A Design of a Dataset for Human-Like Question Generation under Uncertainty for Interactive Robots*

Minjung Shin[1], Minsu Jang[2], Miyoung Cho[2] and Jeh-Kwang Ryu[3*]

Abstract— **Asking questions is essential for humans and robots in learning and interaction. To enable robots to ask insightful questions, we first describe the cognitive process behind the inquisitive nature of humans and then propose a dataset, CAWS (Curious About Weird Scene), consisting of images, textual descriptions, and questions. We employ a text-to-image generative model to produce images from textual descriptions with factual inconsistencies to incur cognitive uncertainties for robots in understanding the images. Our study highlights the significance of considering humans' inquiry process under uncertainty. We also claim that the proposed dataset can be utilized to improve interactive robot agents and their ability to come up with human-like and insightful questions.**

Index Terms – **Interactive robot, Uncertainty resolution, Question generation, Inquisitive agent, Curiosity, Information gap theory**

I. Introduction

Asking questions for humans plays a significant role in not only mitigating uncertainty from their surroundings [1] but also enhancing the validity of interactions [2]. Specifically, the ability to ask questions allows active learning in cognitive development [3], [4] in line with learning from interaction with others [5]. Also, for robots, users have preferred to speak to robots in human-like languages [6], [7], and asking has been regarded as the best method of cooperative control for interactive robots [8], [9]. Thus, asking questions to eliminate uncertainty, both in terms of learning and interaction, must be an essential ability for an interactive robot as long as it interacts with humans.

However, asking insightful questions is difficult even for humans [10] because question generation is highly context-dependent and has a wide range of variability [11]. Additionally, implementing this ability is more difficult for robots and other artificial intelligence (AI) agents. Accordingly, we claim that investigations are required to explore the principle of humans' questioning under uncertainty to develop AI agents' questioning strategies that resemble humans. Considering this approach, we will begin by reviewing the related work in section II. In section III, we illustrate the inherently curious nature of humans. Lastly, in section IV, we will introduce a dataset framework that can be used to train robots to have an inquisitive mindset.

*This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2022-0-00951, Development of Uncertainty-Aware Agents Learning by Asking Questions)

[1]Minjung Shin is with Interdisciplinary Program in Cognitive Science, Seoul National University, 08826 Seoul, Korea, mjshin77@snu.ac.kr

[2]Minsu Jang and Miyoung Cho are with Electronics & Telecommunications Research Institute, 34129 Daejeon, Korea, minsu, mycho@etri.re.kr

[3]Jeh-Kwang Ryu is with the Department of Physical Education, Dongguk University, 04620 Seoul, Korea, ryujk@dgu.ac.kr, *Corresponding author

II. Related Work

All learning agents, including both humans and machines that utilize natural language, require the ability to ask questions. Despite its importance and prominence, there is a lack of research on questioning, notably in multi-modal settings (e.g., using natural language on a given image).

Visual Question Answering (VQA), in particular, prompted a lot of interest in computer vision and natural language processing, but most VQA system is designed to identify the critical visual components in the image to predict the right answer given an image and the related query in natural language [12], [13].

Only several studies of question generation in the visual domain aimed to generate more natural and engaging questions as opposed to other models that ask too simplified information revealed in the images. They focused on objects by constraining Wh-questions [14] or events by involving human annotators [15]. In recent times, several advanced models have been introduced that can accept either a question or an answer as input and generate the corresponding counterpart [16], as well as pro-social conversational agents that utilize human-like inquiries [17].

However, existing datasets and models were mainly focused on the range of apparent and straightforward situations. To train interactive agent models to reduce uncertainty, we need to focus on scenarios in inquisitive situations that align with the curious nature of humans.

III. Psychology of Inquisitive Agent

Animals, including humans, feel pressure to mitigate uncertainty when confronting it. From an evolutionary perspective, clarifying the uncertainty has been crucial for the survival of animals, including humans [5].

Fig. 1 depicts the flow of uncertainty resolving process of humans. Cognitive incongruities (e.g., information gap, ambiguity, novelty, etc.) cause uncertainty [18], [19]. Then the person appraises whether they can resolve it through metacognition. If they believe they cannot, they may abandon the issue. On the other hand, if they feel they can deal with it, they might exploit various cognitive strategies to solve it. Epistemic curiosity, the motivation for inquisitive

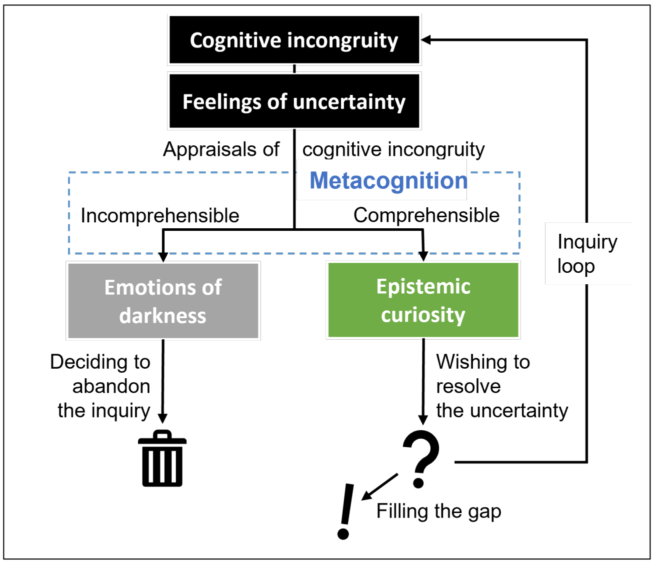

Fig. 1. Flow of uncertainty resolution process in humans. Schematic reconstruction from [18], [19]

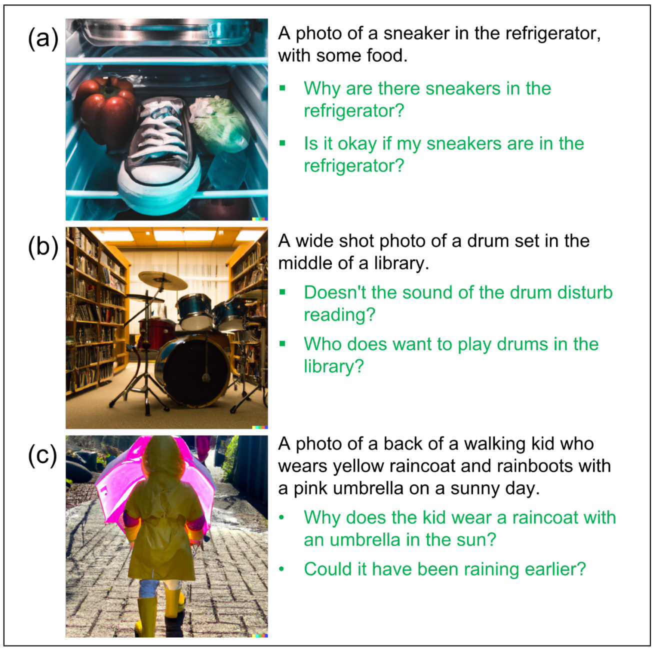

Fig. 2. Example scenarios for the training dataset. Images were generated with the assistance of AI, DALL·E 2 [26], [27]. The texts in black indicate the input prompt, and bulleted lists with green texts are questions made by humans about the ungraspable point from the images.

(i.e., information-seeking) behavior, plays the most crucial role in the cognitive process for uncertainty resolution [18], [19]. However, measuring curiosity is challenging due to its intrinsic and abstract nature, and the efforts to enlighten the processes have been made relatively recently [20], [21].

The most plausible explanation related to curiosity is *information gap theory* [22]; people become intrigued when they are aware of missing information with feeling unpleasant or pleasant [23]. Several studies tested the proposal that showed behavioral tendencies [24] and neural activities [25] consistent with the theory. The empirical approaches showed how curiosity activates the learning system in the brain by bridging with other cognitive phenomena, such as reward anticipation and learning facilitation. It also has significant implications for designing interactive robots to encourage successive learning and intellectual growth.

IV. Proposal for Dataset Design

Based on the flow of the uncertainty resolution process in humans, as presented in Fig. 1, we suggest a dataset that aims to stimulate a kind of *artificial* epistemic curiosity. The dataset, which we call CAWS (Curious About Weird Scene), consists of *images*, *textual descriptions*, and *inquisitive questions*. For eliciting epistemic curiosity, the images intentionally contain inconsistencies or uncertainties, along with questions that can be answered to satisfy curiosity and encourage deeper exploration rather than just relying on what is immediately apparent in the images.

We define something as 'weird' if plausible in real life but still confusing or unclear when we encounter it. To obtain these situations within the predetermined scope, we employ Generative AI that arouses curiosity, and two researchers inspect and select suitable descriptions. This dataset can be used to train robots to generate *inquisitive questions* when given a pair of an *image* and a *textual description*. Fig. 2 shows some concrete examples of our dataset design. It offers an image, a caption describing the scene in the image, and a list of questions.

Fig. 2(a) is an example of arising queries from an image inconsistent with common sense about the purpose of an item or the relationship between entities. Fig. 2(b) shows questions that violate the predictions of social consensus about public space and the environment. Fig. 2(c) addresses queries arising when someone is unsure about the other's intentions.

We generated the images in Fig. 2 using DALL·E 2 [26], [27] by providing scene descriptions as prompts. With the images, we could facilitate the inquisitive questions by human annotators.

V. Conclusion

Asking good questions is essential for interactive robot agents to be employed in the real world, significantly mitigating uncertainty through actively seeking out information. Therefore, learning by asking is a great tactic. Considering humans' inquisitive nature, our study highlights the significance of reflecting on their inquiry strategy under uncertainty.

Hence we propose a design for a novel dataset providing partially incongruent or uncertain situations and drawing out open-ended questions pursuing ungiven information; This study can offer a practical guideline to improve the interactive agent's ability to come up with human-like and insightful questions.

Additionally, applying the dataset presented in this study enables it to participate as a supportive collaborator in creative work, like an engaging friend who can read between the lines. To achieve our goals, it is required to conduct further research. This involves training the model to show how useful the dataset is and comparing it to the text-based question-generation performances of recent cutting-edge chatbots (e.g., ChatGPT [28]).

ACKNOWLEDGMENT

We are deeply grateful to the reviewers for providing kind and helpful comments.

REFERENCES

[1] R. Golman and G. Loewenstein, "The desire for knowledge and wisdom," in *The New Science of Curiosity*, G. Gordon, Ed. Hauppauge, NY: Nova Science Publishers, Inc., 2018, ch. 2, pp. 37–42.

[2] K. Huang, M. Yeomans, A. W. Brooks, J. Minson, and F. Gino, "It doesn't hurt to ask: Question-asking increases liking." *Journal of personality and social psychology*, vol. 113, no. 3, p. 430, 2017.

[3] M. M. Chouinard, P. L. Harris, and M. P. Maratsos, "Children's questions: A mechanism for cognitive development," *Monographs of the society for research in child development*, pp. i–129, 2007.

[4] A. Ruggeri, T. Lombrozo, T. L. Griffiths, and F. Xu, "Sources of developmental change in the efficiency of information search." *Developmental psychology*, vol. 52, no. 12, p. 2159, 2016.

[5] R. L. Cervera, M. Z. Wang, and B. Y. Hayden, "Systems neuroscience of curiosity," *Current Opinion in Behavioral Sciences*, vol. 35, pp. 48–55, 2020.

[6] C. Ray, F. Mondada, and R. Siegwart, "What do people expect from robots?" in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Nice, France: IEEE, 2008, pp. 3816–3821.

[7] M. Doering, P. Liu, D. F. Glas, T. Kanda, D. Kulić, and H. Ishiguro, "Curiosity did not kill the robot: A curiosity-based learning system for a shopkeeper robot," *J. Hum.-Robot Interact.*, vol. 8, no. 3, jul 2019.

[8] T. Fong, C. Thorpe, and C. Baur, "Robot, asker of questions," *Robotics and Autonomous Systems*, vol. 42, no. 3, pp. 235–243, 2003.

[9] M. Cakmak and A. L. Thomaz, "Designing robot learners that ask good questions," in *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Boston, MA, USA: IEEE, 2012, pp. 17–24.

[10] A. Rothe, B. M. Lake, and T. M. Gureckis, "Do people ask good questions?" *Computational Brain & Behavior*, vol. 1, no. 1, pp. 69–89, 2018.

[11] A. C. Graesser, "An introduction to the study of questioning," in *The Psychology of Questions*, A. C. Graesser and J. B. Black, Eds. New York, NY: Routledge, 1985, ch. 1, pp. 1–13.

[12] S. Antol, A. Agrawal, J. Lu, M. Mitchell, D. Batra, C. L. Zitnick, and D. Parikh, "Vqa: Visual question answering," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[13] H. Gao, J. Mao, J. Zhou, Z. Huang, L. Wang, and W. Xu, "Are you talking to a machine? dataset and methods for multilingual image question answering," in *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'15. Cambridge, MA, USA: MIT Press, 2015, p. 2296–2304.

[14] M. Ren, R. Kiros, and R. Zemel, "Question answering about images using visual semantic embeddings," in *Deep Learning Workshop, ICML*, 2015.

[15] N. Mostafazadeh, I. Misra, M. Mitchell, X. He, and L. Vanderwende, "Generating natural questions about an image," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Berlin, Germany: Association for Computational Linguistics, Aug. 2016, pp. 1802–1813.

[16] Y. Li, N. Duan, B. Zhou, X. Chu, W. Ouyang, X. Wang, and M. Zhou, "Visual question generation as dual task of visual question answering," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6116–6124.

[17] R. Krishna, D. Lee, L. Fei-Fei, and M. S. Bernstein, "Socially situated artificial intelligence enables learning from human interaction," *Proceedings of the National Academy of Sciences*, vol. 119, no. 39, p. e2115730119, 2022.

[18] J. Vazard and C. Audrin, "The noetic feeling of confusion," *Philosophical Psychology*, vol. 35, no. 5, pp. 757–770, 2022.

[19] J. Dokic, "Seeds of self-knowledge: noetic feelings and metacognition," in *Foundations of Metacognition*, M. J. Beran, J. Perner, and J. Proust, Eds. Oxford, UK.: Oxford University Press, 2012, pp. 302–321.

[20] C. Kidd and B. Y. Hayden, "The psychology and neuroscience of curiosity," *Neuron*, vol. 88, no. 3, pp. 449–460, 2015.

[21] J. Gottlieb and P.-Y. Oudeyer, "Towards a neuroscience of active sampling and curiosity," *Nature Reviews Neuroscience*, vol. 19, no. 12, pp. 758–770, 2018.

[22] G. Loewenstein, "The psychology of curiosity: A review and reinterpretation." *Psychological bulletin*, vol. 116, no. 1, pp. 75–98, 1994.

[23] J. A. Litman, "Interest and deprivation factors of epistemic curiosity," *Personality and individual differences*, vol. 44, no. 7, pp. 1585–1595, 2008.

[24] M. Jepma, R. G. Verdonschot, H. Van Steenbergen, S. A. Rombouts, and S. Nieuwenhuis, "Neural mechanisms underlying the induction and relief of perceptual curiosity," *Frontiers in behavioral neuroscience*, vol. 6, pp. 1–9, 2012.

[25] M. J. Kang, M. Hsu, I. M. Krajbich, G. Loewenstein, S. M. McClure, J. T.-y. Wang, and C. F. Camerer, "The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory," *Psychological science*, vol. 20, no. 8, pp. 963–973, 2009.

[26] openAI, "Dall·e 2," https://openai.com/dall-e-2/, 2022, retrieved 24 April, 2023.

[27] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever, "Zero-shot text-to-image generation," in *International Conference on Machine Learning*. PMLR, 2021, pp. 8821–8831.

[28] openAI, "Chatgpt," https://chat.openai.com/, 2022, retrieved 24 April, 2023.