

ポール・ハンフリーズの確率因果の理論

網谷祐一*

1 はじめに

因果と説明はとても密接な関係を持っている。航空機事故が起こったと考えてみよう。非常に多くの文脈で、原因(たとえば整備不良)を指摘することは、その事故の説明を構成する。つまり因果の説明——原因に言及することで結果を説明する——である。ポール・ハンフリーズ (Paul Humphreys) は、『説明の可能性』(*Chances of Explanation*) (Humphreys 1989) で、確率因果 (probabilistic causation) を用いた説明の理論を提起した。

彼の理論は (i) 確率因果の理論と (ii) 説明の理論の二つの部分からなる。ハンフリーズは確率因果について「 X が Y の原因であるのは、 X が Y の確率を高めるとき」という正連関の考え方を受け入れる。しかし X が Y に正連関にあるかどうかを知るために必要な確率の比較を「中立状態」(neutral state) と呼ばれる事象との比較で行う (2.1 を参照)。説明については、ハンフリーズは原因を正の原因 (contributing causes) と負の原因 (counteracting causes) に分け、 $P(Y|X)$ の値自体は説明に登場しない点で他の論者と異なっている。また、ハンフリーズはウェスリー・サモンなどと同様に、説明についてのプラグマティックな見方に好意的ではない。

本論文はハンフリーズの議論のうち、主に確率因果についての議論を検討する。確率因果とそれに基づく説明の分野には、長大な哲学的パズルのリスト (疑似因果 (spurious causation), 説明項の関連性など) が存在する。しかしそのすべてに立ち入ることはできないので、ここでは「中立状態」の概念と、いわゆる負連関の問題についてのみ議論する。まず中立状態については反例が指摘されている。本稿では、クリストファー・ヒッチコックの「三項関係としての因果」の考え—— X が Y の確率を高めたかどうかをみるときに、文脈に応じて選ばれた別の事象 Z との比較で判定する——を取り入れると反例自体は回避できるが、ハンフリーズの存在的 (ontic) な説明についての見方に衝突することを示す。さらに、原因が結果に正の連関を持つことを受け入れる多くの確率因果の理論と同じように、ハンフリーズの理論も、負連関 (negative relevance) の問題について満足のいく応答ができていないことを論じる。

本論文の構成は以下の通りである。まず第二節で、ハンフリーズの理論を解説する。第三節では「中立状態」に対する批判を扱う。第四節では「負連関」の問題がハンフリーズの枠組みでも問題を持つことを論じる。結論では、結局のところハンフリーズのどこに問題があるのかを考える。

* プリティッシュ・コロンビア大学哲学科博士課程 yuiami@gmail.com

2 ハンプリーの理論

ハンプリーの確率的因果による説明の理論は、確率因果についての議論と因果的説明の議論にわかれている。本節はそれを順にみていく。

2.1 ハンプリーの確率因果の理論

先に述べたように、ハンプリーは、原因は結果の確率に正の連関を持つと考える点で、スッピース (Suppes) などと軌を一にしている。しかし彼は正連関のアイデアを素直に解釈した、以下の定義を取らない。

X^* は Y の原因である (あるいは Y を妨げる) のは以下のときであり、そのときに限る。つまり $P(Y|X) > P(Y|\neg X)$ (あるいは $P(Y|X) < P(Y|\neg X)$) が成り立つときである。

ハンプリーがこの定義を取らないのは、以下の例が関係している (Humphreys 1989, 41 頁以下)*²。研究者がある新薬の治験を行っているとしよう。事象 E, A, B, C を以下のように定める。

- E : 患者が治癒すること
- A : 患者に新薬を 100mg 投与すること
- B : 患者に新薬を 10mg 投与すること
- C : 患者に偽薬を投与すること

ここで $P(E|A) = 0.9, P(E|B) = 0.4, P(E|C) = 0.2$, また $P(A) = P(B) = P(C) = 1/3$ とする。さてフレッドは新薬の対象の病気の患者で、治療法 B を受けたところ、治癒した。この事象の原因はどうなるだろうか。

$$P(E|\neg B) = \frac{P(E \& (A \vee C))}{P(A \vee C)} = \frac{P(E|A)P(A) + P(E|C)P(C)}{P(A) + P(C)} = .55$$

それゆえ、 $P(E|\neg B) > P(E|B)$ *³。したがって、上の定義が正しいならば、新薬を 10mg 投与されたことはフレッドの治癒の原因ではなく、むしろ治癒を妨げたことになる。この結果は $\neg B$ が実際は A と C という異なった因果因子の選言であることから生じる。これが選言的因果因子 (disjunctive causal factors) の問題と呼ばれるものである。したがってこの場合 $P(E|\neg B)$ と $P(E|B)$ を比べることは適切ではない。ここでハンプリーは「中立状態」という概念を導入

*¹ この定義自体は X が事象トークンか事象タイプであるかにしられないが、後述のように本論文では単一事象の確率因果による説明についてのハンプリーの議論に着目するので、ここでは事象トークンと考える。

*² 以下、本書からの引用はページ数のみで行う。

*³ スペースの関係上詳細にふれることはできないが、確率の哲学的解釈として *Chances of Explanation* では傾向性解釈を採用している (15 頁注 21, および 55 頁を参照)。しかしこれはハンプリーが傾向性解釈を確率の唯一の正しい解釈 (*the interpretation*) と見なしていることを意味しない。Humphreys (1984) を参照せよ。

する。ハンフリーズによれば、中立状態とは、「[問題となる] 変数に対応する性質が全く存在しないような」状態をさす(38頁)^{*4}。たとえば、 B の中立状態を B_0 とすると、ここでの変数は新薬の投与量、それに対応する性質は新薬の薬効だから、じっさいのところ B_0 は C (偽薬の投与) である。したがって $P(E|B)$ と $P(E|C)$ の比較が原因を探るために適切である。そして $P(E|B) > P(E|C)$ だから、 B は E の原因である。負の因果の場合も状況は変わらない。負の因果関係を示すために必要な確率の比較は $P(Y|X)$ と $P(Y|X_0)$ の対である (X の中立状態を X_0 とする)。したがって、ハンフリーズによれば、確率因果の定義は以下のようになる。

ハンフリーズによる確率因果の定義 X が Y の直接の正の原因である (あるいは負の原因である) のは以下のときであり、そのときに限る。

1. X と Y が生じている。
2. $X \& Y$ および $X_0 \& Y$ と物理的に両立可能な (compatible) あらゆる状況 U において、 $P(Y|X) > P(Y|X_0)$ (あるいは $P(Y|X) < P(Y|X_0)$) (ここで X_0 は、 X の中立状態であるとする)。
3. X と Y は論理的に独立である。

(74頁・単純化と記号の変更を行っている)

ハンフリーズは正の原因と負の原因を、 X が X_0 にくらべて Y の確率を高めるか下げるかによって区別する。

2.2 ハンフリーズの説明の理論

ハンフリーズの議論のおもな対象は、単一事象の確率因果による説明である。彼にとって典型的な事例は、前節でふれた正負の原因が両方ともからんだ、次のような事例である。

ヒトが細菌 G に感染し、治療を施されなかった場合、死亡率は 50% である。しかし抗生物質 Q を用いると死亡率は 5% になる。アルパートは G に感染し、医師は Q を投与した。ところが、不幸なことに、彼は G から来る病気によって死亡した。(100頁・単純化してある)

彼の死をどのように説明すればよいだろうか。この事例では G は彼の死 (D) の正の原因—— G は D の確率を中立状態 (感染のない状態) と比べて高める——で、 Q は負の原因である (Q は $P(D|G)$ を下げる^{*5})。この二つの原因を統合するために、ハンフリーズは以下の「説明の標準形式」(canonical form of explanation) と呼ぶものを提案する。

説明の標準形式 Ψ にもかわらず、 Φ によって Y [は生じた]。(101頁・単純化を行っている)

^{*4} ただし、変数に対応する性質が存在しない状態がつねに中立状態となるわけではないことをハンフリーズは認めている(38頁)。しかし「中立状態」についての彼のさらなる定式化がどういう形を取るかは明瞭ではない。

^{*5} このように、対応する正の原因があるときのみ確率に影響を与える場合も負の原因となる(94頁)。

ここで説明項 Φ は正の原因の集合、 Ψ は負の原因の集合である。したがって（もしアルバートの死に上で述べたもの以外の要因が関わっていないとすると）彼の事例の説明は、抗生物質 Q にもかかわらず、細菌 G によってアルバートの死 D が生じた、ということになる。

いくつか注意する点がある。この標準形式においては正当な説明は Φ を必要とするが、必ずしも Ψ は必要としない。「 Φ によって Y 」というのは正当な説明だが、「 Ψ にもかかわらず Y 」というのは正当な説明ではない。つぎに、ヘンペル (Hempel 1965) とは異なり、ハンフリーズは $P(Y|X)$ の値が高い（たとえば 0.9）ことを要求しない。というのは、ハンフリーズにとって正当な説明が成り立つには、 $P(Y|X)$ が $P(Y|X_0)$ より高ければよく、その確率の値自体にはこだわらないからである。ハンフリーズによれば、これによって次の著名な例に対処することが可能になる。

梅毒と麻痺 麻痺 (R) は第三期梅毒の症状であり、第一期、第二期、そして第三期の梅毒 (T) を経た患者が適切な治療を受けないとかかる症状である。パートは未治療の第三期梅毒の患者である。パートが麻痺にかかったとしよう。その原因は T にあるように思える。しかし $P(R|T)$ はただか 25% にすぎない。(Salmon 1992, 28 頁)。

ハンフリーズは、この例では T の中立状態 (T_0) は梅毒(あるいは第三期の梅毒)がない状態であり、原因を判定するにあたり必要な確率の比較は $P(R|T)$ と $P(R|T_0)$ (そして $P(R|T) > P(R|T_0)$) なので、この例に自らの理論はうまく対処できるとする (123, 125 頁)。よってハンフリーズは多くのまれな事象 (improbable events) を説明できることになる。

第三に、ハンフリーズの説明に対する見方は、存在的 (ontic) であって、認識的あるいはプラグマティックなものではない。一般に、認識的あるいはプラグマティックな説明理論は、説明にかかわる人の認識状態 (信念や知識、予期など) や関心を考慮する。それに対して存在的な説明理論は、そうした認識状態や関心などを重要なものとは見なさない。むしろ説明者・被説明者の関心・知識・信念などにかかわりなく、世界の中に存在する真で客観的な因果因子を取り出すことを目指す。ハンフリーズの言葉を用いると、

[客観的な説明は] 因果因子や構造、メカニズム——そのはたらきが問題となる現象の説明の一部になるとみなせるような——を取り出すことに強調点を置く。(98 頁)

したがって、たとえばハンフリーズは真の原因と単なる背景条件をことさら区別しようとする試みに重要性を認めない。たとえば山中で喫煙することと酸素の存在はどちらも山火事 (W) の確率を上げ、ハンフリーズの定義に適用がゆえに W の (正の) 原因のリストに同じ資格で属する。もちろん、山火事を防ぎたいといったわれわれの関心を考えると、 W の説明において酸素の存在に訴えることはまだだが、「そうした客観的なリストが手に入らなければ、こうした人間の関心がかかわる余地はな」く、科学的説明の文脈ではこうした配慮をする必要はない (129 頁、73 頁も参照)。そして、そうしたリストをどうやって手に入れるか、それをどのように (言語的) 説明の中で用いるかを明らかにすることをハンフリーズは目指しているわけである。

3 中立状態

以下の2節ではハンフリーズの理論の問題点から、中立状態と負連関について検討し、ハンフリーズがそれから逃れることができるか考える。

3.1 反例：血圧

ある事象 X が Y の原因であるかどうかを決めるために、ハンフリーズは $P(Y|X)$ と $P(Y|X_0)$ を比較する (X_0 は X の中立状態)。この考えのエッセンスは、原因 X の影響がなくなったとしたら結果 Y (の確率) がどうなるかを考えるところにある。このアイデアは(決定論的・非決定論的)因果についての諸理論、たとえば因果についての反事実説 (counterfactual theory of causation) (Lewis 1973, 1979, 2000)、「構造方程式」(structural equation) アプローチ (たとえば Halpern & Pearl 2001a,b)*⁶、および操作可能性説 (Menzies & Price 1993) と同じ路線を歩んでいる。しかし、ハンフリーズの理論には以下のような反例がある (Hitchcock 1993)。たとえばキャサリンが自分の血圧が彼女の健康に貢献しているのを知りたがっているとしよう。この例においてなにが中立状態だろうか？ 明らかに、血圧がゼロの状態は中立状態ではない。なぜなら、もしそうだと、たとえ彼女の血圧が非常に高い (たとえば最小血圧 130/最大血圧 180) としても、彼女の血圧は彼女が活着していることの正の原因になるだろうからだ。しかし血圧の健康に対する最適値 (たとえば 67/115) を中立状態にするのも都合が悪い。たとえばキャサリンがきわめて高血圧 (たとえば 130/180) だったとしよう。彼女は高血圧患者のための食事療法を取り、血圧を (まだ高めとはいえ) 110/160 まで下げ、高血圧から来る病気を防いだ。しかしもし血圧の最適値が中立状態だとすると、彼女の血圧は依然として彼女が活着していることについての負の原因だということになってしまう。しかしもしそうだとすると、彼女の努力は彼女が現在活着していることを説明しないことになってしまう。もちろん、彼女の努力によって、彼女の生存確率の値自体は上がっているわけだが、そうした事柄は標準形式に則った説明には現れてこない。したがって、結果としての説明には変化はない。ハンフリーズも、ある変数の値自体が因果的効力を持つことに気付いてはいる (95 頁) が、それが「中立状態」に問題を引き起こすことまでは述べていない。

3.2 三項関係としての因果

中立状態の問題にどう対処すればよいだろうか？ ひとつの方策はヒッチコックの「三項関係としての因果」という考えを採用することである (Hitchcock 1993, 1996)。「 X_1 は Y の原因である」という例を取ろう。ヒッチコックも正連関のアイデアを共有する。しかし確率の上昇がみ

*⁶ ここでは「構造方程式」アプローチとしてグリモアらの理論 (Sprites et al. 2001) ではなく、パールらのものを念頭に置いている。

られたかを確かめるときの比較対象 (X_2) は中立状態 (X_0) ではない。むしろそれぞれの関心に応じて X_2 がどういう事象であるかを指定できるのである。すなわち、ヒッチコックによれば、因果関係は X_1 と Y の間に成り立つような関係ではない。むしろ、 X_1 と X_2 と Y の間に成り立つ関係なのである。われわれは「 X_1 は Y の原因であるか」と問うべきではなく、「 X_1 は X_2 とくらべたとき Y の原因であるか」と問うべきであり、 $P(Y|X_1)$ と $P(Y|X_0)$ ではなく、 $P(Y|X_1)$ と $P(Y|X_2)$ をくらべるべきなのである。

この修正案は「中立状態」の問題を回避する。血圧の例を取る。キャサリンの食事療法が彼女の健康に効果的かどうかを知りたいとしよう。食事療法のと、彼女の血圧は最小 130/最大 160(BP_1) から 110/160(BP_2) になり、彼女は生きている (S)。ヒッチコックによれば、この事例では、我々が比較すべき確率は、 $P(S|BP_1)$ と $P(S|BP_2)$ である。医学的知識によれば $P(S|BP_1) < P(S|BP_2)$ なので、 BP_1 との比較でいうと、キャサリンの血圧が BP_2 であることが、彼女が生きのびていることの原因といえる。この修正案は選言的因果因子の問題も逃れている。新薬の例 (2.1) を取ろう。新薬の効果をみるために自然なのは、新薬 (B) と偽薬 (C) の効果を比較することであって、たんに B の効果と $\neg B (= A \& C)$ の効果を比較することでない。したがって研究者はフレッドの治癒を新薬の投与に帰することができるのである。

このように、〈三項関係としての因果〉は中立状態の反例をうまく回避できるように思える。しかし、ハンフリーズがこの案を受け入れられるかはべつの問題である。ヒッチコックのアイデアも問題を抱えている (Hitchcock 1993, を参照) というもののほかに、このアイデアはハンフリーズが批判するプラグマティックな説明理論への傾きを拭いがたく帯びているからである。この案では第三項 (X_2) の取り方によって、簡単に同じ事象 X_1 が結果事象 Y の正の原因になったり負の原因になったりする。したがってこの因果概念にもとづく説明においては、質問者の関心に応じて説明が変わってくる。先の事例の説明で「食事療法の効果を知りたい」という質問者の関心を仮定したことを思い出してほしい。自分の努力を認めたいキャサリン (あるいは、科学の文脈に限定するなら、食事療法の効果を知りたい研究者) にとっては食事療法は生存にとって正の原因だが、彼女の主治医は彼女の血圧にはまだ問題があるというかもしれない。ハンフリーズにとってこれは受け入れがたい帰結のように思われる。もしそうならば、引き続き血圧の事例はハンフリーズへの反例であり続ける。

4 負連関の問題

前節では中立状態の反例をハンフリーズの枠組みの中で回避することが難しいことをみた。本節ではいわゆる「負連関」(negative relevance) の問題を扱い、ハンフリーズがこの問題にうまく対処できるか検討する。原因事象と結果事象の間に負の連関がある事例として、次の著名な例を考えてみる。

仮想的原子崩壊の例 ある原子が励起状態にあるとしよう。励起状態には四つのエネルギーレベルがある。この原子は第四のレベルにあり、中間のレベルを通して第一のレベルに崩壊

するかもしれない。 $P(n)$ を原子が第 n のレベルにある確率、 $P(m \rightarrow n)$ を原子が第 m レベルから第 n レベルに直接崩壊する確率、 $P(m \rightarrow n \rightarrow r)$ を原子が第 m レベルから第 n レベルを経由して第 r レベルに崩壊する確率とする (m, n, r は正の整数で、 $m > n > r$)。すべての可能な崩壊過程の確率が以下のようなものだったとしてみよう。

$$P(4 \rightarrow 3) = P(3 \rightarrow 1) = 3/4, \quad P(4 \rightarrow 2) = P(2 \rightarrow 1) = 1/4$$

そうすると、 $P(1|4) = 10/16$ である^{*7}。しかし、 $P(2 \rightarrow 1) = P(1|2) = 1/4$ だから、ある原子が第二レベルを経由することは、第一レベルに至ることに負の関連しか持たない。しかし、もしこの原子が第四レベルから第二レベルを経由して第一レベルに至ったとしたら、直観的にはその原子が第二レベルにあったことが、第一レベルに至ったことの原因であるように言えそうである。(Salmon 1980, を改変)

ここでは第二レベルに崩壊することが第一レベルに至る確率を下げているにもかかわらず、上の原子の場合は第二レベルにあったことが第一レベルに至ったことの結果のようにみえる。

4.1 可能性についての説明要求か？

ハンフリーズはどのようにこの例に対処するだろうか。彼はこうした「負関連」の事例に対して一項を割り、(i) ある事象そのものに対する説明要求 (explanatory request) と (ii) その事象がどのようにして可能になったか、という可能性についての説明要求 (how-possibly request) を区別することによって対処しようとする (121 頁)。たとえば原子爆弾がどのようにして/なぜ作られたのかについての因果的説明 (例：マンハッタン計画) と、そもそも原子爆弾を作ることがどのようにして/なぜ可能になったのかについての説明 (例：核分裂の発見) は必ずしもおなじではない。そしてこのような負関連の事例で求められているのは、実は (i) 事象そのものについての説明ではなくて (ii) 事象の可能性についての説明に過ぎない。しかし、事象の可能性についての説明は、自らの理論の対象である因果的説明ではない (116 頁)。したがって、梅毒の事例 (2.2) ではパートの麻痺そのものを説明することが求められているのに対し、原子崩壊の事例では、たんに原子崩壊の可能性についての説明——原子が第一レベルに到達することがどのようにして可能なのか——が求められているにすぎず、ハンフリーズの枠組みが適用できなくても問題はない。つまり「原子の第二レベルへの崩壊が第一レベルへの崩壊を可能にした」というのがこの事例における適当な説明 (ただし因果的説明ではない) ということになる。

しかしこれには二つの問題がある。ひとつは、原子の第四レベルから第二レベルへの崩壊は、その原子が第一レベルへ崩壊することを可能にする (make it possible) わけではない。というのははじめからその原子が第一レベルに崩壊することは可能だったからである。むしろ原子の第四レベルから第二レベルへの崩壊は、第一レベルへの崩壊の可能性を保っている (keep it possible)

^{*7} $4 \rightarrow 3 \rightarrow 1$ という経路を経由する確率 (9/16) と $4 \rightarrow 2 \rightarrow 1$ という経路を経由する確率 (1/16) の和。

にすぎない。第二に、このハンフリーズの反論は〈もし原因 X が結果 Y に対して負の関連しか持たないならば、 Y については、その可能性についての説明要求しかない〉ということをも前提としている。しかしこれが常に成り立っているとは思われない。

梅毒と麻痺（改訂版） 先の梅毒の事例で、麻痺（ R ）を患う確率が、未治療の梅毒を患っている以外のすべての場合で、25%を超えているとせよ（たとえば、梅毒患者以外のすべての人は他の病原菌 H に感染しており、 $P(R|H) > 0.25$ である。さらに梅毒の病原菌に感染した場合、 H の活動が完全に抑制されるとする）。パートは梅毒にかかり治療を受けず（ T ）、麻痺を患った。この場合でも、パートの麻痺の原因は未治療の梅毒であるように思える。

この場合では、第一の梅毒の例とは異なり、われわれは $P(R|T) > P(R|X_2)$ となるような X_2 は存在しない。しかしだからといってパートの麻痺の説明要求が急に（麻痺そのものではなく）麻痺の可能性についての説明要求に変わったわけではないように思える。

ハンフリーズはなおも、この改訂された梅毒の例は実際は麻痺の可能性についての説明を要求しているに過ぎないと反論するかもしれない。さらに、確率因果について考える際、たんに二つの事象間の関係だけでなく、それを含んだ系全体を考慮しなくてはならない、と述べるかもしれない。つまりもともとの梅毒の例と改訂された例は異なる系についてのものであって、したがって両者でパートの麻痺の説明に求められるものが異なっても不思議ではないといわれるかもしれない。しかしもしハンフリーズがこの反論を取るならば、与えられた例についての説明要求が (i) 事象自体か (ii) その事象の可能性についてかを判断するのに、なぜわれわれは系全体における諸事象の確率を考慮しなくてはならないのか、という理由を提示する必要があるだろう。そうでなければ、この反論はハンフリーズの枠組みを助けるだけのアドホックなもののみみられるだろう。

最後にふれておく価値があるのは、こうした事例（原子崩壊と改訂された梅毒の例）の因果関係については二つの両立しないように思われる直観があることである。原子崩壊の例でいうと、第二レベルへの崩壊が第一レベルへの崩壊の（正の）原因であると考える人がいる一方、イェルズが示唆しているように（Eells 1991）、原子の第一レベルへの崩壊は、第二レベルへの崩壊にもかかわらず（のために、ではなく）生じたと考える人もいるだろう。こうした複数の直観があることをハンフリーズのもともとの理論では扱うことができないように思われる。

4.2 不可能な事象の確率との比較？

しかし中立状態の概念がここで役立つように思う人もいるかもしれない。先の例で、第四レベルにある原子が第一レベルに崩壊するには、(a) 第三レベル経由および (b) 第二レベル経由の二つの経路がある。ある原子が第四レベルにあるとしよう。さてここで問題となる変数に対応する性質（第二レベルから第一レベルへの崩壊の可能性）がない状態を考えてみよう。すると、(b) の経路が無効になる（この原子は (b) の経路をとれない）。さらに、第四レベルにある原子は (a) の経路も取れないと考えてみよう。もちろん、この事例では、実際の原子は (a)(b) いずれかの経路

を取らざるをえず ($P(4 \rightarrow 3) + P(4 \rightarrow 2) = 1$)、どちらの経路も取れないということはありません。しかしこの、原子崩壊の系において実現不可能な状況 (X^*) を考えると、 X^* において、この原子が第一レベルに崩壊する確率は——原子が崩壊するには (a)(b) の二つの経路しかないの—— $P(1|X^*) = 0$ である。この確率との比較でいうと、ある原子が第二レベルにあることは、第一レベルへの崩壊の確率を高めている (すなわち、 $P(2 \rightarrow 1) = P(1|2) [= 1/4] > P(1|X^*) = 0$)^{*8}。

このように考えると、原子崩壊の事例では、原子が第二レベルにあったことはその原子が第一レベルに至ったことの原因であるといえるようにみえる。同様に、改訂された梅毒の例では、バートの麻痺の原因は彼が第三期の梅毒にあったことである。

この案にはいくつかの利点がある。(i) この案は、少なくとも部分的に中立状態の考えを保っている。(ii) 正連関の概念を保っている。(iii) しかし、先の〈三項関係としての因果〉という考えを受け入れるならば、つねに確率の比較対象としてひとつの事象に縛られなくてもよい。上でわれわれは $P(1|2)$ と $P(1|X^*)$ をくらべたが、必ずしもそうしなくてはならないというわけではない。原子崩壊の例では、 $P(1|2)$ と $P(1|3)$ をくらべて、原子が第二レベルを経由したにもかかわらず第一レベルへの崩壊がおこったと述べることができる (ただしハンフリーズに従うなら、負の原因しかないの、これは正当な「因果的説明」ではない)。したがって、この案は 4.1 の末尾でふれた二つの直観をうまく救うようにみえる。(iv) さらに、この案は、先のハンフリーズの議論——負連関の事例を事象の可能性の説明要求 (how-possibly request) のみを求めているものとして対処する——をうまく位置づけることができる。というのは、この案では、実現不可能な状態 (X^*) における結果事象 Y の確率 ($P(Y|X^*)$) を考える際、原因とみられる事象から結果の事象へ至るすべての経路が閉じられるため、その確率は常にゼロである (つまり不可能である)。それに対してたとえば $P(2 \rightarrow 1) = P(1|2) > 0$ である。この不可能性と可能性の対比 (および原因-結果の間に正連関がないこと) が、原子崩壊の事例が事象の可能性に対する説明要求のみを生み出すように見せていたわけである。

しかしこの案には、ほんとうに実質のある提案なのかという批判があるだろう。というのは、この案では結果事象に至るすべての経路 (a)(b) が閉じた状態を想定するので、結果の確率がゼロになるのは当たり前のように見え、その確率と他の確率をくらべることは無意味にみえるからだ。たとえば、ウェスリー・サモンは、グッドの同様の提案 (Good 1980) を、これでは特に何か明らかになるわけではない (not especially illuminating) として退けている (Salmon 1984a, 202 頁)。しかしここで述べた方策——結果に至る経路がすべて閉じた、という与えられた系においては不可能な状態を考える——は、他の決定論的・非決定論的因果理論ではたびたび試みられている。たとえばパールらによる「構造方程式」アプローチの場合をみてみよう (たとえば Halpern & Pearl (2001a), Hitchcock (2001)^{*9})。予備因果 (preemption) の事例を考えてみる。

8 ただし、一般的な確率の公理系 (たとえばコルモゴロフによる公理系) では、 $P(X^) = 0$ のため条件付き確率 $P(Y|X^*)$ は定義できない。したがってこの案を採用する場合、別の公理系を用いなくてはならない (たとえば Popper 1965, Bartha & Johns 2001, を参照)。

*9 ただし両者の議論は全く同じというわけではない。

見習いの暗殺者がある人物を射殺しようとしている。彼が狙撃に失敗した場合に備えて、彼のコーチも射殺の準備をしている。ところが、見習いは暗殺に成功した。(Hitchcock 2001)

先にも述べたように、「構造方程式」アプローチもデイヴィッド・ルイスらと同様に、事象 X が Y の原因ならば(そして X 以外に Y に因果的効力を持つ事象がないならば)、 X が生じなかったときに Y も生じないことを主張する。この例の問題は、たとえ見習いが暗殺に失敗したとしても、予備的な原因(コーチの狙撃)があるために、結果(暗殺)は変わらないようにみえるところにある。ヒッチコックは、たとえ見習いの狙撃の失敗が結果を反事実的に変えるわけではないとしても、もしももとの原因(見習いの狙撃)と予備原因(コーチの狙撃)のいずれもが生じなかったとき——与えられた系では実現不可能な状態——には暗殺は起こらない、ということを経路に見習いの狙撃は暗殺の原因であると主張する(Hitchcock 2001, 288 頁)。もちろん、ここでは暗殺に至る因果経路はもとの原因と予備原因しかないのだから、双方の経路を封鎖すると結果が生じないのは当たり前であるといわれるかもしれない。しかし、このアプローチについては、少なくともこれらの理論をこの理由のみで退けるような状況にはなっていない(Neuberg 2003, Woodward 2003, Kyberg 2005, Pearl 2003a,b, 2005)*¹⁰。この実現不可能な状態と実現可能な状態を比較することが「構造方程式」アプローチにおいてただトリヴィアルという理由だけで退けられないならば、ここでも先の提案を同じ理由のみで退けることはできない*¹¹*¹²。

したがって、上の案はハンフリーズにとって少なくとも考慮に値するような提案にみえる。ところが、またしても彼の枠組みの下ではこの提案は受け入れられそうもないのである。というのは、ハンフリーズの理論では、系の中で物理法則にてらして実現不可能な事象を考慮することはできないからである。ハンフリーズの確率因果の定義(2.1)では「 $X \& Y$ および X_0 (中立状態) $\& Y$ と物理的に両立可能なあらゆる状況 U 」において確率の比較を行うことになっている。それゆえ、 X_0 がもし物理的に不可能ならば、そうした状況 U 自体が存在せず、 $P(Y|X_0)$ と $P(Y|X)$ を比較することができない。これは因果を三項関係と見なして X_0 の代わりに第三の事象 X_2 を導入しても同じである。実際ハンフリーズは疑似因果(spurious causation)の議論の中で、共通原因 A が結果事象 B と C の必要十分条件である場合、 $B_0 \& C$ や $B \& C_0$ は(物理)法則にてらして生じえないので、 $P(A|B \& C)$ と $P(A|B \& C_0)$ や $P(A|B_0 \& C)$ と $P(A|B_0 \& C_0)$ の比較はできないと述べている(82 頁)。しかし先の提案では、 X^* は系において実現不可能な状態なのだから、そうした状態と物理的に両立可能な状況 U は存在しない。ゆえに、 $P(1|2)$ と $P(1|X^*)$ を比較することはできない。したがって、ハンフリーズはこの提案を受け入れることはできないので

*¹⁰ また、この提案がほんとうになにも明らかにしていないかどうかをきめるには、先に示したこの案のいくつかの利点を考慮する必要がある。

*¹¹ もうひとつの問題はこの案がトークンレベルとタイプレベルの区別に敏感ではないというものである。たとえば原子崩壊の例ではイールズ(Eells 1991)のような論者は第二レベルへの崩壊はタイプレベルでは真の原因であるがトークンレベルでは正の原因なのだ、と述べるかもしれない。しかし、このタイプ/トークンレベルの区別がほんとうにつけられるかどうかには異論がある。たとえば Hausman (2000) を参照。

*¹² また、ルイス流の反実仮想文の解釈では、前件が不可能な事象である場合、後件がトリヴィアルに真になってしまうことも問題である。

ある。

それではなぜハンフリーズは物理的な実現不可能な状況を考慮に入れられないのだろうか。それには彼の存在論が関わっている。ここでは詳細に立ち入ることはできないが、彼の存在論では結果事象はその現象が起こるシステム (S)、対象となる性質 (Z)、および試行 (trial, T)——たとえば実験・観察など——の関数 ($Z(S, T)$) と考えられている。したがってその事象の確率も Z, S, T の性質と見なされる (25 頁, 注 5)。しかし、システム S の中で実現不可能な状態は、実験や観察の対象とはならない。したがってその「結果」事象というものも定義できない。それゆえ、ハンフリーズは物理的に実現不可能な状況とその確率を考慮に入れることができないのである。

5 結論

この論文ではハンフリーズの確率因果による説明理論について議論してきた。ここではハンフリーズの理論について二つの批判を扱った。ひとつは、彼の「中立状態」の概念が血圧のような事例 (3.1) に対応できないということであり、もうひとつは負連関の事例をうまく扱えないというものだった。二つの批判を検討して明らかになったことは、ハンフリーズの存在的な説明理論の枠組みではこうした批判にうまく対処できそうもない、ということである。血圧の例は因果概念を三項関係と考えれば、何とか回避できそうにもみえるが、しかし存在的な説明理論ではなくなってしまう。負連関の事例についても、ハンフリーズのももとの議論——負連関の事例では事象の可能性についての説明要求しか求められていない——は原子崩壊の事例ではうまくいかないことをみた。また、与えられた系 S で実現不可能な事象を前件に持つ条件付き確率を考えると、いう方法 (4.2) も、いくつかの利点をもってはいるものの、これまたハンフリーズの使用論と相容れないことをみた。

これらの批判を通して明らかになってきたことは、

- (ア) 確率因果を正連関からとらえること、および
- (イ) 説明についての存在的な見方

を——ハンフリーズと同じやり方で——両立させることが難しいということである。先に見たように、ハンフリーズは(ア)と(イ)の両方を受け入れている。もちろんハンフリーズも日常的説明が認識状態や関心にかかわることを否定するものではないが、彼は真の客観的な原因を列挙する段階まではそうした関心が何らかの役割を果たすことを否定する (2.2)。だが血圧の事例はこの段階ですでに関心が何らかの役割を果たしていることを示唆する。また原子崩壊の事例では、ハンフリーズは事象自体についての説明を放棄するという形で(ア)を守ろうとするが、問題を解決することはできていない。

これに対して、他の多くの論者は(ア)(イ)いずれかを放棄するか、いずれかの条件をゆるめている。たとえばサモンは(ア)を拒否し (Salmon 1984a,b)、ファン・フラーセンは(ア)を維

持して(イ)を拒否する。またヒッチコックは(イ)に関する条件をハンフリーズよりもゆるめて、 X が Y に対して正負いずれかの因果的連関にあるかどうかについての客観的な区分は存在しないという(Hitchcock 1993, 359頁)。本論文ではどちらかという最後の路線に焦点を当ててきたが、もちろん、どの方策がすぐれているかは別の(長い)考察を必要とする。しかし、ハンフリーズのやり方で(ア)(イ)の連言を受け入れて満足のいく確率因果とそれに基づく説明の理論をたてることは難しいと言えるだろう*¹³。

参考文献

- Bartha, P. & Johns, R. (2001), 'Probability and symmetry', *Philosophy of Science* **68**, S109–S122.
- Eells, E. (1991), *Probabilistic Causality*, Cambridge University Press: Cambridge.
- Good, I. (1980), 'Some comments on probabilistic causality', *Pacific Philosophical Quarterly* **61**, 301–304.
- Halpern, J. & Pearl, J. (2001a), Causes and explanations: A structural- model approach (Part I): Causes, in 'Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence', Morgan Kaufman, San Francisco, CA, pp. 194–202.
- Halpern, J. & Pearl, J. (2001b), Causes and explanations: A structural-model approach (Part II): Explanations, in 'Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI)', Morgan Kaufman, San Francisco, CA.
- Hausman, D. (2000), Causal relations among tokens, types, and variables, Paper presented at University of Maryland, May, 2000, Philosophy of Science Association, Vancouver, November 2000.
- Hempel, C. (1965), *Aspects of Explanation*, Free Press: New York.
- Hitchcock, C. (1993), 'Generalized probabilistic theory of causal relevance', *Synthese* **97**, 335–64.
- Hitchcock, C. (1996), 'Farewell to binary causation', *Canadian Journal of Philosophy* **26**, 267–82.
- Hitchcock, C. (2001), 'The intransitivity of causation revealed in equations and graphs', *The Journal of Philosophy* **98**, 273–99.
- Humphreys, P. (1984), 'Why propensities cannot be probabilities', *Philosophical Review* **94**, 557–570.
- Humphreys, P. (1989), *The Chances of Explanation*, Princeton University Press: Princeton.

*¹³ 本論文は私が受講した二つの因果についての授業——ウェスリー・サモン教授と内井惣七教授の講義(2000年, 於京都大学), およびポール・バーサ教授のセミナー(2005年, 於ブリティッシュ・コロンビア大学)——に多くを負っています。ここに記して深く感謝いたします。

- Kyberg, H. (2005), 'Review of Causality', *Artificial Intelligence* **169**, 174–179.
- Lewis, D. (1973), 'Causation', *Journal of Philosophy* **70**, 556–567.
- Lewis, D. (1979), 'Counterfactual dependence and time's arrow', *Nous* **13**, 455–76.
- Lewis, D. (2000), 'Causation as influence', *Journal of Philosophy* **97**, 182–97.
- Menzies, P. & Price, H. (1993), 'Causation as a secondary quality', *The British Journal for the Philosophy of Science* **44**, 187–203.
- Neuberg, L. G. (2003), 'Review of Causality: Models, reasoning, and inference', *Econometric Theory* **19**, 675–685.
- Pearl, J. (2003a), 'Comments on Neuberg's review of causality', *Econometric Theory* **19**, 686–689.
- Pearl, J. (2003b), 'Reply to Woodward', *Economics and Philosophy* **19**, 341–344.
- Pearl, J. (2005), 'Response to review by Kyburg', *Artificial Intelligence* **169**, 180.
- Popper, K. R. (1965), *The Logic of Scientific Discovery*, Harper Torchbooks: New York.
- Salmon, W. (1980), 'Probabilistic causality', *Pacific Philosophical Quarterly* **61**, 50–74.
- Salmon, W. (1984a), *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.
- Salmon, W. (1984b), Scientific explanation: Three basic conceptions, in P. D. Asquith & P. Kitcher, eds, 'PSA 1984', pp. 293–305.
- Salmon, W. (1992), Scientific explanation, in M. Salmon, ed., 'Introduction to the Philosophy of Science', Hackett Publishing Company, pp. 29–88.
- Sprites, P., Glymour, C. & Scheines, R. (2000), *Causation, Prediction and Search 2nd Edition*, MIT Press: Cambridge MA.
- Woodward, J. (2003), 'Critical notice: Causality by Judea Pearl', *Economics and Philosophy* **19**, 321–340.