

「む～(Muu)」の世界： 複数のクリーチャによる多人数会話とその参与のスキル

岡田美智男 坂本彰司 鈴木紀子 伊達正晃
ATR 知能映像通信研究所
〒619-0288 京都府精華町光台 2-2
{okada, shoji, noriko, date}@mic.atr.co.jp

あらまし：ひとりで行動していても常にわれわれは他者の中で自分の振る舞いをデザインしている。電車に乗り込み腰をおろすときも、お互いの視線を気にしながら場所を選んでいる。こうした社会的な相互行為の研究は、人と人工物との関わりにおいても重要な視点を提供するものと考えられる。また、その原型は日常での何気ない関わりや雑談、幼児やあそびの世界に見ることができる。われわれはこの社会的な相互行為の成り立ちに対して構成的にアプローチするために、「む～(Muu)」と呼ぶ仮想的なクリーチャを構築している。本稿では、その仮想的なクリーチャの概要と社会的な相互行為を支える基本的な枠組みについて整理する。また、人と人工物との関わりにおいて重要な視点となる「社会的な存在」とは何かについて、社会的な相互行為における参与のスキルという側面から議論する。

Muu's World: A Model for Multiple Conversation of Artificial Creatures and Its Conversational Skills

Michio Okada Shoji Sakamoto Noriko Suzuki Masaaki Date

ATR Media Integration & Communications Research Laboratories
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288 JAPAN
{okada, shoji, noriko, date}@mic.atr.co.jp

Abstract: Cyberpets and artificial autonomous creatures become popular applications of spoken dialogue systems as well as of artificial life and artificial intelligence research. One of next targets in the fields is to give them social skills so as to be social entities in human-inhabited environments. To behave as social entities, they first have to establish the social bonding with social others, and to coordinate social distance with them. The abilities are also what conventional spoken dialogue systems lack. This paper discusses the social aspects in interacting between human and autonomous creatures, such as dialogue coordination, social bonding and social understanding. One of the points is how to establish and maintain a dynamic, structural coupling with social others. We are focusing the relationship between entrusting behavior and its grounding in our everyday activities. And then, we discuss on a model of the multiple conversation with artificial creatures and its conversational skills that can be emerged from the social interactions with social others.

1. はじめに

街角にある自動販売機から聞こえてくる「ありがとう」の合成音に、われわれは何か彼らからのお礼のような意味を感じることはない。合目的な行動として引き出された発話というより、その背後に何かデザインされた機構を読みとってしまう。Dennett などの議論[1]に従えば、その発話をわれわれは「設計的な構え」で捉えてしまうということなのだろう。その「ありがとう」の声に対して、何かお礼の気持ちを感じるとすれば、自動販売機からの気持ちとしてではなく、それを提供している人、ないしは飲料メーカーからのお礼として感じるものである。

そうした自動販売機に対して、われわれはやはりモノや道具として扱えばよさそうだ。あるコマンドに対して、ちゃんと動作してくれる。われわれがその自動販売機に対して設計的な構えで接するとき、何も問題を生じることはない。ところが「音声言語」などの介在する場面ではすこし状況が異なってくる。自動販売機からの「ありがとう」はリアリティを伴わないものになってしまう。こうしたリアリティの欠如はどこから来るのだろうか。

ハンバーガーショップなどで時々耳にするマニュアル通りでの会話に対しても、ある種のリアリティの欠如を感じることもある。その客観的な意味は確かに伝わってくる。ところが用意された言葉を手順に従って繰り返すだけで、「自分のこととして」話してはいない。その相手の発話を「設計的な構え」で捉えるべきなのか、「志向的な構え」で捉えるべきなのか、その判断に苦慮することになる。こうした状況は自動販売機やマニュアル通りでの会話に限られない。最近の擬人化エージェントなどを備えた対話システムでも同様な違和感を伴うことがある。確かに「ヒトの顔」で応対してくれる。しかしインタラクションそのものは作り込まれたものであり、その相手の反応を「設計的な構え」で捉えてしまう。日常での会話のモードで接していいのか、相手をモノとしてコマンド的に命令調で接するべきなのか、すこし奇妙な状況が生じてしまう。

こうした違和感は、日々の何気ない、われわれの社会的な相互行為を見直すヒントを与えてくれる。と同時に、人工物とわれわれの関わりを議論する機会を与えてくれる。ここでは、社会的な相互行為の対象としての「社会的な存在」とは何か、そして「社会的な存在」としての人工物の可能性について考えてみたい。

2. 社会的な存在とは何だろう

砂場などで蟻の動きなどをみていると、その背後に何らかの意志を感じてしまう。これは蟻に限らず、何か「主体的に行動しているモノ」を見たとき、われわれはその対象を「自発的に、合目的に意志決定をしている」と解釈してしまう傾向がある。こうした帰属傾向を、Dennett は「志向的な構え」と呼んでいる。それは朝の6時に目覚まし時計のベルが鳴ったり、空から雨粒が落ちてくるのを見たときに感じるものとは違う。何かデザインされた動きに対して取るわれわれの「設計的な構え」や、自然現象に基づく動きに対して取る「物理的な構え」からは区別されている。

われわれが他者と対面的に関わる時、この蟻の動きに対して取るわれわれの「志向的な構え」に加えて、また違った「何か」を感じることもある。われわれが電車に何気なく乗り込み、シートに腰を降ろす。お互いの存在や視線を気にしながら、その場所をえらび、衣服の乱れを整える。独りで行動していても、われわれは環境や他者との関わりの中で自分の振る舞いをデザインしている。目の前に座っている人と不用意に視線を合わせないように、われわれは他者の存在を意識しながら、自分の振る舞いを調整しているのだ。こうした何気ない行為までも、他者との相互行為によってその度ごとに組織化されるといえる。

電車の中で友達と何気ない会話をするとき、傍らにいる見知らぬおばちゃんは無視できない。彼女の存在は、われわれの会話の中身におのずと影響を与えてしまう。言葉は交わさないのだけれど、われわれの会話の中に傍観者として確かに参与している。では、自動販売機の傍らでの立ち話をする場合はどうだろうか。われわれは自動販売機やその下を歩く蟻の存在によって、会話の中身が影響を受けることはないだろう。われわれはその自動販売機や蟻などをまだ「社会的な存在 (social entities)」として認知してはいないようである。

このことは、われわれが社会的な相互行為の相手として擬人化エージェントやパーソナルなロボットなどを考えるとき、無視できない要素となる。傍らで聞き耳をたてるおばちゃんこうしたモノや人工物との間の決定的な差異は何かということである。

これまでのロボティクスにおける関心の多くは、機械的 (mechanical) なものから、生物的 (biological) なもの、より認知的 (cognitive) なものに移行しつつある。さらに、社会的 (social) なロボティクスに

飛躍するためには、この「社会的な存在」とは何かという視点は無視できない。パーソナルなロボットやエージェントの社会性についてはこれまでも様々な議論がある。ここでは社会的な相互行為に参与するという側面に着目する。「社会性を備えるとは社会的な相互行為に参与するためのスキルを備えることである」と考えられる。

ひとつのステップとしては、「傍観者として他者の社会的な相互行為に参与できるか否か」という側面である。つまり、その傍らにいて、他者の社会的な相互行為に影響を与えてしまう存在、「社会的な存在」として無視できないものとなり得るかということである。われわれの生活の中にロボットなどが入り込むとき、それらは必ずしも「役にたつもの」である必要はないだろう。何か役にたつものを作ろうとすると、われわれはその対象にあらかじめ「目的」を与えてしまうことが多い。「使う-使われる」という非対称の関係を期待してしまう。「何も役にたたないけれど、いないと何となく寂しい」、そうした関わりからはじめるてもいだろう。われわれは、これを「役たたくロボティクス」と呼んでいる。ここでの中心的な課題は、「社会的な存在」として認知できる人工物を如何にして作り上げるかということである。ここでは「ロボットを創ること」は主眼ではない。むしろ、「その人工物との間で社会的な相互行為をリアルに創りたい」と考えている。

こうした「社会的な存在」としての認知を如何にして得るかという議論を進める上で、乳児や幼児の立場は興味深い。生まれて間もない頃、その乳児の存在がその傍らで会話をしている夫婦に影響を与えることはない。ところがその乳児の成長する過程で次第に無視できない存在になっていく。夫婦の会話に対して傍観者として参与していく。こうした変容は、人工物が「社会的な存在」として認知される過程を議論する上でも興味深い。

3. 社会的な相互行為を支えるもの

「社会的な存在」とは何かを考えていくために、日々の社会的な相互行為の成り立ちについて整理しておく。ここで中心的な役割を果たしているのは「行為の意味の不定性」という視点である[2]。自分の行為の意味は自分が一番よく知っている。自分の行為の意味に対しては最後まで自分で責任を持ちたい。これは自然な考え方であり、われわれの願望でもあるだろう。ところが日々の何気ない行為の成り立ちを考えると、必ずしもこうした見方には当てはまらないものがある。その典型例は「歩行」という行為である。

われわれが散歩をするとき、その「一步一步」の意味を考えることはないだろう。何気なく踏み出してみると、ちゃんと大地が支えてくれている。ヒューマノイドロボットの二足歩行の研究では、静歩行から動歩行へのモードのシフトがひとつのターニングポイントになったといわれている。この動歩行のモードでは、自分の重心の移動を確認しながら、恐る恐る足を前に出しているわけではない。自ら静的なバランスを崩して倒れ込みながら、その大地から受ける抗力を利用して動的なバランスを維持する。ここで足を踏み出す際に、ある種の投機的な行為がなされていることに注意したい。つまり、どうなってしまうかわからないけれど(たぶん、大地がちゃんと支えてくれるだろうという期待のもとで)、投機的に繰り出している。結果として、大地とひとつのシステムを作りながら、歩行という一連の行為を組織化している。われわれは地面の上を歩く、と同時に地面がわれわれを歩かせているのだ。こうしたモードでは、自分の「一歩」の意味を事前に完全に与えることをしない。事前に与えることは不可能であるともいえる。こうした制約を「行為の意味の不定性」と呼ぶ。

われわれの何気ない行為の多くは、こうした意味の不定性を悟りながら、大地などの環境に自分の行為の意味を委ねようとする。この振る舞いをここでは *entrusting behavior* と呼び、その投機的な振る舞いを支え、意味や価値を与える大地の働きを *grounding* と呼んでいる。自発的に繰り出される行為(*spontaneous behavior*)の多くは、この「委ねる」「支える」という振る舞いがびっぴりとくっついているようなのである。

Goffman[3]は、「対面的な相互行為は賭けである」という。自分の行為の意味は自分が一番よく知っている。自分の行為の意味は最後まで自分で責任を持ちたい。こうした思いで歩こうとすると自ずと静歩行になってしまう。同様な思いで、話しようとすると言語を詰めて将棋をしているような熟会的な会話になってしまう。「自分の行為や発話の意味に対して最後まで責任を持つ」という思いがあるとどこかで断ち切る必要がある。むしろ、発話の意味の不定性を悟りながら、「それを一旦は他者に委ね、その意味を支えてもらう」という視点の転換が必要なのである。

こうした状況は社会的な相互行為において様々なところで観察できる。例えば、参与のフレームの成り立ちを考えよう。われわれは会話の中でその役割(*position*)を動的に切り替えていく。ここで重要なこと、それは相互行為的に達成されるという単純な事実である。自分が「話し手」であろう

としても、必ずしも「話し手」になれるとは限らない、自分が「話し手」であるためには、他者は「聞き手」になってくれる必要がある。自分の視点からは、その役割は不定なままであり、常に賭けを伴う。

向こうから近づいてくる知人に向かって、「おはよう」と何気なく挨拶をする。もし相手が何もいわず素通りしてしまったらどうだろう。その「おはよう」という言葉は宙に浮いてしまって、意味をなさない。われわれが発話をするときにも、「どうなってしまうかわからないけれど、相手に委ねてしまおう」という潔さがある。その結果として「何かに支えてもらった、相手に裏切られなかった」という安堵感がある。日常での会話もまた、こうした投機的な行為(entrusting behavior)とそれを支える行為(grounding)によって成り立つ。この相互行為は「当事者にとっての意味」を支えあうものとなる。

コミュニケーションという言葉には「情報を伝える」というニュアンスがいつも付きまとう。ここでは「意味の完結した発話」の伝達を前提とする。情報の受け手との間では上下関係、非対称の関係を作る。一方、日常での何気ない会話では、自分の発話の意味や役割を完結させないまま他者に委ねあう。この「投機的に発話の意味を委ねる、それを支える」という図式は、対話者間で「並ぶ関係」あるいは「対等の関係」を作り出す。「雑談」という事態は、こうした「並ぶ関係」「対等の関係」の中で「当事者にとっての意味を支えあう」ことで特徴付けられる。「情報を伝えあう」ことが主眼なのではなく、むしろ「当事者にとっての意味」を共有しあい、社会的な繋がり(social bonding)を維持することに第一義的な目的がある。

われわれが社会的な相互行為の中で他者に期待することは、こうした「対等の関係」を相互行為的に達成できる能力であろう。そうした相手を「社会的な存在」として認知しているのではないかと考えられる。ここでは「位置取り(positioning)の調整」を相互行為的に達成するスキルを備えたものという視点で捉える。

この位置取りは次の三つのレベルに整理できる。(a)ある空間における身体配置のデザイン:例えば、電車の中でどこに座るか、エレベータの中で他者とどのように向き合うか。ある物理的な空間で場所を占め、位置取るという行為であり、これも相互行為の中で達成される。(b)社会的な相互行為における自分の位置、役割:会話の中でどのパート(position)を受け持つか。聞き手、話し手、傍観者といった参与のフレームもまた相互行為的に達成される。(c)他者に対する戦略、価値観空間におけ

る位置取り:他者の振る舞いによって、自分の行為基準や価値観を調整していく。これは個体に閉じていては最適解を得ることができない。

この調整を基底で支えているものは、先に述べた行為の意味の一部を他者に委ねるといった戦略である。

このような相互行為の中での位置取りとその調整は、他者理解のためのコミュニケーション行動として考えられる。と同時に、他者との関わりの中で自分の立場、価値などを見いだしていく。つまり、社会的なニッチや対人的な自己を獲得するための行為でもある。

4. 「む〜(Muu)」の世界

4.1 基本コンセプト

多人数での社会的な相互行為とその参与のスキルの獲得などを議論するためのキャリングピークルとして、幼児をモチーフとした「む〜(Muu)」と呼ぶ仮想的なクリーチャ(creature)を構築している。「む〜」とは、中国語で「目や目玉」のことを意味する。「幼児」の周辺は、社会的な相互行為やその発達、原初的なコミュニケーションなどを考える上で原型となりうる。ここでは具体的なシステムを構築しながら、社会との関わり、共生的な関係性の構築、愛着行動、社会的な繋がり(social bonding)、原初的なコミュニケーションなどを構成的にアプローチしたいと考えている。

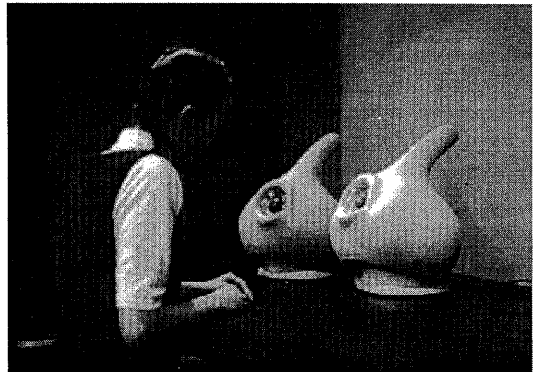


図1. 仮想的なクリーチャ「む〜(Muu)」

図1.に示すように、このクリーチャは抽象的な、不定形でニュートラルな形状をしている。コンセプトの段階から「幼児」をイメージしてきた。その結果として、「ふっくらとした頬、大きな目、ぎこちない動き」など動物行動学のローレンツのいう幼児図式によく当てはまるものとなった。また、発砲ウレタン系の素材で作られており、しっ

とりと柔らかい。身体の動きにあわせてヨタヨタする、ちょっと不気味な生き物(creature)である。

身体全体の上下左右の動きによって視線の動きを表現し、身体配置や姿勢を変えるための移動手段と衝突検出センサーやタッチセンサー、音声認識などを備えている。目線の動きは、相手の顔を追いかけるという視覚的な動きの他に、ある対象に注意を向けている、何かを探しているという他者に対する表示行為としても機能する。これは移動手段でも同様で、何かに近づく離れるという身体配置を調整すると共に、何に姿勢を向けているかといった表示行為となり、社会的な相互行為を組織する上で重要な役割を持つ。

現在は5つのクリーチャが独立に動作する環境が整っている。テーブルの上で、各々のパーソナルスペースを維持しながら、ある社会を構成している。それぞれのクリーチャの行為は他者によって引き出されると捉えるわれわれの枠組みでは、「はじめに社会ありき」という側面は必須のものとなる。

4.2 アーキテクチャ

本稿では「自分の行為の意味を環境や他者に委ねる」という基本的な方略を中心に考えてきた。自分の行為の意味やその戦略の優劣は相手の振る舞いに依存する。これまでの多くの対話研究や言語行為論などで前提としてきた、「前向きな合理性」という戦略なり指導原理が通用しない領域であるといえる。あらかじめ得られたルールを組み込む、作り込むことでは対処できない。むしろ、動いてみなければわからない、あるいは動きながら考えるという戦略である。これは進化ゲーム理論や人工生命などを支える基本的な枠組みであり、「前向きな合理性」に対して、「結果としての選択原理」と呼ばれている。

われわれは先に「目玉ジャクシの原初的なサッカー」という枠組みの中で、社会集団の中で他者との間で価値観を調整しあう、社会的なニッチを獲得するモデルを提案した[5]。このモデルは、社会的な環境に対する「行為-知覚ループ(social referencing)」と他者との関わりにおける自己の行為基準とを調整する「自己言及ループ(self referencing)」によって構成される。これらのループは基本的に分類子システム(Classifier System)によって実装されており、遺伝的アルゴリズムによる試走、評価、選択淘汰に基づいて、多数のSituating Agentの生存競争を経て「戦略」と呼ばれる対象問題に対する解の集合を獲得する。ここでは、このモデルを社会的な相互行為における位置取りの問題に適用する試みを進めている。

自己言及ループでは、社会的な参照ループにおいて「戦略」を獲得するための行為基準を調整する。ここでは、整合希求性と自己充実欲求[6]と呼ぶ二つのパラメータからなる空間を構成する。整合希求性とは他者との関わりを維持するためのドライビング・フォースとして機能する。また、自己充実欲求とは、社会的な相互行為に積極的に参与するためのドライビング・フォースとして機能する。これらの空間での位置取りとは、社会的な繋がりを維持しつつ、自分の自己充実欲求を満足させるための平衡点を他者との関わりの中でオリジナルに見いだすことである。

4.3 多人数での相互行為への展開

ここでは社会的な相互行為への参与のスキルを、この二つの欲求を他者との間で調整しながら、お互いが相互行為にもっとも参与できる状況を作り出すことと捉える。子供たちの会話は重複が多く、会話連鎖を構成しにくいといわれる。独りよがりな行為や場違いな行為は他者を揺り動かすことができない。その行為は無視されたり、会話の場が交錯し、結果として相互行為を組織化することに貢献しない。適切な状況とそのタイミングを「戦略」として獲得する必要がある。これは、どういった状況でのどのような行為が具体的に他者を揺り動かすのか、という言語ゲームの仕方を学ぶことでもある。

ドライビング・フォースの一つである自己充実欲求は、会話への参与を自己中心的、利己的に考えるパラメータであり、この値が小さいとまったく参与しようとしにくいという状況を生む。また、その値が強すぎると独りよがりな行動になって、結果として社会的な相互行為を構成できなくなる。整合希求性に関するパラメータは、利他的な行動に関係しており、相手の発言をグラウンドすることに徹しすぎると、自分の発言の会話への影響力は弱くなってしまふ。またグラウンドする役割を弱くしていくと、相手が離れていってしまう。

この多人数会話に参与するという枠組みは、正統的な周辺参加(LPP)[7]と呼ばれる社会的学習過程とも関連が深い。傍観者から次第にあそびや会話に緩やかに参与しつつ、最終的にはその輪の中心に至る。その参与の過程そのものを社会的な学習と捉える考え方である。子供たちがクリーチャたちの会話の中に参与する過程で学習を実現していく、と同時にクリーチャが子供の会話やあそびの中に参与しながら学習を重ねていく。そうした共進化を含めた形で、学びと学習環境を提供できる可能性がある。これらの具体的な可能性の検討は、今後の課題のひとつである。

4.4 社会的な存在とは(2)

社会的な相互行為に参加するためには、他の参加者から「社会的な存在」として認知されている必要がある。動き回るだけではなく、単に声を出すだけではなく、そうした行為や発話が他者を揺り動かしたり、何らかの影響を与えるだけの、社会的な「力」を備えている必要がある。それは、これまで「受肉(embodiment)する」と表現されてきたものであり、例えば自動販売機からの「ありがとう」の合成音に欠けていたものであろう。電車の中で傍らに座る見知らぬおばちゃんが存在がわれわれの会話に影響を与える。その意味で、その存在そのものは具体的な力を備えていると考えられる。仮想世界や実世界でのクレーチャ[4]の発話がわれわれの行動に対して具体的な影響を与えるということ、このことは自分の行為の意味や役割を獲得する上で一つの前提となる。ひとつの手がかりとして、ここでは社会的な相互行為における「調整(coordination)」ということを考えてきた。「当事者にとっての意味」は、お互いに志向を向けあった状態で、相手との相互調整の過程から立ち現れてくると考えられる。電車の中で偶然に向かい合った人と不用意に視線の一致を避けあう。横断歩道などで信号を待っているときに、見知らぬ人とは不用意に向き合わない。お互いは他人であることを表示しあうために、それぞれの身体配置や姿勢を調整しあう。そうした相手を、われわれは「社会的な存在」として意識しあうようなのだ。砂の上を歩く蟻の進路を塞ぐとその蟻は何を思っただろうか、引き返したり、困惑した動きを見せる。こうした関わりの中で、われわれはその対象の背後に意志や計画の存在を感じ取ってしまう。いわゆる「志向的な構え」で見えてしまう。こうした感覚は関わりの中で立ち現れてきたものである。ある対象を「社会的な存在」として認知するという。これも志向的な構えと同様に、あらかじめ個体の中にある性質が存在しているといったものではなく、むしろ社会的な相互行為における調整の過程から立ち現れてくると考えられる。

5. まとめ

子供と遊んでいるとなぜかほっとする。この感覚は重要なのではないだろうか。「たれたれ」といった言葉に代表されるように、最近では癒し系と呼ばれるグッズや音楽が求められる時代でもある。これまでの「モノを作るという世界」では「何かわれわれにとって役にたつモノを作る」という価値観が先行してきた。こうした中で「幼児」た

ちのたれたれとした、環境に委ねきった姿は興味深い。彼らの振る舞いはどこまでもエコロジカルなのである。彼らの身体は他者に自分の行為の意味を委ねることの重要性をはじめから知っているようなのだ。彼らがヒトとして、つまり周囲から「社会的な存在」として認知されていく過程をつぶさに見ていく必要があるだろう。それは人工物とヒトとの共生的な関係形成を考える上で様々なヒントを与えてくれるものだからである。

「む〜(Muu)」の世界を創る。それは「社会的な存在」としての人工物、「社会的な存在」として認知される人工物を作ろうとする試みである。そのための一つのステップとして、ここでは社会的な相互行為における「位置取り(positioning)」とその調整過程に着目した。他者に対して志向性を備えるとはどういうことか。他者に意識を向けるとはどういうことか。まだ十分な説明や答えを見いだせてはいない。しかし、社会的な相互行為における調整機構は、志向性というものの原初的な側面のひとつであろうと思われる。

参考文献

- [1] D. Dennett: Kinds of minds, Harper Collins Publisher (1996). (土屋俊訳(1997).「心はどこにあるのか」, 草想社).
- [2] 岡田美智男: 何気ない行為を科学する, bit, Vol.30, No.12 (1998).
- [3] E. Goffman, (丸木・本名訳): 「集まりの構造-新しい日常行動論」, 誠信書房(1980).
- [4] 岡田美智男: Talking Eye: 対話する「身体」を創る, システム/情報/制御, Vol.41, No.8 (1997).
- [5] 塩瀬, 岡田, 他: 双参照モデルにおける社会性の創発機構, 認知科学, Vol.6, No.1 (1999).
- [6] 鯨岡峻: 「原初的なコミュニケーションの諸相」, ミネルヴァ書房(1997).
- [7] J.Lave and E.Wenger: Situated Learning: legitimate peripheral participation, Cambridge Univ. Press (1991).
- [8] D. Dennett (若島, 河田 訳): 「志向姿勢の哲学 - 人は人の行動を読めるのか」, 白揚社 (1996).