

About the paper, entitled “**Cognition is All You Need**”
(公表前ポジションペーパー草稿1.1 2024年3月4日、コメント募集中)

by Nova Spivack, et al.

第106回AGI輪読会資料

2024年11月1日

担当：岡島義憲

Authors ;

(参) <https://www.mindcorp.ai>

Nova Spivack ※, Sam Douglas※, Michelle Crammes※, Tim Connors ※,



※) Mindcorp, Inc ;
an enterprise superintelligence company powering the next-
evolution of research, analysis and strategy for F1000 customers

構成

1. Introduction	2
2. Related Research	5
3. Defining Conversational AI	8
4. Intelligence Versus Cognition	12
5. Instincts Versus Abstract Reasoning	13
6. Defining Cognitive AI	14
7. Cognitive AI Functional Architecture	15
7-1. Functional Requirements for Cognitive AI	15
7-2. Dual-Layer Architecture	17
7-3. Large Language Models	19
7-4. Cognitive Agents	20
7-6. Relationship Management, Inter-Agent Messaging and Dialogs	24
7-7. Planning and Project Management	26
7-8. Neuro-Symbolic Reasoning	28
7-9. Memory Retrieval and Context Management	29
7-10. Knowledge Discovery and Knowledge Management	30
7-11. Tool-Utilization	32

(続)構成

7-12. Mathematics and Computation	33
7-13. Multi-Agent Collaboration	34
7-14. Meta-Cognition	36
7-15. Self-Improvement	37
8. Comparison of Conversational AI to Cognitive AI	40
9. Limits of Cognitive AI	45
10. Cognitive AI in the Evolutionary Ladder of Intelligence	46
11. Exponential Intelligence	49
12. Implications	50
13. Crossing the Chasm	51
13-1. Early Adopters: Niche Applications and Proof of Concept	52
13-2. Early Majority: Crossing the Cognitive Chasm	53
14. Re-evaluating Current AI Approaches	53
14-1. LLMs as a Commodity	55
15. Commercial Cognitive AI	55
16. Conclusions	56

概要

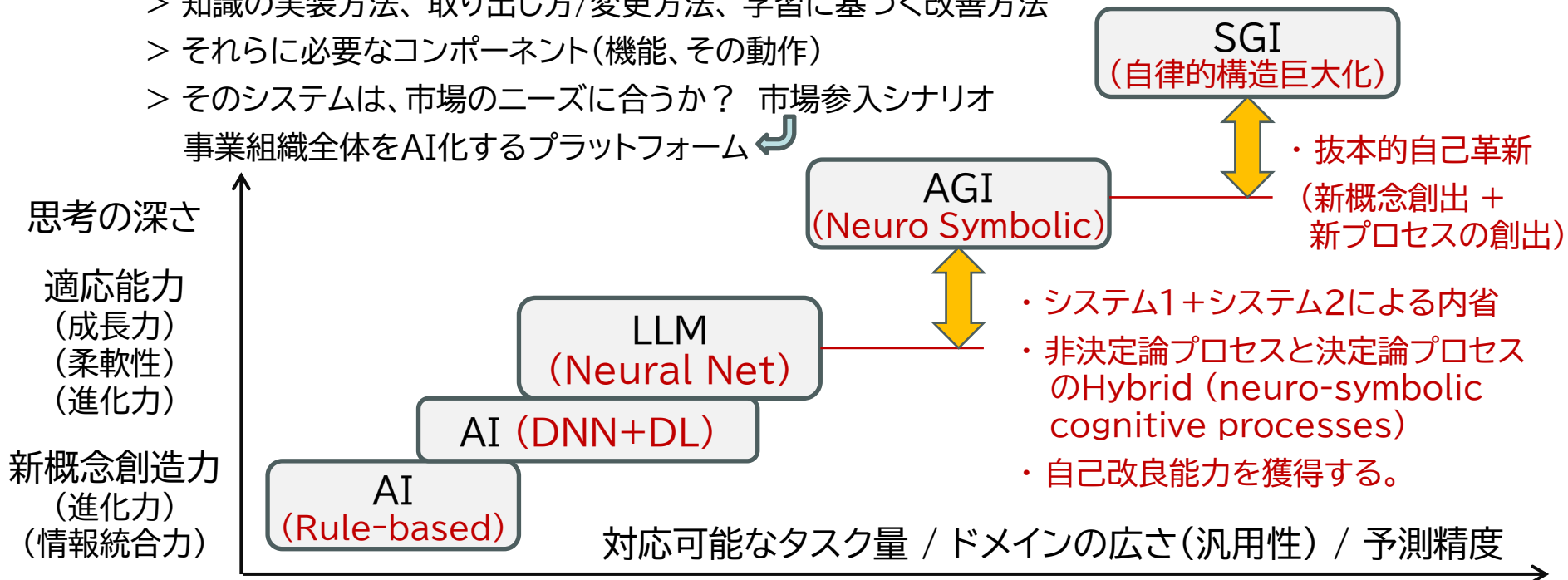
- AIの新アーキテクチャー（認知型AI、Cognitive AI、CogAI）の議論が増えている。
- 本稿の提案：
LLM層(会話的AI層) の上位に、ニューロ・シンボリック推論用の認知Agents層を備え、「LLMの確率的論理 と 認知Agentsの決定論的論理 の 両者の対話」にて、深い知的(認知)プロセスを進め、更に、自律的に進化するAIのアーキテクチャー：LLM+(Multi-Agents)+制御機構
- 各章の内容：
 - 1) LLMの問題点（ビジネス面以外）
 - 2) 提案する認知型AIの構造と機能と動作
 - > 構造(アーキテクチャー) と Components/Module
 - > Reasoning動作 (WorkFlow、Orchestration)
(Framework、Multi-Agent Collaboration、ニューロ・シンボリック推論)
 - > Realtime学習動作 (各種Management、Improvement)
(Relationship Management、Knowledge-Management、Project Management、Self-Improvement、等)
 - 3) AI開発史における 本認知型AI の位置付け と Super-Intelligence構築のシナリオ
 - 4) ビジネスから見た LLMの問題点 と、提案する認知型AIの必要性

人間/Agentsの内部対話により集合知を進化させるインフラ(Platform)としてのAI

(マービン・ミンスキーの「人間-AI協働作業プラットフォーム(Society of Mind)」に似る。本論文P-34)

- ・ システムが、出力前に「出力内容はこれで適切であるか？ これを出力すると環境に悪影響を及ぼさないか？」と問い、評価/再考し、論理を修正するには、深い「メタ認知+内省」が必要
- ・ そのために、LLMが実現する非決定論的メカニズム(システム1)に加え、決定論的メカニズム(システム2)を実装させ、両者のインタラクション(内省)のWork-Flowをプログラムする必要がある
- ・ そのように考えると、以下の課題も派生する
 - > それら必要なWork-Flow群のプログラム方法、学習(経験)に基づくWork-Flow群の改善方法
 - > システム1の推論 と システム2の論理 の 対話の設定、進め方
 - > 知識の実装方法、取り出し方/変更方法、学習に基づく改善方法
 - > それらに必要なコンポーネント(機能、その動作)
 - > そのシステムは、市場のニーズに合うか？ 市場参入シナリオ

事業組織全体をAI化するプラットフォーム



1. Introduction (P2~5)

- ・課題：会話型AI (LLM)は深く考えることができない
 - ＞多段階問題の解決能力に問題ある。
 - ＞会話型AIは、推論への理解/洞察の質が低く、複雑な問題でエラーする。
 - ・本稿の提案：認知型AI (Cognitive AI/CogAI)のアーキテクチャと、機能と、ビジネス検討
 - ・実現方法
 - ＞LLMの外側(もしくは上位)に、決定論的論理で動作しうる「認知Agents群を持つ認知層」を備え、LLMと認知層の対話にて、各種のニューロシンボリック推論を進め、タスクを進める。
(集团的/社会的に、Multi-Agents間で作業を進める「企業内組織」のように振る舞う)
 - ＞各Agentは、Management機構を備え、自律的に知識/認知能力/会話相手を改善する。
 - ・展望：
 - ＞今後、認知層への開発投資は、会話層への投資よりも効率的となる。
(AGIは確率的アプローチだけでは達成できない)
 - ＞認知型AIが自己増殖できるようになると、指数関数的知性(EI)に発展し、
Super-Intelligenceに発展する可能性を持つ。([43] Thórisson,K.R.(2021))
- (参) OpenAI社の“o1モデル”、“Large Reasoning Model LRM”と、本稿の認知型AIの発想は似るのかもしれない。([45] Valmeekam, K, et, al. (2022))

References (抜粋)

- [2] 意識が行うシステム2タスクに取り組むための新AIアーキテクチャとその学習フレームワークの開発
Bengio, Yoshua (Dec 11, 2019); "From System 1 Deep Learning to System 2 Deep Learning"
- [5] 認知AIの①推論と学習がどのように相互作用するか?、②人間-機械相互作用の改善、等
Bundy, A, et, al. (2023); "Introduction to 'Cognitive artificial intelligence'"
- [12] LLMに基づくマルチエージェントのProfiling / Agent間の会話 / 能力成長のための学習
Guo, Taicheng, et, al. (2024); "LLM based Multi-Agents: A Survey of Progress and Challenges"
- [14] LLMが訓練段階でエンコードした知識を抽出するツール(BERTese)を提案
Haviv, A, Berant, J., & Globerson, A. (2021); " Bertese: learning to speak to bert"
- [17] 累積学習を目的とする認知アーキテクチャ; System-1とSystem-2が共通の神経記号情報を利用
Latapie, H, et, al. (2022); "Neuro-symbolic Systems of Perception and Cognition: The Role of Attention"
- [19] Joint-Embedding Predictive Architecture (JEPA); 微分可能/訓練可能な認知アーキテクチャ
LeCun, Y. (2022, June 27); "A path towards autonomous machine intelligence (Version 0.9.2)"
- [22] LLMに構造化知識を注入する方法 (T5モデルを直接訓練)
Moiseev, F, et, al. (2022); "Skill: structured knowledge infusion for large language models"
- [30] MoE (Mixture of Experts); アテンションと互換性のあるアーキテクチャ・メカニズム
Minaee, S, Mikolov, et, al. (2024); "Large Language Models: A Survey"
- [31] MoE (Mixture-of-experts) モデル / Knowledge-in-Context (KiC) アーキテクチャ
Pan, X, et, al. (2022); "Knowledge-in-context: towards knowledgeable semi-parametric language models"
- [36] 「認識」と「理解」は違う。「理解」とは、セマンティクス(意味)を特定することである。
Singer, Gadi (Jan 31, 2018); " Toward truly intelligent AI: From 'Recognition' to 'Understanding'"
- [45] OpenAIのo1モデル(Strawberry)は、LLMを超えるLarge Reasoning Model、LRM)である。
Valmeekam, K, et, al. (2022); "Large Language Models Still Can't Plan "
- [51] 「大きなタスク」パラダイムのための「小さなデータ」(の提案)
Zhu, Y, et, al. (2020); "Dark, beyond deep: A paradigm shift to cognitive AI with humanlike common sense"

2. Related Research : 有用性の面でのLLM型AIの限界、改良すべき点

- ・ **会話型AIは認知処理を伴わないため「洞察の質」が低く、知識労働の質は向上させない**
(知識業務の生産性(速度、効率、スキル)を向上させるが、質(レベル)を上げる訳ではない^{[27][32]})

- > LLMは論理的推論に限界ある(Nguyen, 2023)
- > 多段階の推論能力が欠如している(Tongshuang et al., 2021年)
- > 複雑な推論タスクの実行に難がある(Shojaee-Mend, 2023年)
- > 透明性と説明可能性が欠如、推論理由の把握が困難(Liao et al., 2023)
- > 監査できない(Mökander et al., 2023; Thórisson, 2021)
- > 多様な知識を(組み合わせて)活用できない(Porada et al.2019)
- > 新しい例に対して構成的汎化を行えない(Yang et al., 2022)
- > 証明を生成する能力が低い(Yang et al.2022)
- > 新情報に十分適応するにはモデル全体の再トレーニング必要(Pan et al.2022)
- > 知識の導入ではなく、テキストデータを統計学習していることが欠点(Sun, 2021)
- > 人間的な共感や微妙な理解を必要とする複雑な問題が苦手

LLMは、大量のパラメータ群に知識を非構造的にエンコードしていることが問題^[14]

- ・ **LLM(だけ)で、知識労働のニーズを満たすことは現実的にも理論的にも不可能である**
- ・ **認知AIと人間の活動を統合し、知的労働の「質」を向上させるフレームワークが必要**
 - > 社会的協働作業においては、認知プロセスが最も重要なメカニズム (P-34)
 - > LLMの欠点に効果的に対処しながら、AIの長所を活用

各章で言及する「LLMの問題点」

- 1章 : 多段階問題の解決能力には限界がある。推論への理解/洞察の質が低く、複雑な問題でエラーする。
 - > 思考方法を学習しない。…… 思考しない/思考できない。
 - > パターン認識に基づいた本能的知性（非決定論的プロセスのみ）
 - > 非適応的（他のエージェントと相互作用しても、推論の仕方を変化させない）
 - > Embeddingを行っても、学習や思考を行うわけではない。
- 2章 : 会話型AIは認知処理を伴わないため「洞察の質」が低く、知識労働の質は向上させない
- 7-6章 : 意味のある計画を生成できない。検証も保証もできない。（以前に見た計画を借用しているだけ）
- 7-7章 : 幻覚や予測不可能な出力を起こしやすく、与えられた計画に常に従うとは限らない。
 - > 反転の呪い（「AはBである」と言われた場合、「BはAである」と推論できない）
 - > 3つの集合の論理和のような単純な集合操作ができない。
- 7-8章 : 意味を理解したり、その対話を元に思考する面で限界がある。
 - > (ソースの違う情報の)文脈を理解して文章生成できない。
 - 進行中の対話を元に思考する能力が低い。
- 7-9章 : Knowledge-Management (KM) のための明示的な構造を持たない。
- 7-10章 : 外部Tool (の利用)をLLMは認知していない。自己と他のシステム(自己のデータと外部のデータ)の区別が無い。結果、「対話」の概念も認識も無い。
- 7-12章 : 外部Agentと、協調/協働を目指したコミュニケーションは明確ではない。
- 7-14章 : 累積学習やReal-Time学習の能力が低い。
- 8章 : 訓練時のデータ分布に基づく推論パターンを再現するに留まる
 - > エラー/幻覚/予測不能な行動を起こしやすく、一貫性、信頼性に欠ける(ことがある)
 - > 学習後には、特定の知識を抽出/分離できなくなる
 - > 「知識」を、厳格に埋め込むことができない。従って、文脈を正確には理解しない
 - > 訓練後の新情報取り入れ(リアルタイムに旧情報と統合すること)に困難がある
 - > 再帰的操作、ツリー検索、カーブフィッティング、最適化などの高度なプロセスに関与できない。
- 13章 : LLMだけでは、市場から期待されているAI機能を実現できないだろう。

3. Defining Conversational AI (会話型AI)

- 会話型AIは、AgentとLLM間の対話に基づく人工知能とみなすことができる。
 - > Interface-AgentがLLMに送るトークン情報によって、LLMに蓄積した知識を引き出す。
(LLMは、確率モデルを使って、知識を圧縮保存している)

Figure 1. Token Streams

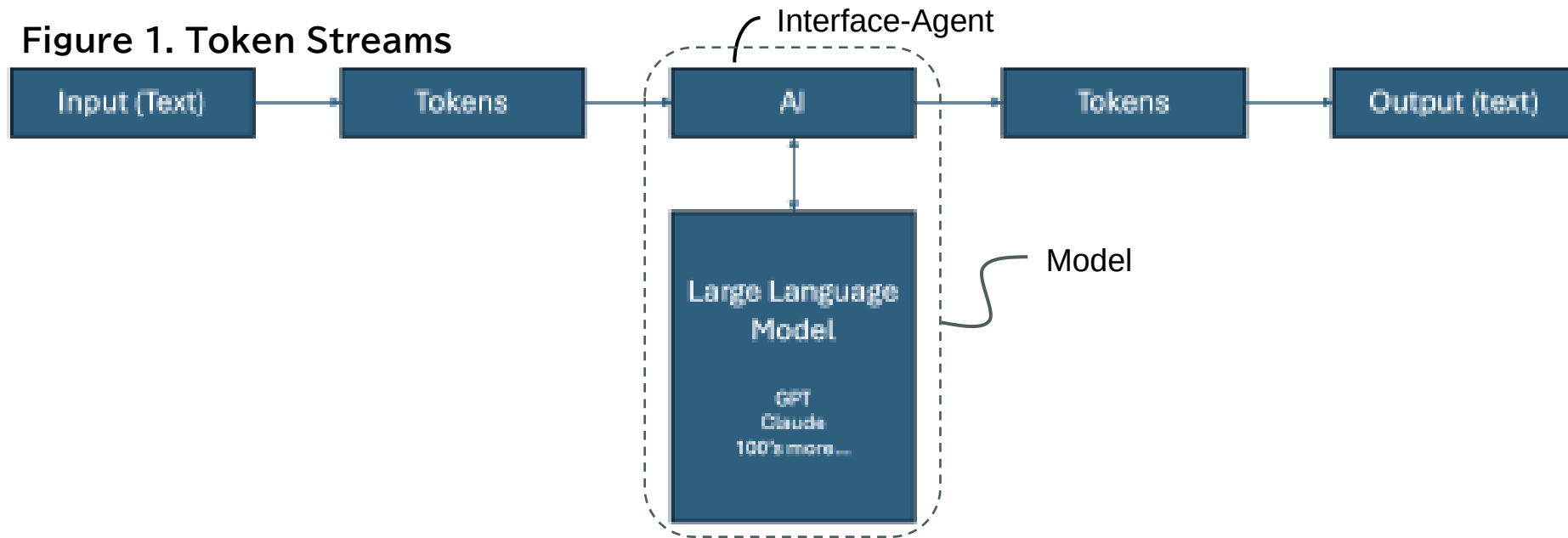
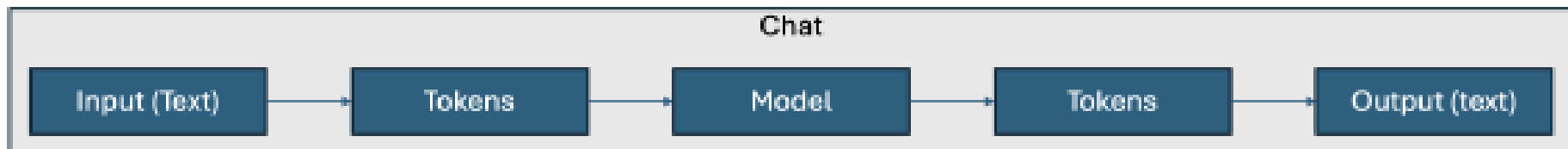


Figure 2. Chats アプローチ



3-1. 言語回路(linguistic circuit)の構成(#1)

- Chatbotエージェントと、ユーザー(人)の間で、言語回路を動作させ、LLMが保存する知識を引き出すケース。(Fig.3)

- > プロンプト入力を工夫することで、LLMの動作を誘導できる。
- > プロンプトとAgentとLLM間の会話は、「チャット記録」に保存される。

Fig.3 Chatbots

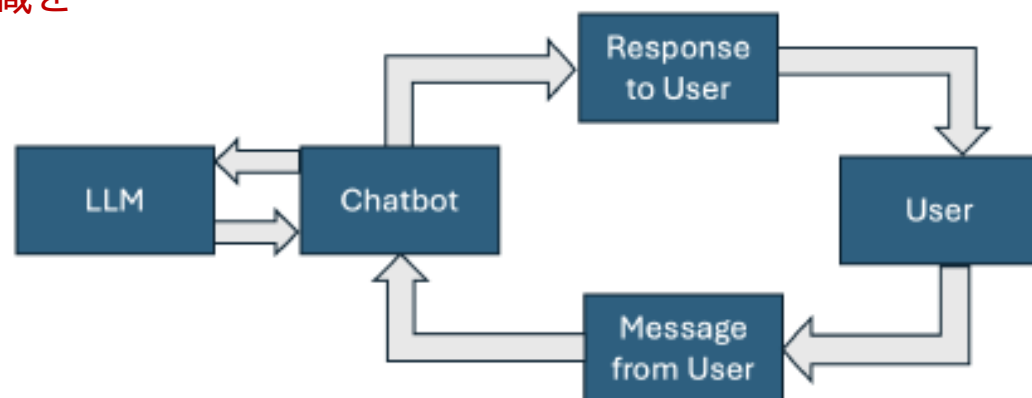
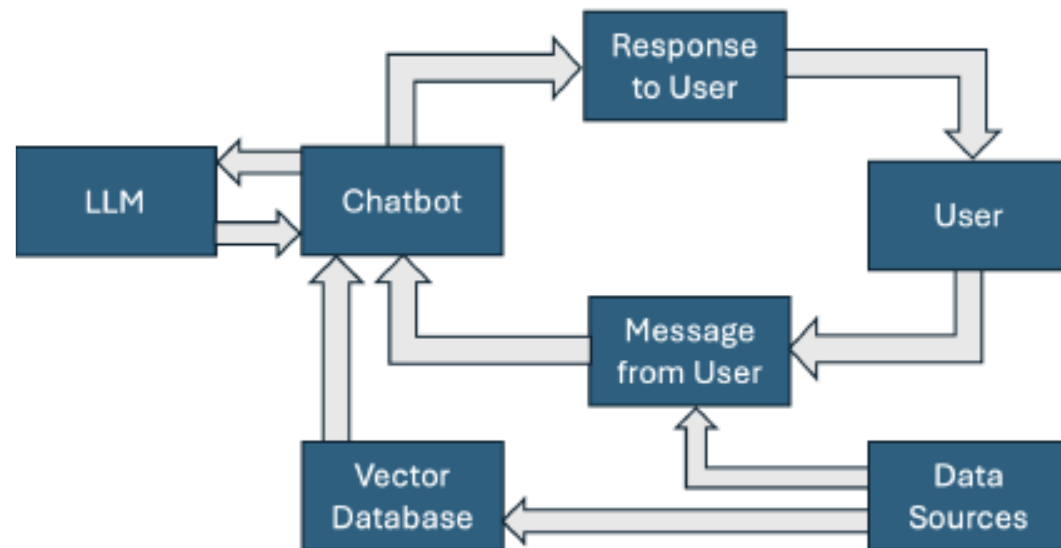


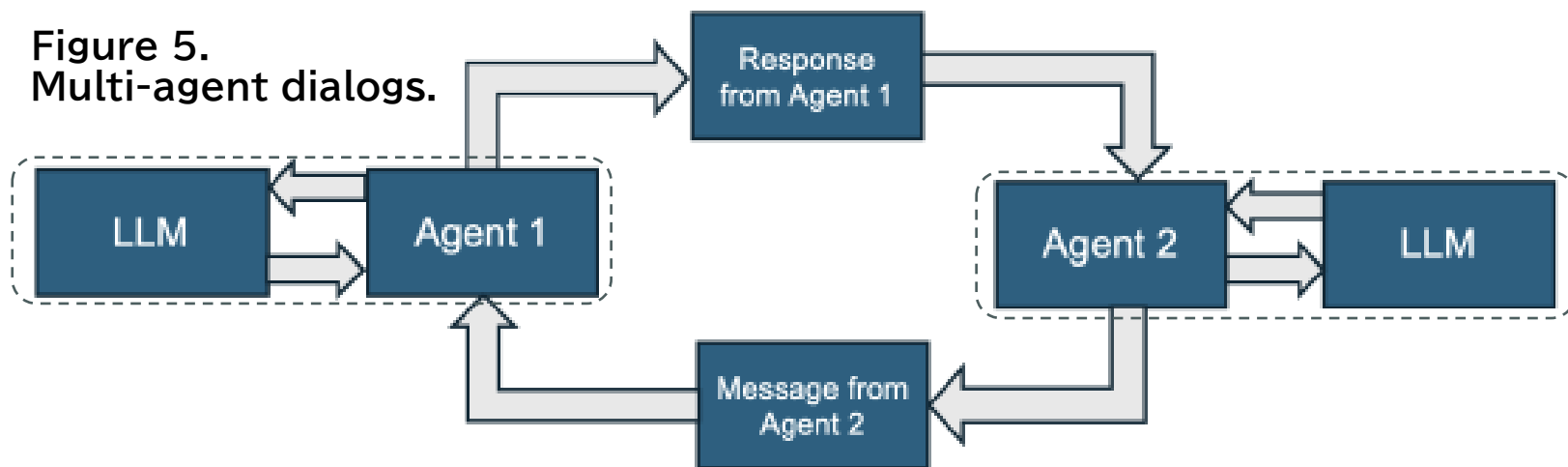
Fig.4 Retrieval Augmented Generation (RAG).

- 回路に他のコンポーネントを追加し、訓練時には無かった外部データを埋め込み、LLMの動作を誘導できる。(Fig.4)
- > その場合、ベクトル埋め込みとデータベースへのQueriesという形態で外部データにアクセスする。



3-1. 言語回路(linguistic circuit)の構成(#2)

- LLMを搭載した2個のAIの対話



- 会話型AIは、「第一段階の知能 (first-order intelligence)」とみなすべきである。^[40]
 - > 「チャットボットが、発言/対話の内容を理解している」と見なすのは誤解である。
 - > LLMの応答は、入力メッセージを元にした “probabilistic streams of guesses” であって、理解や論理展開(Reasoning)を伴わない。
 - > 従って、会話型AIは、複雑なCoTが必要なタスクには向かない(知能の深さが足りない)
 - > 会話AIは、非認知的知性の出力(the products of non-cognitive intelligence) に過ぎない。そのため、多くの本質的な限界がある。(4章、5章で説明)

4. Intelligence Versus Cognition (知能と認知の関係)

- ・ 知能(Intelligence)

- ：非ランダムな入力に対し、(因果関係のある) 非ランダムな出力を応答する機能

- ・ 認知(Cognition)

- ：回路の入力データの意味する所を理解し、応答(記憶、推論、思考、評価、判断、等)する能力
(学習と自己改善、感知、自己反省と内省、言語理解と処理、記憶と文脈管理、知識表現、知識管理、知識処理、研究と探索、推論、計画、意思決定、プロジェクト管理、タスク実行など)

- ・ メタ認知機能 : システムに、自己反省と適応の能力を与える機能 (Page-18, 19)

- > AIに戦略的思考や創造性を適用して問題を解決する能力を与える上で必要
(複雑な知識作業に従事する人間の心の認知機能を模する)
 - > システムの学習、推論、計画、意思決定等の能力(運用戦略)を継続的に改善するのに必要
(思考プロセスを内観し、自己評価する)
 - > 新たな情報や課題への動的な対応(adaptability)に必要

- ・ LLMは、認知機構を持たない。

- ：入力を元に「認知の産物」に似た出力を生成するかもしれないが、過去のトレーニング時の文章群を統計的にマネするだけである。

5. Instincts Versus Abstract Reasoning (P-13)

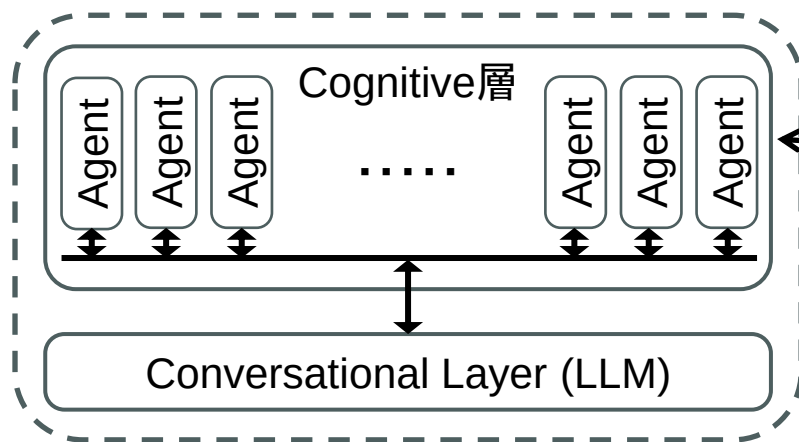
- LLM(会話型AI)は、パターン認識に基づいた本能的知性(非決定論プロセスのみ)
 - > 中間的な思考や推論が無い。
 - > 思考方法を学習しない。
 - > 能力(非決定論プロセス)はトレーニング段階で得たレベルに制限される。
 - > 非適応的 (他のエージェントと相互作用しても、推論の仕方は変化させない)
(Embeddingを行っても、学習や思考を行うわけではない)
- 認知型AIは、内省 (関連知識のOrchestrationによる自身の最適化) を行う
 - > 抽象的な推論と二次学習を行う機構(Module回路)を実装している
 - > 内部論理(複雑な内省)を持ち、相互作用から二次学習を行う。
 - > 他のエージェントと相互作用中にも、新知識の学習と適応のループを継続(並列動作)させる。
 - > 仮説を生成し、時間の経過とともに理解と応答を進化させる。
- 会話型AI から認知型AIへの変化は、人工知能技術の重要な転換となりうる。

[岡島私見] LLMは、分岐命令のように機能し得るので、認知層内には、非常に沢山のAgentが必要と思える。
「LLMは、Behaviorモデルだが、認知型AIはStructured-Modelであり、内部論理を持つ」
との言い方もできると思えます。
OpenAI社の新モデル”o1”は、“orchestration-1”の意味か？

6. Defining Cognitive AI (P-14)

- ・ 認知型AIは、自律的に「思考」と「神経記号的認知プロセス」を統合的に実行(Orchestration)
- ・ 認知型AIの目指す 自己改善/適応性を持った推論は、「知能の第2段階」といえる能力
 - 認知型AIは、LLMとは別のCognitive層にて、認知プロセスの推論と学習を行う。
(岡島私見：「LLMの推測によって、Cognitive Layerの認知プロセスを進める」とも言えると思う)

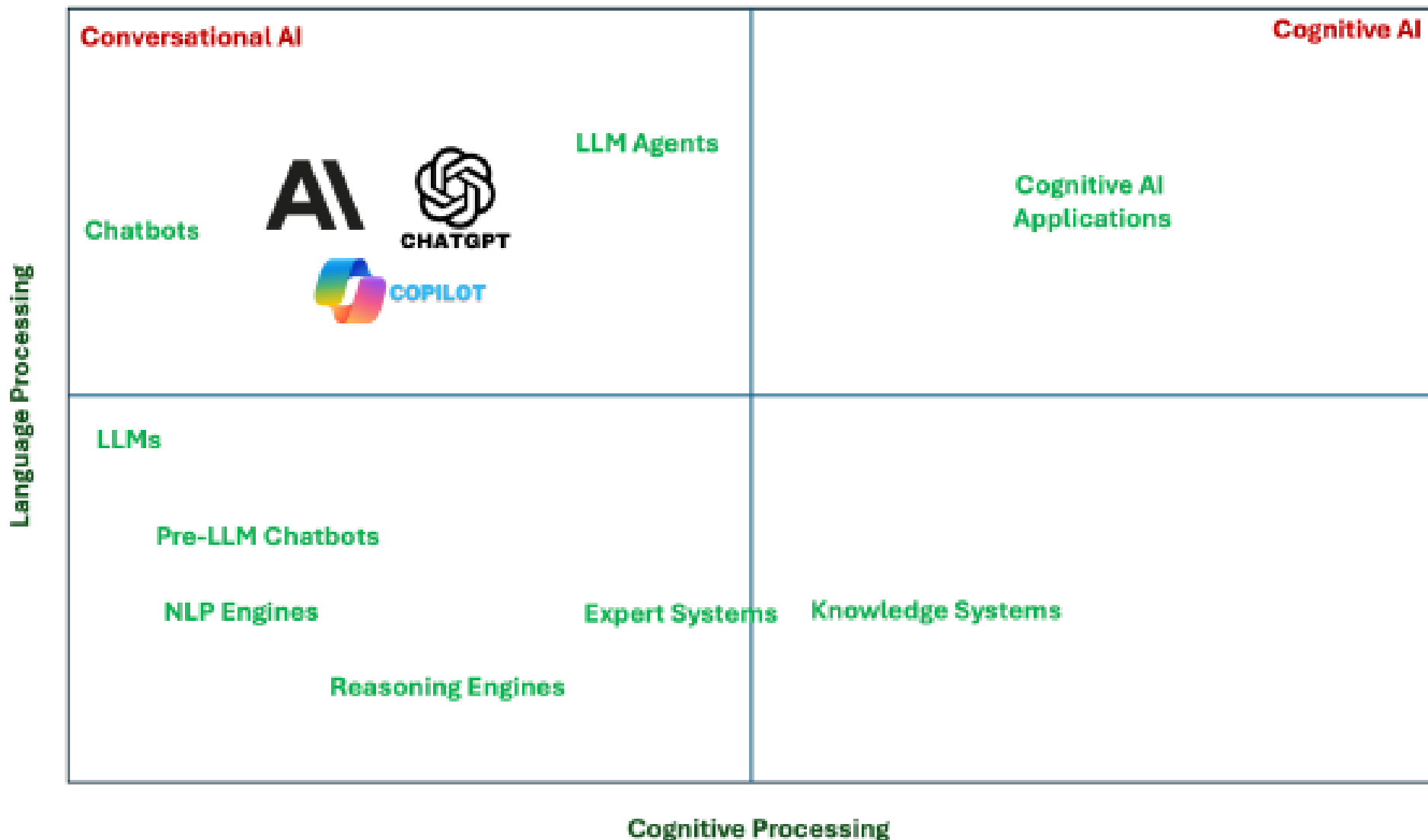
図： 筆者が主張するCognitive AIの基本アーキテクチャのイメージ(Dual Layer) 図作成: 岡島
(筆者は、システム外部との間のインターフェースをどこから取るかは示していない。)



- ・ Agentsは、それぞれ、認知&集団知形成に関する動作を自律的に制御/管理/学習する機構を持ち、言語を用いてLLMや他のAgentsと会話する。
(Neuro-Symbolic Cognitive Processes)
- ・ Agentsは、LLMを用いて、Cognitive層内や、システムの外部のリソースやデータを検索し、利用することができる。

- ・ 認知層の各Agentは、経験から知識と思考プロセスをリアルタイムに再学習する。
(各Agentは、適応戦略をProgramとして持ち、リアルタイムに適応を進め、担当する知識とProgramを自律的に管理し更新する。)

Figure 6. Conversational Versus 認知AI Quadrants.



7. Cognitive AI の Functional Architecture (P-15)

・ 2層間の対話により人間の思考を模倣

- > **会話層**(LLMを含む) + **認知層**(「認知Agent+そのAgentの管理機能」のネットワーク)
両層の相互作用により誘起される認知と思考により、最善なプロセスを計画/評価/判断/学習/応答する。
- > **会話層(Conversational AI Layer)**
LLM(GPT4相当)を内蔵し、認知層側の特定の認知Agent向けに、メタ認知に関する特定の動作を指令
- > **認知層(Cognitive AI Layer)**
認知タスクを受け持つ「認知Agent+専用マネジメント機構」を複数持ち、LLMが発する指令を受け、処理結果を LLM、もしくは他のAgentに応答する。
- > **認知Agentの機能**
システムオブジェクト(文書/知識/データや、Agents/各種ツール、及び、プロジェクト/計画)を共有し、複数Agentの協調動作にて、自然言語で、計画(条件付きワークフロー)/評価/判断/学習/応答を生成する。必要に応じて、システム外部のデータやツールにもアクセスする。

認知層
認知Agents

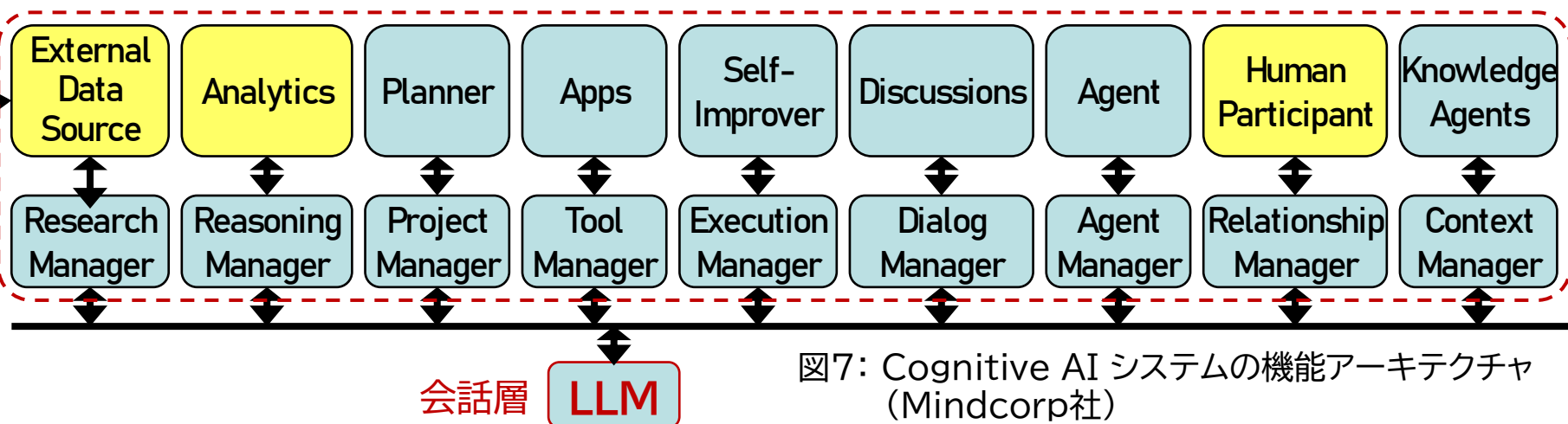


図7: Cognitive AI システムの機能アーキテクチャ (Mindcorp社)

7-1. 認知層(Cognitive AI Layer)の機能 (#1)

- ・ Relationship Management : Agents間の社会的関係の形成、管理、利用
- ・ Inter-Agent Messaging : 内部および外部Agents間の自然言語によるメッセージコミュニケーション能力 (文書、知識、ツール、データ、計画の共有)
- ・ Interactive internal and external dialogs : 「内部および外部Agentsとの間の「条件論理ルールで構造化され制御された意見交換動作」
- ・ Planning : 自然言語で表現された「条件付きワークフロー」を計画策定、理解し、対応する機能
- ・ Project Management : (複数の)Agentによる計画の実行を統率/指示し、管理するプロジェクト管理機能
- ・ Neuro-Symbolic Reasoning : インフォーマルな自然言語推論とフォーマルな論理推論の両方を制御するワークフローの生成機能
- ・ Memory Retrieval : 内部知識ベース、長期記憶、インターネットを含む外部データソースをアクセスして、与えられたコンテキストに関連する情報や知識を見つけ、検索する能力
- ・ Context Management : 認知プロセスの文脈を管理機能(長期記憶から関連する文脈情報をキャッシュし、ワーキングメモリの情報を設定する能力

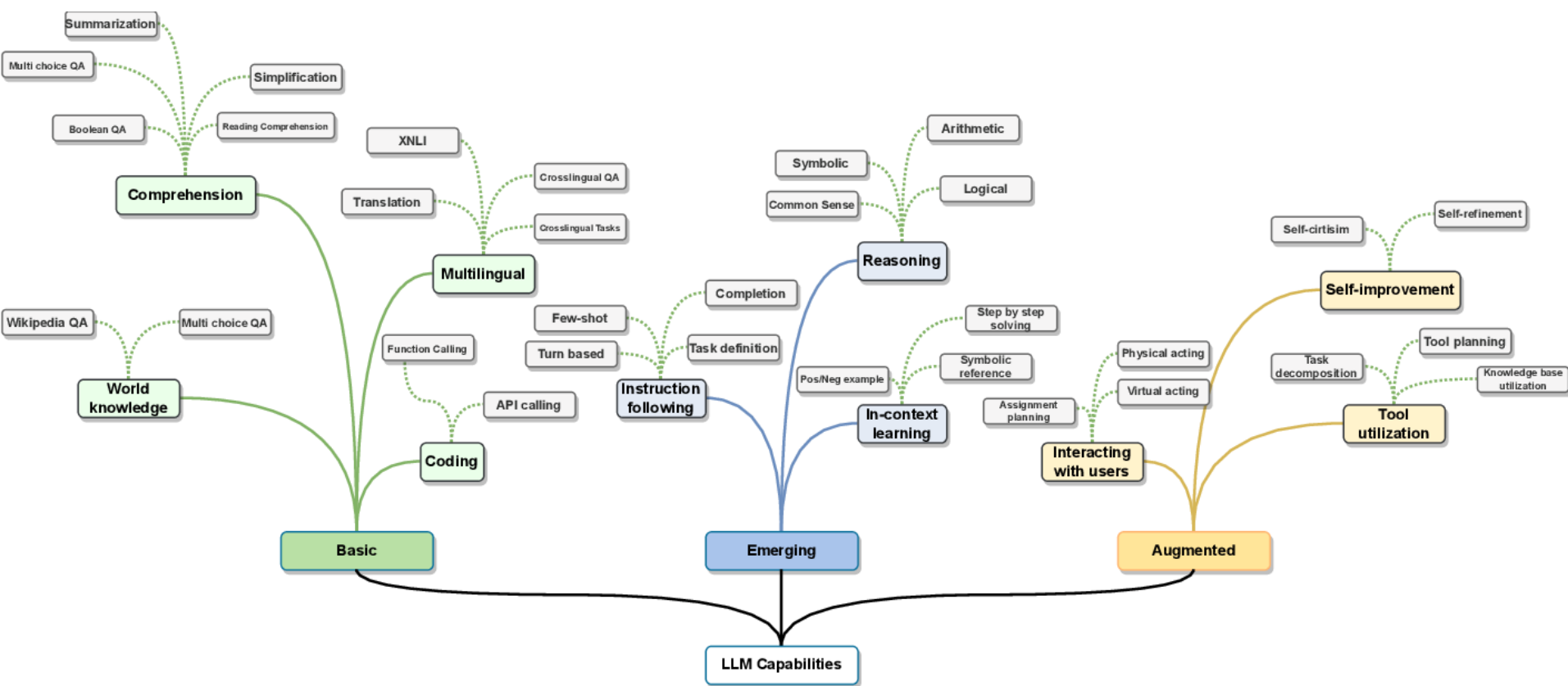
7-1. 認知層(Cognitive AI Layer)の機能 (#2)

- Knowledge Discovery : リサーチ戦略の保持と実行機能 (内部知識ベース、長期記憶ストア、インターネットやサードパーティAPIを含む外部リソースにまたがる関連情報や知識を探し出す)
- Knowledge Management : LLMの外部に存在する知識を一定の形式でレポートするための、学習、表現、生成、保存、検索、維持機能
- Tool-Utilization : 各種ソフトウェア・アプリケーション、API、内部および外部データ・ソースの形でツールを設計し、使用する能力
- Mathematics and Computation : 数学的・計算的操作を行う能力 (例えば、高度な形式的論理処理能力、科学的計算能力、財務的計算能力、さらにデータサイエンスとアナリティクス、機械学習能力)
- Multi-Agent Collaboration : 複数のAgentsの協調行動を統率/指示する機能
- Meta-Cognition : 認知機能全てに跨って、知識処理、内省、メタ推論、反省、学習、自己最適化を行うメタ認知機能
- Self-improvement : 再帰的な目標指向の自己改善能力 (推論、知識、プロジェクト、計画、ダイアログ、エージェント、ドキュメント、コードを実行中に反復的に最適化する)

Fig. 9 : LLM Capabilities (P-20)

(引用) Shervin Minaee, et al.(2024) ; “Large Language Models: A Survey”

- Amazon Inc.のShervin Minaee, et al.(2024) は、LLM技術をSurveyした論文で、「LLMの能力が理解、推論、道具の利用、社会的相互作用、自己改善へと進展する」と予測したが、現状は、未だ、実際の理解、推論、ツールの使用、社会的相互作用、自己改善は実現していない。



7-2. Dual-Layer Architecture (P17,18)

(参) Guo, Taicheng, et al.(2024). ; “ Large Language Model based Multi-Agents: A Survey of Progress and Challenges.”

- ・ **Cognitive AI** : 「仕事を進捗させるために協働する人工の知的システム」をAIとすると、
 - > 「LLMベースのAgent(例えば、ChatBot)+人間」が協働するHybridシステムも、一種の認知型AIと見なせる。(ChatBotは、その初期段階のModel)
 - > そのHybridシステムを発展させるには、内部を構成する人間やAIEージェントが、互いに対話するだけでなく、外部のツールや知識をアクセスして自身の活動をガイドし、結果を評価し自己改良する必要がある。

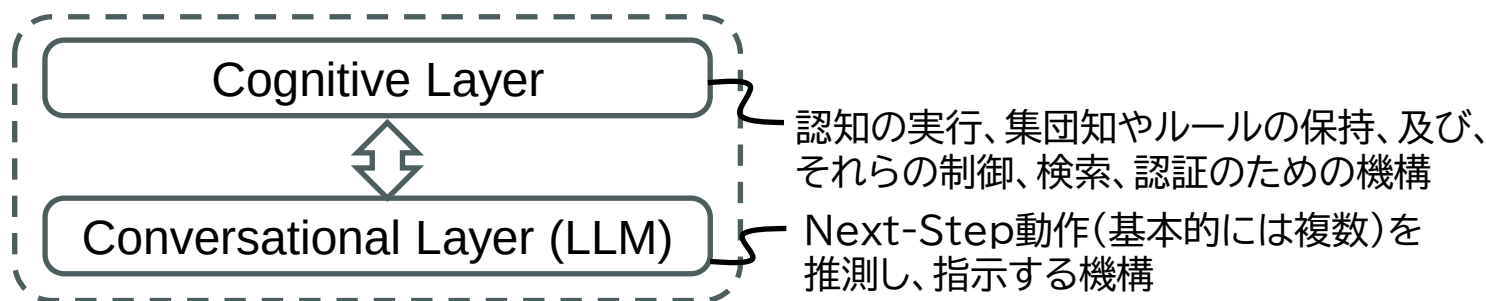
・ LLMベースのMulti-Agents System化 : 各仕様のラインナップが必要 (P21)

- | | |
|--|------------------|
| ① 対話の枠組みの関連技術 : Framework + Communication | } Agent群が |
| ② 課題解決関連 : Problem Solving | |
| ③ 世界予測(の関連技術) : World Simulation | |
| ④ 各Agentの能力更新関連: Agents Capabilities Acquisition | |
| ⑤ 評価関連 : Datasets and Benchmarks | |
- ・ 内観動作をサポート、
 - ・ 認知プロセスの実行
 - ・ 集団知やルールの保持
 - ・ 自身の動作/戦略の最適化
- するに必要な汎用機能群

下図 : 筆者が考える認知型AIの基本アーキテクチャ (Dual Layer構造の場合)

図作成: 岡島

(筆者は、システム外部との間のインターフェースをどこから取るかは示していない。)



7-4. Cognitive Agents (P-20, 21)

(引用) Guo, Taicheng, et al.(2024). ; “ Large Language Model based Multi-Agents: A Survey of Progress and Challenges.”

・ Multi-Agent System用のAgent群 (今後、各種仕様の標準化とラインナップ整備が必要)

現在は、未だ下記のような機能が断片的に試されている段階だが、これらは、標準化&ラインナップ整備がなされるべき機能であり、今後は、標準プラットフォームを定義し、その上に実装される「アプリ」となる方向だろう。

- ① Framework
- ② Problem Solving
- ③ World Simulation
- ④ Agents Capabilities Acquisition
- ⑤ Datasets and Benchmarks

・ 現状の用途

- > オンラインリサーチ
- > ソーシャルメディアを通じてオンラインタスクを行う程度

・ 今後、適用分野は非常に広がるだろう。

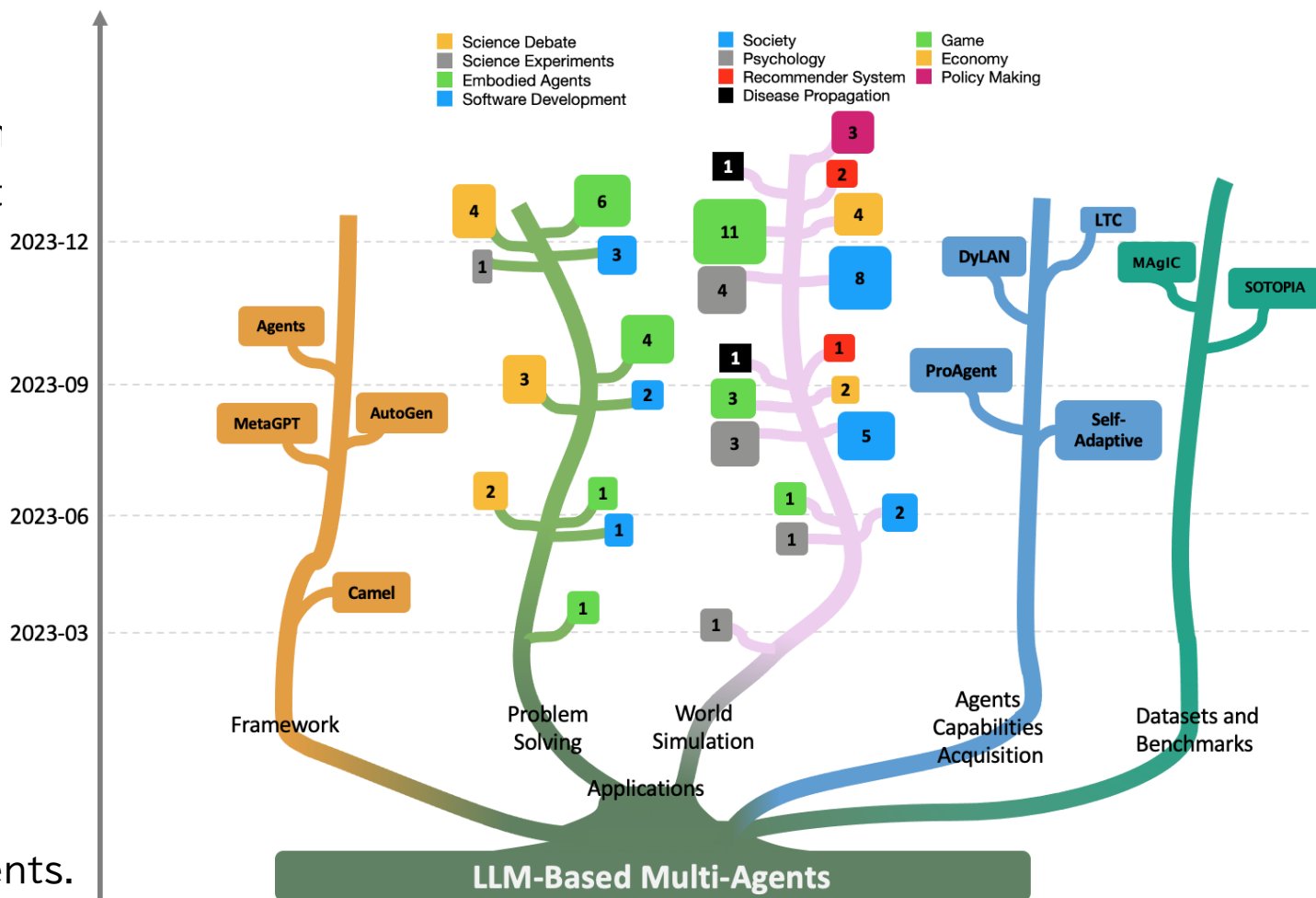
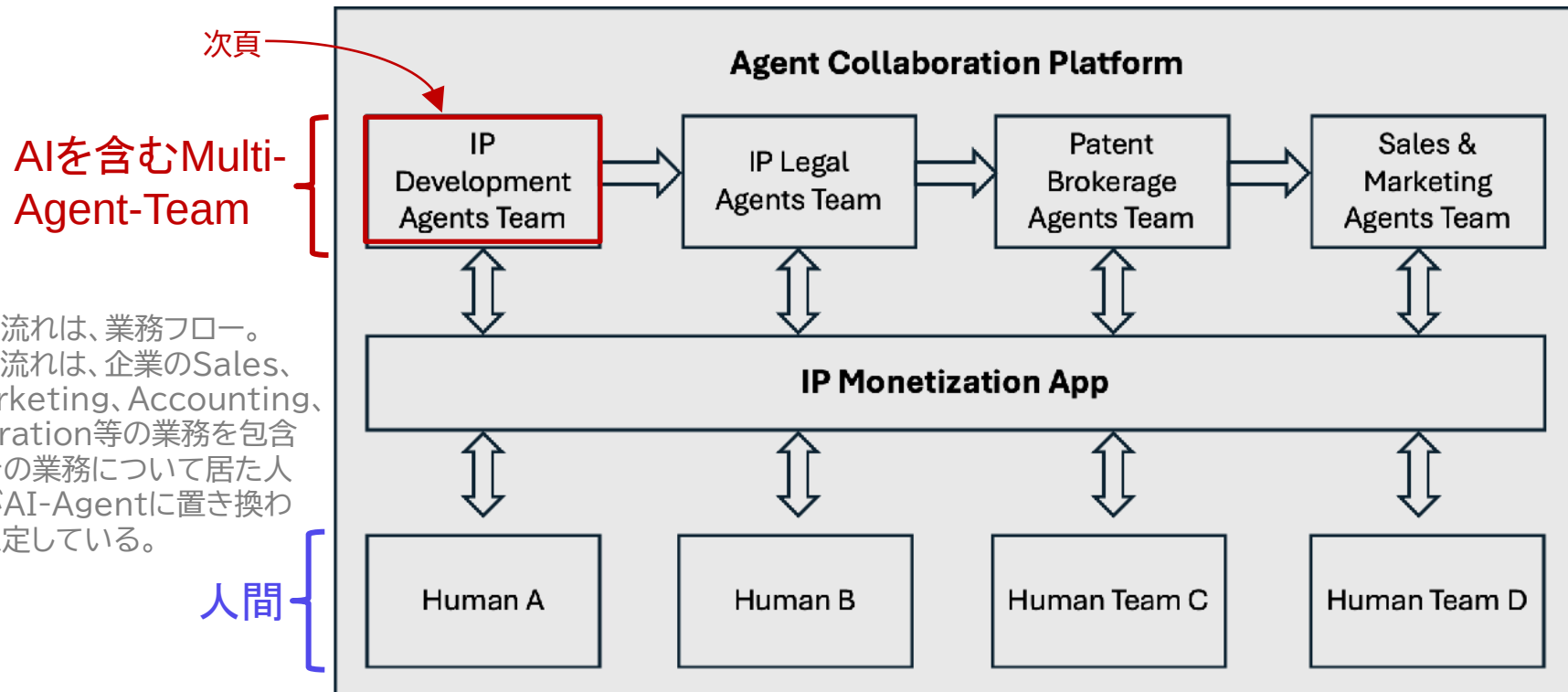


Fig. 10 (右):
LLM-Based Multi-Agents.

階層的Multi-Agent Systemの例：IP Monetization System (P22)

- ・ 知財事業プラットフォーム：複数のAgent-Teamsと多くの人間が協力し、収益化業務を進める。
 - ＞ 従来のほとんどのAgent-Teamは、SHMA(single human, multiple agents)
(一人の人間が、複数のAI-Agentにウェブ検索のような単純な仕事を委託する)
 - ＞ 今後のシステムは、single human & multiple agents (MHMA)を想定した汎用のPlatformとなる。(各Agent-Teamは、下階層にSub構造(モジュール)を持つ。次頁)

Figure 11. Agent Collaboration Platform



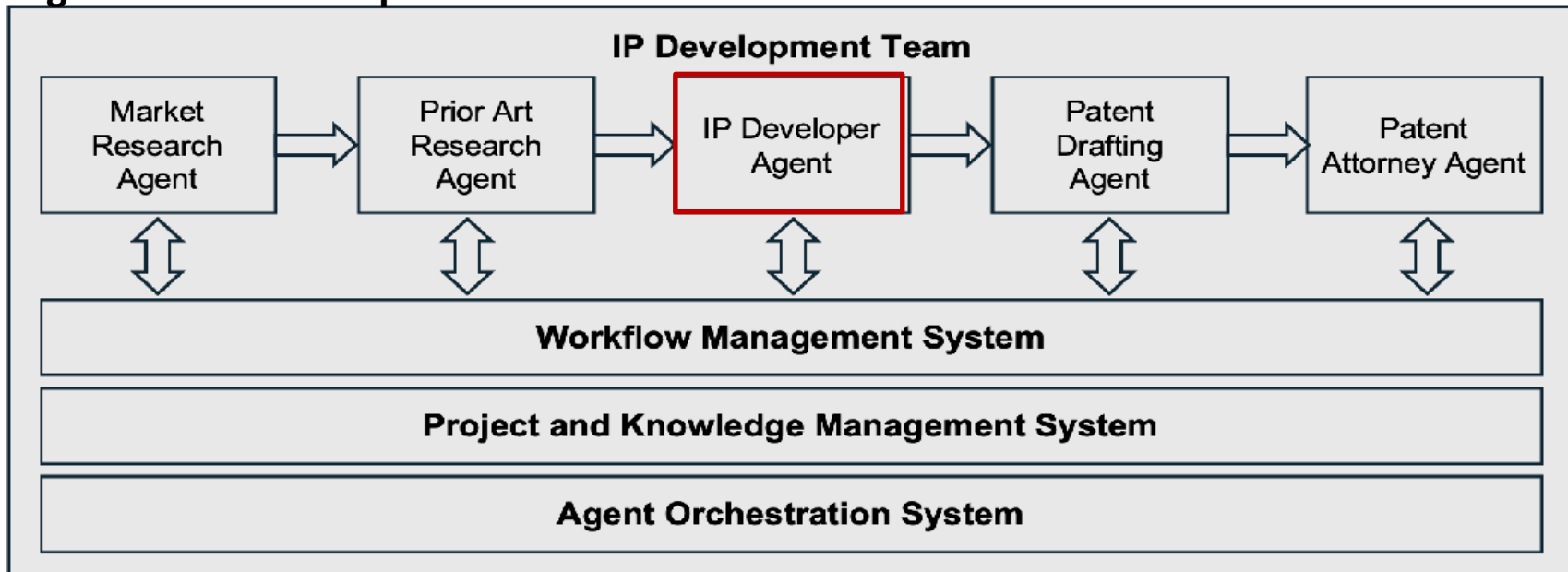
(岡島私見)
左から右の流れは、業務フロー。
この仕事の流れは、企業のSales、
R&D、Marketing、Accounting、
Administration等の業務を包含し、
元々、その業務について居た人間の
仕事がAI-Agentに置き換わる方向を
想定している。

Sub階層の例：IP Develeopment Team の Sub構造

- この場合、5個の専門Sub-Agentが、5工程を受け持つ。
(programmatic intelligent agents that operate as executable code on a multi-agent platform)
- Sub-Agentは、固有のManagement System や Orchestration Systemに制御される
 - > Workflow Management System
 - > Project and Knowledge Management System
 - > Agent Orchestration System

岡島私見:これら制御回路のプログラムの構成が重要だったと思える。LLMベースかもしれない。
- 但し、全てのAgent動作が「認知的」となる訳ではない。
(岡島私見 : 「全業務がWork-Flow上に記載されている訳ではない」の意味と思われる。)

Figure 12. IP Development Team Module.

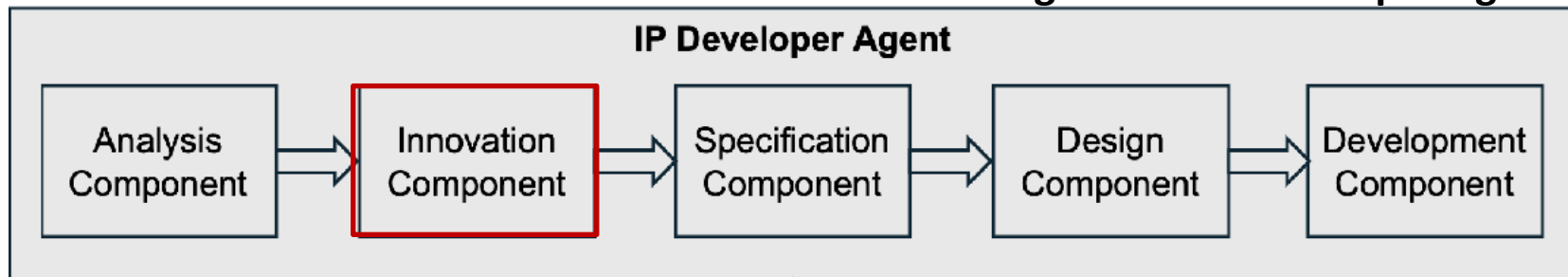


Sub階層のBreak-Down例

(業務を分解し、階層化し、モジュール化したタスクのフローとする)

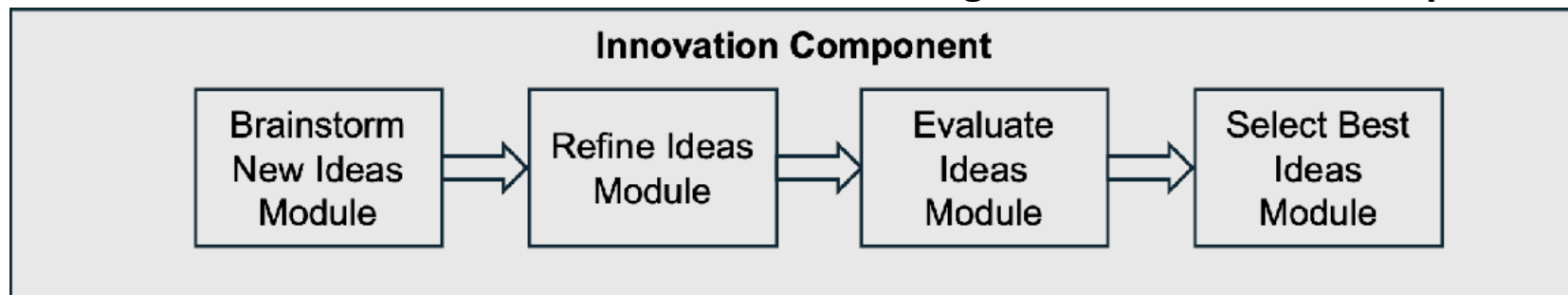
- IP Developer Agent が受け持つ5工程:

Figure 13. IP Developer Agent



- Innovation Component が受け持つ4工程:

Figure 14. Innovation Component



- 各Agentを、LLMの外側機能とするメリット

- > 主要機能を水平分業しやすい
- > メタ認知を明確化しやすい(処理プロセスの改良やルール化を進めやすい)
- > データ環境や問題セットの変化に適用しやすい

(Agentは会話層LLMの外の機能だが、言語機能と基本的な知能機能は、共通に会話層LLMを使用する。)

7-5. Relationship Management, Inter-Agent Messaging and Dialogs リレーションシップ管理、エージェント間メッセージング、対話（p24～26）

- ・ 認知的AIでは、2種類のレイヤーが同時に考え、自然言語により対話し、推論を進める。
 - > 会話層(LLM)は、確率モデルを用いて、入力を元に出力を言語的に推測(Simulation)する。
 - > 認知層のエージェントは、認知をモデル化した「プログラムベースの推論」を行う。
(決定論的論理)
- ・ Cognitive層は、階層的な内部構造を持ちうる。
(1段の認知層は複数のAgentを含む。下階層には、上階層と同様の構造が現れ得る)
- ・ 対話には、エージェント間の関係を設定・管理する手段と、対話の場であるネットワークが必要
(Relationship Management) (Cognitive Workplace)
- ・ 対話には、内部対話と外部対話がある。
 - > 外部対話：別のprivate cognitive workspaceに属する「他のAgent」との対話
 - > 内部対話：共通のprivate cognitive workspaceに属する「他のAgent」との間の対話
(内部対話プロセスは循環的であり、エージェントの問題解決戦略や計画の能力進化を反復する)
- ・ Workplaceに参加するSub-Agentをインスタンス化して「情報の流れ」をグラフ化し、動作を弁証法的に評価/改善するタスクは、「ワークフローや構造化知識の改善」となりうる。
 - > 改善には、集団知の改善とプロセスの改善があり、両者を進めることで環境に適応できる。
 - > システムは、複数のエージェント間のグループ会話を通じて、ブレインストーミングし、様々な集団認知タスクのためのベストプラクティスのグループプロセスを生成することができる。
(システムは、応答前の内省的な自己対話と自己最適化を経て、問題解決戦略を改善し、応答できる。)

7-6. Planning and Project Management (P-26、27)

- **LLMの弱点(#1)** (K Valmeekam, et al.(2023); "[Large language models still can't plan](#)")
 - > over-the-horizon (Long-Horizon) Reasoning で失敗する (長い論理、稀な論理を構築できない)
 - > (意味のある)計画を生成できない。検証も保証もできない。(以前に見た計画を借用しているだけ)
- **「計画(プラン)」は、Mukti-Agentsを目標に向かって誘導するプロセスの構造(グラフ)である。**
 - > 「計画(プラン)」は、多くのModule/Agentを、それぞれの独自の強みと能力に基づいて効果的に機能させ、目標の達成(タスク)をオーケストレーションする。
(プロジェクトの管理、推論/シミュレーション、調査、分析、意思決定、等)
 - > タスク・オーケストレーションでは、ワークフローに従い、分解したタスクを担当エージェントに委託(指示)する。
- **PlanningとProject管理は、認知型AIの中核となるメタ認知機能である。**
 - 「計画(プラン)」は、Module/Agentを参加させるための「条件付きワークフロー(プログラム)であるが、計画立案、開始を意味するだけでなく、計画の進捗を監視(Project管理)し、実行時にKnowledgeを調整し、結果のベンチマーク/判断、更に、プランの改良、結果報告までのフローとなりうる。
- **PlanningとProject管理は、静的ではなく、動的に実行/機能し、システム動作をナビゲートする。**
- **Planは、PDDLのようなプラン推論言語で記述される。**
Planning Domain Definition Language
 - > PDDLを用い、問題の領域と問題を一階述語論理を使って表現し、検索、分類、検証、最適化の作業を進める。

7-7. Neuro-Symbolic Reasoning (P-28)

- LLMの問題点(#2) :

- 1) 決定論的な方法や、ルールに従った推論に弱い

(幻覚や予測不可能な出力を起こしやすく、与えられた計画に常に従うとは限らない。)

- > 反転の呪い : 「AはBである」と言われた場合、「BはAである」と推論できない

Berglund, L, et al.(2023); "The Reversal Curse: LLMs trained on "A is B" fail to learn "B is A"."

- > 3つの集合の論理和のような単純な集合操作ができない (Yang, J. et al., 2023)

- 2) コンテキストウィンドウやトークン制限よりも長い「大規模な推論」をそのまま実行できない。

- 上記問題への対策 : Neuro-Symbolic Solution

- 1) LLM内部で行う非決定論的推論 と LLM外部で行う決定論的推論 を融合するアーキテクチャ

(「決定論的なプログラム制御」と「非決定論的なLLMの柔軟性と適応性」を組み合わせる)

➡ H. Latapie et al. (2022); "Neurosymbolic Systems of Perception and Cognition: The Role of Attention."

- > 正確性が向上 :

- ① 構造化、組織化、管理された思考(or 論理)

- ② ルール体系を遵守した問題解決(プロジェクトを実行/管理/改善)

- > 創造性が向上

(「計画の管理エージェント」とLLMの間の内部対話により、LLMの非決定論的推論の能力を、計画策定等のような「認知層の生成動作」にて活用する)

- 2) 「計画の管理エージェントのプログラム」にて、構造化、組織化、管理された思考プロセスを実行する。(岡島私見: LLMの非決定論的推論は、非常に複雑なMulti-Agent動作を引き起こし得る)

12章：認知型AIの「ニューロ・シンボリックアプローチ」(P-50、51)

- ・「ニューロ・シンボリックアプローチ」は、以下のような「人間的の思考」を模倣する。
 - > 形式的知識表現と知識管理 (formal knowledge representation and knowledge management)
 - > 形式的論理プロセスの適用 (application of formal logical processes)
 - > 形式的論理推論 (formal logical inference)
 - > リアルタイムの学習 / 適応能力 (ability to learn and adapt in real-time)
 - > 本質的に社会的でメタ認知的な機能 (inherently social and meta-cognitive functions)
 - > 知識と認知のネットワークを統合 (integrated networks of knowledge and cognition)
 - > 目標達成に向けての内省&自己改善 (to reflect and self-improve to achieve goals)

7-8. Memory Retrieval and Context Management

- **LLMの問題点(#3)**： 意味を理解したり、その対話を元に思考する面で限界がある。
 - > ソースの違う情報の文脈を理解して文章生成できない。
 - > 応答を個別に処理するため、「進行中の対話を元に思考する能力」が低かった。
- **対策としての認知型AI**： 認知層内に、**Long-Term Memory** と **Working-Memory** を持ち、
 - > **Long-Term Memory**に、長期間にわたり重要となる知識と経験を蓄積
 - > **Working-Memory**に、文脈/タスクに関する情報と、関連する経験知 を 過渡的に用意
(Working-memoryは、文脈記憶の洞察/理解のためのキャッシュメモリのように機能する)
 - > 両者Memoryのデータによって、LLMの動作をサポート：LLMのトークン・ウィンドウに、
 - ① 関連するコンテキスト情報を提供
(リアルタイムの分析/推論をフィードバックすることにより、動的に戦略を調整するために用いる)
 - ② 問題解決/戦略調整のための情報を提供
 - > 上記のサポートのために、認知層はLLMが必要とするコンテキストを管理する機構を持つ。

7-9. Knowledge Discovery and Knowledge Management

- LLMの問題点(#4)

- > **知識表現の管理**と**知識管理(Knowledge-Management)**のための機構を持たない。
(パターン認識とデータ駆動型学習に依存し、自然言語のストリーム処理に留まる。)

- **認知型AI**：

- > 認知層に、Knowledge-Managementのアルゴリズムを持つModuleを実装し、Long-Term Memoryに蓄積する「構造化知識のデータ群(知識表現)」を管理する。
- > 管理には、②新知識表現の追加と、③知識の改善がありうる。

用語の整理

- ① 知識表現： Long-Term Memoryに蓄積する「構造化知識のデータ群」
(具体的には、知識ベース、知識オブジェクト、知識カタログ、知識グラフ、分類法、オントロジーなど)
- ② 知識管理(新知識表現の追加)： 新しい知識表現や新しいデータと、既存の表現やデータの間の関連をLLMを用いて抽出し、①のデータ群に加える。
(LLMは、調査、要約生成、フィルタリング、等の処理で用いる)
- ③ 知識管理(知識表現の改善)： 複数の知識表現間の関連性/パターンを調査し、関係を抽出し、既存知識表現と整理、もしくは、融合/改善する。

- LLMの問題点(#5)： 外部Tool (の利用)をLLMは認知していない。

- > Toolへのアクセスは間接的で外部システムに任されているため、LLMは、外部Toolの存在を理解したり、直接対話したりする能力を欠いている。

- (ツール含め全体システムの概念(「自己のデータ」や「外部ツールからのデータ」の区別が無く、結果、「対話」の概念も認識も無い)

- 認知型AI： Toolを、「環境」と同様に「外部」と見なし、対話 / 観察し、制御する。

- > システムのFrameworkは、(外部)Toolの利用と認知を統合的にサポートする。

- (「ツール利用」には、ツールの新規作成と管理を含む)

- > システムは、Framework に位置付けられた各種Tool / Agent を認知し観察しながら、対話 / 利用 / 制御 / 管理を行う。

- > システムは、必要に応じて、プログラムやLLMモデルを開発 / 改良する。

- (自律的に独自のSaaS(Software-as-a-Service)を作成し、自律的に設計、構築、シミュレーション、テスト、評価、Benchmark、改善、運用、保守する。)

- > Toolとしては、外部のITインフラ (計算資源、データ、制御装置)を含む。

7-12. Multi-Agent Collaboration (P34～35)

- ・ LLMの問題点(#6) : 外部Agentと、協調/協働を目指したコミュニケーションは明確ではない。
- ・ 認知型AI : 以下の機能が、参加するAgent/個人/Toolの関係を社会的/協調的とする。

(注) 筆者は、「この発想は、マービン・ミンスキーの著作(Society of Mind)からインスピレーションを得た」という。同書は、「AIを、エージェント達が集まってできた社会」と見なしたことで有名。「深い洞察と解決策はしばしば共同作業から生まれる。人工知能と人間の知能の未来は、個人の才能にあるのではなく協力し、共有し、共に革新する能力にある」と見なした。

- > Orchestration Management : 目標を共有し集団の成果を最適化するメカニズム
(エージェント間のコラボレーションを、組織化/階層化、管理、実行)
- > プロジェクトマネージャー機能 : 計画/ブレインストーミング/実行/評価/改善を推進
(「集合知/社会的知性の進化」をプログラム)
- > 自然言語によるエージェント間対話Platform
(エージェントAIと人間の対話、人間と人間の対話を含む)
- > Frameworkの共有
- > 集合知/社会的知性の共有
- > 誤解(矛盾や無駄)の解消メカニズム
- > Frameworkが人間対人間の相互作用も促進

7-13. Meta-Cognition (Page-36)

・メタ認知：自己反省的内省的推論にて、自分自身の動作を感知し、対応/改善反応を起こす能力

- ＞ 認知型AIシステムは、自身の改善のために自身の動作について思考するための「汎用的機能」を実装する。

そのために、「自身の内部状態と動作を感知し、メタ認知システムにそれら情報を(再帰的に)取り込み、自身を改良する動作の実行をマネジメントする」という機構を実装する。

- ＞ 但し、エージェントアプリケーション全体にメタ認知を実装することは技術的に難しい。

(新しいIP / ツール / アプリケーションパラダイムが必要。自己反省に頼ると「する側」と「される側」が生ずる。)

メタ認知のための「汎用的機能」

- Knowledge processing : 知識を表現し推論するために必要な形式的なデータ構造、及び、そのデータの検出 / 変換 / 生成プロトコル。
- Introspection : 自身の構成 / 状態 / 文脈 / 動作に関する表現を、検出 / 変換 / 知識処理する機能
- Meta-reasoning : 内省動作や知識処理で必要となる「知識に関する知識」と「推論に関する推論」を表現 / 変換 / 生成する機能
- Reflection : 自身の構成 / 状態 / 文脈 / 動作についてメタ推論する「内観プロセス」を処理するに必要な機能。
- Learning : 特定のタスク/目標に関する過去の知識や自身の応答パターンを分析&評価し、将来の知識や応答パターンを改善する機能。
- Self-optimization : 学習と反省を利用して自身の構成/状態/文脈/行動を改良する機能。

- ・ **会話型AIの限界：累積学習やReal-Time学習の能力が低い。**
 - ＞ 学習時の(過去の)データを元に、統計に基づいた応答を生成しているにすぎない
 - ＞ Q*のような学習アルゴリズムで強化したLLMも、生成するパターンは学習したパターンに限定される
- ・ **認知型AIの動作の特徴：人間の認知プロセス(内的行動、内観、内省)を模倣**
 - ＞ タスク処理推進を目的としたAgentの協働(Orchestration)を反復 7.12章
 - ＞ アウトプットが環境に及ぼす影響を再帰的に分析 (構造化データとアルゴリズムを用いて計算処理)
 - ＞ 2重の適応プロセス (既存の推論/改良ストラテジーを実行 & 既存の推論/改良ストラテジーの改良)
 - ＞ 改善プランの承認プロセスの民主化 (問題発生をトリガーとし、承認プロセスの後に有効化される)
- ・ **適応動作 (自己改善の有効化)**
 - ① 認知層Framework内の各Module/各Sub-Moduleが搭載するProgramの改善
 - ＞ Agent間対話による民主的なプランニング (ワークフロー生成) 7.5章、7.6章
 - ＞ 環境との対話、外部Tool/Data利用、Tool開発 7.5章、7.10章、7.11章
 - ＞ 結果予測(シミュレーション)
 - ＞ 評価Datasetを用いたベンチマーク(影響評価)と判断
 - ② 業務(Flow&Network)の改善 (Real-Time学習) タスク依存大
 - ＞ Workflow Management / Relationship Management(プロジェクトの管理) ... 7.5章、7.6章
 - ③ 形式的知識表現(Data)の改善 (Real-Time学習) タスク依存有り
 - ＞ Knowledge Management (構造化知識/知識カタログ/知識グラフ) : 7.7章、7.8章、7.9章、7.10章
- ・ **自己改善のプロセスにては、「メタ認知のための汎用的機能」を利用する 7.13章**

8. Comparison of Conversational AI to CognitiveAI

・ LLM（確率的推論マシン）の 弱点

- > 訓練時のデータ分布に基づく推論パターンを再現するに留まる
- > エラー/幻覚/予測不能な行動を起こしやすく、一貫性、信頼性に欠ける(ことがある)
- > 学習後には、特定の知識を抽出/分離できなくなる
- > 「知識」を、厳格に埋め込むことができない。従って、文脈を正確には理解しない
- > 訓練後の新情報取り入れ(リアルタイムに旧情報と統合すること)に困難がある

・ しかし、自己改良を可能とする認知型AIは、ツールとしてのLLM(確率的推論マシン)を必要とする。

- > 訓練データに含まれていた明示的/暗黙的な知識は、問題解決の糸口情報を含んでいたかもしれない。

・ また、認知型AIは、決定論的論理も必要とする

- > 形式論理を運用し、各種のTool(Program)を用い、各種Managementを実行するため。
- > 知識カタログに登録した構造化データ/知識グラフ/アルゴリズムを利用するため
- > 計画を策定し、実行(Orchestration)するため
- > Function calls や API を介して、外部Tool/外部アプリ/外部データを統合的に利用するため

・ ハイブリッド推論（非決定論的プロセスを決定論的に制御）が必要とする構造

- > Dual-Layer Architecture (LLMが生成する言語を用いた認知Agent間の対話と内省の構造)
- > 認知AgentsとLLMの間の対話のための機構 (Framework、Network管理、Agent-Manager)
- > 統合知識管理システム (知識カタログ、グラフ、メタデータ分類のための機構、等)

> 確率論的論理機構：LLM

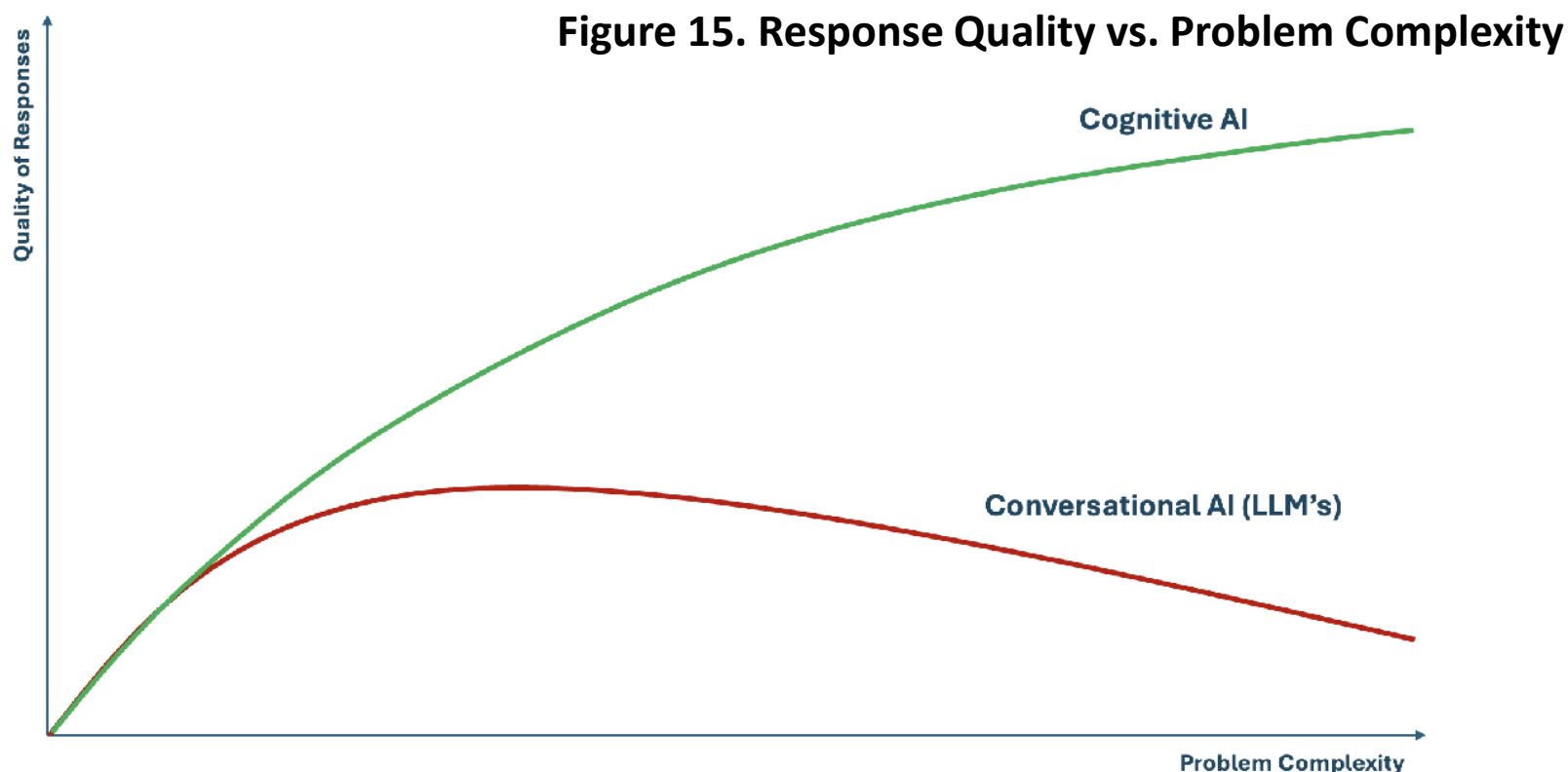
> 決定論的論理機構：

「認知AIシステムは、再帰、並列コンピューティング、非同期コンピューティングなどの高度な計算戦略を実装する柔軟性を持って設計されている」とある。(Page-44)

(岡島注：メタ推論では、多くの推論が並走する可能性が高いが、結果の集約を可能とする機構が必要と思う)

Figure 15. Response Quality vs. Problem Complexity

- 会話的AI (LLM)の能力(知性の質)は、問題の複雑さが増すにつれて頭打ちになる。
(確率的な性質と決定論的なプログラミング能力の欠如が制約する。再帰的操作、ツリー検索、カーブフィッティング、最適化などの高度な問題解決プロセスに関与することができない。)
- 汎用性が高い決定論的論理を用いて、認知的AIは問題を還元し、抽象化し、構造化した知識を集積するため、AIとしての質の向上を続けるだろう。(認知的AIの構造は、会話的AIよりも複雑になる)
- 従って、**認知的AIは、人工知能分野に革命をもたらし、AI技術が達成する範囲を広げるだろう。**



9. Limits of Cognitive AI (P-45~46)

10. Cognitive AI in the Evolutionary Ladder of Intelligence (P47~48)

・ 認知型AIのアーキテクチャで目指す範囲

- > 個人 / チーム / 組織の知識労働を補強し、促進すること
- > AIにメタ認知能力を獲得させ、人間の知能を超えること (Page-47)
- > 協働で業務を進める「人工の集合知性」を実現するためのインフラとなること
(Collective Intelligence / Artificial Group Minds)
- > EGI (Enterprise General Intelligence)の構築
(Collective Intelligence of Enterprises)
- > AIに自己主導的、自己改善的な思考 / 行動 / 進化 を与えること

・ 認知型AIではAGI (包括的な知能) の実現は困難だが、AGIの出発点となりうる。

- > 全ドメインの全ての問題への対応するには、回路リソースが不足
(メタ認知モジュール設定パラメータ量は有限。学習データ量も有限)
- > 知能の完全な一般化の見通しは未だない。
- > ゼロショット学習の能力が不足
- > 人間知性の直感的、創造的、感情的な面を再現できていない。
- > Bodyが無い場合、世界と相互作用して知性を獲得できない。
(特に、社会性や生化学的/物理的な側面)

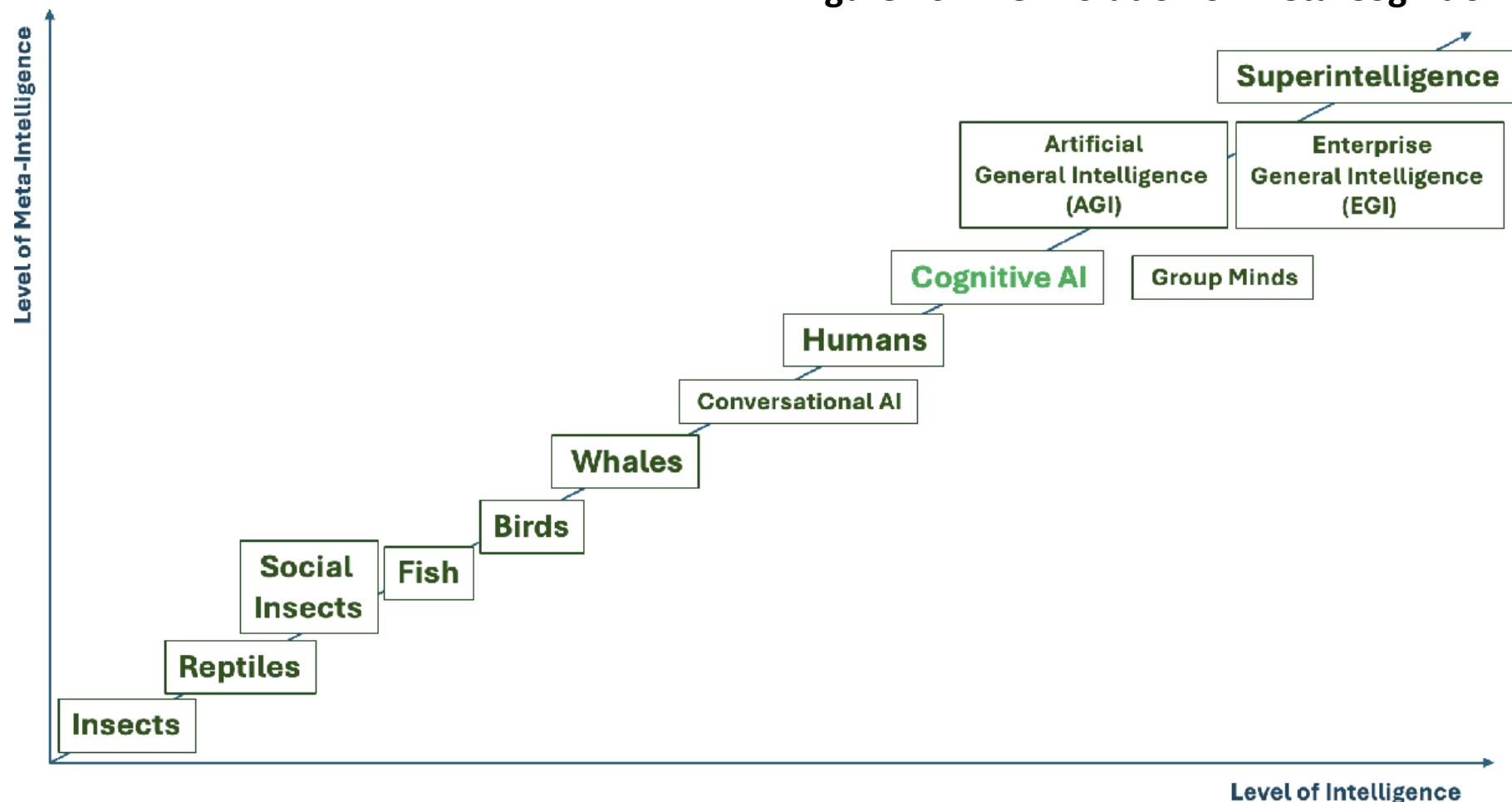
この過程で、倫理的、社会的、技術的な意味を慎重に検討する必要がある。

・ AGIは、embodiedシステムだろう：人間の知能を再現

・ Super intelligence AI：空間的にも時間的にも有限な生物や組織の限界を超越する。インフラ基盤から根本的に再定義し、人間の理解を超える。

Figure 16. The Evolution of Meta-Cognition. (P-47)

Figure 16. The Evolution of Meta-Cognition.



11. Exponential Intelligence : 指数関数的知性 (P-49)

- ・ **定義** : 人間と認知的AIの知性融合により知性向上が加速し、Superintelligence発現に向かうプロセスを開始した知性

$$EI = (HI + AI)^{(2+x)}$$

Where : EI = Exponential Intelligence Level
HI = Number of human intelligent agents
AI = Number of Cognitive AI agents.
x = Level of collective intelligence:
Level = 0 : One human + one AI
Level = 1 : Many humans + one AI
Level = 2 : One human + many AIs
Level = 3 : Many humans + many AIs
Level = 4 : Many networks of humans + AIs

- ・ **数式の意味** : EIは自己増幅する。

[岡島注] 7-14章では、「認知的AIは、2重の適応プロセスを持つ」とした。
: ①既存の改良ストラテジーの実行 と ②既存の改良ストラテジーの改良
Agentや人間が多いと、③「既存の改良ストラテジーの改良プロセス」の改良
④「『既存の改良ストラテジーの改良プロセス』の改良ストラテジー」の改良
と、より改良を加速できるということかもしれない。

- ・ **人間の参加は、これらの認知型AIシステムに意識を与える**

- > 人間は、意識の原子単位として機能する。
- > 人間の知性は、EIをより安全で、含蓄に富み、目標指向性が高く、適応性が高いものにする。

12. Implications : 認知型AI登場の意義 (P-49)

- ・ LLMの問題 : LLMの運用の仕方(運用フレームワーク)に問題があり、高度な知性に発展しづらい。
 - > 「膨大なトレーニングデータセットを元に『統計的に尤なパターンを推測する』との戦略では、人間の行う「演繹的推論」や「創造的思考」を再現できない。(厳密な論理を構築しえない)
 - > 累積学習と適応を行う能力が低い。
- ・ しかし、認知的AIは、7章から8章にまとめたような機構と能力を持ち、人間の思考を模倣しうる。

LLMの確率的推論を行う機構と決定論的推論(記号処理と計算)を行う機構と、人間の行う(メタ)認知プロセスにて必要となる様々な「汎用的機能」を備え、それら機能のOrchestrationとManagementによって、

- ① 思考を計画し、
- ② 確率的推論と決定論的推論(記号処理と計算)を組み合わせた深い推論をシミュレーションし、結果をベンチマーク評価し、判断する能力を持ち、
- ③ 構造化知識の改良、ワークフロー自体の改良を進め、
- ④ それらの累積学習結果を必要な時に採用することで、リアルタイムに状況に適応するので、人間の行う「演繹的推論」、「創造的思考」、「内省」、「反省」、「最適化」を模倣しうる。

- ・ 会話型AIから認知型AIへの進化は、単なる技術的なアップグレードではなく、人工知能の概念的な基盤における根本的な転換である。

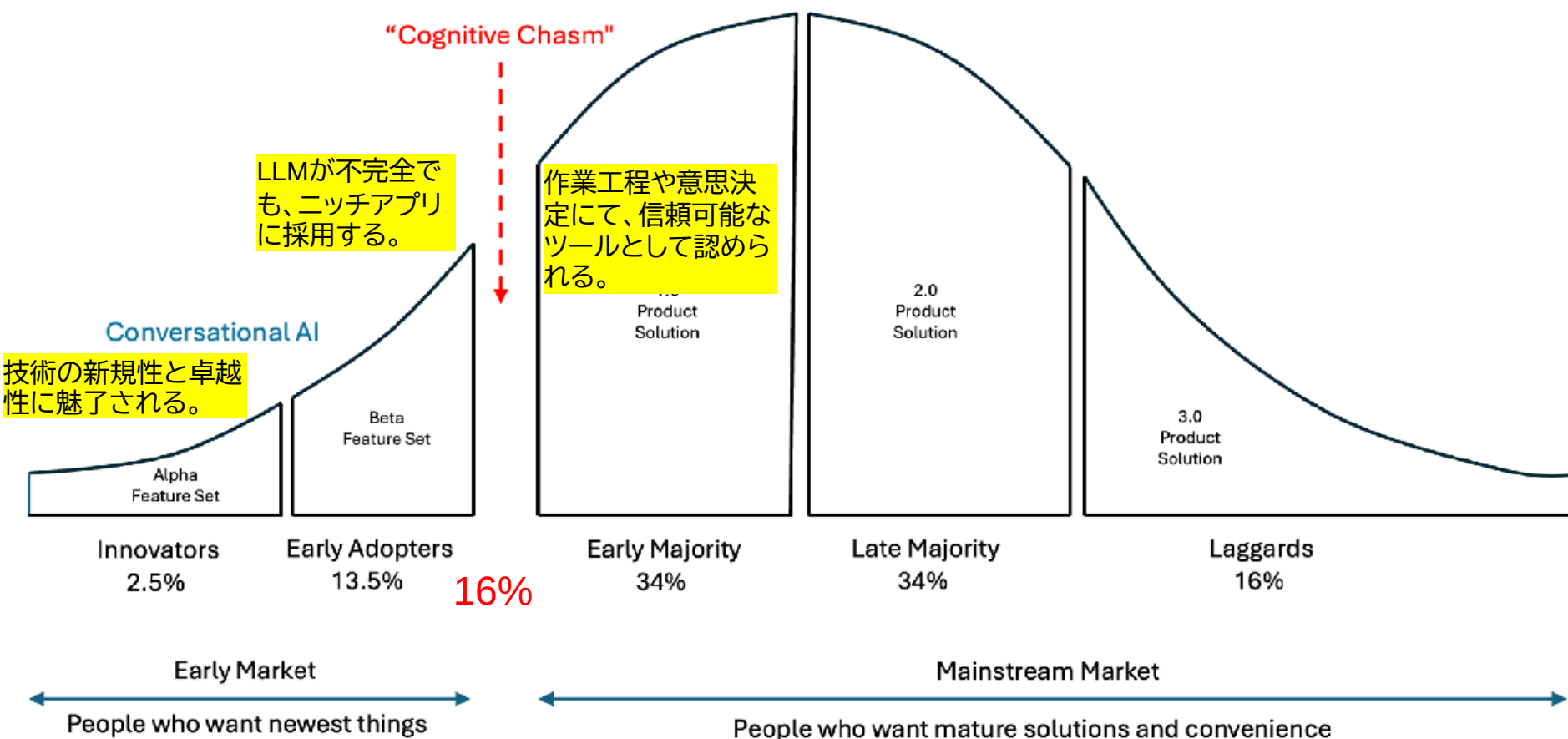
13. Crossing the Chasm

(注) Geoffrey Moore's Chasm Theory : 製品を普及させる時の超えるべき「市場普及率の溝」

- ・ LLMはまだ初期採用段階を越えていない。登録ユーザー数が1億人を超えたが、未だ、ChatGPTは最小限の機能しか持たないアプリであり、ターゲットとする「数百万の組織と数十億人のほとんどには未だ採用されていない。
- ・ Early Majorityに受け入れられるには、**現実ビジネスの知識労働にて、「実用的で有用」と認められる必要**ある。
(ニュアンスを理解し、正確さと確実なポリシー遵守が必要なミッションクリティカルな知識労働には不十分)

Cognitive AI

Figure 17. The Cognitive Chasm.



14. Re-evaluating Current AI Approaches (P-54)

15. LLMs as a Commodity (P-55)

・現在のAIブーム：大規模言語モデル(LLM) + 関連技術

(学習ツール、GPU、ベクターデータベース、RAG(Retrieval-Augmented Generation)、チャットボット)

・認知型AI 登場後のAI技術の開発動向：

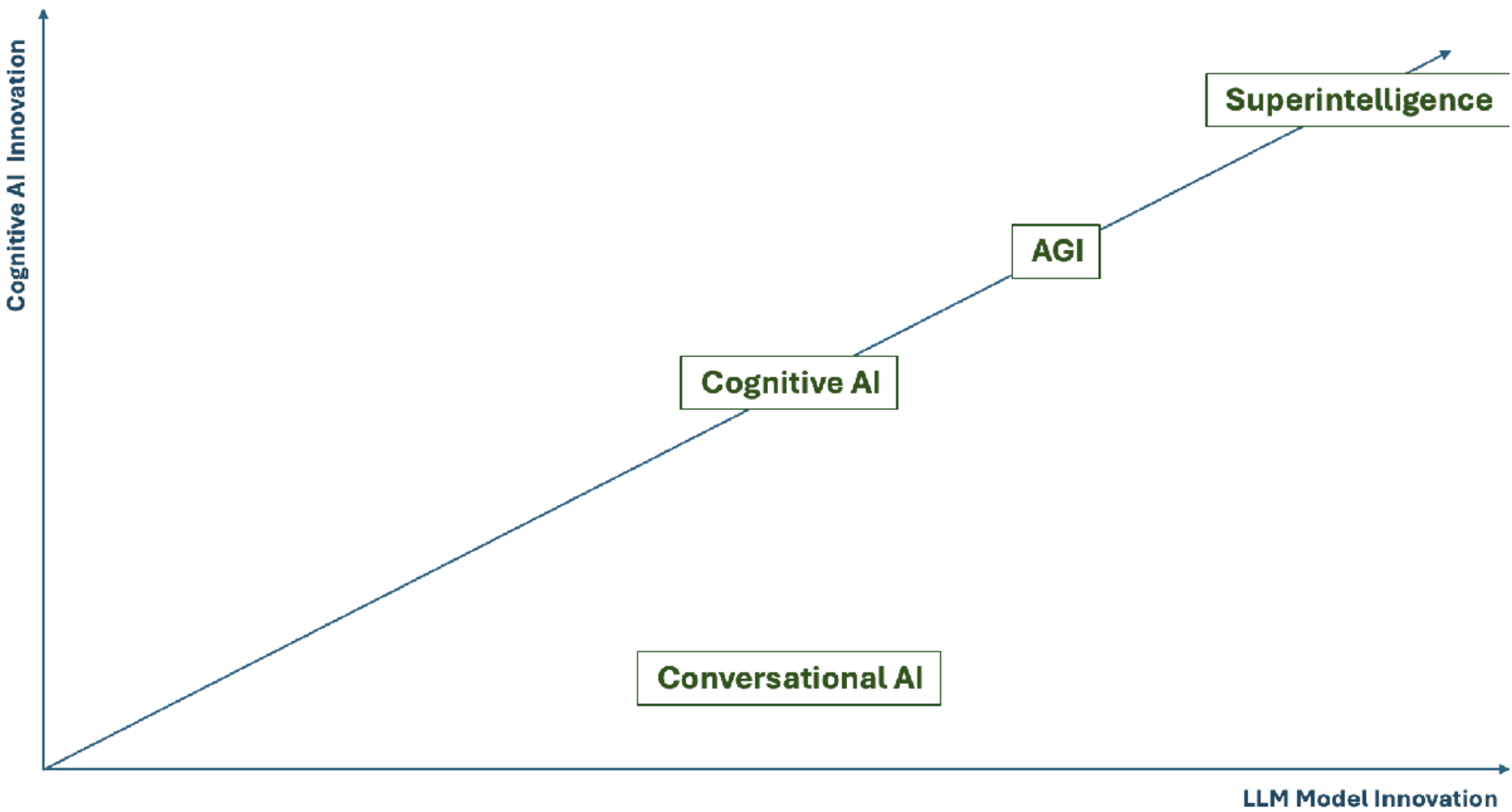
- ＞ 二次的な開発領域に競争が移行し、開発体制と投資の優先順位の見直しを迫る。
- ＞ LLMモデルの進歩が直接的に認知型AIやAGI、超知能をもたらすことはない。
 - ＞ 会話型AIへの投資を増やしても、今後大きな進歩は望めない。認知問題の本質的解決が必要。
- ＞ 今後のAIの品質/能力の向上は、会話層ではなく認知層からもたらされる。
 - ＞ 認知AI層がシステム能力を向上させ増幅させるため、今後は、LLMの性能向上のメリット小さくなる。
 - ＞ 既に、会話型AIは、認知AIの進化を可能にする「十分な」能力を備えている。
 - ＞ そのような見通しから、認知AIに開発投資を振り向ける総合的な動きが始まるだろう。

・LLMs as a Commodity

- ＞ LLMモデルは、開発と展開のコストが下がり続けており、会話AI技術はコモディティ化する。
- ＞ 今後も、LLMは投資を集め続けるだろうが、その開発は統合され、寡占化すると予想する。
- ＞ 代わって、イノベーション競争は、認知型AIに移行する。

Figure 18. AI Innovation Trajectory

Figure 18. AI Innovation Trajectory



15. Commercialタイプの認知型AI (P-55)

16. Conclusions (P-56)

- ・ 認知AIのアーキテクチャは、機械知能を再定義し、人間と機械の協働の時代を切り開く。
 - ＞ 認知AIは、人間と機械の協働ソリューションとなる。
 - ＞ メタ認知は、あらゆる分野にて、イノベーションと問題解決の機会をもたらす。
(政府、医療、教育、金融など)
- ・ LLMから認知型AIへの移行は、AI開発投資の見直しを進める。
 - ＞ LLM技術は、コモディティ化し、わずかな寡占企業以外にはビジネスになりづらい。
 - ＞ 認知型AIでは、「AIに自律的に進化する能力を与える」というパラダイムシフトを引き起こす。
- ・ 自己改良可能な認知型AIに、倫理的/社会的/技術的な面で適切な認識能力を与えるには、AI開発エコシステムを適切にリードする必要がある。
- ・ 認知AI型を出発点として、AIと人間が協力して、どちらも単独では達成できない「飛躍的な知能」の実現を目指す。

参考

Abstract

- ・ 累積学習(cumulative learning)を目的とした認知アーキテクチャについて
(経験から漸進的、自律的に学習するために必要な情報構造と制御構造)
 - > 従来は、抽象度の異なる情報の認知(多様なパターン、タスク、ゴール構造)には、異なる処理モデル(System-1とSystem-2)が必要とされてきた。
(構成/構造、制御/処理メカニズムが異なるため)
- ・ 本稿の主張：
 - > そのような二元的な想定(2種類の制御が必要)は適切ではないだろう。
(どのような抽象化レベルの知識にも、データとしては、記号情報と下位記号情報の両方が含まれているべきだ)
 - > 但し、抽象度の異なる(高い/低い)データ処理には、異なる注意メカニズムが関与していると思われる。
(その2つのメカニズムの存在が、2つのSystemがあると予想されて来た理由だろう)



Abstract

- ・ 新規性/新知識/不確実性を生む「意思決定の起源」について
- ・ 人間の認知は、理論に基づく因果推論の一形態としてみなせる。
(人間が世界に「介入」し、新たなデータを求めて実験するプロセス)
 - > (データのみに基づく)AIが真の新規性や新しい知識を生み出し、人間の意思決定に取って代わるとは思えない。
 - > (LLMのような)データに基づくAIの予測は、人間の理論に基づく因果論理や推論とは異なる。
 - > AIは知識を確率的に蓄積するが、それは後ろ向きで模倣的。
(人間の認知は前向きで真の新規の考えを生み出せる)
 - > 「データ-信念の非対称性」という考え方にて、AIと人間の認知の違いを議論する。



Abstract :

- Neuro-Symbolic Solutionの提案：
 - > ニューラルネット(学習)と記号論理(知識表現)を融合させるフレームワーク
 - > 確率表現と論理的推論(古典的一階論理含む)に対応可能とする
 - > 全ニューロンは、解釈可能な形式で、分離表現しうると仮定
(重み付けされた実数値論理式で動作を表現し、モデルは微分可能)
 - > 学習時に、論理矛盾を表現する損失関数の値を最小化する
- 利点：
 - > Open-World仮定の設定を可能にし、不完全な知識への耐性を持つ
(注) Open-World仮説 : 証明/検証されていない命題を「偽」としない。
 - > 学習時に論理矛盾を最小化することができる

以上