

道徳起源論から進化倫理学へ（三）

内井 惣七

第二部 規範倫理学における還元主義 続

21 合理性——最大化か満足化か？

さて、還元主義のプログラムにとつてまず有望に見えるのは、「あるべき道徳性」を合理性を軸にして構成しようというアプローチである。このアプローチの基本的な構想は次のように述べることができる。

「われわれが現にもつ道徳は、必ずしも体系的ではなく、時には衝突さえ起こしうる規範や義務や価値の寄せ集めである。また、個人のもつ選好には、多くの場合、人それぞれのバイアスがかかっており、それゆえ好みや判断の違いが生じる。しかし、人びとが現にもつ選好ではなく、合理的だといえる選好を基準にして考えれば、バイアスや不一致は相当程度取り除くことができ、規範や価値の体系化も可能になる。あるべき倫理とは、このような合理的見地から受け入れられるという意味で正当化できるものである。」

この構想において、「合理性」の規定が一つのカギになることは明白であるが、その規定のうちに「道徳性」を前提せずには済ませることができれば、還元主義を貫くことは可能に見える。つまり、（あるべき）道徳性は選好その他の非道徳的条件の組み合わせと合理性に還元できると予想されるのである。しかし、規範倫理学で使用すべき「合理性」につ

いては、倫理学者の間でも一般的な合意があるわけではなく、いくつかに立場が分かれる。おそらく、最も重大な立場の分岐は、(1)十分な情報のもとで機能しうる合理性の概念をとるか、それとも(2)十分な情報しか得られない状況のもとで機能しうる、もつと現実的な合理性の概念をとるか、という選択肢で表現できよう。このような分岐を近年問題にし、(2)の選択肢の重要性を力説してきたのは経済学者のハーバート・サイモンであり、彼はそれを「限定された合理性」と名づけて擁護してきた。他方、倫理学者の伝統では(1)の方をとる人々が多いように思われる。例えば、シジウィック、ブランド、ヘアといった功利主義の倫理学者は(1)の立場を採用する。しかし、最近ではギボードやデネットのように(2)の選択肢を追求しようとする動きもある。そこで、まずこの分岐が避けられないものかどうか、調停の可能性があるかどうかを検討してみたい。

ノーベル賞学者サイモンの言う「限定された合理性 bounded rationality」とは、意志決定者のもつ認知的な限界を考慮に入れた合理的選択を表す言葉である。認知的な限界とは、限られた知識および限られた予測能力を意味する。個人や企業の経営行動に主たる関心を持っていたサイモンは、経済学における行動主義的なアプローチをめざし、個人や企業の実際の意志決定のプロセスを解明できるモデルを求めた。

数学的な意志決定の理論で支配的な考え方は、主観的確率と効用の概念を使ったベイジアン（ベイズ主義）の路線であり、

- (a)可能な選択肢のそれぞれについて、
- (b)得られる結果の価値とそれぞれの確率（配分）がわかっており、それに基づいて
- (c)意志決定者の期待効用（結果の価値と確率とを掛けて合計した期待値）を最大化するような選択

が合理的だとされる。

主観的確率の代わりに何らかの「客観的」確率を認める立場でもこの最大化モデルは使用可能である。しかし、この

モデルはしばしば非現実的で、われわれの実際の選択状況に適用するには不適當であるとサイモンは言う。例えば、(a)については、実際の選択に際して可能な選択肢をすべて数え上げることがまず不可能であるし、(b)の確率配分もわからないことが多い。したがって(c)の期待効用の最大化も、不可能であるだけでなく意味さえ持たない場合が多いのである。当然、人々の実際の選択行動を記述するために、このモデルは仮定が強すぎて使えないことが多い。

そこで、サイモンはこれら三つのいずれの条件も、より現実的な条件で置き換えた独自のモデルを提出し、限定された合理性を規定しようとする。

(a')まず、選択肢はあらかじめ決まっているのではなく、少数の選択肢が生み出される過程を考えるべきである。

(b')次に、確率配分はわからないので、それを推定する過程、あるいは確率の知識を前提しないで不確実性に対処する方策を考えるべきである。

(c')そして、最後に、最大化原理ではなく「満足化 *satisficing*」の原理を合理性の核と見なすべきである。

「満足化」とは、最大化（当面は最適化と同義に理解する）が不可能であるにしても計算上のコストが大きすぎる場合に、最善の選択肢ではなく満足のいく選択肢を求めることである。例えば、大きな靴の卸屋で自分のためのスニーカーを一足探したいとき、自分の好みや必要性、耐久性、値段などを考慮して最善の（最も満足のいく）選択を行うことはまず不可能である。選択範囲が有限であるにせよ、最善を求めるためには決まったサイズのスニーカーをすべてのメーカーにわたって調べなければならず、そのためには多大な時間を要するであろう。しかし、半時間程度の限られた時間内でも気に入った商品を探し出すことは可能である。このとき、わたしは、例えば過去の経験から知っている二、三のメーカーの製品に的を絞り (a')および (b)、適当な満足度のレベルを設定し、それをクリアした数個程度の選択肢の中から一つを選ぶ (c') という形で満足のいく選択をおこなえばよいのである。現実的な合理性はこのような形を取るとサイモンは言う。

最大化と満足化とが理論的に必ずしも排他的な関係にあるのではないことは、すでにサイモン自身によって指摘されている。

形式的には、満足化の過程は、探索のコストを考慮に入れ、別の詳細な探索から期待される利益がその探索の機会コストとちょうど等しくなる時点まで探すということにして、いつも最適化の過程に変換することができよう(……)。しかし、この変換は、選択者に、おそらく過重な、情報および計算の負担を強いる。すなわち、探索から期待される限界収益と、機会コストを見積もるといふ負担である。これらの見積もりの問題を解くことは、もとの選択の問題を解くと同じほど、あるいはもっとむずかしいかもしれない。(Simon 197b, 286)

要するに、現在手に入れている選択肢に甘んじてこれ以上の探索を怠ることは、もし探索したなら手に入る余分な利益を失うことになるかもしれない。この(失うであろう余分な)利益が現在の選択肢の「機会コスト」である。そこで、探索を続けるかどうかの決定は、期待される利益と機会コストのバランスで最適化の枠内で扱える。こういった考察で満足化の選択は最適化に変換できる。しかし、満足化がこのようにして理論的に最適化に変換できるとしても、それは実際には(選択の当事者に)余分な計算のコストを導入することになるので、現実的な代替策を提供するわけではなく、有望な現実的な手段は満足のいく選択肢が見つかるまで探すという(満足化の)方法である、とサイモンは論じている。しかし、この引用文のうちに、二種の合理性を調停する重要なカギが含まれている。この点については後に再論する。

22 進化における「最適化」

さて、人々の現実の選択がどのように記述できるかに関して、サイモンの研究はなかなか説得的である。他方、進化倫理学をめざすわれわれにとって、最適化を軸にした合理性の考え方も、なかなか捨てがたい魅力を持っているように見える。なぜなら、自然淘汰による適応はある種の最適化の過程だと解釈しうるからで、進化を扱うのに数学的なゲーム理論が近年成功を収めてきたことがこの解釈を支えている。もちろん、進化の説明は事実に関わる問題であり、選択をいかにすすべきかという選択の合理性という規範的な問題とは別次元の話である。しかし、ベイジアンという最大化と進化論の最適化の論理構造はきわめて類似しているので、少なくともアナロジーとして、進化における最適化は合理性を考えるときの有力なヒントを提供しうる。この点はサイモンにも気になったとみえ、彼がスタンフォード大学で一九八二年に行なった講演（邦訳、キムバク、1987）では、進化の考察にかなりのスペースが割かれている。

前節で解説したように、合理的選択の問題では(a)選択枝の生成、(b)選択枝の評価の仕方、(c)そして選択の原理という三つの条件があった。自然淘汰による進化の過程では、これらに対応して、選択枝に相当するのはグループ内での変異のタイプ（表現型または遺伝子型）であり、選択枝の評価に相当するのはそれぞれのタイプの適応度の違い、そして選択の原理に相当するのは、適応度の違いが長い間には個体数の違いに反映されて最適なタイプが生き残るといふ形の最適化原理である。しかし、サイモンが強調するのは、選択枝の生成には進化でも大きな制約があり、進化の最適化原理はそのような制約のもとでただか局的な最大値に到達するにすぎないという点である（キムバク、1987、70）。したがって、彼によれば、進化の最適化はベイジアンの最大化原理よりも限定された合理性のモデルに近い（キムバク、1987、96）ということになる。以下でこの主張を紹介してみよう。

サイモンの見るところ、（自然淘汰による）進化は、近視眼的な性格を有し、与えられた条件の下で当面の短期的なもの（もちろん、地質学的年代のスケールで）有利さによって働き、適応を成し遂げる。まず、進化の素材となる変異のタ

イプが(a)の選択肢に相当するが、これらの素材は現に種内に蓄積されている遺伝子のコンビネーションと、時たま生じる突然変異とに限定されている。与えられた環境に対して、あらゆる可能性のなかから最適な生物が選り出されるのではなく、現に存在する競争相手に相対的に、このように限定された資源を利用して対応するのが適応の実態にほかならない。かくして、選択肢の生成には限られた情報しか使われていない。次に、その結果達成される最適化は、文字通りの最適化ではなく、このような限定された条件の下での局所的な最適化にすぎない。この点をサイモンは次のようないま比喩で表現している(キヤホン1987, 69-70)。

単純な地形で丘が一つしかない世界では、坂を登り続けて行けばその世界の最高地点に到達する(最適化)。しかし、高台や窪地が多く散在する複雑な地形の世界では、坂を上って到達した頂点は、低い丘の頂点かもしれない。この世界の最高地点である保証はどこにもない(局所的な最適化)。進化は適応度の差を介して盲目的に作用する淘汰である(つまり、どこであれ坂を上るようなもの)から、適応はいつも相対的で局所的でしかない。いったん平衡状態に達していったように見えた状態が、帰化生物が入ると崩れて新たな平衡状態に移るといふ、ダーウィン自身が気づいていた事実もこの点の証拠である、とサイモンは論じる(キヤホン1987, 71-72)。

さて、自然淘汰による進化についてのサイモンの理解に取り立てて問題はない。しかし、このような議論で「最適化」の意義が否定され、ひいては合理的選択の最大化(最適化)モデルの意義が否定されたと見なすのは早計である。サイモン自身の論旨は必ずしもそのような否定を表明しているものではないが、サイモンの権威に訴える論者のうちには、最大化モデルの意義を否定するためにそうする者がいるので、次節でサイモンの議論の意義を検討してみよう。

23 最大化モデルは不要か

まず確認しておかなければならないのは、サイモンが彼の限定された合理性のモデルと対比させた、ベイジアンのも

(あるいはゲーム理論や理論経済学の) 最大化モデルの理解の仕方である。すでに引用した一九八七年の翻訳書では、これに合理性の「全知全能モデル」(サイモンの原語は *olympian model*) という名前がつけられたこともあり、ある種の誤解を助長したと考えられる。数学的な意志決定の理論においては、理論を完結した形にするために、「可能な選択肢」がすべてわかつており、不確実な可能性のすべてについて確率配分が与えられたと仮定する。しかし、この数学的理論を具体的な事例に適用するに際しては、理論あるいは理論的モデルと具体的事例の間の大きなギャップのため、適用条件や初期条件についていくつもの単純化条件や限定条件がつけられる。これは、意志決定の理論に限らず、物理学をはじめとする経験科学では常に行なわれることである。例えば、大砲の弾の弾道を計算するとき、ニュートンの運動方程式を用いるにしても、弾の質量、地球の引力、あるいは空気抵抗などの条件は理論によって決められるわけではなく、理論適用のために別に補わなければならない。

企業や人々の「合理的な意志決定」を記述しようという目的で最大化モデルを用いるときにも事情は同じである。現実の事例では、「全知全能」はもとより、適度に「十分な情報」さえ欠けているので、きわめて大胆な単純化ないし理想化を(理論の外から)補わなければ、最大化モデルを現実で適用できる形にはならない。そこで、明らかに、「理論的な最大化モデル」と「具体例に適用された最大化モデル」とは、異なる二つのレベルに位置している(ニュートン力学の場合もわかり。例えば、空気抵抗が大きくて複雑に変化する場合、弾道計算は相当やっかいである)。そして、サイモンが長年力説してきたように、最大化モデルの具体的事例への「適用モデル」は多くの場合使い物にならないほど不十分であった。その点で、サイモンの「限定された合理性」のモデル(これも適用モデル)の方がはるかに成功を収めてきたことは異論がない。しかし、他方で、サイモンは(21節で指摘したように)このモデルが最大化モデルの枠内で(理論的には)再構成されうることも認めている。この事実を、「再構成」と「具体的適用または使用」とを混同しないかぎり、重要な示唆を与える。

限定された合理性に基づく決定は、その場その場の制約を強く受けるものだから、そういった決定が数多くなされた場合、不調和や矛盾が生じない保証は全くない。まったく同じような状況で、二つの相反する選択がともにその場の「合理的決定」としてなされることもあり得る。しかし、科学という知識の営みと同様、人類が文化的な営みとして展開してきた道徳においても、無矛盾性、一貫性、あるいは普遍性は重要な要件として認められている。したがって、二つの相反する選択がともに合理的と見なされる事態は批判の対象となり、少なくとも改善の余地があるものと見なされる。では、限定された合理性に基づく個々の決定をもう少し体系化し、改善の余地があるものとして許容しうるような「合理性」はどのような形を取りうるだろうか。こういった問いの文脈では、最大化モデルは依然有力な候補である。限定された合理性の「限定」の程度がゆるやかになる（例えば、情報量が増える、計算能力が増える、など）につれ、このモデルは最大化モデルとの類似性を増す。とくに、理論的に限定合理性のモデルが最適化モデルの中で再構成できるといふ事実は、二つが階層構造をなして一つに統合可能であることを強く示唆する。つまり、具体的事例への適用レベルでは最大化モデルが多くの場合使い物にならないとしても、逆に限定された合理性の個々の事例を改善したり体系化するという別のレベルでは、最大化モデルの意義は十分にあり得るのである。

この点は、次のような考察によって擁護することもできる。限定された合理性のモデルにしたがってわたしが選択を行ったとしよう。ところが事後にわかった新しい情報により、実際に選ばれた選択肢Aよりも、考察はされたが選ばれなかつたBの方がより満足のいく結果をもたらしたのであるとわたしは確信した（これはよくある経験である）。このとき、「もつと情報を集めて（あるいは、そうしていたなら）Bを選んだ方がよかつた」とわたしは事後に判断する。そしてこの事後の判断はしばしば将来の指針となることもある。こうした形で、限定合理性に基づく「満足化」には常に改善の余地がある。そして、この「改善」に意味を与えるためには、何らかの形の最大化モデルが適切である。なぜなら、二つの選択肢のうちで「よりよい方を選ぶべし」という規則は最大化原理の別の表現に他ならないからである。

「全知全能」を仮定せず、限られた情報のもとでも、この形の最大化原理の意義を否定するのはきわめて困難である。とくに、人々の選択行動を記述あるいは説明するという文脈から、「いかに選択すべきか」という規範的な文脈に移ったときに、この形の最大化原理の意義と説得性は大きく異なる。規範的な文脈で問題になるのは、当然、あるいは「よりよい」形の選択原理であり、現実を用いられる選択原理に改善の余地があるのに、それを使用すべき原理として推奨するのは「不合理」であろう。

次に、進化についてのサイモンの考察を検討してみよう。彼の指摘のうちで最も重要なのは、自然淘汰による最適化は常に与えられた条件（表現型あるいは遺伝子型と環境）に相対的な局所的最適化であるという点である。したがって、進化の最適化は「全知全能」の仮定のもとでの、無条件の、あるいは全体的な最適化ではない。この点にまったく異論はない。しかし、進化論において、あるいは合理的選択の現実世界への適用において、いったい誰がそのような無条件で全体的な最適化を主張しているのだろうか。

ダーウィニアンは、まず間違いなくサイモンの指摘に同意するであろう。合理的選択の最大化モデルの支持者も、大半は、現実的事例への適用においてサイモンがいうような無条件で全体的な最適化が可能だとは主張しないであろう。すべての可能性が尽くされているというのは数学的な理論の枠内の話である。その理論が多少なりとも具体的な事例に適用され、具体的な最大化モデルが考察されているときは、常にその具体例の限定条件が前提されている。したがって、進化の場合と同様、このモデルでの最大化は、その限定条件内での局所的最大化にすぎない。ただ、限定された合理性のモデルを信奉し、それ以外は必要がないという論者と異なるのは、最大化モデルの支持者は、その限定条件が緩和され、より一般的なモデルが得られた場合には、古いモデルを新しいモデルの中に統合して、改善された選択が可能であり、「よりよい選択」が可能となればそうすべきだ、と主張するにすぎない。

サイモンが言うとおり、局所的な最適化あるいは最大化は全体的な最適化の十分条件ではない。しかし、局所的な最

適化を通じてしか、「改善」という意味での、より広い範囲での最適化はあり得ない。最大化モデルの支持者は、「全知全能」の仮定のもとでの最善を目指せと言っているのではなく、「局所的な最大化を通じて改善を目指せ」と言っているにすぎない。

24 デネットの「倫理的応急処置」

以上的一般論にある程度肉づけするために、限定された合理性を信奉するデネットの倫理学批判を検討し、逆に批判しておきたい。『ダーウィンの危険な考え』(1995)第17章で、デネットは帰結主義(典型的には功利主義)、非帰結主義(例えばカント)のいずれのタイプの倫理学も、実際の倫理問題を解くにはほとんど無力であると痛烈に批判する。

哲学者は、われわれが実際には情報の不足や見落としや偏見によって誤った倫理的結論に達することが多いと認めるが、原理的には理想的倫理学説によってなすべきことが決まるはずだという。しかし、例えば功利原理に基づいて、しかじかの行為がこれだけの善をもたらすだけ、差し引きこれだけの利益があるからこうすべきだ、というたぐいの計算(費用・効用分析)は、いまだかつて倫理学において行なわれたことがない、とデネットは断言する(Darrett 1995, 408)。天文学の原理に基づいて十分な計算を行なえば実用的な航海歴が得られるように、功利原理と人類の長い経験により、何が益をもたらし何が害をもたらすかは多くの場合人々にわかっているのだというミルの答えは、少し真剣に考えたと説得性を失う。例えば、スリーマイル島の原発事故が人々にとって良かったのか悪かったのかという(善の量的比較ではなく、プラスかマイナスかだけの判断に限ってさえ)問いにさえ、答えは確定しない。なぜなら、チェルノブイリほどの大事故ではなかったから、人々に原発の危険性を認識させて予防措置をとらせるのに貢献した点では善であり、ある程度の被害と心配をもたらした点では悪であるにしても、すべてを考慮し、長い目で見た場合全体の善悪のバランスはどうなるのか計算のしようがないからである(「スリーマイル島効果」とデネットは言

う。Denett 1995, 498。これについては異議があるが、議論の本線からはずれるのでここでは論じない。

もつとも、こう言ったからといって、ライバルのカント的非帰結主義の株が上がるわけではない。定言命法の適用も、どのような格率（行為規則）をテストにのせるかについて、外から自由に仮定を持ち込まないと「なすべき」行為の決定には至らないので、功利主義に勝るわけではない。

そこで、デネットの提言は、「実際問題」と「原理問題」とを分けた上で後者のみを扱って前者を無視するという大半の倫理学者のやり方を離れ、現実の倫理問題がどのように扱われて決定に至るか、われわれの実際のやり方を確認してみようではないかということになる。それは、サイモンにならった一つの見識である。ところが、デネットが提出するのは、彼が批判の対象とした哲学者とあまり変わらない、架空の——それもかなり「非現実的な」——事例に即した分析である。ある大学の哲学科が寄付金を得て、国内で公募し、もつとも優れた大学院生に12年間のフェローシップを一つ与えることにしたところ、締め切りまでに二十五万人の応募があった。選考委員会は、この中からどのようにして「最善の」候補者を選ぶか、という問題である。

さて、この「架空の」問題を示されたデネットの友人たちは、大同小異で次のような方策 (Denett 1995, 501) を提案する（ここが「現実的」方策、と言いたいのであろう）。

(a) チェックしやすく、優秀さの目安となる基準を少数選び出す。例えば、平均グレード・ポイント、哲学の履修科目数、応募書類の重さなど（軽すぎるもの、重すぎるものを除外する）の基準によって、最初の選抜を行う。

(b) 残った応募者につき、くじ引きによって、調査可能な数、例えば五〇人とか一〇〇人にまで無作為に選考対象者を絞る。

(c) 残った応募者の書類を委員会で注意深く検討し、委員会の投票で最終候補者を決める。

このような手続きで最善の候補者が選ばれる可能性は少ない。しかし、限られた時間と手間ですべてのいく結果を得る

には、こういったたぐいの手続きがよさそうに思われる。この例は、架空であり幾つかの点で誇張はあっても、現実の倫理問題に対処する場合にみられる特徴を十分に捉えているとデネットは言う。

(1)与えられた、限定された問題についてさえ、「すべてを考慮する」ことは物理的に不可能である。
 (2)さらに考察を限定するため、適当な規則を使った絞り込み(a)が行なわれる。このような規則に誤りがない保証はない。選考のための便宜と規則の信頼性の間で妥協がなされる。

(3)くじ引きの導入は、「選考」の代わりに別の要因による選抜を一時的に導入し、しかしその結果については「責任」をとることを前提している。

(4)問題を過度に単純化した後で、その結果に対してある程度改善を加え、受容可能な最終結果を得ようとする努力がなされる。

(5)得られた最終結果については、再考の余地や後知恵による批判の余地が限りなくあり得る。しかし、結果は結果として、別の課題に向かう。

以上の特徴は、サイモン言うところの「限定された合理性」のデネット版であることが明らかであろう。選択肢の生成は(a)と(b)でなされ、選択の原理は最適化ではなく満足化ではない。さらに、デネットは、「理想的な合理性」をいえるような単一の視点を想定するのは誤りであると主張する(Denett 1995, 501)。われわれは、現にわれわれが使用する合理性をもってしか問題に対処できないので、それが抱える制限をすべてとり払った想定のもとでの解決策は、安直すぎて参考にならない、とデネットは論じる。

では、デネット自身の提言は結局どういうところに落ち着くのだろうか。議論というよりレトリックが目立つ何べいジかを費やして彼が言いたいのは、不完全な「倫理的応急処置」はいくつか異なる処方があつてかまわない、そしてわれわれ自身が新しい処方を書き、倫理的行為者のデザインを改訂しようと務めることで、自他に課された問題によりよ

い解決策を探すことができる (Denett 1995, 510)、「ということらしい(そして、人類は長い歴史を通じてそうしてきたのではないのか?)。しかし、この文脈での「よりよい」とか「改善」とかいう言葉は何を意味するのだろうか。もちろん、これも「全知全能」の観点からみた価値判断ではなく、限定された合理性によるメタレベルの判断にすぎない、しかもこれが最適化に至る保証はどこにもない、とデネットは答えるであろう。

ところが、すでに指摘したように、進化の最適化を主張する人々も、合理的選択の最大化モデルを支持する論者たちも、以上の点に関して実質的な異議を唱える必要はないのである。しかし、メタレベルでの「改善」の判断にコミットし(これは局所的な最大化原理を認めることにほかならない)、下のレベルの判断の調停やながしかの体系化をめざすかぎり、最大化モデルの中心的な意義は復活している。功利主義や帰結主義に対するデネットの時折の嘲笑にもかかわらず、彼がたどり着いた結論(めいたもの)は功利主義と十分に両立する。例えば、彼が「規則尊重」の意義について論じ、それがよいのは「一つの規則または一群の規則が最善であると証明されているから、あるいはいつも正しい答えを出すからではなく、規則をもっていることによりとにかくうまくいく——そして規則をもたない場合には全然うまくいかない——からである」(Denett 1995, 507)と主張するとき、彼の立場はある種の規則功利主義と変わるところがない(うまくいく、とは規則採用の帰結でなくていったい何だろうか?)。最大化原理を現実の事例に適用する場合、人間の限定された能力を考えて規則遵守や習慣づけられた性向(アリストテレスの言う倫理的徳)を通じた行為の規制を考えなければならぬ、というのはミル、シジウィック、ヘア、ブランドらの功利主義者が一貫して行なってきた主張である。「うまくいく」ことだけでなく「よりうまくいく」こと(改善)をめざせば、デネットの主張はたちまち功利主義となる。

以上のように、デネット版の「限定された合理性」の主張をみても、これと最大化原理とは異なるレベルで共存できるだけでなく、限定された合理性の意義を少しでも擁護しようとするれば、ある種の最大化原理がどこかで前提されなけ

ればならないことが判明するのである。これは、サイモンが指摘したように、ダーウィンの進化が局所的な最適化の過程と理解できること、そしてデネットが「ダーウィンの危険な考え」の視点から人間の倫理的意志決定を眺め直そうとしたことを考えあわせると、(わたしに言わせれば)当然の結果であるといえよう。デネットの隠されたメッセージは、おそらく、「倫理的決定の連続によるこの世の改変は、最善に向かつての進歩の過程とは見なせない。それは、局所的な適応による進化が進歩とは基本的に異質であることと同様である。倫理とは、局所的または近視眼的視点からの満足化の過程にすぎない」ということであろう。しかし、これをすべて認めて、デネットと同様に規範的意味での「改善」にコミットするならば、功利主義が主張するような(善の)最大化原理は認められたも同然である。この原理を規範的価値判断として具体的に適用していくためには、常に現実の事例や状況によって課せられた限定を免れることはできない。したがって、合理性の最大化モデルの支持者の現実的な主張は、(われわれに可能な)「局所的改善を通じて最大化を目指す」ということになる。

かくして、以上の考察に誤りがなければ、「合理的」選択を軸にした規範倫理学の還元主義のプログラムは、少なくとも「合理性」概念の選択レベルで挫折することはない。

25 普遍化可能性と合理性

近年広範に論じられている「合理性」について、まだ論じるべきことは尽きていないが、以上の成果をもとにして得られる新たな洞察に話を進めたい。小論17節から19節にかけて普遍化可能性の問題を論じたが、合理性についてのここまでの考察により、普遍化可能性を成立させる条件について、一つの重要な示唆が得られる。

第17節の末尾で、わたしは進化論の網にかかるかぎりでの道徳性は、集団内での「普遍化」しか要求せず、ヘアの言うような厳密な普遍化は文化的な洗練が加えられたものと指摘した(集団内での「普遍化」が個体への言及を含みう

ることに注意を喚起するため、このようにかっこ付きで表現する)。したがって、ヘアが形式的・論理的と見なしたこの条件にも実質的価値判断が潜在しており、その正当性（それが正当化できるとして）を言う必要性がある。少なくとも、わたしが目指す還元主義の規範倫理学ではそういうことになる。ここでは、「べし」あるいは行為の「正しさ」の普遍化可能性に話を絞ろう。もちろん、これがある程度の実質をもつ価値判断を含むことになれば、無条件の正当性を言うことはむずかしい。しかし、どのような条件の下でその正当性が言えるかが明らかになれば、これは哲学的分析における一つの重要な収獲となる。

さて、わたしの見当によれば、合理性、つまり限定された合理性の一つの改善の方向から普遍化可能性（むしろ、「普遍化」の範囲の拡張）が生まれるのである。生物学的「利他性」の起源について、現代の進化生物学の知見に大過がなければ、人間の道徳性は相互的利他性の一つの形である。しかし、この相互性は普遍化可能性にはほど遠い、ただか一集団内に限られた相互性でしかない。その集団の一個体にとって、相互性の考察を身近な仲間や集団の外の人々にまで拡張することには、どのような利益があるのだろうか。サイモンやデネットが強調するように、われわれの日常的な倫理的決定の大多数は近視眼的であり、限定された合理性に基づくにすぎない。ということは、感情移入や共感の対象とされ、その限りにおいてわれわれの倫理的考察に入る人々はきわめて限られているということである。しかし、限定された合理性は改善可能である。少なくとも、きわめて強い限定から、もつとゆるめられた限定のもとで事態を考えることができる程度には改善可能である。サイモンの比喩を借りるなら、小さな丘が一つしか視野に入らない観点から、二つか三つ程度まで見渡せる観点に移ることは、時には可能である。その場合、もつとも近い丘に登るより、もう一つ向こうの丘に登った方が高い地点に到達できることが、しばしばわかるのである。もちろん、このような改善を拒否する選択も可能である。しかし、狭い観点からの満足化だけにこだわることは、相応の機会コストを支払う——観点を広げたなら手に入るであろう余分な利益を失う——ことになる。

非常に図式的にすぎるかもしれないが、社会生活での交渉相手の範囲を広げ、それに伴って倫理的な考察の対象とされる人々の範囲を拡張する、つまり相互性あるいは「普遍化」の考察を広げることには、それに見合う利益が（もちろん条件次第で）期待できるのである。人類の原始的な生活から文明的な生活への移行を可能にしてきたのは、こういった条件ではなかっただろうか。つまり、簡単に言えば、進化によつて獲得された倫理的性向に、人類の長期にわたる経験的な知恵が加われれば、限定された合理性の改善と、それに伴う「普遍化」の範囲の拡張とは、ダーウイン的進化の原理と同じ路線で理解可能となる。したがって、普遍志向の哲学者が、このような過程のもつとも洗練された成果を取り上げて（厳密な）普遍化可能性と見なした条件は、実は程度の差を許容しうる条件となる。このことは、現実のわれわれの倫理においても確認できる。小さな社会や民族の垣根が取り払われて倫理的「普遍化」が拡張される事例と同様、社会的・政治的事態によつては、人々の判断や行動原則の適用範囲が再び縮小され排他性が支配的になる実例は、歴史的にも、現在の世界でも、枚挙にいとまがない（以下の29節を参照）。

もちろん、倫理的考察に入れる人々の「範囲を広げる」とか「排他性」とかいう条件はヘアのいう「普遍化可能性」と概念的にただちに同一視できたり矛盾したりするというわけではない。しかし、例えば(A)肌の色の違いによつてある種の人々を倫理的考察から閉め出す（これは、個体に言及しない基準による線引きだから、厳密に普遍化可能である）という事態と、(B)ある種の人々がわたしを帰属する共同体の外の人だから倫理的考察から閉め出す（これは、個体に言及する線引きだから普遍化可能性に抵触する）という事態とは、実際の倫理的思考（とくに、ホッネのレベルで）ではそれほどきれいに区別されるわけではない。すでに触れたように、あらかじめ普遍化可能性を「倫理的」あるいは「道徳的」という言葉の意味の中に組み込んでしまえば(B)を禁止することは簡単になるが、なぜそうすべきかという問題は「なぜ道徳的であるべきか」という問いのレベルに押し戻されたにすぎない。そこで、わたしは、ここでは(A)(B)二種の事態を区別しないという観点から論じていることを銘記されたい（進化倫理学は類人猿の「倫理」も有意意味だと認める

のである)。それゆえ、「普遍化」の問題が「範囲を広げる」とか「排他性」という言葉に置き換えられて記述されているのである。そして、普遍化の問題を実質的な合理性（単に論理の問題ではなく、実質的内容のある選択が理になっ
ているかどうか）の枠の中で扱おうという姿勢において、わたしのアプローチはヘアよりもシジウィックに近づいてい
る。

26 コミットメント関係と信頼

さて、すでに提示されたわたしの考察が、哲学者風の思弁的なスケッチにすぎないことは認めるにやぶさかでない。
しかし、数少ないとはいえ、こういった考察に光を投げかける分析的・実証的研究も現れている。わたしが重要な示唆
を受けたのは、社会心理学に進化的視点を取り入れた山岸俊男氏の研究（E 卍 1988）である。彼のテーマは人間関係に
おける「信頼」の問題であって、「普遍化可能性」といった抽象的な問題ではない。しかし、倫理的考察の対象とする
人々を、何らかの基準（必ずしも普遍的ではない）に基づいて狭くとるか広くとるかという問題は、社会的交渉相手の
範囲の大小、排他性、開放性といった問題とすぐにつながってくるので、そしてこれらの問題は「信頼」およびその対
立概念と密接につながっているので、彼の考察と成果とは倫理学にも重要なかわりをもつ。

山岸の研究は、人々の間の、あるいは組織の間の協力的な関係を可能にする「信頼」の概念を分析し、日米の比較調
査などの実証的裏づけもとりながら、信頼の意義を明らかにする独自の説、信頼の「解き放ち」理論を打ち出すもので
ある。この説は、「信頼には関係を強化する側面……と同時に、……関係を拡張する側面があること」（E 卍 1998, 55）
を指摘する。この「関係拡張の側面」が、わたしの「普遍化可能性」の問題と関係する部分にはかならない。

さて、この「解き放ち」理論を簡単に紹介しておこう。これは六つの命題よりなる。

(命題1) 信頼は社会的不確実性が存在している状況でしか意味を持たない。つまり、他人に騙されてひどい目にあう可能性が全くない状況では、信頼は必要とされない(三冊1998, 61)。

ここで言われる「社会的不確実性」とは、交渉相手の意図について情報が不足しており、騙されて損をする可能性がある状態を指す。例えば、中古車の市場は、新車の市場よりも一般に社会的不確実性が高い。なぜなら、販売会社の意図について情報不足の程度は同じでも、中古車の方が隠された故障などがある可能性が高く、したがって騙されて損をする可能性が高いからである(画, F10)。また、山岸が問題にする「信頼」とは、自然的秩序に対する期待(例えば、裏山が崩れないと信頼する)ではなく、何らかの倫理的秩序に対する期待の方に分類され(三冊1998, 33)、さらに、社会的関係や社会制度の中で出会う相手がある種の役割を果たすことができるという能力に対する期待ではなく、相手の意図に対する期待の方に分類される(三冊1998, 35)。簡単に言えば、交渉相手が義務や責任を果たすであろうという期待を「信頼」という言葉で表している。

そこで、社会的な不確実性に対処する一つの方法は、相手を信頼して「主観的に」不確実性を低下させることである。もちろん、騙された場合には損が大きくなるかもしれないが、相手が信頼に値する人間であった場合には得られる利益(精神的満足も含めた広い意味)も大きい。

もう一つの対処のしかたは、社会的な不確実性を何らかの仕組みを導入することで「客観的に」取り除こうとすることである。例えば、中古車の場合であれば、一年以内に故障があれば無料で修理をするとか交換するとかの「保証」を販売会社から取り付ければ、買い手は「安心」できる(山岸はこの「安心」を信頼には含めない。三冊1998, 37)。あるいは、商取引などの場合は、(実績のある)特定の相手のみと関係を継続する合意をしておけば、社会的な不確実性を減らすことができる。このような、他の相手を排除して特定の相手のみとつきあいや取り引きを互いに継続するという関係

を、山岸は「コミットメント関係」と呼ぶ。この関係は、好意や忠誠心などの感情的な絆による場合も、外敵に対抗するための内部の結束を強要する「ヤクザ型」の場合もあり得る。いずれにせよ、安定したコミットメント関係の中では、相手に関する情報も増え不確実性が減少し「安心」が得られる。そこで、次の命題2が成り立つ。

(命題2) 社会的な不確実性の生み出す問題に対処するために、人々は一般に、コミットメント関係を形成する
(山岸 1998, 76)。

ここまでの山岸説にとくに目新しいところはない。しかし、命題2が正しいとして、社会的な不確実性を避けるには、コミットメント関係の形成がいつも有効だろうか。実は、そうでないところに「信頼」研究のポイントがある。コミットメント関係は、関係内部での不確実性を減らし、「安心」できる状況を生み出すが、その代償として機会コストを要する。つまり、特定の排他的関係に留まることは、別の相手と新しい関係を持つことで得られる（もしかしたら、より大きな）利益を放棄することを意味し、それ相当のコストを支払っていることになる。そこで、次の命題3が出てくる。

(命題3) コミットメント関係は機会コストを生み出す (山岸 1998, 81)。

機会コストと対比されるもう一つのコストは取り引きコストである。これは、取り引きにかかるコストであり、例えば商取引の場合には、営業社員の人件費とか交通費、書類作成の費用や相手の信用調査などにかかる費用である。社会的な不確実性が大きなところで取り引きする場合には、当然この種のコストが大きくなる。したがって、コミットメント関係の形成は、この取り引きコストを節約する手段と見なすこともできる。そこで、コミットメント関係の形成または

維持が他に比べて有利かどうかは、二種のコストのバランスで決まることになる。

(命題4) 機会コストが大きい状況では、コミットメント関係にとどまるよりも、とどまらない方が有利である(山岸1998, 82)。

小論でのこれまでの議論を想起するなら、この命題は、最適化原理を前提すればトートロジーである。二種のコストが計算できない、あるいは計算がむずかしい場合には、サイモン流の満足化の判断に置き換えられるかもしれない。もちろん、最適化原理を前提するにしても、具体的適用例では常に限定がついている。そして、山岸のポイントも、二種のコストが比較できる状況ではこうなる、ということである。

第五命題に入るには、少々言葉の説明が必要となる。信頼をさらに細かく分類していくと、山岸は一般的信頼と情報依存的信頼とを分ける必要があると考える。情報依存的信頼とは、特定の相手について情報が得られている場合、その情報に基づいて相手を信頼することを指す。これに対し、一般的信頼とは、そのような情報が欠ける場合、判断材料がない場合に他人(一般)を信頼することを指す。要するに、後者は、ある人が「人というものは概して信頼できる」と信じている程度を表すもので、その程度が高い人を「高信頼者」、低い人を「低信頼者」と山岸は呼ぶ。常識的には、「高信頼者」は他人を信じやすいお人好し、「低信頼者」は用心深い気むずかし屋だと考えられるかもしれないが、これは必ずしも当たっていないことを、山岸は実証的に示していく。

(命題5) 低信頼者は、高信頼者よりも、社会的不確実性に直面した場合に、特定の相手との間にコミットメント関係を形成し維持しようとする傾向がより強い。(山岸1998, 84)。

もちろん、この命題が成り立つかどうかは実験等により経験的に示すほかはない。それが、山岸の研究の実証的部分(の二つ)である。さて、以上五つの命題から引き出される一つの重要な帰結は、次の命題6である。

(命題6) 社会的不確実性と機会コストの双方が大きい状況では、高信頼者が低信頼者よりも大きな利益を得る可能性が存在する。(三冊 1998, 84)。

山岸の見るところ、この可能性が「他人を信頼する」ことを促す誘因である。人が他人を信頼するためには、もちろん、その他人が「信頼に値する」(「信頼性」)という人格や行動の特性を持っているのが普通である。しかし、この「信頼される」側の特性を指摘する研究は多いが、「信頼する」側の特性と意義に着目する研究はなかった。山岸の言う信頼の「解き放ち」理論は、そこに着目し、「特定の相手とのコミットメント関係からの離脱を促進することで、既存の関係外部に存在している、より有利な機会へのアクセスを可能とすることから得られる利益」(三冊 1998, 86)が信頼を理解するためのカギであると主張する。そして、この主張に対してすぐに申し立てられそうな誤解に対する彼の注意は、倫理学における功利主義批判に対する注意とまったく同様な論点をついており、きわめて興味深い。

その誤解とは、山岸は「人は自己利益を意図的に追求する手段として他人を信頼するようになる」と主張しているのだという誤解である(三冊 1998, 86)。しかし、信頼に値するような特性を身につけることの利益、あるいは他人を信頼することで得られる利益は、普通の人々にわかる形で存在しているわけではない。信頼と利益との関係ははるかに間接的で、自分では自己利益を無視して行動していると考えている人にとっても、他人を信頼することで結果として利益が得られる可能性がある、と山岸は主張しているにすぎない。一般に、限定された合理性しかもたない人間にとって、すべて計算ずくで信頼と利益との関係を見越して他人を信頼するという選択を行うことなどまず不可能である。しかし、

アリストテレスやミルなども繰り返し指摘するように、ある人格的特性や行動特性を身につけることで、結果としてその人の利益となったり幸福を得たりしうる。そのためには、おそらく、一度自分の利益を度外視する習慣をつけるということさえ必要になるかもしれない。そして、人間の社会には、そのような選択肢をとる人々にとつて、自己利益に専心する人々よりも（結果的に）うまくやっつけていける環境が存在している、というのが山岸説の主張にほかならない。

27 信頼と社会的知性

以上が山岸説の核心部分であるが、これでまだすべてが終わったわけではない。命題6で述べられた、信頼が結果的に利益をもたらすようにする可能性、環境とは、実は社会的環境であり、自然的環境のように比較的に安定しているものではない。信頼がもたらす利益は、環境が変われば変わるし、逆に環境そのものが人々のもつ一般的信頼のレベルによつて変化するという相互のフィードバックのあることが、社会的環境の重要な特性である。では、このような社会的環境の中で、高信頼者の特性、つまり相手の情報が乏しい状況でも概してその人を信頼するという一般的信頼の高さはどうのようにして獲得されたのだろうか。この点に答える補助仮説を補わないと、信頼の「解き放ち」理論は不十分である。そこで山岸が示唆するのは、社会の中での人々の意識的な適応行動の副産物としてこの特性が生まれるという仮説である。

山岸は、まず、高信頼者が単なるお人好しではなく、実は他人の信頼性の指標となる情報に敏感で、より正確に推測できると解釈できる実験事実注意到を喚起する。高信頼者が低信頼者よりも騙されやすいわけではないのである。この点を見極めるために山岸らが工夫した一連の実験の結果は、高信頼者の方が低信頼者よりも、他人の信頼性の指標となりうる情報に関してより敏感に反応し、他人の行動の予測に関してより正確であることを示している（三雅1998, 160-172）。これは、高信頼者が社会的知性において優れていると解釈できる。社会的知性とは、いわゆる知能テストで測

定できるような知性ではなく、日常の社会的な関係の中で自分や他人をうまく扱う能力であり、衝動を制御する能力、他者の心や感情を理解する能力なども含みうる。

そこで、山岸が提出する補助仮説は、一般的信頼の獲得を、次のように二段階に分けて説明しようとする。(1)まず、高信頼者の特性である一般的信頼の高さは、社会的知性の副産物として説明される。「一般的信頼」という限定された特性は、より複雑で一般性の大きな能力から説明される。

次に、社会的知性の発達は、少なくともある程度、個人の努力によって可能だと前提し、そういった努力を行なうこと(自分の周りの人々の行動や感情などに注意を払って学習する)を「認知資源の投資行動」と名づけることにする。そうすると、そのような行動をとることが必要とされる環境の特徴がわかれば、社会的知性の発達を説明できることになる。しかし、そのような環境とは、すでにわれわれが確認したような、社会的不確実性と機会コストが大きいという、命題6で述べられた状況と一致する。そこで、(2)そのような社会環境の中では、認知資源の投資行動が有利な方策となり、社会的知性を発達させる人々が生じることが説明される。

以上二段階の説明のポイントは、社会的知性とそれをもつ個人の利益とのつながりがより直接的でわかりやすいので、こちらは社会的適応行動として説明しやすく、これをつなぎとして、一見したところ有利が見えにくい「信頼」の育成メカニズムが説明されうることであろうか。もちろん、疑問が残らないわけではない。例えば、低信頼者の特性も同じ社会環境の中で「絶滅」しているわけではない、また「絶滅」しそうにもないので、高信頼者の特性だけ説明してすべての話が終わるわけではない。しかし、山岸の指摘は、一般的信頼、信頼性(信頼に値すること)、そして社会的知性が一つのセットとなつて、互いの「適応的価値」(社会環境の中での)を高めあっている(Epstein 1998, 1992)ということである。これに対し、排他的なコミットメント関係にとどまり、関係拡張を求めない(傾向の強い)低信頼者は、新しい関係の中で騙されて損をする可能性が低い、社会的知性を発達させる必要が乏しい。したがって、社会

的知性改善のために認知的資源を投資する必要も乏しいわけである。

28 普遍化可能性に対する示唆

さて、以上二節を費やして紹介した山岸の研究は、倫理判断の普遍化可能性という条件に対して何を示唆するであろうか。ここからは、わたし自身の考察である。何度も断ってきたように、普遍化可能性という抽象的な条件と、信頼や信頼性というかなり具体的な特性とは直接同一視することはできない。しかし、信頼の文脈において普遍化がどのような役割を果たしているかは見ることができ、またこういった具体例の中での普遍化の意義がわからなければ、普遍化自体の意義もわからないであろう。わたしが25節で述べたスケッチを想起していただきたい。社会生活での交渉相手の範囲を広げ、それに伴って倫理的な考察の対象とされる人々の範囲を拡張する、つまり相互性あるいは「普遍化」の考察を広げることには、それに見合う利益が（もちろん条件次第で）期待できる、とわたしは見当をつけた。

他方、山岸説によれば、信頼には関係を強化する側面と同時に、関係を拡張する側面がある。しかも、前節で解説したように、関係の拡張を可能にするのは、不確実性と機会コストが大きな社会環境の中で、認知資源の投資行動によって社会的知性を発達させることであった。山岸はこの過程を主として個人の学習とみなしているようだが、この過程は一方ではそれを可能にする生物学的基盤の考察とつながり、他方では社会的知性のいくつかの要素を洗練または抽象して「道徳性」の本質的な構成要素と見なすという倫理学者の考察（および文化）とつながる。したがって、山岸の考察は、道徳の生物学的基盤（第一部11節参照）から始めて規範倫理学まで考察したいというわたしにとっては、始点と終点をつなぐのに欠けていた「ミッシング・リンク」を（少なくとも一つ）提供してくれることになる。それだけでなく、「交渉相手の範囲を広げ、それに伴って倫理的考察の対象とされる人々の範囲を拡張する」ことに見合う利益がどのようにして生じるかも、山岸説は「信頼」という事例に則して示してくれているのである。

加えて、ここで注意しておきたいのは、山岸の研究から得られる示唆は、狭義の普遍化可能性（あるいは、シジウィックの正義の原理）に対してだけでなく、「善の普遍化可能性」あるいは異なる人々の「善の同等な扱い」（18節、19節）に対しても及ぶということである。社会的知性あるいは信頼は、倫理のいわば形式的な側面だけでなく、実質的な側面（他者の感情や内的状態に気を配る）の方にも大きくかかわっているから、これは当然のことであろう。この点は次節で詳しく触れるが、実は普遍化のより重要な側面はこちらの方である。

さて、以上で「普遍化」と「信頼」との基本的なつながりは確認できたので、もう少し細かい論点に立ち入ってみよう。誤解を避けるためにたびたびお断りしておくが、わたしは厳密な普遍化が行なわれない考察をあらかじめ「道徳」から除外しているわけではないことを銘記されたい。高信頼者と低信頼者とがそれぞれの行動パターンをもって共存していることから明らかなように、社会生活の上で一方が他方より絶対的に有利であるわけではない。それどころか、同一人物がある場面では関係拡張に努力し、別の場面ではコミットメント関係の維持に専念することさえ可能である。にもかかわらず、閉じた関係であろうが開いた関係であろうが、その関係に入る人々に対しては、条件が同じであれば同等の扱いが要求される。これが倫理のいわば最低条件であり、相互利他性の核であるといってもよい（コミットメント関係も相互利他的であることに注意）。サイモンという「限定された合理性」しかもたない人間にとつては、規則あるいは原則によって社会的行動を規制するという方策が一般的である。こういった規則あるいは原則は、同じ条件が成り立つところでは同様に適用される。そして、これが普遍化可能性の核にほかならない。

しかし、この規則や原則が個体に言及しないとか、特定の集団内のみは限定されないとかいう保証はない。山岸のいう低信頼者が頼りやすいコミットメント関係に即して言えば、これはある種の排他性によって社会的不確実性を低減させようとする方策であり、しばしば特定の個人やグループに縛られるので厳密な普遍化は妨げるかもしれない。他方、高信頼者は排他的関係にとどまっていたりしては利用できない可能性を活用するので、彼らの行動原則の適用は、自分や相手

が帰属するグループにこだわらない。これが厳密な普遍化への第一歩である。

この主張に対しては、直ちに予想できる反論がある。それは、コミットメント関係のようなある種の排他的な関係で生じる義務も普遍化可能な原則で表現できるので、「閉じた関係、開かれた関係」という切り方は、普遍化可能性とは無関係である、という反論である。例えば、ヤクザの親分と子分の関係は典型的なコミットメント関係であり、義務と似た「忠誠」や「義理」が生まれる。しかし、この「義理」は「親分は子分の面倒を見、子分は親分の命令に従うべし」という、個体には言及しない普遍的な原則で表現できる。したがって、関係の排他性や拡張は、普遍化とは無関係でなかろうか。この反論は、論理的には非の打ち所がないように見えるが、実は普遍化可能性をすでに前提した立場からの反論であるところにひとつ問題がある。

ついでに、この反論と関連が深い、普遍化可能性のもう一つの問題点もここで指摘しておきたい。日常の道徳的思考において厳密な普遍化を行なわない人々が存在することは、ヘアでさえ認める。しかし、彼らの思考や行動のパターンは、厳密に普遍化を行ない、ヘアの意味での普遍化可能な判断によってみずからの行動を律するという条件を課しても、実質的に再現できるのである。例えば、ポイントを際立たせるために、最も極端なエゴイズムの原則を（一種の思考実験として）考えてみよう。いま、わたしが「すべての人は、このわたしの利益を増進するような行為をすべきである」という原則を奉じるとしよう。この「わたし」とは、固有名で指されるわたし自身であり、この原則はその個体に縛られるので普遍化可能ではない。しかし、この現実世界ではわたしだけにしか当てはまらない普遍的性質の組み合わせを選び出し、その記述をD（例えば、わたしのゲノムの記述）とすれば、Dは普遍的な性質を表し、「すべての人はDが成り立つ人の利益を増進するような行為をすべきである」という原則は普遍化可能な（事実普遍的な）原則となる。そうすると、もとの普遍化できない原則が人々に受け入れがたいのであれば、後の普遍化可能な原則も同様に受け入れがたいことになる。

普遍化可能性は、この「わたし」とほかの人々の立場が逆転した「可能性」をも倫理的考慮に入れさせる条件であるが、現実を支配する因果法則まで変える可能性は要求しないはずである。そうすると、これらの因果法則のもとで同一性と一義性を保つ「普遍的記述」によって「わたし」を同定すれば、わたしは、普遍化可能性は満たしつつ、他者にDが成り立って、わたしがDでなくなる可能性を考慮する必要はなくなる。かくして、普遍化可能性は、普遍性の要求としては強すぎる一方で、妥当な（受け入れ可能な）原則を絞り込む条件としてはあまりに弱すぎるのではなからうか。

以上二つの難点に対するわたし自身の反応は次のとおりである。いずれの場合も、普遍化可能性の条件を適用するレベルを区別する必要がある。さし当たって、(1)限定された合理性のもとでの選択と、(2)それらの選択を評価し、改善するという考察のレベルを分けて考えてみよう（わたしは、ヘアの「直観的レベル、批判的レベル」の区別をここまで相対化していることに注意されたい）。山岸が考察の対象とした低信頼者と高信頼者の選択は(1)のレベルにあり、山岸の「解き放ち」理論は(2)のレベルにある。高信頼者は、山岸説を意識し、それをふまえた上で信頼行動を選ぶわけではないことを銘記されたい。彼らは、「一般的信頼」という特性を身につけているのでそのような選択をする傾向があるのである。同様に、われわれが身の回りで時々出会うエゴイスト（わたしもしばしばエゴイストとなる）は、限定された合理性しかもっていないので、「計算すくで」エゴイストになるとしても、普遍化その他の手続きをすべて踏んだ上でそうするのではない。視野が狭い、他人の都合にまで気が回らない、自分の利害に縛られすぎる（これは善の普遍化の問題と関係してくる）などの特性が作用した結果エゴイストとして振る舞うのである。同様に、ヤクザ型の関係に入る人も、普遍化したうえで入るわけではない。これは（仮に選択だとして）(1)のレベルの選択で、普遍化する必要はないし、できる保証もどこにもない。それに対し、わたしが記述して見せた、ホンネとしての極端なエゴイズムを普遍化可能な原則で表現して擁護しようとするエゴイストは、少なくとも一部(2)のレベルに足を踏み入れている。すなわち、「倫理学者の基準」によれば略式裁判で「倫理」から閉め出される原則を、少なくともその関門は通る形にまで「改

善」する努力を行なっており、これは彼が普通ならしない「認知資源の投資行動」を経た、改善された合理性へと足を踏み出している。また、親分・子分の関係を普遍化して考え、それにふさわしい行動をとる人は、「任侠道」をわきまえた、(2)のレベルに足を踏み入れた「模範的な」(それゆえ映画の主人公となつて観客の共感を呼ぶ)ヤクザなのである。

この区別をふまえたうえで、まず「閉じた関係、開かれた関係という切り方は、普遍化可能性とは無関係である」という反論を考えてみよう。わたしは、「普遍化」のいわば発生条件を考察しようとしているのであつて、すでにわかっている「普遍化可能性」の条件を使って「閉じた関係、開かれた関係」の区別を記述しようとしているのではない。すでに見たドウ・ヴァールの考察(11節)を信用するとしても、チンパンジーが「普遍化可能」な判断を行なつていてまで主張する人は誰もいまい。それどころか、普通の人間でさえ、ホンネとしてはしばしば普遍化可能でない判断に縛られていることさえあると予想できる。そこで、この文脈でドウ・ヴァールが、動物あるいは人間の倫理の必要条件としてあげた諸条件を想起してみよう。

共感と関連した特徴

愛着、援助、感情の伝染

学習によつて、能力に障害がある者あるいは傷ついた者に対して順応し特別な取り扱いをすること

心の中で他者と立場を入れ替える能力、認知的感情移入

規範と関連した特徴

指令的な社会的規則

規則の内面化と罰の予見

相互性

与えること、取り引きすること、復讐すること

相互的な規則を破つた者に対する「道徳的」な攻撃

協調

仲直り、および衝突の回避

共同体への配慮、よい関係の維持

利害の対立を交渉によって調整する

これらの条件のうちに、わたしが「普遍化可能性の核」と名づけた条件は（「相互的な規則」という表現で）含まれている。しかし、協調の項目に含められた「共同体への配慮」は、個体が帰属する特定の共同体だけに限定されたものであり、厳密な普遍化を要求しない。この限定を取り除き、別の共同体のなかでの個体と全体の関係、あるいは異なる共同体をまたがる個体間の関係にまで考察が広がるところに、まさにわれわれの知る人間の倫理の一つの特徴がある（つまり、普遍化可能性は一般に認められており、それゆえへアのような主張が出てくる）。そして、この特徴は、山岸の指摘した「信頼」の特徴と重なり、狭い相互性から開かれた相互性を区別する論理的条件が普遍化可能性にほかならない。この条件を満たせばいつも利益が得られるということではない。それは、信頼が常に有利な結果を生むわけではないのと同様である。しかし、厳密な普遍化可能性が認められ、少なくとも一部の個体に定着するには、それが利益を生むという状況の存在が必要条件である。これは、山岸の命題6に対応する主張である。それとともに、そういった状況に適応するための「認知資源の投資行動」も必要となる。その結果、限定された合理性の「限定」がなくなるわけではないが、「改善」された合理性にはなり得る。こういう形で普遍化可能性の発生あるいは定着条件を指摘し、それ

を限定合理性の改善の話と結びつけるというのが、わたしの真意にほかならない。

次に、もう一つの難点、「厳密な普遍化可能性を満たして、なおかつエゴイズムやコミットメント関係のような排他的な戦略が可能である」という指摘を考えてみよう。これに対するわたしの回答の基本路線は、前述のことから、すでにある程度明らかであろう。限定された合理性のあるレベルにおいて選択された排他的な戦略を、合理性の改善されたレベルで（実践的な内容——どう行為すべきかの指令——は変わらないようにして）表現し直すことはいつも可能である。しかし、普遍化可能性を受け入れて、そのような表現のし直しをする（受け入れないのなら表現のし直しをする必要がないし、わたしは言葉のうえだけの「普遍化」を問題にしているのではない）ことは、個体に縛られた原則から原理的には開かれた原則への移行であり、このステップ抜きで排他性を脱することは不可能である。普遍化可能性を受け入れる、ということとは、単に一つの行動原則の表現方法を変えろということではなく、ほかの原則や価値判断の基準についても影響の及ぶ決定である。そこで、「なぜこの条件を受け入れるのか」という疑問に答えなければならぬ。その答えは、「改善された合理性のもとでは、少なくとも一般的に普遍化から得られる利益が存在するから」というのが有力なものとなる。もちろん、「普遍化の利益」とは、山岸の言う「社会環境」のなかで生じる利益であり、他の人々がとる行動戦略や態度に依存する。また、「普遍化を受け入れる」という個人の決定が、その利益を見越した計算ずくのものである必要もない。

もつとも、普遍化可能性はいわば形式的な条件であるから、これだけで選択を大きく左右することはむしろかしい。したがって、普遍化可能性が妥当な（受け入れ可能な）原則を絞り込む条件としてはあまりに弱すぎるという指摘に対しては争う必要はない。ある原則を受け入れるかどうかの選択は、その帰結の望ましさに依存し、「望ましさ」とは「善悪の判断」にほかならないから、問題は「善悪」の普遍化の問題に持ち越される。しかも、すでに小論の17節から19節にかけて指摘しておいたように、問題の核心は「普遍化」というよりはむしろ「善悪の比較」あるいは「善悪の重みづ

け」の方にかかわるのである。そこで、議論をこちらの問題に向ける時がきた。

29 善の普遍化と重みづけ

他者にとつての善悪を見積もる、あるいは適切に倫理的考慮に入れるという形での善の普遍化可能性（むしろ、シジウィックの意味での「博愛」）が、信頼の場合と同様に関係の拡張を要求するというポイントは、一九九五年のスミソニアン原爆展問題を例にして、具体的に示すことができる。マスコミでも報道され、すでにいくつかの研究書や論集まで出ているこの問題は、第二次世界大戦終結五〇年目を記念して、この終結をもたらした（と見なされる）原爆投下の意義を、戦争のエスカレーション、原爆開発、そして戦後の冷戦と核軍備競争という広い文脈の中に位置づけて示そうというスミソニアンの展示計画が、アメリカ世論の袋だたきにあつて中止に追い込まれたという事件である（ニューズウィーク 1997、並岡 1999、および区井 1999b 参照）。この事件で、退役軍人や空軍協会（空軍支援のロビー団体）など、展示反対派とスミソニアン航空宇宙博物館の企画スタッフの間で重要な争点となつたのは、(1)被爆者の写真や遺物などの被爆資料の展示と、(2)原爆を使わずに日本本土侵攻作戦がとられた場合の死傷者の見積もり数とである。いずれも善の普遍化と関係が深いが、とくに(2)は具体的数値がでてくるのでポイントがわかりやすいと思う。

退役軍人たちの神経を逆なでしたのは、この（架空の）死傷者数が、彼らがこれまで無批判に信じ込んできた「百万人神話」と大きく異なつて見積もられたことである（もちろん、ほかにも要因があろうが、ここではこれに論点を絞る）。「百万人神話」とは、「原爆が使われなければ、本土侵攻作戦が行われ、アメリカ側だけでも約百万人の死傷者がでたはずだ」という信念、ないしは「死傷者」が「死者」にすり変わつて）「原爆は百万人の命を救つた」という信念である。これに対し、スミソニアンが最初の展示台本で歴史的資料に基づいて採用していた数字は、「日本の南の島である九州への侵攻作戦が行なわれた場合の推定死傷者数として……一九四五年六月一八日の会議でマーシャル陸軍元帥、

キング海軍元帥、大統領付き幕僚長を務めるレーヒ提督がトルーマン大統領に提示した数字」(ハーヴェット 1997, 270)であり、マーシャルとキングの推定は「作戦開始後の最初の三〇日で三万人から五万人」である。これは、「死者」の数ではなく「死傷者」の数であることに注意されたい。また、レーヒが示した数字は沖縄戦から推定されており、「九州の作戦全体で最高二十六万八千人の死傷者、そのうち五万人が戦死」と解釈される(後の歴史家の分析によれば下方修正され、死傷者六万三千人となるはず。ハーヴェット 1997, 432)ものであった。

もちろん、いずれの数値も推定でしかないが「死傷者百万」と「死傷者五万」とではあまりに大きな違いである。仮にこのような数字で漠然とではあっても善悪の程度が計れるとするなら、明らかに前の方がより大きな悪である。しかし、ここでのポイントは、こういった比較のうちに、「善の普遍化」を促す契機があるということにある。

まず、「死者一人」の善悪の程度を測る単位が必要である。しかし、この単位が(例えば規約によつて)決められたとしても、「Aさんの死」と「Bさんの死」の価値はどう比較されるのだろうか。さらに、二人の死を合わせた全体の価値はどうなるのだろうか。話を簡単にするために、アメリカ将兵の死はいずれも同等の悪で、全体の悪は個々の悪の算術和で測られるとしよう(これも規約であつてもよい)。この仮定は、(規約だとしても)すでにある種の「普遍化」を導入している。すなわち、「Aさんの悪」と「Bさんの悪」を比較の土俵にのせ、「同等」と見なすという決定である。あるいは、もつとむずかしいのは「わたしの死」と「Bさんの死」の比較と計量であるかもしれない。こういった話が、シジウィックの「自愛」および「博愛」の問題と直結することを確認されたい(18節参照)。

さらに、「アメリカ将兵の死」と「日本人の死」とを比較する段になるとどうだろうか。ここには、「普遍化」のまた別の側面が現れる。二発の原爆により命を落とした日本人(だけではなく外国人も含まれるが)は、一九四五年末までで約二十一万人といわれている。アメリカ将兵(例えば)五万人の死と、日本人二十一万人の死は、どうやって善悪を比較したらよいのだろうか。戦争中は、日本人にとってアメリカ将兵の死は「善」、日本兵の死は「悪」だと見なされ

たのではなかったか。逆もまたしかりで、アメリカ人にとって日本兵の死は「善」、アメリカ兵の死は「悪」と見なされたのではなかっただろうか。少なくとも、敵国人の死は自国民の死と同等の重みを持つ「悪」ではなかったはずである。この事態は、戦時中の（それぞれの国民の）排他的態度を抜きにしては考えられない。ところが、戦後長い時間がたち、原子爆弾の歴史的位置づけが考慮されるようになった文脈では、退役軍人（かつての将兵）にとってさえ、敵国人の死はプラスからマイナスに価値の符号を変えたいらしい。彼らは、原爆で死んだ二十一万の日本人の命と差し引きする数字として、百万ではなく五万人のアメリカ将兵の命ではあまりに軽すぎるとおそらく感じたので、侵攻作戦の死傷者数の見積もりに難癖をつけたのである。これは、彼らにとって、「善悪」のある種の「普遍化」が少なくとも建て前としては受け入れられている証拠となる。そして、この「普遍化」を可能にした少なくとも一つのファクターとして、日米両国の排他的関係が現在はおおむね解消され、「開かれた」関係が可能になっていることを挙げなければなるまい。

もっとも、このような簡略な議論に対しては、直ちに「分析不足だ」という非難が挙げられよう。いわく、「手段としての善と目的としての善の区別が無視されている」、「死の善悪ではなく、死に対する選好にまでさかのぼって善悪を論じるべきである」、「異なる状況での態度や選好を同列に比較すべきではない」など。しかし、こういった分析を加えても、大して結論が変わるようには思われない。例えば、「敵国兵Aが自分の死を嫌悪する度合い」という選好の強さがAにとつての悪の程度を決定するとし、同様に「自国兵Bが自分の死を嫌悪する度合い」がBにとつての悪の程度を決定するとしよう（善悪の個人的尺度の導入）。分析をここまでさかのぼらせても、Aの嫌悪（選好）とBの嫌悪を比較の土俵にのせ、度合いを測る尺度を決めなければならないし、二つの合計や差し引きなどによって全体の価値を出さなければならない点（善悪の個人比較と計量）は同じである。また、第三者のCがAとBの悪を見積もって比較するとしても、まったく同じ問題が現れる。わたしは、ここでこの難問を解決しようとしているのではない。ポイントは、戦時中の状況で人々がこのような比較や合計を試みる場合と、平和で両国に交流がある状況でそうする場合との違いに

着目したいということにすぎない。戦時中には、敵国人の選好と自国人の選好とは、まず同じ重みでは扱われない。そして、気をつけなければならぬのは、「日本人にとって自国兵の死が悪であるのと同様、アメリカ人にとっても自国兵の死は悪である」という形の普遍化は、戦時中でも十分に認められうるということである。にもかかわらず、日本人にとって自国兵の死は悪で、アメリカ兵の死は善であるか、はるかに小さな悪であると判断される（われわれは限定された合理性のもとでの判断を扱っている）。したがって、現在の問題点は、「普遍化」とは明らかに独立であるから、同じ言葉で表現するよりも、「善悪の重みづけ」という言葉で表した方が適切である。そして、まさにこれが排他的な関係と、開かれた関係とでラディカルに変化しうる要素である。これは、わたしの「進化倫理学」の見地からは容易に予測できる事実である。

さて、この程度にまで分析を進めておくと、スミソニアン騒動の一つのポイントは、展示スタッフと、退役軍人らの反対派との間で、この善悪の重みづけをめぐる見解が対立したということになる。もちろん、争点は重みづけに影響しうる事実問題と、（ホンネとしての）重みづけそのものとの双方にまたがる。死傷者数の見積もりは事実問題であるが、もしかすればホンネの重みづけをカモフラージュする争点にされたのかもしれない。もし、事実問題について合意が得られ、重みづけにおいて大差がなければ、原爆投下による犠牲と（死者の大半が非戦闘員であったことの是非はいま無視するとして）、架空の本土侵攻作戦の犠牲者とを比較した善悪の判断（これは限定された判断であり、何十年にもわたるその後の帰結まで考慮する必要はないことに注意）は、日米いずれの観点からも同じになるはずである。したがって、事実問題について合意があつてなお判断が食い違ふとすれば、それは重みづけの違いに帰せられる。例えば、「仮に本土侵攻作戦のアメリカ側の死者が三万人程度だったとしても、二十一人の日本人の犠牲でこの三万人を救つたトルーマンの決定は正しかった」という、かなり極端な見解さえあり得るので、これがアメリカ人の命に格段の重みづけを与えていることは明白である。このような態度はアメリカ側偏重の「排他的」な態度に他ならない（このたぐい

の実例については、並脚・丹脚 1999、467-8 の斉藤道雄の記述を参照)。これが山岸のいう「コミットメント関係」に伴う排他性と同質だというのが、わたしの言いたいことである。もちろん、日本の側の排他性についても同じことが言えるが、煩雑になるのでいちいち断らない。

30 善の重みづけと社会的知性

以上で前節(2)の争点と善の「普遍化」あるいは重みづけの関連は確認できたはずである。次に、(1)の被爆資料の展示に関する争点を話を移そう。こちらも善の重みづけと関係がある。退役軍人やとくに空軍協会が強く反対したのは被爆資料の展示である。これらのなかには、被爆者の顔が写った写真が二十一枚あった。スミソニアン騒動を取材した斉藤道雄が書いているとおり「五十年前、惨禍の中にあつたそれぞれの顔を見ていると、時を越え、国境を越えて迫り来る感情というものがある」(並脚・丹脚 1999、430)。空軍協会は、これを来場者の「感情に訴えようとしている」と見なし削除を要求したのである。しかし、感情に訴えるとうなるのだろうか。この反対をした人々は、展示スタッフと同様、そのことの倫理的効果をよく承知していたに違いない。すなわち、被害者への感情移入や共感が働くと、被害者にとつての善悪が見る人のうちである程度「再現」されるのである。アメリカ人の来場者で、日本人の死や苦しみに小さな重みしか与えていなかった人々においても、この「再現」によってその重みが増す(ドウ・ヴァールが霊長類学者の観点から倫理について真つ先に指摘したのは、「共感と関連した特徴」であつたことを想起されたい)。そこで、「排他的」観点や態度のもとで維持されてきた彼らの価値判断も揺らぐ可能性が大きい。さらに、その効果は「排他的」態度そのものを変えることにつながるかもしれない。

かくして、価値判断や行動原則の普遍化、善の普遍化、重みづけと検討してきたわれわれは、価値判断を変え、排他的態度を変えうる一つの重要なファクターにたどり着いたことになる。ただし、「感情に訴える」という表現につられ

て、その中身を誤解してはならない。倫理判断において重要なのは、個々人にとつての善悪と、それらをもつと広い場で突き合わせるときの重みづけである。個々人の善悪を決定するのは、結局は個人の選好である。また、重みとは、先の度合いと言い換えてよいが、これの核心も「何を何より好むか」という選好にほかならない（結局、一次レベルの選好群の間で優先度合いを決める高次の選好）。ただし、善悪の重みづけが問題になるのは、主として個人に縛られないレベルであり、改善された合理性が働く場においてである。したがって、「感情に訴える」ことが倫理的考察に不合理な要素を持ち込むことになると考えるのは、少なくともこの文脈では誤りである。感情に訴えて変わりうる選好が実は重要なファクターなのであって、これは小論で論じてきた合理的選択の不可欠な構成要素にほかならない。しかも、スミソニアンの例で明らかのように、戦時中に一般の人々にはわからなかったような事情や事実を明らかにし、改善された情報のもとで事態を考察しようという改善された合理性のレベルで変わりうる選好が、改善された合理的判断を支える不可欠の要素となる。

このようにして、事実認識において一致し、善悪の重みづけにおいて歩み寄りが為されたなら、改善された合理性のレベルで意見の違いは小さくなる。そして、倫理的問題の解決を図るには、おそらくこのような方法しかあり得ないのである。

そこで、山岸の「信頼」研究に触発されて展開してきた、わたしの「普遍化可能性」と「善の重みづけ」の考察を簡単に振り返り、わたしの考察が何を付け加えたかまとめておきたい。

まず、「排他的な」コミットメント関係であろうが「開かれた」信頼関係であろうが成り立つ「普遍化の核」がある。これは相互的な規則によって互いの行為を規制するという条件である。これは「排他性」や個体に対する言及を排除しない。

しかし、社会的不確実性と機会コストが大きい状況のもとでは、この「排他性」あるいは「狭い相互性」から抜け出

せる条件が存在し、「開かれた相互性」からより大きな利益が得られる可能性がある。この可能性を利用するためには、社会的知性の改善が一つの必要条件であり、厳密な普遍化可能性はこの改善された知性と合理性によって認められる条件である。わたしは、普遍化可能性の発生条件をいうために、関係の拡張から得られる利益を指摘した山岸のシナリオを応用し、限定された合理性の改善という、サイモンからヒントを得た着想と結びつけて論じた。

しかし、単なる普遍化可能性だけでは、人々の選択や価値判断を大きく変えることはむずかしい。排他性を脱するためのこの形式的条件以上に、選択を実質的に左右する「善悪」の普遍化あるいは重みづけの問題が重要である。善悪の個人的尺度を導入し、善悪の判断の普遍化可能性を認めただけでは、問題が解決するわけではない。善悪の個人間比較と重みづけが何らかの形で導入されないと、行為や原則の帰結を見たうえでの価値判断は決まらない。そこで、わたしが指摘したのは、この重みづけが、判断する人の（国とか共同体とかに縛られた）排他的態度の有無によって大きく左右されうるといふ点である。つまり、善悪の重みづけの問題は、普遍化以上に、山岸のシナリオに適合しているのである。もちろん、限定された合理性の改善という文脈にこの問題が適合することも同じである。

以上のように進めてきた考察でまだ不明確な点があるとすれば、それは合理性の改善と社会的知性の改善とのつながりに関する考察であろう。山岸は、「社会的知性」という言葉を導入するに当たって、「自分自身や他人について理解する能力」、「日常の社会的な関係の中で自分自身や他の人々をうまく扱うのに必要なスキルとしての知能」、あるいは「衝動をコントロールする能力、他人の心を理解する能力、人間関係を円滑にする能力」などを挙げ、これらがある程度訓練によって改善できるという仮定をおいて「知能」ではなく「知性」と呼んだ(三冊198:174-5, 179)。他方、わたしは、サイモンの限定された合理性から出発して、情報収集や計算能力の改善、また選択や判断の改善を目指すことは、最適化の主張と矛盾しないどころか、最適化の枠内に収まることを主張し、限定された合理性の改善という視点を取り入れた。これら二つの試みに関連があることはすでに明らかだが、どういう関連であるかをきちんと特定しておく必要

がある。

山岸の社会的知性は、交渉相手の意図や感情をいろいろな手がかりから推測し、相手の行動を予測するという認知的な働きだけでなく、相手の感情や選好を自分の中で「再現」するという、伝統的な哲学で「共感」と呼ばれた能力も含むはずである。人間関係のなかで「他人の心を理解する」という場合、怒っていると悲しんでいることを知るだけではなく、どの程度怒っているか、悲しみがどんなものか、そして相手が何をどの程度望んでいるかということまで含まなければ「理解」とは言いがたい。俗に言う「相手の立場になって考える」とは、客観的な事実を知るだけではなく、相手の内面で起こっていることを自分の内でシミュレートすることも指す。これの下手な人は「鈍感なひと」とか「思いやりの乏しい人」と呼ばれて、人間関係でつまづくことも多い。ただ、注意しなければならぬのは、感情や選好の程度が、客観的な量として確定しているとはかぎらないという点である。ここに、個人間の比較という難しい問題がある。

それでは、こういった能力を「認知資源の投資行動」によって発達させることにはどういう利益があるのか。一言で言えば人間関係を円滑にするということだが、これにはもちろん信用できない人を見分けて、騙されないようにするという利益も含まれる。しかし、山岸が何度も断っているとおり、経済学の言葉を使っているが、経済的な利益や損得だけが考慮されているわけではない。人間関係が広がる、あるいは深まることから得られる精神的満足も利益に含まれるのである。「信頼」という言葉にはそういった含意もあるはずである。

さて、以上のような社会的知性の改善は、わたしの言う「限定された合理性の改善」の一要素として当然必要である。再びスミソニアンの例に戻るなら、被爆資料の展示の問題がこの点にかかわりが深い。被爆体験をすることは不可能であっても、被爆者の遺品や写真を見て彼らの経験がある程度想像することはできる。わたしの個人的な経験から言っても、小学生のときに見た（当時は画家の名前を知らなかったが、丸木位里・丸木俊夫妻の）「原爆の図」の強烈な印象

はいまだに残っている。少なくとも、こういった経験なしで原爆の善悪は語れないはずである。もちろん、退役軍人の経験についてもしかり。この意味で、スミソニアン展示は、アメリカの人々に対して、戦争と原爆について考えるための「社会的知性」の改善の機会を提供できるはずであった。また、「自国の利害しか見えない」という限定を少しでも取り除いて「合理性を改善する」機会も提供できるはずであった。

さて、限定された合理性の改善のためには、社会的知性の改善の場合と同様、しかるべき「認知資源の投資行動」が必要である。平たく言えば、改善のための努力が必要であり、それにはコストがかかる。これに関して、わたしは山岸が触れていない一側面を指摘しておきたい。これもスミソニアン騒動から啓発的なヒントが得られる点なのであるが、ある種の社会的分業によって合理性を改善することも可能だということを認識する必要がある。例えば、原爆問題について考えようとするとき、「ひとりひとりの市民が合理性のこういった改善に努力すべきだ」という「正論」を唱えることはやさしいが、現実性に乏しい。原爆問題について「改善された合理性」の見地から考察するためには、公文書館へ行って解禁文書を調べる、被爆資料を集める、映画やビデオを作る、退役軍人たちにインタビューする等々の膨大な作業が必要である。「ひとりひとりの市民」がこんなことをできるわけがない。しかし、こういった仕事に情熱をもって当たり、ピューリッツァー賞をもらえるような立派な仕事をする作家がいたり、何人かのスタッフで手分けして何年かかけて仕事をできる博物館や学者の共同研究グループがあったりすることは、人間社会の周知の事実である。合理性の改善のための、または限定された合理性の「限定」を部分的にせよ取り除いていくための努力は、こういった社会的な分業によって、社会全体から見れば効率的に行なうことができる。もちろん、その成果は多くの人々がやがて利用できるようになる。これが、「社会的」知性という言葉が持ちうるもう一つの重要な意味であろう。

31 善の比較と重みづけについての規約主義

さて、小論を終わる前に、これまで議論を避け続けてきた一つの難問に触れておかなければならない。それは、異なる人々の間で、彼らの善悪をどのように比較し、どのように重みづけるかという問題である。⁽¹⁾近年では、この問題は「選好の個人間比較」の問題として知られている。個人の善がその人のもつ選好の序列づけに従って（一定の条件を満たしたうえで）決まるという点についてはおおむね研究者の間で合意があり、わたしもそれを踏襲する。この個人の善が量的に計れるかどうかについては異論もあるが、ある種の理想化を加えれば、基数的な（つまり、序列だけでなく量も決まる）効用関数として数学的に表現できることも周知の事実である。ただし、価値の単位量とゼロ点とは任意であるから、異なる人々の効用関数は、それぞれ異なる単位量の中の「転換比」を決める方法が見つかれば、同一スケールで比較可能となる (Harsanyi 1977, 57)。この「転換比」が、われわれが問題にしてきた重みづけに相当する。要するに、個人の効用関数が決まったとしても、それには「規約的」で任意に選べる要素が入っているので、個人間の比較にはまた別の要素が必要なのである。では、この「別の要素」はどのようにして決まるのだろうか。

限定された合理性に基づく選択では、他者の利害を考慮に入れるとしても、この「別の要素」をハルサーニの「転換比」のような形では考慮しない。むしろ、「他者をどう扱うべきか」という一般的な原則に訴えて、間接的に他者の善を扱うにすぎない。しかし、スミソニアン騒動のように、改善された合理性のレベルで「日本人の死」と「アメリカ兵の死」の値打ちを比較せざるを得ない状況に直面すると、この個人間比較の問題が表面に現れてくる。そこで、わたしが提唱したいのは、「重みづけ」ないしは「転換比」は、差し当たって（個人個人が）規約的に決めてよいという規約主義である。ただし、この規約は、改善された合理性に照らして変更可能でなければならぬ。この規約も（新たな情報とそれによって変わった選好に基づく）合理的選択の対象となりうるのである。

わたしのこの提唱は、ハルサーニのような、拡張された共感（他人がある状況に置かれたと想像し、その人の態度や

感受性をもってその事態を考え、この事態に対する選好を形成する) を使った客観主義とは大きく異なる。彼は、個人個人の態度や選好を決める客観的な因果的変数があつて、これらの変数の値がすべて同じであれば、(同じ因果法則に従つて決定される) 態度も選好も一致するはずだから、拡張された共感のもとで形成される拡張された選好は、変数が同じ値であれば誰にとつても一致すると見なす (Harsanyi 1977, 58-9)。しかし、限定された合理性によつても、改善された合理性によつても、そのような因果的変数の値(を仮に認めるとしても) をすべて知ることが不可能であるから、この方策は実際には役に立たない。

わたしの規約主義は、ハルサーニが理論的に解こうとした問題を、局所的にかつ経験的に解ける形にしていこうという提案である。個人個人が当初勝手な規約によつて定めた重みづけでも、いったん決めれば個人間の比較を可能にする。そして、改善された合理性によつて下された判断を互いに突き合わせる段になつて、異なる重みづけを調整することが可能である。例えば、原爆投下の是非についての判断が好例である。個人の選好レベルにまでおりなくとも、日本人一人の死とアメリカ人一人の死を比較するレベルで重みづけ(規約)を導入してもよい。改善された合理性のもとで、この個人的な規約は、いわば公共の場で試される。強硬な退役軍人でさえ、アメリカ側一辺倒の重みづけでは批判にたえられないし、被害者意識一辺倒の日本人の重みづけも、例えば中国人や東南アジアの人々からの批判にたえられない(もつと具体的な話は、峠脚・丹脚 1993 所収の直野章子の文、510, 528-31 を参照)。社会的知性を動員し、他者の経験を想像することで、それぞれの規約は相当程度「合理的」といえる重みづけに歩み寄れるはずである。そして、これはその場限りのことではない。次の「改善」にまで持ち越すことができる。これは、ドゥ・ヴァールが「協調」のところで挙げた「利害の対立を交渉によつて調整する」という条件に相当する。個人間比較の難問は、このような形でしか扱えないというのが、現在のわたしの考えである。

32 進化倫理学と還元主義のプログラム

合理性についても、還元主義の規範倫理学についても、まだ論じるべきことはあまりに多いが、三回に分けて述べてきたわたしの構想は、少なくとも一つの区切りに達したので、この論文はこれで打ち切る。一言で言えば、わたしの規範倫理学の構想は、シジウィックやヘアの功利主義の路線をおおむね継承するが、進化的視点と「限定された合理性」の考えを取り入れて、局所的な最適化と合理性の改善をつないだ「最大化原理」へと功利原理を限定したところにある。こういう形に改訂すれば、規範倫理学における還元主義のプログラムは可能であるというのがわたしの見通しである。

以下で、今回得られた成果をまとめながら、ささやかな展望も付け加えておきたい。

(1) 倫理的用語の意味から実質的な内容をもつ「道徳性」を導こうとするのではなく、そのような内容が進化と社会的・文化的過程を通じてどのように生成しうるかを示す。「道徳性」を天下りの前提せず、いわばダイナミックな過程を通じて「道徳性」をほかの条件に還元することを目指す。これが進化的アプローチにはかならない。

(2) 進化によって人類に備わった資質として、ドウ・ヴァールが示した条件以上のものは前提しない。とくに個人的あるいは集団的選択の「合理性」については「限定された合理性」しか前提しない。これらを基盤にして、現にわれわれが知る「道徳性」の規範的諸条件に至るリンクは、合理的選択と合理性の改善という二つの過程によってつける。最大化原理は、局所的な改善のレベルでしか使われていないので、人間の能力を超えているわけではなく、われわれは還元的基础を増したわけではない。

(3) 小論である程度立ち入って還元の道筋を示したのは、倫理判断の普遍化可能性という条件と、異なる人々の善の重みづけの過程である。もちろん、すべての人々が普遍化可能性を受け入れるとか、すべての人々が「公平」な重みづけを行なうということを示したわけではない（そんなことは不可能であり、また不必要である）。わたしが示したのは、(2)の基盤から始まった過程の結果、合理的選択の結果として普遍化可能性が少なくとも人々の一部に定着しうることを、

また、われわれが知る「道徳性」によつて要求されるような、善の重みづけの変化が同じく合理的な選択の結果としてもたらされうることにすぎない。しかし、それらの結果が人々に対して規範的な拘束をもたらす理由は、還元主義の路線に従つて明らかにしたはずである。合理的選択の基盤となった選択が人々の行為を拘束するのである。したがつて、合理性を追求しようとするれば「為すべきこと」は受け入れられる。こういった規範に従わない人々が存在することは還元の失敗を意味しない。そういった人々がいるからこそ、「べし」が意味を持ち、道徳の存在理由があるにすぎない。

(4)人間の社会の中で、普遍化可能性を認める人々とそうでない人々とは共存しうる。同じく、善の重みづけについても異なるやり方をとる人々が共存しうる。「重みづけのこうあるべき仕方」とは、その時々で改善された合理性に従つて選択される仕方である、というのが現在のわたしの見当であるから、これは変化しうるので、(おそらくシジウィックが博愛の原理で意図したような)一義性をもつ保証はない。改善されていく合理性による局所的な最大化原理はこのような変化の余地を残す。しかし、これは進化的アプローチの避けがたい帰結の一つであり、われわれ人間にはそれ以上のことは望めない。

(5)異なる人々の善悪の比較の問題はあまり掘り下げられなかったが、これは本来重みづけと切り離しては論じられない。選択を考えるにせよ快楽を考えるにせよ、これらの絶対的な尺度があるとする見方は支持できない。しかし、何らかの形で比較ができなければ多くの倫理問題は解決不可能となるので、比較を可能とするための規約を導入しなければならぬというのが、わたしの見解である。ただし、規約の導入は合理的選択の対象となるので、合理性の改善の過程で、より多くの人々に受容可能な方向に変えていくことは可能だと思われる。

注

(一) この問題は奥野1988a (11.2)´ 1988b (12-19)´ 1999 (10.2) である程度論じられているが、わたしがハルサーニの基本文献

を提供してこの問題の所在を彼女に教えたとはいふ経緯がある。事情を知らない読者が奥野の論文を読み、以下の記述を読んでわたしたち「借用」ならし「剽窃」の嫌疑をかけることのないよう、ここでお断りしておく。

文獻

- Baier, K. (1958) *The Moral Point of View*, Cornell University Press, 1958.
- Barkow, J. H., Cosmides, L., and Tooby, J., eds. (1992) *The Adapted Mind*, Oxford University Press, 1992.
- Cosmides, L. and Tooby, J. (1992) "Cognitive Adaptations for Social Exchange", in Barkow et al. (1992).
- Darwin, Charles. (1871) *The Descent of Man*, Murray, 1871, 2nd ed, 1874.
- Dawkins, R. (1989) *The Selfish Gene*, Oxford University Press, 1989.
- Dennett, D. (1991) *Consciousness Explained*, Little, Brown and Co., 1991.
- . (1995) *Darwin's Dangerous Idea*, Simon and Schuster, 1995.
- . (1996) *Kinds of Minds*, Basic Books, 1996.
- De Waal, F. (1996) *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*, Harvard University Press, 1996.
- Frankena, W. K. (1963) *Ethics*, Prentice-Hall, 1963.
- Gibbard, Allan (1990) *Wise Choices, Apt Feelings — A theory of normative judgment*, Harvard University Press, 1990.
- Hamilton, W. D. (1964) "The Genetical Theory of Social Behavior I, II", *Journal of Theoretical Biology* 7, 1964.
- Harsanyi, J. C. (1977) *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press, 1977.
- Holland, John. (1975) *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, 1975.
- Ruse, Michael. (1986) *Taking Darwin Seriously*, Cambridge University Press, 1986.
- . (1993) "The New Evolutionary Ethics", in *Evolutionary Ethics*, ed. by M. H. Nitecki & D. V. Nitecki, State University

- of New York Press, 1993.
- . (1995) "Evolution and Ethics: The Sociobiological Approach", in *Ethical Theory: Classical and Contemporary Readings*, ed. by L. Pojarn, Wadsworth Press, 2nd ed., 1995.
- Schurman, Jacob Gould. (1887) *The Ethical Import of Darwinism*, Charles Scribner's Sons, 1887. 3rd ed, 1903.
- Sidgwick, Henry. (1902) *Lectures on the Ethics of T. H. Green, Mr. Herbert Spencer, and J. Martineau*, Macmillan, 1902 (edited by E. E. Constance Jones).
- Simon, H. A. (1955) "A Behavioral Model of Rational Choice", *Quarterly Journal of Economics* 69, 1955. Reprinted in *Models of Man, social and rational*, John Wiley and Sons, 1957.
- . (1997a) *Administrative Behavior*, 4th ed., Free Press, 1997 (1st ed., 1947).
- . (1997b) *Models of Bounded Rationality, vol. 3. Empirically Grounded Economic Reason*, MIT Press, 1997.
- Spencer, Herbert. (1879) *The Data of Ethics*, Crowell and Co., 1879.
- Trivers, R. (1971) "The Evolution of Reciprocal Altruism", *Quarterly Review of Biology* 4, 1971.
- Uchii, Soshichi (1997a) "Comments on Professor Ruse's View", http://www.bun.kyoto-u.ac.jp/phisci/Newsletters/newslet_19.html, 1997.
- Uchii, Soshichi (1997b) "Darwin on the Evolution of Morality", <http://www.bun.kyoto-u.ac.jp/~suchii/D.omM.html>, 1997; printed version in Uchii (1999).
- Uchii, Soshichi (1998a) "Sidgwick's Three Principles and Hare's Universalizability", <http://www.bun.kyoto-u.ac.jp/~suchii/sidg&hare.html>, 1998; printed version in Uchii (1999).
- Uchii, Soshichi (1998b) "Utility and Preferences", <http://www.bun.kyoto-u.ac.jp/~suchii/util&pref.E.html>, 1998; printed version in Uchii (1999).
- Uchii, Soshichi (1999) "Three Essays on Ethics", *Memoir of the Faculty of Letters* 38, 1999.

Westermarck, Edward. (1906-08) *The Origin and Development of the Moral Ideas*, 2 vols. Macmillan, 1st ed., 1906, 1908, 2nd ed., 1912, 1917.

ボウラー・P・J (一九九二) 『ダーウィン革命の神話』(松永俊男訳)、朝日新聞社、一九九二。

デネット・D (一九九七) 『心はどこにあるのか』(土屋俊訳)、草思社、一九九七。

デネット・D (一九九八) 『解明される意識』(山口泰司訳)、青土社、一九九八。

ドーキンス・R (一九九二) 『利己的な遺伝子』(日高敏隆他訳)、紀伊國屋書店、一九九二。

ドウ・ヴァール・F (一九九八) 『利己的なサル、他人を思いやるサル』(西田利貞・藤井留美訳)、草思社、一九九八。[De Waal 1996の訳。注と文献表その他が省かれて原著のイメージを大いに損なったのが残念。]

トリヴァース・R (一九九二) 『生物の社会進化』(中嶋康裕他訳)、産業図書、一九九二。

ハーウィット・M (一九九七) 『拒絶された原爆展』(山岡清二監訳)、みすず書房、一九九七。

長谷川寿一 (一九九六) 『児童心理学の進歩』一九九六年版、第1章、金子書房。

長谷川寿一 (一九九七) 『心の進化——人間性のダーウィンの理解』『科学』六七一四、一九九七。

岩垂弘・中島竜美編 (一九九九) 『日本原爆論大系第七巻、歴史認識としての原爆』、日本図書センター、一九九九。

奥野満里子 (一九九七) 『パーフイットの功利主義擁護論』、『哲学研究』五六四号、五六五号、一九九七、一九九八。

奥野満里子 (一九九八) 『シジウィックと現代功利主義』、京都大学大学院文学研究科、博士論文、一九九八。

奥野満里子 (一九九八) 『シジウィックと現代功利主義』、『実践哲学研究』二二、一九九八。

奥野満里子 (一九九九) 『シジウィックと現代功利主義』、勁草書房、一九九九。

サイモン・H・A (一九八七) 『意思決定と合理性』(佐々木恒男・吉原正彦訳)、文真堂、一九八七。

サイモン・H・A (一九八九) 『経営行動』第三版(松田武彦他訳)、ダイヤモンド社、一九八九。

塩野谷祐一 (一九八四) 『価値理念の構造』東洋経済新報社、一九八四。

- 立花隆、利根川進（一九九〇）『精神と物質』、文芸春秋社、一九九〇。
- 内井惣七（一九七四）『倫理判断の普遍化可能性について』『人文学報』三八、一九七四。
- 内井惣七（一九八七）『うそとパラドックス』、講談社現代新書、一九八七。
- 内井惣七（一九八八）『自由の法則・利害の論理』、ミネルヴァ書房、一九八八。
- 内井惣七（一九八九）『真理・証明・計算』、ミネルヴァ書房、一九八九。
- 内井惣七（一九九二）『ダーウィンの自然選択説に関する二つの疑惑』、『実践哲学研究』一五、一九九二。
- 内井惣七（一九九三）『形質分岐の原理——ダーウィンとウォレス』、『京都大学文学部研究紀要』三二、一九九三。
- 内井惣七（一九九六）『進化論と倫理』、世界思想社、一九九六。
- 内井惣七（一九九七）『道徳起源論』、『科学』六七―四、一九九七。
- 内井惣七（一九九八 a）『ダーウィニズムと倫理』、『生物科学』五〇―二、一九九八。
- 内井惣七（一九九八 b）『道徳は進化的に安定な戦略か？——大庭健氏の誤読』、『科学哲学』三一―一、一九九八。
- 内井惣七（一九九八 c）『進化と倫理』、『進化経済学とは何か』（進化経済学会編）、有斐閣、一九九八。
- 内井惣七（一九九八 d）『道徳起源論から進化倫理学へ、第一部』、『哲学研究』五六六号、一九九八。
- 内井惣七（一九九九 a）『道徳起源論から進化倫理学へ、第二部』、『哲学研究』五六七号、一九九九。
- 内井惣七（一九九九 b）『進化的に安定な戦略とは』、『科学哲学』三二―一、一九九九。
- 内井惣七（一九九九 c）『ハーウィット（一九九七）の書評』、『科学哲学ニューズレター』二九号、一九九九年一〇月、http://www.dun.kyoto-u.ac.jp/phisci/Newsletters/newslet_29.html
- ワールドロップ・M・M（一九九六）『複雑系』（田中三彦・遠山峻征訳）、新潮社、一九九六。
- 山岸俊男（一九九八）『信頼の構造』、東京大学出版会、一九九八。

* この論文は、学術振興会（未来開拓学術研究推進事業）の委託研究「情報倫理の構築」(プロジェクト・リーダー水谷雅彦) 道徳起源論から進化倫理学へ (三)

の一環として行なわれた研究成果の一部である。

(筆者) うちい・そうしち 京都大学大学院文学研究科〔科学哲学科学史〕教授

what Socrates says about the right way of education in the *Republic* VII, i.e. *περιαγωγή τῆς ψυχῆς* (518D).

On the basis of these interpretations, I have emphasized subsequently in this paper that the allegory of the 'Cave', there stated, is very closely connected with the process of anamnesis in the *Meno*; that, in order to make clear the significance of the critique of writings in the *Phaedrus*, we must interpret it from Plato's point of view on knowledge, which should be anamnesis from within, performed through our own efforts; and lastly that Plato expresses only a very simple fact, when he, in the *Seventh Letter*, compares the instruction of philosophical knowledge to 'a blaze kindled by a leaping spark' (341D), because this situation of knowledge has been, as I have emphasized, already disclosed clearly at the very first step to geometry in the case of *Meno's* slave-boy.

From the Origin of Morality to the Evolutionary Ethics, part II : Reductionism in the Normative Ethics, continued

by Soshichi UCHII

Professor of Philosophy and History of Science

Graduate School of Letters

Kyoto University

This issue will conclude part II and the whole paper.

My reductionistic program in the normative ethics aims at constructing "the morality as it ought to be" on the basis of rational choices. Its basic idea can be stated as follows: The morality as we actually have is not necessarily systematic; it is a bunch of norms, duties, and values which may sometimes collide with each other; and personal preferences are often biased and therefore may lead to differences of moral preferences and moral judgments. However, if we consider on the basis of *rational* preferences, not on actual preferences, such biases and differences may be removed to a considerable degree, and systematization of norms and values may become possible. The morality as it

ought to be is the one which can be justified, in the sense that it can be accepted, on such a rational basis. Thus if we can use the notion of rationality that does not presuppose morality, reductionism can be maintained.

However, rationality has many interpretations, the most serious difference being between (1) rationality with full information and (2) rationality with limited information and capacity. It is well known that Herbert Simon drew our attention to this distinction, and he emphasized the importance of (2), which he named “bounded rationality”, and in which “satisficing” principle replaces “maximizing” principle in (1). Thus the first task for us is to examine this distinction and see whether there is a possibility for conciliating the two. On my diagnosis, although models of bounded rationality are far more successful as a means for *describing* our actual decisions, the significance of the maximization principle reappears if we try to *improve* bounded rationality or its decisions, and to look for a *systematization* of such decisions which may often be incoherent, fragmentary, or *ad hoc*. Simon himself admits that “satisficing” process can be transformed into “maximization” process if we allow extended framework and extra cost for calculation. Thus my proposal is that we allow improvement of rationality, although we always stick to the condition of “boundedness” of our rationality; maximization principle then works at the level of improvement, prescribing to choose a better alternative between any two available at that level. However, such improvement is always *local*, as is the case with evolution by natural selection; but that is all we can do as a finite being.

I will illustrate the significance of my proposal in terms of my criticism of Dennett’s “moral first aid”, which is his version of the defence of bounded rationality in morals. Since we stick to the condition of boundedness, his criticism against high-minded ethics is well taken. But as long as we are committed to *improvement* in moral matters, as Dennett himself seems to be committed to, the significance of maximization is revived. We should aim at maximization via local improvement, although there is no guarantee of realiz-

ing global maximization.

Now, in order to reconstruct normative ethics on this basis of bounded-but-improvable rationality, we turn our attention to “the universalizability of moral judgment” and “the fair treatment of everyone’s good” (these are two of the most crucial conditions for traditional morality). How can our reductionist program derive these normative conditions or their near-equivalents?

As I see it, the universalizability can be accepted as our rationality is improved in social life. Morality is a form of “reciprocal altruism” as discussed in recent evolutionary biology. But this reciprocity does not demand strict universality, only reciprocity within the group to which one belongs. However, this in-group reciprocity has a room for improvement, as long as there may be a possibility for gaining more by extending social intercourse beyond the group: it has its own “opportunity cost”. Thus by extending our social relationships beyond the group, we can obtain a greater benefit; but for that, we have to extend the scope of moral consideration and this elicits “more” universalizability, and this in turn demands an improvement of our rationality (extending our knowledge, consideration, and calculation a bit further).

This speculation can be supported by recent empirical investigations by T. Yamagishi. He focused on the notion of “trust”, as contrasted with “commitment”, and its role in our social life under uncertainty. He argues that while “commitment”-relation is a means for responding to social uncertainty by in-group favoritism, “trust”-relation is another means in terms of extending our relationships to new partners. And he argues that “trust” can be developed where both social uncertainty and opportunity cost are high, because there are possibilities for obtaining a greater benefit by trusting others under such conditions. He also argues that one needs higher “social intelligence” in order to utilize this “trust”-relation. Thus, although his subject is more concrete and specific than our problem of universalizability, we can apply his scenario to our case, and I point out that the development of “social intelligence” is an important factor for improving our bounded rationality.

We can also apply the same idea to “the fair treatment of everyone’s good”,

which is a more important and substantive condition for morality. The point can be illustrated in terms of the incident of “Enola Gay” exhibit at the Smithsonian Air and Space Museum in 1995. In order to give a fair consideration for everyone’s good (that is, “everyone” involved in a given case), we have to give an appropriate weight to everyone’s good, in addition to universalizing our consideration; and how should we do this? In order to assess the significance of the two atomic bombs exploded on Hiroshima and Nagasaki, we have to give an appropriate weight to a Japanese life and an American soldier’s life, among many other factors (such as the escalation of “strategic bombing” or the post-war armament race between US and USSR) involved in the case; but how should we do this? You can stick to your prejudice (e.g. the myth of “a million lives saved by the bombs”), or you may give a very small weight to 210,000 lives lost in Hiroshima and Nagasaki (by the end of 1945). But a far better alternative is to examine the relevant facts involved (as the Smithsonian staff tried to do), to represent the victims’s sufferings as well as American soldiers’s agony (thus appealing to our imagination and sympathy), and so on so forth, and then, in light of all this, to decide your opinion. Such a decision is made in the light of improved rationality. And don’t say it is impossible for bounded rationality; the Smithsonian staff provided many good materials by their long labor, and that’s a way to improve bounded rationality. And we can expect a greater benefit by going out of in-favoritism and exclusive consideration.

Of course there is a hard question of “interpersonal comparison” of good or preference. For this, I propose a conventionalism coupled with revision in light of improved rationality: initial weights given to different people’s good may be conventionally chosen by each individual, but they can be revised in light of improved rationality. The interpersonal comparison can be made possible by a convention, and if it is unacceptable to many people, it can be changed and hopefully improved, by means of mutual criticism and in light of improved rationality; thus my approach is evolutionary to this extent. A summary and a prospect of reductionism conclude the paper.