Semantic structure in communicative drawings

Kushin Mukherjee
Department of Cognitive Science
Vassar College
kumukherjee@vassar.edu

Robert X. D. Hawkins
Department of Psychology
Stanford University
rxdh@stanford.edu

Judith E. Fan
Department of Psychology
Stanford University
jefan@stanford.edu

Abstract

[jefan: Placeholder abstract: Sets this paper up as being about visual communication, that we use semantic segmentation data to investigate.] Drawing is a versatile tool for communication, spanning detailed renderings and simple sketches. Even the same object can be drawn in different ways, depending on the context. How do people decide how to draw in order to be understood? Here we investigate the semantic structure of drawings as a window into how people deploy both perceptual information and conceptual knowledge to produce communicatively effective drawings in context. We analyzed a dataset containing drawings of real-world objects that were produced in different semantic contexts, and contained both detailed and simpler sketches of each object. We explored the hypothesis that during visual communication, people spontaneously decompose visual objects into semantically meaningful parts (e.g., chairs consist of legs, seat, and back), resulting in a tight correspondence between the organization of this semantic part knowledge and the procedure people use to sketch an object. For example, if someone aims to produce a recognizable sketch of a chair, they produce strokes that represent individually meaningful parts, e.g., seat, armrest, legs. To investigate this, we developed a web-based platform to collect dense semantic annotations of the stroke elements in each drawing. We found that: (1) people are highly consistent in how they interpret what individual strokes mean; (2) single strokes tend to represent a single part category (e.g., leg vs. leg + seat), while multiple strokes may be combined to represent an entire part category (e.g., all the legs on a chair); and (3) strokes representing the same part tend to be clustered in time, suggesting that people tend to start and finish drawing one part of an object before moving onto the next.

Keywords: sketching; cognitive science; perception

Introduction

This is where our introduction will go.

Methods

Dataset

[jefan: Great start! I strongly recommend re-organizing this section. First paragraph should set up criteria our dataset should satisfy (i.e., contain sketches produced in different contexts), and where that dataset came from (i.e., a reference game experiment with the close/far manipulation). Second paragraph should describe properties of the dataset (i.e., how many sketches, how strokes relate to sketches, how sub-stroke elements relate to strokes.)]

We obtained 1198 sketches of a set of 32 real-world objects, consisting of 8 exemplars belonging to 4 basic-level categories: cars, chairs, dogs, and birds. The sketches were made during a two-player Pictionary-style reference game, in which a sketcher aimed to produce sketches of target objects that distinguished them from three distractor objects. The viewer had to guess which of the 4 images the sketch represented.

There were 2 main context conditions in the original experiment - close and far. In the close condition the target image and the distractors belonged to the same basic-level category. In the far condition, the target and each of the distractors belonged to a different basic-level category.

The sketches were represented as scalable vector graphics (SVG) images. We were interested in having participants label strokes in a given sketch. We defined a stroke to be equivalent to a single cubic Bezier curve, i.e., a Bezier curve with two fixed end points and two control points to control curvature. [jefan: Is this right? Each stroke consisted of multiple Bezier curves, no?]

Participants

We recruited a total of 326 participants via Amazon Mechanical Turk (AMT). Participants were paid a base amount of $1 and were given an additional bonus of $0.002 for every stroke they annotated. [jefan: Correction: base amount was $0.35.] In addition to this, they were given a $0.02 bonus for every sketch for which they labeled all strokes.

Annotation Procedure

[jefan: Good first draft b/c it is so comprehensive, as it's often easier to trim down than bulk up sections like this. That said, this section does need to be streamlined quite a bit. In revising, think about what the reader *has* to know in order to understand the mechanics of the task — but leaving aside details of the interface that may make it more user-friendly, but not substantively affect the quality/structure of the data (e.g., how the colors of the selected stroke changed to match the background color of the selected part word). A lot of this information (e.g., running counter, progress meter) should be compactly/effectively conveyed in the task display figure, and not repeated in the main text. In a Methods section, it is important to provide a *rationale* for what we did w.r.t. our scientific objectives — concretely, this means prioritizing information that will help a reader understand the subsequent analyses/results, over (no less critical! but not directly relevant to the science) details of the user interface. For example, one aspect of the task that is important to make sure the reader gets is that annotations were provided at sub-stroke granularity — and this is the reason why we're able to examine the question of whether single strokes spanned or stayed within semantic part boundaries.]

To collect fine-grained annotations of our sketches, we im-

plemented a web-based Javascript annotation tool. Our task consisted of one demonstration trial and 10 annotation trials. We provided participants with a sketch to be annotated on a canvas as well as a category-specific menu of labels, which they were encouraged to use for the annotation task. We also provided them with the option of entering their own labels through a free-response box, which could be accessed by clicking an option called Other on the menu. The participant could see the constituent strokes of the sketch by hovering their cursor over different parts of the canvas. To label a stroke, they could click on it or click and drag their cursor over several strokes. This would highlight the selected strokes. Next, they would choose a label from the menu or submit their own through the free response box. The highlighted strokes would change color to match the background color of the selected part-word in the menu. This was to provide a visual aid to the participant as to what strokes had already been labeled and as what part. Participants were encouraged to conduct their labeling of strokes in bouts they were to highlight all the strokes corresponding to a single instance of a part before selecting a label from the menu. A running counter next to each part would indicate how many bouts of annotation had been completed for that part. A running progress meter below the canvas would inform the sketcher as to how many strokes they had labeled out of the total number of strokes in the sketch. Once they were done with sketch, participants could go on to the next trial by clicking a Next button under the canvas. They could choose to continue to the next trial without labeling every stroke in a sketch, but they would lose out on the completion bonus as well as the amount they would have earned for labeling the remaining strokes. After receiving instructions on how to use the tool, participants were told to test out the interface in a demonstration trial. For this trial alone they could not proceed unless they labeled every stroke in the sketch. The sketch shown for the demo trial was the same for every participant.

In total we collected 3608 annotations

Analysis

Results

Discussion

Acknowledgments

Tables

Figures

References