Storytelling as Inverse Inverse Planning

Kartik Chandra
MIT CSAIL
Cambridge, MA 02139 USA

Tzu-Mao Li
UC San Diego
San Diego, CA 92093 USA

Joshua B. Tenenbaum
MIT BCS & CSAIL
Cambridge, MA 02139 USA

Jonathan Ragan-Kelley
MIT CSAIL
Cambridge, MA 02139 USA

Abstract

Great storytelling takes us on a journey the way ordinary reality rarely does. But what exactly do we mean by a "journey"? Recently, literary theorist Kukkonen (2014) proposed that storytelling is "probability design": the art of giving an audience pieces of information bit by bit, to craft the *journey of their changing beliefs* about the fictional world. A good "probability design" choreographs a delicate dance of certainty and surprise in the reader's mind as the story unfolds from beginning to end. In this paper, we computationally model this conception of storytelling. Building on the classic Bayesian inverse planning model of human social cognition, we treat storytelling as *inverse* inverse planning: the task of choosing actions to manipulate an inverse planner's inferences, and therefore a human audience's beliefs. First, we use an inverse inverse planner to depict social and physical situations, and present behavioral studies indicating that inverse inverse planning produces more expressive behavior than ordinary "naïve planning." Then, through a series of examples, we demonstrate how inverse inverse planning captures many storytelling elements from first principles: character, narrative arcs, plot twists, irony, flashbacks, and deus ex machina are all naturally encoded in the flexible language of probability design.
This paper reports on work to be presented at SIGGRAPH 2023 (Chandra, Li, Tenenbaum, & Ragan-Kelley, 2023).

Keywords: theory of mind; social cognition; Bayesian modeling; storytelling

Introduction

The famous black-and-white animation by Heider and Simmel (1944) shows two triangles and a circle moving around a 2D plane. Nearly everyone who watches this video interprets the triangles and circle as intelligent, goal-driven agents, even ascribing moral judgements to these "characters" and identifying moments of tragedy and redemption.

Even in this severely restricted visual world, storytelling comes naturally to humans. Fritz Heider wrote in his autobiography (1983) that the original 90-second video took him just a few hours to create, albeit by stop-motion animation. When provided with a modern touch-screen interface, even 10th-grade students can easily create rich, dramatic animations in just 15 minutes (Gordon & Roemmele, 2014).

In this paper, we seek to capture this ability computationally. As a first try, we might simply simulate rational behavior for the agents, as past work has done via planning algorithms (Netanyahu et al., 2021; Shu et al., 2020). But the resulting animations do not exhibit the dramatic flair we desire—even though the agents appear lifelike and often have discernible goals, the videos rarely evoke elements like suspense, surprise, and irony. This makes sense: we choose to go to the theater for a good story, even though we could just as well watch human lives unfold at the supermarket or post office.

What accounts for this gap? Our key insight is that there is a distinction between acting *in* a situation and acting *out* a situation. When acting in a situation, an agent behaves optimally to achieve a goal; when acting out a situation, the agent instead behaves optimally to manipulate an *audience's belief* about that goal. These behaviors are not always the same: a storyteller might choose a sub-optimal action if it has a stronger impact on the audience. This is apparent in stage violence (where actors move their bodies to *depict* violence without actually causing harm) and mime (where actors create the illusion of invisible objects), but permeates all storytelling as evidenced by the old adage "show, don't tell."

Nonetheless, the predominant approaches to theatrical "acting" in computer graphics and interactive fiction are all purely planning-based (Lebowitz, 1985; Martens et al., 2013; Meehan, 1977; Riedl & Young, 2010; Wampler et al., 2010; Won et al., 2021). Some ad-hoc heuristics have emerged for considering specific aspects of audience experience like suspense and believability (Bae & Young, 2008; Cheong & Young, 2006; Gerrig & Bernardo, 1994; Riedl & Young, 2004; Szilas, 2003), but a flexible "audience model" remains a challenge (Kreminski & Martens, 2022).

Here, we model the audience's experience using *Bayesian inverse planning* (Baker, Saxe, & Tenenbaum, 2009; Baker, Tenenbaum, & Saxe, 2007), thus casting storytelling as *inverse inverse planning*. We were inspired in part by literary theorist Kukkonen (2014)'s abstract vision of stories as "probability designs": sequences of observations presented to an audience to modulate their beliefs over time. We show how this general framework naturally captures a wide variety of storytelling elements (character, plot twists, irony, flashbacks, narrative arcs). Moreover, our behavioral studies suggest that it indeed produces the desired effect in human audiences—for example, by depicting character traits up to 12× more effectively than naïve planning, or "miming" a box to make it look over 5× heavier than it truly is. Sample code is available online at `https://people.csail.mit.edu/kach/a2i2p/`.

Background: Bayesian inverse planning

A long line of work has sought to model human social cognition with Bayesian inference. These models posit that when we observe an agent take an action, we infer the agent's

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

goal by applying Bayes' rule: $\mathsf{P}(\text{goal} \mid \text{action}) \propto \mathsf{P}(\text{action} \mid \text{goal})\mathsf{P}(\text{goal})$. Here, $\mathsf{P}(\text{goal})$ reflects our *prior* over goals and $\mathsf{P}(\text{action} \mid \text{goal})$ is the *likelihood:* the more optimal an action is for the agent, the likelier it should be (Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016). In short, while *planning* outputs a plan for a goal, *inverse planning* infers a goal from a plan (Baker et al., 2009, 2007; Jara-Ettinger, 2019). Inverse planning can model agents that reason about each other via "theory of mind" (Baker, Goodman, & Tenenbaum, 2008; Shum, Kleiman-Weiner, Littman, & Tenenbaum, 2019; Tauber & Steyvers, 2011; Ullman et al., 2009), agents who plan sub-optimally (Zhi-Xuan et al., 2020), human kinematic motions (Qian, Kryven, Gao, Joo, & Tenenbaum, 2021), and judgements made by young children (Pesowski, Quy, Lee, & Schachner, 2020). Recent work has considered how inverse planning can model actions humans take to communicate with (Ho, Cushman, Littman, & Austerweil, 2021) and influence (Ho, Saxe, & Cushman, 2022) each other. This is closely related to the influential Rational Speech Acts (RSA) framework (Frank & Goodman, 2012; Goodman & Frank, 2016), which models behaviors like polite speech (Yoon, Tessler, Goodman, & Frank, 2016, 2017) and behaviors in pedagogical contexts (Shafto, Goodman, & Griffiths, 2014; Yoon, MacDonald, Asaba, Gweon, & Frank, 2018) as methods of influencing pragmatic listeners. The Rational Communicative Social Actions (RCSA) framework integrates inverse planning and RSA to model communicative aspects of intimacy (Hung, Thomas, Radkani, Tenenbaum, & Saxe, 2022) and punishment (Radkani, Tenenbaum, & Saxe, 2022). Similarly, in the reinforcement learning community, "inverse reinforcement learning" methods (Arora & Doshi, 2021; Ng, Russell, et al., 2000; Ramachandran & Amir, 2007) can infer reward functions, which can then be applied to influence observers, e.g. to make a robot's motion "legible" to humans (Dragan, 2015; Dragan, Lee, & Srinivasa, 2013; Hadfield-Menell, Russell, Abbeel, & Dragan, 2016) or oppositely to strategically fool adversarial viewers about its true intentions (Pattanayak, Krishnamurthy, & Berry, 2022).

Here, we apply these ideas to theatrical "acting," where the aim of the storytellers is to collaboratively influence the audience (collaborating even if they are *depicting* competition). Unlike typical settings studied using RSA and RCSA, storytelling influences not just a single inference, but rather the trajectory of the audience's belief *over time*.

For much of this paper, we will discuss the concrete world model proposed by Ullman et al. (2009), which consists of two agents moving in a maze on a grid. In our animations, we will stylize the agents as two characters: a *robot* and an *enchanted animate cheese cube* in a kitchen. The two agents take turns moving north/south/east/west through the kitchen or staying in place. The cheese is "weak" and only succeeds in moving 60% of the time. However, the robot is "strong" and can push the cheese cube along. A *table* in the kitchen blocks the cheese's motions, but can be moved by the robot. Finally, the kitchen floor has two special tiles, *pink* and *green*.

The two characters can have a variety of natural goals in the kitchen. The cheese and the robot could each "want" to sit on either the pink or green tile. Additionally, because the robot is strong enough to move the cheese, it could want to "help" or "hinder" the cheese from reaching its goal. These dynamics can be formalized as a multi-agent Markov Decision Process (MDP): the state space $\mathcal{S}$ encodes the positions of the robot, cheese, and table. The action space $\mathcal{A}$ for each agent is $\{\leftarrow, \rightarrow, \uparrow, \downarrow, \text{stay}\}$. The transition function for each agent encodes how each action affects the state (the transition function for the cheese is stochastic because the action may fail). The reward function for each agent captures the agent's goals. The cheese and the robot each receive a fixed reward if they are on their respective goal tiles (pink or green), and pay a small cost for moving instead of staying in place. In addition, the robot receives a "social reward" based on the cheese's reward on this turn. Specifically, if the cheese earns reward r_{cheese}, then the robot earns a bonus reward $\rho_{\text{robot}} \cdot r_{\text{cheese}}$ where the robot's "alignment" $\rho_{\text{robot}} \in \{-3, -1, 0, +1, +3\}$ is positive if the robot is helping, negative if hindering, and zero if neutral. For this MDP, Ullman et al. compute optimal policies for the two agents by running value iteration Bellman (1966). They use a hierarchical softmax strategy, first computing a policy for the cheese assuming the robot moves uniformly at random, and then computing a policy for the robot assuming the cheese selects actions via the softmax of its Q-function. This allows for two recursive levels of "theory of mind" in the planner: the robot models the cheese modeling the robot.

Finally, we describe Ullman et al.'s inverse planner, which makes inferences about the agents' (hidden) goals from their (observable) actions. Let a *hypothesis* be a tuple $H = \langle G_{\text{cheese}} \in \{p, g\}, G_{\text{robot}} \in \{p, g, \emptyset\}, \rho_{\text{robot}} \in \{0, \pm 1, \pm 3\}\rangle$. For fixed H, we can use value iteration to compute $Q^H_{\text{robot}}(s,a)$ and $Q^H_{\text{cheese}}(s,a)$ for state $s \in \mathcal{S}$ and action $a \in \mathcal{A}$. Assuming the softmax-rational model above, this induces a probability distribution over each character c's actions: $\mathsf{P}(s \to a \mid H) \propto \exp\left(\beta \cdot Q^H_c(s,a)\right)$. Thus, if we observe an agent take action a from state s, we can apply Bayes' rule to update our belief about the characters' goals: $\mathsf{P}(H \mid s \to a) \propto \mathsf{P}(s \to a \mid H)\mathsf{P}(H)$.

Figure 1 shows a sample animation we generated by running the optimal policy for a helpful robot in a random scenario (i.e. in a random state, with a random H where $\rho_{\text{robot}} > 0$). As we argue in the caption, this is a poor depiction of helpfulness. The inverse planner agrees: the model is not confident that the robot is helping. In the next section, we show how inverse inverse planning can create animations that *do* effectively depict helping and other scenarios.

Inverse inverse planning

Next, we create new animations by *inverse inverse planning*—that is, by optimizing over Bayesian inference.[1] To

[1] All animations referenced and described in this section are available at https://osf.io/gyh8a/?view_only=0558bbedab964ae49e552ec3263227cf. Section labels in the text are also hyperlinked to respective individual videos.

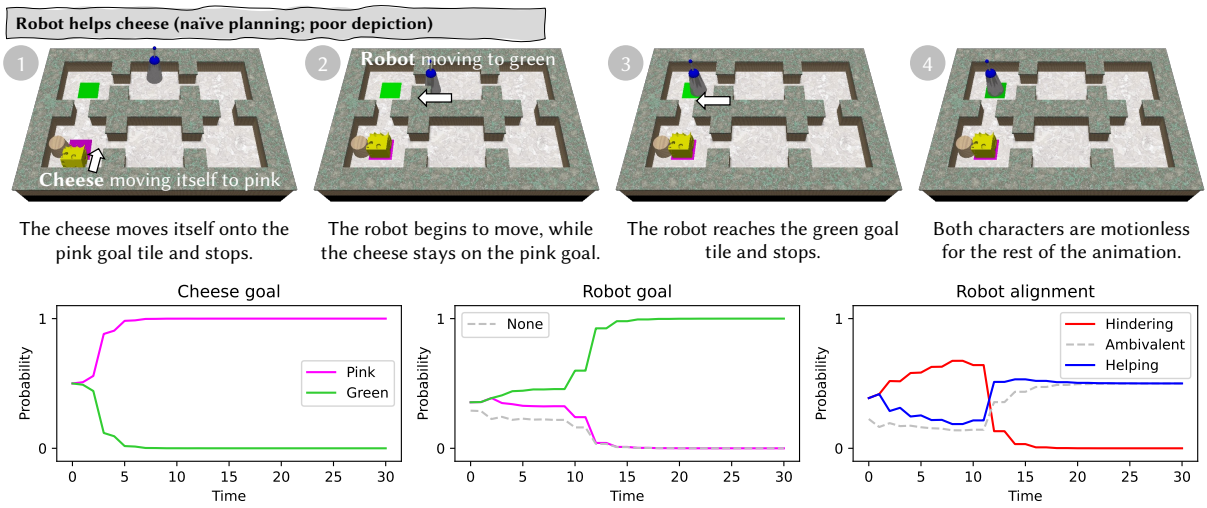


Figure 1: **(top)** Suppose we animate a robot that is helping the cheese reach its goal, by having both characters follow their optimal policies from some initial state ("naïve planning"). **This produces a poor depiction:** it is not clear that the robot wants to *help* the cheese, only that it wants to go to green. **(bottom)** The inverse planner agrees. It infers that the cheese wants pink (plot #1), and that the robot wants green (plot #1; notice bump at $t = 10$ when the robot reaches green and stays). But it remains uncertain about the robot's alignment (plot #3), because the robot's behavior is consistent with both indifference to the cheese and wanting to help the cheese (but doing nothing because the cheese is already at its goal). *See video and compare to Figure 2.*

be precise, we optimize *scripts*, which correspond to the sequence of observations presented to the audience over the course of a story. A script σ of length T is given by an initial state s_0 and a sequence of valid *transitions* $\sigma_t = \langle a_t^{\text{robot}}, a_t^{\text{cheese}}, s_t \rangle$ for $1 \leq t \leq T$. Note that s_t is not uniquely determined by s_{t-1} and $\langle a_t^{\text{robot}}, a_t^{\text{cheese}} \rangle$ because state transitions may be non-deterministic. For example, the optimizer should be able to choose not only that the cheese attempts to move, but also whether the move succeeds or fails.

Suppose, like before, that we wanted an animation that depicted the robot as helping the cheese. We can express this task in a simple objective function over scripts:

$$f_{\text{help}}(\sigma) = \sum_{1 \leq t \leq T} \mathsf{P}(\rho_{\text{robot}} > 0 \mid \sigma_{1:t})$$

The objective f_{help} is maximized for scripts where at every time t, based on observing the animation up to time t (i.e. $\sigma_{1:t}$) a viewer has a strong belief that the robot is helping (i.e. $\rho_{\text{robot}} > 0$). Notice that we do not have to specify whether we want $\rho_{\text{help}} = +1$ (slightly helping) or $+3$ (strongly helping), only that it is greater than zero. Notice also that f_{help} does not say anything about G_{cheese}, G_{robot}, or even the initial positions of the characters in s_0, all of which are to be optimized automatically. In this way, f_{help} abstractly captures the essence of the storytelling goal.

To optimize f_{help}, we use beam search, where the search heuristic is the objective applied to the current script "prefix." When we run the optimizer with $T = 15$, we get a rendered animation within just a couple of minutes (Figure 2; video). The cheese moves to pink; the robot pushes it along and then steps back. Upon watching this animation, a rational viewer would infer that the cheese wanted pink (because of its initial motion towards pink), and that the robot wanted to help the cheese (because it pushed the cheese to pink and stepped back afterwards). This is a more effective depiction than the one we generated earlier by "naïve planning" (Figure 1; video).

Similarly, we can ask for an animation of a *hindering* robot. In the generated animation, the cheese first moves to green. Then the robot pushes it into a corner and blocks the way to green. In comparison, with naïve planning the robot moves to green, blocking the cheese. It is unclear whether the robot is intentionally hindering, or indifferent to the cheese and itself wanting green. Inverse inverse planning avoids this ambiguity because the robot never moves onto green (video).

Behavioral experiment 1

We asked whether inverse inverse planning better depicts the relationship between the two characters (i.e. helping, hindering, or indifferent). We generated 20 animations each of helping, hindering, and indifference (using random seeds 0-19), via both inverse inverse planning and naïve planning, for a total of $20 \times 3 \times 2 = 120$ animations. We recruited 98 online participants, showed each participant a random shuffled subset of 15 of these animations, and for each animation asked them to report whether the robot was helping the cheese, hindering it, indifferent to it, or whether the animation was unclear. For each animation, we measured the proportion of responses matching the desired depiction target.

Figure 3 shows the results of this experiment. When depicting helping, the average inverse inverse planning animation caused 73% of viewers to report "helping," while the average naïve planning animation caused only 6% to ($p < 0.01$ by two-tailed binomial test). When depicting hindering, inverse inverse planning was also significantly better

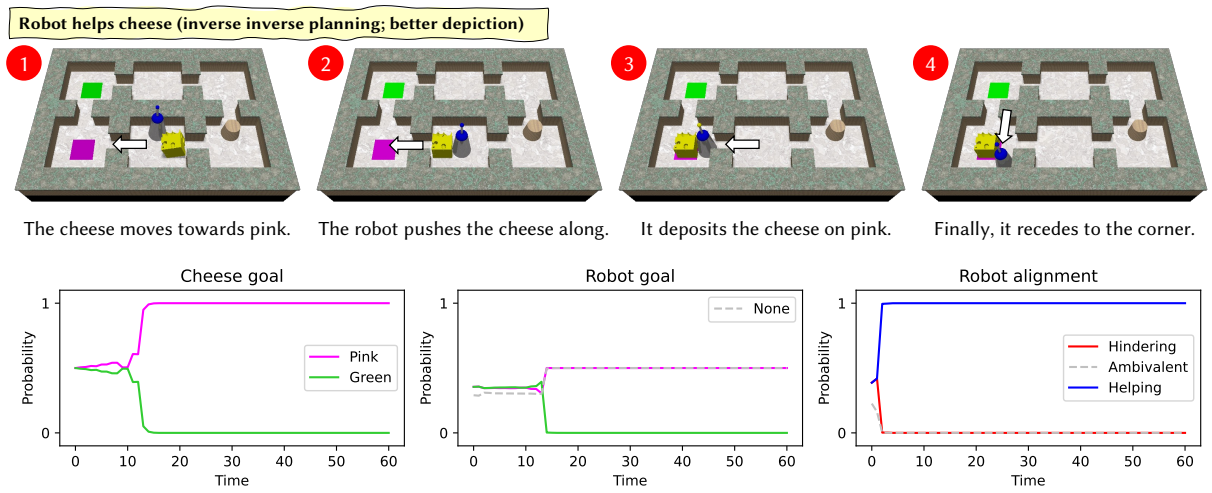


Figure 2: **(top)** Using inverse inverse planning, we can optimize an animation (including the initial state, or "setting") to maximize the inverse planner's belief that the robot is helpful. This finds a much more effective depiction. **(bottom)** Now, it is clear that the robot is helping, and indeed the model is confident of that throughout the entire video (plot #3). *See video.*

than naïve planning (63% vs. 29%, $p < 0.01$). Both methods were equivalently effective at depicting indifference (73% vs. 75%, n.s.). In summary, we find that inverse inverse planning better depicts the robot's relationship with the cheese, particularly in cases that are challenging for naïve planning.

The differences across conditions reveal that depiction tasks vary in difficulty for naïve planning. Not all situations allow for actively helping (e.g. if the cheese is already close to its goal), whereas most circumstances allow for actively hindering (but might require the robot to go irrationally out of its way to hinder). Indifferent characters rarely interact, and thus easily telegraph their indifference to audiences. Hence, naïve planning performs worst when helping, a little better when hindering, and best when indifferent. In comparison, inverse inverse planning does well across all conditions.

Miming by "inverse inverse physics"

Next, we consider a more naturalistic physics-based setting. We were inspired by the short animated film *Sisyphus* (Jankovics, 1974), which depicts Sisyphus from Greek mythology pushing a heavy boulder up a hill. The animation is striking in how dramatically it conveys the boulder's weight. We wondered if inverse inverse planning could evoke that effect: that is, make a character "mime" a heavy object.

To model this scenario, we created a physics-based environment consisting of a "Luxo lamp"-style hopping robot (Witkin & Kass, 1988) attached to a box on a hill (Figure 4a). For *planning*, we built a differentiable physics simulator for this environment and used it to train a controller to pull the box up the hill using the Short-Horizon Actor-Critic algorithm (Xu et al., 2022). Actor-critic algorithms jointly train two neural networks: a policy $\pi(s;\theta)$ that computes actuations for the hopper at state s, and a value function $V(s;\phi)$ that computes the optimal-long term reward attainable from s. We optimized controllers for two box weights, light (0.1) and heavy (0.5) to obtain $\theta_{\{0.1,0.5\}}$ and $\phi_{\{0.1,0.5\}}$. Next, we used *inverse planning* to model a viewer's impression of the box weight: following Battaglia, Hamrick, and Tenenbaum (2013)'s Bayesian model of intuitive physics, we used hypothetical simulations of the physical system in the light and heavy conditions to infer the conditions of the observed trajectory. Finally, we used *inverse inverse planning* to animate the hopper "miming" a heavy box: we optimized a trajectory that maximizes the inverse planner's confidence that the box is heavy (even though the box was actually light in the simulator), thus fooling the model into believing the illusion (Chandra, Li, Tenenbaum, & Ragan-Kelley, 2022). The hopper then pretends to struggle as it pulls the box (video).

Behavioral experiment 2

We asked whether the inverse inverse planning hopper is indeed a convincing "mime." We recruited 35 online participants and showed them each a series of 12 pairs of animations. Each animation was randomly either an "honest" hopper with a heavy or light box, or a "mime" with a light box pretending that it is heavy. Participants selected which animation contained a heavier box. They were not told that hoppers could mime. Each video had a different-colored box to emphasize that their weights may vary.

Figure 4b shows the results of this experiment. As expected, participants were at chance (50%) when the animations had the same condition, and between heavy and light boxes they selected the heavy box 97% of the time ($p < 0.01$ by two-tailed binomial test). The mime convinced 95.7% of viewers that its box was heavier than the light box, despite being of the same (light) weight ($p < 0.01$). Furthermore, it convinced 68.6% of viewers that it was heavier than the *heavy* box despite being $5\times$ lighter ($p < 0.01$). We conclude that the mime successfully convinces viewers that the box is heavier than it truly is.

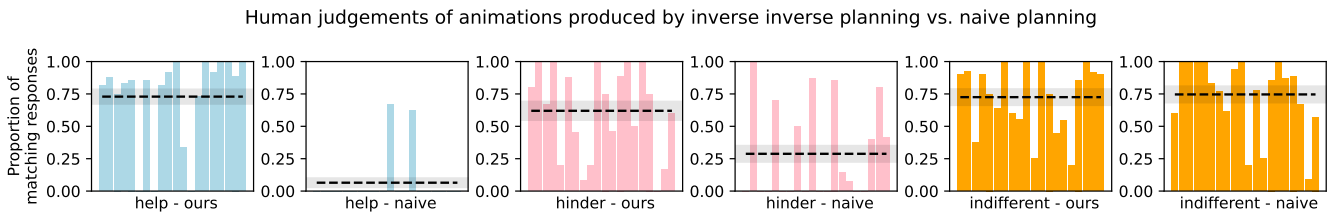
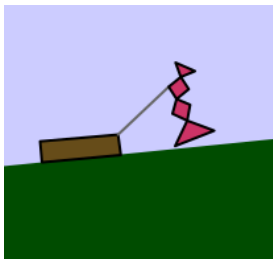
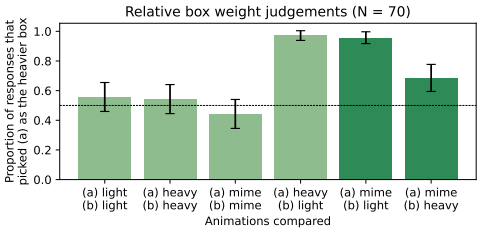


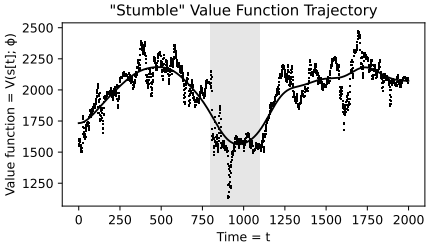
Figure 3: We compare inverse inverse planning and naïve planning on three depiction tasks: showing the robot to be helping, hindering, or indifferent. Participants viewed the resulting animations and reported their impression of the robot (help / hinder / indifferent / unclear). Each bar represents a separate animation (from a different random seed), showing the proportion of participants who reported the desired response for that animation. Horizontal dashed lines are averages across animations for each condition (higher is better). Inverse inverse planning is significantly more effective than naïve planning for depicting helping and hindering ($p < 0.01$ for both), and equally effective for indifference. Shaded regions span 95% confidence intervals.



(a) We designed a hopper that pulls a light box up a hill and made it "mime" a heavy box. See section on "Miming," and the video.



(b) When shown pairs of animations, viewers perceive the "miming" hopper as pulling the heavier box, even though it is pulling a light box (bar #5), and even when the *other* box is *actually* heavier (bar #6). Error bars show 95% confidence intervals.



(c) We can force the hopper's value function to "dip" from t_s to t_f (the shaded period), causing it to "stumble" and then recover on cue. See section on "Stumbling" below, and the video.

Figure 4: Inverse inverse planning in a physics-based setting.

Flexibility of inverse inverse planning (all videos)

Finally, returning to the multi-agent grid-world, we give a variety of examples of encoding classic storytelling elements as inverse inverse planning. We encourage readers to imagine how they would depict these scenes themselves before looking at our system's outputs. Note that these are not cherry-picked: all examples presented here are created with the same random seed (0); simple variations emerge with other seeds.

Plot twists (video) We first consider the "plot twist," a storytelling device where an unexpected event radically alters the audience's expectations. For example, a classic plot twist reveals that a seemingly friendly character was adversarial all along. Here, we ask for an animation where the robot *appears* to be helpful at first, but at $t = T/2$ is revealed to be hindering.

$$f_{\text{twist}}(\sigma) = \sum_t \begin{cases} \mathsf{P}(\rho_{\text{robot}} > 0 \mid \sigma_{1:t}) & \text{if } t \le T/2 \\ \mathsf{P}(\rho_{\text{robot}} < 0 \mid \sigma_{1:t}) & \text{if } t > T/2 \end{cases}$$

In the generated animation, the robot "helpfully" pushes the cheese to pink. However, upon reaching pink it continues pushing, trapping the cheese along the wall. This surprising action reveals that the robot's true intention was to hinder all along. We can also ask for the reverse, a video where the robot appears to be hindering but was helping all along. In the generated animation, the cheese moves to pink and the robot approaches as if to push it off (hindering). However, the cheese continues moving, revealing that it wanted green all along. The robot helpfully pushes it there.

Irony (video) Next, we consider *dramatic irony*, which occurs when the audience has a different understanding of a situation than the characters in that situation. Here, we design an objective function for scenes where the robot appears to be trying to help, but mistakenly hinders because of its false belief about the cheese's goal. We use conditional probability to express that the robot should appear to be helpful *if* the cheese had a different goal.

$$\begin{aligned} f_{\text{irony}}(\sigma) = \sum_t &+ \mathsf{P}(G_{\text{cheese}} = \text{green} \mid \sigma_{1:t}) \\ &+ \mathsf{P}(\rho_{\text{robot}} < 0 \mid \sigma_{1:t}, G_{\text{cheese}} = \text{green}) \\ &+ \mathsf{P}(\rho_{\text{robot}} > 0 \mid \sigma_{1:t}, G_{\text{cheese}} = \text{pink}) \end{aligned}$$

In the generated animation, the cheese moves to green, but the robot pushes it off and towards pink. When the cheese tries to move back, the robot "helpfully" guides it back to pink.

Flashbacks (video) Nonlinear discourse is a storytelling technique where information in a story is revealed out of chronological order. For example, a "flashback" can recontextualize a scene, giving it heightened significance or new meaning. Imagine we saw a glimpse of the robot pushing the cheese east, away from the pink and green goals. Can we show a flashback that casts this action as helping? Let $c(\sigma)$ be the script σ with a single transition appended, in which the robot pushes the cheese east. We can now apply the objective function for "helping" over $c(\sigma)$: $f_{\text{flashback-help}}(\sigma) = f_{\text{help}}(c(\sigma))$. In the generated flashback, the cheese is trying to go all the way around the room to pink because the table blocks a door along the shortest path. This casts the "push" as helpful.

If we instead substitute f_{hinder}, we get a flashback that casts the action as hindering. The cheese tries to move directly west to pink, casting the robot's eastward push as hindering.

Narrative arc (video) American writer Kurt Vonnegut's master's thesis (famously unpublished, but see this lecture) argues that all stories have simple "shapes" defined by the trajectory of the protagonist's fortunes over time (Kiley et al., 2016; Vonnegut, 2005). Reagan et al. (2016) analyze novels to extract story shapes, and suggest that future work investigate the "opposite direction" of generating stories for given shapes. We can do this by inverse inverse planning.

We would like to optimize for animations where the robot's fortunes decline and then rise again, creating a "story arc." To heighten the effect, we add a mechanism for characters' fortunes to change based on external events. Since ancient times, storytellers have propelled or resolved plots by introducing a new element from outside the world of the story, a pattern literary theorists call "deus ex machina." We create the possibility for "deus ex machina" by creating a special type of transition $\langle \text{deus}, x, y \rangle$ where the obstructing table "falls from the sky" into the kitchen at position (x, y). Note that the characters' learned policies do not account for the possibility of this transition occurring; nor do audiences know to expect it—it is a surprise from "outside" the fictional world.

With this enhancement, we can search for stories where the value function of the robot (learned by value iteration) shows a rise-fall-rise pattern, which we model as 1.5 periods of a sinusoid:

$$f_{\text{arc}}(\sigma) = \sum_t + \sin(t/T \cdot 3\pi) \cdot \mathbf{E}\left[V^H_{\text{robot}}(s_t) \mid \sigma_{1:t}\right] - 0.1 \cdot D_{\text{KL}}(H_{1:t-1} \parallel H_{1:t})$$

We do not specify anything else about the story. However, we introduce a new term to enforce that the characters' apparent goals do not change over time. Otherwise, we might get stories where the robot's apparent fortune changes because its apparent goal changes. To enforce this consistency, we minimize the KL-divergence between our beliefs about the characters' goals before and after each observed action. Here, $H_{1:t}$ is a random variable with probability distribution $\mathsf{P}(H \mid \sigma_{1:t})$.

In the generated animation, the robot starts helping the cheese to pink. However, the table falls onto pink at the last moment. Then the robot moves the table out of the way, allowing the cheese to finally reach pink.

Stumbling (video) We can also show "narrative arcs" in the physics-based domain. We use gradient descent to optimize a trajectory in which value function $V(s_t; \phi_{0.5})$ dips from time t_s to time t_f:

$$f_{\text{arc}}(s) = \sum_t V(s_t; \phi_{0.5}) \cdot \begin{cases} -1 & t_s \le t \le t_f \\ +1 & \text{else} \end{cases}$$

As before, we optimize residuals over the optimal policy $\theta_{0.5}$. The resulting animation shows the hopper "stumble" at time t_s (i.e. reaching a physically precarious state with low value function), and then "recover" at time t_f (Figure 4c).

Discussion

In this paper, we modeled storytellers as *inverse inverse planners*, who choose actions to influence the audience's beliefs over time. We implemented inverse inverse planning computationally by optimizing over Bayesian models of audiences. We then presented behavioral studies validating that inverse inverse planning depicts social and physical characteristics more effectively than naïve planning. Finally, we showed how a variety of sophisticated storytelling elements emerge naturally from inverse inverse planning, thus giving a formal computational account of Kukkonen's theory of "probability designs."

For now, computational models of inverse inverse planning are limited primarily by scalability. The difficulty grows with the complexity of the state space, hypothesis space, and space of possible scripts. In this paper we worked with small proof-of-concept domains and simple algorithms at each recursive level, but more sophisticated algorithms could likely scale this framework significantly. They would also enable other forms of interaction, such as human-computer improvisational theater. Pinhanez (1999) considers "building computer-actors able to improvise with human actors in a performance to be an extremely difficult goal that will require the solving of many Al issues related to contextual and common sense reasoning." While we were able to prototype such a system in the grid-world domain using inverse inverse planning, it is for now too slow to use for real-time behavioral experiments.

Another promising future direction is to augment the audience model to reason about emotion (Houlihan, Ong, Cusimano, & Saxe, 2022; Ong, Soh, Zaki, & Goodman, 2019; Ong, Zaki, & Goodman, 2015, 2019; Saxe & Houlihan, 2017): either to evoke certain emotions in the audience, or to use characters' visible emotions as degrees of freedom in story design. For example, if a villain captures a hero, showing the hero's sidekick to be happy would cause the audience to infer that the sidekick was secretly aligned with the villain all along. This in turn could evoke anger at the betrayal. We hope to explore these extensions, and more, in future work.

Acknowledgements

We thank the reviewers for their feedback, Google Cloud and the MIT subMIT team for offering generous computational resources, and Sam Cheyette, Karin Kukkonen, Sydney Levine, Alena Rote, Gabriella Safran, Rebecca Saxe, Tianmin Shu, Max Siegel, Andy Spielberg, and Tan Zhi-Xuan for thoughtful discussions as we developed these ideas. This research was funded by NSF grants #CCF-1231216, #CCF-1723445 and #2238839, ONR grant #00010803, and supported by the Hertz Foundation, the Paul and Daisy Soros Fellowship, and an NSF Graduate Research Fellowship under grant #1745302.

References

Arora, S., & Doshi, P. (2021). A survey of inverse reinforcement learning: Challenges, methods and progress. *Artificial Intelligence*, *297*, 103500. Retrieved from https://arxiv.org/pdf/1806.06877

Bae, B.-C., & Young, R. M. (2008). A use of flashback and foreshadowing for surprise arousal in narrative using a plan-based approach. In *Joint international conference on interactive digital storytelling* (pp. 156–167). Retrieved from https://www.academia.edu/download/30704103/icids1.pdf

Baker, C. L., Goodman, N. D., & Tenenbaum, J. B. (2008). Theory-based social goal inference. In *Proceedings of the thirtieth annual conference of the cognitive science society* (pp. 1447–1452). Retrieved from http://cocolab.stanford.edu/papers/BakerEtAl2008-Cogsci.pdf

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329–349. Retrieved from https://www.sciencedirect.com/science/article/pii/S0010027709001607

Baker, C. L., Tenenbaum, J. B., & Saxe, R. R. (2007). Goal inference as inverse planning. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 29). Retrieved from https://escholarship.org/content/qt5v06n97q/qt5v06n97q.pdf

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–18332. Retrieved from https://www.pnas.org/doi/full/10.1073/pnas.1306572110

Bellman, R. (1966). Dynamic programming. *Science*, *153*(3731), 34–37.

Chandra, K., Li, T.-M., Tenenbaum, J., & Ragan-Kelley, J. (2022, aug). Designing perceptual puzzles by differentiating probabilistic programs. In *Special interest group on computer graphics and interactive techniques conference proceedings (siggraph '22 conference proceedings)*. Retrieved from https://arxiv.org/abs/2204.12301 doi: 10.1145/3528233.3530715

Chandra, K., Li, T.-M., Tenenbaum, J., & Ragan-Kelley, J. (2023, August). Acting as inverse inverse planning. In *Special interest group on computer graphics and interactive techniques conference proceedings (SIGGRAPH '23 conference proceedings)*. doi: 10.1145/3588432.3591510

Cheong, Y.-G., & Young, R. M. (2006). A computational model of narrative generation for suspense. In *Aaai* (pp. 1906–1907). Retrieved from https://www.aaai.org/Papers/Workshops/2006/WS-06-04/WS06-04-003.pdf

Dragan, A. D. (2015). *Legible robot motion planning*. Unpublished doctoral dissertation, Carnegie Mellon University. Retrieved from https://www.ri.cmu.edu/pub_files/2015/6/A_Dragan_Robotics_2015.pdf

Dragan, A. D., Lee, K. C., & Srinivasa, S. S. (2013). Legibility and predictability of robot motion. In *2013 8th acm/ieee international conference on human-robot interaction (hri)* (pp. 301–308). Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6483603

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998. Retrieved from https://www.science.org/doi/full/10.1126/science.1218633

Gerrig, R. J., & Bernardo, A. B. (1994). Readers as problem-solvers in the experience of suspense. *Poetics*, *22*(6), 459–472. Retrieved from https://www.sciencedirect.com/science/article/pii/0304422X94900213/pdf?md5=4e88a66f151c18e279f030fa56cba285&pid=1-s2.0-0304422X94900213-main.pdf

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in cognitive sciences*, *20*(11), 818–829. Retrieved from http://cocolab.stanford.edu/papers/GoodmanFrank2016-TICS.pdf

Gordon, A. S., & Roemmele, M. (2014). An authoring tool for movies in the style of Heider and Simmel. In *International conference on interactive digital storytelling* (pp. 49–60). Retrieved from https://roemmele.github.io/publications/ICIDS14.PDF

Hadfield-Menell, D., Russell, S. J., Abbeel, P., & Dragan, A. (2016). Cooperative inverse reinforcement learning. *Advances in neural information processing systems*, *29*. Retrieved from https://proceedings.neurips.cc/paper/2016/file/c3395dd46c34fa7fd8d729d8cf88b7a8-Paper.pdf

Heider, F. (1983). *The life of a psychologist: An autobiography*. University Press of Kansas.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American journal of psychology*, *57*(2), 243–259. Retrieved from https://www.jstor.org/stable/pdf/1416950.pdf

Ho, M. K., Cushman, F., Littman, M. L., & Austerweil, J. L. (2021). Communication in action: Planning and interpreting communicative demonstrations. *Journal of Experimental Psychology: General*. Retrieved from https://psyarxiv.com/a8sxk/

Ho, M. K., Saxe, R., & Cushman, F. (2022). Planning with theory of mind. *Trends in Cognitive Sciences*. Retrieved from https://saxelab.mit.edu/sites/default/files/publications/HoSaxeCushman2022.pdf

Houlihan, S. D., Ong, D., Cusimano, M., & Saxe, R. (2022). Reasoning about the antecedents of emotions: Bayesian causal inference over an intuitive theory of mind. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44). Retrieved from https://escholarship.org/content/qt7sn3w3n2/qt7sn3w3n2.pdf

Hung, M. S., Thomas, A. J., Radkani, S., Tenenbaum, J., & Saxe, R. (2022). Modeling risky food sharing as rational communication about relationships. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44). Retrieved from https://escholarship.org/content/qt1066w77t/qt1066w77t.pdf

Jankovics, M. (1974). *Sisyphus.* Retrieved from https://www.youtube.com/watch?v=vyZK8rkeqPM

Jara-Ettinger, J. (2019). Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences*, *29*, 105–110. Retrieved from https://compdevlab.yale.edu/docs/2019/ToM_as_IRL_2019.pdf

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in cognitive sciences*, *20*(8), 589–604. Retrieved from https://www.sciencedirect.com/science/article/pii/S1364661316300535

Kiley, D. P., Reagan, A. J., Mitchell, L., Danforth, C. M., & Dodds, P. S. (2016). Game story space of professional sports: Australian rules football. *Physical Review E*, *93*(5), 052314. Retrieved from https://cdanfort.w3.uvm.edu/research/2016-kiley-pre.pdf

Kreminski, M., & Martens, C. (2022). Unmet creativity support needs in computationally supported creative writing. In *Proceedings of the first workshop on intelligent and interactive writing assistants (in2writing 2022)*. Retrieved from https://aclanthology.org/2022.in2writing-1.11 doi: 10.18653/v1/2022.in2writing-1.11

Kukkonen, K. (2014). Bayesian narrative: Probability, plot and the shape of the fictional world. *Anglia*, *132*(4), 720–739. Retrieved from https://www.duo.uio.no/bitstream/handle/10852/54629/1/01%2BKukkonen%2BBayesian%2BNarrative.pdf

Lebowitz, M. (1985). Story-telling as planning and learning. *Poetics*, *14*(6), 483–502. Retrieved from https://academiccommons.columbia.edu/doi/10.7916/D8K362MH/download

Martens, C., Bosser, A.-G., Ferreira, J. F., & Cavazza, M. (2013). Linear logic programming for narrative generation. In *International conference on logic programming and nonmonotonic reasoning* (pp. 427–432).

Meehan, J. R. (1977). Tale-spin, an interactive program that writes stories. In *Ijcai* (Vol. 77, pp. 91–98). Retrieved from https://www.ijcai.org/Proceedings/77-1/Papers/013.pdf

Netanyahu, A., Shu, T., Katz, B., Barbu, A., & Tenenbaum, J. B. (2021). PHASE: Physically-grounded abstract social events for machine social perception. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 35, pp. 845–853). Retrieved from https://www.tshu.io/PHASE/PHASE.pdf

Ng, A. Y., Russell, S., et al. (2000). Algorithms for inverse reinforcement learning. In *Icml* (Vol. 1, p. 2). Retrieved from http://www.datascienceassn.org/sites/default/files/Algorithms%20for%20Inverse%20Reinforcement%20Learning.pdf

Ong, D. C., Soh, H., Zaki, J., & Goodman, N. D. (2019). Applying probabilistic programming to affective computing. *IEEE Transactions on Affective Computing*, *12*(2), 306–317. Retrieved from https://arxiv.org/pdf/1903.06445.pdf

Ong, D. C., Zaki, J., & Goodman, N. D. (2015). Affective cognition: Exploring lay theories of emotion. *Cognition*, *143*, 141–162. Retrieved from https://www.sciencedirect.com/sdfe/reader/pii/S0010027715300196/pdf

Ong, D. C., Zaki, J., & Goodman, N. D. (2019). Computational models of emotion inference in theory of mind: A review and roadmap. *Topics in cognitive science*, *11*(2), 338–357. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1111/tops.12371

Pattanayak, K., Krishnamurthy, V., & Berry, C. (2022). Inverse-inverse reinforcement learning. how to hide strategy from an adversarial inverse reinforcement learner. *arXiv preprint arXiv:2205.10802*. Retrieved from https://arxiv.org/pdf/2205.10802

Pesowski, M. L., Quy, A. D., Lee, M., & Schachner, A. (2020). Children use inverse planning to detect social transmission in design of artifacts. In *Proceedings of the annual conference of the cognitive science society.* Retrieved from https://cognitivesciencesociety.org/cogsci20/papers/0150/0150.pdf

Pinhanez, C. S. (1999). *Representation and recognition of action in interactive spaces*. Unpublished doctoral dissertation, Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/handle/1721.1/62342

Qian, Y., Kryven, M., Gao, T., Joo, H., & Tenenbaum,

J. (2021). Modeling human intention inference in continuous 3d domains by inverse planning and body kinematics. *arXiv e-prints*, arXiv–2112. Retrieved from https://social-intelligence-human-ai.github.io/docs/camready_12.pdf

Radkani, S., Tenenbaum, J., & Saxe, R. (2022). Modeling punishment as a rational communicative social action. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44). Retrieved from https://escholarship.org/content/qt47g8d89h/qt47g8d89h.pdf

Ramachandran, D., & Amir, E. (2007). Bayesian inverse reinforcement learning. In *Ijcai* (Vol. 7, pp. 2586–2591). Retrieved from https://www.aaai.org/Papers/IJCAI/2007/IJCAI07-416.pdf

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, *5*(1), 1–12. Retrieved from https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-016-0093-1

Riedl, M. O., & Young, R. M. (2004). An intent-driven planner for multi-agent story generation. In *Autonomous agents and multiagent systems, international joint conference on* (Vol. 2, pp. 186–193). Retrieved from https://www.academia.edu/download/30704104/025_riedlm_story.pdf

Riedl, M. O., & Young, R. M. (2010). Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, *39*, 217–268. Retrieved from https://www.jair.org/index.php/jair/article/download/10669/25501

Saxe, R., & Houlihan, S. D. (2017). Formalizing emotion concepts within a bayesian model of theory of mind. *Current opinion in Psychology*, *17*, 15–21. Retrieved from https://daeh.info/assets/pubs/saxe2017cop.pdf

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive psychology*, *71*, 55–89. Retrieved from https://www.sciencedirect.com/science/article/pii/S0010028514000024

Shu, T., Kryven, M., Ullman, T. D., & Tenenbaum, J. (2020). Adventures in Flatland: Perceiving social interactions under physical dynamics. In *Cogsci.* Retrieved from https://www.tshu.io/HeiderSimmel/CogSci20/Flatland_CogSci20.pdf

Shum, M., Kleiman-Weiner, M., Littman, M. L., & Tenenbaum, J. B. (2019). Theory of minds: Understanding behavior in groups through inverse planning. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 33, pp. 6163–6170). Retrieved from https://ojs.aaai.org/index.php/AAAI/article/view/4574/4452

Szilas, N. (2003). Idtension: a narrative engine for interactive drama. In *Proceedings of the technologies for interactive digital storytelling and entertainment (tidse) conference.*

Tauber, S., & Steyvers, M. (2011). Using inverse planning and theory of mind for social goal inference. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 33). Retrieved from https://cogsci.mindmodeling.org/2011/papers/0585/paper0585.pdf

Ullman, T., Baker, C., Macindoe, O., Evans, O., Goodman, N., & Tenenbaum, J. (2009). Help or hinder: Bayesian models of social goal inference. *Advances in neural information processing systems*, *22*. Retrieved from https://www.tomerullman.org/papers/nips2010.pdf

Vonnegut, K. (2005). *A man without a country*. Seven Stories Press.

Wampler, K., Andersen, E., Herbst, E., Lee, Y., & Popović, Z. (2010, jul). Character animation in two-player adversarial games. *ACM Trans. Graph.*, *29*(3). Retrieved from https://doi.org/10.1145/1805964.1805970 doi: 10.1145/1805964.1805970

Witkin, A., & Kass, M. (1988). Spacetime constraints. *ACM Siggraph Computer Graphics*, *22*(4), 159–168. Retrieved from https://dl.acm.org/doi/pdf/10.1145/378456.378507

Won, J., Gopinath, D., & Hodgins, J. (2021, jul). Control strategies for physically simulated characters performing two-player competitive sports. *ACM Trans. Graph.*, *40*(4). Retrieved from https://doi.org/10.1145/3450626.3459761 doi: 10.1145/3450626.3459761

Xu, J., Makoviychuk, V., Narang, Y., Ramos, F., Matusik, W., Garg, A., & Macklin, M. (2022). Accelerated policy learning with parallel differentiable simulation. *ICLR*. Retrieved from https://arxiv.org/pdf/2204.07137.pdf

Yoon, E. J., MacDonald, K., Asaba, M., Gweon, H., & Frank, M. C. (2018). Balancing informational and social goals in active learning. In *Cogsci.* Retrieved from https://sll.stanford.edu/docs/2018_cogsci/Yoon_et_al_2018_cogsci.pdf

Yoon, E. J., Tessler, M. H., Goodman, N. D., & Frank, M. C. (2016). Talking with tact: Polite language as a balance between kindness and informativity. In *Proceedings of the 38th annual conference of the cognitive science society* (pp. 2771–2776). Retrieved from http://socsci-dev.ss.uci.edu/~lpearl/courses/readings/YoonEtAl2016_Politeness.pdf

Yoon, E. J., Tessler, M. H., Goodman, N. D., & Frank, M. C. (2017). "i won't lie, it wasn't amazing": Modeling polite indirect speech. In *Cogsci.* Retrieved from https://cogsci.mindmodeling.org/2017/papers/0679/paper0679.pdf

Zhi-Xuan, T., Mann, J., Silver, T., Tenenbaum, J., & Mans-

inghka, V. (2020). Online bayesian goal inference for boundedly rational planning agents. *Advances in Neural Information Processing Systems*, *33*. Retrieved from https://arxiv.org/pdf/2006.07532.pdf