Goethe-Universität
Frankfurt am Main
Institut für Philosophie

Departamento de Filosofia – IFCS
Univers. Federal do Rio de Janeiro
Março de 2010

Teorias de Mudança de Crenças (Dinâmica Doxástica)

André Fuhrmann

Esta série de oito palestras foi realizada entre os dias 3 e 29 de Março de 2010 no programa de Pós-Graduação em Filosofia no IFCS da Universidade Federal do Rio de Janeiro. Gostaria de agradecer aos meus anfitriões, Profs. Wilson Mendonça e Maria Clara Dias, e aos meus alunos no Rio, em particular Idia L. Ferreira e Julia Telles. Fico também grato pelo apoio financeiro recebido pelo CAPES e o DAAD.

As palestras baseiam-se numa contribuição ao *Routledge Companion to Epistemology*, org. por Sven Bernecker and David Pritchard, a ser publicado em 2010.

André Fuhrmann
Rio de Janeiro
Abril de 2010

Conteúdo

Justificação epistêmica

Clara vê o Peter dirigindo um Maserati preto em Frankfurt.
Ela crê que se o Peter dirige um carro em Frankfurt (cidade dele, vamos supor), então ele é dono daquele carro.
Em consequência, Clara crê (ao tempo t) que

(m) Peter é dono de um Maserati preto.

Um problema tradicional

Seja K o conjunto das crenças de Clara no tempo t.

- As crenças em K estão justificadas?
- Em particular, será que a crença (m) está justificada?

A crença (m) depende (deductivamente) nas seguintes crenças:

(a) Foi o Peter quem dirigiu o carro que a Clara viu em Frankfurt.

(b) O carro era um Maserati preto.

(c) Se Peter dirige um carro em Frankfurt, então é o carro dele.

$$\frac{a \qquad b \qquad c\ [= a \wedge b \rightarrow m]}{m}$$

- *Problema*: Será que as crenças (a-c) estão justificadas?
  (O problema de justificar as crenças atuais de uma pessoa)

  $\hookrightarrow$ **Teorias tradicionais de justificação epistêmica.**

Um problema distinto

Clara de repente soube de uma fonte confiável que o Pedro não é dono de um Maserati, i.e. $\neg m$. Sendo assim, ela decide dispensar da sua crença em m (talvez para aceitar $\neg m$ em seguida).

- *Tarefa*: Procede de K para K*-sem-*m!

Mas a presença de m é deductivamente forcada pelas outras crenças da Clara: a, b e c.

- *Segue-se*: Clara deve cancelar pelo menos uma das suas crenças a, b ou c. Naturalmente ao fim de cancelar o comprometimento com m, ela pode desativar muitas outras crenças além de a, b ou c – Clara poderia até cai em meditações cartesianas.

$$\begin{aligned} K - m &= \{\cdots a, b \cdots\} \ ? \ \{\cdots a, c \cdots\} \ ? \ \{\cdots b, c \cdots\} \ ? \\ &= \{\cdots a \cdots\} \ ? \ \{\cdots b \cdots\} \ ? \ \{\cdots c \cdots\} \ ? \ \{\cdots\} \ ? \ \{\mathit{cogito}\}?? \end{aligned}$$

- *Problema*: Dada uma decisão de cancelar m, qual conjunto de crenças $K - m$ qualifica como o legítimo (justificado!) succedor to K?
  (O problema de justificar mudanças de crenças.)

  $\hookrightarrow$ **Teorias "dinámicas" de justificação epistêmica.**

Taking a closer look at the two problems

[Pode-se pular para folha 16: "Voltando ao problema".]

The problem of **static justification** consists in filling the blank in the scheme

(S) $$x \, \mathrm{R} \, K.$$

where R respresents some suitable relation of justification.

The object of justification, K, is an epistemic state, a belief set. Candidates for the justifier x are well known, e.g.:

- fundamental beliefs
  (i.e. a distinguished subset of K),
- coherence
  (i.e. the obtaining of certain relations within K),
- reliable causal anchoring
  (i.e. the obtaining of certain relation between K and the world).

In the **dynamical case** the problem concerns a different schema, viz.

(D) $$x \, \mathrm{R} \, (K, K').$$

The object of justification is now a pair of epistemic states (representing the transition from one belief state, K, to another one, K).

Question:
Can we reduce the dynamic case to the static one or vice versa?

Perhaps ...

Can we reduce the dynamic case (D) to the static one (S)?

1. (K, K') is D-justified iff K' is S-justified.

But the origin state K does not enter the right-hand-side. Thus follow some rather implausible claims:

- $(\rightarrow)$ The existence of an origin K such that the transition from K to K' would be D-justified, suffices for K' to be S-justified.
  - Let $K = K$. The trivial transition (K, K) is presumably always D-justified. But then every belief set K would be S-justified. – Not a good idea.
- $(\leftarrow)$ All transitions (K, K') would be D-justified, given only that K' is S-justified.
  - In effect this denies that there is any interesting notion of justifying belief changes. All belief changes are equally good as long as they are changes to S-justified states. – Radical and implausible as long as more discriminatory theories of belief change remain possible.

Can we reduce the dynamic case (D) to the static one (S)?

2. (K, K') is D-justified iff the successor state K' is justified a time t' when the transition is made from K which was S-justified at some prior time t.

- $(\rightarrow)$ The transition (K, K') may be D-justified without K ever having been S-justified. In effect the proposal makes it impossible to ever leave behing a D-unjustifiable belief state. – Not a good idea.
- $(\rightarrow)$ We may want to allow for the following possibility: The transition (K, K') may be D-justified without K' being S-justified. Perhaps K' is only slightly better than K but not sufficiently better to be called S-justified.

Can we reduce the dynamic case (D) to the static one (S)?

3. (K, K') is D-justified iff K' is better S-justified than K.

- $(\rightarrow)$ Good transition to bad beliefs: The transition could be a D-good one – required under the circumstances – without K' being S-better than K.
  - K' result from K by adding some contingent information. Some such additions must be D-justfiable so let us assume that (K, K') is. But K' is riskier than K. Thus, if we assume that the degree of S-justification decreases as the risk of error increases, then K' is less S-justified than K.
- $(\leftarrow)$ Bad transition to better beliefs: K' could be S-better than K but there is no way of getting from K to K' in a D-justified manner.
  - Let $A \in K$ be a well-(S-)justified belief. Let K' contain only those beliefs that are at least as S-justified as A. The K' is certainly S-better than K. But the transition (K, K') involves a gratuitous loss of information and is therefore not D-justified.

Can we reduce the static case to the dynamic one?

The Cartesian model:

- A belief state is S-justified iff
  the state may be presented as the final point of a series of D-justified transitions taking its origin in the urkorpus (Levi) – cogito (Descartes).

Perhaps ...

A strong thesis (Isaac Levi)

- *Inquiry* is the process by which we “fix” (Peirce) and change our beliefs.
- *Beliefs* are determined by their rôle in inquiry.

- For a belief to play its rôle in inquiry, the agent has to accept his beliefs. To accept a belief is to accept the belief as true.
- As long as an agent accepts a belief he cannot seriously doubt that belief.
- As long as an agent does not seriously doubt a belief, he cannot seriously ask whether that belief is justified.
- Hence, in the course of inquiry an agent presupposes that his current beliefs are **infallible**.
- From the agent’s point of view the attempt to justify his current beliefs is incoherent.
- (This does not rule out that the agent may want to persuade other agents of his belief by offering ”justifying” reasons to adopt that belief. But this is not an attempt at justifying one’s beliefs but an attempt to get others to perform a justified belief *change*.)

- Inquiry presupposes that beliefs are **corrigible**.
- Giving up ("correct") a belief that the agent at some point must take to be infallible needs justification.
- Thus, from the viewpoint of the agent, justifying one's current beliefs is incoherent while passing from one belief state to another does need justification.

The strong thesis:
Static epistemic justification is a bogus problem, while dynamic epistemic justification is not.

[Isaac Levi: *The Enterprise of Knowledge* (1980), *Decisions and Revisions* (1984).]

Whatever one may think of the strong thesis, it is astonishing

- that the static problem of justification has been taken to define the enterprise of epistemology for decades, and it is even more astonishing
- that the distinct problem of dynamic justification has generally gone unnoticed (until ca 1985 and with a few exceptions).

A possible explanation:

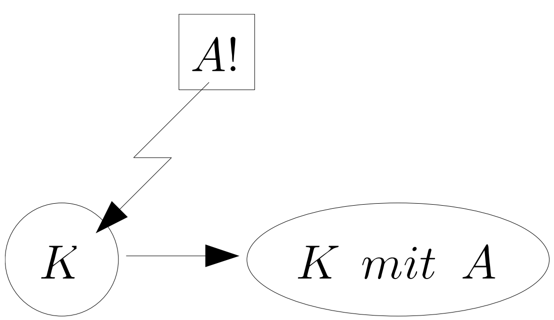
Thomas Kuhn (*The Structure of Scientific Revolutions*, 1970) has presented a theory of belief changes (theory changes). According to that theory large scale belief changes – paradigm shifts – do not bow to rational control. (Though Kuhn also asserted that changes within a paradigm *are* largely governed by the methodological rules of that paradigm.)

- Kuhn's theory (sufficiently vulgarized) has enforced for decades relativistic tendencies in philosophy and – even more so – in neighbouring disciplines. Such tendencies are rarely shared by epistemologists.
  - Belief changes, on the one hand, appeared under the Kuhnian perspective as rationally impredictable and not subject to normative contraints.
  - Most epistemologists agreed that knowledge claims, on the other hand and to the contrary, *are* subject to rational assessment.
- *Consequence*: The static problem remained within traditional epistemology; the dynamic problem – together with its relativistic side effects – was exiled to the History of Science.

Voltando ao problema

Abordagem da forma mais simples:

- Conjunto de crenças.
- Informação nova.
- Decisão (despertada pela informação) de aceitar uma nova crença.
- Como proceder?



Uma abordagem abstrata: três casos, três operações

Estado original K; transição para K'.

Três casos, três operações:

- $K \subset K'$: K' resulta por uma **expansão** de K.
- $K \supset K'$: K' resulta por uma **contração** de K.
- Nem $\subset$ nem $\supset$: K' resulta por uma **revisão** de K.

Uma abordagem mais concreta: Dois caso, três operações

Estado original K; informação nova A!

- *Primeiro caso*: A compatível (= consistente) com K.
  Simples: Vamos *expandir* K por A para obter $K + A$.
- *Segundo caso*: A inconsistente com K, i.e. $\neg A \in K$.
  Difícil: Temos que *revisar* K para incluir A, chegando a $K * A$.
  - Observação quanto ao caso segundo: Primeiro precisa fazer K consistente com A por tirar $\neg A$ de K, assim chegando à *contração* $K - \neg A$. A seguir só precisa expandir o resultado por A. Somamos os dois passos na

    Identitade do Levi:
    $$K * A = (K - \neg A) + A$$
- Assim temos três operações:
  - **Expansão**: $K + A$
  - **Contração**: $K - A$
  - **Revisão**: $K * A$

Programa

Problema inicial:
O que são mudanças racionais ("boas") de crenças?

Tese:
Todas as mudanças boas de crenças podem ser apresentadas como (decompostas em) sequências de boas expansões, contrações ou revisões.

Tarefa(s):
Resolver o problema inicial por buscar uma teoria de boas expansões, contrações e revisões.

Sneak preview:
Veremos que

- o problema das expansões tem uma solução fácil, e
- revisões podem ser definidos em termos de contrações e expansões (Id. do Levi!).

Assim as tarefas reduzem-se a fornecer uma teoria das boas contrações.

Sistemas de crença/Estados epistémicos/Teorias

- Linguagem da lógica proposicional:

$$P, Q, ..., \neg, \wedge, \vee \ \rightarrow, ...$$

$$A, B, C, ...$$

- Obrigações lógicas.
- Fechamento lógico:

$$K = \mathrm{Cn}(K)$$

- Operação de fechamento lógico: Cn

$$X \subseteq \mathrm{Cn}(X);$$
$$\text{se } X \subseteq Y, \text{ então } \mathrm{Cn}(X) \subseteq \mathrm{Cn}(Y);$$
$$\mathrm{Cn}(\mathrm{Cn}(X)) = \mathrm{Cn}(X);$$
$$\text{se } A \in \mathrm{Cn}(X), \text{ então } \exists X' \subseteq X : X' \text{ é finito e } A \in \mathrm{Cn}(X').$$

Forma relacional:

$$X \vdash A \text{ em vez de } A \in \mathrm{Cn}(X) \quad \text{e} \quad A \equiv B \text{ em vez de } \mathrm{Cn}(A) = \mathrm{Cn}(B).$$

- Fechamento de forma relacional:

$$A \in K \text{ sse } K \vdash A.$$

Expansões

Case K é compatível (consistente) com A, i.e.

$$\neg A \notin K$$

$\Rightarrow$

$$K + A := \mathrm{Cn}(K \cup \{A\})$$

Simples!

Revisões e contrações

(m) Pedro é dono de um Maserati.

$$\frac{a \qquad b \qquad a \wedge b \to m}{m}$$

· Escolha: Não somente um problema lógico!

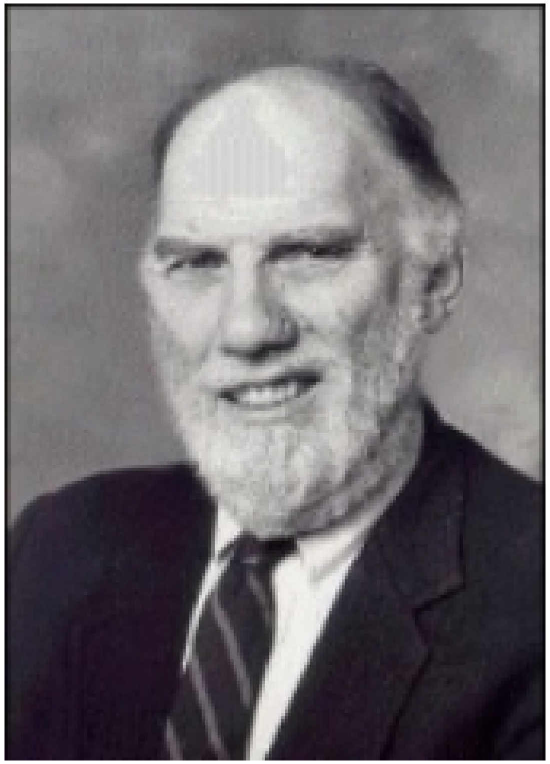
- Núcleo da teoria: identificar as opções e fazer uma escolha racional.

"cognitive decision theory" (Levi) – teoria de escolha cognitiva.

AGM: Postulados e construções

As fontes da teoria de AGM

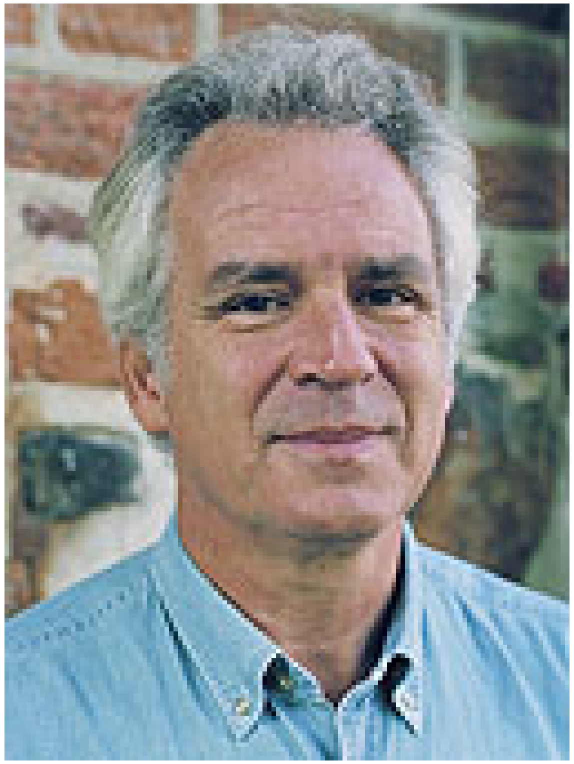
Inquiry **como uma teoria de escolha cognitiva:**
Isaac Levi (1967ff.)



Semântica epistêmica das proposições condicionais:
Peter Gärdenfors (1978ff.)

> If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q.

(Frank P. Ramsey, General propositions and causality, 1929)



Derogação de normas jurídicas:
Carlos Alchourrón e David Makinson (1981ff.)

> Suppose that A is a set of regulations, y is some proposition that is implied by A, and that for some reason a legislative body wants to eliminate y. In such a situation, the body may decide to reject y, with the intention of thereby rejecting implicitly whatever in A implies y, retaining the remainder. This we shall call derogation.

(AM, Hierarchies of regulations and their logic, 1981)

Considerações pré-teóricas

Juntar informaçao interessante/valorosa e confiável:

- ◦ quanto mais interessante tanto mais ariscada,
- ◦ quanto mais confiável tanto menos interessante.
  - · $\mathrm{Cn}(\emptyset)$: o estado mais confiável e o menos interessante!
- Equilíbrio razoável.

- $-/*/+$: teorias $\mapsto$ teorias
- A origem já está num equilíbrio: Não abre mão de informação sem necessidade (Quine: Máxima de Mutilação Mínima)!

Estratégia

- Postulados/condições
- Construções
- Harmonia: resultados des representação
  - construções satisfazem as condições, e
  - operações que satisfazem os condições podem ser tal construidos.
- Estabilidade da teoria

Postulados de AGM para contrações

(C1) Closure	$K - A = \mathrm{Cn}(K - A)$	$(Fechamento)$
(C2) Success	$A \notin K - A$, se $\not\vdash A$	
(C3) Inclusion	$K - A \subseteq K$	
(C4) Vacuity	Se $A \notin K$, então $K - A = K$	
(C5) Recovery	$K \subseteq (K - A) + A$	$(resgate)$
(C6) Congruence	Se $A \equiv B$, então $K - A = K - B$	

Suplementários

(C7) Conjunction 1	$K - A \cap K - B \subseteq K - (A \wedge B)$
(C8) Conjunction 2	Se $A \notin K - (A \wedge B)$, então $K - (A \wedge B) \subseteq K - A$

Postulados de AGM para revisões

(R1)	Closure	$K * A = \mathrm{Cn}(K * A)$
(R2)	Success	$A \in K * A$
(R3)	Inclusion	$K * A \subseteq K + A$
(R4)	Preservation	Se $\neg A \notin K$, então $K * A = K + A$
(R5)	Consistency	Se $\perp \in K * A$, então $\vdash \neg A$
(R6)	Congruence	Se $A \equiv B$, então $K * A = K * B$

Suplementários

(R7)	Conjunction 1	$K * (A \wedge B) \subseteq (K * A) + B$
(R8)	Conjunction 2	Se $\neg B \notin K * A,$ então $K * A) + B \subseteq K * (A \wedge B)$

[BDD-Axiome]

De contrações a revisões e v.v.

Identidade do Levi (LI):

$$K * A = K - \neg A + A$$

Identidade do Harper (HI):

$$K - A = K \cap (K * \neg A)$$

Observação. *Os postulados de contração e os de revisão harmonizam com os outros* módulo *(LI) resp. (HI):*

1. *Dos postulados de contração + LI seguem-se os da revisão; Resgate (Recovery) não entra na derivação.*
2. *Dos postulados de revisão + HI seguem-se os da contração.*

Recovery/Resgate: $K - A + A = K$

- Recovery: evite perder informaçao.
  - Exemplo: apago (Makinson: "withdrawal") satisfaz todos os postulados/C menos Recovery: (para $A \in K$) $K - A = \mathrm{Cn}(\emptyset)$.
- Sob certas condições, Recovery não é plausível (Fuhrmann, Hansson):
  - contrações em bases de crenças.
- Recovery repousa no fato que Cn é determinada pela lógica clássica (com silogismo disjuntivo).

Teste de Ramsey para condicionais (Monotonia)

Gärdenfors: semântica epistêmica para condicionais contrafatuais.

$$A > B$$

Se A fosse o caso [mas talvez não seja], B seria o caso.
Se Pedro fosse convidado, Maria iria embora.

Teste do Ramsey:

$$A > B \in K \text{ sse } B \in K * A$$

Observação. *Ramsey implica em Monotonia para revisões:*

$$\textit{Se } K \subseteq H, \textit{ então } K * A \subseteq H * A.$$

Observação. *(Gärdenfors) Monotonia, Preservation e Consistency são compatíveis somente num caso trivial.*

Prova. Caso não-trivial: Existem A, B, K

- A e B independentes em K $(A \to B, B \to A \notin K)$
- K indeciso quanto a A, B, i.e. $A, \neg A \notin K$ e $B, \neg B \notin K$.

Primeiro: Seja $K' = K + A$ e $K'' = K + B$ ...

$$\dfrac{\dfrac{A \in K'}{A \in K' + \neg(A \wedge B)} \qquad \dfrac{\dfrac{B \notin K'}{A \wedge B \notin K'}}{K' + \neg(A \wedge B) = K' * \neg(A \wedge B)}\;\mathit{Pres.}}{A \in K' * \neg(A \wedge B)} \qquad\qquad \dfrac{\vdots}{B \in K'' * \neg(A \wedge B)}$$

Segundo: Seja $H = K + (A \wedge B)$ ...

$$\dfrac{\dfrac{\dfrac{\dfrac{K' \subseteq H \supseteq K''}{K' * \neg(A \wedge B) \subseteq H * \neg(A \wedge B) \supseteq K'' * \neg(A \wedge B)}\;\mathit{Mono}}{A \in K' * \neg(A \wedge B) \subseteq H * \neg(A \wedge B) \supseteq K'' * \neg(A \wedge B) \ni B}\;(\mathit{de\ acima})}{A \wedge B \in H * \neg(A \wedge B)}}{\mathit{contradição!}}\;\mathit{Cons.}$$

Modalidades reflexivas

Confusão de camadas de reflexão? Compare

$$A > B \in K \text{ sse } B \in K * A$$

com

Teste do Levi (para possibilidade séria)

$$\Diamond A \in K \text{ sse } \neg A \notin K.$$

Observação. *(Fuhrmann) todos os testes deste tipo "reflexivo" são inconsistentes.*

Partial Meet Contraction

(Contração por intersecção parcial)

Construções que capturam o elemento de escolha.

- *Problema*: $K - A$
- *Candidatos*: Os subconjuntos maximais de K, tal que A não se segue.
- *Escolher* ...
  - *todos*: Intersecção dos todos. (Full Meet)
  - *um só*: O melhor (único). (Maxichoice)
  - *uns*: Intersecção dos melhores. (Partial Meet)

C. Alchourron, P. Gärdenfors und D.Makinson, On the logic of theory change: Partial meet functions for contraction and revision, *J. of Symbolic Logic* 50 (1985), 510-530.

A familia $K \perp A$ de **candidatos**

$X \in K \perp A$ sse

1. $X \subseteq K$,
2. $X \nvdash A$, e
3. $\forall Y$: se $X \subset Y \subseteq K$, então $Y \vdash A$.

$K \perp A$: família de conjuntos de "sobreviventes" (remainder set).

Para cada teoria K uma função de **seleção**

$$s_K : \wp(\wp(K)) \longrightarrow \wp(\wp(K)),$$

tal que:

1. $s_K(\mathbb{X}) \subseteq \mathbb{X}, \neq \emptyset$ caso $\mathbb{X} \neq \emptyset$
   [para maxichoice: $s_K(\mathbb{X})$ escolha um conjunto só de $\mathbb{X}$];
2. $s_K(\mathbb{X}) = \{K\}$ caso contrário.

(Não ecreveremos mais o subscripto K se a tal dependência está aparente pelo contexto.)

Contração por full meet:

$$K - A = \bigcap (K \perp A)$$

Contração por maxichoice:

$$K - A = s(K \perp A)$$

(No caso de maxichoice $s_K(K \perp A)$ contém um conjunto $X \in K \perp A$ resp. $= \{K\}$.)

Contração por partial meet, pm:

$$K - A = \bigcap s(K \perp A)$$

Representação para C1-6. *(AGM 1985)*

1. *Toda contração-pm satisfaz C1-6.*
2. *Toda contração que satisfaz C1-6, pode ser representada por uma contração-pm.*

Repräsentation für C1-8. *(AGM 1985)*
Sei $\leq_K$ *eine transitive und reflexive Ordnung auf* $\bigcup_A K \perp A$. *Die Funktion* s *wähle die unter* $\leq_K$ *maximalen Elemente in einer Restefamilie von* K*, d.h.*

$$s(K \perp A) = \{X \in K \perp A : (\forall Y) \textit{ wenn } Y \in K \perp A, \textit{ dann } Y \leq_K X\}.$$

Mit s *so bestimmt, nennen wir* $\bigcap s(K \perp A)$ *eine trpm-Kontraktion (transitively relational pm).*

1. *Jede trpm-Kontraktion erfüllt die Bedingungen C1-8.*
2. *Jede Kontraktion, welche C1-8 erfüllt, kann durch eine trpm-Kontraktion dargestellt werden.*

Prova para a representação de C1-6.

1. *Toda contração-pm satisfaz C1-6.*

Por exemplo *Recovery*: $K = K - A + A$.

$$\textit{A demonstrar}\text{: } K \subseteq \mathrm{Cn}(\bigcap s(K \perp A) \cup A)$$

$$\text{i.e., se } K \vdash B\text{, então } \bigcap s(K \perp A), A \vdash B$$

Caso limite: Se $K \nvdash A$, então $K \perp A = \emptyset$ a assim $s(K \perp A) = \{K\}$—pronto. Seguimos com o caso principal $K \vdash A$. Neste caso temos:

(AM82) $$\bigcap(K \perp A) = K \cap \mathrm{Cn}(\neg A).$$

$$\dfrac{\dfrac{\dfrac{\dfrac{K \vdash B}{K \vdash \neg A \vee B} \qquad \neg A \vdash \neg A \vee B}{K \cap \mathrm{Cn}(\neg A) \vdash \neg A \vee B} \qquad \dfrac{\text{(AM82)}}{K \cap \mathrm{Cn}(\neg A) \subseteq \bigcap s(K \perp A)}}{\bigcap s(K \perp A) \vdash \neg A \vee B}}{\bigcap s(K \perp A), A \vdash B}$$

2. *Toda (C1-6)-contração pode ser construída como uma contração-pm.*

Seja $K - (\ )$ uma contração (K arbitrária) que satisfaz C1-6. Definimos a **função canônica** de selecção σ (para K) assim:

$$(\sigma) \qquad \sigma(K \perp A) = \begin{cases} \{X \in K \perp A : K - A \subseteq X\}, & \text{caso } K \perp A \neq \emptyset; \\ \{K\}, & \text{caso contrário.} \end{cases}$$

A demonstrar: σ (para K) é uma função de selecção no sentido da definição anterior:

(1) σ é bem-definida, i.e. de $K \perp A = K \perp B$ segue-se $\sigma(K \perp A) = \sigma(K \perp B)$.

(2) $\sigma(K \perp A) = \{K\}$, caso $K \perp A = \emptyset$.

(3) $\sigma(K \perp A) \subseteq K \perp A$, caso $K \perp A \neq \emptyset$.

(4) $\sigma(K \perp A) \neq \emptyset$.

(5) $K - A = \bigcap \sigma(K \perp A)$.

(1) $\Leftarrow$ *Congruence* (C6).
(2), (3) $\Leftarrow$ def. (σ).
(4) $\Leftarrow$ def. (σ), se $K \perp A = \emptyset$, i.e. $\vdash A$ (caso limite);
(4) $\Leftarrow$ *Success* (C2) $K - A \nvdash A$ (assim $\exists X \supseteq K - A$ tal que $X \in K \perp A$), se $\nvdash A$ (caso principal).

Ad (5): $K - A = \bigcap \sigma(K \perp A)$.

$\subseteq \Leftarrow$ def. (σ) e *Inclusion* (C3).

$\supseteq$ *A demonstar*:

$$(*) \qquad \text{Se } B \notin K - A, \text{ então } B \notin \bigcap \sigma(K \perp A)$$

Caso (trivial) $A \notin K$. *Vacuity* (C4) dá $K - A = K$ e assim por (σ) temos $K = \bigcap \sigma(K \perp A)$.

Caso $A \in K$. (*) vale trvialmente, caso $B \notin K$.
Então suponhamos que $B \in K$. Temos que encontrar um conjunto X tal que

$$\text{(a) } X \in K \perp A \quad \text{e} \quad \text{(b) } K - A \subseteq X \quad \text{e} \quad \text{(c) } B \notin X.$$

/...

Hipóteses: 1) $A \in K$, 2) $B \in K$, 3) $B \notin K - A$. A encontrar: X t.q

$$\text{(a) } X \in K \perp A \quad \text{e} \quad \text{(b) } K - A \subseteq X \quad \text{e} \quad \text{(c) } B \notin X \ !$$

Lemma (AGM85, 2.4): $\dfrac{X \in K \perp C \quad D \in K \quad X \nvdash D}{X \in K \perp D}$

Ad b)
$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{B \in K}\,\text{hip.2}}{K - A, A \vdash B}\,\textit{Recovery} \qquad \dfrac{}{K - A \nvdash B}\,\text{hip.3}}{K - A, \neg A \nvdash B}}{K - A \nvdash A \vee B}}{\exists X \in K \perp A \vee B : K - A \subseteq X}$$

Ad c)
$$\dfrac{\dfrac{\dfrac{}{X \in K \perp A \vee B}\,\text{b)}}{X \nvdash A \vee B}}{B \notin X}$$

Ad a)
$$\dfrac{\dfrac{}{A \in K}\,\text{hip.1} \qquad \dfrac{}{X \in K \perp A \vee B}\,\text{b)} \qquad \dfrac{\dfrac{\dfrac{}{X \in K \perp A \vee B}\,\text{b)}}{X \nvdash A \vee B}}{X \nvdash A}}{X \in K \perp A}\,\text{Lemma}$$

■

Demonstramos assim:

Representação de C1-6. *Uma operação de contração satisfaz (C1-6), sse ela é uma contração-pm.*

Para a definição de revisões-pm empregamos a Id.Levi:

$$K * A = K - \neg A + A.$$

Assim definimos **revisão-pm**:

$$K * A = \mathrm{Cn}(\bigcap s(K \perp \neg A) \cup \{A\}).$$

Em seguida a harmonia entre os postulados C e R permite a ...

Representação de R1-6. *Uma operação de revisão satisfaz (R1-6), sse ela é uma revisão-pm.*

Sistemas de esferas

A semântica de Lewis para condicionais contrafatuais

Teste do Ramsey

$$A > B \in K \text{ sse } B \in K * A.$$

$K * A$: A revisão da *teoria* K t.q. A está inclusa.

Uma versão epistêmica do **Teste do Stalnaker**:

$$A > B \text{ é verd. em } w \text{ sse } B \text{ é verd. em } (w * A).$$

$w * A$: o mundo mais parecido com w t.q. A é o caso.

Talvez não haja um único mundo mais parecido com w t.q. ... Sendo assim deveriamos preferir a **versão do Lewis** do teste:

$$A > B \text{ é verd. em } w \text{ sse } B \text{ é verd. em todo os mundos em } (w * A).$$

$w * A$: os mundos mais parecidos com w t.q. A é o caso.

es.fe.ra *sf* **1** *Geom* Corpo cujos pontos têm igual distância de um ponto interior (centro).

Minidicionário Escolar da Língua Portuguesa
São Paulo (Melhoramentos), 2000.

Sistemas de esferas para condicionais

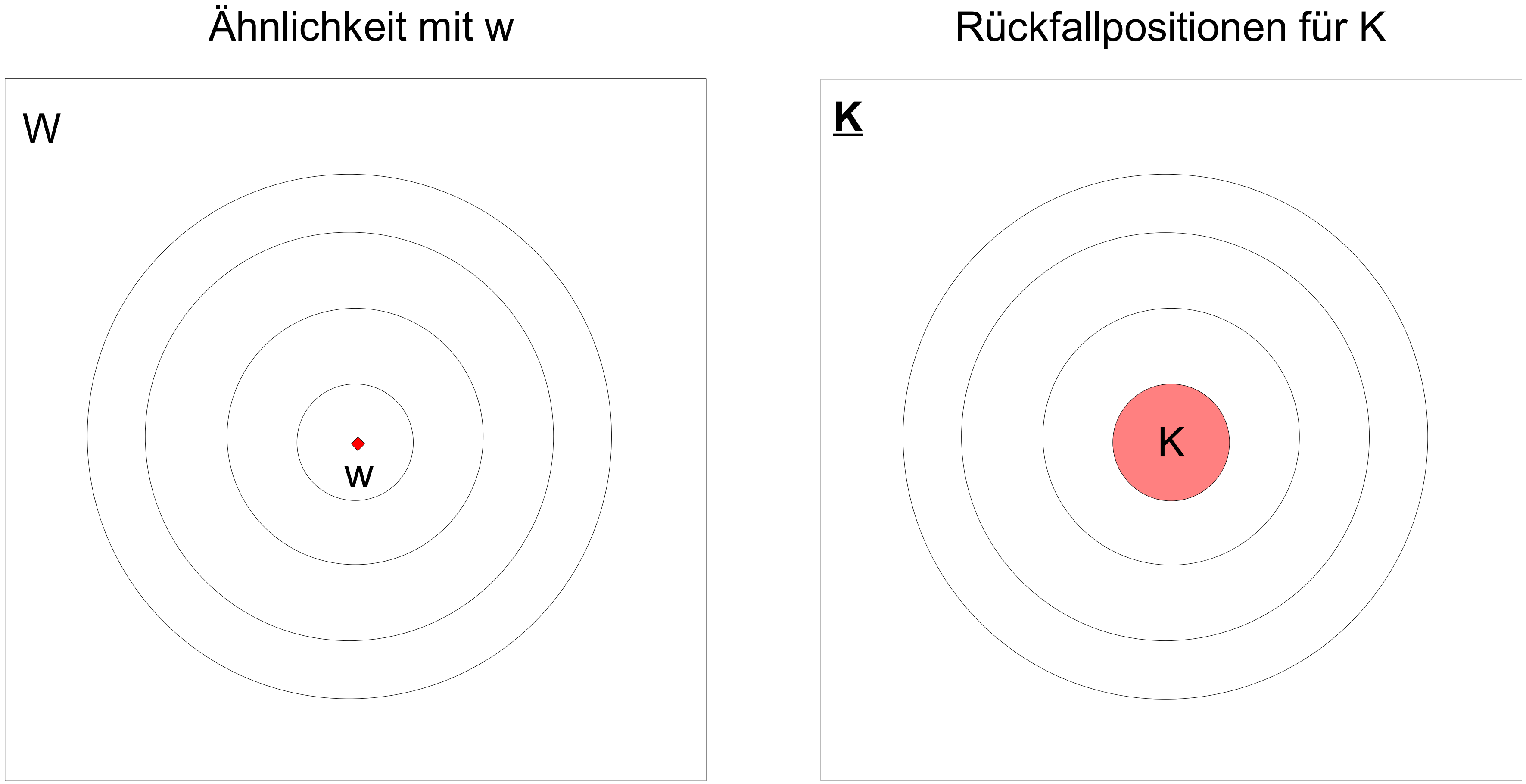
(David Lewis, *Conditionals*, 1973)

- Centrado no mundo atual w, coloque os demais mundos em esferas de semelhança ("proximidade") com w.
- Para avaliar $A > B$, vá à esfera mais próxima a w que admite A e considere todos os mundos-A naquela esfera (se existem mundos-A; caso contrário seja $A > B$ verdadereiro de modo trivial). Agora pergunte, se B for o caso naqueles mundos.

Sistemas de esferas para revisões

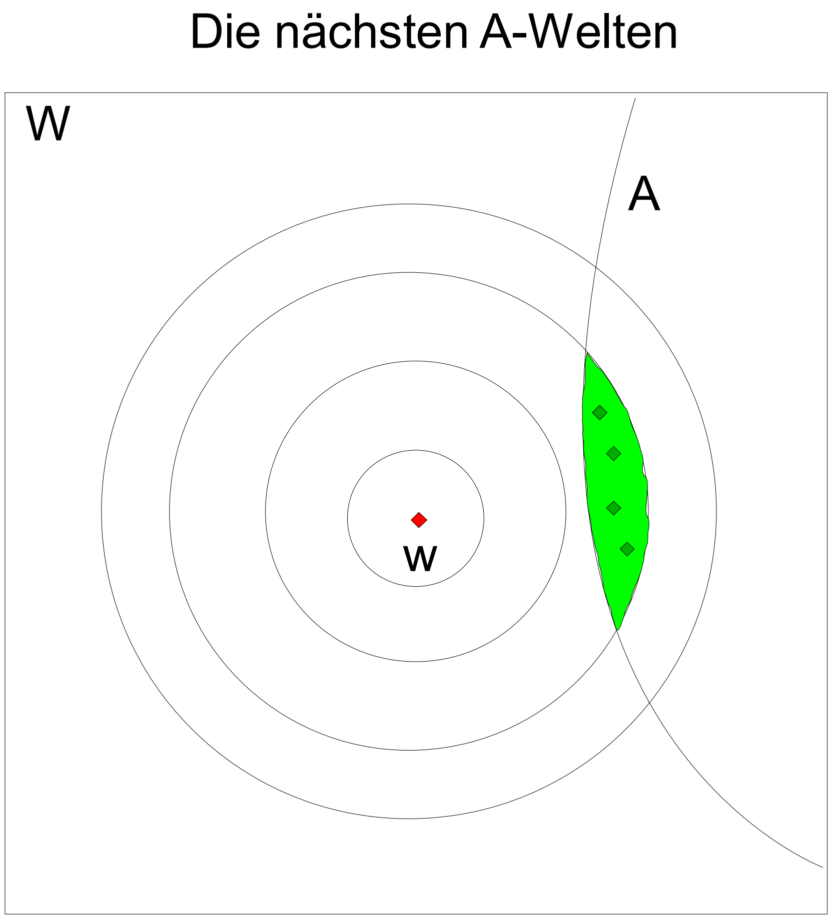
(Adam Grove, Two modellings of theory change, *J. of Philos. Logic* 17 (1988), 157-170.)

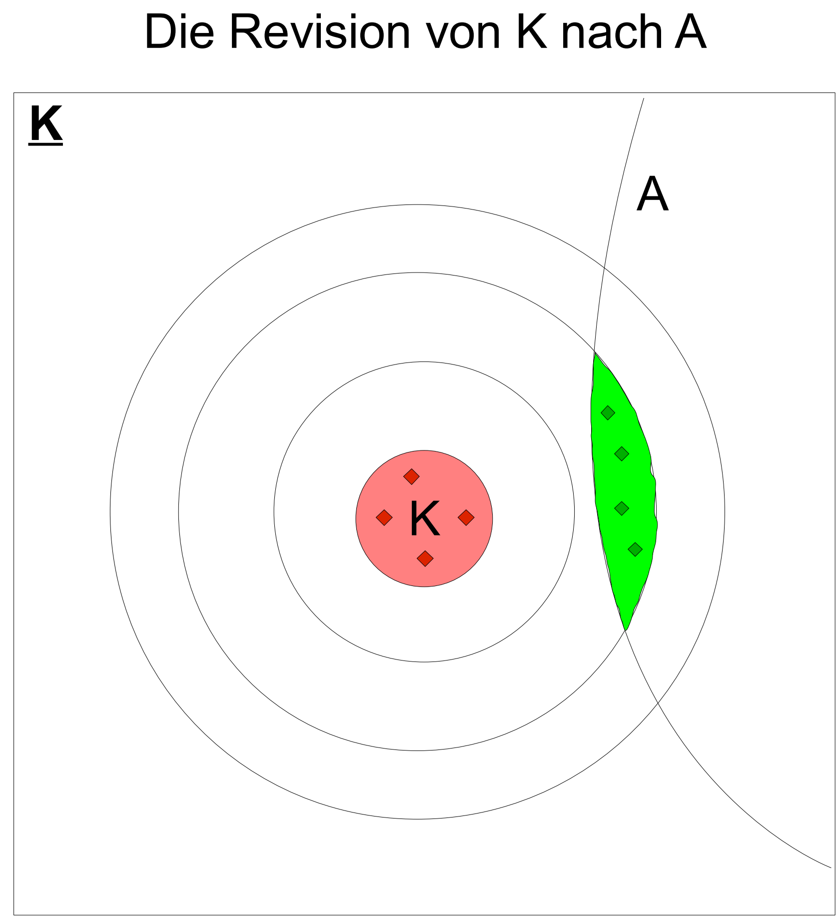
- ○ Representamos teorias não por um mundo só mas por um conjunto de mundos (aqueles mundos/possibilidades que a teoria "deixa em jogo").
- Centrado na teoria atual K, coloque os mundos exluídos por K em esferas de preferência/plausibilidade ("proximidade") do ponto de vista de K
- Para avaliar $B \in K * A$, vá a esfera mais próxima que admite A. A nova teoria $K * A$ será o conjunto dos mundos-A naquela esfera (se existem mundos A; se não $K * A$ será inconsistente ($= L$)). Agora pergunte, se B for o caso naqueles mundos.



Esquerda: Sistema de esferas centrado no mundo w.

Direita: Sistema de esferas ("cebóla de retiros") acerca da teoria K.



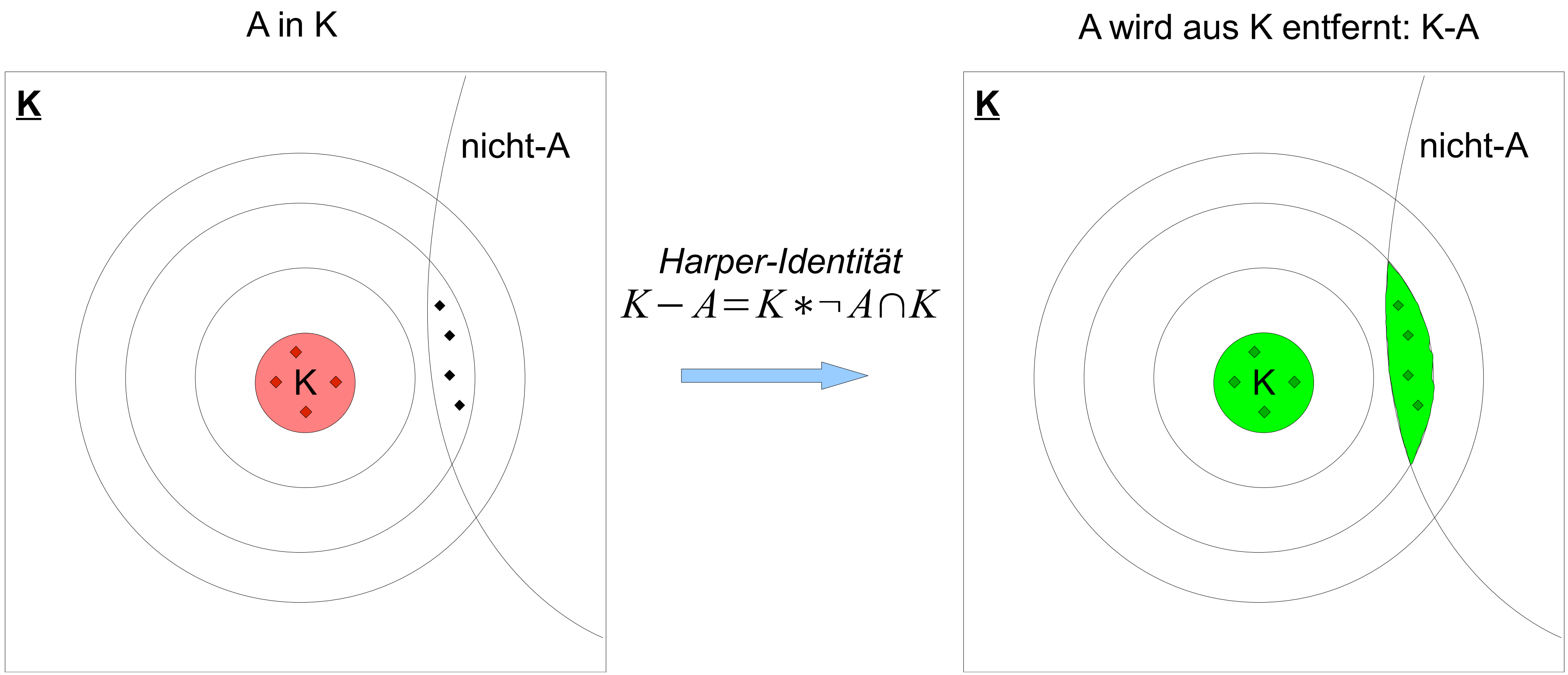


Esquerda: $w * A$ (verde) contém os mundos-A mais próximos de w.

Direita: $K * A$ (verde) contém os mundos-A relativamente mais plausíveis do ponto de vista de K [— na impressão em preto-branco tire o centro K!].

[DDL-Modelle]

Representação de contrações pela Id. do Harper:



Esquerda: K com as suas esferas de plausibilidade relativa.

Direita: Mundos-$\neg A$ juntados a K da maneira preferida segundo K.

Seja K uma teoria e
seja $[K]$ o conjunto das possibilidades abertas ("mundos possíveis") por K, i.e. uma representação de "K" em termos de mundos possíveis.

Sistema de esferas baseado em K (e W):
uma família $\mathfrak{S}_K$ de conjuntos tal que

1. $[K] \in \mathfrak{S}_K$ e $[K] \subseteq X$, $\forall X \in \mathfrak{S}_K$
   — $[K]$ é a esfera menor;
2. $W \in \mathfrak{S}_K$
   — W é a esfera maior;
3. $\mathfrak{S}_K$ está completamente ordenada por $\subseteq$ (i.e. $\forall X, Y \in \mathfrak{S}_K$: ou $X \subseteq Y$ ou $Y \subseteq X$)
   — cébola ou boneca russa;
4. $\forall A$: se $[A] \neq \emptyset$, existe ume esfera mínima $S \in \mathfrak{S}_K$ t.q. $[A] \cap S \neq \emptyset$
   — revisões são (quase) sempre possíveis.

(Simplificado, tirando uma condição "técnica".)

Para determinar $K * A$ num sistema de esferas precisa tomar em conta o *caso limite* que $\neg A \in \mathrm{Cn}(\emptyset)$ (i.e. $\neg A$ é um verdade lógica que não pode ser tirada). (As ilustrações prévias se referiram ao caso principal.)

- Para cada conjunto *não-vazio* $X \subseteq W$ em $\mathfrak{S}_K$ seja $S_X \in \mathfrak{S}_K$ a esfera mínima t.q. a intersecção com X é não-vazio. Assim definimos

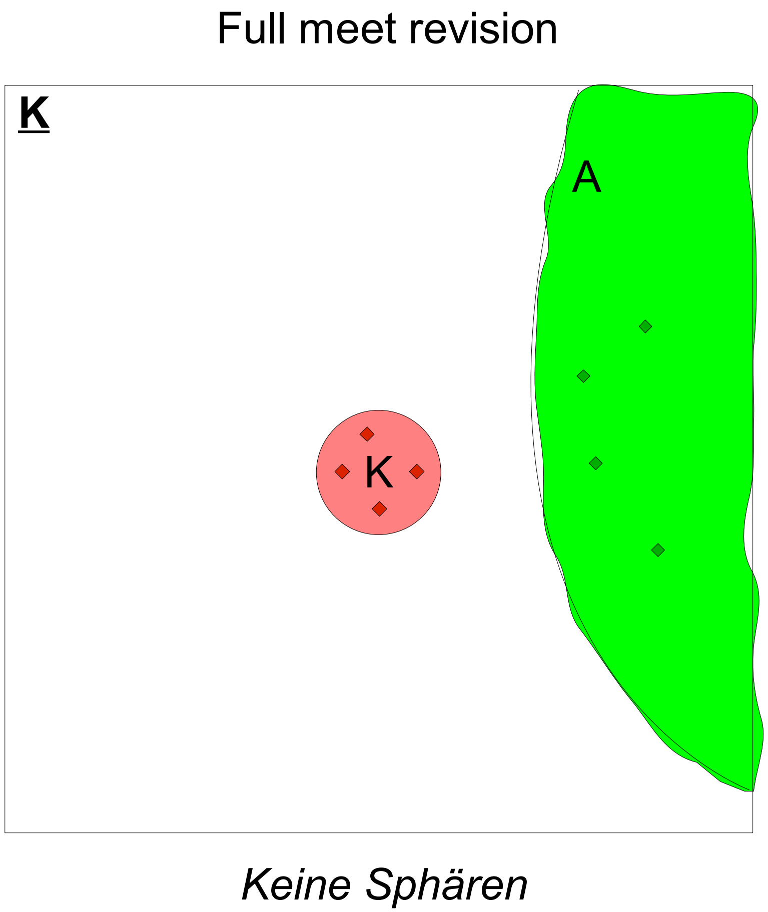
$$(*) \qquad K * A = \begin{cases} S_{[A]} \cap [A], & \text{caso } A \text{ é consitente;} \\ \mathrm{Cn}\,(\bot) & \text{caso contrário.} \end{cases}$$

Representação. *(Grove 1988)*

1. *Revisões segundo (*) num sist. de esferas baseado em K satisfazem os postulos (R1-8), e*
2. *Se uma operação $*$ para K satisfaz $R1-8$, então $K * A$ (para A qualquer) pode ser construído num sist. de esferas baseao em K segundo (*).*

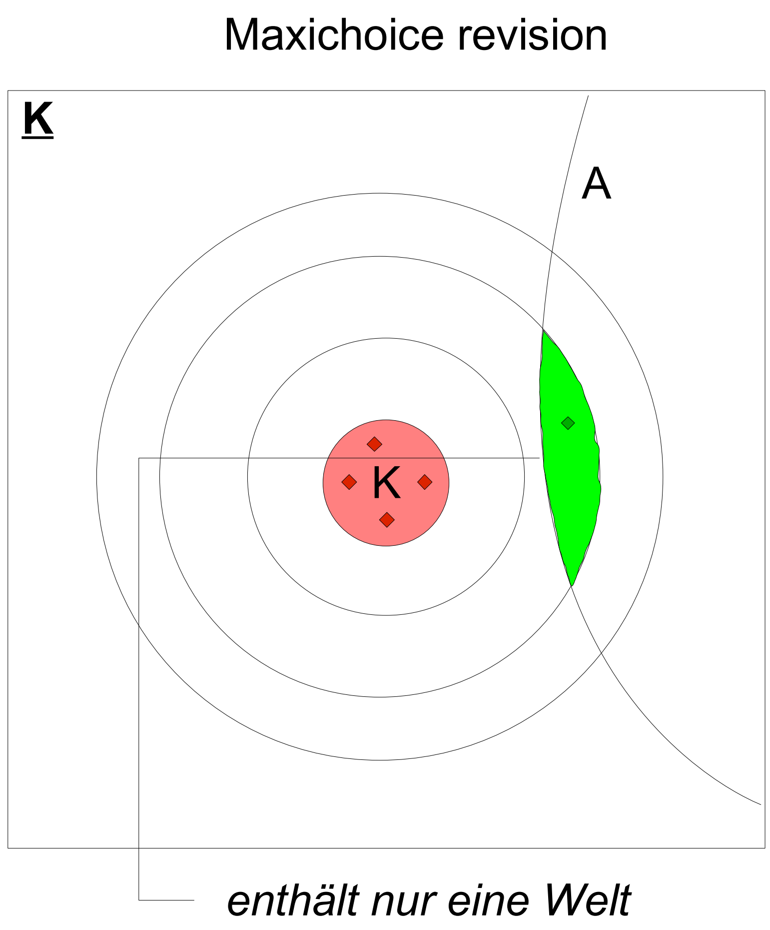
Pela Id. Harper obtemos uma representação análoga para contrações.

Full Meet e Maxichoice



Se $K * A = \mathrm{Cn}(\bigcap(K \perp \neg A) \cup A)$, nenhuma escolha em $K \perp \neg A$ está feita (quer dizer: tudo está sendo "escolhido").

O correspondente sist. de esferas é trivial: $K * A = \mathrm{Cn}(A)$ é pequeno demais!



Se $s(K \perp \neg A)$ contém somente uma teoria maximamente consistente, M, então $K * A = M + A$ também é maximamente consistente, i.e. equivale a um "mundo" só: grande demais!

Isto corresponde a exigência de que cada X-intersecção com uma esfera, $S_X \cap X$, contenha um elemento só. (Análogo à semântica do Stalnaker para os condicionais contrafatuais.)

Esferas e restos (remainders)

Os resultados de representação mostram que há “uma certa correspondência” entre sistemas de esferas e a construção em termos de intersecções parciais.

Já demonstramos isto indiretamente atraves dos postulados. Mas tem uma via mais direta:

- Há uma *bijecção* entre $[\neg A]$ e $K \perp A$. (Consideraremos somente o caso principal: $A \in K$ e A é consistente.)

(Para $X \subseteq W$ seja $|X| = \{A : \forall y \in X : y \models A\}$; $\exists!$ abrevie “existe um único ...”)

  - Das esferas aos restos:
    Para todo $x \in [\neg A]$ temos: $|[K] \cup \{x\}| \in K \perp A$.
  - Dos restos às esferas:
    Para todo $H \in K \perp A$ temos: $\exists! x \in [\neg A]$ tal que $x \in H$.

- Uma esfera$_K$ determina uma escolha em $[\neg A]$ sse uma função de selecção$_K$ faz uma escolha em $K \perp A$.

- Assim existe para cada função de selecção um sistema de esferas tal que as escolhas efetuadas para determinar uma revisão ou contração são as mesmas — e *vice versa*.

AGM e EPDL

$$A \in K * B :$$

Depois o agente revisa K para incluir B, (ele aceita (A))

Forma lógica:

- Proposição A está
- no escopo de um operador epistémico, que está
- no escopo de um operador de agência, que está parametrizado por uma proposição B.

Inidica a combinação seguinte:

- Lógica Epistêmica (ou Doxástica) (com operador **B** de aceitação, "belief") e
- Lógica Dinámica (com "programas" $[a_0], [a_1], [a_2], \ldots$)
- cada programa a_i da forma $*\langle$fórmula$\rangle$ (i.e. revisão) ou $-\langle$fórmula$\rangle$ (i.e. contração).

Lógica Dinámica (resumo rápido)

Lógica multimodal

$$[a_i] \ (i \in \mathbb{N})$$

$$R_i \subseteq W \times W$$

Lógica Dinámica Elementar (EPDL=Elementary Propositional Dynamic Logic)

Cada $[a]$ uma açao / um programa (= L. multimodal sob uma certa interpretação)

$$[a]B:$$

cada vez a termina, B é o caso. ($\langle a \rangle B$: as vezes a termina com B sendo o caso. Suporemos que as ações são deterministas; então $\langle a \rangle = [a]$.)

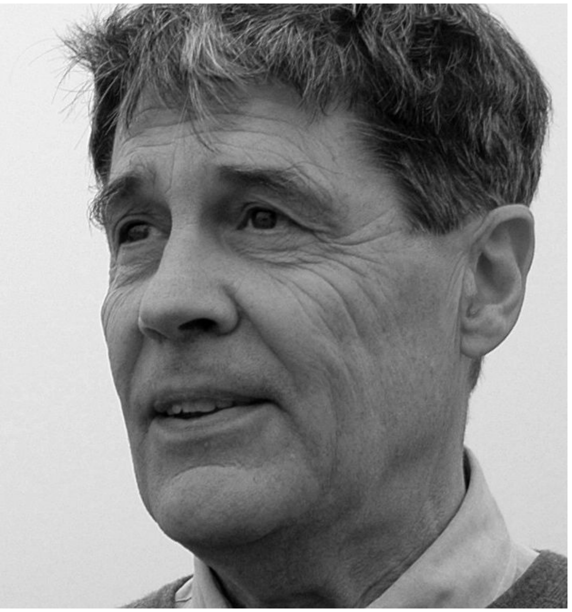
(**Lógica Dinámica** *senso strito*: programas complexos.

- sequências: $[a;b]C$ – b seguido por a resulta em C.
- escolha: $[a \cup b]C$ – ambos, a e b resultam em C.
- loop: $[{}^*a]C$ – Qq repetição (finita) de a resulta em C.

Somente loop acresce mais complexidade. Sem loop ficamos nos limites de EPDL.)

Lógica Dinâmica Doxástica (DDL)

... the driving idea of DDL is that formulæ such as $[*A]B$ are used to express doxastic actions on the same linguistic level on which also the arguments and the outcomes of these doxastic actions are expressed. Since DDL is furthermore motivated by the strong analogy of the operators $[*A]$ with standard modal operators, the intended semantics of DDL will be a possible worlds semantics ... Leitgeb & Segerberg (2007), 169.



Primeiras pesquisas: Fuhrmann (1988, 1989), van Eick (199?), van Benthem (1994).

Vantagens de DDL em comparação com AGM:

- DDL mais expressivo que AGM.
  - Exemplo: $\mathbf{B}[*A]\mathbf{B}B$ — $(B \in K * A) \in K$??
- DDL indica melhoa a estrutura lógica de mudanças de crenças.
  - $A \in K * B$ sugere que K é um parámetro que pode variar tal como A ou B. Na verdade, esta possibilidade é vazia em AGM: K sempre está mantido fixo.
- DDL oferece mais controle em deduções.
  - Por ex. as regras lógicas para $\mathbf{B}$ também aplicam a fórmulas como $\mathbf{B}[*A]\mathbf{B}B$.
- DDL permite transferir teorias já prontas da área da Lógica Modal.

Cuidado:

- Como o teste do Ramsey

$$A > B \in K \Leftrightarrow B \in K * A$$

  trivializa a teoria de AGM

- nenhum esquema da forma

$$\mathbf{B}C(A, B) \leftrightarrow [*A]\mathbf{B}B$$

  (para *qq* A e B) deve sair como teorema de DDL — presupondo que queremos translaçoes de Preservation e Consistency em DDL. ($C(A, B)$: uma fórmula C com subfórmulas A e B.)

A DDL mais simples

(Leitgeb & Segerberg 2007.)

- Somente crenças "fatuais".
- Mudanças somente a respeito de sentenças fatuais.

Sintaxe

Sejam ATM as sentenças atômicas de uma língua, FKT as sentenças fatuais, FML as fórmulas, definidas assim:

1. ATM $\subseteq$ FKT.
2. Se $A, B \in$ FKT, então toda combinaçao booleana de A e B está em FKT.
3. FKT $\subseteq$ FML.
4. Se $A \in$ FKT, então $\mathbf{B}A, \mathbf{K}A \in$ FML.
5. Se $A \in$ FKT e $B \in$ FML, então $[A]B \in$ FML.
6. $[-A]B := B \wedge [*\neg A]B$ ("Harper").

Observação: O *operador de saber* $\mathbf{K}$ ("know") representa crenças irrevisáveis (verdades lógicas, semánticas e talvez mais: o **urcorpus**)

Axiomas

A extensão mínima da Lógica Proposicional (clássica) t.q.
($\Box \in \{\mathbf{B}, \mathbf{K}, \{[*A] : A \in \mathrm{FKT}\}\}$, $\langle *A \rangle := \neg[*A]\neg$)

$$\Box(A \wedge B) \leftrightarrow (\Box A \wedge \Box B)$$

$$\Box\top$$

$$\frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B}$$

Ferner:

Closure	$\mathbf{K}(A \to B) \to ([*C]\mathbf{B}A \to [*C]\mathbf{B}B)$ (redundante)
Success (r2)	$[*A]\mathbf{B}A$
Inclusion (r3)	$[*\top]\mathbf{B}A \to \mathbf{B}A$
Preservation (r4)	$\neg\mathbf{B}\bot \to (\mathbf{B}A \to [*\top]\mathbf{B}A$
Consistency (r5)	$[*A]\mathbf{B}\bot \to \mathbf{K}\neg A$
Congruence (r6)	$\mathbf{K}[A \leftrightarrow B] \to [*A]\mathbf{B}C \leftrightarrow [*B]\mathbf{B}C$
Supp.1 (r7)	$[*(A \wedge B)]\mathbf{B}C \to [*A]\mathbf{B}(B \to C)$
Supp.2 (r8)	$\neg[*A]\mathbf{B}\neg B \to ([*A]\mathbf{B}(B \to C) \to [*(A \wedge B)\mathbf{B}C])$

Mais:

(Função)	$\langle * A\rangle B \leftrightarrow [* A]B$
(KB)	$\mathbf{K}A \rightarrow \mathbf{B}A$
(K*K)	$\mathbf{K}A \leftrightarrow [* B]\mathbf{K}A$

Chamaremos o sistema tal definido de **BDD** ("Basis Doxastic Dynamics").

[AGM-Postulate]

Semântica

Ideia simples: **BDD** é uma lógica multimodal. Assim esperamos para cada operador $\Box_i$ uma interpretação da forma:

$$x \models \Box_i A \text{ sse } \forall y : R_i xy \Rightarrow y \models A$$

No caso que nos interesse, $\Box$ tem estrutura: $[*B]$.

- Esta estrutura deve ter um impacto a determinar os teoremas de **BDD**;
- segue-se que a relação R que acompanha $\Box$ deve representar esta estrutura em alguma maneira:
$$x \models [*B]A \text{ sse } \forall y : R_{[B]} xy \Rightarrow y \models A,$$
onde $[B]$ é a representação semántica da fórmula B.
- *Solução*: Adaptamos os sistemas de esferas que já conhecemos.

Seja W (conjunto dos mundos possíveis) fechado sob união e intersecção finita.

Sistema de esferas $\mathfrak{S}$ baseado em W: uma família de conjuntos em W $[\neq \emptyset]$ t.q.

1. se S é um conjunto não-vazio em $\mathfrak{S}$, então $\bigcap S \in \mathfrak{S}$;
2. $\mathfrak{S}$ está ordenado completamente (linear) sob $\subseteq$ (i.e. $\forall S_1, S_2 \in \sigma$: ou $S_1 \subseteq S_2$ ou $S_2 \subseteq S_1$).
3. $\forall P \subseteq W$: se $P \neq \emptyset$, então existe uma esfera miníma $S \in \mathfrak{S}$ t.q. $P \cap S \neq \emptyset$ (notação $\min(\mathfrak{S} \bullet P)$, veja abaixo).

Estrutura de revisão $(W, \Sigma, R_{P:P\subseteq W})$:

- $W \neq \emptyset$,
- conjunto Σ de esferas em W,
- para cada $P \subseteq W$ ("proposição") uma relação $R_P \subseteq \Sigma \times \Sigma$.

- Condições:

  Seja $\mathfrak{S} \bullet P = \{S \in \mathfrak{S} : S \cap P \neq \emptyset\}$
  (o conunto dos esferas-P; se $\mathfrak{S} \bullet P = \emptyset$, então P in $\mathfrak{S}$não é "concebível", i.e. a negação de P faz parte do urcorpus).

1. Se $R_P \mathfrak{S}\mathfrak{S}'$, então $\bigcap \mathfrak{S}' = \begin{cases} P \cap \min(\mathfrak{S} \bullet P), & \text{caso } (\mathfrak{S} \bullet P) \neq \emptyset \\ \{\emptyset\} & \text{caso contrário .} \end{cases}$ (revisão)
2. $\bigcup \mathfrak{S} = \bigcup \mathfrak{S}'$ $(\forall \mathfrak{S}, \mathfrak{S}' \in \Sigma)$. (urkorpus é constante (**K**).)
3. $\forall \mathfrak{S} \exists \mathfrak{S}' : R\mathfrak{S}\mathfrak{S}'$. (revisões são quase sempre possíveis.)
4. Se $R\mathfrak{S}\mathfrak{S}'$ und $R\mathfrak{S}\mathfrak{S}''$, então $\mathfrak{S}' = \mathfrak{S}''$. (revisões são bem-definidas.)

[Modelos do Grove]

Modelo: Estrutura de revisão com valuação V das fórmulas atômicas.

- **Verdade** num modelo (relativo a um sist. de esferas num mundo):

$$\mathfrak{S}, w \models P \text{ sse } w \in V(P),$$
$$\text{usw. für } \neg, \wedge, \text{ etc.}$$
$$\mathfrak{S}, w \models \mathbf{B}A \text{ sse } \bigcap \mathfrak{S} \subseteq [A],$$
$$\mathfrak{S}, w \models \mathbf{K}A \text{ sse } \bigcup \mathfrak{S} \subseteq [A],$$
$$\mathfrak{S}, w \models [*A]B \text{ sse } \mathfrak{S}', w \models B, \text{ se } R_{[A]}\mathfrak{S}\mathfrak{S}'.$$

- **Validade** num modelo: Verdade relativo aos todos os pares $(\mathfrak{S}, w)$.

TEOREMA (SEGERBERG 2005). *Uma fórmula A é derivável em (teorema de)* **BDD** *sse A é válido em todos os modelos.*