recursion\_writeup\_3

Sam Cheyette

July 9, 2018

Many responses cannot be classified as center-embedded, tail-recusive, or crossing. Indeed, over of all responses, and close to half for the monkeys, do not fall cleanly into one of those categories (notice that the bars in Figures 2 and 3 do not add up to 1). The previous analyses ignored these uncategorized responses, primarily using the relative proportion of crossing and center-embedded responses as a guage of a participant’s learning. While this is entirely valid for its purpose, it does not account for a large portion of the data, and therefore can neither fully explain the variability of the responses nor precisely determine between-participant and between-group differences. So, to better understand the origins of the entire set of responses, we implemented a model to capture the process by which the responses were generated. More specifically, we performed a Bayesian data analysis to jointly infer the strategies used by each participant in the task to make each responses, as well as their noisiness (e.g. mis-presses, memory error, etc…) in implementing those strategies (Gelman et al., 2012). By modeling the strategies that were used by each participant to respond, we we can more precisely describe what participants learned; and by delineating which choices were *intentional* and *unintentional*, we can determine how they were hindered by mistakes.

We formally defined a strategy as a sequence of task-relevant operations . On a given trial, the operations in a strategy are called sequentially until that trial is complete. The three primitive operations we defined are *O*, *C*, and *M*. *O* and *C* choose a random open and closed bracket from the screen; a parameter () determines how biased each one of these is towards a *specific* open or closed bracket. *M* searches through memory for the most recent unmatched bracket and then returns the opposing bracket of the same type. For example, the strategy first chooses an open at random, then another open, then matches the second open, then matches the first open–this strategy correctly outputs only center-embedded recursive sequences. The strategy , on the other hand, is equally likely to generate ([]), ([)], [(]) and [()] since *C* chooses an available “close” at random, regardless of whether it matches the most recent open. We define a *recursive strategy* as one that results in choosing two open brackets and their matching types in order (e.g., ).

The strategies allow for biases towards one specific bracket or another, as well variable levels of noisiness. For instance, could generate center-embedded structures that are biased to begin with “[" rather than a random open bracket; or it could make mistaken bracket-choices frequently. A participant’s “noisiness” refers to the probability that they make a choice inconsistent with the strategy they meant to use on a given trial. This means that for some incorrect responses, such as ([]], the participants have actually intended a correct center-embedded response, such as . Though note that the converse is true as well: some center-embedded responses may have been an accident. By jointly fitting the participants’ intentions, noisiness, and bias together, the model can provide a more complete account of the full range of responses allowing us to more precisely determine what participants have actually learned.

Not including such biases and noise, the hypothesis space over which inference was performed consisted of all strategies that output 4 brackets. Duplicate strategies — those that gave identical responses — were also removed, leaving 12 total strategies. Finally, we considered it likely that many participants used *mixtures* of strategies to produce responses over the course of the task, which would give rise to distributions of responses more complex than that of a single strategy. We deployed a Bayesian inference method to infer what mixture of strategies individuals and populations used.