

# 汎用知能の知識記述長最小化原理仮説の提案

## Proposing minimum knowledge description length principle for general-purpose AI

山川 宏<sup>\*1\*2</sup>  
Hiroshi Yamakawa

<sup>\*1</sup> NPO 法人全脳アーキテクチャ・イニシアティブ      <sup>\*2</sup> (株)ドワンゴドワンゴ人工知能研究所  
The Whole Brain Architecture Initiative, a specified non-profit organization      Dwango Artificial Intelligence Laboratory, Dwango Ltd.

Artificial general intelligence (AGI) is evaluated essentially by the integration of performances for multiple tasks. When tasks for these are fixed, big-switch type AIs, which switch multiple narrow AIs for every task, often show better performance. Therefore, research cannot reach a uniform goal for the AGI to enhance ability to cope with unpredicted situations. This study first defines a standard description length of knowledge. Then, from the point of view of the task domain, the researchers propose the knowledge structure model with a knowledge hierarchy that leads from the concrete to the abstract, citing a scenario study that used this hierarchy. When an ideal AGI and a big-switch type AI have the same level of performance for fixed task sets, the knowledge description length of an ideal AGI is shorter. This reduction is due to the maximal reusability of individual knowledge in an ideal AGI, and this is the foundation of its versatility. Here we call this nature ‘Minimum knowledge description length principle for general-purpose intelligence’.

### 1. はじめに

現状の実用的な人工知能(AI)は特定のタスク領域に特化しており、これは特化型 AI や Narrow AI 等と呼ばれる。これら AI を評価するには特定のタスクに対する性能を評価尺度とすれば十分である。対して、多様なタスクを実行可能な汎用人工知能(AGI)に対しては、どのように評価を行うべきだろうか。チューリングテストや、[市瀬 2015]が最近検討した心理学の知的能力の評価を参考とする方法は、AI が人のように振る舞うことを前提とした、人間中心(anthropocentric)からの評価アプローチである。

しかしここでは、[Legg 2007]において導入された、普遍知能尺度(universal intelligence measure)の立場から検討を行う。この尺度で測定したい知能は、彼らが“intelligence measures an agent’s ability to achieve goals in a wide range of environments”と述べたように、多様な環境(もしくはタスク)における性能を統合した評価尺度で、人間の知性とは無関係に知能の汎用性の評価を目指す。これは非人間中心的(non-anthropocentric)なアプローチとも言える。

環境が多様であるとはいえ任意の環境に対する普遍知能尺度を測定することは難しい。そこで [Hernandez-Orallo 2010]らは生物学的に現実的な環境(もしくはタスク)のクラスとして、{空間、物体、観測と行動、報酬}という概念セットとして普遍知能テスト(universal intelligence test)を導入した。

同様な設定で[Legg 2007]は方策  $\pi$  をもつ効用ベースの知能エージェントについての普遍知能尺度  $I(\pi)$  を定式化している。

$$I(\pi) := \sum_{\mu \in E} 2^{-K(\mu)} V_{\mu}^{\pi} \quad (1)$$

ここでは、環境の全体集合  $E$  から一つの環境  $\mu$  をとりだし、その場合における方策  $\pi$  をもつエージェントの予測未来報酬の総和を  $V_{\mu}^{\pi} := \mathbb{E}(\sum_{i=1}^{\infty} R_i)$  とする。そして環境の複雑さ  $K(\mu)$  に応じて少なめに積算した値を普遍知能尺度としている。ここで  $R$  は方策  $\pi$  をもつエージェントが時刻  $t$  において環境から獲得する報

酬である。そして  $K(\cdot)$  はコルモゴロフ複雑性である。

[Legg 2011]らは、大量の環境を生成的にサンプリングすることでこの評価値の計算を試みた。しかしながら、ロボットやそのシミュレーションなどを用いた現実的な AGI 開発の環境においては、膨大に多様な環境を生成して評価することは難しい。

このため有限個のタスクを用いた評価では、タスクごとに作りこみを行った Big switch 文型の AI の性能が高くなりやすい。すると事前に想定されないタスクに対する性能を高めるという AGI に本来期待される性質を目指した目的から乖離しがちである。そこで、有限のタスクセットに対しても汎用性の高い知能を良いものとして評価できる、修正した普遍知能尺度があれば、AGI 研究を促進できるだろう。

こうした背景から本稿では次の仮説を提案し、その理論的な検討を進めてゆく。

#### 仮説：汎用知能の知識記述長最小化原理

有限個のタスクに対して同レベルの実行性能をもつ複数の知能システムの中では、最小の知識記述長をもつシステムが最大の汎用性を持つ。

#### Hypothesis : Minimum knowledge description length principle for general-purpose intelligence

Among intelligent systems that perform same level for a finite number of tasks, a system with minimum knowledge description length has best versatile ability.

本仮説の提案に至った、以下2つの理由がある。一つ目に、知識にはそれ相応の記述量が必要であり、基本的には記述量が大きなプログラム程良い性能を得られる。しかし Big switch 型 AI の記述長は急増するため、汎用性の高いプログラムは、相対的に小さな記述長となると考えた。二つ目には、「ある事柄を説明するためには、必要以上に多くを仮定するべきでない」というオッカムの剃刀の指針に基づけば、予測性能を高めるためにはコンパクトな知識が望ましいと考えた。

## 2. 知識とその構造

AI の汎用性を議論する準備として、まず知識とその記述長について検討し、さらに適用範囲の観点から知識とそれを組み合わせた知識構造についての仮説を提案する。

### 2.1 知識の内容

知識とは何らかの対象に関する記述であり、データ(事例そのもの)、データを反映したモデルとそのパラメータ、さらにモデルを記述するための表現やアルゴリズムも含む。そして、それらはしばしば互いに書き換え可能である。

たとえば、数字の0と1が交互に並ぶ100文字の数列は、そのままデータとして記述するだけでなく、それを生成するプログラム、その数列を圧縮するプログラムと圧縮結果の対によっても表現できるというように、書き換え可能である。

### 2.2 標準化した知識記述長

知識の記述量による AI システムの性質を議論するためには、知識の量を標準化する必要がある。しかし以上で述べたように、知識は様々な形に変換可能であるし、さらに本質的に同一のアルゴリズムであっても、実装する計算機言語等によって記述の長さは変化する。

そこで、ここでは、知識の標準記述長を以下のように定義する。

**知識記述長：** 知識記述長は、ある知識についてユニークに決定される標準の記述長であり、同一の知識の任意の書き換えを通じて最小に記述される場合の記述長とする。

真の標準知識記述長は、必ずしも正確に求まるとは限らないため、現実的には様々な書き換えを通じて最小となる記述長により近似的に評価することになる。

### 2.3 知識の適用範囲 (適用の前提条件)

知能の本質的な側面は、ある意味で未経験の状況に対して経験に基づいた予測を行う能力である、直接的に経験した事例にのみ問題解決ができてそれ程知的とは言えない。

知識というものが、こうした知能に用いられることを踏まえれば、知識は適用範囲が広く再利用性が高いほど有益である(つまり経験した知識をできるだけ有効に推論に役立てたいのである)。逆にも、本来一体化すべき知識が分断されていると、経験の共有による知識の最適化を促進できないデメリットが生ずる。

各知識の前提条件は、その知識内容を利用しうるかどうかの条件判断を行う知識であり、例えば条件付き確率の条件部に該当する。また前提条件を自動的に抽出するルール抽出技術等も存在する。

### 2.4 一般的な知識構造

議論を進めるために本稿では AI システムがもつ知識の全体としての知識構造をモデル化する。本モデルでは知識は、分解可能な知識の集合体であると仮定する(これは必ずしも自明ではない)。

さらに知識は適用範囲に関わる階層性を持つと仮定する。すると、最も具体的な知識は現実経験する事例である。逆に最も抽象的な知識はシステムが想定する全てのタスクを普くカバーする一般性の高い知識である。

こうした知識構造を図1に示す。図1(a)において、その最上部には、外部から与えられるデータとしての事例を縦棒で示す。

そして知識構造の本体としては、個々の知識が特定のレベルに所属するとする3層の階層構造を仮定している。図中一番上の層となる3層目は、囲碁、医療診断、自動運転などの具体的なタスク領域における知識である。また2層目はより広いゲーム全般や、視覚情報処理全般などといったタスク領域内の常識的な知識である。そして一層目の知識1A はシステムが取り扱う領域全般にわたる常識的な知識(もしくは一般事前知識)である。こうした一般事前知識の例としては、より単純なモデルを選択する傾向(c.f. オッカムの剃刀)といった計算論的な知識や、システムが対処する外界の物理学的原理に由来する制約(例えば、空間や時間の連続性などといった性質)としての知識などが含まれる[山川 2015]。

ここでは、タスク領域に関わる知識構造として典型的な階層性を仮定したが、より一般的には、少なくとも知識間で具体的と抽象的の向きを持つ有向非巡回グラフ(DAG)までは拡張できる。

## 3. 様々な AI の知識構造とその記述長

本節では事前に想定されたタスク(図1中においては、 $\{x_1, x_2, x_5\}$ )においては同等の性能をもつ AGI (a) と Big switch 文型 AI(c)について、事前に想定されないタスクについての能力比較を行う。さらにそれら AI の知識記述長の比較を行う。

### 3.1 理想的な AGI の知識構造 (図 1(a))

前節での知識適用範囲についての議論を踏まえれば、個々の知識の適用範囲が正確かつ最大限に拡大されている方が望ましい。こうした知識で構成される知識構造を備えたシステムが理想的な AGI (汎用性が高い知能)であろう。

図 1(a)において、外界からシステムに対し、異なる領域に属する入力  $\{x_1, x_2, x_3, x_4, x_5\}$  が与えられた場合を考える。たとえば、入力  $x_1$  に対しては知識  $\{3A, 2A, 1A\}$  の階層的知識を利用して高品質に問題解決を行える。同様に  $x_2, x_5$  についても良好な問題解決を行える。対して、入力  $x_3$  については、特化した知識が存在しないために非常に高品質な問題解決は期待できない。しかし、知識 2B と知識 1A を利用することで次善の問題解決は可能である。さらに入力  $x_3$  については利用可能な知識は抽象的な知識 1A のみではあるが、それでも、利用可能な知識が皆無であるよりは優れた問題解決を行える。

以上の説明から明らかなように、個々の知識において適用範囲(再利用性)が可能な限り幅広く設定されていることが AGI システムにとって理想的である。こうしたシステムならば、未経験の状況に対しても、比較的多くの知識を使用して、ある程度は質の高い問題解決が可能であり、これこそが汎用的な知能を発揮した状態である。

次に、この理想的な AGI における知識の記述量を考える。ここで各知識において、知識内容は小文字  $c$  を付して表記し、前提条件は小文字  $p$  を付して表記する。たとえば、図1(a)において知識3Aの内容を3Acで、前提条件は3Apと表記する。

すると理想的な AGI である図1(a)の知識は、 $\{1Ac, 2Ac, 2Ap, 2Bc, 2Bp, 2Cc, 2Cp, 3Ac, 3Ap, 3Bc, 3Bp, 3Cc, 3Cp, 3Dc, 3Dp, 3Ec, 3Ep\}$  である。

以下の議論を単純化するため、図 1(a)における各知識の知識記述長は一定であると仮定し、その長さを1ユニットとする。この例では知識内容が知識の数に相当する9ユニットである。

なお最も抽象的な知識1Aは全ての領域において使用されるため前提条件に基づいた制御の必要はない。よって、前提条件は全部で8ユニットとなる。それゆえ、この理想的な AGI の知識ユニット数は合計で17個である。

### 3.2 特化型 AI の知識構造 (図 1(b))

特化型 AI は特定のタスク領域のみにおいて機能する AI システムである。ここでは図 1(b)に、入力  $x_1$  に対してのみ問題解決できる特化型 AI を図示した。入力  $x_1$  に対しては知識 3Ac, 知識 2Ac, 知識 1Ac が統合された知識を利用することで高品質の問題解決が可能である。ここでは知識 2Ac や知識 1Ac は、他のタスクに使えるような前提条件を持たないため、他のタスクにこの知識を拡大適用できない。そしてしばしば、特化型 AI における3つのユニットにおける知識内容は効率化などのために混合されてしまう。

知識が一旦混ぜ合わされてしまうと、そこから汎用的な知識抽出して適用範囲を拡大することは困難である。このため上記以外の入力  $\{x_2, x_3, x_4, x_5\}$  に対しては問題解決能力を発揮できない。

以上のように入力  $x_1$  に対応する特化型 AI では前提条件の判別を行わず、その判断をするための知識をもたないため、知識内容としての3ユニットだけが備えられている。

### 3.3 Big switch 文型 AI の知識構造 (図 1(c))

多種多様な特化型 AI を束ねて、それらを状況に応じて切り替えて利用する AI は、しばしば Big switch 文型 AI と呼ばれる。

このシステムにおいては、想定されている入力,  $x_1, x_2, x_5$  については、理想的 AGI と同等のパフォーマンスを示す。一方で入力  $x_3$  と  $x_4$  についての問題解決はできない。こうなる理由は、領域ごとに抽象度の異なる知識が混合されているために、そこに潜在的に含まれる一般性の高い知識 (例えば、知識 1A, 2A, 2B, 2C) を、入力  $x_3, x_4$  に対しては利用できないことにある。

次に、この AI における知識記述長に目を向けると、 $\{1Ac \times 5, 2Ac \times 2, 2Ac \times 2, 2Bc \times 2, 3Ac, 3Ap, 3Bc, 3Bp, 3Cc, 3Cp, 3Dc, 3Dp, 3Ec, 3Ep\}$  であり、知識内容が 15 ユニット、前提条件が 5 ユニットで、合計 20 ユニットが設置されている。

この例で知識記述長を理想的 AGI と比較すると、知識内容が  $\{1Ac \times 4, 2Ac, 2Ac, 2Bc\}$  の 7 ユニット増大し、前提条件が  $\{2Ap, 2Bp, 2Cp\}$  の 3 ユニット減少していることから、トータルとして知識記述長は 4 ユニット増大している。

以上のように、本質的に同一の抽象的な知識が一体化されていない場合、知識の重複が発生することで、その記述長が増大する傾向がある。

なお、実際には個々の領域において、レベルの異なる知識が混合されることで、知識記述長は減少しうるが、ここではその影響は考慮していない。

### 3.4 コンパクトな知識記述は AI の汎用性を高めるか

AI の知識構造に関わる以上の議論から、個々の知識の適用範囲を適切に拡大して記述することで再利用性を引き上げられれば、その結果として AI システム全体として汎用性を向上し、同時にシステム全体での知識記述量が減少する。つまりより少ない知識記述長で、同じ程度の性能を得るためには、個々の知識を(再利用可能な形で)有効利用しなければならない。この再利用性の上昇が未知状況での問題解決力を高めるのである。

こうして「汎用知能の知識記述長最小化原理」仮説が支持された。

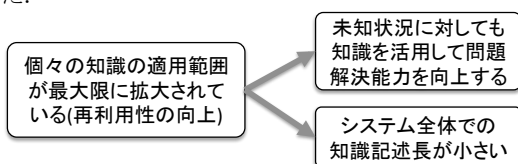
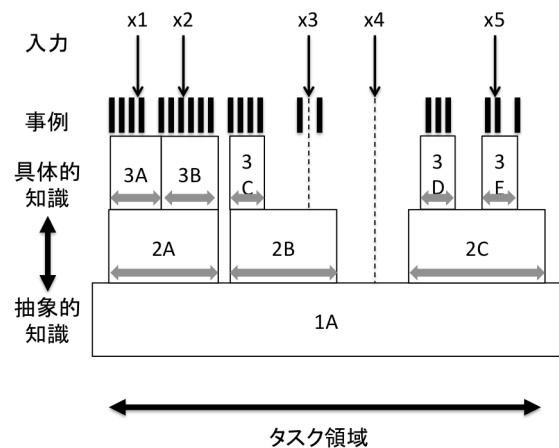
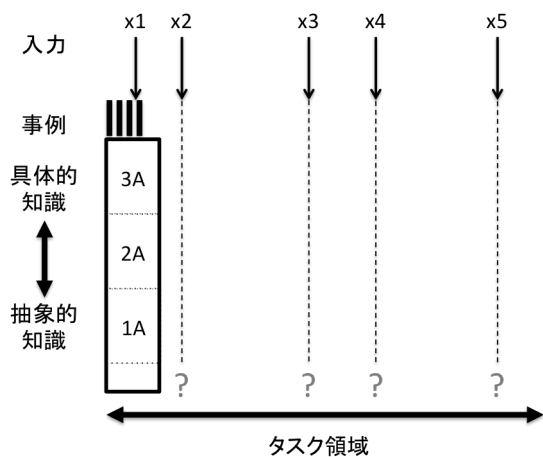


図 2: 知識の記述長と汎用性の関係

(a) 理想的なAGIの知識構造:  
知識内容9ユニット, 前提条件8ユニット, 合計17ユニット



(b) 特化型AIの知識構造:  
知識内容3ユニット, 前提条件0ユニット, 合計3ユニット



(c) Big switch型AIの知識構造:  
知識内容15ユニット, 前提条件5ユニット, 合計20ユニット

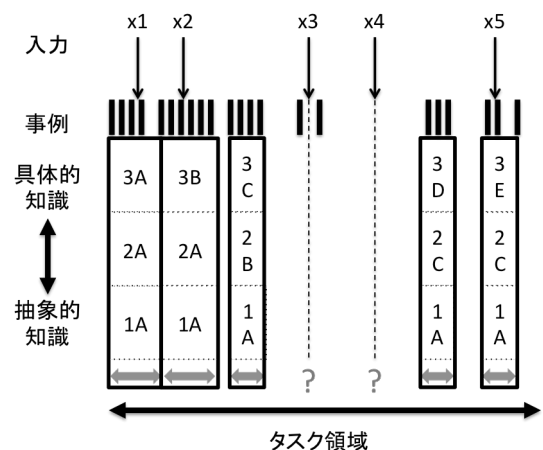


図 1: 人工知能の様々な知識構造

(a) 理想的な AGI, (b) 特化型 AI, (c) Big switch 文型 AI

以上をまとめると、「有限のタスクセットに対する実行性能が同等な人工知能システムを比較した場合、プログラムの記述長がより短い程に、個々の知識の適用範囲がより広く〔再利用性がより高く〕設定されている可能性が高い。それゆえ、十分に知識を準備できないタスク領域においても一定の問題解決能力を発揮しうると期待できる。」

## 4. 議論

### 4.1 知識記述長が汎用性を良い指標とするために

前提として汎用知能の知識記述長最小化原理を認めたとしても、知識の記述長の短さがより汎用性に反映しやすいタスクセットというものがあるだろう。

コンペティション等を通じて AGI 開発を促進する場合に、理想的な AGI と同等の性能を有する Big switch 文型の AI において、より知識記述長が増大しやすいタスクセットを設定できると都合が良い。

先のシナリオにおける検討では、知識を記述するユニット数は以下のように推計できる。

#### 理想的 AGI

- 前提条件ユニット数 = 知識数 - 1
  - 知識内容ユニット数 = 知識数
  - 合計ユニット数 = 知識数 × 2 - 1
- ここで、知識数 > 詳細ドメイン数

#### Big switch 文型 AI

- 前提条件ユニット数 = 詳細ドメイン数
- 知識内容ユニット数 = 詳細ドメイン数 × 階層数
- 合計ユニット数 = 詳細ドメイン数 × (階層数 + 1)

以上より、問題解決に必要な知識構造がより段数の深い階層性を持つタスクセット（環境）を設定した上で、プログラムの記述長を AI の評価尺度に加えれば AGI の研究開発を促進しやすくなると思われる。

### 4.2 プログラミングにおけるリファクタリングとの関連

ソフトウェア工学においては、プログラムの見通しを良くするために、外部からみた動作は変更せずにソースコードの内部構造を整理する。この作業はリファクタリングと呼ばれる。

多数のタスクを行うプログラムにおいては、リファクタリングを行う以前はタスクごとに処理が別れており、Big switch 文型 AI に近い。対してリファクタリングを行うことでメソッド等が整理されるとその再利用性が高まり、ソフトウェア全体として記述長が短くなる。こうした状況が理想的な AGI に近いと考えられる。

逆から見ればリファクタリングはプログラム内における同一の知識を発見する処理である。リファクタリングは実用的な価値が高いにもかかわらず、これは現状で自動化されていない。この事実、自動リファクタリングは AGI の完成にも寄与する残された課題であることを示唆している。

## 5. おわりに

本稿では、人工知能(AI)における知識が簡単な階層モデルであると仮定した。その上で、AI が汎用性を持つことの源泉は、個々の知能の再利用性の高さであり、再利用性の高さが知識の記述量の減少を導くとした。結果として、この仮定のもとでは、システムの知識記述長が短いほど、汎用性が高

まりやすいことが示唆され、つまり「汎用知能の知識記述長最小化原理」が支持された。

本原理の提案は未だ萌芽的であり、その一般性などについての検証が必要である。一方で現実的なプログラムに対する初期的な検証を [坂井 2016]が進めた。ここでは複数のソートアルゴリズムについて、処理速度とプログラム記述長の関係性を分析した。その結果、内部的に複数のサブプログラムの切り替えを行う Big switch 型風のアルゴリズムは処理速度に対して相対的にプログラム長が大きくなることが示された。

今後さらに研究を進め、学習を含む知的システムにおいても、知識記述長による評価を導入したい。それが可能になれば、例えば汎用知能を目指すコンテストや競技、例えば、ロボカップの汎用サービスロボティクスリーグ(GPSR), [GGP], [GVGAI]などのコンテストにおいて、プログラムの短さを競う評価基準を併用することで AGI の研究開発を促進しうらうだろう。

最後に、本検討では考慮しなかったが、AGI においてはデータからの学習する個々の知識についても、適用範囲を拡大したい。関連する技術としては、既にマルチタスク学習や転移学習などの研究もなされている。また AGI における重要な学習技術として、一般事前知識から領域毎に共有される知識を獲得する能力の必要性も指摘されている [山川 2015]。

## 謝辞

本稿の作成にあたり、ドワンゴ人工知能研究所の荒川直哉氏、坂井美帆氏、慶応義塾大学の大澤正彦らとの議論が役立ちました、この場において御礼申し上げます。

## 参考文献

- [市瀬 2015] 市瀬龍太郎, 認知アーキテクチャ評価手法の一考察, 人工知能学会研究会資料, SIG-AGI-001-12, 2015.
- [Legg 2007] Legg, S. and Hutter, M.: Universal intelligence: A definition of machine intelligence. *Minds and Machines*, 17(4):391-444, 2007.
- [Hernandez-Orallo 2010] Hernandez-Orallo, J.: A (hopefully) Non-biased Universal Environment Class for Measuring Intelligence of Biological and Artificial Systems. In E. Baum, M. Hutter, E. Kitzelmann, editor, *Artificial General Intelligence*, 3rd Intl Conf, pages 182-183. Atlantis Press, 2010.
- [Legg 2011] Legg S, Veness J: An Approximation of the Universal Intelligence Measure, *Algorithmic Probability and Friends. Bayesian Prediction and Artificial Intelligence*, Volume 7070 LNCS pp 236-249, 2011.
- [坂井 2016] 坂井美帆, 山川宏, 荒川直哉: 知能の汎用性評価のためのプログラム長の検討:Java ソートプログラムを例に, 人工知能学会研究会資料, SIG-AGI-002-04, 2016.
- [GGPC] General Game Playing, <http://games.stanford.edu/>
- [GVGAI] The General Video Game AI Competition - 2016 <http://www.gvgai.net/>
- [山川 2015] 山川宏: 一般事前知識についての考察, 人工知能学会研究会資料, SIG-AGI-001-11, 2015.