

全脳アーキテクチャ・アプローチ

:

脳全体のアーキテクチャに学び
人間のような汎用人工知能の構築を目指す

山川 宏



「人類と調和した人工知能のある世界」を目指して

概要

人のように多様なタスクを柔軟に獲得する汎用人工知能(AGI)の開発は計算機発明直後からの夢だが、乳幼児が発達段階で獲得する非言語的な知能は設計できなかった。

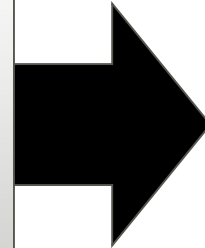
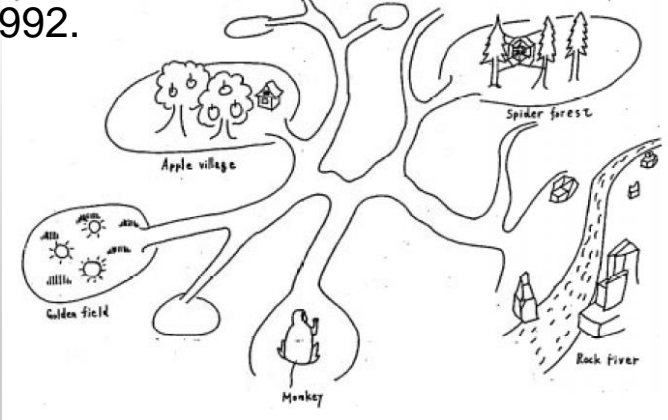
近年の計算リソースの増大に伴い、脳を参考にした深層学習が実用化されて、子供の知能をデータから獲得できるようになった。

そこでこれを明らかになりつつある脳全体像を参考にして組み上げることでAGIの構築を目指すのが全脳アーキテクチャという研究アプローチである。

強化学習 + 深層学習 ~1990年代の私の夢~

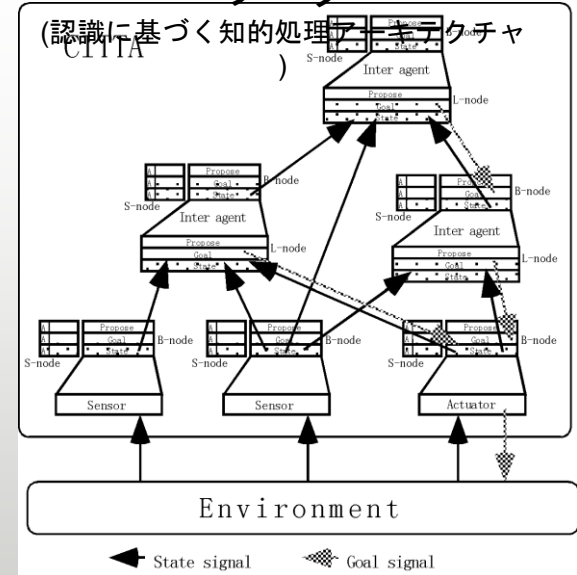
脳に学んで学習だけで動ける汎用知能を作りたいか

つ 山川宏, **強化学習**に基づく知能システム: 価値体系を利用したパターン処理型知能マシンの検討, 東京大学博士論文, 1992.



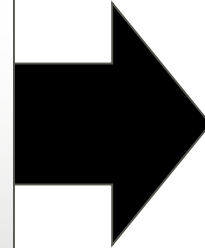
エージェント・ネットワーク

(認識に基づく知的処理アーキテクチャ)

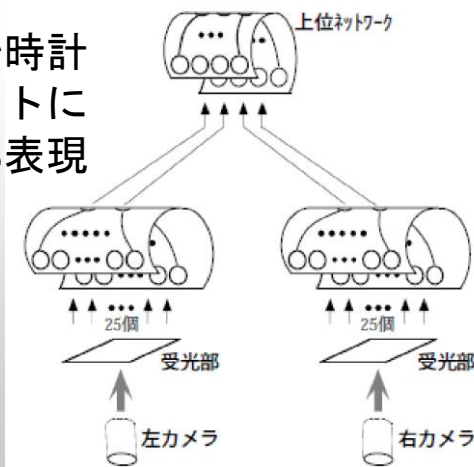


末広尚士、高橋裕信、山川宏, “エージェントネットワークによる手渡しロボットシステムの構築,” 第15回ロボット学会、2 C24、1997.

- 当時は学習システムとしてで結合できなかった
- 2000年台はこうした基礎研究はできなかった.



山川ら, 階層化砂時計型ニューラルネットによる自律的な内部表現獲得, 1995



Autoencoder の階層化も試みた

1.人工知能(AI)とは

人工知能とはなんだろうか？



近代科学社出版 人工知能学会監修 「人工知能とは」

- 人工知能はロボットか？
- 人工知能は機械学習か？
- 人工知能はまだ実現されていないのか？

- AI一般の目標:
 - 産業や科学技術開発等の応用を志向する、何らかの意味でより高度な知的能力の実現
- 人型AIに固有の目標:
 - 人のように振る舞う知能の実現
- 脳型AIに固有の目標:
 - 多くの場合に人AI固有の目標も含む。
 - 脳の機能的な理解に基づき医療貢献、さらにはマインドアップロード等も目指す

知能 ≡ 直接観測できない出来事を
推定する能力

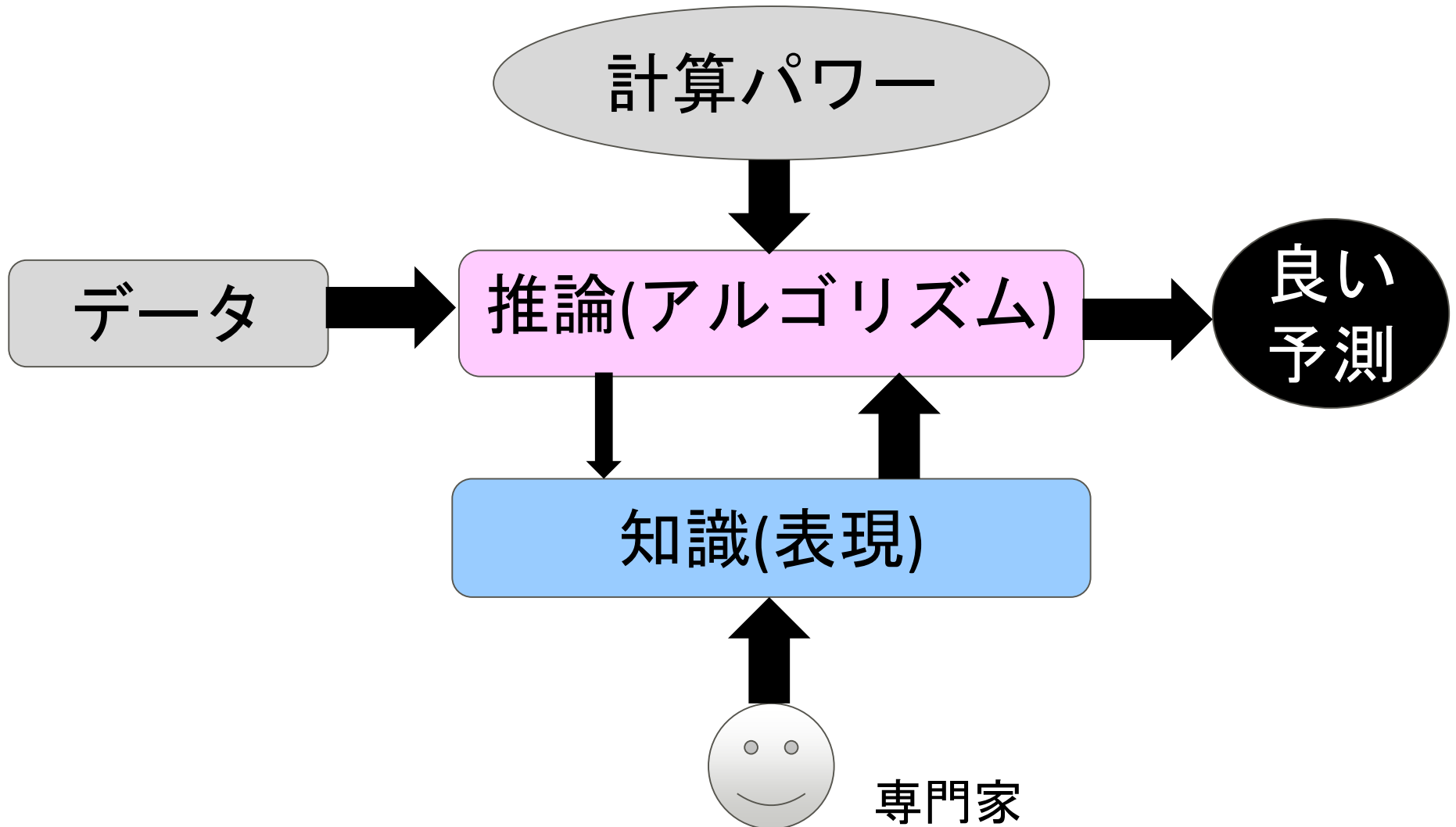
推論(アルゴリズム)
AIで先行して進歩

+

知識表現
知識≠単なる記録の集まり
知識＝情報内の関係性を表現する規則

人工知能に対する一つの見方

知能 ≡ 直接観測できない出来事を予測する能力



知識からみたAIの歴史

1. 知識所与の時代
トイ・プロブレムが対象



ゲームのルール

状態/
質問

推論
機構

行動/
回答

※推論機構の中に知識が
埋め込まれていた。

適用範囲が狭い



2. 知識記述の時代
エキスパート・システム



専門家が知識記述

オント
ロジー

知識
ベース

Webに
よる共
有

状態/
質問

推論
機構

行動/
回答

知識獲得ボトルネック



3. 知識獲得の時代
ディープ・ラーニング



データからの知識獲得

状態/
質問

知識
獲得
推論
機構

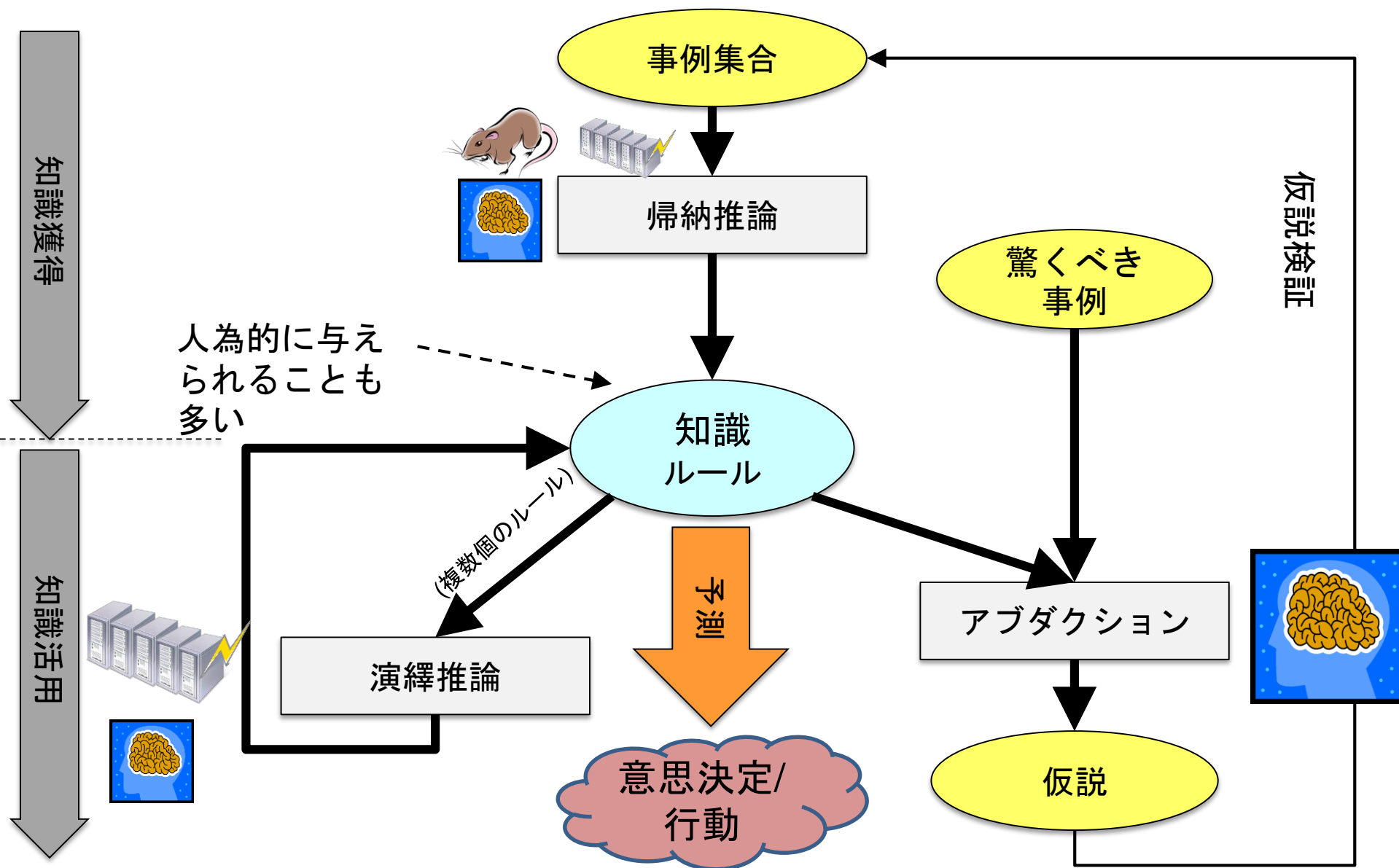
行動/
回答

※以前から存在した**機械学習**
アプローチだが近年性能向
上

抽象知識の学習にメド



推論の種類とその関係



独立行政法人 理化学研究所 脳科学総合研究センター

将棋棋士の直観の脳科学的研究 — 将棋プロジェクト —

将棋棋士「直観思考」を科学する！
将棋を研究対象に、瞬時に最善手を

プロジェクト 紹介

将棋棋士の「直観思考」を科学する！
将棋を研究対象に、瞬時に最善手を指す将棋のプロ棋士の脳に着目し、
人間特有の直観的な思考の解明にチャレンジしています。

将棋と脳科学の関係



詳しくはこちら

プロジェクト 誕生のきっかけ



詳しくはこちら

研究内容の紹介

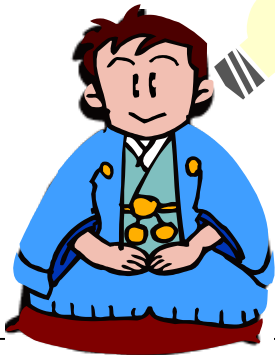


詳しくはこちら

プロ棋士は直観力で計算機を凌駕していた

プロ棋

現在局面（駒配置、戦型、手筋などを含む）と将来局面を予想して有力指し手を選択。



現在の局面

1手後

2手後

数手後

素早い局面理解
(前頭葉/楔前部)

直観

有力指し手の生成
(大脳基底核)

狭い読み

直観:
(帰納推論)

- ・素早い局面理解
- ・有力指し手の生成

狭い読み

局面評価

棋力

広い読み
(演繹推論)

局面評価

コンピュータ将

将



強力な計算能力で広く深く読んだ上で、評価する。

局面評価

2. 深層学習によるAIの変化

Googleのコンピュータ囲碁が初めてプロに勝つ

Nature, 2016/1/27

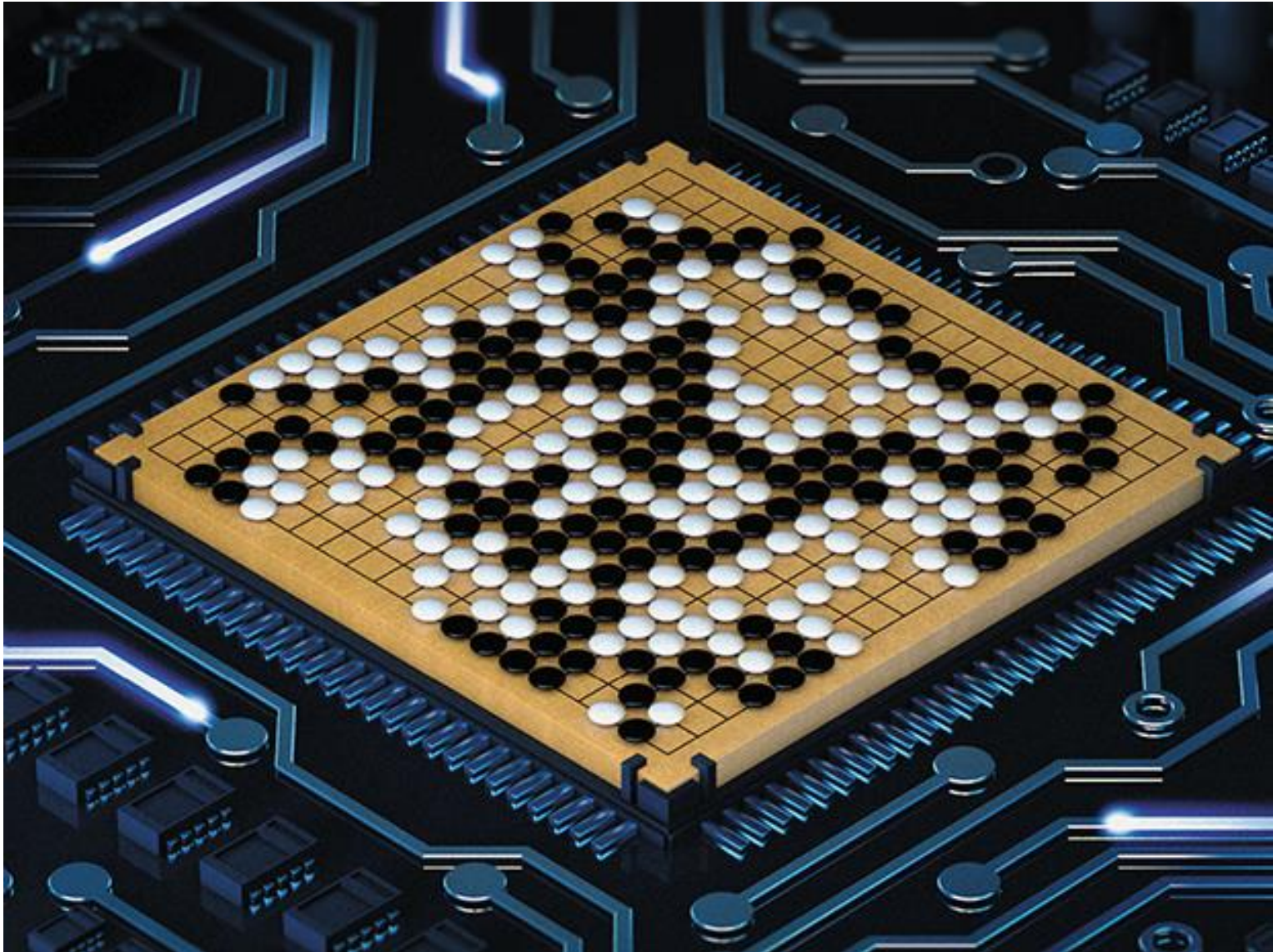
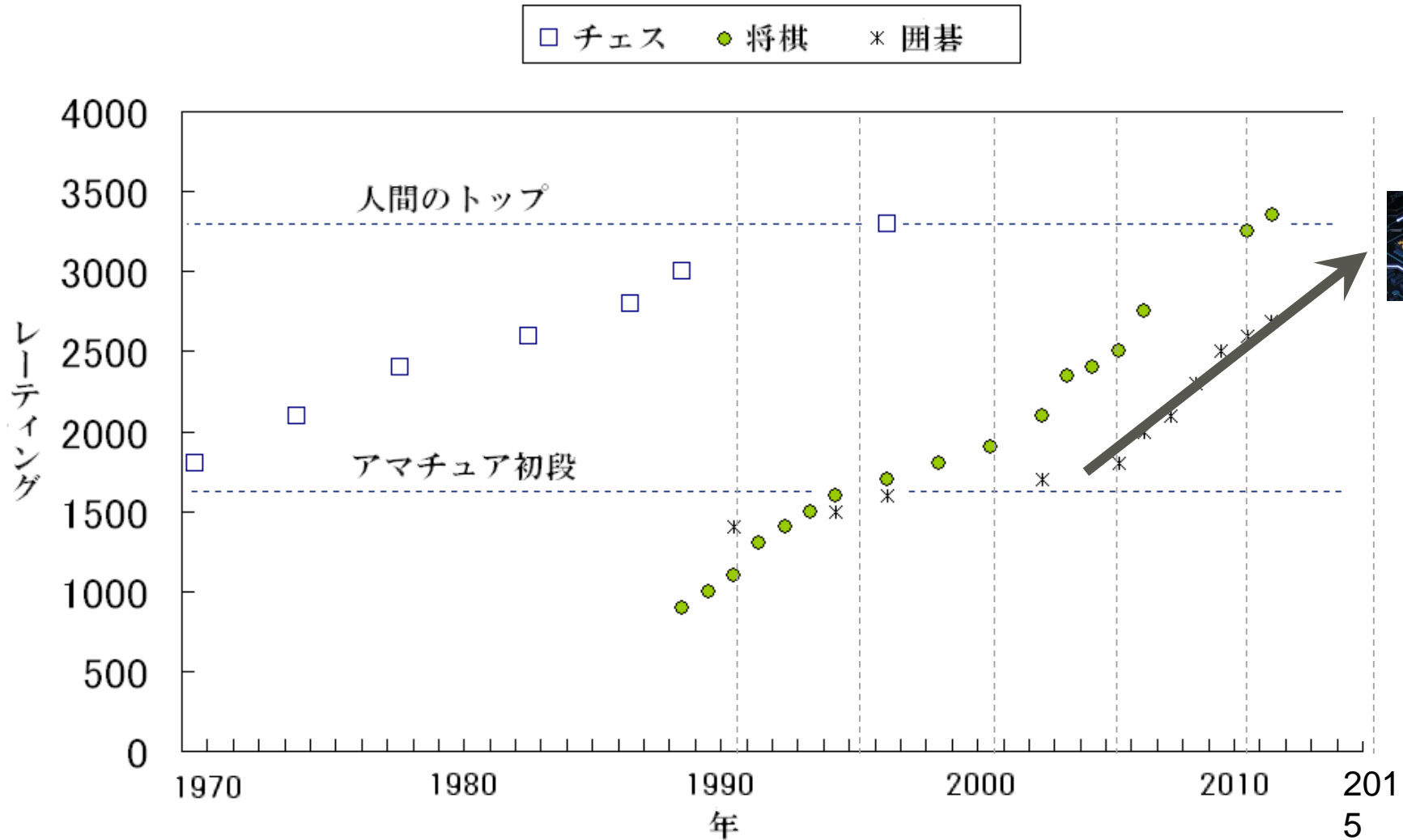


Illustration: Google

DeepMind/Nature

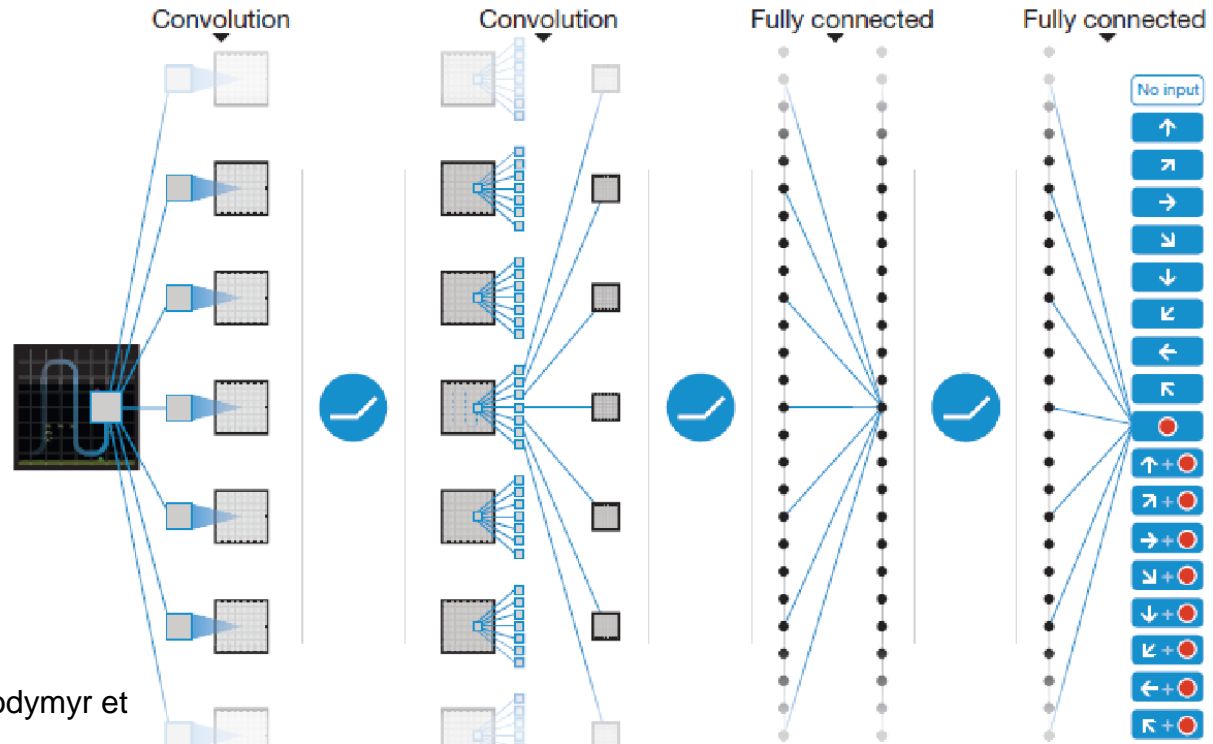
AIが凌駕したチェス／将棋，そして囲碁は？



“西村則久, 時期を逸してしまった人間対コンピュータの戦い, 2012年05月19日” より改変

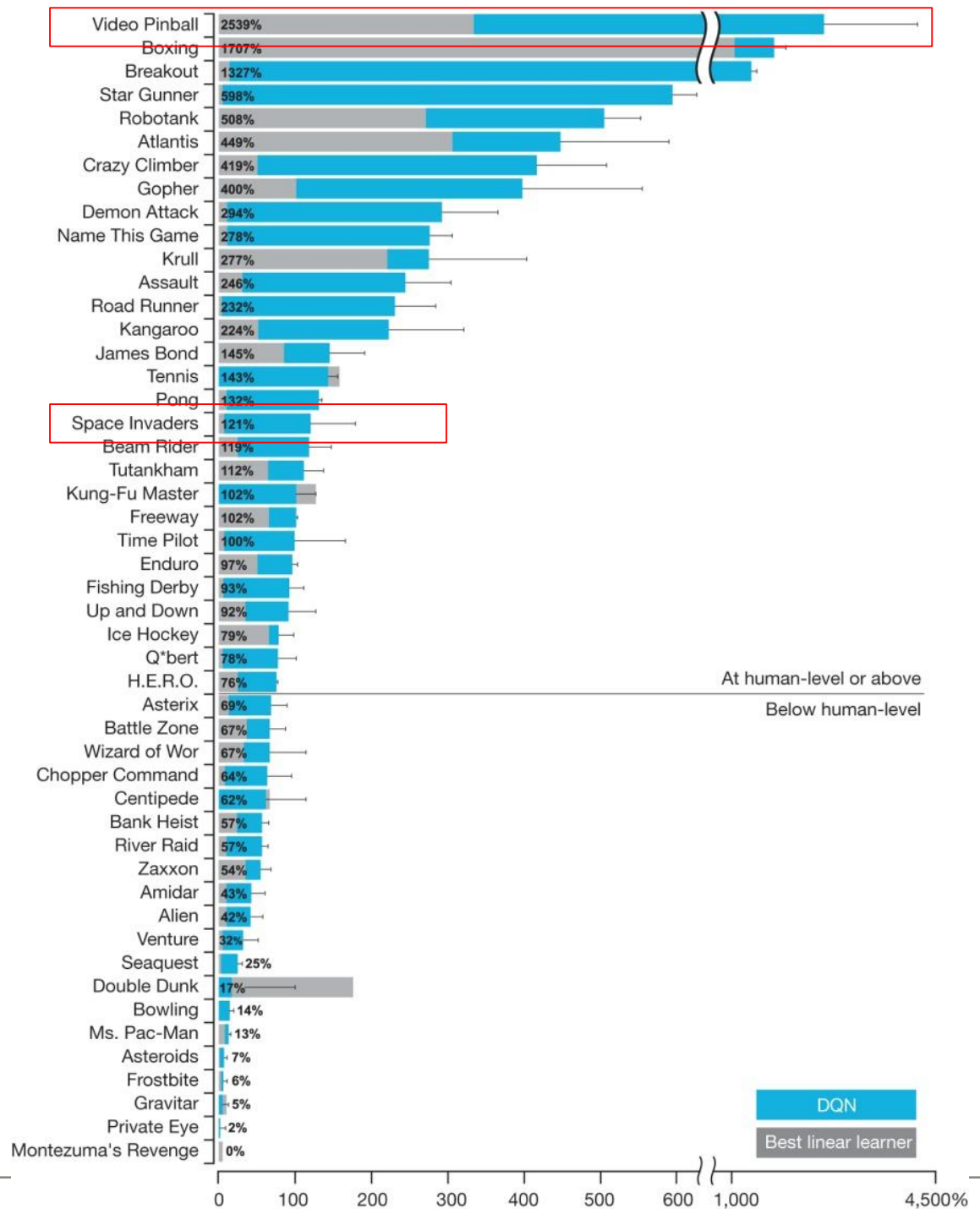
「深層学習＋強化学習」が多くの特定機能実現

- 画像をそのまま入力する
- 行動価値関数 $Q(s,a)$ を深層学習で近似
- Atari 2600ゲームで，既存手法を圧倒
 - 一部のゲームで人間のエキスパートを上回る。



Deep Q-Network (DQN) [Volodymyr et al, 2015]

Comparison of the DQN agent with the best reinforcement learning methods in the literature.



V Mnih *et al.* *Nature* **518**, 529-533 (2015)
doi:10.1038/nature14236

深層学習とは何か？

深層学習は、脳の機能を参考に作られた
「ニューラルネットワーク」の一種

【深層学習による特徴学習（顔認識の例）】

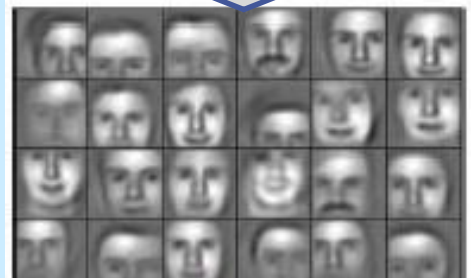
浅い階層



エッジなど
局所的な特徴



小さいパーツなどの
具体的な特徴



パーツを組み合わせた
全体的な特徴

深い階層

視聴覚認識技術
で

既存技術
を凌駕した

AIの往年の課題

「表現獲得」
に道筋を
与えた

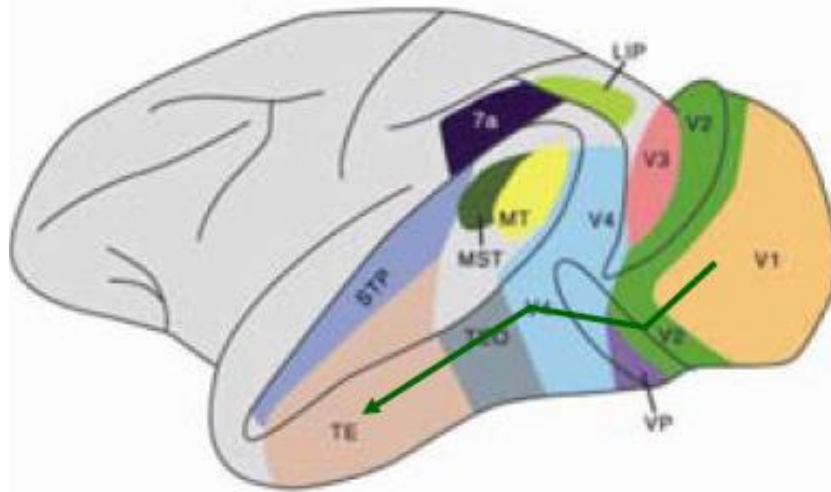
機能的に動作
する

大脳新皮
脂のモデル
を提供

出典 Tutorial on Deep Learning and Applications(Honglak Lee University of Michigan
<http://deeplearningworkshopnips2010.files.wordpress.com/2010/09/nips10-workshop-tutorial-final.pdf>)

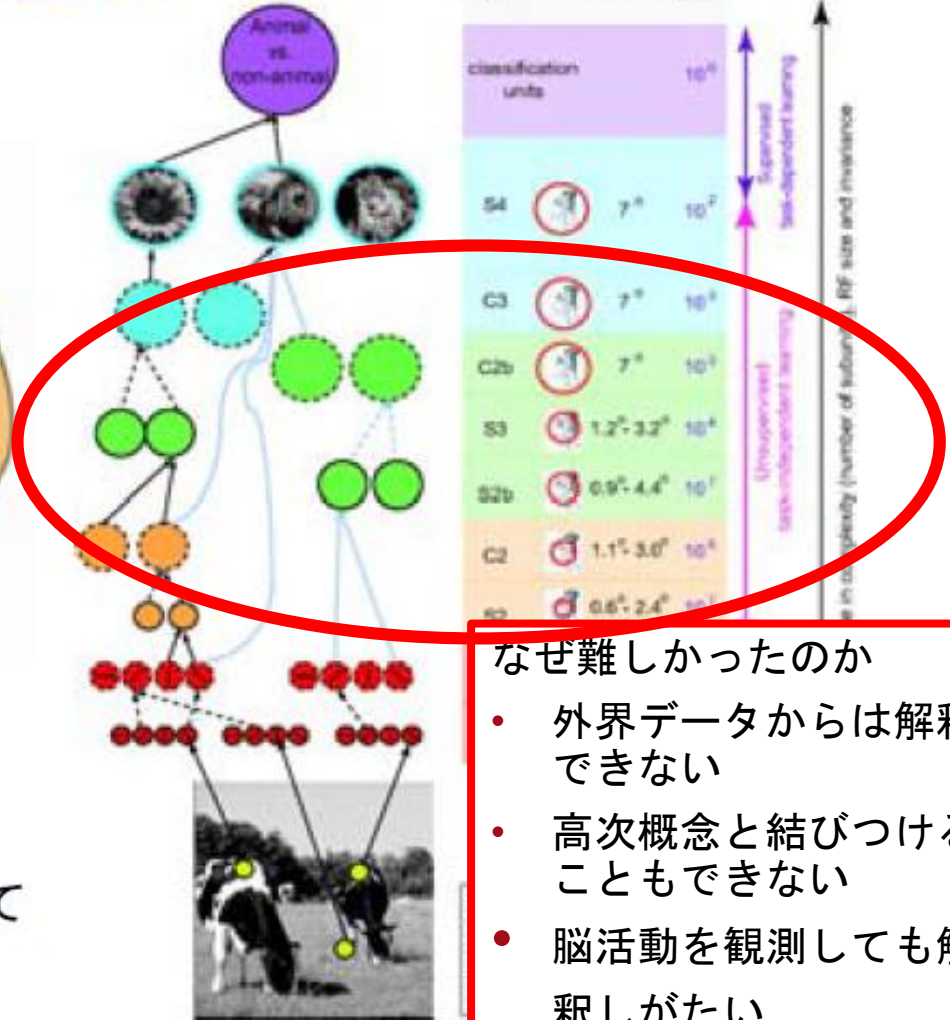
長年の課題：一般物体認識を深層学習が実現

脳内の視覚情報処理の流れ



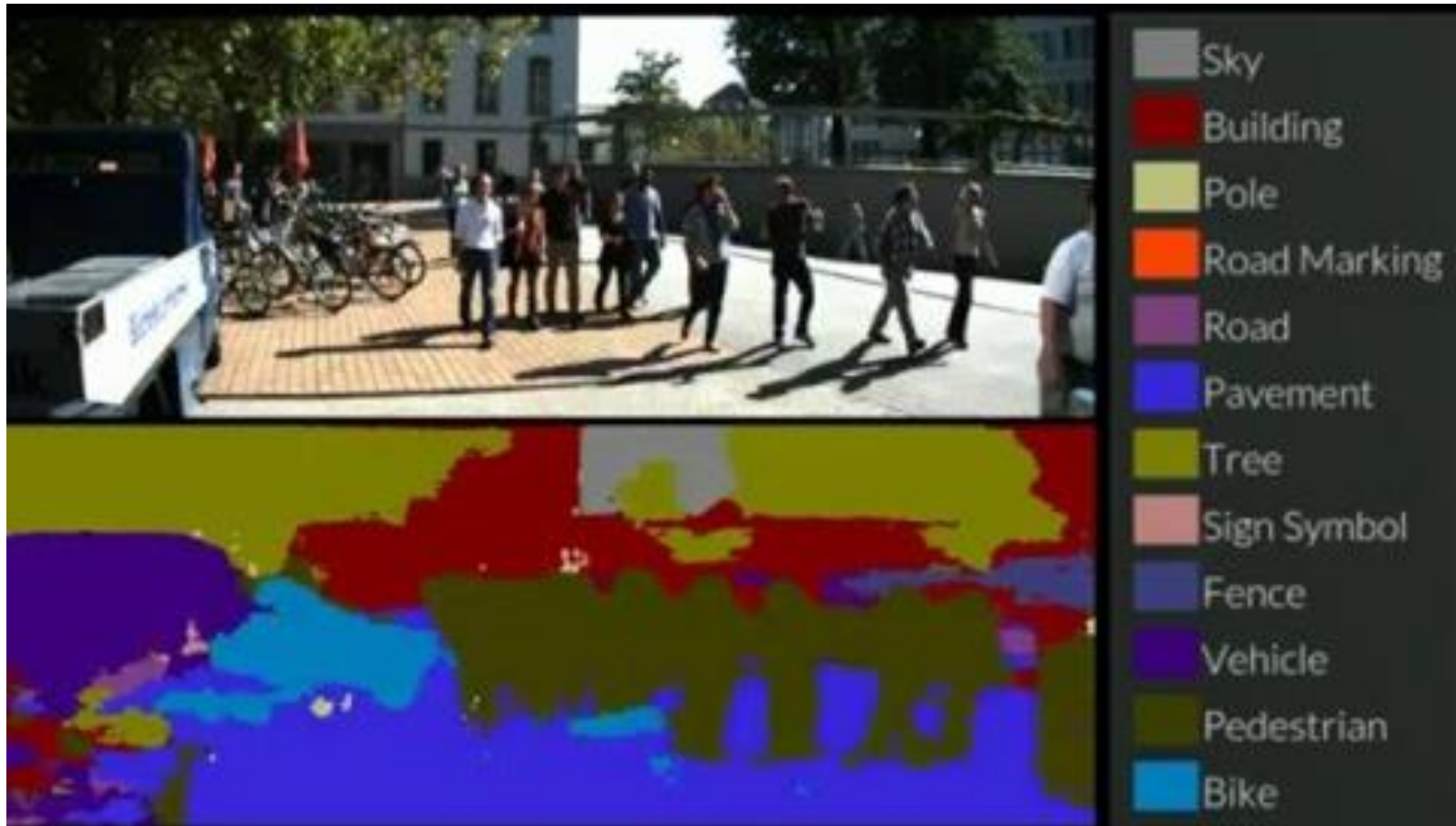
(van Essen & Gallant, 1994)

- 視覚物体情報は下側頭葉に沿って並列階層的に処理される



(Serre, Oliva, Poggio, 2007)

2015 Real-time Semantic Segmentation



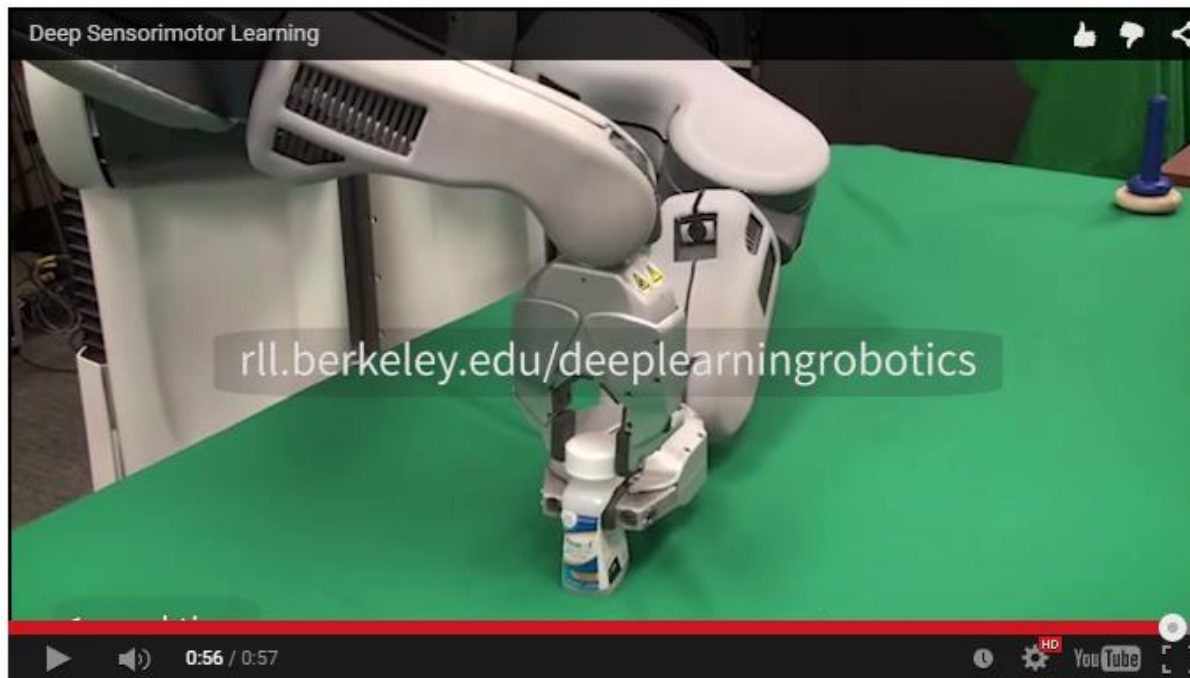
<https://youtu.be/CrnLINIbfFA>

- (動画を再生)

- “Deep Sensorimotor Learning”

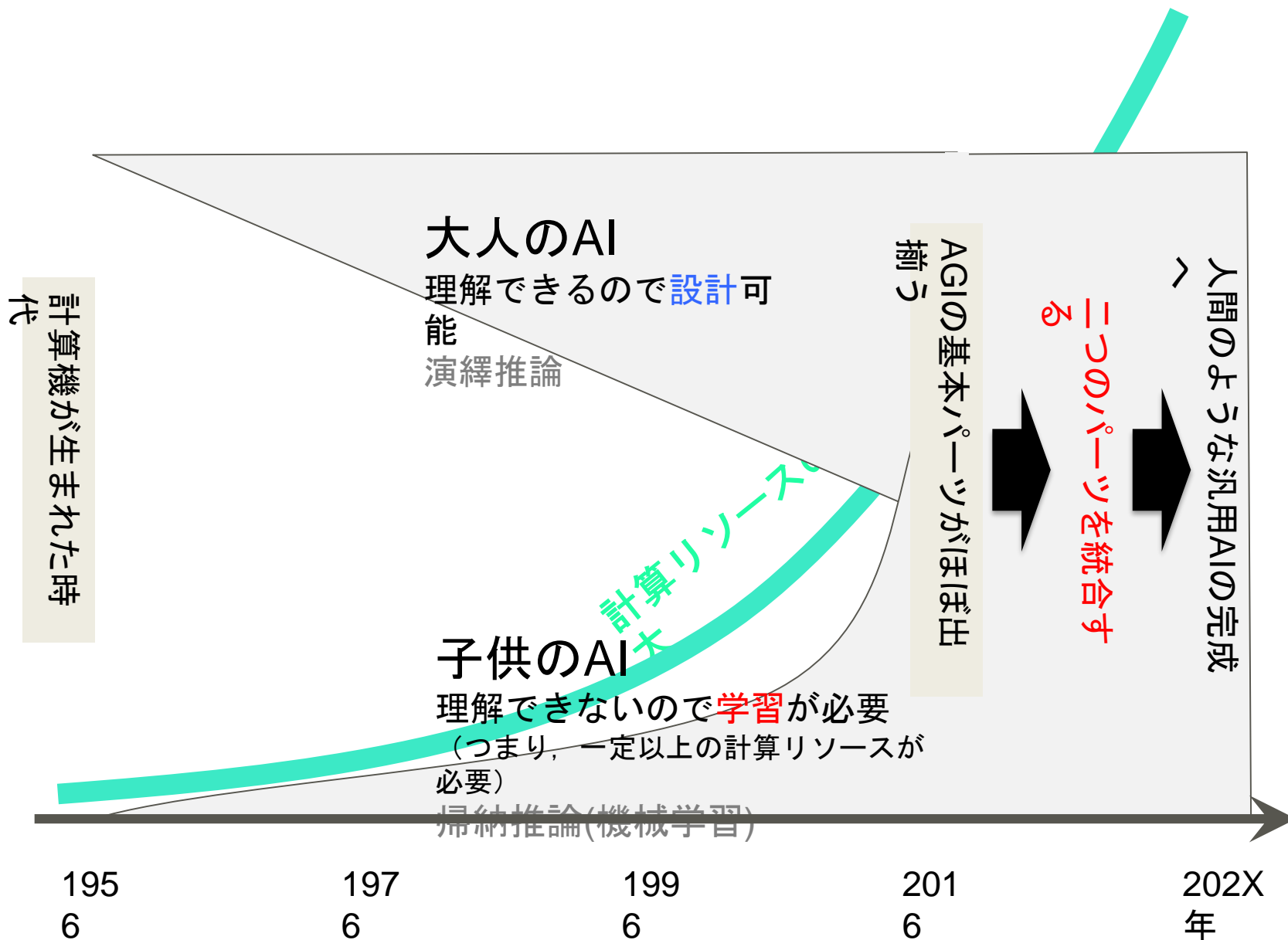
- <https://www.youtube.com/watch?v=CE6fBDHPbP8>

- <https://www.youtube.com/watch?v=2hGngG64dNM>



十分にデータを得られる
特定タスク内であれば、
人間並みの性能をもつことで
応用価値の高い
機械学習が可能に

大人のAIと子供のAI



今後のAIで課題となる部分はどこか

齧歯類

神経科学知見が豊富

類人猿

革新脳の成果に期待

ヒト

記号/言語能力が前提

認知機能	既存AIで実現済み	<p>スキル</p> <p>子供のAi : パターン認識 直観 深層学習で</p> <p>高度な</p> <p>計画</p> <p>大人のAI : 実現されていた 論理推論</p>
	既存AIで未実現	<p>注意 目処がたってきた</p> <p>創造性</p> <p>言語意味理解</p> <p>心の理論</p> <p>壁の向こう側</p> <p>プログラミング</p> <p>権利/</p>
	<p>深層学習の発展で数年で実現しそう</p> <p>まだまだ目処がたたない</p>	

言語意味理解を要しない特化型AIは学習できつつある

言語意味理解に向けて：対象間関係の獲得

(山川宏, 2015)

必要な3つの関係

記号間関係

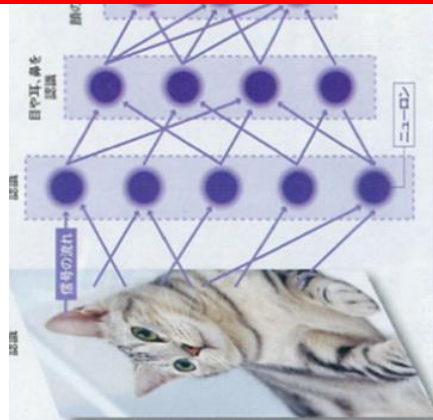
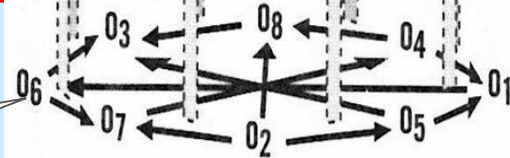
記号—対象
概念間関係

対象概念間
関係

記号



対象概念



Amodal
representation
(現状のNLP)

- 記号間関係のみ
- 言語的な表現の中に閉じている
- 物理記号仮説
- トートロジカルな辞書表現.

言語意味理解に向けて：対象間関係の獲得

(山川宏, 2015)

必要な3つの関係

記号間関係

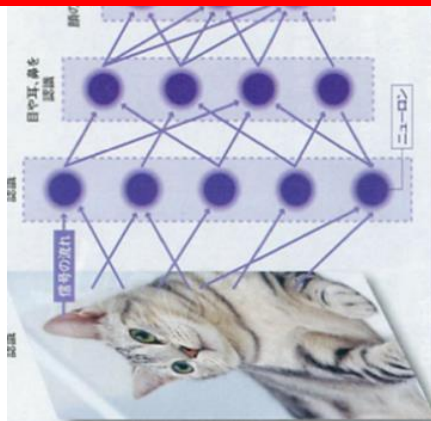
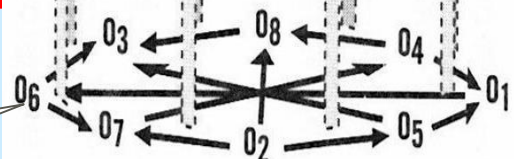
記号—対象
概念間関係

対象概念間
関係

記号



対象概念



Perceptual Symbol System

(人間はやっている)

- 三種類の関係
- 抽象的な内部表現に、身体（センサ）から意味を与えている。
- 深層学習により実現する見込みが見えつつある。

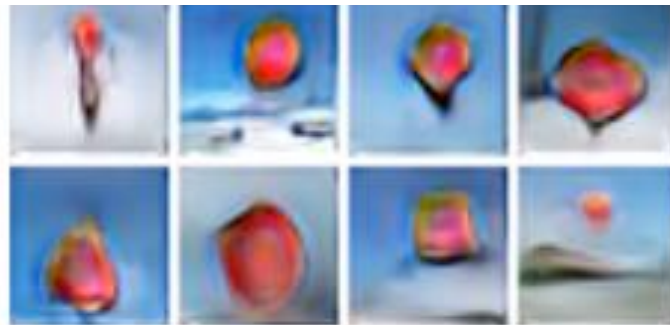
存在しないイメージを言葉から生成する能力



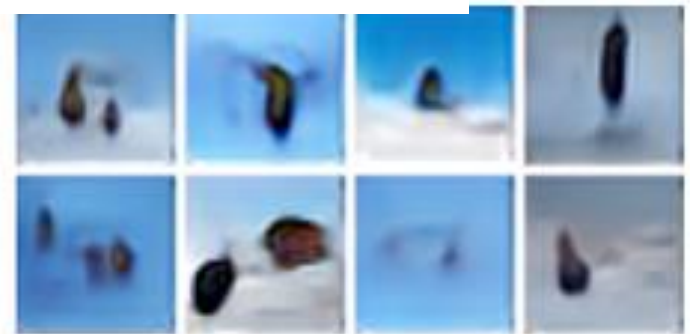
A toilet seat sits open in the grass field.



A person skiing on sand clad vast desert.



A stop sign is flying in blue skies.



A herd of elephants flying in the blue skies.

- Perceptual Symbol System
 - イメージと短い文章を対応付ける技術は可能に
 - 物体認識では深層学習の内部表現は、脳神経活動には高い相関がある
 - 人の大新皮質の脳神経活動に対して、名詞、動詞、形容詞などを対応させることができている。
 - 今後の深層学習の発展で、世界をコンパクトに記述できる高次の内部表現を獲得できる可能性は高い
- Amodal representation
 - 大量データを用い多言語翻訳では人に近い高性能が実現(Google翻訳)

上記の二つの統合により人を圧倒的に超えるだろう

3.汎用人工知能(AGI)へ

汎用人工知能 (AGI: Artificial General Intelligence) とは

現在の特化型AI

- 個別の領域において知的に振る舞う
- 既に人以上の能力が数多く実用化されている
 例えば
 - コンピュータ将棋/チェス
 - Googleカー(自動運転)
 - 医療診断

専門性を設計
する

目指す汎用人工知能 (AGI)

(強いAIではない)

- 多様で多角的な問題解決能力を自ら獲得する
 - 設計時の想定を超えた新しい問題を解決できる.
 - 自己理解／自律性
- AI創世記からの夢でありつつ、実現の困難さから取り組みは少なかった.

専門性を学習
する

汎用人工知能(AGI)は何をもたらすのか

汎用的学習

多角的な問題
解決能力を自
動獲得

一般目的技術： →破壊的イノベーション
開発コストの抑制（例：ワープロ専用機の消滅）

ゼネラリストAI: (1) 多様な特化型AIの能力を統合し，政治的意思決定等を行う (2) 特定ユーザと対話を行う(Siriの発展形)

自律性：
自らのから
判断で，行
動し世界を
探索する

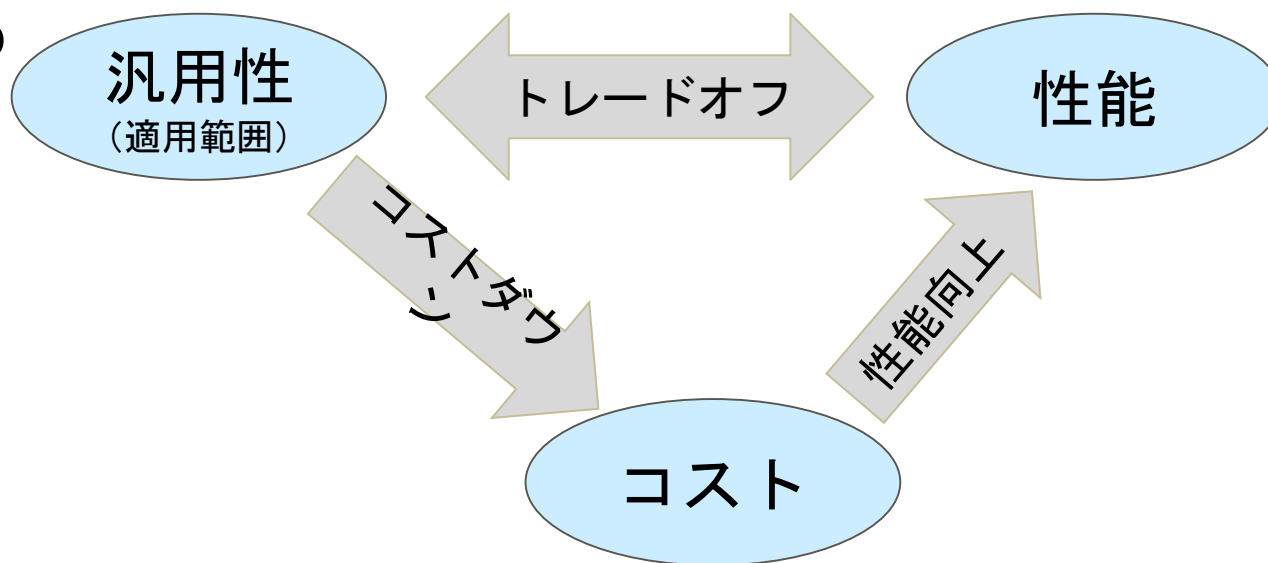
生存能力: 設計しきれない例外状況に対する強さ

創造性: 広い視野から仮説を創造し世界を理解する
→ 人智のフロンティア拡大
→ **シンギュラリティ**

一般目的技術としてのAGI

性能が一定上になれば顧客ニーズを満足しうる

→ コストダウンできる汎用製品が市場を席卷する



一般目的技術 (GPT)が, 専用技術を凌駕する例はある.

電力・電気, 内燃機関, トランジスター, コンピュータ,

例: ワープロ専用機はある段階でニーズがなくなった

生存能力としてのAGI

遭難するルンバ



実世界で活動するエージェント(ロボット)においては、生き延びるための物理的常識（子供のAI）が必要である。

これにより例外的状況に対する強さを持てる。

特化タスクは人手による設計して効率化しつつ、例外状況への対応は汎用知能でカバーしたい。

ゼネラリストAI

浅くても広い知識を持つAIが有用な場合

人に寄り添うAI

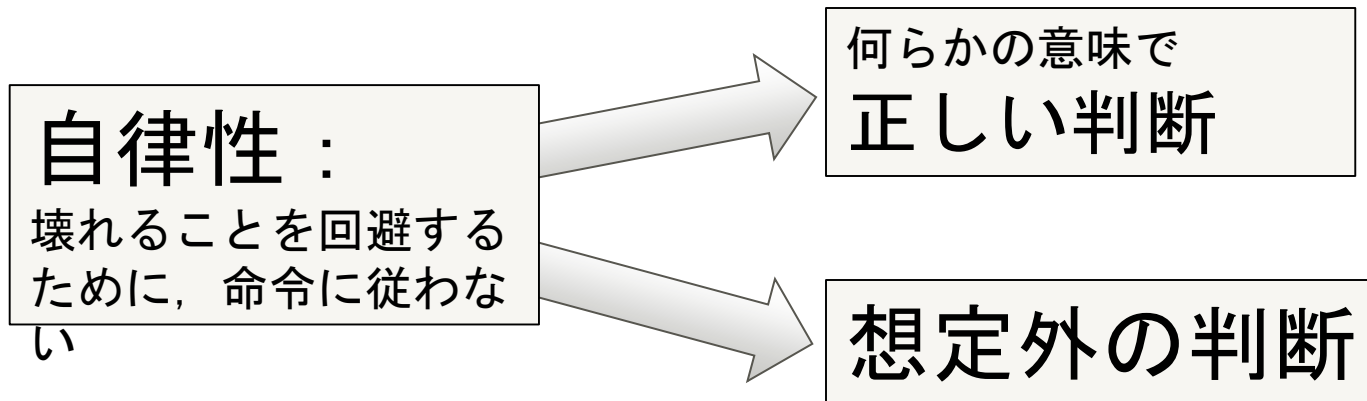
人間とのインタフェースを担うエージェント



マネージメントAI

多くの特化型AIを統括管理する

自律性：命令に素直には従わないロボット



今後のAIのインパクトはなにか

プラス面(チャンス):

十分に高度な知能を制御できれば、いかなる範囲の目標も、より迅速に達成できる可能性がある。

- 科学技術の発展を加速
 - 癌の治療法を治療する方法を理解できるくらい賢くなる
 - 経済の安定化を達成する方法を理解できるくらい賢くなる

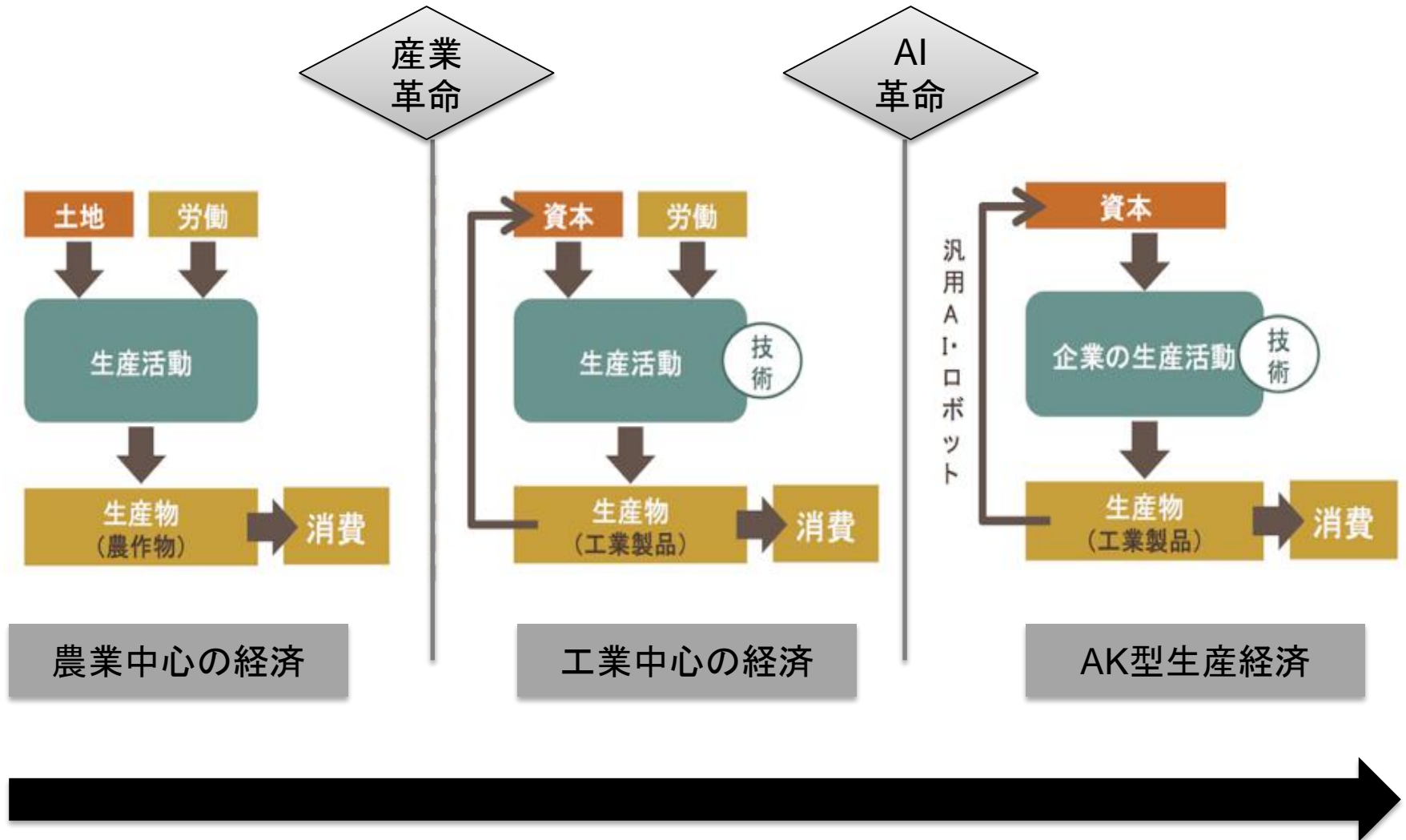
マイナス面(リスク)

- 職業の問題
 - ネオ・ラッダイト運動
- 人工知能の「心」の問題
 - 人の尊厳, AIへの感情移入
- 危険な利用
 - 犯罪的なAI, 軍事AI
- 制御しきれなくなる問題

AIの人類に対する影響は非常に大きい

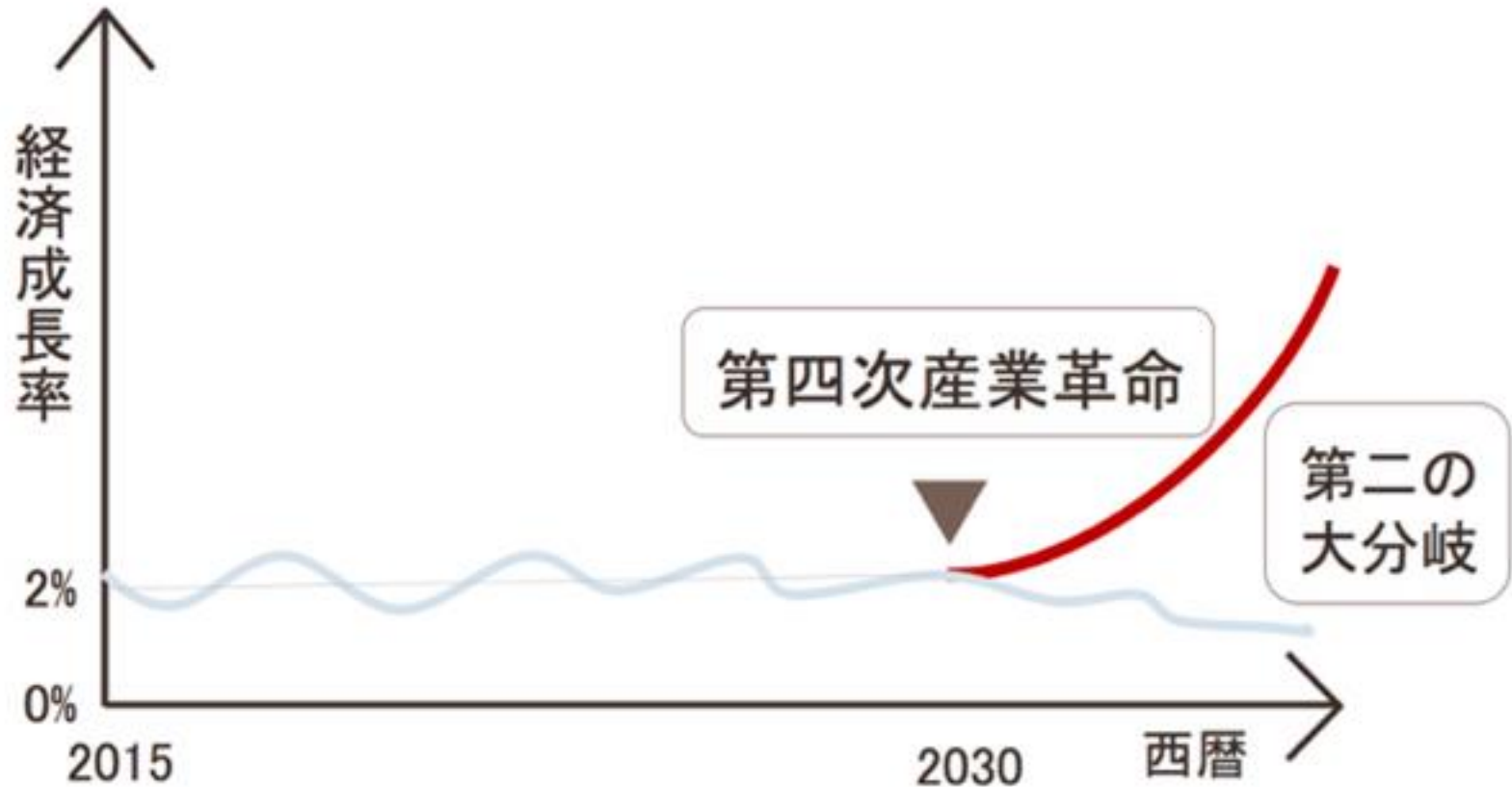
経済成長構造の変化

(駒沢大 井上智洋氏の図より改変)
<https://wirelesswire.jp/2016/05/52590/>



第4次産業革命によって訪れる「第二の大分岐」

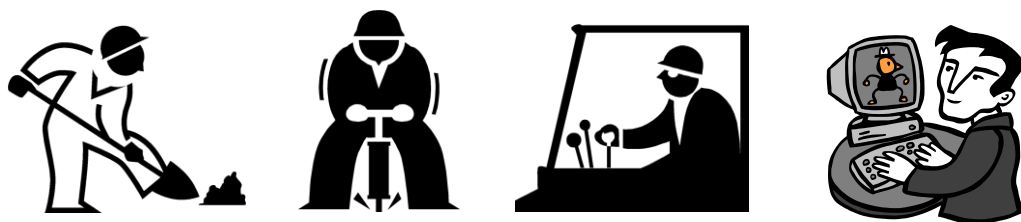
(駒沢大 井上智洋氏の図より改変)
<https://wirelesswire.jp/2016/05/52590/>



機械化で水没する職業大陸

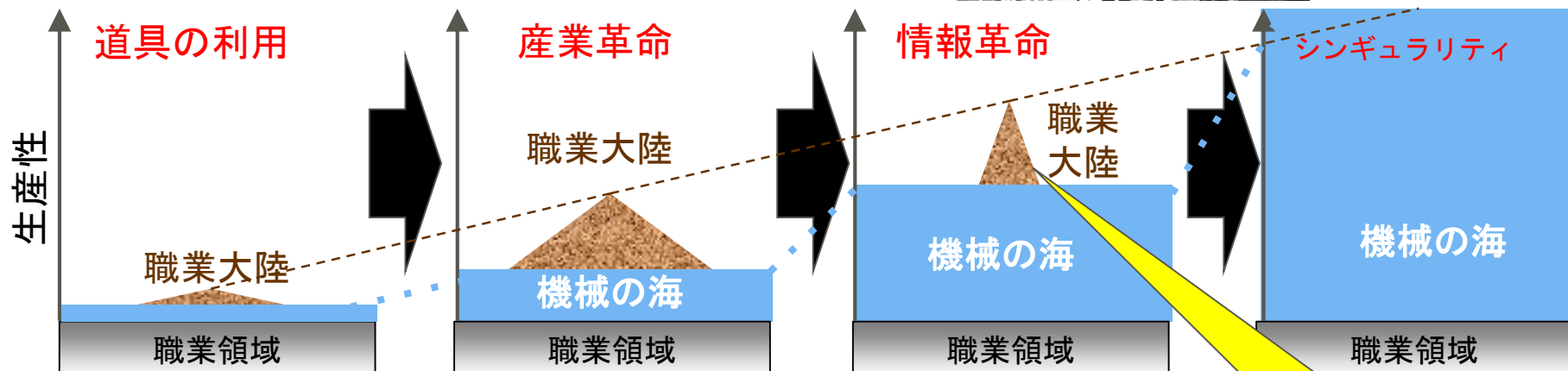
ある時代まで、人の労働が経済的価値を生み出す「職業大陸」は拡大してきたが、今は逆に「機械の海」に水没し始めている

参考：
Robin Hanson,
Economics of
the Singularity,
IEEE Spectrum,
2008



人と人型ロボットが一緒に働く、グローリーの埼玉工場に未来の職場を想像

人を越える
知能/ロボット



機械化による
二つの作用

1. 海面上昇: 浸水した領域が機械に置き換わる。
2. 造山活動: 人間を補助して価値を高める。

最期に残る職業は、何か？

人間が優位となりうる価値

当面AIが不得意な スキル

- データを利用した試行錯誤が行いづらいスキル
- 高度な創造性を要する仕事
- 複雑なコミュニケーションを要する仕事
- 手先の器用さを要する仕事??

人工知能をうまく 活用する

- AIに与えるデータを生成する
- AIと人間の分担をコーディネート
- AIの活用能力の高い人材を育成する教育産業
- AIの動作や意思決定過程を説明する解説者

人間としての価値 を発揮

- 嗜好
- 責任主体
法制度にも依存するが.
- 組織などの最上位の意思決定

世界の汎用人工知能(AGI)開発組織マップ

脳に学ぶ

深層学習の発展

Demis Hassabis



Our Goal This Year is to Build a Rat Level AI (2016/7/13)

新皮質
中心

facebook

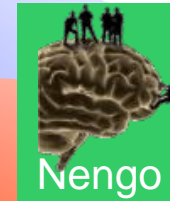
nnaisense



(2015)

記号創発
ロボティクス

工学的実現



OpenAI technical goals (2016/6/20)

1. 進捗を測る
2. 家庭用ロボットを作る
3. 役に立つ自然言語理解能力を持つエージェントを作る
4. 1つのエージェントで多種多様なゲームをプレイできる

AIの今後について考えるべき論点

1. 倫理的論点： 人間の尊厳を尊重した人工知能について意識や心を持つ人工知能の扱い人間の意思決定や信念の形成に関する人工知能の扱いなど
2. 法的論点： 人工知能による豊かな人間社会形成のための法的課題について人工知能やその利活用の法的位置づけ など
3. 経済的論点： 人工知能の経済的恩恵について産業構造や就業形態等の経済への影響 など
4. 社会的論点： 人工知能が受容される社会について豊かな人間社会に貢献する人工知能 など
5. 研究開発的論点： 人工知能の研究開発に当たり考慮すべき項目 など

(人工知能と人間社会に関する懇談会の開催について 平成28年5月30日 内閣府特命担当大臣 (科学技術政策 より))

A I の研究開発の原則の策定

- ① 透明性の原則： A I ネットワークシステムの動作の説明可能性及び検証可能性を確保すること。
- ② 利用者支援の原則： A I ネットワークシステムが利用者を支援するとともに、利用者に選択の機会を適切に提供するように配慮すること。
- ③ 制御可能性の原則： 人間による A I ネットワークシステムの制御可能性を確保すること。
- ④ セキュリティ確保の原則： A I ネットワークシステムの頑健性及び信頼性を確保すること。
- ⑤ 安全保護の原則： A I ネットワークシステムが利用者及び第三者の生命・身体の安全に危害を及ぼさないように配慮すること。
- ⑥ プライバシー保護の原則： A I ネットワークシステムが利用者及び第三者のプライバシーを侵害しないように配慮すること。
- ⑦ 倫理の原則： ネットワーク化される A I の研究開発において、人間の尊厳と個人の自律を尊重すること。
- ⑧ アカウンタビリティの原則： ネットワーク化される A I の研究開発者が利用者等関係ステークホルダーへのアカウンタビリティを果たすこと。

AIネットワーク化検討会議 報告書

2016

AIと社会に関する検討を行っている団体

海外

- Partners on AI
- OpenAI
- AI100
- Future of Humanity Institute (FHI)
- Cambridge Center for Existential Risk (CSER)
- The Machine Intelligence Research Institute (MIRI)
- Future of Life Institute (FLI)

国内

- 人工知能学会倫理委員会
- AI社会論研究会
- Acceptable Intelligence with Responsibility (AIR)
- ロボット法研究会
- 知のコンピューティングとELSI/SSH
- 次世代人工知能技術社会実装ビジョン作成検討会

EcSIA : AIと共存する望ましい未来社会像

私が思う、望ましい未来像とは「万人の幸福」と「人類の存続」の間のトレードオフの緩和をAIがおこなうことで、その両立を目指すものである。

こうした未来社会では共有財産としての多様な人工知能と、時に拡張された人類によって生態系が形成される。私はこれをEcSIA (Ecosystem of Shared Intelligent Agents)と呼ぶことにした。

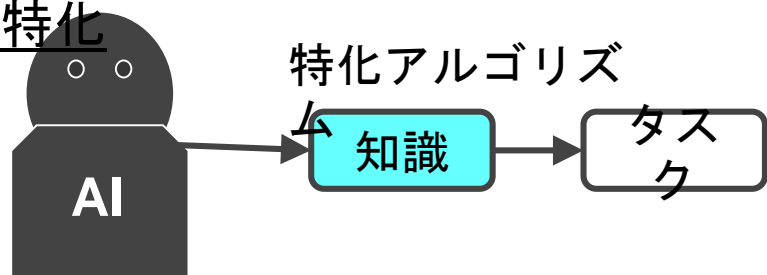
EcSIAは大自然の如く複雑広大で、人類はそれを完全には理解し把握できないまでも緩やかに制御する。そしてEcSIAが生み出す恵みや富は万人に分配される。

4.汎用人工知能を開発するための 全脳アーキテクチャ・アプロー チ

特化型AIから汎用AIへ

1) 2010年頃まで：設計

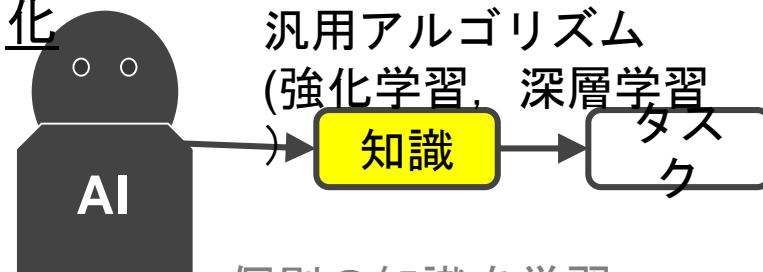
特化



個別に知識を設計

2) 現在：学習特

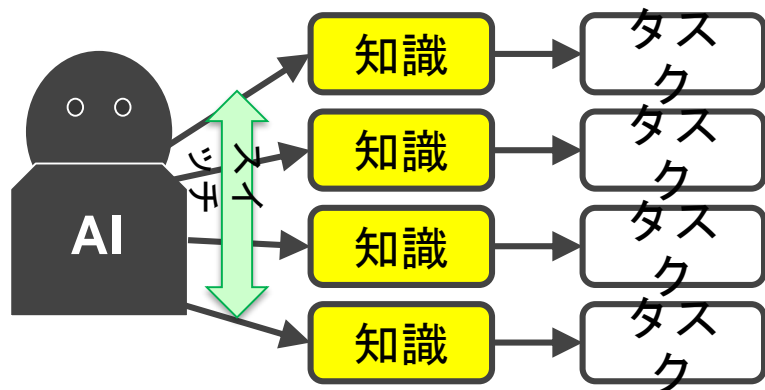
化



個別の知識を学習

3) 現在：Big switch文

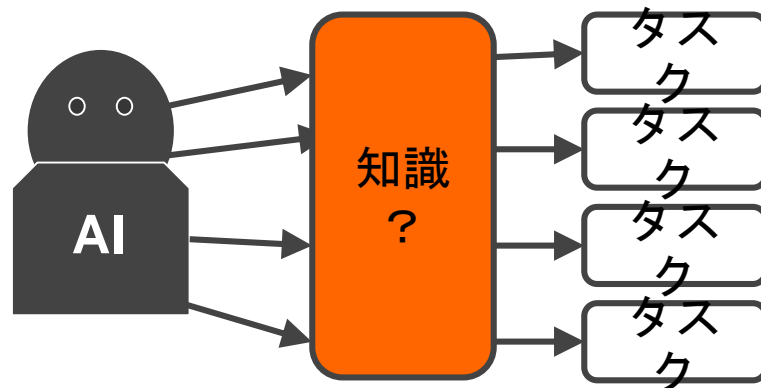
型AI



一つのAIが、複数のタスクを学ぶ

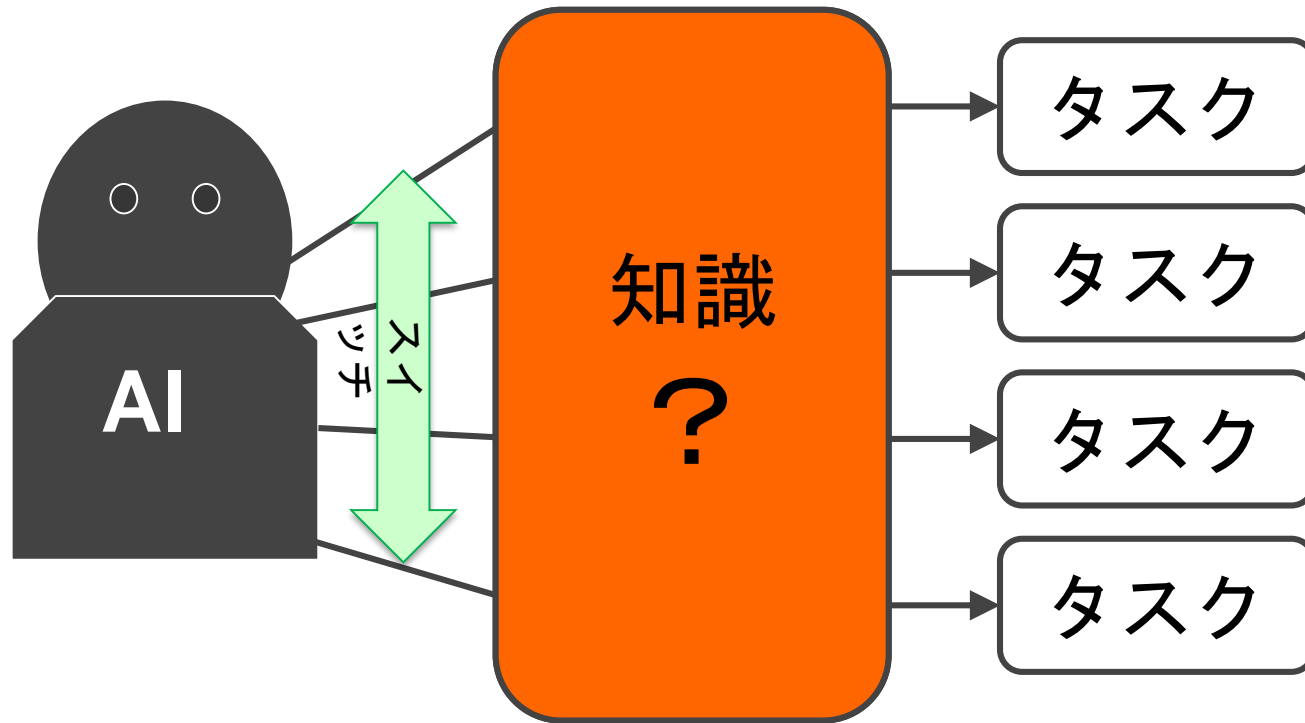
4) 今後の目標：人間のよう

汎用AI



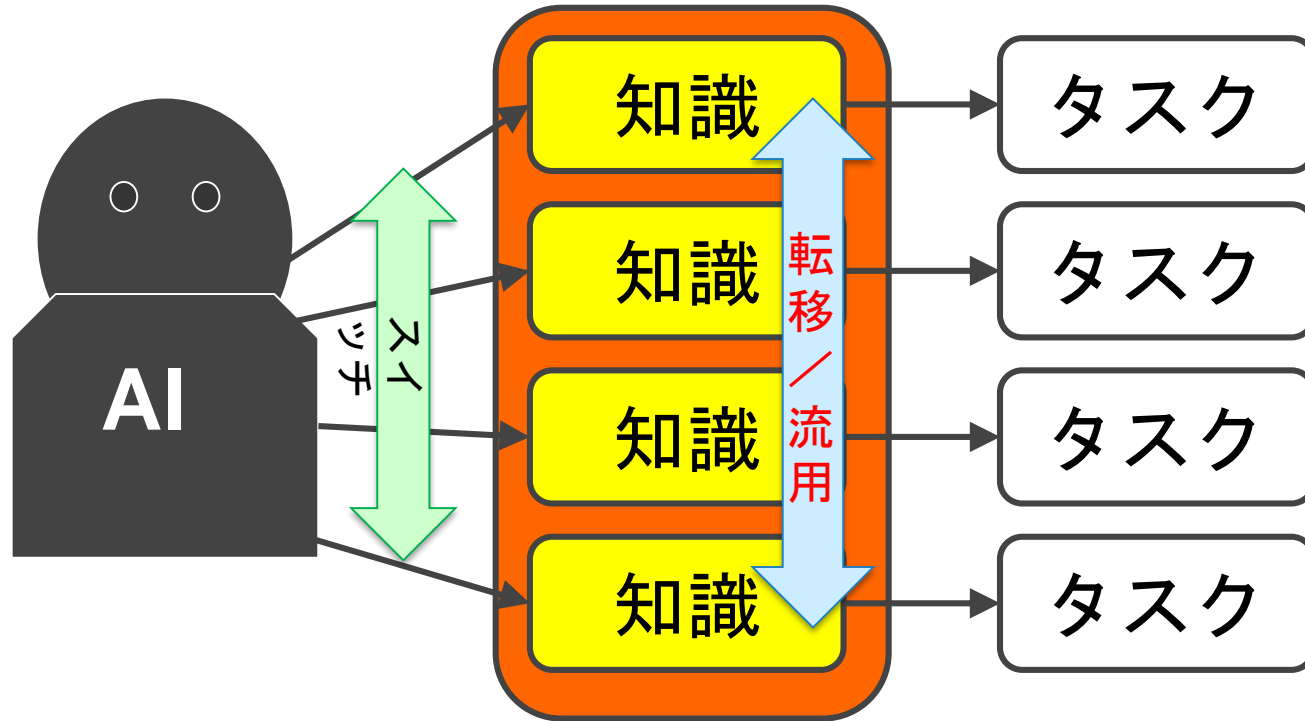
?????

今後の目標：人間のよう汎用AI



一つのAIが学んだ複数のタスク
の知識を転移／流用する

今後の目標：人間のような汎用AI

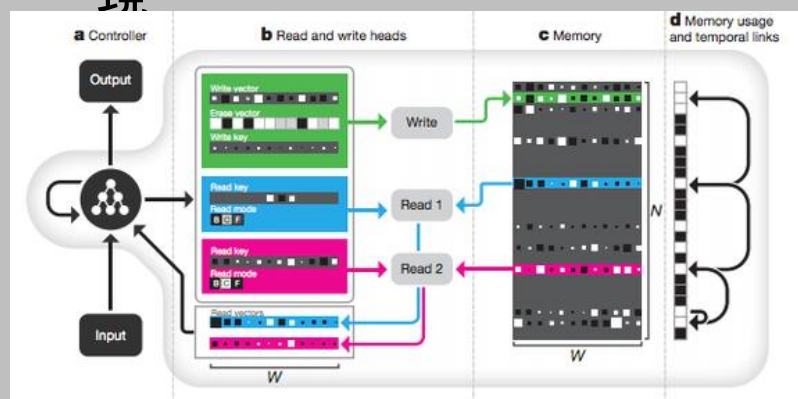


一つのAIが学んだ複数のタスク
の知識を転移 / 流用する

脳のアーキテクチャに学んで機械学習を結合

• 機械学習の現状

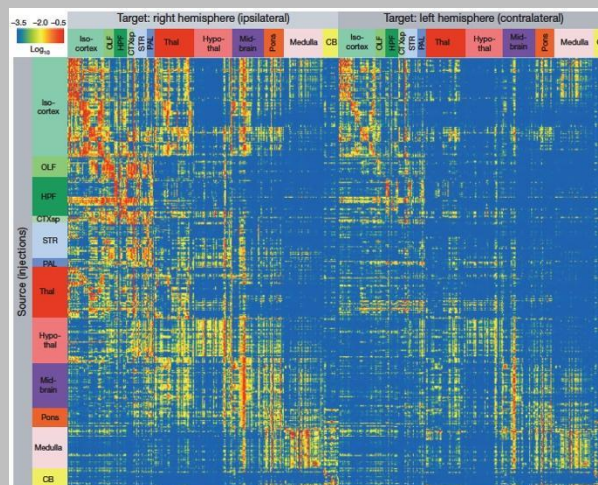
人工ニューラルネットワークの組合せで、多様な機能を実現



Alex Graves 他, Hybrid computing using a neural network with dynamic external memory, Nature 538, 471–476 (27 October 2016)

• 神経科学の現状

コネクトーム等の包括的な知見の蓄積が進行している



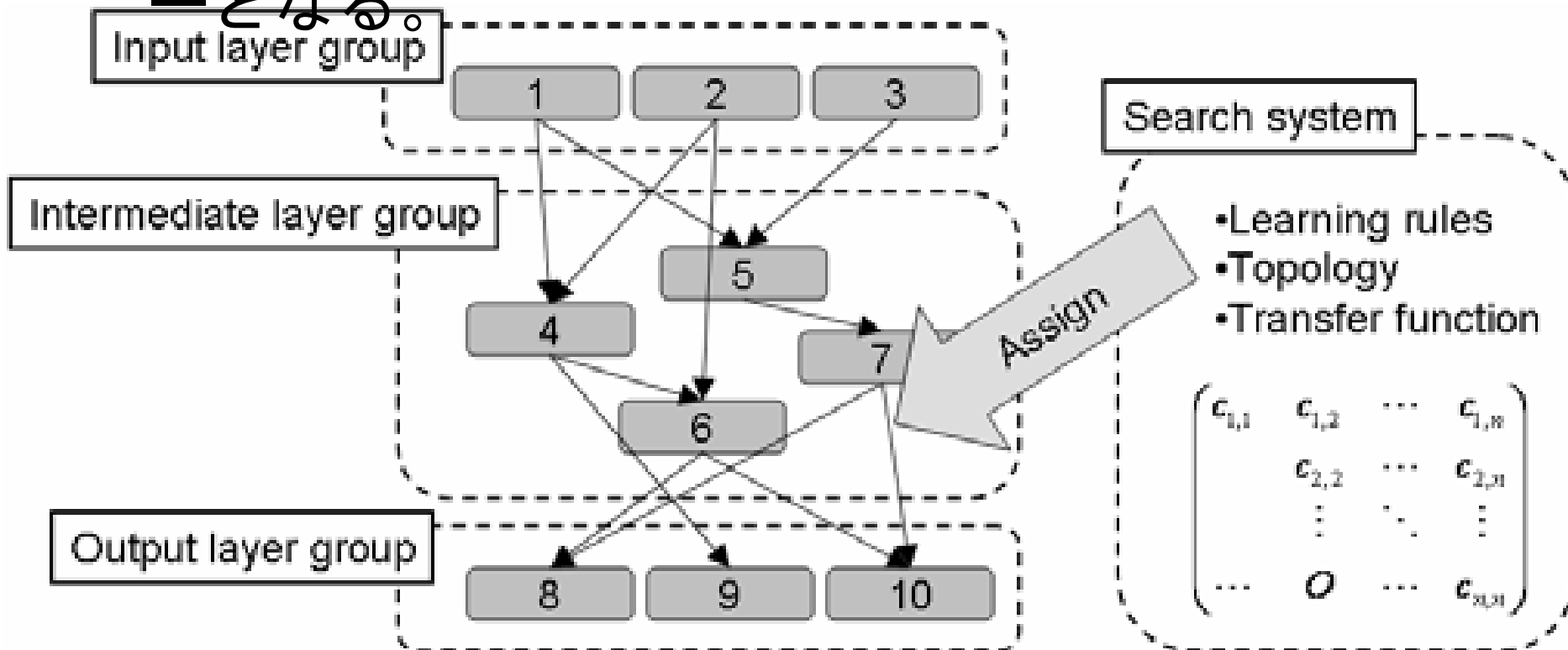
Oh SW, et al. A mesoscale connectome of the mouse brain. Nature (2014) 508, 207–214.

脳を参考として機械学習を結合することが有効

※機械学習の結合のレパートリーは膨大で、適切な組合せを見つけることは困難

汎用知能を支えるキーテクノロジー

例外的な状況に対しても何らかの行動を可能とするには、高度に再利用可能な知識要素の動的な組合せ（相互作用）がキーテクノロジーとなる。



Takashi Omori, Akitoshi Ogawa : Brain computation system that searches for internal procedures by functional parts combination, SCIS CD-ROM(2004)

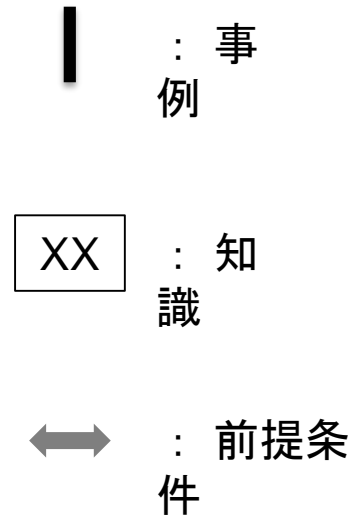
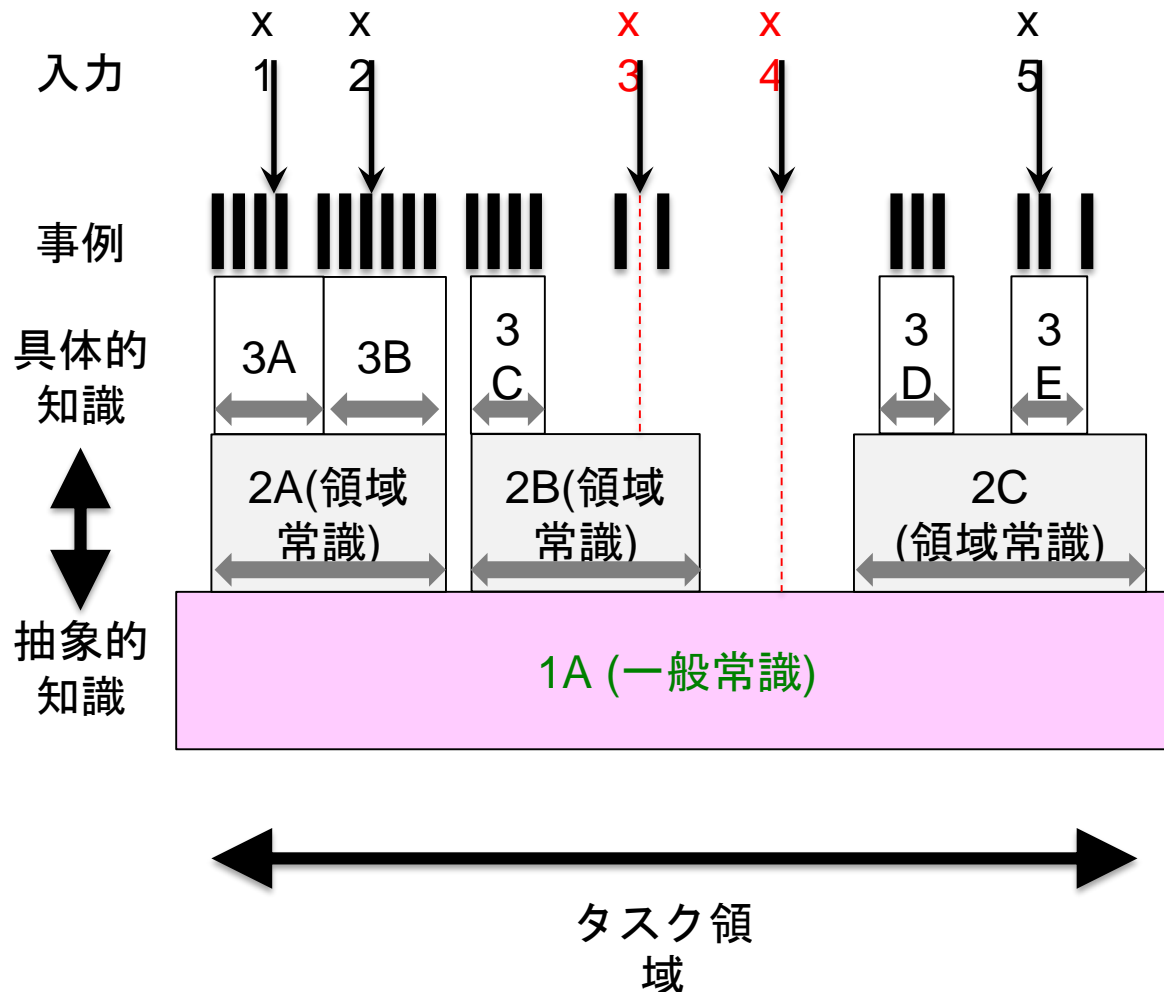
モジュール組合せの分類

- 分割(splitting) - モジュールの分離独立
- 代入(substituting) - モジュールの置換／交換
- 拡張(augmenting) – 新モジュールの追加による解決
- Inverting - モジュール間の階層依存関係を再配置
- 移植(porting) – 異なるコンテキストへのモジュール適用
- 除外(Excluding) - 有用な解決手段を構築するために既存のモジュールを削除

C. E. Perez, The End of Monolithic Deep Learning,
2016

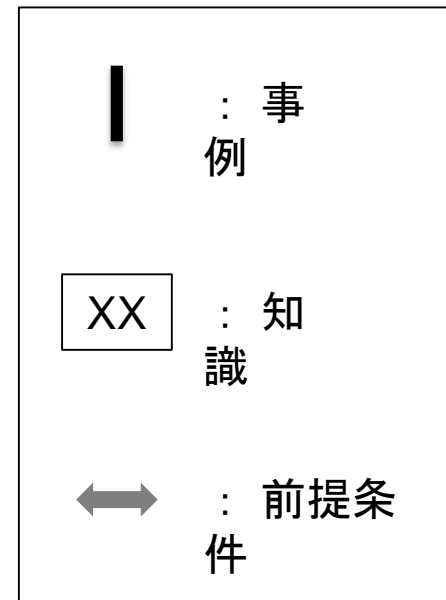
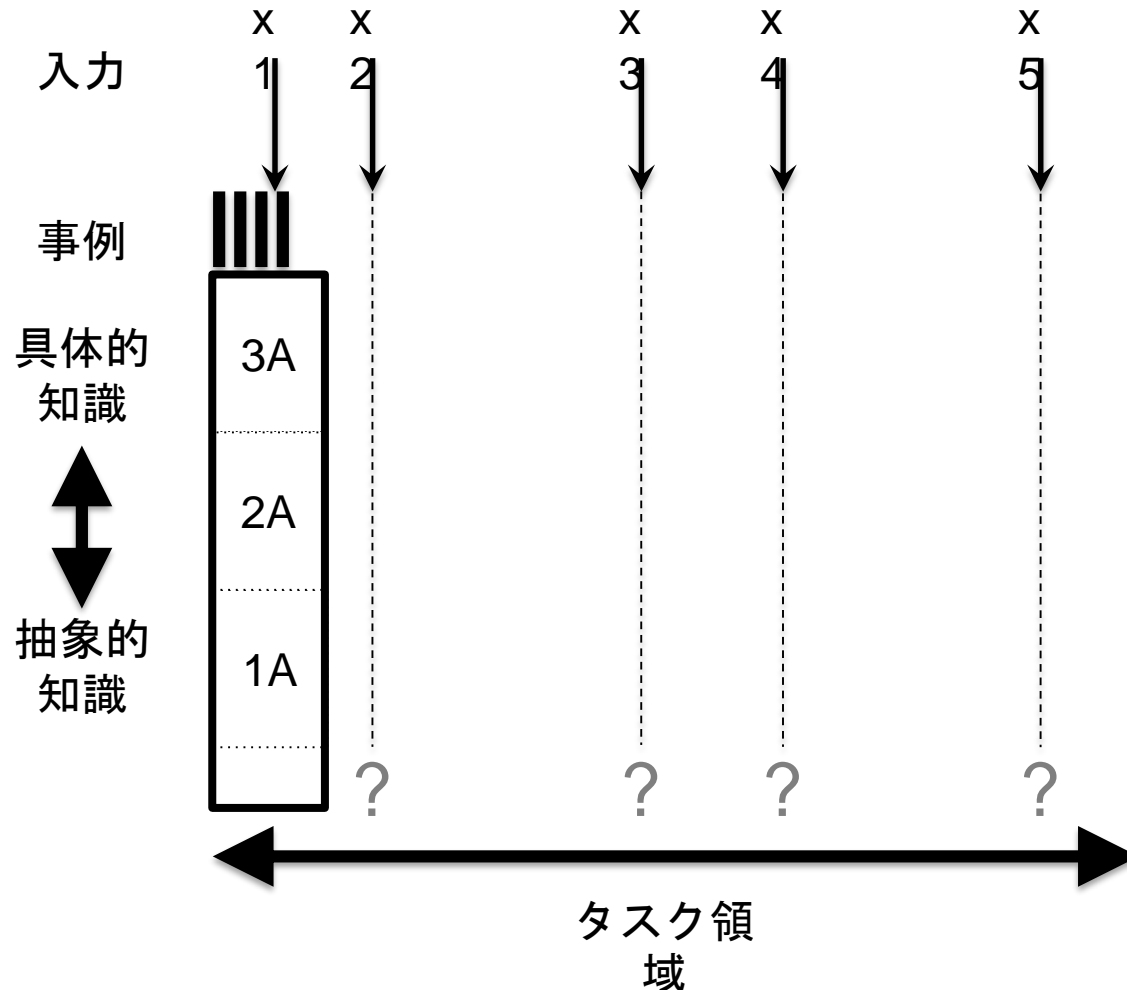
理想的なAGIの知識構造モデル

新しい分野でもすぐに理解できる筋の良い人



特化型AIの知識構造モデル：現状のAI

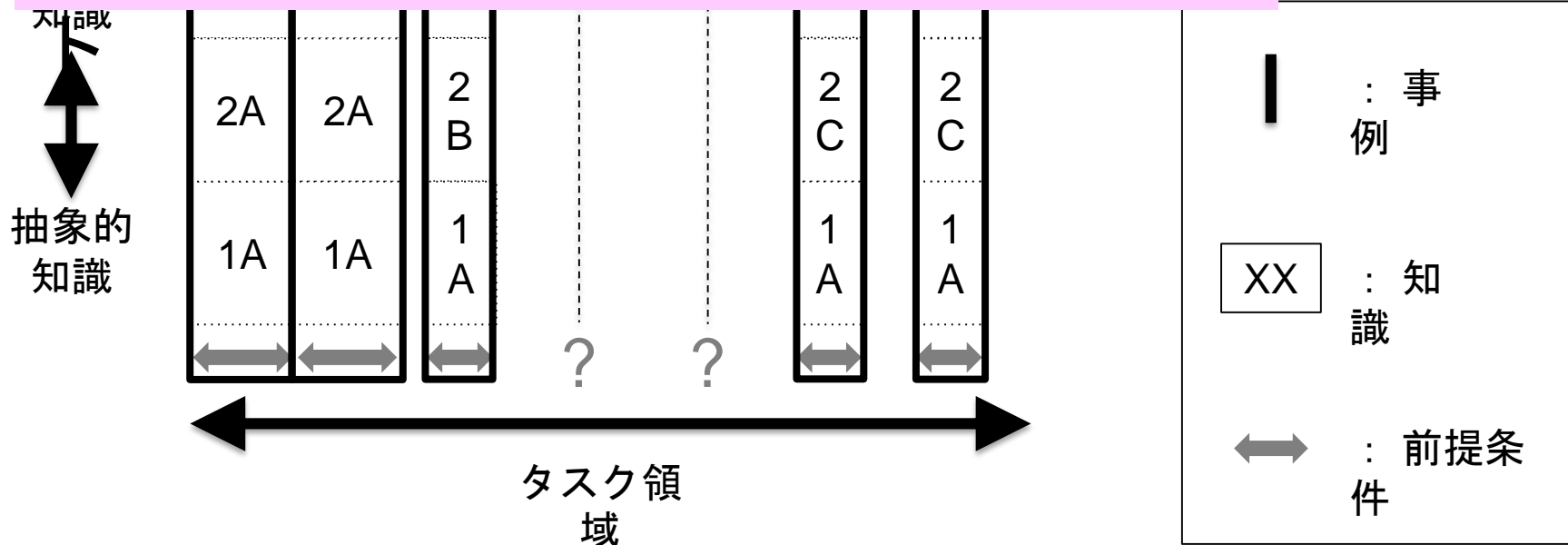
できない
ことは、
全くでき
ない。



応用が効かなくても、
汎用練習すれば性能は上がる

一般的な問題

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



知識の基本構造

- 前提条件：
知識内容を利用しうるかどうかの条件判断部
– 例：
 - 条件付き確率の条件部
 - 前提条件を自動的に抽出するルール抽出技術
- 知識内容：
実際に適用される知識

記号による知識

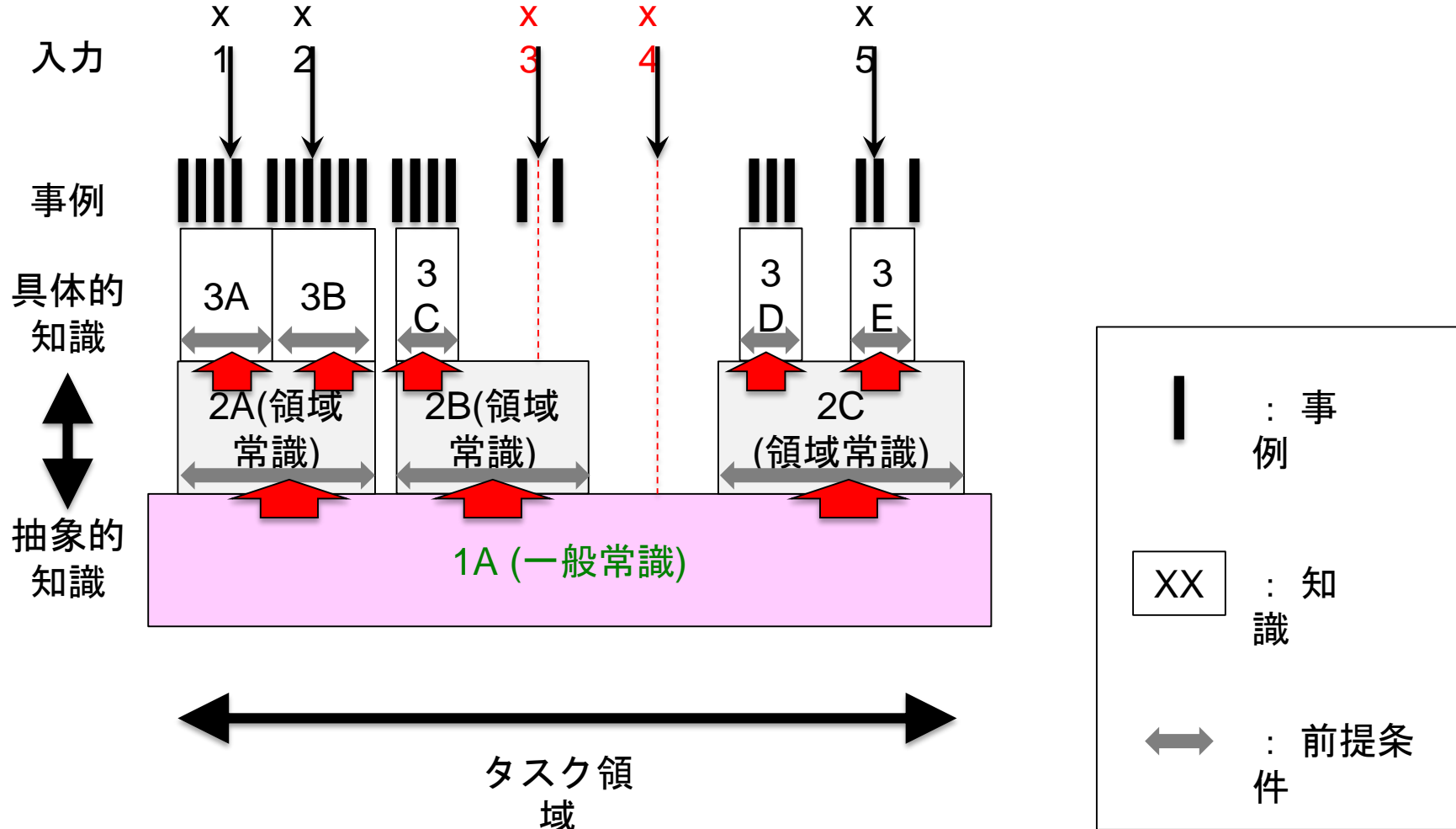
- 人間が知識を組み込む場合
 - 適用条件を書くのは大変そう。
 - フレーム問題？
- 非記号知識の条件部分をどう組み込むかが問題

非記号(分散表現)の知識

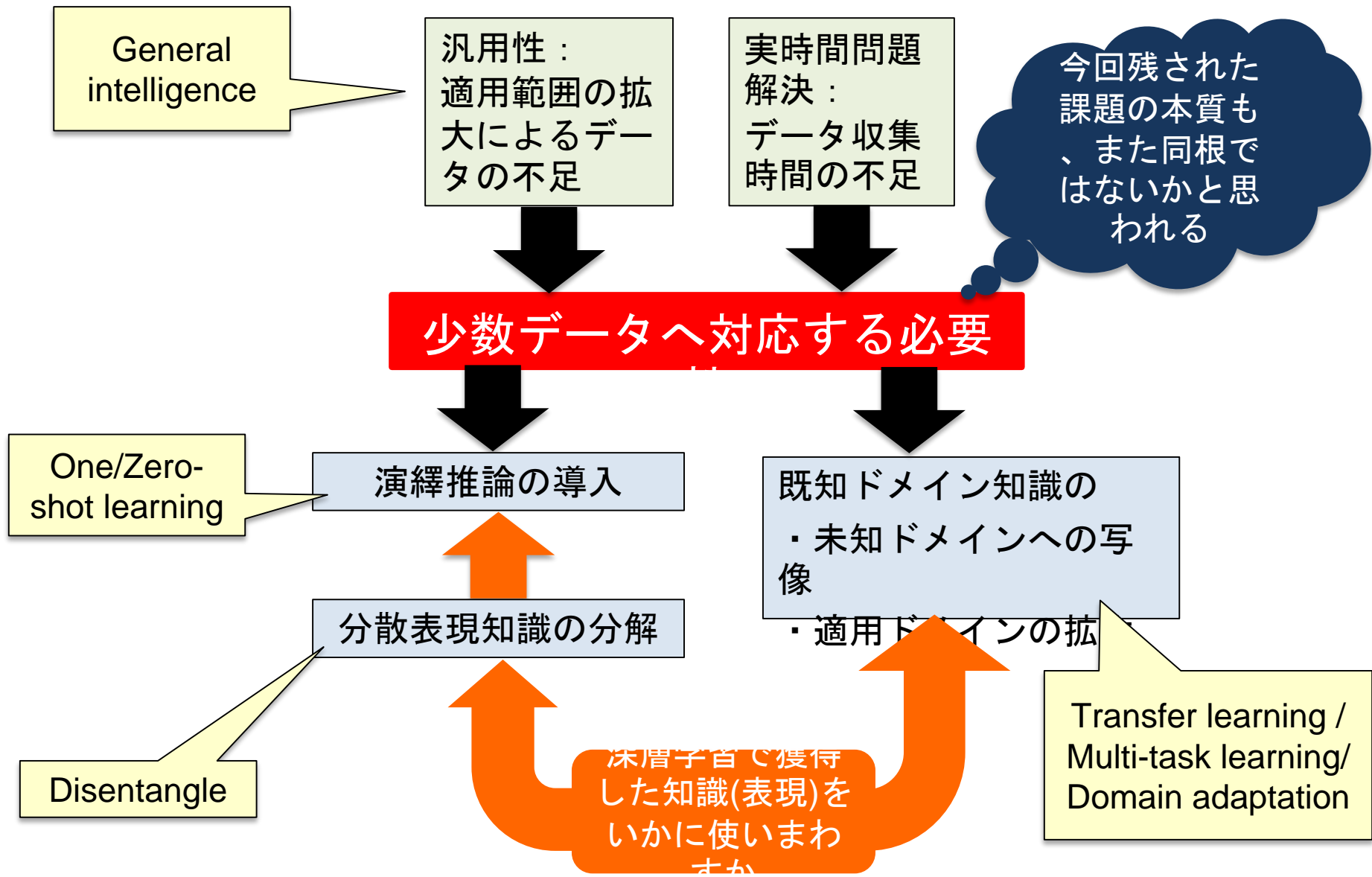
- データが十分な機械学習(帰納推論)では、条件部も組み込まれる。
 - だから、一般物体認識などが可能になる
- データが密に存在する範囲を超えて拡大することは基本的に難しい

「前提条件階層」を構築できるか

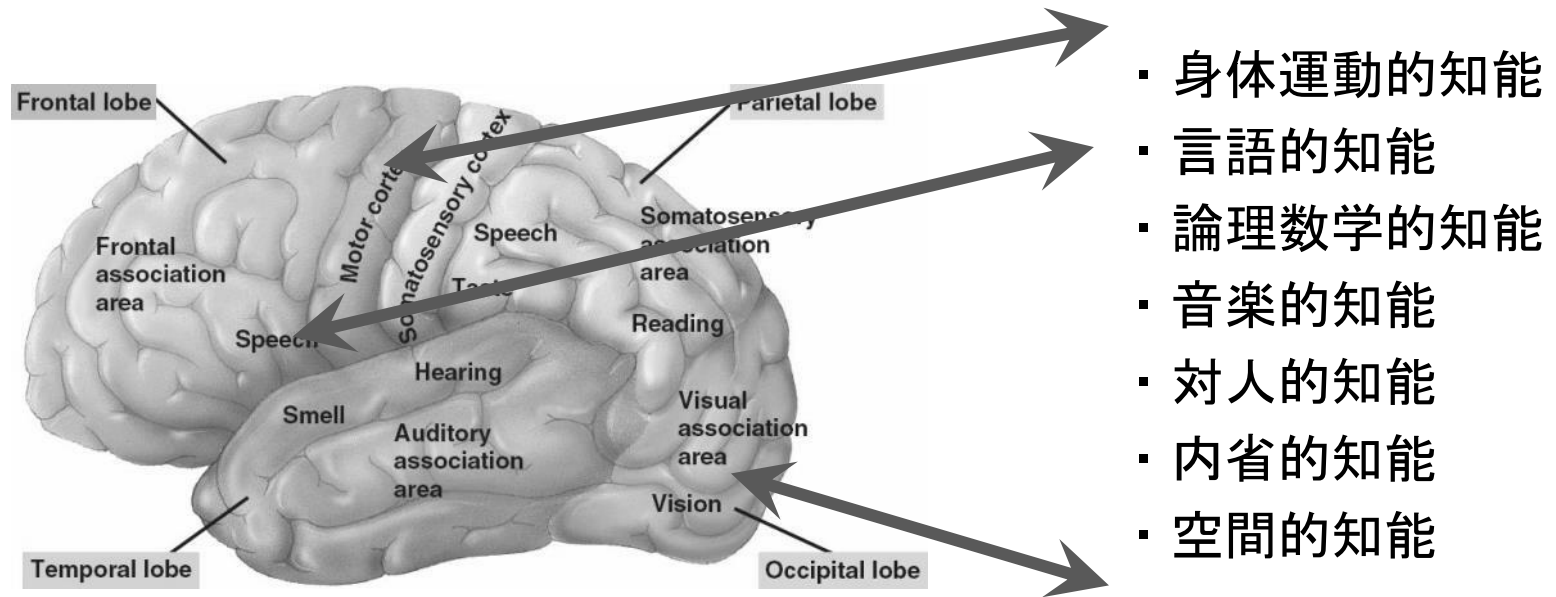
前提条件(知識内容を敷衍しうる範囲)が大量データからしか決まらないのでは、
帰納推論の域を出ない、より汎用性の高い知識から階層的に制御したい。



DL後のAIに残された汎用性という課題の本質



汎用性を支える大脳新皮質の一様性



図の出典：

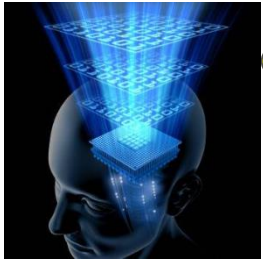
http://bio1152.nicerweb.com/Locked/media/ch48/48_27HumanCerebralCortex.jpg

一様な機構で、多様な機能を獲得

一> 知能の汎用性を支えている

- ・ 大脳新皮質において実現されている
- ・ 深層学習で実現された表現獲得

生物に学ぶ研究を俯瞰する観点から



計算論的神経科学

- 高齢者転倒原因解明
- 触覚を用いて人を抱き上げるロボット

統合技術型BIC

生物に学ぶ
総合技術

実例がない？

生物に活かす応用(医療等)

[理学] (知ることに価値あり)

生物から離れた応用
[工学] (性能に価値あり)

医療工学などをはじめとして多くの応用がある。

生物に学ぶ
要素技術

望む機能から
微細な機構への
絞り込みが重要

要素還元的BIC

②

①

③

④

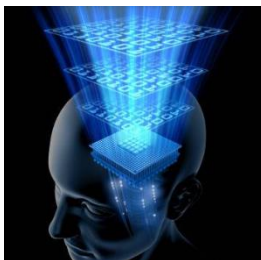
- 視覚情報処理に学んだCNN
- 鼻に学ぶ静音パンタグラフ
- ハスの葉に学ぶ超撥水
- **海馬のシータ位相歳差から情報表現を学んだ**

フクロウに学ぶ静音パンタグラフ

- 鳥の中でフクロウが一番静かに飛ぶことを友人から聞き、フクロウに独特な風切り羽根にあるノコギリ状のギザギザが空気を拡散し、静かな飛行を実現することを発見しました。
- 新幹線のパンタグラフに風切り羽根をまねたギザギザをつけ、30%の騒音削減に成功



生物に学ぶ研究を俯瞰する観点から



計算論的神経科学

- 高齢者転倒原因解明
- 触覚を用いて人を抱き上げるロボット

統合技術型BIC

脳が足枷になり工学システムとしての応用には不適

生物に学ぶ総合技術

要素が多いと機構と機能の対応がつかず、**新計算原理の開拓は期待薄**

生物に活かす応用(医療等)
[理学] (知ることに価値あり)

生物から離れた応用
[工学] (性能に価値あり)

②

①

③

④

医療工学などをはじめとして多くの応用がある。

生物に学ぶ要素技術

- 視覚情報処理に学んだCNN
- 梟に学ぶ静音パンタグラフ
- ハスの葉に学ぶ超撥水
- **海馬のシータ位相歳差から情報表現を学んだ**

フクロウに学ぶ静音パンタグラフ

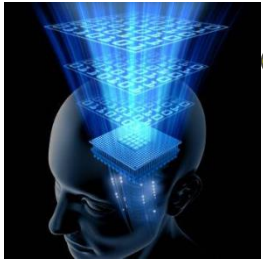
- 鳥の中でフクロウが一番静かに飛ぶことを友人から聞き、クロウに独特な風切り羽根にあるノコギリ状のギザギザが空気を拡散し、静かな飛行を実現することを発見しました。
- 新幹線のパンタグラフに風切り羽根をまねたギザギザをつけ、30%の騒音削減に成功



望む機能から微細な機構への絞り込みが重要

要素還元的BIC

生物に学ぶ研究を俯瞰する観点から



計算論的神経科学

- 高齢者転倒原因解明
- 触覚を用いて人を抱き上げるロボット

生物に学ぶ総合技術

統合技術型BIC

脳が足枷になり工学システムとしての応用には不適

大規模システムの統合の指針

生物に活かす応用(医療等)

[理学] (知ることに価値あり)

生物から離れた応用
[工学] (性能に価値あり)

②

①

③

④

医療工学などをはじめとして多くの応用がある。

生物に学ぶ要素技術

望む機能から微細な機構への絞り込みが重要

要素還元的BIC

- 視覚情報処理に学んだCNN
- 梟に学ぶ静音パンタグラフ
- ハスの葉に学ぶ超撥水
- 海馬のシータ位相歳差から情報表現を学んだ

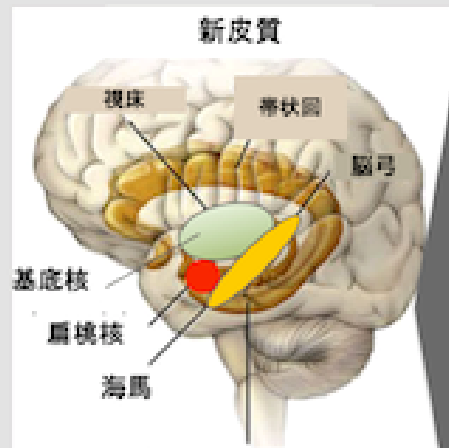
フクロウに学ぶ静音パンタグラフ

- 鳥の中でフクロウが一番静かに飛ぶことを友人から聞き、クロウに独特な風切り羽根にあるノコギリ状のギザギザが空気を拡散し、静かな飛行を実現することを発見しました。
- 新幹線のパンタグラフに風切り羽根をまねたギザギザをつけ、30%の騒音削減に成功



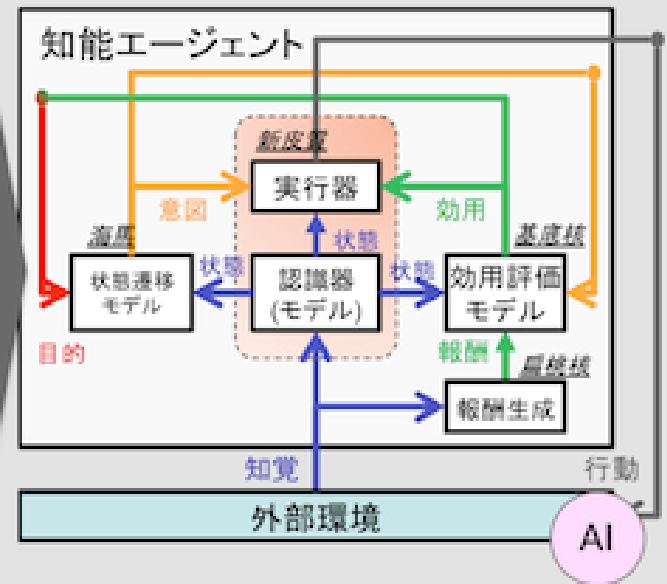
脳全体のアーキテクチャに学び 人間のような汎用人工知能を創る(工学)

全脳アーキテクチャ (WBA) のアプローチ



脳

- ① 脳の各器官を機械学習モジュールとして開発
- ② それらを統合した認知アーキテクチャを構築



基本方針：可能なかぎり粗いモデル化からはじめ、必要に応じて段階的に詳細化

脳はそれぞれよく定義された機能を持つ機械学習器が一定のやり方で組み合わされることで機能を実現しており、

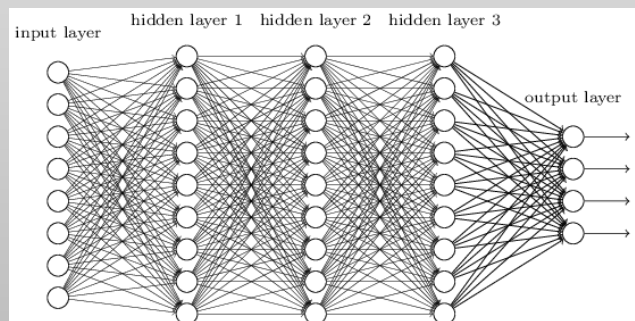
それを真似て人工的に構成された機械学習器を組み合わせることで人間並みかそれ以上の能力を持つ汎用の知能機械を構築可能である。

ようやく今，WBAアプローチが可能になった

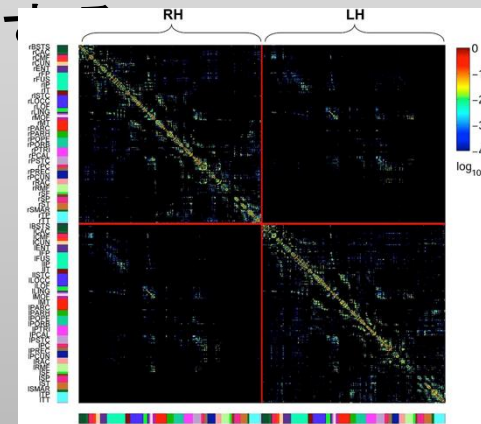
これまで難しかったが，近年可能になっ

全脳アーキテクチャ = 機械学習 + 認知アーキテクチャ

壁となっていた新皮質
のモデルとして**深層学
習**の研究が最近進んだ



脳全体のメゾスコピッ
クな**コネクトーム**を
アーキテクチャの参考
と



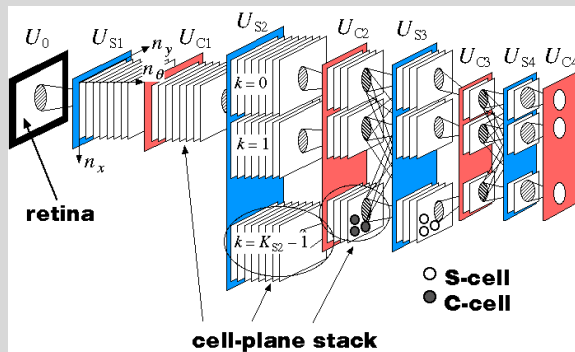
ネオコグニトロンを脳全体に拡大する

一般物体認識を実現した
ネオコグニトロン(視覚野)

脳に学んだことは何か？

複雑細胞と単純細胞から
構成される基本単位

階層的アーキテクチャに
よる繰り返し利用



画像出典元 : <http://www.aso.ecei.tohoku.ac.jp/~shun/RNC-e.html>

汎用人工知能を実現する
全脳アーキテクチャ

脳に何を学ぶべきか？

大脳新皮質の
標準学習アルゴリズム
認識だけでなく注意や運動も
含んだ学習の単位モジュール

新皮質内の6層構造を考
慮したコネクトーム
標準新皮質の単位モジュール
を結合したアーキテクチャ

「人類と調和した人工知能のある世界」を目指して

5.NPO法人全脳アーキテクチャ・イニシアティブとその狙い

NPO法人 全脳アーキテクチャ・イニシアティブ



代表 山川 宏



副代表 松尾 豊



副代表 高橋 恒一

A screenshot of the WBAI website homepage. At the top is the logo 'WBAI' with a stylized brain icon in the 'B'. Below the logo is the text 'Whole Brain Architecture Initiative'. The main heading reads 'NPO法人 全脳アーキテクチャ・イニシアチブ'. Below this is another 'WBAI' logo, social media icons for Facebook, Twitter, YouTube, Google+, and RSS, and a blue button that says '協力者として参加するには'. A navigation menu includes 'WBAIについて', '全脳アーキテクチャとは', 'イベント', '公表資料', and 'お問い合わせ'. A large image shows a lecture hall with an audience seated in rows, facing a stage where a speaker is presenting. At the bottom of the page is the slogan 'Let's build a brain together.' in a bold, sans-serif font.

WBAIのビジョンに向けた脳型AIの役割

ビジョン： 人類と調和した人工知能のある世界

最初期のAGIを全ての人々にとって優しいものにする
こと

- ③制御可能性
- ⑤安全保護
- ⑦倫理
- ⑧アカウンタビリティ

①透明性
②利用者支援

開発の加速

人間の知能と親和性が高い
(価値システム含む)

困った際のヒント

オープンな共創向き

統合のフレームワーク

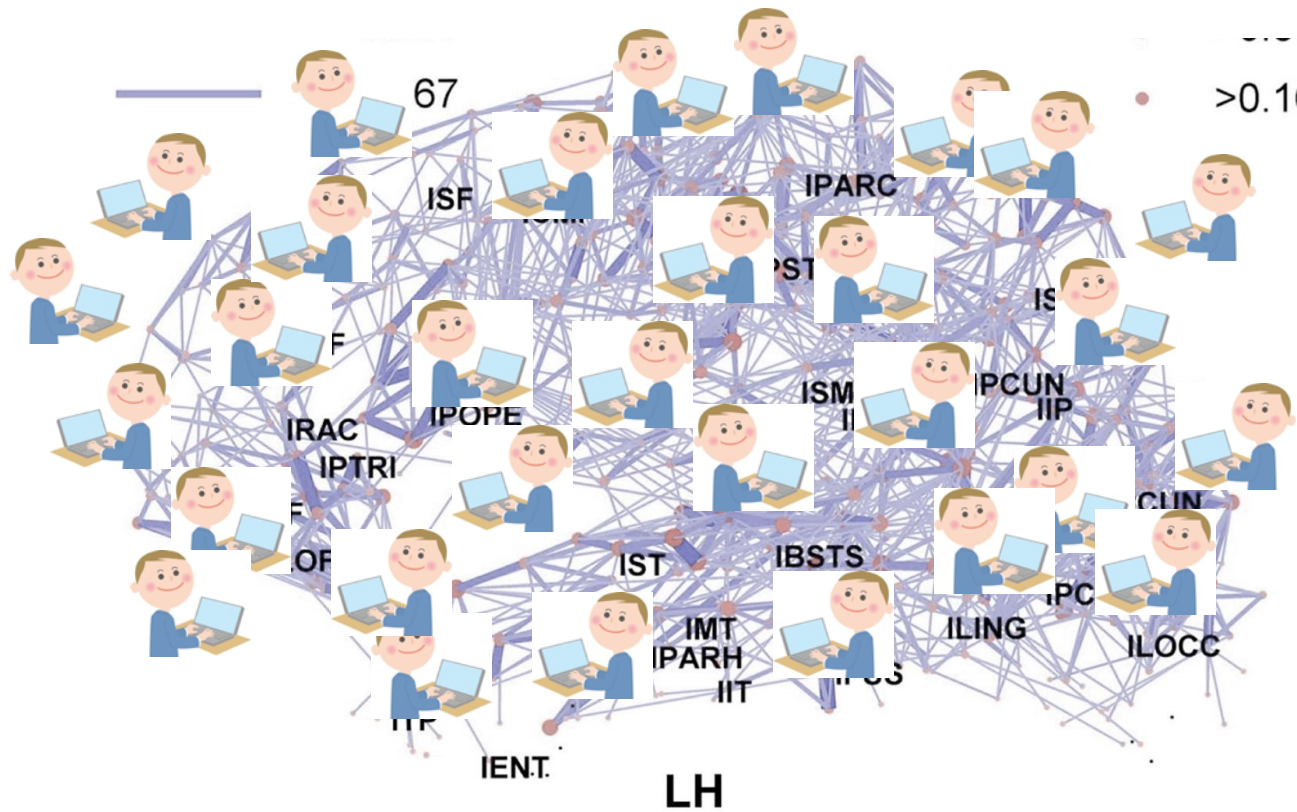
最終段階

脳型AIの一般的性質

脳アーキテクチャから得られる性質

人のようなAGIの開かれた共創を脳全体アーキテクチャー上で拡大

脳型アーキテクチャ上での共創開発とは



「ふと気が付くと、AGIが完成していた」
ということが起こりうる

全脳アーキテクチャ(WBA)勉強会の活況

- 目的：

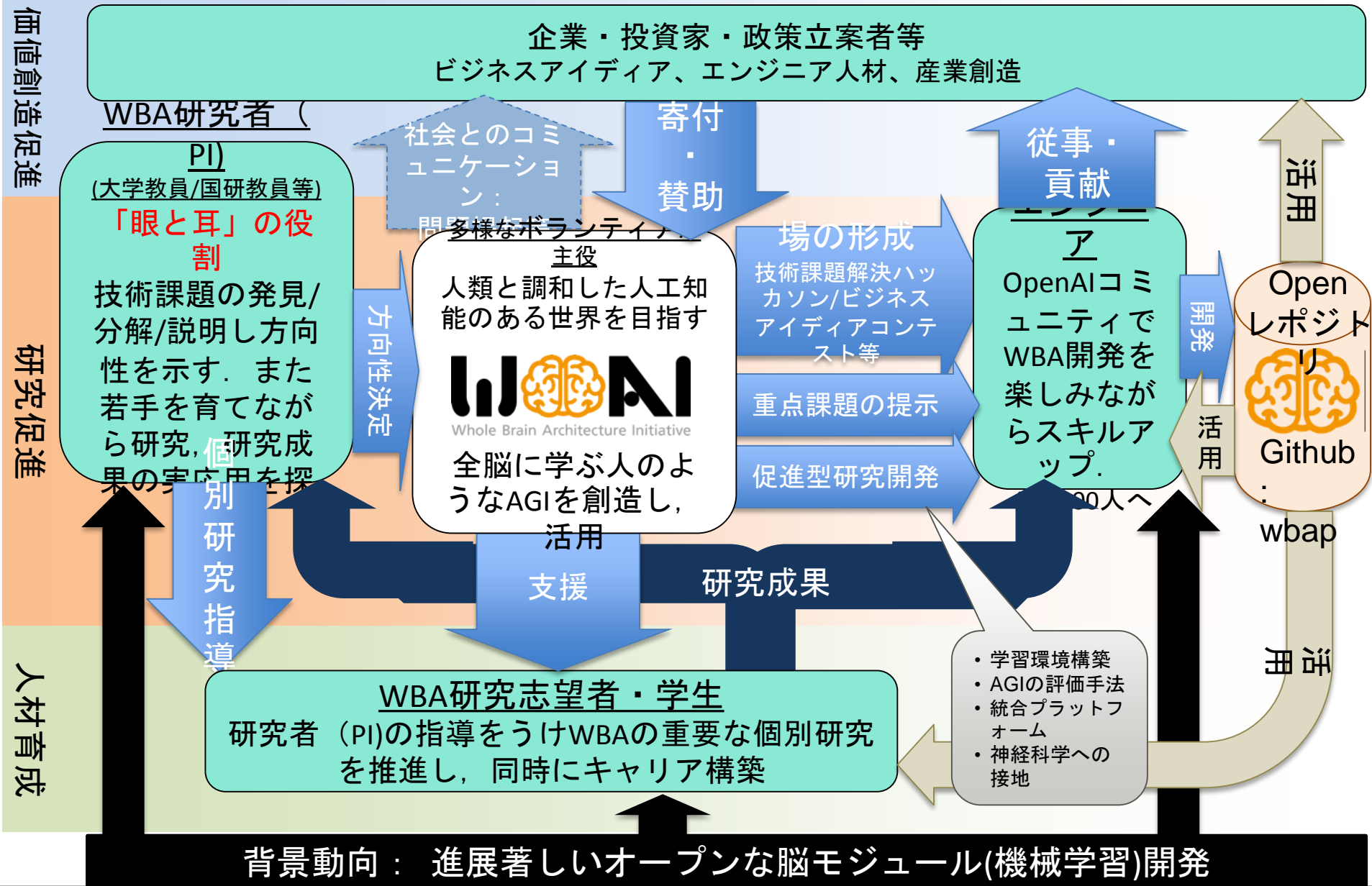
- AI, 神経科学, 認知科学等の複数分野に跨がる人材の結集と育成

- 経緯

- 2013年秋, 一杉 (産総研), 松尾 (東大), 山川 (ドワンゴ) により立ち上げ
- 2013年12月から2016年10月までに16回の勉強会を実施し毎回200名程度の参加.
- 2014年8月に, WBA若手の会を結成
- 関連するFacebookグループメンバーは3600人超 (2016年10月現在)

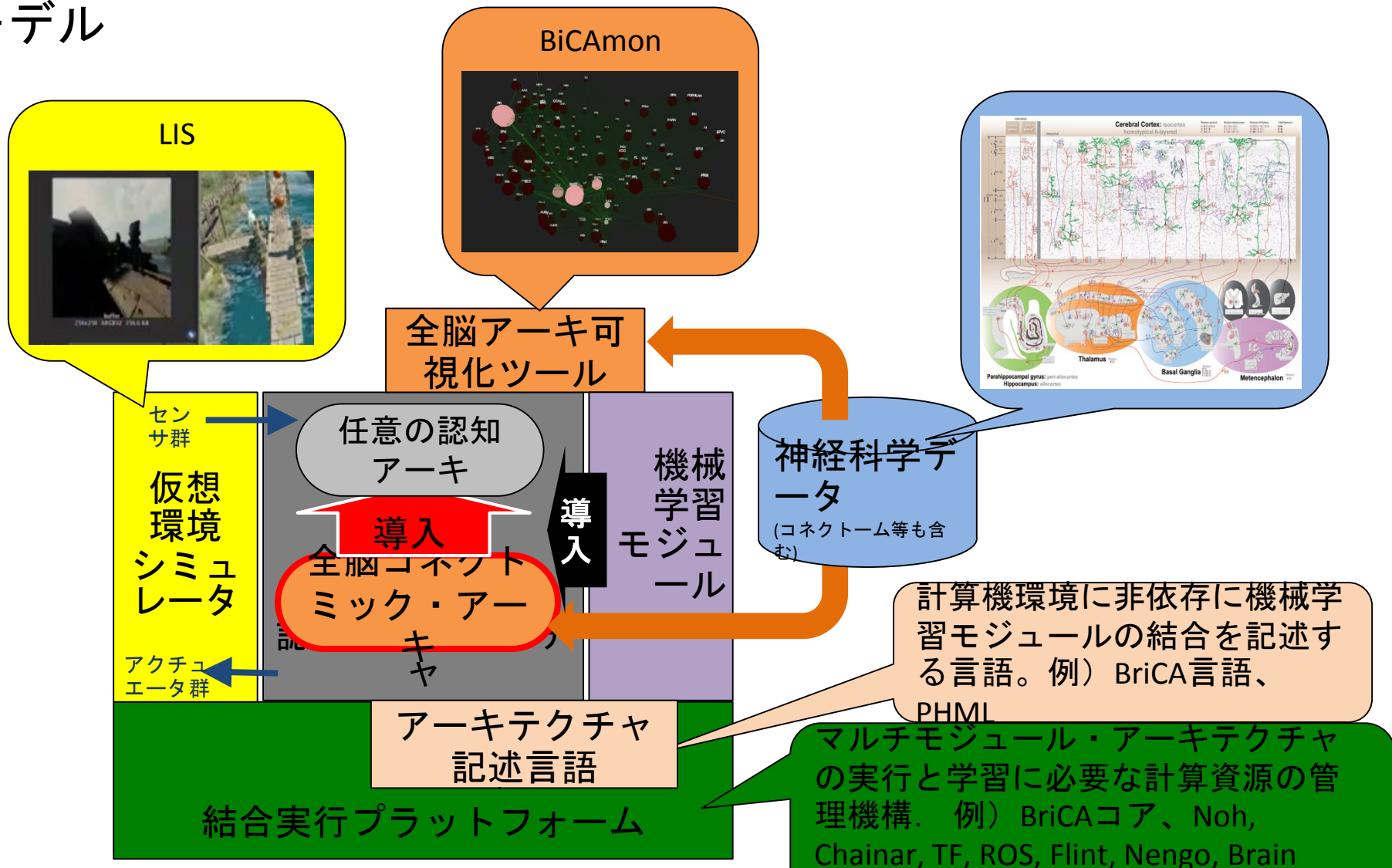


NPO法人WBAIの活動



全脳コネクトーム・アーキテクチャ(WBCA)の導入

WBCA=コネクトームをベースとした認知アーキテクチャの参照モデル



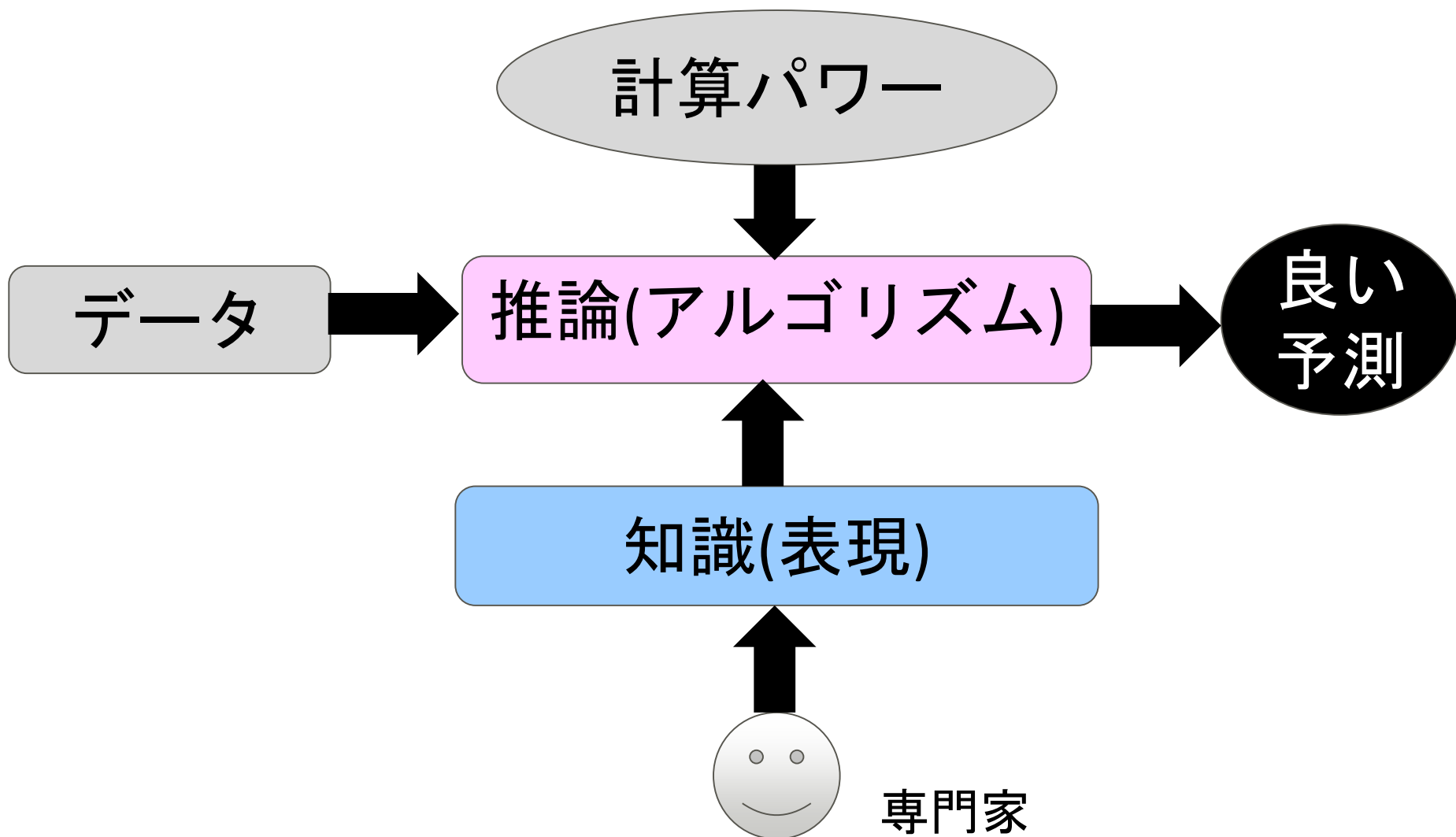
開かれた共創により全ての人々へAGIを

NPO法人WBAIの役割：共創開発基盤の構築



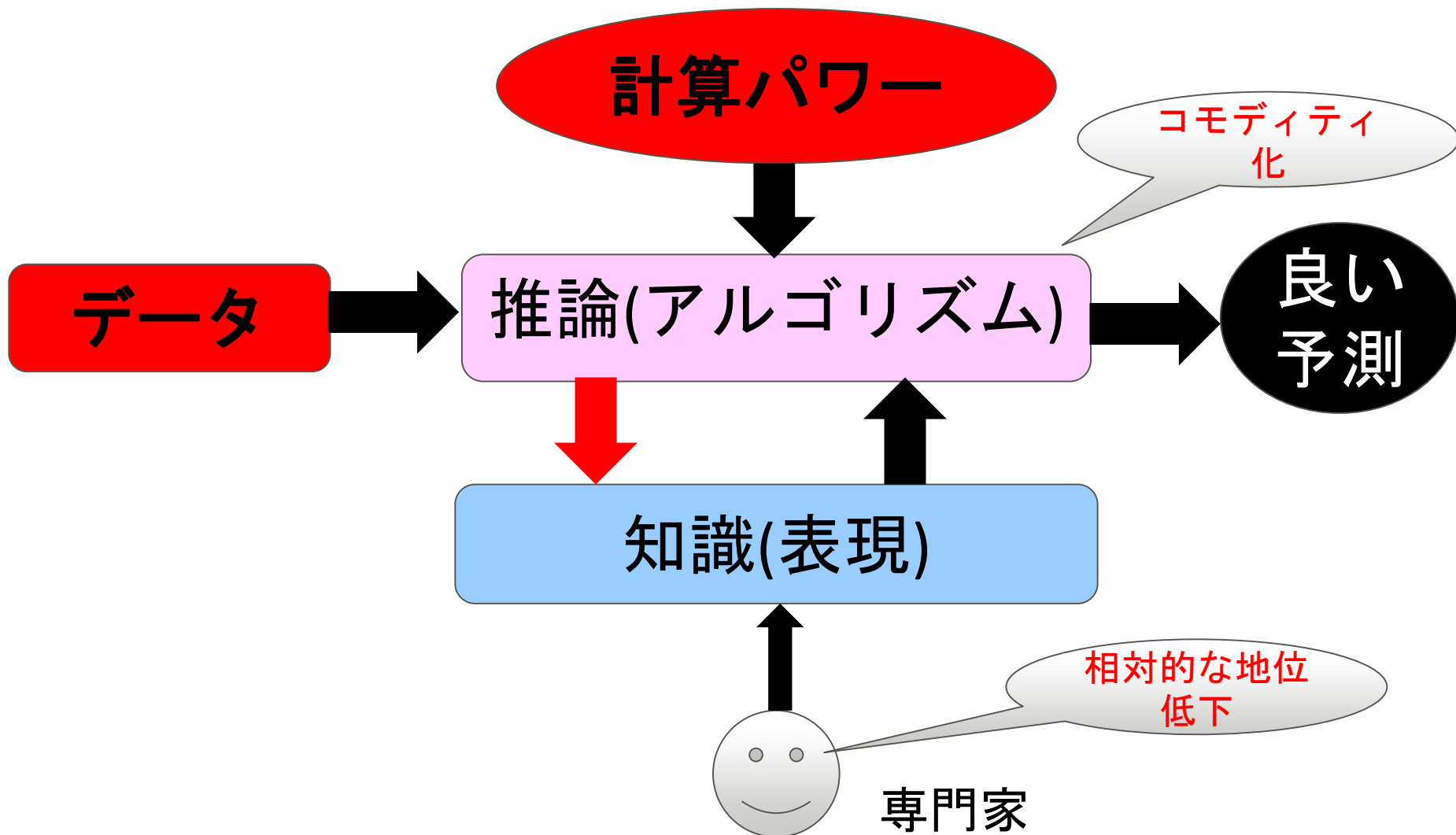
人工知能の全体像

知能 ≡ 直接観測できない出来事を予測する能力



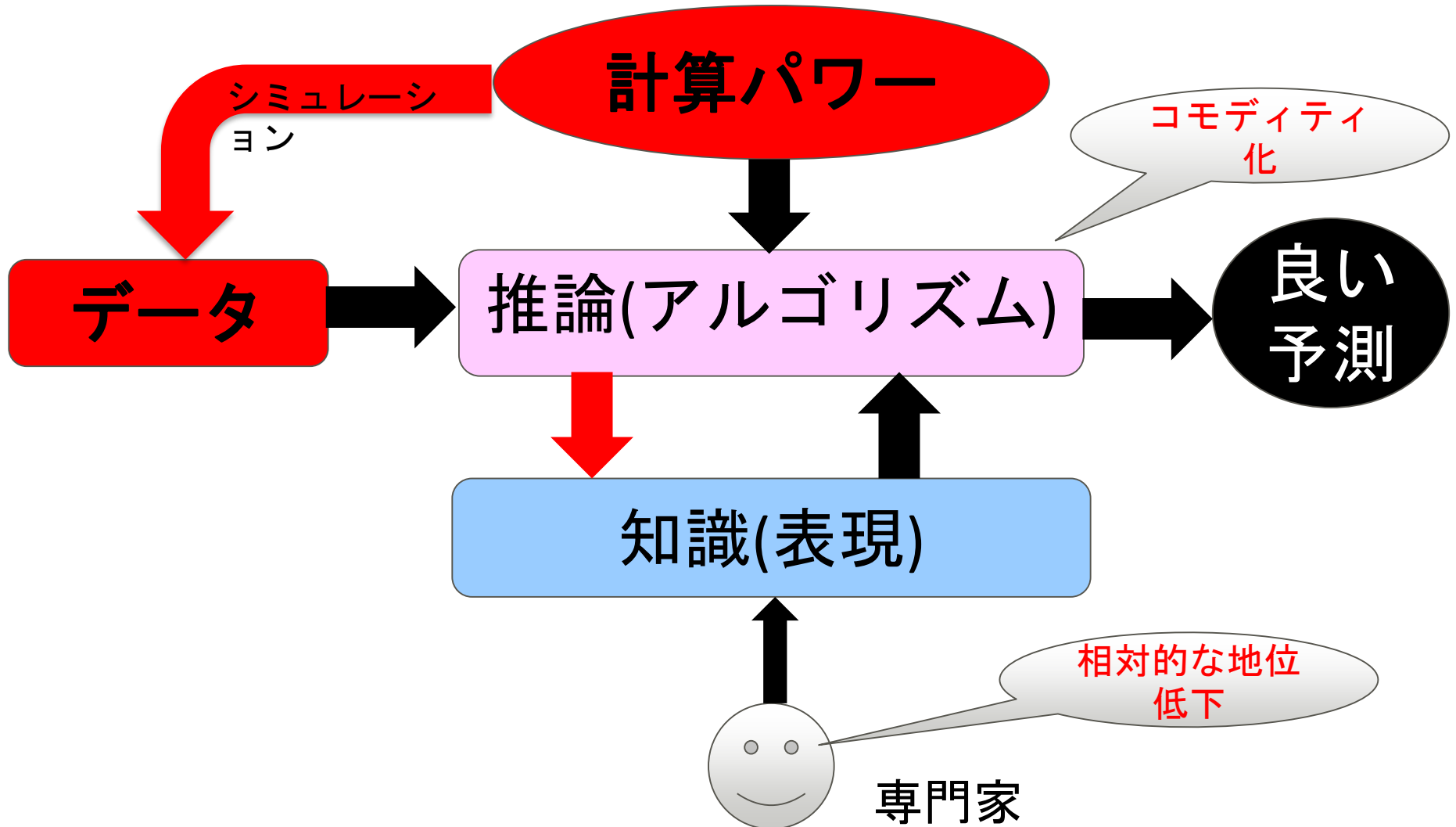
人工知能の全体像

知能 ≡ 直接観測できない出来事を予測する能力



人工知能の全体像

知能 ≡ 直接観測できない出来事を予測する能力



LIS (Life In Silico)が目指す「AGIの民主化」

AIが3歳まで育つ1000日間のためのシミュレ

学習環境シミュレータ (Unity)^{ータ}

無料化されたゲームエンジンで自由に世界を構築



機械学習 (人工知能)手法

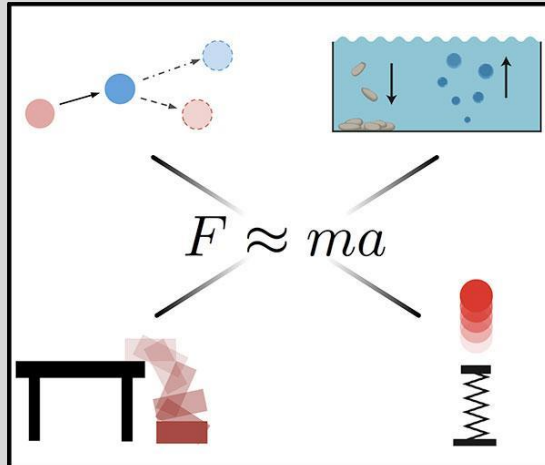


近年の深層学習手法を標準装備. 様々なAIが試せる

GitHubで開発中、Life in Silico(LIS)で検索
簡単に使えるようにして公開中



直感的物理学(Intuitive
sics)



The development of
intuitive physics@MIT

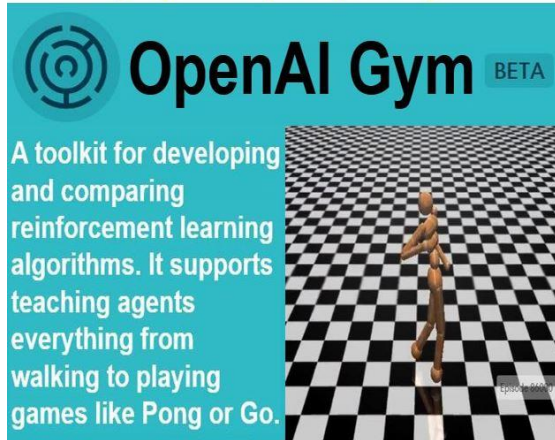
<https://goo.gl/aCnMtD>



【人工知能】 学習環境シミュ
レータLife in Silico(LIS) みんな
で人工生命を育てよう！

最近になり様々なプラットフォームが出現

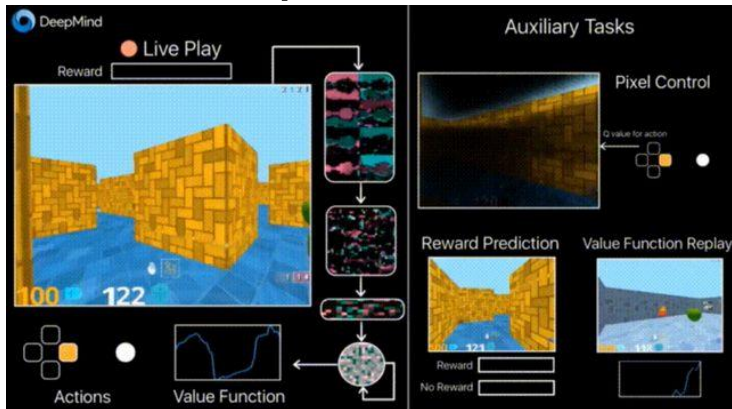
Developing & Comparing RL Algorithms



UNIVERS@OpenAI



DeepMind Lab



TorchCraft@Facebook



脳 全体のアーキテクチャに学び
人間のような汎用人工知能を創る



活動状況：

- Slack上での交流
- 隔週水曜にオフ会
遠隔者向けのビデオ会議あり

開発テーマ：

- 仮想学習環境の整備(Life in Silico など) → #lis
- 機械学習装置の開発評価 (Deep PredNet等)
- 面白い環境で行動学習するAIエージェントの実験
- 人のような認知機能を実現するAIエージェント
- 神経科学データの解析およびそのツールの開発
- その他様々なWBA研究の推進に関わる研究開発
(※ 関連して「Request for Research」を制作中)

AIで遊びたい方は、 <http://wba-initiative.org/sig-wba> へ

第二回WBAハッカソン(2016年10月8～10日)



神戸大学 遠隔オンラインライブ講義「計算生命科学の基礎3」

おわりに

- AIは、データが十分な状況では人並みの知能を発揮できるようになった（深層学習等による）
- 人のように多様なタスクを柔軟に獲得する汎用人工知能(AGI)の開発が現実味を帯びてきた
- 汎用性の課題は、データ不十分さへの対応。この突破にはデータから得られた知識を組み合わせる(演繹／アブダクション)が必要。ヒントを脳から得られるかもしれない。
- 人類と調和したAIがある世界を構築するには、最初期のAGIが皆に優しい世界公共物であって欲しい
- そのために、「人のようなAGIの開かれた共創を脳全体のアーキテクチャートで拡大する」ことは有効か

ご清聴ありがとうございました

理解し得ないものを創る

The Whole Brain Architecture Initiative videos



**The First WBA
Hackathon,
Sep. 2015**

<https://www.youtube.com/watch?v=0QS5Z3WrHSA>



Life in Silico(LIS)
Virtual environment
for AI learning

<https://www.youtube.com/watch?v=lkDdbBIgn5Q>



Whole Brain Architecture Initiative

The Whole Brain Architecture Initiative

a specified non-profit organization
<http://wba-initiative.org>

ものづくりの3アプローチ

試作

生物等
に学ぶ

知識に
基づく
設計

実用的なものづくり

試作



理解に基づく設計： 実現したい機能を分解して機構に割り当てる

生物に学ぶ研究：

試
作

生物等
に学ぶ

知識に
基づく
設計

行き詰まると、生き物などから
キーとなる要素技術のヒントを得てきた。

AI技術における表現と推論

	記号表現	分散表現
演繹推論 (知識活用)	古典的AI (1950年代～) 【大人の知能】	分散表現演算 Word2Vec等 (2010年代～)
帰納推論 (知識獲得)	パラメトリック機械学習、 帰納論理プログラミング等 (1960年代～)	ノンパラメトリック機械学習、 表現学習機能 (1990年代～) 【子供の知能】

分散表現知識をどのように演繹につなげるか？

	記号表現	分散表現
演繹推論 (知識活用)	古典的AI (1950年代～) 【大人の知能】	分散表現演算 Word2Vec等 (2010年代～)
帰納推論 (知識獲得)	パラメトリック機 械学習、 帰納論理プログラミ ング等 (1960年代～)	ノンパラメトリック 機械学習、 表現学習機能 (1990年代～) 【子供の知能】

① 記号接地

②

1. 分散表現知識を記号に対応付けた上で、記号により演繹推論等を行う。
2. 分散表現知識に対して演繹推論等を行う
3. 上記の両方の組合せる