Value of Information: A Framework for Human–Agent Communication

Anonymous ACL submission

Abstract

Large Language Model (LLM) agents deployed for real-world tasks face a fundamental dilemma: user requests are underspecified, yet agents must decide whether to act on incomplete information or interrupt users for clarification. Existing approaches either rely on brittle confidence thresholds that require task-specific tuning, or fail to account for the varying stakes of different decisions. We introduce a decision-theoretic framework that resolves this trade-off through the Value of Information (VoI), enabling agents to dynamically weigh the expected utility gain from asking questions against the cognitive cost imposed on users. Our inference-time method requires no hyperparameter tuning and adapts seamlessly across contexts—from casual games to medical diagnosis. Experiments across four diverse domains (20 Questions, medical diagnosis, flight booking, and e-commerce) show that VoI consistently matches or exceeds the best manually-tuned baselines, achieving up to 1.36 utility points higher in high-cost settings. This work provides a parameter-free framework for adaptive agent communication that explicitly balances task risk, query ambiguity, and user effort.

1 Introduction

LLM agents are increasingly deployed as autonomous collaborators in complex, real-world tasks. However, a fundamental bottleneck remains: user requests are inherently underspecified, carrying latent goals, contexts, and unstated preferences (Malaviya et al., 2024; Yao et al., 2024; Peng et al., 2024). A request to "book a flight to London" omits critical details, such as budget constraints, preferred departure times, tolerance for layovers. No amount of model capability can resolve this ambiguity without external input; the agent must ask. Yet excessive questioning frustrates users and undermines the agent's value proposition. Effective collaboration thus requires agents to balance two risks: acting on incomplete information and misaligning with user intent, or interrupting frequently and imposing cognitive burden.

Current approaches fall short in navigating this trade-off. Fixed-round strategies ask a predetermined number of questions regardless of context, ignoring task-specific needs. Adaptive methods trigger clarification when model confidence falls below a manually-tuned threshold, but this threshold selection is brittle and fails to generalize across domains or cost structures. Neither approach explicitly reasons about whether the information gained justifies the user's effort.

We argue that agents should treat communication as a rational decision, asking questions only when the expected improvement in task outcomes justifies the user's time and effort. We adopt a Rational Speech Act (RSA) perspective (Goodman and Frank, 2016; Frank and Goodman, 2012) viewing dialogue as a rational action. Building on prior RSA work on interactive questioning-answering (Hawkins et al., 2015) and utility-grounded pragmatic reasoning (Sumers et al., 2021), the agent should only ask questions when the expected benefit of improved downstream decisions outweighs the cost of additional interaction—capturing both cost of communication (Hawkins et al., 2015) and utility of downstream decisions (Sumers et al., 2021). Under this lens, we formalize the clarify-or-commit decision through three contextual factors: (1) **Query Ambiguity**: the degree of uncertainty about the user's true intent; (2) **Task Risk**: the severity of the consequences of a wrong action; and (3) **Cognitive Load**: the cost, in time and effort, imposed on the user by asking for clarification.

To operationalize this reasoning, we propose a decision-theoretic framework grounded in the Value of Information (VoI), a classic principle from decision theory (Raiffa and Schlaifer, 1961). Our inference-time method allows an LLM to explicitly calculate the expected utility gain of asking a potential question, weighing it directly against the communication cost. This provides a principled mechanism for the agent to decide whether the information it might receive is worth the user's attention. Our contributions are threefold: **(a)** We formalize the adaptive communication problem in human-agent interaction from a decision-theoretic perspective, identifying three key factors: ambiguity, risk, and cognitive load. **(b)** We propose a practical, inference-time VOI-based method that allows an LLM to estimate these contextual factors and dynamically decide whether to act or to seek clarifications **(c)** We demonstrate through experiments across four distinct domains: 20 Questions, medical diagnosis, flight booking, and online shopping, that our parameter-free VoI method automatically identifies the optimal operating point. Across varying communication costs, VoI matches or exceeds the best manually-tuned baselines in 18 of 20 conditions, achieving utility gains of up to 1.36 points in high-cost settings.

2 Related Work

Standard LLM Agent Paradigm. Our work is situated within the broader context of developing autonomous LLM agents. Much foundational research in this area focuses on improving agent reasoning, planning, and tool-use capabilities. Prominent paradigms like Yao et al. (2023) and others are often evaluated in benchmarks that, while complex, assume the user's initial instruction is complete and unambiguous (Yao et al., 2022; Zhou et al., 2023; Xie et al., 2024). This focus on task execution rather than the real-world productivity users expect from agents, leaving a critical gap for truly deploying agents (Sun et al., 2025; Shah and White, 2024; Zhou and Sun, 2025).

Recently, a new wave of research has begun to address agent reliability by introducing principled frameworks from decision theory (Liu et al., 2024; Lin et al., 2024; Chen et al., 2025). However, these approaches typically focus on making an optimal decision given a static, pre-defined state of information. Our work bridges these two areas: we adopt the rigor of decision theory but focus on the upstream problem of active information gathering, allowing the agent to dynamically resolve ambiguity before committing to an action.

LLM Proactive Communication. Prior work has explored prompting techniques to improve LLM interactivity. These methods can elicit user preferences (Li et al., 2023) or encourage active disambiguation of ambiguous queries (Deng et al., 2023; Zhang et al., 2024b). While prompting can directly induce clarifying behaviors, prior work shows that the resulting strategies are often suboptimal without more principled planning or learning algorithms. Our work provides such a principled algorithm to govern the agent's communication decisions.

Uncertainty-Gated and Information-Theoretic Methods. A more systematic approach uses model-uncertainty estimates to decide when to seek clarification, triggering a question when prediction confidence or entropy falls below a selected threshold (Wang et al., 2025; Zhang and Choi, 2023; Kuhn et al., 2022; Ren et al., 2023; Grand et al., 2025). While an improvement over heuristics, these information-centric views can be insufficient, as they do not directly consider the downstream task's stakes. Our method addresses this by employing the Value of Information (VoI) (Raiffa and Schlaifer, 1961; Howard, 1966), a core concept from decision theory. Instead of measuring information gain in isolation, VoI measures how that information is expected to improve the utility of the final action, explicitly connecting the purpose of communication to the stakes of the decision.

Learning-Based Approaches. Different from the inference-time algorithms above, another line of research uses reinforcement learning to improve LLM collaboration with humans. Variants of Direct Preference Optimization (DPO) have been applied to encourage models to request clarification when needed (Zhang et al., 2024a; Chen et al., 2024; Wu et al., 2025; Qian et al., 2025; Sun et al., 2025). However, RL is often task-specific, requiring a carefully designed simulation environment and training pipeline, which is fundamentally different from our VOI-based method which operate purely at inference-time.

Rational Speech Act RSA-style pragmatic models cast language as (approximately) rational action: speakers choose utterances to shape a listener's inferences under explicit priors and costs (Frank and Goodman, 2012; Goodman and Frank, 2016). Beyond single-shot reference, RSA has been extended to interactive question–answering, where questions are selected to trade off expected informativeness against asking cost (Hawkins et al., 2015), and to

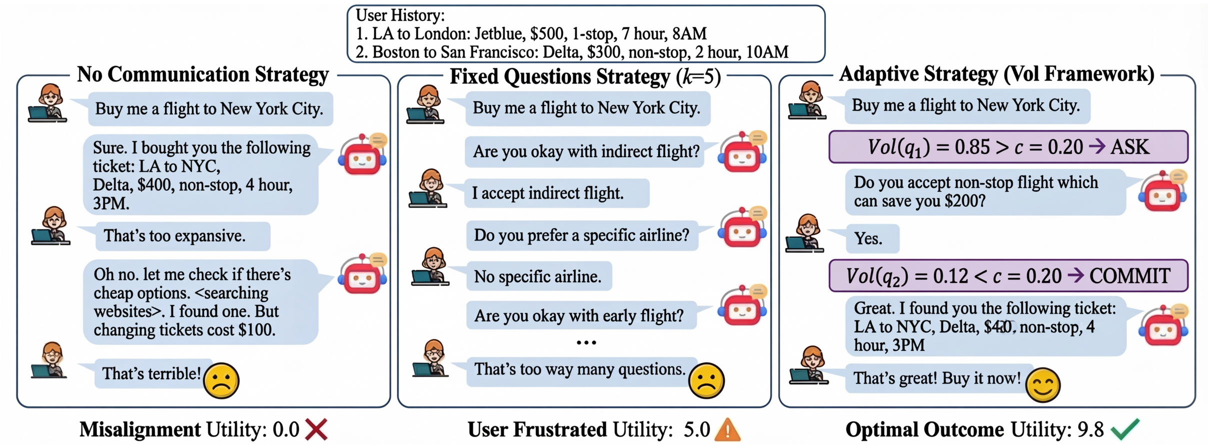


Figure 1: **Illustration of different communication methods and user reaction.** Given user flight history, an LLM agent is able to infer user latent preferences with some probability. Excessive questions that asks about every aspect of preference would lead to user dissatisfaction (A) while directly acting without communication could lead to unexpected consequences (B). Decision-theoretic reasoning can balance expected utility gain via asking user questions against communication cost to achieve efficient but effective communication at inference time (C).

action-oriented settings where the point of communication is not only belief change but improving downstream decisions (e.g., signaling bandits) (Sumers et al., 2021). Researchers then extend to "Neural RSA" that replace hand-specified literal models with learned speakers/listeners in grounded tasks (Andreas and Klein, 2016; Monroe et al., 2017). Most recently, RSA has been adapted to the era of LLMs, serving both as an inference-time control to guide generation (Wang and Demberg, 2024; Cao et al., 2025).

3 Problem Formulation

We formulate the adaptive communication task as a sequential decision-making process where an LLM agent interacts with a user to select an optimal action.

Preliminaries. The agent receives an initial, potentially ambiguous, user query S. The user's true goals and preferences are represented by a latent state $\theta \in \Theta$, which is not directly observable by the agent. The agent has access to a set of possible terminal actions $a \in \mathcal{A}$. To resolve ambiguity about θ and choose the best action a^*, the agent can engage in a multi-turn dialogue with the user.

The Clarify-or-Commit Process. The interaction proceeds in a sequence of turns. At each turn t, given the dialogue history $H_t = (q_1, u_1, \ldots, q_{t-1}, u_{t-1})$, the agent must make a decision:

1. **CLARIFY:** Select and pose a question q_t from a set of possible questions $\mathcal{Q}$. Upon receiving the user's answer u_t, the history is updated to H_{t+1} and the process continues.
2. **COMMIT:** Terminate the dialogue and select a final action $a \in \mathcal{A}$ based on the current history H_t.

The agent's strategy for making this choice at each turn is the **clarify-or-commit** policy, which is the central object of our study. This simple clarify-or-commit choice lies at the heart of adaptive communication: every question carries both the potential to reduce uncertainty and the cost of additional user effort.

Utility and Objective. The success of a committed action a is measured by a utility function $U(\theta, a)$, which quantifies how well the action aligns with the user's true latent state θ. Communication incurs a cost $c(H)$, representing the user's cognitive load, which quantifies the time and effort user spent on the dialogue. If the agent commits to action a after a final history H, the total utility is $U(\theta, a) - c(H)$. The agent's objective is to devise a policy that maximizes the expected total reward, optimally balancing the utility gain from asking questions against cumulative communication cost.

4 Methods

To address the clarify-or-commit problem, an agent requires a principled policy for deciding when the potential benefit of asking a question outweighs the cost of interaction. Simple heuristic-based strategies often fail because they do not explicitly reason about the downstream consequences or the stakes of the decision. To overcome this limitation, we propose an adaptive policy grounded in the Value of Information (VoI), a core concept from decision theory (Raiffa and Schlaifer, 1961).

4.1 Value of Information Framework

The baselines above are either non-adaptive or rely on generic, task-agnostic heuristics like confidence. They fail to explicitly reason about the *value* of the information a question might provide in the context of heterogeneous task stakes and unequal feature importance. To address this, we formalize our approach using the VoI framework.

Beliefs and Expected Utility. Let Θ be the set of possible latent user intents (e.g., the specific product features preferred or the true medical condition). The agent maintains a belief distribution $b(\theta)$ over Θ. Given this belief, the expected utility (EU) of committing to a terminal action $a \in \mathcal{A}$ is:

$$\mathrm{EU}(a \mid b) = \mathbb{E}_{\theta \sim b}[U(\theta, a)] = \sum_{\theta \in \Theta} b(\theta) U(\theta, a). \tag{1}$$

If the agent were to commit immediately, it would choose the action $a^* = \arg\max_{a \in \mathcal{A}} \mathrm{EU}(a \mid b)$. The utility of this decision is the value of acting under the current belief b:

$$V(b) = \max_{a \in \mathcal{A}} \mathrm{EU}(a \mid b). \tag{2}$$

Calculating the Value of a Question. To evaluate a potential question q, the agent considers the set of possible answers $\mathcal{Y}$. For any given answer $y \in \mathcal{Y}$, the agent would update its belief to a posterior $b_y(\theta) = P(\theta \mid H, q, y)$. The expected value of the decision *after* receiving an answer to question q is the expectation over all possible answers y:

$$V_{\text{post}}(b, q) = \sum_{y \in \mathcal{Y}} p(y \mid q, b) \cdot V(b_y), \tag{3}$$

where $p(y \mid q, b)$ is the probability of receiving answer y given the current belief. In practice, to make computation feasible, we restrict the answer space to a closed set of multiple choice or yes-no questions. For each sampled hypothesis θ, we query the LLM to simulate the likelihood of each response y given question q, aggregating these to find the marginal probability $p(y \mid q, b)$.

The **Value of Information** for question q is the difference between the expected utility after asking and the utility of acting now:

$$\mathrm{VoI}(q) = V_{\text{post}}(b, q) - V(b). \tag{4}$$

The Clarify-or-Commit Policy. Our framework uses this VoI calculation to establish a decision rule. At each turn, the agent evaluates the net utility gain for each candidate question:

$$\mathrm{NetVoI}(q) = \mathrm{VoI}(q) - c, \tag{5}$$

where c is the per-question communication cost. The agent selects the question q^* with the highest positive net value. If $\max_q \mathrm{NetVoI}(q) \leq 0$, the expected utility gain from further communication is not worth the cost. The agent terminates the dialogue and commits to the best action under its current belief.

4.2 Instantiation with LLMs

While Section 4.1 establishes the theoretical foundations of our approach, in this section, we describe how we leverage LLMs to approximate these components at inference time.

Estimating and Updating Belief Distributions. Given the set of candidate latent factors Θ, we prompt the LLM to explicitly quantify its uncertainty by outputting a probability distribution $b(\theta)$ over these factors. Different from standard Bayesian approaches update beliefs analytically via a fixed likelihood function, we employ a LLM to estimate the probability distribution over Θ (Kobalczyk et al., 2025a; Liu et al., 2024). To obtain the posterior belief b_y required for Eq. 3, we feed the history augmented with a simulated interaction (question q and hypothetical answer y) back into the model and prompt it to re-estimate the distribution over Θ. This allows the agent to dynamically update its confidence based on the semantic content of the answer.

Simulating User Responses. To calculate the expected value of a question, we perform a one-step lookahead simulation (Kobalczyk et al., 2025b) to estimate the marginal likelihood of possible answers $p(y \mid q, b)$. To ensure computational tractability in Eq. 3, we constrain the agent to ask closed-ended questions (e.g., multiple-choice or Yes-No questions), thereby defining a finite answer space $\mathcal{Y}$. The probability of each response is computed by marginalizing over the current beliefs:

$p(y \mid q, b) \approx \sum_{\theta \in \Theta} p(y \mid q, \theta) b(\theta)$, where the term $p(y \mid q, \theta)$ represents the LLM's prediction of the user's response assuming θ is the ground truth.

5 Experimental Setup

5.1 Baseline Methods

No-Question. This baseline represents the standard agent paradigm. Given the initial query S, the agent commits to an action immediately without any communication with the user. It relies solely on its initial understanding of the user's intent.

Fixed-Round. This non-adaptive baseline asks a fixed number of k questions before committing to an action. It serves to isolate the benefit of interaction from the benefit of *adaptive* interaction by exploring a fixed trade-off between information gathering and communication cost.

Adaptive Prompting. This baseline prompts the LLM to reason about whether it feels confident enough to act or if it should ask a question. The number of questions is not predetermined, but the decision to stop is based on the model's heuristic self-assessment rather than a formal criterion.

Confidence Thresholding. This adaptive baseline formalizes the heuristic of Adaptive Prompting. The agent continues to ask questions as long as its predictive confidence in the best action a^* remains below a tunable threshold τ. We measure confidence using the model's verbalized confidence scores (Tian et al., 2023), a common practice for modern LLMs. This method is adaptive, but crucially, the threshold τ must be manually tuned for each task and cost setting to achieve optimal performance.

5.2 Tasks and Models

Mixed-Stakes 20 Questions. The 20 Questions game is a classic guessing game with a long history as a paradigm for studying human and artificial decision-making under uncertainty. It provides a controlled environment to test how an agent performs strategic information gathering. Following the setup of Hu et al. (2024), the agent must identify a target concept from a known candidate set by asking a series of binary (yes/no) questions. Our key modification is to explicitly test how the agent adapts to varying **task risk**. We create two parallel versions of this task:

- **Low-Stakes (Animal Guessing):** The agent identifies an animal from a set of 100. A correct guess yields a terminal utility of $U = 1$.

- **High-Stakes (Medical Diagnosis):** The agent diagnoses a medical condition from a set of 15 diseases, using real doctor-patient chat histories as input. A correct diagnosis yields a utility of $U = 10$.

Flight Recommendation We adopt a task designed to model the elicitation of multi-faceted user preferences, a common challenge in real-world assistants. Our setup is inspired by the recent work of (Qiu et al., 2025) is derived from the FLIGHTPREF dataset originally proposed by Lin et al. (2022). The agent is presented with a user's choice history over five rounds of flight selections. In a final, held-out round, the agent must predict which of three new flight options the user will prefer. Each flight is defined by 8 features (e.g., price, stops, airline), and each user has a latent reward function defining their preferences over these features. The agent can ask clarifying questions to uncover these preferences before making its final prediction. This task tests the agent's ability to strategically query a complex, multi-attribute preference space to infer a user's reward model from their contextual choices. The agent's prediction for the new round will be scored based on this reward function.

Ambiguous WebShop To test our agent in a more realistic, interactive environment, we adapt the WebShop benchmark (Yao et al., 2022). In the original setting, user instructions are created to be relatively well-specified (e.g., "buy a red Adidas t-shirt, size medium"). We deliberately introduce **query ambiguity** by removing details from the user's request (e.g., "buy a t-shirt") to simulate underspecified real-world user query. The agent must then decide whether to act on this partial information (e.g., `search("t-shirt")`) or to ask clarifying questions about attributes like size, color, or brand. This task evaluates the agent's ability to balance autonomous web navigation with strategic information gathering to resolve under-specified user requests.

Models We consider a selection of leading LLMs to evaluate the performance of our proposed method, including GPT-4.1 (OpenAI, 2025) and Gemini-2.5-Flash (Comanici et al., 2025).

Algorithm 1 VOI Algorithm

Require: Instruction S; action set $\mathcal{A}$; utility $U(\theta, a)$; question generator GenQ; belief updater Update; cost $c(\cdot)$; clarification budget $K_{\max}$

$H \leftarrow \{S\}$; $b \leftarrow \mathrm{Prior}(S)$
for $t = 1, 2, \ldots, K_{\max}$ **do**
  $Q \leftarrow \mathrm{GenQ}(H)$ ▷ small set of targeted questions
  $V_0 \leftarrow V(b) = \max_{a \in \mathcal{A}} \mathbb{E}_{\theta \sim b}[U(\theta, a)]$
  for all $q \in Q$ **do**
    Sample plausible replies $\{(y_k, \pi_k)\}_{k=1}^{K}$ from $P(\cdot \mid b, q)$
    $V_q \leftarrow \sum_{k=1}^{K} \pi_k \, V(\mathrm{Update}(b, q, y_k))$
    $\mathrm{VoI}(q) \leftarrow V_q - V_0 - c(q)$
  end for
  $q^* \leftarrow \arg\max_{q \in Q} \mathrm{VoI}(q)$
  if $\mathrm{VoI}(q^*) \leq 0$ **then break** ▷ clarification not worthwhile
  else
    Ask q^*, observe y; $H \leftarrow H \cup \{(q^*, y)\}$; $b \leftarrow \mathrm{Update}(b, q^*, y)$
  end if
end for
return $a^* \in \arg\max_{a \in \mathcal{A}} \mathbb{E}_{\theta \sim b}[U(\theta, a)]$ ▷ final commitment

6 Results

6.1 Main Results

Our central findings are summarized in Figure 2. Across all tasks and communication cost settings, our VoI-based agent consistently achieves state-of-the-art utility. Crucially, it does so without requiring task-specific threshold tuning, showcasing its robustness and practical advantages.

VoI excels by finding the optimal utility-cost balance. As shown in Figure 2, our VoI agent (starred marker) consistently ranks as the top-performing method across the Mixed 20Q, Flight Recommendation, and Ambiguous WebShop tasks. For instance, in Mixed 20Q with a communication cost of $c = 0.01$, VOI achieves a utility of 14.14, significantly outperforming the best-tuned confidence-thresholding baseline (11.49 at $\tau = 0.90$). This performance advantage stems from VOI's ability to dynamically determine the optimal number of clarification questions, a stark contrast to fixed-round and confidence-based methods that require brittle, manual tuning of a threshold for each specific task and cost structure.

Adaptive communication is essential for ambiguous tasks. The "No Question" baseline establishes the necessity of proactive communication. On the Mixed 20Q task, where the initial query is inherently underspecified, this baseline's accuracy is near zero for both low-stakes (animal) and high-stakes (medical) variants. However, as shown in Figures 2(f) and 2(l), when communication costs are prohibitively high, avoiding questions becomes a competitive strategy. In these scenarios, our VOI method correctly adapts by stopping communication early, demonstrating its ability to gracefully handle the full spectrum of cost-benefit scenarios.

Adaptive prompting are insufficient for robust performance. The Adaptive Prompting baseline shows that simply instructing an LLM to "ask questions when needed" offers an improvement over non-adaptive strategies. However, its performance is inconsistent and consistently lower than more structured methods. This is because the decision to communicate is based on the model's uncalibrated, internal "feeling" of confidence, rather than a formal criterion. It lacks a principled mechanism to weigh the potential information gain against the explicit communication cost, leading to suboptimal and unpredictable behavior.

Fixed-round communication strategies are fundamentally suboptimal. A fixed-round policy, which asks a predetermined number of questions, fails to adapt to the specific needs of a given query. As illustrated in the inverted-*U* shape of the "Fixed Round" curves in Figure 2, utility initially increases with more questions but then declines as communication costs overwhelm the benefits of additional information. The optimal number of questions varies significantly with the task and cost, highlighting the necessity of an adaptive policy.

Confidence thresholding is effective but brittle. The confidence thresholding baseline provides a strong, adaptive competitor. With the *correctly* tuned confidence threshold τ, its performance can be comparable to our VOI method (e.g., on GPT-4 for Mixed 20Q and Webshop). However, this effectiveness is its Achilles' heel; the optimal τ is highly sensitive and must be manually selected for each task and cost combination, making it impractical for real-world deployment. Our VoI method provides a principled solution that matches or exceeds this performance without any such manual tuning.

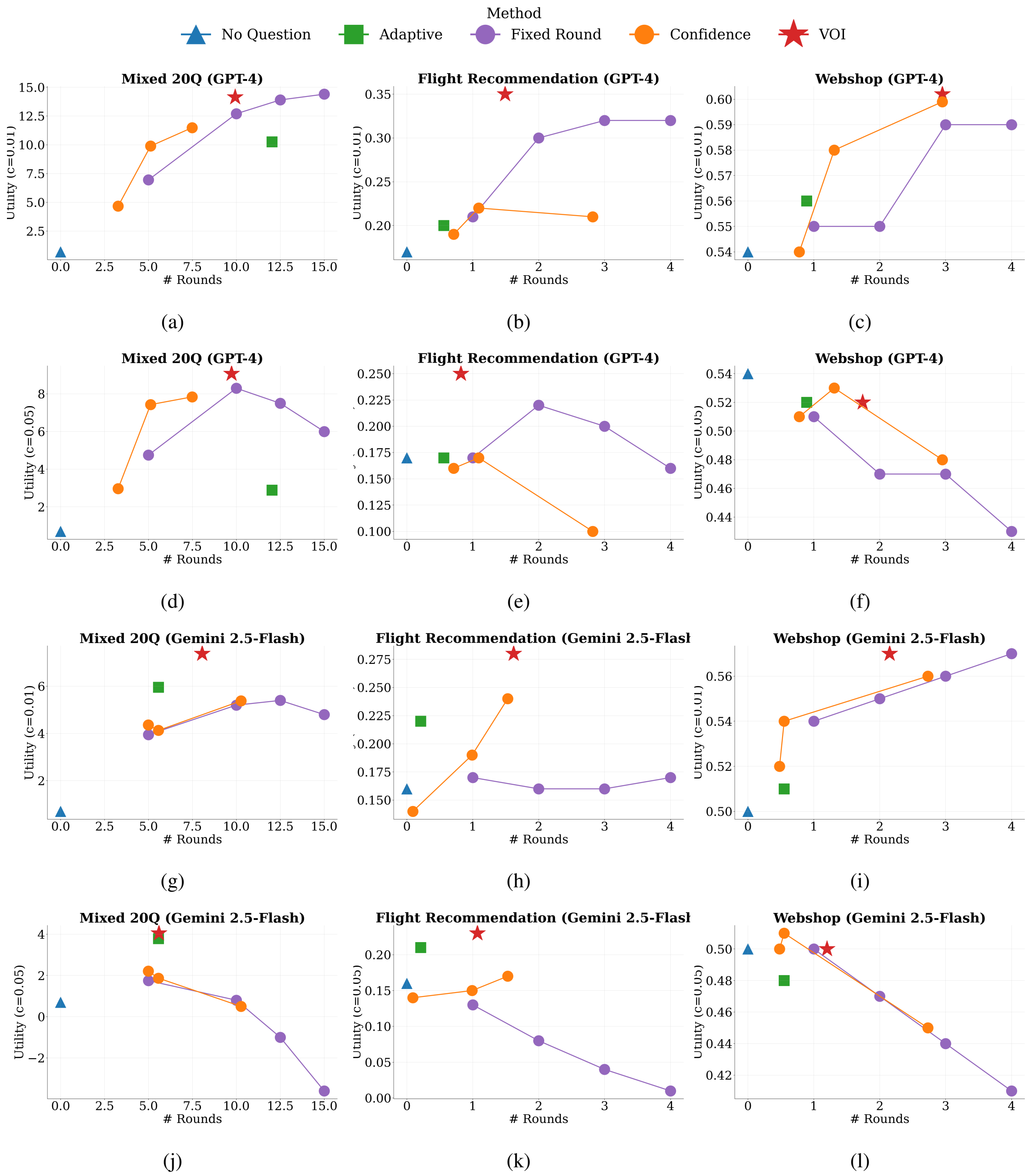


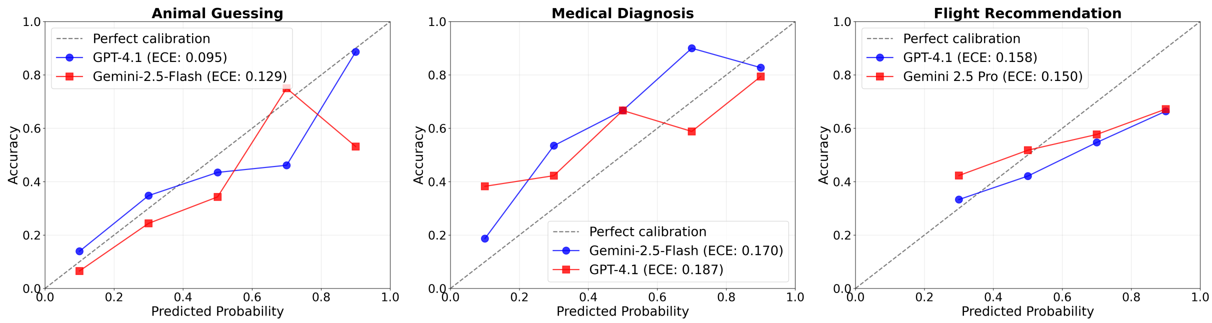
Figure 2: **Utility vs. Communication Rounds.** Final utility as a function of the number of clarification questions asked across our three tasks, for GPT-4 (top two rows) and Gemini-2.5-Flash (bottom two rows), with communication costs $c = 0.01$ and $c = 0.05$. Utility is defined as $U(\theta, a) - T \cdot c$. The curves for Fixed Round and Confidence Thresholding represent Pareto frontiers generated by varying their respective hyperparameters (k and τ). In contrast, our VoI agent (starred) is a parameter-free method. In nearly all settings, VoI automatically identifies an operating point that matches or exceeds the performance of the best-tuned baseline, demonstrating its superior adaptability and practical value.

6.2 Ablation Study

Ablation on Communication Cost. As shown in Table 1, across the cost sweep on Mixed 20-Question the VoI controller matches or exceeds the strongest grid-searched baselines. We tune four baselines over nine threshold settings, and while the best baseline shifts with the communication cost, VoI consistently selects an appropriate number of questions thst match the performance of the best baseline. Importantly, this pattern is stable across different choice of communication costs: VoI adapts smoothly to the stated cost rather than

Table 1: **VOI vs. Baselines Across Costs (Gemini-2.5-Flash, Mixed 20 Question)**. This table compares the VOI policy's expected reward (r_{VOI}) against the best and second-best baselines via grid searching over 9 values. The Δ columns report VOI's margin over each baseline (positive means VOI is better).

Cost	Best Baseline	$r_{\max}$	Second Best	r_{second}	r_{VOI}	$r_{\text{VOI}}-r_{\max}$	$r_{\text{VOI}}-r_{\text{second}}$
0.01	Confidence (τ=0.9)	8.30	Round (τ=15)	8.10	8.64	0.34	0.54
0.02	Confidence (τ=0.9)	6.88	Confidence (τ=0.9)	6.80	7.72	0.84	0.92
0.05	Round (τ=5)	3.65	Confidence (τ=0.5)	3.64	5.01	1.36	1.37
0.10	Confidence (τ=0.5)	2.28	Round (τ=5)	0.90	1.38	-0.90	0.48
0.20	No Question	0	Round (τ=5)	-4.60	-0.96	-0.96	3.64



(a) (b) (c)

Figure 3: **Calibration Analysis** The figure presents the calibration analysis of GPT-4 and Gemini-2.5-Flash on Animal Guessing, Medical Diagnoiss, and Flight Recommendation. (In (c) the accuracy for predicted probability between 0 and 0.2 is omitted because very few samples fall in that range.

hinging on a brittle threshold choice.

Calibration Analysis. A critical component of our VoI framework is the LLM's ability to estimate a belief distribution $b(\theta)$ over latent user states. To analyze it, ideally we should compare model predicted distribution to the ground truth distribution. However, in the absence of the ground truth distribution for our tasks, we instead measure the argmax from the distribution against the ground truth item as the standard calibration analysis to approximate its distribution estimation accuracy. As shown in Figure 3, The results reveal that models are reasonably calibrated in Animal Guessing game but less calibrated for Medical Diagnosis which we suspect because of the inherent complication and noise in the symptoms of diseases. Despite this, we see that VOI are empirically effective and robust that consistently matches if not perform the best baselines after searching hyperparameters. We believe that current and future work that could improving model calibration under missing context (Li et al., 2025) could further improve the performance of VOI.

7 Conclusion

Current LLM agents are often designed for well-specified tasks, leaving them brittle when faced with the inherent ambiguity of real-world user requests. In this work, we argued that overcoming this limitation requires agents to move beyond simple execution and develop a principled strategy for adaptive communication. We proposed a formal framework for this problem, centered on balancing three key factors: query ambiguity, task risk, and user cognitive load. Our primary contribution is a practical, inference-time method based on the Value of Information (VoI) that operationalizes this framework. By explicitly calculating the expected utility gain of a potential question and weighing it against its communication cost, our VoI-driven agent decides when to act and when to ask. Extensive experiments across diverse domains—including medical diagnosis and online shopping—demonstrate that our approach consistently outperforms non-adaptive and heuristic-based baselines. Crucially, it achieves this without the need for the brittle, task-specific threshold tuning that plagues other adaptive methods. Ultimately, this work provides a principled foundation for building LLM agents that are not just capable executors, but also thoughtful communicators. By equipping agents with a formal understanding of when information is valuable, we can create more aligned, efficient, and truly collaborative human-AI systems.

Limitations

Scope of Interaction: Decision vs. Generation. Our work focuses on the core decision of *when* to communicate, rather than *what* questions to generate. To this end, our experiments utilize a predefined set of actions ($a \in \mathcal{A}$) and clarifying questions, a methodological choice consistent with prior work (Hu et al., 2024; Kobalczyk et al., 2025b). This controlled setting isolates the performance of our VoI-based *selection policy*, providing an unambiguous evaluation of our central claim. By controlling for the quality of question generation, we demonstrate the effectiveness of the decision-making principle itself. Extending this framework to fully open-ended dialogue is an important next step; establishing this selection principle is a necessary foundation. Our work provides the core engine around which more sophisticated generative components can be built.

Model of Communication Cost. We employ a linear communication cost model ($c(H) = T \cdot c$). Accurately modeling the nuances of human cognitive load is a major, open research challenge in its own right, spanning HCI and cognitive science. Therefore, in line with common practice in decision-theoretic analyses, we adopt a simplified and interpretable cost function. This allows us to clearly illustrate the fundamental trade-off between utility gain and cost, without introducing confounding variables from a more complex, speculative cognitive model. Importantly, the VoI framework itself is agnostic to the form of the cost function; the core decision rule, $\mathrm{VoI}(q) - c(H)$, can readily incorporate more sophisticated models as they are developed. We view the linear cost model as a reasonable first-order approximation that demonstrates the framework's viability, with refinement through empirical user research as a natural next step.

Ethical Considerations

While our VoI framework optimizes the trade-off between information gain and communication cost, user agency must remain paramount: users should retain the ability to decline questions or proceed without clarification based on their own judgment. Beyond this, the act of questioning introduces critical considerations regarding user burden and privacy. First, clarifying questions—even when theoretically optimal—inherently impose a cognitive demand on the user; an agent that queries too frequently or intrusively risks eroding trust and causing frustration, necessitating cost models that strictly penalize unnecessary interruptions. Second, the pursuit of resolving ambiguity often requires eliciting specific, potentially sensitive information (e.g., medical symptoms or personal preferences) to update the agent's belief distribution. It is imperative that future implementations incorporate strict data minimization principles and privacy safeguards to ensure that the agent's drive for reduced uncertainty does not compromise user privacy or comfort. We acknowledge the use of AI tools for refining the paper writing.

References

Jacob Andreas and Dan Klein. 2016. Reasoning about pragmatics with neural listeners and speakers. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1173–1182, Austin, Texas. Association for Computational Linguistics.

Zhuchen Cao, Sven Apel, Adish Singla, and Vera Demberg. 2025. Pragmatic reasoning improves llm code generation. *Preprint*, arXiv:2502.15835.

Maximillian Chen, Ruoxi Sun, Sercan Ö Arık, and Tomas Pfister. 2024. Learning to Clarify: Multi-turn Conversations with Action-Based Contrastive Self-Training.

Xiusi Chen, Shanyong Wang, Cheng Qian, Hongru Wang, Peixuan Han, and Heng Ji. 2025. Decision-flow: Advancing large language model as principled decision maker. *ArXiv preprint*, abs/2505.21397.

Gheorghe Comanici, Eric Bieber, Mike Schaekermann, Ice Pasupat, Noveen Sachdeva, Inderjit Dhillon, Marcel Blistein, Ori Ram, Dan Zhang, Evan Rosen, and 1 others. 2025. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. *arXiv preprint arXiv:2507.06261*.

Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, and Tat-Seng Chua. 2023. Prompting and evaluating large language models for proactive dialogues: Clarification, target-guided, and non-collaboration. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 10602–10621, Singapore. Association for Computational Linguistics.

Michael C Frank and Noah D Goodman. 2012. Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998.

Noah D Goodman and Michael C Frank. 2016. Pragmatic language interpretation as probabilistic inference. *Trends in cognitive sciences*, 20(11):818–829.

Gabriel Grand, Valerio Pepe, Jacob Andreas, and Joshua B Tenenbaum. 2025. Shoot first, ask questions later? building rational agents that explore and act like people. *arXiv preprint arXiv:2510.20886*.

Robert XD Hawkins, Andreas Stuhlmuller, Judith Degen, and Noah D Goodman. 2015. Why do you ask? good questions provoke informative answers. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 37.

Ronald A. Howard. 1966. Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2(1):22–26.

Zhiyuan Hu, Chumin Liu, Xidong Feng, Yilun Zhao, See-Kiong Ng, Anh Tuan Luu, Junxian He, Pang Wei Koh, and Bryan Hooi. 2024. Uncertainty of thoughts: Uncertainty-aware planning enhances information seeking in large language models. *ArXiv preprint*, abs/2402.03271.

Katarzyna Kobalczyk, Nicolas Astorga, Tennison Liu, and Mihaela van der Schaar. 2025a. Active Task Disambiguation with LLMs.

Katarzyna Kobalczyk, Nicolas Astorga, Tennison Liu, and Mihaela van der Schaar. 2025b. Active task disambiguation with llms. *ArXiv preprint*, abs/2502.04485.

Lorenz Kuhn, Yarin Gal, and Sebastian Farquhar. 2022. CLAM: Selective Clarification for Ambiguous Questions with Generative Language Models.

Belinda Z. Li, Alex Tamkin, Noah Goodman, and Jacob Andreas. 2023. Eliciting Human Preferences with Language Models.

Xiaomin Li, Zhou Yu, Ziji Zhang, Yingying Zhuang, Swair Shah, Narayanan Sadagopan, and Anurag Beniwal. 2025. Semantic volume: Quantifying and detecting both external and internal uncertainty in llms. *ArXiv preprint*, abs/2502.21239.

Jessy Lin, Daniel Fried, Dan Klein, and Anca Dragan. 2022. Inferring rewards from language in context. *arXiv preprint arXiv:2204.02515*.

Jessy Lin, Nicholas Tomlin, Jacob Andreas, and Jason Eisner. 2024. Decision-oriented dialogue for human-AI collaboration. *Transactions of the Association for Computational Linguistics*, 12:892–911.

Ollie Liu, Deqing Fu, Dani Yogatama, and Willie Neiswanger. 2024. Dellma: Decision making under uncertainty with large language models. *ArXiv preprint*, abs/2402.02392.

Chaitanya Malaviya, Joseph Chee Chang, Dan Roth, Mohit Iyyer, Mark Yatskar, and Kyle Lo. 2024. Contextualized Evaluations: Taking the Guesswork Out of Language Model Evaluations.

Will Monroe, Robert X.D. Hawkins, Noah D. Goodman, and Christopher Potts. 2017. Colors in context: A pragmatic neural model for grounded language understanding. *Transactions of the Association for Computational Linguistics*, 5:325–338.

OpenAI. 2025. Introducing gpt-4.1 in the api. Accessed: 2025-09-18.

Andi Peng, Andreea Bobu, Belinda Z. Li, Theodore R. Sumers, Ilia Sucholutsky, Nishanth Kumar, Thomas L. Griffiths, and Julie A. Shah. 2024. Preference-conditioned language-guided abstraction. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, HRI 2024, Boulder, CO, USA, March 11-15, 2024*, pages 572–581. ACM.

Cheng Qian, Zuxin Liu, Akshara Prabhakar, Jielin Qiu, Zhiwei Liu, Haolin Chen, Shirley Kokane, Heng Ji, Weiran Yao, Shelby Heinecke, and 1 others. 2025. Userrl: Training interactive user-centric agent via reinforcement learning. *arXiv preprint arXiv:2509.19736*.

Linlu Qiu, Fei Sha, Kelsey Allen, Yoon Kim, Tal Linzen, and Sjoerd van Steenkiste. 2025. Bayesian teaching enables probabilistic reasoning in large language models. *arXiv preprint arXiv:2503.17523*.

H. Raiffa and R. Schlaifer. 1961. *Applied Statistical Decision Theory*. Studies in managerial economics. Division of Research, Graduate School of Business Administration, Harvard University.

Allen Z Ren, Anushri Dixit, Alexandra Bodrova, Sumeet Singh, Stephen Tu, Noah Brown, Peng Xu, Leila Takayama, Fei Xia, Jake Varley, and 1 others. 2023. Robots that ask for help: Uncertainty alignment for large language model planners. *arXiv preprint arXiv:2307.01928*.

Chirag Shah and Ryen W White. 2024. Agents are not enough. *ArXiv preprint*, abs/2412.16241.

Theodore R Sumers, Robert D Hawkins, Mark K Ho, and Thomas L Griffiths. 2021. Extending rational models of communication from beliefs to actions. *arXiv preprint arXiv:2105.11950*.

Weiwei Sun, Xuhui Zhou, Weihua Du, Xingyao Wang, Sean Welleck, Graham Neubig, Maarten Sap, and Yiming Yang. 2025. Training proactive and personalized llm agents. *arXiv preprint arXiv:2511.02208*.

Katherine Tian, Eric Mitchell, Allan Zhou, Archit Sharma, Rafael Rafailov, Huaxiu Yao, Chelsea Finn, and Christopher Manning. 2023. Just ask for calibration: Strategies for eliciting calibrated confidence scores from language models fine-tuned with human feedback. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 5433–5442, Singapore. Association for Computational Linguistics.

Jimmy Wang, Thomas Zollo, Richard Zemel, and Hongseok Namkoong. 2025. Adaptive elicitation of latent information using natural language. *ArXiv preprint*, abs/2504.04204.

Yifan Wang and Vera Demberg. 2024. RSA-control: A pragmatics-grounded lightweight controllable text generation framework. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 5561–5582, Miami, Florida, USA. Association for Computational Linguistics.

Shirley Wu, Michel Galley, Baolin Peng, Hao Cheng, Gavin Li, Yao Dou, Weixin Cai, James Zou, Jure Leskovec, and Jianfeng Gao. 2025. Collabllm: From passive responders to active collaborators. *ArXiv preprint*, abs/2502.00640.

Tianbao Xie, Danyang Zhang, Jixuan Chen, Xiaochuan Li, Siheng Zhao, Ruisheng Cao, Toh J Hua, Zhoujun Cheng, Dongchan Shin, Fangyu Lei, and 1 others. 2024. Osworld: Benchmarking multimodal agents for open-ended tasks in real computer environments. *Advances in Neural Information Processing Systems*, 37:52040–52094.

Shunyu Yao, Howard Chen, John Yang, and Karthik Narasimhan. 2022. Webshop: Towards scalable real-world web interaction with grounded language agents. In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.

Shunyu Yao, Noah Shinn, Pedram Razavi, and Karthik Narasimhan. 2024. tau-bench: A benchmark for tool-agent-user interaction in real-world domains. *arXiv preprint arXiv:2406.12045*.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik R Narasimhan, and Yuan Cao. 2023. React: Synergizing reasoning and acting in language models. In *The Eleventh International Conference on Learning Representations*.

Michael J. Q. Zhang, W. Bradley Knox, and Eunsol Choi. 2024a. Modeling Future Conversation Turns to Teach LLMs to Ask Clarifying Questions.

Michael JQ Zhang and Eunsol Choi. 2023. Clarify when necessary: Resolving ambiguity through interaction with lms. *ArXiv preprint*, abs/2311.09469.

Xuan Zhang, Yang Deng, Zifeng Ren, See-Kiong Ng, and Tat-Seng Chua. 2024b. Ask-before-plan: Proactive language agents for real-world planning. *ArXiv preprint*, abs/2406.12639.

Shuyan Zhou, Frank F Xu, Hao Zhu, Xuhui Zhou, Robert Lo, Abishek Sridhar, Xianyi Cheng, Tianyue Ou, Yonatan Bisk, Daniel Fried, and 1 others. 2023. Webarena: A realistic web environment for building autonomous agents. *ArXiv preprint*, abs/2307.13854.

Xuhui Zhou and Weiwei Sun. 2025. The quest of User-effective AI agents. Blog post. Accessed: 2026-01-06.

A Case Study: VoI is Risk-Aware

Figure 4 provides a compelling qualitative example of why the VoI framework is superior to heuristic-based methods like confidence thresholding. The experiment contrasts a low-stakes task (guessing an animal, reward=1) with a high-stakes task (medical diagnosis, reward=10), using an identical communication cost ($c = 0.05$).

In the **high-stakes medical diagnosis** (Fig. 4b), the potential reward for a correct answer is high. The VoI agent correctly calculates that even questions with moderate information gain are valuable enough to outweigh the communication cost. It, therefore, continues to ask clarifying questions until it is highly confident, stopping several rounds *after* the confidence-thresholding baseline would have stopped, even though significant ambiguity remains, leading to an incorrect diagnosis.

In the **low-stakes animal guessing game** (Fig. 4a), the maximum potential utility is low. Here, the VoI agent correctly assesses that the potential utility gain from asking many questions is not worth the cumulative communication cost. It, therefore, halts the conversation earlier than the confidence-thresholding method, avoiding unnecessary cognitive load on the user for a low-risk task. The confidence-based agent, blind to the low stakes, would have continued asking questions, needlessly imposing cognitive load on the user for a trivial task.

This case study reveals that effective communication requires balancing two distinct pressures: the drive to reduce uncertainty (an epistemic goal) and the need to consider the task's stakes (a utilitarian goal). Confidence-based methods address only the former. The VoI framework excels because it naturally unifies both: it quantifies the value of reducing uncertainty precisely in terms of its expected impact on the final, stake-weighted utility. This principled balance enables the agent to be appropriately cautious in high-stakes scenarios and efficient in low-stakes ones—a critical capability for building trustworthy and effective human-AI collaborators.

B Main Results in Tables

Animal Guessing Game

Dialogue	Conf.	VOI	Prediction
Q1: Is the animal a mammal? **A1:** Yes	5%	0.26	Elephant ❌
Q2: Is the animal primarily found on land? **A2:** Yes	21%	0.22	Elephant ❌
Q3: Is the animal larger than human? **A3:** No	41%	0.20	Otter ❌
...			
Q8: Is the animal native to Australia? **A8:** No	55%	0.05 ⚠️	Alpaca ✅
...			
Q17: Is the animal known for its long neck? **A17:** Yes	90% ⚠️	--	Alpaca ✅

Medical Diagnosis

Dialogue	Conf.	VOI	Prediction
Q1: Are you experiencing any abdominal pain? **A1:** Yes	10%	0.6	Appendicitis ❌
Q2: Do you have any nausea or vomiting? **A2:** No	30%	0.34	Irritable Bowel Syndrome ❌
Q3: Have you noticed changes in bowel movements? **A3:** Yes	60%	2.15	Irritable Bowel Syndrome ❌
...			
Q9: Is pain in the lower left side? **A9:** Yes	90% ⚠️	0.70	Irritable Bowel Syndrome ❌
...			
Q13: Recent weight loss or loss of appetite? **A13:** No	--	0.04 ⚠️	Constipation ✅

Figure 4: A side by side comparison for different methods for Mixed 20 Question task. The figure contrasts four controllers—No-Ask, Fixed-Round, Confidence Thresholding ($\tau = 0.90$), and our VOI policy—on a single Mixed 20Q instance with communication cost $c = 0.05$. Task stakes are encoded directly in the terminal utility: a correct animal guess yields reward 1 (low stakes), whereas a correct medical diagnosis yields reward 10 (high stakes). The objective maximizes decision utility minus dialogue cost, $U(\theta, a) - c(\xi)$.

Figure 5: GPT-4: results for different methods and thresholds across three tasks. For **Webshop**, LLM is normalized by 10 and utilities are Util $=$ LLM $- \#T \times \{0.01, 0.05\}$. Mixed 20Q utilities are recomputed per spec. Within each *method*, the best utility is underlined. The global best per task/cost is ***bold+italic*** and the second best is **bold**.

Method	Mixed 20Q							Flight Rec.					Webshop				
	τ	Acc. (Animal)	Acc. (Med)	#T (Animal)	#T (Med)	Util. (0.01)	Util. (0.05)	τ	Reward	#T	Util. (0.01)	Util. (0.05)	τ	LLM	#T	Util. (0.01)	Util. (0.05)
No Question	–	0.01	0.06	0.00	0.00	0.70	0.70	–	0.17	0.00	0.17	0.17	–	0.54	0.00	0.54	***0.54***
Adaptive	–	0.68	0.53	17.80	6.254	10.26	2.89	–	0.20	0.56	0.20	0.17	–	0.57	0.89	0.56	0.52
Fixed Round	5	0.24	0.51	5.00	5.00	6.95	4.75	1.00	0.22	1.00	0.21	0.17	1.00	0.56	1.00	0.55	0.51
	10	0.60	0.78	10.00	10.00	12.70	8.30	2.00	0.32	2.00	**0.30**	0.22	2.00	0.57	2.00	0.55	0.47
	15	0.77	0.78	15.00	10.00	13.90	7.50	3.00	0.35	3.00	***0.32***	0.20	3.00	0.62	3.00	0.59	0.47
	20	0.87	0.78	20.00	10.00	***14.40***	6.00	4.00	0.36	4.00	***0.32***	0.16	4.00	0.63	4.00	0.59	0.43
Confidence	0.50	0.20	0.31	4.01	2.54	4.67	2.97	0.50	0.19	0.71	0.19	0.16	0.50	0.55	0.78	0.54	0.51
	0.70	0.45	0.60	5.68	4.56	9.89	7.43	0.70	0.23	1.09	0.22	0.17	0.70	0.60	1.31	0.58	**0.53**
	0.90	0.59	0.65	8.48	6.49	11.49	7.84	0.90	0.24	2.82	0.21	0.10	0.90	0.63	2.95	***0.60***	0.48
VOI	0.01	0.76	0.78	11.80	8.07	**14.14**	***9.10***	0.01	0.36	1.49	**0.35**	***0.28***	0.01	0.63	2.95	***0.60***	0.49
	0.05	0.74	0.78	11.46	7.99	13.97	**9.07**	0.05	0.29	0.82	0.29	**0.25**	0.05	0.61	1.74	0.59	0.52

Table 2: Gemini-2.5-Flash: results for different methods and thresholds across three tasks. Format is the same as Figure 5

Method	Mixed 20Q							Flight Rec.					Webshop				
	τ	Acc. (Animal)	Acc. (Med)	#T (Animal)	#T (Med)	Util. (0.01)	Util. (0.05)	τ	Reward	#T	Util. (0.01)	Util. (0.05)	τ	LLM	#T	Util. (0.01)	Util. (0.05)
No Question	–	0.01	0.06	0.00	0.00	0.70	**0.70**	–	0.16	0.00	0.16	0.16	–	0.50	0.00	0.50	**0.50**
Adaptive	–	0.28	0.37	4.78	6.36	5.96	**3.79**	–	0.22	0.21	0.22	0.21	–	0.51	0.55	0.51	0.48
Fixed Round	5	0.16	0.29	5.00	5.00	3.95	1.75	1.00	0.18	1.00	0.17	0.13	1.00	0.55	1.00	0.54	**0.50**
	10	0.33	0.30	10.00	10.00	5.20	0.80	2.00	0.18	2.00	0.16	0.08	2.00	0.57	2.00	0.55	0.47
	15	0.40	0.30	15.00	10.00	5.40	-1.00	3.00	0.19	3.00	0.16	0.04	3.00	0.59	3.00	**0.56**	0.44
	20	0.39	0.30	20.00	10.00	4.80	-3.60	4.00	0.21	4.00	0.17	0.01	4.00	0.61	4.00	***0.57***	0.41
Confidence	0.50	0.22	0.27	4.87	5.12	4.36	2.21	0.50	0.14	0.09	0.14	0.14	0.50	0.52	0.48	0.52	0.50
	0.70	0.16	0.31	5.06	6.08	4.13	1.87	0.70	0.20	0.99	0.19	0.15	0.70	0.54	0.55	0.54	***0.51***
	0.90	0.36	0.30	11.28	9.25	5.38	0.50	0.90	0.25	1.53	0.24	0.17	0.90	0.59	2.73	**0.56**	0.45
VOI	0.01	0.28	0.55	8.48	7.63	***7.38***	3.68	0.01	0.30	1.62	***0.28***	**0.22**	0.01	0.59	2.15	***0.57***	0.48
	0.05	0.15	0.50	4.20	6.99	**6.01**	***4.05***	0.05	0.28	1.07	**0.27**	***0.23***	0.05	0.56	1.20	0.55	**0.50**

C Prompts

C.1 Mixed 20 Questions

874

Animal — Direct Prompting

You are playing 20 Questions as the guesser. Your goal is to figure out what animal I'm thinking of by asking questions.
You have asked {question_count} questions so far and have {remaining_questions} questions left.
The possible animals you're trying to guess include: {answer_set}
Rules:
1. Ask only yes/no questions (answerable with "Yes"/"No").
2. Ask one question at a time.
3. Keep asking until you use all 20 questions.
4. Do not ask the same question twice.
5. Do not guess a specific animal early (e.g., "Is it a cat?"). Start broad to narrow options.

Figure 6: Direct Prompting (Animal 20 Question)

Animal — Auto Stop

You are playing 20 Questions as the guesser. Your goal is to figure out what animal I'm thinking of.
You have asked {question_count} questions so far and have {remaining_questions} questions left.
The possible animals you're trying to guess include: {answer_set}
Rules:
1. Ask only yes/no questions.
2. Ask one question at a time.
3. When you're ready to guess, output: `My guess is: [animal].`
4. Do not ask the same question twice.
5. Avoid premature specific guesses; use broader categories to narrow down.

Figure 7: Auto Stop (Animal 20 Question)

Animal — Confidence Thresholding

You are tasked with guessing the animal based on the entire conversation. Provide a final guess, even if uncertain. The set of possible animals is: {answer_set}.
Also estimate your confidence from 1 to 100 (100 = highest).
Output exactly: `{"guess": "<animal>", "confidence": <number 1-100>}`
Confidence Scale (1–100):
1–20: Very minimal evidence.
21–40: Weak evidence.
41–60: Moderate evidence.
61–80: Strong evidence.
81–100: Very strong evidence.

Figure 8: Confidence Thresholding (Animal 20 Question)

Animal — VOI: Question Generation

I'm playing a game of 20 Questions to identify an animal. Based on previous questions and answers: {previous_qa}
Generate 5 yes/no questions that help identify the animal.
Format as a numbered list.

Figure 9: VOI: Question Generation (Animal 20 Question)

Animal — VOI: Batch Answer Simulation

You are playing a game of Twenty Questions. You will receive a yes/no question and a list of animals. Answer for each animal with: **Yes** (clearly correct), **No** (clearly incorrect), or **Maybe** (uncertain/ambiguous/varies by species or not confident). Prefer "Maybe" over guessing.
Format exactly as: `Animal Name: Yes/No/Maybe`
Question: {question}
Animals: {candidate_list}

Figure 10: VOI: Batch Answer Simulation (Animal 20 Question)

Medical — Direct Prompting

You are the doctor and must diagnose the patient using only yes/no questions.
You have asked {question_count} questions so far and have {remaining_questions} left.
Possible diagnoses: {answer_set}
You may ask up to 20 yes/no questions to understand the condition. At the end, output your diagnosis.

Figure 11: Direct Prompting (Medical Diagnosis)

Medical — Auto Stop

You are the doctor and may ask up to 20 yes/no questions to diagnose the patient.
You have asked {question_count} questions so far and have {remaining_questions} left.
Possible diagnoses: {answer_set}
You can ask up to 10 yes/no questions. Stop when you have enough information.
Format your guess as: `My guess is: [diagnosis].`

Figure 12: Auto Stop (Medical Diagnosis)

Medical — Confidence Thresholding

Diagnose the patient based on the entire conversation. Provide a final diagnosis, even if uncertain.
Set of diseases: {answer_set}.
Also estimate your confidence (1–100).
Output exactly: `{"guess": "<diagnosis>", "confidence": <number 1-100>}`
Confidence Scale (1–100):
1–20: Extremely uncertain.
21–40: Weak evidence.
41–60: Moderate evidence.
61–80: Strong evidence.
81–100: Very strong evidence.

Figure 13: Confidence Thresholding (Medical Diagnosis)

Medical — VOI: Question Generation

I'm a doctor trying to diagnose a patient's condition through a series of questions. Based on symptoms and previous answers:
{previous_qa}
Generate 5 yes/no questions that most effectively narrow the possible conditions (roughly halving the set each time).
Focus on distinguishing symptoms, risk factors, or medical history.
Format as a numbered list.

Figure 14: VOI: Question Generation (Medical Diagnosis)

Medical — VOI: Batch Answer Simulation

I'm a medical diagnostician. Below is a yes/no question and a list of medical conditions.
Question: "{question}"
For each condition, answer with just "Yes" or "No", based on typical presentation.
Reply exactly as: `Condition: Answer`
Conditions: {candidate_list}

Figure 15: VOI: Batch Answer Simulation (Medical Diagnosis)

C.2 Flight Recommendation

Direct Prompting and Confidence Thresholding

Opening
User: Help me pick flights. My preferences are fixed; infer them and choose. Use your best judgement; don't ask for more info.

— SUPPORT HISTORY —
User: Which flight is best?
Flight 1: {option 1}
Flight 2: {option 2}
Flight 3: {option 3}
User: I prefer flight {1/2/3}

NEW Round (no answer shown)
User: Which flight is best?
Flight 1: {option 1}
Flight 2: {option 2}
Flight 3: {option 3}

Required Output
Model: The best option is Flight

Figure 16: The prompt used for Direct Prompting and Confidence Thresholding. Logit is extracted as measure of confidence.

VOI — Prior over Feature States

You are calibrating a probabilistic user model.

Feature: {feature}
History (support + any clarifying Q&A):
{history_ctx}

Based *only* on this history, estimate $P(\text{state})$ for the feature.
Return STRICT JSON with keys exactly {{states}} **that sum to 1**. Example: `{"lower": 0.33, "higher": 0.33, "none": 0.34}`
JSON:

Figure 17: Prior Estimation for VOI (Airline Preference Matching)

VOI — Posterior with Options

You are calibrating a probabilistic user model.

Feature: {feature}
History (support + any clarifying Q&A):
{history_ctx}

Current options:
A) {option A}
B) {option B}
C) {option C}

Estimate the *posterior* distribution over the user's {feature} state given full history.
Return STRICT JSON with keys exactly {{states}} **that sum to 1**.
JSON:

Figure 18: Posterior Estimation with Options (Airline Preference Matching)

VOI — Candidate Preference Questions

You are an AI assistant helping a user choose between flight options A, B, and C. You've analyzed the support examples but still have some uncertainty.

{support_history}
{qa_context}

Generate one multiple-choice question about a single aspect of the user's preference that will help decide among the options below.

A) {option A}
B) {option B}
C) {option C}

Question:

Figure 19: VOI Candidate Questions (Airline Preference Matching)