Large Language Models as an active Bayesian filter: information acquisition and integration

Sabrina Patania, Emanuele Masiero, Luca Brini, Valentyn Piskovskyi and **Dimitri Ognibene**
University of Milan - Bicocca
dimitri.ognibene@unimib.it

Gregor Donabauer and **Udo Kruschwitz**
University of Regensburg

Abstract

This study investigates Large Language Models (LLMs) as dynamic Bayesian filters through question-asking experiments inspired by cognitive science. We analyse LLMs' inference errors and the evolution of uncertainty across models using repeated sampling.

Building on Bertolazzi et al. (2023), we trace LLM belief states during repeated queries, finding that entropy decreases with each interaction, signaling reduced uncertainty. However, issues like "resurrection" (reassigning probabilities to invalidated outcomes) and "Bayesian apocalypse" (probabilities approaching zero) reveal significant flaws. GPT-4o consistently outperforms GPT-3 in probabilistic reasoning. These results underscore the need for improved architectures for reliability in high-stakes contexts and suggest a link between token-level and task-level uncertainty dynamics that can be leveraged to enhance LLM performance.

1 Introduction

Large Language Models (LLMs) act as reactive agents, primarily engaging in one-step predictions without explicit planning or deliberation mechanisms. This reactivity, often viewed as a limitation (van Lier, 2023; Floridi, 2023; Wu et al., 2024; Mahowald et al., 2024), does not inherently negate the presence of underlying objectives that the agent may pursue (Brooks, 1991). The behaviour of reactive agents is driven directly by their immediate input, thus their ability to find and select information, and deal with uncertainty has been seen as limited (Kaelbling et al., 1998). Yet, they have been shown to be able to determine their inputs in certain cases (Nolfi, 2002; Bonet, 2010). Indeed, reactive systems can perform effective information-seeking behaviours, crucial under uncertainty such as when communicating with hard-to-interpret agents having different knowledge of the interaction context (Paek and Horvitz, 2000; Ognibene and Demiris, 2013), and traditionally associated with more complex deliberative systems and explicit uncertainty reasoning (Beer and Di Paolo, 2023). Studies have shown that even simple reactive systems can engage in exploration and reduce uncertainty through epistemic actions, even without internal simulations or memory (Ognibene et al., 2013). These findings suggest that LLMs, despite lacking explicit internal reasoning capabilities, may still engage in goal-oriented behaviors and possess information-gathering capabilities.

Uncertainty management is crucially connected with information seeking in normative decision systems and also plays an important role in the computations ascribed to the brain (Friston et al., 2015; Kaelbling et al., 1998). However, how task-level uncertainty is processed in multi-layered deep generative models, particularly auto-regressive LLMs (Brown et al., 2020; Radford et al., 2019), and how they compare to normative systems remains largely unexplored. This is particularly interesting given the breadth of information they contain across disparate domains, in contrast to the limited domain variables usually dealt with by normative models.

LLMs learn the conditioned probability distribution of the next token given the sequence of previous tokens (input context) and produce output tokens sampling sequentially from the learnt distribution model (Radford et al., 2018). While uncertainty regarding the next output may be known and represented by the last layer of the model, the long-term evolution of generation or semantic uncertainty is not immediately available (Farquhar et al., 2024; Kuhn et al., 2023; Chen and Mueller, 2023). In fact, when LLMs are used in chatbots, text generation or other non single token output tasks, the stochastic production of a single output is appended to current context input and participates in the generation of successive outputs. Concatenating stochastic processes may result in extended non factual responses or "hallucinations", starting

Proceedings of the 28th Workshop on the Semantics and Pragmatics of Dialogue, September, 11–12, 2024, Trento, Italy.

from one first ambiguous output (Ji et al., 2023; LeCun, 2023).

An important contributing factor is that information about the mental state, knowledge, beliefs, and desires of the writer at the moment of writing the text is not directly available for LLMs during training. This may reduce learning performance (Bianco and Ognibene, 2022) and lead to semantically different next token to have the same probability and contribute to generating hallucinations. Moreover, many LLMs in chatbots appear trained to present overconfident responses even when uninformed and produced in an uncertain state (Chen and Mueller, 2023).

To correct belief tracking (Mrkšić et al., 2017) and uncertainty estimation, i.e. estimating how much an agent knows and does not know about the environment, the domain and the current situation, can be useful to adopt effective uncertainty reduction strategies (Kaelbling et al., 1998; Friston et al., 2015; Taniguchi et al., 2023) such as the generation of questions and clarifications (Varges et al., 2010; Kominis and Geffner, 2017; Tellex et al., 2012). However, models that explicitly reason about uncertainty and lack of knowledge have to face challenging computational complexity due to the expansion of the state space (Kaelbling et al., 1998). Various approximations have been developed also based on neural architectures and reinforcement learning (Ognibene and Baldassare, 2014; Wu et al., 2021; Xu et al., 2022; Wang et al., 2020), and, while these approaches may be particularly data hungry (Schatzmann et al., 2007; Wang et al., 2020), it is worth investigating if and how different LLMs learn similar information-gathering strategies as well as implicitly encode beliefs and uncertainty.

However, unveiling how LLMs may couple information integration and acquisition is challenging due to the limited accessibility and interpretability of LLMs and the stochastic recursive process they use to produce output. Similarly to other approaches (Kuhn et al., 2023; Chen and Mueller, 2023), we use a repeated sampling approach through prompts, or repeated zero shots tests (Brown et al., 2020), to retrieve probabilistic information on the information integration process inside the LLM, i.e. how information in previous dialogue exchanges is integrated into a belief and how this belief determines the output selection strategy. We estimate the evolution of this belief state during the interaction and information acquisition using information theoretic methods, i.e. we measure the entropy of the responses distribution (Ognibene and Demiris, 2013; Friston et al., 2015; Ognibene et al., 2019), an approach already tested analysing information integration in black box models (Lungarella and Sporns, 2006). In other words, we aim to observe if the low-level stochastic process of token production of the LLM can be connected to the dynamic process of task-level information integration. This may later lead to novel and more effective task-level uncertainty management for LLMs.

Building on the experiments from cognitive science on information acquisition strategies (Ruggeri and Lombrozo, 2015) and the analysis of Bertolazzi et al. (2023), our study will computationally trace the belief states of LLMs through repeated queries. By examining the probabilistic responses of these models, we aim to gain deeper insights into their inference processes and uncertainty management. Our approach is inspired by Bayesian filtering, which involves continuously updating the probability distributions of candidate items based on new information from each interaction. This approach allows the models to refine their predictions dynamically, integrating new data to reduce uncertainty over time. When the model is correct, Bayesian models estimations are optimal (Särkkä and Svensson, 2023).

Bayesian filtering, commonly used in dynamic systems for state estimation, applies here as we treat the sequence of interactions as a time series. The model updates its belief state with each dialogue step, combining prior knowledge with new evidence. This method enhances the model's ability to manage and process evolving information, mirroring the cognitive processes involved in human decision-making.

Additionally, we will explore the performance difference between GPT-3 and GPT-4o, investigating how these models handle probabilistic data differently. This study will help frame LLMs as complex systems with significant probabilistic reasoning capabilities, albeit with notable flaws. Addressing these limitations is crucial for enhancing the reliability and accuracy of LLMs, particularly in high-stakes environments such as clinical diagnostics and financial forecasting. In chatbots uncertainty about users' requests, context or domain knowledge affects dialog and could elicit queries aimed at resolving it, but such capabilities are lim-

ited at the moment.

2 Related Work

In recent years, various methods have been proposed to define and quantify uncertainties in the context of Large Language Models (LLMs).

For instance, Chen and Mueller (2023) propose a technique to estimate a numeric confidence score for any LLM output generated by a black-box API. This method involves multiple API calls with varying prompts and sampling temperatures, providing users with a confidence estimate that highlights unreliable outputs. Similarly, Yang et al. (2023) introduce a framework to produce uncertainty-aware LLMs capable of estimating aleatoric, epistemic, or composed uncertainty for each prediction in a model- and data-agnostic manner. Their models learn data-dependent thresholds, enhancing prediction reliability.

Huang et al. (2024) present a unified calibration framework that treats both response correctness and associated confidence levels as distributions. Their approach improves calibration through fine-tuning, integrating relevant documents, and adjusting sampling temperatures. Additionally, Zhang et al. (2024) develop LUQ, a sampling-based uncertainty quantification approach for long texts. LUQ identifies LLMs' lack of confidence in generating factual long texts and proposes the LUQ-ENSEMBLE method, which enhances factuality by ensembling responses from multiple models to select the least uncertain response.

Nevertheless, these methods face limitations in interactive settings due to challenges in measuring the evolution of uncertainty and integrating information throughout interactions (Bertolazzi et al., 2023).

Further contributions addressing these limitations include Ren et al. (2023), who propose a framework for measuring and aligning the uncertainty of LLM-based planners. Their approach ensures that planners recognise their uncertainty and request assistance when necessary, utilising conformal prediction theory to provide statistical guarantees on task completion while minimising human intervention in complex multi-step planning scenarios. This method effectively measures the evolution of uncertainty and integrates information dynamically during interactions. Similarly, Hou et al. (2023) present a method that, instead of ensembling models with different parameters, generates a set of clarifications for the input, processes these through fixed LLMs, and ensembles the corresponding predictions. This approach addresses the integration of information across interactions by generating and processing multiple clarifications.

It is pivotal to highlight that uncertainty handling requires considering both token-level and task-level processes. Information theory approaches, like those discussed by Lungarella and Sporns (2006), can provide valuable insights into how information flow and entropic measures at different levels can be used to analyse and manage these uncertainties effectively.

Our work builds upon these foundational studies by investigating the probabilistic reasoning capabilities of LLMs in dynamic and interactive environments. We aim to fill the gap in understanding how uncertainty evolves throughout interactions and how belief states of LLMs are affected by repeated queries. By comparing the performance of different LLMs, specifically GPT-3 and GPT-4o, we seek to provide deeper insights into their strengths and limitations in managing uncertainty. Our approach leverages repeated sampling and behavioural analysis to develop a comprehensive understanding of LLMs' inference processes, contributing to the broader goal of enhancing the robustness and reliability of LLMs in real-world applications.

3 Experiments

Firstly, we quantified the uncertainty of the model at each step of a dialogue between the Questioner and Oracle. Drawing from the "20 Questions game" (Bertolazzi et al., 2023), we measured the uncertainty associated with both the questions posed and the responses received during the interactions, generated with GPT-3 and GPT-4o.

In order to illustrate a typical interaction in the mentioned game, we provide an example of a prompt supplied to a LLM.

> You will be given of a dialogue of the 20 questions game. You have to list out absolutely all the items from the given candidates set that satisfy each <question, answer> in the given dialogue. The output should strictly use the following template:
> EXPLANATION:

CANDIDATES: item1, item2, item3

Dialogue: target = dalmatian

- *Answerer*: This is the list of candidates: dog, bear, flamingo, hawk, toucan, dalmatian, hippopotamus, chick.

- *Questioner*: Is the item you have assigned an animal?
- *Answerer*: Yes.

- *Questioner*: Is the animal you have assigned a mammal?
- *Answerer*: Yes.

- *Questioner*: Is the mammal you have assigned a carnivore?
- *Answerer*: Yes.

- *Questioner*: Is the carnivorous mammal you have assigned typically found in water?
- *Answerer*: No.

- *Questioner*: Does the carnivorous mammal you have assigned have spots on its body?
- *Answerer*: Yes! That's correct.

Building upon further analyses from the 20 Questions game paper, we implemented an additional method to assess each candidate item's consistency with every question-answer pair throughout the dialogue d_t (with $t \in [0, T]$). Firstly, this methodology allows us to determine which items are systematically excluded at each step t of the dialogue.

To execute this, we employed an additional LLM agent tasked with verifying, given a dialogue d_t up to a certain point t and a candidate item, whether the item was deemed to satisfy all the question-answer pairs of d_t. This query was sampled k times, recording the number of positive occurrences. Consequently, at a given moment t, each i-th item was assigned a probability score computed as follows:

$$p_i(t) = \frac{1}{k} \sum_{j=1}^{k} \delta_{ij}(t) \tag{1}$$

where $\delta_{ij}(t)$ is an indicator function that is 1 if the i-th item satisfies all the question-answer pairs of d_t in the j-th query, and 0 otherwise.

After computing these scores and normalising them, we further calculated the probability distribution across all candidate items. This comprehensive approach provides a dynamic view of the model's uncertainty management and enhances our understanding of the probabilistic reasoning capabilities of LLMs within interactive scenarios, effectively demonstrating the principles of Bayesian filtering by continuously updating beliefs based on incoming data.

Once the probability distributions for each dialogue were calculated, we proceeded to analyse the entropy to assess task-level uncertainty. This analysis involved examining the entropy levels of the distributions at various stages of the dialogue to assess the degree of uncertainty and information gain as the dialogue progressed. The entropy is calculated for the distribution over the items for each dialogue, and then the mean entropy is computed by averaging over all dialogues, resulting in a mean entropy for each step of the dialogue. The mean entropy at each stage t of the dialogue is given by:

$$\overline{H}(t) = \frac{1}{D} \sum_{d=1}^{D} H(p^{(d)}(t)) \tag{2}$$

where $p^{(d)}(t)$ represents the probability distribution of the items at stage t in dialogue d.

By measuring the changes in entropy, we could evaluate how effectively the LLM was processing and refining information through its interactions, and identify any patterns or anomalies in its approach to reducing uncertainty.

We also tested an alternative approach by asking both GPT-3 and GPT-4o to evaluate the validity of items given a dialogue d at step t. Instead of querying each candidate item individually, we presented the entire list of items to the models simultaneously and requested them to identify the valid items (simultaneous approach). This method allows LLM models to consider all options at once, potentially using their comparative reasoning capabilities.

In this approach, the models provided items deemed consistent with the dialogue context up to step t. This method offers a different perspective on assessing candidate items, focusing on the models' ability to process and filter multiple options in parallel.

4 Results

Our analysis identifies a "resurrection" phenomenon, where LLMs reassign non-zero probabilities to outcomes previously deemed invalid. This occurs in about 80% of GPT-3 dialogues. Figures 1 and 2 quantify this for GPT-3. For GPT-4o, Figures 3 and 4 show similar results. Interestingly, the phenomenon appears to be significantly influenced by the approach used, with the simultaneous

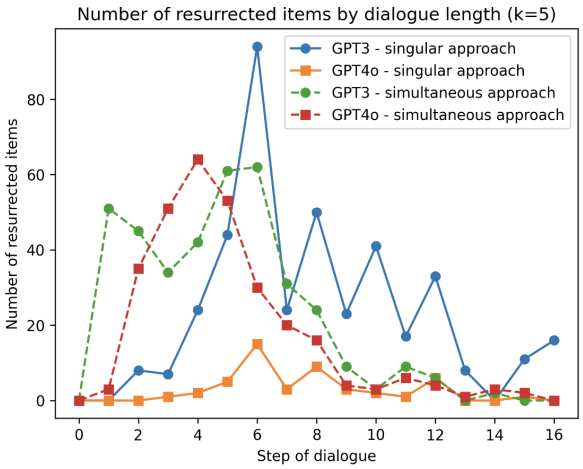

Figure 1: Number of resurrected items for each model and approach as a function of dialogue length on GPT-3 dialogues.

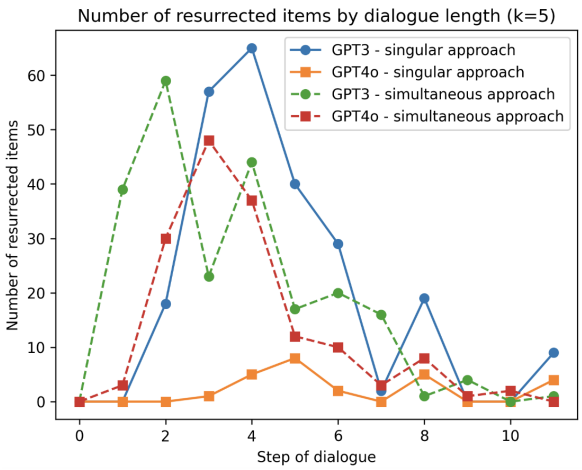

Figure 3: Number of resurrected items for each model and approach as a function of dialogue length on GPT-4o dialogues.

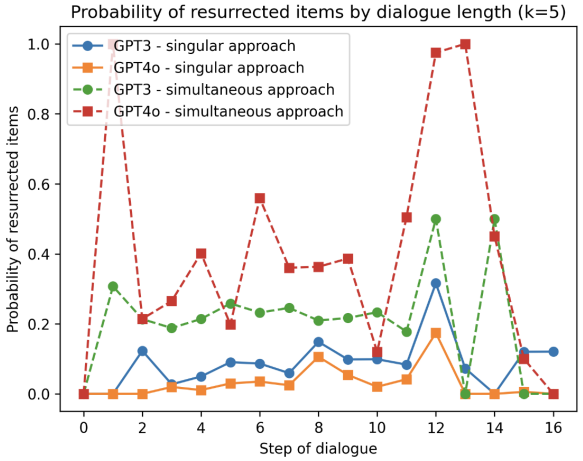

Figure 2: Mean probability absorbed by resurrected items at each step of the dialogue on GPT-3 dialogues.

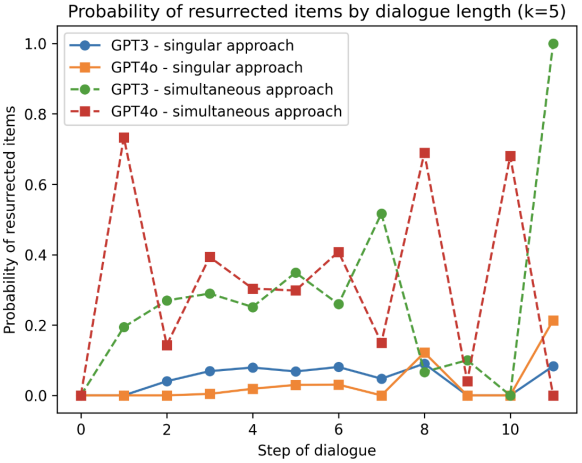

Figure 4: Mean probability absorbed by resurrected items at each step of the dialogue on GPT-4o dialogues.

approach being more affected by the resurrection phenomenon for GPT-4o dialogues.

The duration of the dialogues varies significantly across the dataset. This variation is illustrated in the graph presented in Figure 5, where we compare GPT-3 and GPT-4o for the task, which in this context corresponds to the duration of the dialogue, as a shorter duration implies arriving at a solution with fewer questions.

Figure 6 compares entropy trends for GPT-3 and GPT-4o. The graph includes the ideal entropy curve, which represents the evolution of the probability distribution entropy if the search for the item is carried out optimally, using a binary search approach. GPT-4o outperforms GPT-3, showing less sensitivity to varying k. GPT-3's performance improves with higher k, though it remains more uncertain. Please note that GPT-4o is not tested with $k > 10$ as the results for $k = 5$ and $k = 10$ are very similar, indicating that it is not necessary to increase the sample size.

Figures 7 to 9 detail entropy and cross-entropy results for different models and dialogue sources. While cross-entropy is seemingly a more precise measure of model performance, as it takes into account the correct response, it is important to consider that entropy is a more appropriate measure in our context. This is because we are primarily interested in the overall reduction of uncertainty, and thus in the strategies the model employs to achieve this effectively, rather than its ability to approximate the correct answer. Although, as the figures suggest, these two aspects tend to go hand in hand.

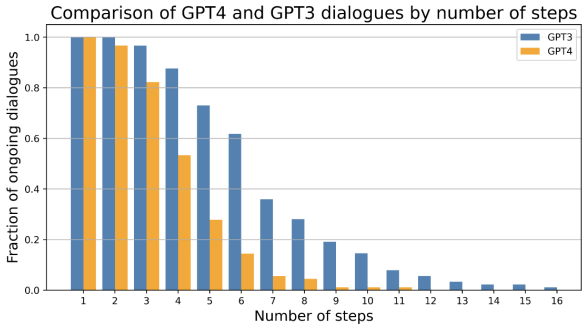

Figure 5: Comparison of dialogue durations for GPT-3 and GPT-4o.

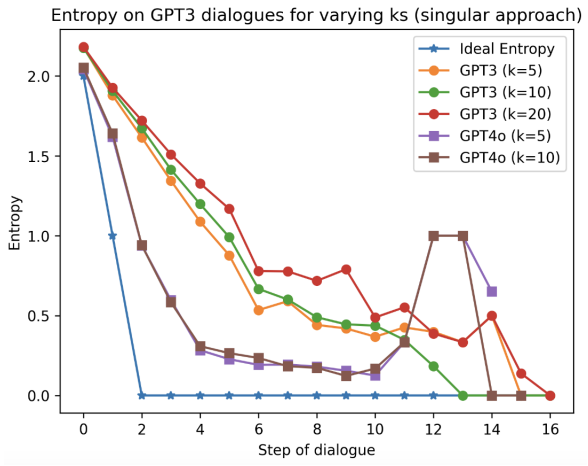

Figure 6: Comparison of entropy trends for GPT-3 and GPT-4o from the beginning of the dialogue, across various k values, with the ideal entropy curve.

Additional insights were gained by analysing dialogue steps in which the model either erroneously confirmed a target (entropy = 0) or generated distributions with only zero values. These instances, observed only in dialogues processed by GPT-4o, were marked by significant ambiguities or errors, often resulting in the incorrect elimination of the true target. This error analysis also extended to GPT-3 and GPT-4o's ability to consistently list valid candidates at each step, revealing differences in their performance throughout the dialogues.

Figures 11 and 12 display the probability of zero distributions by dialogue step for GPT-3 and GPT-4o dialogues, respectively. The results suggest that the phenomenon of Bayesian apocalypse, where all item probabilities approach zero, is more prevalent at specific dialogue steps and is sensitive to the approach used.

Finally, Figures 13 and 14 compare entropy results between simultaneous and singular approaches. The simultaneous approach improves

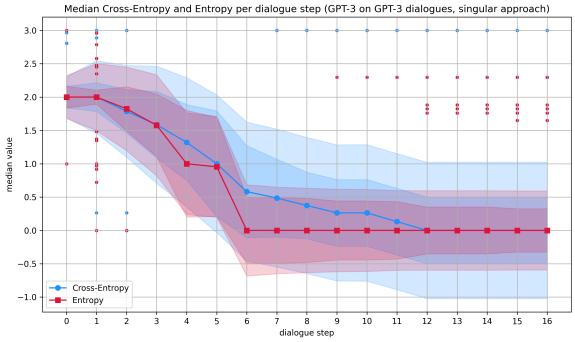

Figure 7: Entropy and Cross-Entropy levels for GPT-3 on dialogues generated by GPT-3.

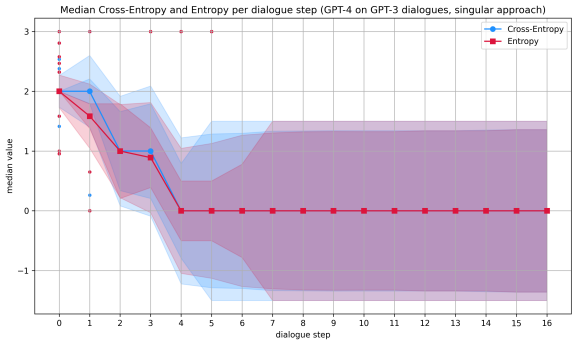

Figure 8: Entropy and Cross-Entropy levels for GPT-4 on dialogues generated by GPT-3.

GPT-3 performance but degrades GPT-4o results at the most significant steps, which are those with the highest number of samples (see Figure 5). GPT-4o consistently outperforms GPT-3, with the simultaneous approach proving to be more beneficial for GPT-3 while negatively impacting GPT-4o performance.

5 Discussion

Our study extends the analysis of Bertolazzi et al. (2023) by computationally tracking LLM belief states through repeated queries. This reveals that entropy decreases with each interaction, and decisions are made when uncertainty is minimised, consistent with normative models (Friston et al., 2015; Ognibene and Demiris, 2013). This suggests a link between the low-level stochastic processes of token production and the integration of higher-level task information. Future work could explore the extraction of uncertainty and information gain predictions from LLM internal states. However, the significant noise in the process suggests that current LLMs, particularly GPT-3, may benefit from targeted training to improve performance.

GPT-4o's better performance compared to GPT-

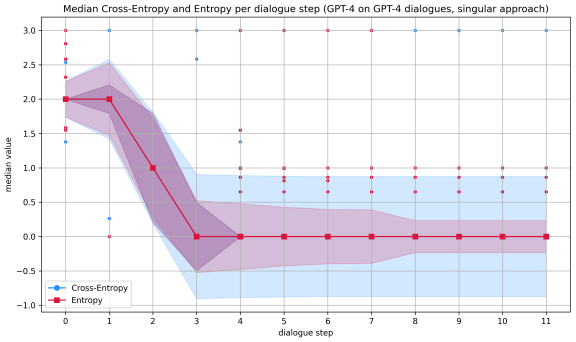

Figure 9: Entropy and Cross-Entropy levels for GPT-4 on dialogues generated by GPT-4.

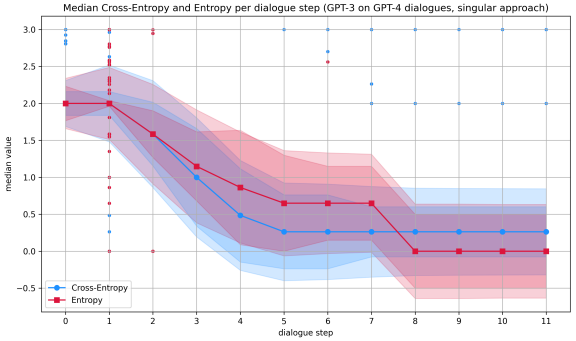

Figure 10: Entropy and Cross-Entropy levels for GPT-3 on dialogues generated by GPT-4.

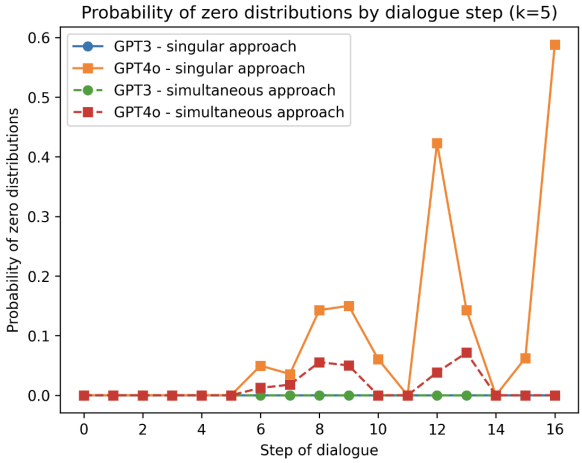

Figure 11: Probability of zero distributions (Bayesian apocalypse) by dialogue length for GPT-3 and GPT-4o on GPT-3 dialogues.

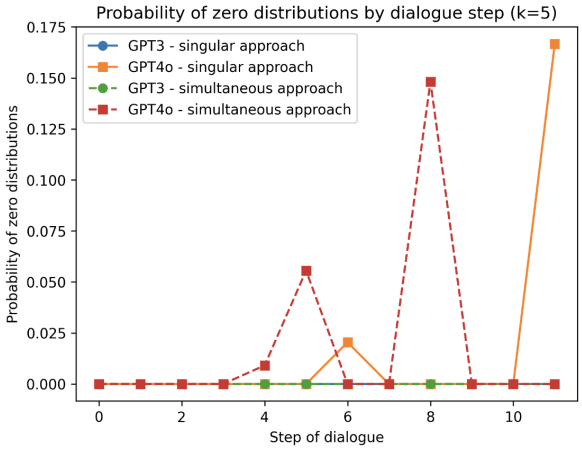

Figure 12: Probability of zero distributions (Bayesian apocalypse) by dialogue length for GPT-3 and GPT-4o on GPT-4o dialogues.

3 on the same dataset likely reflects GPT-3's limitations in retaining and analysing critical information. Although GPT-3 appears to perform better with the alternative approach, this could be misleading. GPT-4o typically reaches solutions faster with shorter dialogues, as shown in Figure 5, where only 25% of GPT-4o dialogues exceed the fifth step, while GPT-3 dialogues often extend to the sixth step. The perceived advantage of GPT-3 in later steps may thus stem from different dialogue lengths rather than actual performance improvements.

The "resurrection" phenomenon, in which LLMs reassign nonzero probabilities to previously invalidated outcomes, is quantified by tracking the frequency and magnitude of these probabilities. As shown in Figure 2, GPT-3's probability for resurrected items slightly decreases over time, indicating how the model handles uncertainty. Figure 4 shows that this phenomenon is less pronounced in GPT-4o. However, the simultaneous approach exacerbates this issue.

The "Bayesian apocalypse", where all probabilities approach zero, results in high uncertainty and challenges in distinguishing valid from invalid hypotheses (Bengtsson et al., 2008). This phenomenon mirrors issues in particle filtering, where insufficient particles fail to represent the true state distribution, leading to similar collapses in probability. This exacerbates model uncertainty and impairs decision-making.

Probability collapses may also indicate hallucinations within dialogues, especially when options are finite. This issue is critical in interactive settings where consistent tracking of candidates is necessary. Hallucinations can lead to incorrect responses and premature elimination of valid options, highlighting the need for better uncertainty management and handling of incomplete or noisy data.

Comparing GPT-3 and GPT-4o, we used both singular and simultaneous sampling approaches. GPT-4o's superior performance likely stems from better information retention and analysis compared

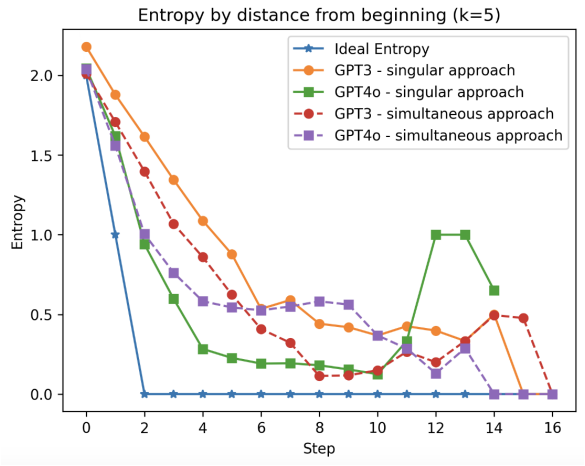

Figure 13: Comparison of entropy trends across models and approaches with GPT-3 dialogues.

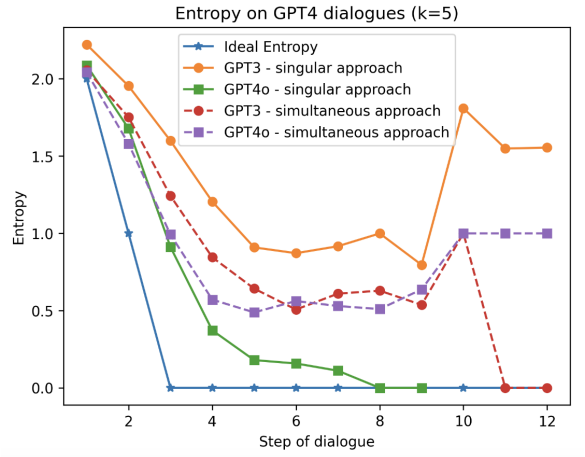

Figure 14: Comparison of entropy trends across models and approaches with GPT-4o dialogues.

to GPT-3. However, GPT-3's apparent advantage with the singular approach after six steps (Figure 13) may be misleading. GPT-4o's shorter dialogues often mean it reaches solutions more quickly, as indicated by Figure 5. Therefore, the perceived benefit of GPT-3 in subsequent steps may reflect differences in dialogue length rather than true performance.

Our findings show that LLMs can effectively explore and identify relevant information without extensive task-specific training, aligning with reactive systems research (Beer and Di Paolo, 2023; Ognibene et al., 2013). Although the entropy reduction approach is effective, current strategies for information integration are still suboptimal. GPT-4o demonstrates more robust performance, but issues in managing probabilistic data and avoiding probability collapses persist. Addressing these challenges is crucial for improving the reliability and accuracy of LLMs, especially for high-stakes applications.

Our results align with Bayesian inference principles, where uncertainty is minimised by updating probability distributions with new evidence. Similarly, LLMs update token predictions based on preceding context, aiming to reduce output uncertainty. Despite lacking explicit task-level uncertainty representation, LLMs dynamically integrate new information, reflecting a Bayesian-like process in their operation.

References

Randall D Beer and Ezequiel A Di Paolo. 2023. The theoretical foundations of enaction: Precariousness. *Biosystems*, 223:104823.

Thomas Bengtsson, Peter Bickel, and Bo Li. 2008. Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems. In *Probability and statistics: Essays in honor of David A. Freedman*, volume 2, pages 316–335. Institute of Mathematical Statistics.

Leonardo Bertolazzi, Davide Mazzaccara, Filippo Merlo, and Raffaella Bernardi. 2023. Chatgpt's information seeking strategy: Insights from the 20-questions game. In *Proceedings of the 16th International Natural Language Generation Conference*, pages 153–162.

Francesca Bianco and Dimitri Ognibene. 2022. Robot learning theory of mind through self-observation: Exploiting the intentions-beliefs synergy. *arXiv preprint arXiv:2210.09435*.

Blai Bonet. 2010. Conformant plans and beyond: Principles and complexity. *Artificial Intelligence*, 174(3-4):245–269.

Rodney A. Brooks. 1991. Intelligence without reason. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence - Volume 1*, IJCAI'91, page 569–595, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Jiuhai Chen and Jonas Mueller. 2023. Quantifying uncertainty in answers from any language model via intrinsic and extrinsic confidence assessment. *arXiv preprint arXiv:2308.16175*.

Sebastian Farquhar, Jannik Kossen, Lorenz Kuhn, and Yarin Gal. 2024. Detecting hallucinations in large

language models using semantic entropy. *Nature*, 630(8017):625–630.

Luciano Floridi. 2023. Ai as agency without intelligence: on chatgpt, large language models, and other generative models. *Philosophy & technology*, 36(1):15.

Karl Friston, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas Fitzgerald, and Giovanni Pezzulo. 2015. Active inference and epistemic value. *Cognitive neuroscience*, 6(4):187–214.

Bairu Hou, Yujian Liu, Kaizhi Qian, Jacob Andreas, Shiyu Chang, and Yang Zhang. 2023. Decomposing uncertainty for large language models through input clarification ensembling. *ICML 2024*.

Yukun Huang, Yixin Liu, Raghuveer Thirukovalluru, Arman Cohan, and Bhuwan Dhingra. 2024. Calibrating long-form generations from large language models.

Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12):1–38.

Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. 1998. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1-2):99–134.

Filippos Kominis and Hector Geffner. 2017. Multiagent online planning with nested beliefs and dialogue. In *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 27, pages 186–194.

Lorenz Kuhn, Yarin Gal, and Sebastian Farquhar. 2023. Semantic uncertainty: Linguistic invariances for uncertainty estimation in natural language generation. *arXiv preprint arXiv:2302.09664*.

Y LeCun. 2023. Do large language models need sensory grounding for meaning and understanding. In *Workshop on Philosophy of Deep Learning, NYU Center for Mind, Brain, and Consciousness and the Columbia Center for Science and Society*.

Max Lungarella and Olaf Sporns. 2006. Mapping information flow in sensorimotor networks. *PLoS computational biology*, 2(10):e144.

Kyle Mahowald, Anna A Ivanova, Idan A Blank, Nancy Kanwisher, Joshua B Tenenbaum, and Evelina Fedorenko. 2024. Dissociating language and thought in large language models. *Trends in Cognitive Sciences*.

Nikola Mrkšić, Diarmuid Ó Séaghdha, Tsung-Hsien Wen, Blaise Thomson, and Steve Young. 2017. Neural belief tracker: Data-driven dialogue state tracking. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1777–1788.

Stefano Nolfi. 2002. Power and the limits of reactive agents. *Neurocomputing*, 42(1-4):119–145.

Dimitri Ognibene and Gianluca Baldassare. 2014. Ecological active vision: four bioinspired principles to integrate bottom–up and adaptive top–down attention tested with a simple camera-arm robot. *IEEE transactions on autonomous mental development*, 7(1):3–25.

Dimitri Ognibene and Yiannis Demiris. 2013. Towards active event recognition. In *IJCAI*, pages 2495–2501.

Dimitri Ognibene, Lorenzo Mirante, and Letizia Marchegiani. 2019. Proactive intention recognition for joint human-robot search and rescue missions through monte-carlo planning in pomdp environments. In *Social Robotics: 11th International Conference, ICSR 2019, Madrid, Spain, November 26–29, 2019, Proceedings 11*, pages 332–343. Springer.

Dimitri Ognibene, Nicola Catenacci Volpi, Giovanni Pezzulo, and Gianluca Baldassare. 2013. Learning epistemic actions in model-free memory-free reinforcement learning: Experiments with a neurorobotic model. In *Biomimetic and Biohybrid Systems: Second International Conference, Living Machines 2013, London, UK, July 29–August 2, 2013. Proceedings 2*, pages 191–203. Springer.

Tim Paek and Eric Horvitz. 2000. Conversation as action under uncertainty. In *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, pages 455–464.

Alec Radford, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, et al. 2018. Improving language understanding by generative pre-training.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Allen Z. Ren, Anushri Dixit, Alexandra Bodrova, Sumeet Singh, Stephen Tu, Noah Brown, Peng Xu, Leila Takayama, Fei Xia, Jake Varley, Zhenjia Xu, Dorsa Sadigh, Andy Zeng, and Anirudha Majumdar. 2023. Robots that ask for help: Uncertainty alignment for large language model planners. In *7th Annual Conference on Robot Learning*.

Azzurra Ruggeri and Tania Lombrozo. 2015. Children adapt their questions to achieve efficient search. *Cognition*, 143:203–216.

Simo Särkkä and Lennart Svensson. 2023. *Bayesian filtering and smoothing*, volume 17. Cambridge university press.

Jost Schatzmann, Blaise Thomson, Karl Weilhammer, Hui Ye, and Steve Young. 2007. Agenda-based user simulation for bootstrapping a pomdp dialogue system. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*, NAACL-Short '07, page

149–152, USA. Association for Computational Linguistics.

Tadahiro Taniguchi, Shingo Murata, Masahiro Suzuki, Dimitri Ognibene, Pablo Lanillos, Emre Ugur, Lorenzo Jamone, Tomoaki Nakamura, Alejandra Ciria, Bruno Lara, et al. 2023. World models and predictive coding for cognitive and developmental robotics: frontiers and challenges. *Advanced Robotics*, 37(13):780–806.

Stefanie Tellex, Pratiksha Thaker, Robin Deits, Thomas Kollar, and Nicholas Roy. 2012. Toward information theoretic human-robot dialog. In *Proceedings of Robotics: Science and Systems*, Sydney, Australia.

Maud van Lier. 2023. Understanding large language models through the lens of artificial agency. *Swedish Artificial Intelligence Society*, pages 79–84.

Sebastian Varges, Silvia Quarteroni, Giuseppe Riccardi, and Alexei V Ivanov. 2010. Investigating clarification strategies in a hybrid pomdp dialog manager. In *Proceedings of the SIGDIAL 2010 Conference*, pages 213–216.

Sihan Wang, Kaijie Zhou, Kunfeng Lai, and Jianping Shen. 2020. Task-completion dialogue policy learning via Monte Carlo tree search with dueling network. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3461–3471, Online. Association for Computational Linguistics.

Yaxiong Wu, Craig Macdonald, and Iadh Ounis. 2021. Partially observable reinforcement learning for dialog-based interactive recommendation. In *Proceedings of the 15th ACM conference on recommender systems*, pages 241–251.

Zhaofeng Wu, Linlu Qiu, Alexis Ross, Ekin Akyürek, Boyuan Chen, Bailin Wang, Najoung Kim, Jacob Andreas, and Yoon Kim. 2024. Reasoning or reciting? exploring the capabilities and limitations of language models through counterfactual tasks. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 1819–1862.

Yunqiu Xu, Meng Fang, Ling Chen, Yali Du, Joey Zhou, and Chengqi Zhang. 2022. Perceiving the world: Question-guided reinforcement learning for text-based games. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 538–560, Dublin, Ireland. Association for Computational Linguistics.

Qi Yang, Shreya Ravikumar, Fynn Schmitt-Ulms, Satvik Lolla, Ege Demir, Iaroslav Elistratov, Alex Lavaee, Sadhana Lolla, Elaheh Ahmadi, Daniela Rus, Alexander Amini, and Alejandro Perez. 2023. Uncertainty-aware language modeling for selective question answering.

Caiqi Zhang, Fangyu Liu, Marco Basaldella, and Nigel Collier. 2024. Luq: Long-text uncertainty quantification for llms.