Title

Diagram-based Problem Solving: Th e Case of an Impossible Problem

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 17(0)

Authors

Narayanan, N. Hari
Suwa, Masaki
Motoda, Hiroshi

Publication Date

1995

Diagram-based Problem Solving: The Case of an Impossible Problem

N. Hari Narayanan
College of Computing & EduTech Institute
Georgia Institute of Technology
Atlanta, GA 30332-0280
narayan@cc.gatech.edu

Masaki Suwa Hiroshi Motoda
Advanced Research Laboratory
Hitachi Ltd.
Hatoyama, Saitama 350-03, Japan
suwa,motoda@harl.hitachi.co.jp

Abstract

Diagram-based problem solving is an activity in which subjects solve problems that are specified in the form of diagrams. Since the diagram contains critical information necessary for problem solving, this is an activity that clearly requires *reasoning with* the diagram. Recent research on diagrammatic reasoning has uncovered many interesting aspects of this process. One such aspect that the authors have been exploring, by means of a set of verbal and gestural protocol analysis experiments, is the role of the diagram in guiding the reasoning process. The trajectory of reasoning is revealed both by the intermediate hypotheses generated, and by the shifts of focus induced from problem solving protocols. In this paper we focus on the protocols collected for a particularly interesting problem, one whose solution is arrived at through a pair of contradictory inferences. We derived the reasoning trajectories of subjects by extracting the temporal order and spatial distribution of their intermediate hypotheses leading toward the final solution. These trajectories indicate that the spatio-temporal order of hypotheses depend on more than the device structure depicted in the diagram and inferred causation of events from the diagram. We propose that subjects employ implicit search strategies which together with their internal goals to verify hypotheses and the need to replenish short term memory influence their reasoning trajectories.

Introduction

More often than not, external representations form integral parts of the representational repertoire utilized by human cognitive processing. However, it is only recently that research has begun to address issues of how cognitive processes operate on distributed representational systems that consist of external and internal representations, and what the representational effects different kinds of representations have on cognitive processes (Zhang & Norman, 1994). Diagrams form an interesting class of external representations, one which is quite often used in activities such as instruction, design and problem solving. Zhang and Norman (1994) argue that external representations can anchor and structure cognitive behaviors. In this vein, how external diagrams facilitate reasoning, visualizations, and problem solving is a topic that has recently received considerable research attention from both computational and cognitive perspectives (for example, see Cheng, 1994; Clement, 1994; Glasgow, Narayanan & Chandrasekaran, 1995; Lindsay, 1994).

Diagrammatic reasoning may be defined as the kind of reasoning in which diagrams are used as external representations. Diagram-based problem solving then is a particular kind of diagrammatic reasoning activity in which problems, specified in the form of diagrams that are annotated with labels and some explanatory text, are solved. Research on diagram-based problem solving with diagrams of mechanical devices has uncovered many interesting aspects of this process. For instance, it has been found that readers who study mixed media descriptions comprising both diagrams and text construct better mental models of the kinematics of a device (Hegarty & Just, 1993) and are better able to solve problems concerning the functioning of the device (Mayer, 1989). Eye-fixation and protocol studies (Hegarty, 1992; Hegarty & Sims, 1994; Narayanan, Suwa & Motoda, 1994) show that subjects decompose the device representation into units corresponding to components or groups of components related by contact and connectivity as perceived from the diagram. These studies also reveal that during problem solving the diagram serves as an external memory, facilitates mental visualization of spatial behaviors of device components, and guides the reasoning process along the direction of causality.

One aspect of diagram-based problem solving that we have been exploring is the role of the diagram in guiding the reasoning process. We have collected and analyzed concurrent think-aloud protocols and gestures of subjects solving a set of device behavior hypothesis problems presented as labeled diagrams. In addition to analyzing verbal protocols, the gestures and marks made by subjects were examined and used to annotate the verbal data. In earlier work we proposed a model of diagrammatic reasoning for this problem solving task, and explained focus shifts induced from the protocol data in terms of perceptual and cognitive processes (Narayanan, Suwa & Motoda, 1994). In this paper we continue that exploration, this time by focusing exclusively on the protocols collected for a particularly interesting problem - an "impossible"[1] problem whose solution is arrived at through a pair of contradictory inferences.

The trajectory of reasoning in diagram-based problem solving is revealed at the macro level by the intermediate hypotheses that subjects generate, hypotheses which lead toward a final solution, and at the micro-level by the shifts of focus that could be induced from protocol data or by tracking eye-movements. Here we derive the reasoning trajectories that subjects traversed in solving this "impossible" problem by extracting, directly from the verbal protocols annotated with

---

[1] This is called an *impossible* problem because unlike the other problems used in this set of experiments, the most intuitive solution to this problem requires one to predict that no motion is possible. Furthermore, this impossibility of motion is revealed only at the very end when two opposite motions get predicted for the same component.

gestural information, the temporal order and spatial distribution of their intermediate hypotheses leading toward the final solution. These trajectories indicate that the spatio-temporal order of subjects' hypotheses depends on the device structure as depicted in the diagram, inferred (from both the diagram and prior knowledge) causation of events, the implicit search strategy employed in traversing branching and merging event chains, internal verification goals, and the need to replenish short term memory.

Experimental Method

The problem solving task used in this study was the following: given the schematic diagram of a mechanical device depicting the spatial configuration of its (labeled) components[2] and an initial behavior, hypothesize the potential behaviors of the device in terms of the behaviors of its components.

Subjects: Five adult high school graduates (three of whom – named subject 1, 2 and 3 – had some vocational training and were employed as technicians whereas the other two – named subject 4 and 5 – were administration employees) volunteered as subjects.

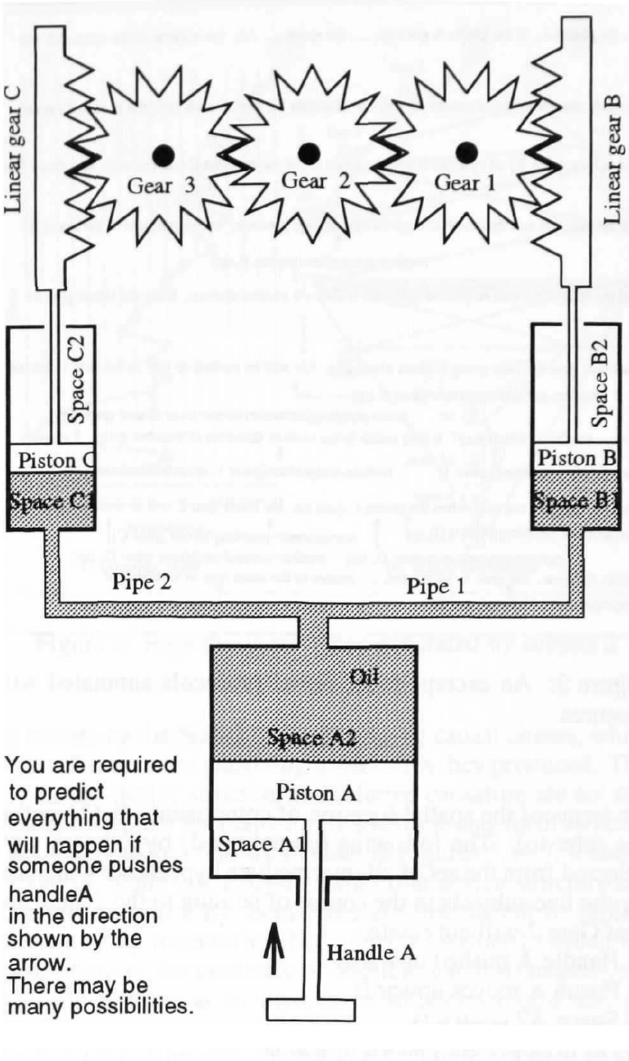
Materials: The subjects were seated at a table and presented with one sheet (per problem) containing a labeled diagram with an initial condition and instructions written below the diagram. A pen was kept on the table. The subjects were told that they could use it to point or draw on the problem sheet.

Procedure: All subjects attended an initial session in which concurrent think-aloud verbal reporting (Ericsson & Simon, 1983) was explained and illustrated by the experimenter. Each subject attended two problem solving sessions lasting approximately 45 minutes each, separated by a week. Subjects were asked not to discuss the experiments among themselves during this period. In each session a subject was first given a general instruction sheet that explained what was expected of them in terms of think-aloud reporting. These instructions followed the guidelines presented in (Ericsson & Simon, 1983) . They were then given three training problems followed by the actual problems. Concurrent verbal and gestural data were collected. Verbal reports (in Japanese) were tape-recorded and gestures with hands and pen were video-taped. The verbal reports were transcribed and translated into English. Gestures and drawings that the subjects made were examined using both the video recording and the problem sheets on which subjects drew. These gestures typically appeared concurrently with verbalizations or during pauses.

The entire study involved five subjects and six problems. Of the thirty protocols collected, twelve (three subjects and four problems) were annotated with gestural data, segmented, encoded and analyzed to arrive at the model and conclusions presented in (Narayanan, Suwa & Motoda, 1994). We consider the protocols of all five subjects on the fifth problem in this paper.

Spatio-Temporal Order of Hypotheses

The "impossible" problem that was given to the five subjects is shown in Figure 1. For the analysis presented here we used their raw verbal protocols directly, annotated with descriptions of accompanying gestures, Figure 2 shows an excerpt from one such annotated protocol.



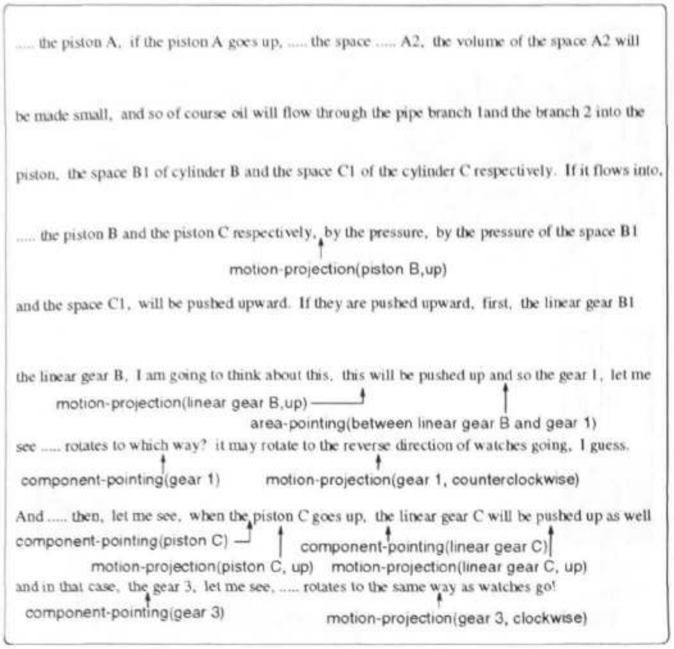
You are required to predict everything that will happen if someone pushes handleA in the direction shown by the arrow. There may be many possibilities.

Figure 1: The Problem

It is fairly obvious that the most intuitive solution of this problem is to predict that Gear 2 will not rotate (and therefore, Handle A will resist being pushed). Since the problem is under-specified (e.g., relevant parameters such as the force with which Handle A is being pushed are not provided), other solutions are certainly possible. In fact, since the problem statement explicitly noted the possibility of multiple solutions, every subject came up with other somewhat more far-fetched scenarios as well (e.g., Gear 2 rotating in one direction, breaking off the teeth of the opposing gear). Interestingly enough, all subjects initially reached the conclusion that Gear 2 will not rotate before considering other possibilities. Upon further reflection on this solution, it may be seen that there is a a sequence of intermediate hypotheses that culminate in this solution and that these hypotheses correspond to a causal chain of events triggered by pushing Handle A. This causal chain splits into two paths at the branching point of the pipe, and later joins together at Gear 2. Therefore, this sequence of hypotheses has both a temporal order (hypotheses regarding earlier events precede those regarding later events in the device) and a spatial distribution over the structure of the device

---

[2] We use the term *components* to mean components, individual parts of components, and substances.



..... the piston A, if the piston A goes up, ... the space ..... A2, the volume of the space A2 will

be made small, and so of course oil will flow through the pipe branch 1and the branch 2 into the

piston, the space B1 of cylinder B and the space C1 of the cylinder C respectively. If it flows into,

..... the piston B and the piston C respectively, by the pressure, by the pressure of the space B1

motion-projection(piston B,up)

and the space C1, will be pushed upward. If they are pushed upward, first, the linear gear B1

the linear gear B, I am going to think about this, this will be pushed up and so the gear 1, let me

motion-projection(linear gear B,up)

area-pointing(between linear gear B and gear 1)

see ..... rotates to which way? it may rotate to the reverse direction of watches going, I guess.

component-pointing(gear 1) motion-projection(gear 1, counterclockwise)

And ..... then, let me see, when the piston C goes up, the linear gear C will be pushed up as well

component-pointing(piston C) component-pointing(linear gear C)

motion-projection(piston C, up) motion-projection(linear gear C, up)

and in that case, the gear 3, let me see, ... rotates to the same way as watches go!

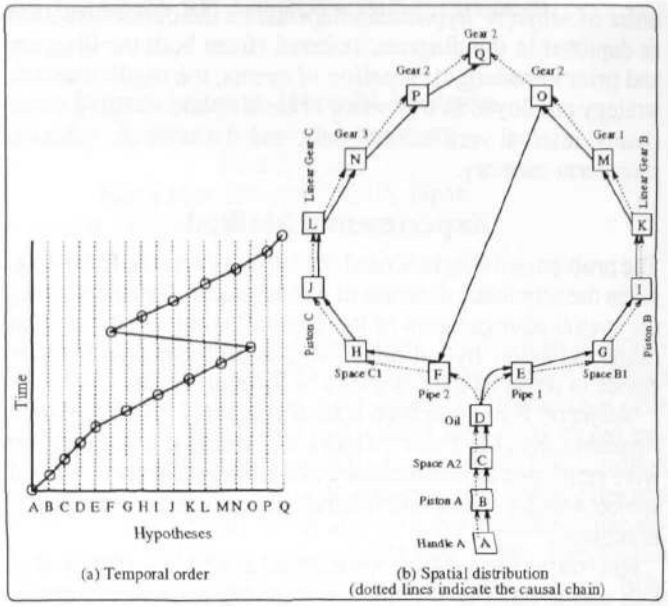
component-pointing(gear 3) motion-projection(gear 3, clockwise)

Figure 2: An excerpt from verbal protocols annotated with gestures

(in terms of the spatial location of components that hypotheses refer to). The following (paraphrased) hypotheses were selected from the set of all intermediate hypotheses generated by the five subjects in the course of coming to the conclusion that Gear 2 will not rotate.

A Handle A pushed upwards.
B Piston A moves upwards.
C Space A2 contracts.
D Oil in Space A2 pushed upwards.
E Oil pressure in Pipe 1 increases.
F Oil pressure in Pipe 2 increases.
G Oil in Space B1 moves upwards.
H Oil in Space C1 moves upwards.
I Piston B moves upwards.
J Piston C moves upwards.
K Linear Gear B moves upwards.
L Linear Gear C moves upwards.
M Gear 1 rotates counter-clockwise.
N Gear 3 rotates clockwise.
O Gear 2 rotates clockwise.
P Gear 2 rotates counter-clockwise.
Q Gear 2 will not rotate at all.

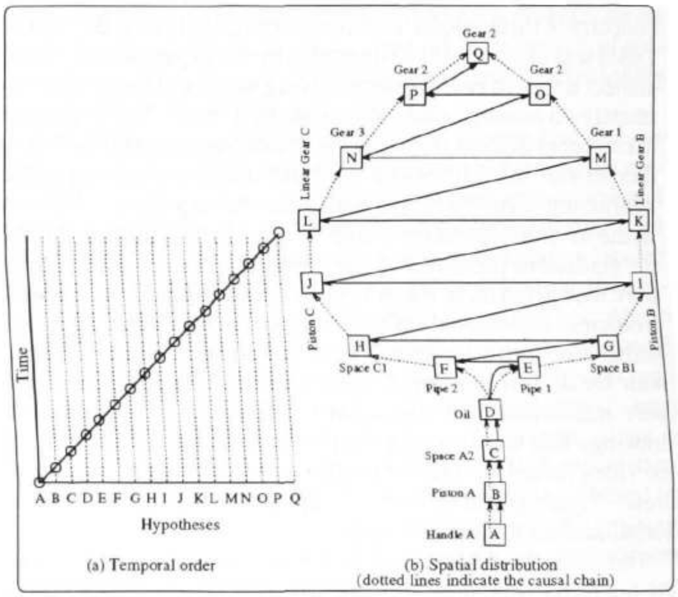
The spatial distribution of these is obvious since each hypothesis refers to one particular component of the device. But what is their temporal order? Clearly, hypotheses A, B, C and D follow in that order. At this juncture the causal chain splits, and the temporal order is no longer unique. The model of diagrammatic reasoning that we proposed in (Narayanan, Suwa & Motoda, 1994) suggests that at the beginning of each cycle of reasoning a hypothesis to focus on is retrieved from short term memory, into which all new hypotheses generated in that cycle is stored at its end. But in what order, if any, are hypotheses retrieved from short term memory? Computationally, there are two standard ways of traversing a branching structure such as a network or tree: depth-first and breadth-first. Figures 3 and 4 illustrate the temporal order and spatial distribution of aforementioned hypotheses if the causal chains were to be followed depth-first or breadth-first respectively (assuming that the right-side branch is attempted first). These figures are provided for purposes of comparison with Figures 5, 6, 7, 8 and 9, which show the temporal order and spatial distribution of the same hypotheses as generated by the five subjects.



(a) Temporal order

(b) Spatial distribution (dotted lines indicate the causal chain)

Figure 3: Depth-first generation of intermediate hypotheses



(a) Temporal order

(b) Spatial distribution (dotted lines indicate the causal chain)

Figure 4: Breadth-first generation of intermediate hypotheses

Discussion

We have illustrated the temporal order and spatial distribution of intermediate hypotheses that subjects generated in the

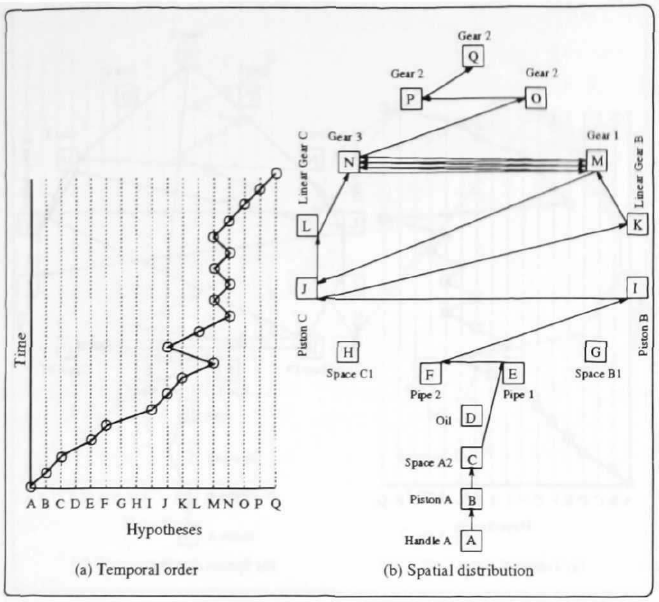


Figure 5: Hypothesis sequence generated by subject 1

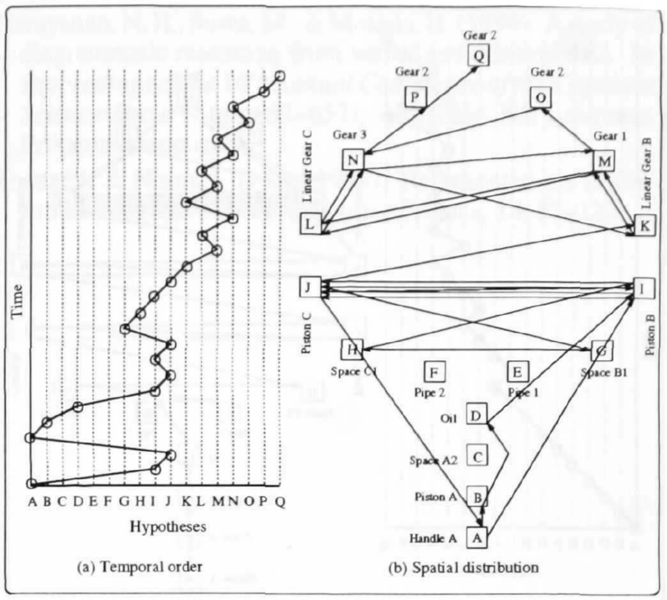


Figure 6: Hypothesis sequence generated by subject 2

course of arriving at an intuitive solution to a given problem of predicting the operation of a diagrammatically specified mechanical device. This problem is interesting because the causal chain of events in the device undergoes a split and a merge, with a contradiction at the merge point alerting the subjects to the solution. The trajectories of subjects' reasoning are derived from the spatio-temporal order in which hypotheses are generated during diagram-based problem solving. What can be said about factors that influence these trajectories? We postulate that reasoning trajectories of diagram-based problem solving are influenced by the following factors.

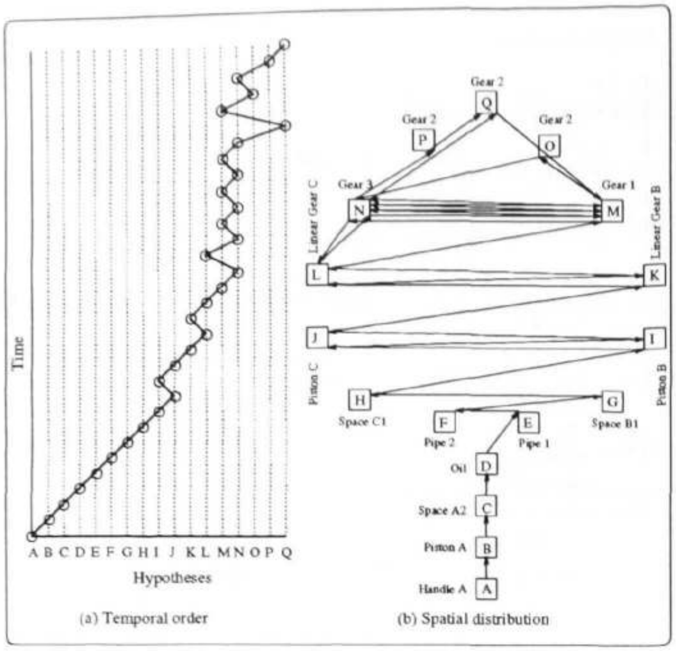
Device Structure: The device structure as depicted in the diagram clearly influences the trajectory of reasoning. This is particularly evident in the problem here because the structure of the device as depicted in the diagram provides the primary clues to the branching and merging of component behaviors.

Inferred Causation: The causal order of events in the operation of the device is inferred from the diagram by applying prior knowledge about mechanical components. Causation is most often inferred by using connectivity, contact or proximity information from the diagram, or by mental visualization of component behaviors using the diagram. As may be verified from Figures 5, 6, 7, 8 and 9, the spatio-temporal order in which hypotheses are generated follows the causal order, except for switches between different branches or to previously generated hypotheses.

Search Strategy: That the device structure as depicted in the diagram can influence one's reasoning trajectory seems intuitive and plausible. So does the conclusion that the flow of causality that subjects infer in the operation of the device will determine their reasoning trajectory to some extent, by influencing the spatio-temporal order in which events are hypothesized. Both of these conclusions are also supported by prior reaction-time, eye-fixation and protocol analysis studies (Hegarty, 1992; Hegarty & Just, 1993; Hegarty & Sims, 1994; Narayanan, Suwa & Motoda, 1994). But we postulate that subjects must have employed some implicit search strategies in traversing the branching and merging causal chains, which also influenced the reasoning trajectories they produced. This is because device structure and inferred causation are not sufficient to explain the many loops and back-and-forth switches between branches that are evident in Figures 5, 6, 7, 8 and 9. However, while our analysis shows that device structure and inferred causation by themselves are insufficient to explain the reasoning trajectories that subjects produced (which led us to postulate the existence of implicit search strategies), the question remains as to what these strategies might be. As the contrast between the reasoning trajectories produced by depth-first and breadth-first searches (Figures 3 and 4) and the actual trajectories of subjects indicates, no subjects consistently employed either one of these strategies.

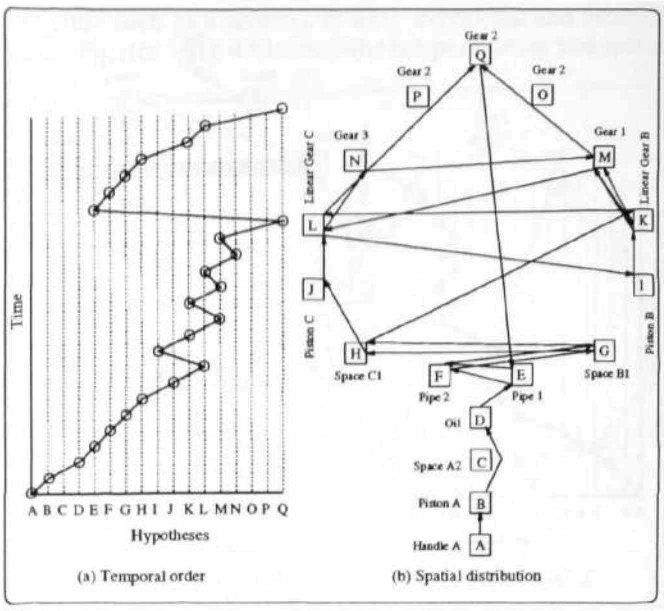
It is also not clear whether the diagram affects the search strategy employed, or whether the influence flows the other way – the search strategy influencing focus shifts among components in the diagram. It is evident that none of the five subjects followed a search strategy (such as depth- or breadth-first) consistently. Instead, the spatio-temporal order of their hypotheses indicate a more or less random combination of the two. This is true in general for all the thirty protocols we have collected. Uncovering the logic behind this apparent randomness by discovering factors that determine the search strategy is an important and unresolved issue requiring further research. A related open question is whether and how the search strategy is affected by the diagram, or whether the influence flows only in the other direction.

Verification goals: Many back-and-forth switches within sub-sequences of hypotheses may be noticed in the hypothesis sequences that the five subjects generated. These can be explained in terms of an internal goal to re-confirm an already generated hypothesis by returning to a previous hypothesis (in the causal chain) and re-deriving the intermediate hypotheses that lead to the hypothesis to be re-confirmed. This behavior is particularly evident among hypotheses closer to the final solution (see the temporal orders of hypothesis generation by



(a) Temporal order (b) Spatial distribution

Figure 7: Hypothesis sequence generated by subject 3



(a) Temporal order (b) Spatial distribution

Figure 8: Hypothesis sequence generated by subject 4

the five subjects).

Replenishing short term memory: It is also possible that some of the shifts across branches that can be observed in these figures are merely for replenishing the set of active hypotheses being maintained in short term memory.

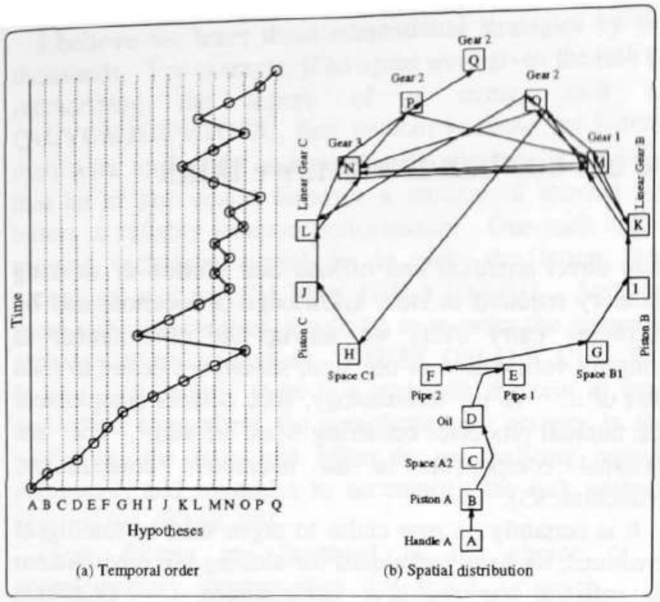
In an earlier paper we did a fine-grained analysis of focus shifts – how the problem solving focus is shifted from component to component – during diagrammatic reasoning (Narayanan, Suwa & Motoda, 1994). This revealed that three perceptual and cognitive processes (noticing connectivity/contacts between components, mental visualization of component behaviors, and search for information) can explain 62% of the total number of focus shifts observed over the twelve protocols considered. The present analysis indicates that some of the remaining 38% focus shifts may be explainable by the search strategy employed, which causes switching among spatially separated components (belonging to different branches of behaviors). Internal goals of verifying (or re-deriving) already generated hypotheses and the need to replenish short term memory from time to time also may generate focus shifts that are not explained by the three factors we considered previously. Thus, the search strategy, internal goals, and replenishing short term memory could be three additional factors which influence focus shifts at the micro-level and spatio-temporal order of hypothesis generation at the macro-level.

Though this study used diagrams of mechanical devices, we believe that our findings have relevance in general to the design of static and dynamic (e.g., animations and other kinds of visualizations) graphical representations to guide reasoning, or to teach phenomena involving the propagation of causality. It is important that such graphical representations (1) clearly show the physical structure of the domain in terms of delineated components, (2) explicate (particularly in case of animations) the causal chains of events that occur, (3) explicitly provide clues about which branches to follow in what order if the behaviors of interest branch and merge instead of being linear, and (4) facilitate the reasoner (or the learner) to go back and forth between earlier and later stages in the reasoning process (e.g., by providing control to the user to "rewind" and "restart" animations).

While previous research by Hegarty and colleagues (Hegarty, 1992; Hegarty & Just, 1993; Hegarty & Sims, 1994) has shown that device structure as depicted in the diagram and inferred causation of events in the operation of a device are significant factors in diagrammatic reasoning, the influence of implicit search strategies and internal goals of reasoning (such as goals to re-confirm hypotheses) on the trajectory of diagrammatic reasoning in the case of devices with branching and merging behaviors has not hitherto been considered. The main contributions of work reported here are in illustrating how the spatio-temporal order of hypothesis generation can be derived from protocol studies of diagram-based problem solving, and how the graphical depiction of this order may provide clues to other factors influencing the trajectory of diagrammatic reasoning. A note of caution, however, is that these conclusions are preliminary, and more research is required to gather additional evidence. What determines the search strategy employed in diagram-based problem solving tasks in which branching and merging behaviors occur? How can diagrammatic representations be tailored to lead reasoners along the right branches, in the right order, and toward the correct conclusions? If returning to previous hypotheses within causal chains is a common phenomenon in diagrammatic reasoning, what implications does it have for the design of animations for teaching students how devices work? These are but a few of the many open questions in need of further research.

Acknowledgments

Financial support for this research was provided by Hitachi Ltd., and the experiments reported here were conducted while



(a) Temporal order

(b) Spatial distribution

Figure 9: Hypothesis sequence generated by subject 5

all three authors were at Hitachi's Advanced Research Laboratory.

References

Cheng, P. C-H. (1994). An empirical investigation of law encoding diagrams for instruction. In *Proceedings of the* 16^{th} *Annual Conference of the Cognitive Science Society* (pp. 171–176). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clement, J. (1994). Imagistic simulation and physical intuition in expert problem solving. In *Proceedings of the* 16^{th} *Annual Conference of the Cognitive Science Society* (pp. 201–206). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ericsson, K. A. & Simon, H. A. (1983). *Protocol Analysis: Verbal Reports as Data.* MIT Press, Cambridge, MA.

Glasgow, J. I., Narayanan, N. H. & Chandrasekaran, B. (Eds.). (1995). *Diagrammatic Reasoning: Cognitive and Computational Perspectives.* Menlo Park, CA: AAAI Press and Boston, MA: MIT Press.

Hegarty, M. (1992). Mental animation: inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 1084–1102.

Hegarty, M. & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717–742.

Hegarty, M. & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and Cognition*, 22(4), 411–430.

Lindsay, R. K. (1994). Understanding diagrammatic demonstrations. In *Proceedings of the* 16^{th} *Annual Conference of the Cognitive Science Society* (pp. 572–576). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research*, 59, 43–64.

Narayanan, N. H., Suwa, M., & Motoda, H. (1994). A study of diagrammatic reasoning from verbal and gestural data. In *Proceedings of the* 16^{th} *Annual Conference of the Cognitive Science Society* (pp. 652–657). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zhang, J. & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, 18, 87–122.