CONFIDENTIAL - FOR PEER-REVIEW ONLY
Image Inference – Inferring actions and desires from static scenes (#26990)

Created: 08/21/2019 11:10 AM (PT)
Shared: 04/13/2021 06:46 PM (PT)

This pre-registration is not yet public. This anonymized copy (without author names) was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) will become publicly available only if an author makes it public. Until that happens the contents of this pre-registration are confidential.

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Research has shown that people can infer an agent's goal by watching how they act (Baker, Saxe, & Tenenbaum, 2009). These models take as input a set of actions, and infer the goal by assuming that the agent was navigating efficiently in space. We hypothesize that this same expectation also allows people to infer agents' goals even in situations where agents are not present and there are no actions to be observed. Specifically, we hypothesize that, when given indirect evidence that an agent was in a certain position in space, people can infer where the agent came from and where they were going by assuming that this state falls within an efficient trajectory in space.

After completing a brief tutorial (see Q8 for tutorial), participants will be presented with two-dimensional grid-worlds with three potential goals, and up to three labeled doors (see Q8 for stimuli). One of the squares in each trial will contain cookie crumbs and participants will be asked to infer (1) the agent's goal (2) the entrance the agent took to enter the room (see Q3 for details on how these are collected). Participants will also have to correctly answer an objective question (Which corner is farthest from Door 1 (there may be more than one)?) on each trial in order to be allowed to continue with the task. The experiment consists of 23 trials presented in random order (see Q4 for design).

3) Describe the key dependent variable(s) specifying how they will be measured.
Goal inference: participants will rate the probability that the agent was going for each of the three goals using continuous sliders ranging from "definitely not" (coded as 0) to "definitely" (coded as 1).

Entrance inference: in trials where there is more than one door, participants will rate the probability that the agent was entering through each door using sliders that are identical to those in the goal inference

4) How many and which conditions will participants be assigned to?
All participants will complete all 23 trials in a random order.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will analyze our data by comparing it to a computational model that infers where the agent entered and what goal they pursued by assuming that the agent was acting efficiently in space. Specifically, our model reconstructs the agent's actions by computing the posterior probability of different sequences of actions given the observed positions of the cookie crumbs through Bayesian inference. The agent's entrance and goal are then computed by integrating over the posterior space of inferred actions.

Our first analysis will consist of correlating mean normalized participant responses with the raw predictions from our computational model . Participant responses will be normalized within-participants, within-trials, and per inference type so that they are valid probability distributions and then they will be averaged across-participants. The model predictions are already valid probability distributions, so they will not undergo further processing. Our second analysis is similar to the first except we will bisect the data by the type of inference participants are making (goal vs. entrance; see Q4). In all of our analyses, we will compute 95% bootstrapped confidence intervals for participant judgments. We will also compare participant responses against a simple alternative model that assumes that the agent entered through the door closest to them, and that they are going to the goal closest to the bread crumbs. We will perform model comparison by computing the correlation difference between our main and our alternative model, along with a bootstrapped 95% confidence interval. We predict that our model will show a reliably higher correlation than the baseline model (operationalized as the 95% confidence interval not crossing 0). This comparison will be performed both on the global analysis (goal and entrance inferences combined) and the split analyses (goal vs. entrance inferences).

If we believe any discrepancies between model predictions and participant judgments are the result of rationality expectations, we will perform post-hoc model adjustment by varying our model's rationality parameter (tau in a softmax equation) to see if the differences between participant responses and model predictions can be attributed to this parameter.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
After reading the task instructions, participants will complete a brief, six-question quiz. Participants that fail the quiz once will be redirected to the

beginning of the instructions and asked to read them again. Participants who fail the quiz twice will not be included in the study. The quiz questions are listed below, with the answers in bold/surrounded by asterisks:

1. How many corners are people walking to?

1 2 3 Not sure

2. Do people always drop their cookie crumbs on their path to/from a corner?

Yes No Not Sure

3. Do people get to choose which door they walk through?

Yes *No* Not sure

4. Do people leave the room out of the same door they entered or the door closest to them?

Same door Closest Door Not sure

5. Can people move diagonally?

Yes *No* Not sure

6. What color are the walls?

White *Gray* Red Not sure

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

Our sample size consists of 40 participants, not counting participants who fail the quiz more than once and are thus not eligible to participate. This sample size was determined by the sample sizes in our previous experiments.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

See the model predictions , procedure, and stimuli at:

https://osf.io/q3ct5/?view_only=4b9b2c68443e4d91b2fb6cf7544055e5