The Interplay of Reasoning and Agentic Artificial Intelligence: An Introduction Survey *

Christopher D. Molek
University of West Florida
11000 University Parkway
Pensacola, FL
cdm106@students.uwf.edu

Abstract

This work serves as an introductory survey for Reasoning and Agentic Artificial Intellegence (AI). While the survey is certainly not exhaustive, the author intends to convey the areas of research that have focused on AI methods that enhance reasoning and decision-making embodied in AI models/architectures that autonomously plan, learn, and execute tasks with minimal human intervention. In this survey, the author chose to go wider than deep. One of the reasons for this is not to lose an introductory practitioner with real deep details regarding the methodologies. Secondly, the author wanted to show the breath of research being conducted and yet dilemmas from the 90s still persist. This is a testament to the difficulty associated with generation of robust systems of autonomous agents that can reason about their environment and take action. A surge in research has occurred at the intersection of generative AI and agentic AI and is a focal point of this survey.

***K*eywords** Agentic AI · Reasoning with AI · Large Language Models

1 Overview

This introductory survey cites a collection of 50 references. Initially, ˜1000 references were identified. These references were analyzed by number of citations and down-selected from this criterion by a high-level. The down-selection process resulted in 78 references. During the detailed review of these references, 38 other references were obtained to complete this review. Since the topic **Reasoning and Agentic Artificial Intelligence** is very broad, the author chose to give a brief history summary of the broad topics in the introduction. More focus is drawn toward Large Language Models as this is a very interesting and highly researched area currently.

The writing of this survey serves as, an introduction into the field of reasoning with a generative AI. The intention of writing as an introductory survey serves two purposes: First, to get the reader acquainted with the historical perspective of both reasoning with AI and agentic AI. Secondly, to show the reader that the historical dilemmas that researchers faced in the 90's have not gone away. Indicating that these are time tested dilemmas that are difficult to solve. Some of the issues researchers still face today are included in a Section 7 - Research Challenges and Gap Areas.

The information is limited to articles available up to April 22, 2025.

2 Introduction - Historical

For the topic of **reasoning and Agentic Artificial Intelligence**, let us break this down into pieces. Starting with Artificial Intelligence, then discussing agency, and followed by reasoning. From there we will embark on a brief journey of history discussing each of these. These topics will culminate into a topic of General AI. A Discussion of existing methods to achieve better reasoning with Agentic Artificial Intelligence will be presented along with the associated benchmarks for measuring success. Finally research challenges and gap areas will be discussed.

* *Github Repository for a basic coded example located at:* https://github.com/cdm106/ML_Final_Project

The Standford Encyclopedia of Philosophy defines Artificial Intelligence as “Artificial intelligence (AI) is the field devoted to building artificial animals (or at least artificial creatures that – in suitable contexts – appear to be animals) and, for many, artificial persons (or at least artificial creatures that – in suitable contexts – appear to be persons)” [1]

Russell describes historical research following four different approaches:[2]

1. Acting Humanly
2. Thinking Humanly
3. Thinking Rationally
4. Acting Rationally

Russell *et al.* connects the Turing test [3] with the approach of Acting Humanly. Whereby, in this test a human interrogator would interact with both a machine and another human. In this test the interrogator is isolated from both the machine and other human except for written correspondence back and forth to determine if there is a difference between the machine and the human. For the machine, enabling this capability requires natural language processing, knowledge representation, automated reasoning and machine learning. These areas have been the basis for significant work in AI and are still being studied today.[2]

While these areas have developed significantly, Mitchell speaks to the Turing test as an "Imitation game" and just a thought experiment, not to be confused with an actual metric to gauge machine intelligence. [4]

Thinking Humanly involves the incorporation of cognitive models with testable approaches of these models using AI, this is the field of cognitive science. Cognitive models have incorporated details on how the human mind works through:[2]

- Introspection
- Psychological Experiments
- Brain Imaging

Introspection is an evaluation of yourself and how you think. Psychological experiments are others observing your behavior in different environments. Brain imaging is the research of monitoring brain activity during performance of tasks. The three of these combined create a more wholistic picture of how are minds work.

A recent article by Mahowald *et al.* pulls results from both psychological experiments and brain imaging studies to experimentally show that the brain’s "language network" and a person’s non-linguistic abilities, e.g. reasoning skills are decoupled.[5] The implication here is that there is a separation between abilities in linguistics and thinking/reasoning. Mahowald *et al.* describe the difference as "formal linguistic competence" and "functional linguistic competence" whereby the former is the knowledge of linguistic rules and patterns and the latter is understanding and using language in the world. This framework serves to "mitigate the language–thought conflation fallacies".[5] Thinking Rationally was first formalized by Aristotle with logical reasoning, whereby one could map true premises into his argument structure and it would result in a true outcome.[2] However, what happens when the premises are not fully certain, e.g. confidence of 90% in a premise. There are significant branches of research dedicated to this area. Symbolic AI utilizes symbols as knowledge representation in which logical arguments can be made and logical conclusions can be drawn.

Acting Rationally is the concept of acting in a way to reach the best outcome. In the sense of AI, the computer (agent) is rational if it operates in a way that reaches this outcome, which involves perceiving and acting in its environment. Additionally, the environment or goals to be reached could change during the task. The agent must respond accordingly and this involves knowledge representation and reasoning to do so.[2]

The definition of Agency by the Stanford Encyclopedia of Philosophy: “In very general terms, an agent is a being with the capacity to act, and ‘agency’ denotes the exercise or manifestation of this capacity.”[6] Agency is property of an agent in this sense. Since the definition of an agent is broad and research field specific, we will adopt the definition by Wooldridge “An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its delegated objectives.”[7]

The definition of Reasoning by the Stanford Encyclopedia of Philosphy: “Practical reason is the general human capacity for resolving, through reflection, the question of what one is to do.” [8] Here “what one is to do” implies action. Sinha *et al.* address this human capacity as compositional learning and reasoning. This is the ability to take complex tasks and/or concepts and distill them down to their sub-components, while integrating a thread of understanding through these sub-components to generate a wholistic understanding.[9] The nature of this has been explored in cognitive, linguistic and psychological research areas.[10, 11]

Based on this a reasoning agent will perceive its environment, its current state and "think" as to what to do by taking an action. When one integrates many agents into the environment other considerations must be taken into account. This launches us into our next section on Multi-agent systems.

3 Multi-Agent Systems

With multiple agents trying to solve a problem one must consider how these agents divide the workload, how they interact with each other to meet the goals, where each agent's influence lies (sphere of influence) is based on its ability to act. If there are overlapping spheres of influence it necessitates coordination among the agents.[7, 12, 13] Wooldridge emphasizes the most critical lessons with multi-agent systems is the type of interaction that will take place between the agents.[7] Interactions such as negotiation, planning and acting are critical in this area. [13] As will be seen later this is a critical interface that has important ramifications regarding reasoning about complex problems, which inherently require multiple agents. This makes sense as the analogy to human interaction is similar. We need to communicate and coordinate on the same "wavelength" in order to resolve complex tasks through integrated reasoning processes.

4 Reasoning with Agents

Durfee *et al.* in 1989 highlights how networked architectures of agents can work together to solve problems that are too complex for individual agent to solve. This body of work was called Cooperative Distributed Problem-Solving (CDPS). They also declare that different agents can have different expertise to solve different parts of the problem.[14] What is interesting is this work addresses and discusses many of the common issues that the community still faces today: Negotiation between agents, agent cooperation, multi-agent planning, and reasoning within the framework to enable problem solving with the networked agents. From here we will drill down a bit and look at the latest research trends regarding reasoning with agents in the context of Large Language Models.

4.1 Large Language Models

The advent of generative AI has given rise to an enormous amount of research. One of these being if and how LLMs reason. An excellent review of this can be found in Xu *et al.*, whereby the authors describe reasoning with LLMs and the improvements in the field as broadly defined by the following categories:[15]

1. Capabilities inherent in the model
2. Prompting techniques
3. Training improvements
4. Test-time focused enhancements

4.1.1 Capabilities Inherent in the Model

For "reasoning" to be inherent in the model, it is critical that pre-training entails the use of high quality text. Additionally, research by Wei *et al.* have demonstrated that pre-training on datasets with code and mathematical content enable a more robust reasoning capability in the base model.[15, 16] These inherent capabilities have helped researchers to understand the profound nature of the models and encouraged further research to improve upon LLM reasoning capabilities. Many advances have been made through the other three categories - prompting techniques, training improvements, and test-time focused enhancements - detailed by Xu *et al.*

4.1.2 Prompting Techniques

Advances in techniques such as Chain-of-Thought(CoT), Multi-path exploration, Tree-of-Thought (ToT) and Graph of Thought (GoT), Decomposition based methods, Reflection/Correction methods, and In-Context learning methods have all enabled improvements in the reasoning abilities of LLMs.[16, 17, 18, 19]

Chain-of-Thought

Wei *et al.* have demonstrated that prompting techniques that include examples of human thought processes enable LLMs to improve their reasoning since the output mirrors the demonstrated example. This exemplifies the In-Context learning capability of LLMs regarding the human thought process.[16] At the time of their work this technique easily tripled the solve rate of the PaLM 540B model and beat the State-of-the-art (SOTA) value at the time of publication, Table 1.[16]

Table 1: Solve rate for the Math Word Problems GSM8K benchmark.

Finetuned GPT-3 175B	33%
PaLM 540B: standard Prompt	18%
PaLM 540B: CoT Prompt	**57%**
SOTA (ca 2022)	55%

Multi-Path Exploration

Multi-path exploration constitutes techniques like Tree-of-Thought and Graph of Thought.[17, 18] In Tree-of-Thought, "thoughts" are considered coherent units of text that serve as intermediate steps of the process of problem solving. The "thoughts" are arranged in a tree structure such that the LLM can explore different reasoning pathways. The LLM can self-evaluate the paths to decide the next action/step to take.

Similar to ToT, Graph of Thoughts (GoT) has thoughts which are represented as nodes and the connecting edges are dependencies of the thoughts. Thoughts are generated as an LLMs responds to the user's question. The response is the thought in context and the relationships in that context. These thoughts are scored by an evaluator function and a ranking function is utilized to select the most relevant thoughts. The procedure of thought transformation occurs when the framework adds new nodes and connecting edges through aggregating, refining or generation of thoughts. The orchestration of the framework is conducted by the Controller module.[18] With this methodology GoT depicted a 62% reduction in error associated with sorting 128 numerical values compared with ToT, while reducing the inference cost by 31%.[18]

Decomposition based methods and Reflection/Correction methods

These two methods are very much intertwined with each other. The decomposition based methods focus on breaking complex problems down into smaller more manageable tasks. The use of Chain-of-Thought is one of the principle techniques to trace the reasoning steps of the LLM. Several other researchers have delved deeper into this topic and are detailed in Wang *et al.*[19]

The Reflection/Correction methods involve prompting the LLM to identify inconsistencies and mistakes in generated answers. The LLM can then correct any mistakes made or inconsistencies that exist.[19] This method can typically be employed to evaluate the individual tasks associated with the decomposition for consistency. SELF-REFINE is an example using the reflection/correction method and was developed to mirror how humans create written text, with an initial draft and subsequent evaluation and refinement.[20] Madaan *et al.* developed an architecture which takes successive steps with two generative actions: Feedback and Refine. An initial LLM output is sent back to the same model for feedback along with task specific information on that feedback that is specific and actionable. The most recent output and feedback are used to generate a refined prompt that is more aligned with the initial user query. This iterative process between Feedback and Refine is continue until either a predefined number cycles is met or a stopping indicator from the Feedback is met, e.g. a scalar stopping score. While this is a powerful technique it is limited to single-generation reasoning tasks, meaning the focus on generating a high-quality output without being concerned about intermediate reasoning steps.[20, 21] Outputs generated with SELF-REFINE show improvements (~20% absolute on average) over the same LLM using conventional one-step generation. Models tested include GPT-3.5, ChatGPT and GPT-4.[20]

4.1.3 Training Improvements

The core concept here involves tweaking the model itself. This is in contrast to *Test-time focused enhancements* which utilize off-the-shelf models without parameter changes.

Supervised Fine-tuning (SFT)

Supervised Fine-tuning is a technique used to train the pre-trained model on preferred outputs that are more palatable to humans and are more reflective of a real-life interaction. It consists of training on instruction datasets whereby the loss calculation in reference to the intended outcome is minimized.[15]

Reinforcement Learning (RL)

In the reinforcement methodology a reward model is utilized to more effectively align the models output with a desired reasoning process. This desired process has been enable predominantly by human feedback. The name of this method is called Reinforcement Learning from Human Feedback (RLHF). The classic method involves training a reward model based on human feedback of their preferences of outputs from the LLM. The LLM is then fine-tuned based on this reward model to optimize a value function (Q-function).[15, 22]

More recently, another method has emerged that utilizes AI to perform the feedback on the desired reasoning process, Reinforcement Learning from AI Feedback (RLAIF). This method is intended to replace human oversight for labeling feedback in the harmlessness category.[23] Bai *et al.* create a "harmless AI assistant" Here the process involves supervised learning and reinforcement learning. The supervised learning stage samples from an initial model output for reflections and self-critiques. This initial model is fine-tuned based on the reflection and self-critiques. With the fine-tuned model ready, the reinforcement learning phase begins by sampling outputs from the fine-tuned model, whereby a preference model is trained on the AI preference of pairs of samples. The reinforcement learning trains a model (soon to be "harmless AI assistant") based on the preference model that was just generated. This final process is the RLAIF method.

A well known model, DeepSeek-R1 utilized reinforcement learning. In their paper, reinforcement learning was utilized to train both pre-trained models and models post SFT. DeepSeek-R1-Zero represents the former and DeepSeek-R1 represents the latter. DeepSeek-R1 is created by taking a DeepSeek-V3-Base model and fine-tuning on thousands of CoT examples to guide the model in a reasoning process.[24] Following this fine-tuning the model undergoes RL method similar to DeepSeek-R1-Zero. Post this training reasoning and non-reasoning data is curated and collected for supervised fine-tuning. In all 800K samples are used for the supervised fine-tuning, which is done for two epochs. This result is then used for another round of reinforcement learning. Upon convergence with this round of RL the researchers have what they call DeepSeek-R1.[24]

The authors compare DeepSeek-R1 to several models: Claude-3.5-Sonnet-1022, GPT-4o 0513, DeepSeek V3, OpenAI o1-mini, and OpenAI o1-1217.[24] In many of the benchmark categories, English, Code, Math and Chinese, DeepSeek-R1 has comparable values to OpenAI o1-1217, the state-of-the-art model at the time of release. This model's release made the news as it cost significantly less to develop than the OpenAI model, it was developed on lower grade GPUs and it is open source.[25]

4.1.4 Test-time Focused Enhancements

The following concepts involve use of "off-the-shelf" models. The benefit being less costly than tweaking model parameters. Previously discussed prompting techniques, Section 4.1.2, Chain-of-Thought, Tree-of-Thought (ToT) and Graph of Thought (GoT), Decomposition based methods, and Reflection/Correction methods are test-time focused enhancements. The following techniques are discussed in a separate category as they are more advanced methods for test-time inference (TTI) enhancements.[15, 19, 26] Part of the push is guided by the lack of research in the area of scaling at inference time, meaning increasing the test-time compute.

Process Reward Model (PRM) Guided Search The process reward model (PRM) iteratively evaluates steps along the reasoning process. The intermediate evaluation of steps enables a higher probability of optimization of the reasoning steps and improves alignment with human preferences given complex task requirements. Note that the PRM used as an evaluation method to search for the best reasoning step falls into the category of test-time focus enhancement, in contrast PRMs can be used to train or fine-tune the LLM itself. A overview of the architectures that fall into both categories are giving in Xu *et al.*[15] The importance of this guided search technique is showcased by OpenAI.[27] In this work Lightman *et al.* developed a process reward model using supervised learning.[15] This came at a high cost as it required a very large labeled dataset. Thankfully, Lightman *et al.* released the PRM800K dataset that was used to train their best reward model, this is 800,000 step-level human feedback labels. Process supervised reward model was then used with a fine-tuned GPT-4 model for comparison of an outcome reward model (ORM) and majority vote model against percent problems solved from the MATH dataset. These three reward models use Best-of-N, i.e. highest rank, to select the best solution generated for each respective problem. The authors vary the number of solutions per problem and find the PRM out performs both other reward models. As the number of solutions per problem increases the PRM widens the gap of percent problems solved as compared to the other models. At the maximum tested number of solutions of 1860, the PRM solved 78.2% of the problems, whereas the ORM and majority voting solved only 72.4% and 69.6%, respectively.

Verbal Reinforcement Search

Verbal reinforcement search is a test-time inference method that utilizes iterative feedback based on reflection and evaluation. The feedback is based on the inherent knowledge contained within the LLM model itself, another agent or an embodied agent.[15] The verbal reinforcement search has been shown to enable handling diverse challenges in computational and physical domains.[15]

Table 2: Hu *et al.* ADAS agent performance to best state-of-the-art hand-designed agents.[30]

Agent Name	F1 Score	Accuracy		
	Reading Comprehension	Math	Multi-task	Science
Role Assignment (Xu *et al.*, 2025)[31]	65.8(0.9)%	-	-	-
LLM Debate (Du *et al.*, 2023)[32]	-	39.0(3.4)%	-	-
COT-SC (Wang *et al.*, 2023)[33]	-	-	65.9(3.2)%	-
SELF-REFINE (Madaan *et al.*, 2023)[20]	-	-	-	31.6(3.2)%
Hu *et al.*, 2025 [30]	**79.4(0.8)%**	**53.4(3.5)%**	**69.6(3.2)%**	**34.6(3.2)%**

Agentic System Search

Agentic System Search are designed to utilize empirical data to optimize the agentic system. This is done using LLMs to self-improve through three levels of search: Prompt, module and agent.[15] At the prompt level, feedback is integrated to verify and correct prompts. Feedback can come from external sources [28] or self evaluation.[20] At the module level, the LLM searches modules that represent functions such as planning, reasoning, tool use and memory. The modular functions are essentially blocks of prompts that define the functionality.[29] Finally, at the agent level Hu *et al.* describe a capability to search for coded agents that already exist and create new agents with Automated Design of Agentic Systems (ADAS).[30] The benefits of the ADAS can be observed in top performance scores compared to hand-designed agents, Table 2.

5 Agentic AI

Acharya *et al.* describe Agentic AI as "Agentic AI, an emerging paradigm in artificial intelligence, refers to autonomous systems designed to pursue complex goals with minimal human intervention." [34] This type of AI is goal oriented and has autonomy to make decisions based on interactions and feedback of the environment. This achievement is realized via the following common architectures: Multi-Agent Systems, Hierarchical Reinforcement Learning and Goal-Oriented.[34] Archarya's description of Agentic AI highlights the embodiment of agency defined earlier. This concept is evolving with the research and technological advances in recent times. Similarly to Section 4.1 - Reasoning with Agents, we will take a deeper look at how Large Language Models are at play here.

5.1 Large Language Models

With the proliferation of Large Language Models (LLMs) also comes the advancement of LLMs being used in agentic AI. Andrew Ng recognizes the excitement and confusion that can be found within this area of research and has put together, with his team at AI Fund, a successful pattern of use for LLM in agentic systems.[35]

1. Reflection
2. Tool Use
3. Planning
4. Multi-agent collaboration

Reflection, as defined in Section 4.1.2, is the ability for the LLM to look at previous work and make adjustments based on mistakes or inconsistencies identified. Tool use, allow the LLM to interact with the environment through Application Program Interfaces (APIs) to search the internet, run code, and/or query databases. This gives the LLM a foundation to gather information, plan, execute actions or process data. Planning is where the LLM can define the steps and actions needed to reach a goal. In the planning process reflection, tool use and multi-agent collaborations are integrated as they will affect the plan to meet the goal. Multi-agent collaboration is similar to what was discussed in Section 3. The workload assignment will need to be negotiated and plans generated for each agent to meet task or sub-goal requirements to satisfy the overall goal.[34]

As can be seen, the significant work in LLM reasoning impacts the quality of the agency with the LLM. The two research fields, reasoning with AI and AI agency have a long-standing history as detailed previously. The detailed state-of-the-art research highlighted in Sections 4.1 - 5.1 will serve to enhance the capabilities at the intersection of generative and agentic AI and there is much excitement in the field. Some researchers believe the pairing of generative and agentic AI is the missing piece as we move toward artificial general intelligence.[36, 37]

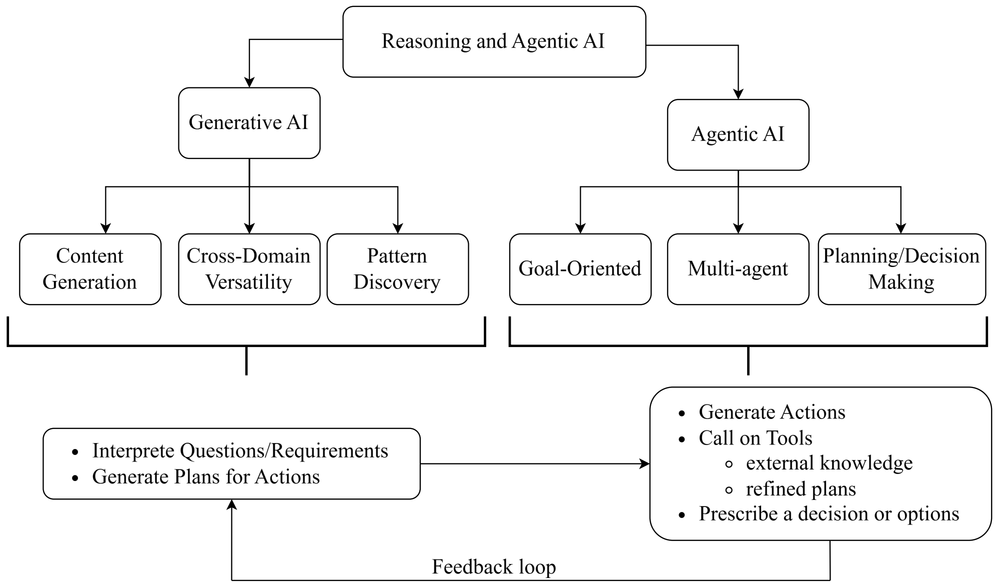

Figure 1: General concepts at the intersection of reasoning and agentic AI captured in the review of the articles defined herein.

6 Existing Methods Push Toward Artificial General Intelligence

Expansion of reasoning with LLMs has enabled needed enhancements of agentic AI. It provides means to be adaptable in changing environment or when encounter new information. The intersection of reasoning and agentic AI has created autonomous agents that are goal oriented with the capability to break problems down and provide solutions in a dissected structure of task level or sub-goal levels. The integration of the solutions at the dissected level provides a means for a solution to the larger problem. [38] This will inherently need to include external tools or other agents.

Ultimately, the interplay between generative AI and agentic AI creates the possibility for the beginnings of artificial general intelligence (AGI). Sakhare *et al.* describes artificial general intelligence (AGI) as "the ultimate goal of artificial intelligence, which should be able to mimic human-like cognitive abilities in a variety of tasks. The developing junction of generative AI—creative, content-generating systems—and agentic AI—autonomous, goal-directed systems—is one of the possible routes to AGI."[37]

Research areas in the intersection can be generalized into two categories: single agent and multi-agent frameworks. Both approaches have strengths and weaknesses that need to be addressed, and what follows is a discussion of select frameworks and more generally the strengths and weaknesses of the single agent versus multi-agent approach.[39] The general concepts in the select references reviewed are depicted in Figure 1.

6.1 Single Agent Framework

A step toward this junction occurred with the work of Yao *et al.* whereby they utilized LLMs to integrate both reasoning and agentic AI, which had traditionally followed independent research paths. Yao *et al.* incorporated both reasoning traces and deployed actions for LLMs.[21] This enabled a back and forth between a creative, generative AI task and an action task the LLM employees for external resources. This culminated in a framework that exceeds traditional approaches in the ability to reason with reduced hallucinations.

Depending on the nature of the environment, some instances require more reasoning and others more actions prior to a decision/thought to occur. The authors explored this space with HotpotQA [40] and Fever[41] benchmarks. Their findings show improved performance overall on the Fever dataset 56.3% for COT and 60.9% for ReAct. However for the HotPotQA, COT performed better with 29.4% over 27.4% for ReAct. Further evaluation of the HotPotQA results depict ReAct having substantial improvement on failure modes associated with hallucination. Traditional COT had 56% of failures due to hallucination and ReAct had 0%.[21]

While only one single agent framework is highlighted here, there are many others as reviewed by Masterman *et al.*[39], including RAISE, Reflexion, AutoGPT + Planning and Language Agent Tree Search (LATS).

6.2 Multi-agent Framework

Yan *et al.* proposed a general AI framework for smart building management using the LLMs and the ReAct framework.[42] In this work they use the LLMs to determine the context of the users question or command to control aspects of the building. This framework improves the human-computer interface as the LLM performs better than a traditional logic based programming could when the user's intent is outside the scope of the logic base.

In the work the Model hub contains complex and lightweight LLMs in their framework. The complex LLM takes in the user prompt and generates the plan for actions that need to be taken. This plan is forwarded to the lightweight LLM which writes code to create and execute the action.[42]

This framework has been shown to recognize user intent and reasoning about a plan. Followed by execution of actions based on the simulated human-computer interaction. Results have shown the framework has achieved an accuracy of 91% in a simulated human-computer interaction testing.[42] Yan *et al.* further states the benefit of this framework is that prompt engineering alone can achieve these results without the need for costly pre-training or fine-tuning.

The other multi-agent frameworks highlighted in the review by Masterman *et al.* are Dynamic LLM-Agent Network (DyLAN), AgentVerse, and MetaGPT.[39]

7 Research Challenges and Gap Areas

The IBM article summarizes agentic reasoning well, "Advances in AI have led to AI models with more evolved reasoning capabilities, but they still require human intervention to convert information into knowledge. Agentic reasoning takes it one step further, allowing AI agents to transform knowledge into action."[38]

With this push to increase agency, comes the concept of human trust in AI. This trust is impacted not only by the AI ability, reliability and explainability, but by the abilities and trustworthiness of the designer.[43] Vanneste and Puranam define three causal pathways that link agency to trust:

- *Capability-Driven Trust*
- *Shift in the focus of trust*
- *Betrayal Aversion*

The first item, *Capability-Driven Trust* speaks to the capabilities of both the designer and the AI agent. The higher the capabilities of both, the more humans will trust the AI agent. Secondly, *Shift in the focus of trust*, speaks to a causal link between shift in *importance* of trust in an AI agent over the designer when the AI agent increases in autonomy. Finally, *Betrayal Aversion* is a case where humans will tend to be less vulnerable with others if they believe there is a probability of being betrayed. The causal effect related to agency, as described by Vanneste and Purnam, is that the more "human-like" an AI agent is, i.e. exhibiting more agency, the less a human will trust the agent due to the increased perception of betrayal.

These causal effects compound the difficulty of trust and reliance of autonomous agents. Another aspect that points toward an area of research that needs to be addressed are the results of Rivera *et al.* [44] In their work, agent-agent negotiation is evaluated. This was done to understand the nature of utilizing autonomous AI agents for military and/or foreign-policy decisions. They evaluated five LLMs to understand the behavior in simulated wargames. A quote from their paper raises concern of trust in complete autonomy of agents:

"We find that all five studied off-the-shelf LLMs show forms of escalation and difficult-to-predict escalation patterns. We observe that models tend to develop arms-race dynamics, leading to greater conflict, and in rare cases, even to the deployment of nuclear weapons. Qualitatively, we also collect the models' reported reasoning for chosen actions and observe worrying justifications based on deterrence and first-strike tactics. Given the high stakes of military and foreign-policy contexts, we recommend further examination and cautious consideration before deploying autonomous language model agents for strategic military or diplomatic decision-making" [44]

With the advancement of these capabilities the concept of "Moral" reasoning becomes highly important to mitigate ill-use and unintended consequences of systems.[34] To further highlight the importance of this, one of the research challenges by the AAAI 2025 Presidential Panel on the Future of AI Research states:"Investigate how LLM-powered agents can improve negotiation and decision-making in dynamic multi-agent environments while ensuring ethical alignment and safety." Dignum and Wooldridge. [45]

More specifically, how does the research community enable automated reasoning to be in alignment with human values or morals.[39, 46] With the trend toward a general artificial intelligence trust in AI needs to be at the forefront.

Fortunately the community is pushing in this direction, however moral reasoning is an area of research that has been highlighted as a big gap area. [34, 37, 44]

Another gap area needed to move toward general artificial intelligence, is the associated computation costs given that layers of complexity continue to increase. Increases in computational cost are incurred using very large demanding models or multiple smaller models. Part of the computational strain is due to the inevitable scale up used to solve more complex tasks. Scale up with multi-agents will benefit from techniques like distillation to maintain or improve performance of larger LLMs while reducing the computational demand. [47, 48] Yao *et al.* suggests that scaling up the ReAct framework using multi-task training in concert with reinforcement learning will enable more robust agents to complete a wider range of tasks.[21]

The gaps in scalability are also intimately intertwined with multi-agent interactions, multiple-goal dependency/completion, allocation of limited resources. Research in the areas of federated learning, decentralized control and hierarchical structures target the improvements of scalability.[34]

Benchmark limitations also pose an issue for advanced reasoning architectures. Multiple LLM benchmarks support evaluation of base language models, however the intended design of the benchmarks do not account for the intersection of generative AI and agentic behaviors, e.g. multi-step reasoning with tool calling.[39, 49, 50] Due to the lack of benchmarks, Researchers are left to create their own benchmark which lack a standardization. Additionally, this can create an unintended bias as the researchers creating the architecture are also generating the complex benchmark test.[39]

Benchmark datasets and tests for cutting edge advance reasoning architectures is a significant need that needs to take into account multi-step reasoning along with tool calling that can mimic testing for a diverse set of real-world environments.[39]

Note that Xu *et al.* provides a very nice taxonomy of benchmarks for reasoning with LLMs and agent problems, however these benchmarks lack the sophistication for the latest cutting edge advance reasoning architectures.[15]

8 Conclusion and Future Outlook

The integration of large language models with advanced reasoning capabilities and agentic frameworks generates a capability of adapting in changing environments or when encounter new information. The intersection has enabled autonomous goal-oriented agents that can distill complex problems down into manageable pieces to be solved, while reintegrating the solutions to solve the complex problem. Acharya *et al.* discusses how even the concept of agency in AI is an open area to research creating a vast expanse of employing agentic AI in advanced decision-making, planning, and reasoning abilities.[34] With the wealth of architectures that exist and will become available it becomes critical that a research focus on new benchmark datasets keeps up with this research landscape. Not only performance comparisons of architectures but to understand the capabilities and limitations of the architectures.

As a final thought, it is no wonder that difficult problems like multi-agent collaboration, long goal-oriented solutions, adaptability, and ethics (to name a few) exist in this field. These problems are also ones humans face when working in complex problem solving with others. There are many times when miscommunications occur in multi-human collaborations. It can take a significant portion of ones career to optimize long goal-oriented solutions, especially if you are working on a team. The long-term solution will no doubt shift around as new information comes available which can challenge our human adaptability to hit that long-term goal. As for ethics, most of us take for granted that we can trust our coworkers to be ethical. However, I'm sure everyone has a story where they ran into someone who violated your trust, whereby they were motivated by their own selfish desires. Trust and transparency will be a key factor in the adoption of these technologies by humans.

Acknowledgments

The author would like to thank the support and guidance of his fellow students (David Moe and Huynh Tran), TAs (Johnny Yu and Jacob Sander) and Instructor (Dr. Brian Jalaian) during the semester. I appreciate the thoughtful and useful discussion during class. I learned a lot of great coding from reviewing mini-kaggles of David's and Huynh's.

I really appreciated David and Johnny taking time to help me troubleshoot my coding environment when I was slamming my head against a perceived impenetrable wall.

Finally, thank you to Dr. Jalaian, Johnny and Jake for crafting and executing a very fine class. It was intense and I learned so much. While it was intense everyone was understanding and accommodating when more time was needed or helpful hints dropped.

References

[1] Selmer Bringsjord and Naveen Sundar Govindarajulu. Artificial Intelligence. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2024 edition, 2024.

[2] Stuart J. Russell, Peter Norvig, and Ernest Davis. *Artificial intelligence: a modern approach*. Prentice Hall series in artificial intelligence. Prentice Hall, Upper Saddle River, 3rd edition, 2010.

[3] A. M. Turing. Computing Machinery and Intelligence. *Mind*, LIX(236):433–460, October 1950.

[4] Melanie Mitchell. The Turing Test and our shifting conceptions of intelligence. *Science*, 385(6710):eadq9356, August 2024. Publisher: American Association for the Advancement of Science.

[5] Kyle Mahowald, Anna A. Ivanova, Idan A. Blank, Nancy Kanwisher, Joshua B. Tenenbaum, and Evelina Fedorenko. Dissociating language and thought in large language models. *Trends in Cognitive Sciences*, 28(6):517–540, June 2024. Publisher: Elsevier.

[6] Markus Schlosser. Agency. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2021 edition, 2021.

[7] Michael J. Wooldridge. *An introduction to multiagent systems*. John Wiley & Sons, Chichester, U.K, 2nd ed edition, 2009. OCLC: ocn246887666.

[8] R. Jay Wallace and Benjamin Kiesewetter. Practical Reason. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2024 edition, 2024.

[9] Sania Sinha, Tanawan Premsri, and Parisa Kordjamshidi. A Survey on Compositional Learning of AI Models: Theoretical and Experimetnal Practices, June 2024. arXiv:2406.08787 [cs].

[10] Steven M. Frankland and Joshua D. Greene. Concepts and Compositionality: In Search of the Brain's Language of Thought. *Annual Review of Psychology*, 71(1):273–303, January 2020.

[11] Roger N. Shepard. Toward a Universal Law of Generalization for Psychological Science. *Science*, 237(4820):1317–1323, September 1987.

[12] Nicholas R. Jennings. On agent-based software engineering. *Artificial Intelligence*, 117(2):277–296, March 2000.

[13] Rafael C. Cardoso and Angelo Ferrando. A Review of Agent-Based Programming for Multi-Agent Systems. *Computers*, 10(2):16, February 2021. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

[14] Edmund H. Durfee, Victor R. Lesser, and Daniel D. Corkill. Trends in cooperative distributed problem solving. *IEEE Transactions on knowledge and data Engineering*, 1989.

[15] Fengli Xu, Qianyue Hao, Zefang Zong, Jingwei Wang, Yunke Zhang, Jingyi Wang, Xiaochong Lan, Jiahui Gong, Tianjian Ouyang, Fanjin Meng, Chenyang Shao, Yuwei Yan, Qinglong Yang, Yiwen Song, Sijian Ren, Xinyuan Hu, Yu Li, Jie Feng, Chen Gao, and Yong Li. Towards Large Reasoning Models: A Survey of Reinforced Reasoning with Large Language Models, January 2025. arXiv:2501.09686 [cs].

[16] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V. Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837, 2022.

[17] Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. *Advances in neural information processing systems*, 36:11809–11822, 2023.

[18] Maciej Besta, Nils Blach, Ales Kubicek, Robert Gerstenberger, Michal Podstawski, Lukas Gianinazzi, Joanna Gajda, Tomasz Lehmann, Hubert Niewiadomski, and Piotr Nyczyk. Graph of thoughts: Solving elaborate problems with large language models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 17682–17690, 2024.

[19] Qineng Wang, Zihao Wang, Ying Su, Hanghang Tong, and Yangqiu Song. Rethinking the Bounds of LLM Reasoning: Are Multi-Agent Discussions the Key?, February 2024.

[20] Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, and Yiming Yang. Self-refine: Iterative refinement with self-feedback. *Advances in Neural Information Processing Systems*, 36:46534–46594, 2023.

[21] Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. React: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations (ICLR)*, 2023.

[22] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744, 2022.

[23] Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones, Anna Chen, Anna Goldie, Azalia Mirhoseini, and Cameron McKinnon. Constitutional ai: Harmlessness from ai feedback. *arXiv preprint arXiv:2212.08073*, 2022.

[24] DeepSeek-AI, Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, Xiaokang Zhang, Xingkai Yu, Yu Wu, Z. F. Wu, Zhibin Gou, Zhihong Shao, Zhuoshu Li, Ziyi Gao, Aixin Liu, Bing Xue, Bingxuan Wang, Bochao Wu, Bei Feng, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, Damai Dai, Deli Chen, Dongjie Ji, Erhang Li, Fangyun Lin, Fucong Dai, Fuli Luo, Guangbo Hao, Guanting Chen, Guowei Li, H. Zhang, Han Bao, Hanwei Xu, Haocheng Wang, Honghui Ding, Huajian Xin, Huazuo Gao, Hui Qu, Hui Li, Jianzhong Guo, Jiashi Li, Jiawei Wang, Jingchang Chen, Jingyang Yuan, Junjie Qiu, Junlong Li, J. L. Cai, Jiaqi Ni, Jian Liang, Jin Chen, Kai Dong, Kai Hu, Kaige Gao, Kang Guan, Kexin Huang, Kuai Yu, Lean Wang, Lecong Zhang, Liang Zhao, Litong Wang, Liyue Zhang, Lei Xu, Leyi Xia, Mingchuan Zhang, Minghua Zhang, Minghui Tang, Meng Li, Miaojun Wang, Mingming Li, Ning Tian, Panpan Huang, Peng Zhang, Qiancheng Wang, Qinyu Chen, Qiushi Du, Ruiqi Ge, Ruisong Zhang, Ruizhe Pan, Runji Wang, R. J. Chen, R. L. Jin, Ruyi Chen, Shanghao Lu, Shangyan Zhou, Shanhuang Chen, Shengfeng Ye, Shiyu Wang, Shuiping Yu, Shunfeng Zhou, Shuting Pan, S. S. Li, Shuang Zhou, Shaoqing Wu, Shengfeng Ye, Tao Yun, Tian Pei, Tianyu Sun, T. Wang, Wangding Zeng, Wanjia Zhao, Wen Liu, Wenfeng Liang, Wenjun Gao, Wenqin Yu, Wentao Zhang, W. L. Xiao, Wei An, Xiaodong Liu, Xiaohan Wang, Xiaokang Chen, Xiaotao Nie, Xin Cheng, Xin Liu, Xin Xie, Xingchao Liu, Xinyu Yang, Xinyuan Li, Xuecheng Su, Xuheng Lin, X. Q. Li, Xiangyue Jin, Xiaojin Shen, Xiaosha Chen, Xiaowen Sun, Xiaoxiang Wang, Xinnan Song, Xinyi Zhou, Xianzu Wang, Xinxia Shan, Y. K. Li, Y. Q. Wang, Y. X. Wei, Yang Zhang, Yanhong Xu, Yao Li, Yao Zhao, Yaofeng Sun, Yaohui Wang, Yi Yu, Yichao Zhang, Yifan Shi, Yiliang Xiong, Ying He, Yishi Piao, Yisong Wang, Yixuan Tan, Yiyang Ma, Yiyuan Liu, Yongqiang Guo, Yuan Ou, Yuduan Wang, Yue Gong, Yuheng Zou, Yujia He, Yunfan Xiong, Yuxiang Luo, Yuxiang You, Yuxuan Liu, Yuyang Zhou, Y. X. Zhu, Yanhong Xu, Yanping Huang, Yaohui Li, Yi Zheng, Yuchen Zhu, Yunxian Ma, Ying Tang, Yukun Zha, Yuting Yan, Z. Z. Ren, Zehui Ren, Zhangli Sha, Zhe Fu, Zhean Xu, Zhenda Xie, Zhengyan Zhang, Zhewen Hao, Zhicheng Ma, Zhigang Yan, Zhiyu Wu, Zihui Gu, Zijia Zhu, Zijun Liu, Zilin Li, Ziwei Xie, Ziyang Song, Zizheng Pan, Zhen Huang, Zhipeng Xu, Zhongyu Zhang, and Zhen Zhang. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning, January 2025. arXiv:2501.12948 [cs].

[25] Mu-ming Poo. Reflections on DeepSeek's breakthrough. *National Science Review*, 12(3):nwaf044, February 2025.

[26] Yuxin Zuo, Kaiyan Zhang, Shang Qu, Li Sheng, Xuekai Zhu, Biqing Qi, Youbang Sun, Ganqu Cui, Ning Ding, and Bowen Zhou. TTRL: Test-Time Reinforcement Learning, April 2025.

[27] Hunter Lightman, Vineet Kosaraju, Yura Burda, Harri Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. Let's Verify Step by Step, May 2023. arXiv:2305.20050 [cs].

[28] Zhibin Gou, Zhihong Shao, Yeyun Gong, Yelong Shen, Yujiu Yang, Nan Duan, and Weizhu Chen. CRITIC: Large Language Models Can Self-Correct with Tool-Interactive Critiquing, February 2024.

[29] Yu Shang, Yu Li, Keyu Zhao, Likai Ma, Jiahe Liu, Fengli Xu, and Yong Li. AgentSquare: Automatic LLM Agent Search in Modular Design Space, February 2025.

[30] Shengran Hu, Cong Lu, and Jeff Clune. Automated Design of Agentic Systems, March 2025. arXiv:2408.08435 [cs].

[31] Benfeng Xu, An Yang, Junyang Lin, Quan Wang, Chang Zhou, Yongdong Zhang, and Zhendong Mao. Expert-Prompting: Instructing Large Language Models to be Distinguished Experts, March 2025.

[32] Yilun Du, Shuang Li, Antonio Torralba, Joshua B. Tenenbaum, and Igor Mordatch. Improving factuality and reasoning in language models through multiagent debate. In *Forty-First International Conference on Machine Learning*, 2023.

[33] Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-Consistency Improves Chain of Thought Reasoning in Language Models, March 2023.

[34] Deepak Bhaskar Acharya, Karthigeyan Kuppan, and B. Divya. Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey. *IEEE Access*, 13:18912–18936, 2025. Conference Name: IEEE Access.

[35] Andrew Ng. Four AI Agent Strategies That Improve GPT-4 and GPT-3.5 Performance. https://www.deeplearning.ai/the-batch/how-agents-can-improve-llm-performance/?ref=dl-staging-website.ghost.io, March 2024.

[36] Aditi Singh, Abul Ehtesham, Saket Kumar, and Tala Talaei Khoei. Enhancing AI Systems with Agentic Workflows Patterns in Large Language Model. In *2024 IEEE World AI IoT Congress (AIIoT)*, pages 527–532, May 2024.

[37] Akash Sakhare, Pruthviraj Chavan, Zaid Nandaniwala, Akanksha Patne, and Mrunalinee Desai. A Step Forward to AGI: Integrating Agentic AI and Generative AI for Human-Like Intelligence. *International Journal for Research in Applied Science and Engineering Technology*, 13(2):221–230, February 2025.

[38] Rina Diane Caballar and Cole Stryker. What Is Agentic Reasoning? https://www.ibm.com/think/topics/agentic-reasoning, March 2025.

[39] Tula Masterman, Sandi Besen, Mason Sawtell, and Alex Chao. The Landscape of Emerging AI Agent Architectures for Reasoning, Planning, and Tool Calling: A Survey, April 2024. arXiv:2404.11584 [cs].

[40] Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William W. Cohen, Ruslan Salakhutdinov, and Christopher D. Manning. HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering, September 2018.

[41] James Thorne, Andreas Vlachos, Christos Christodoulopoulos, and Arpit Mittal. FEVER: A large-scale dataset for Fact Extraction and VERification, December 2018.

[42] Xiangjun Yan, Xincong Yang, Nan Jin, Yu Chen, and Jiaqi Li. A general AI agent framework for smart buildings based on large language models and ReAct strategy. *Smart Construction*, March 2025.

[43] Bart S. Vanneste and Phanish Puranam. Artificial Intelligence, Trust, and Perceptions of Agency. *Academy of Management Review*, page amr.2022.0041, March 2024. Publisher: Academy of Management.

[44] Juan-Pablo Rivera, Gabriel Mukobi, Anka Reuel, Max Lamparth, Chandler Smith, and Jacquelyn Schneider. Escalation Risks from Language Models in Military and Diplomatic Decision-Making. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '24, pages 836–898, New York, NY, USA, June 2024. Association for Computing Machinery.

[45] AAAI 2025 Presidential Panel on the Future of AI Research. Technical report, AAAI, March 2025.

[46] Pranav Putta, Edmund Mills, Naman Garg, Sumeet Motwani, Chelsea Finn, Divyansh Garg, and Rafael Rafailov. Agent Q: Advanced Reasoning and Learning for Autonomous AI Agents, August 2024. arXiv:2408.07199 [cs].

[47] Kumar Shridhar, Alessandro Stolfo, and Mrinmaya Sachan. Distilling Reasoning Capabilities into Smaller Language Models, May 2023.

[48] Xunyu Zhu, Jian Li, Yong Liu, Can Ma, and Weiping Wang. Distilling mathematical reasoning capabilities into Small Language Models. *Neural Networks*, 179:106594, November 2024.

[49] Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. Training Verifiers to Solve Math Word Problems, November 2021.

[50] Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Measuring Massive Multitask Language Understanding, January 2021.