

Dennett の論考による三つのスタンスの検証

寺田 和憲^{†a)} 岩瀬 寛[†] 伊藤 昭[†]

Verification of Three Stances Discussed by Dennett

Kazunori TERADA^{†a)}, Yutaka IWASE[†], and Akira ITO[†]

あらまし 哲学者 Dennett は人間が他者の振舞いを理解し予測するために意図, 設計, 物理の三つのスタンスを使い分けているとし, 哲学的論考によってその妥当性を示した. しかし, 人間が本当にそのようなスタンスを使い分けているかどうかは定かではなく心理学的研究によってその存在が証明されているわけではない. そこで, 本研究では, スタンスを科学的に定義し, 実際に振舞い理解において用いられているかどうかを検証するために三つの心理実験を行った. まず, Dennett のスタンスをアニメーション化し, アニメーションに対する被験者の印象記述を分析することで, 振舞い理解のための四つの言語的概念カテゴリーを明らかにした. 次に, 60 個の様々な対象の振舞いを参照基準として用いることで, 四つのうち三つの言語的概念カテゴリーが Dennett のスタンスと近いものであることが分かった. しかし, Dennett の主張のような原理帰属による理解がなされているという直接的証拠は得られなかった. 議論から, 実際の振舞い理解においては振舞いの性質が注目され, 意図, 決定, 受動, 複雑の四つの概念的カテゴリー化が行われていると結論づけた.

キーワード Dennett, 意図スタンス, 設計スタンス, 物理スタンス, 心的姿勢

1. ま え が き

哲学者 Dennett は人間が他者の振舞いを理解し予測するために用いる認知バイアスとして意図スタンス (intentional stance), 設計スタンス (design stance), 物理スタンス (physical stance) と呼ばれる三つの心的姿勢の存在を指摘し, 哲学的論考によってその妥当性を示した [1]. 物理スタンスとは主体の物理的組成, 物理的性質, 物理法則に基づいて振舞いを予測する戦略である. 設計スタンスとは物理スタンスで想定される物理的組成などの細部を無視し, 主体が設計意図に基づいて作られていることを前提として, 様々な状況において設計どおりに振る舞うと予測する戦略である. 意図スタンスとは主体の振舞いが意図, 信念, 願望などの心的状態に基づいて合理的に生成されているという前提のもとに, 振舞いの起源となる心的状態を帰属した上で, 振舞いを予測する戦略である.

スタンスの存在は直感的には妥当だと思われるが, 人間が本当にそのようなスタンスを使い分けているか

どうかは定かではなく^(注1), 心理学的研究によってその存在が証明されているわけではない. これは, Dennett の論考によって導かれた概念を科学的手法によって検証可能なものとして定義することが難しいからである. そこで本研究では, スタンスを定義し, 実際に振舞い理解において用いられているかどうかを検証するために三つの実験を行った. なお, 本研究では Dennett の論考を検証するという立場から, 三つのスタンスの存在を前提とした上で話を進める.

2. 実験 1: スタンスを表す概念

Dennett の論考は哲学的考察であり, それぞれのスタンスについて科学的根拠に基づいて客観的に定義されているわけではない. また, その大部分が意図スタンスについての説明に割かれており, 設計スタンスと物理スタンスは意図スタンスを対照化するために用いられているにすぎない. そこで, 振舞いの観察者の視点からスタンスを定義するために, Dennett の論考をできるだけ忠実に再現したアニメーションを作成し, それぞれのアニメーションに対する被験者の印象記述を分析する実験を行った.

[†] 岐阜大学工学部, 岐阜市

Faculty of Engineering, Gifu University, 1-1 Yanagido, Gifu-shi, 501-1193 Japan

a) E-mail: terada@info.gifu-u.ac.jp

(注1): 意図スタンスと同義である心の理論研究は多数行われている.

2.1 アニメーションによるスタンスの表現

Dennett の論考と認知心理学の知見に我々の考察を追加し三つのアニメーションを作成した。アニメーションを作成する際には次の点を考慮した。まず、外観による目的や機能の連想を避けるために、対象の形状を抽象的な形状である円にした。また各アニメーションにおいて同一のサイズの円になるようにし、注目すべき対象であることを示すために円のみを赤くした。更に、視点を横からの視点、すなわち画面の上下が重力の方向になるように統一した。振舞いに関しては、できるだけ具体性のない振舞いでかつ各スタンスをうまく表現するような典型的なものになるようにした。図 1 に、意図 (図 1(a)), 設計 (図 1(b)), 物理 (図 1(c)) のそれぞれのスタンスを採用することを想定して作成したアニメーションの概略を示す。以下では、各スタンスを表現するアニメーションの詳細を示す。

(a) 意図的な振舞いを表すアニメーション

意図スタンスの採用に寄与する性質として、目的となる地点へ向かう目的志向性 [3], 外部からの力によらず動作を行う自己推進性 [4], 重力などの物理現象に逆らった動作を行うエネルギー法則違反 [5], 異なる方法によって単一の目的状態に達する等終極性 [6] などが挙げられている。Gergely ら [2], [7] は Dennett の主張に基づき、合理性の知覚が意図スタンスの採用の基本となっていると考え、目的と状況に応じて効率の良い動作を行う合理的な動作を知覚できるか否かによって乳児が心の理論の萌芽を有しているかどうかを調べた。我々の用いたアニメーションは Gergely らの用

いたアニメーションとほとんど同じである。このアニメーション中の主体の振舞いを理解するためには、円が「反対側へ行きたい」という意図を帰属しなければ合理的な理解はできない。

(b) 設計された主体の振舞いを表すアニメーション

このアニメーションでは主体の振舞いが分別というタスクを遂行するために設計されたものであることを表している。意図スタンスを示すアニメーションにおいては、円が試行錯誤を経て目的に到達する様子を表現することで被験者を円の内部状態 (意図) に注目させることを目標とした。一方、設計スタンスでは対象の振舞いを設計者が規定した機能に基づいて解釈する [1], [8] ため、このアニメーションでは対象そのものが意図をもつのではなく、特定の入力に対して特定の出力を行うという「設計原理」によって振舞いが規定されていることを表している。

この振舞いの特徴は等速での移動、対象物の形状に応じた移動であるため、周期的、規則的、アルゴリズム的である。また、より高い抽象レベルで見ると、分別という機能が特徴となる。同じ振舞いを繰り返すという意味では固執であるが、数回の失敗の後に成功に至るといような目的志向的な固執ではない。

(c) 物理的振舞いを表す主体のアニメーション

物理スタンスを表現するアニメーションは、円が左上から出現し、坂を弾みながら落ちてくる。このアニメーションでは物理的組成までを知ることはできないが、重力という物理法則や壁と円の相対的な弾性係数の違いという物理的性質を知ることができる。

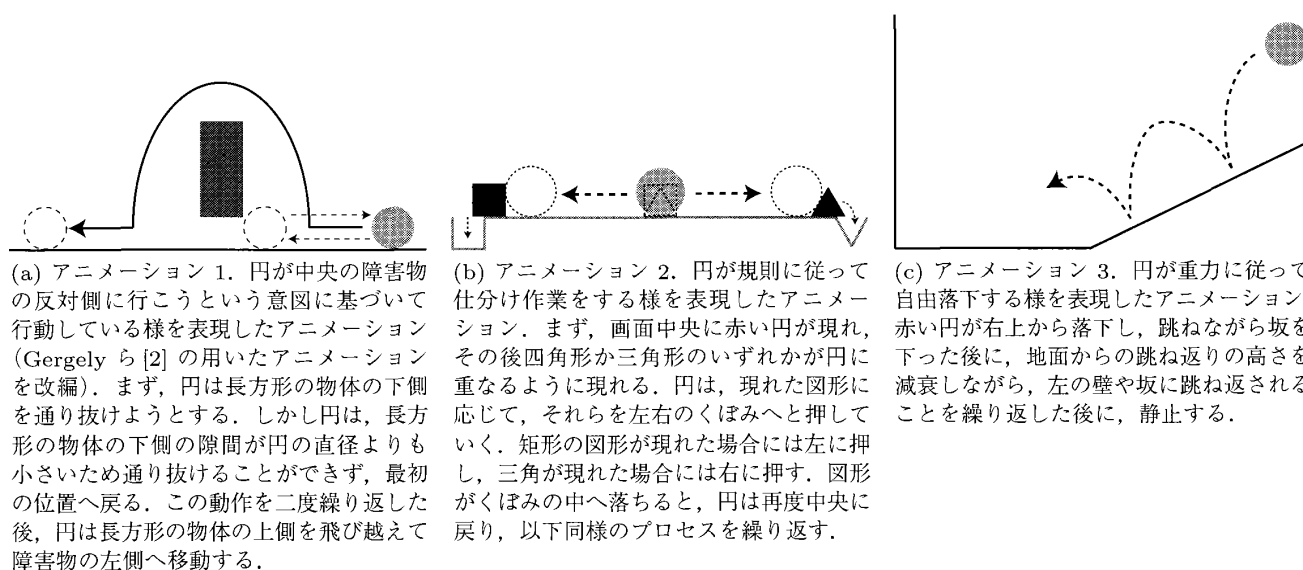


図 1 実験で用いた三つのアニメーション

Fig. 1 Three animations used in our experiment.

2.2 被 験 者

工学系の学部生と大学院生 32 名（男性 23 名，女性 9 名：平均年齢 22.3 歳）が実験に参加した。

2.3 実験方法

被験者は Web ブラウザ上で三つのアニメーションを順に見せられた。アニメーションは繰り返し再生され，それぞれのアニメーションを見る時間は十分与えられた。三つのアニメーションを見終わった後に，アニメーション一つに対して文字入力フォーム一つがあるページを示し，アニメーションに対する印象を入力することを求めた。このときのアニメーションの提示順は被験者ごとにランダムであった。

2.4 実験結果

ほとんどの被験者が印象を一文で記述していた。各アニメーションに対する被験者の印象を定量化するために以下の処理を行った。まず，形態素解析システム Sen を用いて形態素に分割し，名詞，動詞，副詞，連体詞のみを取り出した。次にそれらの語から，「感じ」や「思う」など被験者自信が主語になっている単語，「表現」や「様子」など，対象や状況を記述するための単語，「動き」や「動作」など対象の動きを説明するための単語を削除した。この操作によってアニメーション 1，2，3 それぞれに対して 107，86，99 語が抽出された。抽出された全 292 語のうち振舞いの性質を形容する単語のみを取り出した。この操作によって約半数の 161 語が抽出（アニメーション 1，2，3 それぞれに対して 51，57，53）された。これは重複を許した数なので重複する語を除くと 79 語が残った。

結果を分かりやすく示すために 79 個の単語を意味の類似性に基づいて 16 のカテゴリに分類した。カテゴリ作成とカテゴリ名の決定は我々が行った。カテゴリを表す単語とカテゴリを構成する単語の関係を表 1 に示す。また，それぞれのカテゴリの単語が三つのアニメーションそれぞれに対して使用された頻度を図 2 に示す。

2.5 考 察

分析によって得られた単語と Dennett の論考におけるスタンス，心理学的知見を対比する。

2.5.1 アニメーション 1 に対する印象語

Dennett によると，意図スタンスでは心的状態（意図）を帰属した上で振舞いを理解する。アニメーション 1 の作成時に我々が想定した心的状態は「障害物の反対側へ行きたい」という意図である。被験者がこの意図を想定していたと考えられるかについて検討する。

表 1 アニメーションの印象記述に用いられた単語
Table 1 Words for representing impressions of three animations.

カテゴリ	単語
感情	憤る，いらいら，沈静，いらだつ
知性	思考，意識，意思，見出す，学習，知能，頭脳，かしこい
判断	判断，問題，解決，考える，決める
努力	試行錯誤，克服，乗り越える，挑戦，チャレンジ，困難
達成	成功，失敗，達成，クリア
単調	単調，機械，淡々と，単純，リズム
仕事	仕事，作業，こなす，はたらく
法則	法則，規格，則る，摂理
正確	正確，きちんと，きっちり
規則	繰り返し，繰り返す，反復，規則正しい，一定
予定	予め，決める，あらかじめ，決まる，準備，プログラム
識別	識別，選択，整理，分ける
複雑	複雑，意外，不規則
自然	自然，ありのまま
受動	受動，流れ，任せる，任す，逆らう，逆らえる，まま，なす，無抵抗
物理	物理，収束，落下，落ちる，重力，ぶつかる，飛ぶ，飛び越える

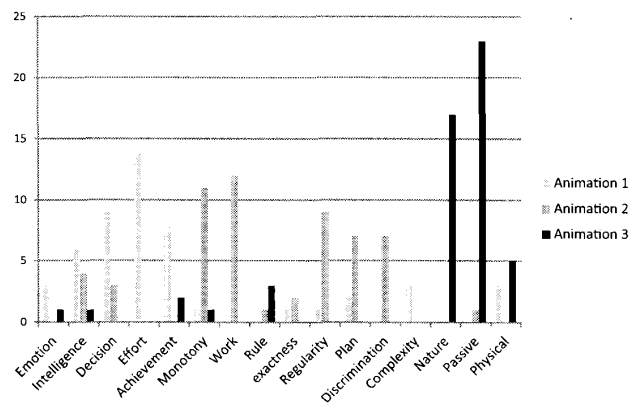


図 2 アニメーションの印象記述に用いられた単語の分布
Fig. 2 Distribution of words for representing impressions of three animations.

アニメーション 1 の印象記述に用いられたのは 16 カテゴリのうち 11 カテゴリの単語である。その中で，アニメーション 1 に対しての使用頻度が他の二つのアニメーションよりも高いものは感情，知性，判断，努力，達成，複雑の 6 カテゴリの単語である。これら 6 カテゴリに属する単語の総利用数のうち，80%がアニメーション 1 に対して用いられていた。これらの中で意図を想定していたことを示唆する単語は努力カテゴリに属するものである。努力カテゴリの語句はアニメーション 1 に対してのみ用いられ，アニメーション 1 に対して用いられたものの中で最も頻度が高い。このカテゴリに属する単語は，試行錯誤，

克服, 乗り越える, 挑戦, チャレンジ, 困難である. これらのうち, 困難以外の単語は目標状態を想定した上で, 目標状態に対する近接, 到達の様子を説明するものである. したがって, 「障害物の反対側へ行く」という意図が想定されていたものと考えられる. また, 達成カテゴリーの成功, 失敗, 達成, クリアなどの単語は目的の達成を評価するための単語であるので, 何らかの意図を想定したと考えることができる. ただし, これらは振舞いの結果を評価するためのものであり, 振舞いの起源である意図に言及するものではない. 知性, 判断カテゴリーの知性や判断力は目的志向的な主体が合理的な行為を生成するために必要な能力であるが, 直接目的を想定するものではない. 複雑カテゴリーに属する単語は意図的な主体が生成する逸脱的行動を説明する単語であるが, 目的を直接想定している根拠とはならない.

以上のことから, アニメーション 1 の理解のために, 意図の想定を前提とする単語や目的志向性から派生する単語が用いられたことから, Dennett の考える意図スタンスと類似した概念がアニメーション 1 から知覚されたといえる.

2.5.2 アニメーション 2 に対する印象語

設計スタンスとは設計原理を帰属して振舞いを理解し予測する戦略である. 設計原理とは, 設計者の実装した目的とそれを実現するためのアルゴリズムやメカニズムなどの法則である. アニメーション 2 に対して目的や法則に関する概念に相当する単語が用いられたかどうかについて考察する. アニメーション 2 の印象記述に用いられたのは 16 カテゴリーのうち 10 カテゴリーの単語である. その中で, アニメーション 2 に対しての使用頻度が他の二つのアニメーションよりも高いものは単調, 仕事, 正確, 規則, 予定, 識別の 6 カテゴリーの単語である. これら 6 カテゴリーに属する単語の総利用数のうち, 89%がアニメーション 2 に対して用いられていた.

目的に関して抽象的に言及する単語は努力と達成カテゴリーの語句であるが, アニメーション 2 に対して用いられることはなかった. しかし, 具体的な目的である識別に関する単語は用いられた. また, アルゴリズムやメカニズムに関する単語として, 法則カテゴリーの単語があるが, アニメーション 2 に対しては 1 回用いられただけであった. しかし, アルゴリズムやメカニズムに従った振舞いの性質である「あらかじめ決まっている」ことに関する単語は使用された (予定

カテゴリー). また, 単調や規則, 正確の法則に従う振舞いの表面的属性に関する単語が用いられた. 仕事カテゴリーに属する単語は, その振舞いを行うこと自体が目的ではなく, 他者から強いられ, 金銭等を得るといった別の目的の手段になっていることを意味するものである. 多くの設計された物が使用者の意図を遂行するための手段として用いられることを考えると, 仕事カテゴリーの単語が設計された振舞いの記述に用いられたことは妥当である.

以上のことから, アニメーション 2 の理解のために, 設計原理を帰属していたという直接的な証拠は得られなかったが, 設計原理を帰属することによって派生的に得られる印象を感じていたという間接的な証拠が得られたために, Dennett が考える設計スタンスと類似した概念がアニメーション 2 から知覚されたといえる.

2.5.3 アニメーション 3 に対する印象語

物理スタンスとは物理法則を帰属して振舞いを理解し予測する戦略である. 物理法則を帰属して振舞いを理解していたかどうかについて検討する. アニメーション 3 の印象記述に用いられたのは 16 カテゴリーのうち 8 カテゴリーの単語である. その中で, アニメーション 3 に対しての使用頻度が他の二つのアニメーションよりも高いものは法則, 自然, 受動, 物理の 4 カテゴリーの単語である. これら 4 カテゴリーに属する単語の総利用数のうち, 91%がアニメーション 3 に対して用いられていた. これらの中で, 法則と物理カテゴリーの単語は物理法則や物理的性質に言及したものであり, Dennett の物理スタンスの考え方と一致する. しかし, それらの割合は少なく, 代わって自然や受動カテゴリーの単語が多かった. これらの概念による振舞い理解は Dennett の考えには入っていない. 自然や受動カテゴリーの単語は, アニメーション 1, 2 の場合と異なり, 間接的に振舞いの原理を想定していたことを示すものではない. このことは, 物理的な振舞いをする主体の理解に, 必ずしも振舞いの原理である物理法則を帰属する必要がなく, 自然や受動という理解をする方が有効であることを示唆する.

なされるがままやありのままという概念は意図的主体の振舞いと対照的なものである. 意図的主体の振舞いの特徴として重力などの物理法則違反が提案されている [5]. 言い換えると, 非意図的な主体は物理法則に逆らわないような振舞いをすると考えてもよい.

2.5.4 三つの概念的カテゴリーをなしているか

図 2 を見ると, 三つのアニメーションそれぞれに対

して用いられた単語はある程度の分布の偏りがあり、三つの山を形成していることが分かる。すなわち、アニメーション 1 では努力、判断、達成、知性に関連する単語の頻度が高く、アニメーション 2 では仕事、単調、規則に関連する単語の頻度が高く、アニメーション 3 では受動、自然に関連する単語の頻度が高い。しかし、それぞれのカテゴリーに属する単語は必ずしも三つのアニメーションに対して排他的に用いられているわけではない。これは、三つのアニメーションに対してそれぞれ固有の概念的理解がなされたことを示唆するが、三つの山を形成している単語が実際に三つのスタンスに相当する概念の形成に寄与していることを直接説明するものではない。例えば、アニメーション 1 では努力、判断、達成、知性に関連する単語の頻度が高かったが、それらの単語が意図スタンスという単一の概念を形成する特徴となっているかどうかは明らかではない。そこで、得られた単語が三つの概念的カテゴリーを形成しているかどうかを確認するための実験を行った。

3. 実験 2：60 個の動画を参照基準とした概念カテゴリーの抽出

実験 1 では三つのスタンスを表すアニメーションそれぞれに対して特徴的に用いられる単語群が存在することが示唆された。本実験では、それらの単語群が三つのスタンスに相当する概念を表象するための特徴となっているかどうかを調べた。調査方法には因子分析を採用した。79 個の特徴を表現するための潜在的な因子が三つであり、なおかつ、因子の内容に妥当性があれば、79 語の背後に三つのスタンスに相当する概念が存在していると考えられる。実際には、79 語から意味の類似した単語を除いて得られた 26 個の語句を特徴として用いた。26 個の語句を評定するための参照基準として、様々な主体の振舞いが記録された 60 個の動画を用いた。動画を見せるという実験設定を用いたのは、三つのスタンスが振舞い理解のために用いられる心的姿勢であり、振舞い理解時における判断を求めるためである。

3.1 被験者

工学系の学部生と大学院生 20 名（男性 14 名、女性 6 名：平均年齢 20.1 歳）が実験に参加した。いずれの被験者も実験 1 に参加していない。

3.2 実験方法

まず、実験 1 で得られた表 1 の単語を網羅するよう

表 2 26 個の語句

Table 2 26 words for impression assessment.

カテゴリー	語句
感情	感情的
知性	意識的、考えている
判断	判断している、臨機応変
努力	努力している
達成	成功、達成（しようと）している
単調	単調、単純
仕事	義務的
法則	法則に従っている、アルゴリズムに従っている
正確	正確
規則	規則的
予定	予め決められている
識別	—
複雑	予測できない、変化がある、複雑
自然	自然な、ありのまま
受動	なされるがまま、身を任せている、受動的、逆らっている、能動的
物理	—

な 26 個の語句を決定した（表 2 参照）。決定には次の規則を用いた。

- 各単語が抽象度の低い特定の現象を表現する単語で、かつ、共通する上位の概念が存在する場合はそれに置き換える。
- 類似した単語はそれらの中から代表する語句を選ぶ、若しくは他の適切な語句に置き換える。
- 単語に共通する特徴を表す単語に置き換える。

識別と物理のカテゴリーに属する単語は特定の動作や物理現象に関する単語であったため特徴として採用しなかった。

次に、26 の語句を用いて 60 個の動画を評定する被験者実験を行った。アンケートは html で作成し、実験は全て Web ブラウザ上で行った。被験者には動画一つと印象評定のための 26 の語句が表示されているページが順に示された。アンケートは全 60 ページである。動画の提示順は被験者ごとにランダムにした。回答に制限時間はなく、被験者が入力を行った後に自動的に次のページへのリンクをクリックすることで次のアンケート画面が提示されるようにした。被験者には各動画に対して 26 個の語句がそれぞれの動画を説明するものとしてどの程度適切であるかを 7 段階のリッカートスケール（1. 全く適切でない-7. とても適切である）によって評定することを求めた。

表 3 の第 1 列に 60 個の動画で映された主体と振舞いを示す。60 の動画の選定基準は次のとおりである。

- 刺激動画の振舞いの主体が明らかに特定できる。

表 3 実験に用いた動画と因子得点, アニメーションによる評定値の平均
Table 3 Factor scores and mean of animation similarity ratings for sixty video stimuli.

振舞い	因子得点				評定値		
	因子 1	因子 2	因子 3	因子 4	アニメ 1	アニメ 2	アニメ 3
1. 男性が固く閉じた缶を開ける	2.00	-0.61	-0.72	-1.11	4.70	1.90	2.00
2. ロボットアームが溶接作業を行う	-0.47	1.80	-1.11	-0.18	2.60	4.90	1.70
3. 火が紙を燃やして消える	-1.47	-0.49	0.72	0.19	2.00	1.80	3.75
4. 和菓子職人が精密な動作で繰り返し生地を伸ばす	1.26	1.17	-0.87	-1.40	2.35	5.30	2.60
5. ダンサーがロボットダンスを行う	0.62	0.31	-1.31	-0.38	2.95	3.70	1.75
6. 軍隊が行進する	0.31	1.89	-0.51	-1.23	2.20	3.95	2.50
7. 人型ロボットがルービックキューブを解く	1.22	1.66	-0.89	1.47	3.30	5.60	2.15
8. 歩行中の BigDog がバランスを崩すも立て直す	0.30	-1.08	-0.30	1.57	3.70	2.95	3.10
9. 6 足ロボットがぬいぐるみを掴み体に載せる	0.54	1.30	-1.30	0.89	3.35	5.10	1.95
10. 少年が一回転しながら川に飛び込む	1.38	-0.70	0.01	0.52	4.15	1.90	4.60
11. スキーヤーが斜面を滑り降りてくる	1.39	0.48	0.19	0.25	3.35	4.05	4.10
12. 人が丘の上から滑って転がってゆく	1.39	0.48	0.19	0.25	3.40	2.00	5.70
13. 自転車に乗った人がジャンプに失敗し転倒する	0.21	-1.62	0.24	-0.32	3.95	2.15	5.15
14. ラフティング中のボートが転覆し流される	0.88	-0.99	0.92	1.98	4.20	2.50	4.50
15. 寝そべっていた猫が台の上から落下する	-0.52	-1.34	0.38	-0.01	4.10	1.80	4.25
16. 犬が眠気に耐えきれず転倒する	-0.78	-1.66	0.46	-0.34	3.35	1.80	3.65
17. ししおどしの筒が上下に動く	-1.33	0.99	1.06	-1.43	3.20	3.55	4.55
18. 水車が回転している	-1.23	1.53	1.40	-1.39	2.05	3.65	4.35
19. 複数の鳥型のカイトが風で動く	-1.32	-0.31	0.89	1.02	2.50	2.50	3.20
20. 人型やじろべえが左右に揺れる	-1.37	1.21	0.65	-1.68	2.55	4.35	3.25
21. ASIMO が階段を登る途中で転倒する	-0.81	0.01	-0.13	1.41	4.00	3.60	3.70
22. 進行中の自動車が溝に落下する	-0.92	-1.45	-0.08	0.35	2.80	2.45	3.90
23. 男性がウィングスーツで滑空する	0.34	-0.43	1.14	0.73	2.60	2.35	3.80
24. 女性が迎いを見回している	0.02	-1.78	-0.46	-0.61	3.15	2.45	2.65
25. 幼稚園児たちがお遊戯をしている	1.00	-0.47	0.48	0.76	3.25	3.25	2.35
26. 男性が工場で仕分け作業を行う	1.05	1.01	-0.49	-0.52	2.95	5.75	2.50
27. カラスが針金を使って筒から餌を取り出す	2.04	-0.22	-1.11	1.23	4.30	3.40	1.95
28. 熊が木の棒を振り回し遊んでいる	1.09	-1.72	-0.14	0.08	3.70	3.25	2.45
29. オウムが棒を使って体を掻いている	1.90	-0.66	-0.62	0.15	3.35	3.55	1.95
30. ササゴイが水面に葉を落としおびき寄せた魚を捕らえる	1.99	-0.37	-0.17	0.66	4.00	3.95	2.20
31. 猫がゴム製の輪で遊んでいる	1.28	-1.49	0.28	0.48	4.80	2.70	3.65
32. 犬が風船で遊んでいる	0.89	-1.92	-0.18	-1.04	3.65	2.75	2.10
33. イルカが気泡で作った輪を追いかける	1.39	-0.74	-0.22	0.21	3.40	2.75	2.70
34. 猫が餌を食べている	0.96	-0.92	0.29	-1.88	2.50	2.75	2.00
35. 犬が眠気をこらえている	-0.09	-1.37	0.41	-0.37	3.15	2.30	2.45
36. 朝顔の蔓が棒に巻きつく (早送り映像)	0.51	0.04	0.73	0.35	3.00	2.55	2.75
37. 花が風に揺れている	-1.65	-1.11	1.52	-0.31	2.05	1.95	3.90
38. 花のつぼみが開く (早送り映像)	0.23	0.28	0.42	0.62	3.05	2.35	3.30
39. ススキが風に揺れている	-1.90	-1.13	1.68	-0.35	2.30	2.15	3.70
40. 人型ロボットが横たわった姿勢から反動をつけて起き上がる	-0.05	1.08	-0.79	0.75	4.30	2.75	2.55
41. 人型ロボットがバイオリンを弾く	-0.33	1.42	-0.76	0.56	3.25	3.95	2.10
42. 二足ロボットが足踏みし方向転換する	-0.76	0.45	-1.08	0.74	2.95	3.25	1.95
43. フェービーが表情を動かし左右に揺れる	-0.53	0.39	-0.98	0.84	2.95	2.80	2.05
44. ライトレースロボットが線上を進む	-0.29	1.82	-0.58	-0.37	3.35	4.60	2.70
45. 倒立振り子ロボットが倒れた状態から起き上がり, バランスを取りながら移動する	-0.65	-0.95	-0.75	1.54	4.20	3.55	2.70
46. ベルトコンベアがテニスボールの仕分けを行う	0.34	1.64	0.10	0.55	2.95	5.45	2.55
47. 複雑に組まれた木製の棒上でビー玉が次々と転がってゆく	-1.36	1.31	0.62	0.23	2.95	4.15	4.85
48. 踏切の遮断機がバーを降ろす	-0.90	1.53	-0.64	-1.10	2.00	3.55	2.95
49. 自動車がドリフトする	1.16	-0.09	-0.81	1.44	3.20	3.50	2.70
50. 四輪駆動車が坂を登ろうとするがスリップしなかなか登れずにいる	0.87	-1.22	-0.76	0.29	4.85	2.30	3.35
51. バスが停留所に到着する	0.63	0.67	-0.93	-0.70	2.90	3.45	2.85
52. 電車が走ってくる	-0.72	1.27	-0.43	-0.73	2.30	2.55	3.65
53. ロープウェイが乗り場に到着する	-0.90	1.53	-0.64	-1.10	2.40	2.95	3.95
54. 旅客機が飛び立つ	-0.14	0.91	-1.07	-0.13	2.50	2.40	2.70
55. 船が海上を進む	0.03	0.42	-1.05	-0.50	2.55	3.10	2.55
56. 逆立ちゴマが回転しながらひっくり返る	-0.85	0.16	0.87	0.29	2.75	2.15	4.10
57. 風鈴が風に揺れている	-1.73	-0.76	1.86	0.36	2.40	2.35	4.05
58. 竹トンボが回転している (スロー映像)	-1.11	0.66	0.46	-1.33	1.95	2.25	3.30
59. 雲が空を流れてゆく	-1.69	-1.09	1.25	0.08	1.80	1.70	3.60
60. 滝で水が流れている	-1.41	0.59	1.22	-1.58	1.95	1.55	4.80

表 4 因子負荷量
Table 4 Factor loading.

	因子 1	因子 2	因子 3	因子 4	共通性
意識的	.826	-.100	-.129	.030	.709
考えている	.799	-.066	-.111	.184	.689
判断している	.736	.099	-.148	.138	.592
努力している	.711	-.009	-.114	.265	.588
能動的	.679	-.139	-.294	-.088	.575
臨機応変	.664	.049	.026	.359	.572
達成（しようと） している	.640	.240	-.186	.212	.546
受動的	-.566	.201	.406	.119	.539
感情的	.530	-.414	.073	.068	.539
逆らっている	.400	-.242	-.080	.165	.252
予め決められてい る	-.184	.776	-.097	-.140	.664
正確	.087	.771	-.145	-.155	.647
規則的	-.117	.720	.052	-.356	.661
法則に従っている	-.242	.717	.103	-.128	.599
アルゴリズムに従 っている	-.076	.614	-.010	.069	.387
義務的	.083	.570	-.073	-.093	.346
成功	.334	.547	-.112	.039	.425
予測できない	.124	-.524	.107	.377	.444
自然な	-.086	-.285	.708	-.127	.606
ありのまま	-.150	-.201	.682	-.190	.563
身を任せている	-.498	.071	.603	-.041	.618
なされるがまま	-.505	.065	.588	-.031	.605
複雑	.167	-.031	-.121	.725	.570
単純	-.097	.219	.290	-.571	.468
単調	-.177	.451	.186	-.546	.567
変化がある	.189	-.262	.076	.523	.384
固有値	7.44	4.84	1.92	1.66	
累積寄与率 (%)	20.5	37.2	45.9	54.3	

- 長さが 10 秒前後で、無音声である。
- 三つのスタンスそれぞれが典型的に採用されるであろう対象を網羅する。
- 人間の機械的振舞いなど、主体そのものに対する典型的スタンスと振舞いに対するスタンスが不一致なものを含む。

具体的には次のとおりである。1 から 3 の動画はそれぞれ、意図、設計、物理の三つのスタンスによって解釈されると我々が考える典型的な振舞いを映した動画である。4 から 22 の動画は、主体そのものに対する典型的スタンスと振舞いに対するスタンスが不一致なものである。すなわち、二つのスタンスによる解釈が可能なものである。例えば、4 は意図的主体である人間が機械的振舞いをしているために、主体のみに注目すると意図スタンスを採用することになるが、振舞いのみに注目すると設計スタンスを採用することになる。各動画のスタンスの組合せは次のとおりである。4 から 9 が意図スタンス若しくは設計スタンス、10 から 16 が意図スタンス若しくは物理スタンス、17 から

22 が設計スタンス若しくは物理スタンスである。23 から 60 の動画は人間・動物・植物・ロボット・機械・乗り物・単純物・自然物の 8 種類のカテゴリーに属する主体が様々な振舞いを行うものである。

3.3 実験結果

評定値に対して因子分析を行った。因子の抽出は重み付けのない最小二乗法を用いた。因子数は固有値 1 以上の規準を設け 4 因子とした。バリマックス回転後の因子パターン、共通性、固有値、累積寄与率を表 4 に示す。また、表 3 の 2 列から 5 列に各動画に対する因子得点の全被験者の平均を示す。

3.4 考察

以下では抽出された四つの因子がどのような概念を表しているかについて考える。第 1 因子で因子負荷量の高い語句は感情、知性、判断、努力、達成、能動（受動の反意）のカテゴリーに属するものである。これらは、実験 1 において、アニメーション 1 を記述する際に多く用いられた単語とほぼ一致している。このことから、因子 1 はアニメーション 1 を見ることによって得られる概念に相当していると考えられる。第 2 因子で因子負荷量の高い語句は仕事、法則、正確、規則、予定のカテゴリーに属するものである。これらは、実験 1 において、アニメーション 2 を記述する際に多く用いられた単語とほぼ一致している。このことから、因子 2 はアニメーション 2 を見ることによって得られる概念に相当していると考えられる。第 3 因子で因子負荷量の高い語句は自然、受動カテゴリーに属するものである。これらは実験 1 において、アニメーション 3 を記述する際に多く用いられた単語とほぼ一致している。このことから、因子 3 はアニメーション 3 を見ることによって得られる概念に相当していると考えられる。アニメーション 1 から 3 はそれぞれ意図、設計、物理スタンスを採用することを想定して作られたものなので、因子 1 から因子 3 はそれぞれ意図スタンス因子、設計スタンス因子、物理スタンス因子と考えることができる。

第 4 因子で因子負荷量の高い語句は複雑カテゴリーに属するものである。これらの語句は実験 1 において、アニメーション 1 を記述する際に用いられた単語であるので、因子 4 はアニメーション 1 を見ることによって得られる概念に相当していると考えられる。しかし、因子 4 が因子 1 に含まれなかったことから、因子 1 の概念とは別の概念を表していると考えられる。因子 1 と因子 4 の違いについては総合考察で述べる。

以上のことから, Dennett の論考による三つのスタンスと近い三つの概念的カテゴリーが存在することが分かった. すなわち, 意図的主体の振舞いの特徴である感情や知性, 判断, 努力といった概念が意図スタンスという単一の因子に, 設計された主体の振舞いの特徴である, 単調, 仕事, 法則, 制約, 規則, 予定といった概念が設計スタンスという単一の因子に, 物理的な振舞いの特徴である自然, 受動という概念が物理スタンスという単一の因子に縮約されることが分かった.

以上で示されたのは, 振舞い理解において4種類の言語的概念カテゴリーが存在し, そのうちの三つはDennettのスタンスと類似しているということである. しかし, 実験1で考察したように, 三つのスタンスに相当する概念カテゴリーはDennettが主張する概念を忠実に表しているわけではない. どのような違いがあるかについては総合考察で述べる.

4. 実験3: アニメーションから得られる印象と言語的概念の相関

実験1, 2ではアニメーションの印象を言語によって記述したものがDennettの三つのスタンスの概念を含む4カテゴリーに分節されることが示された. このことはアニメーションを見ることによって得られる三つの概念と四つの言語的概念のうち三つが一致していることを示唆する. このことを別の視点から検証するために, 三つのアニメーションを見ることによって得られる概念とそれぞれの因子の間にどれぐらい相関があるかを調べる実験を行った. この検証は, 実験2で用いた動画を再度参照基準として用い, 同一の動画に対する語句による評定(実験2の結果)とアニメーションとの類似度の評定を対比することで行った.

4.1 被験者

工学系の学部生と大学院生20名(男性16名, 女性4名; 平均年齢21.9歳)が実験に参加した. いずれの被験者も実験1, 2に参加していない.

4.2 実験方法

実験方法は, アニメーションと動画の類似性を評定するという点を除いて実験2と同じである. アンケートのページには動画一つとアニメーション三つが表示され, 被験者はアニメーションそれぞれに対して7段階のリッカートスケール(1. 全く類似していない-7. とても類似している)によって類似性を評定することを求められた.

表5 因子得点とアニメーションによる評定値の相関係数
Table 5 Correlation coefficients between factor scores and scale ratings for animations.

	因子1	因子2	因子3	因子4
アニメーション1	0.598**	-0.379**	-0.313*	0.475**
アニメーション2	0.251	0.628**	-0.455**	0.043
アニメーション3	-0.378**	-0.129	0.646**	-0.063

4.3 実験結果

各動画に対する評定値の全被験者の平均を表3の6列から8列に示す. 各動画ごとに一元配置分散分析を行い, 有意なものに対してTukeyの方法による下位検定を行った. 検定の結果, 単一のアニメーションに対する評定値が5%以下で有意に高いものを太字で示した. また, 表5に全動画を対象として, 実験2で得られた因子得点とアニメーションに対する評定値の相関を求めたものを示す. 各因子についてアニメーションとの相関が最も高かったものを太字で示した. また, 無相関検定の結果, 5%で有意なものには*, 1%で有意なものには**を数値の右側に表示した.

4.4 考察

まず, 動画の性質に対するアニメーションの選択傾向について概観する. 表3の評定値をアニメーションごとに見ると, 評定値の高さと主体の種類に関係がないことが分かる. 例えば, アニメーション1に対する評定値が他のアニメーションより有意に高いものは1, 40, 50の動画であるが, 主体は人間, ロボット, 自動車と一貫していない. この傾向はアニメーション2, 3でも同様である. 一方で, 評定値の高いアニメーションの振舞いには共通性がある. アニメーション1の評定値が高い動画の振舞いに共通するのは行為の目標状態が明確に理解できるという点である. アニメーション2では振舞いには決められた手順で特定のタスクをこなすという共通性がある. アニメーション3では表面的な振舞いの類似性や共通する物理法則はないが, 物理現象という共通点がある. 例えば動画3は燃焼(酸化), 動画37, 39, 57などは風力(流体の運動)が振舞いを支配している法則である.

本実験における類似性の判定は7段階のリッカートスケールを用いて行った. 評定値の4は「どちらでもない」を表す. 評定値の平均を見ると, 平均が5(少し類似している)以上のものは4%(7/180)しかない. 一方で, 2(類似していない)以上3(あまり類似していない)未満は42%, 3以上4未満は30%である. すなわち, 被験者の多くは全体的に類似性を否定的

に捉えていたといえる。上記で考察した、単一のアニメーションに対する評定値が有意に高いものの中には 3 以上 4 未満の値のものがある。これらの動画では、消去法的にアニメーションが選ばれたものと考えられる。動画とアニメーションの間では振舞いそのものの見た目やタスク、状況の類似性が低いために、このような結果になったと考えられるが、消去法による選択であっても、選択の際に最も類似性が感じられるものを残す必要があるため、動画とアニメーションの間に何らかの類似性を感じていたものと考ええる。

次に、アニメーションを見ることによって得られる概念と言語的概念の相関について考察する。表 5 によると因子 1 とアニメーション 1, 因子 2 とアニメーション 2, 因子 3 とアニメーション 3 の間の相関が比較的高いことが分かる。以上のことから、因子 1 から 3 までの言語的概念カテゴリーとアニメーション 1 から 3 を見ることによって得られるそれぞれの概念は一致しているといえる。

因子 4 と最も相関の高いアニメーションはアニメーション 1 であることが分かる。因子 4 とアニメーション 1 は因子 1 ほど高い相関はないが、他の二つのアニメーションに比べると高い相関をもっているため、因子 4 はアニメーション 1 から得られる概念に相当していると考えられる。これは実験 2 の結果と一致している。

結果を全体として見ると、因子 1 から 3 の因子得点とアニメーション 1 から 3 の評定値の間に相関があることが確認された。しかし、個別の動画に注目すると、齟齬が見られるものがある。因子 1 から 3 の中で因子得点が最大である因子の番号と評定値が最大であるアニメーションの番号が一致していない動画は 20 個である（該当する動画は表 3 の振舞いの記述欄に網掛けで表示）。この原因を考える。

例えば、4 の動画は和菓子職人が生地を伸ばす振舞いである。この動画に対して、言語的理解では意図スタンス、アニメーションとの類似性による判定では設計スタンスが採用されたといえる。また、40 の動画は人型ロボットが反動をつけて起き上がる振舞いである。この動画に対して、言語的理解では設計スタンス、アニメーションとの類似性による判定では意図スタンスが採用されたといえる。これらの齟齬に共通するのは、言語的理解では、人間なら意図スタンス、機械なら設計スタンスのように、その対象に対して取られる典型的スタンスが採用されているのに対し、アニメーションとの対比では、振舞いそのものに注目したスタンス

が採用されていることである。この一因としては、アニメーション中の主体が円という抽象的な図形であったために、主体そのものよりも振舞いに対して注意が向けやすかったことが考えられる。20 個中 12 個 (4, 5, 10, 11, 12, 18, 26, 36, 40, 43, 47, 58) の動画における齟齬はこれと同一の説明が可能である。

5. 総合考察

Dennett のスタンスの考え方は、振舞いの起源（意図）や振舞いを制御している要因（設計原理、物理法則）といった原理を想定した上で振舞いの理解と予測を行うというものである。しかし、実験 1 で明らかになった 16 種類の内容は原理そのものについて言及するものよりも、原理から発生する振舞いの性質を表現した単語によって構成されていた。被験者らが実際に原理を想定した上で振舞いを理解していたかどうかは明らかではないが、このことは、具体的な原理を想定しなくても振舞いの性質だけから振舞い理解（スタンスの選択）がなされることを示唆する。実験 2 では 16 種類の内容が四つの抽象的概念に縮約され、四つの内容のうち三つは Dennett の論考によるスタンスとおおむね一致していることが示された。これは、表面的な振舞いの性質による理解と Dennett のいう振舞いの原理に基づく理解がほぼ一致していることを意味する。このことから、Dennett の心的姿勢のカテゴリー化がある程度の妥当性をもったものであることが示されたと考えられる。

しかし、Dennett のいう原理的理解と我々の結果が示す振舞いの性質による理解の間には乖離がある。実験 2 の考察では、因子 1 が意図スタンス、因子 2 が設計スタンス、因子 3 が物理スタンスに相当するものであると考えた。因子 1 で因子負荷量の高い語句は意図的な主体が生成する振舞いの多様な性質を表しており、原理そのものについて言及しているわけではないが、Dennett の考える意図スタンスに近い概念であると考えられる。しかし、因子 2 は設計された主体が生成する振舞いの性質を表しているものの、設計原理についての概念を表しているというよりは、法則による決定性を表していると考えられる。因子 3 に関しては、実験 1 の考察で述べたように、物理法則を想定したものではなく、単に振舞いの主体の受動性を表す概念である。

因子 4 は実験 2 と 3 の結果からアニメーション 1 から得られる概念を表していることが分かったが、因子

1とは独立であるために、意図的主体だけが生成する振舞いの性質を表しているのではないと考えられる。実際に、表3の因子得点を見ると、必ずしも意図的主体の振舞いにおいて高い値を示しているわけではない(例えば8, 14, 45)。これらの動画中の主体の振舞いに見られる共通点はその振舞いが複雑であったり、予測が困難だということである。そのために、因子4はDennettのスタンスの中に含まれない、予測できない複雑な振舞いに関する概念を表していると考えられる。予測困難さや複雑さを発生する原因としては、偶然や乱数、カオス力学、意図が存在するが、因子4ではそれらが分離されずに一つの因子に縮約されているものと考えられる。

これらの議論に基づいて、Dennettの考えるスタンスは論理的に妥当ではあり、そのような分類がなされていることを示唆する間接的結果が得られたが、実際の振舞い理解には異なる戦略が採用されている(カテゴリー化がなされている)と結論づける。以下に我々の研究によって明らかになった戦略を示す。Dennettの考えに倣ってスタンスと表記する。

意図スタンス 意図のような心的状態をもつ主体が生成する多様な振舞いを理解し予測する戦略。

決定スタンス 法則などによって振舞いが決定論的に決定する主体の振舞いを理解し予測する戦略。

受動スタンス それ自身に振舞いを決定する原因や法則をもたず、外部要因によって振舞いが決定するような主体の振舞いを理解し予測する戦略。

複雑スタンス 振舞いを理解、予測できないものとして捉える戦略。理解も予測もしない。

6. む す び

本論文ではDennettが哲学的論考で指摘している三つのスタンスが心理学的に妥当であるかについて検討するために三つの実験を行った。まず、Dennettのスタンスをできるだけ忠実に再現したアニメーションを作成し、アニメーションに対する被験者の印象記述を分析することで、アニメーションから得られる概念を言語化する実験を行った。実験の結果、三つのアニメーションの理解のために16種類の特徴が用いられていることが分かった。また、16種類の特徴が三つのスタンスに相当する概念的カテゴリーを形成していることを示唆する結果が得られた。これを検証するために、16種類の特徴が三つの概念に縮約されるかどうかを確認する実験を行った。これは60個の動画を

参照基準として用い、16種類の特徴を構成する26個の印象語句が適切であるかどうかの評定を求める被験者実験によって行った。評定結果に対して因子分析を行ったところ、4因子が抽出され、そのうち3因子はDennettの三つのスタンスと同等のものであることが分かった。これはアニメーションから得られる三つの概念と四つの言語的概念のうち三つが一致していることを示唆する。このことを別の視点から検証するために、三つのアニメーションから得られる概念とそれぞれの因子の間の相関を調べる実験を行った。この実験によって因子1から3までの言語的概念カテゴリーとアニメーション1から3を見ることによって得られる概念は一致していることが分かった。また、因子4はアニメーション1から得られる、因子1とは異なる、概念に相当していることが分かった。

以上の実験によって、表面的な振舞いの性質による理解とDennettのいう振舞いの原理に基づく理解がほぼ一致していることが分かった。しかし、Dennettの主張のような原理帰属による理解がなされているという直接的証拠は得られていない。そこで、実験結果について考察を行い、実際の振舞い理解においては振舞いの性質が注目され、意図、決定、受動、複雑の四つの概念的カテゴリー化が行われていると結論づけた。

Dennettのスタンスは振舞い予測のための戦略である。一方で我々が取った手法は印象記述や印象、類似性の評定に基づいた概念的カテゴリーの抽出である。実験においても、振舞いを予測する戦略とは異なる特徴に着目して、被験者がアニメーション・動画を評価した可能性が考えられる。更に、上記で示した四つの概念的カテゴリーもスタンスと称したものの予測戦略として機能することを確認したわけではない。四つの概念的カテゴリーが予測戦略の分類として妥当なものであるかどうかを調べるのは今後の課題である。

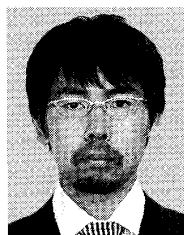
文 献

- [1] D.C. Dennett, *The Intentional Stance*, Cambridge, Mass, Bradford Books/MIT Press, 1987.
- [2] G. Gergely, Z. Nádasdy, G. Csibra, and S. Bíró, "Taking the intentional stance at 12 months of age," *Cognition*, vol.56, no.2, pp.165-193, 1995.
- [3] D. Premack and A.J. Premack, "Motor competence as integral to attribution of goal," *Cognition*, vol.63, no.2, pp.235-242, 1997.
- [4] D. Premack, "The infant's theory of self-propelled objects," *Cognition*, vol.36, no.1, pp.1-16, 1990.
- [5] B.J. Scholl and P.D. Tremoulet, "Perceptual causality and animacy," *Trends in Cognitive Sciences*, vol.4,

no.8, pp.299-309, 2000.

- [6] F. Heider, The Psychology of Interpersonal Relations, Lawrence Erlbaum Associates, 1958.
- [7] G. Gergely and G. Csibra, "Teleological reasoning in infancy: The naïve theory of rational action," Trends in Cognitive Sciences, vol.7, no.7, pp.287-292, 2003.
- [8] A. Matan and S. Carey, "Developmental changes within the core of artifact concepts," Cognition, vol.78, pp.1-26, 2001.

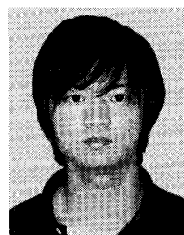
(平成 23 年 4 月 4 日受付, 8 月 1 日再受付)



寺田 和憲

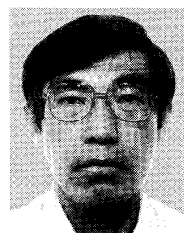
1995 阪大・工・精密卒. 2001 奈良先端科学技術大学院大学情報科学研究科博士後期課程了. 2000 独立行政法人通信総合研究所特別研究員. 2002 岐阜大学工学部応用情報学科助手. 2007 同助教. 博士(工学). 知性に関わる哲学的問題に興味をも

つ. 人工知能学会, 日本ロボット学会, 情報処理学会, ヒューマンインタフェース学会各会員.



岩瀬 寛

2011 岐阜大・工・応用情報卒. 同大学院工学研究科博士前期課程在学中.



伊藤 昭 (正員)

昭 47 京大・理・物理卒. 昭 54 同大学院理学研究科博士卒(理博). 同年郵政省電波研究所(現通信総合研究所)入所. 平 10 山形大学工学部教授. 平 12 岐阜大学工学部教授, 現在に至る. 人工知能, マルチエージェント学習, コミュニケーションの

認知科学の研究に従事. 情報処理学会, 人工知能学会, 日本認知科学会, 日本ソフトウェア科学会各会員.