

# 身体化人工知能

月本 洋  
(株) 東芝 研究開発センター

人工知能には記号主義とコネクショニズムがあるが、各々欠点があり、それを統合する記号パターン研究がなされている。本論文では身体化人工知能(Embodied AI:EAI)を提唱する。これは、コネクショニズムと記号主義を仮想的身体運動である想像を通して統合することを特徴とする。EAIでは、記号処理は感覚運動用に訓練された人工神経回路網から構成される。想像が身体を必要とするので、EAIは、ロボットでは実現できる。EAIは、身体のない計算機では、想像の言語的側面であるメタファー等に基づく人工知能(Metaphor Based AI:MBAI)として実現される。

## Embodied AI

Hiroshi Tsukimoto  
Corporate Research & Development Center, Toshiba Corporation

The integration of connectionism and symbolicism is studied in order to solve their defects. This paper presents a new paradigm called Embodied AI(EAI) which is an integration of connectionism and symbolicism through imaginations. EAI insists that symbolic processings should consist of artificial neural networks trained for bodily movements. Imaginations need bodies. Robots can realize imaginations, and so can realize EAI. Current computers do not have bodies, and so cannot realize imaginations. Since the main linguistic aspects of imaginations are metaphors, EAI is realized as Metaphor Based AI (MBAI) in the current computers.

### 1 はじめに

現在、計算機で人間の心(知能)を模倣する試み(これを人工知能と呼ぶ)が行われているが、その主な二つの手法は記号主義と結合主義(コネクショニズム)である。記号主義の欠点を解決するという触れ込みで登場したコネクショニズムは、脳の神経回路網を模倣したモデルを用いるが、そのモデルは非線形関数モデルであり、記号主義で扱える述語を表現できない等の問題がある。そこで、記号主義とコネクショニズムを統合する記号パターン統合という研究がなされている[10]。筆者もその研究を行っているが、人工神経回路網を非古典論理で推論できること、すなわち人工神経回路網も記号であることを示した。これは、コネクショニズムと記号主義は峻別するようなものではないことを意味する。この二つの主義は、知能の本質である身体性を無視している。本稿では、身体性を考慮に入れた Embodied AI を提唱する。Embodied AI の特徴を極めて簡単に言えば、仮想的身体運動であるところの想像(イメージ)をこころ(知能)の中心的存在として扱うということである。この Embodied AI は、身体のない現在の計算機では、イメージの主な言語的側面であるメタファーに基づく人工知能になるが、これは(言語)記号と人工神経回路網を、

メタファーを通して統合するという記号パターン統合の一手法になる。Embodied AI は、ロボットでは、イメージを中心にしたロボットの構成原理を提示することになる。また、Embodied AI は、記号を、イメージを通して身体にそして世界に基礎付けるので、記号定礎問題(Symbol grounding problem)の解にもなっている。本稿では、コネクショニズムの哲学的問題としては、分散表現と心の自然化についても簡単に述べる。

### 2 コネクショニズムについて

#### 2.1 人工知能の手法について

記号主義は、主に離散的な記号の形式的操作、即ち計算によって知能を実現しようとする主義である。常識推論等を、古典論理を拡張することによって実現しようとするいわゆる論理主義はその典型である。典型的な形式的知識表現方法は、一階述語古典論理に基づいた知識表現方法である。この表現方法は厳密なため、人工知能が取り扱う現実世界の記述や常識の記述には不向きであり応用上種々の不具合がある。また、規則(ルール)を用いて知識処理を行おうとするエキスパートシステムであ

るが、その規則を獲得するのが困難であることが分かり、それは知識獲得問題として知られているが、現在も有効な解決方法が存在していない。記号主義に対して結合主義（コネクショニズム）という立場がある。ここではコネクショニズムを人工神経回路網と PDP のアプローチと同義に用いる。これは主に神経回路網のモデルを用いて各種計算を実行することによって心（知能）を実現しようとする主義である。並列処理や分散処理を特徴としている。

## 2.2 コネクショニズムの実際

神経回路網のモデルもホップフィールド型、ボルツマン型などいくつかあるが、情報処理に最も使われているのは3層フィードフォワード人工神経回路網である。3層とは、入力層、中間層、出力層である。しかし、これは、脳の神経回路網を余りにも単純化して、実際の脳の神経回路網のモデルと呼んで良いかどうか疑問である。たとえば、実際の脳の信号はパルスであるが、上記の人工神経回路網の信号は連続である、という違いもある。また学習方法であるが、最も良く使われているのは、誤差逆伝搬法であるが、これは人間の神経回路網の学習方法とは大きく異なっている。実際の神経回路網を忠実に人工神経回路網でモデル化する研究もあるが、脳の（認知）機能（通常は非常に多くの神経回路で実現される）を（簡単な）人工神経回路網でモデル化しようとするものが多い。したがって、多くの人工神経回路網は、人間の脳の神経回路網のモデルというよりは、心（知能）の非線形関数による回帰モデルである。次に、人工神経回路網の能力であるが、記号処理に関して、代入ができない、全称記号を用いた述語論理式が表現できない、等の問題がある。また、人工神経回路網は暗箱であるという問題も有る。コネクショニズムと記号主義の欠点を、解決しようとする研究が行われており、これらは以下の3種類の研究に分類できる。1) 構造化：人工神経回路網を構造化して述語処理、自然言語処理を行う研究 2) 分散表現：構造化せずに、分散表現のままで種々の情報処理を行う研究 3) 記号パターン統合：記号処理と人工神経回路網（パターン）処理を統合して高度な情報処理を行う研究筆者は3番目の研究を行っているが、それを以下に簡単に紹介する。

## 2.3 暗黙知の言語化（数式から言語へ）と人工神経回路網の論理的推論

（人工）神経回路網に非言語的に貯えられている知識を暗黙知と言い、暗黙知の言語化とは、人工神経回路網から記号を抽出することである。ここで言う記号とは、現状では、古典命題論理の命題である。数学的にはブール関数と線形関数の両者を包含する多重線形関数空間が基本になる。多重線形関数とは2変数でいえば  $axy+bx+cy+d$  (但し  $x, y$  は変数、 $a, b, c, d$  は実数。) である。多重線形関数空間で、人工神経回路網をブール関数で近似することができる。定義域が連続の場合は、人工神経回路網を連続ブール関数で近似する。これによ

り、暗箱であるという人工神経回路網の欠点を言語的知識の抽出で解決している [15]。次に、人工神経回路網の論理的推論であるが、これは人工神経回路網を論理的に推論することである。これが可能なのは、人工神経回路網を非古典論理の論理的命題と見做せるからである。古典論理では人工神経回路網は推論できないので、非古典論理で人工神経回路網を推論することを考える。非古典論理には種々あるが、中間論理 LC、Lukasiewicz 論理、product 論理の完全な連続値論理で基本的に人工神経回路網を推論できる。その理由は以下の通りである。人工神経回路網が離散定義域の時は多重線形関数である。多重線形関数空間は上記の非古典（連続値）論理の代数モデルである。別の言い方をすれば、上記の3論理は、多重線形関数空間に対して完全である。したがって、上記の3個の非古典（連続値）論理で人工神経回路網が推論できる [16]。人工神経回路網が論理的に推論できれば、記号主義的手法の一問題である知識獲得問題が部分的に解消できる。このように、人工神経回路網は論理的に推論できるのであるから、人工神経回路網にもとづくモデルを用いるコネクショニズムを記号主義から峻別することはできなくなる。コネクショニズムは、記号主義に対峙するものとして扱われる場合が多いが、このような状況を考えると、コネクショニズムは、通常記号主義が主に離散的な記号を扱うのに対し、実数を扱うことから来る連続性を特徴としている違いはあるが、記号主義とコネクショニズムの違いはそれほど大きくないと思われる。

## 2.4 人工神経回路網は、非線形関数のグラフ表現である。

人工神経回路網は（現状では）最も優れた（万能）非線形回帰モデルとすることを考慮すれば、実際に人工神経回路網で研究されていることは機能主義に近い。と言える。人工神経回路網は脳の神経回路網のモデルと言う側面もあるので、低水準の認知モデルにはなりうる。高水準の認知モデルにはなり得る場合もあるが、それは、人工神経回路網が脳の神経回路網のモデルであるからではなく、（万能）非線形回帰モデルであるからである。このような見方をすれば、脳は（万能）非線形回帰モデルの一実現形態である、とも言える。コネクショニズムは、非線形関数である人工神経回路網のグラフ表現の結線由来するものであるが、線形関数も結線を用いてグラフ表現ができるので、線形関数を多用している統計学もコネクショニズムと言わねばならないが、そのようには呼ばない。すなわち、「コネクショニズム」と言う語は、コネクショニズムの特徴を適確につかんでないと言える。コネクショニズムの特徴はコネクション（結合）ではなく、非線形関数にある。

## 2.5 学習の重要性について

また、人工神経回路網はデータを用いた学習を必要とする。実際の神経回路網も学習が必要である。先天的な神経回路の回路構成も、長い進化の過程で、人間が集団

で外界のデータを学習してきたものである。このように考えると、神経回路の機能は、神経回路に帰着されると言うよりは、学習データが存在する(した)外界に帰着される性質のものである。なぜそのような機能になったかは、素材である神経回路を調べるよりは、学習データを調べるべきであろう。神経回路はその機能を実現する素材を提供しているが、これは、学習を可能にしていると同時に、学習の制約としても働いている。したがって、脳が半導体から構成されていたとすれば、外界が同じ学習データを提供してきたとしても、別様の脳に仕上がっていたであろう。

### 3 コネクショニズムの哲学的問題について

#### 3.1 分散表現について

最初に一つ注意しておきたいが、分散表現は、脳に関することであり、心に関するのではない。心に関する「分散表現」はイメージであるが、これは後述する。コネクショニズムの特徴の一つに分散表現があるが、分散表現は良いことなのであろうか。英語が上手な日本人は、英語を処理する部分と日本語を処理する部分が異なり、英語が下手な日本人は、同じところで処理する、とのことである。機能分化しているほうが優れているのである。また、たとえば、擬音は聴覚野と言語野で反応するが、これは現在の人間の脳の言語野が聴覚野から分離したと推察される。すなわち、昔の人間の脳は、聴覚野と言語野に機能分化していない。すなわち、分散表現されている。脳は分散表現から機能分化へと進化してきた。分散表現が良いと言うならば、言語機能のない昔の人間の脳の方が良いことになるが、これはおかしいであろう。さらに進化が進めば、現在の脳が分散表現されている部分も機能分化するであろう。コネクショニズムの分散表現に対する高い評価は、現在の人間の脳が、進化の最終段階であれば、そうであろうが、進化の途上であれば、分散表現は、単に現在の人間の脳が未分化すなわち不完全であることを意味しているだけであろう。筆者は、現在の人間の脳が進化の最終状態であるとは考えないので、現在の人間の脳の分散表現は、その脳が未分化すなわち不完全であることを意味していると、考える。

#### 3.2 心の自然化について

##### 3.2.1 相互作用としての心

人間は環境と他の人間との相互作用の中で生きている。このように相互作用は二つに分類できるから、心も、対自然の相互作用の部分と対人間の相互作用の部分の二つに分けて議論したほうが良いのではないだろうか。対自然の部分は、たとえば触覚などがそうであろうか。対人間の部分は、意識などがそうであろうか。一般に相互作用を記述する時には、その相互作用に関して閉じている系について記述するのが基本である。意識、心を相互作用とするならば、一人の人間の心、一人の人間の意識を

議論するというのは議論の仕方として、片手落ち、もしくは、そのような記述の仕方は不適切なのではないか。少なくとも、二人はほしい。相互作用として記述するならば、意識は、一人間の意識ではなく、人間集団の意識、すなわち人間集団の構成員間の相互作用である、と記述すべきであろう。意識は、自分を意識するのが基本ではなく、他人を意識するのが基本である。自分の意識は、他人への意識の仕方から学ぶ。心も、同様に、自分の心の理解の仕方は他人の心の理解の仕方の模倣である[6]。私は、一人でいて、一人で考えているときも、他人と話すのと同じ方法でしか自分自身と話せない。その話し方は人間集団(社会)から学んだものである。社会はこのように意識、心の中に入り込んでいる。心は社会的である。意識も社会的である。他人の心の理解の仕方でも自分の心を眺めたのが一人称的な心である。心という相互作用が、対人間の相互作用と対自然の相互作用の二つから構成されているとすれば、対人間の相互作用である心(意識など)で対自然の相互作用の心(感覚)を理解しようとしても理解しきれないものが存在するのは当然かもしれない。それが感覚質などであろうか。二つの相互作用は、同じ心と呼ぶには、異質なもののなのかもしれない。

##### 3.2.2 心/脳と重力/電磁力の並行的な議論

電磁力で結合している二つの物質間に働く重力を記述するのに、一つの物質だけを取り上げて記述することも可能であるが、基本的な記述の単位は、二つの物質である。質点系の力学を一質点に注目して記述することの意味はあるのか。心も同様ではないだろうか。次のような対応を取れるであろう。人間集団/物質集団、人間/物質、心/重力、脳/電磁力(物質は、基本的に電磁力で結合している。)ふたりの人間間の相互作用(心)がその人間の脳に作用する。これは重力が、その重力が作用している二つの物質の電磁力的結合に作用するのとおなじである。重力はその物質を構成する分子の結合状態に影響を及ぼす。心もその人間の脳の状態に影響を及ぼす。逆に、物質を構成する分子の電磁力的結合状態に変化があれば、その変化はその物質が関与している重力の相互作用に影響を及ぼす。心に関しても、同様である。脳の状態が変われば、心にも変化が生じる。

##### 3.2.3 重力/電磁力の並行的議論による心/脳に関する哲学的問題の素描

心と脳の因果関係は、心が原因の場合もあれば、脳が原因の場合もある。電磁力が原因の場合もあれば、重力が原因の場合もある、のと同様である。意識などの心の機能が人間集団の構成員間の相互作用であるとするならば、その存在論的身分は、電磁力で結合している二つの物質間に働く重力の存在論的身分にぞらえて議論できるのではないだろうか。重力はどこに存在するのであろうか。それが作用している物質に存在するのであろうか。そうではないであろう。重力そのものは見えないが、見えないからといって、存在しないと誰も考えない。心

も、同様に考えられないのだろうか。複数の人間間に作用する相互作用であり、それがどれか一つの人間の中に存在するとかしないとかの議論は、あまり意味がないのではなかろうか。心を脳に還元する説に関してであるが、心を人間集団の構成員間の相互作用とする立場からすれば、心を脳に還元するという話は、重力を電磁力に還元する話と同じである。重力が電磁力に還元できないのと同じように、心も脳に還元できない。心を神秘的なものとしてではなく、自然なものとして把握したいということに関しては、従来の自然化は、心を脳に還元するという形で実現しようとしているが [8]、心を人間集団の構成員間の相互作用と捉えることも、自然化の一つの実現形態である。心を脳に帰着させることで、神秘性がなくなり、自然化されるというのは、脳が物理的なもしくは可視的な存在だからである。それと同様に、心を人間集団(社会)に帰着させることで、神秘的でなくなり、自然化される。というのは、人間集団(社会)が物理的なもしくは可視的な存在だからである。重力の存在は、物資が動くことで確認できる。心という相互作用の存在も、人間が動く事で確認できる。愛憎などで殺人、自殺などが起こる。

### 3.2.4 心を脳に還元したいのはなぜ？

心を脳に還元したいのは、心を神秘的なものとして扱いたくないからである。脳に還元できれば、見えるものになる、からである。見えるものに帰着させることで、自然化し、安心するのである。それは、人間が、すべてのものを超越したいという欲求とほぼ同じなのでないだろうか。人間は中間者(人間集団の構成要素)なのですべてのものを超越するなどということは無理であろう。心は人間から見れば神秘的であるが、人間集団から見れば神秘的ではない。人間集団は実在している。帰着もしくは還元は、通常、あるものからそのものの構成要素に対して行われる。したがって、心を人間集団に帰着させることで、自然化すると言うのに対しては、違和感を持たれる読者は多いであろうが、帰着を構成要素でないものに対して行っても良いであろう。人間が心を脳に帰着させて可視化したいというのは、重力作用を受けている物質が、重力を電磁力で結合しているその物質に帰着させようとしているようなものではないであろうか。

## 4 記号主義とコネクショニズムの想像による統合-Embodied AI-

記号主義は心のモデルとしては、言語(社会)よりのモデルであり、コネクショニズムはその逆に、脳よりのモデルであると言える。記号主義とコネクショニズムを統合すると言うことは、言語(記号)と脳のモデルを統合すると言うことを意味する。言語(記号)と脳のモデルを統合するために、人間がどのように言語(記号)を脳で処理しているかを見てみよう。まず、言語を中心とした記号の理解を考える。記号の理解とは、記号の意味の理解である。ところで、人間の知能の特徴、すなわち、

人間が他の動物より知的であると思われる部分は何であろうか。種々の意見が有ろうが、それは言語であろう。したがって、知能の中心的なものとして言語を扱うことにする。なお、以降では、現在の計算機で処理できる人工知能を記号的人工知能と呼ぶ。ロボットと対比させるためである。

### 4.1 理解について

「意味の理解」ということばは、あいまいであるが、結構異なる二つの“意味”がある。この二つの“意味”は明確に識別する必要がある。意味に関しては、諸説有り、万人が合意するような定義は存在しないが、一つは、イメージ、表象である。これは、記号論の言う記号内容である [11]。もう一つは、用法である [18]。この二つに対応して、意味の理解も大きく二つに分けられる。一つは、想像可能性であり、もう一つは、記号操作可能性である。想像可能性を一人称的理解と呼ぶ。どのように想像できたかどうかはその本人以外に観測できないから、このように呼ぶ。これに対して、記号操作可能性を三人称的理解と呼ぶ。記号操作できるかどうかは、他人でも観測できるからである。たとえば、5次元空間は、記号操作ができるが、想像できない。これに対して、3次元空間は記号操作もできるが、想像もできる。意味を理解するとは、この二つの事を指すが、この二つを混用している。たとえば、「坊主が屏風に書いた坊主が屏風に書いた坊主の絵を書いた」の文は、「坊主が屏風に(((坊主が屏風に書いた)坊主が屏風に書いた)坊主が屏風に書いた)坊主の絵を書いた」と構造化され、それなりに理解できる。「坊主が屏風に書いた」は文法的には何回でも挿入できる。われわれはそのような文を文法的に正しい文として受け入れねばならない。即ち、我々はそれを理解せねばならない。しかし、あまりにもその挿入の数が多くなると、我々は理解できなくなる。我々はそのような文を理解できると言うべきなのであろうか、理解できないと言うべきなのであろうか。この問題は、そのような文は三人称的には理解(記号操作)できるけれど、一人称的には理解(イメージ)できないということで整理できる。記号操作可能性は想像可能性に基づく。想像可能性無き記号操作可能性は、すぐ消える。たとえば、数学の公式の丸暗記である。イメージ抜きの理解は、浅い理解である。2節で、人工神経回路網は、代入、全称記号の述語論理式の処理ができないと述べたが、それは、人間の「脳もできない」、すなわち、想像できない、のであるから、脳のモデルとしてはそれで良いのである。

### 4.2 想像について

最近の脳の実験では、想像する場合と実際の場合で脳と同じ部位を使っていることが分かってきた [9]。たとえば、指を動かすのを想像する場合と、実際に指を動かす場合では、同じ脳の部位が活性化されることが多くの実験で確認されている。正確に言うと、想像の場合には、筋肉からのフィードバックが無いが、この事実は、

身体運動の想像とは仮想的な身体運動であることを意味する。画像（たとえば三角形）を思い浮かべるときも、眼球が運動していることが確認されている。したがって、われわれが身体運動を想像するときには、仮想的に身体を動かしているのである [3]。これが、「想像には身体が必要である。」ということである。身体運動以外の想像の時には、複数の身体運動の想像を組み合わせ、それに何かが付加されている。たとえば、暗算するときには、言語野、前頭葉等に加えて、運動野、聴覚野、視覚野が動く。すなわち、暗算するときには、音か絵を用いている。音を用いるときには、耳や舌を仮想的に使い、絵を用いているときには、目を仮想的に使う。ある特定の画像の描像機能は、それ専用のプログラムと装置を製作すれば、実現できるかもしれないが、一般的な画像の描像機能は、眼球運動と同等のプログラムと装置を製作せねばならないであろう。別の言い方をすれば、一般的な画像の描像機能を実現しようとすると、出来上がるものは、眼球運動を模倣したプログラムと装置になるであろう。

### 4.3 メタファーに基づく人工知能

想像には身体に基づいているので、身体がない計算機には、想像はできない。したがって、記号的人工知能では想像可能性（一人称的意味の理解）は実現できない。これが、記号的人工知能の限界である [17]。記号的人工知能で実現できる範囲は、記号操作可能性（三人称的意味の理解）の範囲内であるが、もっとも自然知能（人間）に似ている、すなわちもっとも想像可能性に近いことができる記号的人工知能は、想像の言語的側面であるメタファー（比喩、隠喩）等の記号操作に基づくものである。想像力の言語的側面はメタファーだけではなく、メトニミー（換喩）、シネクドキ（提喩）とあるが [12]、主要なのはメタファーなので、メタファーに注目する。これがメタファーに基づく人工知能である [13][14]。すなわち、メタファーに基づく人工知能は、計算機による想像の模倣である。記号的人工知能は想像そのものは実現できないが、想像の形式操作は実現できる。想像の形式は、経験の形式の拡張である [4][7]。想像の形式は、想像の記号的側面であるメタファー（など）の形式である。メタファーに基づく人工知能はメタファーの基本領域と応用領域と投射から構成される。包含等の単純な基本領域は形式化ができるが、他の多くの基本領域は、このように形式化できず、人工神経回路網等で構造化されることになる。メタファーに基づく人工知能から、数理論理学を眺めると、命題論理は、包含の領域の形式体系になる。命題論理の論理積、論理和、否定は、他の一般の述語と同様に述語表現しても良いわけであるが、多分人間にとって包含はもっとも基本的なメタファーなので、この包含に関しては、述語表現せずに、それらの記号操作を具体的に書き下した（これを命題化、命題表現と呼ぶ）のであろう。述語表現では何も（三人称的に）理解できないが、命題表現では（三人称的に）理解できる、すなわち記号操作できる。他の述語も、包含という「述語」が「命題化」されているように、「命題化」される（べきである）。記号的人工知能というと、論理による人工

知能を思い描く人が多いと思う。論理というと、数理論理学を意味する場合が多い。しかし、論理とは、辞書には、ものの考え方、と書かれている。論理とは、ものの考え方なのだから、想像力をどう使うかも、論理になるであろう。すなわち、想像力の形式を論理とみなしても良いであろう。筆者は、この想像力の形式の方が、（人工）知能の論理としては、記号論理学より、適切であると考ええる。記号論理学は、数学のための論理学である。想像力の形式であるが、想像それ自体は、各人の外に出せないで、科学的に扱えない。そこで、想像力に関して、科学的に、公共的に、客観的に扱える、範囲となると、それが、想像力の言語記号的側面であるメタファー（等）の形式となるのである。メタファーに基づく人工知能は、人間の理解に必須である想像を中心に、現在の論理を、再編成してみようという試みともいえる。自然言語に関しても、想像を核とした理論を構築しようという試みであり、認知言語学の考え方と同じである。メタファーに基づく人工知能とコネクショニズムの関係であるが、人工神経回路網は、心（知能）のモデルというよりは、脳のモデルなので、言語処理等の脳の高度な機能に関しては、単に言語処理現象を非線形帰属しているだけであり、本来の、脳のモデルで心のモデルを作成するというにはなっていない。コネクショニズムには、脳のモデルである人工神経回路網で心（知能）のモデルを構成する基本指針がない。メタファーに基づく人工知能は、コネクショニズムに対して、低次の認知（運動）などの人工神経回路網を、イメージを通して構築して、認知や言語処理モデルを作成するという基本方針を提示する。

### 4.4 ロボット

現在の計算機に限定しなければ、想像の実現方法として、ロボット [1] が考えられる。ロボットは身体があるので、想像ができる可能性が有る。その場合、（知能）ロボットの言語処理の部分は、メタファーに基づく人工知能と同様なものとなる。前に、人間が言語理解を仮想的な身体運動であるイメージを用いて行っていることを述べた。したがって、知能ロボットの構成原理は、「言語理解のプログラム（回路）は、身体運動のプログラム（回路）を一部分として含む。もしくは、複数の身体運動のプログラム（回路）に何らかのプログラム（回路）を付加したものが、言語理解のプログラム（回路）になる。」となる。これは、イメージ処理を通して言語処理と身体運動を統合するようなロボットの構成原理である。計算機の場合は、炭酸飲料水が飲めないで、その炭酸飲料水の味は想像できないが、ロボットの場合には、何か飲むという行為は基本的に可能なので、その味を想像できる可能性がある。しかし、ロボットがガソリンを飲む場合に、この味は人間には理解できない。人間とロボットが、どれほどお互いに理解しあえるのであろうか。これは、難しい問題である。このような問題はすでに存在している。たとえば、人間と猿がどのくらいお互いに理解し合っているのであろうか。人間と猿の相互理解の問題と、人間とロボットの相互理解の問題は基本的に同じ

であろう。味に関しては、計算機が味覚を分からなくても、もしくは、ロボットの味覚が人間の味覚と同じでなくても、それなりに会話が成立するであろう。しかし、身体が無ければ、空間（存在）そのものが理解できない。身体が無ければ、存在そのものが理解できないので、味覚とは大きく違い、意味の理解の根幹そのものが成立しないであろう。しかし、この場合の身体とはどこまでを指すのであろうか。難しい問題である。最後に、意識、志向性についても、少し言及したいと思う。現在の計算機での意識の実現は不可能であると思う。ロボットでは可能であると思う。しかし、知的なロボット（知能ロボット）ができるには、もう一つ条件が必要であろう。それは、ロボットの社会、集団である。3.2で考察したように、人間の心が人間間の相互作用であるならば、ロボットの心もロボット間（もしくは、ロボットと人間間）の相互作用である。したがって、ロボットに心を持たせるには、ロボットの集団、社会が必要なのである。知能ロボットは、群知能ロボットとしてしか存在し得ない。メタファーに基づく人工知能、イメージを中心に構成した知能ロボットは、イメージと主な言語的側面であるメタファーを通して記号を身体にそして世界に基礎付けるから、総称して Embodied AI と呼ぶ。日本語では身体化人工知能であろうが、「身体化」はあまり適切ではないと考える。読者から適切な日本語を教えていただければ幸いである。いわゆる記号定礎問題 (Symbol grounding problem)[5] もイメージ（メタファー）をそして身体を通して世界に基礎を定めることができると考える。そして、本節の冒頭の、言語（記号）モデル中心の記号主義と脳モデル中心のコネクションイズムは、イメージ（メタファー）を通して統合されるのである。

## 5 おわりに

本稿では、コネクションイズムの技術的側面と哲学的側面について述べ、コネクションイズムを超えるために、Embodied AI を提示した。これは、記号が記号表現と記号内容から構成され、人間がその記号内容であるイメージを用いて記号処理を行っており、さらにそのイメージが仮想的な身体運動（とその拡張）であることに注目して、現在の計算機では、イメージの記号的側面であるメタファー（等）に基づく人工知能を構築することを、知能ロボットでは、イメージ（想像）を中心にした構成原理を、提唱するものである。

## 参考文献

- [1] R.A. Brooks: Intelligence without Representation, in *MIT Teche Report*, 1988.
- [2] P.S. Churchland: *Neurophilosophy: Toward a unified science of the mind-brain*, MIT Press, 1986.
- [3] 藤岡喜愛: イメージと人間、日本放送出版協会、1974.
- [4] M. Johnson: *The Body in the Mind*, The University of Chicago Press, 1987 (心の内の身体、菅野他訳、紀伊国屋書店、1991)
- [5] S. Harnad: The Symbol Grounding Problem, *Physica D*, Vol.42, pp.335-346,1990.
- [6] N. Humphrey: *The Inner Eye*, Faber and Faber, 1986. (内なる目、垂水訳、紀伊国屋書店、1993)
- [7] G. Lakoff and M. Johnson: *Metaphors We Live By*, University of Chicago Press, 1980. (レトリックと人生、渡部他訳、大修館書店、1986)
- [8] 信原幸弘: 心の現代哲学、勁草書房、1999.
- [9] CA. Porro, MP. Francescato, V. Cettolo, ME. Diamond, P. Baraldi, C. Zuiani, M. Bazzocchi, and PE. di Prampero: Primary motor and sensory cortex activation during motor performance and motor imagery: A functional magnetic resonance imaging study, *Journal of Neuroscience* 16(23), pp.7688-98, 1996.
- [10] RWC 情報統合ワークショップ'95, 1995.
- [11] F.D. Saussure: *Cours de linguistique générale* (一般言語学講義、小林英夫訳、岩波書店、1972).
- [12] J.R. Taylor: *Linguistic Categorization*, Oxford University Press, 1995 (認知言語学のための14章、辻幸夫訳、紀伊国屋書店、1996) .
- [13] 月本 洋: メタファに基づく人工知能、日本ソフトウェア学会チュートリアル「メタファとインタラクティブソフトウェア」、1999.
- [14] Hiroshi Tsukimoto: Embodied Ontologies: Ontologies for Real Agents, The First Asia-Pacific Conference on Intelligent Agent Technology, pp. 162-171, 1999.
- [15] Hiroshi Tsukimoto: Extracting Rules from Trained Neural Networks, *IEEE Transactions on Neural Networks*, Vo.11, No.2, pp.377-389, 2000.
- [16] Hiroshi Tsukimoto: Symbol pattern integration using multilinear functions, *Deep Fusion of Computational and Symbolic Processing*, T. Furuhashi, S.Tano, and H.A.Jacobsen eds., Springer-Verlag, 2000. (To appear)
- [17] 月本 洋: 記号的人工知能の限界、足立自朗、渡辺恒夫、石川幹人、月本洋編: 心とは何か-心理学と諸科学の対話、北大路書房、2000. (出版予定)
- [18] L. Wittgenstein: *Philosophische Untersuchungen* 1953. (藤本訳、哲学探究(ウイトゲンシュタイン全集8) 大修館書店、1968)