

情報概念の真理性は正当化できるか

Can we justify that the concept of information requires truth?

榎本 啄杜

Abstract

The concept of information is ambiguous and has different meanings depending on the context in which it is used. On the other hand, there is a strong belief that “information” that generates knowledge must encapsulate truth. Luciano Floridi is the defender of this idea, but in his argument, there seems to be some leaps in logic and inaccuracies in formula. In this paper, I try to reconstruct his argument and interpret it as reasonable one. However, some questions still remain, so I shall mention them at the end.

(1) 研究テーマ

知識論の分野において、情報概念がどのように使用されているかを検討することを通じて、情報概念の分析・精緻化を試みる。

(2) 研究の背景・先行研究

情報の哲学における主要な問題として、情報概念で知識をどう説明するかというものが挙げられる。たとえば、F. ドレツキ（1981）は知識と情報の関係について、以下のように述べている。

この意味¹では、誤情報は情報の類ではない。（中略）情報とは知識を生み出すことができるものであり、知識が真理を要求する以上、情報も真理を含意するはずである。（Dretske 1981, p.45）

ここでは、知識のもつ真理性は、それを引き起こす情報の真理性に由来していると考えられている。ここで問題となるのは、知識にとって必要とされる情報である「意味論的信息（semantic information）」が真理性をもつことは必ずしも自明ではなく、当該分野においても論争が続いているということである。

これを受けて L. フロリディ（2011）は、意味論的信息が真理性をもつことを示すために、形式的な議論を行っている（Chapter 4）。しかし、この議

論は本来あるべき前提や推論過程が省略されており、一見すると真理性を獲得するまでの過程に飛躍があるように思える。また、論理式に不正確な部分があることによって、解釈に困難をきたす箇所も見受けられる。本稿では、これらの不足部分を補うことを通じて当該議論を再構成し、筋の通ったものとして解釈し直すことを試みる。

フロリディの議論を概観する前に、前提となる背景を確認しておく²。情報科学やそれに関連する分野³で広く受け入れられている意味論的情報の定義として、GDI (General Definition of Information)⁴がある。これはデータ概念に基づくもので、意味論的情報は「(i) 1つ以上のデータで構成されており (ii) well-formed で⁵ (iii) 意味をもつ」という3つの条件で定義される。定義からも明らかのように、GDI は真理値に関する条件をもたない。フロリディによると、GDI は真理中立的であるために誤情報やトートロジーも情報の一種として認められることになってしまうため、意味論的情報の必要条件ではあるが、十分条件にはなっていない。これを解決するためには、GDI の定義に「(iv) 真である」を付け加えなければならない。

上記を達成するために、背理法を用いて以下のような意味論的議論 (semantic argument) がなされる。使用する語彙は以下の通り。

$D = \{p_1, \dots, p_n\}$; D は空でもよい命題の集合 (GDI の集合)⁶

ϕ, ψ ; D 上の命題変数

$S = \{i_1, \dots, i_n\}$; S は空でもよい意味論的情報の具体例の集合

x, y ; S 上の命題変数

$t(\phi)$; ϕ は真である (事実式)

$f(\phi)$; ϕ は偽である (事実式)

$t / f(\phi)$; ϕ は真である、または偽である (事実式)

$T(\phi)$; ϕ はトートロジーである

$C(\phi)$; ϕ は自己矛盾式である

$H(\phi)$; ϕ の情報量⁷

$P(\phi)$; ϕ の確率⁸

なお、 S は D の部分集合であり、意味論的情報はすべて命題の形をとる。また、彼によると、意味論的情報のモデルは①P.1~P.4 の4つの原理を必ず実装し、②R.1~R.3 の3つの要請をできるだけ満たすことが求められる。4つの原理 (P) と3つの要請 (R) は、以下の通りである。

P.1 : 情報量の非負性

$$(\forall x H(x) \geq 0)$$

P.2 : 情報量の加法性

$$(\forall x \forall y ((x \neq y) \rightarrow (H(x \cup y) = H(x) + H(y))))$$

P.3 : 逆関係原理 (Inverse Relationship Principle (IRP))

$$(\forall \varphi (P(\varphi) = 1) \rightarrow (H(\varphi) = 0)) \quad ^9$$

P.4 : 情報量が 0 の命題は意味論的情報ではない

$$(\forall \varphi ((H(\varphi) = 0) \rightarrow \neg(\varphi \in S)))$$

R.1 : 情報量は物理的なダメージがなくてもシンタクスの的に減少しうる

R.2 : 情報リポジトリに任意の命題 (事実式) を追加し続けても情報量は増加しづらい

R.3 : 情報量は物理的にもシンタクスの的にも意味論的にも減少しうる

なお、P.1 と P.2 は変数 x, y が用いられているため S の議論領域、P.3 と P.4 は変数 φ が用いられているため D の議論領域であると想定されている。前述したように S は D の部分集合であるため、 D の議論領域である原理は S にも適用可能であるが、その逆は成り立たない。

まず、 S と D を同じ集合だと仮定する。つまり、すべての命題は意味論的情報であると仮定する。すると、P.3 と P.4 から次のように言える。

$$T(\varphi) \rightarrow (P(\varphi) = 1)$$

$$(P(\varphi) = 1) \rightarrow (H(\varphi) = 0)$$

$$\underline{(H(\varphi) = 0) \rightarrow \neg(\varphi \in S)}$$

$$T(\varphi) \rightarrow \neg(\varphi \in S)$$

S と D は同じ集合であるのに、トートロジーは意味論的情報ではないという事実が導かれた。また、次のようにも言える。

$$(P(\varphi) = 0) \rightarrow (P(\varphi) < 1)$$

$$(P(\varphi) < 1) \rightarrow (\varphi \in S)$$

$$\underline{C(\varphi) \rightarrow (P(\varphi) = 0)}$$

$$C(\varphi) \rightarrow (\varphi \in S)$$

この結論は Bar-Hillel-Carnap Paradox¹⁰を引き起こす。P.3 (IRP) によ

れば、出来事の生起確率が低くなればなるほど情報としての価値は増すため、確率が 0 になる自己矛盾式は非常に大きな情報量をもつ。しかし、 $p \wedge \neg p$ が非常に大きな情報量をもった情報として認められるというのは、避けられるべき事態である。GDI は確率についての前提を含まないために問題とならなかったが、この 2 つの難点は確率的アプローチをとると不可避なものである。

S と D を同一視するこの考え方、つまり GDI がこそが意味論的情報であるという考え方は、前述のように様々な分野で受け入れられている。しかし、いくつかの原理を適用するだけで上記のような矛盾が生じてしまうため、この仮定を棄却し、これらの難点を排除するような修正を加える必要がある。これは、GDI に以下の規則を付加することで解決できる。

$$[1] \quad \forall \varphi ((T(\varphi) \vee C(\varphi)) \rightarrow \neg(\varphi \in S))$$

これにより、すべての命題のうち、トートロジーと自己矛盾式以外の命題が意味論的情報になる。つまり、真偽はどうあれ、事実式でなければ意味論的情報としては認められない。こうして導かれた GDI+ [1] のモデルは、意味論的情報が真理性をもたないとする論者でも受け入れられるものである。

ところで、現時点のモデルにおいて、P.1 と P.2 から次のように言える¹¹。

$$[2] \quad \forall \varphi \forall \psi ((\varphi \neq \psi \wedge t/f(\varphi) \wedge t/f(\psi)) \rightarrow (0 < H(\varphi) < H(\varphi \cup \psi) > H(\psi) > 0))$$

$$[3] \quad H(U_1^n \varphi) < H(U_1^{n+1} \varphi)$$

[3] は [2] を一般化したものである。これによると、任意の命題をとり続ければいくらでも総情報量が大きくなり続け、物理的にダメージを与えてもしない限りは、減少することはない。

しかし、ある命題のすぐ後にそれを否定する命題をとる操作を想定すれば、現在のモデルは、シンタクスの情報量が減少してしまう現象を説明できないことに気が付く。たとえば、命題 p を受け取った後に命題 $\neg p$ を受け取ったとすると、どちらの方が信頼できる命題であるかを判断できない以上、新しい命題 $p \vee \neg p$ は情報量をもたない。4 つの原理だけではこの現象に太刀打ちできないため、現在のモデル GDI+ [1] は、さらに R.1 と R.2 の 2 つの要請を満たすべきであると考えられる。

R.1: 情報量は物理的なダメージがなくてもシンタクスの減少しうる

$$[4] \quad \diamond(H(U_1^n \varphi) > H(U_1^{n+1} \varphi)) \quad (\text{physically and syntactically})$$

R.2: 情報リポジトリに任意の命題（事実式）を追加し続けても情報量は増加しづらい

$$[5] P(H(U_1^n \varphi) \geq H(U_1^{n+1} \varphi)) \geq P(H(U_1^n \varphi) < H(U_1^{n+1} \varphi))$$

R.2 は R.1 をエントロピー的に修正したものである。これによって、今や GDI はトートロジーや自己矛盾式、さらには複数合わさることで矛盾してしまうような事実式の集合を排除できるようになった。

ところが、この改訂版 GDI ではシンタクスの情報量の減少は説明できるが、意味論的な減少（semantic loss）はまだ説明できていない。元の真なる命題に誤った言明を加えてしまえば、元の命題が情報量のストックに貢献していた分の情報量は減少してしまうにもかかわらず、手持ちの規則だけでは情報量が増加することになってしまう。偽である命題はいくらでも作ることができるので、このような仕方では情報量が単調に増加してしまう場合もまた排除しなければならない。これを考慮すると、GDI はさらに R.3 の要請を満たすべきである。

R.3: 情報量は物理的にもシンタクスのにも意味論的にも減少しうる

$$[6] \Diamond(H(U_1^n \varphi) > H(U_1^{n+1} \varphi)) \text{ (physically, syntactically and semantically)}$$

R.3 は R.1 と R.2 を含んでいるため、R.3 を満たす規則を付加すれば、すべての要請を満たすことができる。こうして得られる規則が、次である。

$$[7] \forall \varphi ((\varphi \in S) \rightarrow t(\varphi))$$

[7] は定義から自明に [1] を排除しており、R.1 と R.2 は R.3 に含まれているため、[7] だけでこれまでの問題をすべて解決することができる。よって、意味論的情報のあるべきモデルは GDI+ [7] である。

(3) 筆者の主張

以上がフロリディの議論の概要である。[1] を導くまでは比較的明快に議論が進むが、3つの要請を持ち出してから [7] を導き出すまでの過程に、一見すると飛躍があるように思える。そのような感覚は、ここに論理式の誤りと推論過程の省略があるために生じるものである。これを補ってやれば、フロリディは当初の宣言通り背理法を用いて、段階を踏んだ議論を行っていることがわかる。

まず、[2] と [3] は D 上の変数を用いて記述されていたが、正しくは S 上の変数を用いて、以下のように記述されるはずである。

$$[2'] \quad \forall x \forall y ((x \neq y \wedge t/f(x) \wedge t/f(y)) \rightarrow (0 < H(x) < H(x \cup y) > H(y) > 0))$$

$$[3'] \quad H(U_1^n x) < H(U_1^{n+1} x)$$

この段階 ($GDI + [1]$) で意味論的情報として認められているのは、トートロジーと自己矛盾式を除くすべての命題であり、真偽を問わず事実式の集合である。[2] と [3] は事実式について述べているのに加えて、議論領域が S である $P.1$ と $P.2$ を適用していることは明らかであることから、 S 上の変数である x と y を用いて [2'] [3'] のように記述することが求められる。

さらに、フロリディの記述には以下の推論が欠落している。現在、 $GDI + [1]$ と S は同じ集合であると仮定している。 S (つまり $GDI + [1]$) では $P.1$ と $P.2$ から [2'] と [3'] が成り立っているはずである。しかし、この集合の一部の要素において、 $R.1$ と $R.2$ ([4] と [5]) で説明されるような現象 (S 上の変数を用いて $\diamond(H(U_1^n x) > H(U_1^{n+1} x))$) が実際に成り立っており、[2'] [3'] と矛盾してしまう。具体的には、非負性をもつ情報量の加法性が $GDI + [1]$ の内部では明らかに成り立っていない。これは、 $GDI + [1]$ と S を同一視する仮定によって生じた矛盾であるため、この仮定は誤りである。つまり、 S は $GDI + [1]$ の集合よりもさらに小さい集合であり、その S は、上記の矛盾を引き起こした原因である、合わさると矛盾している事実式を $GDI + [1]$ から除いたものと一致する。

$R.3$ についての議論も同様に言える。上で導かれた暫定版 S を仮に S' とすると、 S' と S は同じ集合であると仮定できる。 S' では [2'] と [3'] が成立するはずだが、これと矛盾するはずの $R.3$ ($\diamond(H(U_1^n x) > H(U_1^{n+1} x))$) が S' 内で成り立ってしまい、矛盾する。これは S' と S を同一視する仮定によって生じた矛盾であるため、この仮定は誤りである。つまり、 S は S' よりもさらに小さい集合であり、その S は、上記の矛盾を引き起こした原因である、偽の事実式を S' から除いたものと一致する。

このように背理法を適用し続けて最後まで残ったのは、以下のように真の事実式だけである¹²。

① トートロジー	② 合わさると矛盾する事実式	
① 自己矛盾式	③ 偽の事実式	④ 真の事実式

この図全体で D の集合、つまり GDI である。①まずトートロジーと自己矛

盾式が含まれることで矛盾が生じたため、これらを D から除いた。この時点で残っているのは事実式の集合である。導かれた集合を D' とすると、②合わさると矛盾してしまう事実式が含まれることで矛盾が生じたため、これも D' から除いた。導かれた集合を D'' とすると、③偽の事実式が含まれることで矛盾が生じたため、これも D'' から除いた。④すると残るのは真の事実式だけである。これを S と同一視してもこれ以上の矛盾は特に生じないため、 S の要素であれば、それは真の事実式であると結論してもよい。

フロリディの議論を再構成する作業は以上である。後半から飛躍していたように思えた当該議論も、4つの原理と3つの要請の正当性を受け入れる限り、一貫して背理法を用いた合理的な論証になっていると言ってよいだろう。また、当該議論は集合 S と D の議論領域の違いに大きく依拠していることもわかった。もしも [2] [3] が原文の通り D の議論領域であるとすれば、同じ D 上で [2] [3] と R.1~R.3 を同時に受け入れることになってしまい、議論の出発点で矛盾が生じてしまう。当該議論における背理法は、 S と D の議論領域の違いがなければ成り立たないものだと言える。

(4) 今後の展望

「筆者の主張」において、「4つの原理と3つの要請の正当性を受け入れる限り（中略）合理的な論証になっている」、「当該議論は集合 S と D の議論領域の違いに大きく依拠している」と述べた。これはつまり、前提の正当性が失われることがあれば、合理的な論証とは言えなくなることを示唆している。この点において、以下の疑問が残る。

たとえば P.1 情報の非負性について、当該議論においては $H(\varphi)$ の計算方法が明記されていないため、これだけを見ていては、どのような値をとるのが不明である。確率的アプローチをとるのであれば対数関数を用いた自己情報量の計算方法¹³が一般的だが、この計算方法であれば、どのような命題についても定義上自明に負の値をとらない。また、当該議論がなされている章の次の章（Chapter 5）で真理値に基づいた情報量の計算方法が提案されている¹⁴が、この計算方法でも情報量は負の値をとることはない。すると、P.1 は S 上の議論領域であるだけでなく、 D 上の議論領域でもあるということになる。その他の原理及び要請についても、その正当性を考察すべき点が今後の課題として残されている。

(5) 参考文献

Bar-Hillel, Y., and Carnap, R. 1953, 'An Outline of a Theory of Semantic

- Information' repr. in Bar-Hillel (1964), 221-274.
- Dretske, F. I. 1981, *Knowledge and the Flow of Information* (Oxford: Blackwell).
- Floridi, L. 2005, 'Is Semantic Information Meaningful Data?' *Philosophy and Phenomenological Research*, 70(2), 351-370.
- Floridi, L. 2010, *Information: A Very Short Introduction* (Oxford: Oxford University Press).
- Floridi, L. 2011, *The Philosophy of Information* (Oxford: Oxford University Press).

¹ 「この意味」とは、直前の文章である以下を指している。「出来事の状態は、それにあたることによって適切な観測者が X について学ぶことができるちょうどその程度において X についての情報を含む。(中略)これを『情報』という用語の核となる意味として言及しようと思う。」

² (Floridi 2011, p.83-84)

³ 情報管理、データマイニング、データベース設計や意思決定論などの分野が挙げられている (Floridi 2011, P.83)

⁴ フロリディ (2005) では同じものが SDI (Standard Definition of Information) と呼ばれている。

⁵ ここで言う well-formed とは、必ずしも文法的なものでなくてもよい。たとえば、電気の回線が物理的にうまく繋がっていても well-formed である。これにより、命題以外のものでも意味論的情報として認められる。

⁶ ここでは簡単のため、GDI の代わりに命題を用いて議論が進められるが、議論の進め方と結論的には本質的な影響はない。

⁷ 「情報量」という言葉における「情報」は、ある事象の確率によって決まる数学的な量のことであり、フロリディが真理性をもたせようとしている意味論的情報とは異なる。 D 上の変数についても情報量を計算できるのはそのためである。

⁸ 原文 (Floridi 2011, p.98) では S 上の変数 x を用いて $P(x)$ と記載されているが、確率は D 上の変数についても計算できる (実際にフロリディも D 上の変数について確率を用いている) ので、 $P(\phi)$ の誤植だと思われる。

⁹ この論理式は、IRP から導かれはするが IRP そのものではない。本議論ではこの論理式しか用いないためにこう記載していると考えられる。なお、IRP は $1 - P(\phi)$ で表現されるので、 $P(\phi)=1$ ならば情報量は 0 になる。

¹⁰ Bar-Hillel-Carnap Paradox: 理想的な受信者が受け取ることのないような自己矛盾文は、最も大きな情報量を運ぶことになる。(Bar-Hillel and Carnap 1953, p.229)

¹¹ 「筆者の主張」で詳しく述べるが、[2] と [3] は本来であれば S 上の命題変数である x と y を用いて記述されるべきである。この誤りが議論をわかりにくくしている要因の一つであると筆者は考える。

¹² この図では便宜上互いに排反であるかのように記載しているが、②合わさると矛盾する事実式と③偽の事実式④真の事実式には重複があり、実際には排反なわけではないことに注意。

^{1 3} 自己情報量は $I(\varphi) = -\log P(\varphi)$ で計算される。

^{1 4} (Floridi 2011, p.123-127) 参照。対象となる命題が、言及しようとしている現実 w からどの程度離れているかの度合いに基づいて計算される。

(関西大学)