

学習に基づくドレッキ表象論の検討 —意味の因果的役割の問題に関して—

勝亦佑磨 (Yuma Katsumata)

東京大学・日本学術振興会

本発表では、心の哲学において志向性の自然化の問題に取り組む立場の一つである「目的論的機能主義」に焦点を当てる。より具体的には、ドレッキが『行動を説明する』(Dretske, 1988) で展開した学習に基づく表象論が、「意味の因果的役割の問題」に対して提示している解決策をみて、それが妥当なものであるかどうかを検討する。

意味の因果的役割の問題とは、次のような問題である。日常的な理解では、ある行動の時点における心的状態は、その行動に因果的役割を果たしていると考えられる。例えば、ある男性の「ビールを飲みたい」という欲求と「冷蔵庫の中にビールがある」という信念は、「冷蔵庫を開ける」という行動に因果的役割があると考えられる。それでは、なぜ信念や欲求のような心的状態は、行動に因果的役割があると考えられるのだろうか。ドレッキによれば、こうした心的状態が「意味 (表象内容)」を持つゆえに、ある行動が引き起こされる。つまり、こうした心的状態の持つ意味は、行動に因果的役割を果たしていると考えられるのである。例えば、「ビールを飲みたい」という欲求と「冷蔵庫の中にビールがある」という信念が意味を持つゆえに、「冷蔵庫を開ける」という行動が引き起こされる。そして、これらの心的状態の持つ意味は、「冷蔵庫を開ける」という行動に因果的役割を果たしていると考えられるのである。ところが、非物理的な心的状態の意味が物理的な身体を動かすということは不可解なことである。というのも、私たちの身体を実際に動かしているのは脳状態などの物理的な状態であるとする、非物理的な意味が行動に因果的に関与する余地がなくなってしまうからである。そうだとすると、心的状態の意味はいかにして行動に因果的役割を持ちうるのだろうか。

ドレッキは、こうした問題を以下のように解決する。まず、心的状態が意味を持つとは、生物学的な内的状態 C が外部の状態 F を表象することであり、 C が F を表象するとは、「 C が F を表示する機能を持つ (C が F についての情報を運ぶ機能を持つ)」ことにほかならない。そして、動物の内的状態 C (例えば脳状態) は、行動における学習を通じて F を表示する機能を獲得し、 F についての表象になる。例えば、信号が赤のときにバーを押すと餌が与えられるような環境におけるハトの学習を考えてみよう。学習過程において、信号が赤のとき (F) にハトに生じる内的状態 C はバーをつつく際の運動 M を引き起こし、このことがハトにとって有益な結果 (餌) をもたらした。その結果、ハトには、「 F である際に生

じる C が M を引き起こす」ようなメカニズムが形成される。こうして、ハトの内的状態 C は F を表示する機能を獲得し（すなわち F についての表象になり）、F という意味を持つ。そして、この意味が現在のハトの行動において因果的役割を持つと考えられるのである。

ここで重要なのは、ドレッキにとって意味とは、非物理的なものでもなければ単に現在生じているようなものでもなく、動物の過去の学習過程において生じた様々な物理的な事象の総体のことであるという点である。すなわち意味とは、「F である際に生じる C が M を引き起こす」ようなメカニズムが形成されるまでに生じてきた脳状態や環境などを含むようなものである。このようにドレッキは、意味を物理的に還元することによって、意味の因果的役割の問題に一定の解決策を与えていると考えられる。

ところが、以上にみたようなドレッキによる説明には、次のような批判がある。第一に、ドレッキによる歴史的説明（すなわち動物の学習過程を用いた説明）に対する批判が挙げられる（Horgan, 1991）。ドレッキの説明にしたがえば、現在の行動に因果的役割があるのは、現在の内的状態の持つ意味というよりはむしろ過去に生じてきたものであるということになる。だが、これはわれわれの直観に反する。というのも、われわれの直観では、現在の行動に因果的役割があるのは、現在の内的状態の持つ意味であると考えられるからである。第二に、そもそも意味に因果的役割を認めることはできないという立場からの批判が挙げられる（Kim, 1991）。キムによれば、生物の持つ内的状態の意味論的性質とは脳神経科学的性質に付随するものであり、行動に因果的役割があるのは後者である。そうすると、内的状態の意味論的性質は因果的役割を持たないことになる。

それでは、こうした批判に対し、ドレッキはどう応答しうるだろうか。本発表では、以上のような議論をふまえて、ドレッキ表象論が意味の因果的役割の問題を解決できているかどうかという問題を検討したい。

《主な参考文献》

Dretske, Fred. 1988, *Explaining Behavior: Reasons in a World of Causes*, Cambridge, MA: MIT Press.

Horgan, Terence. 1991, "Actions, Reasons, and the Explanatory Role of Content," *Dretske and His Critics*, edited by Brian P. McLaughlin, Oxford, Basil Blackwell, pp.73-101.

Kim, Jaegwon. 1991, "Dretske on How Reasons Explain Behavior," *Dretske and His Critics*, edited by Brian P. McLaughlin, Basil Blackwell, pp.52-72.