Causal descriptions and collaborative physical design

G. C. Borchardt
Artificial Intelligence Laboratory
Massachusetts Insitute of Technology
Cambridge MA 02139
USA

Abstract

In the design of physical artifacts, written causal descriptions serve as an important medium for communication among members of a design team. This paper describes recent research in getting programs to read and reason on the basis of such descriptions and explores the use of this technology in constructing programs that support collaborative design by (1) cataloging and comparing alternate specifications of physical behavior relating to a designed artifact, and (2) managing an exchange of written causal descriptions among members of a design team.

1 Introduction

This paper describes recent work in getting computers to understand written causal descriptions of physical behavior, with an emphasis on the use of this technology in supporting teams of humans engaged in the design of physical systems such as mechanical or electronic artifacts. This work is described more fully in (Borchardt, 1992a), (Borchardt, 1992b) and (Borchardt, 1993) and involves three related thrusts: (1) a characterization of the problem of understanding written causal descriptions—referred to as the *causal reconstruction* problem, (2) the development of a computational approach to performing this task, involving a new representation for physical behavior called *transition space,* and (3) the demonstration of this approach in a program called PATHFINDER. Several new capabilities related to collaborative design are expected to emerge from the application of this technology, including: the ability of programs to maintain large knowledge bases of specifications of physical behavior for portions of a designed artifact, such that inconsistencies and possible interactions between the behavioral specifications may be identified; and the ability of programs to manage an exchange of causal descriptions between design team members, selectively paraphrasing descriptions or answering questions in response to the comprehension requirements of particular recipients.

*This article describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-91-J-4038.

Human designers exchange causal information in a number of contexts, offering explanations of desired physical behavior for systems being designed, characterizing the behavior of previously-designed systems or modules of these systems, explaining faults in designed systems, explaining desirable or undesirable interactions between modules in a system, and so forth. Causal descriptions serve as the medium for these exchanges. In the context of this research, a "causal description" is taken to be a verbal description—for simplicity, a written description—composed by a human for the purpose of conveying knowledge of the causal workings of a particular physical system to other humans or to a computer program. Such descriptions appear naturally in sources such as encyclopedias, reports and user manuals. In the operation of PATHFINDER, descriptions of this sort are first rendered in simplified English in such a way as to retain most of their original vocabulary, yet exclude complicated syntactic devices, referential devices and so forth.

In the context of collaborative design, the use of simplified English as a medium for computer-assisted exchange of causal knowledge draws a useful compromise between the extent of current technology and the needs of human designers. Unrestricted natural language text is sufficiently opaque to computer processing as to hinder the ability of programs to perform many important functions requiring an understanding of the *content* of a communication. On the other hand, to require humans to enter knowledge in a specialized symbolic representation offers its own problems: the use of symbolic representations is difficult to standardize (whereas language is standardized through daily use by a population), and furthermore, many symbolic representations lack the overall expressiveness of language regarding the range of physical behavior comprehended by humans.

2 Causal Reconstruction and Transition Space

Listed below are a few simple examples of written descriptions of the sort processed by PATHFINDER. (Processing for a more complex description concerning the exposure of film in a camera is described in the above-mentioned references.) A primary difficulty in processing such descriptions involves figuring out precisely how the stated events fit together into larger activities, since written descriptions often do not explicitly include such information.

```
The trigger moves.  The hammer is released.

Object 1 slides.  Object 2 is scratched.  Object 3 is struck.

The steel table is hot.  The copper bar rubs against the steel
table.  The copper bar becomes hot.

The wheel rolls on the concrete.  The wheel is pushed by the axle.
The wheel spins on the axle.  The wheel stops spinning on the axle.
```

Given suitable background information of a generic nature concerning types of objects and events involved and ways of restating activities, PATHFINDER is able to read descriptions such as the above and answer non-trivial questions concerning relationships among events and the temporal sequencing of changes in individual attributes of the participating objects. As an example, for the first description above, some of PATHFINDER's responses to questions are as follows:

```
How could the trigger moving cause the hammer to hit the firing pin?

     The trigger moving could cause the trigger to unlatch the
     hammer, which could cause the releasing of the hammer,
     which could cause the hammer to hit the firing pin.

What happens to the restraint of the hammer by the trigger?

     First, as the trigger moves, the restraint of the hammer by
     the trigger does not disappear.  Next, as the hammer is
     released, the restraint of the hammer by the trigger
     disappears.  Next, the restraint of the hammer by the
     trigger does not appear.
```

Central to PATHFINDER's processing of causal descriptions is its ability to recognize unstated associations between events, this accomplished by a heuristically-guided matching process that looks for overlaps among the sets of individual changes implied by various events. The basis for this matching process is a representation called *transition space,* which depicts physical events in terms of *transitions,* or sets of changes expressible in everyday language. The use of language as a grounding for the representation of changes distinguishes this approach from related representations employed in qualitative reasoning (Forbus, 1984) (de Kleer and Brown, 1984) (Kuipers, 1986). Drawing on psychological research regarding language and perception (e.g., (Miller and Johnson-Laird, 1976)), it is possible to capture a wide range of verbally-expressible changes using a few simple varieties of sentences. The following are examples of such sentences (attributes appear in boldface, while characterizations of change appear in italics):

The **contact** between the bolt and the plate *appears.*
The **position** of the trigger *changes.*
The **temperature** of the bearing *does not increase.*
The rod *becomes* **bent.**
The surface *remains* **sticky.**
The structure *becomes* **covered** by the extinguishing foam.

Generic definitions for physical events are supplied to PATHFINDER using structured combinations of such statements, as in the following definition for the event "to unlatch":

```
Object 21 unlatching object 22 translates to the following event.
First, the position of object 21 changes, the speed of object 21
does not disappear, the heading of object 21 does not disappear, the
distance between object 21 and object 22 does not appear, the
contact between object 21 and object 22 does not disappear, and the
restraint of object 22 by object 21 does not disappear.  Next, the
position of object 21 changes, the speed of object 21 does not
disappear, the heading of object 21 does not disappear, the distance
between object 21 and object 22 appears, the contact between object
21 and object 22 disappears, and the restraint of object 22 by
object 21 disappears.
```

The transition space representation encodes individual statements of change as predicate logic assertions, using a set of ten predicates covering a range of change possibilities for boolean, qualitative and quantitative attributes.[2] These ten predicates are listed below, accompanied by graphic symbols used in diagrams depicting the representations for events.

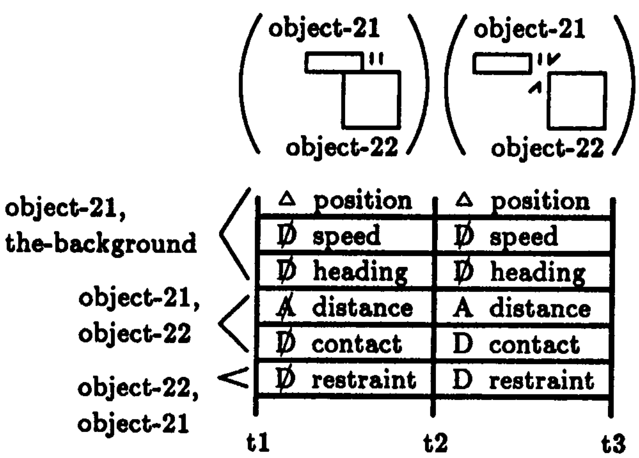
— (presence versus absence) —

for boolean attributes	[A]	APPEAR	[A̸]	NOT-APPEAR
	[D]	DISAPPEAR	[D̸]	NOT-DISAPPEAR

— (specializations of NOT-DISAPPEAR) —

for qualitative attributes	[△]	CHANGE	[△̸]	NOT-CHANGE
for quantitative attributes	[+]	INCREASE	[+̸]	NOT-INCREASE
	[−]	DECREASE	[−̸]	NOT-DECREASE

The above definition for the event "to unlatch" produces the following transition space encoding, as depicted in graphic form. The event consists of a sequence of two transitions ("First, ..." versus "Next, ..." in the above definition), each depicted as a column of entries specifying individual changes for attributes. Relevant objects and time points for the individual changes are depicted to the left and below, and a drawing of the activity (for human inspection only) appears at the top. In this event, "object 21" first moves while maintaining contact with and restraint of "object 22." Then, "object 21" continues moving while the contact and restraint disappear.



Object 21 unlatches object 22.

PATHFINDER recognizes unstated associations between various events in a causal description by looking for partial matches—possible overlaps—among the transition space representations for those events, also considering partial matches with various precedent events supplied to the program. Several heuristics are used to choose among competing

[2]The precise form of these logical assertions is described in the above-mentioned references.

partial matches identified by the program. In processing the above-listed description "The trigger moves. The hammer is released.", these heuristics identify partial matches with respect to the unlatching event described above, such that the moving trigger is drawn into correspondence with the moving object of the unlatching event, and the hammer in the releasing event is drawn into correspondence with the remaining object of the unlatching event. This association is then used as a basis for answering questions, as evidenced in the above-listed answer to the question "How could the trigger moving cause the hammer to hit the firing pin?".

While matching between event representations forms the basis for PATHFINDER's enumeration of associations between events—this ultimately leading to its capability to answer questions such as illustrated above—there are also a number of additional components in the comprehension process which increase its overall effectiveness in meeting special circumstances. Inference is used in two capacities: to augment the sets of assertions included within event representations and thereby provide a broader basis for matching with other event representations, and as a means of checking matches between events for logical consistency. Secondly, various transformations are applied to the event representations to generate alternate representations at different levels of abstraction or in terms of different underlying metaphors. These alternate representations, when matched with the representations of other events in a description, serve to bridge discontinuities arising from the writer's use of abstraction or analogy within a description. Thirdly, explicit declarations of inter-event associations appearing in a description are used to constrain the overall matching process.

3 Opportunities for Collaborative Design

In the operation of PATHFINDER, simplified English plays two distinct roles. At a lower level, it provides a grounding for the transition space representation, enabling generic definitions for events to be supplied to the program (e.g., the above definition for unlatching). At a higher level, it serves as the medium by which causal descriptions are presented to the program and a question/answer cycle is performed to test comprehension. These two uses of simplified English in PATHFINDER translate to two distinct areas of application for this technology in the support of collaborative design, as described below.

3.1 Managing Multiple Specifications of Designed Behavior

First, putting aside the question of how a program might assist in collaborative design by managing an exchange of actual *causal descriptions* as processed by PATHFINDER, we might first envision a simpler system whereby a program merely keeps track of many transition space accounts of relevant behaviors for a designed artifact, these entered either graphically in a format resembling the above illustration for the unlatching event, or in a written form resembling the preceding written definition for the unlatching event. For this scenario, we might assume that a standard vocabulary of object types and attributes has been previously settled upon, so that different specifications of behavior may be compared directly through matching. By incorporating a matcher of the sort used in PATHFINDER, such a program could perform consistency checks between alternate specifications of the same behavior, as well as identify new associations between specifications of different behaviors (e.g., between an existing behavior for part of the artifact and an undesirable faulty behavior conceivably continuing from that point). Following is a list of four dimensions along which alternate specifications of behavior might be distinguished in such a system. For each dimension, important opportunities exist for the application of matching techniques such as used in PATHFINDER.

Different hypothetical instantiations of a given behavior. Different behavioral specifications might be maintained by the system in order to model a particular physical behavior in its intended form (desired behavior), in various alternate and undesirable forms (faulty behaviors), as indicated by an engineering model of the artifact (predicted behavior), and as actually found in a prototype version of the artifact (observed behavior). Of course, different design team members might be responsible for entry of each type of specification. Through matching, the program could be expected to identify discrepancies between predicted behavior and desired behavior, or between observed behavior and predicted behavior, and additionally, the program could be expected to identify cases where faulty behaviors constitute portions or conceivable continuations of predicted or observed behaviors.

Specifications associated with the past, present, or future. By maintaining a knowledge base of behavioral specifications associated with previously designed artifacts and their components, the system might assist team members in identifying possible design precedents for a current situation.[3] Additionally, by maintaining and distinguishing between specifications of desired behavior for current versus future versions of an artifact, it should be possible to identify instances in which predicted behavior for the current design anticipates part or all of a future behavior envisioned for the artifact.

Temporal granularity of specifications. By employing a composition table for transition space change characterizations (e.g., an INCREASE followed by a NOT-CHANGE corresponds to an overall characterization of INCREASE), compatibility checks may be made between behavioral specifications at different levels of temporal granularity.

Alternate viewpoints regarding the design. A similar check for compatibility may be made when two subgroups of a design team submit behavioral specifications for an interface between their respective modules of the overall design. Similarly, when a submodule forms part of the external interface for its parent module, the corresponding behavioral specifications at each level in the design may be compared for compatibility.

3.2 Managing an Exchange of Knowledge at the Level of Causal Descriptions

Progressing to the more complex scenario in which a program is set to the task of managing an exchange of causal descriptions between members of a design team, we might envision the above set of opportunities extended to include the following new capabilities.

Paraphrasing of causal descriptions. In PATHFINDER, transformations applied to event representations prior to matching permit the program to bridge a range of discontinuities arising from a writer's use of analogy and abstraction. A modification

[3] Ruecker (1992) has been investigating the use of the transition space representation as one component in a design documentation system supporting this sort of precedent retrieval.

of this technique may enable a program to construct and evaluate the comprehensibility of *new* descriptions employing abstraction and analogy. Given background knowledge regarding the information requirements of different participants in the design process, this capability could be used to tailor specific descriptions to specific recipients—summarizing, elaborating or paraphrasing behaviors in terms of analogies where suited to the needs of those recipients.

Question answering. Similarly, a question answering capability of the sort appearing in PATHFINDER could be used to provide an additional degree of flexibility in tailoring the transfer of causal knowledge to the needs of particular recipients. For instance, given a detailed causal description of a particular physical behavior as provided by one member of a design team, other members could then elect to be presented with a short summary of the supplied description, this summary to be selectively elaborated where desired through question answering capabilities provided by the managing program.

Support for hybrid man-machine design environments. Given a program interface supporting an exchange of causal descriptions between human designers, a further step could be taken in extending the design environment to include specialized machine agents targeting subtasks such as enumeration of alternative designs for particular submodules, or critiquing of designs offered by various team members. A possible approach to constructing this extended environment would have machine agents provided with an ability to communicate via causal descriptions and questions, this shielding other agents—both human and machine—from the intricacies of particular representations and reasoning algorithms employed by these agents.

References

Borchardt, G. C. (1992a). *Causal Reconstruction: Understanding Causal Descriptions of Physical Systems,* Ph.D. Dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA.

Borchardt, G. C. (1992b). Understanding causal descriptions of physical systems, *Proc. AAAI Tenth National Conference on Artificial Intelligence,* San Jose, CA: 2–8.

Borchardt, G. C. (1993). Causal reconstruction, A.I. Memo 1403, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

de Kleer, J. and Brown, J. S. (1984). A qualitative physics based on confluences, *Artificial Intelligence* **24**:7–83.

Forbus, K. D. (1984). Qualitative process theory, *Artificial Intelligence* **24**:85–168.

Kuipers, B. (1986). Qualitative simulation, *Artificial Intelligence* **29**:289–338.

Miller, G. A. and Johnson-Laird, P. N. (1976). *Language and Perception,* Harvard University Press.

Ruecker, L. (1992). *mSIBYL: A Design Documentation and Management System,* Master's Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA.