

2017 年度
哲学若手研究者フォーラム個人研究発表

認識状態のモデルを応用した「なぜ-疑問」モデルの構築
—科学哲学における2つの説明理論を手がかりに—

大阪大学大学院 人間科学研究科

小川文紀

[発表者]

- ・氏名：小川文紀
- ・所属：大阪大学大学院 人間科学研究科 (博士前期課程2年)
- ・専門：分析形而上学 [主に「なぜ何もないのではなく何かがあるのか？」という問いを研究]
- ・なぜ-疑問研究のモチベーション：なぜ-疑問モデル (理論) を用いて上記の問いにアプローチしたい

発表要旨

目的：なぜ-疑問モデルの構築

本発表の目的は、なぜ-疑問文によって表現される「なぜ-疑問 (Why-question)」のモデル化である。すなわち、なぜ-疑問という抽象的対象がどのような構造を持つのかを明らかにする (図 1)。



図 1: なぜ-疑問文によって表現されるなぜ-疑問

結論：なぜ-疑問モデルは2つの時点の認識状態モデルから成る

「なぜ E なのか？」によって表現されるなぜ-疑問のモデル Q' [$Q' = \langle K, K_E^- \rangle$] は、 E を学んだ直後の (質問者の) 認識状態モデル K [$K = \langle W, \{P_w \mid w \in W\}, B \rangle$] と、 E を学ぶ直前の (質問者の) 認識状態モデル K_E^- [$K_E^- = \langle W_E^-, \{P_w \mid w \in W_E^-\}, B_E^- \rangle$] の2つの要素から成る (図 2)。

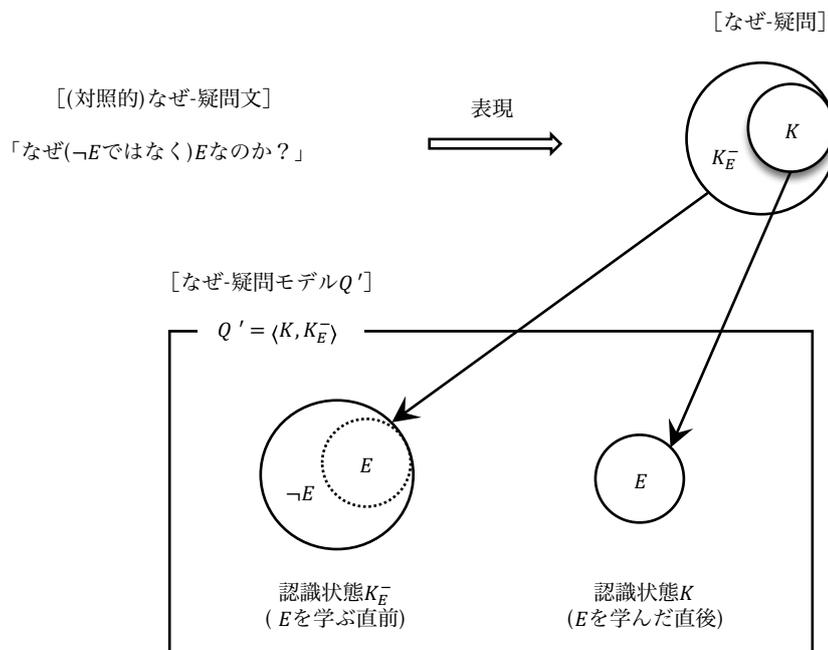


図 2: なぜ-疑問モデル Q' の構造

方法：2つの説明理論の応用

2つのプラグマティックな説明理論、(1) 科学哲学者の Bas van Fraassen の説明理論 (van Fraassen 1980: ch. 5) と、(2) 認知科学者の Peter Gärdenfors の説明理論 (Gärdenfors 1988: ch. 8) を応用する。説明理論とはどのような場合に説明が成り立ちまた成り立たないのかを明らかにする理論であるが、彼らはそうした説明の成否が文脈に左右されると考えるため、両者の説明理論には文脈的要素が組み込まれている。

こうした共通点を通じて (1) の文脈的要素を (2) の文脈的要素に置き換える。具体的には、(1) で用いられるなぜ-疑問モデル $Q [Q = \langle P_k, X, R \rangle]$ の要素であるトピック P_k と対照クラス X に、(2) で用いられる認識状態モデル $K [K = \langle W, \{P_w | w \in W\}, B \rangle]$ と認識状態モデル $K_E^- [K_E^- = \langle W_E^-, \{P_w | w \in W_E^-\}, B_E^- \rangle]$ をそれぞれ移植し、関連性関係 R を切除する (図3)。これにより、なぜ-疑問モデル $Q' [Q' = \langle K, K_E^- \rangle]$ を得る。

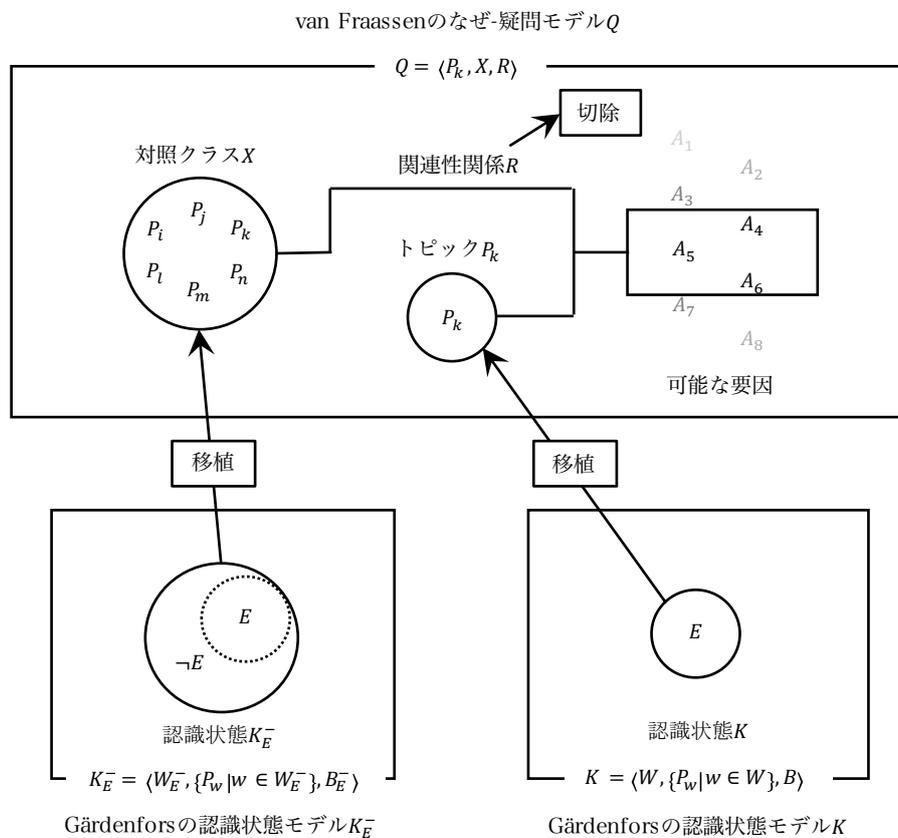


図3: なぜ-疑問モデルへの認識状態モデルの移植

先行研究

van Fraassen の説明理論

—説明＝なぜ-疑問の答え—

van Fraassen は説明とはなぜ-疑問の答えであり、説明の理論はなぜ-疑問の理論でなければならないと主張する (van Fraassen 1980: 134, 156)。そのため、彼の説明理論は説明となるかどうかをなぜ-疑問の答えとなるかどうかによって評価するなぜ-疑問の理論となった。

—なぜ-疑問 (説明の要求) の文脈依存性—

なぜ-疑問の答えとなるかどうかは、問われているなぜ-疑問と相対的に評価しなくてはならない (*Ibid.*: 156)。だとすれば、答えの評価を行うために、まずはどのようななぜ-疑問が問われているのかを特定する必要がある。彼によると「なぜ-疑問とは説明の要求である」(*Ibid.*: 126, 156)。したがってどのようななぜ-疑問が問われているのかを特定するには、そのなぜ-疑問がどのような説明を要求しているのかを特定しなければならない。

ところが全く同じなぜ-疑問文「なぜ P なのか？」によって問われているにもかかわらず、文脈が異なることで要求する説明が違ってくることが往々にして起こる (*Ibid.*: 156)。彼はそうした事例を説明に関する過去の議論から拾い集めている。その過去の議論とは、(1) 説明における因果の議論 (*Ibid.*: ch. 5, §2.5-7) と、(2) 説明におけるなぜ-疑問の議論 (*Ibid.*: ch. 5, §2.8) である。

(1) 説明における因果の議論 (関連性関係 R)

彼は説明における因果性に関して、David Lewis の反事実的条件法による因果の分析 (*Ibid.*: ch. 5, §2.5) と、Wesley Salmon による因果的メカニズムモデルを概観 (*Ibid.*: ch. 5, §2.6) した後、説明をおおよそ次のように捉え直す: 各々の出来事は科学が記述する因果の網目に置くことができ、なぜある出来事が起こったのかの説明とは、その出来事へと至る数多くの要因の中から「**際立った要因**」を示すことである (*Ibid.*: 124)。しかし彼によると何が際立った要因となるかは文脈に依存する。例えば、ある交通死亡事故の原因について、医者は「出血多量」、弁護士は「ドライバーの不注意」、技術者は「ブレーキの欠陥」、都市計画家は「高い植え込み」と答えるだろう。つまり何が際立った要因となるかは、その人の志向や関心といった文脈的なものに依存している *¹(*Ibid.*: 125)。

彼はさらに、これらの際立った要因は組み合わせることはできないと注意を促す。技術者は「周囲の環境を固定して」視界を遮る植え込みがあってもブレーキが良ければ事故は起こらなかったと考え、都市計画家は「車の構造を固定して」ブレーキに欠陥があっても植え込みが視野を遮らなければ事故は起こらなかったと考えている。つまり志向や関心による各々の際立った要因の選択の背後には、類似性をどのように捉えるかについて「対立する反事実的条件法」があるのである (*Ibid.*: 125)。

*¹ここで、なぜ-疑問 (説明の要求) の観点から少々補足しておきたい。ある「交通死亡事故」を因果の網目に置いた時、そこへ至る諸要因はたくさんある: 「出血多量」、「ドライバーの不注意」、「ブレーキの欠陥」、「タイヤの劣化」、「シートベルトの誤作動」、「高い植え込み」など。説明とはこれらの諸要因からどれか1つの要因を示すことである。さて、技術者に「なぜこの交通死亡事故が起こったのか？」と聞けば、彼は自分の志向や関心からその問いによって**技術的**な説明が求められていると理解するだろう。つまり彼にとって「なぜこの交通死亡事故が起こったのか？」という問いは、先述の**全ての要因** (「出血多量」、「ドライバーの不注意」、「ブレーキの欠陥」、「タイヤの劣化」、「シートベルトの誤作動」、「高い植え込み」) の中からどれか1つの要因を選ぶことを要求するものではなく、**技術的**な要因 (「ブレーキの欠陥」、「タイヤの劣化」、「シートベルトの誤作動」) の中からどれか1つの要因を選ぶように要求するものなのである。彼はそうして選出された要因を「**際立った要因**」ないし「**可能な要因** (特定の観点 (志向・関心) から見て説明 (答え) として挙げる事が可能な要因)」と呼ぶ。なぜ-疑問の役割はこうした「際立った要因」を選択するところまでである。技術者は技術的要因の中から最終的に「ブレーキの欠陥」を説明 (答え) として選択したが、そうした「際立った要因」の中からの選択はなぜ-疑問の行うところではなく、それとは別に「答えの評価」の中で行なわれる。

(2) 説明におけるなぜ-疑問の議論 (対照クラス X)

彼は説明におけるなぜ-疑問に関して、Bengt Hansson のなぜ-疑問の研究を手がかりに次のような事例を挙げる: 「なぜアダムはリンゴを食べたのか?」という疑問は、「なぜアダムはリンゴを(イヴに返すのではなく)食べたのか?」と「なぜアダムは(ナシではなく)リンゴを食べたのか?」の2通りに解釈可能であるが、「空腹だったからだ」という答えは前者にとっては答えでも後者にとっては答えとならない (*Ibid.*: 127)。これら2通りの解釈は要求する説明(答え)が違っているのであるが、その違いは明らかに、説明を求めている出来事の代わりに起こっていたかもしれない出来事(対照的選択肢)をどのように設定するのかに起因している。例えば後者の解釈では、「アダムがリンゴを食べた(P_k)」の代わりに「アダムがナシを食べた(P_i)」や「アダムがミカンを食べた(P_j)」などが対照的選択肢となるだろう。彼はこうした選択肢の集合($\{P_k, P_i, P_j\}$)を対照クラス X と名付けた^{*2}(*Ibid.*: 127)。そしてどのような対照クラスになるかもまた、文脈に依存すると主張する (*Ibid.*: 129)。

なぜ-疑問は説明の要求であるので、説明の要求が異なれば、それは異なるなぜ-疑問である。そのため彼は、なぜ-疑問の一般的形式は、単に「なぜ P_k なのか?」なのではなく、「なぜ X (他のメンバー)と対照的に P_k であるのか?」であると主張する (*Ibid.*: 127)。しかし対照クラス X は文脈の上では明らかであるので、一般的に明示化されることはないとされる (*Ibid.*: 128)。

—なぜ-疑問の評価と答えの評価の文脈依存性—

なぜ-疑問の答えとなるかどうかは問われているなぜ-疑問と相対的に評価しなくてはならない。そのために、まずどのようななぜ-疑問が問われているのかを特定する必要があった。ではなぜ-疑問の特定がされたあと答えはどのように評価されるのか。彼は Michael Friedman の説明理論をもとに (*Ibid.*: ch. 5, §2.3)、説明(答え)とされるものは仮定された背景 K と相対的に評価されなければならないと考える (*Ibid.*: 110)。「背景 K と相対的に評価する」とは、背景 K + 答え A が説明すべき事実 P を含意するかといったことを調べることだが、背景 K は通常すでに P を含意するため、その場合には P を含意しないよう背景 K の部分である $K(Q)$ を用いて評価を行わなくてはならない (*Ibid.*: 147)。彼はこの背景 K と K の部分である $K(Q)$ の両方を用いて様々な評価を行う。評価項目は大別して2つ、(3) そもそもそのなぜ-疑問が生じるかどうかの評価と (*Ibid.*: 145-6)、(4) 答えがより良いものかどうかの評価である (*Ibid.*: ch. 5, §4.4)。

(3) そもそもそのなぜ-疑問が生じるかどうかの評価 (背景 K)

なぜ-疑問が立てられた文脈においては、質問者とそれを聞いている人によって受け入れられている背景理論と背景情報から成る背景 K がある。この背景 K もまた、質問者とそれを聞いている人が誰なのかに依存するため文脈的なものである。そしてなぜ-疑問が生じるかどうかは、この背景 K がそこで問われたなぜ-疑問の前提を適切に含意するかどうかによって定まるとされる (*Ibid.*: 145-6)。

(4) 答えがより良いものかどうかの評価 (背景 K の部分 $K(Q)$)

すでに述べた通り、背景 K + 答え A が説明すべき事実を含意するかどうかによって答え A を評価する場合、説明すべき事実(トピックが真であることと対照的選択肢が偽であること)を含意しないよう背景 K の部分である $K(Q)$ を用いなくてはならない。しかしこの $K(Q)$ が背景 K のどの部分であるのかということもまた文脈に依存するとされる (*Ibid.*: 147)。彼はこうした $K(Q)$ を用いて、 $K(Q)$ + 答え A が説明すべき事実(トピックが真であることと対照的選択肢が偽であること)を含意するかどうか、または確率的に $K(Q)$ のもとでの説明すべき事実の確率と $K(Q)$ + 答え A のもとでの説明すべき事実の確率の比較を行うことで、その答え A がより良いものであるかどうかを評価する (*Ibid.*: 147-8)。

^{*2}この段階では、 P_k は対照クラス X に属してもよいし属さなくてもよいとされる。

—なぜ-疑問の理論—

van Fraassen のなぜ-疑問の理論は、説明における上記 (1) から (4) の文脈的側面を反映したものとなるが、その形式化にあたっては Nuel Belnap の疑問の論理に依拠しており、「直接的な答え」や「疑問の前提」などの重要な諸概念はそこから得られている (*Ibid.*: ch. 5, §4.2)。彼のなぜ-疑問の理論の概要は以下のようになる: (A) なぜ-疑問の特定 [モデル化・直接的な答え・前提] (*Ibid.*: 141-5)、(B) なぜ-疑問の評価 [なぜ-疑問の発生] (*Ibid.*: 145-6)、(C) 答えの評価 [答えの良さ] (*Ibid.*: ch. 5, §4.4)。

(A) なぜ-疑問の特定 (モデル・直接的な答え・前提)

モデル なぜ-疑問文「なぜ (X の他のメンバーとは対照的に) P_k なのか?」によって表現されるなぜ-疑問 Q は、3つ組 $\langle P_k, X, R \rangle$ と同一視される。すなわち $Q = \langle P_k, X, R \rangle$ である。ただし、トピック P_k は命題 (説明が求められている出来事) であり、対照クラス X [$X = \{P_1, \dots, P_k, \dots\}$] は命題のクラス (選択肢の集合) であり、関連性関係 R は P_k と X の対 $\langle P_k, X \rangle$ と命題 A (可能な要因) の間の関係とする^{*3} (*Ibid.*: 141-3)。これによって、なぜ-疑問 Q の文脈を反映したモデル化がなされる。

直接的な答え 直接的な答えは以下の文形式を持つ: 「 A なので X (の残り) と対照的に P_k である」。この表現を持つ命題を B とすると、直接的な答えは以下のように定義される: 命題 B がなぜ-疑問 Q の直接的な答えである $\stackrel{\text{def}}{\iff}$ 対 $\langle P_k, X \rangle$ と関係 R を持つ命題 A があり、かつ B が真 iff (P_k , かつ $i \neq k$ である全ての i について $\neg P_i$, かつ A) が真 (*Ibid.*: 143-4)。この定義によりなぜ-疑問 Q の直接的な答えの集合が定まり、なぜ-疑問 Q の特定がなされる。またこの直接的な答えを用いてなぜ-疑問 Q の前提が得られる。

前提 なぜ-疑問 Q の前提は以下の3つである: (i) P_k が真、(ii) X の中で P_k のみが真、(iii) $\langle P_k, X \rangle$ と R を持つ少なくとも一つの A が真。そして (P_k , かつ $i \neq k$ である全ての i について $\neg P_i$) という命題は、なぜ-疑問 Q の中心的前提とされる (*Ibid.*: 144-5)。これにより、なぜ-疑問 Q が生じるかどうか評価される。

(B) なぜ-疑問の評価 (なぜ-疑問の発生)

なぜ-疑問の発生 なぜ-疑問 Q が立てられた文脈においては背景理論と背景情報から成る背景 K がある。この時、その文脈においてなぜ-疑問 Q が発生する iff (K が中心的前提を含意し、かつ K は全ての前提の否定を含意しない)。したがって、もし K が Q の中心的前提を含意しないか、またはいずれかの前提の否定を含意するならば、なぜ-疑問 Q は「拒絶」される (*Ibid.*: 145-6)。

(C) 答えの評価 (答えの良さ)

答えの良さ 答えの良さは以下の3点で評価される: (a) 答え A 自体が受け入れ可能か [背景 K が $\neg A$ を含意しないかどうか]、(b) 答え A がどのくらいトピック P_k を対照クラス X の他のメンバー $P_i (i \neq k)$ より有利にするか [$K(Q) +$ 答え A がトピック P_k と $\neg P_i (i \neq k)$ を含意するかどうか、あるいは $K(Q)$ 単独よりも $K(Q) +$ 答え A のもとの P_k と $P_i (i \neq k)$ の各々への確率配分が、どのくらい $P_i (i \neq k)$ の各々の確率よりも P_k の確率を上げているか]、(c) 答え A と他の可能な答えとの3つ観点からの比較 [(c-1) 背景 K において答え A の確率が他の可能な答えよりも高いか、(c-2) 答え A が他の可能な答えよりもトピック P_k を有利にしているか、(c-3) A が他の可能な答えによって関連を失わない (濾過されない) か] (*Ibid.*: ch. 5, §4.4)。これらの条件をもとに、なぜ-疑問の答えが評価される。

—van Fraassen の説明理論 (なぜ-疑問の理論) による分析の一例—

次項はなぜ-疑問の理論による技術者の事例の分析の一例である (図 4)。

^{*3}先の脚注で述べた通り、技術者にとっては「ブレーキに欠陥があること」、「タイヤの性能が悪かったこと」、「シートベルトが誤作動したこと」は説明 (答え) となりうる要因だが、「出血多量があったこと」、「ドライバーに不注意があったこと」、「植え込みが高かったこと」は説明 (答え) となりうる要因ではない。関連性関係 R はこうした説明における文脈的側面を反映したものである。つまり関連性関係 R は、なぜ-疑問 Q に対しどのような種類の要因を説明 (答え) として挙げてよいのか (説明的関連性) を外延的に定める関係である。そしてそうした関係を持つ諸要因を彼は「可能な答え」とも呼ぶ。

「なぜこの交通死亡事故が起こったのか？」

(A)なぜ-疑問の特定

トピック $P_k = P_1$

$$Q = \left\{ \begin{array}{l} \text{対照クラス } X = \{P_1, \neg P_1\} \\ \text{関連性関係 } R = \left\{ \begin{array}{l} \langle \langle P_1, \{P_1, \neg P_1\} \rangle, A_1 \rangle, \\ \langle \langle P_1, \{P_1, \neg P_1\} \rangle, A_2 \rangle, \\ \langle \langle P_1, \{P_1, \neg P_1\} \rangle, A_3 \rangle \end{array} \right\} \end{array} \right\}$$

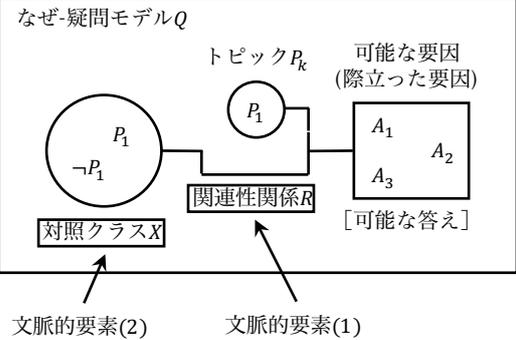
同一視

$$Q \text{ の直接的な答えの集合} = \left\{ \begin{array}{l} A_1 \text{ だから } \neg P_1 \text{ ではなく } P_1, \\ A_2 \text{ だから } \neg P_1 \text{ ではなく } P_1, \\ A_3 \text{ だから } \neg P_1 \text{ ではなく } P_1 \end{array} \right\}$$

含意

Qの前提: (i) P_1 (ii) $\neg P_1$ (iii) $A_1 \vee A_2 \vee A_3$
Qの中心的前提: $P_1 \wedge \neg \neg P_1$

P_1 : この交通死亡事故が起こった
 A_1 : ブレーキに欠陥があった
 A_2 : タイヤが劣化していた
 A_3 : シートベルトが誤作動した



(B)なぜ-疑問の評価

背景K: 背景理論T + 背景情報C

$$\begin{array}{l} K \models P_1 \wedge \neg \neg P_1 \\ K \not\models \neg P_1, \neg \neg P_1, \neg (A_1 \vee A_2 \vee A_3) \end{array} \quad \left[\begin{array}{l} \text{なぜ-疑問} Q \\ \text{発生} \end{array} \right]$$

(C)答えの評価

(a) 答え A_1, A_2, A_3 が受け入れ可能か

$$K \models \neg A_3 \quad K \not\models \neg A_1, \neg A_2 \quad \left[\text{答え } A_3 \text{ 脱落} \right]$$

(b) 答え A_1, A_2 が $\neg P_1$ より P_1 を有利にするか

$K(Q)$: 背景Kの部分 (ただし、 $K(Q) \not\models P, \neg P$)

$$K(Q), A_1 \models P_1, \neg \neg P_1 \quad K(Q), A_2 \not\models P_1, \neg \neg P_1$$

答え A_2 の再調査
ただしPは確率関数

$$\left(\begin{array}{l} P(P_1 | K(Q) \wedge A_2) = 0.8 \\ P(\neg P_1 | K(Q) \wedge A_2) = 0.2 \end{array} \right)$$

(c) 答え同士の比較

(c-1) 背景Kにおける答え A_n の確率

$$P(A_1 | K) = 1 \quad P(A_2 | K) = 0.7 \quad [P(A_1 | K) = 0]$$

(c-2) どの答えが $\neg P_1$ より P_1 をさらに有利にしているか

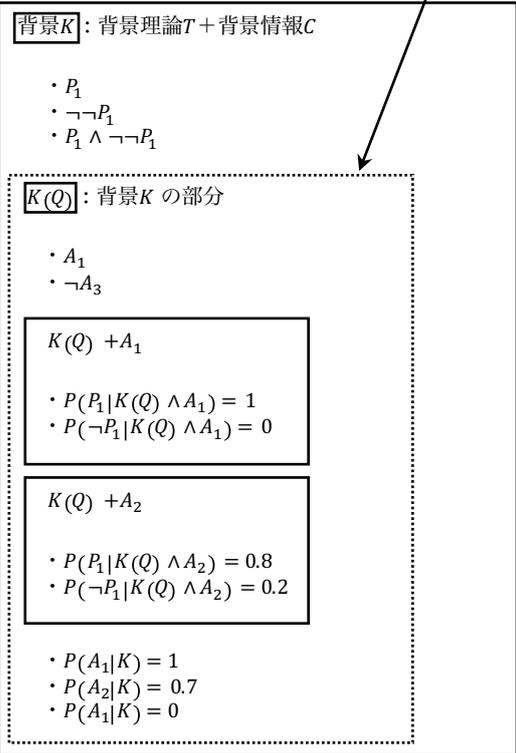
$$A_1[1] > A_2[0.6] \quad \left[\text{ただし } [x] \text{ の } x \text{ は何らかの方法で算出} \right]$$

(c-3) 他の答えによって関連を失わないか(濾過しないか)

$$A_2 \text{ が } A_1 \text{ を } P_1 \text{ から濾過しない} \quad \left[\text{答え } A_2 \text{ 脱落} \right]$$

文脈的要素(3)

文脈的要素(4)



「なぜならブレーキに欠陥があったからだ」

図 4: なぜ-疑問の理論による技術者の事例の分析の一例

Gärdenfors の説明理論

— 説明 = 驚きを減らすこと —

Gärdenfors は説明とは驚きを減らすことであると主張する。彼によると「なぜ E なのか？」と説明を求めるのは、予期していなかった出来事 E が起こって驚いたからである。そしてそれに対する説明 $T&C$ は、明らかに E に対する驚きを減らす効果を持っている (Gärdenfors 1988: 167, 176, 178)。彼はこの考えに基づき、独自の認識状態モデルである「認識状態の二階確率モデル」*4 (Second-Order Probabilistic Models of Epistemic States) と AGM 理論を用いて説明理論を構築した。

— 認識状態の二階確率モデル K (対象言語・構造・定義) —

対象言語 対象言語は一階言語。事実を表わす単称文 (e.g., Qa) と一般的知識や帰納的信念を表わす確率文 (e.g., $p(Qx/Rx) \geq 0.5$) のみ。量化文 (e.g., $\forall xQx$) は排され、確率文 (e.g., $p(Qx) = 1$) で代替する。

構造 認識状態 K は $\langle W, \{P_w \mid w \in W\}, B \rangle$ と同一視される $[K = \langle W, \{P_w \mid w \in W\}, B \rangle]$ *5。ただし、可能世界の集合 W は同じ個体集合を持った可能世界の集合、確率測度 P_w は W の全ての可能世界 w に対して w の個体集合のベキ集合上で定義される確率測度、信念関数 B は W のベキ集合上で定義される確率測度である。なお、所与の可能世界 w に対し、個体定項と述語記号が指示する w における個体と個体集合を伝える解釈関数 I が仮定される (e.g., 個体定項 a なら $a_w = I(a)$, 述語記号 Q なら $Q_w = I(Q)$)。

定義 [真理の定義] 原子文の真理の定義は以下の通り: 単称文 Qa が w で真 $\stackrel{\text{def}}{\iff} a_w \in Q_w$ 、確率文 $p(Qx/Rx) \geq r$ が w で真 $\stackrel{\text{def}}{\iff} P_w(Q_w/R_w) \geq r$ 。結合子の真理条件は通常なされる通り。この真理の定義により文 A の解釈が生成される ($I(A) = \{w \mid A \text{ が } w \text{ で真}\}$)。[信念の定義] 認識状態 K において文 A が知識として受け入れられている (信じられている) $\stackrel{\text{def}}{\iff} K$ の構成要素である可能世界の集合 W の全ての世界において A が真 (図 5)。(Ibid.: §8.3)

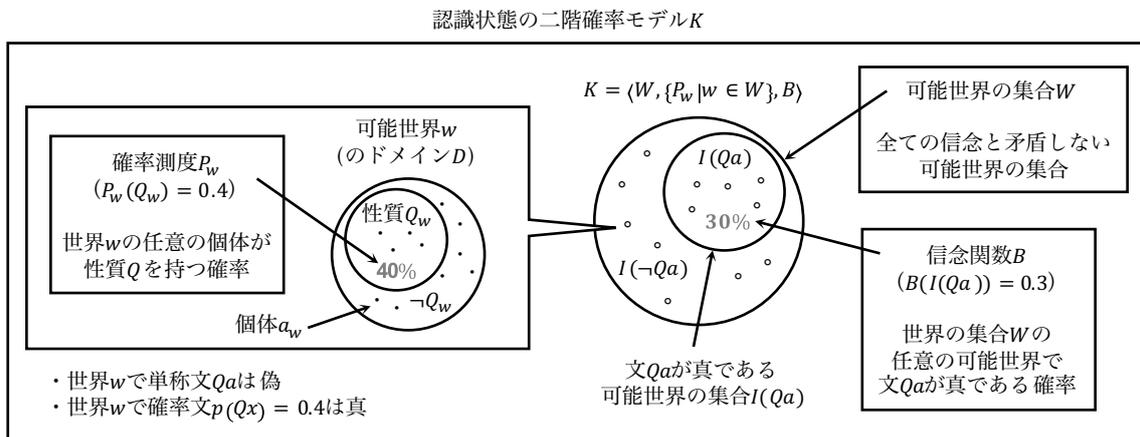


図 5: 認識状態の二階確率モデル K の一例

認識状態の二階確率モデルの大きな特徴の一つは、信念関数 B によって文に対する**信念の度合い** (信念値) を再現できる点にある。これにより、認識主体がどの程度出来事 E が起こると信じているのか、つまり認識主体が出来事 E に対して予測する可能性の高さを知ることができる。

*4彼の認識状態の二階確率モデルは、可能世界の集合 W のベキ集合上で定義される確率関数から成る確率的可能世界モデルの一種であり、ベイジアンモデルの一種である (Ibid.: 40)。

*5正確には、おそらく $K = \langle W, D, \{P_w \mid w \in W\}, \{I_w \mid w \in W\}, B \rangle$ のようになると思われる。

—AGM 理論—

AGM 理論を用いることで、認識状態の二階確率モデル K の信念変更後のモデルを得ることができる (*Ibid.*: 174-5)。しかしその場合、認識的コミットメント関数は通常のように K に対してではなく、その構成要素である可能世界の集合 W と信念関数 B に対して適用され、得られた W' と B' (および $\{P_w \mid w \in W'\}$) を構成要素とする認識状態 K' [$K' = \langle W', \{P_w \mid w \in W'\}, B' \rangle$] が信念変更後のモデルとなる^{*6}。

例えば、認識状態の二階確率モデル K に信念 E を加えた (i.e., K の E によるエクспанション) 信念変更後のモデル K_E^+ を得たいなら、 K の構成要素である W から信念 E と両立可能な可能世界の集合 W_E^+ を取り出し、同じく K の構成要素である B を用いて W_E^+ の全ての部分集合 V について $B_E^+(V) = B(V/W_E^+)$ と定義することで信念関数 B_E^+ を得て、そうして得られた W_E^+ と B_E^+ (および $\{P_w \mid w \in W_E^+\}$) を構成要素とする認識状態 K_E^+ [$K_E^+ = \langle W_E^+, \{P_w \mid w \in W_E^+\}, B_E^+ \rangle$] が信念変更後のモデルとなる。

以下の図 6 は認識状態の二階確率モデル K から信念 E を除去した (i.e., K の E によるコントラクション) 信念変更後のモデル K_E^- の構成の概略である^{*7}。

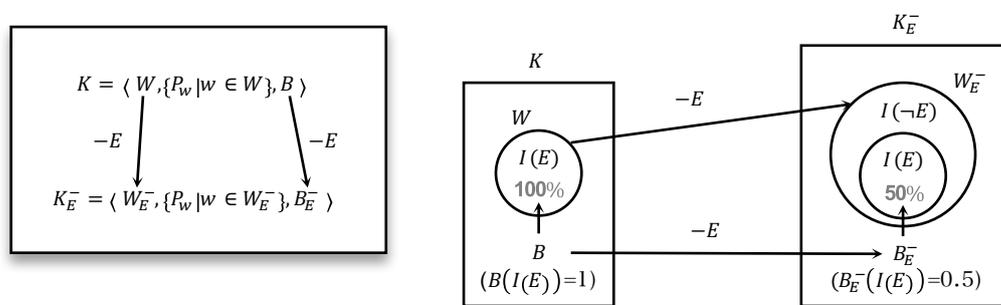


図 6: 認識状態の二階確率モデル K の E によるコントラクション

^{*6}念のため、どのようにして認識状態の二階確率モデル K から AGM 理論を利用して信念変更後のモデル K' が得られるのかを補足しておきたい。認識的コミットメント関数とは、認識状態と認識の入力を独立変数とし、認識状態を従属変数とする 2 変数関数である。AGM 理論では、認識状態に情報を加えたり引いたりした時にどのように認識状態が変化するか (信念変更) をこうした関数で捉える (*Ibid.*: ch. 1)。信念変更は基本的に 3 種類ある: エクспанション (e.g., $K \not\models A$ から $K \models A$)、コントラクション (e.g., $K \models A$ から $K \not\models A$)、リヴィジョン (e.g., $K \models \neg A$ から $K \models A$)。これら 3 種類の信念変更を認識的コミットメント関数で表す時、各々の信念変更が満たすべき関数の性質を各々の信念変更ごとに公理としてまとめたものが AGM 基本原理であり、そうした各々の公理を満たす関数は、それぞれエクспанション関数、コントラクション関数、リヴィジョン関数と呼ばれる。

認識状態モデルには、論理的帰結関係で閉じた文集合 K を認識状態と見なす信念集合 (*Ibid.*: §2.2) や、言語 L の文集合上に定義される確率関数 P を認識状態と見なすベイジアンモデル (*Ibid.*: §2.7) などがある (これらの文に解釈関数を適用すれば、可能世界の集合 W を認識状態と見なす可能世界モデル (*Ibid.*: §2.4) や、可能世界の集合 W のベキ集合上に定義される確率関数 P を認識状態と見なす確率的可能世界モデル (*Ibid.*: 40) となる)。信念集合の信念変更を認識的コミットメント関数で表す場合、その関数は文集合 K と認識的入力 A を独立変数とし、文集合 K を従属変数とする 2 変数関数となり、一方でベイジアンモデルの信念変更を認識的コミットメント関数で表す場合、その関数は確率関数 P と認識的入力 A を独立変数とし、確率関数 P を従属変数とする 2 変数関数となる。両モデルの本質的な違いはこうして信念変更にも及ぶため、それぞれ個別に信念集合用の AGM 基本原理 (*Ibid.*: ch. 3) とベイジアンモデル用の AGM 基本原理が (*Ibid.*: ch. 5) 用意されている (可能世界モデルや確率的可能世界モデルに対しては、基本的にそれぞれ解釈関数を外すことで信念集合やベイジアンモデルの AGM 基本原理を応用することができるが、このうち可能世界モデルに対しては、それとは別に Lewis の反事実的条件文のモデルである圏域システムを応用した Adam Grove の圏域システムを用いた信念変更のモデル化 (*Ibid.*: §4.5) がなされている)。

認識状態の二階確率モデル K から信念変更後のモデル K' を得るやり方は以下の通り: (1) 認識状態の二階確率モデル K の構成要素である W に対し、信念集合用の AGM 基本原理を満たす認識的コミットメント関数を利用して (または Grove の圏域システムを利用して) 信念変更後の W' を得て、(2) 認識状態の二階確率モデル K の構成要素である B に対し、ベイジアンモデル用の AGM 基本原理を満たす認識的コミットメント関数を利用して信念変更後の B' を得て、(3) 得られた W' と B' (および $\{P_w \mid w \in W'\}$) を構成要素とする認識状態の二階確率モデル K' [$K' = \langle W', \{P_w \mid w \in W'\}, B' \rangle$] を構成すれば、それが信念変更後のモデルとなる。

^{*7}コントラクション関数はコントラクションに関する AGM 基本原理を満たさなければならないが、コントラクションに関する AGM 基本原理はコントラクション関数を一意に定めるのに十分ではない (*Ibid.*: 61)。そのためどのコントラクション関数が実際に用いられるのかについては決定されず、したがってどのような信念変更後のモデルが得られるのかは分からない。その反面、必ずどれかのコントラクション関数が用いられなければならないので、何らかの信念変更後のモデルが得られなければならない。

— 驚きの形式化 —

彼の説明理論では説明となるかどうかは驚きが減っているかどうかで決まる。したがって彼の説明理論においては、どの程度の驚きが生じているのかを定量的に扱うことのできる驚きの形式化が必要である。驚きとは予期しない出来事が起こることだ。つまり出来事 E に驚くとは、 E が起こったと分かる直前まで E は起こりにくいと思っていたということである。そしてその時点で E が起こりにくいと思っていたらほど、 E に対する驚きの強さは強くなる (逆相関する)。彼はこれを認識状態モデル K_E^- を用いて形式化する。

認識状態 K が E を信念として受け入れた直後の認識状態とすると、 K から E を除去した認識状態 K_E^- は E を信念として受け入れる直前の認識状態に相当する *⁸(*Ibid.*: 168, 176)。すると E が起こったと分かる直前まで E は起こりにくいと思っていたことは認識状態 K_E^- によって再現することができ、また E に対する驚きの強さは、その時点で E が起こりにくいと思う強さ (すなわち K_E^- における E の信念の度合い $B_E^-(I(E))$) に逆相関するので、 $B_E^-(I(E))$ によって測ることができる *⁹(*Ibid.*: 167-8)。以下の図 7 は、驚きの強さと K_E^- における E の信念の度合いとの逆相関の関係の概略である。

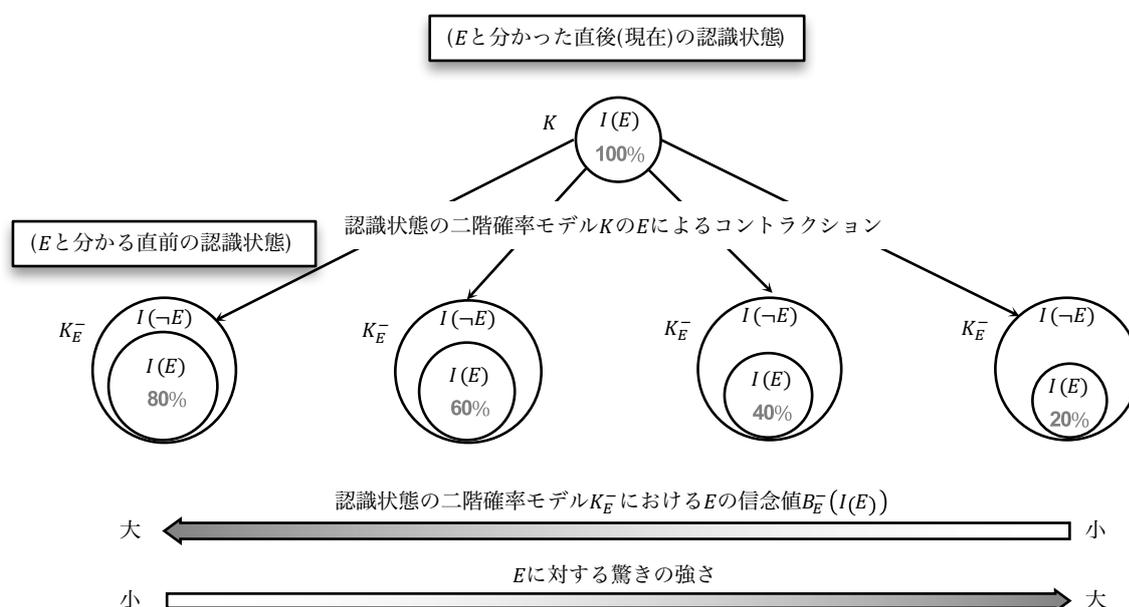


図 7: 驚きの強さと信念値の逆相関

つまり、 K_E^- における E の信念値 $B_E^-(I(E))$ の値が小さければ小さいほど驚きは強くなり、 $B_E^-(I(E))$ の値が大きければ大きいほど驚きは弱くなる。

*⁸ コントラクション関数自体にはそうした時間的な概念は含まれておらず、単に認識状態から特定の信念を除去するにすぎない。

*⁹ 「驚き」と「驚きの強さ」について少々補足しておきたい。彼はこれらの概念をそれほど形式的に明確化していない (単に K_E^- における E の信念の強さで驚きの強さが測れるとしか言わない) ので、以下ではこれらの概念を敷衍して論じたい。出来事 E に驚くとは、(1) E が起こったと分かる直前までは起こりにくいと思っていたのに、(2) なぜかは知らないが実際には E が起こったと分かったということである。明らかに驚きが生じているのは (2) の時点 (認識状態 K) である。しかし、その驚きが E のどんな側面に対するものなのか (どんな出来事 (対照的選択肢) ではなく E が起こったことに驚いているのか)、そしてその強さはどれくらいなのかということは、(1) の時点 (認識状態 K_E^-) を見てみないと分からない (特に K_E^- における E の信念値 $B_E^-(I(E))$)。したがってそれは、認識状態 K_E^- と K の組み合わせによって捉えることができるかもしれない。一方で、 E に対する驚きの強さ $S(E)$ [$0 \leq S(E) \leq 1$] は、(1) において E が起こると思う強さ (= K_E^- における E の信念の度合い) $B_E^-(E)$ [$0 \leq B_E^-(E) \leq 1$] に逆相関するので、例えば単純に「 $S(E) = 1 - B_E^-(E)$ 」などで測ることができるように思われる。

— 驚きを減らす = 直前の認識状態における信念の度合いを上げる —

出来事 E に対する驚きの強さは K_E^- における E の信念の度合い $B_E^-(I(E))$ に逆相関していた。したがって、出来事 E に対する驚きを減らすには $B_E^-(I(E))$ を高くすればよい。もし、 $T\&C$ のもとでの E の信念値 $B_E^-(I(E)/I(T\&C))$ が $B_E^-(I(E))$ より高い (i.e., $B_E^-(I(E)/I(T\&C)) > B_E^-(I(E))$) のであれば、つまり E が起こったと分かる直前の認識状態 K_E^- に $T\&C$ を信念として加えた認識状態 $(K_E^-)_{T\&C}^+$ における E の信念値 $(B_E^-)_{T\&C}^+(I(E))$ が、 K_E^- に $T\&C$ を信念として加えない場合の E の信念値 $B_E^-(I(E))$ よりも高くなる (i.e., $(B_E^-)_{T\&C}^+(I(E)) > B_E^-(I(E))$) のであれば、 $T\&C$ を信念として K に加えれば E に対する驚きは弱まる *10(図 8)。(*Ibid.*: §8.4)

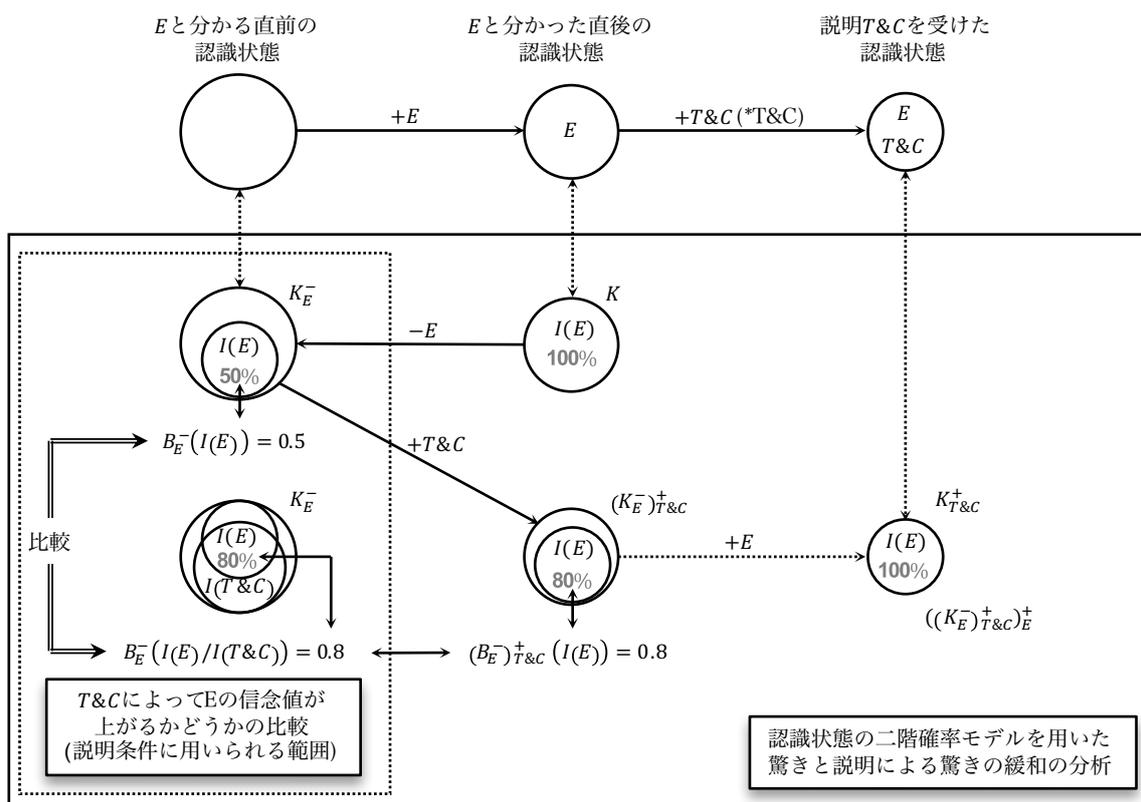


図 8: 認識状態の二階確率モデルを利用した説明による驚きの減少の分析例

— Gärdenfors の説明理論における説明の必要条件 —

ゆえに文 $T\&C$ が説明となるためには、 $B_E^-(I(E)/I(T\&C)) > B_E^-(I(E))$ とならなければならない。そのため彼の説明理論においては、以下のように説明の必要条件が与えられる *11(*Ibid.*: 178): 文 $T\&C$ が認識状態 K と相対的な単称文 E の説明である \implies (i) T は確率文の有限集合であり、(ii) C は単称文の有限集合であり、(iii) $B_E^-(I(E)/I(T\&C)) > B_E^-(I(E))$ を満たし、(iv) $B(I(T\&C)) < 1$ を満たす (i.e., $K \not\models T\&C$) *12。

*10 驚き是一种の認知的不協和と見なすことができる (*Ibid.*: 169)。つまり E であることが分かって信念として受け入れたものの (認識状態 K)、 E を除いた K の残りの部分では E は起こりにくいと考えている (認識状態 K_E^-) 状態、つまり K において信念同士が対立している状態である。こうした認知的不協和は驚きを減らすことで緩和される。

*11 彼が説明の必要十分条件ではなく必要条件しか与えないのは、彼の説明条件にはまだ付け加えるべきものがあると彼自身が考えているからである (*Ibid.*: 189)。

*12 最後の条件 (iv) は「 E だから E 」という自明な説明を避けるために与えられる (*Ibid.*: 178)。

— 単一試行確率 —

彼の説明理論において重要なのは、認識状態 K_E^- における被説明項 E の信念値 $B_E^-(I(E))$ である。その値は認識状態 K_E^- における被説明項 E の期待度であり、主観的な単一試行確率であると見なせるだろう。被説明項 E は単称文であり、その主観的な単一試行確率は K_E^- で受け入れられている確率文に依存する (*Ibid.*: 177)。彼はこうした単称文と確率文の間の繋がりを (1) 期待確率 P_V (*Ibid.*: 173-4) と (2) 条件 SCP (*Ibid.*: 176-8) によって形式化し、それらを用いて認識状態 K_E^- における被説明項 E の信念値 $B_E^-(I(E))$ を求める。

(1) 期待確率 P_V

$$P_V(Q) \stackrel{\text{def}}{=} \sum_{w \in V} P_w(Q_w) \cdot B(\{w\}) / B(V) \quad [\text{ただし } B(V) \neq 0, V \subseteq W] \quad *13.$$

(2) 条件 SCP

R を K において Ra であることが真として受け入れられている (i.e., $I(Qa) = W$ である) 最も強い述語とする。もし a について Qa と関連のあることが他に何も知られていないならば、 $B(I(Qa)) = P_W(Q/R) \quad *14$ 。

— 参照クラス —

条件 SCP における述語 R は**参照クラス**と呼ばれる。p. 4 に既出の Hansson は、「ジョーンズはヘビースモーカーだからだ」が、「なぜ私の友人達の中でジョーンズがガンになったのか？」の答えとなるが、「なぜヘビースモーカー達の中でジョーンズがガンになったのか？」の答えにはならないことに気付いた。そして彼は、両者の相違点となるなぜ-疑問の一部「私の友人達」や「ヘビースモーカー達」を参照クラス ^{*15} と呼んで、これが説明にとって決定的な要素であると主張し、なぜ-疑問の観点から説明の解明に取り組んだ。これに対して Gärdenfors は、そうした参照クラスは認識状態に組み込まれていると主張する (*Ibid.*: 182-3)。

^{*13}期待確率 P_V は認識状態の二階確率モデルにおいて定義される概念である。 P_V は一般化のために W の部分集合 V を用いて定義されるが、 V に W を取れば K において任意の個体が性質 Q を持つのはどのくらいの確率だと期待されているのかを知ることができる。その概要は次のようにまとめることができるだろう。認識状態の二階確率モデル K の中では確率測度 P_w と信念関数 B によって二階の確率分布が生じている。例えば K の W において、確率文 $p(Qx) = 0.5$ が真であるような可能世界の集合 $I(p(Qx) = 0.5)$ があったとする。その場合、その要素となる全ての可能世界 w について $P_w(Q_w) = 0.5$ が成り立っているため、それらの世界では任意の個体が性質 Q を持つ確率は 50% である。信念関数 B は、そうした可能世界の集合 $I(p(Qx) = 0.5)$ へさらに確率を与える。例えば $B(I(p(Qx) = 0.5)) = 0.8$ なら、 K において確率文 $p(Qx) = 0.5$ が 80% 信じられているということなので、任意の個体が性質 Q を持つ確率が 50% だと 80% 信じられていることになる (このように確率に対して確率が与えられているので、「二階確率モデル」と呼ばれている)。さて目下の認識状態 K では、確率文 $p(Qx) = 0.5$ が (100%) 信じられている訳ではないので、 K では任意の個体が性質 Q を持つ確率についてどのような値を持つと想像されているのか分からない。こうした状況下でも K において任意の個体が性質 Q を持つ確率についての信念があたかもあるかのように (特定の値を持つように)、二階の確率分布を一階の確率へ変換するのが期待確率 $P_W(Q)$ である。 K の W の要素となる全ての可能世界 w は確率測度 P_w によって任意の個体が性質 Q を持つ確率について特定の値を持っており、そうした特定の値を信念関数 B が与える確率の重みで平均する。要するに期待確率 $P_W(Q)$ とは、 W の任意の可能世界 w で任意の個体が性質 Q を持つ確率の期待値あり、その値が r だとすれば、 K において任意の個体が性質 Q を持つ確率が r だと期待されていると言おうということである。この期待確率により K ではどんな述語に対しても、任意の個体とその述語によって表現されている性質を持つ確率について特定の値を与えることができるようになる。しかし K において確率文 $p(Qx) = 0.5$ が信じられているなら期待確率は $P_W(Q) = 0.5$ となるが、その逆が成り立たないことには注意する必要がある。 (*Ibid.*: 173-4)

^{*14}条件 SCP は Hans Reichenbach の「最も狭い参照クラスの使用の原理」に依拠するとされている。その狙いは以下のようにまとめることができる。例えば認識状態 K において Nietzsche $[a]$ が梅毒性全身麻痺 $[Q]$ になったことが信じられていない時、それが起こる確率 $B(I(Qa))$ はどの程度だと期待されているのだろうか。これを単に $B(I(Qa)) = P_W(Q)$ とすると不都合が生じる。なぜなら $P_W(Q)$ は任意の個体が性質 Q を持つ期待確率だからだ。任意の個体には水や砂など、そもそも全身麻痺にならない個体も含まれる。そのため $P_W(Q)$ の値はかなり低い (おそらくほぼ 0 である) だろう。なぜこんなことになるのかというと、Nietzsche が人間である $[H]$ という情報を反映していないからである。この情報を反映して、 $B(I(Qa)) = P_W(Q/H)$ としてみよう。 $P_W(Q/H)$ は任意の人間が梅毒性進行麻痺となる期待確率である。もし K で確率文 $p(Q/H) = 0.0001$ が信じられているなら、その値は 0.0001 となり、先の $P_W(Q)$ の値を大きく上回る。しかし認識状態 K においては、Nietzsche が梅毒に罹患している $[R]$ という事も信じられているかもしれない。そうした情報も加味して、 $B(I(Qa)) = P_W(Q/R)$ としてみよう。 $P_W(Q/R)$ は任意の梅毒患者が梅毒性進行麻痺となる期待確率である。もし K で確率文 $p(Q/R) = 0.1$ が信じられているなら、その値は 0.1 となり、先の $P_W(Q/H)$ の値をまたも大きく上回る。このように、 K において Qa が起こる確率 $B(I(Qa))$ がどの程度だと期待されているのかを求めるには、 K において信じられている a に関するすべての情報を考慮しなくてはならない。これを正確に述べたのが条件 SCP である。重要なのは K における a に関する情報の有無が、 K における Qa の信念値 $B(I(Qa))$ に決定的な影響を与える点であり、ひいては説明となるかどうかにかんして決定的な影響を与えるという点にある。 (*Ibid.*: 176-8)

^{*15}van Fraassen の対照クラスが**命題のクラス**であったのとは対照的に、Hansson の参照クラスは**個体のクラス** (性質) であることに注意しなければならない。本発表は、参照クラスはもちろん対照クラスもまた認識状態に組み込まれていると主張するものである。

—Gärdenfors の説明理論による分析の一例—

以下は彼の説明理論を用いて、認識状態の違い (参照クラスの違い) が説明の成否を分ける事例を分析したものである (図 9)。

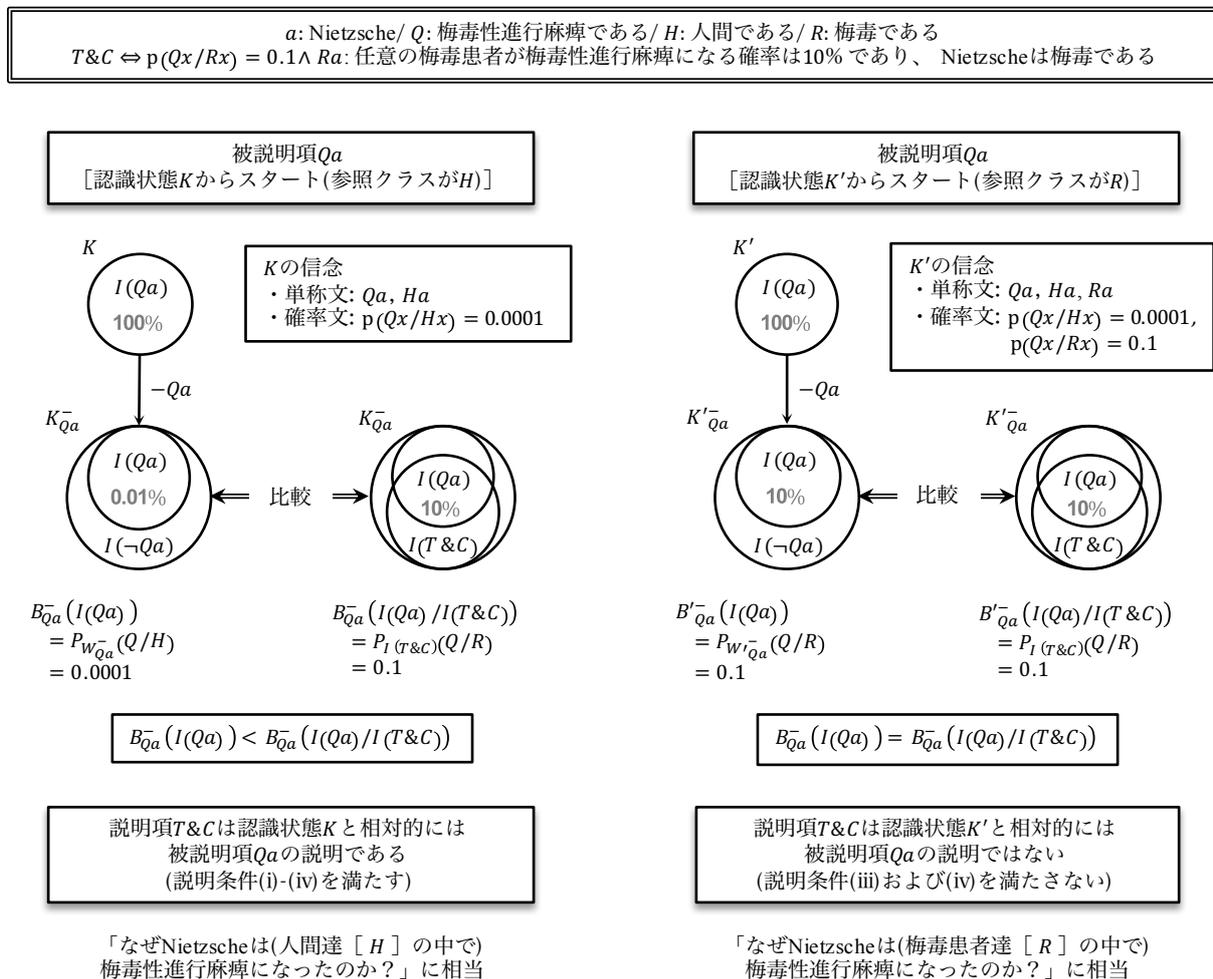


図 9: 認識状態の違い (参照クラスの違い) が説明の成否を分ける事例の分析の一例

—説明の文脈依存性—

彼の説明理論では、開始点となる認識状態 K がどのようなものであるのかは説明となるかどうかにとって決定的である。さらにたとえ開始点となる K が同じであっても、用いられるコントラクション関数が異なれば^{*16} 信念変更後の認識状態 K_E^- が異なるものとなり、これも説明となるかどうかにとって決定的となる (*Ibid.*: 188)。したがって、(1) 認識状態 K および (2) 認識状態 K_E^- は文脈的要素であると言えるだろう。

彼はこのような説明の文脈依存性について van Fraassen の説明理論との一致を見るが、van Fraassen は基礎的な文脈的要素を形成する認識状態の構造については何ら説明を与えていないと批判している (*Ibid.*: 170)。

^{*16} コントラクションに関する AGM 基本原理を満たすコントラクション関数は一意ではないので、実際に用いられるコントラクション関数がどういったものかは一意に定まらない。

なぜ-疑問モデル Q' の構築

— 2つの説明理論の文脈的要素 —

上記のように、van Fraassen と Gärdenfors の両者は説明の成否が文脈に左右されると考えるため、彼らの説明理論には共に文脈的要素が組み込まれている。彼らの説明理論に用いられる文脈的要素をリストアップしてみよう。

・ van Fraassen の説明理論における文脈的要素

[F-1] 関連性関係 R [F-2] 対照クラス X [F-3] 背景 K [F-4] K の部分 $K(Q)$

・ Gärdenfors の説明理論における文脈的要素

[G-1] 認識状態 K [G-2] 認識状態 K_E^-

両者の説明理論における文脈的要素には明らかな類似点がある: (1) 背景 K と認識状態 K 、(2) K の部分 $K(Q)$ と認識状態 K_E^- 、(3) 対照クラス X と認識状態 K_E^- 、(4) トピック P_k と認識状態 K 。(最後の (4) におけるトピック P_k は文脈的要素とされていないが、これが認識状態 K と類似点を持つことを後で論じる。)

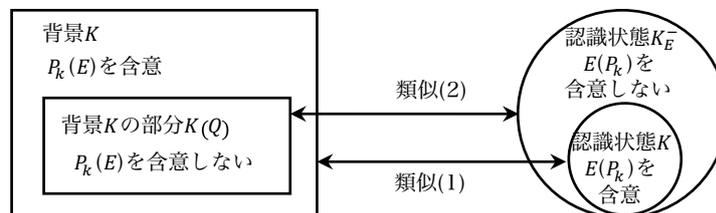
— 2つの説明理論における文脈的要素の類似点 —

(1) 背景 K と認識状態 K

これらの文脈的要素が類似していることは一見して明らかである。[F-3] 背景 K は、質問者とそれを聞いている人によって受け入れられている背景理論と背景情報から成り、説明すべき事実であるトピック P_k を含意しなくてはならない。一方で [G-1] 認識状態 K も、事実を表す単称文と一般的知識や帰納的信念を表わす確率文から成り、被説明項 E を含意しなければならない (図 10)。

(2) K の部分 $K(Q)$ と認識状態 K_E^-

これらの文脈的要素も一見して明らかな類似がある。[F-4] K の部分 $K(Q)$ は、説明すべき事実であるトピック P_k を含意しないような背景 K の部分であり、説明 (答え) の評価に用いられる。一方で [G-2] 認識状態 K_E^- も、認識状態 K から被説明項 E を除去した (それゆえに $K_E^- \not\models E$ となるような) 認識状態であり、説明の評価に用いられる^{*17}(図 10)。



背景 K および $K(Q)$ は文集合、認識状態 K および K_E^- は可能世界モデルであるため、各々の包含関係は逆転している

図 10: 背景 K と認識状態 K および K の部分 $K(Q)$ と認識状態 K_E^- の類似

^{*17} van Fraassen はこれら文脈的要素の類似に関して言及している (van Fraassen 1980: 225)。

(3) 対照クラス X と認識状態 K_E^-

[F-2] 対照クラス X は命題のクラスである。しかしそれは単なるクラスではなく選択肢の集合、すなわち説明を求めている命題 (トピック P_k) とその代わりに真であったかもしれない命題 (対照的選択肢) の集合である^{*18}。一方で [G-2] 認識状態 K_E^- の W_E^- も、説明を求めている命題 E が真である可能世界とその代わりに真であったかもしれない命題が真である可能世界 ($\neg E$ が真である可能世界) の集合である (図 11)。

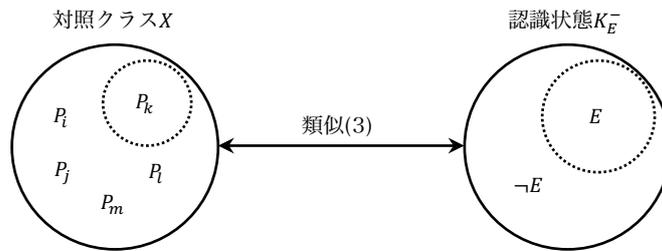


図 11: 対照クラス X と認識状態 K_E^- の類似

(4) トピック P_k と認識状態 K

[F-*] トピック P_k は説明を求めている命題である。それは対照クラス X の要素でもある。一方で [G-1] 認識状態 K の W も説明を求めている命題 E が真である世界の集合である。そしてそれは認識状態 K_E^- の W_E^- に包含される可能世界の集合でもある (図 12)。

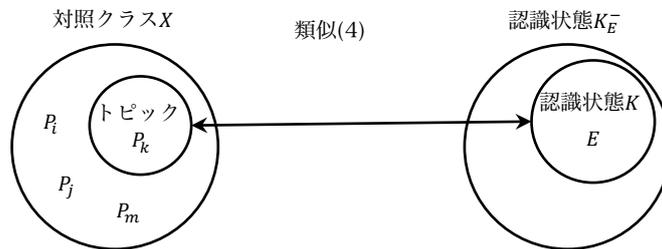


図 12: トピック P_k と認識状態 K の類似

^{*18}対照クラス X のメンバー間にどのような関係があるのかについては、何人かの論者達によって議論がなされている。実は van Fraassen 自身その関係について多くを語っていない (彼はなぜ-疑問の前提において対象クラス X の要素のうちトピック P_k のみが真で、それ以外の要素 $P_i (i \neq k)$ は偽であるとしか言わない)。これに対して Dennis Temple はこれでは対照クラス概念が未定義なままだと van Fraassen を批判し、van Fraassen の挙げる実例を通してその関係を「その状況における排他的選択肢 (両立不可能)」と解し、そこには何らかの様相的關係があると見ている (Temple 1988: 144-7)。しかしこの見解は、トピック P_k と対照的選択肢 $P_i (i \neq k)$ との関係であり、トピック P_k 以外の対照的選択肢同士 P_i と $P_j (i \neq j \text{ かつ } i, j \neq k)$ の間にも両立不可能な関係があるのかどうかは分からない。もしそれらの間にも両立不可能な関係があるならば、次項図 14 のようになる。個人的には対照クラス X のメンバー間の関係にはこうした条件を付け加える必要があるのではないかと考えるが、正直なところ今はよく分からないので、本発表では Temple の主張に則って次項図 13 のようにトピック P_k とそれ以外の対照 $P_i (i \neq k)$ との間には両立不可能な関係があるということだけを仮定して話を進めたい。念のために付け加えると、Temple は、なぜ-疑問の一般的形式を「なぜ Q ではなく P なのか?」と考える van Fraassen に代表されるような立場を「対照説」と呼び、単に「なぜ P なのか?」と考える Carl Hempel に代表されるような立場を「命題説」と呼んで、自身は「命題説」の立場を取っている。つまり Temple は van Fraassen と異なる立場を取っている。

—対照クラス X と認識状態 K_E^- の相違点—

対照クラス X と認識状態 K_E^- には異なる点もある。対照クラス X はトピック P_k 以外の要素がそれぞれ独立した命題であるのに対して、認識状態 K_E^- の場合は被説明項 E 以外が単に E の否定となっている。しかしこの違いは、対照クラス X のトピック P_k 以外の要素が認識状態 K_E^- における E の否定の**具体的な内容**を示していると見なすことができるかもしれない。例えば「なぜアダムはミカンでもナシでもなくリンゴを食べたのか？」によって表現されるなぜ-疑問 Q があったとする。この場合、トピックは「アダムがリンゴを食べた (P_1)」であり、対照クラスは $X = \{ \text{「アダムがリンゴを食べた } (P_1)\text{」, 「アダムがミカンを食べた } (P_2)\text{」, 「アダムがナシを食べた } (P_3)\text{」} \}$ となる。認識状態 $K_{P_1}^-$ においては、「アダムがリンゴを食べた (P_1)」の対照的選択肢として「アダムがリンゴを食べなかった ($\neg P_1$)」しか取らないのに対して、対照クラス X ではそれを「アダムがミカンを食べた (P_2)」と「アダムがナシを食べた (P_3)」に限定して、例えば「アダムがイヴにリンゴを返した」などの対照的選択肢を除外していると考えられる (図 13)。

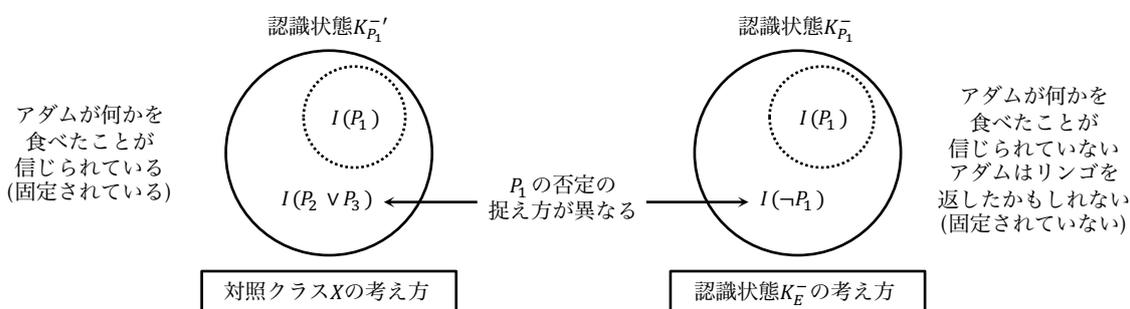


図 13: 対照クラス X の考え方を反映した認識状態 K_E^-

上記の対照クラス X の考え方を反映した認識状態 $K_{P_1}^-$ では、「アダムがリンゴを食べた (P_1)」か、あるいは「アダムがミカンかナシを食べた ($P_2 \vee P_3$)」のどちらかであると信じられている (P_1 と P_2, P_3 は排他的であり両立不可能)。したがって「アダムが何かを食べた」ということが信じられて (固定されて) いる。このように対照クラス X のトピック P_k 以外の対照的選択肢 $P_i (i \neq k)$ は、**トピック P_k の否定の具体的な内容**を示すことで認識状態 K_E^- において**保持すべき信念を固定する**という役割があるように思える。これは認識状態 K_E^- にはない特徴なので、以後の議論ではこの特徴を対照クラスのアドバンテージと見なして取り入れていくことにしたい。

上記ケースでは、「アダムがリンゴとミカンを食べた」という可能性は除外されるが、「アダムがミカンとナシの両方を食べた」という可能性は除外されない。もし「アダムはリンゴかミカンかナシのどれか一つだけを食べた」ということが信じられているなら、 $I(P_1)$ と $I(P_2)$ と $I(P_3)$ は互いに素となる (図 14)。

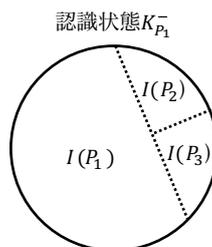


図 14: 対照クラスの各要素が各々両立不可能なケース

— 関連性関係 R はどこへ行ったのか —

類似点 (1) と (2) は van Fraassen の説明理論の背景に関係するのに対し、類似点 (3) と (4) はなぜ-疑問モデル Q に関係する。このことは、Gärdenfors の認識状態モデルによって van Fraassen のなぜ-疑問モデル Q が構築できるかもしれないことを示唆しているように思える。van Fraassen のなぜ-疑問モデル Q の構造は $Q = \langle P_k, X, R \rangle$ だった。そこで類似点を通じ、 P_k, X をそれぞれ K, K_E^- に置き換えて新しいなぜ-疑問モデルを構築しようとすると、関連性関係 R が足りないことに気が付く。これまで見た限りでは、Gärdenfors の説明理論の中には van Fraassen のなぜ-疑問モデルの関連性関係 R に相当するものが見当たらなかった。一体なぜ見当たらないのだろうか (図 15)。

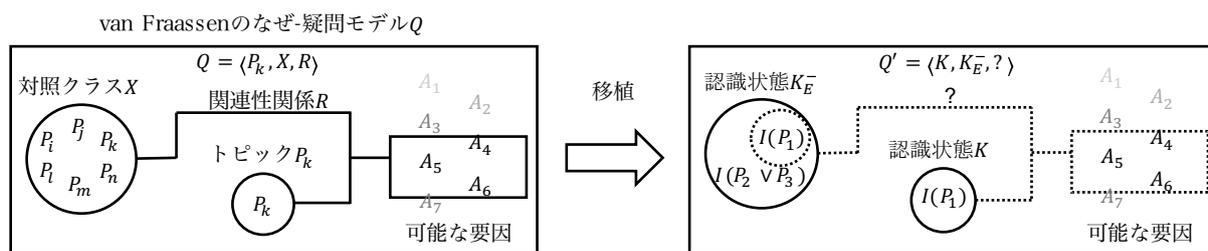


図 15: 関連性関係 R がない

結論から言えば、 K と K_E^- のみで関連性関係 R の役割がカバーされるからである。つまり**関連性関係 R は不要**だということだ。なぜそれが不要であるのかは、彼らの理論を構成する文脈的要素の間の類似点からではなく、彼らの理論が依拠する道具の類似点から、すなわち、しばしば因果の定義^{*19}に用いられる反事実的条件法のモデル「Lewis の圏域システム」と、Lewis の圏域システムを応用した信念変更のモデル「Grove の圏域システム」の類似点から理解することができる。

*19 「原因」と「際立った要因」の違いについて補足しておきたい (van Fraassen 1980: 115-6)。しばしば原因は反事実的条件文を用いて次のように定義される: A が E の原因である $\stackrel{\text{def}}{\iff}$ もし $\neg A$ だったなら $\neg E$ だっただろう。そして Lewis の圏域システムにおけるその真理条件は以下である: 反事実的条件文「もし $\neg A$ だったなら $\neg E$ だっただろう」が現実世界で真 \iff (1) 圏域システムのどの圏域 S_i においても $\neg A$ が真である世界が属していない、または (2) $\neg E$ が真である世界を少なくとも一つ含む最小の圏域 S があり、かつその圏域 S の全ての世界で $\neg A \rightarrow \neg E$ が真。van Fraassen が因果と反事実的条件法をセットで話題にする時、おそらく念頭にあるのは上記のような定義と真理条件だと思われる。しかし彼はこの因果の定義は**際立った要因** (because で示されるような要因) の定義としては必要だが十分ではない、つまり反事実的条件文が真なのに際立った要因とは見なしにくいケースがあると考え (際立った要因が多くの諸要因の中から選ばれていた、つまり諸要因の一部だったことを思い出そう)。彼が示す反例は以下のようものだ: 「目覚ましがならなかったら目が覚めなかっただろう」が真であれば「目覚ましがなったから (because) 目が覚めた」は真となるように思えるが、「夜に一切寝なければ目が覚めなかっただろう」が真であっても「夜に寝たから (because) 目が覚めた」は真とは思えない。目覚まし時計も睡眠も目覚めの原因ではあるが、睡眠は because で示されるような際立った要因ではない。では何がそうした違いを生み出しているのかというと、「なぜ A さんは目が覚めたのか?」と問う人は、大抵「A さんは寝た」ということを暗黙的に**固定** (前提) しているので、「夜に寝たから (because) 目が覚めたんだ」という返答を奇妙に思うわけである (「ビッグバンが起こったから目が覚めたんだ」という返答と同類である)。これを Lewis の圏域システムから見れば (次項の図を参照)、「目が覚めない ($\neg E$)」が真である世界を少なくとも一つ含む最小の圏域 S において、 S の全ての世界で「寝た (A)」が真である (固定されている) のだが、そのために S の全ての世界で「寝なかった ($\neg A$)」が偽となり、結果として S の全ての世界で「寝なければ目が覚めない ($\neg A \rightarrow \neg E$)」が**空虚**に真となってしまうことが問題なのだと思う (対照的に目覚まし時計の場合は「目覚ましがならなければ目が覚めなかった」が真であっても、**空虚**に真ではない)。要するに、圏域 S と $\neg A$ が真である可能世界の集合と**共通部分を持つかどうか**が、際立った要因となるかどうかにかんして決定的に作用しているのだろう (原因として認められてもそうした共通部分を持たなければ際立った要因とならない)。そこで、反事実的条件文を経由せずに直接 Lewis の圏域システム上で**際立った要因**を定義することにしよう: A が E の際立った要因である $\stackrel{\text{def}}{\iff}$ $\neg E$ が真である世界を少なくとも一つ含む最小の圏域 S があり、かつ**その圏域 S は少なくとも一つ $\neg A$ が真である世界を含み**、かつその圏域 S の全ての世界で $\neg A \rightarrow \neg E$ が真。以降の議論ではこうして定義された「際立った要因」を用いることにする。

— 関連性関係 R のおさらいと Lewis の圏域システム —

すでに述べた通り、関連性関係 R はその文脈において答えと見なしうる要因 (際立った要因) を定める関係である。そして van Fraassen によれば、そういった答えと見なしうる要因の選択は志向や関心といった文脈的なものに依存しており、その背後には「対立する反事実的条件法」があるとされていた: 技術者は「周囲の環境を固定して」視界を遮る植え込みがあってもブレーキが良ければ事故は起こらなかったと考え、都市計画家は「車の構造を固定して」ブレーキに欠陥があっても植え込みが視野を遮らなければ事故は起こらなかったと考える。これを Lewis の圏域システムで再現すれば以下になるだろう (図 16)。

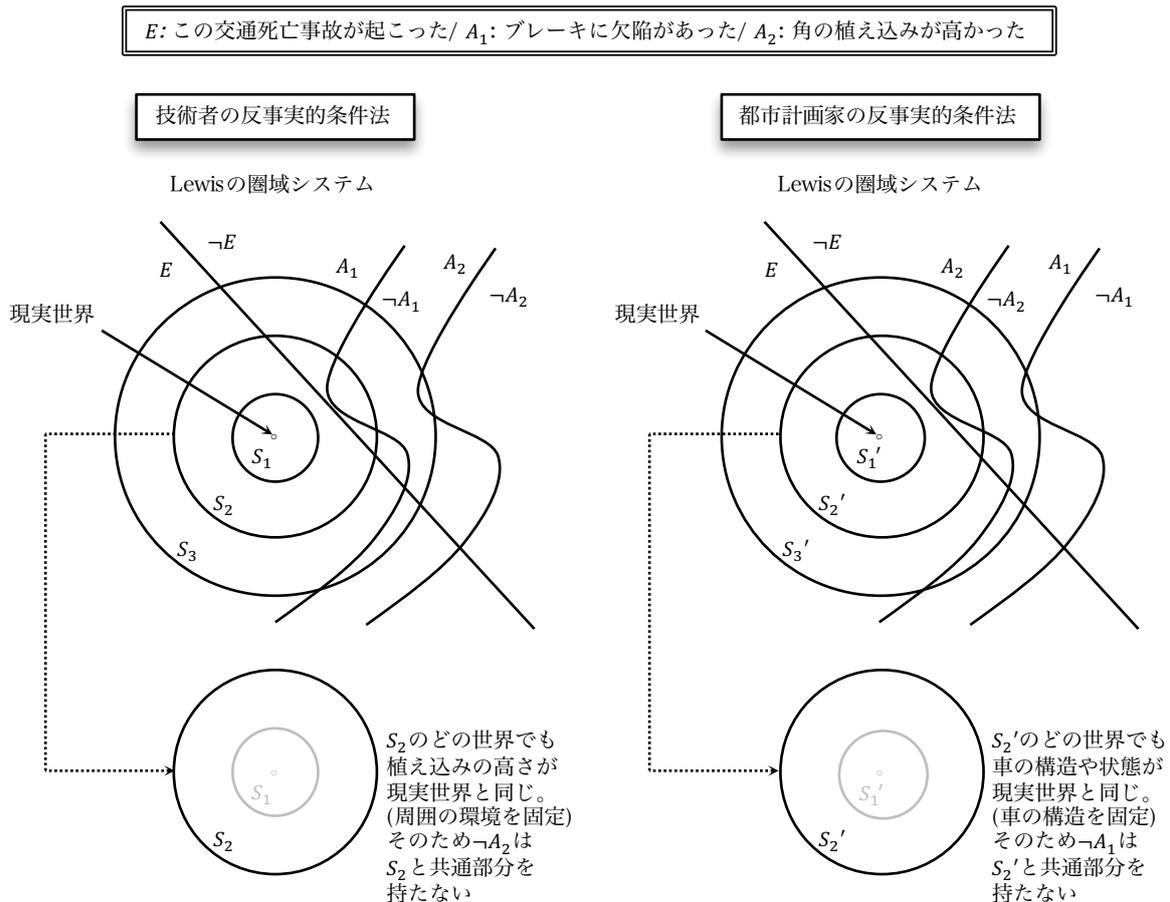


図 16: Lewis の圏域システムで見る「対立する反事実的条件法」

例えば上記の技術者の圏域システムにおいては、 $\neg E$ と共通部分を持つ最小の圏域 S_2 内の全ての可能世界で植え込みの高さが現実世界と同じ高さとなるため (周囲の環境が固定されるため)、文「角の植え込みが高くない」が真である世界は S_2 に一切含まれず、そのため「角の植え込みが高い」ことは「この交通死亡事故が起こった」ことの原因とはなっても**際立った要因**にはならないので^{*20}、「なぜこの交通死亡事故が起こったのか?」という問いに対して、「なぜなら (because) 角の植え込みが高かったからだ」と答えることができなくなってしまう (つまり「角の植え込みが高かった」ことは答えとしての関連を失ってしまう) のである。同様のことは都市計画家の圏域システムについても言える。「対立する反事実的条件法」とはこのことを指していると思われる。こうした対立が起こるのは明らかにそれぞれの圏域システムにおいて何を固定するか (何を現実世界と同じとするか) が異なっているから、つまり技術者と都市計画家の考える諸可能世界間の類似関係が異なっているからである。

^{*20} 「際立った要因」の定義については、前頁の脚注を参照されたい。

—コントラクションと Grove の圏域システム—

Grove の圏域システムは Lewis の圏域システムと形式的にはほぼ同じであるが、Lewis の圏域システムが反事実的条件文の意味論を与えるために用いられるのに対して、Grove の圏域システムは信念変更のモデルとして用いられる。重要な違いは、Lewis の圏域システムが一つの可能世界 (多くは現実世界) に中心化されるのに対して、Grove の圏域システムは認識状態 K (可能世界の集合) に中心化される点にある。下図 17 は、Grove の圏域システムによる、認識状態 K からコントラクションを行うことで得られる信念変更後の認識状態 K_E^- の構成である。しかし Lewis の圏域システムと同様、信念変更後の認識状態 K_E^- がどのようなものとなるかは、圏域システムにおける諸可能世界間の類似関係 (エントレンチメント) に依存する。

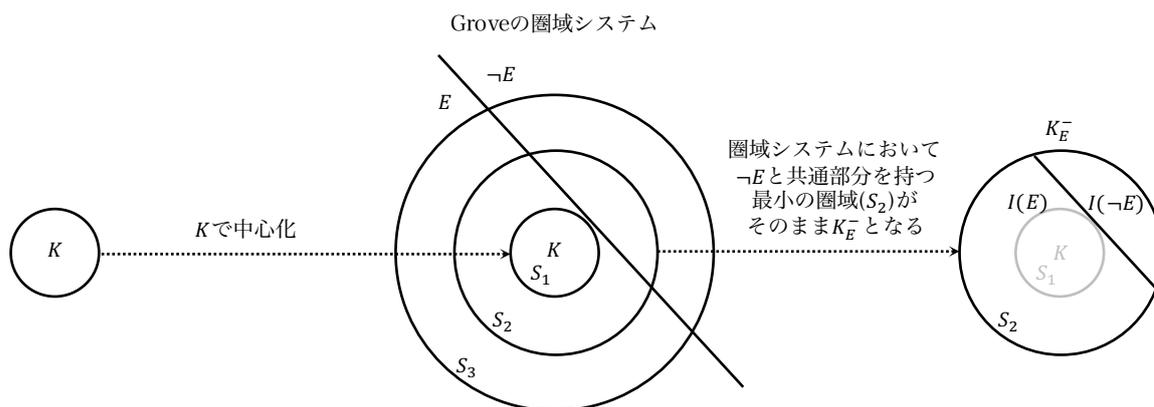


図 17: Grove の圏域システムにおける認識状態 K のコントラクション

—関連性関係 R が不要である理由—

p. 15 の考察では、対照クラス X におけるトピック P_k 以外の対照的選択肢 $P_i (i \neq k)$ は、トピック P_k の否定の具体的内容を示すことで、認識状態 K_E^- において保持すべき信念を固定するという役割があるように見えた。これを Grove の圏域システムから見れば、対照クラス X は圏域システムにおける諸可能世界間の類似関係に制限をかけて、認識状態 K_E^- で保持すべき信念を固定しているのだと見なすことができる。さてそうした Grove の圏域システムにおける諸可能世界間の類似関係 (エントレンチメント) への制限は、そのまま Lewis の圏域システムへと持ち込むことができ、図 16 のようにどのような要因が答えと見なしうるのか (際立った要因) を定める (図 18)。要するに**対照クラス X さえ定まれば関連性関係 R は自動的に定まるので、関連性関係 R が不要になるのである。**

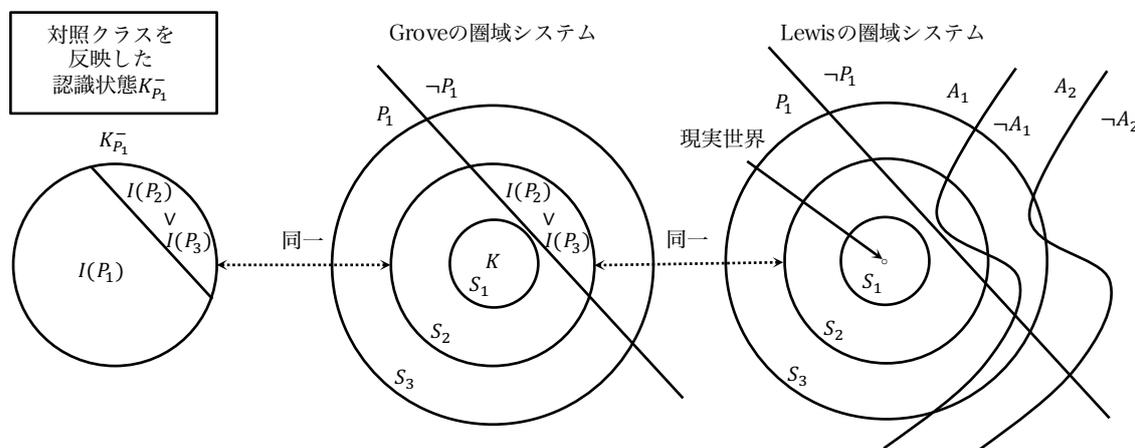


図 18: 対照クラスによる諸可能世界間の類似関係の制限

—なぜ-疑問モデル Q' —

以上の考察に基づいて、なぜ-疑問文によって表現されるなぜ-疑問のモデル化を行う。構造については、van Fraassen のなぜ-疑問モデル $Q = \langle P_k, X, R \rangle$ を元に、(1) P_k には Gärdenfors による認識状態の二階確率モデル K を移植し、(2) X には K の P_k によるコントラクションを行って得られる信念変更後の認識状態の二階確率モデル $K_{P_k}^-$ を移植し、(3) R の除去を行った。なぜ-疑問の定義については関連性関係 R の不足を補うことを目指して対照クラス X の考え方を導入しているが、それはあくまで関連性関係 R の不足を補おうとしているだけであり、なぜ-疑問の実態に即した定義を与えるためにはもっと強い条件を追加する必要があると考えるため、必要十分条件ではなく必要条件としている。

構造 なぜ-疑問モデル Q' は $\langle K, K_{P_k}^- \rangle$ と同一視される [$Q' = \langle K, K_{P_k}^- \rangle$]。ただし、 K は認識状態の二階確率モデルであり、 $K_{P_k}^-$ は K の P_k によるコントラクションを行って得られる信念変更後の認識状態の二階確率モデルである。

定義 なぜ-疑問モデル Q' において (対照的) なぜ-疑問文「なぜ (P_i, \dots, P_n ではなく) P_k なのか?」が問われている \implies (1) K において P_k が信じられている、かつ (2) $K_{P_k}^-$ の $W_{P_k}^-$ において $I(\neg P_k) = I(P_i \vee \dots \vee P_n)$ 。
(図 19)

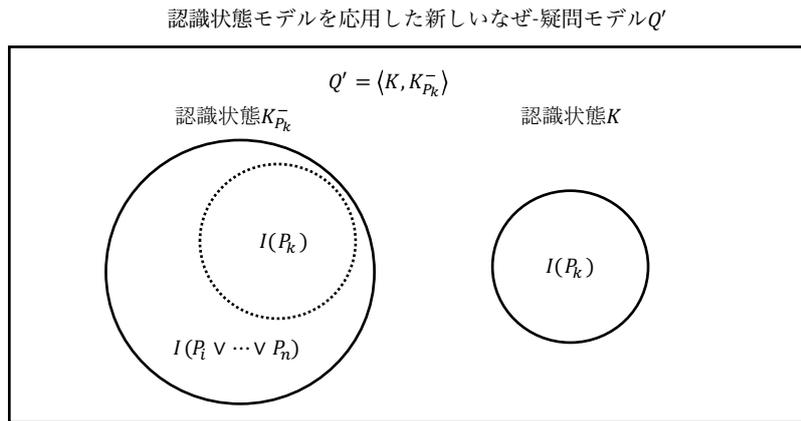


図 19: 認識状態の二階確率モデルを応用したなぜ-疑問モデル Q'

まとめ

—以上の考察が正しいなら一体何が得られるか—

・なぜ-疑問文は「なぜ？」と疑問を感じている認識状態を表現している

なぜ-疑問モデル Q' は2つの認識状態モデル $K, K_{P_k}^-$ から成っている。 K は P_k を信じたばかりの現在の認識状態とされ、一方で $K_{P_k}^-$ は P_k を信じる直前の認識状態とされていた。 P_k を信じる直前に P_k が起こると予想していなければいけほど驚きは大きくなり、「なぜ？」と思う気持ちは大きくなるだろう。なぜ-疑問文はこのように、疑問を感じている認識状態を表現していると考えられる。

・認識状態と信念変更の観点から見たなぜ-疑問と説明のプロセス

おそらく「なぜ Q ではなく P なのか？」と問う人は、 P ではなくて Q が起こると思っていたのだと思われる。しかし実際には P が起こってしまった。この時彼の心の中では認知的不協和が生じる。 P であることは分かっているが、彼の信念の P 以外の部分は Q を支持しているのである。このままでは同じ状況下で再び Q を予測し、それが外れて P が起こってしまうかもしれない。そのため信頼性のある予測力を手に入れるために自分の信念に変更を加える必要がでてくる。そこで P に驚いていない人に「なぜ Q ではなく P なのか？」と尋ねるわけである。というのも P に驚いていない人は Q ではなく P が起こると思っていたはずなので、 Q ではなく P が起こる理由を知っている可能性が高いからだ。一方 P に驚いていない人は、そのなぜ-疑問文を聞くと選択肢(圏域)を P と Q の二択に絞り込み、 Q ではなく P が起こる理由(答え)を質問者に的確に答えることができる。こうして答えを聞きだした質問者はその答えを信念に組み込めば、再び同じ状況下で Q ではなく P を予測できるようになり、予測力が向上する。認識状態と信念変更の観点から見ると、なぜ-疑問と説明は以上のようなプロセスの中の一部として位置づけることができるように思われる。

・対照クラス X が関連性関係 R を定めている

もっとも興味深いのは対照クラス X と関連性関係 R の間の関係がこれまでよりかなり明確になったことである。上記のプロセスを見ると、全ての始まりは質問者が P ではなく Q を予測していたことだ。しかしその予測が外れて Q ではなく P になり、その理由が知りたいと考えて、質問者は選択肢を P と Q の二択に絞り込むことを強要する対照的ななぜ-疑問「なぜ Q ではなく P なのか？」を尋ねる。つまり質問者が P ではなく Q を予測していたことが対照クラス X を定めるのである。一方で対照的ななぜ-疑問「なぜ Q ではなく P なのか？」を尋ねられた返答者は、 P と Q の二択になるような圏域(可能性の範囲)へと自らの認識状態をコントラクションし、質問者の望む(適切で関連のある)答えをそうした認識状態を通じて与えることができる。つまり今度は対照クラス X が関連性関係 R を定めるわけである。予測のミス → 対照クラス → 関連性関係という一連の流れの明確化はこれまでの考察から得られる重要な成果と言える。

・背景となぜ-疑問が一体化している

説明理論ではしばしば背景の取扱いに困難を抱えるが、p. 13-4 の類似点の考察から [認識状態 $K =$ 背景 $K =$ トピック P_k] が成り立ち、また [認識状態 $K_E^- =$ 背景 $K(Q) =$ 対照クラス X] が成り立つので、なぜ-疑問モデル $Q' = \langle K, K_{P_k}^- \rangle$ には背景を含め、説明に必要なと思われる全ての要素がすでに組み込まれていることになる。これにより将来なぜ-疑問モデル Q' 上で説明条件を与えて説明理論を構築する場合、比較的シンプルな説明モデルが構築できる可能性がある。

—以上の考察に問題はないか—

• 対照クラス X の読み

本発表では対照クラス X の要素であるトピック P_k とそれ以外の要素の間に両立不可能な (排他的) 関係があることを前提として話を進めたが、van Fraassen 自身はそこまで述べてはいない (p. 14 の脚注 18 参照)。Temple はそうした両立不可能性を認めるが、そうした読みが許されるのかどうかは分からない。

• 際立った要因の読みと定義が合っているかどうか

「際立った要因」は van Fraassen の説明理論において重要な役割を果たすが、その概念に対して明確な定義は与えられておらず、本発表ではその概念の明確化を行うために Lewis の圏域システム上での定義を行った (p. 16 の脚注 19 参照)。この定義はこれまでの考察において重要な役割を果たしており、これが覆されるとこれまでの考察も覆される可能性がある。しかしこの定義もまた許されるものかどうかは分からない。

• 検証を行っていない

新たに構築したなぜ-疑問モデル Q' が期待通りに振る舞うか、具体的に言えばなぜ-疑問モデル Q' が関連性関係 R の役割をカバーしているかどうかを確かめようと試みたが、十分な時間がなく断念した。したがってこのモデルが上手く機能するかどうかは現在のところ分からない。

参考文献

- Gärdenfors, Peter (1988). *Knowledge in flux: Modeling the dynamics of epistemic states*. The MIT press.
- Temple, Dennis (1988). The contrast theory of why-questions. *Philosophy of Science* 55 (1):141-151.
- Van Fraassen, Bas C. (1980). *The Scientific Image*. Oxford University Press.