Active Reasoning in an Open-World Environment

Manjie Xu[1], Guangyuan Jiang[2], Wei Liang[1,3], Chi Zhang[4], Yixin Zhu[2]
[1]School of Computer Science & Technology, Beijing Institute of Technology, [2]Institute for AI, Peking University
[3]Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing, China, [4]National Key Laboratory of General Artificial Intelligence

NEURAL INFORMATION PROCESSING SYSTEMS

Motivation

Active interaction with the environment is fundamental to human understanding of the world around us.

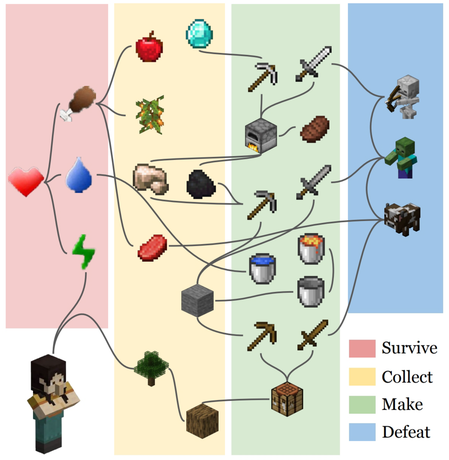
- Both neural and behavioral studies indicate that through active engagement with their surroundings, humans garner critical insights and foster a profound understanding of complex phenomena.
- When confronted with partial or ambiguous data, our innate response is to seek supplementary evidence, hypothesize, and put forth possible explanations, sometimes even reevaluating initial assumptions.

We present Conan to capture the dynamic and exploratory essence of abductive reasoning—termed herein as active reasoning.

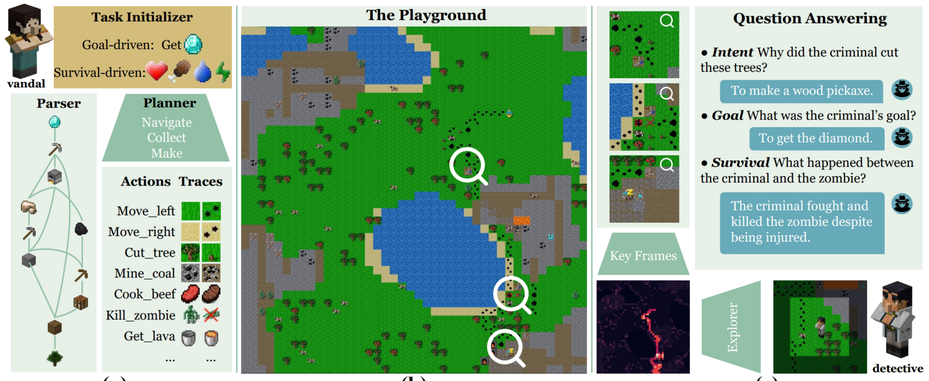
Standing head and shoulders above traditional single-round passive reasoning benchmarks, Conan boasts an open-world arena, urging agents to actively probe surroundings and engage in multi-round abductive inferences, all while leveraging in-situ collected evidence alongside pre-existing knowledge.

The Playground

The playground is where the agents are created and traces left persist.



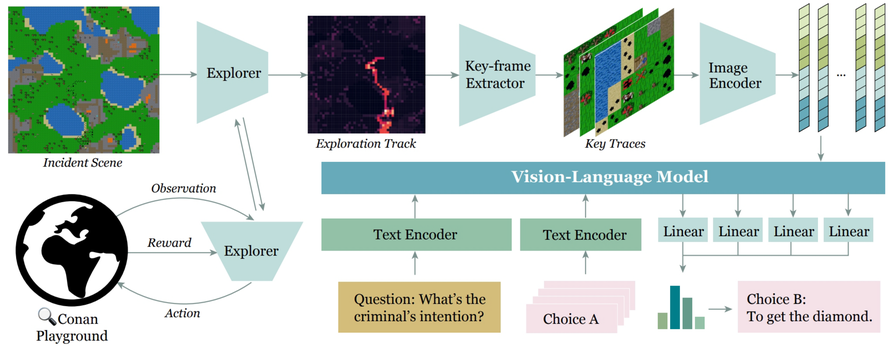
- Conan offers an extensive assortment of interactive items: food, materials, mobs, and tools, each tied to specific actions.
- Advancing from its predecessor, the original Crafter, Conan now features 32 distinct traces covering all agent actions such as crafting, collecting, defeating, eating, drinking, and incurring injuries.
- This enhancement enables the design of 60 varied abductive reasoning tasks within the scene.



Conan Overview

Conan is conceived as a detective game, transmuted into a question-answering challenge.

- The detective is tasked with a query and an "incident scene" riddled with traces left by a vandal.
- Given the initial paucity of conclusive information, the detective must embark on an in-depth exploration of the scene.
- Conan to span various levels of abstraction, from localized intentions (Intent) to overarching objectives (Goal) and survival states (Survival).



The Detective

Conan casts the abductive reasoning challenge as a detective game, necessitating a detective to efficiently explore and gather information from the environment to deduce plausible explanations (i.e., answers) for the given question. Building on previous work that utilizes hierarchical models for task decomposition, our pipeline is structured into two primary phases: an exploration phase for trace collection, followed by an abductive reasoning phase.

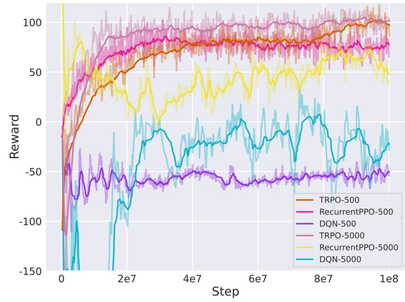
Abduction from Deduction (AfD)

"Set a thief to catch a thief"

For a skillful detective to abduce what a vandal does, it needs an in-depth grasp of vandals' modus operandi, motivations, and decisionmaking process. We articulate the problem of abductive reasoning based on evidence and knowledge from known deductive transitions.

$$P(g \mid O) = \mathop{\mathbb{E}}_{P(S|O)}[P(g \mid S, O)] = \mathop{\mathbb{E}}_{P(S|O)}[P(g \mid S)],$$

$$P(g \mid S) \propto P(S \mid g) \propto \prod_i \pi(a_i \mid s_i, g),$$



The Explorer

We compare DQN , TRPO, and RecurrentPPO as the explorer. TRPO and RecurrentPPO manifest similar performance in terms of rewards following a substantial number of steps, markedly surpassing the DQN explorer.

The VL-Reasoner

We employ a multi-choice question-answering paradigm to solve Conan. Specifically, the model is presented with a question, its corresponding exploration frame sequence, and each potential answer choice, subsequently generating a score for each choice. We evaluate several well-established multimodal models: Vanilla-Trans, FrozenBiLM, and Flamingo-Mini. Our reasoning models are tested under three different settings: Standard, Ideal Explorer, and AfD.

	Standard				Ideal Explorer				AfD			
	I	*G*	*S*	*O*	*I*	*G*	*S*	*O*	*I*	*G*	*S*	*O*
Vanilla-Trans	32.9	25.0	24.5	28.8	64.0	**78.4**	58.1	66.1	24.8	23.3	24.5	24.3
F-BiLM-BERT	72.6	**44.4**	**54.4**	61.0	87.5	59.5	61.5	74.0	82.8	**42.9**	**55.5**	66.0
F-BiLM-DeBERTa	82.9	43.1	52.2	65.3	**87.7**	71.8	**63.9**	**77.8**	82.9	41.9	53.8	65.4
Flamingo-Mini	**86.2**	43.3	49.5	**66.3**	85.8	47.8	56.6	69.0	**84.9**	42.5	52.2	**66.1**

Conclusion

We introduce Conan, a benchmark tailored to evaluate and assess models' active reasoning ability in addressing incomplete-information questions in an interactive environment.

Benchmarking the efficacy of contemporary machine learning models on Conan, we elucidate the model limitations in interacting with the environment that leads to failure in higher-level, longer-term abductive reasoning.