

## 『論語』の公理化 $\implies$ 孔子の A I 化へ：その背景

日大理工・数 高橋英之 ( Hideyuki Takahashi )

### I 導入：目標・目的、テーマの位置付け、なぜ論語か

(1) 目標：人工知能への論理主義的アプローチをとる。コンピュータは近年、計算機というより 推論機 になってきているから、一つの論理体系がほとんどそのまま一つのプログラムとなりうる。論語の公理化がもしできれば、その論理体系は原理的にコンピュータ上に載る。つまり 孔子 (ないし孔子的精神) が コンピュータ中に実現 できる。それにより 孔子との Q A (質問応答) も可能となる。このテーマは今や完全に 射程距離圏内 にある。

目的：人工知能論としては、人間の 精神 (およびその働きである思想の本質) の 解明 をはかりたい。この課題は最新の技術と最古の思想との結合であり、両者はともに 精神 をめざす。これは工学ではなくて 理学 である。精神の解明それ自体が目的。／本研究は、筆者の基本テーマである「思想の数学化、思想の情報科学的研究」の一環である一文献(1)(2)。

／これは“価値についての価値自由な議論”である。既成の分野としては、様相論理の一種として欧米で研究されている義務論理の話と関連がある。判断において事実判断と価値判断とを区別するのが基本である。事実vs価値は、形容詞としては真偽vs善悪（よい・わるい）の対比となる。真偽の論理学に対して善悪の論理学（広くは価値の論理学）が存在すべきだ。前者は自然科学や数学の論理であり、よく研究されている。後者は我々の日々の行動を決める論理であり、あまりにも身近な論理なのに、さして研究されていない。「よい・わるい」の論理学を確立するのが本研究の核心となる。

（2）なぜ論語か：①ほぼ確定したテキストが、これほど論議され尽くした文書は、他に殆ど例を見ないこと。約五百の短い文章—これは定理集であり、多くの注釈書がその証明を与えていると見なせる。②現実に長期にわたって人々を支配した実績。③その現実主義・論理的整合性（つまり合理主義）と、にも拘らずその奥底にある天命などの非合理主義。その思想が現代の必要に応える所があること。

## II 方法論から枠組みへ：善悪の論理学と実践的推論

（1）論語を公理化する方法。自然科学では自然現象から帰納して法則（仮説）を見出だし、そこから逆に演繹して元の現象や新現象を導出する。我々の前には論語の約五百の文章

があり、これが現象。そこから帰納してその根底にある信念＝公理系を見出だし、それから各種演繹によって元の文章を導出すること。同時にそこで使われている論理の種類を明らかにすること。これが公理化である。公理化ができれば、論語を超えて、もっと多くの命題を導けるが、残念ながら自然科学の〈検証〉にあたるステップはない。／この論理体系をコンピュータ化すれば、孔子とのQ & Aが可能になる。昔の人間を復活再生させるのに、バイオ技術で遺伝子によるクローンを作る方法もあるが、こんな情報科学的な方法もありうる、と言いたい。夢のある研究だと思う。

規範の領域における言葉の〈意味〉の問題。これは、まずは数学と物理学の関係と同じである。初めに基本術語と実世界との対応づけがなされる。理論内部で純形式的に推論が行われ、その結論は基本術語の間の構造的関係となっているから、その関係が実世界と突き合わされる。／但し、現実と理論との間に“食違い”があるときの対処が異なる。食違いがあれば「理論の方を変えるべきだ」というのが物理学であり、食違いがあれば「現実の方を改めるべきだ」というのが倫理である。裁く主体が現実の方か言葉の方かの違いがある。

(2) 「推論の形式化」の歴史—道徳と倫理に関して。アリストテレスは様相の論理も述べたが、〈実践的推論〉の定式

化も行い「規範事例型」と「目的－手段連関型」という2種のあることを指摘した。／人工知能論で例えばSchank&Riesbeck 著『自然言語理解入門』のアプローチ、特にPOLITICSは、価値観の問題を真正面から取り上げていて示唆的である。／近年、様相論理学(modal logic)の一分野として義務論理学(deontic logic)が研究されている。これは、必然( $\Box p = \sim \Diamond \sim p$ )、可能( $\Diamond p = \sim \Box \sim p$ )、不可能( $\sim \Diamond p = \Box \sim p$ )という3つの様相演算子に丁度対応して、人間の義務の領域において、

$p$ であることは義務である ( $O p = p$  is obliged =  $\sim P \sim p$ 、つまり  $p$ でないことは許されない)、

$p$ であることは許される ( $P p = p$  is permitted =  $\sim O \sim p$ 、つまり  $p$ でないことは義務ではない)、

$p$ であることは禁じられる ( $F p = p$  is forbidden =  $\sim P p = O \sim p$ 、つまり  $p$ であることは許されない、または、 $p$ でないことは義務である)

という三つの様相演算子を導入し、類似の公理を立てて論理展開を行うものである。これはまだオモチャのレベルで、しかも規範を〈様相〉として捉えるのが本当に適切かどうかという根本的な疑問がある。

(3) 新しい「善悪の論理学」が存在すべきだ、というのが

筆者の主張である。これは既存の「真偽の論理学」とは別物だが、ある程度までそのアナロジーをたどって行くことになる。その特徴を順次、述べてみよう。

①真偽の論理は基本的に二値の論理なのに対して、善悪の論理は基本的に量的（アナログ的）であり、絶対的に善い、善い、無記、悪い、絶対的に悪い、などの等級が付く。このことは命題  $p$  に対して  $-\infty \sim +\infty$  の実数値を与える評価関数  $f$  を考えることで把握できる。いま  $k$  を非負実数（0でもよい）とし、 $O p : f(p) > k$ ,  $P p : f(p) \geq -k$ ,  $F p : f(p) < -k$ ,  $f(\sim p) = -f(p)$  と定義したとすると、

$$\sim P \sim p \Leftrightarrow \sim (f(\sim p) \geq -k) \Leftrightarrow \sim (-f(p) \geq -k) \Leftrightarrow$$

$$\sim (f(p) \leq k) \Leftrightarrow f(p) > k \Leftrightarrow O p$$

$$\sim O \sim p \Leftrightarrow \sim (f(\sim p) > k) \Leftrightarrow \sim (-f(p) > k) \Leftrightarrow$$

$$\sim (f(p) < -k) \Leftrightarrow f(p) \geq -k \Leftrightarrow P p$$

となり、義務論理学では公理と置かれる式が結果として導出されることになる。規範よりも先に善悪がある、という考え方には根拠がある。／論語においては殆どの場合、一つの事物なり行為なりに対して、それを肯定する価値基準と、それを否定する価値基準とが並存する。言いかえれば、為すべき行為を決めようとするとき人には互いに相反する二つの方向へ力が働く。その二つの力の「中」（中正）を取る、という

形で人の行為が決まる。礼もその精神で定められている。

②真偽の論理においては矛盾が出たら終りだが、善悪の論理では矛盾（衝突という）は当り前、矛盾の解決こそが問題である。例えば、孝の倫理と、親の悪との衝突。解決法としては、一つの価値基準ではどちらの評価値が上かを見る。複数の価値基準に対してはプライオリティ（優先順位）を設定する、などが代表的である。

③ここが肝腎なのだが、真偽の論理では真理性は前件から後件へと流れ出すのに、善悪の論理では善性は後件から前件へと溯る（溯行の原理）。ある命題が〈真〉であるという性質は、「AならばBである、しかるにAは真である、故にBは真である」というふうに、公理集合に源を発して、演繹的な結論のほうに流れ出ていく。“泉”から“水”が流れ出すように。／それと対照的に、ある行為や事物が〈善〉であるという性質は、「AならばBである、しかるにBは善である、故にAは善である」というふうに逆向きに進む。“海”から“河”へと、善なる性質が溯行する。このA→Bには、(a) 原因→結果、あるいは(b) 手段→目的、あるいは(c) 個別→一般、などが来る。(a)の場合のAは予備的善、(b)は手段的善、(c)は個別的善、元々のBは本来的善、と呼びたい。／悪とは善なる状態を損ない、あるいはそこから離反するこ

とである。上と同様に「AならばBである、しかるにBは悪である、故にAは悪である」と、悪性の溯行もある。／ある知人が筆者に「善悪を論じてはいけない」と忠告した。「善悪を論じれば必ず争いになる。争いはよくない。だから善悪を論じるのはよくない」という論法で、それが正に善悪の論理である。／善悪の論理の一体系においては、公理で一つの事物（例えば人々の幸福）に善（プラス）の値を与え、あとは溯行の原理によって次々に多くのものに善悪（プラス・マイナス）の値を付与していく、という形を取る。

④真偽の論理では通常、単一の論理体系・推論方法を用いるが、行動を決めるための論理である善悪の論理は、必然的に総合的となって多数の論理を用い、その統合こそが問題となる。／種類とは別に、精神には「道徳的理想の論理」（スーパーエゴ）、「現実的認識の論理」（エゴ）、「感情の論理」（イド）など、論理を司る主体（モジュール）の違いからくる区別がある。感情は非論理的な何かではなく、感情には感情の論理がある： +の価値を得る（+）と→喜びなどの+の反応がある。+の価値を失う（-）と→悲しみなどの-の反応がある。-の価値を得る（+）と→否定的な-の反応がある。-の価値を失う（-）と→肯定的な+の反応がある。この土の関係は、+と-の掛け算の法則と同じである。

⑤真偽の論理（数学）では数学的構造を研究するが、善悪の論理では精神、社会、世界の構造とその同型を論じる。

⑥既存の記号論理学、数理科学に対して、記号倫理学、論理科学が在る必要がある。

善悪から規範への移行はどうなされるか。「Aは善であり、且つ、 $\sim A$ は善でない」という平叙文から→「Aを為すべきである」という規範文を導出する推論規則を設けたい。この規則には一般的妥当性がある： 英語の構文で、意味的に等価な言い換えに、次のものがある。 Don' t do A, or B. = If you do A, then B. これは右辺の「行為－反価値」結合から、「禁止命令－違反した場合の悪しき結果」という結合を導いている。これに似た議論が法学の分野で見られる：「人を殺せば、三年以上の刑に処す」（要件－効果）とは言い換えれば「人を殺すな、さもなくば三年以上の刑に処するぞ」（禁止命令－違反者への制裁）の意味でもある。なお、日常言語ではく論理的な裏が暗黙のうちに言われていることに留意すべきだ。

（4）各種実践的推論のサーヴェイ： 価値と事実との組み合わせには4通りある。①「事実－価値」結合、②「価値－事実」結合、③「事実－事実」結合、④「価値－価値」結合である。この4つとも論語には出てくるが、中でも大事なものは

①と②、特に①である。／①の「事実—価値」結合は事物に対して+の態度、または-の態度を付与する。この型には「刺激—反応」対や「状況—行動」対が含まれる。②は①の逆であり、「目的—手段」対が含まれる。③は因果的連鎖を与え、④は例えば価値と価値の比較である。／①と②という対比は多くの分野に見出される（詳細は略）。例えば法学での「要件—効果」対を用いた三段論法は①の型、法的道具主義での「目的—手段」思考は②の型である。動物行動ではレスポナント条件反射が①の型、オペラント条件反射が②の型である。これら二つの型は人間にも動物にも通底しており、動物も人間も基本は同じか、と思わせられる。

孔子思想の中心である仁と礼は②の型に属する。礼とは制度慣習法、儀式や礼節、今でいえば法律に近いものである。孔子において、根本的な善とは「人民の均しい幸福」であり、それを実現しようとする意志が仁である。ただ、人々の幸福のためには「秩序」が必要であり、秩序を与えるものとして礼がある。つまり〈目的〉は仁、その〈手段〉としての礼、という論理構造となっている。

### Ⅲ 「論語の論理」の実際—具体的な文章に沿って。

(1) 論語の基本的な命題構造は、「事実—価値」結合である！ 學而第一の1「子曰く、學びて而して時に之を習ふ。

亦説ばしからずや。／朋遠方より來る有り。亦樂しからずや。／人知らずして慍みず。亦君子ならずや」。これは三つの節からなり、各節の基本構造は「事実—価値」対、つまり事実を価値に結びつけるものとなっている。第一節「学習は悦ばしい」にあって、学習は単なる行為であり事実領域に属する。悦ばしいとは価値感情であり、価値の領域に属する。この文においてそれら事実と価値とを結合しているのである。結合の資格は何か。まずは経験であり、私は悦ばしい、あなたもそうではありませんか、と勧誘する。この感情を、+の価値を獲得（+）することは+の感情を引き起こす、という「感情の論理」から説明もできる。さらには論語の体系の中でこの「事実—価値」結合を導出することもできる。／細かい分析に入るよりも前に、エレメンタリーな構造が「事実—価値」結合にほかならないことに注目したい。これこそが論語解明の糸口である。論語注釈の二千年の歴史で、いまだこの観点が提出されたことがないのは不思議な感じがする。この認識が筆者に「論語の現代的研究」の意義を確信させた。

（2）「同型」の原理。その例：學而第一の2「有子曰く、其の人と爲りや孝弟にして、而して上を犯すことを好む者は鮮し。上を犯すことを好まずして亂を作すことを好む者は、未だ之有らざるなり。君子は本を務む。本立ちて道生ず。孝

弟やは、其れ仁の本と爲すか」(弟子たちの発言も孔子思想を表すものと考えられる)。家族という階層秩序と、社会の階層秩序との、順序構造としての同型性が認識されている。人の生い立ちにおいては家族から社会へ、と進む—これが人間愛の発展である。価値付けの論理としては逆向きに、社会から家族へ、と進んでいる—すなわち、社会的秩序に価値がある、故にそれを守り従う心性(mentality)に価値があり、それを用意する手段として、家族内の倫理=孝が善しとされている。///以下、略。詳細は別の機会に述べたい。

終わりに この「論語の情報科学的研究」は徐々に完成に向かっており、これから段々と発表してゆきたい。

参考文献 (1) 高橋英之著『コンピュータの中の人類』

(御茶の水書房) …ソフト宇宙論=キリスト教的世界観の数学的基礎。これについては92年8月の日本SF大会で招待講演を行い、好評を得た。/(2) 高橋英之著『思想のソフトウェア』(法藏館) 93年4月に出版予定—これは古今東西の思想宗教を並べるなかから、思想の“周期律表”を見出し、その“元素分析”に進み、さらに“思想の合成法”にまで至ろうとする試みである。この本によって大きな(トップレベルの)見取り図ができたと考える。

(3) 簡野道明『論語解義』明治書院、ほか多数。