MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 1403 February, 1993

Causal Reconstruction

Gary C. Borchardt

Abstract

Causal reconstruction is the task of reading a written causal description of a physical behavior, forming an internal model of the described activity, and demonstrating comprehension through question answering. This task is difficult because written descriptions often do not specify exactly how referenced events fit together. This article (1) characterizes the causal reconstruction problem, (2) presents a representation called *transition space,* which portrays events in terms of "transitions," or collections of changes expressible in everyday language, and (3) describes a program called PATHFINDER, which uses the transition space representation to perform causal reconstruction on simplified English descriptions of physical activity. PATHFINDER works by identifying partial matches between the representations of events and using these matches to form causal chains, fill causal gaps, and merge overlapping accounts of activity. By applying transformations to events prior to matching, PATHFINDER is also able to handle a range of discontinuities arising from a writer's use of analogy or abstraction.

Copyright © Massachusetts Institute of Technology, 1993

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.

1 Introduction

Humans often learn about the causal workings of physical systems by reading written descriptions of the sort appearing in encyclopedias, reports and user manuals. This article presents research in getting programs to read and reason on the basis of such descriptions [7, 8].

This task is both important and difficult. It is important because written causal descriptions constitute not only an abundant resource for use in constructing intelligent user manuals, design documentation systems and planning/diagnosis systems, but also a convenient medium for interaction with these systems during operation. Moreover, these descriptions cover a wide range of phenomena humans find difficult to describe by other means: complex interactions such as combustion and phase changes, intuitive concepts such as sounds, paths and collections, and metaphorically-modeled activities such as radio signals "spreading" in space.

The task is difficult because written causal descriptions rely heavily on the reader's ability to supply missing objects and events, identify relations between parts of the description, and perform inference from the information provided. One especially difficult subtask is that of determining how the events in a description fit together into causal chains or overlapping accounts of activity, as often these relationships are left implicit by the description.

As an example, consider the following excerpt taken from the opening paragraph of the entry for "camera" in the *Encyclopedia Americana* [21]:

> **CAMERA.** The basic function of a camera is to record a permanent image on a piece of film. When light enters a camera, it passes through a lens and converges on the film. It forms a latent image on the film by chemically altering the silver halides contained in the film emulsion.

Given this description, a human previously unfamiliar with the operation of a camera should be able to answer non-trivial questions such as the following:

"What happens to the distance between the light and the film?" (. . . This distance decreases, then disappears as the light converges on the film.)

"How does the light 'converging on the film' relate to the light 'forming the image on the film?'" (. . . The former causes the chemical alteration of the silver halides, which change appearance, thus constituting the latter.)

"How could a building reflecting light into the camera cause the light to converge on the film?" (...This event ends with light entering the camera, from which it passes through the lens and converges on the film.)

"Does the light come into contact with the film emulsion?" (...Yes. The light contacts the silver halides while chemically altering them, and as these are a part of the emulsion, the light must contact the emulsion.)

To answer these questions, the reader must identify unstated associations among the referenced events and use these associations to assemble a unified account of the activity. The first question requires knowledge of the temporal sequencing and overlapping of events, so that the progression of changes in a particular attribute can be recounted. The second question requires a summarization of the indirect relationship between two events. The third and fourth questions introduce new events to be associated causally with existing events or as a re-description of part of the activity.

Loosely, we may define the *causal reconstruction* problem as follows. As input, we are given a written causal description, taken to be a body of text composed by a human for the purpose of conveying knowledge of the causal workings of a particular physical system to another human or computer program. Within the causal description, several types of statements may appear. (The next section presents a simplified account of causal descriptions as involving three types of statements: references to the occurrence of particular events, static background statements, and "connecting statements" describing general relationships between referenced events.) Given this input, the task is to form an internal model of the described activity—this model not directly accessible externally—and use the model to demonstrate comprehension of the input description through question answering.

What resources are available to the reader in completing this task? First, there are several explicit features of causal descriptions that can provide assistance in causal reconstruction. Considering only the event references, narrative ordering can be taken as a clue to chronological or even causal ordering of events. In the following simple description, narrative ordering correctly indicates a causal relationship between the first and second events.

> The steam rises. The steam contacts the metal plate.

Similarly, roles played by the objects in the events provide clues to inter-event relationships. In the example below, a causal chain from the first event to the second event is suggested by both narrative ordering and the fact that "the staple" appears as the patient in the first event and then as the agent in the second event.

> The metal tab presses against the staple. The staple pierces the paper.

On the other hand, there are also many cases where these simple heuristics do not work, as illustrated by the following two examples:

> The hand holds the bolt. The bolt remains between the first finger and the second finger.

> The metal table melts the ice cube. The ice cube lands on the metal table.

In the first example, the events do not form a chain as suggested by both narrative ordering and object role-playing; rather, the events are concurrent, with the second merely restating part of the activity of the first. In the second example, the narrative ordering heuristic would suggest a causal chain from the first event to the second. In this case, however, the only plausible interpretation is that the second event, landing on the metal table, has caused the first event, melting of the ice cube.

Another source of assistance to the reader are explicit declarations of inter-event relationships appearing in a description ("connecting statements," as described in the next section). These statements assert, for example, that one event causes another, precedes another, or is a part of another.

While such declarations do provide assistance in causal reconstruction, they do not provide so much help as to trivialize the task. First, these statements simply do not appear with enough regularity in written descriptions to provide dependable indications of inter-event relationships. Indeed, in some cases we could require on the order of N^2 such assertions for a set of N events, since by knowing that event A is a part of event C and that event B is also a part of event C, we still do not know how events A and B relate to one another.

There is a more fundamental shortcoming of explicit declarations of inter-event associations, however, and this shortcoming applies also to the narrative ordering and object role-playing heuristics in those cases where they provide useful clues. The problem is that all of these devices provide only *general* indications of how particular events are related to one another. In the above situation involving events A, B and C, what we really need to know is which parts of event C correspond to each of the events A and B. Similarly, if we are told that event I causes event J, we still do not know what part of I leads to the initiation of J—is it the middle of I? The end of I? We need to know this in order to answer questions about the time sequence of changes for particular attributes, such as the first question listed above for the camera description. As a result, the reader must take recourse in his or her knowledge of *what happens* during particular types of events in order to work out specific inter-event relationships consistent with the general framework of indications provided by devices such as connecting statements, narrative ordering and object role-playing.

The core of the approach advanced in this article is a representation called *transition space,* capturing knowledge about what happens during physical events. Two important insights are embodied in this approach. First, by representing events primarily in terms of transitions—or sets of changes occurring during the temporal unfolding of events—a wide range of associations between events may be recognized by looking for partial matches, or overlaps, between the representations of individual events. The use of *changes* as a basis for matching draws motivation from two general sources:

- Research in psychology characterizes perceived causality as an association between consecutive changes in a scene—as, for example, when a person observes an ongoing sequence of physical activity [46, 47].

- When describing activity—especially where qualitative changes are involved—it is often necessary to refer explicitly to changes rather than their component states as causal antecedents and consequences. For example, if two objects come into contact and one breaks, it is not the state of *being* in contact that causes the breakage, but rather a *transition* from non-contact to contact. Similarly, the causal effect is not one of *being* broken, but *becoming* broken.

The second insight is that transition-based representations of events are easy to generate from simple, stylized verbal accounts of what happens during these events. This would seem to be because language inherently stresses information about transitions when used to describe the temporal unfolding of physical activity.

Use of the transition space representation for performing causal reconstruction has been tested in PATHFINDER, a 20,000 line program coded in Common Lisp and run on a Symbolics 3640 Lisp Machine. The primary component of this program is a set of utilities for representing, matching and conducting inference and transformations on events in transition space. Additional components include a parser operating on a simple context-free skeleton of English grammar, a simple language generation capability, and a set of supervisory routines for conducting causal reconstruction. PATHFINDER has been applied to over 60 causal descriptions, including a simplified version of the camera description appearing above. The remaining descriptions processed by PATHFINDER involve mostly 2–4 events each and span a wide range of physical domains including: interaction between solid objects and liquids, condensation and melting, combustion, radio signals, light, chemical reactions and electric currents.

Sections 2, 3 and 4 present the three principal components of this research: the causal reconstruction problem, the transition space representation, and the PATHFINDER program. Related literature is discussed in Section 5, and conclusions of the research appear in Section 6.

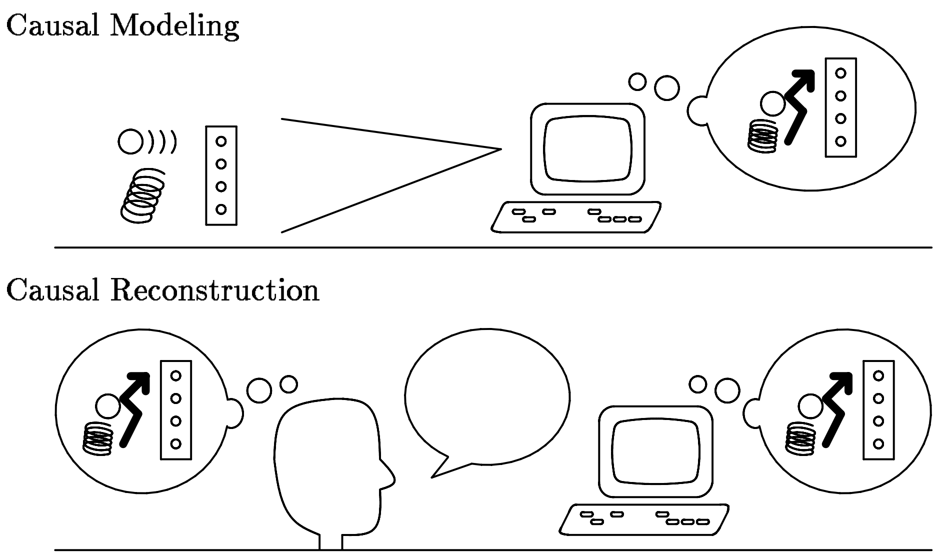


Figure 1: Causal modeling versus causal reconstruction.

2 Causal Reconstruction

The previous section defined causal reconstruction as the task of reading a written causal description, forming an internal model of the described activity, and demonstrating comprehension through question answering. This section characterizes the task more precisely.

Causal reconstruction is related to causal modeling [17, 53] in that both tasks involve *acquisition* of knowledge needed to perform causal reasoning. The tasks differ with respect to the source of this knowledge. Figure 1 illustrates the two tasks schematically. In causal modeling, causal knowledge is obtained from direct observation of the environment. In particular, the input data has not been organized or filtered by an intermediate agent prior to entry into the program. In contrast, causal reconstruction involves input expressly supplied by an intermediate agent already possessing causal knowledge of a situation. While the research presented here concerns textual input only, in general, the input might vary between text, diagrams, animations, numerical data, equations, and other types of information.

The intermediate agent present in the causal reconstruction task affects the nature of this task profoundly in two ways. First, this agent is assumed to be committed to the objective of efficiently communicating the desired causal information. Thus, we can characterize that agent as a "cooperative" participant in the communication process and expect the agent to comply where possible with conversational maxims as outlined by Grice [27]:[1]

The maxim of quantity. Provide as much information as is required for the current purposes of the exchange, but do not provide more information than is required.

The maxim of quality. Try to make your contribution one that is true.

The maxim of relation. Be relevant.

The maxim of manner. Be perspicuous. (Avoid obscurity of expression, avoid ambiguity, be brief and be orderly.)

From the comprehender's perspective, these constraints sanction inferences in certain cases. For example, from the maxim of relation, the comprehender may assume that since a piece of information has been included in the input text, it is indeed relevant and must be related to some other piece of information in the input text.

The second consequence of inserting an intermediate agent in the process is that in effect, the objective for the comprehender is not one of modeling a causal system in the world—for that matter, the input description could conceivably concern a fictitious system—but rather to replicate a causal model known to the intermediate agent (hence, the choice of the term "reconstruction"). This in turn affects what may be taken as a successful completion of the task.

Assuming for the moment that we have at our disposal an effective means of inspecting the comprehender's causal model, we might as a first approximation require that the comprehender's model match precisely the causal model held by the agent composing the description. But we can easily see that this is too restrictive. For example, if the described system is a mechanical one, the comprehender's model could certainly be at a different scale in terms of physical size, or oriented differently in space, or it could contain parts having different shapes or composition than the corresponding parts in the writer's model. A better criterion is to insist that the new causal model be consistent with the writer's causal model to the extent of the information appearing in the description. Another way of saying this is that the comprehender's causal model *must also be describable using the input description.* Viewed in this way, we may make a second appeal to Grice's Maxims of Conversation, this time in the context of the input description taken as an account of the comprehender's newly-constructed causal model. If, by these maxims, we find that the original description is unacceptable as an account of this new model, we may then claim that the new model is faulty and that the comprehender has failed at the task of causal reconstruction. Specifically, Grice's

[1] Although these constraints are advanced with respect to spoken interaction, we would expect writers of causal descriptions to obey such constraints as well.

Maxims motivate use of the following criteria when evaluating the comprehender's understanding of the described activity.

From the maxim of quantity:

1.) Does the comprehender's model introduce new objects or events not motivated in the causal description? (If so, the comprehender's model is unacceptable, because the description is required to be adequately informative in motivating such objects or events.)

From the maxim of quality:

2.) Does the new model disagree in any way with the description? (If so, the new model is unacceptable, because the description is required to be truthful.)

3.) Is the new model physically unrealizable? (If so, it is unacceptable, because we take the description to provide an account of a physically realizable behavior.)

From the maxim of relation:

4.) Does the new model fail to incorporate any information supplied in the description? (If so, the new model is unacceptable, because the description is required to contain only relevant information.)

5.) Does the new model fail to associate any pair of events? (If so, we may disqualify the new model on the grounds that a causal description is expected to relate—causally or otherwise—a set of events.)

From the maxim of manner:

6.) Does the new model make any component of the description redundant, such that the description could be condensed without loss of informativeness? (Suppose the new model interprets two referenced events as describing the same identical activity. It may then be rejected on the grounds that the description is required to be brief and may not contain redundant statements.)

Human evaluation is inherent in assessing success at causal reconstruction, since causal descriptions are at present only identifiable and understandable to humans. The criteria listed above help transform this overall assessment into a set of specific judgments that are less open to dispute.

Returning to the issue of how the comprehender's model might be inspected, we see a further opportunity for increasing the precision of the evaluation process. By requiring a demonstration on the part of the comprehender, we can better assess the actual *value* of the comprehender's newly gained knowledge, as it may be argued that new facts in a target representation have value only relative to the agent's ability to put that knowledge to use. In this article, we consider such demonstration through question answering; however, in a broader sense, we might take the demonstration to include activities such as paraphrasing the input causal description, acting out or constructing an animation of the described scenario.

A final precaution relevant to the assessment process concerns the standardization of conditions under which the entire causal reconstruction process takes place. Motivating this standardization is the fact that Grice's Maxims depend on the intended audience for an utterance; for example, in describing a situation to a child, one must include more information. For automated comprehension of causal descriptions, it is entirely unclear what supporting knowledge a program should have available to it for use in comprehension. For simplicity, then, we assume an absence of relevant background knowledge on the part of the program. As a result, we require definitions for events, static properties of objects, rules of inference and so forth to accompany an input description. This simplifies the task, but by no means trivializes it. Given the pieces of a puzzle, the program must still determine how these pieces fit together.

2.1 Task Restrictions for PATHFINDER

In constructing the PATHFINDER program, several simplifications to the general causal reconstruction task were required. This section outlines these simplifications. First, PATHFINDER sidesteps a number of difficult issues in natural language processing, accepting input descriptions consisting solely of three types of statements:

event references, noting the occurrence of specific events (e.g., "The rocket pushes away the jet of exhaust." or "The iron bar rusts."),

background statements, describing static properties and relationships that hold for the duration of the described activity (e.g., "The liquid is combustible." or "The hub is a part of the wheel."), and

connecting statements, specifying explicit relationships between events appearing in the description (e.g., "The bar pushing on the beam causes the beam to move." or "The component moving is a part of the structure expanding.").[2]

In addition, Grice's maxims are raised to the level of *requirements* for the writer, rather than admonitions open to compromise in delicate communication contexts, as originally described by Grice.

[2]In naturally-occurring causal descriptions, devices such as causal and temporal connectives serve in this capacity.

In compliance with the maxim of quantity relative to the program's assumed blank knowledge base, the following components of supplementary information are also accepted by PATHFINDER as part of an input specification. Examples of these constructs appear in Section 4.1:

additional background statements, supplying object types and static relationships between objects,

event definitions, describing generic occurrences of events,

precedent events, possibly required for comprehension,

rules of inference, as relevant to the activity, and

rules of restatement, relating alternate descriptions of the same activity.

Regarding the product of causal reconstruction, PATHFINDER is targeted to demonstrate comprehension by answering four types of questions. These questions relate to two levels of comprehension applicable to causal descriptions:

the event level, concerning properties, relationships and other attributes of *events* occurring in the described situation, and

the object level, concerning properties, relationships and other attributes of *objects* involved in those events, and how these change during the unfolding of events.

The four types of questions fielded by PATHFINDER are listed below. Questions of type 1 address comprehension at the object level, and the remaining types address comprehension at the event level:

Type 1. Questions concerning the time-course of changes in particular attributes of objects. For example:

> What happens to the position of the jet of exhaust?
> When is the light inside the camera?

Type 2. Questions concerning relationships between particular events referenced in the input description. For example:

> How does the light entering the camera relate to the light passing through the lens?
> How does the transmitting of the radio wave into space relate to the strength of the radio wave decreasing?

Type 3. Questions concerning possible causal relationships between events referenced in the description and new events introduced at the time of questioning. Such questions may be used to explore simple predictions or explanations beyond the immediate context of a causal description. For example:

> How could the water boiling cause the steam to condense on the metal plate?
> How could the trigger moving cause the hammer to hit the firing pin?

Type 4. Questions concerning the possible paraphrasing of a portion of the activity in terms of a newly-supplied event. For example:

> Does the steam convert to a liquid?
> Is the wheel pushed?

3 Transition Space

Section 1 argued that in order to successfully perform causal reconstruction, a program must have knowledge of what happens during particular types of physical events. Since the program is replicating a human skill, a good source of constraint for the representation of what happens during events is the psychological literature regarding events. Several useful guidelines may be extracted from this literature, as outlined below.[3]

Representation in terms of objects and attributes. "Objects" may be taken to include any quantities whose properties, relationships or other attributes may be described; e.g., physical objects, spaces, paths, events, systems, and so forth. Attributes may be physical (e.g., distance, size and shape) or conceptual (e.g., being a member of a collection). In [47], Miller and Johnson-Laird present a good summary of objects and attributes used by humans in describing the world. These authors characterize attributes as typically unary (e.g, size) or binary (e.g., distance), and also as either qualitative (e.g., color or contact) or quantitative (e.g., length or pressure).

Time as a sequence of moments. Investigation by Newtson and his colleagues [50, 51] indicates that time is perceived—at least at the level of conscious awareness—as a sequence of discrete moments rather than a continuum of activity. Further research by these investigators indicates that events are delimited by specific "breakpoints"—time points at which significant changes are perceived by the observer.

Qualitative comparisons and changes. Following from the superiority of humans at relative—as opposed to absolute—estimation of attribute values, it is natural to represent both static comparisons and dynamic changes in a qualitative manner [47].

[3]Several of these guidelines also serve to motivate work in qualitative physics [15, 22, 37].

Causation as an association between changes. Experimental evidence characterizes perceived causality as an association between consecutive changes in a scene [46]. Concurring accounts appear in the cognitive development literature [43, 63] and elsewhere in psychology [47].

The final point above is significant. While a number of representations include *states* as possible antecedents or consequents of causality, it would seem that *changes* are most suitable in the perceptually-oriented realm of physical activity.

Intuition would seem to concur with this characterization of causality. Regarding causal antecedents, if states are to do the causing, then we might ask why a causal effect occurs precisely when it does and not earlier; thus, we are led to suspect that some additional ingredient has fallen into place just prior to the causal effect. This final change may then be ascribed as the antecedent of causality. Regarding causal effects, if there is no change in a scene, then causation contributes nothing to the reasoning process; we can reason just as well with what we knew to be true beforehand.

Certainly, representations permitting states as causal antecedents and consequents have utility. For example, we may usefully speak of a "state" of low blood sugar resulting in various effects if we choose a level of abstraction factoring out ongoing circulatory and metabolic processes in the body. In other cases, we may incorporate dynamic information into the specification of states, as when we refer to "moving" objects or characterize instantaneous changes in quantities as "increasing," "decreasing" or "steady," as in the qualitative physics literature.

However, such techniques are not always applicable. In particular, instantaneous directions of change can be specified only for quantitative attributes, like "temperature" or "elevation," which can be differentiated with respect to time, but not for qualitative attributes, like "contact," "support," and "inside," as often appear in verbal accounts of activity. For example, if two objects come into contact, with one object breaking as a result, there is no "state of coming into contact" to serve as the causal antecedent. At one instant the objects are not in contact, and this certainly does not cause the breakage, and at a subsequent instant they are in contact—but this cannot be attributed as the antecedent either, because they could have been simply resting, in contact, for a long time. A better characterization identifies the causal antecedent as the change from non-contact to contact between the two objects.[4] Likewise, the causal consequent in this example is more suitably characterized as a change from unbroken to broken for the second object, rather than simply a state of being broken.

Representing causality as an association between changes—where these changes

[4] Additionally, of course, there are other less articulable factors contributing to the breakage: sufficient momentum on the part of the first object, sufficient brittleness for the second object, sufficient inelasticity of the collision, and so forth.

may involve quantitative or qualitative attributes appearing in verbal accounts of activity—we are led to the notion of transition space. In transition space, individual points—or transitions—correspond to possible combinations of changes in relevant attributes of objects participating in a described causal scenario. Events may be thought of as short paths in this space, corresponding to sequences or simple directed, acyclic graphs of transitions. As such, events comprise simple causal explanations which may be combined to produce larger causal explanations—larger "paths" in transition space—that serve as models of described activities.

Important to the construction of this representation is its basis in verbal accounts of activity. Attributes and their changes are taken directly from simple, stylized English statements. Thus, the representation has not only an established semantics (grounded in that of the English statements on which it is based), but as well, individual assertions in the representation may be converted back into their corresponding verbal form when necessary in order to discuss the suitability of representing particular events in particular ways.

Equally important, it should be noted that in grounding the representation in verbal accounts of activity, those aspects of human knowledge about events that are less easily expressed verbally are omitted. For instance, the representation is not intended to capture spatial knowledge of a non-propositional nature, as used in estimating shapes, directions, textures and so forth. Also, it is not intended to capture knowledge underlying human ability to classify objects and situations from visual or other sensory perceptions, or knowledge underlying human ability to estimate likelihood of various causal circumstances. These types of knowledge are doubtlessly required for processing some descriptions; however, as illustrated in this article, there is also a range of simple descriptions that can be processed solely on the basis of articulable knowledge about what happens during events.

Listed below are several examples of the sorts of statements of change grounding the transition space representation (attributes appear in boldface, characterizations of change appear in italics).

> The **contact** between the steam and the metal plate *appears*.
> The **concentration** of the solution *increases*.
> The **appearance** of the film *changes*.
> The pin *becomes* **a part of** the structure.
> The water *remains* **inside** the tank.

The change characterizations in these expressions are limited in number. Assuming the existence of a "false" or "absent" value in the range of each attribute, then if a particular attribute of one or more objects is specified as either present or absent at each of two time points, one of which follows the other, and we include possible information concerning the qualitative relationship between the attribute values at the two time points, then the following set of ten change characterizations covers the range of alternate circumstances based on this information:

(presence versus absence)

for boolean attributes	APPEAR	NOT-APPEAR
	DISAPPEAR	NOT-DISAPPEAR

(specializations of NOT-DISAPPEAR)

for qualitative attributes	CHANGE	NOT-CHANGE
for quantitative attributes	INCREASE	NOT-INCREASE
	DECREASE	NOT-DECREASE

The ten change characterizations are depicted as predicates taking four arguments: an attribute of concern, an object or tuple specifying multiple objects, a first time point and a second time point. The assertions below correspond to each of the English statements above.

APPEAR(contact, <the-steam, the-metal-plate>, t1, t2)
INCREASE(concentration, the-solution, t3, t4)
CHANGE(appearance, the-film, t5, t6)
APPEAR(a-part-of, <the-pin, the-structure>, t7, t8)
NOT-DISAPPEAR(inside, <the-water, the-tank>, t9, t10)

Two grammatical forms are currently used in PATHFINDER for entering the English statements grounding the transition space representation.[5] The first form is as follows:

<attribute-expression> *<verb-group>*

where

<attribute-expression> ::=
`the` *<attribute>* *<preposition>* *<noun-phrase>*
{ { *<preposition>* | `and` } *<noun-phrase>* }*

and the verb group indicates one of the ten change characterizations. The first three sentences above are of this form. The second form, illustrated in the last

[5] The grammar used by PATHFINDER is a simple context-free semantic grammar [31]. Other grammatical forms for specifying changes could be included in an extended grammar (e.g., specifications such as "The water increases in temperature.").

two sentences above, is applicable only to certain boolean attributes and involves a single object followed by a restricted set of verbs ("becomes," "becomes not," "remains" and "remains not") and a predicate modifier expression:

<object> *<verb-group>* *<predicate-modifier>*

where

<predicate-modifier> ::=
    <attribute> [[*<preposition>*] *<noun-phrase>*
        { { *<preposition>* | `and` } *<noun-phrase>* }*]

A variant of this second form covers the "is-a" attribute and includes statements such as "The water becomes a vapor." and "The solution remains an acid.".

Together, the grammatical forms for entering statements regarding changes cover a wide range of phenomena, as illustrated by the following examples:

The **interaction** between the wrench and the bolt *disappears.*
The **frequency** of the disturbance *decreases.*
The **usefulness** of the device for the task *disappears.*
The surface *remains* **sticky**.
The plate *becomes* **fastened** to the housing.
The structure *becomes* **covered** by the extinguishing foam.
The rod *remains not* **bent**.
The weather vane *becomes* **directed** from the west to the east.
The liquid *remains* **frozen**.
The sponge *becomes* **dry**.

While the ten change characterization predicates of the representation are at an appropriate level of abstraction for association with simple English assertions of change, a number of overlaps exist between these predicates and thus they are unsuitable as primitives in the representation (six of the predicates are specializations of "NOT-DISAPPEAR," and some pairs of these six are compatible while others are not). By constructing definitions for these predicates in terms of two primitive five-place predicates "EQUAL" and "GREATER," plus their negations, the processes of matching and inference on event representations are simplified considerably. The following paragraphs develop the entire representation in a bottom-up manner, from these primitives to the representation of events.

The following five types of symbols appear in the representation:

Objects, both perceptual and conceptual. For example: solids; quantities of liquid, gas, fire, etc.; spaces, surfaces, paths and edges; events and sequences; collections—ultimately, anything that may participate in an event.

Attributes, both perceptual and conceptual. For example: length, width, depth, size, weight and color; position, elevation, orientation, speed, heading, direction, distance; "insideness," pressure, contact, restraint; "is-a," "a-kind-of," "made-of," "part-of," "before" and "after." As in everyday language, a degree of overlap appears in the set of attributes.

Time points, as needed to distinguish points of comparison within events.

Reference standards, used as fixed points of reference for comparison. For example: object types ("solid," "event," "collection," etc.), colors, substances, numbers, qualitative directions such as "up" and "forward," a fixed frame of reference for motion ("the-background"), and a quantity "null" representing the "false" or "absent" state for all attributes.[6]

Predicates, used in assertions comparing attribute values. For qualitative attributes, "EQUAL" and "NOT-EQUAL" serve as primitive predicates. For quantitative attributes, these plus "GREATER" and "NOT-GREATER" serve as primitives.

Reference standards have two distinct functions: they serve as unchanging points of reference for comparisons, and they span different description discourses. As unchanging points of reference, reference standards support certain types of inference. Suppose the color of an object is specified as matching the color "green" at one time point, and is specified as not changing in color over an interval from that time point to a later time point. We would like to conclude that the object is still green; however, this requires an assumption that the color green has not changed during the course of that same interval: this assumption is provided by classifying "green" as a reference standard. The other aspect, spanning of description discourses, distinguishes reference standards from objects, which must be individualized so that their mention in two different discourses serves to generate two distinct symbols, not one (e.g., "the-pin-1" versus "the-pin-2").

The primitive predicates "EQUAL," "NOT-EQUAL," "GREATER" and "NOT-GREATER" take five arguments: an attribute for comparison, a first object and associated time point, and a second object and associated time point. For binary attributes, tuples of objects are used in the second and fourth positions. Additionally, for unusual attributes such as "between," a nesting of tuples is employed. Three simple assertions involving these predicates are given below.

GREATER(length, object-1, t1, object-2, t1)
NOT-EQUAL(position, <object-1, the-background>, t1,
    <object-1, the-background>, t2)
EQUAL(distance, <object-1, object-2>, t1, null, t1)

[6] Reference standards are called "prototypes" in [6].

The first assertion states that "object-1" has greater length than "object-2" at time "t1." The second, that the position of "object-1" relative to "the-background" is different at time "t1" and time "t2" (i.e., "object-1" has moved). The third, utilizing a comparison to the "null" reference standard, states that there is no distance between "object-1" and "object-2" at time "t1."

From the primitives, we first define a set of six predicates for making assertions at a single time point. In these definitions, variables are denoted by labels beginning with "?". Also, note that several of the "NOT-" forms are not strict logical negations: they assume some of the same information as the positive forms, in line with common usage of these terms in language.

PRESENT(?attribute, ?object, ?t1) $\Longleftrightarrow$
  NOT-EQUAL(?attribute, ?object, ?t1, null, ?t1)

NOT-PRESENT(?attribute, ?object, ?t1) $\Longleftrightarrow$
  EQUAL(?attribute, ?object, ?t1, null, ?t1)

MATCH(?attribute, ?object-1, ?object-2, ?t1) $\Longleftrightarrow$
  PRESENT(?attribute, ?object-1, ?t1) AND
  PRESENT(?attribute, ?object-2, ?t1) AND
  EQUAL(?attribute, ?object-1, ?t1, ?object-2, ?t1)

NOT-MATCH(?attribute, ?object-1, ?object-2, ?t1) $\Longleftrightarrow$
  PRESENT(?attribute, ?object-1, ?t1) AND
  PRESENT(?attribute, ?object-2, ?t1) AND
  NOT-EQUAL(?attribute, ?object-1, ?t1, ?object-2, ?t1)

EXCEED(?attribute, ?object-1, ?object-2, ?t1) $\Longleftrightarrow$
  PRESENT(?attribute, ?object-1, ?t1) AND
  PRESENT(?attribute, ?object-2, ?t1) AND
  GREATER(?attribute, ?object-1, ?t1, ?object-2, ?t1)

NOT-EXCEED(?attribute, ?object-1, ?object-2, ?t1) $\Longleftrightarrow$
  PRESENT(?attribute, ?object-1, ?t1) AND
  PRESENT(?attribute, ?object-2, ?t1) AND
  NOT-GREATER(?attribute, ?object-1, ?t1, ?object-2, ?t1)

For comparisons across time points, the following definitions specify the ten change characterizations introduced above. (In these definitions, "null" is also used to designate an irrelevant time point argument.)

APPEAR(?attribute, ?object, ?t1, ?t2) $\Longleftrightarrow$
  PRESENT(after, <?t2, ?t1>, null) AND
  NOT-PRESENT(?attribute, ?object, ?t1) AND
  PRESENT(?attribute, ?object, ?t2)

```
NOT-APPEAR(?attribute, ?object, ?t1, ?t2) ⟺
        PRESENT(after, <?t2, ?t1>, null)  AND
        NOT-PRESENT(?attribute, ?object, ?t1)  AND
        NOT-PRESENT(?attribute, ?object, ?t2)

DISAPPEAR(?attribute, ?object, ?t1, ?t2) ⟺
        PRESENT(after, <?t2, ?t1>, null)  AND
        PRESENT(?attribute, ?object, ?t1)  AND
        NOT-PRESENT(?attribute, ?object, ?t2)

NOT-DISAPPEAR(?attribute, ?object, ?t1, ?t2) ⟺
        PRESENT(after, <?t2, ?t1>, null)  AND
        PRESENT(?attribute, ?object, ?t1)  AND
        PRESENT(?attribute, ?object, ?t2)

CHANGE(?attribute, ?object, ?t1, ?t2) ⟺
        NOT-DISAPPEAR(?attribute, ?object, ?t1, ?t2)  AND
        NOT-EQUAL(?attribute, ?object, ?t1, ?object, ?t2)

NOT-CHANGE(?attribute, ?object, ?t1, ?t2) ⟺
        NOT-DISAPPEAR(?attribute, ?object, ?t1, ?t2)  AND
        EQUAL(?attribute, ?object, ?t1, ?object, ?t2)

INCREASE(?attribute, ?object, ?t1, ?t2) ⟺
        NOT-DISAPPEAR(?attribute, ?object, ?t1, ?t2)  AND
        GREATER(?attribute, ?object, ?t2, ?object, ?t1)

NOT-INCREASE(?attribute, ?object, ?t1, ?t2) ⟺
        NOT-DISAPPEAR(?attribute, ?object, ?t1, ?t2)  AND
        NOT-GREATER(?attribute, ?object, ?t2, ?object, ?t1)

DECREASE(?attribute, ?object, ?t1, ?t2) ⟺
        NOT-DISAPPEAR(?attribute, ?object, ?t1, ?t2)  AND
        GREATER(?attribute, ?object, ?t1, ?object, ?t2)

NOT-DECREASE(?attribute, ?object, ?t1, ?t2) ⟺
        NOT-DISAPPEAR(?attribute, ?object, ?t1, ?t2)  AND
        NOT-GREATER(?attribute, ?object, ?t1, ?object, ?t2)
```

A *transition* is a set of assertions at and between two ordered time points. Events are sequences, or more generally, directed acyclic graphs of transitions. The representations for events are called *event traces*, highlighting their correspondence to simple paths in transition space.

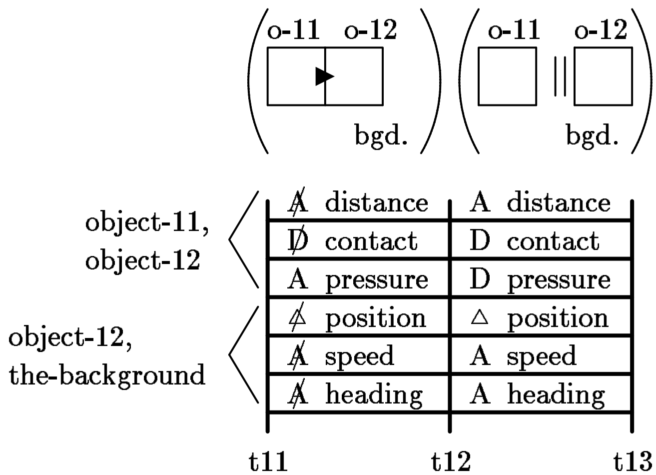


Figure 2: Event trace depicting the event "push away."

Throughout this article, a special graphic format is employed for illustrating transitions and event traces. This format highlights dynamic information, with static information listed simply as assertions beneath the graphic representation if necessary. In these diagrams, the ten change characterizations are coded as follows:

$\boxed{\text{A}}$	APPEAR	$\boxed{\not\text{A}}$	NOT-APPEAR
$\boxed{\text{D}}$	DISAPPEAR	$\boxed{\not\text{D}}$	NOT-DISAPPEAR
$\boxed{\triangle}$	CHANGE	$\boxed{\not\triangle}$	NOT-CHANGE
$\boxed{+}$	INCREASE	$\boxed{\not+}$	NOT-INCREASE
$\boxed{-}$	DECREASE	$\boxed{\not-}$	NOT-DECREASE

Figure 2 illustrates the graphic representation for an event trace depicting the event "push away." This event has two transitions, each specified as a column of coded assertions between vertical lines depicting time points. The time points are labeled at the bottom, and arguments to each attribute are provided at the left. Also, for expository purposes, a drawing is included above each transition: these drawings are not part of the actual representation as used by PATHFINDER.

The first transition of this event trace may be read as follows: the distance between "object-11" and "object-12" does not appear, contact between the two objects does not disappear, pressure between the objects appears, the position of "object-12" does not change relative to "the background," and speed and heading (direction of motion) of "object-12" with respect to the background do not appear. In the second transition, distance appears between the two objects, contact

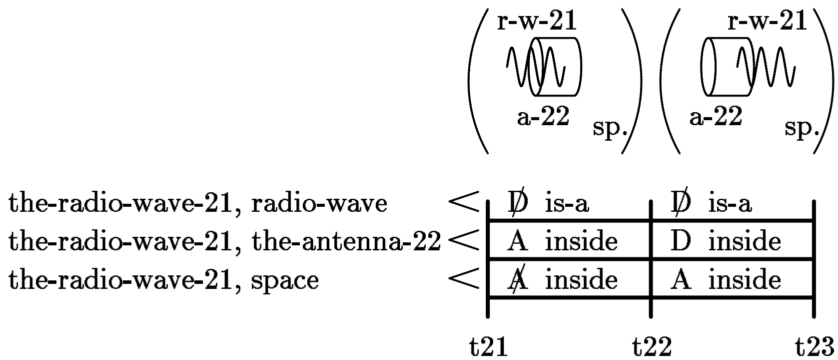


Figure 3: Transmission of a radio wave into space.

and pressure disappear, the position of “object-12” changes with respect to “the background,” and speed and heading appear for “object-12.”

Contrasting with the event trace in Figure 2 is the event trace illustrated in Figure 3, representing the transmission of a radio wave into space. This trace incorporates a conceptual object (the radio wave) and a metaphorically-applied attribute (“inside”). “Space” is represented as a reference standard, as this is a discourse-independent, relatively unchanging quantity.

In general, a particular event may be represented in a range of ways in transition space, depending on such choices as: (1) the selection and granularity of participating objects for the event, (2) the selection and granularity of particular attributes for describing those objects, and (3) the granularity of time points. The transition space representation explicitly provides this latitude and furthermore explicitly addresses the issue of how such alternate descriptions of events can be matched with one another. This is an inherent part of the causal reconstruction task, as the comprehension of causal descriptions inescapably involves a reconciliation of events often described at different levels of abstraction or possibly in terms of different underlying metaphors.

In the other direction, several intuitive guidelines help constrain the use of transition space in representing particular events expressible in language. The following list summarizes these guidelines.

- The granularity of temporal intervals is generally not so fine as to produce adjacent, identical transitions (e.g., each attribute specifying something like an “INCREASE” followed by an “INCREASE”).

- On the other hand, intervals are sufficiently subdivided so as to indicate sets of changes that cause other sets of changes, and the temporal granularity may vary within an event trace. Extremely fine temporal granularity may be used to represent such situations as appearance of pressure from one object to a second leading to return pressure from the second to the first.

- In all cases, we should be able to say that there exists a causal relationship between the set of changes in one transition and the set of changes in an immediately following transition.

- Unless an event depicts non-activity (e.g., "not moving"), the representation should always include at least one definite change (APPEAR, DISAPPEAR, CHANGE, INCREASE or DECREASE) in its opening transition.

- All attributes and assertions needed to discriminate a *typical* instance of an event from other types of events should be included. For example, the representation for "enter" should include not only a first object appearing "inside" a second, but movement for the first as well. Otherwise, the specification would also account for the event "move to surround."[7]

- On the other hand, we may choose to omit particular attributes or change characterizations where we wish to accommodate variations within the set of typical occurrences for an event. For example, when one object hits a second object, information on motion of the first object following the instant of contact may be excluded to allow for possible bouncing, stopping, etc. of that object.

3.1 Matching Between Event Traces

Given a set of event traces for the events in a causal description, simple inter-event associations may be detected by identifying partial matches between the traces. As an example, consider the following simple description:

```
The board is dented.  The wrench is dropped.
```

Note that while this description is certainly understandable—that is, we as humans can determine how the two referenced events might fit together—it is not as directly stated as it could be. For example, the description could have been presented as "The wrench is dropped. The wrench dents the board.". However, supplied in the above-listed form, the description illustrates a simple situation where neither temporal ordering of event references nor role-playing by objects involved in the events can help in determining how the events fit together. That

[7]In the examples run on PATHFINDER, some liberty has been taken with respect to this guideline; for example, object type information is often excluded where an event involves only physical objects.

is, we really need to know what happens during the stated events in order to determine how these two events might be related to one another.

Figure 4 illustrates two event traces corresponding to the "denting" and "dropping" events in the above description. These traces are derived from the generic event definitions appearing below. The definitions are written in stylized English as accepted by PATHFINDER and serve to generate an initial set of pattern event traces which are then mapped to the particular circumstances of the events referenced in the description.[8]

```
Object 1 denting object 2 translates to the following event.
First, the position of object 1 changes, the speed of object 1
does not disappear, the heading of object 1 does not
disappear, the distance between object 1 and object 2
decreases, and the contact between object 1 and object 2 does
not appear.  Next, the position of object 1 changes, the speed
of object 1 disappears, the heading of object 1 disappears,
the distance between object 1 and object 2 disappears, and the
contact between object 1 and object 2 appears.  Next, space 3
becomes a dent, space 3 becomes a part of object 2, and object
1 becomes inside space 3.

Object 11 dropping object 12 translates to the following
event.  First, the distance between object 11 and object 12
appears, the contact between object 11 and object 12
disappears, object 11 becomes not in control of object 12, the
support of object 12 by object 11 disappears, object 12
becomes not supported, the position of object 12 does not
change, the speed of object 12 does not appear, the heading of
object 12 does not appear, and the elevation of object 12 does
not change.  Next, the distance between object 11 and object
12 does not disappear, the contact between object 11 and
object 12 does not appear, object 11 remains not in control of
object 12, the support of object 12 by object 11 does not
appear, object 12 remains not supported, the position of
object 12 changes, the speed of object 12 appears, the heading
of object 12 appears, and the elevation of object 12
decreases.  Next, the distance between object 11 and object 12
does not disappear, the contact between object 11 and object
12 does not appear, object 11 remains not in control of object
12, the support of object 12 by object 11 does not appear,
```

[8]For ease in specifying changes in "position," "speed," "heading," "orientation," "angular speed" and "angular heading," PATHFINDER behaves as if the phrase "with respect to the background" has been included where no second object has been specified as a frame of reference.

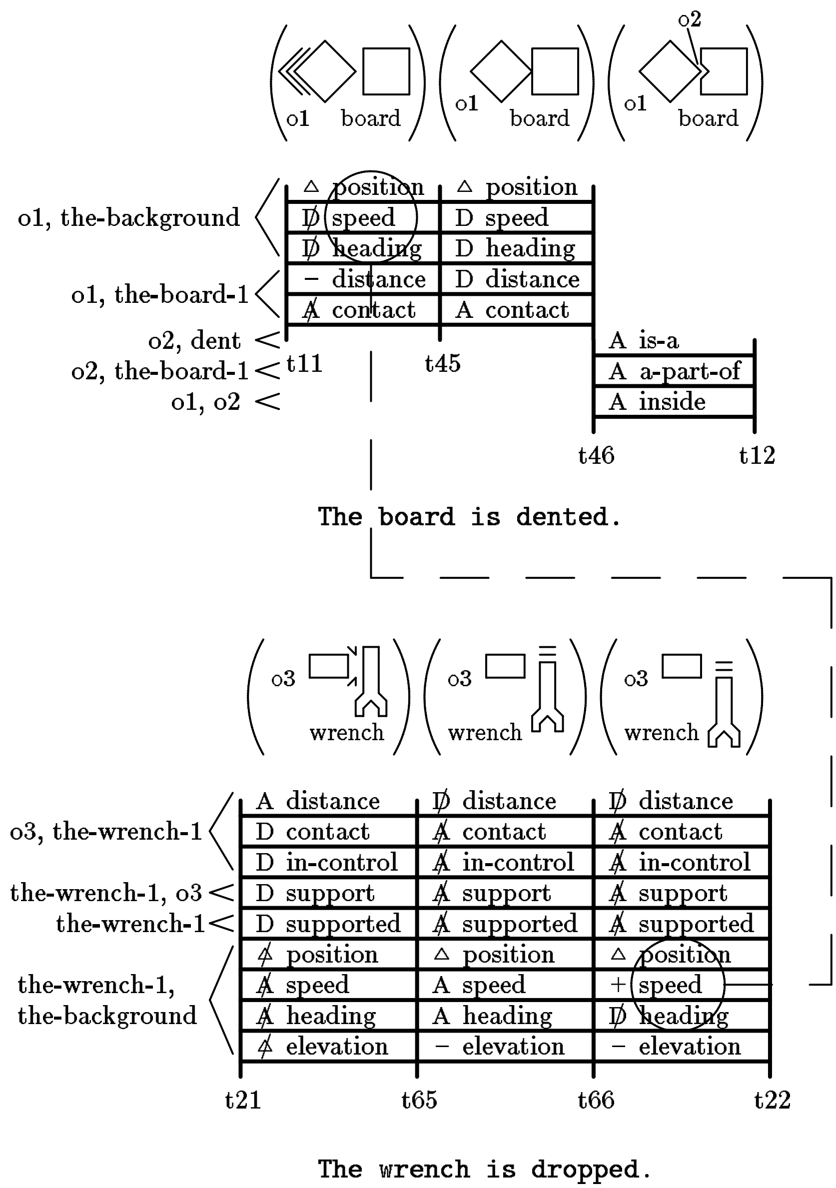


Figure 4: Event traces for the description: "The board is dented. The wrench is dropped." A partial match between the two traces is indicated.

```
object 12 remains not supported, the position of object 12
changes, the speed of object 12 increases, the heading of
object 12 does not disappear, and the elevation of object 12
decreases.
```

In mapping the generic event definition traces to the description-specific versions appearing in Figure 4, several hypothesized objects must be introduced, as both event references in the original description appear in passive voice with no agent specified. (Also, the implicit dent appears as a hypothesized object "o2.") For the "denting" event, a hypothesized object approaches and then contacts the board, followed by the appearance of a dent in the board. For the "dropping" event, a hypothesized object comes out of contact and relinquishes control of the wrench, followed by the wrench starting and continuing to fall.

The supplied event definitions are intended to be generic in nature, not referring to the specific circumstances of the described situation. In particular, the definition for denting—and thus PATHFINDER's depiction of the event "The board is dented."—leaves open the issue of whether the object causing the denting comes to rest at the completion of the denting or continues its motion, acquiring new speed and heading in the interval from "t46" to "t12." Separately, because the dent does not exist prior to "t46," no information is provided concerning it prior to this point.[9]

Given the event traces appearing in Figure 4, the next step is to enumerate partial matches between these traces. Here, we distinguish between two classes of partial matches. In the first class, *partial chaining* matches, one event continues the activity originating in another event. In the second class, *partial restatement* matches, two events match in some other way. Partial chaining matches can be taken as an indication of a plausible causal connection between two events; partial restatement matches simply indicate an overlap in the activity described by two events. More precisely, the following definition is used to distinguish partial chaining from partial restatement matches:

> A *partial chaining match* between two event traces is a single interval match involving at least one definite change (APPEAR, DISAPPEAR, CHANGE, INCREASE or DECREASE) and situated such that exactly one of the traces is begun by the matching interval and continues beyond the matching interval.

In this definition, two extra provisions have been added to the initial characterization appearing in the above text. The first is that the match consist of a

[9]The numbering of time points in PATHFINDER does not always reflect their temporal sequencing. This is an artifact of the implementation.

single interval. Intuitively, it would seem that events overlapping in more than one successive complex of changes are more restatements of one another than causally associated events; however, this rather subjective restriction could alternatively be omitted. The second provision is that at least one definite change be involved in the match. This restriction is included so that events involving no changes—for example, continued support of a block by a table—may not be taken as causal antecedents of other events—for example, the block sliding off the end of the table.

Given the event traces depicted in Figure 4, six partial matches may be identified. All but one of these matches are unacceptable, as they either result in inconsistencies if inference is conducted on the combined assertions of the two events (e.g., one event may state that speed "appears" while the other that it "does not disappear") or equate the distinct objects "the wrench" and "the board." Thus, in this simple example, no comparative ranking of partial matches is required to determine how the events might be related to one another. In the general case, of course, many acceptable partial matches may exist between the various pairs of events referenced in a description. Section 4 describes the heuristics used by PATHFINDER in choosing among alternative partial matches.

The remaining acceptable partial match is illustrated in Figure 4. Note that with regard the attribute "speed" in this match, an assertion involving the predicate INCREASE in the "dropping" trace has been matched with an assertion involving the predicate NOT-DISAPPEAR in the "denting" trace. This particular matching of assertions relies on the fact that the an INCREASE assertion *covers* a NOT-DISAPPEAR assertion (i.e., its expansion at the level of (NOT-)EQUAL and (NOT-)GREATER assertions is a superset of the corresponding expansion for the NOT-DISAPPEAR assertion).

The match identified here is a *partial association,* as the transitions in question do not match completely. The two transitions involved in the match are really distinct points in transition space: they involve distinct time points and objects, and furthermore, each contains additional assertions not appearing in the other. In the interest of performing subsequent question-answering, it is useful to work out in greater detail exactly how the two events are associated with one another. This is accomplished by a process of *elaboration,* involving the application of simple transformations to the two event traces in order to bring the concerned transitions into a complete match. Here, we distinguish between two classes of transformations: *information-preserving* and *non-information-preserving.* Transformations of the first type belong to inverse pairs of transformations on event traces in transition space; transformations of the second type do not.

The choice of transformations applied in the elaboration process is based on information contained in the partial match. This information is presented below for the partial match of concern between the denting and dropping events. The three matching assertions are indicated as they appear in the "denting" activity, with the bindings providing a mapping to equivalent assertions in the context of the "dropping" activity.

match type: partial chaining
first trace: trace-48 ("The wrench is dropped.")
second trace: trace-17 ("The board is dented.")
matched assertions:
CHANGE(position, <o1, the-background>, t11, t45)
NOT-DISAPPEAR(speed, <o1, the-background>, t11, t45)
NOT-DISAPPEAR(heading, <o1, the-background>, t11, t45)
bindings: { [t66 / t11], [t22 / t45], [the-wrench-1 / o1] }

Figure 5 illustrates how the partial match between the original event traces may be elaborated to produce a chain of three transformations and a complete chaining match. Traces A and B are the original "denting" and "dropping" traces, respectively. Trace B is transformed by a *reduction* operation (a non-information-preserving transformation) which removes all assertions except those involved in the partial match or tracking the same attribute-object combinations—"position," "speed," and "heading" of "o1" with respect to "the-background"—for other intervals. This reduction transformation produces trace C. Trace A is first transformed by an *equivalence* operation (an information-preserving transformation), replacing time points "t11" and "t45" with their matched equivalents "t66" and "t22" and replacing the hypothesized object "o1" with its matched equivalent "the-wrench-1." This transformation produces trace D. Trace E is then produced by a reduction operation on trace D, removing all assertions except those concerning "position," "speed," and "heading" of "the-wrench-1" with respect to "the-background," as involved in the partial match. Finally, traces C and E are associated by a complete chaining match, as the third transition in trace C is identical to the first transition in trace E.

A set of event traces linked by *complete associations*—transformations and complete chaining or restatement matches—is called an *association structure*. Association structures can be diagramed using a second, more abstract graphic format corresponding to a stylized three-dimensional characterization of transition space. Figure 6 illustrates this format for the sequence of associations indicated in Figure 5. Here, transitions appear as points, with event traces depicted as arrows, or more generally as directed acyclic graphs. Heavy arrows depict original traces for events referenced in a description; lighter arrows depict intermediate traces formed in elaboration of a partial match. Finally, associations are represented by alignment of the event traces in three dimensions: horizontal for complete chaining matches, vertical for non-information-preserving transformations, and depthwise for information-preserving transformations (including complete re-

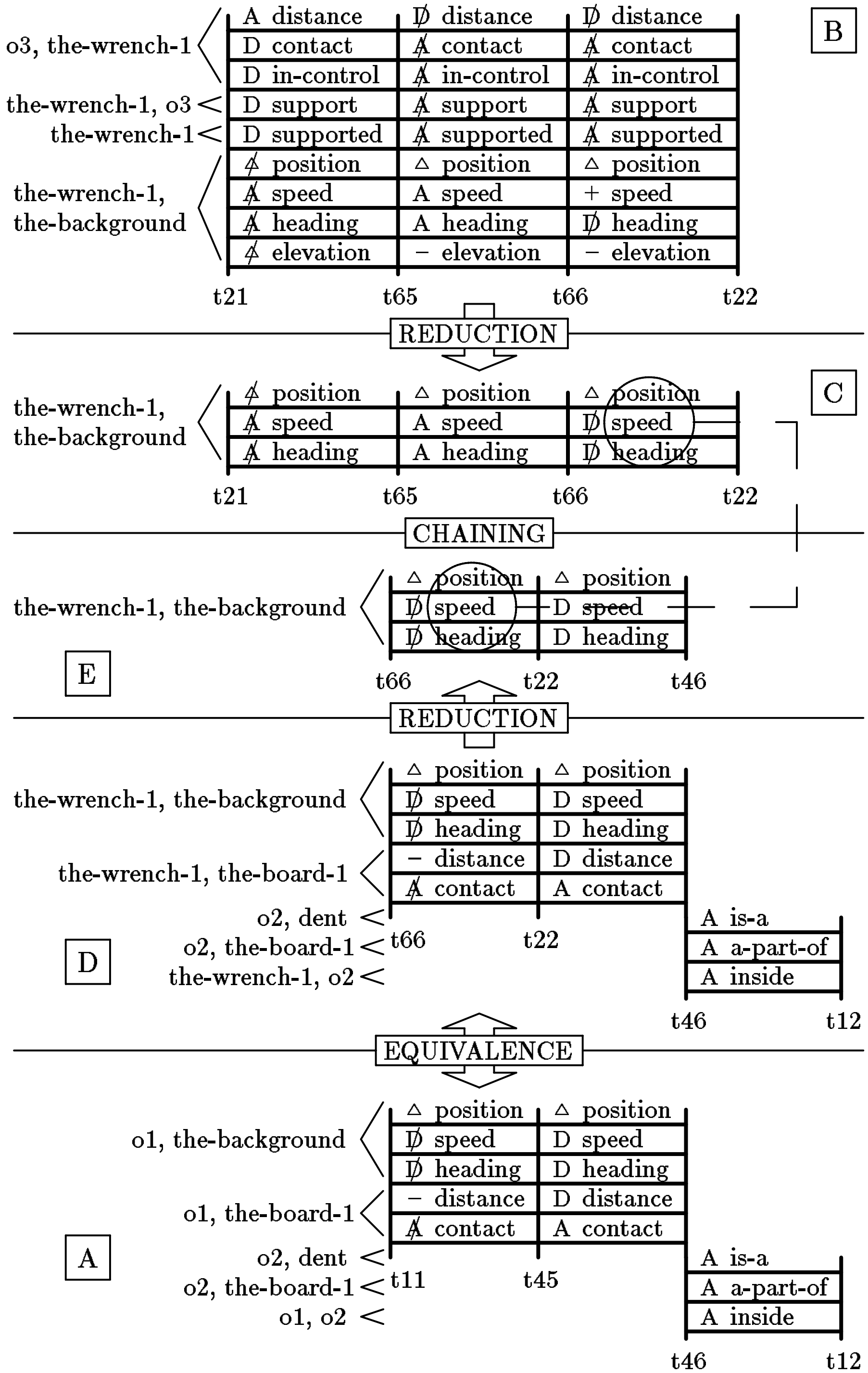
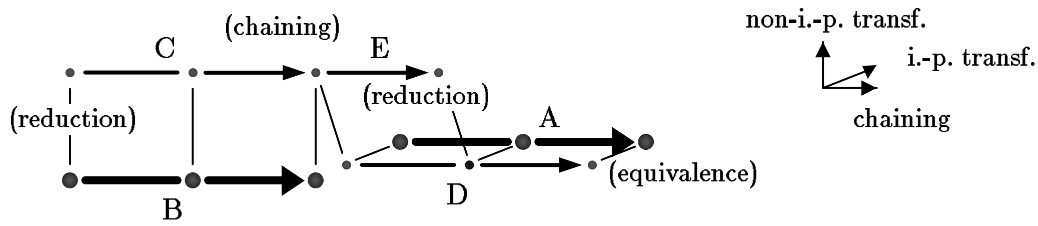


Figure 5: Transformations bringing the "denting" and "dropping" traces into a complete chaining match.



A: `The board is dented.`
B: `The wrench is dropped.`

Figure 6: Association structure for the denting/dropping example.

statement matches, which involve a similar mapping of quantities from one trace to the other).

Construction of an association structure by identifying and elaborating partial matches between events constitutes most of the work done by PATHFINDER in the course of performing causal reconstruction. The remaining subsections in this section outline additional mechanisms enhancing this basic approach. Once an association structure has been constructed, questions of the sort fielded by PATHFINDER can be answered. This part of the processing is illustrated in Section 4.1, in the specific context of the camera example introduced at the beginning of the article.

3.2 Inference and Background Statements

The first extension has to do with the inclusion of deductive inference. While partial matches serve as the basis for identifying associations between events, inference can be used to augment event traces with relevant new assertions, this providing a broader basis for detecting partial matches among the traces. Background statements supplied as part of a causal description (e.g., "The water is inside the tank.") may be included in the inference process.

A second use of inference concerns the checking of partial matches for consistency. Given time point and object equivalences generated by a partial match, it may be the case that the combined set of assertions for the matched event traces is logically inconsistent. When this happens, the partial match in question must be discarded or expanded into a set of reduced partial matches, each of which omits one of the indicated equivalences.

Orthogonal to the question of *when* inference is to be used, there is also a question of what *types* of inference may be of use. Two types of inference present themselves: inference concerning *predicates* in the transition space representation, and inference concerning *attributes* in the representation.

As defined above, each of the predicates used in our realization of transition space decomposes into a set of assertions involving only the primitives EQUAL, NOT-EQUAL, GREATER and NOT-GREATER. Taking advantage of this fact, we may conduct inference over predicates by focusing entirely on relationships between these four primitive predicates. Higher-level predicates may be converted into their corresponding EQUAL/GREATER-level expansions prior to inference, and following inference, the resulting set of assertions may be condensed where possible to redescribe the activity in terms of the higher-level defined predicates. It is a rather simple matter to enumerate logical relationships between EQUAL, NOT-EQUAL, GREATER and NOT-GREATER, and such rules are directly implemented within the PATHFINDER program. Following are a few examples of these rules:

EQUAL(?attribute, ?object-1, ?t1, ?object-2, ?t2)
$\Longrightarrow$ EQUAL(?attribute, ?object-2, ?t2, ?object-1, ?t1)

EQUAL(?attribute, ?object-1, ?t1, ?object-2, ?t2) AND
GREATER(?attribute, ?object-2, ?t2, ?object-3, ?t3)
$\Longrightarrow$ GREATER(?attribute, ?object-1, ?t1, ?object-3, ?t3)

NOT-GREATER(?attribute, ?object-1, ?t1, ?object-2, ?t2) AND
NOT-GREATER(?attribute, ?object-2, ?t2, ?object-1, ?t1)
$\Longrightarrow$ EQUAL(?attribute, ?object-1, ?t1, ?object-2, ?t2)

In addition, PATHFINDER uses a similar set of rules to detect logical inconsistencies. Two examples of these rules are as follows:

NOT-EQUAL(?attribute, ?object-1, ?t1, ?object-1, ?t1)
$\Longrightarrow$ ...inconsistency

EQUAL(?attribute, ?object-1, ?t1, ?object-2, ?t2) AND
GREATER(?attribute, ?object-1, ?t1, ?object-2, ?t2)
$\Longrightarrow$ ...inconsistency

In contrast to inference involving predicates, inference involving attributes relies on a potentially limitless set of inference rules. These rules can express symmetric or transitive relationships for binary attributes:

PRESENT(contact, <?o1, ?o2>, ?t1)
$\Longrightarrow$ PRESENT(contact, <?o2, ?o1>, ?t1)

PRESENT(above, <?o1, ?o2>, ?t1)
$\Longrightarrow$ NOT-PRESENT(above, <?o2, ?o1>, ?t1)

PRESENT(inside, <?o1, ?o2>, ?t1) AND
PRESENT(inside, <?o2, ?o3>, ?t1)
$\Longrightarrow$ PRESENT(inside, <?o1, ?o3>, ?t1)

or they may express arbitrary relationships between different attributes:

PRESENT(contact, <?o1, ?o2>, ?t1)
$\Longrightarrow$ NOT-PRESENT(distance, <?o1, ?o2>, ?t1)

PRESENT(inside, <?o1, ?o2>, ?t1) AND
PRESENT(above, <?o2, ?o3>, ?t1)
$\Longrightarrow$ PRESENT(above, <?o1, ?o3>, ?t1)

A particular implementation strategy has been found to improve considerably the time efficiency of the inference process in PATHFINDER. By this strategy, assertions involving the predicate "EQUAL" are used to construct *equivalence classes* of attribute–object–time-point triples, as, for example:

{ (speed, <object-1, the-background>, t11),
(speed, <object-1, the-background>, t12),
(speed, <object-1, the-background>, t13) }

Assertions involving "NOT-EQUAL," "GREATER" and "NOT-GREATER" are then maintained with respect to these equivalence classes rather than with respect to particular members of the classes, resulting in a corresponding reduction in size for the base of assertions over which inference is conducted.

Background statements supplied in the input to a program can also contribute to the inference process. The following are examples of simple background statements expressible using the input grammar for PATHFINDER.

The block is a solid object.
The beam is not flexible.
The support of object 1 by object 2 is present.
The elevation of the ceiling exceeds the elevation of the floor.

These statements are translated into simple assertions involving the predicates PRESENT, NOT-PRESENT, MATCH, NOT-MATCH, EXCEED and NOT-EXCEED, as defined above. Due to the time-invariant nature of these assertions, the value "null" is again used as a time point argument. The following assertions represent each of the background statements listed above.

PRESENT(is-a, <the-block, solid-object>, null)
NOT-PRESENT(flexible, the-beam, null)
PRESENT(support, <object-1, object-2>, null)
EXCEEDS(elevation, <the-ceiling, the-floor>, null)

Once a base of background assertions has been set up, rules of inference can be applied to extend the set of background assertions, producing new assertions of the same time-independent nature. Later, whenever inference is to be applied to a specific event trace or set of event traces, the background assertions can be instantiated to each specific time point of the involved event traces, augmenting the set of assertions from which inference can draw new consequences. In this manner, background statements can contribute both to the extension of event traces in preparation for matching and the checking of partial matches for consistency.

3.3 Exploratory Transformations

As a second important extension, transformations of both varieties may be applied in an exploratory manner to event traces, forming alternate characterizations of events at different levels of abstraction or in terms of different underlying metaphors. An event trace together with its transformed images forms a small "cluster" of traces, all of which participate in the matching process. In this manner, a program may bridge discontinuities arising from the writer's use of analogy or abstraction.

Figure 7 illustrates a simple exploratory information-preserving transformation of type *substitution*. If we are told that an object "slides to a stop," it is natural to represent this by the first trace. For a rotating object like a wheel, however, a substitution of attributes taking us into the domain of spinning objects may be more appropriate. By including both traces in the association process, we can determine by matching which interpretation is correct. Beneath the event traces, the association structure fragment produced by the transformation is illustrated. By convention, event traces produced by exploratory transformation are depicted in outline.

A second example appears in Figure 8. This transformation is also of type substitution, but it is used in a different way. Suppose a statement in an input description asserts that a radio wave "spreads thin" over space. Using a literal account of this event as involving a liquid spreading thin over a supporting surface, an inconsistency arises from multiple, conflicting type specifications: the radio wave is specified as a radio wave from background information and as a liquid from the spreading event. In this case, a suitable transformation to the domain

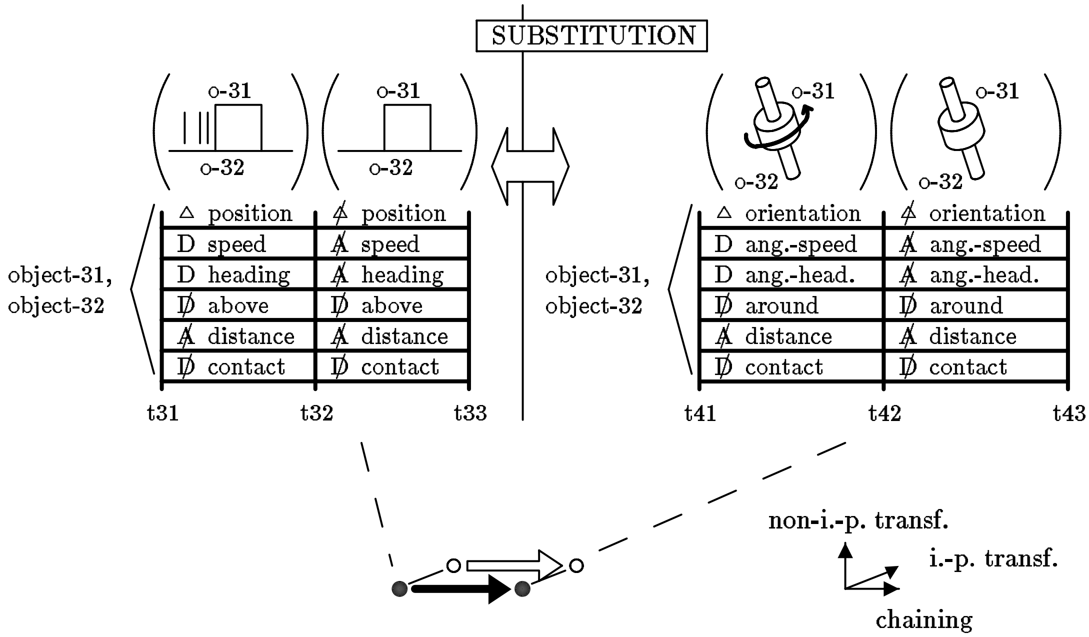


Figure 7: An exploratory information-preserving transformation of type substitution.

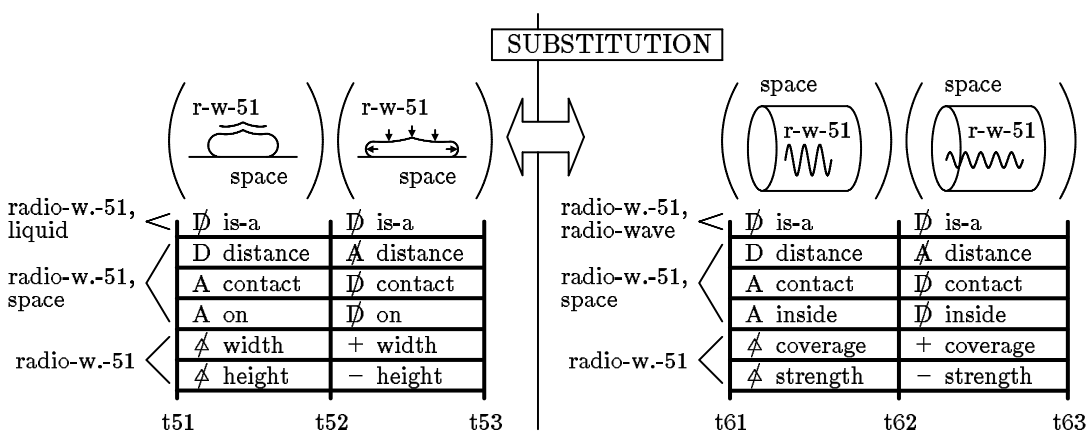


Figure 8: A second substitution transformation.

of radio waves removes the conflict, providing an initial consistent account of the event suitable for matching. As described in Section 4.1, similar transformations are required to process events of the camera description characterizing "the light" as if it were a physical object entering and passing through other objects.

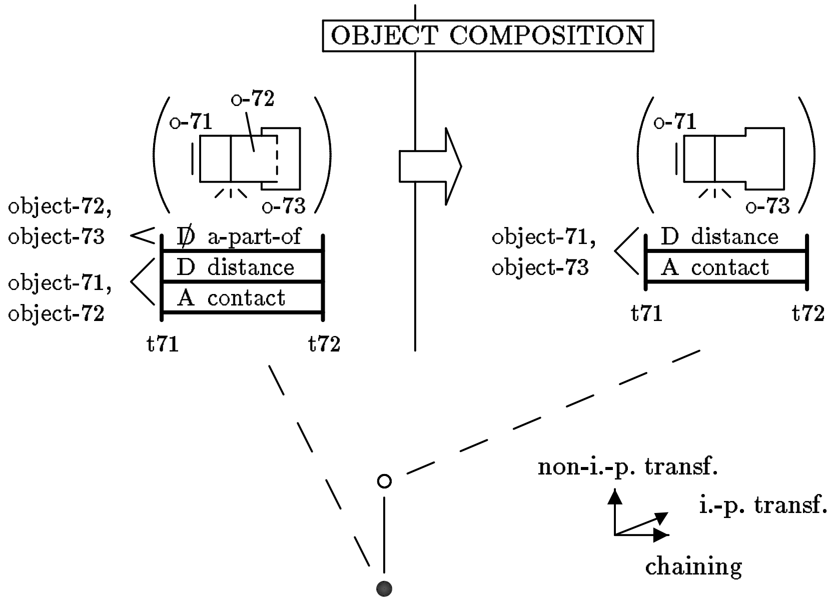


Figure 9: An exploratory non-information-preserving transformation of type object composition.

Figure 9 illustrates an exploratory non-information-preserving transformation of type *object composition*. Suppose one object is specified as coming into contact with a second object, and the second object is a part of a third object. An alternate, more abstract specification of the event portrays the first object as simply coming into contact with the third object. Such a situation arises in processing the camera description: light comes into contact with the silver halides as part of chemically altering them, yet this must be matched with light "converging on the film" which contains the silver halides.

As a second example, a non-information-preserving transformation of type *attribute-object reification* is illustrated in Figure 10. This transformation restates a decrease in "speed" for "object-81" as a decrease in "size" for a new object, "the-speed-of-object-81." In doing so, the initial application of a specific attribute to a specific object has been *reified* as a new object in the representation (see, for example, [23]). This particular transformation serves to bridge discontinuities arising from the initial, literal interpretation of events specifying, for example, a "reducing" or "trimming" of the speed of an object. Taken literally, such events specify a decrease in *size* for an affected quantity.[10]

[10] As a means of anchoring the otherwise arbitrary symbol "the-speed-of-object-81," the left side of the transformation includes a specification of this object as a measurement of "speed" for "object-81."

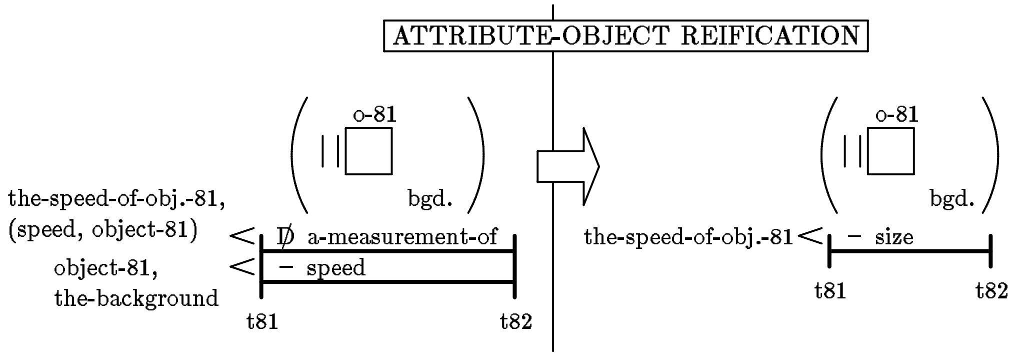


Figure 10: An attribute-object reification transformation.

The following list summarizes the range of exploratory transformations investigated in the course of the examples run on PATHFINDER. Details of these transformations appear in [7].[11]

Information-Preserving Exploratory Transformations:

Equivalence. Replacement of a time point, object, reference standard or attribute with a *synonym* quantity, producing an alternate description of the same event (e.g., replacing the attribute "support" with "supported-by," in keeping with the equivalence of statements like "The support of the book by the table does not disappear." and "The book remains supported by the table.").

Substitution. Replacement of time points, objects, reference standards or attributes with *different, but parallel* quantities in a new context, producing a description of an event distinct from the original event, yet parallel in the types of changes involved.

Non-information-Preserving Exploratory Transformations:

Generalization. Replacement of a reference standard (e.g., an object type such as "container") with a new, more general reference standard (e.g., "physical object").

[11]The complete set of *association types* explored in PATHFINDER includes these plus three additional varieties: complete chaining associations and reduction transformations—these used only in the elaboration of partial matches—and inference transformations—these considered separately from the exploratory transformations.

Interval Composition. Merging of two adjacent time intervals into a single composite interval, with changes specified according to the composition of the changes in the original two intervals.

Attribute Composition. Reexpression of activity originally involving a set of related attributes (e.g., "height," "width" and "depth") as activity involving a single, encompassing attribute (e.g., "size").

Object Composition. Reexpression of activity originally involving the parts of an object as activity involving the whole object.

Attribute-Object Reification. As illustrated above—replacement of activity involving a particular attribute applied to a particular object with assertions involving a newly-created object representing the attribute applied to its argument.

Event-Attribute Reification. A transformation replacing part of an event trace (e.g., changes involved in light striking a surface) with assertions involving a newly-introduced attribute applied to one of the participating objects (e.g., "illuminated" applied to the surface).

Event-Object Reification. A transformation replacing part of an event trace (e.g., changes involved in a collision) with assertions involving a newly-created object representing the replaced activity (e.g., a new object representing the collision, with other objects "engaged-in" the collision object).

In all cases, exploratory transformations produce *plausible* redescriptions of events—redescriptions that leave room for error. For example, the above object composition transformation only speculates that coming into contact with a part of an object might be redescribed as coming into contact with the whole, since it is conceivable that "object-71" might already be in contact with some other part of "object-73" before coming into contact with "object-72." For this reason, transformed images of events are granted a probationary status, to be accepted only when elaborated partial matches link these images to other known occurrences.

Implementationally, exploratory transformations take the form of bidirectional rules acting to replace sets of assertions appearing in event traces with other sets of assertions. These rules may be combined to form rule clusters corresponding to more comprehensive transformations. Each rule in such a cluster describes an operation to take in a particular circumstance (presence of a phenomenon, absence, appearance, change, etc.). As a simplification, rules involving a complete match between the two sides of the rule (e.g., the substitution transformations illustrated above) are augmented by a binding list mapping time points, objects, reference standards and attributes between the two sides of the rule. Whenever such a rule is found to be applicable to an event trace, the event trace is additionally transformed by a substitution of terms according to this binding list. In many cases, this simplification permits a single rule covering, say, presence of a concerned phenomenon, to take the place of a large cluster of rules detailing alternate circumstances.

3.4 Connecting Statements

The third extension has to do with compliance with "connecting statements," as described in Section 2.1. Connecting statements abstract such devices as causal/temporal connectives and adverbial phrases when used to indicate explicit inter-event relationships. A good enumeration of such devices appears in [39]. As input to PATHFINDER, connecting statements involve two nominalized event references connected by a verb construction indicating a temporal relationship (e.g., "occurs after"), causal relationship ("causes"), analogical relationship ("is equivalent to," "parallels"), or abstraction relationship (e.g., "summarizes"). The following are examples of connecting statements accepted by PATHFINDER:

The ball moving occurs after the hitting of the block.
The structure expanding causes the component to move.
The pile growing summarizes the pile increasing in height.
The electric current traveling from the first junction to the second junction
  is equivalent to the electric current passing through the filament.

In many cases, it is possible to tell before elaboration of a partial match that it will either: (1) guarantee compliance with a particular connecting statement, or (2) guarantee a failure of compliance with a particular connecting statement. In the heuristics for choosing among competing partial matches, PATHFINDER rewards partial matches in the first case and abandons them in the second case. After a partial match has been elaborated, a more stringent test for compliance with connecting statements is possible and may result in removal of the chain of full associations generated from the partial match.

4 PATHFINDER

As stated in Section 1, PATHFINDER is a 20,000 line program coded in Common Lisp and run on a Symbolics 3640 Lisp Machine. It contains a set of utilities for representing, matching and conducting inference and transformations on events in transition space, a parser operating on a simple context-free skeleton of English grammar, a simple language generation capability, and a set of supervisory routines for conducting causal reconstruction. PATHFINDER has been applied to over 60 causal descriptions, most involving 2–4 events, in a wide range of physical domains including: interaction between solid objects and liquids, condensation and melting, combustion, radio signals, light, chemical reactions and electric currents.

All input to PATHFINDER consists of statements in simplified English. The input grammar for PATHFINDER is a semantic grammar in that it includes meaning-based categories such as "attribute," "object," "relative time expression" and so forth (see, for example, [31]). This grammar grounds out in a set of five non-terminal categories: object, reference-standard, attribute, verb-group and preposition. The lexicon for PATHFINDER is initialized with a large set of prepositions and attributes related to physical events, plus a range of verb groups providing a foundation for the representation (e.g., "changes," "does not disappear," and so forth). New members of all five categories are inserted into the lexicon following queries to a human operator during the parsing process.

Input quantities supplied to PATHFINDER were described in Section 2.1, and examples of many of these quantities appear in Section 4.1. First, PATHFINDER is given a causal description, consisting of (1) *event references* ("The light enters the camera."), (2) *background statements* ("The head is a part of the nail.") and, in some cases, (3) *connecting statements* ("The device starting to move causes the lever to start to move."). Next, a set of supplementary information is provided, possibly including (1) additional *background statements*, (2) *event definitions*, (3) *precedent events*, which may be of use in reconstructing the activity, (4) *rules of inference,* and (5) *rules of restatement*, including specifications of analogical mappings and rules of abstraction.

Given input in this form, PATHFINDER performs causal reconstruction in four phases, as outlined in Figure 11 (a). In the first phase, it parses the input text and uses the supplied event definitions to form event traces for all events referenced in the description, as illustrated in the context of the denting/dropping example presented in Section 3.1. Background statements, rules of inference and rules of restatement are also translated to transition space representations during this step.

In the second phase, PATHFINDER extends the event traces through inference and applies exploratory transformations—these motivated by rules of inference and rules of restatement in the input—producing for each event a cluster of traces describing that activity in different ways. The process begins with inference carried out on each of the original event traces depicting referenced events and precedents. The images of these traces under inference are next transformed where possible via exploratory transformations, with the resulting images of transformation again extended through inference. This process continues recursively with newly-generated image traces subjected to further transformation and inference where possible. The process is bounded in a simple way by permitting only one application of any particular transformation—applied in either the forward or reverse direction—to appear along any path of successively-applied transformations

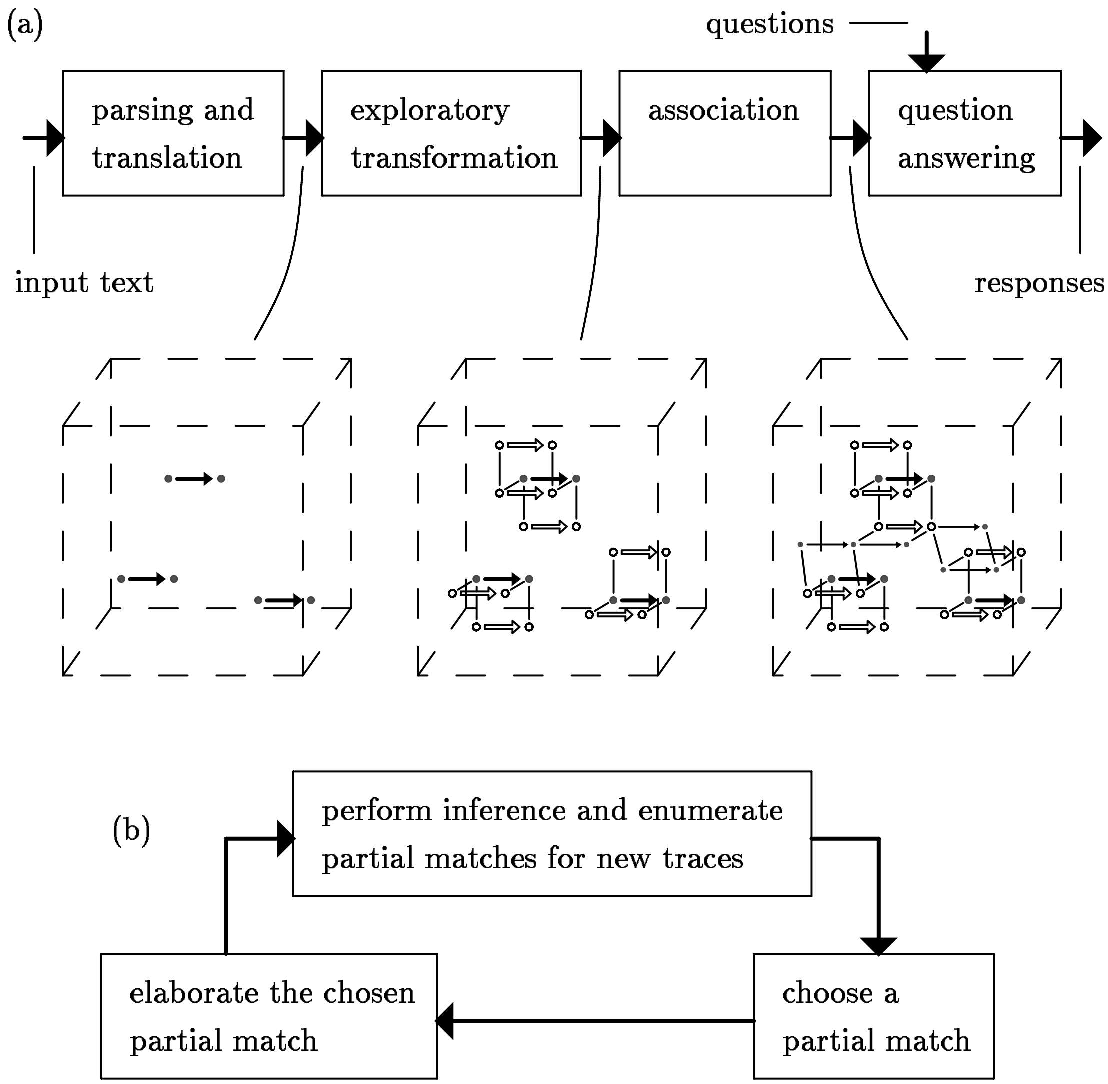


Figure 11: (a) Four phases of execution for PATHFINDER when performing causal reconstruction, (b) the association cycle.

emanating from one of the original event traces.[12]

In the third phase of operation, PATHFINDER constructs an agenda of partial

[12]Two alternate methods of bounding the transformation process are: (1) placing a depth limit on the number of transformations performed on each initial event trace, and (2) abandoning paths of successive transformation that yield duplicates or reductions of other event traces produced thus far. The first approach suffers from arbitrariness, while the second does not bound sequences of transformation that introduce new objects (e.g., a sequence of object composition transformations postulating new subparts of parts of an object).

matches between event traces in different clusters. By iteratively choosing the top-ranked partial match and elaborating it in the manner outlined in Section 3.1, PATHFINDER associates the clusters together.

Figure 11 (b) details the iterative cycle of the third phase. This process is seeded with the set of event traces produced in phase two of PATHFINDER's operation; on subsequent iterations, event traces formed during the previous elaboration of a partial match are also included. At each iteration, event traces newly added to the process are subjected to inference if this has not already been done, and partial matches are enumerated between these new event traces and all traces involved in the association process.[13] The generated partial matches are then inserted as new entries in the agenda, and heuristics are used to score all agenda entries, leading to the selection of a new partial match for consideration. This match is then elaborated into a sequence of full associations, possibly generating new intermediate event traces in the process. Also during elaboration, several checks are conducted (as outlined below) to ensure that the chosen partial match is not defective in some way. The entire cycle repeats until a single tree of associations is produced, connecting all events referenced in the input description.

Several heuristics are used to rank the partial matches contained in PATHFINDER's agenda. A detailed account of the heuristic function used by PATHFINDER appears in Appendix A. The first heuristic listed below is given strongest influence over the selection process.

Matching between transitions. Definite changes (APPEAR, DISAPPEAR, CHANGE, INCREASE, DECREASE) are weighted most, other dynamic assertions next and static assertions least.

Proximity to description events. Penalties are introduced for matches involving precedent events or exploratory transformations of events.

Narrative ordering. Preference is given to chaining matches between events referenced consecutively in the description.

Current status of the association structure. Penalties are introduced for matches providing a second antecedent or consequent for an event, matches between events already connected via associations, and matches involving hypothesized objects (e.g., a conjectured part of a physical object).

Types of associations. Partial restatement matches are penalized slightly relative to partial chaining matches. Also, matches not fulfilling any connecting statements are penalized slightly, and matches violating a connecting statement are penalized heavily.

[13]Actually, only traces marked "active" are used for matching. This set excludes traces that have been extended via inference, traces that are reductions of other traces, and one of a pair of traces linked by an equivalence association—these traces yielding only redundant or degenerate partial matches relative to the matches for other traces not excluded from the active set.

Following the selection of a partial match for elaboration, several additional tests are conducted, as outlined below. Failure to pass one of these tests forces the selected partial match to be discarded, possibly to be replaced in the agenda by a set of reduced matches, each excluding one of the symbol pairs in the binding list of the original match (e.g., "[t11 / t23]"), along with all matched assertions requiring that particular symbol binding.

Equivalences between description objects. Partial matches must not equate different objects named in the input description (e.g., "the hammer" and "the nail"). However, such objects may be equated with hypothesized objects, such as an unnamed actor of an event.

Logical inconsistency. The combined assertions of two matched traces must not be logically inconsistent, given the object and time point equivalences generated by the matcher.

Violated connecting statements. A partial match must not violate an inter-event relationship specified by a connecting statement in the description or supplementary information.

In its fourth phase of operation, PATHFINDER fields questions of the four varieties listed in Section 2.1. These questions are answered either by inspecting the association structure completed in the third phase, or by performing one or more additional cycles of association followed by inspection of the resulting association structure. Specific procedures used in question answering are illustrated in the context of the example presented in the following section.

4.1 Processing the Camera Description

This section describes PATHFINDER's processing of the camera description introduced in Section 1. An abbreviated session transcript for the example appears in Figures 12 and 13; excerpts from the input file appear in Figure 12 and question answering excerpts appear in Figure 13. The complete input file for this example plus additional samples of question answering appear in [7].

The causal description, shown in Figure 12 (a), contains six event references, as follows:

```
The camera records the image on the film.
The light enters the camera.
The light passes through the lens.
The light converges on the film.
The light forms the image on the film.
The light chemically alters the silver halides.
```

(a) *(the causal description in simplified English)*

```
The camera records the image on the film.  The recording of the image
is a function of the camera.  The light enters the camera.  The light
passes through the lens.  The light converges on the film.  The light
forms the image on the film.  The light chemically alters the silver
halides.  The silver halides are contained in the emulsion.  The
emulsion is a part of the film.
```

(b) *(an event definition for "entering," involving physical objects)*

```
Object 11 entering object 12 translates to the following event.
Concurrently, object 11 remains a physical object, object 12 remains a
physical object, object 12 remains hollow, the position of object 11
changes, the speed of object 11 does not disappear, the heading of
object 11 does not change, and object 11 becomes inside object 12.
```

(c) *(a precedent event: change of appearance during chemical transformation)*

```
Object 61 changes appearance from chemical transformation.
```

```
Object 61 changing appearance from chemical transformation translates
to the following event.  Concurrently, object 61 remains a physical
object, object 61 becomes not made of substance 62, object 61 becomes
made of substance 63, and the appearance of object 61 changes.
```

(d) *(a restatement rule: light viewed as a physical object w.r.t. "contact")*

```
Concurrently, quantity 141 is a beam of light, object 142 is a
physical object, and the contact between quantity 141 and object 142
is present.  The following statement parallels the preceding
statement.  Concurrently, object 151 is a physical object, object 152
is a physical object, and the contact between object 151 and object
152 is present.
```

(e) *(a restatement rule: contact with a part summarized as contact with the whole)*

```
Concurrently, object 201 remains a part of object 202, the distance
between object 203 and object 201 disappears, and the contact between
object 203 and object 201 appears.  The preceding statement is
summarized by the following statement.  Concurrently, the distance
between object 203 and object 202 disappears, and the contact between
object 203 and object 202 appears.
```

Figure 12: Input text for the camera description (partial).

(a) *(a question of type 1: describing the time-varying status of a selected attribute)*

What happens to the distance between the light and the film?

> First, as the light passes through the lens, the distance between the light and the film decreases. Next, as the light converges on the film, the distance between the light and the film disappears. Next, as the camera records the image on the film, the distance between the light and the film does not appear.

(b) *(a question of type 2: describing the interrelationship between two events)*

How does the light converging on the film relate to the light forming the image on the film?

> The light converging on the film causes the light to chemically alter the silver halides, which ends with the silver halides changing appearance from chemical transformation, which occurs at the end of the light forming the image on the film.

(c) *(a question of type 3: identifying a plausible causal connection)*

How could the building reflecting the light into the camera cause the light to converge on the film?

> The building reflecting the light into the camera could end with the light entering the camera, which could cause the light to pass through the lens, which could cause the light to converge on the film.

(d) *(a question of type 4: restating a portion of the activity)*

Does the light come into contact with the emulsion?

> Yes. The light coming into contact with the emulsion is a part of the light converging on the film.

Figure 13: Question answering for the camera description.

Also appearing in the causal description are three background statements:

```
The recording of the image is a function of the camera.
The silver halides are contained in the emulsion.
The emulsion is a part of the film.
```

Each statement in the input description to PATHFINDER finds motivation in the original description extracted from the *Encyclopedia Americana* and listed in Section 1, as a comparison with that description will confirm.[14]

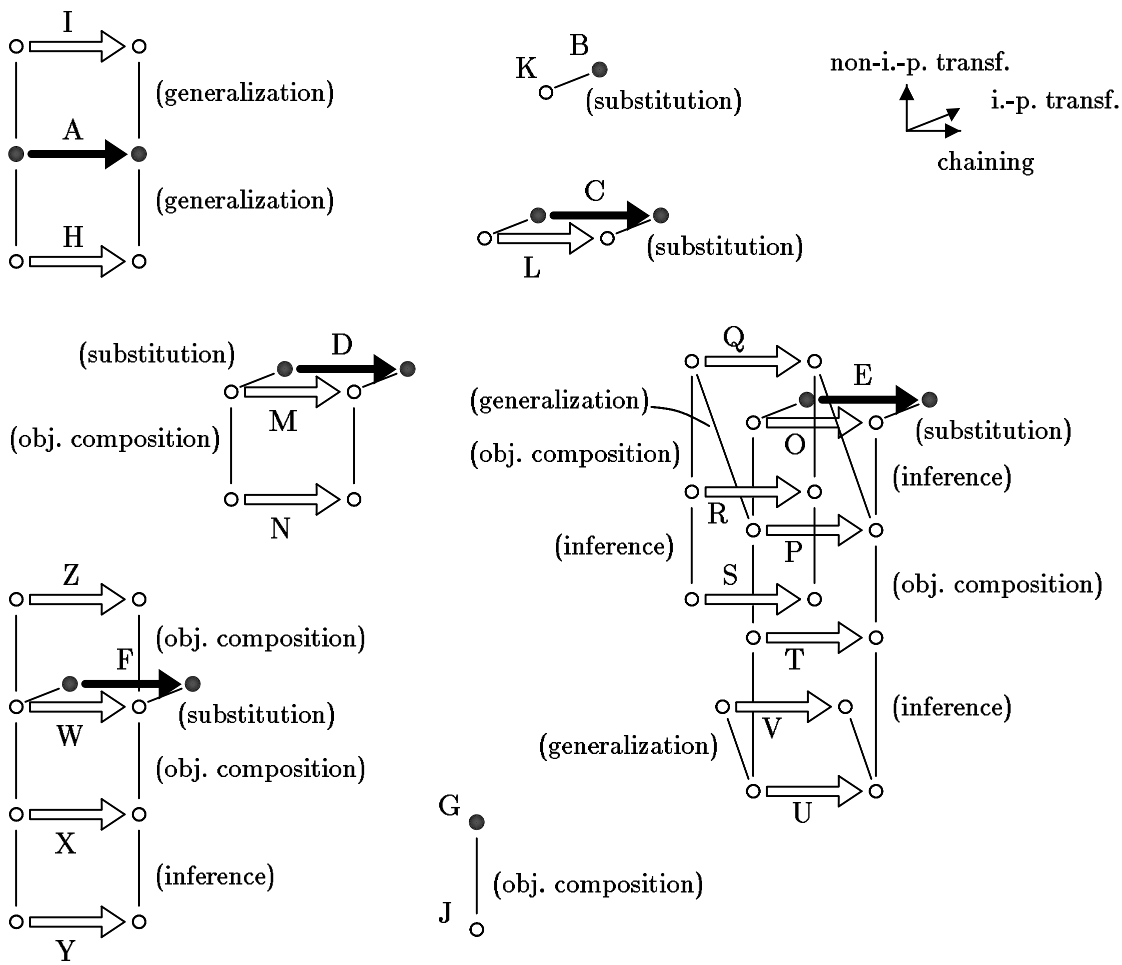
Following the causal description, the input file contains several segments of supplementary information, including additional background information concerning the types of each object (e.g., "The light is a beam of light."), generic definitions for each of the six referenced events (e.g., Figure 12 (b)), one precedent event—"Object 61 changes appearance from chemical transformation."—(Figure 12 (c)), five rules of inference, three rules of restatement concerning light viewed metaphorically as a physical object (e.g., Figure 12 (d)), and five rules restatement concerning abstraction relationships (e.g., Figure 12 (e)).

The presentation here chronicles PATHFINDER's processing of the example first in overview, followed by a more detailed consideration of a particular portion of the generated association structure. Further discussion of PATHFINDER's processing for this example may be found in [7].

During the first phase PATHFINDER parses the input text and uses the supplied event definitions to construct event traces for the six referenced events plus the one precedent event. Rules of inference and restatement in the input text are used to form inference and exploratory transformations, and these are used to form alternate characterizations of the events in phase two. Figure 14 illustrates the generated clusters of event traces at the end of phase two, using the graphical format for association structures. As before, initial representations of events are depicted using heavy arrows/dots, and transformed images are depicted in outline. Traces A through F in Figure 14 depict the six events referenced in the description, while trace G depicts the precedent event.

In the third phase, PATHFINDER executes six iterations of its association cycle, working out associations between traces in different clusters of the partially completed association structure. On the first iteration of the association cycle, PATHFINDER chooses among 120 candidate partial matches; on the remaining iterations, similar numbers of partial matches are considered. At the end of the association process, the seven clusters of traces have been associated together, resulting in the completed association structure shown in Figure 15. (Unused transformed images of traces are not shown in this figure.)

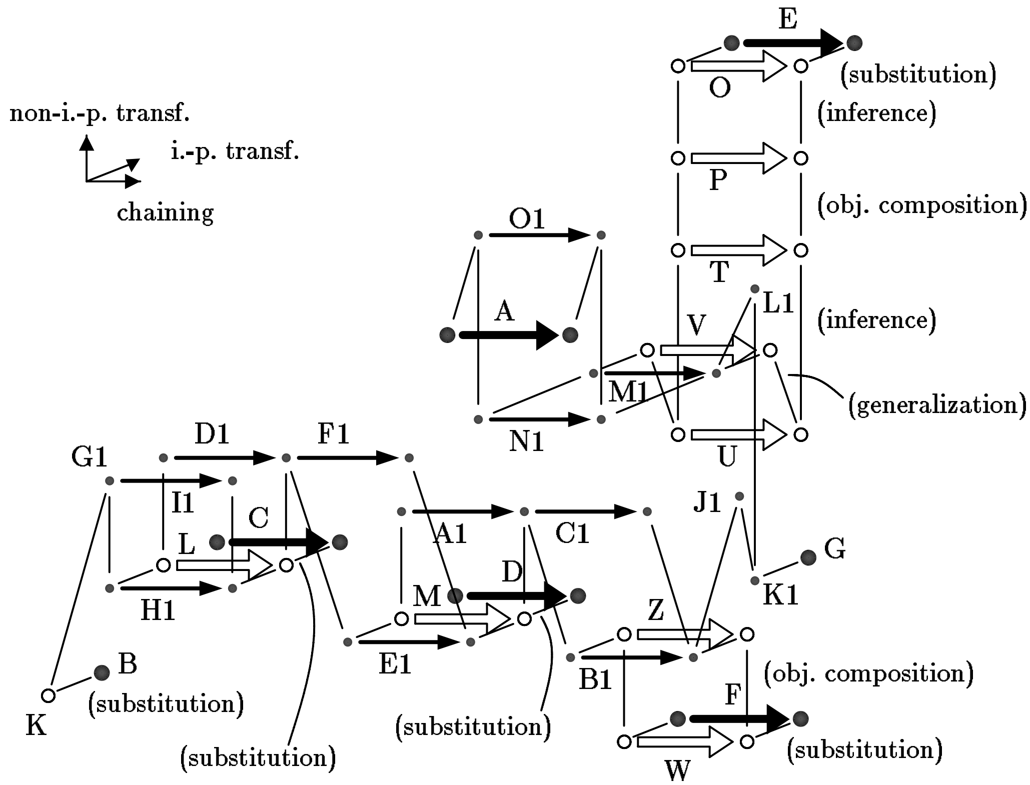
[14] Also, the use of "by" in the original description could be taken as motivation for a connecting statement "The light forming the image on the film summarizes the light chemically altering the silver halides.". PATHFINDER can process the description with or without this statement, and since the latter situation is more difficult, it is this situation that is presented here.



A: `The camera records the image on the film.`
B: `The light enters the camera.`
C: `The light passes through the lens.`
D: `The light converges on the film.`
E: `The light forms the image on the film.`
F: `The light chemically alters the silver halides.`

G: `Object 61 changes appearance from chemical transformation.`

Figure 14: Partial association structure at the end of phase two of PATHFINDER's operation.



A: `The camera records the image on the film.`
B: `The light enters the camera.`
C: `The light passes through the lens.`
D: `The light converges on the film.`
E: `The light forms the image on the film.`
F: `The light chemically alters the silver halides.`

G: `Object 61 changes appearance from chemical transformation.`

Figure 15: Status of the association structure at the end of the association phase for the camera description.

The following discussion focuses on a portion of the association structure bridging two accounts at different levels of abstraction—namely, the portion relating the two events "The light converges on the film." (trace D) and "The light chemically alters the silver halides." (trace F). Figure 16 illustrates the contents of the traces marked D, M, F, W and Z in Figure 15. As noted above, traces M, W and Z have been formed by exploratory transformations of the original two traces, D and F.

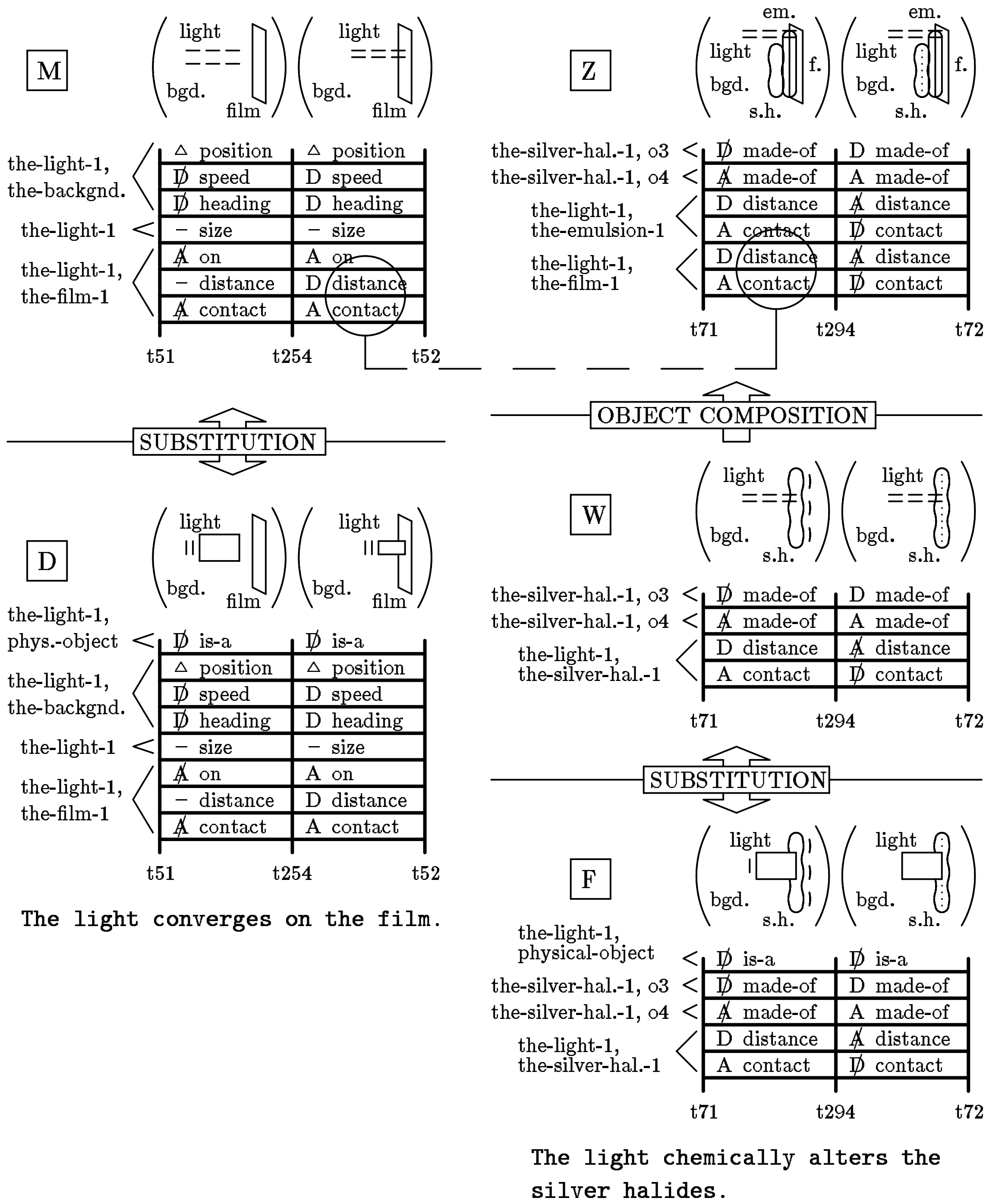


Figure 16: Original traces and transformed images for a portion of the camera description. A partial match is identified between traces M and Z.

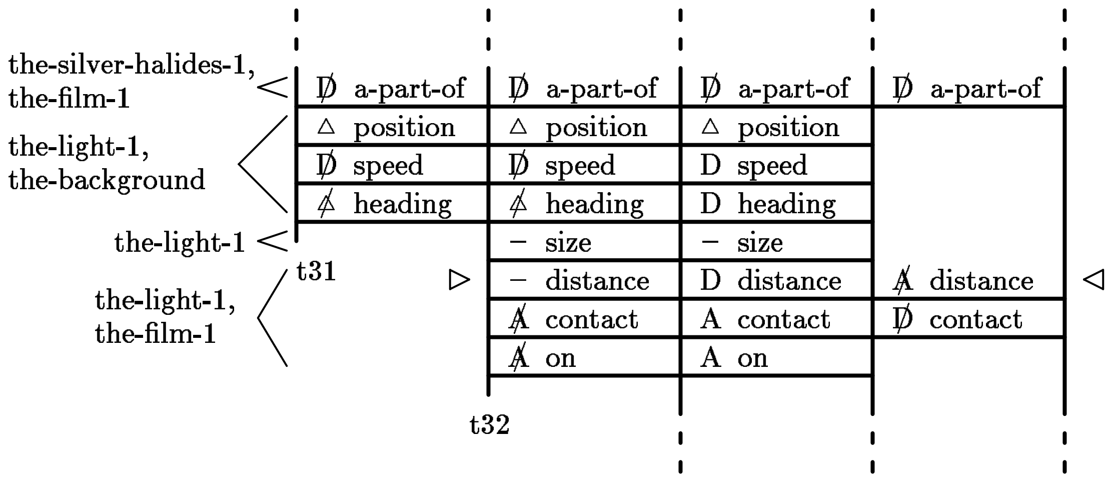


Figure 17: Part of the composite trace for the camera description. The indicated row is used in answering the submitted question.

For M and W, a substitution transformation derived from the rule of restatement shown in Figure 12 (d) has recast activity involving a physical object with activity involving a beam of light.[15] For Z, an object composition transformation derived from the rule of restatement shown in Figure 12 (e) has recast light contacting the silver halides as light contacting the film. A partial chaining match has been identified between traces M and Z, both of which specify the light contacting the film. Finally, traces A1, B1 and C1 of Figure 15 have been produced by elaboration of this partial match (involving an equivalence mapping from Z to B1, removal of assertions from M to A1 and from B1 to C1, and chaining from A1 to C1).

After the events have been associated, PATHFINDER proceeds to its fourth phase, question-answering. As described in Section 2.1, the first type of question answered by PATHFINDER concerns the time-varying status of particular attributes of objects; e.g., "What happens to the distance between the light and the film?" To answer this, PATHFINDER forms a *composite trace* of the activity by first mapping traces in the association structure through indicated equivalence associations so that they involve a consistent set of objects and time points, then merging the sets of assertions in these traces in order to produce a composite account. Figure 17 illustrates part of the composite trace formed for the camera description. Given this trace, PATHFINDER extracts the row relevant to the submitted question and expresses this fragment in simple English, also inserting phrases to situate the changes within the context of particular events within which

[15] PATHFINDER routinely removes from event traces all assertions matched by background assertions. Thus, following the substitution operations creating traces M and W from D and F, all assertions concerning "the light" remaining a "beam of light" are removed from M and W.

Priority	Association type	As expressed in English
1.	SUBSTITUTION	"...`parallels`..."
2.	CHAINING	"...`causes`..." / "...`is caused by`..."
3.	REDUCTION	"...`is a part of`..." / "...`involves`..."
4.	GENERALIZATION, COMPOSITION, and REIFICATION	"...`summarizes`..." / "...`is summarized by`..."
5.	INFERENCE	"...`implies`..." / "...`is implied by`..."
6.	EQUIVALENCE	"...`is equivalent to`..."

Table 1: Prioritization and English translations for association types, as used to describe sequences of associations in PATHFINDER.

they occur. PATHFINDER's response for this question is listed in Figure 13 (a).

The second type of question concerns inter-event relationships; e.g., "How does the light converging on the film relate to the light forming the image on the film?" To answer this type of question, PATHFINDER extracts the sequence of associations connecting the specified events in the association structure and describes this sequence in simple English, highlighting important associations as it goes. For the above question, the relevant sequence is a chain of 14 associations involving the description events "The light converges on the film." and "The light forms the image on the film.", plus a transformed image of the description event "The light chemically alters the silver halides." and a transformed image of the precedent event "Object 61 changes appearance from chemical transformation.". This sequence of associations is summarized below.

The light converges on the film.
$\longrightarrow$ SUBSTITUTION $\longrightarrow$ REDUCTION
$\longrightarrow$ CHAINING $\longrightarrow$ REDUCTION (inverse)
(The light chemically alters the silver halides.)
$\longrightarrow$ REDUCTION $\longrightarrow$ REDUCTION (inverse)
(The silver halides change appearance from chemical transformation.)
$\longrightarrow$ REDUCTION $\longrightarrow$ REDUCTION (inverse)
$\longrightarrow$ EQUIVALENCE $\longrightarrow$ GENERALIZATION (inverse)
$\longrightarrow$ INFERENCE (inverse) $\longrightarrow$ COMPOSITION
$\longrightarrow$ INFERENCE (inverse) $\longrightarrow$ SUBSTITUTION
The light forms the image on the film.

To summarize sequences of associations, PATHFINDER uses a prioritization of association types. Between each pair of description/precedent events represented along the sequence, the single association of highest priority is used to describe the relationship between those events. Table 1 depicts the prioritization scheme found to be most effective over the set of descriptions processed by PATHFINDER.

If two associations of equal priority and opposite in orientation appear along such a path segment, PATHFINDER defaults to a temporal characterization of the inter-event relationship:

> "`...occurs at the beginning of ...`" / "`...begins with ...`"
> "`...occurs at the end of ...`" / "`...ends with ...`"

or, if these temporal characterizations do not fit, a catch-all characterization:

> "`...coincides with ...`"

For the above question "How does the light converging on the film relate to the light forming the image on the film?", the answer produced by PATHFINDER is listed in Figure 13 (b).[16]

The third and fourth types of questions also ask about inter-event associations, but require PATHFINDER to do further association first. Otherwise, these questions are handled in the same manner as questions of the second type. If necessary, supplementary information such as event definitions may be provided with these questions. The third type of question involves plausible causal associations, as might be used in making simple predictions or explanations; e.g., "How could the building reflecting the light into the camera cause the light to converge on the film?". For this particular question, PATHFINDER identifies a partial restatement match between the building reflecting the light into the camera and the previously-specified event of the light entering the camera. After elaborating this partial match, PATHFINDER traces the sequence of associations between the two indicated events and answers as listed in Figure 13 (c).

The fourth type of question asks if a new event may be used to paraphrase a part of the activity; e.g., "Does the light come into contact with the emulsion?". For this particular question, PATHFINDER identifies a partial restatement match between the new event, the light coming into contact with the emulsion, and a transformed version of the light converging on the film. PATHFINDER's response is listed in Figure 13 (d).

5 Related Literature

The work presented in this article relates to a rather diverse set of research efforts in artificial intelligence, psychology, linguistics and philosophy. This section outlines some of these relationships.

[16]The original representations for both of the concerned events are inconsistent, specifying "the light" as a physical object. In answering questions, PATHFINDER ignores associations that transform inconsistent accounts into consistent ones, and thus the substitution associations in the extracted sequence are ignored.

Perhaps the most closely allied research to that described here is the work in reasoning about physical systems, and in particular, work in qualitative reasoning [15, 22, 37] and model-based reasoning [13, 14]. The approach described here draws important motivation from these efforts and shares with them an emphasis on representing physical activity in terms of time-varying parameters of participating objects, plus, in the case of qualitative reasoning, an emphasis on qualitative characterization of changes in these parameters.

Due to the particular nature of the causal reconstruction task, there are also some differences between the approach described here and a significant fraction of the work in qualitative reasoning and model-based reasoning. The most central difference is a grounding of the representation in *language,* rather than in a scientific model of a target physical behavior. This has an effect of bringing into the representation a range of phenomena recognized and articulated by humans, yet normally not included in scientific accounts of activity: objects such as beams of light, radio signals, sounds, dents, paths and so forth, and attributes like "support," "inside," "appearance," and "shape," which are often fundamentally qualitative in nature and rougher in granularity than the quantitative phenomena supporting many scientific models of behavior. One result of this is that the transition space representation tends to be somewhat more macroscopic, with events often spanning several qualitative states or operating regions of a device. Other differences with some the work in qualitative reasoning and model-based reasoning include a greater emphasis on representation at different levels of abstraction and in terms of different underlying analogies, and the use of heuristically-guided matching rather than constraint propagation as a mechanism for reasoning.

A number of other research efforts concerned with reasoning about physical systems have drawn wholly or in part on cognitive motivation [9, 29, 30, 41, 49, 57]. One example is the work in naive physics [29, 30]. Transition space shares with this approach an ontological viewpoint that includes in the description of physical behavior many quantities appearing in everyday language—edges, surfaces, quantities of liquid, spaces and so forth. Contrasting with this approach, however, are the extended modes of reasoning employed in transition space, from heuristic matching to the incorporation of analogy and abstraction. Separately, the research described here also shares a basic orientation towards representation found in the work on CYC [41], including a grounding out in language and a reliance on analogy and abstraction to fill in missing information (also articulated as the Breadth Hypothesis in [42]). However, the tasks being addressed are somewhat different: acquiring causal knowledge solely through language versus acquiring and organizing knowledge of all variety of things via language and/or direct entry of symbolic information.

Also very relevant to the work described here is the literature on reasoning about time, including research in reasoning about temporal ordering information

[1, 60, 61] and reasoning about events and causality (e.g., [2, 45]). Following the observation by Hanks and McDermott [28] that several non-monotonic logic formalisms are insufficiently constraining for the formation of simple predictions about the future, several preference mechanisms have been offered for ordering the set of extensions to a logical theory as required to make such predictions (e.g., [48, 58]). However, there remains a concern that no general preference criteria may exist for this task [19]. An alternative, of course, is to employ heuristics in selecting among candidate sequences of events. In particular, the work presented here would suggest that in the specific context of postulating causal sequences connecting events referenced in a written description, the heuristic of combining events based on partial matches between their implied changes would appear to be extremely useful.

Two research efforts in the area of causal modeling are those of Doyle [17, 18] and Pearl and Verma [53]. While both fit the characterization of causal modeling outlined in Section 2, there are some differences between the two approaches. Doyle's approach is largely symbolic, consisting of a set of device models for individual physical mechanisms, plus a set of heuristics for assembling the devices into candidate models of a system, to be checked against input/output data using constraint propagation. Pearl and Verma's approach is grounded in probability calculations, and in this case, the causal modeling program is allowed to interact with the system being modeled. Both approaches lend themselves to possible interaction with the approach described here in targeting the combined causal reconstruction/causal modeling problem occurring when a program is simultaneously presented with a causal description and a demonstration of a target physical behavior.

The work described here is also related to research in analogy (e.g., [10, 24, 64]), and possible interactions arise in this context as well. As many analogical reasoning programs represent individual events by the equivalent of atomic formulae (e.g., "COLLIDE(o1, o2)"), these fitting into larger *systems* of events to be associated analogically, a promising approach would involve the use of transition space representations to add further discriminatory power to such programs in deciding which particular pairs of events ought to be matched in the course of associating larger systems of events.

The work of Rieger [55] constitutes an important precedent for the work described here. While this approach was never fully implemented and tested, it carried a significant intuitive appeal in modeling language comprehension of the sort involved in causal reconstruction as involving search through a discrimination network for candidate device models, each supporting simple reasoning about physical behavior. Some possible drawbacks of Rieger's approach include the presence of several overlapping varieties of causality, this making it difficult to establish canonical representations for devices, the absence of an explicit notion of time, and, perhaps most significantly, the absence of a mechanism for combining previously unassociated physical mechanisms—that is, pairs of device models for which no explicit relationship has been inserted into the knowledge base.

Providing an original motivation for the research described in this article is the Event Shape Diagram representation of Waltz [62]. This representation describes events in terms of various functions and assertions plotted against time (these roughly equivalent to individual rows in the diagrams for event traces), with additional assertions added to specify relationships between plottings. Event Shape Diagrams can be used to draw distinctions between closely related events (e.g., "eat," "swallow," "nibble" and "gulp"), or to process metaphorical references to events (e.g., a person "eating up" compliments). The work described here takes this idea one step further in additionally using detailed knowledge of the temporal unfolding of events to distinguish descriptions at different levels of abstraction and to detect plausible causal associations between events. Also in this regard, the work described here serves to extend another representation inspired by Event Shape Diagrams: the Event Calculus representation [4, 5] used to summarize events occurring in an observed scene given a time-log of object positions, orientations and other simple properties as might be output from a computer vision system.

In the area of natural language understanding, the transition space approach exhibits some similarity to the Conceptual Dependency representation and other work in knowledge-based language comprehension [20, 36, 40, 56]. In common between these approaches and the transition space representation is the notion of representing events/actions as stereotypical instances to be matched with one another. The transition space approach may be thought of as augmenting this process to include matching of fine-grain information concerning the sequencing and simultaneity of individual changes expressible in language.

Also relevant in the area of natural language understanding is the body of work in spreading activation (e.g., [3, 12, 52, 54]. Much of the work in spreading activation has dealt with narrative understanding—that is, stories about human actions and their consequences. A possible difficulty for this approach concerning physical causality is that physical events are very context-dependent. For example, if a hand lets go of an object, the object will fall only if it is not otherwise supported. To represent this context in a semantic network, event nodes must be split into subnodes depicting special cases, and this subdivision becomes rather unwieldy if we consider the variety of possible contextual variations of each event and the fact that only some of variants of one event may be linked to variants of another event. Transition space captures this context directly as added assertions in an event trace. These added assertions are then incorporated into the matching and inference processes used to determine if two events may be related in a particular way. A separate possible difficulty concerning the use of spreading activation is that this approach places a heavy burden on the knowledge engineer in charge of the network to explicitly enter all possible inter-event relationships—a task which becomes progressively more difficult as the size of the network increases. For the transition space representation, the inter-event associations are *implicit.* To enter a new event, we need only specify what happens during the event, resulting in an automatic association of that event with other events known to the system.

Section 3 listed several results/analyses from psychology relevant to the representation of events and causality. Also relevant to the work described here is the research in mental models [26, 34]. Of interest regarding the causal reconstruction task are two types of results from this research: (1) documentation of cases where humans successfully employ analogy to reason about abstract domains in terms of simpler or more accessible domains (e.g., reasoning about electricity as if it were flowing water or a moving crowd [25], and (2) documentation of cases where human reasoning fails due to a reliance on faulty models (e.g., taking force as responsible for motion rather than acceleration [11, 16, 44]). Both types of results suggest components of supplementary knowledge useful for the comprehension of causal descriptions—the latter being useful if we wish to compensate for faulty explanations offered by the writers of particular causal descriptions.

Separately, research by Kahneman *et al.* [35] appeals to many of the same concerns as mental models research. This work outlines several heuristics employed by humans in simple reasoning tasks. Of these, the Availability Heuristic is perhaps most relevant to the task of causal reconstruction, as by this heuristic, a program with a limited knowledge base of events, rules of inference and rules of restatement may draw conclusions by generalizing over the possible reconstructions given its limited repertoire of knowledge.

Concerning research in linguistics, Grice's Maxims of Conversation were mentioned in Section 2. Also relevant is the work in natural language semantics, as, for instance, the work Jackendoff [32] and Talmy [59]. Jackendoff proposes a semantic representation consisting of a few major ontological categories (e.g., thing, place, direction, action, event, manner and amount) and a few major types of events (e.g., go, be and stay), these specialized as necessary to accommodate the circumstances of particular semantic fields (e.g., spatial, temporal, possessive, identificational, circumstantial and existential). Talmy proposes a "cognitive semantics," including a central semantic category of "force dynamics" concerning the interaction of quantities with respect to force. By analogy, a range of phenomena may be viewed in terms of such interactions: mechanical activities, enabling and prevention, social influence, wanting and refraining, modal operators such as "can" and "must," and so forth. Both Jackendoff and Talmy thus offer a degree of encompassing organizational structure for the construction of a broad knowledge base of events supporting causal reconstruction. In particular, these theories suggest the possibility of performing causal reconstruction in a range of domains using only a limited set of core events.

The philosophical position advanced by Lakoff and Johnson [33, 38] also supports the idea of performing causal reconstruction in a range of domains using a limited, core set of events. This work advances the view that metaphorical language appears not only for stylistic reasons, but for reasons of underlying *comprehension* of particular domains in terms of other domains. As illustrations, the authors present a number of extended analogies permitting not only redescription of one activity in terms of another, but as well, reasoning about one activity in terms of another. Examples of these analogies are: an-argument-is-war, time-is-a-limited-resource, the-mind-is-a-container, ideas-are-objects, love-is-a-journey and theories-are-buildings. In [33], an extended examination is given to such analogies arising from basic functions and experiences of the human body. In combination with the work of Jackendoff and Talmy, this theory would seem to suggest that a suitable source of core events supporting causal reconstruction in a range of domains might be a combination of bodily-kinesthetic and simple mechanical interactions.

6 Conclusions

There are three main contributions of this work. The first is a characterization of the causal reconstruction problem, including a set of criteria for evaluating performance of a program at the task and a sample program interface for testing out particular approaches. Second is the transition space representation and its accompanying machinery for associating events and answering questions about causal descriptions. As highlighted in previous parts of this article, key aspects of the transition space representation are its focus on information about *changes* as a basis for determining how the events in a causal description might fit together, and its grounding in simple statements articulated in everyday language. The third component is the PATHFINDER program, including a specific set of association heuristics for use in causal reconstruction, and a range of simple examples processed by the program.

Several immediate extensions of the current approach are possible, as well as some broader extensions. One immediate extension involves expanding the association mechanism of PATHFINDER to produce a *graph* of associations between the referenced events in a description, rather than simply a tree of associations. This would be useful for processing descriptions in which one event summarizes a chain of two or more events, and descriptions involving a fork in the activity which later rejoins (e.g., a nail pushes back each of two bundles of wood fibers, which then spring back to grip the nail). To implement this extension, new heuristics would be required to evaluate the usefulness of adding further associations to an already-connected association structure in order to provide matches for particular assertions not yet matched via inter-event associations. Other immediate extensions involve the addition of a backtracking mechanism or interactivity during the association process in order to counteract or avoid initially-desirable associations which become untenable in light of further association, and an extension of the inference component to generate new assertions and detect conflicts on the basis of all event traces involved in the processing, rather than just individual traces or pairs of traces as in the current version of PATHFINDER.

A related immediate extension concerns the informed application of exploratory transformations. This would involve either the development of suitable heuristics for deciding when and in which manner to transform event traces in order to increase the likelihood of generating partial matches with other traces, or, perhaps more suitably, the use of metric information to gauge the degree of granularity for particular event descriptions with respect to temporal durations, physical size of participating objects, specificity of attributes and so forth. Given this metric information, it should then be possible to determine a suitable level of abstraction and metaphorical specification for all events, then apply only those transformations that take events closer to this common "arena" in transition space.

Moving to broader extensions, the following areas of exploration merit further investigation:

Natural language input. PATHFINDER currently finesses a number of issues in generalized natural language processing and could benefit from the incorporation of additional techniques for handling such things as lexical ambiguity, syntactic and semantic ambiguity, reference, tense and aspect, metonymy, ellipsis, and focus. It is possible, however, that the transition space representation may be able to provide assistance in addressing some of these issues. Concerning reference resolution, for example, it is conceivable that one important constraint governing the association of antecedents with pronouns might be whether or not the particular event occurrences implied by such associations fit together with other events in the sense explored here.

Other types of input. The previous section mentioned the possibility of combined causal reconstruction/causal modeling carried out by a program accepting a written description of an activity while exposed to a demonstration of the same activity. A related extension concerns the processing of written descriptions coupled with diagrams. For both tasks, it may be the case that a more elaborate treatment of *spatial* aspects of physical behavior is required—that is, beyond the propositional treatment currently employed in PATHFINDER. Separately, it may also be useful to explore causal reconstruction involving a combination of text and equations as input, as when a written description is used to supply intuitive background knowledge surrounding the use of a highly-technical model of a physical process.

Incremental causal reconstruction. It may be worthwhile to explore incremental reconstruction of causal scenarios, such that an initial attempt is made to associate the first two events referenced in a description, then incorporate the next referenced event, and so forth. This would appear to be more in line with human comprehension of causal descriptions. This approach was not taken in the construction of PATHFINDER because it requires the development of yet another set of heuristics, these regarding decisions of when to proceed with the best current match for the sentences read so far, versus when to forgo a match in hopes of obtaining a better match once more sentences are read.

Disjunctions of behaviors. Many descriptions of the sort appearing in encyclopedias simultaneously specify several variants of a target behavior through the use of conditional statements, hypothetical situations, or enumeration of alternate behaviors. To handle such descriptions, a mechanism for representing and reasoning about disjunctions of behaviors is required. As part of this, a facility for reasoning about relative *likelihoods* of alternatives would presumably be required. Also related to disjunctive behaviors are repetitive events and feedback cycles. While transformations of the event-object reification variety can be constructed to abstract sequences of repeated activities into simple declarations of repetition of particular events, the ability to handle disjunctions of behaviors could support reasoning about specific numbers of iterations (if small) for these types of events.

Accumulated supporting knowledge. Also mentioned in the previous section was the possible incorporation of a preset knowledge base of core events, rules of inference, rules of restatement and so forth, supporting causal reconstruction in a range of domains. One way to construct such a knowledge base is to perform causal reconstruction in a *cumulative* manner, maintaining supplementary knowledge between processing sessions and perhaps weeding out knowledge used only sparingly in the future. As suggested in the previous section, a good place to start might be with causal situations involving bodily-kinesthetic events or simple mechanical events, then progress to other, less perceptually vivid domains (e.g., radio signals, electricity, biological processes) by relying heavily on analogies to these events. Some of the examples explored on PATHFINDER provide initial support for this hypothesis.

Appendix A. PATHFINDER Heuristics

New event traces are added to the association structure when: (1) the initial set of referenced events and precedents is added, (2) inference and exploratory transformations generate images of these traces, and (3) elaboration of partial matches generates intermediate traces. As characterized in Section 4, inference, reduction and equivalence associations in the association structure determine which traces are considered "active" for matching. New active traces are then matched with all existing active traces not belonging to the same cluster of traces in the association structure.

Partial matches enumerated by the program are labeled as either partial chaining matches or partial restatement matches, using the rule specified in Section 3.1. Partial restatement matches may involve more than a single matching interval. For partial chaining matches, a record is made of assertions appearing in the consequent trace that are covered by assertions in the antecedent trace. For partial restatement matches, a record is made of assertions appearing in the trace previously existing in the association structure that are covered by assertions in the trace newly added to the association structure.

Given a partial match determined in this manner, PATHFINDER first forms a raw score for the partial match, as follows:

((10 × [the number of definite changes—APPEAR, DISAPPEAR, CHANGE, INCREASE and DECREASE—covered in the match]) +
(3 × [the number of other dynamic assertions (NOT-APPEAR, etc.) covered in the match]) +
(1 × [the number of static assertions (PRESENT, etc.) covered in the match]))
÷
[the number of temporal intervals involved in the match]

This raw score is then reduced in response to applicable penalties from the following list (e.g., for two 10% penalties and one 30% penalty, the raw score is multiplied by $(0.9) \times (0.9) \times (0.7)$).

Regarding proximity to description events:

10% penalty	[if the first trace has been formed by exploratory transformation and does not lie along a path of association between two description or precedent events]
10% penalty	[if the second trace has been formed by exploratory transformation and does not lie along a path of association between two description or precedent events]
30% penalty	[if the first trace depicts a precedent event]
30% penalty	[if the second trace depicts a precedent event]

Regarding narrative ordering of event references:

10% penalty	[if the match is not a partial chaining match between successively-referenced events in the description]

10% penalty [if the match is a partial chaining match from one referenced event to the immediately preceding event]

Regarding the current status of the association structure:

10% penalty [if the match involves a hypothesized object]
30% penalty [if the match equates two hypothesized objects]
30% penalty [if the match provides a redundant antecedent for an event]
30% penalty [if the match provides a redundant consequent for an event]
100% penalty [if the match involves two events already connected by a path of associations]

Regarding types of associations:

10% penalty [if the match is a partial restatement match rather than a partial chaining match]
30% penalty [if the match does not complete a path of association fulfilling an explicitly-entered connecting statement]
100% penalty [if the match completes a path of association that conflicts with an explicitly-entered connecting statement]

Once a partial match has been selected for elaboration (and removed from the agenda, three additional tests are made, possibly resulting in a cessation of processing for the partial match and insertion of new, reduced partial matches in the agenda. As described in Section 4, these tests concern (1) equivalences between description objects, (2) logical inconsistency, and (3) violated connecting statements.

Acknowledgement

I would like to thank my thesis supervisor, Patrick Winston, and the members of my thesis committee, Randall Davis, David Waltz and Susan Carey, for helpful guidance and criticism in the course of this research. Also, this research owes much to the comments and suggestions of many members of the Learning Group at the MIT Artificial Intelligence Laboratory, including Rick Lathrop, Jintae Lee, Lukas Ruecker, Sajit Rao, Michael de la Maza, Sigrid Unseld, Zakia Zerhouni, Jonathan Amsterdam and David Kirsh.

References

[1] J. F. Allen, Maintaining knowledge about temporal intervals, *Commun. ACM* **26** (11) (1983) 832–843.

[2] J. F. Allen, Towards a general theory of action and time, *Artif. Intell.* **23** (2) (1984) 123–154.

[3] R. Alterman, A dictionary based on concept coherence, *Artif. Intell.* **25** (2) (1985) 153–186.

[4] G. C. Borchardt, A computer model for the representation and identification of physical events, Master's Thesis, Report T-142, Coordinated Science Laboratory, University of Illinois, Urbana, IL (1984).

[5] G. C. Borchardt, Event calculus, *Proceedings IJCAI-85* Los Angeles, CA (1985) 524–527.

[6] G. C. Borchardt, Transition space, AI Memo 1238, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA (1990).

[7] G. C. Borchardt, Causal reconstruction: understanding causal descriptions of physical systems, Ph.D. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA (1992).

[8] G. C. Borchardt, Understanding causal descriptions of physical systems, *Proceedings AAAI-92* San Jose, CA (1992) 2–8.

[9] T. Bylander and B. Chandrasekaran, Understanding behavior using consolidation, *Proceedings IJCAI-85* Los Angeles, CA (1985) 450–454.

[10] J. G. Carbonell, Learning by analogy: formulating and generalizing plans from past experience, in: R. S. Michalski, J. G. Carbonell and T. M. Mitchell, eds., *Machine Learning: An Artificial Intelligence Approach* (Tioga, Palo Alto, CA, 1983).

[11] J. Clement, A conceptual model discussed by Galileo and used intuitively by physics students, in: D. Gentner and A. L. Stevens, eds., *Mental Models* (Lawrence Erlbaum, London, 1983) 325–340.

[12] A. M. Collins and E. F. Loftus, A spreading-activation theory of semantic processing, *Psychol. Rev.* **82** (6) (1975) 407–428.

[13] R. Davis, Diagnostic reasoning based on structure and behavior, *Artif. Intell.* **24** (1–3) (1984) 347–410.

[14] R. Davis and W. Hamscher, Model-based reasoning: troubleshooting, in: H. E. Shrobe, ed., *Exploring Artificial Intelligence* (Morgan Kaufmann, San Mateo, CA, 1988) 297–346.

[15] J. de Kleer and J. S. Brown, A qualitative physics based on confluences, *Artif. Intell.* **24** (1–3) (1984) 7–83.

[16] A. A. diSessa, Phenomenology and the evolution of intuition, in: D. Gentner and A. L. Stevens, eds., *Mental Models* (Lawrence Erlbaum, London, 1983) 15–33.

[17] R. J. Doyle, Hypothesizing device mechanisms: opening up the black box, Ph.D. Thesis, Technical Report 1047, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA (1988).

[18] R. J. Doyle, Reasoning about hidden mechanisms, *Proceedings IJCAI-89* Detroit, MI (1989) 1343–1349.

[19] J. Doyle and M. P. Wellman, Impediments to universal preference-based default theories, MIT/LCS/TM-416, Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, MA (1989).

[20] M. G. Dyer, *In-Depth Understanding* (MIT Press, Cambridge, MA, 1983).

[21] *The Encyclopedia Americana,* International Edition (Grolier, Danbury, CT, 1989).

[22] K. D. Forbus, Qualitative process theory, *Artif. Intell.* **24** (1–3) (1984) 85–168.

[23] M. R. Genesereth and N. J. Nilsson, *Logical Foundations of Artificial Intelligence* (Morgan Kaufmann, San Mateo, CA, 1987).

[24] D. Gentner, Structure-mapping: a theoretical framework for analogy, *Cogn. Sci.* **7** (2) (1983) 155–170.

[25] D. Gentner and D. R. Gentner, Flowing waters or teeming crowds: mental models of electricity, in: D. Gentner and A. L. Stevens, eds., *Mental Models* (Lawrence Erlbaum, London, 1983) 99–129.

[26] D. Gentner and A. L. Stevens, eds., *Mental Models* (Lawrence Erlbaum, London, 1983).

[27] H. P. Grice, Logic and conversation, in: P. Cole and J. L. Morgan, eds., *Syntax and Semantics, Volume 3: Speech Acts* (Academic Press, London, 1975) 41–58.

[28] S. Hanks and D. McDermott, Default reasoning, nonmonotonic logics, and the frame problem, *Proceedings AAAI-86* Philadelphia, PA (1986) 328–333.

[29] P. J. Hayes, The second naive physics manifesto, in: J. R. Hobbs and R. C. Moore, eds., *Formal Theories of the Commonsense World* (Ablex, Norwood, NJ, 1985).

[30] P. J. Hayes, Naive physics I: ontology for liquids, in: J. R. Hobbs and R. C. Moore, eds., *Formal Theories of the Commonsense World* (Ablex, Norwood, NJ, 1985).

[31] G. G. Hendrix, E. D. Sacerdoti, D. Sagalowicz and J. Slocum, Developing a natural language interface to complex data, *ACM Trans. Database Syst.* **3** (2) (1978) 105–147.

[32] R. Jackendoff, *Semantics and Cognition* (MIT Press, Cambridge, MA, 1983).

[33] M. Johnson, *The Body in the Mind* (University of Chicago Press, 1987).

[34] P. N. Johnson-Laird, *Mental Models* (Harvard University Press, Cambridge, MA, 1983).

[35] D. Kahneman, P. Slovic, and A. Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases,* (Cambridge University Press, 1982).

[36] J. L. Kolodner, Maintaining organization in a dynamic long-term memory, *Cogn. Sci.* **7** (4) (1983) 243–280.

[37] B. Kuipers, Qualitative simulation, *Artif. Intell.* **29** (3) (1986) 289–338.

[38] G. Lakoff and M. Johnson, *Metaphors We Live By,* (University of Chicago Press, 1980).

[39] G. Leech and J. Svartvik, *A Communicative Grammar of English* (Longman, Essex, UK, 1975).

[40] W. G. Lehnert and C. L. Loiselle, An introduction to plot units, in: D. Waltz, ed., *Semantic Structures: Advances in Natural Language Processing,* (Lawrence Erlbaum, London, 1989) 125–165.

[41] D. Lenat, M. Prakash and M. Shepherd, CYC: using common sense knowledge to overcome brittleness and knowledge acquisition bottlenecks, *AI Mag.* **6** (4) (1986) 65–85.

[42] D. B. Lenat and E. A. Feigenbaum, On the thresholds of knowledge, *Proceedings IJCAI-87* Milan, Italy (1987) 1173–1182.

[43] A. M. Leslie and S. Keeble, Do six-month-old infants perceive causality? *Cogn.* **25** (1987) 265–288.

[44] M. McCloskey, Naive theories of motion, in: D. Gentner and A. L. Stevens, eds., *Mental Models,* (Lawrence Erlbaum, London, 1983) 299–324.

[45] D. McDermott, A temporal logic for reasoning about processes and plans, *Cogn. Sci.* **6** (2) (1982) 101–155.

[46] A. Michotte, *The Perception of Causality,* translated by T. and E. Miles from French edition, 1946 (Methuen, London, 1963).

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

[48] L. Morgenstern and L. A. Stein, Why things go wrong: a formal theory of causal reasoning, *Proceedings AAAI-88* Saint Paul, MN (1988) 518–523.

[49] N. H. Narayanan and B. Chandrasekaran, Reasoning visually about spatial interactions, *Proceedings IJCAI-91* Sydney, Australia (1991) 360–365.

[50] D. Newtson and G. Engquist, The perceptual organization of ongoing behavior, *J. Exp. Soc. Psychol.* **12** (1976) 436–450.

[51] D. Newtson, G. Engquist and J. Bois, The objective basis of behavior units, *J. Pers. and Soc. Psychol.* **35** (12) (1977) 847–862.

[52] P. Norvig, Marker passing as a weak method for text inferencing, *Cogn. Sci.* **13** (4) (1989) 569–620.

[53] J. Pearl and T. S. Verma, A theory of inferred causation, *Proceedings 2nd International Conference on Principles of Knowledge Representation and Reasoning,* Cambridge, MA (1991) 441–452.

[54] M. R. Quillian, The teachable language comprehender: a simulation program and theory of language, *Commun. ACM* **12** (8) (1969) 459–476.

[55] C. Rieger, An organization of knowledge for problem solving and language comprehension, *Artif. Intell.* **7** (2) (1976) 89–127.

[56] R. C. Schank and C. K. Riesbeck, *Inside Computer Understanding* (Lawrence Erlbaum, London, 1981).

[57] V. Sembugamoorthy and B. Chandrasekaran, Functional representation of devices and compilation of diagnostic problem-solving systems, in: J. Kolodner and C. Riesbeck, eds., *Experience, Memory, and Reasoning* (Lawrence Erlbaum, London, 1986) 47–73.

[58] Y. Shoham, Chronological ignorance: time, nonmonotonicity, necessity and causal theories, *Proceedings AAAI-86* Philadelphia, PA (1986) 389–393.

[59] L. Talmy, Force dynamics in language and cognition, *Cogn. Sci.* **12** (1) (1988) 49–100.

[60] M. B. Vilain, A system for reasoning about time, *Proceedings AAAI-82* Pittsburgh, PA (1982) 197–201.

[61] M. Vilain and H. Kautz, Constraint propagation algorithms for temporal reasoning, *Proceedings AAAI-86* Philadelphia, PA (1986) 377–382.

[62] D. L. Waltz, Event shape diagrams, *Proceedings AAAI-82* Pittsburgh, PA (1982) 84–87.

[63] P. A. White, Causal processing: origins and development, *Psychol. Bull.* **104** (1) (1988) 36–52.

[64] P. H. Winston, Learning new principles from precedents and exercises, *Artif. Intell.* **19** (3) (1982) 321–350.