Towards Inquisitive Artificial Intelligence: Autonomous Questioning Large Language Models for Continuous Learning

Sunehildeep Singh
Conceptual Paper
Email: sunehildeep@gmail.com

***Abstract*—Current Large Language Models (LLMs), while exhibiting remarkable proficiency in language tasks, are primarily passive responders to user prompts. They lack intrinsic motivation to explore, question, and autonomously expand their knowledge base. This paper proposes a novel conceptual framework for *Autonomous Questioning Large Language Models (AQ-LLMs)* designed for continuous learning and proactive knowledge acquisition. The core innovation lies in endowing LLMs with the capacity to generate and pursue their own questions, driven by live data feeds and interaction with users and potentially other agents. I outline this conceptual framework for AQ-LLMs, detailing key components such as question generation mechanisms (anomaly detection, novelty seeking, uncertainty-driven inquiry), user/environment interaction for feedback, and learning modules for knowledge refinement. I discuss potential implementation strategies, evaluation metrics, and the broader implications of such architectures for moving towards more inquisitive and generally intelligent AI systems. This work aims to shift the paradigm from reactive LLMs to proactive learners, drawing inspiration from human curiosity and continuous learning processes.**

***Index Terms*—Large Language Models, Artificial General Intelligence, Continuous Learning, Autonomous Questioning, Curiosity-Driven Learning, Active Learning, Inquisitive AI, Neuro-Symbolic AI, Reinforcement Learning.**

I. Introduction

Large Language Models (LLMs) like Deepseek R1 and OpenAI o3 have demonstrated unprecedented capabilities in natural language processing, generation, and understanding. They excel at tasks ranging from text summarization and translation to creative content generation and question answering. However, despite these advancements, current LLMs remain fundamentally passive tools, primarily responding to explicit user prompts. They lack the inherent drive to independently explore, learn, and expand their knowledge in a manner akin to human curiosity and continuous learning.

This limitation hinders the progress towards Artificial General Intelligence (AGI). Truly intelligent agents should not just react to external stimuli but proactively engage with the world, identify knowledge gaps, and autonomously seek information to fill those gaps. Drawing inspiration from human cognitive development, particularly the inquisitive nature of children and their continuous learning through questioning, I propose a novel **conceptual framework** for *Autonomous Questioning Large Language Models (AQ-LLMs)*.

The central premise of AQ-LLMs is to equip LLMs with the capacity to generate and pursue their own questions, enabling them to become active learners rather than passive responders. This is achieved by integrating mechanisms that allow the LLM to:

- **Process Live Data Feeds:** Consume real-time information from diverse sources (text, audio, potentially visual data) to ground their understanding in the current state of the world and identify novel or anomalous information.
- **Generate Autonomous Questions:** Develop internal mechanisms to formulate questions based on perceived knowledge gaps, anomalies in data, uncertainty, or a drive for novelty and deeper understanding.
- **Interact for Feedback:** Engage in dialogues with users (humans), other AI agents, or environmental sensors to seek answers to their generated questions and refine their knowledge.
- **Continuously Learn and Adapt:** Integrate feedback and newly acquired information to update their knowledge base, improve question generation strategies, and refine their overall understanding of the world.

II. Related Work

This research is related to several key areas within Artificial Intelligence and Cognitive Science, including Large Language Models, Active Learning, Reinforcement Learning with Intrinsic Motivation, Curiosity-Driven Learning, Cognitive Architectures, and Neuro-Symbolic AI. These fields provide foundational concepts and motivation for the proposed Autonomous Questioning LLM Agent framework.

III. Conceptual Framework for Autonomous Questioning LLM Agents (AQ-LLM Agents)

This section outlines a conceptual framework for the Autonomous Questioning LLM Agent (AQ-LLM Agent). The core components, described below, are envisioned to enable autonomous question generation, interaction, learning, and continuous knowledge expansion. A visual representation of a possible architecture is omitted in this conceptual paper to maintain focus on the framework itself, but could be developed in future work.

The key modules and their functionalities are detailed below:

A. LLM Core

This is the foundational component, leveraging a state-of-the-art Large Language Model. It provides the base language understanding, generation, and knowledge representation capabilities upon which the autonomous questioning mechanisms are built.

B. Live Data Input Module

This module is responsible for ingesting and pre-processing real-time data streams from various sources. Data types could include:

- **Textual Data:** News feeds, scientific publications, social media (filtered and ethically sourced).
- **Audio Data:** Ambient sound, speech-to-text transcriptions (filtered, ethically sourced).
- **(Future Extension) Visual Data:** Image and video feeds (filtered, ethically sourced and analyzed).

The module filters, cleans, and formats the incoming data for further processing by the AQ-LLM Agent.

C. Question Generation Module

This is the core innovation of the AQ-LLM Agent. This module generates questions autonomously based on:

- **Anomaly Detection:** Identifying deviations from expected patterns or established knowledge in the live data stream. Questions arise to investigate these anomalies.
- **Novelty Seeking:** Recognizing and focusing on new or rare information in the data, driven by a curiosity to explore the unknown.
- **Uncertainty-Driven Inquiry:** Identifying areas where the LLM's internal knowledge is weak or uncertain when processing data. Questions are formulated to reduce this uncertainty.
- **Goal-Oriented Questioning:** If the AQ-LLM has a defined learning goal (e.g., "understand climate change"), questions are generated to actively pursue that goal, driven by information from live data.

The question generation mechanism can employ various techniques, including:

- **Prompt Engineering within the LLM Core:** Utilizing carefully designed prompts to guide the LLM to generate questions based on data analysis.
- **Hybrid Neuro-Symbolic Approaches:** Combining neural network-based anomaly/novelty detection with symbolic reasoning to formulate more structured and logical questions.
- **Reinforcement Learning for Question Strategy Optimization:** Training a separate RL module to learn effective question generation strategies that maximize information gain or learning progress.

D. User/Environment Interaction Module

This module manages the communication interface for the AQ-LLM Agent to ask questions and receive feedback. Interaction modalities could include:

- **User Interface (App/Web):** For direct interaction with human users, presenting questions via notifications and receiving text-based answers and feedback.
- **Agent-to-Agent Communication:** If multiple AQ-LLM agents are deployed, this module enables them to engage in dialogues, ask questions to each other, and share knowledge.
- **Environmental Sensors (Potentially):** For interaction with sensors in physical or simulated environments, receiving sensory feedback relevant to their questions (e.g., for embodied AQ-LLMs).

E. Learning and Adaptation Module

This module is responsible for integrating the feedback received through the interaction module and adapting the AQ-LLM Agent's knowledge and behavior. Learning mechanisms could include:

- **Feedback Integration:** Processing user answers, agent discussions, or environmental feedback to update the LLM's internal knowledge representation. This could involve techniques like knowledge base updates, fine-tuning the LLM (with caution and data curation), or using external memory mechanisms.
- **Curiosity-Driven Reinforcement Learning:** If an RL approach is used for question generation, this module would update the RL policy based on the effectiveness of generated questions in leading to new information and knowledge gain.
- **Meta-Learning for Questioning Strategies:** Exploring meta-learning techniques to enable the AQ-LLM to learn better question generation strategies over time, improving its ability to ask insightful and informative questions.

F. Control and Monitoring Module

This module oversees the overall operation of the AQ-LLM Agent, managing resource allocation, controlling the frequency and type of questions generated, and monitoring the agent's learning progress and behavior. It ensures the agent operates efficiently and ethically.

IV. Proposed Learning Mechanisms

Several learning mechanisms are envisioned to empower the AQ-LLM Agent:

A. Feedback Integration and Knowledge Update

Developing robust methods to integrate diverse forms of feedback (user answers, agent discussions, environmental responses) into the LLM's knowledge is crucial. Techniques like knowledge graph augmentation, semantic parsing of feedback into structured knowledge, and selective fine-tuning of the LLM based on high-quality feedback will be explored.

B. Curiosity and Intrinsic Motivation through Reinforcement Learning

To drive autonomous question generation, I propose incorporating Reinforcement Learning principles with intrinsic reward signals. The reward function could be designed to incentivize:

- **Novelty:** Rewarding questions that lead to the discovery of new information or concepts.
- **Information Gain:** Rewarding questions that maximally reduce uncertainty or expand the agent's knowledge.
- **Engagement and Informative Feedback:** Rewarding questions that elicit detailed and informative responses from users or other agents.

C. Meta-Learning for Questioning Strategies

To further enhance the agent's questioning capabilities, meta-learning approaches could be explored. This would involve training the AQ-LLM to learn optimal strategies for question generation, adapting its questioning style and focus based on its learning experiences and the types of feedback received. This could lead to increasingly sophisticated and effective question-asking abilities over time.

V. Evaluation Framework

Evaluating the effectiveness of AQ-LLMs requires developing novel metrics beyond standard language benchmarks, which primarily focus on passive response tasks. I propose a multi-faceted evaluation framework:

A. Qualitative Evaluation

Human evaluation will be critical to assess the quality, relevance, and insightfulness of the questions generated by the AQ-LLM Agent. User studies can be conducted to evaluate:

- **Question Relevance and Clarity:** Do the questions make sense in context and are they clearly formulated?
- **Novelty and Insightfulness:** Do the questions explore new or interesting aspects of a topic?
- **Learning Value:** Do the questions effectively elicit information that helps the agent learn and expand its knowledge?
- **User Engagement:** Are the questions engaging and interesting for human users to answer?

B. Quantitative Metrics (Exploratory)

Developing quantitative metrics for evaluating autonomous questioning is challenging but crucial for systematic progress. Potential metrics that could be explored include:

- **Question Diversity and Coverage:** Measuring the breadth of topics covered by the agent's questions.
- **Novelty Score (Data-Driven):** Quantifying the "novelty" of questions based on their statistical rarity or distance from the agent's existing knowledge representation.
- **Information Gain Proxy:** Developing metrics to estimate the potential information gain expected from asking specific questions (e.g., based on uncertainty reduction or expected answer informativeness).
- **Learning Progress Metrics:** If integrated with a knowledge base or learning system, tracking the rate of knowledge acquisition or improvement in performance on downstream tasks as a result of autonomous questioning.

VI. Expected Results and Discussion

I anticipate that AQ-LLM Agents, by virtue of their autonomous questioning capabilities, will demonstrate:

- **Enhanced Knowledge Acquisition:** The active pursuit of questions will enable AQ-LLMs to learn more efficiently and comprehensively compared to passively trained LLMs.
- **Improved Adaptability and Robustness:** Continuous learning from live data will allow AQ-LLMs to adapt to changing information landscapes and handle novel situations more effectively.
- **Greater Inquisitiveness and Proactive Behavior:** AQ-LLMs will exhibit a more proactive and curious behavior, actively seeking to understand and explore the world around them, moving beyond mere response generation.
- **Potential for Deeper Language Understanding:** The process of formulating and refining questions may lead to a deeper, more nuanced understanding of language and the world, as the agent actively seeks to resolve ambiguities and fill knowledge gaps.

However, significant challenges are anticipated:

- **Complexity of Implementation:** Designing and integrating the various modules, particularly the question generation and learning mechanisms, will be technically challenging.
- **Evaluation Difficulty:** Developing robust and comprehensive evaluation metrics for autonomous questioning and continuous learning is an open research problem.
- **Ethical Considerations:** Autonomous questioning agents, especially if interacting with live data streams, raise ethical concerns regarding data privacy, bias amplification, and responsible AI development. These aspects will require careful consideration and mitigation strategies.

VII. Conclusion and Future Work

This paper proposes a conceptual framework for Autonomous Questioning Large Language Models (AQ-LLMs), aiming to move beyond passive response generation and towards more proactive, inquisitive, and continuously learning AI systems. By endowing LLMs with the capacity to generate and pursue their own questions, driven by live data feeds and user/environment interaction, I envision a significant step towards more generally intelligent AI.

Future work will focus on:

- **Detailed Design and Implementation:** Developing and implementing specific algorithms and architectures for the proposed modules, particularly the question generation and learning mechanisms.

- **Experimental Validation:** Conducting simulations and user studies to evaluate the performance of AQ-LLM Agents and refine the architecture and learning strategies.
- **Exploration of Embodiment and Multimodality:** Extending the AQ-LLM framework to incorporate embodied interaction and multimodal data processing, further grounding their intelligence in real-world experiences.
- **Ethical and Societal Impact Assessment:** Thoroughly analyzing and addressing the ethical implications of autonomous questioning agents and developing guidelines for responsible development and deployment.

Acknowledgment

I thank the hypothetical AI research community for inspiring discussions and ongoing advancements in the field.