Modeling Decision Problems for Relevant Answers to Polar Questions

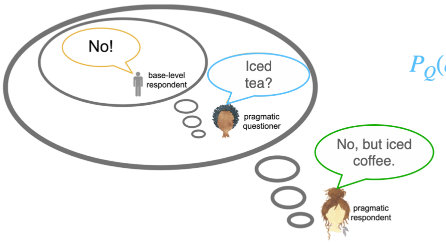
Polina Tsvilodub, Michael Franke, Robert Hawkins
polina.tsvilodub@uni-tuebingen.de

Imagine you are working as a barista at a coffeeshop. A customer asks "Do you have iced tea?" but you've run out. They have asked a yes-no (or *polar*) question, so you should respond "no", as suggested by classic accounts of questions in linguistics (Hamblin, 1973). However, this minimal answer is intuitively unsatisfying. Instead, you may prefer to say something like "No, I'm afraid we're out of iced tea *but we do have iced coffee*", mentioning a *relevant alternative* (Clark, 1979).

In previous work, we proposed a novel cognitive model of pragmatic overinformative question answering (the PRIOR-PQ model) and empirically evaluated some of its key predictions (anonymous, n.d.). We formulated our model in the tradition of the Rational Speech Act (RSA) framework (Frank & Goodman, 2012), couching it in an *action-oriented* definition of relevance hinging on the questioner's *decision problem* (DP) (van Rooy, 2003). However, one limitation of many cognitive models like RSA is the necessity to elicit auxiliary intuitive world knowledge in costly human experiments. One potential alternative to human data are predictions of SOTA large language models (LLMs). Yet to maintain the quality of the cognitive model, careful testing of LLM-supplied data in the context of the model is needed.

In this work, we test predictions of `gpt-4o-mini` for intuitive information about the DP in PRIOR-PQ. PRIOR-PQ captures a cooperative answerer that chooses an answer increasing the expected utility of the questioner's future actions under their DP. The DP is a tuple consisting of a set of world states, a set of actions, a *utility function*, and a probability distribution capturing the questioner's prior beliefs about the world states (see Fig. 1). The model is presented in formal detail in Fig. 1. We compared the predictions of PRIOR-PQ to human data in two experiments (case study 2, 3). In both, a polar question about a target appeared in a context presenting available options (but not the target) which varied in terms of their practical utility for the questioner (example vignettes are below). We elicited free production responses from humans ($N = 162$ and $N = 130$). To model the inference about the likely questioner DP in PRIOR-PQ and predict the optimal answer, we modeled four or five types of DPs, one corresponding to each of the available options (see example). Each DP was associated with different utilities, or, payoffs for each other option, given a target option. Supplied with utility ratings elicited in human experiments (slider ratings, $N = 453$ and $N = 130$), the model's predictions aligned well with human data, particularly capturing the preference for overinformative *competitor* responses mentioning only a relevant option (Fig. 2A).

Here, we explore whether DP utilities sampled from `gpt-4o-mini` given the prompt from human experiments align with human ratings. The utilities were sampled with temperature $\tau = 0.1$, given the additional instruction to produce ratings from 0 to 100 instead of a slider, for ten iterations, for each option pair. The text predictions were cast to numbers. Figure 2(B) shows predicted utilities for each item, averaged over runs, against results from the human experiments for both case studies, indicating high correlation ($R^2 = 0.92$ and 0.87). The order of preferences for different alternatives (e.g., "iced coffee" vs. "Chardonnay"), given a target ("iced tea"), corresponds to intuitions for our vignettes for both human and LLM results. These results provide a promising avenue for ongoing work in which we integrate LLM utility predictions into PRIOR-PQ simulations. Including LLMs in PRIOR-PQ, given careful comparison of LLM and human results, provides a promising avenue towards scaling up rational cognitive models.



Decision problem: $D = \langle W, A, U, \pi_Q^W \rangle$

- $W =$ {iced coffee, soda, Chardonnay}
- $A =$ {iced coffee, soda, Chardonnay, --}
- $U : W \times A \to \mathbb{R}$

Questions: Q = {w? wh?}

Responses: R = {—, yes, no, no + W, no + $w \in W$}

$$P_{R_0}(r \mid w, q) \propto \begin{cases} 1 \text{ if } r \text{ is true in } w \text{ and safe for q} \\ 0 \text{ otherwise} \end{cases}$$

[literal & safe respondent]

$$P_Q(q \mid D) \propto \text{SM}[\mathbb{E}_{w\sim\pi}\mathbb{E}_{r\sim P_{R_0}(\cdot|w,q)}(V(D^{|r,q}) - C(r)) - C(q)]$$

[question P proportional to expected usefulness of responses for updated DP]

$$V(D^{|r,q}) = \mathbb{E}_{a\sim N^D}\mathbb{E}_{w\sim\pi_Q^{W|r,q}} U(w, a)$$

[expected utility given stochastic policy and response from R_0]

$$N^D(a) \propto \text{SM}(\mathbb{E}_{w\sim\pi_Q^W} U(w, a))$$

[stochastic policy]

$$P_{R_1}(r \mid q) \propto \text{SM}(\mathbb{E}_{D\sim\pi^D_{R_1}(\cdot|q)} U(r, D))$$

[response maximizing utility for inferred DP]

$$\pi^D_{R_1}(D \mid q) \propto \pi^D_{R_1}(D) P_Q(q \mid D)$$

[inferred DP given question]

$$U(r, D) = (1-\beta)(-KL(\pi^{W|r,q} \,||\, \pi^W_{R_1})) + \beta V(D^{|r,q}) - C(r)$$

[informativity] [DP utility]

Figure 1: PRIOR-PQ model overview. The pragmatic answerer R_1 reasons about a questioner Q who selects a question according to the utility of the information for their DP that it is likely to elicit from a safe and true base respondent R_0.

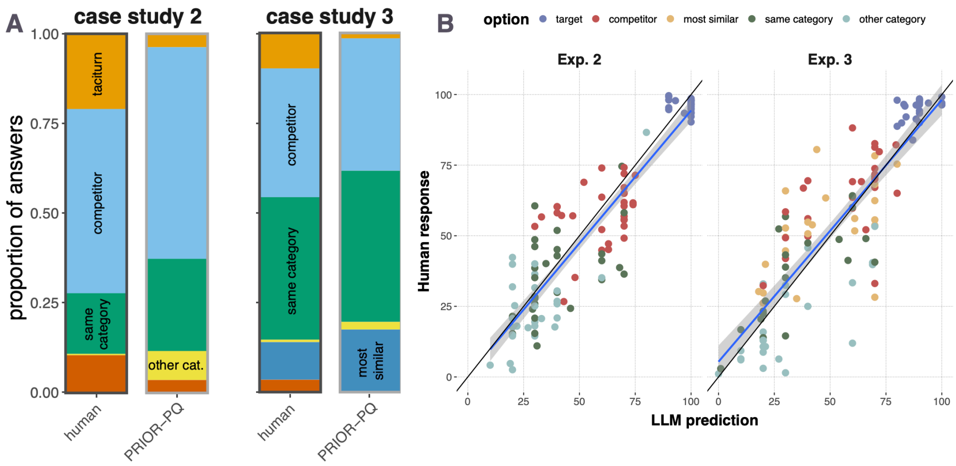


Figure 2: **A**: Proportions of responses mentioning different alternatives (color) in the two experiments, produced by humans and predicted in simulations by PRIOR-PQ. **B**: Mean by-item utilities of different options (color) when the target option (e.g., iced tea) is the goal, predicted by `gpt-4o-mini` against human ratings.

Example vignette from Exp. 2: You are a bartender in a hotel bar. The bar serves only soda (same category), iced coffee (competitor) and Chardonnay (other category). A woman walks in. She says: "Do you have iced tea?" (target) **Example vignette from Exp. 3:** Context 1: Your friend is having a sleepover with some friends on the weekend. [...] Context 2: Your roommate [...] has a large mirror that she needs to pack for transportation. Shared options and question: You have the following items at home that you could spare for some time: some bubble wrap (competitor 2), a pillow (most similar), a sleeping bag (competitor 1) and a carpet (other category). Your friend asks: "Do you have a blanket?" (competitor i was "same category" in other context)

References: Clark, H.H. (1979) Responding to indirect speech acts. *Cognitive Psychology*. Frank, M.C. & Goodman, N.D. (2012) Predicting Pragmatic Reasoning in Language Games. *Science*. Hamblin, C. (1973) Questions in Montague English. *Foundations of Language*. [Redacted for anonymity] (under review) Relevant answers to polar questions. van Rooy, R. (2003) Questioning to resolve decision problems. *Language & Philosophy*