Hypothetical Drawing in Embodied Spatial Reasoning

Atsushi Shimojima (ashimoji@mail.doshisha.ac.jp)
Faculty of Culture and Information Science, Doshisha University
Kyotanabe, Kyoto 610-0394 JAPAN

Yasuhiro Katagiri (katagiri@fun.ac.jp)
Department of Media Architecture, Future University - Hakodate
116-2 Kameda-Nakano, Hakodate, Hokkaido 041-8655 JAPAN

Abstract

Inference by hypothetical drawing is a reasoning process in which imaginary drawings of diagrams are employed for problem solving. We argue, in this paper, that inference by hypothetical drawing is one of the core cognitive mechanisms for realizing embodied spatial reasoning in higher cognition. Based on an eye-tracking experiment on logical inference tasks, we present evidence for the occurrence and the pervasiveness of inference by hypothetical drawing in diagrammatic reasoning. We provide its characterization as a mode of inference for higher cognition, which effectively utilizes inherent characteristics of external space and our capabilities for its manipulation.

Keywords: spatial reasoning; diagrammatic reasoning; embodied reasoning; hypothetical drawing; visual indexing; eye-movement measurement

Introduction

When we try to solve visually posed problems, such as geometry proof, our reflection tells us that it is often useful to draw imaginary lines and other figures on the visual presentation of the problem. We also feel that these hypothetical drawings work as if they were real drawings in the problem solving processes. We call this mode of inference involving imaginary drawings inference by hypothetical drawing. The phenomena of inference by hypothetical drawing have been studied in the context of finding out mental processes involved in visual inference (Schwartz, 1995; Trafton & Trickett, 2001; Yoon & Narayanan, 2004). We believe that it is one of the core cognitive mechanisms for mediating internal mental processes and external resources in higher cognition. Thus, its theoretical and empirical underpinning is significant in elucidating how cognition incorporates spatial and embodied mode of reasoning into our repertoire of higher cognitive processes.

Two directions can be noted in which extension and revision of the orthodox conception on human cognition, discrete symbol manipulation (Newell & Simon, 1972), have been pursued. Prevalence and effectiveness of diagrammatic representations (maps, graphs, flow charts) in human reasoning have inspired studies on diagrammatic reasoning. Characterization of the diagrammatic representation systems in terms of the spatial constraints utilized, and comparison of representational power of diagrammatic and propositional representation systems have been pursued (Shimojima, 1996; Stenning & Oberlander, 1995), with an obvious interest in their place in human problem solving capabilities.

Careful examination of human inference practices have also led to the emphasis on the importance of interaction with environment. Schön (1983) pointed out that a central part of creative thinking consists in the interplay of externalization of ideas in sketches or notes and re-reading of them. Ballard et al. (1997), based on psychological experiments involving visuo-motor coordination tasks, reported that eye-fixations work as a kind of deictic channel that connects internal cognitive processes with external visual environment, through which we access information incrementally as need arises. These and other findings unanimously suggest the need to rethink the strict division between the domains of continuous signals and discrete concepts, and encourage a picture with tighter connection between perception and cognition. Barsalou (1999) even argued that the nature of symbol system itself should be grounded on perception.

Special mechanisms have been proposed for these embodied visual reasoning. The idea of visual routines (Kosslyn, 1994; Ullman, 1984; Chapman, 1989) is to explicate human vision processing in terms of manipulation of a set of pointers to the visual scene. Pylyshyn (2003) proposed the notion of visual index as a fundamental and general mechanism for visual inference and mental imagery.

These studies have mostly looked at the boundary between visual perception and cognition, and little research effort has been directed toward characterization of reasoning mechanisms and elucidation of their roles in higher cognition. We provide evidences for and provides a characterization of inference by hypothetical drawing through an eye-tracking experiment on diagrammatic reasoning, and argue that it is one of the core mechanisms for incorporating embodied spatial reasoning in higher cognition.

Inference by Hypothetical Drawing

Conditions for inference by hypothetical drawing

The efficacy of drawings in diagrammatic reasoning is ascribed to the following two conditions:

- Effective utilization of spatial constraints

Once an adequate diagram is prepared, conclusions can often be read off without performing any inference steps; for example, when point A is placed above B, and B above C, then the fact that A is above C is already in the diagram and need not be inferred. This is because spatial medium inherently imposes transitivity constraint for certain spatial relations such as above.
• Drawing and observation as basic inference operations

Addition of items to and reading information off from drawings are two major steps for inference.

Both of the conditions presuppose that drawings are real, e.g., physically realized in external medium which people have to interact with through perception and action.

Even though it seems clear and reasonable that the process of inference by hypothetical drawing constitutes one of the central processes, on par with diagrammatic reasoning with real drawings, it is necessary to explicate, our intuitions aside, how these two conditions are satisfied with imaginary drawings. Only with such an explication, we can establish the process of inference by hypothetical drawing as a viable mechanism in higher cognition.

Computational model for inference by drawing

In order to substantiate our intuition on the common nature in the processes involving real and hypothetical drawings, and to experimentally investigate the plausibility and characteristics of inference by hypothetical drawing, we begin by stipulating a uniform computational model for the inferences with drawings (see Figure 1).

![Diagram](attachment:image.png)

Figure 1: A scheme for inference by drawing.

The components of the model are the following:

**Higher Cognition Module (HCM):** We assume that the higher cognition module (HCM) operates solely on propositional symbolic representations.

**Procedure Translator (PT):** The procedure translator (PT) accepts propositional information from HCM and translates it into a complex unit of procedures for accessing and manipulating the external environment.

**External Environment (ENV):** The external environment (ENV) provides a real spatial medium in which inference by drawing takes place. Not only typical 2D surface for drawing but also our surrounding 3D space itself can function as external environment.

**Deictic Index (INDEX):** Deictic index (INDEX) is an attention pointer, such as eye-fixation, to objects or location in space. INDEX is a computational resource for procedures, and provides perceptual and motor interface to the external environment. Since we can attend simultaneously to several objects, we assume an index pool which holds a set of indices for objects readily available for perception and action at the present time.

Consider an example of reading off a piece of information that A implies B from a Euler diagram with two concentric circles. Procedure translator first converts the domain predicate ‘implies’ to a spatial relation ‘IS-CONTAINED-IN’ by semantic mapping. It then constructs a procedure for visually inspecting the diagram, e.g., an interaction with the environment. The procedure, when executed, searches for a circle labeled ‘A’ in the environment and places an INDEX to it. Another circle ‘B’ is located in a similar way, and the spatial relation ‘IS-CONTAINED-IN’ will be checked to see if it holds for the two INDICES.

We assume the following set of base procedures:

- **place-object:** A procedure for introducing a new object by placing it at a new location that satisfies the prescribed relationship with objects already present and in focal attention. Given a spatial RELATION and a set of INDICES, assign an INDEX to the new LOCATION (an imaginary object) that satisfies the RELATION together with the INDICES, add the INDEX to the index pool, and return the INDEX.

- **identify-object:** A procedure for searching and identifying an object by its label and other properties within the objects which are already present and attended to. Given a PROPERTY, locate an object in the environment which has an index in the index pool, assign an INDEX to it, and return the INDEX.

- **check-relation:** A procedure for testing if a spatial relationship holds between objects which are in focal attention. Given a spatial RELATION and a set of INDICES, see if the RELATION holds for the INDICES, and return TRUE/FALSE.

Stipulation of a computational model provides us with a solid basis on which to conduct experimental investigations and try out different hypotheses on the phenomena of hypothetical drawing. The point of the indexing mechanism is that a deictic index can be assigned not only to an object explicitly drawn in the scene but also to a location which has no accompanying objects, as when people fixated on a blank point in a scene. Eye movement to a blank point in a scene, if observed, can be interpreted as a positive evidence for index assignment to an empty location, and, hence, for an occurrence of hypothetical drawing.

**Task**

Shimojima and Fukaya (2003) used an eye-tracking method to obtain an evidence for hypothetical drawing in the sense just explained. The present study also uses an eye-tracking method, but it is designed to verify the existence of hypothetical drawing in a yet another way, by introducing *indefiniteness* in the diagrammatic inference task.

In our experiment, the subjects were presented a number of four-term transitive inference problems along with simple position diagrams. Each problem was presented in four steps. The following is a sample procedure:
**Step 1** An audio-recording, “A is cleaner than B,” is played, while the diagram in Figure 2-(a) is presented on a computer display at the same time.

**Step 2** Another audio-recording, “C is cleaner than A,” is played, while the diagram on the display is unchanged.

**Step 3** Another audio-recording, “A is cleaner than O,” is played, while the diagram is unchanged.

**Step 4** An audio recording, “Is C cleaner than O?” is played, while the diagram is unchanged.

![Figure 2: A sample diagram used in the experiment, with rough locations of hypothetically drawn symbols.](image)

From the premise that C is cleaner than A (step 2) and the premise that A is cleaner than O (step 3), it follows that C is cleaner than O. Thus, the correct answer to the question in step 4 is “yes” in this particular problem. The premise given in step 1 is not relevant to the solution.

In this procedure, the subject is instructed to interpret the diagram in step 1 as expressing the same information expressed by the audio recording played at the same time. Thus, in our sample problem, the subject should interpret Figure 2-(a) to mean that A is cleaner than B, assuming the semantic rule that a symbol’s being above another symbol means that the referent of the first symbol is cleaner than the referent of the second symbol.

Throughout the task, the diagram only expresses the first premise presented in step 1. The second and the third premises supplied in steps 2 and 3 remain unexpressed in the diagram. However, if we assume that the subject uses hypothetical drawing in solving our sample problem, where would he or she “draw” symbols?

The second premise says that C is cleaner than A. According to the semantic rule for the diagram at hand, a symbol’s being above another symbol means the referent of the first symbol is cleaner than the referent of the second. Thus, to reflect the second premise in the present diagram, it is sufficient to “draw” the square symbol [C] in the blank area above the square symbol [A]. Thus, this area is the hypothetical drawing position (HDP) in step 2 of the present problem. The greyed symbol [C] in Figure 2-(b) indicates this fact. Thus, one way of verifying the existence of hypothetical drawing is to analyze eye-movements in and out of HDPs: if eye-movements to an area have a stronger tendency when it is an HDP than when it is not, it counts as an evidence that the subjects are engaged in hypothetical drawing.

This much of the experimental idea has been adopted by Shimojima and Fukaya (2003). Our current method allows us to test the existence of hypothetical drawing in yet another way. Consider how the third premise in step 3 might be expressed in the present diagram. It says that A is cleaner than O. Given the semantic rule of the current diagrams, expressing this information would require that the symbol [A] be above the symbol [O]—that is, that the symbol [O] be below the symbol [A]. In order to place [O] below [A], however, one need make a difficult choice on where to place [O] relative to the existing symbol [B]. It would be a mistake to place [O] above [B] since it would carry the information that O is cleaner than B, which is not implied by the given premises. It would be also a mistake to place [O] below [B] since it would carry another unwarranted information that B is cleaner than O. The position diagrams only allow the symbols to be ordered linearly without overlap, there is no definite way of placing [O] below [A] without thereby expressing an unwarranted piece of information. The drawing position of the symbol [O] is indefinite in this sense. Figure 2-(c) indicates this indefiniteness by placing the greyed [O] beside [B].

This is an instance of the very common property of diagrams, known as “specificity” (Stenning & Oberlander, 1995) or “particularity” (Kulpa, 2003). Varieties of diagrams, including maps, geometry diagrams, Euler diagrams, prohibit the exclusive expression of information in this manner, enforcing the choice of additional, unwarranted information to express. This property is a characteristic weakness of diagrams, and as such, it intervenes in a cognitive process only when the cognitive process involves a drawing on a diagram. Thus, if we could find an evidence that the subjects’ performance are affected in the contexts that requires the “drawing” of indefinite symbols, then we could count it as a new kind of evidence for the existence of hypothetical drawing.

**Method**

**Problem types**

All the problems used in our experiment adopted the same four-step procedure explained in the previous section. A half of the problems are “vertical problems,” using a diagram with vertically arranged square symbols (see Figure 2-(a) for example). The other half of the problems are “horizontal diagrams,” using a diagram with horizontally arranged symbols. A half of vertical problems use “cleaner than” as the relational predicate, while the other half use “dirtier than.” Similarly, a half of horizontal problems use “east to” and the other half use “west to” as the relational predicate.

Also, the contents of the premises used in problems significantly vary, so that different problems may prescribe different positions and timings for hypothetically drawn objects on the given diagrams. Figure 3 shows six different types (v1, v2, v5, v6, v7, and v8) of vertical problems, classified on this basis. Here, the number in a grey square indicates the timing (the step number) in which the symbol in question is supposed to be hypothetically drawn on the diagram.
For example, the configuration of type v2 has a greyed square numbered 2 in the uppermost area, meaning that problems of of this type prescribes a hypothetical drawing in that area in step 2. The sample problem described in the previous section is of type v2, and as the reader may recall, it indeed prescribes a hypothetical drawing of a symbol [C] in the uppermost area in step 2. Also, the sample problem prescribes a hypothetical drawing of an indefinite symbol [O] in step 3, and the configuration of v2 indicates this fact by the greyed square numbered 3 placed beside the lower black square. Thus, one can see from Figure 3 that problems of types v1, v2, v7, and v8 introduce indefinite symbols in some timing, whereas problems of types v5 and v6 never do so.

In the similar vein, horizontal problems are divided into six types (h1, h2, h5, h6, h7, and h8), which are completely analogous to their vertical counterparts in their prescriptions of the timings and positions of hypothetical drawing.

Predictions on eye-movements

Our computational model for inference by drawing distinguishes the drawing operation (place-object) and the searching operation (identify-object) on diagrams even when the same object is being addressed. The following hypotheses on the working of these operations can be generated naturally from the model:

- When drawing an object, people move their eyes to the location where the object should be placed to assign an INDEX to it.
- When searching an object in the scene, people move their eyes over to all the possible candidate locations to find the appropriate INDEX of it.
- Searching takes place (and results in failure) before drawing is initiated to find out that the target is a new object to be added to the diagram.

Compare the drawing of C in Figure 2-(b) and the drawing of O in Figure 2-(c). In Figure 2-(b), eyes would move to as high a location as grey [C] to (hypothetically) draw the object C. In Figure 2-(c), eyes would also move to as high a location as grey [C] in searching for the object O before the search fails and turns to (hypothetically) draw grey [O] beside [A].

Applying the above set of hypotheses, we obtain a systematic predictions on eye movements in each step of a problem of a given type. Due to space considerations, however, we will confine our analysis to step 3 of each problem.

Look back to the list of six types of vertical problems in Figure 3. What can be predicted on eye-movements in the third step of each type? For problem types v1 and v6, we predict eyes move to relatively high positions to “draw” the symbol numbered 3. Eyes would also move to higher positions in problem types v2 and v5, but this time the symbol numbered 2 drawn in step 2 is checked during the “search” for the symbol numbered 3. The higher limits of eye movements in problem types v7 and v8 would be relatively low, because there is nothing to draw or search in higher positions.

Thus, we can categorize vertical problems according to what can be predicted for the higher limits of eye-movements in the third step. We will call problems of types v1 and v6 “higher-hi,” with the subscript “D” indicating that eyes would move to higher positions to draw a symbol (rather than to search a symbol). Problems v2 and v5 will then be called “higher-hs,” with “S” indicating searching. Problems v7 and v8 will be simply called “lower-hi.”

We can similarly categorize vertical problems according to what can be predicted for the lower limits of eye-movements in the third step. Problems of type v5 and v7 are “lower-lo,” v6 and v8 are “lower-lo,” and v1 and v2 are “higher-low.”

We can get analogous categories of horizontal problems. Specifically, we can categorize horizontal problems into “letter-left,” “letter-left,” and “righter-left” according to what are predicted for the leftmost limits of eye-movements. The same set of problems can be re-grouped into “righter-right,” “righter-right,” and “letter-right” in the context of evaluating the rightmost limits of eye-movements.

Predictions on response latency

Position diagrams used in our experiments have a weakness in expressive flexibility. Consequently, the problem of positioning an indefinite symbol occurs in certain contexts. Look back to Figure 3 again. The grey squares placed beside other squares are indefinite symbols, and the numbers inside the squares indicate the timings (step numbers) in which they must be positioned in the relevant diagrams. Thus, the second steps of problem types v1 and v7 would be more difficult than the second steps of other problem types. Also, the third steps of problem types v2 and v8 would be more difficult than the third steps of other problem types. We predict that this difference would be reflected in the difference in latency.

Analogous considerations apply to horizontal problems. Thus, we predict that latency in second steps would be greater in problem types h1 and h7 than in other problem types, and that latency in third steps would be greater in problem types h2 and h8 than in other types.

Procedures

A total of 26 undergraduate students (17 females and 9 males) participated in the experiment for monetary reward. The diagram was presented on a 19-inch computer display at the resolution of 1280 × 1024 pixels. The subject’s jaw was placed.
Table 1: Vertical position (pixel coordinate) of upper and lower limits of eye-movements.

<table>
<thead>
<tr>
<th></th>
<th>higher-hi</th>
<th>higher-lo</th>
<th>lower-hi</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>upper</strong> limits</td>
<td>663.2 (S.D.: 75.8)</td>
<td>695.7 (S.D.: 92.3)</td>
<td>600.5 (S.D.: 56.5)</td>
</tr>
<tr>
<td><strong>lower</strong> limits</td>
<td>355.6 (S.D.: 62.2)</td>
<td>341.0 (S.D.: 77.8)</td>
<td>378.0 (S.D.: 65.9)</td>
</tr>
</tbody>
</table>

Table 2: Horizontal position (pixel coordinate) of leftmost and rightmost limits of eye-movements.

<table>
<thead>
<tr>
<th></th>
<th>lefter-left</th>
<th>righter-left</th>
<th>lefter-right</th>
<th>righter-right</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>leftmost</strong> limits</td>
<td>510.5 (S.D.: 55.9)</td>
<td>490.5 (S.D.: 56.4)</td>
<td>566.2 (S.D.: 50.2)</td>
<td></td>
</tr>
<tr>
<td><strong>rightmost</strong> limits</td>
<td>806.1 (S.D.: 65.7)</td>
<td>812.2 (S.D.: 68.7)</td>
<td>741.5 (S.D.: 56.0)</td>
<td></td>
</tr>
</tbody>
</table>

on a fixed support. The distance between the subject’s eyes and the display was approximately 21.7 inches, and the viewing angles were approximately 38 degrees (horizontal) and 31 degrees (vertical). The eye-movements were tracked and recorded by NAC EMR-AT (model ST-600). The subject was instructed to solve the problems as quickly and accurately as possible. The subject used a response pad (Cedrus RB-530) to control the progress of all four steps in a problem, and the latency in each step was recorded. After 8 practice problems, the subject solved 56 problems, with a break in every 14 problems.

Results

Eye movements

Each of 26 subjects solved 56 problems, totaling up to 1456 trials. A total of 1243 trials (85.4 %) were answered correctly, and we analyzed eye-movements in those trials only. Due to calibration failure, the data of one subject had to be excluded from analysis. Due to space considerations, we will focus on eye-movements in third steps of the problems.

Table 1 shows the average vertical coordinate values (“Y-values” henceforth) of the upper limits and the lower limits of eye-movements during the third steps of trials. As predicted, eyes reached significantly higher in higher-high problems than in lower-high problems \( (F(2, 22) = 25.11, p < .01) \). Also, eyes reached significantly lower in lower-low problems than in higher-low problems \( (F(2, 22) = 4.74, p < .05) \), although the pair-wise difference between lower-low problems and higher-low problems did not reach significance.

Table 2 shows the average horizontal coordinate values (“X-values” henceforth) of the leftmost limits and the rightmost limits of eye-movements during the third steps of trials. Again, eyes moved as predicted: the X-values of the leftmost limits of eye-movements is significantly smaller in lefter-left problems than in righter-left problems \( (F(2, 22) = 31.29, p < .01) \), while the X-values of the rightmost limits of eye-movements is significantly larger in righter-right problems than in lefter-right problems \( (F(2, 22) = 19.35, p < .01) \).

Response latency

Table 3 shows the average response latency in second and third steps, comparing the contexts with and without indefinite symbols.

<table>
<thead>
<tr>
<th>Problem types</th>
<th>Mean (msec)</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second v1, v7, h1, h7</td>
<td>4074</td>
<td>925.7</td>
</tr>
<tr>
<td>steps other types</td>
<td>3756</td>
<td>744.2</td>
</tr>
<tr>
<td>Third v2, v8, h2, h8</td>
<td>4448</td>
<td>1387.4</td>
</tr>
<tr>
<td>steps other types</td>
<td>3873</td>
<td>830.7</td>
</tr>
</tbody>
</table>

Evidence for inference by hypothetical drawing

Significant differences found in both directions for both X-values and Y-values shown in Tables 1 and 2 count as evidences that inference by hypothetical drawing is a real mental process operating in human reasoning with diagrams.

Figure 4 shows the positions of the average vertical limits and horizontal limits of eye-movements as revealed in Tables 1 and 2, relative to the two square symbols in vertical and
horizontal diagrams. In higher-high problems, eyes tend to reach higher than the top side of the upper square symbol (by 50 pixels on average), while they tend to remain lower than that side (by 28 pixels on average) in lower-high problems. On the other hand, eyes tend to reach lower than the bottom side of the lower square symbol (by 42 pixels on average) in lower-low problems, while they remain closer to that side (by 30 pixels on average) in higher-low problems. A similar tendency can be observed for the horizontal problems.

Coping with indefiniteness
Table 3 indicated that even for hypothetical objects, weakness in expressive flexibility of diagrammatic representations adversely affected on the drawing performance. This implies that the process of inference by hypothetical drawing is an actual “drawing,” in the sense that it shares the essential characteristics of actual drawing, and it is subject to constraints inherent in space. If it were not drawing, it would not be affected in a way drawing would be affected. In other words, this finding provides us with another type of evidence for the reality of inference by hypothetical drawing. It also indicates that inference by hypothetical drawing is a pervasive process, which is invoked even when it is not fully up to the task.

Uniformity across real and hypothetical objects
Consider the upward eye-movement patterns when the symbol numbered 3 is hypothetically drawn in the problems v5 and v8 in Figure 3. In both of the cases, eyes would scan all the objects already introduced into the diagram before they notice the symbol must be newly drawn in the diagram. If real and hypothetical objects are uniformly treated in the process of inference by hypothetical drawing, eyes will move up toward the high position of the hypothetical object for the symbol numbered 2 in v5, which should be higher than the position of the top real object, which is the target of the upward eye-movement in v8. If, on the other hand, the process is only sensitive to real objects in searching, then there will be no significant difference in the upward eye-movement patterns in v5 and v8.

Difference between “higher-hi3” and “lower-hi” positions shown in Table 1 suggests, and a further examination of the data confirmed, that there is a significant difference in upward eye-movements between the two problems. The same difference is also observed for all directions. These findings provide a strong support for the putative characterization of the process of inference by hypothetical drawing that it is uniform across real and hypothetical objects.

Conclusions
The phenomena of hypothetical use of drawings in diagrammatic problem solving were investigated. A computational model was proposed for the detailed examination of the processes involved. Based on the measurement of eye movement and response latency in a reasoning task experiment, strong evidences were found confirming our intuition that inference by hypothetical drawing is a real and pervasive mental process. It is successfully employed even in cases where use of space is inadequate, and it treats real and hypothetical objects uniformly without procedural distinctions. The environment in which inference by hypothetical drawing takes place can be extended to 3D space, and it constitutes a general mechanism for embodied spatial reasoning for higher cognition.

Acknowledgments
The work reported in this paper was partially supported by Japan Society for the Promotion of Science Grants-in-aid for Scientific Research (B) 18300052, and Grants-in-aid for Scientific Research (C) 18500206.

References