fossem compatriotas, Bizet seria italiano" são asseríveis a partir do estado de crenças B representado acima.

O procedimento de revisão do sistema RMS-C, em relação aos demais, deixa mais claro o que deve ser assumido pelo processo de revisão e o que pertence ao processo de revisão. O sistema RMS-C não assume que um condicional contrafactual $a > c$ é razoavelmente inferível a partir de um estado de crenças B para mostrar que $a > c$ é inferível a partir de B. Naturalmente é necessário que certas distinções provindas de intuições pré-formais sejam assumidas e no sistema RMS-C isso se dá com relação à distinção entre nós-fato, nós-fato-independentes, regras e definições. A pré-suposição da distinção de tais conceitos em um estado de crenças B é mais plausível que a pré-suposição (como é o caso da abordagem de contrafactuais de Gärdenfors e Hansson) da existência de uma função de revisão para B que somente efetue mudanças de crenças razoáveis. Essa assunção implicitamente toma como dada a distinção entre aqueles contrafactuais que são razoavelmente inferíveis a partir do estado de crenças B e aqueles que não são razoavelmente inferíveis a partir de B.

Não obstante, se a distinção pressuposta no sistema RMS-C entre nós-fato, nós-fato-independentes, regras e definições depender em última instância da noção de condicional contrafactual, o nosso consolo é que ao menos essa dependência será menos direta que aquela que existe nas outras abordagens que investigamos.

Notas

[47]Our rule thus reads that a counterfactual is true if and only if there is some set S of true sentences such that S is compatible with C and with $\neg C$, and such that $A \cdot S$ is self-compatible and leads by law to C; while there is no set S' compatible with C and with $\neg C$, and such that $A \cdot S'$ is self-compatible and leads by law to $\neg C$.

[48]"A counterfactual conditional is, in effect, nothing other than a conditional which draws a consequence from an antecedent that is in fact a belief-contravening hypothesis."

[49][...] the solution of the Problem of Counterfactual Conditionals lies in making and supporting a distinction, within the group of logically eligible alternatives, between "natural" and "unnatural" ways of effecting a reconciliation between a belief-contravening hypothesis on the one hand, and on the other the entire set of residual beliefs which continue to be collectively inconsistent with it.

[50]We have used the mode of formulation of the law statement (and thus an element over and above the "assertive content" of the law) as a guide for establishing the relative modal categoricity of the fact-statements relevant to the context of the counterfactual hypothesis and needed for its analysis.

[51]Now, when we begin "If Jones were in Carolina, ..." we are abstracting from his actual geographic location, and putting him in a presumably non-actual location, while leaving other this as far as possible, the same. It is plain that the "other things" which may be included in the relevant conditions, S, *must not include anything about his actual location* at the time of the antecedent.

Capítulo 5

Considerações Finais

As investigações na teoria de revisão de crenças que tiveram início na década de 1970 tiveram um papel importante no estudo dos condicionais contrafactuais. A linha de pesquisa proposta por Goodman nessa época já praticamente havia caído em desuso e as abordagens de contrafactuais por meio de mundos possíveis recém propostas, como [Sta68] e [Lew73] começavam a ganhar popularidade. As propostas feitas por Rescher [Res64], embora bastante pertinentes e inovadoras para a abordagem de Goodman, não tiveram grande repercussão na literatura sobre contrafactuais. Uma possível razão era que os sistemas obtidos por meio da noção de mundo possível eram muito sofisticados em termos formais e, não obstante sua tenra idade, bastante desenvolvidos. Isso deve ter atraído a atenção dos pesquisadores inclinados à abordagens mais formalísticas a problemas conceituais. O modelo AGM veio a calhar no âmbito de investigação dos contrafactuais porque ao mesmo tempo que conseguia reproduzir os resultados formais das abordagens de mundos possíveis, utilizava conceitos iniciais muito mais razoáveis.

O teoria AGM de revisão de crenças foi desenvolvida com a intenção de modelar o processo de mudança de crenças de "agentes ideais" ou logicamente oniscientes, que acreditam em todas as consequências lógicas de suas crenças. A ideia basilar é que, quando um agente de tal feitura muda suas crenças, ele o faz de uma maneira mínima. Isto é, para remover uma crença mantém o máximo de crenças possível e para adicionar uma crença adiciona o mínimo de crenças possível. Para tentar capturar essa ideia, denominada neste texto "critério de mudança mínima", Alchourrón *et al.* propuseram uma série de postulados com a intenção de regular cada tipo de mudança de crenças de modo que elas efetuassem mudanças mínimas. Neste sentido, foram propostos postulados como $^{+}$**-vacuidade** e $^{+}$**-minimalidade** para a operação de expansão, ***-vacuidade**, ***-suplementar**$_1$ e ***-suplementar**$_2$ para a operação de revisão e **–vacuidade**, **–suplementar**$_1$, **–suplementar**$_2$ e **–recuperação** para a operação de contração. Dentre esses postulados, o postulado que mais fortemente asseguraria que as mudanças incidentes nos estados de crenças fossem mínimas é **–recuperação**. Não obstante, uma função de contração de K por a, denominada "contração de intersecção total", cujo resultado é nada mais que os elementos de K que implicam $\neg a$ (todos outros elementos eventuais não relacionados nem com a

nem com $\neg a$ de K são eliminados), satisfaz **–recuperação**, como mostra o teorema 1.7.2 (seção 1.7 acima). Isso é razão suficiente para concluir que o postulado não consegue assegurar que a operação de contração efetue mudanças mínimas.

No artigo [AGM85] foi demonstrado (seção 1.6.1 acima) que uma função de contração para K e a satisfaz os postulados **–fecho/–suplementar**$_2$ se e somente se é uma função de contração por intersecção parcial, em que esse último termo significa que a função é obtida por meio de uma seleção dos subconjuntos maximais por inclusão de K que não implicam a. Posteriormente, (seção 1.6.2 acima) Gärdenfors e Makinson [GM88] mostraram que uma função (–) é uma função de contração construída a partir de um ordenamento de arraigamento epistêmico obedecendo os postulados **EE1/EE5** se e somente se (–) é uma função de contração por intersecção parcial. Portanto, grande parte dos resultados do modelo AGM gira em torno da noção de intersecção parcial. No entanto, os postulados e conceitos (para a contração, revisão e arraigamento epistêmico) conectados com essa noção de intersecção parcial não parecem plausíveis, mesmo se considerados no âmbito de "agentes ideais". No caso dos postulados para as operações de mudança de crenças, o critério de recuperação (seção 1.7 acima), em primeiro lugar, não cumpre seu papel (ver teorema 1.7.1), em segundo lugar, mesmo se o cumprisse de alguma forma, poderia ainda ser argumentado em favor de sua inadequação material. No caso da contração por arraigamento epistêmico, construída a partir de um ordenamento de arraigamento epistêmico que obedece certos postulados (**EE1/EE5**), a situação não é melhor, porque os postulados **EE1/EE5** implicam que o ordenamento de arraigamento epistêmico que os obedece é conectado (teorema 1.6.1), uma característica que intuitivamente não é válida em geral. Ademais, os postulados **EE1/EE5** foram propostos com a intenção de capturar alguns aspectos qualitativos da relação de importância epistêmica, mas, como argumentamos na seção 1.6.2, não é possível pretender razoavelmente que esses postulados capturem essa noção.

Tomando como teoria subjacente o modelo AGM e especificamente a operação de revisão, Gärdenfors propôs uma abordagem para condicionais contrafactuais (capítulo 2 acima). Nessa abordagem, ele propôs por meio do teste de Ramsey axiomas correlatos a cada postulado para a operação de revisão. De modo surpreendente, Gärdenfors mostra que a sua abordagem é equivalente, em termos de potência expressiva, à lógica "oficial" de Lewis [Lew73] para contrafactuais, que é baseada no conceito de mundo possível. Notamos que o formalismo oferecido pelo modelo AGM, adaptado à investigação dos contrafactuais por Gärdenfors, apresenta um ganho importante com relação às teorias de contrafactuais baseadas em mundos possíveis, como a de Lewis, que tornaram-se predominantes depois do artigo de Stalnaker [Sta68]. Isto ocorre por ao menos dois motivos muito fortes: (i) as teorias de mundos possíveis não elucidam o que é um mundo possível; (ii) não bastando a obscuridade da própria noção de mundo possível, ainda se presume que existe um ordenamento dos mundos possíveis em termos de similaridade. Em suma, o suposto *explicans* faz uso de um aparato conceitual para a elucidação do *explicandum* tal que o aparato conceitual é muito mais complexo que o próprio *explicandum*. Ademais[i], as

[i]Para não deixar de utilizar um argumento *ad verecundiam*.

abordagens em geral de contrafactuais surgiram da semântica de mundos possíveis de Kripke, e o próprio Kripke tenta [Kri80] explicar o que é um mundo possível em termos de situações contrafactuais, além de considerar a apropriação da noção de mundos possíveis pelos filósofos como geradora de pseudo-problemas e enganadora, conforme as seguintes evidências textuais:

> Até então nós estamos falando do mundo atual. Agora considere um mundo possível. Considere uma situação contrafactual em que, digamos, ouro de tolo ou pirita de ferro fosse realmente encontrada em várias montanhas dos Estados Unidos, ou em áreas da África do Sul e da União Soviética [Kri80, pág. 124].[52]
>
> Em qualquer situação contrafactual onde as mesmas áreas geográficas fossem preenchidas de tal substância, elas não teriam sido preenchidas com ouro (*ibid.*).[53]
>
> Eu às vezes utilizei 'situação contrafactual' no texto; Michael Slote sugeriu que 'estado possível (ou história) do mundo' poderia ser menos enganador que 'mundo possível'. É melhor ainda para evitar confusão não dizer 'Em algum mundo possível, Humphrey teria ganho' mas, simplesmente, 'Humphrey poderia ter ganho'. O aparato de mundos possíveis tem sido (eu espero) muito útil pelo menos na medida em que a teoria dos modelos conjuntista da lógica modal é considerada, mas encorajou pseudo-problemas filosóficos e imagens enganadoras (*ibid.*, pág. 48 nota 15).[54]

Gärdenfors, utilizando a noção de estado de crenças e a noção de mudança mínima entre estados de crenças, conseguiu criar uma lógica para contrafactuais equivalente à de Lewis, que utiliza a noção de mundos possíveis. Portanto, Gärdenfors conseguiu reduzir drasticamente a complexidade do *explicans* em relação ao *explicandum*, que é a noção de condicional contrafactual. Apesar disso, alguém que gostaria de guiar ou entender melhor o processo de inferência de condicionais contrafactuais seguramente não avançaria muito ao analisar os axiomas para a lógica de contrafactuais de Gärdenfors, nem os postulados para a operação de revisão do modelo AGM. Um sintoma de que havia algo errado em relação à abordagem para os condicionais contrafactuais, baseada do teste de Ramsey e os postulados do modelo AGM é o teorema de impossibilidade de Gärdenfors (teorema 2.3.2). O teorema de impossibilidade mostra que assunções do modelo AGM relativamente fracas como ***-vacuidade** são incompatíveis com o teste de Ramsey.

Em um artigo de 2006, Tennant mostra (teorema 3.2.1) que dado um conjunto de crenças K e uma sentença a, para "praticamente" qualquer conjunto J é possível encontrar uma função de revisão (*) obedecendo os postulados AGM ***-fecho/*-suplementar**$_2$, tal que $K_a^* = J$. O advérbio "praticamente" da frase anterior tem um sentido exato: o conjunto J apenas deve ser consistente e conter a sentença a. Isto é, se um conjunto J é consistente e $a \in J$, então J é qualificado pelo modelo AGM como uma revisão de K por a. O teorema de degeneração de Tennant (teorema 3.2.1) mostra que os postulados AGM são demasiado amplos para captar alguma noção razoável de mudança de crenças. Dado que a operação de revisão está tão pouco limitada no modelo AGM, haja vista o teorema 3.2.1, mostramos que condicionais contra-intuitivos como aqueles propostos por Quine (apresentados na seção 3.4 acima), são inferíveis na lógica de contrafactuais de Gärdenfors (teorema 3.4.4). Isso nos levou à procura de abordagens para contrafactuais que tenham como teoria subjacente uma teoria de revisão de crenças mais restrita que o modelo AGM. Assim, analisamos (seção 3.6 acima) a teoria de revisão por base de crenças de Hansson e sua abordagem para contrafactuais.

A abordagem de Hansson de contrafactuais em termos de revisão em bases de crenças novamente apresenta um avanço sobre a abordagem de Gärdenfors. Isso é o caso pois o teorema 3.2.1 já não se aplica, a quantidade das possíveis revisões que satisfazem os postulados para a revisão em bases de crenças é muito menor que no modelo AGM e a operação de revisão é mais intuitiva que aquela. No entanto, também mostramos (corolário 3.6.7) que no sistema de Hansson é possível construir funções que satisfaçam os postulados e que resultem em operações de revisão que validem contrafactuais implausíveis.

Portanto, fez-se necessária uma volta aos primórdios da investigação dos contrafactuais (capítulo 4 acima). Neste sentido, é curioso ver quão maior é o esforço de Goodman, Parry e Rescher de elucidar efetivamente o conceito de contrafactual, comparado às investigações posteriores. Isto é perceptível pois os critérios propostos por esses autores buscavam nos guiar na prática inferencial de contrafactuais. Por exemplo, o critério oferecido por Goodman no sentido de que não pode haver dois contrafactuais opostos que sejam sustentados por premissas de mesmo grau epistêmico (seção 4.2 acima); ou o princípio de retenção de Rescher (seção 4.3.1 acima), que tenta explicar por que certas informações devem ser removidas em detrimento de outras ao se levantar uma hipótese contrafactual; por fim, o critério de Parry (seção 4.4 acima), de que não podemos utilizar como premissa informações que dependam daquilo do qual se abstrai ao propor uma hipótese contrafactual.

A partir dessas considerações, tomamos uma teoria de revisão de crenças com uma estrutura bastante rica, o sistema RMS de Doyle (seção 4.4.1 acima), e propomos algumas alterações estruturais do sistema RMS, obtendo o esboço de uma variante denominada RMS-C. Mostramos (seção 4.4.2 acima) que o sistema RMS-C consegue lidar de forma adequada os contra-exemplos para as abordagens dos condicionais contrafactuais de Gärdenfors e Hansson.

Acreditamos que a representação dos estados de crenças como estruturas capazes de reproduzir as relações de justificação de forma razoavelmente rica, como o sistema RMS-C, é capaz de oferecer uma melhor elucidação da noção de informação relevante e do processo de inferência contrafactual.

Notas

[52] So far we are speaking of the actual world. Now consider a possible world. Consider a counterfactual situation in which, let us say, fool's gold or iron pyrites was actually found in various mountains in the United States, or in areas of South Africa and the Soviet Union.

[53] In any counterfactual situation where the same geographical areas were filled with such a substance, they would not have been filled with gold.

[54]I have sometimes used 'counterfactual situation' in the text; Michael Slote has suggested that 'possible state (or history) of the world' might be less misleading than 'possible world'. It is better still, to avoid confusion, not to say, 'In some possible world, Humphrey would have won' but rather, simply, 'Humphrey might have won'. The apparatus of possible words has (I hope) been very useful as far as the set-theoretic model-theory of quantified modal logic is concerned, but has encouraged philosophical pseudo-problems and misleading pictures.

Referências Bibliográficas

[AGM85] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[AM82] Carlos Alchourrón and David Makinson. On the logic of theory change: contraction functions and their associated revision functions. *Theoria*, 48:14–37, 1982.

[Cha03] Oswald Chateaubriand. *Logical Forms- Part II*, volume II. Coleção CLE, 2003.

[Chi46] Roderick M. Chisholm. The contrary-to-fact conditional. *Mind*, 55(220):289–307, October 1946.

[Coo57] John Cooley. Professor goodman's "fact, fiction & forecast". *Journal of Philosophy*, 54(10):293–311, 1957.

[Doy79] Jon Doyle. A truth maintenance system. *Artificial Intelligence*, 12(3):231 – 272, 1979.

[Fuh91] André Fuhrmann. Theory contraction through base contraction. *Journal of Philosophical Logic*, 20(2):175–203, 1991.

[Gär78] Peter Gärdenfors. Conditionals and Changes of Belief. *Acta Philosophica Fennica*, (30):381–404, 1978.

[Gär84] Peter Gärdenfors. Epistemic importance and minimal changes of belief. *Australasian Journal of Philosophy*, 62(2):136–157, 1984.

[Gär86] Peter Gärdenfors. Belief revisions and the ramsey test for conditionals. *The Philosophical Reveiw*, 95(1):81–93, 1986.

[Gär88] Peter Gärdenfors. *Knowledge in Flux.* Mit Press, 1988.

[GM88] Peter Gärdenfors and David Makinson. Revisions of knowledge systems using epistemic entrenchment. In *Proceedings of the 2nd conference on Theoretical aspects of reasoning about knowledge*, pages 83–95. Morgan Kaufmann Publishers Inc., 1988.

[Goo47] Nelson Goodman. Counterfactual conditionals. *Journal of Philosophy*, 44(5):113–128, 1947.

[Goo57] Nelson Goodman. Reply to an adverse ally. *Journal of Philosophy*, 54(17):531–535, 1957.

[Han91] Sven Ove Hansson. Belief contraction without recovery. *Studia Logica: An International Journal for Symbolic Logic*, 50(2):251–260, 1991.

[Han92a] Sven Ove Hansson. In defense of base contraction. *Synthese*, 91(3):439–445, 1992.

[Han92b] Sven Ove Hansson. In defense of the ramsey test. *Journal of Philosophy*, 89(10):522–540, 1992.

[Han99] Sven Ove Hansson. *A Textbook on Belief Dynamics*. Kluwer Academic Publishers, 1 edition, 1999.

[Han06] Sven Ove Hansson. Logic of belief revision. In *The Stanford Encyclopedia of Philosophy*. 2006.

[Kri80] S.A. Kripke. *Naming and Necessity*. Harvard University Press, 1980.

[Lev03] Isaac Levi. Counterexamples to recovery and the filtering condition. *Studia Logica: An International Journal for Symbolic Logic*, 73(2):pp. 209–218, 2003.

[Lew73] David Lewis. *Counterfactuals*. Blackwell, 1973.

[LR91] Sten Lindström and Wlodek Rabinowicz. Epistemic entrenchment with incomparabilities and relational belief revision. In *The Logic of Theory Change*. Springer, 1991.

[LR95] Lindstrom and Rabinowicz. The ramsey test revisited. In CROCCO, FARINAS DEL CERRO, and HERZIG, editors, *Condidionals: From Philosophy to Computer Science*. Claredom Press - Oxford, 1995.

[Mak85] David Makinson. How to give it up: a survey on the logic of theory change. *Synthese*, 62:347–363, 1985.

[Mak87] David Makinson. On the status of the postulate of recovery in the logic of theory change. *Journal of Philosophical Logic*, 16:383–394, 1987.

[Mak97] David Makinson. On the force of some apparent counterexamples to recovery. In Valdés, Krawietz, von Wright, and Zimmerling, editors, *Normative Systems in Legal and Moral Theory*, pages 475–481. Duncker and Humbolt, 1997.

[Mor92] Michael Morreau. Epistemic semantics for counterfactuals. *Journal of Philosophical Logic*, 21(1):33–62, 1992.

[Par57] William Parry. Reëxamination of the problem of counterfactual conditionals. *Journal of Philosophy*, 54(4):85–94, 1957.

[Qui82] W.V.O Quine. *Methods of Logic*. Cambridge, 1982.

[Ram50] Frank Ramsey. General propositions and causality. In Braithwaite, editor, *The Foudations of Mathematics and Other Logical Essays*, pages 237–255. Routledge and Kegan Paul LTD, 1950.

[Res61] Nicholas Rescher. Belief-contravening suppositions. *The Philosophical Review*, 70(2):176–196, April 1961.

[Res64] N. Rescher. *Hypothetical reasoning*. Studies in logic and the foundations of mathematics. North-Holland Pub. Co., 1964.

[Sta68] Robert Stalnaker. A theory of conditionals. In N. Rescher, editor, *Studies in Logical Theory*, pages 98–112. Oxford, 1968.

[Ten6a] Neil Tennant. On the degeneracy of the full agm-theory of theory-revision. *Journal of Symbolic Logic*, 71(2):661–676, 2006a.

[Ten6b] Neil Tennant. New Foundations for a Relational Theory of Theory-revision. *Journal of Philosophical Logic*, 35(5):489–528, October 2006b.

[Ten94] Neil Tennant. Changing the theory of theory change: Towards a computational approach. *The British Journal for the Philosophy of Science*, 45(3):865–897, 1994.

[Ten97] Neil Tennant. Changing the theory of theory change: Reply to my critics. *The British Journal for the Philosophy of Science*, 48, 1997.

[Ten08] Neil Tennant. Belief-revision, the ramsey test, monotonicity, and the so-called impossibility results. *The Review of Symbolic Logic*, 1(04):402–423, 2008.

[Zen09] F. Zenker. *Ceteris Paribus in Conservative Belief Revision: On the Role of Minimal Change in Rational Theory Development*. Peter Lang Pub Incorporated, 2009.

Índice Remissivo