Learning in social context

Erica J. Yoon*, **Kyle MacDonald***, **Mika Asaba**, **Hyowon Gweon**, and **Michael C. Frank**
{ejyoon, kylem4, masaba, hyo, mcfrank} @stanford.edu
Department of Psychology, Stanford University
*These authors contributed equally to this work.

Abstract

...

Keywords: Learning; social context; information gain; OED; self-presentation; goal tradeoff

Introduction

Learning takes place in social contexts.

A simple case study of situations where people have to decide between information gain and self-presentation,

Computational model

A key assumption underlying inferences in recent formal models of social cognition is that people act approximately optimally given a utility function (e.g. Goodman & Frank, 2016; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016). Our model operates under the same principle of an approximately optimal agent, who reasons about the utility function that represents a combination of multiple goals. In Yoon, Tessler, Goodman, & Frank (2017), the utility function comprised a weighted combination of multiple utilities (goals) considered by the speaker, reflecting a principled tradeoff between different communicative goals (e.g. to be informative, to make the listener feel good, and to appear to be a helpful speaker). We use a similarly structured utility function that reflects different goals that a learner has in a social learning context. Specifically, we model how a person may make a decision to act based on his desire to learn how a toy works (*learning utility*), to make the toy operate and perform a given function (*performance utility*), or to present himself as a competent individual who knows how to make the toy work (*presentational utility*; see the model diagram in Figure 1).

First, the *learning utility* symbolizes the goal to learn new information, which in our paradigm specifically is associated with figuring out how a given toy works. The learning utility is formally represented by an OED model (Lindley, 1956; "Optimal Experiment Design"; Nelson, 2005), which quantifies the *expected utility* of different information seeking actions. Here we follow the mathematical details of the OED approach as outlined in Coenen, Nelson, & Gureckis (2017) that was implemented in our model. The set of queries, each realized through taking an action, is defined as $Q_1, Q_2, ..., Q_n = Q$. The expected utility of each query ($EU(Q)$) is a function of two factors: (1) the probability of obtaining a specific answer $P(a)$ weighted by (2) the usefulness of that answer for achieving the learning goal $U(a)$.

$$EU(Q) = \sum_{a \in q} P(a)U(a)$$

There are a variety of ways to define the usefulness function to score each answer (for a detailed analysis of different approaches, see Nelson (2005)). One standard method is to use *information gain*, which is defined as the change in the learner's overall uncertainty (difference in entropy) before and after receiving an answer. This information gain is then the usefulness of the answer to the query, and thus is equal to the learning utility:

$$U_{learning} = U(a) = ent(H) - ent(H|a)$$

where $ent(H)$ is defined using Shannon entropy[1]. MacKay (2003), which provides a measure of the overall amount of uncertainty in the learner's beliefs about the candidate hypotheses.

$$ent(H) = -\sum_{a \in A} P(h) log_2 P(h)$$

The conditional entropy computation is the same, but takes into account the change in the learner's beliefs after seeing an answer.

$$ent(H|a) = -\sum_{h \in H} P(h|a) log P(h|a)$$

To calculate the change in the learner's belief in a hypothesis $P(h|a)$, we use Bayes rule.

$$P(h|a) = \frac{P(h)P(a|h)}{P(a)}$$

The learner performs the expected utility computation for each query in the set of possible queries and picks the one that maximizes utility. In practice, the learner considers each possible answer, scores the answer with the usefulness function, and weights the score using the probability of getting that answer. In our paradigm, a learner thinking about the learning utility considers acting on the toy one way over another, and computes how informative a given answer should be in reducing uncertainty about how the toy works.

Second, the *performance utility* is the utility of successfully making the toy operate. Specifically within our current paradigm, the performance utility is the expected utility of music playing (m) given the learner's action a.

[1] Shannon entropy is a measure of unpredictability or amount of uncertainty in the learner's probability distribution over hypotheses. Intuitively, higher entropy distributions are more uncertain and harder to predict. For example, if the learner believes that all hypotheses are equally likely, then they are in a state of high uncertainty/entropy. In contrast, if the learner firmly believes in one hypothesis, then uncertainty/entropy is low.

$$U_{performance} = \ln(P_L(m|a))$$

Thus, performance utility is maximized by taking an action that is most likely to make the toy "go" and play music, which is the operation of interest.

When there is no observer present, the learner considers the tradeoff between the learning utility and performance utility, and he determines his action based on a weighted combination of the two utilities:

$$U(a;\phi;obs = no) = \phi \cdot U_{learning} + (1-\phi) \cdot U_{performance},$$

where ϕ is a mixture parameter governing the extent to which the learner prioritizes information gain over making the toy play music.

When there is another person present to observe the learner's action, this observer O reasons about the competence c of the learner L which is equal to whether the learner was able to make the toy work.

$$P_O(c) \propto P_L(m|a)$$

The learner thinks about how the observer infers the learner's competence, and his *presentational* utility is based on maximizing the apparent competence inferred by the observer.

$$U_{presentation} = \ln(P_O(c))$$

Thus, when there is an observer present, the learner considers the tradeoff between the learning utility and presentational utility:

$$U(m;a;\phi;obs = yes) = \phi \cdot U_{learning} + (1-\phi) \cdot U_{presentational}$$

Based on the utility functions above, the learner (L) chooses his action a approximately optimally (as per optimality parameter λ) given his goal weight and observer presence.

$$P_L(a|\phi, obs) \propto \exp(\lambda \cdot \mathbb{E}[U(a;\phi;obs)])$$

Experiment

Method

Participants FIXME participants with IP addresses in the United States were recruited on Amazon's Mechanical Turk. We excluded FIXME partcipants who failed to answer manipulation check questions correctly (See Procedure section for details on the manipulation check questions), and thus the remaining FIXME participants were included in our final analysis.

Stimuli and Design We asked participants to imagine they were children's toy developers. We presented three different toys that look very similar but each work in different ways, and provided instructions for them. *The ButtonMusic toy* instructions were: *"Press the button on the right to play music. Pull the handle on the left to turn on the light. Doing both produces both effects."* "*The HandleMusic toy* instuctions were: *"Pull the handle on the left to play music. Press the button on the right to turn on the light. Doing both produces both effects."* and"*the BothMusicLight toy* instuctions were: *"Pull the handle on the left AND press the button on the right to turn on the light and play music at the same time. The button press or handle pull on its own doesn't produce any effect."* Each toy had a label showing its name.

We presented a story to the participants that their boss motivated a goal the participants must achieve by acting on the toy. Importantly, the toy was missing its label, such that partcipants could not know whether the toy was a ButtonMusic, HandleMusic, or BothMusicLight toy. In the *learning* condition, the boss said "That must be one of the new toys that you've been working on. But it looks like you forgot to put on the label! Can you figure out whether this toy is a ButtonMusic toy, HandleMusic toy, or BothMusicLight toy?"; in the *performance* condition, the boss said "That must be one of the new toys that you've been working on. I want to hear the music it plays."; and in the *presentation* condition, the boss said "That must be one of the new toys that you've been working on. How does it work?" followed by the prompt "... you only had one chance to impress your boss and show that you're competent ..." Then we asked participants to select one action to try out on the toy: to "press the button", "pull the handle", or "press the button and pull the handle." Each participant was randomly assigned to one of three goal conditions, and shown a randomized order of actions to choose from.

Procedure Participants were first introduced to the task and shown a picture of a toy with labels on its parts. Then they read instructions for each of the three toys, after which they were asked what they would do to make the toy play music and to make it turn on the light, to make sure they understood the instructions. Participants who answered these manipulation check questions incorrectly were excluded from our final analysis. We then asked participants to rate prior likelihood that an unknown toy is a ButtonMusic toy, HandleMusic toy, or BothMusicLight toy. Participants read a scenario for one of the three goal conditions, and the following instruction: "If you only had one chance to try a SINGLE action to [goal], which action would you want to take? You will get a 10 cent bonus after submitting the HIT if you [goal]." After selecting a response out of three possible actions, the participants were asked again to rate the likelihood for which toy the unlabeled toy was. The experiment can be viewed at `https://langcog.stanford.edu/expts/EJY/soc-info/goal_actions_ver2/soc_info_goals.html`.

Results

Discussion

Acknowledgements

This work was supported by NSERC postgraduate doctoral scholarship PGSD3-454094-2014 to EJY . . . FIXME.

References

Coenen, A., Nelson, J. D., & Gureckis, T. (2017). Asking the right questions about human inquiry.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The nave utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 986–1005.

MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

Nelson, J. D. (2005). Finding useful questions: On bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *112*(4).

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 *Proceedings of the thirty-ninth annual conference of the Cognitive Science Society*.

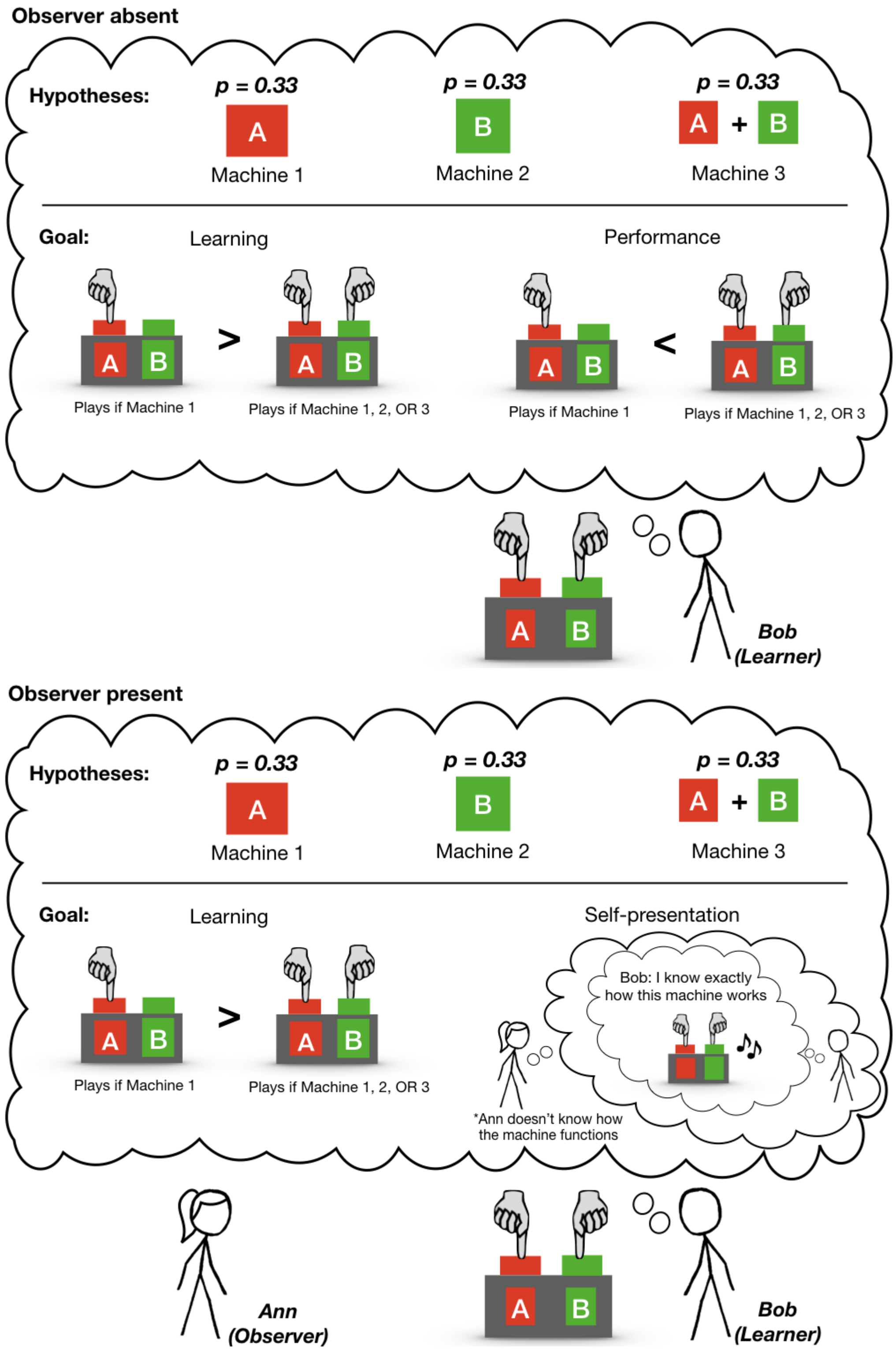

Figure 1: Diagram of the computational model: The learner considers possible hypotheses and his contextual goals. When an observer is absent, he considers his learning goal (to maximize information gain) and performance goal (to play music) and decides on an action. When an observer is present, his decision for an action is based on his learning goal vs. presentational goal (to have the observer infer his competence).