1O5-OS-22b-1

情報理論から意味の理論へ

From Information Theory to a Theory of Sense-making

日高　昇平 *1
Shohei Hidaka

*1北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

A long-standing dream in research on artificial intelligence (AI) is to build a *strong* AI, which *understands* and processes the input, unlike a *weak* AI which just processes it as programmed. Toward realization of this dream, we need a mathematical formulation on what understanding is. In the present study, starting off by revisiting Shannon's mathematical theory of communication, I argue that it is a model of information transmission but not that of information understanding, because of its common codebook shared by the sender and receiver. I outline the steps to build a model of information understanding, by seeking possibilities of decoding without the shared codebook. Given the model of information understanding, I discuss its relationship to other known problems in AI research, such as the symbol grounding problem and frame problem.

1.　理解：強い AI と弱い AI をわかつもの

1.1　中国語の部屋

人工知能 (以下 AI) の研究における哲学は、大きく二つに分けられる。一つの流儀は、ある特定の機構をもった機械あるいは計算プログラムを作り、人あるいはその他動物の行動や心の働きとその機械を比較し、人の心の働きと同種の機構を実現する一例として機械をみなす方法論である。もう一つの流儀は、人を含め「知能」をもつものは、特定の機能の実現によって定義できるとし、逆にそうした機能実現こそが「知能」の本質であると考える方法論である。前者は対象となる認知過程の機能実現可能性の証明としての AI、後者は機能が実現できる限り人と機械を区別せず、高い機能をもった AI により人と同等の「知能」を人工的に実現できると考える。Searle [Searle 80] は前者を弱い AI、後者を強い AI と呼び、強い AI の思想に懐疑的な批判を行った。

Searle の強い AI に対する批判は、「中国語の部屋」と呼ばれる思考実験に集約される。中国語の部屋の思考実験では、中国語を知らない Searle 自身が、小窓からスクリプトの出し入れが可能である部屋に入り、彼にとっては理解できない中国語のスクリプトおよび問いを受け取り、Searle の母国語である英語で教示されるルールにしたがって中国語の答えを返す。もし英語で書かれたルールが十分に洗練されていれば、Searle 自身が中国語をまったく理解できないとしても、部屋の外に出てくる中国語の答えは、外の観察者からは中国語話者と違いが判別できないであろう。したがって、この中国語の部屋における、英語の教示 (プログラム)、そしてそれを実現する Searle 自身 (機械・AI) は中国語を理解できずに、しかし Turing test [Turing 50]*1 に合格するはずである。多くの AI プログラムは、このように書かれたプログラムどおりに入・出力を対応付けるものであるから、そこに「理解」は必要でなく、「理解」をする人の知能と同等には扱えない (つまり、強い AI になりえない) とする。一方、「英語の部屋」に Searle が入り、外から与えられた教示なしに自身の英語の理解に基づいて英語のスクリプトに回答することができるのなら、それは知能と呼んでよいとする。

以上の Searle の中国語の部屋論法から察せられるとおり、強い AI と弱い AI をわける鍵となるのは、「意味」あるいは「理解」である。Searle の批判は、単に特定の機能あるいは入出力の対応関係の実現だけにとどまらない「理解」なるものが知能の本質であるという点にある。しかし、Searle 自身は「理解」が何であるか明確な定義を与えておらず、直感的な事例の列挙による説明にとどまっている。上記の中国語の部屋論法において、Searle 自身が部屋に入る理由は、Searle が自身の中国語および英語の理解について明確に宣言できるからである *2。一方、他者が中国語の部屋に入った場合、Searle がその部屋が知能をもつか否か (部屋の主が外からの教示を必要とするか否か) を明確に判別する手段はない。つまり、中国語の部屋論法は知能の定義を理解に置き換えており、理解の定義は「私が胸を張って理解できると言えるか」という Searle 自身の主観的な判断に委ねられている。それでも中国語の部屋論法が広く受け入れられたとすれば、それは論理ではなくむしろ人の共感あるいは「理解」に依拠していると考えられる。

1.2　マネ碁 AI と情報の伝達

本稿において私は、「理解」とは何かという問いに対し、直観や共感に訴えることなく定式化を与えたい。まずは直観に訴えるために作られた中国語の部屋論法を、その本質を失わずより簡潔にしよう。

次のような思考実験を考える：
チェスや将棋ではトップレベルの人間プレイヤーと同等あるいはそれ以上の強さをもつ AI がすでに存在する。一方、まだ本稿執筆時点で囲碁では人のほうがまだ強い *3。そこで囲碁において、名人と同等以上の強さをもつ AI プログラムが作れたなら囲碁版の Turing test に合格と認めよう。それには、以下

連絡先: 日高　昇平，北陸先端科学技術大学院大学，石川県能美市旭台 1-1

*1　言語的な応答能力によって知能を定義する試み.

*2　Searle によれば、Searle が英語の教示をすべて「理解」し (understand everything)、中国語のスクリプトを全く「理解」しない (understand nothing)。さらに、英語教示に従う (program として) ことと、部屋として「理解」することには関係がないとする。これが明確に宣言できるのは Searle のみであろう。

*3　改稿時点で、DeepMind の開発した Alpha Go[Silver 16] がトップレベルの棋士である李セドルを公開対局で破った。しかし、この思考実験に本質的な影響はない。

のようなごく簡単なプログラムを作ればよい。碁盤 A では名人と AI プレイヤーが、碁盤 B では名人と肩を並べる挑戦者と AI プレイヤーが対局する。碁盤 A では AI プレイヤーが後手番となり、碁盤 B では AI プレイヤーが先手番とする。碁盤 A で名人の第一手をまち、それと同じの手を AI プレイヤーは碁盤 B で打つ。碁盤 B での挑戦者の第二手をまち、それと同じてを AI プレイヤーは碁盤 A で打つ。以下これを繰り返すと、AI プレイヤーは二局とも引き分けあるいは、一勝一敗の結果を得るはずである。従って、この AI プレイヤーは名人と同格であり、囲碁版の Turing test を合格する。

このマネ碁をモチーフにした思考実験は、中国語の部屋の外にいる中国語話者を名人、また Searle に英語の教示を与える翻訳者を挑戦者とすると、本質的に中国語の部屋論法と同型である。しかし、Searle の「理解」の直観的な判断を待たずに、このマネ碁 AI が囲碁を理解する必要がないことは明白であろう [*4]。マネ碁論法における囲碁 AI は、囲碁に関する知識や計算を必要とせず、碁盤 A と碁盤 B の間の通信あるいは情報の伝達しているに過ぎない。これは中国語の部屋論法において、Searle が中国語話者と翻訳者の問答を媒介をしているのと同じである。従って、中国語の部屋論法は、本質的に情報の伝達と理解の違いを指摘しようとしているのである。重ねて言えば、弱い AI とは AI の設計者と AI が機能する対象者の間の情報伝達を行う媒介といえる。

1.3 情報の 3 問題

マネ碁論法から、情報の伝達と理解の違いに本質があることが明らかになった。情報の伝達に関しては、すでに Shannon によって定式化された情報理論があり、伝達可能性を与える通信容量定理により明確な定義が存在する。これに加えて、情報理論は Searle の指摘より 30 年以上前に定式化されているが、その定式化には情報の伝達と理解の違いを表現するための洞察が含まれている。なぜなら、情報の理解という問題の定式化を巧妙に避けて、情報の伝達のみに焦点絞って数学的に表現したのが情報理論といえるからである。実際、Shannon と Weaver は、明確に情報の伝達・理解・因果性の 3 つを区別し、情報理論によって伝達の問題だけを解くことを宣言している [Shannon 49]。

本稿の 2. では、まず情報理論を概観し Shannon の情報伝達モデルを再考し、情報の理解を定式化の基礎とする。Shannon の情報伝達モデルの再考から得られる洞察は、伝達される符号列と元の信号との対応表である符号器が、情報の送り手と受け手の間で**事前に共有されているという仮定**が、この定式化において重要な役割を果たしているという点である。この仮定により、最も重要な伝えられた符号列の「解釈」あるいは「理解」の問題に悩まされることなく符号列が誤りなく伝達可能な条件にのみ焦点を絞って数学的な議論を展開できる。本稿では、この共有符号器の仮定を緩和することで、情報の理解の数学的な定式化を模索していく。

2. 情報の伝達モデル再考

そもそも情報とは何だろうか。Shannon の通信モデル (1) では、ある事象 (信号) が確率的に発生し、それに予め割り当てられた符号を、通信路 (Channel) を通じて逐次送り、受け取った出力符号を基に、元の事象を推定する。事象に符号を割り当てる写像を符号器 (Encoder)、符合の集合を Codebook、符号から事象への写像を復号器 (Decoder) と呼ぶ。Shannon の通信モデルでは、送り手を受け手は Codebook あるいは符号器を共有し、受け手はこの共有符号器により Decoder を構成する。

送り手が 2^r 通りの事象のそれぞれに n 文字の符号列を割り当て (事象の集合 $E = \{1, \ldots, 2^r\}$ に対し符号器 $e : E \mapsto \{0,1\}^n$ を用いる)、そのうち 1 つの符号列 $c \in \{0,1\}^n$ を通信路を用いて送る。受け手は出力列 $\hat{c} \in \{0,1\}^n$ に対し符号器 $d : \{0,1\}^n \mapsto E$ を使い元の事象を推論する。すべての信号に対して、元の事象と推論された事象が一致、つまり $d(\hat{c}) = e^{-1}(c)$ であるとき [*5]、送り手と受け手の間で通信が可能であると定義する。このとき、Shannon の通信容量定理は以下を主張する。

Theorem 1. $r < \max_{P(X)} I(X;Y)$ を満たす限り、十分に大きな n に対して $d(\hat{c}) \neq e^{-1}(c)$ である確率を任意に小さくできる符号器と復号器が存在する。また逆に $r \geq \max_{P(X)} I(X;Y)$ なる符号を送る場合、どんな符号器と復号器を使っても通信には無視できない誤り確率が存在する。

この定理の厳密な証明は他に譲るとして、ここではこの定理の前提として通信する送り手・受け手は符号器 (Codebook) を共有知識として持つ必要がある事を強調したい。つまり、受け手が出力符号から元信号を推論するために、送り手の使った符合器を利用する。具体的に、Shannon の証明では、ランダム符号と呼ばれるランダムに発生させた文字列を符合として利用する符号器と、大数の弱法則を利用した復号器を使い、通信容量の限界を満たす

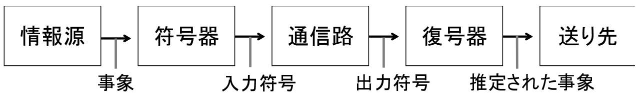

図 1: 情報伝達モデル.

3. 情報理解の定式化

3.1 共有符号器のない通信

前述の Shannon の情報伝達モデルにおいて、誤りを最小化しながら復号するためには、受け手の持つ符号器 C が肝要である。しかし、符号器 C 自体も単なる符号列と同等であることを考慮すれば、符号器 C なる符号がいかに送り手と受け手の間で共有されたのかという素朴な疑問が残る。符号器 C は、Shannon モデルで通信をする「前」に，送り手から受け手に (あるいは逆に)” 伝達” しなければならないはずである。Shannon モデルでは送り手と受け手の C は同等であるので、C は誤りのない通信路を通って送り手と受け手の間で共有されたことになる。しかし、これは奇妙である。もし符号器を誤りなしに C を伝達する手段があるのなら、Shannon モデルで想定するように誤りのある通信路を使う必要がない。

この奇妙な前提を理解するには、Shannon モデルが主に通信機器あるいはアルゴリズムを設計する電気工学的な立場から利用されてきた背景を考慮する必要がある。この立場では、送り手・受け手は、設計者により設計時に一度だけ同一の符号器を与えられており、それ以降、「設計者」という誤りのない通信路を使わず独立に通信することが想定されている。つまり、ここには、通信主体とその設計者という分離がある。通信主体

*4 しかも、全く同じ AI でチェスや将棋などを含むいかなる 2 人ゼロ和完全情報ゲームを用いるすべての Turing test にも合格できる。

*5 厳密には、$d(\hat{c}) \neq e^{-1}(c)$ である確率が任意に小さくできるとき

は与えられた符号器を使って決められた手順で通信を行うのみであり、符号の解釈 (信号) は送り手・受け手の設計者によってなされる。言い換えれば、Shannon の情報伝達モデルは巧妙に解釈の問題を避け、通信容量および符号器・復号器の設計可能性の問題を切り出した点で秀逸であるといえるだろう。

一方、こうした工学的立場を離れれば、“通信” は必ずしも完全に互いに符号器を共有して行われるものではない。たとえば、我々の最も身近な符号である言語は、元信号である我々意図および意味を反映するべく設計されているはずであるが、互いに完全な符号器を共有して会話をしていることは稀なはずである。あるいは、そもそも我々が言葉を交わす目的の一つは、互いの符号器を推定しあうことであるとも言える。ここで、符号として言葉あるいは符号器として言語は、音声言語だけでなく、ジェスチャー、スクリプトの交換などのあらゆる知覚可能な記号を広く指しており、当然ながら中国語の部屋もその一種である。また、人と人の通信だけでなく、神経細胞間や、その他生体の間の相互作用もやはりある種の通信とみなせる。こうした生体間の通信では、一定程度遺伝的に固定された性質があり、共有知識とみなせる可能性はあるものの、必ずしも遺伝的に与えられた符号器だけではその挙動すべてを説明できない。もし、遺伝的に固定された符号器だけに基づいて通信をするならば、その範囲を超えた新たな事象の表現が困難であり、柔軟性に乏しく適応的ではないだろう。

従って、完全なる符号器の共有を前提とせず、しかし符号列から、符号器の性質そのものを推測あるいは学習するにはどうするべきかという問いがたつ。もし受け手が符号器を知らず、符号列だけから符号器を推定できる場合、この符号器の推定過程を「理解」(understanding) として定義できるのではないだろうか。

3.2 メタ復号問題

これまでの議論で、符号器が陽に与えられていない場合に、それを推定する問題を考え、その基本要素に対応する概念を割り当てた。これは、符号器そのものを推定しなければならないため、言い換えればそれに対応する復号器を動的に構築する過程である。従って、我々はこれをメタ復号問題 (復号器の復号化) と呼び、いかにこのメタ復号化が可能であるかを問いたい。一つの素朴な案は、「メタ符号器」の共有を前提として、同じ Shannon の通信モデルを使い、符号器を伝えるというものだろう。しかし、明らかにこの論法は「ではメタ符号器の通信は ⋯ 」という無限後退に陥る。従って、単純に Shannon の通信モデルを入れ子にするだけではこの問題は解決できない [*6]。一方、いかに生体など自然物であっても、生まれもって何も共有していないわけではなく、物理的・生物的になんらかの制約をもっている。従って、符号器そのものではないとして、どの程度のゆるい条件を共有していれば、後天的・適応的に符号器を推定可能であるかを問うのが妥当だろう。よって戦略的にいくつかの問いに分けて問うべきである。

1. どのような条件・制約でメタ復号可能であるか
2. 符号器のどのような性質であればメタ復号可能であるか
3. 既知の符号器から、未知なる符号器をいかに推定するか

本稿では、これらすべての問題を取り上げるのではなく、これらすべてに共通する最も基本的な問題として、まずそもそもメタ復号化の可能性あるいはその候補となるモデル・定式化について論じたい。

*6 入れ子になった Shannon の通信モデルでは、メタモデルの通信容量が入れ子モデルの通信容量を決める。従って、同じメタ符号器問題へと帰着する。

3.3 メタ符号化

どの程度の条件がメタレベルの情報を伝えるために必要であるか考察するために、まずごく簡単なケースを考え、その条件を緩和していく。

まずメタ復号問題において、符号器は未知であるものの、仮に答えである事象が与えられたとしよう。この場合、この事象に対応する出力符号を繰り返し観察することで、この出力符号の集合と事象を対応づけることができる。では、次に、どの事象に対応する出力符号であるか不明であるが、同じ事象に対する出力符号が繰り返し観察できるとする。この場合、この出力符号をどの事象 (整数 $i = 1, \ldots, k$) に対応付けるかわからないが、仮にこれを j としておけば、j と j 以外の事象が異なる出力符号に対応する限りこの j の復号は可能である。これは一見自明なことのように思えるが、つまり、事象の同一性さえわかれば、整数の交換に関する対称性を除いて、元の事象と復号した事象の間に 1 対 1 の対応 (i.e., a bijective map $\{1, \ldots, k\} \mapsto \{1, \ldots, k\}$) をつけることができる。情報の本質が 1-1 対応付けであることを考慮すれば、交換に関する対称性は本質的ではなく、こうした写像が構築可能ならば、メタ復号が達成されたと定義する。また、この議論から明らかなように、メタ復号は事象の同一性を伝えることによって達成可能である。

では事象の同一性をどのように伝達すればよいのだろうか。ここでも素朴に思いつく戦略から考察しよう。まず思いつくのは、そもそも 1 つの事象しか送らない場合である。あるいは、受け手の立場では、すべての出力符号は 1 つの事象を表していると解釈する場合である。この場合、復号可能であるのは自明であるが、情報量は 0 となる。次に、2 つの事象 1, 2 に関する符号をランダムに選び送る場合を考える。この場合、どの符号が事象 1 あるいは事象 2 に対応するか、受け手にはわからない。そのため誤りが任意に小さい対応付けができず、やはり情報量は 0 となる。

こうした素朴な議論から、一見、事象の同一性そのものを意味ある形で送るのは不可能に見える。これに対し、事象 1 のみを送る戦略と事象 1 と 2 を送る戦略の二つが識別可能であるか考えよう。前者では、1 つの出力符号の集合が観測されるのに対し、後者では 2 つの出力符号の集合が観測される。従って、後者のほうが出力符号の確率測度のエントロピーは高く、2 つの戦略は潜在的に区別可能である。しかし、2 つの戦略を確率的に混合して事象 1 または 2 を送る場合、それは戦略 2 と同じになってしまい、やはり受け手の側から区別がつかないように見える。

そこで、条件を緩和して、送り手も受け手も以下の事前知識を共有すると仮定する : 送り手が戦略 1(事象 1 だけを送る) と戦略 2(事象 1 と事象 2) のいずれかをある時間 t にとると、時間 $t+1$ に同じ戦略を繰り返す確率は 1/2 より高いもしくは低い。この場合、連続する戦略系列が同じ (もしくは異なる) 戦略である確率が高く、戦略 1 と戦略 2 とが区別できる可能性が出てくる。この仮定は、送り手のとる戦略の時間的依存性を意味する。こうした時間的依存性の下、上記のように異なる数の符号を混合して符号を送ることを、仮に一つのメタ符号化としておこう。

3.4 メタ復号器

次の問題は、仮にこうしたシステムの時間依存性があるとして、時間変化する未知なる戦略の系列をいかに復号するか、

である。すでに述べたように、異なる数の符号を送る戦略をとっている一定期間は、異なるエントロピーを示すはずである。従って、未知なるシステムについて、かつそのシステムが変化するより高い時間解像度で、各区間のエントロピーを求めることができればよい。送り手は、符号の内容はさておき、時間的に送る事象の数を変化させ、これによりメタレベルの符号化とする。とりうる事象の範囲および数を動的に変化させることによる符号化をメタ符号化と呼ぼう。これに対するメタ復号化の可能性を問うのが、時間依存性のあるシステムのメタ復号問題である。

この議論における、システムのとりうる事象の数は、一般に(ある種の) 次元によって与えることができる [*7]。つまり、メタ復号問題のコアとなる問題は、動的な次元推定問題である。上記の議論のとおり、メタ復号問題を動的次元推定問題として捉えるために必要な条件は、ある種のシステムの時間的な一貫性(局所的に同一) および長期的にはシステムが変化すること (変化しないシステムにはメタレベルの情報がない) である。こうした、複数の「次元」が混在しうる集合あるいは測度を扱うのに十分な性質を備えた量として、点次元 (pointwise dimension あるいは local dimension) がある。点次元は、ルベーグ測度 0 の集合の「大きさ」を扱うのに用いられる Hausdorff 次元や、容量次元 (Capacity; Box-counting dimension) などの複数の種類の次元と関係があり、点次元の分布は、他の次元の上限および下限を与える [Cutler 93]。この意味で、点次元は他の次元より次元の異なる測度の混合に対して、より詳細な情報を持っている。

前節で導入したメタ符号は、符号の数に対応して点次元が時間的に切り替わる系列であるとみなせる。従って、このメタ符号に対応するメタ復号器を構築するには、ある十分に長い有限の符号から、その符号の点次元の変化を任意の精度で推定できればよい。

3.5 まとめ：メタ符号化/復号化

以上の議論から、共有符号器を前提とせず、私の考えうる最も弱い制約の下で情報の伝達可能性のある一つの手順は以下のとおりである。

1. (メタ符号) 送り手は事象 $i \in \{1, \dots, k\}$ のそれぞれに対し、$(i+1)$ 個の記号からなる符合を生成する。あるいは、i 次元乱歩列を符号とする。
2. (誤りのない未知の変換が行われる通信路) 入力列に対する未知の bi-Lipschitz 変換を出力列とする。
3. (メタ復号) 受け手は出力符号に対して点次元推定を行い、同一の次元を持つ列が同一の事象を表すとみなす。

最尤法など適当な推定法を用いることで、任意に符号長を大きくすれば、復号誤りの確率を任意に小さくできるため、受け手は点次元の異なる k 種類の信号を受け取ることが可能である。この通信において、受け手は送り手がどのような空間を利用してデータを符号化するか事前に知る必要はなく、また通信路としても任意の未知なる bi-Lipschitz 変換を仮定している。この条件でも、事象の番号の交換に対する対称性を除き、点次元推定により事象の同一性を復号することが可能である。

*7 Shannon entropy, Lyapnov exponent, Hausdorff dimension の関連性については [Young 82] を参照。

4. 総合考察

Searle の定義する強い AI と弱い AI をわかつ壁を乗り越えることは、人工知能研究者の長年の夢である。あるいは、それは Shannon が情報伝達モデルとして切り離すことであきらめた情報の理解の数学的な定式化でもある。

本稿では、中国語の部屋論法と情報の伝達モデルの本質的な同型性を指摘し、Shannon が情報伝達モデルを考察した。この情報伝達モデルは、送り手を受け手が同一の符号器を共有できるとの仮定から、情報の伝達と理解を切り分けた巧妙な通信モデルである。この考察から、情報の理解の定式化として、共有されていない符号器そのものを受け手が復号するというメタ復号問題を考えた。これは、Shannon の定式化の範疇を超えており、単なる伝達のみならず、元の事象と符号の対応関係を再構築することである。本稿ではこのメタ復号を情報の「理解」のモデルとして提案する。

最後に、本稿で提案する情報理解のモデルに関連するいくつかの問題に触れて、その関連性を指摘しておきたい。本稿では、受け手が符号器を知らず、符号列だけから符号器を推定できる可能性を議論し、この符号器の推定過程を「理解」と定義した。この問いは情報伝達モデルを基礎においた記号接地問題 (Symbol Grounding Problem)[Harnad 90] の定式化とも言えるだろう。

これに対し、本稿の定式化では、対応付けの可能性に関して二つの条件に分ける。一つは符号器が与えられ、送り手と受け手の間での対応づけが容易な場合 (情報の伝達モデル) であり、もう一つは符号器自体も推定の対象となる場合 (情報の理解モデル) である。いずれの場合も、対応がつく対象あるいは対応付けそのものはいわゆる「記号」の基礎的な単位として扱える。

参考文献

[Cutler 93] Cutler, C. D.: *A Review of the Theory and Estimation of Fractal Dimension*, Vol. 1, pp. 1–107, World Scientific (1993)

[Harnad 90] Harnad, S.: The symbol grounding problem, *Physica D: Nonlinear Phenomena*, Vol. 42, No. 1, pp. 335–346 (1990)

[Searle 80] Searle, J. R.: Minds, brains, and programs, *Behavioral and brain sciences*, Vol. 3, No. 03, pp. 417–424 (1980)

[Shannon 49] Shannon, C. E. and Weaver, W.: *The mathematical theory of communication*, University of Illinois press (1949)

[Silver 16] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Driessche, van den G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al.: Mastering the game of Go with deep neural networks and tree search, *Nature*, Vol. 529, No. 7587, pp. 484–489 (2016)

[Turing 50] Turing, A. M.: Computing machinery and intelligence, *Mind*, pp. 433–460 (1950)

[Young 82] Young, L.-S.: Dimension, entropy and Lyapunov exponents, *Ergodic Theory and Dynamical Systems*, Vol. 2, pp. 109–124 (1982)